Craniofacial Identification

The promotion of CCTV surveillance and identity cards, along with ever-heightened security at airports, immigration control and institutional access, has seen a dramatic increase in the use of automated and manual recognition. In addition, several recent disasters have highlighted the problems and challenges associated with current disaster victim identification.

Discussing the latest advances and key research into identification from the face and skull, this book draws together a wide range of elements relating to craniofacial analysis and identification. It examines all aspects of facial identification, including the determination of facial appearance from the skull, comparison of the skull with the face and the verification of living facial images. With sections covering the identification of the dead and of the living, it provides a valuable review of the current state of play along with the latest research advances in this constantly evolving field.

Caroline Wilkinson is Professor in Craniofacial Identification at the Centre for Anatomy and Human Identification at the University of Dundee. She is an expert in facial anthropology and author of *Forensic Facial Reconstruction* (Cambridge University Press, 2004).

Christopher Rynn is a Postdoctoral Researcher at the Centre for Anatomy and Human Identification at the University of Dundee. He is an expert in craniofacial approximation/reconstruction and anthropology, and in forensic image enhancement and analysis.

Craniofacial Identification

Edited by

Caroline Wilkinson and Christopher Rynn
Centre for Anatomy and Human Identification
University of Dundee

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS
Cambridge, New York, Melbourne, Madrid, Cape Town,
Singapore, São Paulo, Delhi, Mexico City

Cambridge University Press
The Edinburgh Building, Cambridge CB2 8RU, UK

Published in the United States of America by Cambridge University Press, New York

www.cambridge.org
Information on this title: www.cambridge.org/9780521768627

First published 2012

Printed in the United Kingdom at the University Press, Cambridge

A catalogue record for this publication is available from the British Library

Library of Congress Cataloguing in Publication data
Craniofacial identification / edited by Caroline Wilkinson and Christopher Rynn.
p. cm.
Includes bibliographical references and index.
ISBN 978 0 521 76862 7 (hardback)
1. Facial reconstruction (Anthropology) 2. Face – Identification. 3. Face – Imaging. 4. Biometric identification. I. Wilkinson, Caroline, 1965– II. Rynn, Christopher. II. Title.
GN74.C73 2012
599.9′48–dc23 2011034917

ISBN 978 0 521 76862 7 Hardback

Contents

The colour plates are to be found between pages 88 and 89.

Contributors

Jeremy N. Bailenson
Virtual Human Interaction Lab
Stanford University

Tatiana Balueva
Russian Academy of Science

Vicki Bruce
School of Psychology
Newcastle University

Natalie Butcher
School of Social Sciences and Law
Teesside University

Peter Claes
Centre for Processing Speech and Images
ESAT–PSI
Future Health Department – IBBT
Katholieke Universiteit, Leuven

John Clement
Dental School
University of Melbourne

Josh P. Davis
Department of Psychology and Counselling
University of Greenwich

Sven De Greef
Forensic Odontology
Department Oral Health Services
Katholieke Universiteit, Leuven

Charlie Frowd
University of Central Lancashire

Stuart Gibson
School of Physical Sciences & Visionmetric Ltd
University of Kent, Canterbury

Peter J. B. Hancock
Department of Psychology
University of Stirling

Catriona Havard
School of Psychology
University of Aberdeen

David Hunter
School of Psychology
University of St Andrews

Karen Lander
School of Psychological Sciences
University of Manchester

Carrie Leonetti

Gregory Mahoney

Matthew Maylin
School of Physical Sciences & Visionmetric Ltd
University of Kent, Canterbury

Amina Memon
Royal Holloway University of London

Joe Mullins
Forensic artist

David Perrett
School of Psychology
University of St Andrews

Stephen Richmond
Dental School
Cardiff University

Christopher Rynn
Centre for Anatomy & Human Identification, CLS
University of Dundee

Kathryn Y. Segovia
Virtual Human Interaction Laboratory
Stanford University

Chris Solomon
School of Physical Sciences & Visionmetric Ltd
University of Kent, Canterbury

Paul Suetens
Centre for Processing Speech and Images
ESAT–PSI
Future Health Department – IBBT
Katholieke Universiteit, Leuven

Bernard Tiddeman
Department of Computer Science
Aberystwyth University

Amy Tillotson
Centre for Anatomy and Human Identification
University of Dundee

Arshed Toma
Dental School
Cardiff University

Tim Valentine
Department of Psychology
Goldsmiths, University of London

Dirk Vandermeulen
Centre for Processing Speech and Images
ESAT–PSI
Future Health Department – IBBT
Katholieke Universiteit, Leuven

Ludo Vermeulen
Forensic artist

Elizaveta Veselovskaya
Russian Academy of Science

Caroline Wilkinson
Centre for Anatomy & Human Identification, CLS
University of Dundee

Guy Willems
Forensic Odontology
Department Oral Health Services
Katholieke Universiteit, Leuven

Mineo Yoshino
National Research Institute of Police Science
Chiba, Japan

Alexei Zhurov
Dental School
Cardiff University

Chapter 1

Familiar face recognition

Vicki Bruce

1.1 Introduction

The task of the police officer investigating a crime, or of the forensic anthropologist working with human remains, is to establish an identity of a criminal or of a victim. This often involves working with images of faces – building a composite image from the memory of a witness to the crime, seeking CCTV images of the person or persons who might have committed the crime, or building a model of the face of an unknown person from their skull. Many of the chapters in this volume describe the processes involved in such reconstructions.

Once an image of a face is obtained, however, it needs to be identified. TV programmes such as Crimewatch in the UK (see www.bbc.co.uk/crimewatch) often display witness-generated facial composites or CCTV images of people associated with crimes which sometimes result in people volunteering new information; Richard Neave's constructed image of the unidentified victim 115 of the London Underground tube station fire at Kings Cross in 1987 (Chambers 2007) was displayed widely in newspapers and posters. CCTV images of the London 'nail bomber' in 1999 led to hundreds of responses, including one phone call from a man who named a work colleague, David Copeland, as resembling the images. This crucial lead quickly led to the arrest of David Copeland, unfortunately not soon enough to prevent the third devastating bomb attack on The Admiral Duncan, a London pub.

While the problems of building or using images of unfamiliar faces are the subject of many other chapters in this volume, here I focus on relevant issues about the identification of familiar faces, and also elaborate on the contrast between familiar and unfamiliar face recognition.

1.2 Familiar versus unfamiliar face recognition

When police investigate crimes the identity of the villain or villains is often unknown, and the police will question witnesses who may be asked in different ways to assist with establishing the identity of a person who was unfamiliar to them before the criminal incident. When a witness to a crime is asked to try to identify the criminal from photographs or a line-up, or to build an image of the face using a face composite system, this is typically difficult and error-prone. In contrast, most of us manage to recognise the faces of familiar people in our everyday lives or on the television or other media reasonably accurately, despite the enormous variation shown between different appearances of the same person. Error-prone unfamiliar face memory is transformed to a generally very reliable performance with known faces.

There is some neuropsychological evidence that somewhat different processes underlie the recognition of unfamiliar faces and the recognition of familiar ones. 'Prosopagnosic' people have difficulties recognising faces, sometimes as a result of brain injury. They may be unable to recognise even their closest relatives from their faces, and even their own face in the mirror may appear unfamiliar to them (see Harris and Aguirre, 2007, for a brief overview of this condition, and Duchaine and Nakayama, 2006 for one of many recent papers on congenital or developmental prosopagnosia). But some prosopagnosic patients can manage to match images of unfamiliar faces, albeit using laborious and

Craniofacial Identification, ed. C. Wilkinson and C. Rynn. Published by Cambridge University Press.

sometimes time-consuming processes. Young and his colleagues (1993) investigated residual deficits in face processing in a group of 34 brain-injured war veterans. Amongst this group they found one man who was impaired on familiar face recognition, but unimpaired at unfamiliar face matching, and another who was impaired on unfamiliar face matching but recognised familiar faces quite normally. This pattern of 'double dissociation' is consistent with the idea that different brain areas and/or processes are involved in the two tasks of recognising familiar faces and matching unfamiliar ones.

Further evidence for the dissociation between familiar and unfamiliar face processing is the observation that familiar face recognition seems to depend on rather different kinds of information from unfamiliar recognition. Lander and Butcher (Chapter 11) review work suggesting that representations of familiar faces capture something about their characteristic patterns of motion as well as static form. Even within static form, there appear to be different emphases in the visual representations of familiar and unfamiliar faces. Ellis *et al.* (1979) first demonstrated that unfamiliar face recognition was dominated by the external features of the face, including the hair. People find it very difficult to recognise once-viewed faces if their hair is concealed or changed. However, when faces are familiar their internal features are more important in their recognition. This dominance of the external over internal features for unfamiliar faces has been much-replicated (e.g. Young *et al.*, 1985*b*) and it is found in tasks that involve just matching unfamiliar faces without any memory load (Bruce *et al.*, 1999). O'Donnell and Bruce (2001), using newly familiarised faces, showed that it seemed to be the eyes in particular that became better represented in recently familiarised faces.

Why should representations of familiar faces become shifted towards internal and away from external features? For an unfamiliar face, at least in Western Europe where hair is generally visible and very variable in style and colour, the outer features of the face probably convey the most information that will be useful for matching and memory. As we see a face more frequently, however, we will attend more to its internal features, to see direction of gaze, expression and lip movements, which are all important for other social functions (see Bruce and Young, 1998 for an introductory overview). Moreover, hairstyle will vary from one occasion to another, and people may sometimes wear hats or scarves. Thus it is not surprising that the visual memory system begins to weight the internal features more strongly. Consistent with this, Megreya and Bindemann (2009) showed that when face recognition was tested in Egypt, adults, but not children, showed better memory for internal features for unfamiliar as well as familiar faces, probably because the commonplace wearing of head scarves means that attention is always oriented more towards internal than external features in that culture.

The relative importance of the internal features of a familiar face, particularly the eyes, has implications for the effectiveness of different kinds of disguises. The true identities of Batman, and The Lone Ranger, were effectively concealed by their wearing of masks covering their eyes, in addition to hats that covered their hair. Sunglasses and hats can allow celebrities anonymity in a crowd.

Burton and his colleagues (2005) have suggested that representations of familiar faces are built up by simply averaging together individual instances of seen faces. The differential weightings of internal features during familiarisation, and for cultures where external features are not often viewed, will likely arise as a result of selective attention weighting areas of the face differently. An averaging mechanism with the addition of selective weighting through attentional mechanisms would allow the representation of a face to develop in a way which allows the face to be well-recognised, despite variations in expression and viewpoint which are extremely detrimental to recognition of unfamiliar faces (e.g. Bruce, 1982). Burton and his colleagues have shown that people are faster to recognise familiar faces when shown an average of 20 different photographic instances of them, than when shown individual instances alone. Moreover, in an interesting practical extension of this work, Jenkins and Burton (2008) showed that when an average of 20 instances of each of a set of famous face targets was used as a probe to match against a database of images of over 3000 celebrities, performance increased from 51% correct matching obtained from using individual image probes, to 100% correct when the averages were used as probes. From the evidence of this preliminary study Jenkins and Burton suggest that average images should be used on identity documents such as passports to improve both human and computer use of such images.

So, our representations of unfamiliar faces make it very difficult to generalise to novel views, expressions and contexts, while our representations of familiar faces transcend these limitations, perhaps because of consolidation of different instances to an average. We can recognise familiar faces well, even from low-quality images. Harmon (1973) showed good recognition of familiar faces in pixellated images where only very coarse-scale information about identity was preserved. And Burton and his colleagues (1999) showed that students could recognise with high accuracy very poor quality CCTV images of their lecturers *provided* that images of the face area were visible, even though these were very degraded. Indeed we are so good at recognising familiar people from moving degraded images that there must be real concern about the efficacy of attempts to conceal identities in TV clips showing witnesses or children. TV editors working on filmed footage are unfamiliar with the people depicted, and it is possible that the degree of blurring or pixellation applied, which seems satisfactory to them, will be insufficient to render these faces unrecognisable to familiar observers (Lander *et al.*, 2001).

The representation of familiar faces emphasises internal features, particularly the eyes. Several recent studies have demonstrated that prosopagnosic people do not look at or use the eyes/upper face features when looking at faces and trying to recognise them (Caldara *et al.*, 2005; Bukach *et al.*, 2008). This is likely to be a symptom rather than a cause of their deficit, however, since prosopagnosic people are often completely normal at recognising facial expressions and perceiving eye gaze (e.g. Young *et al.*, 1993; Duchaine *et al.*, 2009), both of which also use information from the eye regions of the face.

In many other respects, though, the perception of familiar faces seems very similar to that of unfamiliar faces. Both familiar and unfamiliar faces appear to be processed 'configurally' (see Chapter 2) though more piecemeal processing may be beneficial in certain kinds of matching task (see later). The recognition of both familiar and unfamiliar faces is impaired by unusual transformations such as inversion and negation, or by changes in lighting direction which go beyond those usually experienced. Such extreme variations in orientation, contrast and luminance create changes well beyond the usual range of variation encountered for an individual face and which would therefore not be incorporated in the internal representation for a familiar one. Johnston and Edmunds (2009) provide a recent review of factors affecting unfamiliar and familiar face recognition.

1.3 Individual differences in face recognition

In recent years there has been renewed interest in the theoretical and practical implications of individual differences in face recognition. As with all other psychological abilities, tests of a reasonable sample of participants will reveal a range of performance scores. Recent interest has focused on understanding reasons for this variation (general ability? or something specific to face processing?). Woodhead and Baddeley (1981) noted a wide variation in abilities in face-recognition tests ranging from performance in discriminating familiar from novel faces at or near chance (d' = 0) to performance which was near perfect (d' = 6.8). Testing a group of relatively good and relatively poor recognisers again later, they found a significant difference in face recognition abilities remained, and the groups also differed slightly in their recognition memory ability for non-face pictures, but there was no difference in their verbal memory abilities.

Differences are observed in face-matching ability when there is no memory component at all. Megreya and Burton (2006) used a face-matching task introduced by Bruce *et al.* (1999), where participants must decide which, if any, of 10 faces in an array matches a good-quality target photograph. Performance on such tasks is surprisingly prone to error, but also quite variable (Megreya and Burton found mean performance of 82% correct and a standard deviation of 12%). An even simpler task, the Glasgow Face Matching Task (GFMT) requires participants to decide if two face images are the same or different people (see Figure 1.1). Using the short form of this test with 194 volunteers, Burton *et al.* (2010) report a mean of 81.3% and standard deviation of 9.7 on this task, which has been replicated in Newcastle.[1] In this sample (Burton *et al.*, 2010), and in the samples tested in Newcastle about 10% of participants scored between 51% (around chance) and 70% correct, while the best 10% scored above 95% correct on the test.

Megreya and Burton (2006) examined what factors correlated with performance in the array-matching task. There were significant positive correlations between unfamiliar face matching and recognition memory for unfamiliar faces, and the matching task also correlated with a number of other tasks of visual memory and processing. The matching task did *not* correlate, however, with participants' abilities to

Figure 1.1 Are the two people shown the same or different? This can be a surprisingly difficult task. Figure courtesy of Mike Burton, Glasgow University, and also published in Burton *et al.* (2010).

recognise famous or recently familiarised faces – unless these were turned upside down! Indeed the best predictor of performance on matching unfamiliar faces when these were shown *upright*, was conducting the same or similar tasks with unfamiliar or familiar faces when these were *inverted*. Now we know that when faces are inverted, participants must rely more on analysing faces in a piecemeal way since the 'configural' processing that characterises upright face perception is impaired when faces are inverted (e.g. see Young *et al.*, 1987; Leder and Bruce, 2000). Megreya and Burton's finding suggests that matching unfamiliar faces well requires the ability to analyse local features and ignore the overall impression gained by more configural processing.

This suggestion that unfamiliar face matching requires a specific analytic processing style does not, however, account for the broader range of abilities to recognise faces from memory. Prosopagnosic people, who have difficulties recognising familiar faces in their everyday lives, seem to have specific deficits in processing 'configural', not local, face features. More recently, Russell *et al.* (2009) have studied volunteers who are particularly good at recognising faces: 'Super-recognisers'. These score extremely highly on the Cambridge Face Matching, as well as the Cambridge Face Memory tests (tests of unfamiliar face processing), but are not particularly skilled at tasks of inverted face recognition.

This suggests that skill with recognising familiar faces and remembering unfamiliar faces in some tasks relies on expertise in configural processing, but that other tasks involving faces, particularly matching unfamiliar faces, *also* require the ability to analyse local features well. Indeed these observations themselves reinforce earlier theoretical suggestions that there is a distinction between the processing of familiar and unfamiliar faces (Bruce and Young, 1986). If prosopagnosia is largely an impairment of the 'configural' processing of faces – on which most people are expert – then the observation that some prosopagnosics can perform well on unfamiliar tasks that might rely more on piecemeal processing is less surprising.

1.4 Stages in person recognition

When we meet someone we know well, and are expecting to see, then person recognition is usually effortless and complete. 'Hello Vicki' my father used to say, immediately, even though at the age of 90 he had very poor vision. But this facility in familiar person recognition belies a number of discrete stages in the process that can be revealed through systematic recording of everyday behaviour or more laboratory-based study.

Young *et al.* (1985*a*) collected a large number of instances of errors and difficulties in person recognition in a study that they imaginatively titled 'The faces that launched a thousand slips'. They asked 22 volunteers to note down any difficulties they experienced when recognising people for a period of 8 weeks, and they analysed the 922 recorded incidents from the 7 weeks following the initial familiarisation week. The first thing to note is that this equates to almost one incident per day, on average – showing that everyday person recognition is far from problem-free. Next was the grouping of incidents into distinct types. Most common (about a third of all the records) were when a person was misidentified – one person was confused with someone else, or a stranger was taken to be someone familiar, e.g. '. . . I saw a person with a dog and I thought it was a dog owner I sometimes see there. It was the wrong type of dog: I thought he must have got a new one!' Next most common – approximately 25% of all records – occurred where someone appeared to be familiar but the volunteer was unable to remember why they were familiar. Sometimes this situation was resolved by a protracted process, e.g. by the person in the bank who 'saw a person and I knew there was something familiar immediately. After a few seconds I realised she was from a shop on campus or a secretary of one of the departments. I eventually remembered by a process of elimination'. And the third most common type – 21% of records – was where it proved difficult or impossible to retrieve a full identification of an encountered person – often it was the name that proved elusive, e.g. someone recognised an actress from a poster and 'knew what films the actress was in and knew she does a lot of Beckett, but it was another minute before I could remember her name'.

These three types – misidentification, familiar only, and difficulty in retrieving full details together account for almost 80% of the incidents. While not all the records involved face recognition (because there were some concerning, for example, failures to recognise voices), these three kinds of difficulty map well onto a broad 'flow model' of the stages involved in recognising someone from their face, shown in Figure 1.2.

Figure 1.2 A simple three-stage model of person identification from the face.

This flow model describes three broad stages in person recognition. Step 1 is to match the incoming visual pattern against a stored visual representation of what that person looks like. If there is a match at this level then the pattern may be recognised as 'familiar'.[2] Step 2 involves retrieving information about why the person is familiar – where you know them from, what they do for a living and so forth. The final step, according to this simple 'stage' model, involves retrieving their name.

The flow model shows the steps occurring in this strict order. There is no route directly from the pattern-matching stage to the name, for example. There is a good deal of evidence that appears consistent with this proposal that name retrieval cannot occur before something else about the person is known. In the diary study (Young *et al*., 1985*a*), there was not a single recorded error or difficulty in person recognition where a face was named but there was no knowledge of who they were. People never say things like 'I know that face – it's John Lennon – but I have no idea who John Lennon is or where I have seen this face before'. A more formal test of this strict sequence was conducted by Brennen *et al*. (1990) in an experiment which required that people answer a series of questions about celebrities from definitions – as they might if they were playing Trivial Pursuit, for example (e.g. What's the name of the person who played the nervous man in the shower scene in Hitchcock's *Psycho*?). For some questions, people knew the answers. Others, they knew they didn't know. But sometimes they felt they knew the answer but the name was 'on the tip of their tongue' and they just couldn't get to it. In this case, according to the sequence of steps, the participant is at Step 2, and stuck on the process of getting beyond that to the name. So, what happens if you see a picture of the person whose name you are trying to retrieve? Brennan *et al*. compared the correct names retrieved from tip-of-the-tongue states when a picture of the unnamed person was shown with when the original question was just repeated. Seeing the face gave no advantage at all. Indeed it appeared to annoy participants, who would say, through gritted teeth, such things as 'I know what he *looks* like, I just can't remember the name!' (Anthony Perkins, by the way.)

So it looks as though identifying a familiar person has (at least) three stages and problems can arise at any one of them. A person may be mistakenly not recognised at all, or they may be misidentified because the visual pattern gets matched to the wrong stored visual representation. (For example, I have recently discovered that I had a long-standing confusion in my own mind between two actresses – Zoe Wanamaker and Barbara Flynn – who I thought were both Zoe Wanamaker. Indeed it was only when my partner denied that we were watching Zoe Wanamaker in an old episode of the crime thriller *Cracker*, leading me to search google web images, that I was convinced that these really were two different women. So, in terms of our flow-chart, all the times I looked at Barbara Flynn and thought she was Zoe Wanamaker I was matching the pattern of her face against the wrong visual representation. Moreover, through doing this, I was creating an internal representation for these two women which was some kind of average of the two of them.) In a more important forensic context, there is plenty of scope for a police officer, or a member of the public viewing TV images, to mistakenly identify someone as a person they know, given a certain resemblance and, perhaps, some context (see later).

At the second stage, a person may seem familiar, correctly, but the details of why they are familiar may be elusive, or misremembered. One of the more intriguing things about the 'super-recogniser' study (Russell *et al*., 2009) was their claimed ability not just to remember faces, despite a considerable lapse of time, but to know *why* they seemed familiar too. One of the super-recognisers remarked 'I've learned to stop surprising people with bizarre comments like, "Hey, weren't you at that so-and-so concert last fall . . . I recognise you"' and another said 'I do have to pretend that I don't remember, however, because it seems . . . that they mean more to me than they do when I recall that we saw each other once walking on campus four years ago in front of the quad!'. Remembering why someone is familiar may involve retrieving specific *episodic* information, such as this (I saw him then and there) or it may be more general '*semantic*' information – that's the person who reads the news on TV. The simple stage model outlined here does not adequately distinguish between these two different kinds of knowledge we may have of why a face is familiar to us. These differences are important in a forensic context. We might see a CCTV image of a person wanted in connection with a crime and 'know' that is our next-door neighbour, or the person we buy our papers from, or we may 'know' this is the same person we interviewed in connection with a similar incident last year. But when it comes to the provision of an alibi, for example, we may need to remember not just who a particular person is, when questioned, but that we definitely did see them on their regular train to work that morning at the same time they were supposed to be committing a crime somewhere else.

The most difficult thing about identifying familiar people is remembering their names, even when these are highly familiar to us. There's something rather strange about names, even when they are short, simple and concrete. McWeeney *et al*. (1987), for example, discovered that people found it harder to learn the same lexical items when presented as names ('Mr Baker') than when presented as occupations ('a baker'). The simple stage model places names as the final stage in a sequence of identification steps, and there is certainly plenty of evidence consistent with the 'last stage' account. First, as noted above, names are never retrieved without knowing something else about the person. Second, tasks which require names to be retrieved or verified are conducted more slowly than tasks that require other kinds of information to be used. For example, Burton, Jenkins and McNeill (2002; see Calderwood and Burton, 2006) repeatedly showed pictures of just four famous faces – two pop stars and two politicians, two called Peter and two called Paul. Even after extended practice participants were slower to say the name than the occupation of each face that they saw. A control study showed that other participants were no slower to read the names than the occupations out loud suggesting it is not because of difficulties articulating these labels. Johnston and Bruce (1990) asked participants to make speeded judgements about whether pairs of faces shared the same first name, or shared properties such as nationality or being dead or alive (e.g. if on a trial the participants saw pictures of John Lennon and James Dean they would respond 'no' if the question was whether they shared a name; and 'no' if they were asked if they matched nationality or not, but 'yes' to the question of matching on dead–alive). Again, despite the use of a set of just eight repeated, highly familiar faces, the decisions which involved names were made more slowly, even than decisions about whether the two depicted individuals were both dead, or both alive.

However, the slowness and error-prone nature of name retrieval need not mean that they are reached 'after' other person information, as in Figure 1.2. It may just be that names are generally available later than such information. Burton and Bruce (1992) demonstrated, using an 'interactive activation with competition' (IAC) model, that empirically observed effects could be modelled without the need to posit an extra stage for names. The IAC model of person recognition supposes that faces (or voices, or names) are recognised as familiar when sufficient activation is present within 'person identity nodes' (PINS) which pool activity from modality-specific pattern processors that respond separately to faces, voices and so forth. Personal information, including such things as occupations, nationalities and so forth, is retrieved via the PINS. Burton and Bruce demonstrated that if proper names were treated just the same way as other things known about people, and placed within the same pool of personal information, they naturally became activated more slowly than other kinds of information. This is because a name is a unique item of information about identity. We know only one Tony Blair and one Margaret Thatcher. Occupations, however, even rare ones like Prime Minister, are shared by several, sometimes many, people. Thus the PINS for Tony Blair and Margaret Thatcher (and Winston Churchill and Gordon Brown, etc.) will all have a link to the piece of information about 'Prime Minister'. These multiple links make activation in shared nodes rise more quickly than in unique nodes, simulating neatly the differential ease of determining occupations and other types of information compared with names.

While the IAC model incorporates activation between different pools there is inhibition between all units within a single pool (there is 'competition' as well as 'interactive activation'). In the case of names, though, the original Burton and Bruce (1992) proposal cannot be quite right. Bredart *et al.* (1995) demonstrated that there cannot be a *single* pool of undifferentiated information covering names, occupations, nationalities, favourite foods and so forth, because if this was the case then it would be harder to retrieve information about people we know a lot about compared with people we know rather little about, because additional information would inhibit the activation of any particular semantic unit, including that for the name. Bredart and his colleagues (1995) suggest instead that there are separate pools for distinct kinds of person identity – a pool for nationality, a pool for occupation, a pool for names – and inhibition occurs within each pool separately.

The suggestion that names are at the same stage as other personal information and retrieved more slowly due to their connection patterns implies that in certain contexts, if the connection to a name is strong enough, it might be retrieved more quickly than, say, occupation. Calderwood and Burton (2006) have demonstrated this with personally familiar people. In line with our intuitions, perhaps, it is indeed easier to recall your partner's name than his/her occupation.

Although the issue about memory for names seems to be a rather theoretical one, there are implications of this for investigative work. It is important to note that remembering names, even of quite familiar people, is difficult and prone to error. While the police seek a name for a person seen at a crime, images from cameras or composites may trigger less specific cues to identity – 'the bloke I used to work with' – and these should be seen as important leads even when names are not known or remembered.

1.5 Contextual factors in person recognition

We are often surprised to meet people we know in unexpected places, and we may fail to recognise them completely under certain circumstances. Australian psychologist Don Thomson (1986), contrived a situation in which parents even failed to recognise their own daughter, by exploiting her unexpected presence in London when the parents were visiting there, and instructing her to stand near their hotel with an unfamiliar accomplice. The daughter was told to show not a single sign of recognition when her parents rushed up to greet her, and as a result the parents apologetically moved on. Although this seems unlikely, imagine you saw Paul McCartney in your local launderette – you would probably assume it wasn't him, but was just a striking resemblance. Don Thomson's friends clearly recognised their daughter, but then decided instead that she was someone who *looked like* (incredibly like) her. It is not quite clear, though, whether an inappropriate or unlikely context can make an otherwise familiar face seem completely unfamiliar and it would be difficult to contrive an experimental test of this.

Context can play some strange tricks, however. The same Don Thomson was once accused of a serious sexual assault and this seems to have been because he was appearing on the television while the unfortunate woman was being raped. His face therefore became associated with the incident. The current Labour MP for Neath and former government minister, Peter Hain, was accused of robbing Barclays Bank in 1975 – when he was a postgraduate student and leader of the Young Liberals and an anti-apartheid activist. He had been briefly in the vicinity of the bank at the time it was robbed, as he had been buying a new typewriter ribbon, and so it would have been possible for passers-by to note his presence out that day. There have even been suggestions that the robbery may have been deliberately staged using a look-alike of Peter Hain, whose political activities were not welcomed by the (then) apartheid regime in South Africa. Whatever the truth of such suggestions, newspapers published his picture with a headline about Hain's arrest *before* the main witness – the bank cashier – was shown a line-up in which he appeared, making it extremely likely that the witness could have found his face familiar, but not from the crime scene itself. Indeed even without the newspaper headline, someone who has some degree of celebrity from media coverage might be particularly susceptible to being mistakenly associated with a crime by virtue of their apparent familiarity.

The research by Young and his colleagues (1985*a*) on everyday errors and difficulties in person recognition, and other laboratory research by Hay, Young and Ellis (1991) and Hanley and associates (e.g. Hanley and Cowell, 1988) has demonstrated clearly that people quite often judge a face to be familiar without remembering why. Approximately 5–10% of attempts to identify a celebrity (Hanley and Cowell, 1988) led to judgements that the face was familiar, with no further information forthcoming. It is therefore very important that eyewitnesses asked to scrutinise faces in a line-up should not have been exposed to the person's face earlier, and that people who could be familiar to an eyewitness through general media exposure or from otherwise innocent contexts should not be placed in a line-up at all. In several cases reviewed by the Devlin committee (Devlin, 1976) witnesses were asked to observe a line-up having already been shown photographs of the suspect. For example, in the case of George Ince, who was tried in 1973 for murder following an armed robbery, the key witness who picked him out from an identity parade had already been shown photographs of him prior to the line-up.

Appropriate contextual information can also facilitate or 'prime' recognition of highly familiar faces in a way that speeds up familiarity decisions. Bruce and Valentine (1986) first demonstrated that faces from closely associated pairs – such as Stanley Laurel from 'Laurel and Hardy' were judged familiar more quickly if they immediately followed an image of their associated partner's face. This priming effect suggests that the representations of different people in memory are interconnected in some way. The IAC model introduced earlier provides one account of such priming effects.

Sinha and Poggio (1996) discussed a contextual priming effect possibly related to this. At the time it was published, Bill Clinton was President and Al Gore Vice-President of the USA. Viewers of the picture at that time quite readily recognised Al Gore standing beside Bill Clinton, even though the image shown actually has Bill Clinton's internal face features pasted into Al Gore's hairstyle. This demonstration depended upon the relative unfamiliarity of Al Gore at that time (hence the dominance of external face features given the appropriate context for his appearance). One way of explaining this would be to suggest that the recognition of Clinton's face produced activation of the recognition unit for Gore's which therefore required less evidence (and in this case, less accurate evidence) to become active too. Such demonstrations suggest that we might be convinced we have seen someone we expect to see at a particular time or place when actually we have only seen someone who resembles them.

It is clear that contextual factors can be very important in our recognition of familiar people. In a criminal context it may be extremely important that people are able to remember when or where they saw someone. A witness to a crime must know that the person in the line-up is familiar from the scene of the crime (and not from the newspapers, as in Peter Hain's case). And an alibi may be asked to testify that they saw a particular acquaintance in a place (not the crime scene) at a specified time, so they too must remember the context of a recent encounter with a known individual. While familiar people's identities are often bound up with places where they are seen ('are you on the telly?', or 'that's the man I see walking his dog') we know rather little about how well we are able to remember specific episodes involving particular people. I see the same elderly gentleman walking his greyhound most weekends, when my own walks are later than during the week. But was it last Saturday or Sunday that I saw him last? I couldn't say.

1.6 Conclusions

We are good at familiar face recognition. Members of the public, witnesses or police who claim to recognise a person in a CCTV image or other context as someone they know should be taken very seriously. This is why getting a good image of a face, via composite construction or CCTV, and broadcasting it widely, is so important to investigations. But we are not perfect. We may confuse two people, be unable to remember their names, or be misled by context. A witness must know that someone was familiar from the scene of the crime, rather than some other place. And the alibi must do more than recognise someone, they must be able to swear they were in a particular place at a particular time. This memory for time and place is a quite different ability, and one which we know less about, but may be absolutely critical in an investigation.

References

Bredart, S., Valentine, T., Calder, A. J. and Gassi, L. (1995). An interactive activation model of face naming. *Quarterly Journal of Experimental Psychology*, **48**, 466–486.

Brennen, T., Baguley, T., Bright, J. and Bruce, V. (1990). Resolving semantically induced tip-of-the-tongue states for proper nouns. *Memory and Cognition*, **18**, 339–347.

Bruce, V. (1982). Changing faces – visual and non-visual coding processes in face recognition. *British Journal of Psychology*, **73**, 105–116.

Bruce, V., Henderson, Z., Greenwood, K. *et al.* (1999). Verification of face identities from images captured on video. *Journal of Experimental Psychology: Applied*, **5**, 339–360.

Bruce, V. and Valentine, T (1986). Semantic priming of familiar faces. *Quarterly Journal of Experimental Psychology*, **38A**, 125–150.

Bruce, V. and Young, A. W. (1986). Understanding face recognition. *British Journal of Psychology*, **77**, 305–328.

Bruce, V. and Young, A. (1998). *In the Eye of the Beholder: The Science of Face Perception*. Oxford: Oxford University Press.

Bukach C. M., Le Grand, R., Kaiser, M. D., Bub, D. N. and Tanaka, J. W. (2008). Preservation of mouth region processing in two cases of prosopagnosia. *Journal of Neuropsychology*, **2**, 227–244.

Burton, A. M., and Bruce, V. (1992). I recognize your face but I can't remember your name – a simple explanation. *British Journal of Psychology*, **83**, 45–60.

Burton, A. M., Bruce, V. and Johnston, R. A. (1990). Understanding face recognition with an interactive activation model. *British Journal of Psychology*, **81**, 361–380.

Burton A. M., Jenkins, R., Hancock, P. J. B. and White, D. (2005). Robust representations for face recognition: the power of averages. *Cognitive Psychology*, **51**, 256–284.

Burton A. M., White, D. and McNeill, A (2010). The Glasgow Face Matching Test. *Behaviour Research Methods*, **42**, 286–291.

Burton, A. M., Wilson, S., Cowan, M. and Bruce, V. (1999). Face recognition from poor quality video: evidence from security surveillance. *Psychological Science*, **10**, 243–248.

Caldara, R., Schyns, P., Mayer, E. *et al.* (2005). Does prosopagnosia take the eyes out of face representations? Evidence for a defect in representing diagnostic facial information following brain damage. *Journal of Cognitive Neuroscience*, **17 (10)**, 1652–1666.

Calderwood, L. and Burton, A. M. (2006). Children and adults recall the names of highly familiar faces faster than semantic information. *British Journal of Psychology*, **97**, 441–454.

Chambers, P. (2007). *Body 115: The Mystery of the Last Victim of the King's Cross Fire*. Chichester: Wiley.

Devlin, Lord (1976). *Report to the Secretary of State for the Home Department of the Departmental Committee on Evidence of Identification in Criminal Cases*. London: HMSO.

Duchaine, B., Jenkins, R., Germine, L. *et al.* (2009). Normal gaze discrimination and adaptation in seven prosopagnosics. *Neuropsychologia*, **47**, 2029–2036.

Duchaine, B. C. and Nakayama, K. (2006). Developmental prosopagnosia: a window to content-specific face processing. *Current Opinion in Neurobiology*, **16**, 166–173.

Ellis, H. D., Shepherd, J. W. and Davies, G. M. (1979). Identification of familiar and unfamiliar faces from internal and external features – some implications for theories of face recognition. *Perception*, **8**, 431–439.

Hanley, J. R. and Cowell, E. S. (1988) The effects of different types of retrieval cues on the recall of names of famous faces. *Memory and Cognition*, **16**, 545–555.

Hanley, J. R. and Damjanovic, L. (2009). It is more difficult to retrieve a familiar person's name and occupation from their voice than from their blurred face. *Memory*, **17**, 830–839.

Harmon, L. D. (1973). The recognition of faces. *Scientific American*, **229** (5), 70–82.

Harris, A. M. and Aguirre, G. K. (2007). Prosopagnosia. *Current Biology*, **17**, R7–R8.

Haslam, C., Kay, J., Hanley, J. R. and Lyons, F. (2004). Biographical knowledge: modality-specific or modality-neutral. *Cortex*, **40**, 451–466.

Hay, D. C., Young, A. W. and Ellis, A. W. (1991). Routes through the face recognition system. *Quarterly Journal of Experimental Psychology*, **43A**, 761–791.

Jenkins, R. and Burton, A. M. (2008). 100% accuracy in automatic face recognition. *Science*, **319**, 435.

Johnston, R. A. and Edmunds, A. J. (2009). Familiar and unfamiliar face recognition: A review. *Memory*, **17**, 577–596.

Johnston, R. A. and Bruce, V. (1990). Lost properties? Retrieval differences between name codes and semantic codes for famous people. *Psychological Research*, **52**, 62–67.

Lander, K., Bruce, V. and Hill, H. (2001). Evaluating the effectiveness of pixelation and blurring on masking the identity of familiar faces. *Applied Cognitive Psychology*, **15**, 101–116.

Leder, H. and Bruce, V. (2000). When inverted faces are recognised: the role of configural information in face recognition. *Quarterly Journal of Experimental Psychology*, **53**, 513–536.

McWeeney, K. H., Young, A. W., Hay, D. C. and Ellis, A. W. (1987). Putting names to faces. *British Journal of Psychology*, **78**, 143–149.

Megreya, A. M. and Bindemann, M. (2009). Revisiting the processing of internal and external features of unfamiliar faces: the headscarf effect. *Perception*, **38**, 1831–1848.

Megreya, A. M. and Burton, A. M. (2006). Unfamiliar faces are not faces: evidence from a matching task. *Memory and Cognition*, **34**, 865–876.

O'Donnell, C. and Bruce, V. (2001). Familiarisation with faces selectively enhances sensitivity to changes made to the eyes. *Perception*, **30**, 755–764.

Russell, R., Duchaine, B. and Nakayama, K. (2009). Super-recognisers: people with extraordinary face recognition ability. *Psychonomic Bulletin and Review*, **16**, 252–257.

Sinha, P. and Poggio, T. (1996). I think I know that face. . .. *Nature*, **384**, 404.

Thomson, D. M. (1986). Face recognition: More than a feeling of familiarity? In H. D. Ellis, M. A. Jeeves, F. Newcombe and A. Young (Eds.), *Aspects of Face Processing*, pp. 118–122. Dordrecht: Martinus Nijhoff.

Woodhead, M. M. and Baddeley, A. D. (1981). Individual differences and memory for faces, pictures and words. *Memory and Cognition*, **9**, 368–370.

Young, A. W., Hay, D. C. and Ellis, A. W. (1985*a*). The faces that launched a thousand slips: Everyday difficulties and errors in recognising people. *British Journal of Psychology*, **76**, 495–523.

Young, A. W., Hay, D. C., McWeeny, K. H *et al.* (1985*b*). Matching familiar and unfamiliar faces on internal and external features. *Perception*, **14**, 737–747.

Young, A. W., Hellawell, D. J. and Hay, D. C. (1987). Configural information in face perception. *Perception*, **16**, 747–759.

Young, A. W., Newcombe, F., De Haan, E. H. F., Small, M. and Hay, D. C. (1993). Face perception after brain injury: selective impairments affecting identity and expression. *Brain*, **116**, 941–959.

Notes

1. The individual differences in face matching appear to be highly robust. In unpublished undergraduate projects supervised by me in Newcastle, Lizzie Smith and Charles Okell in 2009 tested 100 participants on the short form of the GFMT and found a mean of 82.1 and standard deviation of 10.0. Claire Chandler and Joanne Sweeney tested a new sample of 77 participants in 2010 and found a mean of 81.3 and a standard deviation of 8.3.

2. Burton *et al.* (1990), elaborated the original model schematised here, and suggested familiarity results from a stage of pooling information across different input routes such as face and voice. However, some observations of differential access to semantics from faces, voices or names otherwise equated for familiarity are challenging for this model (see Haslam *et al.*, 2004; Hanley and Damjanovic, 2009).

Chapter 2

Unfamiliar face recognition

Peter J. B. Hancock

2.1 Introduction

'Unfamiliar face recognition' here means the task of recognising someone previously unfamiliar, usually seen only once before. The classic test methodology is to show a sequence of face images and later ask which of a larger set were shown before, but this chapter will also consider face matching, where two or more faces are presented simultaneously. This is one of the more common practical applications, for identity verification. Every time a photo-ID is presented, such as a passport, someone has to match the photograph to the person in front of them. The specific issue of identification parades, where an unfamiliar person has been seen and identification is later attempted from a line-up, is addressed in Chapter 9.

Mostly, face recognition appears effortless, which it usually is, for people that we know well. Where individuals claim to be bad at face recognition, they usually mean that they fail to recognise someone seen perhaps only once before, which is embarrassing, especially if they have recognised you. There is in fact wide variation in ability, from those unable to recognise even highly familiar people, to a few who can recognise someone seen only once, some years earlier (Russell *et al*., 2009).

A word of caution, when reading a published memory study: check whether the images used at study are different from those used at test. Early studies often used the same pictures and the practice continues surprisingly often. In the limit, there may be no face recognition involved, since an image may be remembered by some photographic blemish even if care has been taken to remove cues such as background and clothing. The data below on the difficulty of matching different pictures of the same person, without any memory load, indicates the effect of a change of image.

2.2 Matching studies

One of the clearest demonstrations of the inadequacy of unfamiliar face matching comes from a field study by Kemp *et al.* (1997). It is worth looking at this in some detail, as the results cast serious doubt on the utility of photographic identity cards. Their study was conducted in a supermarket, using real credit cards with photographs less than 6 weeks old. Shoppers presented cashiers, who knew they were being observed, with a card, which carried an incorrect photograph on half the occasions. The card was presented in an opaque wallet, such that the shopper was unaware of whether the card had a correct photograph. There were four different cards for each shopper. One carried a photograph of the shopper from the time of the study, one carried a photograph of them looking different, such as a change of hairstyle, or removal of spectacles, one was a 'similar' foil and the fourth was a dissimilar foil. The similar foils were selected by eye; from a set of 75 photographs for each sex, 25 similar-looking pairs were chosen. One would be the shopper, the other their similar foil. The remaining 25 images were allocated to the shoppers on the grounds that they looked dissimilar.

The experiment took place after normal closing in a supermarket. Shoppers presented items for purchase as usual and paid with one of the cards. The cashier checked the signature and the photograph, and called a supervisor if they wished to reject either. Overall, they made the correct decision on 67.4% of occasions, i.e. one third of the transactions were incorrect. Nearly 7% of the correct photographs, with the shoppers looking as they did at the time, were rejected, while almost 64% of the similar foil cards and 34% of the unmatched foils were accepted. To reiterate this point: on a third of occasions when shoppers presented a card bearing a photograph selected to look unlike them, cashiers who

Craniofacial Identification, ed. C. Wilkinson and C. Rynn. Published by Cambridge University Press.

knew they were under test, though unaware of exactly how, failed to reject them.

Subsequently, Bruce and her colleagues (1999) tested matching performance under more controlled laboratory conditions. They used a still image from a VHS-quality video as the target, to be matched to 10 photographs in a line-up. The images were all of young male police officers, with videos and photographs taken in the same sitting, there were therefore no changes in appearance or lighting. The complete set of 120 photographs was rated for similarity and the 10 most similar to each target used for the line-up. Half of the time, one of the images was replaced by the true target photograph. The initial experiment was intended as a control, with subsequent variations testing the effects of changing viewpoint, colour etc. However, performance in this baseline condition, matching neutral expression, full-face video stills to photographs, was only 70% correct. Subsequent studies have replicated and extended this basic finding, for example, even if the correct match is always present, and participants are told this, accuracy is only 79% (Bruce *et al.*, 2001). When the task is simplified to a single pair of faces, which either match or do not, performance is still only 78% on match trials and 84% on mismatches (Megreya and Burton, 2006). Interestingly, a test in our laboratory of a current automatic face recognition system got just less than 70% correct on the same test set, rather to the surprise of its developers. However, it is not something peculiar to this image set: Burton *et al.* (2010) developed a new matching task, again using pictures taken at the same time, from the same distance, just with different cameras, and found an overall matching accuracy of 90%. This may sound good, but remember, there is no memory load, and pictures were taken at the same time from the same viewpoint. Translate a 10% error rate into the process of checking passports to understand the consequences.

Lee *et al.* (2006) tested the effects of time pressure and an additional task on face matching. The faces used were the FERET set (Phillips *et al.*, 1998), comprising multiple monochrome images taken from different views and on different occasions. The task was simply to decide whether two images portrayed the same person, in either 6 or 15 seconds. Half the participants had an additional multiple choice task, requiring reference to biographical information about the person, before the decision task could be made. Under the easiest condition (long time and no task) participants averaged 89% correct rejections, falling to 65% under stress. However, the false rejection rate was high, at around 60%. In other words, at least 10% of false matches were accepted, despite rejecting more than half of those who did actually match.

It is apparent that people are really rather bad at matching unfamiliar face images. There are two bits of good news. The first, that some people seem naturally much better at it than others, is considered in the next section. The second is that experts are better than the lay public. Wilkinson and Evans (2009) tested the ability to match typically fuzzy CCTV images to a photographic line-up of either five or six faces. Lay participants averaged 61% of hits and 42% correct rejections, while two facial image analysts averaged 92% hits and 75% correct rejections. This was when the CCTV view showed the whole head; obscuring the hair with a cap caused the general public to do quite a lot worse, from an average across conditions of 53% to 37%, while the two experts only declined from 85% to 81%. Rather worryingly, in the hat condition, false acceptances became the most common result from the public, suggesting that jurors might be more likely to convict someone who is innocent than reject someone who is guilty based on CCTV evidence. The authors therefore argue for the merit of providing expert analysis in court to assist a jury.

2.3 Individual differences

The rather low average performance in the image-matching studies above conceals large individual variation in performance. Most of the work to date has studied prosopagnosia, where individuals may be unable to recognise even highly familiar people such as members of their family. There is an interesting double dissociation in neurological conditions: some prosopagnosic individuals demonstrate ability to match unfamiliar faces while having problems recognising familiar faces; other conditions manifest good familiar face recognition, but particularly poor matching of unfamiliar faces (Malone *et al.*, 1982). This adds weight to the notion that familiar and unfamiliar face processing differ in some fundamental way.

On the other hand, the boundary between familiar and unfamiliar faces is rather variable. Russell *et al.* (2009) report on four individuals with exceptional abilities at face recognition. These people reported being able to recognise someone seen only once before, perhaps in college or at a concert several months or even years earlier. In fact, they had learned to pretend not to recognise people, since those thus recognised found it unsettling. In objective tests, these four indeed performed far above the normal range, two or three standard deviations better than control participants. 'Familiar' is therefore a relative concept; it depends how good your memory is. The authors suggest that face-matching tests might usefully be given to those who need to do it professionally.

Megreya and Burton (2006) sought to identify whether individual differences in face matching correlated with other perceptual abilities. They found some significant correlations between face-matching accuracy and, most strongly, a measure of visual short-term memory and speed of matching line-drawn pictures. Matching ability was also strongly correlated with ability to match the faces when inverted. The implication is that unfamiliar faces are in fact being treated as pictures and that it makes little difference which way up they are. Megreya and Burton did find a correlation with performance on matching familiar faces, but crucially, only when the familiar faces were inverted. When upright, familiar face matching is a very different task from unfamiliar matching. The task becomes one of identification – who is the target person; is that person in the line-up? This tactic isn't possible for unfamiliar faces, so viewers have to rely on visual similarities. Their results lead Megreya and Burton to conclude, as the title of their paper, that 'unfamiliar faces are not faces'.

2.4 Mechanisms of face recognition

So, how are people processing unfamiliar faces, and how might that differ from familiar ones? Ellis *et al.* (1979) found that, while recognition of familiar faces was best when the inner parts of the face were presented, unfamiliar faces were as well recognised by their external features, such as hair and ears. For recognising a familiar face, hair is not generally very helpful: amongst White people, hair typically varies rather a lot and people may change their style from one meeting to another, therefore hair is not diagnostic of identity over the longer term. Hair varies much less between individuals of other races (e.g. most East Asian people naturally all have rather straight, black hair), and will not distinguish unfamiliar individuals as effectively. For unfamiliar faces, the variability of White people's hair is an asset, as it offers a distinctive feature to latch onto, while the actual face is learnt. Bruce *et al.* (1999) found that while matching on unfamiliar faces did deteriorate when only external features were presented (from 84% to 73%) it was much worse when only internal features were present (49%). Brooks and Kemp (2007) tested sensitivity to the movement of the main facial features for both familiar and unfamiliar faces. Changes to eye and nose position were significantly more detectable if the faces were familiar, but there was no difference for the mouth or the ears; in fact participants were essentially unable to detect quite large shifts in ear location even in familiar faces. On the whole, external features are much more important for discrimination of unfamiliar individuals than for familiar or cross-race identification, and while we are extremely sensitive to any variation within the internal features of familiar faces, the less diagnostic external information is not particularly well utilised.

The problem with recognising faces is that, to a first approximation, they all look alike; pairs of eyes and brows above a central nose and mouth. In detail they differ in a variety of ways: the individual features vary in shape and appearance; their relative positions vary; and there are overall differences in head shape and in skin colour and texture. The generally accepted view is that we can recognise familiar faces both by the individual features and by their positioning, the latter generally referred to as the configuration of the face. For example, Schwaninger *et al.* (2004) tested recognition of faces which were either blurred, cut up into their constituent features, or both. Blurred faces can still be recognised by the configuration of the face; disconnected features can also be recognised, but if these are blurred, both the configuration and the features are unrecognisable. The configural route appears to be a hallmark of processing of familiar faces. Initially it was reported that children tend to use a more feature-based method of identification even for familiar faces (e.g. Campbell *et al.*, 1999), but more recently evidence has been found that although face processing is of a poorer standard in general, even very young children use configural processing (Bonner and Burton, 2004; Crookes and McKone, 2009), indicating that differences in ability may owe more to the development of skill rather than different forms of processing per se.

Inverting a face appears to remove or at least reduce our ability to see the configuration of a face. This is most famously demonstrated in the Thatcher illusion (Thompson, 1980), where the eyes and mouth of a face are cut out and inverted. Seen the correct way up, the face is grotesque, but when inverted, the alteration is barely visible. A number of studies have indicated that changes to configuration are hard to detect when a face is inverted. Searcy and Bartlett (1996) compared participants' ability to detect whether two faces were the same or different, given either changes to features (such as blacking out teeth) or to configuration (such as moving the mouth down). When the faces were inverted, configural changes were much less detectable, with only around half being identified within the three seconds allowed. By contrast, featural changes remained highly detectable even when inverted, at around 94%. Leder and Bruce (2000) tested some more subtle variations of features and the relationships between them and again found that it was specifically the relationship information that suffered when the faces were inverted. Goffaux and Rossion (2007) reported that inversion particularly affects detection of changes in vertical relationships, such as the distance between nose and mouth, rather than horizontal ones, such as eye width. This is consistent with Dakin and Watt's proposition that vertical spacings in a face carry much of the information about identity (Dakin and Watt, 2009).

There are clear indications that we cannot help seeing a face as a whole, and it can therefore be hard to detect if elements are altered, for example, a slimmer nose might be detected as the eyes being further apart. This holistic processing of faces was demonstrated by Young *et al.* (1987). Using celebrity photographs halved horizontally below the eyes, they demonstrated that when the top half of one is aligned with the bottom half of another, it becomes difficult to identify the constituent parts. Identification of the parts was easier when the composite images were inverted, or if the parts were misaligned. The authors concluded that when the face halves were joined, the perception was of a novel face image that affected the appearance and recognition of each part. Although Young and his associates employed famous faces, the composite effect is also apparent with unfamiliar faces (Hole, 1994; Rossion and Boremanse, 2008).

Some studies have challenged the notion that inversion differentially affects perception of the configuration of a face. Riesenhuber *et al.* (2004) used a 3D morphing tool to generate faces that differed in positioning of features, or in the features themselves. Using a sequential same/different task, they found a similar inversion effect for both types of face change. Also, Sekuler *et al.* (2004) used a response classification task to isolate the regions of a face used to decide which of two faces had been displayed, when obscured by the addition of noise. They found that very similar regions of the face (the eyes) were used to make the decision whether the face was upright or inverted, leading them to conclude that inversion causes only quantitative and not qualitative differences. However, Rossion (2008) argues convincingly that Sekuler and colleagues' classification image technique is incapable of demonstrating a qualitative difference between upright and inverted faces. The debate as to exactly how faces are represented and what is lost when they are inverted is a very live one; see for example Riesenhuber and Wolff (2009), Yovel (2009) and the reply by Rossion (2009). My take on it is this: we do automatically integrate a face into a whole, but we don't explicitly represent the relationships between features, except as a means of recognising known faces. An unfamiliar face can therefore be represented either at a rather descriptive level, or as a pattern of activation across neurons that respond to known faces (e.g. Jiang *et al.*, 2006). Unless an unknown face happens to be rather similar to one we know, this pattern doesn't actually mean anything useful, so we are left with the rather featural, descriptive level representation. With experience, we build a more complete representation of someone's face, but a brief view is not enough for most of us.

2.5 Hemispheric specialisation

It is generally held that face recognition is predominantly performed in the right hemisphere of the brain (Haxby *et al.*, 2000). There is a specific region known as the fusiform face area (FFA) that responds strongly to face stimuli. Arguments persist as to whether this is specific to faces, or a more general area of expertise for within-object identification used, for example, to distinguish between cars or dogs (e.g. Gauthier *et al.*, 2000). A variety of lines of evidence, reviewed by Bourne *et al.* (2009) suggests that the left hemisphere may be specialised for featural processing, while the right hemisphere handles configural information. Bourne and colleagues found that blurred faces produced a repetition priming effect if presented to the right hemisphere (left visual field) while cut-up faces, similar to those used by Schwaninger *et al.* (2004), produced priming in the left hemisphere. Since featural processing is associated with unfamiliar face processing and configural with familiar faces, this would suggest a tendency for the lateral separation of the two forms of processing, with unfamiliar faces being more a left-hemisphere activity. Consistent with this, Gobbini *et al.* (2004) found stronger activation in the left fusiform area for unfamiliar faces, compared with famous ones. It does appear to be the case that recognition of familiar faces and matching of unfamiliar faces dissociate: e.g. Malone *et al.* (1982) report two cases where one function appeared in the normal range while the other was impaired.

2.6 Lighting and colour

The appearance of a face is determined by its shape, by its surface reflectance and by the lighting on it. By far the majority of face perception studies have been done with photographic or other two-dimensional images. We are able to infer some shape from the shading information on such images, with the assumption that lighting is coming from above. The strength of this assumption is given by the hollow face illusion (Gregory, 1970) where a face mask that is actually inverted in depth is perceived as projecting towards the viewer. Photographic negation of a face image makes recognition very difficult (Galper, 1970), but if they are lit from below, then negation has less of an effect (Johnston *et al.*, 1992). Since negation inverts the shading due to shape, this result suggests that at least some of the effects of negation come from disruption of the apparent shape of the face. However, Russell *et al.* (2006) directly compared discrimination ability for faces that differed in shape, in surface appearance (pigmentation) or both. These faces were generated using a 3D face model, but presentation to participants was as a 2D image. They found that photographic negation had little effect on matching performance for faces that differed only in shape, but produced a large decrement, from about 78% down to about 64%, when faces differed in pigmentation. Performance for faces that differed in both was better overall, but still affected by negation. Their conclusion was that negation disrupts our ability to use such pigmentation cues to identity. While we would never see a face in photographic negative under normal circumstances, we might well see one with an unusual direction of lighting. Because our perceptual system assumes lighting from above, a bottom-lit face looks very strange and it can be hard to recognise even very familiar faces. Matching of unfamiliar faces is also significantly affected: using 3D face surfaces, which could be lit artificially, Hill and Bruce (1996) showed that matching was impaired when faces were bottom lit; and was even worse when one was top and the other bottom lit.

Pigmentation cues should not be confused with colour information: under certain lighting conditions, such as sodium street lighting, which is essentially monochromatic, or lighting too dim for our cones to be active, there will be no colour information in a seen face. Naively it might be expected that colour would be better for identification than monochrome but in tests the effect is usually either non-existent, or reversed. Thus incorrect colour can hinder an accurate match; Bruce *et al.* (1999) tested line-ups in both monochrome and colour and found that hit rates decreased from an average of 67% to 59%. Consistent with this, Russell *et al.* (2007) found greater sensitivity to changes in pigmentation when faces were presented in colour than in monochrome; it becomes easier to reject a wrong match. The problem is that two images of the same face may differ a great deal in apparent colouration. ‘Accurate’ colour requires complex and careful calibration and therefore results differ between cameras and especially with lighting; simply taking pictures of people with different cameras appears to be enough to cause difficulties with matching (Burton *et al.*, 2010). While there is an advantage for CCTV, in that it enables identification of the colour of clothing, it therefore has little advantage for faces. In fact, the actual colour of a face has remarkably little effect for recognition of familiar faces. Kemp *et al.* (1996) compared inversion of the hue and luminance components of an image and found that while negating the luminance had a large effect, inverting the hue, so that red turns green and vice versa, had little effect on recognition. However, recognition of pictures of unfamiliar faces was affected by hue changes, indicating again that unfamiliar faces are stored more as a picture.

2.7 Distance: resolution and spatial scale

A distant face appears smaller, obviously enough, with the consequence that only relatively coarse details are visible. This is the logical equivalent of blurring, so it reduces visibility of feature-level information while maintaining an impression of the configuration. The consequence is that it is possible to recognise highly familiar people at a distance, but unfamiliar face processing suffers. Much work has been aimed at identifying the spatial scales required for face recognition (reviewed in Ruiz-Soler and Beltran, 2006). Early studies produced rather confusing results, because of a mixture of testing methods, including both familiar and unfamiliar faces, but suggested that something like 8–16 cycles per face (cpf) is optimal (Costen *et al.*, 1996). Goffaux *et al.* (2005) constructed faces that differed in either configuration or in features and performed a matching task, where one of two faces was identical to a third, displayed simultaneously. The faces were either intact, or filtered to show only high (> 32 cpf) or low (< 8 cpf) frequency information. As expected, error rates were lowest with intact faces, but high spatial frequencies alone were almost as good for identifying featural changes. Filtering had little effect on error rates for the configural changes (which at about 40% were not much better than chance), but reaction times were significantly faster for low-frequency filtered faces than for the other two. Thus although high frequencies can support configural information, it is processed more efficiently in lower frequencies.

2.8 Distance: perspective changes

The distance from which a photograph is taken, or from which a person is seen, will affect the appearance of their face. Close to, the nose looks relatively big, and ears tend to disappear around the side of the head. This affects memory: Liu and Chaudhuri (2003) found that while people recognised around 90% of faces seen at the same distance at study and test, this fell to between 40–50% if there was a big change in apparent distance. In common with most face research to date, this study used flat images, while real faces are 3D objects. Adding stereo information (i.e. full 3D depth) helps to overcome some of the effects of varying distance. Liu and Ward (2006) found an accuracy of about 87% for study and test at the same distance, dropping to 63% with a large change; adding stereo information increased this to 70%. Given that in some unfortunate circumstances, people may have seen a face very close to, there would appear to be a case for enabling subsequent recognition attempts to take place at the same apparent distance. It would therefore be useful to test whether a future replacement for the current UK VIPER system, which enables line-ups to be conducted by studying videos, might helpfully be a '4D' system (i.e. moving 3D), that allows viewing of the recorded faces from any apparent distance.

2.9 Distinctiveness

Things which are distinctive tend to stick in our memories. Some faces are rather 'typical', near to an overall average in appearance, while others may be distinctive in terms of individual features or overall shape. Light *et al.* (1979) confirmed that faces rated as typical were less well remembered, due to being more alike and therefore confusable. Wickham *et al.* (2000) showed that the correlation between distinctiveness and hit rate is stronger after a 5-week delay than for an immediate test. Distinctive faces are also less likely to be incorrectly identified; there is therefore a negative correlation between distinctiveness and false positives and a positive one with hit rate, and yet false positive and hits do not correlate (Bruce *et al.*, 1994; Hancock *et al.*, 1996). It is not the case that a face that is easy to remember will necessarily be easy to reject; Lewis and Johnston (1997) confirmed that this is because there are different underlying mechanisms at work – distinctive faces are remembered because they stand out, typical faces are falsely recognised because they tend to resemble someone you know. Since the latter is idiosyncratic to each observer, correlation with hit rates is low. There remain some faces that people tend to say yes to, whether or not they were seen in the study phase, while others tend to be rejected in both cases (Hancock *et al.*, 1996). Clearly, it is better to have the latter if intent on a life of crime.

Caricature emphasises the aspects of a face that are distinctive and well-known to cartoonists; their drawings provide a shorthand way to recognise a famous person. Computer-generated caricatures, which systematically exaggerate faces from a specified mean or average face, have been much studied, with some evidence that they allow better recognition of known faces. Deffenbacher *et al.* (2000) tested the effects of caricaturing unfamiliar faces both at study and test: at study, a caricatured face is simply more distinctive and, as expected, better recognised at test. However, for faces studied in their original form, caricature at test decreased hit rates, along with false alarm rates to those unstudied. They therefore found no evidence of caricaturing doing anything other than increasing the distinctiveness of an unknown face. Caricature has also shown beneficial distinctiveness effects in face-matching tasks; tests in our laboratory demonstrated that caricaturing all of the faces in a matching task can help with accuracy, especially reducing false matches. Overall accuracy on a hard subset of the Bruce *et al.* (1999) line-up set climbed from 60.5% to 69% (McIntyre *et al.*, submitted).

2.10 Viewpoint and expression

It is well established that changes in viewpoint affect unfamiliar face recognition and matching: Bruce (1982) found that hit rates for 24 unfamiliar faces fell from 90% when identical views (identical images, in fact) were shown at test, to 76% when the images varied either in viewpoint or expression, and to 61% when both viewpoint and expression changed. Viewpoint changed between frontal and three-quarter view, and expression between neutral and smiling. O'Toole *et al.* (1998) presented faces for study and test in either frontal, three-quarter or profile views. As expected, each type was best recognised when tested in the same orientation as it was studied, with the three-quarter view highest. Faces studied at three-quarter view were also relatively well recognised when tested at frontal or profile view. With no memory load, a change in orientation reduces matching performance; Bruce *et al.* (1999) found that hit rates dropped from 70% when their video still showed the same frontal view as their line-up images, to 60% when it showed a 30° view. Overall therefore, it seems that a three-quarter view affords the most information about a face and that recognition or matching performance tails off with increasing deviation from the studied view. Changes in expression have the same effect; by changing the image, recognition decreases. Chen and Liu (2009) tested whether learning a face from multiple views might help with generalisation to a new expression or vice versa. They found that training with multiple viewpoints helped with subsequent recognition, independent of any change of expression, but training with multiple expressions did not aid recognition.

2.11 Race

The difficulties of remembering or matching unfamiliar faces increase when they are of an unfamiliar race; the effect is robust and is termed the own race bias (ORB), or other race effect. Meissner and Brigham (2001) performed a meta-analysis of 91 studies from 39 papers and found that on average, participants were 1.4 times more likely to correctly recognise an own-race face, and 1.56 times more likely to incorrectly identify an other-race face. There are a variety of putative explanations for why other races should be harder to identify. The simplest, known as the contact hypothesis, is based on differential experience; Meissner and Brigham found supportive evidence for contact being a factor in the size of the effect, though this begs the question of what it is that changes with experience of another race. Learning about your own race apparently takes place in the first year of life: Kelly *et al.* (2007) tested infants' ability to discriminate faces from their own and other racial groups and found that while 3-month-olds could differentiate all races, by 9 months they were restricted to their own. There is, however, evidence of some flexibility. Sangrigoli *et al.* (2005) tested Korean-born adults who had been adopted into European families between the ages of 3 and 9, and found they showed the same pattern of ORB as native French controls, though the effect sizes on their matching task were quite small. De Heering *et al.* (2010) tested a larger group of Asian adoptees and found that they performed equally well with Asian and White faces, between 6 and 14 years after adoption. So while early life contact evidently has some effect, it is modifiable with further experience.

The way in which a face is classified also affects how well it is remembered. MacLin and Malpass (2001) produced faces that might be regarded as either Hispanic or Black, depending on the hairstyle that was added. This allowed them to test memorability of exactly the same faces in same-race and cross-race conditions; Hispanic participants remembered faces with Hispanic hairstyles better than those with Black hairstyles, principally because of higher false alarm rates with Black hairstyles. However, social grouping is sufficient to produce this effect. Bernstein *et al.* (2007) found that merely classing exactly the same faces as being either in- or out-group was enough to cause a bias. Faces were presented either on a red background, to signify the home university, or a green one, to signify a rival establishment. A control group were not told of any significance of the coloured backgrounds and their memory was unaffected by them, while the experimental group remembered significantly more of the supposed in-group.

It does not seem to be the case that participants simply attend more to their own race. Humphreys *et al.* (2005) found that both Caucasian and Asian participants detected changes to their own race of face faster in a change-blindness study, where the pictures shown contained two people of each race. They argued that, since this crossover effect did not occur when other body parts change, it was not caused by people attending more to the own-race faces. Yoriko Hirose and I decided to check this more directly, using the same paradigm with an eye-tracker to confirm exactly where participants were looking (Hirose and Hancock, 2007). We found that both races actually looked at White faces sooner, but that it simply took longer from first fixating the face that was changing to noticing the change when it was of the other race.

More recently, we have been studying possible ways to improve matching of other-race faces by using computer graphics to alter the appearance of the faces. This work has drawn on perceptual expertise explanations of ORB, that discrimination ability is a function of the faces encountered, and limited exposure to other-race faces means the appropriate range of variation to distinguish them is never learned (MacLin and Malpass, 2001). If you average together a set of African American faces and a set of White faces, you get an average for each race. This allows computation of the average difference, both in terms of face and feature shapes and in skin tone. Applying this difference to a given image gives a passable approximation of what that person would look like if they were of the other race. We reasoned that if you transform both the target and line-up faces from the 'other' race to a familiar one, the images would vary on learned own-race dimensions, which would significantly enhance discrimination and improve perceptual matching. Skin-tone transformation gave visually unacceptable results, but it was possible to transform the shape successfully. In contrast to our predictions, where all images were presented together in a simultaneous array, the transformation impaired face matching. The parallel line-up encourages featural comparisons and a relative judgement strategy, leading us to speculate that sequential presentation might encourage a more holistic style of processing within an absolute judgement task. Transformation of other race faces towards own race variation was successful within the sequential matching procedure, providing support for theories of perceptual expertise, and improving overall accuracy from 58% to 71% (McIntyre *et al.*, in preparation).

2.12 Age

Just as we are better at recognising others of our own race, there is evidence that we are better at recognising those of our own age. Wright and Stroud (2002) tested young (18–25) and older (35–55) adults for identification of a culprit in four simulated crime videos. Two of the adults were 'young' and two 'older' and participants were better at identifying those of their own age group, though the difference was significant only for younger participants. Anastasi and Rhodes (2005) tested children aged 5–8 and adults aged 55–89 for memory of 32 faces in four age groups (5–8, 18–25, 35–45 and 55–75 years). They were tested on a set of 64 faces, half of which had previously been studied. The results showed a full crossover interaction, with both children and the adults best remembering people their own age. Anastasi and Rhodes argued against a contact explanation for this effect, preferring Sporer's (2001) in-group/out-group model. However, Harrison and Hole (2009) provide evidence in favour of a contact explanation. They compared a group of trainee teachers with a group of similar-aged controls who had little contact with children. At study, they were shown 32 photographs, half of children aged 8–11, the other half of adults of a similar age to the participants. At test, they were shown different photographs of the same people, plus 32 others. The trainee teachers showed no difference in recognition score between the two age groups (though with a trend for better performance on the children, a 85% hit rate, cf. 82%), while controls were significantly poorer on the child images (78%, cf. 85%). This may seem surprising, since the control adults must once have been children, with, presumably, plenty of contact with other children. The authors suggest that either a contact hypothesis is correct, and our face-representation systems are updated to reflect the faces that we typically see, or that teachers are more motivated to differentiate children's faces.

2.13 Motion

People move in idiosyncratic ways, and so familiar faces are associated with particular mannerisms and styles of movement, which aids their identification. Tests on the effects of motion for unfamiliar faces have produced mixed results. Pike *et al.* (1997) presented faces for study either as videos, with the head rotating from one profile to the other; as five still images from the same sequence; or as single still images. Tested on an unseen frontal image with different lighting, the videos produced the best memory, followed by the five images. It is unclear from this whether the advantage is motion, or simply more views of the face. In a series of experiments, Christie and Bruce (1998) equated the number of frames shown in each condition, with the only significant difference being an advantage for studying still frames. O'Toole *et al.* (2002) speculate that one reason might be that motion carries with it other information, such as expressions, speech production and eye-gaze, which may distract from learning identity. However, motion also draws attention, and Lander and Bruce (2003) suggest this is why they did find an advantage for learning from moving images exceeding that for multiple stills. Overall, it seems that seeing a moving face allows better learning through affording more views, and through holding attention better, but that learning someone's characteristic movements to the point where they aid identification comes only with greater familiarity.

2.14 Delay

Most tests of face recognition are laboratory-based and involve minimal delay, perhaps a few minutes during which some filler task is given. Longer delays are procedurally difficult, and while two early studies showed rather little effect, most find that longer delays reduce recognition, as might be expected. Krouse (1981) found that recognition of the same picture declined from 75% with no delay to 59% after a week, where chance performance would be 25%. When the image changed, there was less of a difference, with performance dropping from 57% to 53%. Wickham *et al.* (2000) tested recognition after a 5-week delay. Hit rate dropped from 67% with no delay to 56%, while false positives climbed from 27.5% to 38%, where chance performance is 50%. Thus 62% of responses were correct after the delay. Memon *et al.* (2003) tested relatively young (16–33) and old (60–82) participants on a line-up task after a delay of a few minutes and after one week. Overall, their young group actually performed a bit better after the delay, partly through being more conservative and offering fewer false positives; the older group deteriorated significantly, from 66% to 35% correct. For younger adults at least, deterioration in performance is not substantial, and certainly remains above chance even a month after exposure.

2.15 Verbal descriptions

Activity before the recognition task has a perhaps surprising influence on accuracy; Schooler and Engstler-Schooler (1990) discovered that giving a verbal description of the previously seen face significantly impaired performance on a subsequent line-up task. Accuracy in picking from 8 pictures fell from 64% to 38%. There is more on this 'verbal overshadowing' effect in Chapter 9, but there are other activities that may affect performance. One is viewing Navon letters; large letters made up from different small letters (e.g. a capital F created by suitably arranging many small capital Es). Macrae and Lewis (2002) showed a video, then gave participants a filler task of reading, or naming Navon letters, before attempting to identify the 'culprit' from a line-up. They found that performance increased from 30%, when participants were asked to read the small letters, to 83% when asked to read the large ones. Subsequent work has suggested the main effect is a reduction in performance when reading the small letters (Perfect, 2003). One explanation is that, as with verbal descriptions, people reading Navon letters are asked to concentrate on small details, rather than the bigger picture that is good for recognising a face. However, Lewis (2006) tested some other filler tasks, and found that while reading, doing a Sudoku puzzle or a simple crossword had little effect on recognition, doing a cryptic crossword reduced performance. He speculated that cryptic crosswords, reading small Navon letters and verbal descriptions all involve suppression of an obvious answer as a common factor, though why this might affect recognition is unclear.

2.16 Tests

Various tests of unfamiliar face processing have been developed. Two commonly used early tests were the Warrington Recognition Memory for Faces (RMF) and the Benton Facial Recognition Test (BFRT). However, because both show whole faces, it is possible to perform well just by looking at things like the hairline. In fact, Duchaine and Weidenfeld (2003) showed that it is possible to score in the normal range, even when the internal face features have been obscured. Duchaine and Nakayama (2006) therefore developed the Cambridge Face Memory Test and demonstrated that it separates those with prospagnosia more effectively than the other two. The Cambridge test contains six target faces and 46 distracters. A larger set, containing over 200 faces and more suitable for general experiments with unfamiliar faces, is available from Glasgow (Burton *et al.*, 2010).

2.17 Adaptation

Stare at an average male face for a few seconds, and a briefly glimpsed androgynous face will look female. Study a female face, and the same androgynous face will look male. This is the phenomenon of adaptation, first studied in faces by Webster and MacLin (1999). The general effect has been long known in various guises, for example the waterfall illusion, whereby if you look at a waterfall for a few seconds and then at the bank, the patch corresponding to the retinal location of the waterfall will appear to drift upwards. Another familiar example comes from travelling, especially backwards, in a train; when it stops, the platform may appear to be moving. The directly retinotopic nature of this type of adaptation indicates that its origin is at a relatively low level in the visual stream. Face adaptation is almost invariant to a two-fold change in size between adapting and test stimuli (Rhodes *et al.*, 2004) but is still retinotopic (Afraz and Cavanagh, 2009), suggesting that it is operating at some intermediate level. On the other hand, it is possible to adapt individual identities (Carbon *et al.*, 2007), implying the location of action is relatively high, presumably in the fusiform face area. There has been an explosion of research in face adaptation in the last decade, since it promises to give insights into how we represent faces. The general mechanism appears to relate to the way we tune our perception to the statistics of the world. An evident example is the variations in aspect ratio caused by the switch from 4:3 to widescreen television. Initially the incorrect format is very obvious, but after a while it becomes much less worrisome.

An intriguing aspect of face adaptation is the ability to adapt different classes of faces in opposite directions simultaneously. Little *et al.* (2005) showed that it is possible to adapt female faces in one direction (e.g. making the eyes appear too far apart), while simultaneously making male faces appear to have their eyes too close together. This appears to argue for separate populations of neurons for male and female faces. This may be possible, but then there would need to be separate populations for upright and inverted faces (Rhodes *et al.*, 2004) and for different races (Ng *et al.*, 2006). However, it seems that the categories need to be naturally occurring, such as male and female. Bestelmeyer *et al.* (2008) used computer graphic techniques to produce faces that were all perceived as female, but differed along the male–female continuum as much as average male and female faces. They could not induce opposite adaptation in the two different groups of female faces, implying that the adapting neurons are sensitive to sex differences, not just differences in appearance. Adaptation studies therefore have the potential to help us understand how faces are represented in our brains, which should lead to better models of our face recognition, and possibly better automated systems.

Since it is unlikely that an unfamiliar face would be encountered while in some adapted state, other than in a psychology laboratory, the real-world effects of adaptation on unfamiliar recognition and matching should be limited. However, there may well be implications in the area of facial composite production (see Chapters 3 and 4). Here, the task is to try and recreate a likeness of a person seen during the course of some crime. Whatever system is used, the face image should become steadily more like that held in memory. However, the witness should adapt to the face being shown. Just as staring at a male face causes other faces to look less masculine, so staring at a given identity will tend to make that identity weaker. This may lead the witness to produce a caricature of the person, over-emphasising the appearance in a bid to counter adaptation. We are currently testing this prediction and possible ways to address it in our laboratory.

2.18 Summary

In contrast to our excellent abilities with familiar faces, we are in general really rather poor at processing unfamiliar ones. A simple change of camera is enough to make matching two pictures of someone unreliable, while attempting to match someone to their photograph is highly error-prone. We tend to rely on superficial properties, such as the hair, or surface marks such as moles. Changes of lighting, viewpoint or expression all affect recognition performance, with abilities becoming even worse if the face is of a different race, age and perhaps even sex from ourselves. However, relatively modest training experience of a face, such as watching a video of the person for a minute or two, is enough to start the transition towards a more familiar style of face processing and improve performance. In the limit, is it the case that 'unfamiliar faces are not faces' (Megreya and Burton, 2006)? That seems a little strong: unfamiliar faces are subject to holistic processing (Hole 1994), but it seems that we are truly expert, not in face processing generally, but in familiar face processing.

References

Afraz, A. and Cavanagh, P. (2009). The gender-specific face aftereffect is based in retinotopic not spatiotopic coordinates across several natural image transformations. *Journal of Vision*, **9**(10), 1–17.

Anastasi, J. S. and Rhodes, M. G. (2005). An own-age bias in face recognition for children and older adults. *Psychonomic Bulletin and Review*, **12**(6), 10–43.

Bernstein, M. J., Young, S. G. and Hugenberg, K. (2007). The cross-category effect: Mere social categorization is sufficient to elicit an own-group bias in face recognition. *Psychological Science*, **18**(8), 706–712.

Bestelmeyer, P. E. G., Jones, B. C., DeBruine, L. M. *et al.* (2008). Sex-contingent face aftereffects depend on perceptual category rather than structural encoding. *Cognition*, **107**(1), 353–365.

Bonner, L. and Burton, A. M. (2004). 7–11-year-old children show an advantage for matching and recognizing the internal features of familiar faces: Evidence against a developmental shift. *Quarterly Journal of Experimental Psychology A: Human Experimental Psychology*, **57**(6), 1019–1029.

Bourne, V. J., Vladeanu, M. and Hole, G. J. (2009). Lateralised repetition priming for featurally and configurally manipulated familiar faces: Evidence for differentially lateralised processing mechanisms. *Laterality: Asymmetries of Body, Brain and Cognition*, **14**(3), 287–299.

Brooks, K. R. and Kemp, R. I. (2007). Sensitivity to feature displacement in familiar and unfamiliar faces: Beyond the internal/external feature distinction. *Perception*, **36**(11), 1646–1659.

Bruce, V. (1982). Changing faces: visual and non-visual coding processes in face recognition. *British Journal of Psychology*, **73**(1), 105–116.

Bruce, V., Burton, A. M. and Dench, N. (1994). What's distinctive about a distinctive face? *Quarterly Journal of Experimental Psychology A: Human Experimental Psychology*, **47**(1), 119–141.

Bruce, V., Henderson, Z., Greenwood, K. *et al.* (1999). Verification of face identities from images captured on video. *Journal of Experimental Psychology: Applied*, **5**(4), 339–360.

Bruce, V., Henderson, Z., Newman, C. and Burton, A. M. (2001). Matching identities of familiar and unfamiliar faces caught on CCTV images. *Journal of Experimental Psychology: Applied*, **7**(3), 207–218.

Burton, A. M., White, D. and McNeill, A. (2010). The Glasgow Face Matching Test. *Behavior Research Methods*, **42**, 286–291.

Campbell, R., Coleman, M., Walker, J. *et al.* (1999). When does the inner-face advantage in familiar face recognition arise and why? *Visual-Cognition*, **6**(2), 197–216.

Carbon, C., Strobach, T., Langton, S. *et al.* (2007). Adaptation effects of highly familiar faces: immediate and long lasting. *Memory and Cognition*, **35**(8), 1966–1976.

Chen, W. and Liu, C. H. (2009). Transfer between pose and expression training in face recognition. *Vision Research*, **49**(3), 368–373.

Christie, F. and Bruce, V. (1998). The role of dynamic information in the recognition of unfamiliar faces. *Memory and Cognition*, **26**(4), 780–790.

Costen, N. P., Parker, D. M. and Craw, I. (1996). Effects of high-pass and low-pass spatial filtering on face identification. *Perception and Psychophysics*, **58**(4), 602–612.

Crookes, K. and McKone, E. (2009). Early maturity of face recognition: No childhood development of holistic processing, novel face encoding, or face-space. *Cognition*, **111**(2), 219–247.

Dakin, S. C. and Watt, R. J. (2009). Biological "bar codes" in human faces. *Journal of Vision*, **9**(4), 1–10.

Deffenbacher, K. A., Johanson, J., Vetter, T. and O'Toole, A. J. (2000). The face typicality-recognizability relationship: encoding or retrieval locus? *Memory and Cognition*, **28**(7), 1173–1182.

De Heering, A., de Liedekerke, C., Deboni, M. and Rossion, B. (2010). The role of experience during childhood in shaping the other-race effect. *Developmental Science*, **13**(1), 181–187.

Duchaine, B. C. and Nakayama, K. (2006). The Cambridge Face Memory Test: results for neurologically intact individuals and an investigation of its validity using inverted face stimuli and prosopagnosic participants. *Neuropsychologia*, **44**(4), 576–585.

Duchaine, B. C. and Weidenfeld, A. (2003). An evaluation of two commonly used tests of unfamiliar face recognition. *Neuropsychologia*, **41**(6), 713–720.

Ellis, H. D., Shepherd, J. W. and Davies, G. M. (1979). Identification of familiar and unfamiliar faces from internal and external features: some implications for theories of face recognition. *Perception*, **8**(4), 431–439.

Galper, R. E. (1970). Recognition of faces in photographic negative. *Psychonomic Science*, **19**(4), 207–208.

Gauthier, I., Skudlarski, P., Gore, J. and Anderson, A. (2000). Expertise for cars and birds recruits brain areas involved in face recognition. *Nature Neuroscience*, **3**(2), 191–197.

Gobbini, M., Leibenluft, E., Santiago, N. and Haxby, J. V. (2004). Social and emotional attachment in the neural representation of faces. *Neuroimage*, **22**(4), 1628–1635.

Goffaux, V., Hault, B., Michel, C., Vuong, Q. C. and Rossion, B. (2005). The respective role of low and high spatial frequencies in supporting configural and featural processing of faces. *Perception*, **34**(1), 77–86.

Goffaux, V. and Rossion, B. (2007). Face inversion disproportionately impairs the perception of vertical but not horizontal relations between features. *Journal of Experimental Psychology: Human Perception and Performance*, **33**(4), 995–1002.

Gregory, R. L. (1970). *The Intelligent Eye*. London: Weidenfeld and Nicolson.

Hancock, P. J. B., Burton, A. M. and Bruce, V. (1996). Face processing: human perception and principal components analysis. *Memory and Cognition*, **24**(1), 26–40.

Harrison, V. and Hole, G. J. (2009). Evidence for a contact-based explanation of the own-age bias in face recognition. *Psychonomic Bulletin and Review*, **16**(2), 264–269.

Haxby, J. V., Hoffman, E. and Gobbini, M. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences*, **4**(6), 223–233.

Hill, H. and Bruce, V. (1996). Effects of lighting on the perception of facial surfaces. *Journal of Experimental Psychology: Human Perception and Performance*, **22**(4), 986–1004.

Hirose, Y. and Hancock, P. J. B. (2007). Equally attending but still not seeing: an eye-tracking study of change detection in own- and other-race faces. *Visual Cognition*, **15**(6), 647–660.

Hole, G. J. (1994). Configurational factors in the perception of unfamiliar faces. *Perception*, **23**(1), 65–74.

Humphreys, G., Hodsoll, J. and Campbell, C. (2005). Attending but not seeing: the 'other race' effect in face and person perception studied through change blindness. *Visual Cognition*, **12**(1), 249–262.

Jiang, X., Rosen, E., Zeffiro, T. *et al.* (2006). Evaluation of a shape-based model of human face discrimination using fMRI and behavioral techniques. *Neuron*, **50**(1), 159–172.

Johnston, A., Hill, H. and Carman, N. (1992). Recognising faces: effects of lighting direction, inversion, and brightness reversal. *Perception*, **21**(3), 365–375.

Kelly, D., Quinn, P., Slater, A. *et al.* (2007). The other-race effect develops during infancy – evidence of perceptual narrowing. *Psychological Science*, **18**(12), 1084–1089.

Kemp, R., Pike, G., White, P. and Musselman, A. (1996). Perception and recognition of normal and negative faces: the role of shape from shading and pigmentation cues. *Perception*, **25**(1), 37–52.

Kemp, R., Towell, N. and Pike, G. (1997). When seeing should not be believing: photographs, credit cards and fraud. *Applied Cognitive Psychology*, **11**(3), 211–222.

Krouse, F. L. (1981). Effects of pose, pose change, and delay on face recognition performance. *Journal of Applied Psychology*, **66**(5), 651–654.

Lander, K. and Bruce, V. (2003). The role of motion in learning new faces. *Visual Cognition*, **10**(8), 897–912.

Leder, H. and Bruce, V. (2000). When inverted faces are recognised: the role of configural information in face recognition. *Quarterly Journal of Experimental Psychology*, **53A**, 513–536.

Lee, M. D., Vast, R. L. and Butavicius, M. A. (2006). Face matching under time pressure and task demands. In *Proceedings of the 28th Annual Conference of the Cognitive Science Society*, pp. 1675–1680.

Lewis, M. B. (2006). Eye-witnesses should not do cryptic crosswords prior to identity parades. *Perception*, **35**(10), 1433–1436.

Lewis, M. B. and Johnston, R. A. (1997). Familiarity, target set and false positives in face recognition. *European Journal of Cognitive Psychology*, **9**(4), 437–459.

Light, L. L., Kayra Stuart, F. and Hollander, S. (1979). Recognition memory for typical and unusual faces. *Journal of Experimental Psychology: Human Learning and Memory*, **5**(3), 212–228.

Little, A. C., DeBruine, L. M. and Jones, B. C. (2005). Sex-contingent face after-effects suggest distinct neural populations code male and female faces. *Proceedings of the Royal Society B: Biological Sciences*, **272**(1578), 2283–2287.

Liu, C. H. and Chaudhuri, A. (2003). Face recognition with perspective transformation. *Vision Research*, **43**(23), 2393–2402.

Liu, C. H. and Ward, J. (2006). The use of 3D information in face recognition. *Vision Research*, **46**(6), 768–773.

MacLin, O. H. and Malpass, R. S. (2001). Racial categorization of faces: the ambiguous race face effect. *Psychology, Public Policy, and Law*, **7**(1), 98–118.

Macrae, C. N. and Lewis, H. L. (2002). Do I know you? Processing orientation and face recognition. *Psychological Science*, **13**(2), 194–196.

Malone, D. R., Morris, H. H., Kay, M. C. and Levin, H. S. (1982). Prosopagnosia: a double dissociation between the recognition of familiar and unfamiliar faces. *Journal of Neurology, Neurosurgery and Psychiatry*, **45**(9), 820–822.

McIntyre, A., Hancock, P. J. B. and Kittler, J. Reducing the own race bias in perceptual face matching. In preparation.

McIntyre, A., Hancock, P. J. B. and Kittler, J. Improving face matching with caricature. Submitted to *Applied Cognitive Psychology*.

Megreya, A. and Burton, A. M. (2006). Unfamiliar faces are not faces: evidence from a matching task. *Memory and Cognition*, **34**(4), 865–876.

Meissner, C. A. and Brigham, J. C. (2001). Thirty years of investigating the own-race bias in memory for faces: a meta-analytic review. *Psychology, Public Policy, and Law*, **7**(1), 3–35.

Memon, A., Bartlett, J., Rose, R. and Gray, C. (2003). The aging eyewitness: effects of age on face, delay, and source-memory ability. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, **58**(6), 338–345.

Ng, M., Ciaramitaro, V. M., Anstis, S., Boynton, G. M. and Fine, I. (2006). Selectivity for the configural cues that identify the gender, ethnicity, and identity of faces in human cortex. *Proceedings of the National Academy of Sciences*, **103**(51), 19552–19557.

O'Toole, A. J., Edelman, S. and Bulthoff, H. H. (1998). Stimulus-specific effects in face recognition over changes in viewpoint. *Vision Research*, **38**(15–16), 2351–2363.

O'Toole, A. J., Roark, D. A. and Abdi, H. (2002). Recognizing moving faces: a psychological and neural synthesis. *Trends in Cognitive Sciences*, **6**(6), 261–266.

Perfect, T. J. (2003). Local processing bias impairs lineup performance. *Psychological Reports*, **93**(2), 393–394.

Phillips, P., Wechsler, H., Huang, J. and Rauss, P. (1998). The Feret database and evaluation procedure for face-recognition algorithms. *Image and Vision Computing*, **16**(5), 295–306.

Pike, G. E., Kemp, R. I., Towell, N. A. and Phillips, K. C. (1997). Recognizing moving faces: the relative contribution of motion and perspective view information. *Visual Cognition*, **4**(4), 409–437.

Rhodes, G., Jeffery, L., Watson, T. *et al.* (2004). Orientation-contingent face aftereffects and implications for face-coding mechanisms. *Current Biology*, **14**(23), 2119–2123.

Riesenhuber, M., Jarudi, I., Gilad, S. and Sinha, P. (2004). Face processing in humans is compatible with a simple shape-based model of vision. *Proceedings of the Royal Society B: Biological Sciences*, **271**(Suppl. 6), S448–S450.

Riesenhuber, M. and Wolff, B. S. (2009). Task effects, performance levels, features, configurations, and holistic face processing: a reply to Rossion. *Acta Psychologica*, **132**(3), 286–292.

Rossion, B. (2008). Picture-plane inversion leads to qualitative changes of face perception. *Acta Psychologica*, **128**(2), 274–289.

Rossion, B. (2009). Distinguishing the cause and consequence of face inversion: the perceptual field hypothesis. *Acta Psychologica*, **132**(3), 300–312.

Rossion, B. and Boremanse, A. (2008). Nonlinear relationship between holistic processing of individual faces and picture-plane rotation: evidence from the face composite illusion. *Journal of Vision*, **8**(4), 113.

Ruiz-Soler, M. and Beltran, F. S. (2006). Face perception: an integrative review of the role of spatial frequencies. *Psychological Research/Psychologische Forschung*, **70**(4), 273–292.

Russell, R., Biederman, I., Nederhouser, M. and Sinha, P. (2007). The utility of surface reflectance for the recognition of upright and inverted faces. *Vision Research*, **47**(2), 157–165.

Russell, R., Duchaine, B. C. and Nakayama, K. (2009). Super-recognizers: people with extraordinary face recognition ability. *Psychonomic Bulletin and Review*, **16**(2), 252–257.

Russell, R., Sinha, P., Biederman, I. and Nederhouser, M. (2006). Is pigmentation important for face recognition? Evidence from contrast negation. *Perception*, **35**(6), 749–759.

Sangrigoli, S., Pallier, C., Argenti, A., Ventureyra, V. and de Schonen, S. (2005). Reversibility of the other-race effect in face recognition during childhood. *Psychological Science*, **16**(6), 440–444.

Schooler, J. W. and Engstler-Schooler, T. Y. (1990). Verbal overshadowing of visual memories: some things are better left unsaid. *Cognitive Psychology*, **22**(1), 36–71.

Schwaninger, A., Wallraven, C. and Bulthoff, H. (2004). Computational modeling of face recognition based on psychophysical experiments. *Swiss Journal of Psychology*, **63**(3), 207–215.

Searcy, J. H. and Bartlett, J. C. (1996). Inversion and processing of component and spatial-relational information in faces. *Journal of Experimental Psychology: Human Perception and Performance*, **22**(4), 904–915.

Sekuler, A. B., Gaspar, C. M., Gold, J. M. and Bennett, P. J. (2004). Inversion leads to quantitative, not qualitative, changes in face processing. *Current Biology*, **14**(5), 391–396.

Sporer, S. L. (2001). Recognizing faces of other ethnic groups: an integration of theories. *Psychology, Public Policy, and Law*, **7**(1), 36–97.

Thompson, P. (1980). Margaret Thatcher: a new illusion. *Perception*, **9**(4), 483–484.

Webster, M. A. and MacLin, O. H. (1999). Figural aftereffects in the perception of faces. *Psychonomic Bulletin and Review*, **6**(4), 647–653.

Wickham, L. H. V., Morris, P. E. and Fritz, C. O. (2000). Facial distinctiveness: its measurement, distribution and influence on immediate and delayed recognition. *British Journal of Psychology*, **91**(1), 99–123.

Wilkinson, C. and Evans, R. (2009). Are facial image analysis experts any better than the general public at identifying individuals from CCTV images? *Science and Justice*, **49**(3), 191–196.

Wright, D. B. and Stroud, J. N. (2002). Age differences in lineup identification accuracy: people are better with their own age. *Law and Human Behavior*, **26**(6), 641–654.

Young, A. W., Hellawell, D. and Hay, D. C. (1987). Configurational information in face perception. *Perception*, **16**(6), 747–759.

Yovel, G. (2009). The shape of facial features and the spacing among them generate similar inversion effects: a reply to Rossion (2008). *Acta Psychologica*, **132**(3), 293–299.

Chapter 3

EFIT-V: Evolutionary algorithms and computer composites

Chris Solomon, Stuart Gibson and Matthew Maylin

3.1 Introduction

In the event of a crime, police officers often rely on a witness to provide a comprehensive account of the incident. In some circumstances, the witness has to convey a description of the perpetrator based only on a brief encounter. The pertinent question is how is it possible to accurately convey the perpetrator's face when the image only exists as a memory in the witness's mind? This corresponds well to the typical circumstances under which a trained police officer will subsequently work with the victim (or other witness to a crime) in an attempt to produce a facial likeness or facial composite of the perpetrator. Unless the witness is a gifted artist it is unlikely that he or she will be able to provide a reliable sketch of the offender/perpetrator. Typically, assuming of course that the attacker is unknown to the witness, he or she will first be asked to provide a detailed verbal description of the attacker and to recount the incident in as much detail as possible. When the interview is complete, an attempt is then made to produce a likeness under the guidance of a specially trained operator. Whilst sketch artists are still used widely in the USA, this process will most likely (in the UK at least) use some form of computerised facial composite system. A facial composite system is a tool designed to allow the expression of the facial appearance retained in the witness's memory in some tangible form, such as a digital image or computer print-out. The desired outcome is that the generated composite be of sufficient accuracy that subsequent display to members of the public will result in direct recognition and that the details of the suspect will be supplied to the police. In many cases, a generated composite may not be accurate enough to produce a definite 'hit' but may nonetheless provoke members of the public who recognise basic similarities to provide the names of possible suspects. In most cases, it is the combination of the composite with other basic information such as age, build, domicile and the type of crime that results in the provision of suspect names.

The process by which a witness and a composite operator arrive at a final facial composite is a complex interplay of computer imaging and human cognitive function, and the final result depends on a number of factors. The overall success of the composite process is, first and foremost, reliant on the witness's ability to retain some memory of the face in question. Undoubtedly some people are better equipped to perform this task than others. Other factors such as the witness's state of mind (i.e. they may be in various degrees of shock as a result of the crime), the period of time over which the crime took place, the proximity of the perpetrator to the witness during the crime and the time elapsed between the crime and the composite construction will also affect the memory. From a scientific and technological perspective, there are critical aspects to consider in the design of an effective composite system. It should provide sufficient flexibility of use and image quality to meet the needs of different witnesses and operators and should be constructed, as much as possible, to match their normal cognitive processes.

Although systematic methods for remembering faces have been outlined by Penry (1971), the inventor of the original PhotoFIT system, it is unreasonable and impractical to expect that potential witnesses and victims will be well trained in these techniques. Rather, the emphasis must be on the design of composite systems and associated interview techniques

Craniofacial Identification, ed. C. Wilkinson and C. Rynn. Published by Cambridge University Press.

which allow the best evidence to be produced by ordinary members of the public. The relative weakness of human beings at the process of recall and description of faces as contrasted with their remarkable capacity for face recognition is well documented (Tanaka and Farah, 1993). In simple terms, to recall and describe a face accurately, even that of a family member or a close relative, is cognitively difficult for many people.

Until recently, the facial composite systems used by international police forces were exclusively based on a construction process in which individual facial features (eyes, nose, mouth, eyebrows, etc.) are selected one at a time from a large database and then electronically 'overlaid' to make the composite image. Such systems are often referred to as *feature-based* since they rely on the selection of individual features in isolation. However, after a long period of research and development work conducted largely within British universities, systems based on a rather different principle are finding increasing use by police forces (Frowd *et al.*, 2010; Gibson *et al.*, 2009; Solomon *et al.*, 2009). These systems may be broadly described as holistic or global, in that they primarily attempt to create a likeness to the suspect through an evolutionary mechanism in which a witness's response to groups of complete faces (not just facial features) converges towards an increasingly accurate image. This chapter will focus on a system which had its origins in research undertaken by Kent University and the Open University and is now commercially established under the name EFIT-V (www.visionmetric.com). EFIT-V may best be described as a hybrid system, in that it attempts to exploit holistic information whilst retaining some of the feature-based processing employed in most systems of the previous generation and which are effective under certain conditions.

The chapter is divided into four sections. In the remainder of this section, the historical background to facial composite techniques will be described, and selected research which bears on the development of the new generation of composite systems will be reviewed. In section 3.2, the central mathematical concepts and foundations on which the EFIT-V system is based will be outlined, and the conceptual implementation of this approach to composite construction will be described. Also, the preferred evolutionary strategy that is employed to achieve convergence in the system will be described. In section 3.3, the systematic tools and functions provided in EFIT-V will be described, along with the empirical adaptations which have proved necessary and effective when working with real witnesses. In the final section, an overview of the operational procedure employed with EFIT-V will be offered, and current and future directions will be outlined.

3.1.1 Historical background to holistic composite systems

Although there are many differences of detail in the various available commercial composite systems, until 5 years ago all commercial systems operated with the same feature-based philosophy, whereby the individual features of the face are selected from databases of examples which have been suitably categorised. The commercial systems *E-FIT*, *Pro-FIT*, *Identikit*, *Comphotofit* and *Faces* all fall into this category. One of the earliest innovations, which aimed partly to address the limitations associated with a finite database of candidate features was the experimental system developed by Brunelli and Mich (1996) named 'Spotit!'. This system relied on a Principal Component Analysis (PCA) model for each class of feature, achieving a reduction in the dimensionality of the problem and providing a basis from which novel features can be constructed. The 'pre-face' image or starting point in this system was the mean face into which the facial features are set/blended. The appearance of each facial feature is then controlled by seven sliders, where each slider corresponds to a principal component or mode of variation. The range of composites that can be produced using this technique is fundamentally limited by the finite size of sample used in the PCA. However, Brunelli and Mich (1996) included a tool that allowed the operator to manually distort the shape of a chosen feature. In this sense there is an unlimited set of feature shapes that can be achieved. One of the major weaknesses would appear to be that the sliders incorporated in the interface control changes in appearance defined on a mathematical premise, and as such do not correspond to a specific perceptual meaning (e.g. 'a turned-up nose'). Therefore, any prospective witness/operator will find it difficult to locate the optimum slider positions required for a good likeness to the target face.

An 'intelligent' search procedure is required to overcome the difficulty of selecting the most appropriate features from an almost unlimited sample.

Genetic algorithms (GAs) (Holland, 1975) offer a conceptually pleasing solution to the search problem. Evolutionary techniques based on Darwinian theory that simulate complex structures, textures, and motions for use in computer graphics and animation were previously described as early as 1991 (Sims, 1991). DiPaolo (2002), also working in the computer graphics arena, describes such an algorithm for facial appearance, based on an aesthetic selection process in which faces are represented by genotypes comprising 25 parameters. However, the first recorded use of a GA for generating facial composite imagery for police use was Caldwell and Johnston (1991). The GA is initialised with a population of 20 faces which are constructed from individual facial features in a style reminiscent of earlier systems. Faces are displayed to the operator, who is required to assign a fitness score to each face depending on its similarity to the target. Parent faces are chosen from the initial population according to their associated fitness score and bred with each other using the standard principles of crossover and mutation.

Attempts have been made to incorporate information concerning the configuration of facial features into feature-based systems such as EFIT, and the effectiveness of these ad hoc 'pseudo holistic' approaches has been examined (Stegmann, 2000). However, a more elegant and possibly more effective approach is to model facial variation as a whole. To the best of our knowledge, the earliest work in this direction was that of Hancock (Hancock *et al.*, 1996; Hancock, 2000) who describes a developmental system that utilises both global PCA face models and a GA. Critically, this design appears to be the first to allow composite images to be created by adjusting global/holistic properties of facial appearance, in a way that is not too demanding of the witness. Unlike previous systems this approach is truly global, relying on whole face templates (the principal components) rather than a database of facial features. In this context the principal components are often referred to as eigenfaces, which had been explored in the computer vision literature and originated with the seminal work of Sirovich and Kirby (1987). Hancock used two separate PCA models, one for face shape and another for pixel intensity values. Using two independent models overcomes problems associated with head pose and blurring which would otherwise degrade the composite images. The operator was presented with a selection of 18 faces to which fitness ratings were to be assigned on a scale of 0 to 10. The genetic algorithm selected faces with a high rating (fitness proportionate selection) as parents. Parameters defining an offspring's appearance were selected at random from the parents (uniform crossover) and a mutation applied to some of the parameter values. This procedure was performed 18 times to form a new generation of faces. Hancock's original PCA model was built on a limited sample of 20 female faces. The system has been subsequently refined and developed by Frowd and Hancock in a series of publications and is now known as Evo-FIT (Frowd *et al.*, 2010). Another experimental system based on an evolutionary/PCA method has been developed by Tredoux *et al.* (1999). The EFIT-V system with which this chapter is primarily concerned was conceived independently (the research system operated under the name EigenFIT; Gibson *et al.*, 2003) and assumed commercial form under the name EFIT-V (Gibson *et al.*, 2009; Solomon *et al.*, 2009) in 2007. This system also employs PCA model building and evolutionary search techniques though it differs somewhat in basic approach and functional details. The EFIT-V system is built on two core elements – the construction of a statistical appearance model of human faces and a stochastic search algorithm. In the following sections, we describe these core elements of the EFIT-V system.

3.2 Mathematical model of facial appearance

The mathematical model used to represent facial appearance in EFIT-V is that of a shape–texture appearance space model. The construction of such a model comprises three necessary elements or parts:

(1) The **Training** or generation of the appearance model from a population sample of faces.
(2) The **Decomposition** of a given face in digital form into its appearance model parameters.
(3) The **Synthesis** of a face from its appearance model parameters (i.e. the reverse of decomposition).

The three elements of training, decomposition and synthesis respectively enable us to *model* human facial appearance, *reduce* a digital representation of a human face to its most compact (parametric) form and conversely to *reproduce* the facial appearance from its parametric form. These steps in these three stages may be summarised as follows.

3.2.1 Training – the generation of the facial appearance model

(1) The faces in the training set are first hand marked at a number of control points delineating the main features of the face to form a set of shape model vectors $\mathbf{S}_i$. The Procrustes aligned mean of the shape vectors, $\overline{\mathbf{S}}$ is calculated. We refer to this as the *prototype* shape.
(2) A principal component analysis (PCA) is carried out on the ensemble of aligned shape vectors – that is, we find a linear combination of the mean-subtracted shape vectors $\mathbf{P}_S = (\mathbf{S} - \overline{\mathbf{S}})\mathbf{B}_S$ which satisfies the required orthogonality relationship – $\mathbf{P}_S^T\mathbf{P}_S = \Lambda_S$ where $\boldsymbol{\Lambda}_s$ is a diagonal matrix and $\mathbf{P}_S$ is the matrix containing the principal components. The required diagonalising matrix $\mathbf{B}_s$ can be found by standard eigenvector analysis.
(3) The corresponding texture map vectors $\mathbf{T}_s$ for each face in the training sample are warped using standard Delaunay triangulation to the prototype shape. The resulting texture values are then referred to as the *shape-free* texture maps.
(4) A PCA is carried out on the shape-free texture maps. That is to say, we again find a diagonalising matrix $\mathbf{B}_T$ such that $\mathbf{P}_T = (\mathbf{T} - \overline{\mathbf{T}})\mathbf{B}_T$ with $\mathbf{P}_T^T\mathbf{P}_T = \Lambda_T$.
(5) It is important to recognise that the shape and texture in a human face are correlated. In the final stage, we therefore combine the separate linear models by decorrelating the shape and texture. We form a block matrix, $\mathbf{B}$:

$$\mathbf{B} = \begin{bmatrix} \mathbf{W}\mathbf{B}_S \\ \mathbf{B}_T \end{bmatrix}$$

where the upper element of the block contains the eigenvectors which diagonalise the shape covariance and the lower element comprises the eigenvectors which diagonalise the texture (shape-normalised) covariance. The matrix $\mathbf{W}$ is a diagonal matrix of weights which is required to make the shape and texture parameters, which have different units, commensurate (Cootes *et al.*, 1995). This is achieved by scaling the total variance associated with the shape so that it matches the total variance associated with the texture. In this way, equal weighting is ascribed to the shape and texture. This process may be described mathematically as

$$\mathbf{W} = r\mathbf{I} = \begin{bmatrix} r & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & r \end{bmatrix}$$

$$r = \frac{\lambda_g}{\lambda_s}, \lambda_g = \sum \lambda_{gi}, \lambda_s = \sum \lambda_{si}$$

where λ_{gi} is the variance associated with ith texture principal component and λ_{si} is the variance associated with the ith shape principal component. We apply a further PCA on the columns of $\mathbf{B}$: namely, we seek an orthogonal matrix $\mathbf{C}$ such that

$$\mathbf{C} = \mathbf{Q}^T\,\mathbf{B}$$

where the columns of $\mathbf{Q}$ are the eigenvectors and $\mathbf{C}$ is the matrix of appearance parameters for the training sample. The key result here is that each column of $\mathbf{C}$ provides a parametric description of the corresponding face in the training sample which is optimally compact in the linear, least-squares sense.

In Figures 3.1 and 3.2, we show the first few principal components of facial shape and texture according to the procedure outlined above.

3.2.2 Decomposition of a face into appearance parameters

Decomposition of a given face into its appearance parameters proceeds by the following stages –

(1) The facial landmarks are placed and the Procrustes aligned shape vector $\mathbf{S}$ of the face is calculated.
(2) $\mathbf{S}$ is projected onto the shape principal axes $\mathbf{P}_S$ to yield the *decoupled* shape parameter vector, $\mathbf{b}_S$.
(3) The face texture is warped to the prototype or ‘shape-free’ configuration.
(4) The shape-free texture map is projected onto the texture principal axes $\mathbf{P}_T$ to yield the decoupled texture appearance parameters.
(5) The appearance parameters are calculated using the eigenvector matrix

$$\mathbf{C} = \mathbf{Q}^{\mathrm{T}}\mathbf{b} = \begin{bmatrix} \mathbf{Q}_{\mathrm{S}}^{\mathrm{T}} & \mathbf{Q}_{\mathrm{T}}^{\mathrm{T}} \end{bmatrix} \begin{bmatrix} \mathbf{W}\mathbf{b}_{\mathrm{S}} \\ \mathbf{b}_{\mathrm{T}} \end{bmatrix}.$$

3.2.3 Synthesis of a face from its appearance parameters

The reconstruction of the separate shape and (shape-free) texture vectors of a sample face from its

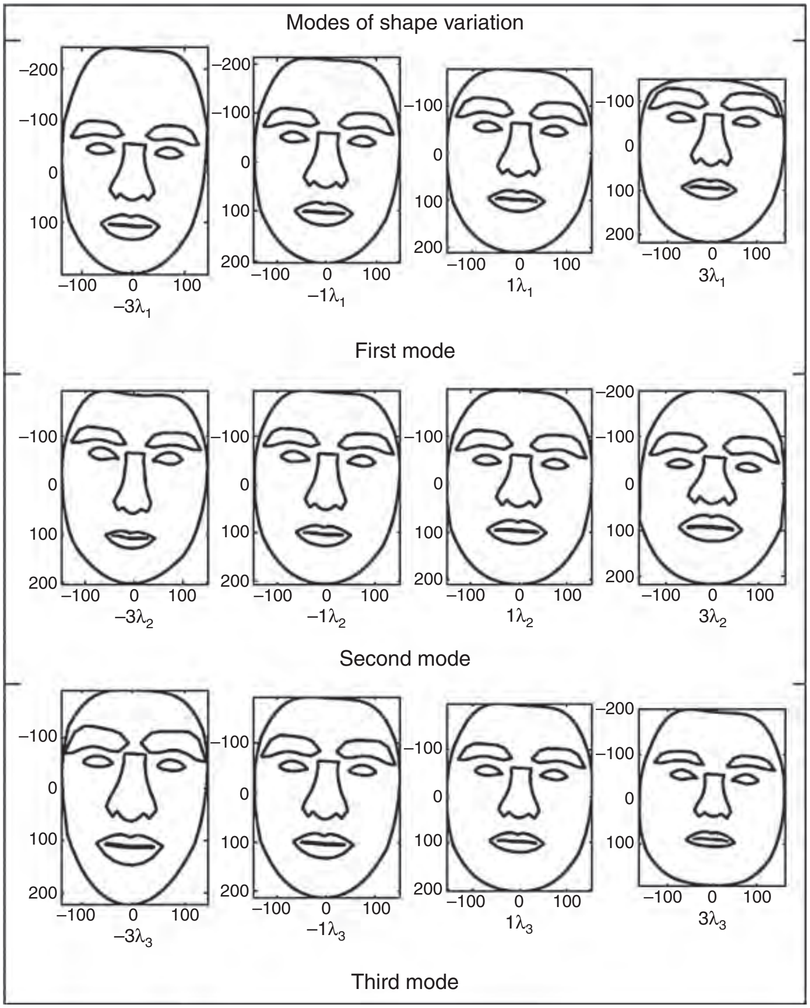


Figure 3.1 Illustration of the shape principal components or shape 'modes' extracted from the training sample of faces. Negative and positive multiples of each of the three modes are added to the average shape to indicate how they affect the face shape. Note how the modes affect multiple aspects of the face simultaneously – e.g. the third mode significantly affects head shape ('long and pointed' to 'square'), eyebrows ('thick' to 'thin') and mouth shape ('full and curved' to 'thin and straight').

appearance parameters c is calculated through the linearity of the model according to the equations

$$\mathbf{S} = \overline{\mathbf{S}} + \mathbf{P}_S \mathbf{W}_S^{-1} \mathbf{Q}_S \mathbf{c} \qquad \mathbf{T} = \overline{\mathbf{T}} + \mathbf{P}_T \mathbf{Q}_T \mathbf{c}$$

where $\overline{\mathbf{S}}$ and $\overline{\mathbf{T}}$ are the mean shape and shape-free textures, $\mathbf{P}_S$ and $\mathbf{P}_T$ are the shape and texture principal components and $\mathbf{Q}$ is the eigenvector matrix separable into shape and texture block form as $\mathbf{Q} = \begin{bmatrix} \mathbf{Q}_S \\ \mathbf{Q}_T \end{bmatrix}$. The decoupled shape and texture appearance parameters are given by

$$\mathbf{b}_S = \mathbf{W}_S^{-1} \mathbf{Q}_S \mathbf{c} \quad ; \quad \mathbf{b}_T = \mathbf{Q}_T \mathbf{c}.$$

Warping the shape-free texture to the required shape completes the facial synthesis. We refer the reader to the work of Cootes and colleagues (Cootes *et al.*, 1992, 1995, 1998; Cootes and Taylor, 2001) for the comprehensive mathematical details of appearance models.

3.2.4 Summary of appearance model properties

Despite the relative mathematical complexity, we can easily summarise the key properties of appearance space:

- Any face can be described as a (parametric) vector of coefficients, $\mathbf{c} = [\, c_1 \quad c_2 \quad \cdots \quad c_{n-1} \quad c_n \,]^T$, providing the extension of the face along each of the appearance space axes.

Modes of texture variation

$-3\lambda_1$ $-1\lambda_1$ $1\lambda_1$ $3\lambda_1$

First mode

$-3\lambda_2$ $-1\lambda_2$ $1\lambda_2$ $3\lambda_2$

Second mode

$-3\lambda_3$ $-1\lambda_3$ $1\lambda_3$ $3\lambda_3$

Third mode

Figure 3.2 Illustration of the texture principal components or texture 'modes' extracted from the training sample of faces. Negative and positive multiples of each of the three modes are added to the average texture to indicate how they affect the facial appearance. Note that the textures are shape normalized and that the class model in this example is mixed race.

- Appearance space is a multidimensional vector space, the axes of which correspond to specific shape-texture facial characteristics/features. As such, the appearance parameters control the amount of each global, shape-texture principal component in the face. All dimensions in face space represent *commensurate* quantities which describe characteristics of the face as a whole. Such a representation is mathematically and intuitively satisfying.
- Appearance space is an optimally compact space. The combined linear PCA on shape and texture ensures that the resulting matrix of appearance parameters is optimally compact in the linear, least-squares sense. Thus, a representative training sample will enable us to reconstruct both in and out-of-sample images, to within a given least-squares error, using a minimum number of dimensions in appearance space. **The reduction in the dimensionality of the problem is the major advantage of the PCA**.

These properties mean that the appearance space provides a compact, parametric representation of facial appearance which is inherently *global or holistic*. This is so because the alteration of any one of the parameters will in general affect the entire facial appearance. Since the sample distribution of the

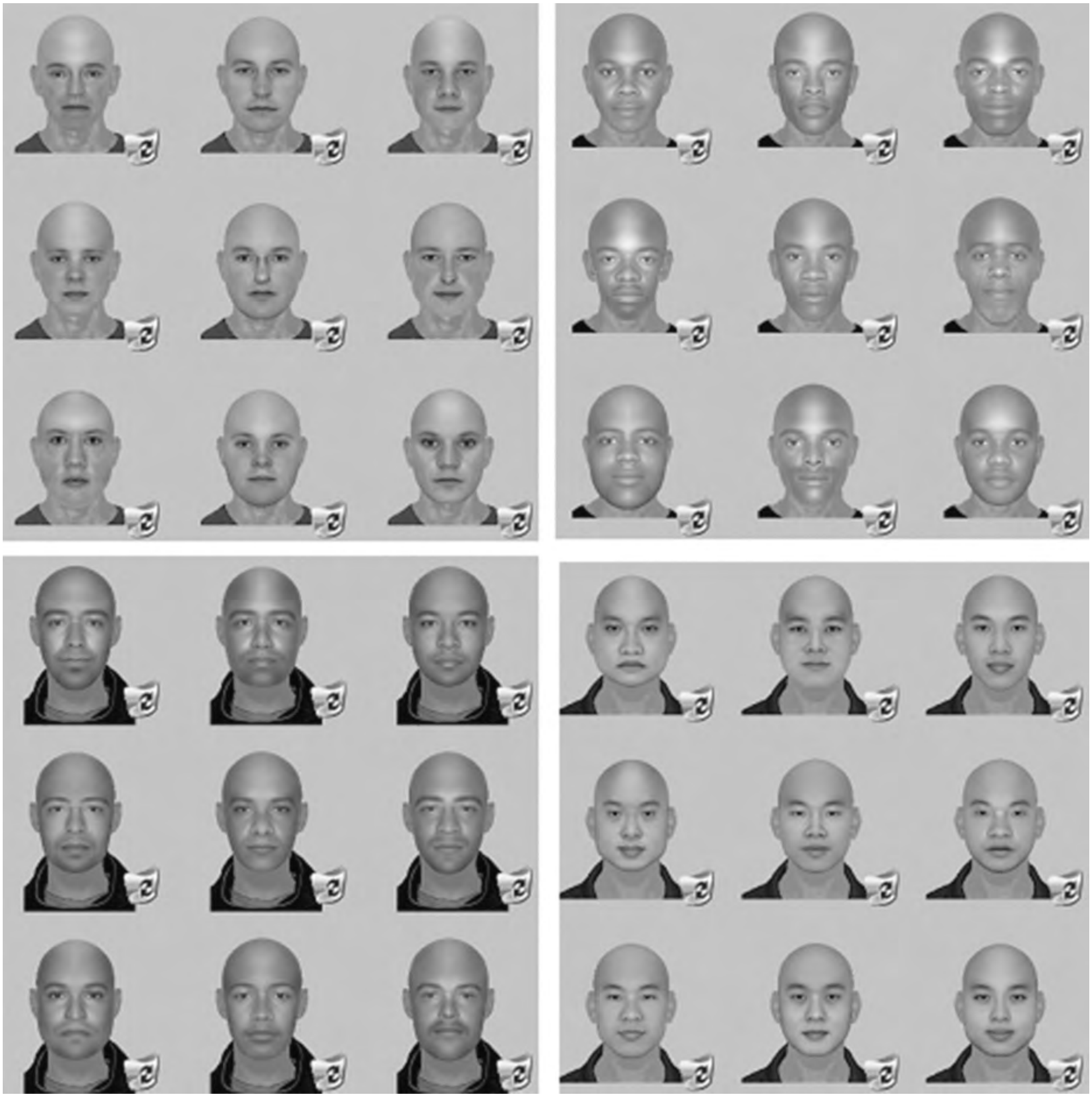


Figure 3.3 Random generation of faces in EFIT-V through sampling a statistical model of facial appearance. Examples are shown for each of the class models *White male*, *Black male*, *Hispanic male* and *Oriental male*.

appearance parameters may be well approximated by a multivariate Gaussian, it is a simple matter to randomly sample the distribution to generate plausible new examples of faces. Figure 3.3 shows examples of such randomly generated faces in EFIT-V for four different class models (White male, Black male, Oriental male, Hispanic male).

3.2.5 Significance of appearance model representation

The construction of a statistical appearance model of the human face for a facial composite system first requires a sample of facial images which are broadly representative of the faces that one will subsequently wish to synthesise. As such, they should show appropriate variation in age and facial characteristics. Earlier versions of the EFIT-V system constructed a single appearance model over a broad range of ethnicities. Whilst this has certain advantages in terms of mixed race representations and mixed gender characteristics, more recent versions have employed separate appearance models for each broad ethnic group which are referred to as *class models*.

The essential significance of the appearance model in our specific context is four-fold:

(1) It produces an optimally compact, parametric representation of the face for a selected class model.
(2) By manipulation of these parameters, it is possible (at least in the case of a sufficiently large training sample) to accurately approximate any face belonging to the given class even if it was not used as part of the training process.
(3) The random generation of examples from the model forms the basic mechanism by which the recognition capacity of the witness is brought to bear on the creation of the composite image.
(4) This representation allows us to consider the manipulation of the underlying parameters as a search in a multidimensional face space in which each parameter defines a dimension and has a certain extension within it.

The consideration of how to manipulate the underlying parameters in such a way as to converge towards the solution has been researched in some detail. The method employed in EFIT-V devised by Pallares-Bejarano (2006) is briefly summarised in the following section.

3.2.6 Stochastic search algorithm

The appearance model described in the previous sections provides the means for synthesising plausible face images and we find that an adequate approximation to any face can be typically obtained from an appropriate selection of 60 independent appearance model parameters. In principle, the parameter values could be determined using many different approaches. For instance, a naive approach would be to construct a user interface in which each parameter value is controlled by a slider, and the resulting composite face displayed to the witness. Another approach would be to implement a purely random search of the parameter space in which a large number of candidate faces are synthesised from which the best likeness to the target face is selected. Although, in theory, both of these methods are capable of achieving a likeness, neither method takes into account 'ease of use' nor the time required to achieve a likeness.

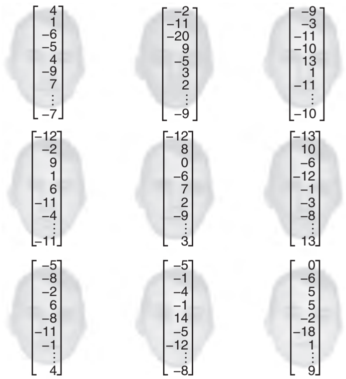
It is crucial to recognise that the optimum search procedure for this task must be an algorithm that is a suitable compromise between human usability and speed of convergence (i.e. the required number of faces seen and rated by the user before a satisfactory composite is achieved). The stochastic search procedure in EFIT-V is an asexual evolutionary process termed the SCM algorithm (select-clone-mutate) whose steps are as follows:

(1) The process is initialised by using a pseudo-random number generator to obtain nine vectors each containing 60 double precision random numbers (decision variables) drawn from a standard normal distribution:

$$\mathbf{c} \sim N(0,1)$$

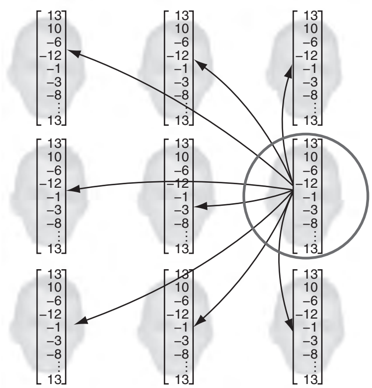
Each of the nine vectors thus constitutes a single genotype, representing an encoded face. Collectively, the nine genotypes are referred to as the initial population. A decoded face image is termed a phenotype. The purpose of the initial population is to seed the evolutionary algorithm: thereby providing a starting point from which a likeness to a target face can be evolved.
(2) A transformation is applied to each genotype vector. The transformation maps the standard, normal decision variables to appearance model parameters that follow the multivariate normal distribution relating to the chosen class model.
(3) From each of the appearance parameter vectors, a face image is synthesised as described in section 3.2. Figure 3.3 illustrates typical examples of nine such phenotype face images for a given generation.
(4) From the array of nine faces, the witness is required to select the single face that most closely resembles the suspect (Figure 3.4) and to (optionally) reject one or more of the weaker examples. The selected face is the fittest phenotype, also referred to here as the stallion. It is the only face in the current generation from which genetic code is propagated into the next generation.

Figure 3.4 The main steps in the SCM (select-clone-mutate) search algorithm.

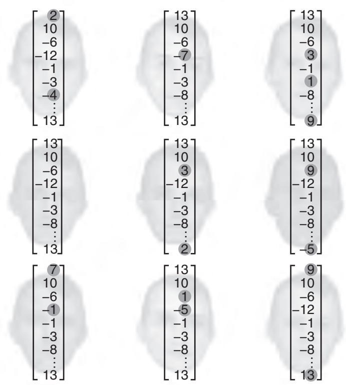


(a) **Initial population:** Genotypes generated using PRNG

(b) **Initial population:** Phenotypes synthesised from the genotypes - fittest phenotype (as selected by the witness) circled in red.



(c) **Cloning (Multiply):** Genotype corresponding to fittest phenotype cloned nine times.



(d) **Mutation:** Random mutations on eight of the nine clones. The stallion remains unaltered, although its position in the new generation is randomised.

(e) **New generation:** Based on mutated genetic material from selected face.

The rejected faces can be used to bias the subsequent generation of faces away from regions of the search space occupied by previously rejected faces.

(5) The genotype corresponding to the stallion is duplicated or cloned nine times, thereby copying the genetic code of the selected face into a new generation of nine faces.

(6) Eight of the cloned genotypes are mutated to produce variations on the selected stallion image. The remaining clone is left unaltered and is positioned randomly in the new array of nine faces. From these genotypes nine new phenotypes are constructed. Thus a new generation of faces is produced.

(7) Steps are repeated until an acceptable likeness to the suspect's face is achieved.

The essential steps in the SCM algorithm are outlined in Figure 3.4.

3.3 Systematic operations in the EFIT-V system

In section 3.2, the elements of the EFIT-V system necessary for a stochastic or evolutionary mode of operation were presented. In that mode of operation, the witness is simply required to make simple decisions in response to the facial stimuli. The conceptual simplicity of this method and its inherent appeal to facial recognition capacity is a major strength. However, at this juncture, we must stress two things. Firstly, many important aspects of facial appearance which appear to be closely related to identity *cannot be modelled* effectively using the statistical approach outlined above. These aspects include hairstyle, effects bearing on the complexion of the skin (such as wrinkles, freckles, acne and pockmarks) as well as distinctive features, such as dimpled chins, that are not captured by the model. Secondly, feedback from police forces using early versions of the system and subsequent operational feedback strongly support the notion that witnesses should have the freedom to alter the appearance deterministically at any time. Thus, although evidence has been marshalled to support the notion that observers achieve face recognition tasks through configurational cues, it is nonetheless common for a witness to remember particular, distinctive features and to sometimes encounter profound difficulty in progressing the composite until the desired feature is incorporated. Under such circumstances, it is clearly less than ideal to request that a witness 'wait until a suitable feature is evolved'.

In this section, we outline some of the means provided in EFIT-V for exploiting this kind of witness information. The attempt to integrate feature-based (a large, hooked nose) and semantic information ('he looked really sleazy') with a holistic-evolutionary mode of operation thus results in a hybrid approach, providing maximum flexibility to the witness and operator.

3.3.1 Locking facial features

A relatively common occurrence for a witness operating in the evolutionary mode is to identify a specific facial feature within a face that is deemed satisfactory and which they do not wish to change. EFIT-V allows the operator to lock a facial feature by selecting the corresponding region on a schematic face image. Once the feature has been locked, it appears highlighted in the schematic image to inform the user that no further shape deformation of the selected feature will occur during subsequent generations. In terms of our underlying mathematical model, the process is expressed by a vector addition comprising the current stallion shape S_t and a snap shot of a previous stallion S_{t_0} captured at time t_0. Here, we use *time* to refer to a particular generation number, with $t > t_0$.

$$s'_t = s_t\left[\mathrm{I} - \mathrm{W}_f\right] + s_{t_0}\mathrm{W}_f$$

where I is the identity matrix and W_f is a diagonal matrix with elements equal to one or zero. W_f is referred to as the feature selector since it effectively extracts all of the coordinates from S_{t_0} corresponding to the fixed feature. We can extend the equation to include multiple features, locked at different instances (generations).

$$s'_t = s_t\left[\mathrm{I} - \mathrm{W}_{f1} - \mathrm{W}_{f2} - \mathrm{W}_{f3} \cdots - \mathrm{W}_{fN}\right] + s_{t_1}\mathrm{W}_{f1} + s_{t_2}\mathrm{W}_{f2} + \cdots s_{t_N}\mathrm{W}_{fN}$$

$\mathrm{W}_{f1}, \mathrm{W}_{f2}, \cdots, \mathrm{W}_{fN}$ are the feature selectors for the 1st, 2nd and nth features respectively (i.e. nose, mouth ... etc.) and $s_{t_1}, s_{t_2}, \cdots s_{t_N}$ are snapshots of the stallion taken at times $t_1, t_2 \cdots t_n$. Hence one or more features may be locked at once. If the user wishes to evolve a single feature in isolation, all other features can be locked.

3.3.2 Feature clone tool

A closely related function to feature locking, and one that was developed as a result of feedback from police operators, is the feature clone function. This function allows a witness to identify an individual feature within one of the faces in the given generation and then copy that feature across the whole generation. This functionality is particularly useful when a witness is able to accurately recognise a feature but does not deem the face in which it is found to be the most accurate. Feature cloning copies the shape only to avoid problems with blending the texture between features. Rather than attempt to impose the feature in terms of the global appearance model which may not span the vector space, we adopt a simpler ad hoc approach. This is essentially similar to that used in the manipulation of individual feature shapes and is therefore described in section 3.5.

3.3.3 Blend tool

The simplicity of the SCM algorithm employed in EFIT-V excludes the option for carrying over facial characteristics from more than one phenotype to the next set of nine faces. This can be achieved in some measure through systematic blending of one or more faces in the current generation. This is achieved by forming a weighted combination of the genotypes (appearance model parameter vectors $\{\mathbf{c}_i\}$) comprising the current generation. Faces for which $\mathbf{c}_i = 0$ do not contribute to the blended face. If the blended face is considered to bear a better likeness to the target than the current stallion, the current stallion is replaced by the blended face. The updated stallion is reconstructed from the appearance model parameter vector in the usual way.

3.3.4 Facial attribute manipulation

Descriptive semantic labels are often used to describe facial appearance. For instance, a witness may describe a perpetrator as 'more masculine', less 'healthy-looking' or 'older' with respect to the current composite image. Traditional methods for producing composite imagery are incapable of allowing direct control of such perceived attributes. Conversely, this is relatively easy to implement in the EFIT-V system framework and is achieved by identifying directions in the parameter space that can be directly related through statistical methods to a perceived increase or decrease in a specific attribute.

We outline here a simple procedure for defining such a direction through the appearance model parameter space, corresponding to maximum variation in a specified facial attribute. A similar approach to modifying facial appearance has previously been described by Burt and Perrett (1995) and Benson and Perrett (1993), in which shape and texture were treated separately. An appearance model offers a more elegant solution in which facial appearance can be modified by perturbing a single vector of parameters that simultaneously control both shape and texture.

3.3.4.1 Training process

To manipulate a chosen facial attribute, a prior training procedure is required in which a relationship is sought between the attribute of interest and each appearance parameter. For attributes in which a dichotomy exists, such as the sex attribute, the simplest approach is to separate the training examples into two classes A and B (e.g. males and females). Prototypes (constructed from class means) can then be formed by determining the mean vector of appearance model parameters for each class, $\bar{\mathbf{c}}_A$ and $\bar{\mathbf{c}}_B$. Having formed two prototypes, a direction in appearance space is simply calculated as the difference vector between them,

$$\Delta\mathbf{c} = \mathbf{c}_{att} = \bar{\mathbf{c}}_A - \bar{\mathbf{c}}_B \,.$$

We refer to the vector of appearance model parameters $\mathbf{c}_{att}$ as the attribute vector. For attributes that vary continuously (albeit on a discrete scale, e.g. age) with respect to variations in appearance model parameter values, an alternative approach is required. One such approach is to perform a multiple regression analysis, relating the attribute to the parameter values using a regression equation. Regression methods are closely related to the techniques and procedures employed in classification problems. Ramanathan and Chellappa (2005) use probabilistic eigenspaces and a Bayesian classifier to determine the age difference indicated by two images of the same subject's face, where the time interval between the first and second image was in the range 1–9 years. Details of alternative approaches to manipulating age (a subtly different problem to classification) can be found in Scandrett *et al.* (2006), and Lanitis *et al.* (2002). However, a simpler approach that has been found to produce visually acceptable results in our application is to arrange the attribute

Figure 3.5 Automatic ageing in EFIT-V using the learned ageing direction in appearance space.

values (scores) in ascending order and perform a median cut, thereby forcing a dichotomy. In this case, each sample face decomposed into its appearance model parameters and assigned to either class A or B as follows:

$$\mathbf{c}_i = \left\{ \begin{array}{c} class\,A\;\; if\;\; rank(i)\; \leq Median \\ class\,B\;\; if\;\; rank(i) > Median \end{array} \right\}.$$

Prototypes and the attribute vector can then be constructed exactly as previously described. The current commercial version of EFIT-V currently allows the transformation of age and distinctiveness and it is likely that other attributes on which there is a high level of agreement amongst observers will be included in the near future.

A face is aged or rejuvenated by adding or subtracting a scalar multiple of the ageing attribute vector to the appearance model parameters representing the composite face, and then reconstructing the image. In EFIT-V, the attribute manipulation is always performed on the current stallion. The attribute transformation process can be represented using vector notation as:

$$\mathbf{c}'_i = \mathbf{c}_i + \alpha \mathbf{c}_{att}$$

where the scalar α controls the degree of ageing or, for negative values of α, the rejuvenation of the face. An example of facial ageing of an EFIT-V image using this method is shown in Figure 3.5.

3.3.5 Local feature manipulation

Although one of the main strengths of the PCA model is its capacity to generate complete and plausible face images, with facial features properly related in the context of the whole face, there are instances when this holistic approach can be a disadvantage. One important case in which the global manipulation method proves inadequate is when the witness has remembered something distinctive about a particular facial feature and wishes to make a change to a localised region of the composite image. Localised modifications of this nature cannot be accommodated by the appearance model because in the global PCA framework, alterations to individual features are always accompanied by uncontrollable changes to the face as a whole.

EFIT-V allows the operator to manipulate the aspect ratio, position and overall size of the individual facial features, as well as to effect a trapezoidal transform on the coordinates. The procedure involves shape modifications only, thereby avoiding any potential problems associated with blending of texture patches. This straightforward method provides a powerful tool for making adjustments to the ongoing composite image that is not too onerous for the witness. Statements regarding the width, height and position of facial features are easily interpreted and do not require any complex vocabulary that could be misconstrued. As with any user-defined change in facial appearance, the operator themselves may also place implicit constraints on the deformation. With these constraints in place, face shapes that are modified using this local feature tool retain a realistic appearance despite the fact that, in general, they lie outside the span of the shape principal components. This implies that the local feature tool provides a means for introducing new plausible shape variation.

3.3.5.1 Managing feature manipulation in appearance space

The construction of the statistical appearance model means that the core elements of the face are represented by a vector $\mathbf{c} = (c_1, c_2 \cdots c_n)$ of global appearance parameters. These parameters are global in the sense that altering a single parameter alters both the shape and texture of the entire facial appearance. Distinct parameter vectors $\mathbf{b}_s$ for the shape and $\mathbf{b}_T$ for the texture can be obtained directly from the appearance model via the following equations:

$$\begin{aligned}\mathbf{b}_s &= \mathbf{Q}_s\,\mathbf{c}\mathbf{W}^{-1} \\ \mathbf{b}_T &= \mathbf{Q}_T\mathbf{c}\end{aligned}.$$

In turn, the actual global shape vector **x** of the face and the shape-normalised texture-map **g** of the face can be generated as:

$$\begin{aligned}\mathbf{x} &= \overline{\mathbf{x}} + \mathbf{P}_s\mathbf{b}_s \\ \mathbf{g} &= \overline{\mathbf{g}} + \mathbf{P_T}\,\mathbf{b_T}\end{aligned}.$$

Consider now some arbitrary change introduced into the coordinates of **x** by a local manipulation of a feature shape such as the nose or mouth so that:

$$\mathbf{x} \mapsto \mathbf{x}' = \mathbf{x} + \Delta\mathbf{x}.$$

To represent this vector in the global shape space, spanned by the principal components, we must project the vector onto the shape principal components contained in the columns of matrix $\mathbf{P}_s$ as:

$$\mathbf{b}'_s = \mathbf{P}_s^T(\mathbf{x}' - \overline{x}).$$

In general, the localised manipulation of coordinates can result in a new shape vector, which *does not lie* within the span of the shape space and the new shape vector $\mathbf{b}'_s$ will not, in general, enable exact reconstruction of the shape vector. Rather we have:

$$\mathbf{x}' = \overline{\mathbf{x}} + \mathbf{P}_s\,\mathbf{b}'_s + \vec{\varepsilon}_s$$

where $\vec{\varepsilon}_s$ defines a residual component of the modified shape that cannot be constructed using a linear combination of the shape principal components $\mathbf{P}_s$, since it is orthogonal to the vector space spanned by the columns of $\mathbf{P}_s$. It is clear that to maintain an accurate parametric representation of the generated composite, it is necessary to keep a record of both the global appearance vector c and the component of the shape deformation that is perpendicular to shape principal components. The dimensionality of $\vec{\varepsilon}_s$ is high compared with the very compact representation provided by the appearance vector (or indeed its associated shape and texture parameters). One approach to the book-keeping of composite production is to incorporate any local deterministic changes introduced by the operator and witness through repeated projection of the shape and texture vectors onto their respective principal components, thus, obtaining the nearest global representation together with the component $\vec{\varepsilon}_s$. In this work, a simpler but equally effective approach was taken, which is summarised as follows:

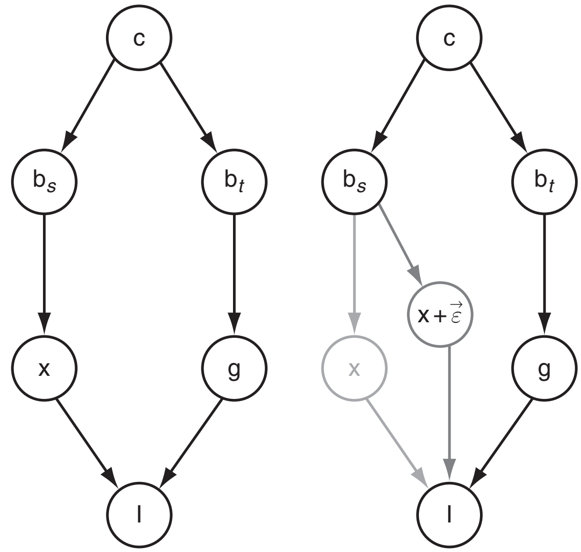


Figure 3.6 Left: Construction of a composite image within the confines of the learned appearance space. The appearance model parameters $\mathbf{c} = (c_1, c_2 \ldots c_n)$ are decoupled into shape and texture parameters $\mathbf{b}_s$ and $\mathbf{b}_T$. These in turn are used to generate the shape-free texture **g** and shape **x** before the final image **I** is created. Right: The same sequence but systematic offsets which lie outside the span of the appearance space are added to the shape vector before the final image is created.

(1) Allow the core evolutionary procedure to continue as in the standard mode of operation, resulting in instances of whole face variations on the current stallion.
(2) Any deterministic changes to the shape and texture introduced by the witness are recorded and treated independently of the global model as offset vectors. In effect, the net deviation of the shape and texture from that predicted by the global model is updated on each occasion that deterministic changes are made.

The major advantage of this approach is that it removes the computational overhead which is associated with recalculation of the global model and simplifies the implementation. The approach is schematically summarised in Figure 3.6.

3.3.6 Applying fine details to a composite

In general, given an appropriate training sample, the AM is capable of capturing most of the natural variation in facial appearance. However, the finer structures such as wrinkles, freckles, pockmarks, etc. that often occur in real faces are underrepresented in face images synthesised from the model. This shortfall is particularly apparent when generating faces of old subjects since fine facial detail becomes more prevalent with an increase in age. The problem is intrinsic to the process by which new examples of faces are synthesised, whereby a new example of a face is constructed by a weighted average of existing face images. Inevitably, this averaging procedure results in a certain degree of smoothing in the synthesised face. Fine details that exhibit low spatial correlation between observations (sample faces) tend to be 'averaged out' by this process. Although wrinkles are often more common in specific regions of the face (for example 'crow's feet' appear at the outer corners of the eyes), their prominence, exact position and frequency of occurrence vary from subject to subject. Landmarking these delicate features is not practicable, therefore a different approach is required. The issue of enhancing fine detail in averaged face images has been investigated in previous work. In Tiddeman *et al.* (2001) a prototype face, formed by averaging a sample of face images, was decomposed using a wavelet analysis and the high-order details boosted to compensate for the inherent loss of high spatial frequency information.

EFIT-V employs an empirical but effective method for applying fine facial detail to a target face. A detailed mathematical description is a little lengthy and so we restrict the explanation here to the essential idea: namely, to extract the fine facial details from a sample face A with the intention of applying these details directly to some target face B in a plausible fashion. The procedure breaks down into five basic steps:

(1) First identify a face image containing the detailed structures (e.g. wrinkles).
(2) Perform a piece-wise affine warp on the landmark coordinates to transform the shape to a reference (typically, the average shape).
(3) Perform a high-pass Fourier filtering of the image to extract only the detailed structure.
(4) Add some (adjustable) fraction of the high-pass filtered image to the Fourier domain version of the shape-normalised, target image.
(5) Inverse Fourier transformation and reverse the affine warp.

An example image is shown in Figure 3.7.

Figure 3.7 Application of high-frequency image structure.

3.4 The use of EFIT-V – operational procedure

It has been outlined how EFIT-V combines stochastic search methods with systematic tools for altering facial appearance to form an essentially hybrid system. At the time of writing, the system has been in routine use for 4 years (it has undergone continuous and significant developments over this time) by approximately 45 police forces. EFIT-V is designed to be flexible, so that the approach to face construction can be adapted to the witness and their abilities and preferences. For this reason, no rigid operational procedure is recommended. However, the following sequence of events is typical of an EFIT-V interview.

(1) An initial interview is first conducted. Current ACPO (Association of Chief Police Officers) guidelines recommend an interview based on the cognitive model initially devised by Geiselman and Fisher. We note in passing that alternative interview models are being researched which may better suit the methodology of systems such as EFIT-V.
(2) On system start-up, a PACE-compliant form is displayed into which details that identify the composite are entered. Fields are provided for the witness's forenames, surname, date of birth and also the operator's rank and number. This information is combined with the current date to generate a unique reference number that can be used to identify the composite in the future.
(3) In an attempt to accurately seed the starting point for the stochastic search, they are asked to select from a sample of basic face shapes displayed on the screen. An 'unsure' option exists in which case the

first faces are generated as variations about the average face shape.

(4) The witness then proceeds to select the hairstyle, or in the event that this was not visible, the headwear worn by the suspect. From a perceptual point of view it is sensible to ask the witness to select an appropriate hairstyle first as the external features are more salient in unfamiliar faces and there is evidence to suggest that facial features should be selected in order of decreasing significance. The user can scroll through the available hairstyles using a slider, with each increment in slider position displaying nine more hairstyles in the familiar three-by-three configuration. Hairstyles are mapped onto a blank head-shape so that the witness may view the hair in some context. A filter is provided to sort the hairstyles into categories describing the length and colour. Thus the number of candidate hairstyles that the witness must examine can be greatly reduced by marking the appropriate checkboxes provided in the filter.

(5) At this point, the stochastic search begins and a first generation of nine faces is presented. The witness is asked (but not forced) to select that face from the nine which best resembles the suspect. If unable to do this because none bears sufficient resemblance, the witness may request to see another group until a face appears that constitutes a satisfactory starting point. Virtual faces that exhibit a poor likeness to the target may be removed from view. This de-clutters the field of view making it easier for the witness to form an opinion on the suitability of the remaining visible faces. Once the witness has made a choice, the initial population is then replaced by nine more faces comprising the first generation and the process of selecting a face is repeated.

In this mode, the user has the option to lock the shape of a particular facial feature that exhibits a good likeness to the corresponding feature in the target face. This is achieved by choosing a region in an iconic face located on the right hand side of the interface and then selecting one of the nine virtual faces. Placing the cursor over a feature in the iconic face and using a single click of the left mouse button turns that region from grey to blue, indicating that the shape of the chosen feature will be fixed through subsequent generations. Deselecting a region of the iconic face re-introduces shape variation in the previously locked feature. If required, more than one feature may be fixed at any given instant. The available choice of features that can be locked are the eyebrows (left/right pair), eyes (left/right pair), nose, mouth and face shape or any combination of these.

(6) Typically, after a witness has selected the preferred face from 5 or 6 generations, he or she will be prompted to provide feedback on the composite image so far and whether there are any definite characteristics or features that they would like to alter. The operator will then respond using the array of systematic tools at their disposition in an attempt to 'progress' the composite towards a better likeness. This step is an important and characteristic one in the use of EFIT-V since the operator should resist (in general) any attempt to create a final product at this stage but only to use the information provided by the witness to progress the composite likeness and confirm whether (or not) the modified composite now represents a better likeness.

(7) If the witness confirms an improved likeness, the system immediately reverts to the evolutionary mode and takes this likeness as the starting point for the subsequent stochastic variations. Again, the witness is then asked to select the best example from a sequence of generations, 3–6 being typical.

(8) Further systematic changes may be requested at this stage, including the use of paintwork using an associated graphics package such as Adobe Photoshop or Corel Paintshop. Paintwork can be added to EFIT-V images at any time and the system used subsequently in the normal way but it is generally advisable to defer certain forms of artistic enhancement to late in the process due to the problems of image registration.

(9) An *Undo* function (1 step back at a time) and a *load* function (back to an arbitrary step in the construction) functions are available. Typically, these are used when the witness feels that a poor choice of face has been made, or a face has been selected unintentionally and another image is preferred.

(10) Once an acceptable likeness has been obtained, the current stallion image (best likeness) is saved both as a graphics file and as a file internal to EFIT-V which can be loaded into the program at any point in the future. A complete audit trail is

kept by EFIT-V so that the time, date, personnel involved and all steps in the construction process are saved to the hard disk of the computer. The entire contents can also be burned directly to CD from within EFIT-V as an exhibit for subsequent evidential procedures.

3.4.1 Use of EFIT-V and future development

A meaningful evaluation of the effectiveness of any facial composite system is surprisingly complex. From a policing perspective, the primary goal is very simple: namely, the provision of a correct name for the suspect. Secondary goals also exist: namely, speed and ease-of-use which can have direct financial implications for police forces. The attempt to reach the primary goal is affected by many factors:

- The inherent capability of the composite system to create good likenesses (i.e. its *imaging* capability).
- Its inherent methodology (i.e. the degree to which the system matches the cognitive processes and needs of the witness and operator).
- The capability and skills of the police operator. Although standardised training courses exist, operator skill and experience vary widely.
- The capability, willingness and emotional state of the witness.
- The nature of the crime or offence. There is, for example, some provisional evidence that offences in which the witness has time or presence of mind to consciously attempt to commit the face to memory result in a different encoding of the face to offences in which no conscious attempt is made (e.g. a distraction burglary in which the victim will have no reason to consciously attempt to remember the face of the perpetrator *at the time of the incident*).
- The effective use of the composite once it has been created. In simple terms, a composite image may represent a good likeness to the target subject but if the subsequent procedures do not result in a sufficient or appropriate cross-section of the police or public seeing the image, it will fail in its basic objective. Strangely, this aspect and its impact on success rates are often overlooked.
- The supply of relevant supplementary information along with the composite image can affect the ability and willingness of someone to offer a name. For example, an image supplied with supplementary information such as 'a white male aged 30–35 years, with a strong northern accent and scruffily dressed' is likely to have better chances of success than the image alone.

At the time of writing, EFIT-V is in routine use by just over half of the UK's police forces and in seven other countries. As a commercial system, the evaluation of its effectiveness to date has largely been gauged by feedback from police forces using the system routinely

Figure 3.8 Construction of famous faces from memory using the EFIT-V system. The images were created during training of police operators. Proceeding from the top row, left to right, the target subjects are Carlos Tevez (footballer), Alex Ferguson (football manager), Eric Cantona (former footballer and now actor), Gordon Brown (former Prime Minister), Bruce Lee (martial arts expert), John Major (former Prime Minister).

rather than through academic studies. At the E-FIT user conference in 2009, West Yorkshire police reported a 40% naming rate over an 18-month period from May 2008 to November 2009 which encompassed more than 1000 interviews. In 2010–11, a naming rate of 55% was reported, covering just under 300 interviews over an 8-month monitoring period. This indicates that when the factors mentioned above are properly considered, performance can be excellent. Functional developments of the system have also been (and will continue) to be driven by the real-life requirements of operators and witnesses but also by advances in our understanding of facial processing and image processing.

Figure 3.8 shows an array of famous faces, all of which were created by trainees on various training courses delivered in 2010. These images were not produced under conditions of strict forensic validity, but were produced from memory and indicate EFIT-V's inherent ability to produce accurate likenesses of the subject. A novel use of EFIT-V has recently been reported by Cheshire Police. An elderly man suffered a fatal cardiac arrest in Davenham, Cheshire but the police were unable to identify him. The forensic services' post-mortem photographs of the man were not considered suitable for public release and likely to cause distress. Accordingly, EFIT-V was used to create a likeness to the subject, a procedure lasting only 1 hour, which was then printed in the local press. The man was subsequently identified from the image shown in Figure 3.9.

Figure 3.9 Construction of an EFIT-V using post-mortem photographs as a target (image courtesy of Cheshire Police).

References

Benson, P. J. and Perrett, D. I. (1993). Extracting prototypical facial images from exemplars. *Perception*, **22**, 257–262.

Brunelli, R. and Mich, O. (1996). SpotIt! An interactive identikit system. *Graphical Models and Image Processing: GMIP*, **58**(5), 399–404.

Burt, D. M. and Perrett, D. I. (1995). Perception of age in adult caucasian male faces: computer graphic manipulation of shape and colour information. *Royal Society Proceedings B*, **259**, 137–143.

Caldwell, C. and Johnston, V. S. (1991). Tracking a criminal suspect through facespace with a genetic algorithm. In *Proceedings of the Fourth International Conference on Genetic Algorithms*, pp. 416–421. Morgan Kaufmann.

Cootes, T. F., Cooper, D., Taylor, C. and Graham, J. (1995). Active shape models – their training and application. *Computer Vision and Image Understanding*, **61**(1), 38–59.

Cootes, T. F., Edwards, G. J. and Taylor, C. J. (1998). Active appearance models. In H. Burkhardt and B. Neumann (Eds.), *Proceedings of the European Conference on Computer Vision*, *Vol. 2*, pp. 484–498. Berlin: Springer.

Cootes, T. F. and Taylor, C. J. (2001). Statistical models of appearance for computer vision. *Technical Report, Imaging Science and Biomedical Engineering*. University of Manchester, UK, September 2001.

Cootes, T. F., Taylor, C. J., Cooper, D. H. and Graham, J. (1992). Training models of shape from sets of examples. In *British Machine Vision Conference*, p. 918.

DiPaolo, S. (2002). *Investigating Face Space*. SIGGRAPH presented paper (sketch).

Frowd, C. D., Hancock, P. J. B., Bruce, V. *et al.* (2010). Giving crime the 'evo': catching criminals using EvoFIT facial composites. In G. Howells, K. Sirlantzis, A. Stoica, T. Huntsberger and A. T. Arslan (Eds.), *2010 IEEE International Conference on Emerging Security Technologies*, pp. 36–43.

Gibson, S. J., Solomon, C. J., Maylin, M. I. S. and Clark, C. (2009). New methodology in facial composite construction: from theory to practice. *International Journal of Electronic Security and Digital Forensics*, **2**, 156–168.

Gibson, S. J., Solomon, C. J. and Pallares-Bejarano, A. (2003). Synthesis of photographic quality facial composites using evolutionary algorithms. In R. Harvey and J. A. Bangham (Eds.),, *British Machine Vision Conference 2003*, *Vol. 1*, pp. 221–230.

Hancock, P. J. B. (2000). Evolving faces from principal components. *Behaviour Research Methods, Instruments and Computers*, **32**(2), 327–333.

Hancock, P. J. B., Burton, A. M. and Bruce, V. (1996). Face processing: human perception and principal components analysis. *Memory and Cognition*, **24**(1), 26–40.

Holland, J. H. (1975). *Adaptation in Natural and Artificial Systems*. Ann Arbor, MI: University of Michigan Press.

Lanitis, A., Taylor, C. and Cootes, T. F. (2002). Toward automatic simulation of aging effects on face images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **24**(4), 442–455.

Pallares-Bejarano, A. (2006). *Evolutionary Algorithms for Facial Composite Synthesis*. PhD thesis, University of Kent.

Penry, J (1971). *Looking at Faces and Remembering Them: A Guide to Facial Identification*. London: Elek Books.

Ramanathan, N. and Chellappa, R. (2005). Face verification across age progression. *IEEE Computer Vision and Pattern Recognition*, **2**, 462–469.

Scandrett, C. M., Solomon, C. J. and Gibson, S. J. (2006). A person-specific rigorous aging model of the human face. *Special Issue Pattern Recognition Letters: Vision for Crime Detection and Prevention*, **27**(15), 1776–1787.

Sims, K. (1991). Artificial evolution for computer graphics. *Computer Graphics*, **25**(4), 319–328.

Sirovich, L. and Kirby, M. (1987). Low dimensional procedure for the characterization of human faces. *Journal of the Optical Society of America*, **4**(3), 519–524.

Solomon, C. J., Gibson, S. J. and Maylin, M. I. S. (2009). A new computational methodology for the construction of forensic, facial composites. *Computational Forensics, Lecture Notes in Computer Science*, **5718/2009**, 67–77.

Stegmann, M. (2000). *Active Appearance Models: Theory, Extensions and Cases*. Master's thesis, Department of Mathematical Modelling, Technical University of Denmark, Lyngby, Denmark.

Tanaka, J. W. and Farah, M. J. (1993). Parts and wholes in face recognition. *Quarterly Journal of Experimental Psychology*, **46A**, 225–245.

Tiddeman, B. P., Burt, M. and Perrett, D. I. (2001). Prototyping and transforming facial textures for perception research. *IEEE Computer Graphics and Applications*, **21**(5), 42–50.

Tredoux, C., Rosenthal, J., Da Costa, L. and Nunez, D. (1999). *Reconstructing Faces with an Eigenface Composite System*. SARMAC III, Boulder, Colorado, July 1999.

Chapter 4

Facial recall and computer composites

Charlie Frowd

4.1 Introduction

Imagine, if you will, that you are sitting quietly outside a café sipping your favourite hot beverage when someone rushes past and snatches your mobile phone, which you left on the table, as you often do. You were able to get a good look at the person's face, albeit for a short time. Your next hour is spent speaking with a police officer giving a description of what happened and what the offender looked like.

It is likely that you could describe accurately what happened. You will probably also be able to describe the perpetrator's build and clothing. There should be no trouble in saying what was the sex of the person and his or her ethnicity; you should be reasonably accurate at estimating the age, height and weight. You could probably remember some details of the person's face.

Incidents such as these are known as 'volume' crime. They occur frequently, often without physical assault to the victim, and their seriousness, at least from a legal perspective, is fairly minor. Due to limitations in police resources, many perpetrators of volume crime are never caught, although time spent locating particularly prolific offenders can be worthwhile.

Crimes involving these repeat offenders, and other crimes of a more serious nature including murder, arson and rape, are generally given higher priority in police investigations. It is in these cases that eyewitnesses (witnesses or victims) may be asked to engage in a range of tasks to assist in the detection and later conviction of the offender. When the police have a *suspect*, they may be asked to take part in an identification parade. (Further details about this are the focus of a separate chapter.) Alternatively, eyewitnesses may be shown photographs of previously arrested criminals for identification, sometimes referred to as *mugshots*. In the absence of a suspect, CCTV footage or other evidence, witnesses may be called upon to externalise an offender's face. The aim is to create a visual image based on remembered information, so that it can be shown to other people for identification. Such images are known as *facial composites* and are seen in the newspapers and on TV, for example, BBC CrimeWatch. The idea is that someone who is familiar with the face will name it to the police and, in doing so, will provide new lines of enquiry.

The focus of the current chapter is on the construction and the recognition of facial composites produced by modern software systems. A separate chapter in this volume details how composites are created by sketch artists. This chapter describes and evaluates typical software programs that the police use to construct faces. It will be demonstrated that the traditional approach used with eyewitnesses is generally ineffective for producing identifiable images, and that alternatives are required if composites are to be effective in the battle against crime. Several successful developments are described. The final section looks to the future and asks what might be on the horizon for producing even more effective faces.

4.2 Introduction to face systems and cognitive processes

The traditional procedure for externalising a face from memory involves a police officer asking the eyewitness to first describe what they remember of the offender's face and then, based on that information, selecting individual facial features from a 'kit' of parts. Features include hair, face shape, eyes, nose, brows, mouth and ears. The resulting image is an assemblage or 'composition' of components, hence the term facial *composite*. Note, however, that this term also refers to

Craniofacial Identification, ed. C. Wilkinson and C. Rynn. Published by Cambridge University Press.

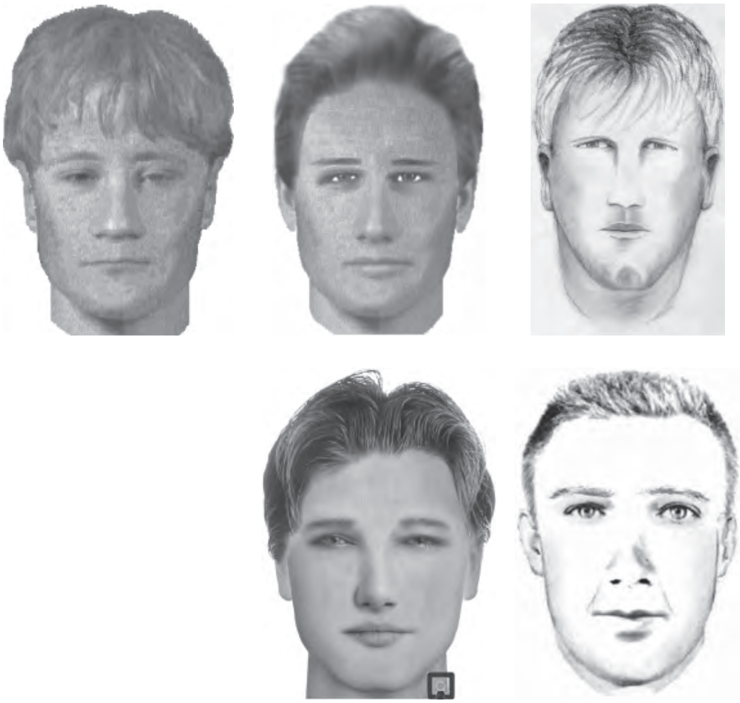


Figure 4.1 Examples composites from (left to right) E-FIT, PRO-fit, Sketch, EvoFIT, FACES 3.0 and Identikit 2000. These were constructed from different people's memories of the actor, Ben Affleck (Frowd *et al.*, 2005*a*, 2007*d*).

images produced by sketch artists, and the newer third-generation systems described below. Examples from various systems are presented in Figure 4.1.

Irrespective of the technique used to synthesise the face, one of the aims of a composite is that someone who is familiar with the face will identify it. Perhaps somewhat unexpectedly, most identifications are made by police officers of repeat offenders. Some, however, emerge as part of a public appeal for information. In either case, with the aim of improving recognition rates, composites are often accompanied by details of both the crime and the offender: age, build, clothing, accent, etc.

Eyewitnesses usually see an offender once, at the time they witness the crime, and so building a composite involves the perception of an *unfamiliar* face. In contrast, identification of the image later by another person involves *familiar* face perception, which is based on an established, long-term visual memory of the face. There is a wealth of evidence to suggest that these two types of perceptual process are very different to each other (e.g. Ellis *et al.*, 1979; Bruce, 1982; Burton *et al.*, 1999); they are even carried out in different hemispheres of the brain (Schooler, 2002). As a consequence, one would expect that aspects of a face that are important for face construction would be different to those that are important for composite identification. This turns out to be the case.

The police ask witnesses to describe, or recall, an offender's face. Next, they see example features matching their description, decide upon the best matches (a second recall exercise), and resize and reposition each feature on the face as appropriate with the aim of creating the best likeness. Face construction in this way, by the selection of individual features, is traditionally thought to be one of information *recall*. In contrast to recall, as described in earlier chapters, face recognition is assumed to be more holistic in nature, emerging from the parallel processing of individual features and their position, or *configuration*, on the face.

The processes of face recognition and face recall are nicely illustrated in a study by Wells and Hryciw (1984). Participants first made 10 judgements about a target face. One group was asked to base these on physical attributes, such as length of nose or thickness of brows, while the other was asked to rate on perceived personality traits such as honesty and intelligence. After this target *encoding* phase, participants either constructed a composite using the American Identikit system, or attempted to identify the target from among six alternative faces. Those who made physical judgements were found to construct a better quality Identikit than those who made personality judgements, but those asked to judge personality traits were more successful at picking out the target face from among alternatives. The research demonstrates that encoding a face by its physical attributes is beneficial when the subsequent task is of a similar nature: whole face encoding, which arises from personality attribution, is best for recognition.

It turns out that the method we use to remember or encode a face is based somewhat on our expectation of what we will be asked to do subsequently. Olsson and Juslin (1999) found that a holistic encoding was preferred by 64% of participants who were unaware of an ensuing memory task, so-called *unintentional* learning, but feature encoding was preferred (62%) following *intentional* learning (Laughery *et al.*, 1986). In conjunction with the above findings of Wells and Hryciw, and others (Frowd *et al.*, 2007*b*), the implication is that composite quality is somewhat dependent on how a face is encoded in the first place. It also divides eyewitnesses into two fairly broad groups. Firstly, there are those who suppose that a description may be required of them, and deliberately use *feature encoding*; they do this by silently creating 'mental notes' about the features of the face. Secondly, there are eyewitnesses who are surprised by the crime, such as the mobile phone theft described above, or domestic distraction burglaries, and encode the face in a more natural, holistic way.

4.3 Facial recall

Irrespective of the method used to encode the face, eyewitnesses who construct composites by selecting individual features are required to give a detailed description of the face, and with good reason. Contained within modern computerised systems are hundreds of examples per feature, far too many for an eyewitness to review in their entirety. The software operative therefore uses the given description to limit the number of features shown. In the PRO-fit system, for example, there are 219 noses in the White-Male database, but a more manageable set of 24 that are classified with a 'short' length and an 'average' width. Similarly, when working with artists who construct sketched composites, witnesses are directed to example features from catalogues of reference materials that match their verbal description.

There are potential problems, however, with using verbal descriptions as part of constructing faces. Ellis *et al.* (1980) found that when matching facial descriptions to target photographs, the longer the time between seeing the perpetrator and giving the description, the more matching errors were made. This indicates that information recall is of less value with time, an established property of human memory (Baddeley, 1990). They also found that the number of descriptors recalled was significantly less after one day, which is of forensic importance since eyewitnesses typically recall faces several days after a crime; for some, this can be even longer.

Laughery and colleagues (1986) similarly asked participants to describe faces from memory. Their data suggested that greater attention is generally given to the upper than the lower part of the face and, more generally, that this proceeds top-down from hair to chin. Participants in general had difficulty in describing facial features and also tended to use relative rather than absolute terms, thereby making comprehension difficult; example adjectives include: *small*, *medium*, *large*; *long*, *wide*, *short*; *broad*, *average*, *narrow*; *dark*, *light*. The authors note that the problem with describing faces from memory may simply be because we do not practice doing this in our everyday lives. The exception was hair: presumably we use this feature to describe other people; we also instruct barbers and hairdressers in our desired coiffure.

The above research findings are worrying as the adjectives used for a given face may vary from witness-to-witness, but are a fundamental component of traditional composite construction. Considerable effort, however, has been spent on improving the ability to recall such information from co-operative eyewitnesses and suspects. The work was carried out initially in the 1980s by Ron Geiselman, Ron Fisher and their colleagues (Geiselman *et al.*, 1986; Fisher *et al.*, 1989), resulting in the *Cognitive Interview* (CI). This is a set of interviewing techniques aimed at eliciting the most complete and accurate recall of information: details of a crime, details of the people involved, descriptions of faces, etc. The CI has been extensively evaluated and revised (see Milne and Bull, 1999; Wells *et al.*, 2007, for reviews; for a meta-analysis, see Köhnken *et al.*, 1999).

One of the principles of the CI is that a complete description of an event, a person's appearance, or a face, is unlikely to be given in a single exhaustive attempt. This can be improved, however, by making several such recall attempts, each of which will result in a different path being taken through our memories and, in doing so, trigger new information. Human memories are not stored in a logical order, much as we would like, but are based on factors such as their relevance or salience, event position, level of attention (arousal), type of encoding, etc. Memories tend to decay, as mentioned above, making access more difficult, but this can be improved by *context reinstatement*: asking eyewitnesses to visualise smells, sounds, personal feelings and the environment (Davies and Thomson, 1988; Vervaeke *et al*., 2002) – all potential cues that can trigger memories.

Research has also shown that recall is reduced under high levels of physical arousal (state anxiety) (Brigham *et al*., 1983; Valentine and Mesout, 2009). This is relevant to when a crime is being witnessed, but also when the person recalling the event is feeling anxious. To overcome this latter issue to some extent, the CI has an initial 'rapport-building' stage to help eyewitnesses relax.

The CI has been further adapted for investigative interviews as part of an interviewing *framework* that UK police officers follow. This is known as the PEACE CI, a mnemonic that defines each stage of the process: Planning and preparation, Engage and explain, Account, Closure and Evaluation. The interview is also combined with further techniques called *Conversational Management* that have the aim of further enhancing recall. See Dando *et al*. (2009) for a recent review.

4.4 Computerised feature systems

Face-production systems emerged in the 1970s, were non-computerised and composed of facial features printed onto rigid card (Photofit) or transparencies (Identikit). Several serious problems were identified with each, including limitations in the range, sizing and placement of facial features (Davies, 1983; Shepherd and Ellis, 1995). Modern software systems have attempted to overcome these deficiencies.

There are now many software products available to build feature composites. Police forces in the UK rely on two such systems, E-FIT and PRO-fit (Frowd *et al*., 2005*b*); those in the US have greater choice: FACES, Identikit 2000, ComPhotofit, CRIMES, Compusketch, CD-FIT, E-FIT and FaceKit (McQuiston-Surrett *et al*., 2006). Each system contains a large collection of individual features, normally electronically 'cut' from photographs of faces, and classified in terms of size, shape, colouring, etc. In the UK, only a sample of features is taken from each photograph, to prevent the subjects' faces from being reconstructed. PRO-fit databases, for example, sample five features per face.

It is normal for composite systems to have a range of databases that span different ethnic backgrounds, gender and age. Most produce composites in greyscale, as our face perception ability is as accurate in this mode as it is in colour (Kemp *et al*., 1996; Davies and Thasen, 2000). Attempts to design a colour system seem to confirm this notion (Davies, 1986; Frowd *et al*., 2006*b*). Some systems, such as Identikit 2000 and Compusketch, present features in a sketch-like format, as illustrated in Figure 4.1.

The aim with these methods is to produce an identifiable *likeness* rather than a facsimile. This is partly due to the face being constructed from memory, but also because databases do not contain all possible combinations of feature shapes, textures, etc. To compensate for the latter deficiency, eyewitnesses are given the opportunity for artwork to be applied to the composite face, to add shading, wrinkles, facial marks, etc. While such enhancements require expertise on the part of the police software operative, it can improve the identifiability of a composite (e.g. Gibling and Bennett, 1994). In a small (unpublished) project carried out by the author, for example, people were asked to name composites produced by Frowd *et al*. (2005*b*), or the same images after removal of artwork, as Figure 4.2 illustrates. People's ability to correctly name the composites halved with the artwork removed.

Figure 4.2 Artistic enhancement: the image on the left was constructed using E-FIT of the pop singer, Noel Gallagher; the final image, after artwork was applied, is shown on the right.

In an attempt to increase the effectiveness of software feature systems, a 'cognitive' approach is the standard method used to build the face. This is based on considerable research suggesting that we are better able to select individual features when they are embedded in a complete face than when we see them as isolated parts (Davies and Christie, 1982; Tanaka and Farah, 1993). For example, a nose is selected with greater accuracy when seen in an intact face than when seen on its own. As we perceive faces as complete entities – that is, we process faces holistically – the cues provided by each feature help trigger other memories and improve feature selection. In modern software systems, witnesses now assess individual features as they are switched in and out of an intact face.

4.5 Composite system performance

How effective are computerised composites? This question could be answered by auditing their effectiveness in criminal investigations, perhaps via the number of arrests or criminal convictions resulting directly from composite images (e.g. Frowd *et al*., 2011). While attractive, the approach suffers from the normal problems associated with field experiments: lack of control. Many factors are likely to affect both the quality of a composite and whether it is correctly identified. Example factors include the encoding of the face, the level of anxiety experienced, the number of people who see the composite (circulation) and the presence of additional information on wanted posters: modus operandi, physical descriptions of an offender, clothing, etc.

System evaluations have therefore focused on the controlled environment of the laboratory to maintain good internal validity (Ellis *et al*., 1975; Koehn and Fisher, 1997; Davies *et al*., 2000; Frowd *et al*., 2005*a*; Brace *et al*., 2006). This was the approach taken by the author: asking the question, 'How well can people construct composites in the laboratory when procedures are used that follow those of 'real' witnesses in the UK?'

Frowd and colleagues (2005*b*) evaluated the performance of five composite systems. These included E-FIT, PRO-fit and Photofit. There was also a sketch artist, who drew the face by hand using pencils, and a system in development called EvoFIT, which is discussed in detail in a separate section below. In the research, participant-witnesses looked at a good quality colour photograph of an unfamiliar target for one minute. They did so in the knowledge that a composite would be required, to encourage a potentially optimal (feature) encoding strategy. Three to four hours later, they met with an experienced composite operator, in this case one of four experimenters, or a police sketch artist, who worked with them to construct a composite. Each person received a cognitive interview (CI), to prompt recall of the face's appearance, were shown appropriate features that matched this description (except EvoFIT) and attempted to produce the best likeness possible in their own time. Systems were used as stipulated by the manufacturers, and the use of artwork techniques was offered to each person to enhance the likeness of the face.

The targets were 10 celebrity faces, half of which had been previously rated by other people as being distinctive in appearance, and the other half as average. Each person made a single composite of one of these targets with one of the five systems. Fifty images were produced in total. See Figure 4.3 for examples.

Figure 4.3 Composites constructed by Frowd *et al*. (2005*b*) using (left to right): E-FIT, PRO-fit, Sketch, Photofit and EvoFIT. Can you name the celebrities? They are listed at the end of the chapter.

Another group of participants evaluated the composites, primarily by naming (but a matching task was also used, as discussed later). Images from E-FIT and PRO-fit were named equivalently with a mean at 18.0%, compared with sketches at 9.2%, Photofit at 6.2% and EvoFIT at 1.5%. Thus, naming was fairly good for the two computerised feature systems currently used in the UK, and both were better than the Photofit system that they had replaced; however, naming from sketch and EvoFIT was surprisingly low. The study found a large effect of target distinctiveness: composites of an unusual face were named about three times more successfully overall than those of a more average appearance. This facial distinctiveness effect is observed more generally, when *recognising* unfamiliar faces (Shapiro and Penrod, 1986), and here suggests that offenders will be much more identifiable from a composite if their face has an unusual appearance.

Other research projects have found a similar naming rate for the software feature systems when the delay-to-construction is short (Davies *et al.*, 2000; Bruce *et al.*, 2002; Frowd *et al.*, 2004, 2007*a*, 2007*b*, 2008*a*; Brace *et al.*, 2006). A different story emerges, however, when the target delay is much longer. Frowd and colleagues (2005*a*) followed the same basic design, with the same operators controlling E-FIT, PRO-fit and sketches, but the delay was two days, which is the norm for real witnesses. Examples are presented in Figure 4.1. This time, participant-witnesses made composites that were correctly named at only 3.2% overall, with sketches emerging as the best method, at 8.1%. A sorting task was used as an alternative (proxy) to naming, involving additional participants matching composites to target photographs. There is a tendency when completing the task to compare individual features between composites and targets, and so it provides a broad measure of *feature* quality in the composites. Based on the rate of successful matching, sketches also emerged as the best method of face construction. Taken together, the Frowd *et al.* studies suggest that while sketching is unable to produce composites as effectively as the software feature systems when the delay is short, it is more effective when the memory is older and weaker.

The above research employed faces of celebrities as targets, which may not be representative of faces of criminals. Follow-up work using a similar design, including a long delay but non-famous targets, found that while a good-quality face was produced occasionally, naming rates remained low overall for the computerised feature systems (Frowd *et al.*, 2005*c*, 2007*b*, 2010). An evaluation of other methods – FACES 3.0 and Identikit 2000 – has been carried out, with similarly disappointing results (Frowd *et al.*, 2007*d*). Other researchers tell a similar story (Koehn and Fisher, 1997; Kovera *et al.*, 1997).

It is perhaps worth mentioning that the participant-witnesses in the above projects were generally given an excellent opportunity to remember the face, engaged in techniques believed to construct the face in the best way, and yet their composites were rarely identified successfully by other people. Such a result is worrying since these techniques are used to detect offenders. What then appears to be the problem?

4.6 What could be going wrong?

Frowd *et al.* (2007*a*) evaluated the quality of the internal and external features of composites constructed in Frowd *et al.* (2005*a*) – i.e. those produced after a two-day delay. Examples are presented in Figure 4.4. Two proxies to naming were used, including the sorting

Figure 4.4 Internal features, external features and complete composite of the actor, Ben Affleck, used in Frowd *et al.* (2007*a*).

task described above, with participants inspecting composites of internal features, external features or unaltered images. The study revealed similar matching for external and complete composites, but both were superior to internal composites. A follow-up experiment replicated this internal composite *disadvantage*; it also indicated that hair was an important exterior feature.

The study indicates a general ineffectiveness of composite systems in constructing the internal features, but it is this region of the face that is important for later recognition by another person. Also, that there is an emphasis on the exterior part, in particular the hair, during construction. These findings are consistent with the above research, indicating a bias towards verbal recall for the exterior region. Further research by Frowd and Hepton (2009), which focused on the EvoFIT system described below, has also demonstrated the importance of the internal features in a recognition (rather than a matching) task, a forensically more valid measure. This work has also shown that the amount of *identifying* information in the external features is rather low, as is the case with photographic (non-composite) faces (Ellis *et al.*, 1979).

There is a more fundamental problem: the tasks of face description and feature selection are simply contrary to the way faces are seen, as wholes. This observation dates back over 30 years, and was noted long before the emergence of the modern systems (Davies *et al.*, 1978). It is clear that advances have been made, since a more identifiable face can be constructed using today's technology (Frowd *et al.*, 2005*b*), but performance remains poor when construction take place following long retention intervals (as is the case with witnesses to crime).

There is a hint from sketch production about how to improve the software systems. While sketching is still based on the selection of individual features, the initial focus is on *configural* information – the placement of features on the face – and then on increasing the detail in groups of features. This procedure would appear to encourage more natural, holistic face perception (Davies and Little, 1990; Frowd *et al*, 2005*a*), an approach similar to modern software systems that require feature selection in the context of a complete face. In fact, a holistic approach to face production has been successfully applied to each stage of the process: to the initial interview, system and finished composite. Some of these are discussed below; each is in UK police use.

4.7 Improving computer composites

4.7.1 Morphing

Bruce and colleagues (2002) carried out one of the first projects to successfully improve the performance of a modern face system. The research mirrored the practical situation where an offender had been seen by multiple eyewitnesses, and tried to answer the question as to which of these observers should construct a composite. The problem is that no test currently exists to reliably predict who would produce the best quality image. Their solution was to ask each person to construct a face, and then to average the individual attempts into a single 'morphed' image. The authors argued that as the composites were created from different people's memories, any errors therein would not be correlated and so would cancel in the production of a morph. They demonstrated that the morphed composite was more identifiable than the average individual image, and was sometimes better than the best individual instance.

As a result of their work, the UK police guidelines on facial identification have been modified to permit construction of multiple composites of the same offender, for the purpose of producing a morph for public appeals (ACPO, 2009). Alternative approaches to morphing are required, however, for situations involving a single eyewitness. While it is possible for an observer to create multiple composites of the same offender (see Frowd *et al.*, 2006*c*, for such a case) the norm is to create a single image. Alternative approaches are presented below for this more general case.

4.7.2 Interviewing

As mentioned above, considerable effort has improved the effectiveness of the CI, to recover the most accurate and complete description of a crime, a criminal's face, etc. The interview concerns information recall, which is why it is of value at the start of face construction (Frowd *et al.*, 2005*c*): to enable subsets of features to be located within a composite system. Next, witnesses identify the best matching features. When doing this, they are also engaging in *face recognition* processes. Improving these processes should therefore improve the accuracy of feature selection and the overall identifiability of a composite.

An established procedure to improve unfamiliar face recognition is to attribute holistic judgements at encoding, as described above (Wells and Hryciw, 1984). Another is to make these judgements prior to a recognition attempt (Berman and Cutler, 1998). This latter method was developed into a 'holistic' CI, or H-CI, for face construction (Frowd *et al*., 2005*c*, 2007*b*, 2008*a*). The procedure commences as normal with a CI. Next, witnesses are asked to think about the personality of the face silently for one minute and make seven whole-face judgements (e.g. intelligence, friendliness, honesty, pleasantness, athleticism, trustworthiness and distinctiveness) about the face, rating each on a three-point Likert scale (low/medium/high) before constructing the face as normal.

Frowd and colleagues (2008*a*) made composites following a CI, which were correctly named at 9%, while those created after an H-CI were very much better named, at 41%. The research indicated that specific holistic judgements do not appear to be important for the interview to be effective, allowing police operatives to select items appropriate for each investigation. It would be inappropriate, for example, to ask victims of sexual assault about the 'friendliness' of an offender's face, but this may be acceptable for confidence crimes. A list of suitable adjectives may be found online at HCI (2009).

4.7.3 Caricature

One reliable finding in the literature is that faces of an unusual appearance are recognised more accurately than more average faces (Shapiro and Penrod, 1986). As mentioned above, such target distinctiveness effects extend to face construction for both manual and computerised systems (Frowd *et al*., 2005*b*). Frowd *et al*. (2007*c*) argued that composites tend to be quite bland in appearance and so their recognition might be improved by artificially inflating the level of distinctiveness. While artists do this effectively, to produce very recognisable renditions, caricaturing involves considerable skill and is somewhat idiosyncratic (Goldman and Hagen, 1978). Various commercial software programs, however, are now available that produce more consistent results. The PRO-fit system itself, for example, includes such a utility. Each works by comparing features' shapes and relationships in a facial image – be it a photograph of a face or a composite – with respect to an average face (Benson and Perrett, 1991) and then exaggerating any differences to produce a positive caricature: they can also de-emphasise differences to produce a representation that is more similar to the reference, a negative caricature.

In a series of experiments, a fixed level of positive caricature, such as any of those illustrated in Figure 4.5, only slightly improved participants' ability to identify the composite face (Frowd *et al*., 2007*c*). In follow-up work, another group of people adjusted the level of caricature, positive to negative, to produce the most identifiable image. There were two surprising results. Firstly was a general preference for a *negative* caricature for the three types of systems tested. This appears to have emerged as the composites contained errors, which became reduced, so rendering the image more acceptable, when the face was made to appear more average. The outcome is somewhat similar to a morphed composite: a reduction in error and a face that better resembles the intended target. Secondly, there were large individual differences in preferences: some people preferred a moderate positive caricature, while others a slight negative one.

Figure 4.5 A composite of the former UK Prime Minister Tony Blair that has been progressively caricatured as part of Frowd and colleagues (2007c): −50%, −25%, 0% (veridical), +25% and +50% caricature. Naming of the target identity was found to substantially improve when participants saw a sequence of 21 such images.

These results led to an important finding: naming increased by about 50% overall when the composite was seen to change in small (5%) steps from –50% to +50% caricature. Also, the sequences were effective for the three types of system tested, but the best improvement emerged for images that were poorly named initially – typical of those produced in criminal investigations. More specifically, 13 of the composites were poorly named, with a mean between 0 and 10%, an overall mean of 3.0%, but this substantially rose to 31.2% when people were presented with the 21 frame sequence. The sequences therefore increased correct identification ten-fold!

The most convenient format for publishing these composite sequences is via an animated GIF image; they can be used on wanted person's web pages, or on TV. An example is available for viewing online at www.EvoFIT.co.uk. More recent research (Frowd *et al.* under revision-*a*) confirms that the sequence is mainly effective over the positive caricature range: that is, by enhancing facial distinctiveness. This result (along with other findings presented in the paper) suggests that we tend to create composites that are anti-caricatures.

4.7.4 System

The above techniques all suffer from one problem: face construction is problematic if an eyewitness is unable to recall an offender's face in detail. Without good recall, a subset of individual features cannot be identified and so there will be too many items for a witness to inspect. In the UK, for this reason, about 70% of eyewitnesses are denied the opportunity of constructing a composite from a feature system or sketch.

One obvious way forward is to break the dependence on verbal descriptions. However, if a single face is still used, eyewitnesses may still engage in recall to comment upon the accuracy, size and position of individual features. The solution taken by several system designers has been to present multiple faces concurrently and ask witnesses to base selections on the *overall* appearance. The face construction task should then become one of face recognition. Doing so should have the advantage of more stable performance over time, compared with systems based on face recall (Ellis *et al.*, 1980; Davies, 1983; Shepherd, 1983), offering the possibility of accurate face construction even from weak memories.

Several third-generation or 'recognition' systems have now emerged that are premised in this way. In the UK, there are EvoFIT (Frowd *et al.*, 2004) and EFIT-V (Gibson *et al.*, 2003), and in South Africa, ID (Tredoux *et al.*, 2006); there is also a German system in development (Blanz *et al.*, 2006). The systems have a software module that is able to generate a large number of faces, each with a specified set of face coefficients. The basic approach is to present witnesses with screens of randomly generated faces to select. Selected items are 'bred' together, by combining the underlying coefficients, to produce more items for selection. Repeating this procedure a few more times allows the face set to become more similar to each other and more similar to the offender's face in the memory of the eyewitness. The best likeness is ultimately taken as the 'composite'. The underlying mechanism for generating faces using PCA models, and the typical method for combining items using an evolutionary algorithm (EA), is the subject of Chapter 3.

All composite systems are essentially engaged in a search problem: to locate a specific identity. For the recognition types, the search is within a high-dimensional face 'space' and each presented face represents a potential 'solution'. In principle, the more faces on which witnesses give feedback, the more thorough the search will be and the greater the chance will be of locating an identifiable face. (See Frowd *et al.* 2004, for an evaluation of 'population' size to support this idea.) However, system designers need to be careful to allow convergence on an appropriate likeness before witnesses suffer fatigue.

In the following section, EFIT-V is outlined, followed by EvoFIT and the developments that have been necessary to render it effective. As with the feature types, the software is controlled by experienced personnel.

4.7.5 System: EFIT-V

In the EFIT-V system, a database of race and gender is first selected, to reflect the background of the offender as remembered by the eyewitness. The system is flexible in use, but witnesses generally start by selecting an appropriate facial shape – round, square, oval and so forth – and a hairstyle, before being presented with screens of nine faces that change in appearance by both 'shape' (feature shapes and placement on the face) and 'texture' (pixel colourings of the eyes, brows, mouth, etc.). Witnesses select examples that resemble an offender, or reject others, and an EA breeds the relevant items together. The process is repetitive, and there are software tools for manipulating individual features and whole-face attributes such as age and masculinity. A paint program is available to add lines, wrinkles, shading, etc. See Chapter 3 for a more detailed description of EFIT-V.

4.7.6 System: EvoFIT

To construct an EvoFIT, an appropriate database for age, gender and race is first selected. Witnesses then choose an appropriate set of external features, in particular hair, and are presented with arrays of complete faces, 18 per page. During the early stages of development, users preferred to view facial shape and facial texture separately (Frowd *et al.*, 2004) and so the interface was developed to present information in this way. In practice, they are shown four screens of shape and select the two best likenesses per page up to a maximum of six, then similarly for texture. Also, to assist with selecting the latter, the textures themselves are presented on a specified face shape, one that the user believes to be the best shape at this stage. To assist in the conversion of a good likeness, witnesses next choose from the best combination of selected shape and texture, and identify a so-called 'best' face, which is subsequently given twice the number of breeding opportunities in the EA. EvoFIT tended to converge on a good facial type after three breeding cycles – the initial generation plus two breeding cycles.

Frowd *et al.* (2007*b*) evaluated the effectiveness of this version of EvoFIT using the 'gold standard' described above: unfamiliar targets to the people constructing the composites, a two-day delay, use of a CI, etc. Under target feature encoding, composites were named better from EvoFIT (11%) than from a typical modern 'feature' system (4%). While this represents an improvement in identifiability relative to the traditional, overall performance was not impressive.

Two further developments have been effective. The aim of the first was to address the issue raised previously: the external features have high importance for unfamiliar face perception – composite construction here – but it is the internal features that are important for recognition – composite naming. In EvoFIT, a Gaussian or

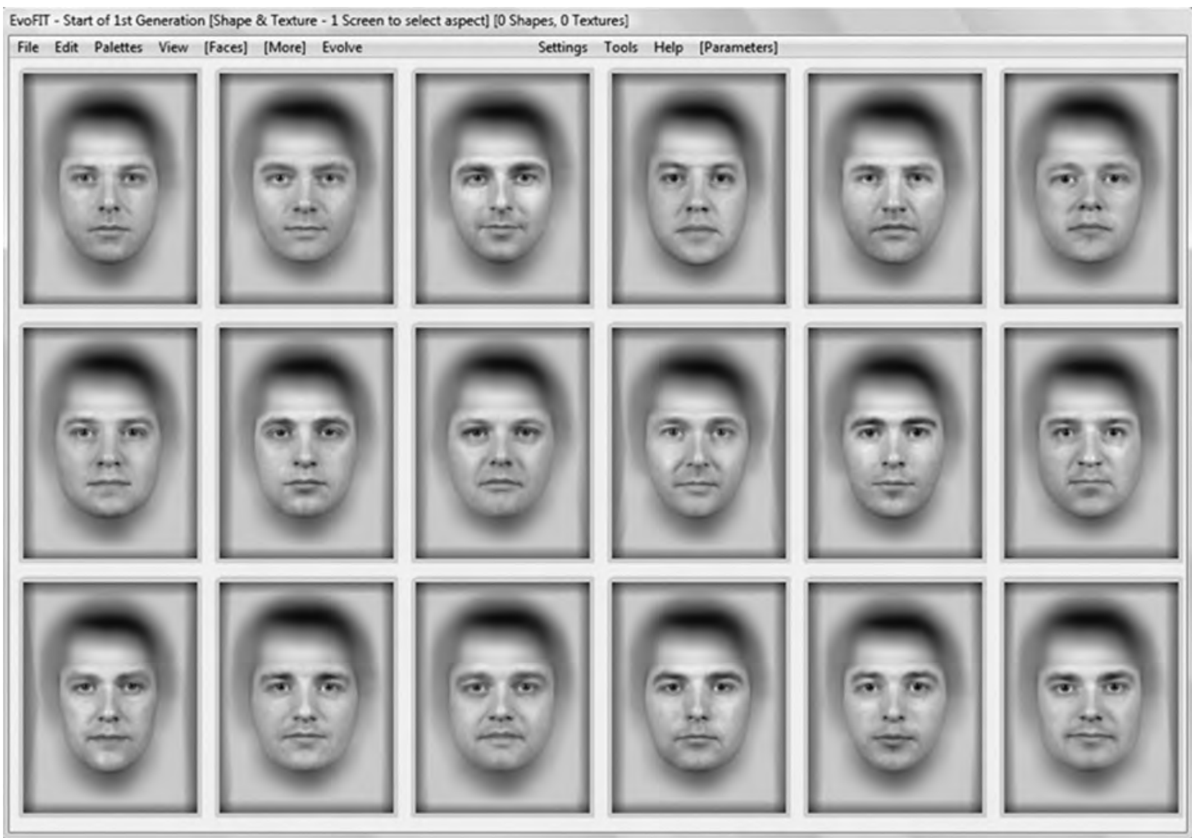


Figure 4.6 An example EvoFIT screen. The external features of the faces are 'blurred', to allow eyewitnesses to focus on the central part of the face. At the end of evolving, the blur is removed to allow the final image to be clearly seen.

'blur' filter is applied to the external features after they had been chosen. This image filtering allows selections to be based on the central part of the face, as Figure 4.6 illustrates. After evolving, blurring is disabled. Two projects have shown that the selective blurring of faces in this way promotes a more identifiable composite when the target delay is short (Frowd *et al.*, 2008*b*) and long (Frowd *et al.*, 2010).

The second development allowed better convergence of age and other holistic properties of a face. This was achieved in part by limiting the age capability of the face generators, specifically by building (PCA) models from faces of a specified age range. There are now separate models built from people in their 20s, 30s, etc. However, an incorrect aged face was still produced sometimes (although the age match was now closer than before). The basic problem here is the complexity of the task: witnesses may select faces that are accurate in one aspect (e.g. honest-looking) but not another (attractiveness), and these choices will be reflected with evolution. This issue was addressed by allowing users to change their evolved face along a number of psychologically sensible 'holistic' dimensions. These include age, face weight, masculinity, attractiveness, threatening and honesty. The method used to create these scales is described in Frowd *et al.* (2006*a*); an example manipulation is shown in Figure 4.7.

Figure 4.7 The face on the left was evolved of the footballer David Beckham from memory in Frowd *et al.* (2006*a*); on the right, after holistic tool use. In this case, the perceived level of health and attractiveness were increased.

Example composites evolved with EvoFIT and these two developments are presented in Figure 4.8. Using the gold standard construction procedure, which is described above, including a two-day delay, the system's effectiveness was recently assessed (Frowd *et al.*, 2010). The enhanced software was used, except that witnesses also selected a facial *aspect ratio*, an appropriate face width and height, after choosing the hair. Selecting a facial aspect ratio allows items to be shown with the ratios constrained, with the aim of helping face selection and the production of an identifiable likeness. Second, due to improvements brought about by blurring, holistic tools and facial aspect ratios, two rather than three breeding cycles were used. This has the obvious advantage of presenting witnesses with fewer faces to choose from. Results of the evaluation were that blurring and holistic scales were effective on their own, but best performance emerged when used in conjunction with each other: mean naming was 25% correct from EvoFIT and 5% from a 'feature' system.

EvoFIT has turned out to be something of an enigma, and attempts to improve its effectiveness have not always been successful (Frowd *et al.*, 2007*b*, 2008*b*). While the third-generation systems are supposed to be based on face recognition rather than recall, this idea is too simplistic (at least there is evidence of this from the development of EvoFIT). For example, one might expect that the holistic component of the holistic CI (H-CI) would be effective on its own, to improve a user's face recognition ability and thus his or her accuracy in selecting whole faces from the presented arrays. In recent work (Frowd *et al.*, under revision-*b*), this holistic-attribution component actually promoted worse quality composites than the face recall part of the CI; curiously,

Figure 4.8 Example EvoFITs constructed using the system plus recent developments. The celebrity identities shown are listed at the end of the chapter.

correct naming of EvoFITs from H-CI (40%) was superior to those from CI (25%). Face recall therefore allows selections to be made with more accurate features, which is beneficial to the evolution process in the long run. But, detailed descriptions produce over-emphasis on individual features, and so holistic attribution after the CI provides a shift towards whole-face selection. That said, there is a twist in the tale, as a previous EvoFIT procedure with feature manipulations made early on gave inferior results for the H-CI. So, the type of interview administered and the procedure used to evolve the face have a somewhat complex relationship. (See also Frowd *et al.*, 2011, for the value of interviewing techniques in police field-trials as well as evidence that EvoFIT leads to the arrest of a suspect in about 40% of criminal cases.)

The performance of EvoFIT continues to improve. Most recently, Frowd *et al.* (in press-*a*) found that while blurring external features in the face arrays was a step in the right direction (as illustrated in Figure 4.6), the mere presence of the exterior context remains a distraction to the person constructing the face. In the most recent incarnation of EvoFIT, the external features are masked until the internal features have been constructed in their entirety. Relative to external-features blurring, this technique doubled naming rates to 45% using the gold-standard construction procedure.

There is also recent evidence that when we recognise a finished composite, for example when seeing it in the newspapers, we process an image from EvoFIT more holistically, and with a greater contribution from the internal features of the face, than we do for a composite produced from a feature system (Frowd *et al.*, in press). These findings highlight parallels with recognising facial composites from EvoFIT and recognising faces in general.

4.8 The future

One might ask whether facial composites really have a role to play in identifying offenders. About five years ago, indications were that modern computer feature systems consistently failed to produce good-quality images when the target delay was fairly long (Frowd *et al.*, 2005*a*). This is particularly worrying as many police forces rely on this type of technology to detect offenders. It is interesting to note the historical pattern that has emerged with composite systems: a system is produced, adopted by police forces and then found to be ineffective. This was the case with the non-computerised systems, Photofit and Identikit, and likewise with their modern computerised decedents. When developing EvoFIT, we were keen to avoid making the same mistake; about 10 years of intensive development were required to reach satisfactory performance (Frowd *et al.*, 2010).

At least one of the above developments has general benefits: caricature animation is not only effective for feature systems, but also for sketches, which the evidence would suggest are hard to recognise even when the target delay is short (Frowd *et al.*, 2005*b*); for EvoFIT, indications are that the animation boosts mean naming levels by a further 15%. In addition, while the H-CI is beneficial to the latest EvoFIT procedure, and perhaps to all recognition-based systems, it will be interesting to see whether the new interview is of value to sketch artists; indications are that it should be, given that sketching involves the selection of individual features, which should similarly be improved following improvements to an observer's face recognition ability. It is also possible that the H-CI may be even more effective here due to the inherently holistic nature of sketch production (Davies and Little, 1990).

Sketches tend to be qualitatively different to the other systems in one respect: they contain less shading information – see Figures 4.1 and 4.3. This would appear to result from our inability to recall the texture of the face in sufficient detail for accurate rendering on the page, and therefore major regions are often left blank, or with minor shading. One curious possibility is that a sketched face may, in some circumstances, be more accurate overall by virtue of there being less incorrect information!

In a small project, the effect of reducing visual detail was explored (Frowd *et al.*, 2008*b*). This was done by simply increasing the brightness level in a set of composites. Results found significantly better naming for such enhanced images relative to veridical. What this demonstrates is that information reduction can be useful, as some of the inaccurate information will be removed, to allow a perceiver's cognitive system to 'fill in the gaps'. Examples are presented in Figure 4.9. While a less detailed representation does not help face *construction* by individual features (Frowd *et al.*, 2005*a*, 2007*d*), it does for EvoFIT (Frowd *et al.*, 2008*b*). Indeed, a database of this type may be useful in situations where the race of the face is unfamiliar to the eyewitness and recognition abilities are challenged even further (Meissner and Brigham, 2001). It may also be valuable for observers who have viewed an offender with *unintentional* encoding, perhaps due to the sudden nature of the crime, and so have fairly limited detail of the face on which to draw.

Figure 4.9 Reducing the detail of a composite by increasing brightness levels. The left pair (a, b) are of Tony Blair, while the right pair (c, d) are of the actor Nicholas Cage.

Software designers are still producing a 'one-size fits all' solution: a single system for all eyewitnesses. This simply may not be optimal. While it is normal for witnesses to produce a poor likeness using feature selection following long delays, they sometimes produce a very good one – Figure 4.9's PRO-fit of Nicholas Cage is a case in point; the same idea applies to sketched composites (Frowd *et al.*, 2005*a*). Perhaps one of the challenges facing psychologists today is to understand individual differences between observers for the purpose of matching them to the face construction technique. This turns out to be a particularly hard task, one that I myself have attempted. Given the role that face recall and recognition play in face construction, which also appears to change according to the system used, perhaps a combined recall-recognition diagnostic test might be a fruitful avenue of research.

4.9 Concluding comments

About five years ago, face construction by the selection of individual features was shown to be a generally ineffective method, especially when involving long delays. Considerable research effort, spanning some three decades, has led to important developments to the system, interview and image presentation format. It is now possible that a face built under forensically relevant conditions is well recognised and thus of value to law enforcement, whether this is from a modern feature system or from the newer EvoFIT (and perhaps even one of the other recognition types). The future for facial compositing is promising: there is clearly headroom for improvement and some interesting new avenues for research and development.

Answers

The identities shown in Figure 4.3 are of Brad Pitt (actor), Robbie Williams (musician), David Beckham (footballer), Noel Gallagher (musician) and Michael Owen (footballer). In Figure 4.8, they are of Simon Cowell (music manager), George W. Bush (former US president), David Tennant (actor) and Noel Gallagher (musician).

Acknowledgement

The author would like to thank Dr Chris Atherton at the University of Central Lancashire, Preston, for her insightful comments on a draft version, and Emma Walker for proofreading a near-final version.

References

ACPO (2009). *Facial Identification Guidance*. Produced by the National Policing Improvement Agency. Unpublished document.

Baddeley, A. (1990). *Human Memory: Theory and Practice*. London: Lawrence Erlbaum Associates.

Benson, P. J. and Perrett, D. I. (1991). Perception and recognition of photographic quality caricatures: Implications for the recognition of natural images. *European Journal of Cognitive Psychology*, **3**, 105–135.

Berman, G. L. and Cutler, B. L. (1998). The influence of processing instructions at encoding and retrieval on face recognition accuracy. *Psychology, Crime and Law*, **4**, 89–106.

Blanz, V., Albrecht, I., Haber, J. and Seidel, H. P. (2006). Creating face models from vague mental images. *Eurographics*, **25**. Electronic document.

Brace, N., Pike, G. E., Allen, P. and Kemp, R. (2006). Identifying composites of famous faces: investigating memory, language and system issues. *Psychology, Crime and Law*, **12**, 351–366.

Brigham, J. C., Maass, A., Martinez, D. and Whittenberge, G. (1983). The effect of arousal on facial recognition. *Basic and Applied Social Psychology*, **4**, 279–293.

Bruce, V. (1982). Changing faces: visual and non-visual coding processes in face recognition. *British Journal of Psychology*, **73**, 105–116.

Bruce, V., Ness, H., Hancock, P. J. B, Newman, C. and Rarity, J. (2002). Four heads are better than one. Combining face composites yields improvements in face likeness. *Journal of Applied Psychology*, **87**, 894–902.

Burton, A. M., Wilson, S., Cowan, M. and Bruce, V. (1999). Face recognition in poor quality video: evidence from security surveillance. *Psychological Science*, **10**, 243–248.

Dando, C. J., Wilcock, R. and Milne, R. (2009). The Cognitive Interview: The efficacy of a modified mental reinstatement of context procedure for frontline police investigators. *Applied Cognitive Psychology*, **15**, 679–696.

Davies, G. M. (1983). Forensic face recall: the role of visual and verbal information. In S. M. A. Lloyd-Bostock and B. R. Clifford (Eds.), *Evaluating Witness Evidence*, Ch. 6. Chichester: John Wiley and Sons.

Davies, G. M. (1986). Recall and reconstruction of faces; Implications for theory and practice. In H. D. Ellis, M. Jeaves, F. Newcome and A. Young (Eds.), *Aspects of Face Processing*, Dordrecht: Martinus Nijhoff.

Davies, G. M. and Christie, D. (1982). Face recall: an examination of some factors limiting composite production accuracy. *Journal of Applied Psychology*, **67**, 103–109.

Davies, G. M., and Little, M. (1990). Drawing on memory: exploring the expertise of a police artist. *Medical Science and the Law*, **30**, 345–354.

Davies, G. M., Shepherd, J. and Ellis, H. (1978). Remembering faces: acknowledging our limitations. *Journal of Forensic Science*, **18**, 19–24.

Davies, G. M. and Thasen, S. (2000). Closed-circuit television: How effective an identification aid? *British Journal of Psychology*, **91**, 411–426.

Davies, G. M. and Thomson, D. M. (1988). *Memory in Context: Context in memory*. Chichester: John Wiley and Sons.

Davies, G. M., van der Willik, P. and Morrison, L. J. (2000). Facial composite production: a comparison of mechanical and computer-driven systems. *Journal of Applied Psychology*, **85**, 119–124.

Ellis, H. D. Shepherd, J. W. and Davies, G. M. (1975). Use of photo-fit for recalling faces, *British Journal of Psychology*, **66**, 29–37.

Ellis, H. D., Shepherd, J. W. and Davies, G. M. (1979). Identification of familiar and unfamiliar faces from internal and external features: some implications for theories of face recognition. *Perception*, **8**, 431–439.

Ellis, H. D., Shepherd, J. W. and Davies, G. M. (1980). The deterioration of verbal descriptions of faces over different delay intervals. *Journal of Police Science and Administration*, **8**, 101–106.

Fisher, R. P., Geiselman, R. E. and Amador, M. (1989). Field test of the cognitive interview: enhancing the recollection of actual victims and witnesses of crime. *Journal of Applied Psychology*, **74**, 722–727.

Frowd, C. D., Atherton, C., Skelton, F. C. *et al.* (under revision-*a*). Understanding the animated-caricature advantage for facial composite images. *Visual Cognition*.

Frowd, C. D., Bruce, V., McIntyre, A. and Hancock, P. J. B. (2007*a*). The relative importance of external and internal features of facial composites. *British Journal of Psychology*, **98**, 61–77.

Frowd, C. D., Bruce, V., McIntyre, A. *et al.* (2006*a*). Implementing holistic dimensions for a facial composite system. *Journal of Multimedia*, **1**, 42–51.

Frowd, C. D., Bruce, V., Ness, H. *et al.* (2007*b*). Parallel approaches to composite production. *Ergonomics*, **50**, 562–585.

Frowd, C. D., Bruce, V., Plenderleith, Y. and Hancock, P. J. B. (2006*b*). Improving target identification using pairs of composite faces constructed by the same person. *IEE Conference on Crime and Security*, pp. 386–395. London: IET.

Frowd, C. D., Bruce, V., Ross, D., McIntyre, A. and Hancock, P. J. B. (2007*c*). An application of caricature: how to improve the recognition of facial composites. *Visual Cognition*, **15**, 954–984.

Frowd, C. D., Bruce, V., Smith, A. and Hancock, P. J. B. (2008*a*). Improving the quality of facial composites using a holistic cognitive interview. *Journal of Experimental Psychology: Applied*, **14**, 276–287.

Frowd, C. D., Bruce, V., Storås, K., Spick, P. and Hancock, P. J. B. (2006*c*). An evaluation of morphed composites constructed in a criminal investigation. *Proceedings of the 16th Conference of the European Association of Psychology and Law*, pp. 59–66. London: IP-PA Publishing.

Frowd, C. D., Carson, D., Ness, H. *et al.* (2005*a*). Contemporary Composite Techniques: the impact of a forensically-relevant target delay. *Legal and Criminological Psychology*, **10**, 63–81.

Frowd, C. D., Carson, D., Ness, H. *et al.* (2005*b*). A forensically valid comparison of facial composite systems. *Psychology, Crime and Law*, **11**, 33–52.

Frowd, C. D., Hancock, P. J. B., Bruce, V. *et al.* (2011). Catching more offenders with EvoFIT facial composites: lab research and police field trials. *Global Journal of Human Social Science*, **11**, 46–58.

Frowd, C. D., Hancock, P. J. B. and Carson, D. (2004). EvoFIT: a holistic, evolutionary facial imaging technique for creating composites. *ACM Transactions on Applied Psychology* (*TAP*), **1**, 1–21.

Frowd, C. D. and Hepton, G. (2009). The benefit of hair for the construction of facial composite images. *British Journal of Forensic Practice*, **11**, 15–25.

Frowd, C. D., McQuiston-Surrett, D., Anandaciva, S., Ireland, C. E. and Hancock, P. J. B. (2007*d*). An evaluation of US systems for facial composite production. *Ergonomics*, **50**, 1987–1998.

Frowd, C. D., McQuiston-Surrett, D., Kirkland, I. and Hancock, P. J. B. (2005*c*). The process of facial composite production. In A. Czerederecka, T. Jaskiewicz-Obydzinska, R. Roesch and J. Wojcikiewicz (Eds.), *Forensic Psychology and Law*, pp. 140–152. Krakow: Institute of Forensic Research Publishers.

Frowd, C. D., Nelson, L., Skelton, F. *et al.* (under revision-*b*) Interviewing techniques for Darwinian facial composite systems. *Journal of Experimental Psychology: Applied*.

Frowd, C. D., Park., J., McIntyre, A. *et al.* (2008*b*). Effecting an improvement to the fitness function. How to evolve a more identifiable face. In A. Stoica, T. Arslan, D. Howard, T. Higuchi, and A. El-Rayis (Eds.), *2008 ECSIS Symposium on Bio-inspired, Learning, and Intelligent Systems for Security*, pp. 3–10. NJ: CPS. (Edinburgh).

Frowd, C. D., Pitchford, M., Bruce, V. *et al.* (2010). The psychology of face construction: giving evolution a helping hand. *Applied Cognitive Psychology*, **25**, 195–203.

Frowd, C. D., Skelton, F., Atherton, C. *et al.* (in press-*a*). Recovering faces from memory: the distracting influence of external facial features. *Journal of Experimental Psychology: Applied*.

Frowd, C. D., Skelton, F., Butt, N., Hassan, A. and Fields, S. (in press-*b*). Familiarity effects in the construction of facial-composite images using modern software systems. *Ergonomics*.

Geiselman, R. E., Fisher, R. P., MacKinnon, D. P. and Holland, H. L. (1986). Eyewitness memory enhancement with the cognitive interview. *American Journal of Psychology*, **99**, 385–401.

Gibling, F. and Bennett, P. (1994). Artistic enhancement in the production of photofit likeness: an examination of its effectiveness in leading to suspect identification. *Psychology, Crime and Law*, **1**, 93–100.

Gibson, S. J., Solomon, C. J. and Pallares-Bejarano, A. (2003). Synthesis of photographic quality facial composites using evolutionary algorithms. In R. Harvey and J. A. Bangham (Eds.), *Proceedings of the British Machine Vision Conference*, pp. 221–230.

Goldman, M. and Hagen, M. A. (1978). The forms of caricature: physiognomy and political bias. *Studies in the Anthropology of Visual Communication*, **5**, 30–36.

HCI (2009). Holistic-Cognitive Interview. www.uclan.ac.uk/scitech/psychology/research/evofit/holistic_cognitive_interview.php

Kemp, R., Pike, G., White, P. and Musselman, A. (1996). Perception and recognition of normal and negative faces: the role of shape from shading and pigmentation cues. *Perception*, **25**, 37–52.

Koehn, C. E. and Fisher, R. P. (1997). Constructing facial composites with the Mac-a-Mug Pro system. *Psychology, Crime and Law*, **3**, 215–224.

Köhnken, G., Milne, R., Memon, A. and Bull, R. (1999). The cognitive interview: a meta-analysis. *Psychology, Crime and Law*, **5**, 3–27.

Kovera, M. B., Penrod, S. D., Pappas, C. and Thill, D. L. (1997). Identification of computer generated facial composites. *Journal of Applied Psychology*, **82**, 235–246.

Laughery, K. R., Duval, C. and Wogalter, M. S. (1986). Dynamics of facial recall. In H. D. Ellis, M. A. Jeeves, F. Newcombe and A. Young (Eds.), *Aspects of Face Processing*, pp. 373–387. Dordrecht: Martinus Nijhoff.

McQuiston-Surrett, D., Topp, L. D. and Malpass, R. S. (2006). Use of facial composite systems in U.S. law enforcement agencies. *Psychology, Crime and Law*, **12**, 505–517.

Meissner, C. A. and Brigham, J. C. (2001). Thirty years of investigating the own-race bias in memory for faces: a meta-analytic review. *Psychology, Public Policy, and Law*, **7**, 3–35.

Milne, R. and Bull, R. (1999). *Investigative Interviewing: Psychology and Practice*. Chichester: John Wiley and Sons.

Olsson, N. and Juslin, P. (1999). Can self-reported encoding strategy and recognition skill be diagnostic of performance in eyewitness identification? *Journal of Applied Psychology*, **84**, 42–49.

Schooler, J. W. (2002). Verbalization produces a transfer inappropriate processing shift. *Applied Cognitive Psychology*, **16**, 989–997.

Shapiro, P. N. and Penrod, S. D. (1986). Meta-analysis of facial identification rates. *Psychological Bulletin*, **100**, 139–156.

Shepherd, J. W. (1983). Identification after long delays. In S. M. A. Lloyd-Bostock and B. R. Clifford (Eds.), *Evaluating Witness Evidence*, pp. 173–187. Chichester: John Wiley and Sons.

Shepherd, J. W. and Ellis, H. D. (1995). Face recall – methods and problems. In S. L. Sporer, R. S. Malpass and G. Koehnken (Eds.), *Psychological Issues in Eyewitness Identification*, pp. 87–115. Hillsdale, NJ: Lawrence Erlbaum.

Tanaka, J. W. and Farah, M. J. (1993). Parts and wholes in face recognition. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, **46A**, 225–245.

Tredoux, C. G., Nunez, D. T., Oxtoby, O. and Prag, B. (2006). An evaluation of ID: an eigenface based construction system. *South African Computer Journal*, **37**, 1–9.

Valentine, T. and Mesout, J. (2009). Eyewitness identification under stress in the London Dungeon. *Applied Cognitive Psychology*, **23**, 151–161.

Vervaeke, G., Vanderhallen, M., Van Koppen, P. and Goethals, J. (2002). *Much Ado About Crime: Chapters on Psychology and Law*. Brussels: Uitgeverij Politeia NV.

Wells, G. L. and Hryciw, B. (1984). Memory for faces: encoding and retrieval operations. *Memory and Cognition*, **12**, 338–344.

Wells, G., Memon, A. and Penrod, S. D. (2007). Eyewitness evidence: improving its probative value. *Psychological Sciences in the Public Interest*, **7**, 45–75.

Chapter 5

Facial ageing

David Hunter, Bernard Tiddeman and David Perrett

5.1 Introduction

In this chapter we outline current research into face ageing and explore both computational methods and perceived results. We describe methods to digitally alter an image of a human face of known age, such that the image becomes an estimation of the same person apparently aged by a set amount. This has many useful applications, from locating missing persons to improving identification systems to take ageing into account. It is also of interest to the cosmetics industry, to simulate and evaluate the effects of anti-ageing treatments, as well as the entertainments industry, for example to age an actor.

Previous research into synthetic face ageing has concentrated on transforming a two-dimensional (2D) image. These methods work by applying a shape and colour change to the input image, often based on a learned statistical model. Early methods such as cardioidal strain, were non-statistical and relied on the similarity between certain mathematical functions and large-scale biological changes (Pittenger and Shaw, 1975; Pittenger *et al.*, 1975; Mark and Todd, 1983; Bruce *et al.*, 1989). More recent researchers have used statistical modelling methods to derive a model from a set of training images (Lanitis *et al.*, 2002; Scandrett *et al.*, 2006). The primary variations in these methods have been the functions and methods used to train the model.

Much of the previous research can be broken down into two major categories; ageing simulation and age estimation. Ageing simulation is the process of synthesising a face image such that it resembles an input face aged by a specified number of years. Age estimation is the reverse of age simulation, using statistical models to estimate the age of a person based on their physical appearance. Although this chapter concentrates on ageing simulation, many of the ideas and principles behind age estimation are still relevant. Also, some methods of automated age estimation have been used to perform ageing simulation, where image parameters are altered in order to match the recognised age to the desired age (Lanitis *et al.*, 2002).

5.1.1 Desirable features of an ageing system

In describing and analysing the ageing systems we will be assessing them on their performance on two main criteria. Namely the ability to:

- *Accurately model age changes*. That is, the ability to accurately predict and synthesise the changes in face appearance resulting from the effects of ageing.
- *Retain identity*. Not only must the resulting face image appear to be of the target age but it must also be recognisable. That is, the viewer of the image must be able to see that it is the original person aged by a set amount.

In 2D and 3D face-image processing, the effects of ageing can be simulated by altering the images in a number of different ways. The main changes that can be applied to images are:

- *Shape changes*, i.e. changes in overall appearance resulting from bone and cartilage growth, or from movement effects, such as sagging.
- *Colour changes*. Overall changes in face colour resulting from ageing.
- *Textural changes*. Localised changes in face appearance, such as those caused by wrinkling.

Craniofacial Identification, ed. C. Wilkinson and C. Rynn. Published by Cambridge University Press.

5.2 Non-statistical methods

Cardioidal strain has been used by a number of researchers (Pittenger and Shaw, 1975; Pittenger *et al.*, 1975; Mark and Todd, 1983; Bruce *et al.*, 1989) Cardioidal strain approximates shape changes caused by bone growth, making the head smaller, elongating the chin and raising the position of the nose and eyes. It was found by Pittenger and Shaw (1975) to positively affect how humans perceived the age of an outline of the face. Similarly, Mark and Todd (1983) found that applying cardioidal strain to a 3D model of a 15-year-old female's head also positively affected perceived age. However, as reported by Bruce *et al.* (1989), many of the observers did not see the faces (transformed to be younger) as younger.

Cardioidal strain has proved effective for simulating the large-scale shape changes caused by ageing, but is less suited to modelling smaller local changes, which do affect how ageing is perceived by a viewer (Burt and Perrett, 1995).

Ramanathan and Chellappa (2006) used a modified version of the cardioidal strain whereby the parameters of the modified cardioidal strain were adjusted such that the shape changes produced corresponded to the changes in ratio of a number of anthropometric measurements taken at key feature points on the face. These measurements were taken at a number of different age ranges and prototypes generated at 2, 5, 8, 12, 15 and 18 years. The cardioidal strain model was fitted to this set of prototypes. The results were evaluated using a recognition experiment based on eigenfaces (Turk and Pentland, 1991) and found that correct identification rates improved with their method with 58% accuracy as opposed to 44% without over 109 test images. Although this method can be adapted for individualised shape transforms it is not applicable to colour transforms.

5.3 Facial composites

In order to understand how computerised face-ageing algorithms work, it is worth reviewing the methods used to analyse and modify face images. In order to properly compare two images of a face, it is necessary to generate a one-to-one correspondence between the two faces. This is normally achieved by placing a set of landmarks on the face images that specify significant features, e.g. nose, eyes, mouth, etc. Thus, for example, the corner of the left eye in one image is matched to the corner of the left eye in all other correctly landmarked images. Correspondence of points between landmark positions can be estimated by interpolation. By warping the features of one face such that they match the locations of the other, the colours of the faces can be properly compared (Figure 5.1).

5.3.1 A face description model

In this section we define a face description model that will be used throughout the rest of this chapter. The shape is converted to a single long vector formed by concatenating the 2D-positions of the landmarks:

$$\mathbf{s} = \{x_1, y_1, x_2, y_2, \ldots, x_n, y_n\}, \quad (5.1)$$

where n is the number of landmarks. The shape of all the sample images is normalised by warping into a standard shape, usually the average (mean) shape for the set. The colour can then be described in a similar way to the shape, by treating the rgb (red-green-blue) values of the pixels sampled from the shape-normalised image as a single long vector. Hence the colour vector is given by:

Figure 5.1 An example of landmarking a face image. A set of landmarks are specified by hand on an image. The original image is shown on the left and the landmarked image is in the centre. The image is then warped, to place the landmarks into the same positions as those of a reference shape, as shown on the right. This figure is produced in colour in the plate section.

Figure 5.2 Two prototypes. The left image is built from a set of 'East-Asian' males aged 13–18 years. The right image is built from a set of 'East-Asian' males aged 50–54 years.

$$\mathbf{c} = \{r_1, g_1, b_1, r_2, g_2, b_2, \ldots, r_m, g_m, b_m\}, \qquad (5.2)$$

where m is the number of colour samples in the shape-normalised image and r,g,b are the *red*, *green*, and *blue* values of the samples. By using a common template of landmarks and a common reference shape for colour sampling, both n and m are constant over a set of face images and **s** and **c** have a consistent ordering.

5.3.2 Building a prototype

Using the above face description it is possible to build a composite face made of many exemplar images. Given a set of landmarked and registered faces, with the shape of the ith face image vectorised as $\mathbf{s}_i$, and colour as $\mathbf{c}_i$. A composite of these faces can be constructed by taking the mean values of **s** and **c** over the set of images.

$$\hat{s} = \sum_{i=1}^{k} \mathbf{s}_i, \qquad \hat{c} = \sum_{i=1}^{k} \mathbf{c}_i \qquad (5.3)$$

where k is the number of face images used to form the composite.

For the most accurate results, the position, orientation and scale of all the shape templates should be normalised to the average shape before constructing the average. This can be achieved using an Eigen-analysis based technique or by iteratively estimating the mean by aligning to the current estimate of the mean (initialised as a single individual), averaging and repeating until convergence. Normalisation of the image intensity is also possible (e.g. to normalise the brightness and standard deviation of the pixel brightness across the face area), but this can mask desired effects (e.g. the face becomes darker with age) so controlled lighting is a better solution where available.

Figure 5.3 Examples of face images aged using prototypes. The top row shows a student age female aged to a target age of 55. The bottom row shows a male student aged to a target age of 55. This example uses separate prototypes for male and female faces, with each set containing only White faces.

These composite images can be used to form *prototypical* faces that display a particular characteristic. For example, if we take a set of female faces we can create a composite that shows the average characteristic of female faces and thus forms a female prototype.

This method can also be used to create prototypical face images for a particular age group, by building a composite from exemplar faces of the desired ages (Figures 5.2 and 5.3). The pixel and shape differences between two age prototypes can be used to define an ageing trajectory. Applying this trajectory as a transform to a face image results in a face that appears to have aged, while retaining the identity of the original subject. The trajectory is applied in four stages:

- Orient the start and end prototype landmarks to the subject's landmarks.
- Calculate the differences between the normalised landmark locations and add them to the subject's landmarks to define the new shape.
- Warp the start and end prototype images and the subject's image into this new shape.
- Add the pixel colour difference between the warped prototypes to the warped subject image at each pixel.

5.3.3 The efficacy of prototype face images

Prototype face images aim to capture the specific characteristics of a group. Here we are studying ageing, so we wish to extract and display characteristics related to age. A key question is whether a prototype constructed from a group of faces of similar age appears to a human observer to be approximately the same age as the faces that make up the group. Or alternatively whether human observers detect, with some degree of accuracy, an age difference between young and old prototypes.

Rowland and Perrett (1995) generated two prototype images using the averages of 20 young male faces (between 25 and 29) and 20 older faces (also males between 50 and 54). The differences were added to a target face, using image warping to produce the appearance of ageing. It was noted that both shape and colour changes separately produce an increase in perceived age, although the age difference produced by the combined shape and colour transform was significantly less than the 25-year age gap. It was postulated that this was caused by the algorithm blurring out textural detail such as wrinkles. Importantly it was demonstrated that the transform maintains the identity of the person, thus the resulting image not only looks older but look like the *same* person, older.

Burt and Perrett further investigated the process of ageing using these facial composites and transform algorithms (1995). Face images of 147 White males between 20 and 62 were collected and divided into seven sets, each spanning five years. An average for each group was calculated along with a population average made by combining the groups. It was found that the perceived age of the composite average of each group was consistent with the average perceived age of the individuals that made up the group, but noted that raters tended to underestimate the age of the composite images. This underestimation was greater in the older age groups than in the younger age groups. It was concluded that the warping and blending process retained most of the age-related information and suggested that the underestimation was due to a loss of textural detail in the blending process. In the same paper, two different ageing transforms were described: one based on colour caricatures and another based on the vector difference between the oldest and the youngest groups. Colour caricatures were created by doubling the colour difference (in rgb space) between the average of the 50–54 age group and the population average. In the second transform the difference in the shape and colour, between the oldest and the youngest age group was calculated. The shape and colour differences were then superimposed onto a target image. Experimental evaluation showed that both techniques produced a significant increase in the perceived age, although significantly less than the age difference between the original groups used to train the transform. It is not unreasonable to suppose that the consistent underestimation of the prototype's age is a result of textural smoothing in the averaging process, smooth skin being a sign of relative youth.

5.4 Statistical representation of faces

For many applications, such as face feature location, face recognition or face ageing it is useful to have a compact, parametric description of the face. A statistical method first developed for the purpose of face recognition (Turk and Pentland, 1991) is commonly used as a face description that can be built automatically from a set of delineated face images. Based on the Karhunen–Loève theorem, the space of human face images is described as a weighted linear combination of basis vectors, with the weights acting as a parameterisation of the face.

The algorithm for generating a new face-space from a set of images begins in a similar manner to the prototyping method. A set of human face images is manually landmarked, with the landmark positions forming a shape description, $\mathbf{s}'^{(i)}$, where i is the ith image in the set. The mean of the landmarked feature positions over the set is defined as $\hat{\mathbf{s}}$. The images are then warped using a suitable warping function into the shape of $\hat{\mathbf{s}}$. We call the vectorised ith image $\mathbf{c}'^{(i)}$. The mean shape and colour are subtracted from both the shape and colour components, which are then used as the columns of matrices S' and C'. Then principal components analysis (PCA) is used to find a set of vectors that best describe the variance in the dataset. The first principal component is the vector that maximises the explained variance in S', or put another way minimises the mean-squared error of approximating the data. The second principal component maximises the remaining variance explained once the variance from the first principal component is removed, and so on.

The original faces in the training set can be exactly reconstructed as a weighted sum of all the principal components. Using all the components does not provide a compact description of the face set: it has the same dimensions as the training set. In addition the higher-order components will not generalise well to out-of-set images, whereas the lower-order components represent the main variations within the sample set and usualy generalise well to the rest of the population. Often the number of components to use for a face description is chosen by having the model explain a certain amount of the variance in the face space (usually around 95%). Principal components analysis also provides the standard deviation associated with each component. We define the cumulative energy for the ith component as $g_l = \frac{\sum_{i=1}^{l} \sigma_i}{\sum_{i=1}^{m} \sigma_i}$, i.e. the cumulative sum of the standard deviations up to and including the ith principal component divided by the total sum of standard deviations. We find the first g_l such that $g_l \geq \varepsilon$, where ε is the amount of variance the model is required to explain, and truncate the PCA model to the first l principal components. We call these truncated components S and C respectively. A new face image can be constructed given a set of shape and colour parameters, $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ as,

$$\mathbf{s}^{(new)} = \hat{s} + \sum_{i=1}^{l} \alpha_i \mathbf{s}_i, \mathbf{c}^{(new)} = \hat{c} + \sum_{i=1}^{l} \alpha_i \mathbf{c}_i, \qquad (5.4)$$

Here $\mathbf{s}_i$ and $\mathbf{c}_i$ are the ith principal components in the shape and colour description. The probability distribution of the shape descriptor is:

$$p(\mathbf{s}) : e^{-\frac{1}{2}\sum_i \frac{\alpha_i^2}{\sigma_{\mathbf{s},i}}} \qquad (5.5)$$

From this we can now define a metric measuring the difference between two sets of face parameters and thus their original images, using the Mahalanobis distance (i.e the distance in a space scaled so that the standard deviation of all components is 1).

The model described thus far does not learn the correlations between shape and appearance. Such correlations undoubtedly exist in faces, for example the facial shape and colour differences associated with different ethnicity, or the shape changes associated with ageing (mostly sagging skin) and the corresponding colour changes (greying hair, etc). A second application of PCA can be used to learn these correlations from the (separate) shape and colour PCA parameters. The shape parameters are scaled to match the variance between the shape and colour parameters. Then the shape and colour parameters are placed into a single vector and PCA analysis is applied to these. Analysis of an input image is performed by first estimating the separate shape and colour parameters, combining them (with appropriate weighting) into a single vector and then estimating the weights of the combined model. In a similar fashion a face can be synthesised from a set of combined weights by using them to rebuild the combined shape and colour vector, splitting the vector into separate shape and colour components and rebuilding the separate shape and shape free image.

5.5 Ageing using principal components analysis

Many methods in modern ageing research stem from the work of Lanitis *et al.* (2002). Ageing functions were generated by fitting polynomial curves through a set of faces parameterised using PCA. A PCA model was generated from a set of landmarked face images, and the faces were then parameterised using a truncated set of eigenvectors.

Given a set of faces of a number of subjects at various age points, Lanitis was able to generate a series of age functions through the PCA face space that describe ageing, and investigate the ability of polynomial functions to estimate the age of a face given a set of parameters ($\mathbf{p}$),

$$f(\mathbf{p}) = age = C + \mathbf{a}\mathbf{p} + \mathbf{b}\mathbf{p}^2 \qquad (5.6)$$

Here $\mathbf{a}$ and $\mathbf{b}$ are weights that parameterise the polynomial function and C is a constant. Lanitis and colleagues also investigated polynomials of degree 1 and 3, but found polynomials of degree 2 to be superior at estimating the subjects' age.

As the function implicitly assumes that all the faces in the set age in the same manner, Equation 5.6 can be considered to be a *Universal Ageing function*. In general, however, individuals age differently. A key insight of Lanitis's paper was that people of similar appearance age in a similar manner. As such, examining the relationship between the parameters of facial appearance and the parameters of the ageing path for a particular person they could generate ageing functions tailored to an unseen individual. An ageing path for a specific individual is generated by fitting a quadratic polynomial curve to the facial parameters from face images of the same person at different age points in that person's life. It was found that the facial appearance parameters and ageing function parameters had a correlation coefficient of 0.55, suggesting that faces with similar appearance do indeed age in a similar manner. Using this, they were able to generate an individualised ageing function for an unseen individual as a weighted sum of the ageing functions for similar individuals in the dataset. If the ageing function for an individual in the training set, *i*, is denoted f_i then the ageing function for an unseen individual is formulated as

$$f_{new} = \sum_i \lambda_i f_i \tag{5.7}$$

where λ is a set of weights for each individual in the training set. These weights were based on the similarity between the two faces, estimated using the probability distribution generated from the construction of the PCA model.

Lifestyle information about the individuals in the dataset was also gathered, such as gender, socio-economic factors, weather exposure, etc. by asking those volunteering facial images to fill in questionnaires. The lifestyle information was vectorised and scaled such that the total variance in lifestyle information equalled the total variance of the facial parameters. In this way a new ageing function can be generated by weighting the ageing functions in the dataset by the combined appearance–lifestyle probabilities. This produced a higher correlation coefficient of 0.72 suggesting that lifestyle has a significant impact on the visual effects of ageing. By comparing the estimated ages of the face images to the known ages using a leave-one-out method, it was demonstrated that individualised age models produce a more accurate estimate than global ageing functions. This was the case for both appearance-based weighting and combined lifestyle–appearance weighting. However, this method relies upon the existence of similar faces in the training set, otherwise the age function tends towards the global age function, as Lanitis and colleagues showed by attempting to estimate the ages of faces from a different ethnic group than that used to train the age model.

Their work also covered the area of synthesising facial ageing, generating the aged face images using the inverse of the polynomial functions used in age estimation. The results of the age synthesis were evaluated both quantitatively and perceptually. The parameters of the aged faces were compared with the parameters of a face image of the same individual at the target age using the Mahalanobis distance. The rendered face images were also shown to a set of human raters, who were asked to judge whether the synthesised, age-progressed image looked older than the original image, and whether the rendered image was more similar to the target individual than the original. It was concluded from both the quantitative and perceptual results that both global and individual ageing functions produced suitably aged individuals, but that the individualised method was superior.

5.6 Trajectories

In addition to appearance and lifestyle information, we often have access to other information about a subject whose image we wish to age. For example we might have images of the individual over a number of years, or we might have images of their parents. Scandrett *et al.* (2006) investigated ageing functions using combinations of ageing trajectories. Like Lanitis *et al.* (2002), Scandrett used a face model parameterised using PCA, and aged the model through this PCA space. In order to eliminate variations caused by pose and expression, the horizontal and vertical rotations as well as the amount of smiling were weighted subjectively by human observers and then defined in the face space as the sum of score-weighted face parameters. Weighted multiples of these vectors were then used to alter the pose and expression so that they became uniform. Even in small rotations parts of the face that are occluded become visible; the textures are unknown and must be approximated. Scandrett *et al.* (2006) achieved this by reflecting the normalised image about its vertical axis. Like other 2D ageing methodologies, Scandrett found that lighting variations reduced clustering of face parameters of around the same age and thus the quality of aged textures. This effect was particularly pronounced with trajectories derived from an individual's history using fewer face samples resulting in less smoothing of errors.

Each of the trajectories were designed to extract different factors that affect ageing, such as personal history, sex, how parents aged etc. The trajectories

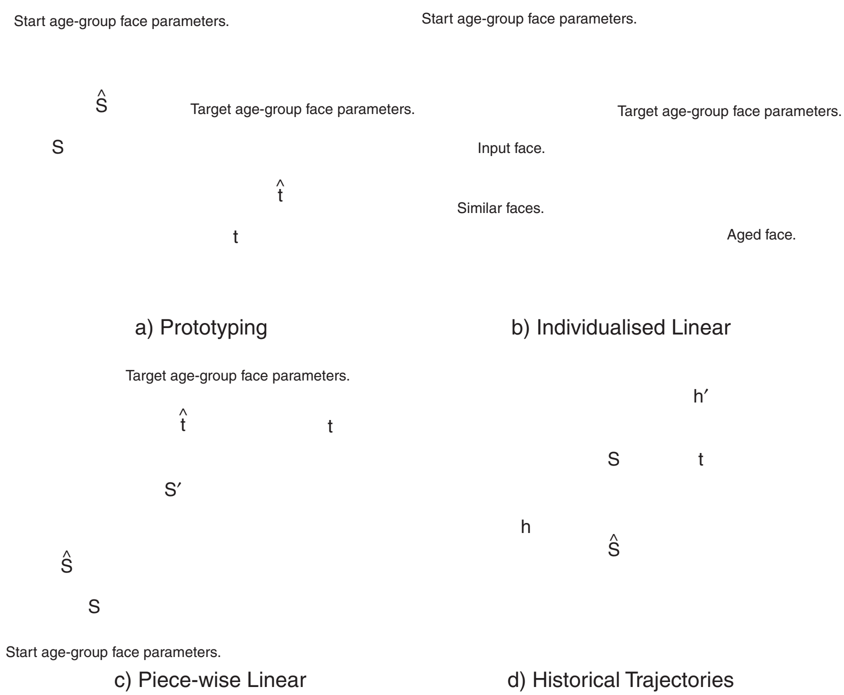


Figure 5.4 A comparison of two ageing functions. The Prototyping function, (a), describes a transform of an input face (**s**) to (**t**). **ŝ** and **û** represent the source and target Prototypes respectively. The individualised function (shown in its linear incarnation for simplicity) in (b), shows how a new ageing trajectory is created from the training set. Those transforms that originate 'near' the unseen face have the strongest influence on its final trajectory. Diagram (c) shows the piece-wise linear ageing used by Scandrett *et al.* (2006). A trajectory for each age group is calculated as an age-weighted average (shown in green). A face can be aged by stepping from group to group along the trajectories defined for each group. Diagram (d) shows how an historical ageing trajectory works (Scandrett *et al.*, 2006). Given two faces one (**s**) the starting point of the ageing, and one (**h**) of the same person at a younger age. The trajectory predicts a face at **h′**. This trajectory is modified by the trajectory from an overall age group **ŝ** to create a target at **t**. This figure is produced in colour in the plate section.

were defined as the sum of the face parameters centred on the group mean weighted by the mean shifted age of each face.

$$\mathbf{v} = \sum_{i=1}^{k} (y_i - \hat{y})(\mathbf{p}_i - \hat{\mathbf{p}}) \tag{5.8}$$

Here the mean age of the age-group $\hat{y}$ is subtracted from the age of the *i*th individual, y_i and used as a weight on the vector produced by subtracting the mean parameters $\hat{\mathbf{p}}$ from the parameters of the *i*th individual.

Face images were aged by altering their parameters in the direction of one or more combined ageing trajectories until the target age was reached. Each individual update step is similar to the update step for prototype images, however the prototype updates are applied directly to the images. Scandrett *et al.* (2006) applied modifications to PCA parameters which were then converted back to specific shape and colour combinations. They also differ in how they compute the update directions, Prototyping transforms are between group averages, trajectories are within group.

As males and females are known to age in different ways, sex-specific ageing trajectories were produced for each group. An input face was then compared with these ageing trajectories to determine the comparative influence of the male and female trajectories using a ratio of distances from the face to each trajectory. It is often the case that multiple images of an individual are available covering a range of ages, all of which pre-date the 'start' age of an ageing function. The 'start' age is the age of the most recent image and therefore the closest to the target age. Scandrett *et al.* (2006) used these images to construct what they called an 'historical' ageing trajectory, using the age-weighted average of the images in the same manner as they had for other groups of images. This historical trajectory could then be combined with the average ageing trajectory, to produce an aged face image. The trajectory was weighted according to a maximum-likelihood metric, constraining the *trajectory* to be typical of those between the source and target age groups, and also constraining final *appearance* to be a typical member of the set of faces at the target age.

The results were analysed using the root-mean-squared error, both on the shape vertices and per pixel, between the resulting face image and a known ground truth image of the individual at the target age. The faces had been converted to greyscale and normalised to have a mean intensity of zero and standard deviation of one, in order to remove some of the effects of lighting on the results. It was found that in general the root-mean squared shape and texture errors were lower when compared with the ground truth image than with other images in the target age set, and concluded that the ageing methods both aged the individuals appropriately and retained identity through the transform. It was found that the most accurate method of ageing varied between individuals and so they could not conclude which method had the best performance.

5.7 3D

Scherbaum *et al.* (2007) fitted a 3D morphable model to a database of laser-scanned cylindrical depth-maps. A database of 200 adult images and 238 children and teenagers was used. The latter group ranged in age from 96 months to 191 months. In order to improve the resolution of the face texture map, the textures were reconstructed from three photographs taken at three separate angles. The parameters of the model and the age of the subject were used to train a Support Vector Regression (SVR) model. The SVR formed a mapping from the high-dimensional parameter space of the model to the R space of the subject's age. This was used to estimate the age of the subject once the parameters of the morphable model had been found. A new face model could be synthesised from a given set of parameters by 'stepping' through the curved SVR space using a fourth order Runge–Kutta algorithm, using the parameters and an estimated age as the starting point.

Scherbaum *et al.* (2007) did not use multiple time-space images of the same individual in building the model. Their claim to individualisation is the observation that, based on the mean angles between the support vector gradients, the SVR produced different ageing trajectories for different individuals and could therefore be said to be individualised. While this is true, the variation is derived from a large number of single 'snapshots', i.e. it describes the variations within a population. It may not necessarily capture the variations due to ageing in a particular individual.

Park *et al.* (2008) performed a similar experiment to the authors, fitting a 3D morphable model to a set of delineated faces using point data. Ageing was performed by calculating a set of weights between an input face and exemplar faces in the same age group. These weights were then used to build an aged face as a weighted sum of the corresponding faces at the target age. The results were compared using Cumulative Match Characteristic curves to other ageing methods and reported similar results to other methods. It was observed that shape modelling in 3D gave improved performance in pose and lighting compensation.

Hunter and Tiddemann (2009) also fitted 3D morphable models to a set of face images, using image colour as well as delineated point data. They used a partial least squares (PLS) based method to face models (see section 5.8).

5.8 Partial least squares

In the previous sections we have discussed ageing of *unstandardised* facial images using *unstandardised* training data. The unstandardised images have many sources of variation due to external factors such as head rotation, lighting and expression. Scandrett *et al.* (2006) attempted to standardise the images by using subjective rating of factors such as lighting, expression and rotation, with some success. An alternative is to try and mathematically find the components of the face model which are related to ageing and separate them from those that are not.

Principal components analysis decomposes the data based on within-set variances only, without reference to the variables that we are trying to model using the data. This can result in loss of details that do not explain much of the variance in the input data but are of significance in regard to a particular feature that we wish to model. An alternative decomposition known as *Partial Least Squares* (PLS) attempts to describe a set of dependent variables from a set of predictors. Using PLS we can produce a set of vectors that approximately spans the space of face models and are highly correlated with the age.

Hunter and Tiddeman (2009) used PLS on 3D morphable models to produce aged 3D models of the faces. The parameters of the morphable models were updated using an individualised linear transform, in a similar manner to that described in section 5.3, through PLS space instead of PCA. Using human raters to determine the perceived age of the 3D model, it was found that PLS outperformed both prototyping and PCA-based linear methods. It was also determined how well humans were able to recognise the face-model after the transformation as being the same person. In this test the PCA-based transform resulted in faces that were more easily recognised than those from the PLS transform. This is a result of the transformation produced by the PLS method being larger than using PCA.

5.9 Fine detail analysis

Up to this point we have mainly discussed methods to analyse and synthesise age-related changes in the shape of a face. Some methods have included colour either as a combined model, e.g. a combined shape and colour PCA model of Lanitis *et al.* (2002), or have applied the same process to separate colour parameters (Burt and Perrett, 1995; Scandrett *et al.*, 2006; Hunter and Tiddeman, 2009). When modelling shape changes we expect a reasonably smooth transform that will not introduce noise or discountinuity, but when altering the facial colours this is not quite such a desirable property. The problem is that facial features such as wrinkles and spots are not particularly consistent in number, appearance or location. The process of generating a prototype or a PCA decomposition smooths over these features, resulting in a poor depiction of them in the aged image. In the case of PCA the features can even be missing in the lower-dimensional approximation used to train the ageing paths.

These smoothing effects can be seen as a low-pass filter on the face image that removes any high-frequency information, i.e. fine detail. The approaches at correcting for this phenomenon therefore concentrate on recording and reproducing the high-frequency information otherwise discarded.

Hussein (2002) synthesised wrinkles by attempting to align the surface normals of two faces, an older and a younger using the relationship between pixel intensity and surface orientation. Under the assumption that the two surfaces shown in the image are co-incident and, under the same lighting conditions, surface details such as wrinkles would become the primary changes in intensity. The ratio of the two images was used,

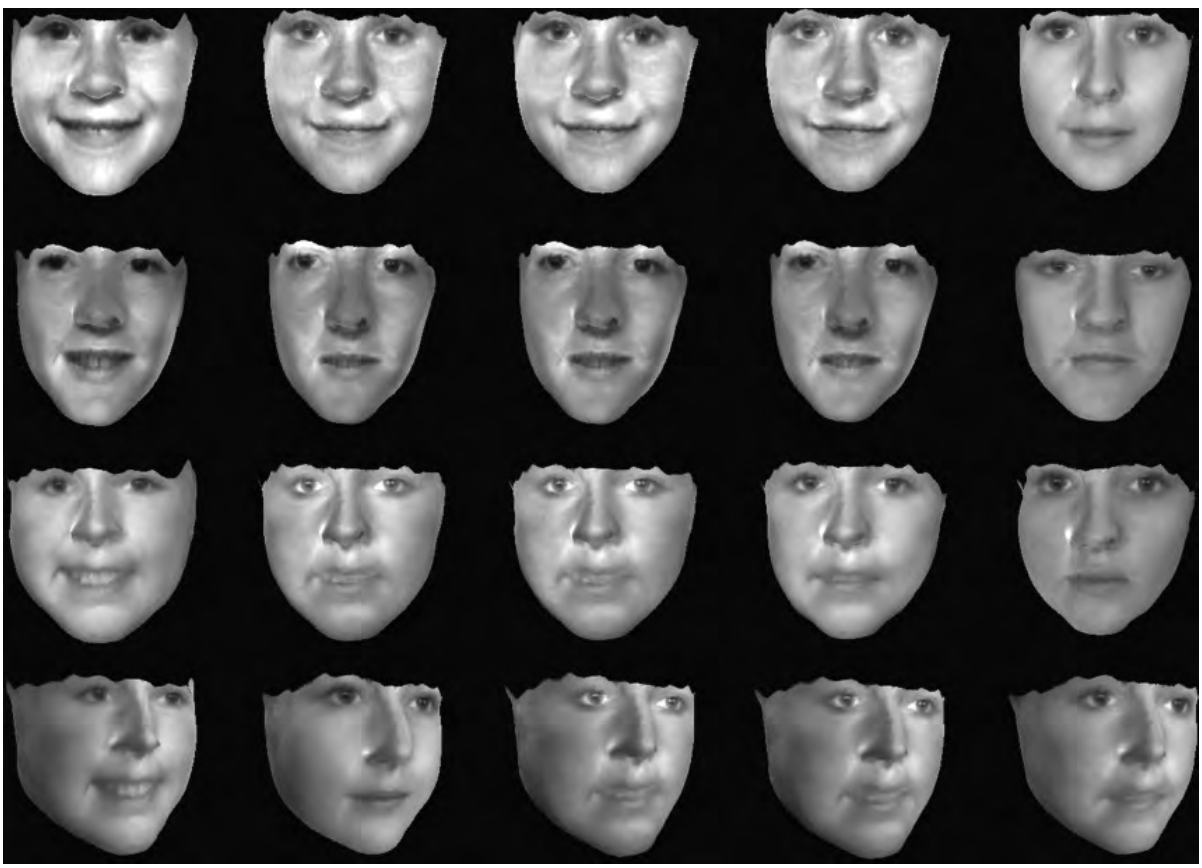


Figure 5.5 Examples of aged face images. Each row contains a different individual. The columns show, from left to right, the original mid-child face model for each individual, the mid-child face model aged to student age using the Prototyping method, the Individual Linear method and the PLS method. The right-most column shows the original face-model for the individual at student age. The final row shows the same individuals as the 3rd, rotated to show a three-quarter profile.

smoothed with a Gaussian filter multiplied with one of the images, so that the fine detail of the other was applied to it. Their method suffered from two main drawbacks: firstly, it could not be used under varying lighting conditions; secondly, the age was defined from only one image and thus would generally not produce a convincing ageing result for an arbitrary individual.

Gandhi (2004) used an Image Based Surface Detail Transfer (Liu *et al.*, 2004) procedure to map the high-frequency information from an older prototype to a younger, and vice versa using a Gaussian convolution as a low-pass filter. The idea here was to take the high-frequency details of the input image and replace them with the target's details. The Gaussian convolution produced two images, one the smoothed original containing the low-frequency large-scale details and the other, the result of applying a standard boost filter, containing the high-frequency fine-scale details. An aged image was synthesised by combining the high frequency of a prototype with the low frequency of the image. Varying the width of the kernel would vary the size of details captured and thus the perceived age of the person. The prototypes at each age were created by averaging all the images in an age group. Smoothing problems were avoided by combining the high-frequency parts of the training images with the combined average to retain fine detail.

Tiddeman *et al.* (2001) used a Gabor wavelet function to detect edges in the image and decompose it into a pyramid of images containing edge information at varying spatial scales The edge magnitudes were then smoothed with a B-spline filter to give a measure of edge strength about a particular point in each sub-band. Prototypes at each age were generated using the technique of Benson and Perrett (1993) and the wavelets were then amplified locally to match the mean of the set. The values of the input wavelet images were modified to more closely match those of the target prototype. These were tested perceptually and found to reduce the gap between the perceived age of the image and the intended age. The method was then extended using Markov Random Fields (Cross, 1980). An individual was aged using the prototyping method of Burt and Perrett (1995), described above. Detail was added to the resulting face by decomposing the image into a wavelet pyramid and scanning across the sub-bands using the MRF model to choose wavelet coefficients that match the cumulative probability of the input values. It was found that human raters found the resulting image more closely matched the target age of the older group than either the wavelet method on its own or the prototyping method; it also succeeded in the rejuvenating test where wavelets failed. It was also found that the images were rated as more realistic than those generated using wavelets alone (Tiddeman *et al.*, 2005).

More recently Suo *et al.* (2007) presented a method for simulating ageing using a mixture of a global face-ageing model and separate facial components aged according to a Markov process. The low-resolution global changes are performed by a simple merging with an appropriate image of the target age, which has obvious drawbacks (i.e. it could alter the identity). Facial components (such as eyes, nose, etc.) are chosen from a selection of facial components of the target age with similar appearance to the original using a probabilistic sampling model. Wrinkles are added to the face model as the age is increased using a similar process. The process is attractive in that it produces very realistic output images and can produce a range of plausible aged images, reflecting the increasing uncertainty of appearance over extended time periods. On the other hand, the probabilistic models need to be learned carefully to avoid biasing the results.

5.10 Conclusion

In this chapter we have outlined the current state of the art in facial ageing and age estimation, demonstrating how apparently simple methods such as linear regression and prototyping are surprisingly effective in simulating and synthesising the effects of ageing on human face images. However, human raters consistently rate the resulting images as younger than the target age. The most sophisticated methods all attempt to close this gap by focusing on different aspects of the problem. In general, authors have concentrated on either improving statistical shape models of structural changes due to ageing, or on using high-pass filters and pyramid decomposition to model fine-textural changes.

References

Benson, P. J., and Perrett, D. I. (1993). Extracting prototypical facial images from exemplars. *Perception*, **22**, 257–262.

Bruce, V., Burton, M. and Doyle, T. (1989). Further experiment on the perception of growth in three dimensions. *Perception and Psychophysics*, **46**(6), 528–536.

Burt, D. M., and Perrett, D. I. (1995). Perception of age in adult caucasian male faces: computer graphic manipulation of shape and colour information. *Proceedings of Royal Society of London*, **B-259**, 137–143.

Cross, G. R. (1980). *Markov Random Field Texture Models*. PhD thesis, East Lansing, MI, USA.

Gandhi, M. R. (2004). *A Method for Automatic Synthesis of Aged Human Facial Images*. Technical report, Master's thesis, McGill University, Canada.

Hunter, D. W. and Tiddeman, B. P. (2009). Visual ageing of human faces in three dimensions using morphable models and projection to latent structures. *VISAPP 2009: Proceedings of the Third International Conference on Computer Vision Theory and Applications*. Lisboa, Portugal.

Hussein, H. K. (2002). Towards realistic facial modeling and re-rendering of human skin aging animation. *SMI '02: Proceedings of the Shape Modeling International 2002*. Washington, DC: IEEE Computer Society.

Lanitis, A., Taylor, C. J. and Cootes, T. F. (2002). Toward automatic simulation of aging effects on face images. *IEEE Trans. Pattern Anal. Mach. Intell.*, **24**(4), 442–455.

Liu, Z., Zhang, Z. and Shan, Y. (2004). Image-based surface detail transfer. *IEEE Computer Graphics and Applications*, **24**(3), 30–35.

Mark, L. S., and Todd, J. T. (1983). The perception of growth in three dimensions. *Perception and Psychophysics*, **33**(2), 193–196.

Park, U., Tong, Y. and Jain, A. K. (2008). Face recognition with temporal invariance: a 3d aging model. *Proceedings of International Conference on Automatic Face and Gesture Recognition, Amsterdam, Netherlands*, pp. 1–7.

Pittenger, J. B., and Shaw, R. E. (1975). Aging faces as visco-elastic events: implications for a theory of nonrigid shape perception. *J. Experimental Psychology: Human Perception and Performance*, **1**(4), 374–382.

Pittenger, J. B., Shaw, R. E. and Mark, L. S. (1975). Perceptual information for the age level of faces as a higher order invariant of growth. *Journal of Experimental Psychology: Human Perception and Performance*, **5**(3), 478–493.

Ramanathan, N. and Chellappa, R. (2006). Modeling age progression in young faces. *Computer Vision and Pattern Recognition Conference* (CVPR '06), pp. 387–394.

Rowland, D. A., and Perrett, D. I. (1995). Manipulating facial appearance through shape and color. *IEEE Computer Graphics and Applications*, **15**(5), 70–76.

Scandrett, C. M., Solomon, C. J. and Gibson, S. J. (2006). A person-specific, rigorous aging model of the human face. *Pattern Recognition Letters*, **27**(15), 1776–1787.

Scherbaum, K., Sunkel, M., Seidel, H.-P. and Blanz, V. (2007). Prediction of individual non-linear aging trajectories of faces. European Association for Computer Graphics, 28th Annual Conference, EUROGRAPHICS 2007, Prague, Czech Republic. *Computer Graphics Forum*, **3**, 285–294.

Suo, J., Min, F., Zhu, S. C., Shan, S. G. and Chen, X. L. (2007). A multi-resolution dynamic model for face aging simulation. *IEEE Computer Vision and Pattern Recognition*, **2007**, 1–8.

Tiddeman, B., Burt, M. and Perrett, D. I. (2001). Prototyping and transforming facial textures for perception research. *IEEE Computer Graphics and Applications*, **21**(5), 42–50.

Tiddeman, B., Stirrat, M. and Perrett, D. I. (2005). Towards realism in facial image transformation: results of a wavelet MRF method. *Computer Graphics Forum*, **24**(3), 449–456.

Turk, M. and Pentland, A. (1991). Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, **3**(1), 71–86.

Chapter 6

Age progression and regression

Joe Mullins

6.1 Introduction

An explanation of the process of age progression and regression should begin with a definition of each. Age progression is the process of modifying a photograph of a person to represent the effect of ageing on their appearance. Digital image processing is the most common current technique, although artists' drawings are often utilised. Age progression is most often employed as a forensic tool by law enforcement officers to show the likely current appearance of a missing person predicted from a photograph that may be many years out of date. Age regression is defined as the modification of a photograph of a person to simulate their appearance at a younger age. Each of these processes are useful tools for law enforcement to assist in the recovery of long-term missing children, identify fugitives and assist in criminal investigations. The process for creation of these images will be described in this chapter.

6.2 Age progression

There are two categories of age progression: juvenile and adult. Juvenile age progression is used to help find abducted and missing children. A face changes significantly throughout childhood and age progression images have proven very useful in the recovery of these children. In the case of Jaycee Dugard, who was abducted from a bus stop and held captive for 18 years, law enforcement issued many images depicting how Jaycee would currently appear while her family searched for her over the years. A third-generation age progression was close to how Jaycee appeared upon her rescue. In this case the aged image did not aid in the recovery but it illustrates how these images can be effective and still allow recognition many years after a child goes missing. For more examples of age-progressed images, see the National Center for Missing and Exploited Children website (NCMEC, *2011a*).

Adult age progressions are commonly used in law enforcement to update fugitive images. Since convicted criminals are only photographed on entry to a prison facility, an escaped convict or previously convicted fugitive may appear quite different to the last recorded photograph. Updated age progression images, presented on wanted posters or in the media, are effective tools to help generate new leads for the capture and conviction of fugitives.

6.3 Juvenile age progression

The task of age-progression image creation often falls to a specialised forensic artist, who is able to manipulate the photograph of an individual at a younger age, either manually by sketching or digitally using photo-editing software programs, to artificially age the face and portray his or her appearance at an older age (the current age). The artist must combine artistic talent with an understanding of the growth and development of the face. Family reference pictures are crucial to produce a more accurate depiction of the child. Images of the biological parents or siblings at or around the target age of the missing child are preferable. If a child went missing around the age of 6 years with a target age of 12 years, the artist would request images of the parents or siblings at around the age of 12 years.

All the facial features inherited from the parents should be visible in the provided photographs. The shared characteristics are important in age progression (Figure 6.1). Since siblings are genetically closer to one another than either parent to the child, sibling photographs can be effective to help the artist maintain a reliable likeness. The artist is not the expert on the individual child's face and the biological parents

Craniofacial Identification, ed. C. Wilkinson and C. Rynn. Published by Cambridge University Press.

Figure 6.1 Reference photographs of the child's biological parents. Image on the left is father at age 12 years. Centre image is subject at age 6 years. Image on the right is subject's mother at age 12 years.

can be very helpful in providing characteristic age-related information about their child. When families do not have photographs available the artist must then use other more general reference images. Ageing a White male from 10 to 15 years would require reference photographs of other 15-year-old White males to match the child's facial features as closely as possible. A good tip is to use five or six photographs and only utilise small portions of each reference image; taking the eyes from one image, for example, and the nose from another. These templates can be resized and manipulated to fit, like a puzzle, using the younger child's face as a guide to placing the pieces. The completed image should still resemble the child and should not be recognisable as one of the reference templates. The artist must retain each child's unique facial identity in the age-progressed image.

The challenge for the artist is not to alter the image to the extent that the individual characteristics of the child's face are no longer recognisable. For example, if a child has a particular facial feature that is unique, the artist needs to ensure that is still visible in the age-progression image. A mole, birthmark or scar should always be visible as these are identifying marks important for recognition and recovery. Erasing these landmarks creates a fictitious face and will not maintain the child's unique identity.

So, how do you age a child's face? What makes a child's face look older? How can you tell the difference between a 6-year-old and a 3-year-old? How do you handle such things as clothing or hairstyles?

Deciduous or baby teeth are useful to indicate age. There are predictable patterns for tooth emergence in females and males and these patterns allow anthropologists to determine the age of a child. By the age of 12 years most children have all their permanent teeth, with the exception of the wisdom teeth, which do not typically erupt until 17–21 years (Figure 6.2). So it is possible to predict the age of a child by assessing an image of a smiling child, especially when some missing teeth are visible (Figure 6.3).

Ageing a child's face is accomplished in reference to the developmental stages of growth. Ageing an infant younger than 3 years is very difficult as infants tend to look a lot like each other and have not yet developed characteristic features (Figure 6.4). The procedure involves the elongation of the lower two-thirds of the child's face to depict age-related changes (Figure 6.5). This can be carried out by pulling down the area of the face just beneath the eyes.

Once the face is stretched the steps employed to age a child's face include reshaping of the mouth, addition of darker shadows along the sides of the nose, addition of smile lines, lengthening the child's neck, removal of baby fat and sharpening the angles of the lower jaw. When ageing juvenile males the addition of an Adam's apple helps to mature the appearance. The Adam's apple (the thyroid cartilage) is present at birth, but becomes more developed during puberty. The clothing and hairstyle applied to an age-progression image will be entirely subjective, so the artist must choose styles that are appropriate to the age, gender, culture, religion and personal circumstances of the individual.

If the ears are visible in the original photo it is important for the artist not to alter the inner pattern of the ear. The soft tissue of the ear is unique and should be shown in the progression image.

A child's smile may also be unique, so no drastic changes should be made at the mouth and lips. If teeth are visible, deciduous teeth must be replaced with permanent teeth where appropriate. Eyebrows will thicken with age and facial hair will become more apparent (Figure 6.6).

The older the child, the easier it is to complete an age progression. A 10- or 11-year-old has most of the characteristics in their face defined, so the artist has most of the information already in the image and is less dependent on acquiring family reference photographs.

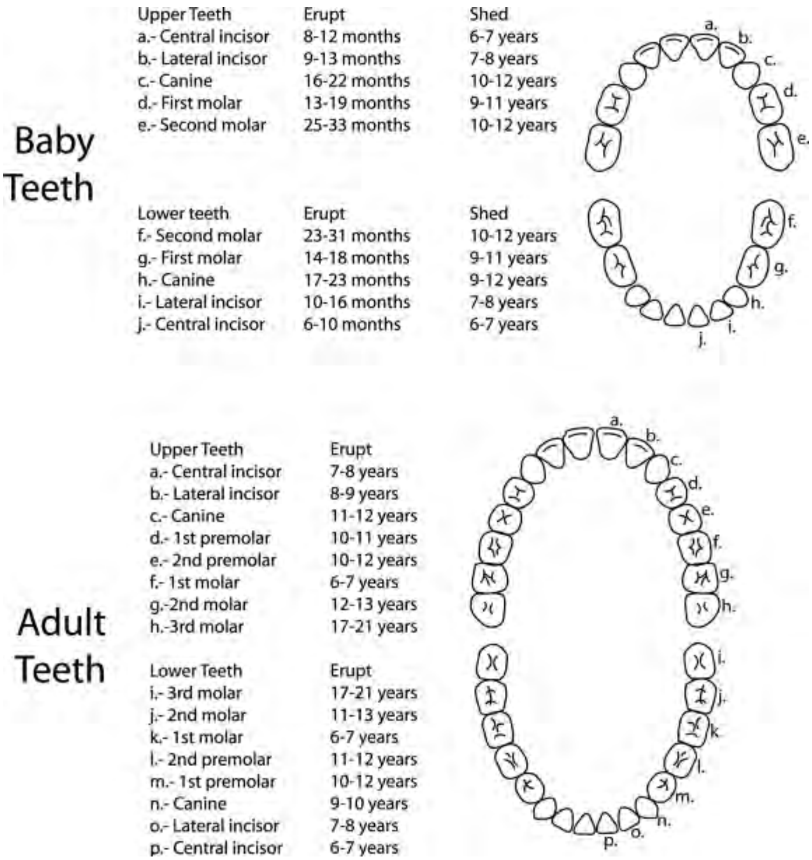


Figure 6.2 Dental eruption chart. From Scheuer and Black (2000) after Ubelaker (1978).

Figure 6.3 Female juvenile missing canine teeth indicating her age between 10 to 12 years and male juvenile missing lateral incisor indicating his age between 7 to 8 years.

The main purpose of an age progression is to generate leads for law enforcement officials. The artist need only create an image that sparks recognition. Since the process is subjective and the face is unique, it is impossible for an artist to produce an exact image of the target appearance. The success of age progression will be dependent on the maintenance of the child's unique facial identity and whether or not the right person sees the image and recognises the child.

These age progression techniques have been used by the National Center for Missing and Exploited Children's Forensic Services Unit since 1990 (http://www.missingkids.com). In that time over 900 missing children's cases have been resolved as a result of age-progression images.

6.4 Ageing adults

Ageing of an adult face is much easier than for a child. An adult face has all the defined features . . . moles, scars, permanent teeth, hairlines etc. Where possible the artist should request reference photographs from law enforcement officials and similarities between biological parents and siblings should be clearly visible (Figure 6.7).

Figure 6.4 Female infant. Note the dramatic changes in the face from 3 months to a year and a half.

Figure 6.5 Elongation of the lower two-thirds of the face for growth depiction. This is only carried out for children older than 3 years.

Having good-quality reference photographs to work with will always produce a more accurate age progression. When family likeness is apparent in the reference photographs, it is possible for the artist to create an image based solely on template references. Figure 6.8 illustrates the process of ageing an adult female from age 18 to 34 years using reference pictures of siblings. The artist was able to create an age-progression image that was easily recognisable as the subject.

Regarding the artificial ageing of adults, the task of the forensic artist is no longer to apply growth patterns to the face, since the growth process is completed by around 18 years. As we get older, the skin shows signs of age faster than any other organ of the body. It loses its elasticity, forming lines and wrinkles. Bones shrink away from the skin, causing sagging. There is a loss of the fat beneath the skin, which causes cheeks and eye sockets to hollow. This also leads to a loss of firmness around the neck. According to the American Academy of Dermatology (2010) there are two distinct types of ageing. Ageing caused by the genetic makeup we inherit is called intrinsic (internal) ageing. The other type of ageing is known as extrinsic (external) ageing and is caused by environmental factors, such as exposure to the sun's rays (Helfrich *et al.*, 2008).

Figure 6.9 shows an example of an extreme ageing of an adult covering a span of 50 years. No family reference photographs were used: only the application of the effects of ageing using the original photograph as a guide. Figure 6.10 shows an example of how the forensic artist can simply modify an existing image of a female to project how she may appear around age 60 years.

When artificially ageing an adult face, it is not necessary to be as precise: if the final image appears to be approximately the target age it should be successful. When ageing a face from 25 to 45 years, it is impossible to predict how the face would appear at exactly 45 years, since the aging of the adult face self-evidently varies with lifestyle and environment, rather than following a relatively predictable pattern of growth and development.

Keeping all of these factors in mind the forensic artist must apply these characteristics to the face to portray to the best of his or her ability how this individual will look in the future. According to NCMEC

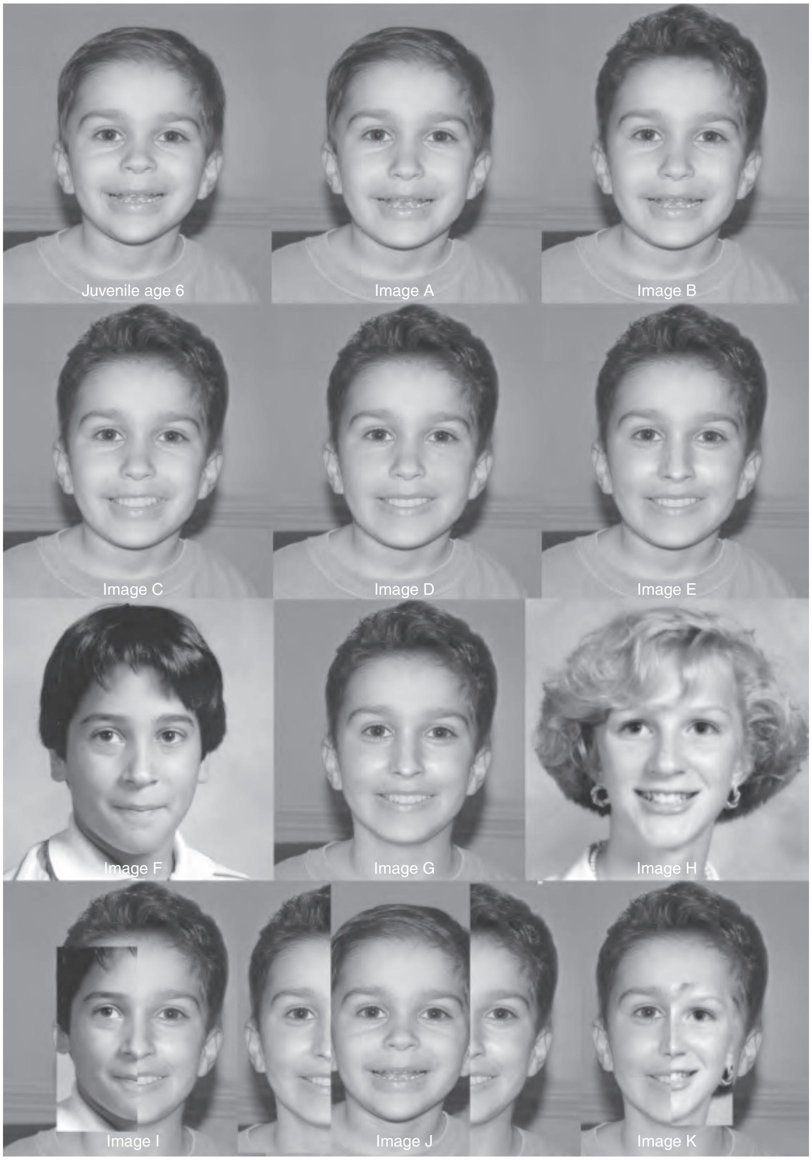


Figure 6.6 Age progression of juvenile male from age 6 to 12 years.
Image A: Applying lower two-thirds growth.
Image B: Change hairstyle.
Image C: Adding more mature teeth.
Image D: Thinning of cheeks (removing baby fat).
Image E: Adding length and shadowing to nose.
Image F: Biological father at age 12 years.
Image G: Lengthening of neck. Child aged to 12 years.
Image H: Biological mother at age 12 years.
Image I: Age progression image merged with father.
Image J: Age progression image merged with original image.
Image K: Age progression image merged with mother.

(2011*b*) an average of 2185 children are reported missing daily. Since 1984, NCMEC has assisted law enforcement to recover more than 157 720 children: some of these were missing for many years, so whether it is five, 10 or 30 years later, these images have been proven a useful resource for law enforcement.

6.5 Age regression

Reverse ageing of an individual to remove the mature features to make them appear younger is also possible. This process is known as age regression and involves the shrinking of the vertical facial proportions, removing wrinkles, changing the hairstyle, removing facial hair, increasing the size of the iris, shortening the nose and elevating the tip of the nose.

As illustrated in Figure 6.11 the process of artificial ageing also works in reverse. It is possible to modify, manipulate and tweak any image of a face, whether it is ageing a fugitive, ageing a child who has been missing for 10 years or shaving years off a possible suspect's appearance, the role of the Forensic Artist is diverse.

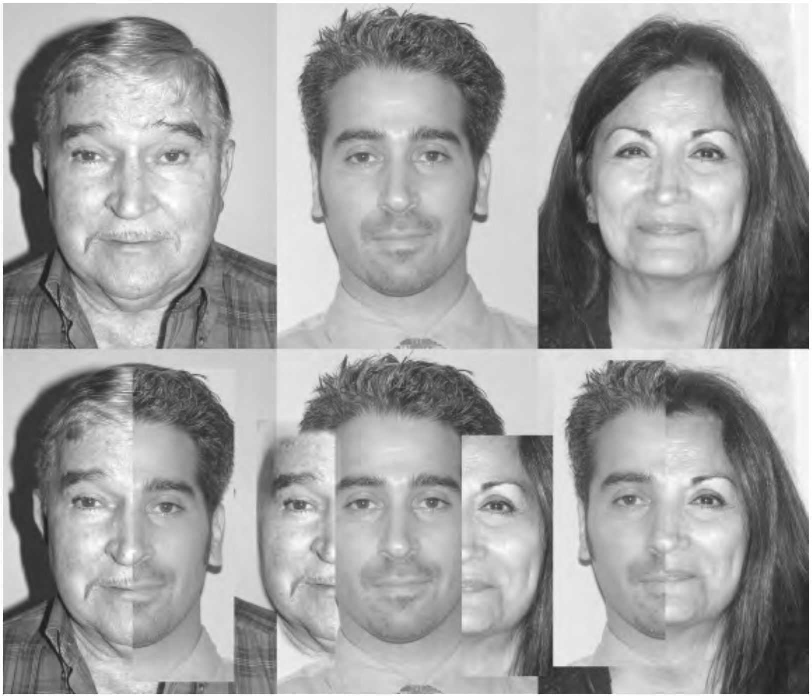


Figure 6.7 Images of biological parents (top left and right) and son (top centre) Note the strong resemblance of the son and father.

Figure 6.8 Age progression of adult female from age 18 to 34 years using siblings' features.

Figure 6.9 Adult male age progression shown (from left to right) at ages 40, 50, 70 and 90 years.

Figure 6.10 Adult female age progression.

Figure 6.11 Age regression of adult female.
Step 1: Increase size of iris.
Step 2: Make neck thinner and remove bags under eyes.
Step 3: Move lower two-thirds up.
Step 4: Adjust upper lids of each eye.
Step 5: Add younger hairstyle.
Step 6: Modify nose and add younger teeth.

Combining expertise in fine art and an extensive knowledge of growth, development and anatomy brings together art and science to serve law enforcement.

References

American Academy of Dermatology (2010). Aging Skin Net [online]. Available at: http://www.skincarephysicians.com/agingskinnet/basicfacts.html

Helfrich, Y. R., Sachs, D. L. and Voorhees, J. J. (2008). Overview of skin aging and photoaging. *Dermatology Nursing*, **20**(3), 177–183.

National Center for Missing and Exploited Children (NCMEC) (2011*a*). *Age Progression and Facial Reconstruction Services*. [online] Available at: http://www.missingkids.com/missingkids/servlet/ServiceServlet?LanguageCountry=enUS&PageId=174 [accessed 10 Jan 2011].

National Center for Missing and Exploited Children (NCMEC) (2011*b*). *FAQ: Statistics*. [online] Available at: http://www.missingkids.com/missingkids/servlet/PageServlet?LanguageCountry=en_US&PageId=2810 [accessed 10 Jan 2011].

Scheuer, L. and Black, S. M. (2000). *Developmental Juvenile Osteology*. London: Academic Press.

Ubelaker, D. H. (1978). *Human Skeletal Remains: Excavation, Analysis and Interpretation*. Chicago, IL: Aldine Transaction Publishers.

Chapter 7

Computer-assisted age progression

Stuart Gibson

7.1 Introduction

Age progression is a technique for increasing the perceived age of subjects' faces for the purpose of locating missing persons. This is normally accomplished by the manual manipulation of an archive image depicting the subject's face that was captured at some time in the past. Age progressions may be achieved by hand drawing or by skilful use of photo-editing computer software. Although forensic age progression has historically resided entirely in the domain of the forensic artist, potential improvements in accuracy[1] and execution time can be achieved by combining automated computer-based methods (developed by the researchers in the field of computer vision) with the traditional approach. It is unlikely that age progression could ever be totally automated as the role of the forensic artist extends beyond the task of transforming the face according to generic ageing traits. The forensic artist may use photographs of older siblings for guidance, be required to adapt the age-progressed image to reflect the lifestyle of the subject (e.g. living rough or abusing drugs) or apply bespoke additions such as distinctive clothing or facial hair.

This chapter describes a framework for computer-assisted age progression in which the perceived age of a face image may be increased by manually setting the values of three progression-control parameters that effect (i) changes to facial structure that are associated with generic ageing, (ii) genetic bias (due to similarity to blood relatives) and (iii) fine facial details including wrinkles. The method presented here has been implemented in the Matlab programming language. Future work may include a plug-in for a popular photo-editing package which would provide a valuable tool for forensic artists. In section 7.3, a basic method for generic age progression is outlined. Section 7.4 describes a method for adding genetic bias to the age progression when supporting images of relatives' faces are available for reference. Section 7.5 shows how the fine facial details associated with ageing may be extracted from a 'donor' face image and then applied to an age-progressed face, thereby enhancing the perceived age. Typical examples of age-progressed faces, obtained using the method described, are presented in section 7.6. The effectiveness of our approach in the context of the objectives of this work is discussed in the concluding section.

7.2 Background

A wide number of important applications exist for methods that accurately age images of the human face. Examples include the production of age-progressed images for tracing missing persons, predicting the effects of smoking and drug use on facial appearance, special effects in animation and film and cosmetic surgery planning. Automated ageing methods have also been used in fourth generation, facial composite systems (Davies and Valentine, 2007) based on evolutionary principles (Gibson *et al.*, 2003; Frowd *et al.*, 2004).

Chronological ageing refers to the gradual degeneration of tissues occurring throughout the body during an individual's lifetime, and is physically apparent in the skin, organs and organ systems. These physiological changes manifest at predictable stages, but also include inherited tendencies. Extrinsic ageing is caused by a number of external factors that are related to an individual's lifestyle. The abuse of drugs can accelerate the ageing process. A number of different effects can take place depending on the substance of

Craniofacial Identification, ed. C. Wilkinson and C. Rynn. Published by Cambridge University Press.

abuse. By smoking, an individual's body and skin is exposed to toxins. These toxins can affect the collagen and elastin present in lung tissue, and is believed to cause elastotic changes in skin. The skin may also become dehydrated (dry texture), and it has also been reported (Solly, 1856) that smokers generally have a more pale complexion, compared with non-smokers. This could be due to the fact that the cutaneous microvasculature is constricted by acute and long-term smoking, leading to ischaemia. Research by Ippen and Ippen (1965) showed that 79% of smokers had this complexion although 19% of non-smokers also exhibited this skin type. Another study by O'Hare *et al.* (1999) has shown that smoking was a main contributor to the development of wrinkles, however these only occurred in 6% of the sample, and there was no significant difference between smokers and non-smokers. Stress as well as sleep deprivation can accelerate the ageing process and cause extended premature wrinkle development. Anxiety and/or stress cause muscular tension, resulting in wrinkles becoming more pronounced, especially in areas such as the forehead and eyebrows. This type of extrinsic ageing can be related to the ageing caused by drugs and alcohol, as these substances can have an effect on the sleeping pattern of the individual. Again, it is difficult to assess the extent of ageing caused, due to factors being interrelated.

Quantitative approaches to age progression began with the work by Pittenger and Shaw (1975, 1979) on the cardioidal strain transform. This is a non-linear topological transformation which gives a plausible mathematical model for the global changes in human face shape due to growth. A revised cardioidal strain transformation model has since been proposed by Ramanathan and Chellappa (2006) which is capable of predicting changes in facial appearance due to increasing age and performing face recognition independent of age. Burt and Perrett (1995) proposed an ageing method which operated with both the shape and texture of the face. In this work, a simple 'prototype' algorithm was implemented in which the shape and colour differences between prototypes of a young age group and an older age group were added to the subject images to increase the perceived age. In a medical study of craniofacial development, Hutton *et al.* (2003) defined an ageing trajectory through a model space produced by the construction of a shape model based on 3D facial meshes. Kernel smoothing was used to estimate average faces at given ages and the trajectory was defined by the path between these averages. Using this trajectory, examples within the space were reconstructed at a target age. Lanitis *et al.* (2002), Scandrett *et al.* (2006); Hill *et al.* (2004, 2005) and Blanz and Vetter (1999) have all employed methods based on the generation of appearance models in which the age of the subject and its location within the model space allow a variety of statistical learning or fitting methods to be applied to estimate an ageing function. Once a suitable ageing function has been established, the model parameters may be modified to calculate a new set of parameters at the target age. Gandhi (2004) trained a Support Vector Machine (SVM) to derive an accurate relationship between a facial image and the age of the subject, which is referred to as the age prediction function. Wang and Ling (2004) also used SVMs to automatically synthesise the ageing effect on human face images. A comprehensive, comparative evaluation of automatic age-progression methodologies is provided by Lanitis (2008). Despite the considerable advances made in the work cited above, an automatic system for age progression, which can perform the diverse range of graphic enhancement tasks normally executed by a forensic artist, has yet to make a large impact within the domain of forensic age progression.

The majority of the methods for automatic age progression 'learn' the changes in facial appearance associated with increasing age by observing the average trend over a sample of example faces of various ages. Ageing models that rely on averaging sample images do not adequately represent the fine facial details, such as wrinkles and skin blemishes, associated with the ageing process. The problem can be attributed to the poor subject-to-subject spatial correlations of fine facial details. The body of work that attempts to address the issue of enhancing fine detail is limited. Tiddeman *et al.* (2001) took a prototype face, formed by averaging a sample of face images, and decomposed it using a wavelet analysis. The resulting high-order details were boosted to compensate for the inherent loss of high spatial frequency information. Gibson *et al.* (2009) show how the fine facial details associated with ageing may be extracted from a 'donor' face image (which they refer to as a wrinkle-map) and then applied to an age-progressed face thereby increasing the realism and perceived age. The technique described is a variation on the high-boost filter that is frequently used in digital image processing applications and is similar in principle to a practice adopted by some forensic artists in which fine facial details are copied from one face into another using a graphics package. This procedure is acceptable since the identity of a subject is, to a good approximation, determined by the structure of a face and its internal features rather than the modest blemishes and wrinkles that can affect the skin.

7.3 Computer-assisted method for ageing faces

The approach taken is to 'learn' how faces age by observing a general trend over a sample of training images, depicting subjects of different ages. Initially, shape and texture models are constructed using a method that closely resembles the work by Cootes and Taylor (2001). Here the term *texture* is used to describe a set of pixel intensity values obtained from a shape-normalised digital image of the suspect's face. The term *shape* refers to a vector of landmark points that delineate the perimeter of a subject's face and the boundaries of the main facial features (see Figure 7.1). The texture of a face can be represented by a vector **t** that points to a location in a high dimensional pixel-space. Making use of the similarities in basic facial geometrical structure between subjects, a more compact representation of texture can be obtained using principal components analysis (PCA) in which the texture can be written as a vector, **g**, of much smaller dimension than **t**. This is an optional procedure that reduces computer-processing time. There is some evidence which suggests that PCA-based models of facial appearance share similarities with observations made on the way in which the brain encodes faces (Hancock *et al.*, 1996). The procedure is also used to obtain a compact representation, **b**, of a subject's face shape. Prior to applying the PCA, each face shape is aligned, scaled and rotated to the mean face shape through an iterative process called Procrustes alignment (see Stegmann 2000 for details).

Any shape vector within the training set may then be described as

$$\mathbf{x} = \bar{\mathbf{x}} + \mathbf{P}_s\mathbf{b} \qquad (7.1)$$

where $\mathbf{x}$ is the shape data vector, $\bar{\mathbf{x}}$ is the mean shape data vector, the columns of the matrix $\mathbf{P}_s$ are the orthogonal modes of variation or principal components (PCs) and $\mathbf{b}$ is a column vector of shape model parameters that combine the PCs in the correct proportion. To a good approximation, the shape of any face that is not included in the training sample may also be expressed by Equation 7.1.

The texture model is created by first warping all training example images to the mean face shape using a piece-wise, affine transform. By warping to the mean face shape (shape-normalised frame of reference), a near-perfect registration is achieved between corresponding facial features over the sample images. The pixel values from each training example are then extracted to form an ensemble of shape-normalised texture vectors (Figure 7.2). PCA is again applied to produce a set of texture principal components. Any

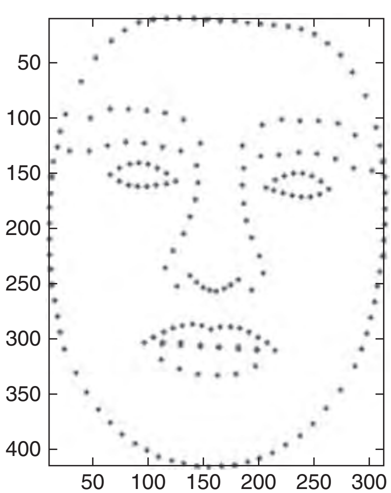


Figure 7.1 Points used to delineate the face shapes of the training sample.

(a) example face

(b) example face in the shape-normalised reference frame

Figure 7.2 The texture model is constructed from shape-normalised example images. Each sample image is warped to the mean face shape using Delaunay triangulation and a piece-wise affine transformation.

example of a texture vector within the training set may then be described as a sum of principal components using an equation similar in form to Equation 7.1:

$$\mathbf{t} = \bar{\mathbf{t}} + \mathbf{P}_t\mathbf{g} \tag{7.2}$$

where **t** is the texture data vector, $\bar{\mathbf{t}}$ is the mean texture data vector, the matrix $\mathbf{P}_t$ contains the texture PCs and **g** is a set of texture model parameters that combine the PCs in the correct proportion. The two sets of model parameters, **b** and **g**, are relatively compact and enable us to conveniently visualise a given facial appearance as a point in a multidimensional vector space of effective dimension < 100.

To calculate an ageing axis, $\mathbf{V}_{age}$, through the model space (applies separately to shape and texture models), the model parameters of an individual face are weighted by the corresponding mean-subtracted age and a sum formed over all training examples.

$$\mathbf{V}_{age} = \frac{\sum_{i=1}^{N}\left(W_i - W_\mu\right)\left(\mathbf{b}_i - \mathbf{b}_\mu\right)}{\sum_{i=1}^{N}\left(W_i - W_\mu\right)} \tag{7.3}$$

where N is the number of training examples, W_i is the age of the ith example, W_μ is the mean age of all training examples, $\mathbf{b}_i$ is the set of shape parameters of the ith face and $\mathbf{b}_\mu$ is the set of mean shape parameters. A simplified, schematic, 2D representation of the ageing axis is shown in Figure 7.3. A scalar multiple, α, defining the increase in age in years, of the ageing axis $\mathbf{V}_{age}$ may then be added to the model parameters, **b**:

$$\mathbf{b}' = \mathbf{b} + \alpha\mathbf{V}_{age} \tag{7.4}$$

where $\mathbf{b}'$ is the set of aged shape parameters. Replacing **b** with $\mathbf{b}'$ in Equation 7.1 gives the aged face shape. Hence, if the subject is depicted at y_1 years old in the input then the perceived age after progression will be $y_2 = y_1 + \alpha_1$. We will refer to α as the progression-control-parameter that determines an increase in age in units of years. Negative values of α will cause the subject's face to appear younger.

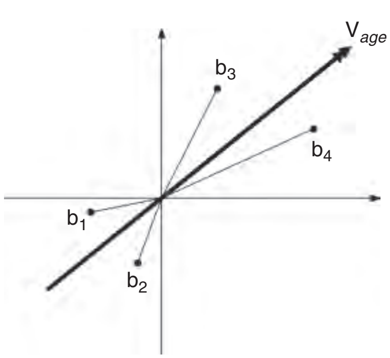


Figure 7.3 The shape vectors, $\mathbf{b}_1$ to $\mathbf{b}_N$, are weighted by their corresponding ages and summed to form the ageing axis, $\mathbf{V}_{age}$.

Sexual dimorphism in the ageing of humans (Enlow and Hans, 1996) would seem to indicate a possible need to consider the two sexes separately. Hence, separate ageing axes should be defined for male examples and female examples within the shape space. This dimorphic approach is also applied to the texture. An individual is then aged using the either the male axis or female axis according to their sex. Once ageing has proceeded according to these models, the resultant aged texture is then warped from the shape-normalised frame of reference to the aged shape using a Delaunay-based, piece-wise, affine transformation (Stegmann, 2000) thereby producing an output face aged in both shape and texture. Scandrett *et al.* (2006) offer an alternative method that uses a weighted combination of both male and female axes.

7.3.1 Stages of facial development – piece-wise axes

The effects of ageing on facial appearance do not occur in a uniform fashion (Hutton *et al.*, 2003). For example, puberty causes significant changes to the face over a short period of time (Enlow and Hans, 1996). Bogin identifies four stages of development: infancy (0 and 3 years – a period of rapid growth); childhood (between 3 and 7 years) when more moderate growth takes place; the juvenile stage (between 7 and 10 years), characterised by slower growth; and, finally adolescence (between 10 and 20 years) when periods of rapid growth occur (Bogin, 1999). To account for differences and rate of growth between childhood and adulthood, we considered a comparative technique in which 'piece-wise' ageing axes were calculated using faces restricted to specified age ranges. Two axes were calculated for each sex and the model parameters defined by adapting Equation. 7.4 as follows:

$$\mathbf{b}' = \begin{cases} \mathbf{b} + \alpha_1\mathbf{V}_{6\to20}, & y_1 \geq 6, \quad y_2 < 20 \\ \mathbf{b} + \alpha_1\mathbf{V}_{6\to20} + \alpha_2\mathbf{V}_{20+}, & 6 \leq y_1 < 20, \quad y_2 \geq 20 \\ \mathbf{b} + \alpha_2\mathbf{V}_{20+}, & y_1 \geq 20, \quad y_2 \geq 20 \end{cases} \tag{7.5}$$

where $\mathbf{V}_{(6\to20)}$ is the ageing trajectory for facial development between the ages of 6 and 20 and $\mathbf{V}_{20+}$ is the trajectory for age increases during adulthood. The sum of the progression-control parameter α and the age of the subject, y_1, as depicted in the input image must not exceed 20. Accordingly, for progressions in which the subject is to be aged from childhood to $y_2 > 20$, the additional conditions $\alpha_1 = 20 - y_1$ and $\alpha_2 = y_2 - 20$ apply.

7.4 Extension of the ageing model to reflect genetic bias

In order to age a face in the most accurate manner, we must consider not just average tendencies over a peer-group sample but also person-specific factors. Person-specific factors include, for example, previous (i.e. historical) facial development trends for the given individual, which may reasonably be expected to have some correlation with future growth and development patterns, and the facial appearance and development of close relations (in particular, parents and siblings). Person-specific factors such as these are ultimately biological in origin but no attempt to build a predictive model based on directly measurable biological factors such as DNA is currently possible. Instead, a weighted combination is formed comprising the aged face (obtained using the method described in section 7.3) and an age-adjusted photograph of a biological relative.

$$\mathbf{x}_{aged} = (1-\beta)(\bar{\mathbf{x}} + \mathbf{P}_s\mathbf{b}') + \beta\mathbf{x}_{bio} \quad 0 \le \beta \le 1 \quad (7.6)$$

$$\mathbf{t}_{aged} = (1-\beta)(\bar{\mathbf{t}} + \mathbf{P}_t\mathbf{g}') + \beta\mathbf{t}_{bio} \quad 0 \le \beta \le 1 \quad (7.7)$$

β is a progression-control parameter that determines the relative contributions from the ageing axis and the biological relative; $\mathbf{x}_{bio}$ and $\mathbf{t}_{bio}$ are the shape and texture contributions respectively from the relative that have first been aged to the target age using the method described in section 7.3. Here we suggest that an appropriate value for β is chosen by the forensic artist to reflect the similarity between the subject and their relative although the weighting may alternatively be determined automatically using statistical means (Scandrett *et al.*, 2006). It is desirable to use a relative of the same sex and of a similar age to the target age for the progression.

7.4.1 Lifestyle-specific ageing

In section 7.3 no reference was made to the lifestyle led by the subjects who comprised the training sample from which the shape and texture models were built. We may expect that by replacing the training sample with images of subjects who have led a particular lifestyle that our model will represent lifestyle-specific ageing traits. However, such data are scarce and therefore difficult to obtain. Where images of subjects are available before or at early stages of exposure to the extrinsic ageing factor and also after prolonged exposure, before and after prototypes can be formed using a slight variation on the method originally described by Burt and Perrett (1995) which can be considered a simplification of the method outlined in section 7.3. Here the lifestyle-ageing progression is calculated directly on the face shape and texture without using PCA.

Clarke *et al.* (2010) used this method to model the effects of prolonged methamphetamine use on facial appearance as follows: Twenty images (eight female, 12 male) supplied by the Multnomah County Sheriff's Office were used as training examples to create prototype face images. The 'before' prototype was constructed from the images of 10 subjects in the early stages of addiction. An 'after' prototype was constructed from images of the same 10 subjects after prolonged use of the drug spanning a few months or years. The apparent exposure of a veridical (true/unaltered) face of any healthy subject to the extrinsic ageing factor (in this example drug use) was then realised by adding to it some proportion of the difference between prototypes, and $\Delta\bar{S}_p$ and $\Delta\bar{T}_p$ are the differences between prototypes for shape and texture respectively. If the veridical face-shape and face-texture of the subject are denoted by S_k and T_k respectively, then the lifestyle-specific aged face-shape and face-texture are given by

$$S_{(k,aged)} = S_k + \alpha\Delta\bar{S}_p$$
$$T_{(k,aged)} = T_k + \alpha\Delta\bar{T}_p$$

where α is a scaling parameter that is controlled manually and determines the strength of the effect. In the final step the morphed face-image, $I_{(p,s)}$, was constructed by warping $T_{(k,aged)}$ to the modified face-shape, $S_{(k,aged)}$. An example is provided in Section 7.6.

7.5 Applying fine details to an aged face

The method described so far in this paper is capable of generating age progressions that characterise the large-scale facial geometry, changes associated with ageing. However, the fine structures such as wrinkles and liver spots are underrepresented in age-progressed images. This shortfall is particularly apparent when progressions to old age are required since fine facial detail becomes more prevalent with an increase in age. The problem is intrinsic to the process by which the ageing vector $\mathbf{V}_{age}$ is defined. Although the occurrence of wrinkles is more common in specific regions of the face (for example 'crow's feet' appear at the outer corners of the eyes), their exact positions vary from subject to subject and are therefore 'averaged out' when $\mathbf{V}_{age}$ is calculated.

7.5.1 Extracting wrinkle-maps from sample images

In the remainder of this section, an ad hoc method for applying fine facial detail to a target face is discussed. The processing is assumed to take place in the shape-normalised reference frame. The idea is to extract the fine facial details from a sample, or 'donor', face A with the intention of applying these details directly to a target face B. Let T_A be the texture (shape-normalised face-image) of subject A and let T_{Ad} be a detail-image containing the fine facial detail from T_A. If A is an elderly subject, then texture T_{Ad} will be likely to contain facial wrinkles that are associated with the ageing process. The term *wrinkle-map* will be used when referring to T_{Ad}. A wrinkle-map is obtained by applying an image processing filter to T_A that extracts the high spatial frequencies within the image, or equivalently by subtracting a smoothed version of the image from itself as follows

$$T_{Ad}(x,y) = T_A(x,y) - \bar{T}_A(x,y) \qquad (7.8)$$

where $\bar{T}_A(x,y)$ is a blurred version of T_A. The procedure is illustrated schematically in Figure 7.4.

7.5.2 Applying details to a target face

Once extracted from a 'donor' subject (for example subject A in Figure 7.4), a wrinkle-map can be applied to a target face thereby increasing its fine detail content. This technique is particularly useful for increasing the apparent age of a subject. The method can be considered as a hybrid form of the standard *high-boost filtering* method commonly used in digital image processing (here the definition of high-boost filtering provided by Gonzalez and Woods, 2002 is used):

$$T_{Bhb} = (\gamma - 1)\,\bar{T}_B + T_{Bd},\ \gamma \geq 1 \qquad (7.9)$$

$$T'_{Bhb} = (\gamma - 1)\,\bar{T}_B + T_{Ad},\ \gamma \geq 1 \qquad (7.10)$$

In the normal implementation of high-boost filtering, a detail-image is constructed as defined by Equation 7.8 and added to a fraction of the original image (Equation 7.9). This procedure makes the fine details more prominent in the high-boost image T_{Bhb}. Equation 7.10 is a modified version of the standard high-boost procedure, whereby the detail image is not formed from the original image, but from a different image instead. As the value of the scalar A increases beyond 1, the contribution of the detail-image becomes less important. In practice $\gamma \approx 2$, and the filter kernel used to form $\bar{T}_A$ in Equation 7.8 should be small enough that only fine details are copied to subject B. If these precautions are *not* adhered to the identity of subject B may be altered when the wrinkle-map is applied. The *hybrid high-boost* image that results, T'_{Bhb}, comprises the low-medium spatial frequencies of image T_B and the high spatial frequencies belonging to image T_A. A schematic overview of this hybrid high-boost method is provided in Figure 7.5. Superposition of fine details from *both* subjects can produce undesirable ghosting artefacts in the hybrid high-boost image T'_{Bhb}. Applying an image-smoothing filter to T_A prior to adding the detail image overcomes this problem. For a practical implementation we propose that an appropriate wrinkle-map is chosen by the forensic

Figure 7.4 Wrinkle-map T_{Ad} extracted from a sample face A using image subtraction 1.8. In this example, T_{Ad} was created by using an 8×8 averaging filter to form a smoothed image $\bar{T}_A$, which was then subtracted from the $340 \times 267 \times 3$ original T_A.

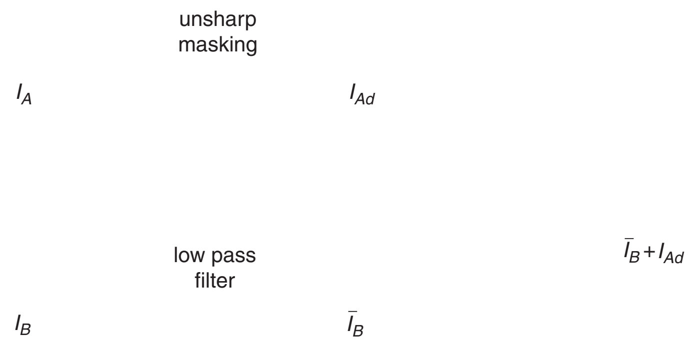


Figure 7.5 Application of a wrinkle-map obtained from subject *A* to a target face *B*.

(a) $A = 2$, filter size $= 8 \times 8$ (b) $A = 2.5$, filter size $= 8 \times 8$

Figure 7.6 In the interest of ease of implementation, the prominence of each wrinkle-map is determined by the relative proportions of target face and wrinkle-map. The relative proportions can easily be controlled by varying the scalar *A* in the hybrid high-boost Equation 7.10.

artist from a database of candidate wrinkle-maps and applied to the age-progressed image (Figure 7.6). In this context $\bar{T}_B$ is the output image obtained by following the method described in section 7.3 (and section 7.4 when appropriate). We define γ as the third progression-control parameter that determines the strength of the wrinkle-map withing the age-progressed image.

7.6 Examples of age-progressed faces

Consider first Figures 7.7–7.9. For each sex the $\mathbf{V}_{6\to 20}$ axis was determined using $N = 30$ example faces in the age range $6 \to 20$ according to Equation 7.3. Similarly, $\mathbf{V}_{20+}$ ageing axes were determined for males and females using $N = 170$ and $N = 250$ examples respectively. All sample faces were drawn from the population of White Europeans, reflecting the ethnicity of the subjects to be aged. Figures 7.7, 7.8 and 7.9 depict age progressions of three different subjects from 7 years old to 17 years old (19 for subject A) using the $\mathbf{V}_{6\to 20}$ axis for males and females as appropriate. In each example the age progression was influenced by the facial appearance of the blood relative shown. The Genetic bias (progression-control parameter, β) was chosen according to visual assessment of the facial similarity between the subject and their relative. The relatives were initially age progressed/regressed to 17/19 years old using the method described in section 7.3. In these examples the age progression is determined solely by the progression-control parameters α and β and no wrinkle-maps are required to achieve a plausible and accurate result. Veridical (actual) images of the sujects at 17/19 were obtained after the age progressions were generated and these have been included in Figures 7.7, 7.8 and 7.9 for comparison.

The age progression of a subject from 28 to 60 years using the $\mathbf{V}_{20+}$ axis alone is shown in Figure 7.10b. Differences between the face shape and subtle differences in tone between this image and the image of the subject at age 28 can be observed (Figure 7.10a). However, the perceived age of the subject in Figure 7.10b is less than 60 years due to under-representation of the fine facial details associated with ageing. The addition of a wrinkle-map, as described in section 7.5, addresses this problem and the progression achieves a plausible representation of the subject at 60 years old. In the final aged image (Figure 7.10c), grey streaks have been added to the hair manually using a proprietary graphics package.

(a) subject A age 7 (b) brother of subject A age 19 (c) age progression of subject A to 19 (d) veridical image of subject at 19

Figure 7.7 Subject aged to 19 years old using linear trajectory and genetic bias. Progression-control parameters are $\alpha_1 = 10$, $\beta = .25$ and $\gamma = 0$.

(a) subject B age 7 (b) brother of subject B age 14 (c) age progression of subject B to 17 (d) veridical image of subject at 17

Figure 7.8 Subject aged to 17 years old using linear trajectory and genetic bias. Progression-control parameters are $\alpha_1 = 10$, $\beta = .20$ and $\gamma = 0$.

(a) subject C age 7 (b) mother of subject C age 18 (c) age progression of subject C to 17 (d) veridical image of subject at 17

Figure 7.9 Subject aged to 17 years old using linear trajectory and genetic bias. Progression-control parameters are $\alpha_1 = 10$, $\beta = .30$ and $\gamma = 0$.

Figure 7.10 Age progression using the V_{20+} axis for female faces. Progression-control parameters are $\alpha_2 = 32$, $\beta = 0$ and $\gamma = 1.5$.

Figure 7.11 Lifestyle specific ageing: (a) Veridical face; (b) Face modified to mimic prolonged methamphetamine use.

Figure 7.11a shows a true likeness of a healthy individual. Figure 7.11b contains an image of the same subject modified, using the lifestyle-specific ageing model, to suggest prolonged use of methamphetamine. The age/lifestyle-progressed image exhibits a more gaunt appearance, and paler complexion, than the subject's true appearance but identity is retained. Fine structures such as wrinkles and other forms of heterogeneous complexion (e.g. sores) associated with methamphetamine users are underrepresented, causing the skin complexion of the prototype face to be unrealistically homogeneous for the duration of exposure.

7.7 Summary and conclusions

In this paper, we have presented a computer-assisted method for changing the perceived age of face within a 2D digital image. Piece-wise ageing trajectories were calculated and these were shown to provide a realistic mechanism for achieving age progressions from childhood to adulthood. Lifestyle-specific ageing can also be modelled using the same principle. The method was enhanced by the addition of wrinkle-maps that captured the fine facial details associated with increasing age during adulthood. We have shown, indirectly, that fine facial details such as wrinkles play an important role in the perception of age but do not affect identity. In this work manual control over the age progression can be attained by selecting appropriate values for the progression-control parameters, α, β and γ. We speculate that these parameters may be controlled by sliders in future computer software based on the method described. The current method does not support automatic changes to hair thickness, style and colour or indeed any alterations to any region of the image that lies between the perimeter of the face and the perimeter of the image (see the neck region of the subject depicted in Figure 7.10). Therefore we refer to this approach as computer assisted and suggest that effects such as greying or thinning of hair is achieved by the forensic artist using a proprietary photo-editing package. However, we anticipate that a software tool based on this work could be used by forensic artists to significantly reduce the amount of time required to produce age progressions.

References

Blanz, V. and Vetter, T. (1999). A morphable model for the synthesis of 3D faces. In *SIGGRAPH'99 Conference Proceedings*, pp. 187–194.

Bogin, B. (1999). *Patterns of Human Growth*, 2nd Edn. Cambridge: Cambridge University Press.

Burt, D. M. and Perrett, D. I. (1995). Perception of age in adult caucasian male faces: computer graphic manipulation of shape and colour information. *Royal Society Proceedings B*, **259**, 137–143.

Clarke, D., Riggs, M. and Gibson, S. (2010). Prototyping perceptions of health for inclusion in facial composite systems. In G. Howells, K. Sirlantzis, A. Stoica, T. Huntsberger and T. Arslan (Eds.), *International Conference on Emerging Security Technologies* (EST 2010). IEEE Conference Publishing Services.

Cootes, T. and Taylor, C. (2001). Statistical models of appearance for medical image analysis and computer vision. In *SPIE Medical Imaging*, pp. 236–248.

Davies, G. M. and Valentine, T. (2007). Facial composites: forensic utility and psychological research. In *Handbook of Eyewitness Psychology. Vol. 2: Memory for People*, pp. 59–83. Mahwah, NJ: Lawrence Erlbaum Associates.

Enlow, D. H. and Hans, M. G. (1996). *Essentials of Facial Growth*. London: W.B. Saunders Company.

Frowd, C. D., Hancock, P. J. B. and Carson, D. (2004). Evofit: A holistic, evolutionary facial imaging technique for creating composites. *ACM Transactions in Applied Perception*, **10** (1), 19–39.

Gandhi, M. (2004). *A Method for Automatic Synthesis of Aged Human Facial Images*. Master's thesis, Department of Electrical and Computer Engineering, McGill University, Montreal, QC, Canada.

Gibson, S. J., Scandrett, C., Solomon, C., Wilkinson, C. J. and Maylin, M. I. S. (2009). Computer assisted age progression. *Journal of Forensic Science, Medicine and Pathology*, **5**, 174–181.

Gibson, S. J., Solomon, C. J. and Pallares Bejarano, A. (2003). Synthesis of photographic quality facial composites using evolutionary algorithms. In R. Harvey and J. A. Bangham (Eds.), *British Machine Vision Conference 2003*, Vol. **1**, pp. 221–230.

Gonzalez, R. C. and Woods, R. E. (2002). *Digital Image Processing*, 2nd Edn. Upper Saddle River, NJ: Prentice Hall.

Hancock, P. J. B., Burton, A. M. and Bruce, V. (1996). Face processing: human perception and principal components analysis. *Memory and Cognition*, **240**(1), 26–40.

Hill, C. M., Solomon, C. J. and Gibson, S. J. (2004). Plausible aging of the human face using a statistical model. *Proceedings of the 7th International Conference on Computer Graphics and Imaging*, **270**, 56–59.

Hill, C. M., Solomon, C. J. and Gibson, S. J. (2005). Aging the human face – a statistically rigorous approach. In *IEE International Symposium on Imaging for Crime Detection and Prevention*, pp. 89–94.

Hutton, T. J., Buxton, B. F., Hammond, P. and Potts, H. W. W. (2003). Estimating average growth trajectories in shape-space using kernel smoothing. *IEEE Transactions on Medical Imaging*, **220**(6), 47–753.

Ippen, M. and Ippen, H. (1965). Approaches to a prophylaxis of skin aging. *Journal of the Society of Cosmetic Chemists*, **16**, 305–308.

Lanitis, A. (2008). Comparative evaluation of automatic age-progression methodologies. *EURASIP Journal of Advances in Signal Processing*, **8**(2), 1–10.

Lanitis, A., Taylor, C. and Cootes, T. (2002). Toward automatic simulation of aging effects on face images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **240** (4), 442–455.

O'Hare, P., Fleischer, A., D'Agostino, R., Feldman, S., Hinds, M., Rassette, S., McMichael, A. and Williford, P. (1999). Tobacco contributes little to facial wrinkling. *Journal of the European Academy of Dermatology and Venereology*, **12**, 133–139.

Pittenger, J. and Shaw, R. (1975). Aging faces as visual-elastic events: Implications for a theory of nonrigid shape perception. *Experimental Psychology: Human Perception and Performance*, **10**(4), 374–382.

Pittenger, J. and Shaw, R. (1979). Perceptual information for the age level of faces as a higher order invariant of growth. *Experimental Psychology: Human Perception and Performance*, **50**(3), 478–493.

Ramanathan, N. and Chellappa, R. (2006). Modeling age progression in young faces. In *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference*, pp. 387–394.

Scandrett, C. M., Solomon, C. J. and Gibson, S. J. (2006). A person-specific, rigorous aging model of the human face. *Special Issue Pattern Recognition Letters: Vision for Crime Detection and Prevention*, **270**(15), 1776–1787.

Solly, S. (1856). Clinical lectures on paralysis. *Lancet*, **2**, 641–643.

Stegmann, M. (2000). *Active Appearance Models: Theory, Extensions and Cases*. Master's thesis, Department of Mathematical Modelling, Technical University of Denmark, Lyngby, Denmark.

Tiddeman, B. P., Perrett, D. I. and Burt, D. M. (2001). Prototyping and transforming facial textures for perception research. *IEEE Computer Graphics and Applications*, **210**(5), 42–50.

Wang, J. and Ling, C. X. (2004). Artificial aging of faces by support vector machines. In A. Y. Tawfik and S. D. Goodwin (Eds.), *Advances in Artificial Intelligence, Lecture Notes in Computer Science*, **3060**, 499–503. Berlin: Springer.

Note

1. Although numerical measures of accuracy are not necessarily the same as perceptual accuracy.

Chapter 8

Facial recognition from identification parades

Catriona Havard and Amina Memon

8.1 Introduction

Identification parades are one of the most common means of identifying a perpetrator of a crime and can be powerful evidence in securing convictions in criminal cases. In an identification parade (also known as a line-up) a suspect is placed amongst a number of similar-looking people (foils) and the task for the witness is to either select the person they recognise as being the culprit, or state the culprit is not there. Unfortunately, witnesses do not always correctly identify the culprit from a line-up and in some cases innocent people are wrongly identified. This issue has been investigated by the Innocence project, who at the time of writing, have been involved in 251 exonerations based on DNA evidence in the USA. Of these cases approximately 75% of those convicted were cases of mistaken identity (Innocence Project at: http://www.innocenceproject.org/). Several similar organisations have developed worldwide such as the UK Innocent Network, the Australian Innocence Network and the Innocence Project New Zealand, all with the aim to overturn convictions of those who have been wrongfully imprisoned. One real life case of someone who was exonerated by the Innocent Project is Calvin Willis; he was wrongly accused of rape and served 22 years before DNA evidence proved he was innocent. Even though one of the police reports stated the victim did not see her attacker's face, she was shown a line-up and said she was told to pick the man without the full beard. She later testified that she did not choose anyone, although the police said she picked Willis. Willis's name only came into the investigation because the victim's neighbours had mentioned Willis's name when discussing who might have committed the crime. Even with all these flaws in the investigation the jury still convicted Willis and sentenced him to life (for more details of this case see Innocence Projects webpage at: http://www.innocenceproject.org/Content/297.php). This illustrates how powerful eyewitness evidence can be in the courtroom, even when there are obvious flaws in the way it has been obtained.

This chapter will explore a number of issues that can influence how accurately witnesses make decisions when viewing a line-up. These issues can be split into two separate strands: system variables and estimator variables (Wells, 1978). System variables are factors that are under control of the police conducting the parade, and include the method of foil selection and the identification format e.g. photo or video line-up. They also include whether the line-up presentation is sequential or simultaneous, instructions given to the witness, and whether they view a mugbook or make a composite prior to viewing a line-up. Estimator variables are factors associated with the witnesses and their view of the crime and are not under any control from the police or the judicial system. These include the witness's age and eyesight, how long they saw the culprit for and also whether the culprit was the same age or ethnic background as the witness. We will examine how system and estimator variables can influence facial identification from line-ups, but first the method used to study the factors influencing eyewitness identification accuracy will be described.

8.2 The eyewitness paradigm

Studies investigating how eyewitnesses make identifications from line-ups usually follow the same methodology. Initially the witnesses view some type of staged event (e.g. a non-violent theft), or a video. Witnesses

Craniofacial Identification, ed. C. Wilkinson and C. Rynn. Published by Cambridge University Press.

are usually unaware of the true purpose of the study until after it has taken place. Then having viewed the event they may carry out a filler task, or be interviewed, or a longer delay may be implemented prior to carrying out a line-up. Researchers typically use two types of line-up: target present (TP) line-ups where the culprit is present and target absent (TA) line-ups where the culprit is not present. Target absent line-ups are used to simulate the situation where the police have arrested the wrong person and also to examine choosing biases (Wells, 1993), which will be explored in more detail later. In a TA line-up the culprit is replaced by someone who fits the culprit's description, or general appearance. For TP line-ups there are three possible responses: a correct identification where the witness identifies the culprit; an incorrect rejection (miss) where the witness states the culprit is not there; and a foil identification, where the witness chooses someone other than the target. For TA line-ups there are only two responses that can be made: a correct rejection where the witness states that the culprit is not there, or a false identification, where the witness chooses someone from the line-up.

In their review of eyewitness testimony, Wells and Olson (2003) suggest that when the police carry out a line-up it is similar to the way researchers conduct an experiment. 'Crime investigators begin with a hypothesis (that the suspect is the culprit), create a design for testing the hypothesis (embed the suspect among fillers), carry out a procedure (e.g., provide pre-line-up instructions and present the group to an eyewitness), observe and record the eyewitness's behavior (witness decision), and then interpret and revise their hypothesis (whether the suspect is the culprit)' (p. 285). They then go on to suggest that there are various things that can go wrong with the experiment that can bias the line-up or the witness; some of these issues will be explored later in this chapter.

8.3 Line-up formats

In the USA, the majority of line-ups are presented using static photographs (Wells *et al.*, 2007). There are a number of advantages of using photographic line-ups as compared with live line-ups. Firstly, there is less chance that the line-up will be cancelled due to either the suspect or witness failing to attend, or due to a lack of volunteers serving as foils in the line-up. Secondly, a database of images can be accumulated so that foils can be more closely matched to the appearance of any suspect, rather than trying to find suitable stand-ins available on the day. Thirdly, showing a photographic line-up may be less stressful for a witness than facing the accused in a live setting (Pike *et al.*, 2000).

Photographic line-ups can be presented either simultaneously or sequentially. The more traditional form is the simultaneous line-up (Wogalter *et al.*, 2004), where a number of face images are shown simultaneously in an array, and the witness views all the images at the same time and then makes a decision about whether the culprit is one of them. It has been suggested that the simultaneous line-up procedure encourages the witness to make a relative judgement and choose the person who looks most like the culprit, and thereby increase the false identification rate if the culprit is not present in the line-up (Wells, 1984; Lindsay and Wells, 1985; Lindsay *et al.*, 1991 Lindsay and Bellinger, 1999; Dysart and Lindsay, 2001; Kneller *et al.*, 2001). Evidence for the relative judgement comes from 'a removal without replacement procedure', where witnesses are shown a staged crime and see either a TP line-up, or a TA line-up which has exactly the same line-up members, except that the culprit has been removed. For the TP line-up, 54% of witnesses correctly identified the culprit, however in the TA line-up where the culprit had been removed, 68% of witnesses still incorrectly chose someone from the line-up (Wells, 1993).

To reduce the use of relative judgements and thereby reduce false positive choices, Lindsay and Wells (1985) designed the sequential line-up procedure. In the sequential line-up procedure, the witness is unaware of the number of line-up members he or she is about to see and views one line-up member at a time. When presented with each line-up member the witness must decide whether the person is the perpetrator prior to seeing the next person. Once a decision is made they are not allowed to view that person again. The premise is that by looking at each line-up member individually, this should discourage the witness simply choosing the person that looks most like the perpetrator as they have to make more of an 'absolute' judgement to each face, e.g. 'is this the perpetrator or not'. The line-up finishes either when the witness identifies one of the line-up members as being the culprit, or when all the line-up members have been viewed. Additional features were suggested in the design of the sequential line-up, such as only single suspect line-ups, double-blind testing, confidence statements prior to any feedback and using only foils that matched the suspect's description.

A number of studies have compared identification rates for simultaneous versus sequential line-ups.

Research that compared the false identification rate for TA line-ups, found that they were significantly greater for simultaneous line-ups (43%) as compared with sequential line-ups (17%). The correct identification rate for the TP line-ups was not significantly lower for the sequential (50%) versus the simultaneous line-ups (58%) (Lindsay and Wells, 1985). Other research has also reported a reduced false identification rate for sequential line-ups as compared with simultaneous line-ups, without any reduction in the correct identification rates (Sporer, 1993; Kneller *et al.*, 2001). However, not all research has found a consistent advantage for sequential line-ups over simultaneous line-ups. Children are more likely than adults to make a false identification from a TA line-up (Parker and Ryan, 1993; Lindsay *et al.*, 1997; Pozzulo and Lindsay, 1998; Memon and Gabbert, 2003), and sequential line-ups do not reduce false identification rates for children (Lindsay *et al.*, 1997; Pozzulo and Lindsay, 1998).

A meta-analysis by Steblay *et al.* (2001) found that although false identifications for TA line-ups were reduced for sequential line-ups (28%) as compared with simultaneous line-ups (51%), correct identifications from TP line-ups were significantly lower for sequential line-ups (35%) when compared with simultaneous line-ups (50%). However, once moderator variables that approximated real-world conditions were considered, this difference disappeared and the authors concluded that correct identification rates for simultaneous and sequential line-ups were likely to be small or non-existent. Other researchers have challenged this meta-analysis claiming that it may be more accurate to conclude that in some respects sequential line-ups are better than simultaneous line-ups as they can reduce false identifications, but in other respects they are worse, as they can also reduce correct identifications. Additionally the advantage of sequential line-ups has not been replicated in the field (McQuiston-Surrett *et al.*, 2006). The Illinois Pilot Project was a field study carried out in three different states in the USA, comparing the traditional simultaneous line-up with the sequential line-up. It reported that witnesses who viewed the simultaneous line-ups were more likely to identify the suspect than those who viewed a sequential line-up and that witnesses who viewed sequential line-ups were less likely to choose a line-up member, than those who viewed simultaneous line-ups (Malpass, 2006).

There is still ongoing debate over which line-up method should be implemented by police forces. Some researchers argue that sequential line-ups should be implemented as they reduce the risk of mistaken identity and wrongful conviction (Lindsay *et al.*, 2009*a*, 2009*b*), whilst Malpass *et al.* (2009*a*, 2009*b*) argue that simultaneous line-ups should be used because the advantages of sequential line-ups are offset by the decrease in correct identification as compared with simultaneous line-ups (for the full debate see Lindsay *et al.*, 2009*a*, 2009*b*; Malpass *et al.*, 2009*a*, 2009*b*).

Traditionally, line-ups in the UK were carried out using live parades; however, since the implementation of recent legislation (Police and Criminal Evidence Act in England and Wales, 1984; Vulnerable Witness Act in Scotland, 2004), nearly all line-ups are now presented via video. There are two types currently being used, VIPER (video identification procedure electronic recording) and PROMAT (profile matching). Both formats show a series of images (on average nine) depicting the head and shoulders of each line-up member. Initially each person faces forward then rotates their head from left to right showing both profiles before facing forward again. There are a number of advantages of using video line-ups over live parades. Field research that compared live parades and video parades (VIPER) found that only 5.2% of 403 video parades had to be cancelled, whereas 46.4% of 680 live parades had to be cancelled (Pike *et al.*, 2000). There is also evidence that video parades can be less biased than live parades. Valentine and Heaton (1999) asked mock witnesses to pick out a suspect using a verbal description from either a photograph from a live parade, or a video parade. They found that 25% of witnesses chose the suspect in the live line-ups, whereas only 15% chose the suspect from the video line-up. The proportion choosing the suspect from video line-ups was not significantly different than that expected by chance. However, choosing from a live line-up was statistically greater than expected by chance.

By their very nature video line-ups are presented sequentially with one member being seen at any one time. Experimental research has found that in some circumstances video line-ups can yield greater results as compared with static photographic line-ups. Research that has compared video and sequential photographic line-ups[1] reported fewer false identifications for TA line-ups that were video line-ups, as compared to photo line-ups for adult witnesses (Valentine *et al.*, 2007) and also adolescent witnesses (Havard *et al.*, 2010). However, not all studies have found an advantage for TA video line-ups. Darling *et al.* (2008) found no differences in correct rejection rates. Furthermore, when it comes to correctly identifying a culprit from a TP line-up the majority of studies have found no significant differences in the number of correct identifications from video or photo line-ups (Cutler and Fisher, 1990; Valentine *et al*., 2007; Darling *et al*., 2008; Havard *et al*., 2010).

Research has also been carried out in the field to investigate the identification rates from video parades. A survey of all VIPER parades conducted in Scotland in 2008 reported that suspect identifications were made 44% of the time (Memon *et al*., in press); this figure is similar to that found in other field studies examining identification from live parades (Wright and McDaid, 1996; Valentine *et al*., 2003). However, VIPER parades were found to produce foil identifications on 42% of occasions, which is higher than that reported previously by live parades (Wright and McDaid, 1996; Valentine *et al*., 2003). Memon and her colleagues suggest several explanations for the high choosing rate for foil identifications. One explanation is that witnesses may feel more compelled to choose someone from a video line-up, as compared with a live parade. Alternatively, the foils used in the VIPER parades are obtained from a large database so therefore it is possible that the foils may bear a stronger resemblance to the suspect than the stand-ins that were previously recruited for live parades. Another explanation is that all video parades shown in the UK are viewed twice prior to the witness making an identification. The increase in choosing rates is an issue that is explored in the next section.

8.4 Viewing the line-up once or twice

In the UK, as the majority of line-ups are conducted using sequentially presented video images, there is the opportunity for the witness to view the line-up a second time. Indeed, in Scotland the Lord Advocate's guidelines on Visual Identification Parades (2007) state that 'the witness should normally view the whole set of images at least twice before confirming that he or she wants to view the images or any part of them again. Only where the identification is unequivocal at the first viewing, and further viewing is likely to cause distress to the witness, should this practise be departed from' (Appendix C). Additionally, in England and Wales PACE (Police and Criminal Evidence Act) code of practice (1984) also suggests that a video parade should be viewed twice before the witness makes a decision. However, there is little research that supports the practice of viewing a line-up twice, although this procedure may have originated from viewing live line-ups.

One study by Pike *et al.* (1999) investigated the influence of multiple viewings of line-ups and showed witnesses a staged crime and 4 weeks later the witnesses viewed either a target present (TP), or target absent (TA) line-up, either once or twice. For the TP line-ups, performance improved with additional viewing, however for the TA line-ups those who were allowed to rewind the tape at will, were more accurate in their decisions than those who viewed the parade once or twice. Another study with child witnesses (aged between 6–11 years) found that viewing a video line-up twice had no significant effect on correct identification rates when the culprit was present in the line-up. However, when the culprit was absent, witnesses were significantly more likely to make a false identification from the line-up if they had viewed the line-up twice (58%) as compared with only once (38%). Havard *et al.* (2008) suggested that showing the line-up to children twice increased the likelihood that someone from the line-up would be chosen whether the culprit was present or not.

Under some circumstances, the procedure of viewing the line-up twice does not appear to have any detrimental effects. Valentine *et al.* (2007) compared VIPER line-ups administered in two different ways to adult witnesses to a staged event; strict sequential viewing (where the individual line-up members are only viewed once) and the conventional practice of viewing the line-up twice before making an identification. Valentine *et al.* (2007) found that for the strict sequential condition, there was a lower rate of correct identifications in TP line-ups compared with the standard VIPER procedure of viewing the line-up twice. There were no differences in false identification rates between the strict sequential and the standard VIPER procedure in TA line-ups.

8.5 Line-up biases

The following section explores a number of issues that the police should be aware of when constructing and administering a line-up, to ensure that the witness chooses the person that they recognise as being the perpetrator rather than the only viable option. Furthermore, researchers investigating eyewitness identification should also be aware of these issues to ensure that their participants are choosing a person from the line-up in relation to the variable they are trying to test (independent variable) and not through extraneous variables, such as biases in the line-up.

8.5.1 Selection of foils

When placing a suspect in a line-up, how the foils are selected and the extent to which they resemble the suspect can strongly influence a witness's decision. Line-ups have the potential to be extremely biased where there may be only one plausible match for the culprit. One example of a real case is when a Black suspect was placed in a line-up with White foils (Ellison and Buckhout, 1981). More recently a line-up was deemed as being biased against the suspect as all the foils had wire-rimmed glasses, whereas the suspect wore thick dark-rimmed glasses and the witness had indicated that the perpetrator had 'Buddy Holly' glasses (MacLin *et al*., 2006). The Innocent Project has revealed several occasions when the line-up has been biased against the suspect, such as the suspect's photo being the only one in colour, the suspect and only one other member of the line-up were shirtless (the perpetrator had been described as shirtless) and a line-up where the suspect was the only member of the line-up to appear in an orange prison jumpsuit (Innocence Project Report, 2009).

To ensure that the line-up is fair, the foils can be chosen in two different ways. They can either be chosen because they bear a physical resemblance to the suspect, as suggested in the PACE code of practice (1984) in England and Wales, or foils can be selected who fit the witness's description of the perpetrator, as suggested by Luus and Wells (1991) and the US Department of Justice's Guide for Law Enforcement (1999). Luus and Wells argue that by choosing foils that match the suspect's appearance, the foils may bear too much resemblance to the suspect and then make it more difficult for the witness to correctly identify the suspect amongst a group of doppelgangers.

One study that investigated the different methods of foil selection showed a group of participants a staged theft of a cash box and then asked them to give a written description of the thief (Luus and Wells, 1991). Witnesses then either saw a line-up with foils that were selected on the basis of that description, or they were selected because they bore a physical resemblance to the culprit. Witnesses were significantly more accurate at correctly identifying the target when the foils matched the witness's description (67%), compared with line-ups where the foils matched the culprit's appearance (22%). False identifications when the culprit was absent (TA) from the line-up were also lower when foils were chosen because they matched the description (32%), as compared with the suspect resemblance line-ups (47%), however this difference was not found to be statistically significant. The advantage of using the witness's description for foil selection has been found in several other studies (Lindsay *et al*., 1994; Juslin *et al*., 1996). However, some studies have failed to find any advantage for using suspect descriptions over the general resemblance, as a means for foil selection for TA or TP line-ups (Tunnicliff and Clark, 2000; Darling *et al*., 2008). Furthermore, using the witness's description of the suspect is only useful when the suspect's appearance matches the description; if the witness gives a description that does not match the suspect's appearance, e.g. black hair when the suspect has brown hair, and foils chosen are on this basis, then they may not resemble the suspect and the line-up may be biased. When the witness's description does not match the suspect then it is recommended that the foils be chosen in relation to the suspect's appearance for the features that do not match (Valentine, 2006).

8.5.2 Instructions prior to line-up

One issue that can influence whether someone is chosen from a line-up is the instructions given to the witness prior to viewing a line-up. Archival studies of actual eyewitnesses to serious crimes show that on average, eyewitnesses select someone from a line-up who is a known innocent foil 30% of the time (Slater, 1994; Wright and McDaid, 1996; Behrman and Davey, 2001; Valentine *et al*., 2003; Behrman and Richards, 2005; Memon *et al*., in press; Wells *et al*., 2007; Wright and Skagerberg, 2007). Pre-line-up instructions are one factor that can significantly influence whether the witness chooses someone from the line-up. Instructing the witness that the culprit may or may not be present can reduce the pressure the witness might have to choose someone from the line-up, whereas not being told that the culprit might not be there may make the witness assume the suspect is in the line-up. Research findings have shown that false identifications from TA line-ups are significantly higher when the witness is not given the pre-line-up instruction than when the witness is given the pre-line-up instruction (Malpass and Devine, 1981; Brewer and Wells, 2006; see meta-analysis by Steblay, 1997; more recent meta-analysis by Clark, 2005). Wells and Quinlivan (2009) also suggest that additional information relayed to the witness such as 'the culprit has been found, that the police know who did the crime, or that they already have plenty of evidence against the person' (p. 7) may suggest to the witness that the culprit is in the line-up and increase the pressure to choose someone from the line-up.

The US Department's research report (1999) suggests that witnesses are informed that the person who committed the crime may or may not be present prior to seeing either a live line-up or photo array. In the UK, policies have even been implemented (PACE in the England and Wales, 1984; Lord Advocates Guidelines in Scotland, 2007), so that prior to seeing a video line-up the witness is told that the person 'may or may not be there'. However, even being told that 'the person may or may not be there' does not necessarily prevent a witness from choosing someone from a line-up. Children in particular are more likely to make a positive choice when presented with a line-up and thereby more likely to falsely identify someone when the culprit is absent from the line-up (Parker and Carranza, 1989; Parker and Ryan, 1993; Beal *et al.*, 1995; Dekle *et al.*, 1996; Lindsay *et al.*, 1997; Pozzulo and Warren, 2003; Pozzulo and Balfour, 2006). A number of studies have tried to reduce the false identification rates made by children; they have included providing an additional response that children can use when picking a line-up member e.g. Mr Nobody (Davies *et al.*, 1989), or silhouette (Zajac and Karageorge, 2009) or a 'not there' response printed on a card (Beal *et al.*, 1995). These studies highlight the importance of unbiased instructions that allow witnesses to 'not' choose someone if they feel the suspect is not in the line-up.

8.6 Activities conducted prior to viewing a line-up: describing the face, mugbooks and composite construction

This section describes a number of activities that the witness may be asked to carry out by the police prior to viewing a line-up and how they may affect identification.

8.6.1 Verbal overshadowing effect

When a person witnesses a crime one of the things they are usually asked to do is describe the person they saw, and up until the end of the last century little was thought about the consequences of this task. Then a series of experiments by Schooler and Engstler-Schooler (1990) showed participants a staged crime video and asked them to either describe the facial features of the culprit, or carry out a filler task (control condition) before being presented with a line-up. They found that those who were in the description condition were significantly less able to identify the culprit than those in the control condition – they called this the *verbal overshadowing effect*. Since the first series of studies several researchers have replicated these results; however, several more have failed to find the verbal overshadowing effect (for reviews see Meissner and Memon, 2002). Meissner and Brigham (2001) carried out a meta-analysis and found a small effect of verbal overshadowing, which was more likely to occur if the identification task was preceded immediately by the description (less than 10 minutes) rather than after a delay (over 30 minutes). Verbal overshadowing was also more likely if participants were asked to give elaborate descriptions, rather than a free recall.

There are a number of explanations as to why giving a description of a face may impair later recognition. Schooler *et al.* (1997) suggest that when carrying out the act of describing a face, verbal processing needs to be engaged and this may later interfere with trying to retrieve the visual code of the face from visual memory. Meissner *et al.* (2001; see also Meissner and Brigham, 2001) have proposed a 'misinformation account' whereby non-veridical information elicited by the description impacts unfavourably upon memory for the described face. This misinformation may alter the memory of the face in relation to the verbal description. However, in real eyewitness situations verbal overshadowing is unlikely to occur as its effects are not long-lasting (Finger and Pezdek, 1999; Meissner and Brigham, 2001), and it is doubtful that after witnessing a crime an eyewitness would be asked to describe a perpetrator less than 30 minutes prior to viewing a line-up.

8.6.2 Viewing mugbooks/mugshots

In some investigations, prior to seeing a line-up a witness is shown a series of facial images called mugshots or an album called a mugbook. It has been argued that when a witness sees a mugbook it may bias their decision when faced with a line-up, and actually make them less accurate. Both the US Supreme Court and the UK Police and Criminal Evidence Act (PACE) guidelines recommend that a witness who has seen mugshots should not subsequently view a line-up, in case the witness's decision at the line-up stage is biased by their prior exposure to the mugshots. This has been supported by a number of studies that have shown that when witnesses view a mugbook it can have a detrimental effect on subsequent recognition (Brown *et al.*, 1977; Davies *et al.*, 1979; Gorenstein and Ellsworth, 1980; Brigham and Cairns, 1988; Dysart *et al.*, 2001; Hinz and Pezdek, 2001; Memon *et al.*, 2002; Pezdek and Blandon-Gitlin, 2005). Additionally, a meta-analysis concluded that exposing witnesses to mugshots negatively impacted on subsequent line-up identification rates (Deffenbacher *et al.*, 2006). One of the problems with viewing a mugbook is if the culprit is not present in the mugbook, but one or more foils shown in the mugbook are presented in a subsequent line-up. In these situations witnesses can often choose a person they have seen previously in the mugbook, rather than the person they saw commit the offence – this has been called the *transference effect*.

8.6.3 Constructing composites

If the police do not have a suspect in mind for a particular crime and the witness can give a good description of the culprit, then a facial composite may be made in the hope that someone from the general public will recognise the face and come forward (for details on composite construction see Chapter 4). If a suspect is arrested as a result of the composite, the person may be placed in a line-up to see if the witness can identify them. There have been inconsistent results from studies that have investigated the effect of composite construction prior to line-up identification. Some early studies found that building an Identikit could increase later recognition accuracy (Mauldin and Laughery, 1981). Meissner and Brigham (2001) compared the results of eight studies and found that participants who were asked to create a facial composite performed more accurately than those who were not asked to create a composite. However, other studies showed that building a composite could make witnesses reluctant to make an identification when the target was later present (Yu and Geiselman, 1993). In a more recent study, Wells *et al.* (2005) found that building a composite decreased correct identification of a target face from a line-up, however, when shown a TA line-up building a composite did not increase the false identification rate. They suggested that the impairment was a result of breaking down the face into individual features which was required for the composite task, and that this could be a similar process to verbal overshadowing.

8.7 Estimator variables that influence identification from line-ups

This section will focus on issues that are not under the control of the justice system and relate to the witness and their experience of the incident. However, these are factors that should also be taken into consideration when conducting line-ups and when identification evidence is presented in court.

8.7.1 Witness/suspect age and own-age biases

Factors that can influence whether an identification is made from a line-up are the age of the witness and whether it is similar to that of the suspect. Research investigating how children perform with line-ups has found that when identifying a culprit from a TP line-up, children over the age of 5 years can perform as accurately as adults (Goodman and Reed, 1986; Parker and Carranza, 1989; Parker and Ryan, 1993; Lindsay *et al.*, 1997; Pozzulo and Lindsay, 1998; Pozzulo and Balfour, 2006). However, when faced with a line-up when the culprit is absent, children are significantly more likely to make a false identification as compared with adults (Parker and Carranza, 1989; Parker and Ryan, 1993; Beal *et al.*, 1995; Dekle *et al.*, 1996; Lindsay *et al.*, 1997; Pozzulo and Warren, 2003; Pozzulo and Balfour, 2006). A similar pattern has been found for older adults (over 60 years); they too can perform equally to young adults on TP line-ups (Yarmey and Kent, 1980; Memon *et al.*, 2002, 2003; Memon and Gabbert, 2003). However, when presented with a TA line-up, older adults are significantly more likely to make a false identification as compared with younger adults (Yarmey and Kent, 1980; Memon *et al.*, 2002, 2003; Memon and Gabbert, 2003; Rose *et al.*, 2003, 2005; Wilcock *et al.*, 2005; Havard and Memon, 2009). For a review of the literature on older adult witnesses see Bartlett and Memon (2007). There are several explanations as to why these groups perform more poorly on TA line-ups than young adults; they include cognitive deficits such as a feeling of familiarity for a face resulting in a false identification and also meta-cognitive and social explanations, such as forgetting the instruction that the person may or may not be there and feeling that they must choose one of the line-up members.

Another explanation for the high choosing rates made by children and older adults for TA line-ups is that the majority of studies use young adults as targets, and this puts these groups at a disadvantage as there is an own-age bias for face recognition (for more details on face recognition see Chapter 2). There is some evidence to support this from laboratory face recognition studies (Bäckman, 1991; Anastasi and Rhodes, 2005, 2006; Perfect and Moon, 2005). There are also a few eyewitness studies where an own-age bias has been found for line-up identifications. Memon *et al.* (2003) showed younger (16–33 years of age) and older (60–82 years of age) adult participants videos that depicted staged crimes by older and younger criminals and then asked them to identify the targets from TA and TP line-ups. They found that overall, older adults more likely to make a false identification on a TA line-up, compared with the younger adults, and they were especially more likely to make false identification with the younger adult line-ups. Havard and Memon (2009) showed older (aged 61–83 years) and younger (aged 18–35 years) adults two films that depicted a staged theft that were identical apart from the fact that one had a young target and the other an older target. The witnesses then had to try to identify both culprits from separate line-ups. They found that older adults performed more poorly for both the TP and TA line-ups and showed no own-age bias; however, the young adults showed an own-age advantage for the TA line-ups. Pozzulo and Demsey (2009) showed a group of young adults a film that depicted either a young adult thief or a child thief, prior to carrying out a line-up. They found that young adults were better at identifying a child culprit from a TP line-up, but made fewer false identifications from a TA line-up if it was made up of adults compared with children.

Some research has not found an own-age bias for TA line-ups, but one has been observed for TP line-ups. Wright and Stroud (2002) showed younger (18–25 years old) and older (35–55 years old) adults four simulated crime videos. In two videos the culprit was a young adult and in two the culprit was an older adult. They found that the younger adults and older adults were better at identifying the own-age culprit from a TP line-up, however there was no effect of age for the TA line-ups. In another study, Perfect and Harris (2003, Experiment 3) also found that older adults (mean age 66.6 years) were better at identifying own-age target faces from a line-up, as compared with younger target faces, but no pattern was found for young adult participants (mean age 22 years).

The own-age bias has not always been consistently reported for eyewitness studies. Three similar studies (Rose *et al.*, 2002, 2005; Wilcock *et al.*, 2005) showed younger and older witnesses a videoed event with both younger and older adult targets and then showed the participants TP and TA line-ups. Although for all three studies the older age group had poorer performance than the younger age group, none of the studies showed a significant own-age bias for either group. For one of the studies (Rose *et al.*, 2005) a reversed age bias was reported, where the younger participants were less likely to falsely identify the older culprit from a TA line-up, as compared with the younger culprit. Rose *et al.* (2005) suggest that the older faces may have been more distinctive because of physical differences such as scars and age lines and the young faces may have been a more homogeneous set. This returns back to the issue of foil selection as it may be easier to select appropriate foils for younger faces as compared with older faces, due to them having fewer distinctive features. Explanations for the own-age bias are similar to those for the own-race bias and will be described in the next section.

8.7.2 Witness/suspect race and own-race bias

Another issue that can influence identification from a line-up is whether the witness is the same race as the suspect. A number of face recognition studies (see Chapter 2) have shown that individuals are more accurate at identifying someone of their own race, than they are at identifying someone of a different race. This has become known as the own-race bias, cross-race or the other-race effect, and has been supported by a number of meta-analyses (see Bothwell *et al.*, 1989; Anthony *et al.*, 1992; Sporer, 2001). One meta-analysis by Meissner and Brigham (2001) found that individuals were not only worse at recognising others from another race, but were more likely to make a false identification of someone of another race. For a review of the literature see Brigham *et al.* (2007). Although there have been many studies from the face recognition literature investigating the cross-race effect very few have been conducted using the eyewitness paradigm. In their meta-analysis Meissner and Brigham (2001) found less than 10% had used line-ups. Furthermore, although the area has received little research from the eyewitness field, the Innocent Project found that 53% of mistaken identification cases involved cross-race identification (Innocent Project Report, 2009), therefore making it a noteworthy area of research.

There have been a few studies that have used line-ups to investigate the cross-race effect. Smith *et al.* (2004) asked White participants to try and identify either a White or Afro-Caribbean target they had previously seen from either a TP or TA simultaneous line-up. They found that participants were more accurate at identifying the same-race targets, as compared with the other-race targets and more likely to falsely recognise other-race foils as compared with same-race foils. In another study White and First Nation participants were presented with a series of First Nation and White faces and then 12 line-ups (six TP and six TA). They found that all participants were significantly more accurate with their own race, however, all participants were also more likely to choose from the First Nation line-up as compared with the White line-ups, especially the White participants (Jackiw *et al.*, 2008).

Kask and Bull (2009) investigated identification of multiple suspects of different ethnicities. They presented White adults with four target faces of different ethnicities (White, Afro-Caribbean, Latino and Turkish). Then, after a delay, they either saw a sequential line-up for each ethnicity (2 TP and 2 TA), or 24 faces shown sequentially, and the participants either made a decision to each face, or once all the faces had been shown. They found that other-race faces were more often falsely identified as compared with own-race faces; however, in contrast to other studies they found that own-race faces were less often correctly recognised as compared with other-race faces. They claimed the reason for this was because participants used more liberal criteria for other-race faces as compared with own-race faces, which resulted in the biased identification of other-race faces.

The cross-race effect can not only influence witnesses' decisions when trying to make an identification from a line-up, but can also influence the construction of the line-up. Brigham and Ready (1985) tasked a group of Afro-Caribbeans and Whites to select five foils that were similar in appearance to a target photo, a task similar to that when constructing a line-up. They found that both groups went through more photos when looking for own-race matches than when looking to match the foils to the other-race target, and also used a stricter criterion for similarity for the own-race photos. Several other studies have found that race can influence the fairness of line-ups (Brigham *et al.*, 1990; Lindsay *et al.*, 1999) and Brigham *et al.* (2007) suggest that line-ups should be constructed by operators who are the same race as the suspect, to increase fairness and remove any bias.

There are a number of explanations as to why we have own-race and own-age biases when it comes to recognising faces. One theoretical account that has been used to explain the own-race bias can also help to explain the own-age bias and relates to the amount of contact one has with same-race individuals (see Brigham and Malpass, 1985; Slone *et al.*, 2000). According to the contact hypothesis, we gain expertise in processing same-race faces as they are more frequently encountered, leading to a processing and retrieval advantage for own-race faces. This expertise leads to a configural or holistic processing mode, where the face is processed as a whole for own-race faces. Other-race faces are processed in a less efficient manner using a featural strategy, where the features are examined in a piecemeal fashion, which can lead to poorer encoding (Hancock and Rhodes, 2008). This could also be the case for own-age faces as we may have more contact with other people our own age. This could explain why the own-age bias is not always found, as some people may have contact with people from a variety of different age groups.

An alternative, yet similar, explanation for the own-age bias was offered by Anastasi and Rhodes (2005, 2006) in relation to Sporer's in-group/out-group model of face processing. This model suggests that in-group faces are processed automatically and with expertise, whereas out-group faces first need to be categorised as belonging to the out-group and hence are not processed with the same expertise as in-group faces. Rodin (1987) suggests that when encountering new people decisions are made about whether the person is suitable for social inclusion and age can be used as a criterion to disregard an individual for social inclusion. Due to this cognitive disregard, faces categorised as belonging to the out-group may be cognitively ignored (Rodin, 1987) and deemed as deserving less attention leading to worse recognition of out-group faces (Bernstein *et al.*, 2007). It may also be important if the faces belonging to the out-group are perceived as positive role models, or disliked individuals, as this can in turn influence automatic attitudes (Dasgupta and Greenwald, 2001) that may influence face recognition (Meissner and Brigham, 2001). The various accounts of the own-race and own-age bias are overlapping and predict that witnesses are more likely to recognise a perpetrator who is the same race and of a similar age.

8.7.3 The weapon focus effect and stress

Another issue that may influence how accurately an eyewitness can identify the culprit of a crime is whether the perpetrator had a weapon. A number of studies have found that when the perpetrator of a crime holds a weapon, there is less chance that an eyewitness will be able to identify them (for a review see Steblay, 1992). One study monitored witnesses' eye movements and found that the presence of a gun can draw a witness's attention away from other things, such as the face of the person holding the gun (Loftus *et al.*, 1987). There is still a debate about the cause of the weapon focus effect – some researchers suggest that it is due to stress (Öhman *et al.*, 2001*a*, 2001*b*) whereas other researchers suggest that it is the novelty of weapons that attracts attention (Pickel, 1998, 1999). Some recent research found that regardless of the novelty of items being held by actors, the presence of a gun produced poorer recognition scores and culprit descriptions (Hope and Wright, 2007). However, in a field study of real witnesses in London, the presence of a weapon had no effect on the likelihood of identifying a suspect (Valentine *et al.*, 2003). In fact, witnesses were more likely to identify a foil if a weapon was not present.

There is mixed evidence as to whether stress or high emotional arousal can impair line-up identification. Tollestrup *et al.* (1994) examined real-world identifications of suspects according to crime type. They found that identifications were highest for victims of robberies (46%), and were lower for robbery witnesses (33%) and lowest for victims of fraud (25%). Tollestrup *et al.* interpreted this as showing that stress could enhance face recognition memory. Deffenbacher *et al.* (2004) suggest that high levels of stress can either impair or enhance eyewitness identification. In their meta-analysis they found that high levels of stress impaired the accuracy of eyewitnesses' line-up identifications, as compared with low levels of stress (Deffenbacher *et al.*, 2004). This was confirmed by a novel study by Valentine and Mesout (2009). They took participants to the London Dungeon, where an actor dressed in a robe would step out in front of them and block their path. They were later shown a line-up and asked to try to identify the person and their state of anxiety was measured using a questionnaire and also through a heart monitor. Those who showed higher state anxiety made fewer correct identifications from the line-ups and also fewer correct descriptions of the target person. All this research illustrates that the findings are mixed and in some circumstances high stress can have a negative impact on line-up identification, but this is not always the case. There are numerous extraneous factors which could influence how a witness responds in a stressful situation including how the threat is perceived by the witness, their individual coping strategy, and the reactions of others. These are all issues that have yet to be explored in future research.

As mentioned in the opening of this chapter, eyewitness identification can be powerful evidence in court, however, it is not always accurate and many people have been wrongfully sentenced as a result of misidentification. This chapter has explored a number of issues that can affect identification from a line-up. Issues under police control are system variables, and those that relate to the witness and the crime itself are estimator variables. The police and researchers investigating eyewitness identification should be aware of these issues to try to make the procedure of conducting a line-up objective and completely unbiased. When foils are selected for a line-up they should resemble the suspect sufficiently, so that the suspect does not stand out in any way. The instructions and briefing that witnesses are provided with prior to a line-up should make it clear that the perpetrator may or may not be present in the line-up. The UK has moved from use of live parades to video and the evidence so far suggests that this method can produce correct identification rates comparable with photo line-ups. The situation is less clear with respect to the effect of viewing the entire line-up twice on foil identification rates, with some indications that witnesses may feel under greater pressure to choose with a video line-up. Activities the witness can be asked to do prior to viewing a line-up, such as viewing a mugbook or making a facial composite, can also influence whether the perpetrator is identified. Estimator variables are not under any control, however, the police conducting the line-up and researchers should be aware of the potential impact of biases arising from the age and race of both the witness and suspect. Finally, there is some evidence to suggest the presence of emotionally arousing stimuli (e.g. weapons) and how a witness reacts to environmental stressors could impact identification rates under some circumstances.

References

Anastasi, J. S. and Rhodes, M. G. (2005). An own-age bias in face recognition for children and older adults. *Psychonomic Bulletin and Review*, **12**(6), 1043–1047.

Anastasi, J. S. and Rhodes, M. G. (2006). Evidence for an own-age bias in face recognition. *North American Journal of Psychology*, **8**, 237–253.

Anthony, T., Cooper, C. and Mullen, B. (1992). Cross-racial facial identification: a social cognitive integration. *Personality and Social Psychology Bulletin*, **18**, 296–301.

Australian Innocence Network. Innocence Project. Available at: http://www.griffith.edu.au/law/innocence-project/partnerships-collaboration/australian-innocence-network [accessed on 16 March 2010].

Bäckman, L. (1991). Recognition memory across the adult life span: the role of prior knowledge. *Memory and Cognition*, **19**, 63–71.

Bartlett, J. C. and Memon, A. (2007). Eyewitness memory in young and older adults. In R. Lindsay, R. Ross, D. Read and M. Toglia (Eds.), *Handbook of Eyewitness Psychology. Vol. II: Memory for People*, pp. 309–338. Mahwah, NJ: Lawrence Erlbaum.

Beal, C. R., Schmitt, K. L. and Dekle, D. J. (1995). Eyewitness identification of children. effects of absolute judgements, nonverbal response options, and event encoding. *Law and Human Behavior*, **19**, 197–216.

Behrman, B. W. and Davey, S. L. (2001). Eyewitness identification in actual criminal cases: an archival analysis. *Law and Human Behavior*, **25**, 475–491.

Behrman, B. W. and Richards, R. E. (2005). Suspect/foil identification in actual crimes and in the laboratory: a reality monitoring analysis. *Law and Human Behavior*, **29**, 279–301.

Bernstein, M. J., Young, S. G. and Hugenberg, K. (2007). The cross-category effect: mere social categorization is sufficient to elicit an own-group bias in face recognition. *Psychological Science*, **18**, 706–712.

Bothwell, R. K., Brigham, J. C. and Malpass, R. S. (1989). Cross-racial identification. *Personality and Social Psychology Bulletin*, **15**, 19–25.

Brewer, N. and Wells, G. L. (2006). The confidence–accuracy relationship in eyewitness identification: Effects of lineup instructions, foil similarity and target-absent base rates. *Journal of Experimental Psychology: Applied*, **12**, 11–30.

Brigham, J. C., Bennett, L. B., Meissner, C. A. and Mitchell, T. L. (2007). The influence of race on eyewitness memory. In R. C. L. Lindsay, D. F. Ross, J. D. Read and M. P. Toglia (Eds.), *Handbook of Eyewitness Psychology. Vol. II: Memory for People*, pp. 257–281. Mahwah, NJ: Lawrence Erlbaum.

Brigham, J. C. and Cairns, D. (1988). The effect of mugshot inspections on eyewitness identification accuracy. *Journal of Applied Social Psychology*, **16**, 1394–1410.

Brigham, J. C. and Malpass, R. S. (1985). Differential recognition for faces of own- and other-race persons: what is the role of experience and contact? *Journal of Social Issues*, **41**, 139–155.

Brigham, J. C. and Ready, D. R. (1985). Own-race bias in lineup construction. *Law and Human Behavior*, **9**, 415–424.

Brigham, J. C. Ready, D. J. and Spier, S. A. (1990). Standards for evaluating the fairness of photograph lineups. *Basic and Applied Social Psychology*, **11**, 149–163.

Brown, E., Deffenbacher, K. and Sturgill, W. (1977). Memory for faces and the circumstances of encounter. *Journal of Applied Psychology*, **62**, 311–318.

Clark, S. E. (2005). A re-examination of the effects of biased lineup instructions in eyewitness identification. *Law and Human Behavior*, **5**, 575–604.

Cutler, B. L. and Fisher, R. P. (1990). Live lineups, videotaped lineups and photo arrays. *Forensic Reports*, **3**, 439–448.

Darling, S., Valentine, T. and Memon, A. (2008). Selection of lineup foils in operational contexts. *Applied Cognitive Psychology*, **22**, 159–169.

Dasgupta, N. and Greenwald, A. G. (2001). On the malleability of automatic attitudes: combating automatic prejudice with images of admired and disliked individuals. *Journal of Personality and Social Psychology*, **81**, 800–814.

Davies, G., Shepherd, J. and Ellis, H. (1979). Effects of interpolated mugshot exposure on accuracy of eyewitness identification. *Journal of Applied Psychology*, **64**, 232–237.

Davies, G. M., Tarrant, A. and Flin, R. H. (1989). Close encounters of the witness kind – children's memory for a simulated health inspection. *British Journal of Psychology*, **80**, 415–429.

Deffenbacher, K. A., Bornstein, B. H., Penrod, S. D. and McGorty, K. (2004). A meta-analytic review of the effects of high stress on eyewitness memory. *Law and Human Behavior*, **28**, 687–706.

Deffenbacher, K. A., Bornstein, B. H. and Penrod, S. D. (2006). Mugshot exposure effects: retroactive interference, mugshot commitment, source confusion, and unconscious transference. *Law and Human Behavior*, **30**, 287–307.

Dekle, D. J., Beal, C. R., Elliot, R. and Huneycutt, D. (1996). Children as witnesses: a comparison of lineup versus showup identification methods. *Applied Cognitive Psychology*, **10**, 1–12.

Dysart, J. E. and Lindsay, R. C. L. (2001) A preidentification questioning effect: serendipitously increasing correct rejection. *Law and Human Behavior*, **25**, 155–165.

Dysart, J., Lindsay, R. C. L., Hammond, R. and Dupuis, P. (2001). Mugshot exposure prior to lineup identification: interference, transference and commitment effects. *Journal of Applied Psychology*, **86**, 1280–1284.

Ellison, K. W. and Buckhout, R. (1981). *Psychology and Criminal Justice*. New York: Harper and Row.

Finger, K. and Pezdek, K. (1999). The effect of the cognitive interview on face identification accuracy: release from verbal overshadowing. *Journal of Applied Psychology*, **84**, 340–348.

Goodman, G. S. and Reed, R. S. (1986). Age differences in eyewitness testimony. *Law and Human Behavior*, **10**, 317–332.

Gorenstein, G. W. and Ellsworth, P. C. (1980). Effect of choosing an incorrect photograph on a later identification by an eyewitness. *Journal of Applied Psychology*, **65**, 616–622.

Hancock, K. and Rhodes, G. (2008). Contact, inversion and the other-race effect in face recognition. *British Journal of Psychology*, **99**, 45–56.

Havard, C. and Memon, A. (2009). The influence of face age on identification from a video line-up: a comparison between older and younger adults. *Memory*, **17**, 847–859.

Havard, C., Memon, A., Clifford, B. and Gabbert, F. (2008) Obtaining evidence from child witnesses: the advantage of VIPER parades. Paper presented at the Scottish Institute of Policing Research (SIPR) Evidence and Investigation Network Seminar, *Obtaining Evidence from Vulnerable Witnesses*, Aberdeen, 15 October 2008.

Havard, C., Memon, A., Clifford, B. R., and Gabbert, F. (2010). A comparison of video and static photo lineups with children and adolescent witnesses. *Applied Cognitive Psychology*, **24**, 1209–1221.

Hope, L. and Wright, D. (2007). Beyond unusual? Examining the role of attention in the weapon focus effect. *Applied Cognitive Psychology*, **21**, 951–961.

Hinz, T. and Pezdek, K. (2001). The effect of exposure to multiple lineups on face identification accuracy. *Law and Human Behavior*, **25**, 185–198.

Home Office (2002). *Achieving Best Evidence in Criminal Proceedings: Guidance for Vulnerable or Intimidated Witnesses, including Children*. London: Home Office Communication Directorate.

Innocent Network UK (2010) *Innocent Projects*. Available at: http://www.innocencenetwork.org.uk/innocenceprojects.htm [accessed 16 March 2010].

Innocence Project (2010). *The Faces of Exoneration*. Available at: http://www.innocenceproject.org/ [accessed 16 March 2010].

Innocence Project New Zealand. Available at: http://www.victoria.ac.nz/ipnz/ [accessed 16 March 2010].

Innocent Project Report (2009) *Re evaluating Lineups: Why Witnesses make Mistakes and How to Reduce the Chance of Mistaken Identification*. Available at: http://www.innocenceproject.org/docs/Eyewitness_ID_Report.pdf [accessed 16 March 2010].

Innocent Project (2010). *Understanding the Causes of Misidentification*. Available at: http://www.innocenceproject.org/understand/Eyewitness-Misidentification.php [accessed 16 March 2010].

Jackiw, L. B., Arbuthnott, K. D., Pfeifer, J. E., Marcon, J. L. and Meissner, C. A. (2008). Examining the cross-race effect in lineup identification using Caucasian and First Nation samples. *Canadian Journal of Behavioural Science*, **40**, 52–57.

Kask, K. and Bull, R. (2009). The effects of different presentation methods on multi-ethnicity face recognition. *Psychology, Crime and Law*, **15**, 73–89.

Juslin, P., Olsson, N. and Winman, A. (1996). Calibration and diagnosticity of confidence in eyewitness identification: Comments on what can be inferred from the low confidence-accuracy correlation. *Journal of Experimental Psychology: Learning, Memory and Cognition*, **22**, 1304–1316.

Kneller, W., Memon, A. and Stevenage, S. (2001). Simultaneous and sequential lineups: decision processes of accurate and inaccurate eyewitnesses. *Applied Cognitive Psychology*, **15**, 659–671.

Lindsay, R. C. L. and Bellinger, K. (1999). Alternatives to the sequential lineup: the importance of controlling pictures. *Journal of Applied Psychology*, **84**, 315–321.

Lindsay, R. C. L., Lea, J. A., Nosworthy, G. F. *et al.* (1991) Biased lineups: sequential presentation reduces the problem. *Journal of Applied Psychology*, **76**, 796–802.

Lindsay, R. C. L., Mansour, J. K., Beaudry, J. L., Leach A. M. and Bertrand, M. I. (2009*a*). Sequential lineup presentation: patterns and policy. *Legal and Criminological Psychology*, **14**, 13–24.

Lindsay, R. C. L., Mansour, J. K., Beaudry, J. L., Leach A. M. and Bertrand, M. I. (2009*b*). Response. Beyond sequential presentation: misconceptions and misrepresentations of sequential lineups. *Legal and Criminological Psychology*, **14**, 31–34.

Lindsay, R. C. L., Martin, R. and Webber, L. (1994). Default values in eyewitness descriptions: a problem for the match-to-description lineup foil selection strategy. *Law and Human Behavior*, **18**, 527–541.

Lindsay, R. C. L., Pozzulo, J. D., Craig, W., Lee., K. and Corber, S. (1997). Simultaneous lineups, sequential lineups and showups: eyewitness identification decisions of adults and children. *Law and Human Behavior*, **21**, 391–404.

Lindsay, R. C. L., Ross, D. F., Smith, S. M. and Flanigan, S. (1999). Does race influence measures of lineup fairness? *Applied Cognitive Psychology*, **13**, 109–119.

Lindsay, R. C. L. and Wells, G. L. (1985). Improving eyewitness identification from lineups: simultaneous versus sequential lineup presentation. *Journal of Applied Psychology*, **70**, 556–564.

Loftus, E. F., Loftus, G. R. and Messo, J. (1987). Some facts about weapon focus. *Law and Human Behavior*, **11**, 55–62.

Lord Advocates' guidelines – The conduct of visual identification procedures (2007). Available at: http://www.copfs.gov.uk/Resource/Doc/13547/0000269.pdf [accessed 5 October 2007].

Luus, C. A. E. and Wells, G. L. (1991). Eyewitness identification and the selection of distracters for lineups. *Law and Human Behaviour*, **15**, 43–57.

MacLin, M. K., MacLin, O. H. and Albrechtsen, J. (2006). Using image manipulation to construct fair lineup: the case of the Buddy Holly glasses. *Canadian Journal of Police and Security Services*. Available at: http://findarticles.com/p/articles/mi_m5EQH/is_1_4/ai_n24988961/?tag=content;col1 [accessed 13 April 2010].

Malpass, R. S. (2006). Notes on the Illinois pilot program on sequential double-blind identification procedures. *Public Interest Law Reporter*, **11**, 5–47.

Malpass, R. S. and Devine, P. G. (1981). Eyewitness identification: lineup instructions and the absence of the offender. *Journal of Applied Psychology*, **66**, 482–489.

Malpass, R. S., Tredoux, C. G. and McQuiston-Surrett (2009*a*). Public policy and sequential lineups. *Legal and Criminological Psychology*, **14**, 1–12.

Malpass, R. S., Tredoux, C. G. and McQuiston-Surrett (2009*b*). Response to Lindsay, Mansour, Beaudry, Leach and Bertrand's 'Sequential lineup presentation: Patterns and policy'. *Legal and Criminological Psychology*, **14**, 25–30.

Mauldin, M. A. and Laughery, K. R. (1981). Composite production effects on subsequent facial recognition. *Journal of Applied Psychology*, **66**, 351–357.

McQuiston-Surrett, D., Malpass, R. S. and Tredoux, C. G. (2006). Sequential vs. simultaneous lineups: a review of methods, data and theory. *Psychology, Public Policy and Law*, **2**, 137–169.

Meissner, C. A. and Brigham, J. C. (2001). Thirty years of investigating the own-race bias in memory for faces: a meta-analytic review. *Psychology, Public Policy, and Law*, **7**, 3–35.

Meissner, C. A., Brigham, J. C. and Kelley, C. M. (2001). The influence of retrieval processes in verbal overshadowing. *Memory and Cognition*, **29**, 176–186.

Meissner, C. and Memon, A. (2002). Verbal overshadowing: a special issue involving theoretical and applied issues. *Applied Cognitive Psychology*, **16**, 869–872.

Memon, A. and Gabbert, F. (2003). Improving the identification accuracy of senior witnesses: do pre-lineup questions and sequential testing help? *Journal of Applied Psychology*, **88**, 341–347.

Memon, A., Havard, C., Clifford, B., Gabbert, F. and Watt, M. (in press). A field evaluation of the VIPER system: A new technique for eliciting eyewitness identification evidence. *Psychology, Crime, and Law*.

Memon, A., Hope, L., Bartlett, J. and Bull, R. (2002). Eyewitness recognition errors: the effects of mugshot viewing and choosing in young and old adults. *Memory and Cognition*, **30**, 1219–1227.

Öhman, A., Flykt, A. and Esteves, F (2001*a*). Emotion drives attention: detecting the snake in the grass. *Journal of Experimental Psychology: General*, **130**, 466–478.

Öhman, A., Lundqvist, D. and Esteves, F. (2001*b*). The face in the crowd revisited: a threat advantage with schematic stimuli, *Journal of Personality and Social Psychology*, **80**, 381–396.

Parker, J. F.and Carranza, L. E. (1989). Eyewitness testimony of children in target-present and target absent line-ups. *Law of Human Behavior*, **13**, 133–149.

Parker, J. F. and Ryan, V. (1993). An attempt to reduce guessing behavior in children's and adults' eyewitness identification. *Law and Human Behavior*, **17**, 11–26.

Perfect, T. J. and Harris, L. J. (2003). Adult age differences in unconscious transference: source confusion or identity blending? *Memory and Cognition*, **31**, 570–580.

Perfect, T. J. and Moon, H. (2005). The own-age effect in face recognition. In J. Duncan, L. Phillips and P. McLeod (Eds.), *Measuring the Mind: Speed, Control, and Age*, pp. 317–340. Oxford: Oxford University Press.

Pickel, K. L. (1998). Unusualness and threat as possible causes of 'weapon focus'. *Memory*, **6**, 277–295.

Pickel, K. L. (1999). The influence of context on the 'weapon focus' effect. *Law and Human Behavior*, **23**, 299–311.

Pezdek, K. and Blandon-Gitlin, I. (2005). When is an intervening lineup most likely to affect eyewitness identification accuracy? *Legal and Criminological Psychology*, **10**, 247–263.

Pike, G., Kemp, R., Brace, N., Allen, J. and Rowlands, G. (2000). The effectiveness of video identification parades. *Proceedings of the British Psychological Society*, **8**, 44.

Pike, G., Rowlands, G., Towell, N. and Kemp, R. (1999). Video identification and simultaneous vs. sequential lineups. Paper presented at the *1st Joint Conference of the American-Law Society and the European Association of Psychology and Law*, Dublin, Ireland. Cited in Kemp, R., Pike, G. E. and Brace, N. A. (2001). Video-based identification procedures; combining best practice and practical requirements when designing identification systems. *Psychology, Public Policy and Law*, **7**, 802–807.

Police and Criminal Evidence Act (1984)/Codes of Practice (2008). Available at: http://police.homeoffice.gov.uk/

operational-policing/powers-pace-codes/pace-code-intro/ [accessed 4 April 2009].

Pozzulo J. D. and Balfour, J. (2006). Children's and adults' eye witness identification accuracy when a culprit changes his appearance: comparing simultaneous and elimination line up procedures. *Legal and Criminal Psychology*, **11**, 25–34.

Pozzulo, J. D. and Demsey, J. (2009). Could target age explain identification accuracy differences between child and adult eyewitnesses? *Psychiatry, Psychology and Law*, **16**, s137–144.

Pozzulo, J. D. and Lindsay, R. (1998). Identification accuracy of children versus adults: a meta-analysis. *Law and Human Behavior*, **22**, 549–570.

Pozzulo, J. D. and Warren, K. L. (2003). Descriptors and identifications of strangers by youth and adult eye witnesses. *Journal of Applied Psychology*, **88**, 315–323.

Rodin, M. J. (1987). Who is memorable to whom: a study of cognitive disregard. *Social Cognition*, **5**, 144–165.

Rose, R. A., Bull, R. and Vrij, A. (2003). Enhancing older witnesses' identification performance: context reinstatement is not the answer. *Canadian Journal of Police and Security Services*, **1**, 173–184.

Rose, R. A., Bull, R. and Vrij, A. (2005). Non-biased lineup instructions do matter – a problem for older witnesses. *Psychology, Crime, and Law*, **11**, 147–159.

Schooler, J. W., Engstler-Schooler, T. Y. (1990). Verbal overshadowing of visual memories: some things are better left unsaid. *Cognitive Psychology*, **22**, 36–71.

Schooler, J. W., Fiore, S. M. and Brandimonte, M. A. (1997). At a loss from words: verbal overshadowing of perceptual memories. In D. Medin (Ed.), *The Handbook of Learning and Motivation*, pp. 293–334. San Diego, CA: Academic Press.

Slater, A. (1994). Identification parades: a scientific evaluation. *Police Research Award Scheme*. London: Police Research Group, Home Office.

Slone, A. E., Brigham, J. C. and Meissner, C. A. (2000). Social and cognitive factors affecting the own-race bias in Whites. *Basic and Applied Social Psychology*, **22**, 71–84.

Smith, S. M., Stinson, V. and Prosser, M. A. (2004). Do they all look alike? An exploration of decision-making strategies in cross-race facial identifications. *Canadian Journal of Behavioural Science*, **36**, 146–154.

Sporer, S. L. (1993). Eyewitness identification accuracy, confidence, and decision times in simultaneous and sequential lineups. *Journal of Applied Psychology*, **78**, 22–33.

Sporer, S. L. (2001). Recognizing faces of other ethnic groups: an integration of theories. *Psychology, Public Policy, and Law*, **7**, 36–97.

Steblay, N. (1992). A meta-analytic review of the weapon-focus effect. *Law and Human Behavior*, **16**, 413–424.

Steblay, N. (1997). Social influence in eyewitness recall: a meta-analytic review of lineup instruction effects. *Law and Human Behavior*, **21**, 283–297.

Steblay, N., Dysart, J., Fulero, S. and Lindsay, R. C. L. (2001). Eyewitness accuracy rates in sequential and simultaneous lineup presentations: A meta-analytic comparison. *Law and Human Behavior*, **25**, 459–473.

Technical Working Group for Eyewitness Evidence (1999). *Eyewitness Evidence: A Guide for Law Enforcement*. Washington, DC: United States Department of Justice, Office of Justice Programs.

Tollestrup, P. A., Turtle, J. W. and Yuille, J. C. (1994). Actual victim and witnesses to robbery and fraud: An archival analysis. In D. F. Ross, J. D. Read and M. P. Toglia (Eds.), *Adult Eyewitness Testimony: Current Trends and Developments*, pp. 144–160. Cambridge: Cambridge University Press.

Tunnicliff, J. L. and Clark, S. E. (2000). Selecting foils for identification lineups: matching suspects or descriptions. *Law and Human Behavior*, **24**, 231–258.

Valentine, T. (2006). Forensic facial identifcation. In A. Heaton-Armstrong, E. Shepherd, G. Gudjonsson and D. Wolchover (Eds.), *Witness Testimony; Psychological, Investigative and Evidential Perspectives*. Oxford: Oxford University Press.

Valentine, T., Darling, S. and Memon, A. (2007). Do strict rules and moving images increase the reliability of sequential identification procedures? *Applied Cognitive Psychology*, **21**, 933–949.

Valentine, T. and Heaton, P. (1999). An evaluation of the fairness of Police line-ups and video identifications. *Applied Cognitive Psychology*, **13**, S59–S72.

Valentine, T. and Mesout, J. (2009). Eyewitness identification under stress in the London Dungeon. *Applied Cognitive Psychology*, **23**, 151–161.

Valentine, T., Pickering, A. and Darling, S. (2003). Characteristics of eyewitness identification that predict the outcome of real lineups. *Applied Cognitive Psychology*, **17**, 969–993.

Vulnerable Witness Act Scotland (2004). Available at: http://www.opsi.gov.uk/legislation/scotland/acts2004/asp_20040003_en_1 [accessed 4 April 2009].

Wells, G. L. (1978). Applied eyewitness-testimony research: system variables and estimator variables. *Journal of Personality and Social Psychology*, **12**, 1546–1557.

Wells, G. L. (1984). The psychology of lineup identification. *Journal of Applied Social Psychology*, **14**, 89–103.

Wells, G. L. (1993) cited in Wells, G. L., Small, M., Penrod, S., Malpass, R. S., Fulero, S. L. and Brimacombe, C. Eyewitness identification procedures:

recommendations for lineups and photospreads. *Law and Human Behavior*, **22**, 603–646.

Wells, G. L., Charman, S. D. and Olson, E. A. (2005). Building face composites can harm lineup identification performance. *Journal of Experimental Psychology: Applied*, **11**, 147–157.

Wells, G. L., Memon, A. and Penrod, S. (2007). Eyewitness evidence: improving its probative value. *Psychological Science in the Public Interest*, **7**, 45–75.

Wells, G. L. and Olson, E. (2003). Eyewitness testimony. *Annual Review of Psychology*, **54**, 277–295.

Wells., G. L. and Quinlivan, D. S (2009). Suggestive eyewitness identification procedures and the Supreme Court's reliability test in light of eyewitness science: 30 years later. *Law and Human Behavior*, **33**, 1–24.

Wilcock, R. A., Bull, R. and Vrij, A. (2005). Aiding the performance of older eyewitnesses: enhanced non-biased lineup instructions and lineup presentation. *Psychiatry, Psychology, and Law*, **12**, 129–140.

Wogalter, M. S., Malpass, R. S. and McQuiston, D. E. (2004). A national survey of U.S. police on preparation and conduct of identification lineups. *Crime Psychology and Law*, **10**, 69–82.

Wright, D. B. and McDaid, A. T. (1996). Comparing system and estimator variables using data from real lineups. *Applied Cognitive Psychology*, **10**, 75–84.

Wright, D. B. and Skagerberg, E. M. (2007). Post-identification feedback affects real eyewitnesses. *Psychological Science*, **18**, 172–178.

Wright, D. B. and Stroud, J. S. (2002). Age differences in lineup identification accuracy: people are better with their own age. *Law and Human Behavior*, **26**, 641–654.

Yarmey, A. D. and Kent, J. (1980). Eyewitness identification by elderly and young adults. *Law and Human Behavior*, **4**, 359–371.

Yu, C. J. and Geiselman, R. E. (1993). Effects of constructing Identi-Kit composites on photospread identification performance. *Criminal Justice and Behavior*, **20**, 280–292.

Note

1. It should be noted that these studies did not use the sequential procedure as described by Lindsay and Wells (1985), but photographs were presented one at a time for the line-up rather than simultaneously as they are in the UK (Home Office, 2002).

Chapter 9

Virtual human identification line-ups

Kathryn Y. Segovia, Jeremy N. Bailenson and Carrie Leonetti

9.1 Introduction

The legal system has long relied on eyewitnesses of crimes to help identify and prosecute criminals. Eyewitness identification is arguably one of the most influential forms of evidence that can be presented to a judge or jury. The general idea behind eyewitness identification is that a witness to a crime, whether a victim or bystander, can later establish the perpetrator's identity. It is assumed that if the witness had a good view of the crime and was paying attention to the physical characteristics of the perpetrator, then the witness's memory will be a valid indicator of identity, particularly if the witness is certain about his or her identification. In this chapter, we will examine the current state of traditional eyewitness identification line-ups, present an overview of a recently proposed idea to utilise virtual environments and virtual humans in identification line-ups, and finally present a comprehensive review of how virtual human identification line-ups may offer advantages over traditional line-ups.

9.2 Traditional eyewitness identification line-ups

According to Wells and Olson (2003), a line-up is a procedure in which a criminal suspect (or a photograph of the suspect) is placed among other people (live line-up) or photographs (photo array), commonly referred to as fillers or foils. Line-ups and photo arrays typically contain at least six individuals or photographs, comprised of the suspect and at least five fillers. The individuals or photographs are then presented to an eyewitness, either sequentially or simultaneously, for identification. Eyewitness identification evidence is powerful; more than 75 000 people each year become criminal defendants on the basis of eyewitness identifications (National Science Foundation, 1997). Although eyewitness identification is one of the most compelling types of evidence to which a jury or judge is exposed, experimental research (Wells, 1993; Wells and Loftus, 2003) and cases of DNA exoneration (Wells and Quinlivan, 2009) have prompted scholars and practitioners to question the accuracy, confidence levels and procedures of eyewitness identification.

9.3 Experimental research in eyewitness identification

Over 100 years ago, Professor Hugo Munsterberg, chair of Harvard's psychology laboratory, undertook the large task of persuading legal scholars, legal professionals and the general public that even confident and honest individuals could deliver mistaken eyewitness identifications (Doyle, 2005). Armed with extensive experience and knowledge in psychological research, Munsterberg pushed psychologists and legal scholars to further investigate the reliability and accuracy of eyewitness identification. He challenged the legal system's complacent acceptance of eyewitness testimony, to which legal scholars responded harshly with their own counterattacks. Despite Munsterberg's strong efforts and experimental research evidence, his attempt to inform the field was described as a 'miserable failure' (Doyle, 2005: 10).

9.3.1 Accuracy

Since the days of Munsterberg, many psychologists have undertaken the, now more readily received, task of studying the reliability of eyewitness identification. Some scholars have created mock crime scenarios in

Craniofacial Identification, ed. C. Wilkinson and C. Rynn. Published by Cambridge University Press.

the laboratory (see Wells, 1993 for a review) or conducted field experiments (Brigham *et al.*, 1982) followed by eyewitness identification tasks. Other scholars have conducted archival analysis of actual court cases involving eyewitness identification (Behrman and Davey, 2001). Results from experimental and archival studies suggest that eyewitness identification is quite prone to error (Wells *et al.*, 1979*a*, 1980, 1981; Lindsay *et al.*, 1981; Behrman and Davey, 2001; Wells and Olson, 2003). One recent study showed that the rate at which eyewitnesses select someone other than the suspect from both photo and live line-ups is approximately 20% (Valentine *et al.*, 2003). In studies where researchers stage crimes, false identifications occur at rates as high as 90% (Wells, 1993).

Despite mounting evidence of the inaccuracy of traditional eyewitness identification procedures, they remain among the most commonly used tools in criminal investigations. Meanwhile, social scientists have wrestled with why honest eyewitnesses mistakenly identify the wrong person as the perpetrator of a crime.

9.3.2 Confidence

The problem of frequent misidentifications in line-ups is compounded by the fact that identifications (correct or incorrect) are often asserted with varying levels of confidence. Experiments by Wells and colleagues in the 1970s showed that witnesses who make false identifications are not appreciably less confident than witnesses who make accurate identifications (as cited in Wells *et al.*, 1979*b*). In a study by Lindsay *et al.* (1981), participants viewed cross-examinations of witnesses who had accurately identified the culprit in the line-up and cross-examinations of witnesses who had falsely identified the culprit. Participants were unable to detect, as measured by their belief in the witness, which witnesses had given false identifications and which witnesses had given accurate identifications.

In a more recent analysis of the confidence–accuracy relationship in eyewitness identification studies, scholars found that while accuracy and confidence are only modestly correlated in broad comparisons of witnesses, the type of identification decision and the conditions under which memories are stored can significantly moderate the confidence–accuracy relationship. For example, witnesses making positive identifications (identifying a person in the line-up) are more likely to have a strong correlation between confidence and accuracy than those witnesses making 'line-up rejection decisions' (failing to identify a suspect from the line-up; Sporer *et al.*, 1995). Furthermore, a study by Lindsay *et al.* (1998) concluded that both accuracy and confidence tend to be higher under conditions that lead to good memory than under conditions that lead to poor memory (such as longer versus shorter viewing times). Decades of psychology research have come to confirm what Hugo Munsterberg so vehemently believed: '[T]he feeling of certainty stands in no definite relation to the attention with which the objects are observed. . . . The correlations between attention, recollection, and feeling of certainty become more complex the more we study them' (as cited in Doyle, 2005: 43).

Furthermore, eyewitness confidence is malleable and susceptible to influence and suggestion, particularly when the administrator of the line-up provides confirming or disconfirming post-identification feedback (Bradfield *et al.*, 2002). One study that examined the effects of post-identification feedback found that it produced 'strong effects' on the witnesses' reports on a range of factors, including certainty and memory clarity (Wells and Bradfield, 1998). Unfortunately, research has consistently shown that an eyewitness's degree of confidence in identification at trial is the single greatest factor affecting whether jurors believe that the identification is accurate (Bradfield and Wells, 2000).

9.4 DNA exonerations

While the results of the aforementioned research slowly became known and tested in the academic community, the criminal justice community did not embrace the warnings regarding the inaccuracies of eyewitness identification until the 1990s when DNA tests began exonerating innocent prisoners (Wells and Olson, 2003). Since its first debut, post-conviction DNA testing has exonerated more than 200 individuals. Furthermore, a recent estimate shows that of the 232 DNA exonerations through 2008, 75% included at least one misidentification (Innocence Project, 2008). '[N]umerous analyses over several decades have consistently shown that mistaken eyewitness identifications are the single largest source of wrongful convictions' (Wells and Seelau, 1995). Furthermore, scholars hypothesise that identifications may continue to be highly inaccurate, because procedures used in the past to obtain eyewitness identifications are similar to those used in jurisdictions today (Yacona, 2005).

9.5 Case law governing eyewitness identification

Throughout the last century, the United States court system has made rulings regarding the procedures that surround eyewitness identification. For example, the United States Supreme Court ruled that during a line-up the defendant has a right to counsel (*United States v Wade*, 1967) and that several factors must be considered before a pre-trial identification can be considered admissible in the court of law (such as the time lapse between the crime and the line-up and the eyewitness's opportunity to view the perpetrator; *Neil v Biggers*, 1972; *Manson v Brathwaite*, 1977).

In addition, in *United States v Downing* (1985) the United States Court of Appeals recognised that the rules of evidence may permit expert testimony about eyewitness identification. Such experts generally testify about issues involving the reliability and susceptibility of human memory. Finally, in lieu of allowing expert testimony, some courts give a cautionary jury instruction encouraging jurors to consider whether, or to what extent, the witness had the ability and the opportunity to observe the perpetrator at the time of the offence and make a reliable identification later. Still, commentators have noted the insufficiency of both expert testimony and cautionary jury instructions to prevent mistaken identification (Koosed, 2002).

Although the data overwhelmingly show that eyewitness identifications are plagued with challenges and risks, it is highly unlikely that the legal system will abandon the use of eyewitness identifications. Identifications are powerful pieces of evidence and can, in the right context, be a reliable means for bringing justice. This situation begs scholars and practitioners to work together in evaluating and improving the identification line-up. Indeed, many scholars have focused on questions of how manipulations to line-up procedure and structure may affect accuracy and confidence (for examples see Wells, 1993; Malpass and Lindsay, 1999; Semmler *et al.*, 2004).

Throughout the rest of this chapter, we will explore the affordances of Immersive Virtual Environment Technology (IVET) as a way of improving identification line-ups. Not only will we describe the technology involved in such a line-up procedure, but we will address multiple affordances of IVET that could improve the current state of identification line-ups. Finally, we will explore recent empirical research on the topic of virtual human identification line-ups.

9.6 Immersive virtual environments

Traditionally eyewitness identification line-ups, as described above, have been conducted using live individuals or photographs of individuals. In practice, the photographic line-up is used most often, largely because this method is less costly and time-consuming (Kerstholt *et al.*, 2004). However, courtroom processes in general are relying more and more on information technology (Goodman *et al.*, 1998), and emerging digital technology may change the way we conceive of and conduct line-ups. Immersive Virtual Environments (IVEs) have existed for decades (Lanier, 2001) and have been widely utilised in a variety of contexts (see Fox *et al.*, 2009 for a review), including clinical counselling (Riva *et al.*, 1998), medical (Reznek *et al.*, 2008) and educational (Bailenson *et al.*, 2008*b*) settings among others. Recently scholars and academic journals have been directing growing levels of attention toward utilising IVEs in a new arena – the legal system. *Presence: Teleoperators and Virtual Environments*, the leading virtual reality journal for over two decades, dedicated a complete journal issue to exploring how virtual environments may be utilised in the legal arena (Barfield *et al.*, 2005).

9.7 Virtual environments

A virtual environment (VE) is 'a digital space in which a user's movements are tracked and his or her surroundings are rendered, or digitally composed and displayed to the senses, in accordance with those movements' (Fox *et al.*, 2009: 95). Cellular phones and traditional videogame consoles are rudimentary examples of virtual environments; the user's key presses are tracked and then those movements are rendered allowing the user to navigate menus or game scenes. More *immersive* forms of virtual environments utilise a combination of software and interactive hardware devices to present the user with the perceptual experience of being surrounded by a virtual, computer-generated environment (Loomis *et al.*, 1999). There are many unique hardware setups that can be utilised to make the user feel surrounded by a virtual world. For example, the CAVE® (computer-assisted virtual environment) system utilises multiple cameras and projection screens that surround the user

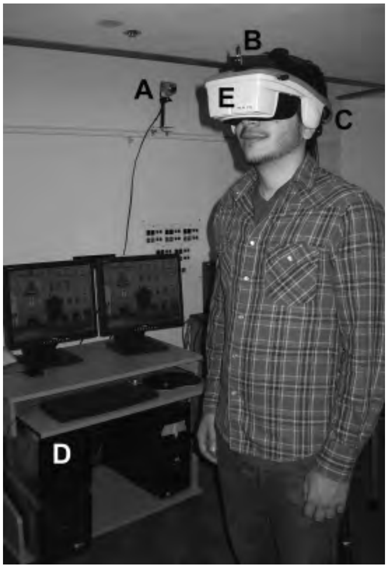


Figure 9.1 An example of a virtual environmental setup. Cameras (A) track an optical sensor (B) indicating the participant's position in the room. An accelerometer (C) gathers information about the participant's head movements. This information is relayed to the computer (D), which determines how the environment is rendered and what the participant sees in the head-mounted display (E).

in an enclosed room. As the user moves, his or her motions are tracked by computer vision operating from the cameras, and the images on the surrounding screens are projected to constantly adjust to the user's new position and perspective.

In another example, as shown in Figure 9.1, the user experiences an immersive virtual environment by donning a virtual reality stereoscopic head-mounted display (HMD); the HMD consists of two miniature display screens – one for each eye, creating a field of view that is wider than a desktop display (commercial products currently range from 20 degrees to 120 degrees in field of view). Each eye receives a slightly different rendering or image at about a 10-centimetre displacement, resulting in the experience of stereoscopic or binocular vision. In addition to the three-dimensional (3D) images presented via the HMD, the immersive virtual environment technology also provides its user with a high level of interactivity. The LED light connected to the HMD is monitored by multiple cameras located around the room. An accelerometer is attached to the HMD and tracks the participant's head movements by a three-axis orientation system. The location of the LED light and the position of the accelerometer are utilised to render an image specific to the user's position in the HMD. The image in the HMD is updated more than 60 times per second.

9.8 Virtual humans

Scholars have been creating virtual human models and using them in psychological research studies for over three decades (Ellis *et al.*, 1975). There are many unique processes for generating virtual models of human busts. Some scholars use 3D scanners which use optical methods to directly measure the shape and albedo (or ability to reflect light) of a person's head and face. In a simpler fashion, 3D models of human busts can be constructed using slider bars to adjust the hairstyle, hair colour, skin colour, facial characteristics and more of a generic human bust, much like it is possible to design personalised virtual humans online and in modern video games. Finally, it is also possible to build 3D busts using photographs and modelling software (Figures 9.2 and 9.3). In Figure 9.2, the front and profile photographs are blended using photogrammetric software to create a replica of the individual's head. The software creates an underlying structure of the head by collocating feature points on the two photographs and then stitches together the images to create a unified texture to wrap around the mesh. This process takes approximately 20 minutes. The 3D model can then be animated to move, blink, speak and generate a multitude of facial expressions. An alternative method is to use a 3D scanner to create a virtual bust by creating a point cloud of the head using a variety of cameras that scan for distance and colour information (Bernardini and Rushmeier, 2002).

The 3D head can then be attached to a virtual body and imported into an IVE using commercially available software. The virtual body can be selected or manipulated on a variety of characteristics such as sex, height, weight, build, race and posture to match the physical human model or parameters desired. Finally, the virtual human busts and bodies can be imported into virtual environments such as a classroom, supermarket, courtroom or park. Imposing photographs of an actual environment, such as a specific alley (Figure 9.4, Panel B), it is possible to build a virtual environment that closely represents a physical world environment. The opportunities for manipulating the virtual environment, the virtual human, and his or her virtual movements and expressions are limited only by the imagination of the programmer.

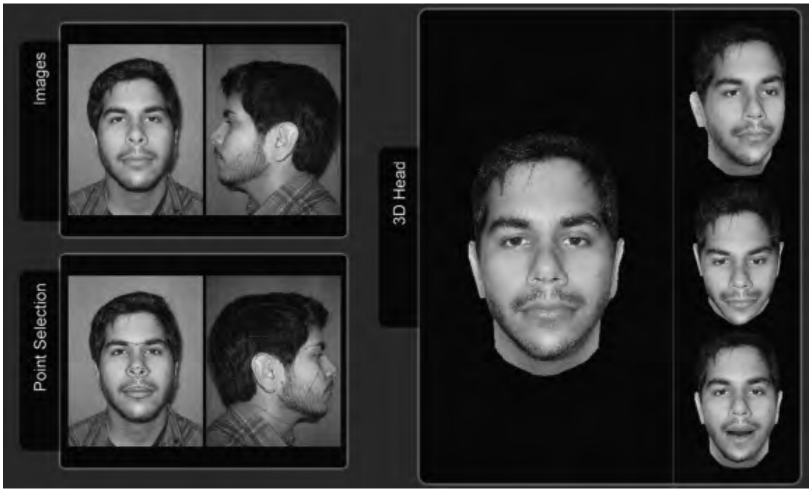


Figure 9.2 The images and modelling techniques used in creating 3D heads. The top left panel shows two photographs of a person, and the bottom left shows the anchor points manually placed on the photographs with photogrammetric software. The right panel shows the resulting 3D head, which is a mesh model with texture wrapped around it.

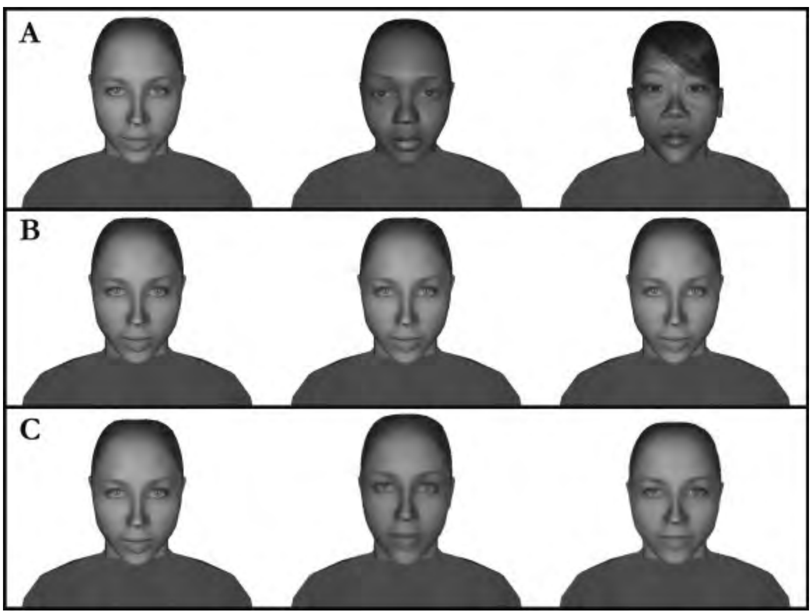


Figure 9.3 Examples of line-ups containing a representation of the suspect (White female on the left in each panel) and foils (all images to the right of the suspect) that could be created using 3D digital modelling techniques. Virtual human identification line-ups with digitally created foils can help avoid unfair line-ups (where the foils are too dissimilar to the suspect such as in Panel A, or conversely where the foils are too similar to the suspect as in Panel B) by digitally creating foils that are similar but not identical to the suspect (Panel C).

9.9 Virtual human line-ups and recent research

Research on the use of virtual humans and immersive virtual environments in eyewitness identification procedures began less than a decade ago (Bailenson *et al.*, 2003*a*) and has already encompassed several unique areas of research.

9.9.1 Identifying virtual busts

Are 3D virtual busts easily identified? In order to answer this question, Bailenson and colleagues conducted a series of studies to see if humans were indeed capable of identifying digital busts of real human beings (Bailenson *et al.*, 2003*b*, 2004). Participants in these studies trained for their identification tasks by watching videos or viewing photographs of the

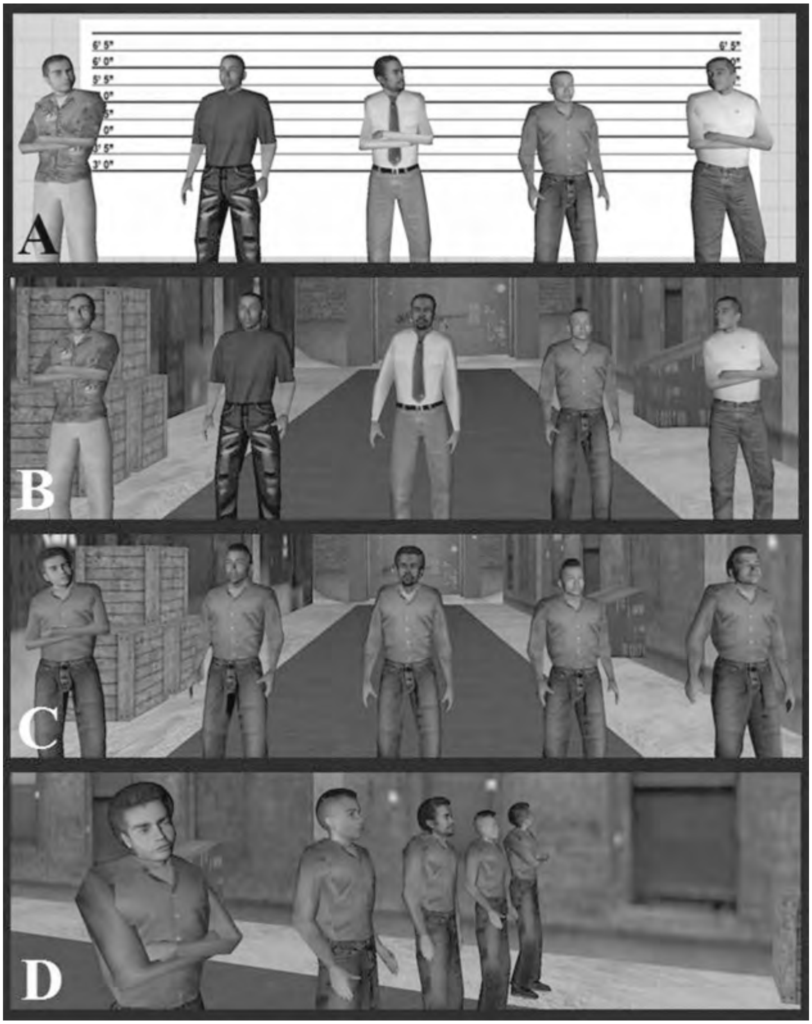


Figure 9.4 An example of a virtual human identification line-up (A). Virtual identification line-ups allow many affordances such as the ability to place line-up members in a virtual version of the actual crime location, such as an alley (B), the ability to give all line-up members similar characteristics such as clothing and hair styles among others (C), and the ability to permit witnesses to move around the line-up members and view them from varying distances and angles (D).

suspects. Afterwards, the participants were asked to identify the humans they had just observed. A test for each human being was conducted using either (1) a pair of photographs of two different individuals – the perpetrator and a foil, or (2) a pair of still-images of digital busts – again of the perpetrator and a foil. The studies revealed that people are quite good at recognising digital busts.

Overall, the results of over a dozen studies involving over 300 experimental participants demonstrate that people are slightly better at recognising photographs of faces than images of virtual busts (as cited in Bailenson *et al.*, 2006). However, in the foreseeable future, as the technology improves, this slight advantage for photographs will likely disappear, if not be completely surmounted. Virtual human busts, in comparison to photographs, can present 3D information concerning depth and be easily portrayed from variable distances and angles (Bailenson *et al.*, 2006).

9.9.2 Virtual suspects and foils

Bailenson *et al.* (2006) recently conducted a study where participants witnessed a staged crime and then examined line-ups within an IVE that contained 3D virtual busts of the culprit and six foils. Using the tracking system described above, participants were given either unlimited viewpoints of the busts (in terms of angle and distance) or a single viewpoint at a single angle and distance (which was either a chosen or assigned viewpoint). Results demonstrated that unlimited viewpoints improved accuracy in culprit-present line-ups but not in culprit-absent line-ups. Additionally, across conditions, post-hoc measurements showed that when the chosen view of the suspect during the line-up was similar to the view during the staged crime in terms of distance, identification accuracy improved. These findings generate optimism about the usefulness of virtual line-ups even though higher fidelity, digital humans may be necessary before such line-ups can be implemented in the legal arena.

9.9.3 Virtual administrators

Virtual humans can be utilised as the targets and foils in a line-up as well as the virtual humans or administrators that conduct the line-up. In a recent study (Cutler *et al.*, 2009), students watched videotaped crimes and attempted to identify the perpetrators from photograph line-ups conducted by either a research assistant playing the role of an investigator or a virtual administrator. The virtual administrator was a 3D representation of a plain-clothed human detective and was displayed on a 19″ flat-panel monitor. Furthermore, the virtual administrator was programmed in such a way that it was able to verbally lead the witnesses through the identification task while engaging in normal communication protocol with turn-taking and feedback (through a set of speakers and a microphone). For more information on the virtual administrator technology, see Daugherty *et al.* (2008).

The participants in the virtual administrator and human administrator conditions performed similarly on the line-up identification task along several dimensions, such as number of correct identifications and number of foil identifications. The successful use of virtual human administrators in line-up procedures is an important finding because virtual administrators can be used to provide a unique advantage over human administrators. Virtual administrators are experimentally *blind*, in that they do not know which individual in the line-up is the suspect. Being blind to which person in the line-up is the suspect or culprit is important, because it eliminates the possibility that an administrator who knows which individual is the suspect or culprit, may intentionally or inadvertently give subtle signals in words or actions that may make the eyewitness more likely to identify the suspect. A human administrator can provide subtle influences that lead eyewitnesses toward a suspect, away from fillers, and provide post-identification feedback to the eyewitness influencing his or her confidence in the identification during trial testimony.

Scientific research has consistently shown that test subjects are influenced by the expectations of those who perform the tests (Adair and Epstein, 1968; Aronson *et al.*, 1990; Phillips *et al.*, 1999; Garrioch and Brimacombe, 2001). A recent experiment demonstrated that witnesses were more likely to make decisions consistent with line-up administrator expectations when the level of contact between the administrator and the witness was high as compared with when it was low (Haw and Fisher, 2004). A meta-analysis of decades of identification data documented this ‘interpersonal expectancy effect’ (Rosenthal and Rubin, 1978). Additionally, a meta-analysis of the level of bias in instructions given to eyewitnesses participating in line-ups or photo arrays has also documented a higher rate of false positives when the instructions provided are biased (Steblay, 1997). In fact, in a series of rules that Wells and colleagues present for minimising false identifications, the first rule states that ‘the person who conducts the line-up or photospread should not be aware of which member of the line-up or photospread is the suspect’ (Wells *et al.*, 1998: 627).

9.10 Advantages of immersive virtual line-ups for future work

Wells and Luus suggest that ‘one of the guiding heuristics for exploring possible improvements to line-up methods is the analogy between conducting a line-up and conducting a psychology experiment’ (as cited in Wells, 1993: 557). A few parallels between a line-up and an experiment include the fact that investigators have a hypothesis (that the suspect is the culprit), questions are asked (such as ‘Do you see the culprit among these people?’), responses are recorded, and the original hypothesis is reevaluated in light of the new data (Wells, 1993). However, as Wells (1993) notes, actual line-ups generally lack the kinds of controls that are essential to making clear inferences from the data.

IVET has been described as an impressive tool for improving the study of social psychological processes (Blascovich *et al.*, 2002). More specifically, IVET helps to improve the trade-off between making experimental scenarios as realistic and similar to actual world scenarios as possible while still maintaining a high level of control, so that effects and their causes can be more easily parsed out. Accordingly, Bailenson and colleagues suggest that immersive virtual line-ups allow a greater level of control than traditional photograph or live line-ups, because they permit every single user (jurors, witnesses or suspects) to experience the same sensory input during the line-up procedure (Bailenson *et al.*, 2006).

IVET can, in experiments and line-ups alike, simultaneously maximise the control and realism of the situation by creating virtual environments and virtual humans that are highly realistic and yet highly controlled. Below we discuss several ways in which virtual human line-ups can add additional control to traditional line-up procedures as well as other advantages of virtual human line-ups.

9.10.1 Line-up composition

In order for a line-up to be fair, the fillers must be reasonable alternatives to the suspect vis-à-vis the description provided by the eyewitness (Brigham *et al.*, 1999). The goal of a line-up should be to truly test the eyewitness memory, rather than to employ a composition that suggests an identification of the suspect. Immersive virtual environments offer an opportunity for increased fairness in identification line-ups.

Research conducted specifically on the usage of foils in identification line-ups suggests that if a line-up is to be a fair test of an eyewitness memory it should contain foils that match the eyewitness description of the perpetrator, loosely resemble the suspect, are different enough to be discernible from one another, and possess features that reduce the likelihood of guessing or the use of deductive reasoning. When an innocent suspect is the only person to fit the description provided by the eyewitness (or more closely matches the perpetrator than the fillers), the eyewitness develops a preference for the suspect and his corresponding confidence level is greater.

When fillers that match the eyewitness description are selected, identification accuracy rates go up and false identifications (and corresponding inflated witness confidence) can be minimised (Wells *et al.*, 1993). Accordingly, fillers should be selected based on their resemblance to the description provided by the eyewitness, but should not unduly stand out from the suspect. For example, if the culprit is described as a tall White female, and the line-up consists of four non-White females and only one White female, the identification task will not be a challenging enough test of the witness's memory (see Figure 9.3, Panel A for a digitally created example). On the other hand, if the line-up consists of five tall, White females that are nearly identical, the identification task may be too difficult or challenging for the witnesses (see Figure 9.3, Panel B for a digitally created example).

In a series of studies (Doob and Kirshenbaum, 1973; Wells *et al.*, 1979*b*; see Lindsay and Wells, 1980 for a review), unsuspecting participants witnessed a crime and were asked to identify the criminal from either a fair or unfair line-up that either contained the culprit (culprit-present line-up) or contained an innocent individual who resembled the culprit (culprit-absent line-up). Results from the studies show that high-similarity line-ups, in comparison to low-similarity line-ups, yielded fewer identifications of the criminal and of the innocent suspect. However, it is noteworthy to point out that the reduction in identifications of the innocent suspect was much more significant than the reduction of identifications of the guilty criminal (Lindsay and Wells, 1980). Furthermore, an analysis of the data by Wells and Lindsay concluded that 'identification evidence obtained from relatively fair high-similarity line-ups is superior to similar evidence obtained from relatively unfair low-similarity line-ups' (as Lindsay and Wells, 1980: 303).

IVE line-ups have the advantage of allowing for a more controlled, fair use of foils, in that investigators and police would not have to rely on recruiting actual human beings who are similar to the description of the culprit. Instead, foils could be drawn from a large database of digital busts (Bailenson *et al.*, 2006: 260). In addition, Bailenson *et al.* (2006) point out that the underlying digital mesh and texture that composes each virtual bust (see Figure 9.2) can be easily manipulated using morphing software (such as Magic Morph). Thus it becomes quite easy to digitally create foils by manipulating the underlying mesh or by pigmenting the texture map of the suspect to whatever degree of dissimilarity is desired. Furthermore, scholars have been studying the means and procedures by which computers can help identify faces or select faces that are similar to a target image for several decades (Ellis *et al.*, 1989; Troje and Bulthoff, 1996).

Foils of differing levels of similarity to the suspect can be created with relative ease using digital modelling, morphing and texturing software. For example, Panel C of Figure 9.3 shows digital representations of five line-up members that have been created by morphing (or blending) the suspect's head (the image on the far left) with other images of White females to create line-up members that are similar, but not identical, to the suspect. In this respect the line-up is both a fair and challenging test of the witness memory.

9.10.2 Contextual matching

Through decades of research, scholars have outlined a variety of conditions that affect the accuracy of eyewitness identification. One condition that is shown to hinder accuracy in line-up identifications is the mismatch of contextual or background surroundings that the eyewitness experiences between the crime and the line-up procedure (Malpass and Devine, 1981; Krafka and Penrod, 1985). Previous research shows that this mismatch of contexts between encoding and recall can be improved by conducting line-ups in the same physical location as the crime occurred (Davies and Milne, 1985), by merely showing the eyewitness photographs of the environment where the crime took place (Cutler *et al.*, 1987), or by showing the witness objects from the physical environment (Krafka and Penrod, 1985). In fact, Memon and Bruce (1983) demonstrated that in certain situations, the background context of a photograph can have a greater impact on recognition than the facial features themselves.

The affordances of IVET would allow line-ups to improve upon previous methods of matching encoding and retrieval contexts by creating a controlled and safe retrieval context that is highly similar to that of the encoding context. For example, using digital photographs of the actual crime scene, programmers and graphic artists can model a full size, 3D version of the actual crime scene (see Figure 9.4, Panel B). Additionally, rendering the specific details of certain weather conditions, specific amounts of daylight and the precise location of both objects and individuals in the environment becomes possible. The richness of the environment would far surpass what is possible from photographs alone; 3D virtual environments can convey unique information relating to depth and multisensory input. The eyewitness could once again feel immersed in the environment where the crime occurred (without actually interacting with the physical suspect in that context) which has very powerful implications for enhanced recall accuracy.

9.10.3 Mode

Previous studies show that faces presented in one mode (such as two-dimensional line drawings or movie clips) are more likely to be recognised when they are presented for identification in the same mode during both the testing and training phases (compared with when the two modes were different in training and testing). For example, faces studied in photographs during the learning phase are more likely to be recognised in the test phase when they are presented again via photographs as opposed to detailed line drawings (Davies *et al.*, 1978). In addition, Patterson (1978) found that faces studied in photographs during the learning phase are more likely to be recognised in the test phase when they were again presented via photographs as opposed to several seconds of film.

Bruce *et al.* (1999) revealed an even more striking finding when they asked participants to match an image of a target face taken from high quality video against an array of still photographs of faces (that either included or did not include the target face). The task did not require the participant to remember the target face – only match the target face to other visual stimuli. Still, participants were correct a mere 70% of the time, or even less when there was variation in the target facial expression or pose. According to this research and in order to maximise correct identifications, a line-up should be presented in a 3D form (using either live human beings or 3D models of human beings as can be seen in Figure 9.4) so that the members of the line-up are viewed in a mode very similar to the realistic, 3D mode in which they were originally witnessed during the crime.

9.10.4 Adaptable distance and angle

Research shows that matching the viewpoint between encoding (at the time of the crime) and recall (during the identification line-up) is an important factor in increasing correct identifications. For example, Bruce (1982) found that correct identification rates in recognition memory for identical pictures were 90% correct when images were presented from similar viewpoints across presentation and testing sessions as compared with 72% correct when viewpoint was changed between presentation and test.

Immersive virtual human line-ups would allow witnesses to adjust the distance and angle with which they choose to view the virtual suspects and foils. As explained by Bailenson *et al.* (2006), one unique capability of IVEs is that users may actively and freely move about an environment to examine the suspects and foils. Line-ups may be improved by allowing the eyewitness to recreate the distance and angle of viewpoint from which she witnessed the crime. For example, if during a crime the witness's only view of the culprit was an eye-level profile view, he or she may wish to recreate this view. Figure 9.4, Panel D depicts a view from the virtual line-up that recreates the witness's viewpoint during the crime. In an immersive virtual human line-up, the witness can move freely about the environment, adjusting his or her body and head position to create the exact distance and angle of view experienced at the crime scene.

Allowing a witness to approach a suspect or foil from any angle or distance in a virtual world is an important capability that is not currently available by any other means. In a live line-up, forcing a witness to approach and examine possible suspects could be dangerous or emotionally traumatic. In addition, this would compromise the witness's anonymity. In photograph and video line-ups, the captured information is constrained by the degree to which they can be presented in varying distances and angles, and it is unrealistic to take large arrays of photographs and video-feeds of suspects and foils from every distance and angle (Bailenson *et al.*, 2006).

9.10.5 Adaptable facial expressions and attire

In 1982, Bruce found that hit rates in recognition memory for identical pictures were 90% correct compared with 81% correct when expression was changed from smiling to unsmiling. Using the animation and morphing software, it is possible to depict virtual humans with specific facial expressions such as anger, sadness, or astonishment. In this way, the facial expression of the criminal during the crime can be digitally manipulated, replicated and assigned to each virtual human in the line-up.

Along similar lines, virtual line-up members can be depicted in a variety of clothing styles and hairstyles that is unparalleled by the physical world (Bailenson *et al.*, 2008*a*). In Panel C of Figure 9.4, all of the line-up members are presented with the same clothing (red button-down shirts and dark jeans) and similar hairstyles (variations of the afro). Lindsay *et al.* (1987) conducted a study where participants witnessed a staged crime and then participated in a photographic line-up procedure. Line-up attire was manipulated such that (1) only the perpetrator wore clothing similar to that worn during the crime; (2) all line-up members wore different attire; or (3) all line-up members were dressed alike. Results showed that while the rate of identifications of the guilty individual was not influenced by line-up attire, innocent suspects were most likely to be misidentified from a line-up where only the perpetrator wore clothing similar to that worn during the crime. While further research on the effects of clothing style in line-ups needs to be conducted, IVET allows both a context for this research and an easier means for manipulating diverse clothing styles in virtual line-ups than in traditional line-ups.

9.10.6 Motion and animation

The benefits of motion in facial identification contexts are much less clear, especially for unfamiliar faces. Some studies have found that seeing unfamiliar faces (faces that have only been seen for a short amount of time) in motion improves recognition memory during subsequent identification tasks, while other studies have reported no advantages (see Bruce, 2009 for a review). For familiar faces, Lander and Bruce (2000) suggest that the benefits of motion may arise because characteristic patterns of motion are stored in memory and can help activate the appropriate person identity when the static visual form system is deficient.

In virtual line-ups, the suspects and foils can be animated to display generically programmed movement or specific movement recorded from an actual person's bodily motions (Bailenson and Yee, 2005). While Figure 9.4 is a combination of still shots from a virtual human identification line-up, it is still possible to see that the virtual humans are slowing shifting their weight from side to side, looking around, and crossing their arms occasionally. While the motion of suspects and foils in line-ups has not been thoroughly explored, more general research demonstrates that the processes governing perception of human faces have a substantial spatial gestalt component (Tanaka and Sengco, 1997; Farah *et al.*, 1998). These studies provide evidence that not only are the local features of a face important in recognition but also the global configuration of those local features, which shifts throughout movement at different angles and distances (Bailenson *et al.*, 2006).

9.10.7 Recording of the identification procedure

The Innocence Project recommends the creation of an electronic record of the identification procedure in order to preserve the precise communications, verbal and non-verbal, made by the eyewitness and the line-up administrator, including subtle clues provided by the administrator that could affect the witness's identification or confidence level and the witness's specific statement of certainty made upon identification (Innocence Project, 2008). In addition to preserving the results of all identifications and non-identifications (including eyewitness identification of fillers), a virtual line-up would also preserve identification information, such as the sources of all virtual busts used, the virtual busts themselves, the manner in which they were presented to the eyewitness, and the way in which the eyewitness observed the virtual line-up members (the amount of interpersonal distance maintained, the viewing angles and viewing time) for the jury to view at trial.

9.10.8 Portability

> A major factor contributing to the high incidence of miscarriages of justice from mistaken identification has been the degree of suggestion inherent in the manner in which the prosecution presents the suspect to witnesses for pretrial identification. . . . There is grave potential for prejudice, intentional or not, in the pretrial lineup, which may not be capable of reconstruction at trial (United States v. Wade, 1967).

Traditionally, line-ups are conducted during pretrial investigation. However, using immersive virtual line-ups, legal teams could bring line-ups (or recordings of line-ups) into the courtroom. The counsel, jury, judge and others could view 2D images of what the eyewitness experiences inside the HMD, or perhaps in a technologically advanced courtroom each individual might have his or her own HMD to view the virtual line-up. Transporting the line-up identification procedure from a physically distant room and distant time to the courtroom could foster high-impact demonstrations of identifications (Bailenson *et al.*, 2006), and give the judge or jurors a first-hand perspective of how the eyewitness comes to make an identification decision.

Leonetti and Bailenson (2010) state that the use of an IVE during a jury trial can 'be viewed as merely another point along a line of technological progression, from scene viewing to photography to video evidence to virtual evidence. Employing an IVE during trial would be no different in substance than the admission of other types of testimonial, photographic, and demonstrative evidence that courts have permitted for decades' (p. 1119). As scholars and professionals, we are nearing a time when IVEs may become an influential and pervasive form of technology in the legal arena.

9.10.9 Unobtrusive measures

In IVEs, a large variety of unobtrusive measures become easily accessible. Campbell and colleagues defined 'unobtrusive measures' as measures that are made without the awareness of them being measured (Webb *et al.*, 2003). Unobtrusive measures are very applicable to research conducted using IVET. IVET is capable of collecting many different types of non-verbal behaviour (such as body position, head position, gaze direction and others). Another example of an unobtrusive measure that is readily accessible via IVET is the distance a user maintains between himself and other individuals in the virtual environment. The LED on top of the HMD (see Figure 9.1) constantly monitors the user position in the virtual environment and allows continual recording of the interpersonal distance the user maintains between herself and other virtual humans.

It is quite possible that such unobtrusive measures could be gathered and analysed to reveal telling patterns in eyewitness non-verbal behaviours (Bailenson *et al.*, 2003*b*). In a recently conducted study, Yee *et al.* (2011) followed 76 students over a 6-week period as they engaged in social interactions in a virtual world called Second Life. The linguistic and behavioural output of each of the 76 students was recorded and analysed. Results from the study show that unobtrusively recorded virtual behavioural data can be used to predict certain characteristics (such as personality qualities) about the real-life humans behind the virtual interactions. Recording and analysing the behavioural output of witnesses in virtual human identification line-ups may prove to be an extremely informative line of research. For example, the eyewitness may subconsciously maintain greater interpersonal distance between himself and the actual culprit than between himself and the foils in the line-up.

9.11 Conclusion

Immersive virtual human identification line-ups certainly will not ameliorate all of the challenges associated with traditional, physical world line-ups. The quality and realism of virtual reality graphics and image rendering, which continually improve over time, still have more room for improvement. However, we do believe that the advantages afforded by immersive virtual line-ups create a compelling case for further research, testing and possible implementation in future court cases.

Virtual line-ups have the capability to aid scholars in identifying the positive and negative aspects of live and photographic line-ups, because they allow the isolation of variables such as viewpoint, emotional expressions, and context. Aside from serving as a research tool for conducting controlled experiments comparing line-ups procedures and structures, virtual line-ups present a feasible, perhaps even superior, alternative to live and photographic line-ups. It is imperative that as technology advances, we continue to consider and evaluate the practicality of using virtual line-ups as a part of our modern-day legal system.

References

Adair, J. G., and Epstein, J. S. (1968). Verbal cues in the mediation of experimenter bias. *Psychological Reports*, **22**, 1045–1053.

Aronson, E., Ellsworth, P. C., Carlsmith, J. M. and Gonzales, M. H. (1990). *Methods of Research in Social Psychology*. New York: McGraw-Hill.

Bailenson, J. N., Beall, A. C. and Blascovich, J. (2003*a*). Using virtual heads for person identification: an empirical study comparing photographs to photogrammetrically-generated models. *Journal of Forensic Identification*, **53**(6), 722–728.

Bailenson, J. N., Beall, A. C., Blascovich, J. and Rex, C. (2004) Examining virtual busts: are photogrammetrically-generated head models effective for person identification? *PRESENCE: Teleoperators and Virtual Environments*, **13**(4), 416–427.

Bailenson, J. N., Blascovich, J., Beall, A. C. and Loomis, J. M. (2003*b*). Interpersonal distance in immersive virtual environments. *Personality and Social Psychology Bulletin*, **29**, 1–15.

Bailenson, J. N., Blascovich, J., Beall, A. C. and Noveck, B. (2006). Courtroom applications of virtual environments, immersive virtual environments, and collaborative virtual environments. *Law and Policy*, **28**(2), 249–270.

Bailenson, J. N., Davies, A., Blascovich, J. *et al.* (2008*a*). The effects of witness viewpoint distance, angle, and choice on eyewitness accuracy in police lineups conducted in immersive virtual environments. *Presence*, **17**(3), 242–255.

Bailenson, J. N., and Yee, N. (2005). Digital chameleons: Automatic assimilation of nonverbal gestures in immersive virtual environments. *Psychological Science*, **16**(10), 814–819.

Bailenson, J. N., Yee, N., Blascovich, J. *et al.* (2008*b*). The use of immersive virtual reality in the learning sciences: digital transformations of teachers, students, and social context. *Journal of the Learning Sciences*, **17**, 102–141.

Barfield, W., Lauria, R., Mann, S. and Kerr, I. (2005). Special section: legal, ethical, and policy issues associated with virtual environments and computer mediated reality (Guest editors' introduction). *Presence: Teleoperators and Virtual Environments*, **14**(6), iii–v.

Behrman, B. W., and Davey, S. L. (2001). Eyewitness identification in actual criminal cases: an archival analysis. *Law and Human Behavior*, **25**(5), 475–491.

Bernardini, F. and Rushmeier, H. E. (2002). The 3D model acquisition pipeline. *Computer Graphics Forum*, **21**(2), 149–172.

Blascovich, J., Loomis, J., Beall, A. *et al.* (2002). Immersive virtual environment technology as a methodological tool for social psychology. *Psychological Inquiry*, **13**, 103–124.

Bradfield, A. L., and Wells, G. L. (2000). The perceived validity of eyewitness identification testimony: a test of the five Biggers criteria. *Law and Human Behavior*, **24**, 581–594.

Bradfield, A. L., Wells, G. L. and Olson, E. A. (2002). The damaging effect of confirming feedback on the relation between eyewitness certainty and identification accuracy. *Journal of Applied Psychology*, **87**, 112–120.

Brigham, J. C., Maass, A., Snyder, L. D. and Spaulding, K. (1982). Accuracy of eyewitness identifications in a field setting. *Journal of Personality and Social Psychology*, **42**(4), 673–681.

Brigham, J. C., Meissner, C. A. and Wasserman, A. W. (1999). Applied issues in the construction and expert assessment of photo lineups. *Applied Cognitive Psychology*, **13**, S73–S92.

Bruce, V. (1982). Changing faces: visual and non-visual coding processes in face recognition. *British Journal of Psychology*, **37**(1), 105–116.

Bruce, V. (2009). Remembering faces. In J. R. Brockmole (Ed.), *The Visual World in Memory*, pp. 66–88. Hove: Psychology Press.

Bruce, V., Henderson, Z., Greenwood, K. *et al.* (1999). Verification of face identities from images captured on video. *Journal of Experimental Psychology: Applied*, **5**(4), 339–360.

Cutler, B. L., Daugherty, B., Babu, S., Hodges, L. and Van Wallendael, L. (2009). Creating blind photoarrays using virtual human technology. *Police Quarterly*, **12**(3), 289–300.

Cutler, B. L., Penrod, S. D. and Martens, T. K. (1987). The reliability of eyewitness identifications: the role of system and estimator variables. *Law and Human Behavior*, **11**, 233–258.

Daugherty, B., Babu, S., Van Wallendael, L. R., Cutler, B. L. and Hodges, L. (2008). A comparison of virtual human vs. human administration of police lineups. *IEEE Computer Graphics and Applications*, **28**(6), 65–75.

Davies, G., Ellis, H. D. and Shepherd, J. (1978). Face recognition accuracy as a function of mode of representation. *Journal of Applied Psychology*, **63**(2), 180–187.

Davies, G. and Milne, A. (1985). Eyewitness composite production: a function of mental or physical reinstatement of context. *Criminal Justice and Behavior*, **12**(2), 209–220.

Doob, A. N., and Kirshenbaum, H. (1973). Bias in police lineups – partial remembering. *Journal of Police Science and Administration*, **1**, 287–293.

Doyle, J. M. (2005). *True Witness: Cops, Courts, Science, and the Battle against Misidentification*. New York: Palgrave Macmillan.

Ellis, H., Shepherd, J. and Davies, G. (1975). An investigation of the use of the Photo-fit technique for recalling faces. *British Journal of Psychology*, **66**(1), 29–37.

Ellis, H. D., Shepherd, J. W., Flin, R. H., Shepherd, J. and Davies, G. M. (1989). Identification from a computer-driven retrieval system compared with a traditional mug-shot album search: a new tool for police investigation. *Ergonomics*, **32**, 167–177.

Farah, M. J., Wilson, K. D., Drain, M. and Tanaka, J. N. (1998). What is 'special' about face perception? *Psychological Review*, **105**(3), 482–498.

Fox, J., Arena, D. and Bailenson, J. N. (2009). Virtual reality: a survival guide for the social scientist. *Journal of Media Psychology*, **21**(3), 95–113.

Garrioch, L. and Brimacombe, C. A. E. (2001). Lineup administrators' expectations: their impact on eyewitness confidence. *Law and Human Behavior*, **25**, 299–315.

Goodman, G. S., Tobey, A. E., Batterman-Faunce, J. M. *et al.* (1998). Face-to-face confrontation: effects of closed-circuit technology on children's eyewitness testimony and jurors' decisions. *Law and Human Behavior*, **22**(2), 165–203.

Haw, R. M., and Fisher, R. P. (2004). Effects of administrator-witness contact on eyewitness identification accuracy. *Journal of Applied Psychology*, **89**, 1106–1112.

Innocence Project (December, 2008). *Eyewitness Identification Resource Guide*.

Koosed, M. M. (2002). The proposed innocence protection act won't – unless it also curbs mistaken eyewitness identifications. *Ohio State Law Journal*, **63**, 263.

Kerstholt, J. H., Koster, E. R. and van Amelsvoort, A. G. (2004). Eyewitness: a comparison of live, video and photo line-ups. *Journal of Police and Criminal Psychology*, **19**(2), 15–22.

Krafka, C. and Penrod, S. (1985). Reinstatement of context in a field experiment on eyewitness identification. *Journal of Personality and Social Psychology*, **49**(1), 58–69.

Lander, K. and Bruce, V. (2000). Recognizing famous faces: exploring the benefits of facial motion. *Ecological Psychology*, **12**, 259–272.

Lanier, J. (2001). Virtually there. *Scientific American*, **24**, 66–75.

Leonetti, C. and Bailenson, J. N. (2010). High-tech view: the use of immersive virtual environments in jury trials. *Marquette Law Review*, **93**(3), 1073.

Lindsay, D. S., Read, J. D. and Sharma, K. (1998). Accuracy and confidence in person identification: the relationship is strong when witnessing conditions vary widely. *Psychological Science*, **9**(3), 215–218.

Lindsay, R. C., Wallbridge, H. and Drennan, D. (1987). Do the clothes make the man? An exploration of the effect of lineup attire on eyewitness identification accuracy. *Canadian Journal of Behavioural Science*, **19**(4), 463–478.

Lindsay, R. C., and Wells, G. L. (1980). What price justice? Exploring the relationship of lineup fairness to identification accuracy. *Law and Human Behavior*, **4**(4), 303–313.

Lindsay, R. C., Wells, G. L. and Rumpel, C. M. (1981). Can people detect eyewitness identification accuracy within and across situations? *Journal of Applied Psychology*, **66**, 77.

Loomis, J. M., Blascovich, J. J. and Beall, A. C. (1999). Immersive virtual environment technology as a basic research tool in psychology. *Behavior Research Methods, Instruments, and Computers*, **31**(4), 557–564.

Malpass, R. S., and Devine, P. G. (1981). Realism and eyewitness identification research. *Law and Human Behavior*, **5**, 299–301.

Malpass, R. S., and Lindsay, R. C. (1999). Measuring lineup fairness. *Applied Cognitive Psychology*, **13**(S1), S1–S7.

Manson v Brathwaite, 432 U.S. 98 (1977).

Memon, A. and Bruce, V. (1983). The effects of encoding strategy and context change on face recognition. *Human Learning: Journal of Practical Research and Applications*, **2**(4), 313–326.

National Science Foundation (January 3, 1997). False identification: new research seeks to inoculate eyewitnesses against errors. *National Science Foundation Press Release*. Retrieved 25 May 2010, http://www.nsf.gov/pubs/stis1997/pr971/pr971.txt

Neil v Biggers, 409 U.S. 188 (1972).

Patterson, K. E. (1978). Person recognition: More than a pretty face. In M. M. Grunneberg, P. E. Morris and R. N. Sykes (Eds.), *Practical Aspects of Memory* (pp. 227–235). London: Academic Press.

Phillips, M. R., McAuliff, B. D., Kovera, M. B. and Cutler, B. L. (1999). Double-blind lineup administrations as a safeguard against investigator bias. *Journal of Applied Psychology*, **84**, 940–951.

Reznek, M., Harter, P. and Krummel, T. (2008). Virtual reality and simulation: training the future emergency physician. *Academic Emergency Medicine*, **9**(1), 78–87.

Riva, G., Wiederhold, B. K. and Molinari, E. (Eds.) (1998). *Virtual Environments in Clinical Psychology and Neuroscience: Methods and Techniques in Advanced Patient–Therapist Interaction*. Amsterdam: IOS Press.

Rosenthal, R. and Rubin, D. B. (1978). Interpersonal expectancy effects: the first 345 Studies. *Behavioral and Brain Sciences*, **3**, 377–386.

Semmler, C., Brewer, N. and Wells, G. L. (2004). Effects of postidentification feedback on eyewitness identification and nonidentification confidence. *Journal of Applied Psychology*, **89**(2), 334–346.

Sporer, S. L., Penrod, S., Read, D. and Cutler, B. (1995). Choosing, confidence, and accuracy: a meta-analysis of the confidence-accuracy relation in eyewitness identification studies. *Psychological Bulletin*, **118**(3), 315–327.

Steblay, N. M. (1997). Social influence in eyewitness recall: a meta-analytic review of lineup instruction effects. *Law and Human Behavior*, **21**, 283–298.

Troje, N. F., and Bulthoff, H. H. (1996). Face recognition under varying poses: the role of texture and shape. *Vision Research*, **36**(12), 1761–1771.

Tanaka, J. W., and Sengco, J. (1997). Features and their configuration in face recognition. *Memory and Cognition*, **25**, 583–592.

United States v Downing, 753 F.2d 1224 (3d. Cir. 1985).

United States v Wade, 388 U.S. 218 (1967).

Valentine, T., Pickering, A. and Darline, S. (2003). Characteristics of eyewitness identification that predict the outcome of real lineups. *Applied Cognitive Psychology*, **17**(8), 969–993.

Webb, E. J., Campbell, D. T., Schwarz, R. D. and Sechrest, L. (2003). The use of archival sources in social research. In M. Bulmer, (Ed.) *Sociological Research Methods: An Introduction*, pp. 113–130. Piscataway, NJ: Transaction Publishers.

Wells, G. L. (1993). What do we know about eyewitness identification? *American Psychologist*, **48**(5), 553–571.

Wells, G. L., and Bradfield, A. L. (1998). 'Good, you identified the suspect:' Feedback to eyewitnesses distorts their reports of the witnessing experience, *Journal of Applied Psychology*, **83**, 360–376.

Wells, G. L., Ferguson, T. J. and Lindsay, R. C. (1981). The tractability of eyewitness confidence and its implication for triers of fact. *Journal of Applied Psychology*, **66**, 688–696.

Wells, G. L., Leippe, M. R. and Ostrom, T. M. (1979*b*). Guidelines for empirically assessing the fairness of a lineup. *Law and Human Behavior*, **3**(4), 285–293.

Wells, G. L., Lindsay, R. C. and Ferguson, T. J. (1979*a*). Accuracy, confidence, and juror perceptions in eyewitness identification, *Journal of Applied Psychology*, **64**, 440–448.

Wells, G. L., Lindsay, R. C. and Tousignant, J. P. (1980). Effects of expert psychological advice on juror judgments in eyewitness testimony. *Law and Human Behavior*, **4**, 275–285.

Wells, G. L., and Loftus, E. F. (2003). Eyewitness memory for people and events. In A. M. Goldstein (Ed.), *Handbook of Psychology: Forensic Psychology*, pp. 149–160. New York, NY: John Wiley and Sons.

Wells, G. L., and Olson, E. A. (2003). Eyewitness testimony. *Annual Review in Psychology*, **33**, 277–295.

Wells, G. L., and Quinlivan, D. S. (2009). Suggestive eyewitness identification procedures and the Supreme Court's reliability test in light of eyewitness science: 30 years later. *Law and Human Behavior*, **33**, 1–24.

Wells, G. L., Rydell, S. M. and Seelau, E. P. (1993). The selection of distracters for eyewitness lineups. *Journal of Applied Psychology*, **78**, 835–844.

Wells, G. L. and Seelau, E. P. (1995). Eyewitness identification: psychological research and legal policy on lineups. *Psychology, Public Policy, and Law*, **1**, 765–791.

Wells, G. L., Small, M., Penrod, S. *et al.* (1998). Eyewitness identification procedures: recommendations for lineups and photospreads. *Law and Human Behavior*, **22**(6), 1–39.

Yacona, R. (2005). Manson v. Brathwaite: The Supreme Court's misunderstanding of eyewitness identification. *John Marshall Law Review*, **39**, 539–561.

Yee, N., Harris, H., Jabon, M. and Bailenson, J. N. (2011). The expression of personality in virtual worlds. *Social Psychology and Personality Science*, **2**, 5–12.

Chapter 10

Computer-generated face models

Bernard Tiddeman

10.1 Introduction

The rapid progress in digital image capture and computer processing power has provided new opportunities for the analysis and synthesis of facial shape and appearance. New algorithms have been developed to exploit these improvements and have found applications in health, security and forensics. In this chapter some of the main techniques that have been developed to capture, model and process facial image data are reviewed.

10.2 Data capture

10.2.1 Photography

The simplest technique for capturing facial image data is photography. There have been rapid improvements in the resolution of digital cameras, so that the quality of direct digital capture equals or even exceeds chemical film. While two-dimensional (2D) images present the simplest method for facial image capture, 2D data present a number of problems when it comes to analysing and processing the data. Faces are inherently 3D in nature and the projection to 2D can introduce variation in the appearance that is not intrinsic to the faces themselves. Even small changes in the facial orientation can produce a vastly different projected 2D facial shape. Changes in lighting are also difficult to model and remove using purely 2D methods.

10.2.2 Structured light systems

Several techniques have been devised to capture the 3D shape of the facial surface. One of the earliest methods, and still popular, is the laser-video capture system (Jarvis, 1983; Besl, 1988). Typically a laser is fanned into a plane of light by passing it through a cylindrical lens. The intersection of the laser plane and the surface produces a contour that can be observed using a video camera. The laser plane can be easily segmented from the video image, and by using knowledge of the geometry of the system the 2D locations in the video image can be converted to 3D points. The laser-stripe and video camera are moved across the surface either by hand or by a motor-driven system to capture the entire surface. The data from laser-video systems is typically high resolution but can suffer from a number of drawbacks:

(1) The capture is relatively slow, and so it is possible for the subject to move.
(2) Faces captured from such systems often exhibit low-resolution oscillations on the surface. It is not clear if this is due to vibrations of the capture system or is due to the pulsing of blood in the face being captured.
(3) The laser-stripe can be obscured by other parts of the surface. This problem can be reduced by using multiple stripes and/or video cameras.
(4) Specular reflections of the laser stripe can cause problems, but this is not a major issue with facial capture and can be further mitigated by using polarised light and filters (Clark *et al.*, 1997).

Other systems have been developed to improve the capture speed. For example, phase measuring profilometry (PMP) uses a projected fringe pattern, usually a sinusoidal intensity profile (Srinivasan *et al.*, 1984). A minimum of three captures of the shifted fringe pattern are required in order to remove the effects of ambient lighting and the underlying reflectivity of the surface to give the fringe line intensity (or equivalently the angle of the intensity profile sine wave). This still leaves the problem of deciding which fringe a particular point belongs to. This aliasing problem can be reduced by an unwinding algorithm, but problems can arise due to occlusions. More recently, multiple

Craniofacial Identification, ed. C. Wilkinson and C. Rynn. Published by Cambridge University Press.

coloured stripes have been segmented from a projected stripe pattern to develop a real-time face-scanning system (Fechteler and Eisert, 2008).

10.2.3 Stereoscopic systems

Stereo matching algorithms have long been of interest to the computer vision community (Grimson, 1985; Kanade, 1994) and several companies have developed stereoscopic systems for facial capture, e.g. 3dMD (http://www.3dmd.com/) and Dimensional Imaging (http://www.di3d.com). The idea is to capture the face from two slightly different angles and use knowledge of the location of the cameras and the internal camera parameters to allow any points detected in both images to be converted to a 3D location. The difficulty lies in reliably matching points between the two images. Early systems used (by today's standards) low-resolution cameras, which were unable to find features suitable for matching on many areas of the face (such as the forehead and cheeks). To reduce this problem a pattern can be projected on to the face, such as a random speckle pattern. More recent systems are able to use the skin texture in high-resolution images to perform matching without the need for a projected speckle pattern, reducing the complexity of the capture system. This allows real-time capture of moving faces, although reconstruction of the 3D surface at each frame requires additional offline processing. The quality of the captured facial surface can also vary widely, depending on the stereo matching algorithm used (often these are proprietary and the details are not publicly available). Some systems take a parsimonious approach and reject pixels where the matching confidence is low. Others accept far more points, but these may be inaccurate or not connected to the face (e.g. background pixels). Interpolation algorithms can be used to estimate missing values, but little can be done about incorrect values without more information about the surface under study.

10.2.4 Volumetric (CT and MRI)

Finally, there are volumetric medical capture systems which capture the external and internal structure of the head or other organs (Udupa and Herman, 2000). Computed tomography (CT) uses collimated X-rays which are projected through the patient from a number of different directions; these radial projections can be reconstructed into a slice through the patient using one of a number of algorithms (such as convolution back projection). These slice images show highest contrast for bone, but can also reveal the skin and internal soft tissues at lower contrast. Due to the dangers associated with ionising radiation (a routine head CT delivers approximately 15 times the dose of a single chest X-ray), CT is not considered safe for non-clinical use (Shrimpton *et al.*, 2003). Recent advances in CT capture such as helical or cone-beam systems allow faster capture and so lower dosage (Mozzo *et al.*, 1998).

Magnetic resonance imaging (MRI) uses a strong magnetic field to align the spin of protons in hydrogen nuclei (largely in water or fatty tissues) (Lauterbur, 1974; Damadian *et al.*, 1977; Hinshaw *et al.*, 1977). A radio frequency (RF) pulse can be used to knock the direction of alignment of these protons into a higher energy state. As these decay back to their original state they emit an electromagnetic signal which the scanner can detect. By applying additional gradients to the magnetic field, different slices of the subject can be selected and the location of the emitted radiation can be inferred to build up a map of the density of the nuclear magnets. MRI does not use high-energy ionising radiation so is considered safe for general use, although the equipment remains expensive. MRI shows highest contrast for fatty tissues (such as those just below the skin) and internal soft tissues, but low contrast for hard tissues such as bones.

Volumetric data can be visualised directly using volume-rendering techniques (Drebin *et al.*, 1988; Levoy, 1988), but often requires segmentation for visualisation or processing of specific organs and tissues. Surfaces can be extracted, for example using the marching cubes algorithm (Lorensen and Cline, 1987). Connected areas can be grouped using connected components algorithms and morphological operations (such as erode and dilate) (Drebin *et al.*, 1988). Simple thresholding can be used to extract the skin or bone voxels in CT data.

10.3 Statistical modelling

10.3.1 Shape PCA analysis of points

Once the data have been captured, often the first requirement in medical, forensic or research applications is to analyse the statistics of the captured shapes. Measuring lengths, angles, etc. between manually or automatically identified landmark points can have its uses, but says little about which points have moved or how. An alternative approach is to analyse the global shape variations, using all the available landmarks. The shape variation is usually expressed relative to the average shape of the sample set. Calculating the mean shape itself is not trivial, as first the extrinsic sources of variation due to the global position, orientation and scale of the landmarks need to be removed. The optimal alignment between two shapes is achieved when the sum of squared distances between matched landmarks is minimised. For two shapes this is reasonably straightforward – the centre of mass of the points are aligned, the scale is adjusted so that the root mean square distance from the centre of mass is the same and the orientation can be normalised using a singular value decomposition based technique (Besl and McKay, 1992). When constructing an average the optimal target shape for the alignment is the mean itself. Herein lies the difficulty; to calculate the mean requires alignment to the mean, which hasn't yet been calculated. Various solutions to this problem have been proposed – the simplest is to iterate the process. At each iteration, all the samples are aligned to the current estimate of the mean, which can be initialised as one sample in the set. The mean is recalculated and the process is repeated (Dryden and Mardia, 1998).

Once the mean has been estimated, all the samples can be aligned to the mean. The residual differences between the aligned shapes represent the differences in shape. Typically the next step is to calculate the main axes of variation in shape. Each sample is treated as a point in a high dimensional space by converting the (x,y) or (x,y,z) coordinates of the shape's points into a single long vector. Principal component analysis (PCA) can be used to find a set of orthogonal direction vectors that span the space of the sample points, and give the variance along each direction. This analysis implicitly assumes that the space spanned by the samples is linear, which is in general not true (due the alignment process), but is a reasonable approximation when variations are small. For greater accuracy non-linear versions of PCA are available, such as kernel PCA (Schölkopf *et al.*, 1999) and Principal Geodesic Analysis (PGA) (Fletcher *et al.*, 2004).

PCA of shape vectors can be used to analyse or synthesise instances of the class of shapes. For any given shape the distance the shape lies along each of the principal component directions gives a compact description of the shape. These weightings can be used in further analysis (such as regression, clustering or as input to other machine learning algorithms). A set of weightings can be used to synthesise examples in the set. An example of the use of shape PCA is in the Active Shape Model algorithm for object delineation. After some suitable initialisation the algorithm alternates optimisation of matching edge profiles with a projection into the PCA shape space.

10.3.2 PCA of appearance

In a similar manner to the above, appearance (i.e. surface texture colours) can be processed to find the mean and principal variations (Cootes *et al.*, 2001). Typically the face images are first aligned to the mean shape using image warping. This can be either feature-based (e.g. piece-wise linear interpolation over a triangulation of the feature points, or thin-plate spline-based warping (Bookstein, 1989) or automatically (e.g. using optical flow: Vetter and Poggio, 1997). In the latter case the determination of the mean shape needs to be interleaved with the feature matching in an iterative process. The mean can be calculated directly as the average colour of each pixel independently. In order to calculate the principal components the usual direct method would require finding the Eigenvectors (and Eigenvalues) of a matrix the size of the number of pixels squared. This is usually an extremely large matrix, but a mathematical trick can be used to convert the problem to a much more tractable one. This involves forming the N×N matrix of dot products between image residuals (the image minus the mean) and finding the Eigenvalues and Eigenvectors of that much smaller matrix. The Eigenvalues are the same as the non-zero Eigenvalues of the original matrix, and the Eigenvectors are formed as a weighted sum of the image residuals, using the Eigenvectors to supply the weights.

In the description above the images have not had any explicit normalisation. Normalisation (such as normalising the mean pixel value within each image to zero and the standard deviation of pixels values to 1) will usually project the points into a non-linear subspace. Non-linear versions of PCA (such as PGA or kernel PCA) are more appropriate in these cases.

Analysis of a face image from a PCA model is simply formed by subtracting the mean image and finding the dot product of this residual with each of the principal components. Synthesis is equally simple and is a matter of forming a weighted sum of the components and adding them to the mean. This image can be warped from the mean shape into the shape of a PCA-synthesised shape model to vary the shape. Using separate shape and colour models ignores the correlations between the two. A combined model can be formed by concatenating each shape and colour PCA weight vectors into a single longer vector, with the shape components scaled appropriately to avoid dominance by the colour components (Cootes *et al.*, 2001). A further PCA can be applied to these vectors to produce a combined shape and appearance model. Synthesis is performed on the combined vectors, the separate shape and appearance vectors are extracted from the synthesised vector and used to generate a shape and an appearance image. Finally the image can be warped from the mean shape into the target shape to complete the synthesis.

Applications of appearance models are similar to those of PCA shape models. The PCA parameters can be used as a compact description of the shape in machine learning algorithms, expression analysis etc. or synthesis of face models can be used in computer vision to locate faces by finding the parameters of the face model that minimise the differences with the target face (Cootes *et al.*, 2001). The method has also been extended to use 3D models, in which the lighting and 3D orientation also need to be estimated (Blanz and Vetter, 1999) (Figure 10.1). These have found applications in animation, face recognition and facial ageing synthesis (Hunter and Tiddeman, 2009).

PCA models have a number of limitations. Firstly, they do not separate different sources of information, such as lighting, expression and identity. This means that analysis, such as face recognition, is not robust to variations in lighting, expression etc. Synthesis will also suffer, as it will only be possible (for example) to synthesise an individual under certain lighting or with a particular expression. The second problem is that the warping and blending does not fully align small-scale details, such as wrinkles, hair, stubble etc. To capture these details requires additional processing.

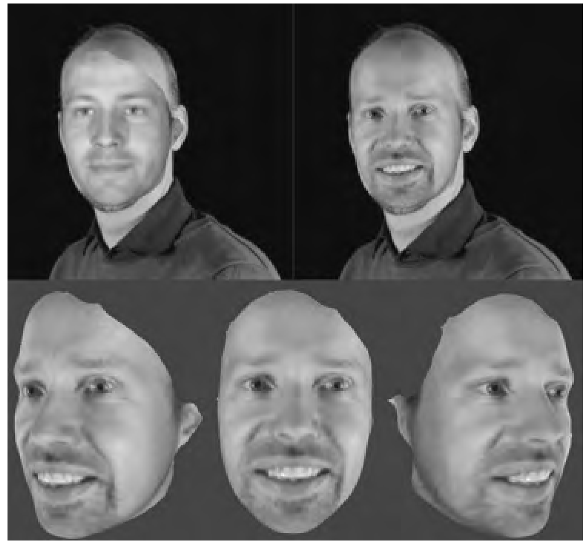


Figure 10.1 An example of fitting a 3D morphable (PCA) model to an image. The target image is shown on the top right and the estimated 3D model is rendered with the estimated pose and lighting overlaid on the target image on the top right. The same model is shown under uniform lighting shown from two angles underneath.

10.3.3 Separating sources of variation

The ability to separate different sources of variation in facial images can improve both face recognition and face synthesis. In the face recognition literature, Fisherfaces have been proposed, which use Fisher's Linear Discriminant Analysis to separate intrinsic and extrinsic sources of variation (Belhumeur *et al.*, 1997). Orthogonal directions are found that maximise the ratio of within-class variance to the between-class variance. Fisherfaces have shown superior performance in face recognition to PCA-based Eigenfaces, being far less sensitive to extrinsic variations such as lighting.

Fisherfaces do not usually assign specific labels to all examples of within-class variation, such as lighting. In some applications it is desirable to be able to assign parameters to a range of different sources of variability independently, e.g. face identity, expression, lighting, age, etc. Multi-linear methods have been developed that allow analysis of a set of data (e.g. faces) along multiple linear axes (Alex *et al.*, 2002). The basis vectors themselves are combined as a linear combination, but the weights associated with each basis vector are non-linear, derived from an outer-product tensor of the weights along each axis (Figure 10.2). These multi-linear models have proved successful in face recognition (Tiddeman *et al.*, 2007), expression synthesis (Macedo *et al.*, 2006) and video-puppetry (Vlasic *et al.*, 2005), as they capture individual variations in the form of expressions.

10.3.4 Texture analysis/synthesis

The second limitation of PCA-type models is their inability to model realistic texture detail. Warping each image and blending the results together reduces

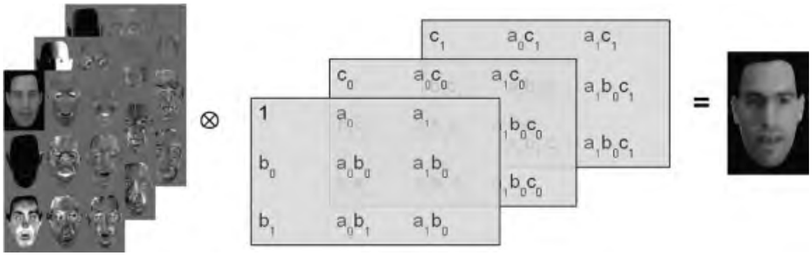


Figure 10.2 Multi-linear models can be used to separate sources of variation in images. In the example above the tensor is learnt from faces varying in identity, expression and lighting. An image can be reconstructed from the outer product of weights for each parameter.

the amplitude of high-frequency detail information, leading to unrealistically smooth appearance of skin, hair, stubble etc. This is particularly noticeable in the average face, but also affects many of the principal components. Several algorithms have been developed for capturing the textural details in facial models, many derived from earlier algorithms for texture synthesis. Amongst the simplest is to amplify the local frequency information of the mean face to match the typical variance of the sample (Tiddeman *et al.*, 2001) (Figure 10.3). As the term 'local' is dependent on image resolution and so forth, the operation is carried out at all scales (along a dyadic sequence) using a resolution appropriate for that image scale. This simple algorithm has also been extended to transforming images (e.g. face ageing). The results are a dramatic improvement in the appearance of 'unstructured' textures (such as stubble and some wrinkles) but it performs less well with more structured textures, such as long hair.

In order to improve the appearance of structured textures, methods have been developed based on the local image statistics in a neighbourhood of each pixel (Tiddeman *et al.*, 2005). Using the assumption that the probability distribution for a particular pixel's intensity is dependent on its neighbours, the pixel probabilities across the image can be maximised to form a more realistic prototype. The optimisation can be performed on a multi-scale basis, starting from the lowest resolution and working towards the highest. Applying this method often leads to locking on to a particular individual at a low resolution and copying large parts of their appearance into the prototype. To make the prototype more like the mean image the smoothing of the probability distribution can be increased at coarser scales. Working in a wavelet space can also improve the appearance of the protoypes and make the distributions less dependent on image intensity and more on the local variations in intensity, which are perceptually more important. These methods have been extended to the transformation of faces, where the processing of textural details (e.g. to add or remove wrinkles or stubble) is even more important visually (Figure 10.4). A similar method has been developed for synthesising faces from a combination of PCA and blending of image appearance using Poisson editing (Mohammed *et al.*, 2009).

Figure 10.3 Averages formed by warping and blending tend of loose fine detail from the textures (left), additional processing such as amplifying the edges at multiple scales (right) can improve the appearance of the textures and so capture age better.

10.4 Parameterisation issues

When constructing shape averages it is important to accurately match corresponding features in different 2D face images or 3D models. A number of points can be reliably located, for example at corners, but many more points are available along edges and across surfaces, where reliable matching is more difficult. In many cases simple heuristics are used, such as parameterising curves by the fraction along their length. In 3D, face surface models can be parameterised by making an approximate alignment using feature points, and then projecting the normals of one surface into the other and using these points to improve the alignment. Such methods can often produce reasonable results in practice, but require careful selection of the

Figure 10.4 Transforming images based on prototypes. Top row gender change, bottom row age change. From left to right, original images, shape and average colour change with no texture processing, amplifying intensity changes across multiple scales and using local neighbourhood information at multiple scales.

initial feature points, which often include pseudo-landmarks that cannot be reliably located.

Automated methods for matching surfaces have received a great deal of attention in computer vision. Often the problem is to match one surface to another. Iterative closest point (ICP)-based methods can often produce good results (Amberg *et al.*, 2007), but the range of deformations needs to be restricted, e.g. by using a model of the allowed deformations, rigidity or smoothness constraints. Settings for rigidity of the models or smoothness of the mapping necessarily require some heuristics to set the values appropriately. Use of models such as linear PCA to constrain the matching requires a set of matched examples to build the PCA model, so is appropriate for automation (to reduce manual input) but not for improving the quality of matching.

In the field of morphometrics, sliding semi-landmarks have been proposed (Green, 1996; Bookstein, 1997) which allow points to slide along curves or across surfaces under certain constraints. The constraints can include allowing the points to slide tangentially to the curve or surface whilst minimising the bending energy of the warp between the two samples (this favours smooth deformations), or the minimising of Procrustes distance between the samples (Sheets *et al.*, 2004), ignoring any component of the displacement tangential to the reference curve or surface (favours proximity). The sliding semi-landmark approach follows our intuition that shape variations between two samples should be smooth and nearby points should be matched, but it does have some drawbacks. Firstly, the metrics used (e.g. bending energy or perpendicular Procrustes distance) are not symmetrical, so matching A to B will typically give different results from matching B to A (although the measure can be performed in both directions and summed to produce a symmetric measure). Secondly, the choice of error metric can produce different results (Perez *et al.*, 2006). Thirdly, care needs to be taken when extending to groupwise alignment as described below.

Groupwise alignment aims to automatically match a set of 2D or 3D shapes or images to each other while simultaneously building a model of the main shape deformations. The aim is to produce a model that is specific to the class of objects but that also generalises well to other examples of the same class of object. The sliding semi-landmarks technique can be extended to allow iterative adaptation to the mean, minimising the warp-bending energy or the perpendicular Procrustes distances of all samples from the current estimate of the mean, rebuilding the model and repeating. This of course still suffers from the problems described above, and does not guarantee to produce the most compact and specific model for the class under study. To automatically produce a model that does satisfy these conditions the problem is reformulated as that of transmitting the model and all the examples in the most compact way possible, which leads to the minimum description length (MDL) approach (Davies *et al.*, 2008*a*). In order to formulate the MDL of a set of matched models, the quantisation of the model, the model parameters and the errors of each example from

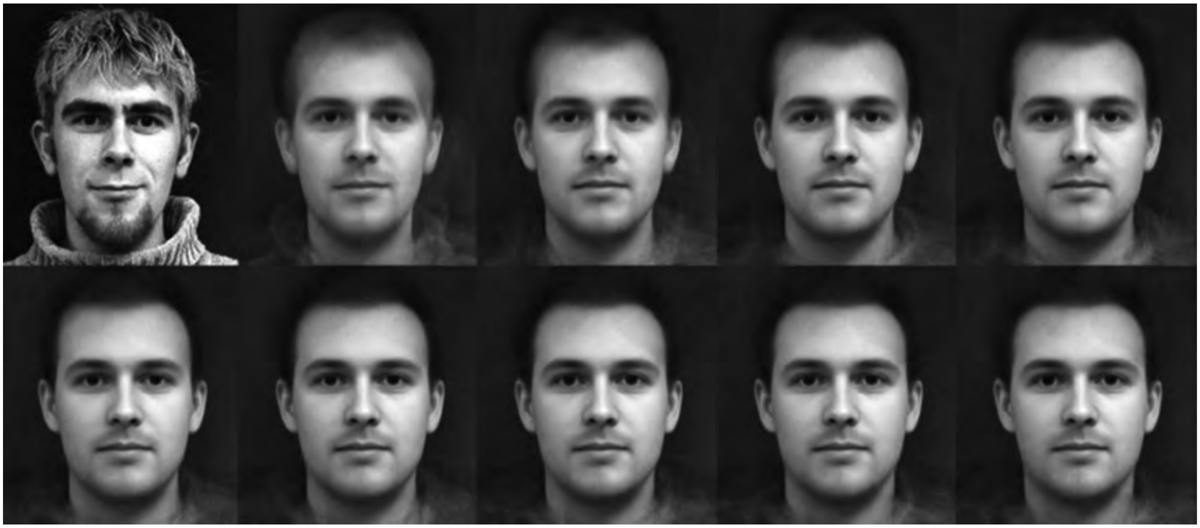


Figure 10.5 An example of groupwise registration. A set of male faces are automatically aligned to the current estimate of the average built using a leave one out strategy. The average is initialised using a single example (top left) and iterations increase from left to right and top to bottom.

the model description need to be considered. This leads to a mathematical formulation of the description length for a given set of parameterisations (i.e. the sliding semi-landmarks describing the surfaces or contours), which can be optimised iteratively. Regularisation can help improve the speed of the optimisation whilst converging to the same minima (Davies *et al.*, 2008*b*). Often simplifications of the full MDL are used for groupwise alignment (Figure 10.5). For example, a leave-one-out strategy, where the model is built from N-1 samples, the parameterisation of the Nth model is optimised, then the model is inserted back into the set and the process is repeated with a different model (Sidorov *et al.*, 2009).

10.4.1 Physical modelling

The modelling of facial changes, such as expression, the effects of surgery or growth, can be performed by statistical/geometrical methods as described in the previous sections, provided sufficient training data are available. An alternative is to model the physical behaviour of tissues in the face directly. Terzopoulos and Waters' early work performed facial animation using a three-layer model of the skin, including elasticity, volume preservation and damping terms (Terzopoulos and Waters, 1990). To animate the model 'muscles' were added that applied forces to various parts of the face model. The model was animated using finite difference methods to numerically integrate the solution through time. Similar methods have been used for prediction of the outcomes of facial surgery (Keeve *et al.*, 1996). These have been extended to use finite element methods (Larrabee *et al.*, 1986; Pieper *et al.*, 1995; Koch *et al.*, 1996) of the facial soft tissues. The finite element method divides the study volume into discrete elements and models the behaviour within each element while ensuring continuity across element boundaries. More recent work in this area has focused on, for example, more accurate representation of the facial anatomy (Barbarino *et al.*, 2008), better integration with other treatment planning tools (Zachow *et al.*, 2006) and real-time interaction (Berkley *et al.*, 2004).

Finite element methods can also be used to model the stresses and strains in facial hard tissues and provide evidence for the changes in facial structure through growth and evolution. Current theories of bone growth and (re)modelling suggest that bone deposition is induced by mechanical stimulation (e.g. high stress/strain) and bone resorption by low mechanical stimulation. Stress-strain loading has been shown to have played a role in the development of various facial structures in humans and other species, such as the locations of sutures in lizards (Moazen *et al.*, 2009), the supra-orbital torus in macaques (Kupczik *et al.*, 2009) and development of the human jaw (Gröning *et al.*, 2009).

10.5 Summary and conclusions

This chapter provides a broad overview of facial modelling using computer graphics techniques. When selecting facial modelling, analysis and synthesis techniques for a particular application consideration needs to be given to the availability of data and the appropriateness of the particular computational tools. The choice of data capture will depend on availability, cost and quality of the data available. A range of statistical tools for the analysis of shape have been described, from simple landmark analysis through to automatic groupwise modelling and analysis of surfaces and volumes. For facial synthesis and animation, facial models built from statistical data have proved effective, provided care is taken in the treatment of textural detail and the modelling of different sources of variation (such as identity and expression). Finally the main alternatives to statistical modelling, i.e. physical modelling using finite difference or finite element methods, with applications in facial surgery simulation and analysis of growth and evolutionary development were briefly discussed.

References

Alex, M., Vasilescu, O. and Terzopoulos, D. (2002). Multilinear analysis of image ensembles: TensorFaces. In *Proceedings of the European Conference on Computer Vision* (ECCV). *Lecture Notes in Computer Science*, **2350**, 447–460.

Amberg, B., Romdhani, S. and Vetter, T. (2007). Optimal step nonrigid ICP algorithms for surface registration. *Computer Vision and Pattern Recognition* (CVPR), **2007**, 1–8.

Barbarino, G., Jabareen, M., Trzewik, J. and Mazza, E. (2008). Physically based finite element model of the face. *Biomedical Simulation, Lecture Notes in Computer Science*, **5104**, 1–10.

Belhumeur, P. N., Hespanda, J. and Kriegeman, D. (1997). Eigenfaces vs. Fisherfaces: recognition using class specific linear projection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **19**(7), 711–720.

Berkley, J., Turkiyyah, G., Berg, D., Ganter, M. and Weghorst, S. (2004). Real-time finite element modeling for surgery simulation: an application to virtual suturing. *IEEE Transactions on Visualization and Computer Graphics*, **10**(3), 314–325.

Besl, P. J. (1988). Active, optical range imaging sensors. *Machine Vision and Applications*, **1**, 127–152.

Besl, P. J., and McKay, N. D. (1992). A method for registration of 3-D shapes. *IEEE Transactions Pattern Analysis and Machine Intelligence*, **14**(2), 239–256.

Blanz, V. and Vetter, T. (1999). A morphable model for the synthesis of 3D Faces. SIGGRAPH'99 Conference Proceedings, pp. 187–194.

Bookstein, F. L. (1989). Principal warps: thin plate splines and the decomposition of deformations. *IEEE Transactions Pattern Analysis and Machine Intelligence,* **11**, 567–585.

Bookstein, F. L. (1997). Landmark methods for forms without landmarks: localizing group differences in outline shape. *Medical Image Analysis*, **1**, 225–243.

Clark, J., Trucco, E. and Wolff, L. B. (1997). Using light polarization in laser-scanning. *Image and Vision Computing*, **15** (2), 107–117.

Cootes, T. F., Edwards, G. J. and Taylor, C. J. (2001). Active appearance models. *IEEE Transactions Pattern Analysis and Machine Intelligence*, **23**(6), 681–685.

Damadian, R., Goldsmith, M. and Minkoff, L. (1977). NMR in cancer: XVI. Fonar image of the live human body. *Physiological Chemistry and Physics*, **9**, 97–100.

Davies, R., Twinning, C. and Taylor, C. (2008*a*). *Statistical Models of Shape – Optimisation and Evaluation*. Berlin: Springer.

Davies, R. H., Twining, C. J. and Taylor, C. (2008*b*). Groupwise surface correspondence by optimization: Representation and regularization. *Medical Image Analysis*, **12**, 787–796.

Drebin, R. A., Carpenter, L. and Hanrahan, P. (1988). Volume rendering. *Computer Graphics*, SIGGRAPH88.

Dryden, I. L., and Mardia, K. V. (1998). *Statistical Shape Analysis*. Chichester: John Wiley and Sons.

Fechteler, P. and Eisert, P. (2008). Adaptive color classification for structured light systems. *Proceedings of Computer Vision and Pattern Recognition* (CVPR) 2008.

Fletcher, P. T., Conglin L, Pizer, S. M. and Sarang, J. (2004). Principal geodesic analysis for the study of nonlinear statistics of shape. *IEEE Transactions on Medical Imaging*, **23**(8), 995–1005.

Green, W. D. K. (1996). The thin-plate spline and images with curving features. In K. V. Mardia, C. A. Gill and I. L. Dryden (Eds.), *Image Fusion and Shape Variability*, pp. 79–87. Leeds: University of Leeds Press.

Grimson, W. E. L. (1985). Computational experiments with a feature based stereo algorithm. *IEEE Transactions Pattern Analysis and Machine Intelligence*, **7**(1), 17–34.

Gröning, F., Liu, J., Fagan, M. J. and O'Higgins, P. (2009). Validating a voxel-based finite element model of a human mandible using digital speckle pattern interferometry. *Journal of Biomechanics*, **42**, 1224–1229.

Hinshaw, D. S., Bottomley, P. A. and Holland G. N. (1977). Radiographic thin-section image of the human wrist by nuclear magnetic resonance. *Nature*, **270**, 722–723.

Hunter, D. W., and Tiddeman, B. P. (2009). Visual facial ageing using PLS: Visual ageing of human faces in three dimensions using Morphable Models and projection to latent structures. *International Conference on Computer Vision Theory and Applications* (VISAPP).

Jarvis, R. A. (1983). A perspective on range finding techniques for computer vision. *IEEE Transactions Pattern Analysis and Machine Intelligence PAMI-5*, **2**, 122–139.

Kanade, T. (1994). A stereo matching algorithm with an adaptive window: Theory and experiment. *IEEE Transactions Pattern Analysis and Machine Intelligence*, **16**(9).

Keeve, E., Girod, S. and Girod, B. (1996). Craniofacial surgery simulation. *Proceedings Visualization in Biomedical Computing, Lecture Notes in Computer Science*, **1131**, 541–546.

Koch, R. M., Gross, M. H., Carls, F. R. *et al.* (1996). Simulating facial surgery using finite element models. Proc. SIGGRAPH '96, pp. 421–428.

Kupczik, K., Dobson, C. A., Crompton, R. H. *et al.* (2009). Masticatory loading and bone adaptation in the supraorbital torus of developing macaques. *American Journal of Physical Anthropology*, **139**, 193–203.

Larrabee, W. F.Jr. and Sutton, D. (1986). A finite element model of skin deformation. II. An experimental model of skin deformation. *Laryngoscope*, **96**(4), 406–412.

Lauterbur, P. C. (1974). Magnetic resonance zeugmatography. *Pure and Applied Chemistry*, **40**, 149–157.

Levoy, M. (1988). Display of surfaces from volume data. *IEEE Computer Graphics and Applications*, **8**(3), 29–37.

Lorensen, W. E., and Cline, H. E. (1987). Marching cubes: a high resolution 3D surface construction algorithm. *Computer Graphics*, **21**(4), 163–169.

Macedo, I., Brazil, E. V. and Velho, L. (2006). Expression transfer between photographs through multilinear AAM's. *Computer Graphic and Image Processing*, SIBGRAPI-**06**, 239–246.

Moazen, M., Curtis, N., O'Higgins, P. *et al.* (2009). Assessment of the role of sutures in a lizard skull – a computer modelling study. *Proceedings of the Royal Society, Part B*, **276**, 39–46.

Mohammed, U., Prince, S. J. D. and Kautz, J. (2009). Visio-lization: generating novel facial images. *ACM Transactions on Graphics* (*Proceedings SIGGRAPH 2009*), **28**(3), 57:1–57:8.

Mozzo, P., Procacci, C., Tacconi, A., Tinazzi Martini P. and Bergamo Andreis I. A. (1998). A new volumetric CT machine for dental imaging based on the cone-beam technique: preliminary results. *European Radiology*, **8**(9).

Perez, S. I., Bernal, V. and Gonzalez, P. N. (2006). Differences between sliding semi-landmark methods in geometric morphometrics, with an application to human craniofacial and dental variation. *Journal of Anatomy*, **208**(6), 769–784.

Pieper, S., Laub, D. and Rosen, J. (1995). A finite element facial model for simulating plastic surgery. *Plastic Reconstructive Surgery*, **96**(5), 1100–1105.

Schölkopf, B., Smola, A. and Müller, K. R. (1999). Kernel principal component analysis. In B. Schölkopf, C. J. C. Burges and A. J. Smola (Eds.), *Advances in Kernel Methods – Support Vector Learning*, pp. 327–352. Cambridge, MA: MIT Press.

Sheets, H. D., Keonho, K. and Mitchell, C. E. (2004). A combined landmark and outline-based approach to ontogenetic shape change in the Ordovician Trilobite *Triarthrus becki*. In: A. Elewa A (Ed.), *Applications of Morphometrics in Paleontology and Biology*, pp. 67–81. New York: Springer.

Shrimpton, P. C., Miller, H. C., Lewis, M. A. and Dunn, M. (2003). Doses from Computed Tomography (CT) examinations in the UK – 2003 Review. National Radiological Protection Board.

Sidorov, K., Richmond, S. and Marshall, D. (2009). An efficient stochastic approach to groupwise non-rigid image registration. *Proceedings of Computer Vision and Pattern Recognition* (CVPR), **2009**, 2208–2213.

Srinivasan, V., Liu, H. C. and Halioua, M. (1984). Automated phase-measuring profilometry of 3-D diffuse objects. *Applied Optics*, **23**, 3105–3108.

Terzopoulos, D. and Waters, K. (1990). Physically-based facial modeling, analysis, and animation. *Journal of Visualization and Computer Animation*, **1**(4), 73–80.

Tiddeman, B., Burt, D. M. and Perrett, D. (2001). Computer graphics in facial perception research. *IEEE Computer Graphics and Applications*, **21**(5), 42–50.

Tiddeman, B., Hunter, D. and Yu, M. (2007) Fibre centred tensor faces. *British Machine Vision Conference*, pp. 449–458.

Tiddeman, B., Stirrat, M. and Perrett, D. (2005). Towards realism in facial transformation: results of a wavelet MRF method. *Computer Graphics Forums*, **24**(1–5).

Udupa, J. K., and Herman, G. T. (2000). *3D Imaging in Medicine*, 2nd edn. Boca Raton, FL: CRC Press.

Vetter, T. and Poggio, T. (1997). Linear object classes and image synthesis from a single example image. *IEEE*

IEEE Transactions Pattern Analysis and Machine Intelligence, **19**(7), 733–742.

Vlasic, D., Brand, M., Pfister, H. and Popović, J. (2005). Face transfer with multilinear models. *ACM Transactions on Graphics* (*TOG*), **24**(3), 426–433.

Zachow, S., Hege1, H-C. and Deuflhard, P. (2006). Computer-assisted planning in cranio-maxillofacial surgery. *Journal of Computing and Information Technology – CIT*, *14*(**1**), 53–64.

Chapter 11

Recognising and learning faces in motion

Karen Lander and Natalie Butcher

11.1 Introduction

This chapter reviews the role of motion when recognising already familiar faces and when learning new faces. This issue is interesting since, historically, face perception research has utilised static images of faces, with little consideration for the potential role of dynamic information created by seeing a face in motion. Much recent research has suggested that dynamic information contributes important additional information both when recognising familiar faces and when learning new ones. In this chapter we also introduce some of our own recent research which aims to investigate the importance of seeing own-race and other-race faces moving at encoding and testing. The practical and theoretical implications of our results are discussed.

Face recognition in humans is a highly developed visual perceptual skill. Face processing is performed by a distributed cognitive system likely to have evolved in the very distant past in evolutionary history (Pascalis and Kelly, 2009). Faces are naturally intricate structures composed of multiple complex features (e.g. eyes), which are themselves constructed from multiple lower-level features (e.g. contrast, frequency, orientation), located and orientated according to a unique configuration (Peterson *et al.*, 2009). Despite this, in everyday life, the task of recognising and identifying individuals from their face is undertaken with relative ease and with little apparent effort (Christie and Bruce, 1998). We recognise faces from all directions under many different viewing conditions (Christie and Bruce, 1998). Variations in lighting and viewpoint, among other non-optimal viewing conditions, are encountered in everyday life yet recognition remains highly accurate (Hill *et al.*, 1997; Braje *et al.*, 1998). We are also adept in identifying particular characteristics from unfamiliar faces, such as their age, sex, race and emotional state with incredible accuracy (McGraw *et al.*, 1989; Montepare and Zebrowitz, 1998). Accordingly, psychologists have long been interested in when face perception is optimal, and in particular have been keen to understand the cognitive processes that occur during face perception.

One area of research that has been investigated by cognitive psychologists is that of the role of motion in recognising already familiar faces (Christie and Bruce, 1998) and learning new faces (Knight and Johnston, 1997). This research is important for both practical and theoretical reasons. Practically, determining the importance of motion for face recognition and learning is interesting for forensic applications. For example, we can address whether CCTV footage is best viewed in motion or static, in order to determine the identity of the suspect shown (Bruce *et al.*, 2001). Further practical applications include moving face identity parades (VIPER – Video Identification Parade Electronic Recording; Kemp *et al.*, 2001) and security systems that are used to verify identity (for example, Vetter *et al.*, 1997). Motion is also important in the creation of realistic animated avatars and other gaming technologies (Zhang *et al.*, 2004). Theoretically, determining the role of motion should help us understand how motion is linked to identity processing and other face-perception tasks (for example, expression, visual speech processing). Here, several models have aimed to map the cognitive and neural processing of both changeable and non-changeable aspects of the face (see Haxby *et al.*, 2002; O'Toole *et al.*, 2002).

Specifically, in this chapter we describe research investigating the recognition advantage for moving faces. This research utilises faces already known – either personally familiar or famous. At this point it

Craniofacial Identification, ed. C. Wilkinson and C. Rynn. Published by Cambridge University Press.

is important to make a distinction between rigid and non-rigid facial motion. By rigid motion, we refer to movement of the whole head, for example when nodding or shaking. In contrast non-rigid motion refers to movement of the facial features themselves, as shown when talking or expressing. The majority of this chapter is spent describing the impact of motion when learning new faces. Here we describe research recently conducted in our laboratory, which was designed to investigate the importance of motion when learning new faces from the same or another race.

11.2 Recognising familiar faces

Early findings by Bruce and Valentine (1988), using point-light displays (Johansson, 1973), suggested that isolated dynamic information may act as a cue to identity when static cues are diminished. More convincing evidence that movement is important in the recognition of individual faces comes from Knight and Johnston (1997). Knight and Johnston (1997) presented famous faces either in a negative (contrast-reversed) format or upside down, and compared recognition performance from moving and static sequences. Results indicated that moving famous faces were recognised significantly more accurately than static ones, but only when these were shown as upright-negative images. Knight and Johnston (1997) proposed that seeing the face move may provide evidence about its three-dimensional structure, compensating for the degraded depth cues available within a negative image (Bruce and Langton, 1994). Alternatively, Knight and Johnston (1997) suggest that known faces may have characteristic facial gestures, idiosyncratic to the individual viewed.

Follow-up research (see Lander *et al.*, 1999, 2001) showed that the recognition advantage for moving faces is not specific to upright-negated images. Instead motion confers benefits through a range of image manipulations, including thresholding (one-bit per pixel black-and-white images), pixelation and Gaussian blurring. Here what seems important for the demonstration of a motion recognition advantage is not the nature of the image manipulation but that recognition performance is below ceiling, allowing higher recognition rates to be found. In these experiments the moving sequences show famous faces, pictured from the shoulders up, talking and expressing. Thus, these experiments clearly demonstrate that non-rigid movement adds useful information when recognising the identity of famous faces shown in difficult viewing conditions. It is important to note at this point that the motion recognition advantage is not solely due to the increased number of static images shown when the face is in motion. Indeed when the number of images was equated across static and moving presentation conditions (Lander and Bruce, 2000) there was still an advantage for viewing the face in motion.

Furthermore, the viewing conditions in which facial motion is observed have been shown to affect the extent to which motion aids recognition. Research indicates that disruptions to the natural movement of the face can influence the size of the motion advantage in facial recognition. Lander and colleagues (Lander *et al.*, 1999, Lander and Bruce, 2000) found lower recognition rates of famous faces when the motion was slowed down, speeded up, reversed or rhythmically disrupted. Thus, seeing the face move naturally provides the greatest advantage for facial recognition. A further demonstration of this point comes from research using both natural and artificially created smiling stimuli (Lander *et al.*, 2006). To create an artificially moving sequence, Lander *et al.* (2006) used a morphing technique to create intermediate face images between the first and last frames of a natural smile. When shown in sequence, these images were used to create an artificially moving smile that lasted the same amount of time, and had the same start and end point as the natural smile for that individual. Results found that familiar faces were recognised significantly better when shown naturally smiling compared with a static neutral face, a static smiling face or a morphed smiling sequence. This further demonstrates the necessity for motion to be natural in order to facilitate the recognition advantage.

Lander *et al.* (2006), whilst investigating the effects of natural versus morphed motion, also found a main effect of familiarity, revealing that the nature of the to-be-recognised face can influence what effect motion has on facial recognition. It was suggested that the more familiar a person's face is, the more we may be able to utilise the movement of their face as a cue to identity. Indeed, characteristic motion patterns may become a more accessible cue to identity as a face becomes increasingly familiar. In very recent work, Butcher (2009) found a significant positive correlation between rated face familiarity and the size of the recognition advantage for moving compared with static faces. This research was conducted using famous faces and found that the more familiar the famous face was rated to be, the larger the recognition advantage for viewing that face in motion.

Another factor of the to-be-recognised face that has been shown to be important in understanding what mediates the motion advantage is distinctiveness. Facial recognition research has demonstrated a clear benefit for faces that are thought to be spatially distinctive, as findings indicate that distinctive faces are better recognised than faces that are rated as being ‘typical’ (Light *et al.*, 1979; Valentine and Bruce, 1986; Valentine, 1991; Valentine and Ferrara, 1991; Vokey and Read, 1992). It has also been established that a larger motion recognition advantage is attained from distinctive motion than typical motion (Lander and Chuang, 2005; Butcher, 2009). Lander and Chuang (2005) found that motion was a more useful cue to identity for ‘distinctive movers’ compared with more ‘typical movers’ (also see Butcher, 2009). It is important here to consider that distinctiveness in facial motion may refer to:

(1) A motion characteristic or typical of a particular individual.
(2) An odd motion for a particular individual to produce.
(3) A generally odd or unusual motion.

Also, it may be that spatial and temporal distinctiveness of faces are in some way related. For instance, spatially distinctive faces might naturally have more distinctive movements, a notion that should itself be addressed in future research.

Finally, in terms of familiar face recognition and motion, studies with prosopagnosic patients have shown a number of interesting findings. For example, a possible dissociation has been revealed between the ability to recognise a face from the motion it produces, and the ability to recognise it from a static image. In this work, Steede *et al.* (2007) reported the case study of a developmental prosopagnosic patient, C.S. Despite C.S. being impaired at recognising static faces, he was able to effectively discriminate between different dynamic identities, and demonstrated the ability to learn the names of individuals on the basis of their facial movement information (at levels equivalent to both matched and undergraduate controls). This case study indicates a possible dissociation between the cognitive mechanisms involved in the processes of recognising a static face and those involved in recognising a dynamic face. This research is supported by neuroimaging studies that have demonstrated functional separation of motion and structural aspects of face perception in humans (Haxby *et al.*, 2002). Haxby *et al.* (2002) found that facial movements activated the superior temporal sulcus (STS) area while the more shape-based aspects of the face activated the fusiform gyrus. Further neuropsychological studies have revealed differential activations for processing motion and static face information (Schultz and Pilz, 2009).

Previous attempts to determine whether an individual who is impaired at static face recognition can use facial motion as a signal to identity have revealed conflicting findings (Lander *et al.*, 2004). Lander and her colleagues (2004) tested a prosopagnosic patient, H.J.A. on his ability to use motion as a cue to facial identity. In Experiment 1, H.J.A. attempted to recognise the identity of dynamic and static famous faces. H.J.A. was found to be severely impaired in his ability to recognise identity, and was not significantly better at recognising moving faces compared with static ones. In order to test H.J.A.’s ability to learn face–name pairings a second experiment was conducted using an implicit face-recognition task. In this experiment H.J.A. was asked to try to learn true and untrue names for famous faces, which were shown in either a moving clip or a static image. H.J.A. found this a difficult task and was no better with moving faces or true face–name pairings. Some prosopagnosic patients have previously found it easier to learn true face–name pairings more accurately and efficiently than untrue ones (covert recognition; De Haan *et al.*, 1987). From this case study Lander *et al.* (2004) concluded that H.J.A. was not able to overtly or covertly use motion as a cue to facial identity. Despite this, a third experiment demonstrated that H.J.A. was able to decide whether two sequentially presented dynamic unfamiliar faces had the same or differing identities. His good performance on this matching task demonstrates that H.J.A. retains good enough motion-processing abilities to enable him to match dynamic facial signatures, yet insufficient abilities to allow him to store, recognise or learn facial identity on the basis of facial movements. Thus, any possible cognitive dissociation between the processing of facial identity from a moving face compared with a static one requires further investigation and discussion (see Lander *et al.*, 2004 for elaboration and implications for models of face recognition).

Having clearly demonstrated a moving recognition advantage it is important to investigate the theoretical basis of this effect. Two theories have been proposed by O'Toole *et al.* (2002), namely the representation enhancement hypothesis and the supplemental information hypothesis. The representation enhancement hypothesis (O'Toole *et al.*, 2002) suggests that facial motion aids recognition by facilitating the perception of the three-dimensional structure of the face. It posits that the quality of the structural information available from a human face is enhanced by facial motion, and this benefit surpasses the benefit provided by merely seeing the face from many static viewpoints (Pike *et al.*, 1997; Christie and Bruce, 1998; Lander *et al.*, 1999). As this mechanism is not dependent on any previous experience with an individual face it seems to predict that motion should aid recognition of previously unfamiliar faces, a notion discussed next in this chapter.

In contrast, the supplemental information hypothesis (O'Toole *et al.*, 2002) assumes that we represent the characteristic facial motions of an individual's face as part of our stored facial representation for that individual. For the particular individual's characteristic facial motions to be learned, experience with that face is needed – the face must be familiar. 'Characteristic motion signatures' are learned over time, allowing a memory of what facial motions a person typically exhibits to be stored as part of their facial representation. Therefore, when motion information for an individual has been integrated into the representation of their face, this information can be retrieved and used to aid recognition of that face. Support for the idea that facial motion becomes intrinsic to a familiar individual's face representation comes from a study conducted by Knappmeyer *et al.* (2003). This study used two synthetic heads, each animated by the movement of a different volunteer. Participants viewed and thus became familiar with either head A with motion from volunteer A, or head B with motion from volunteer B. In the test phase an animated head constructed from the morph of the two synthetic heads (A and B) was produced. Participants were asked to identify whose head was shown. It was found that participants' identity judgements were biased by the motion they had originally learnt from head A or B, demonstrating that representations of an individual's characteristic facial motions are learned and are inherent to that individual's face representation (also see Lander and Bruce, 2004).

11.3 Learning unfamiliar faces

In contrast to these beneficial effects on the identification of familiar faces, the effect of motion on the learning of unfamiliar faces is much less clear. Early work conducted by Christie and Bruce (1998) found no advantage for either rigid or non-rigid motion in face learning. In this incidental learning task, participants were shown faces either as a moving computer-animated display or as a series of static images, and were asked to decide whether they thought each person shown studied arts or science subjects at university. The number of frames in the moving and static conditions was equated in the learning phase. The motion involved was either non-rigid (expression changes) or rigid (head nodding or shaking). At test, participants saw either moving sequences or a single static image (the number of frames was not equated at test) and were asked which were the faces of people seen earlier in the arts/science task. Results indicated there was no benefit for studying faces in motion on the subsequent recognition task. However, there was a slight benefit for testing faces in (rigid) motion, compared with static images. In line with this finding, Schiff *et al.* (1986) found an advantage for testing recognition memory for unfamiliar faces using a moving sequence rather than a static 'mugshot' photograph. These findings can be compared with those using familiar faces (Knight and Johnston, 1997; Lander *et al.*, 1999, 2001) who also found a beneficial effect of motion at test.

Despite this early work a number of studies have found an advantage for learning faces from moving sequences. For example, Pike *et al.* (1997) filmed actors rotating on a motorised chair, which was illuminated from a single light source. In the learning phase, participants were asked to try to learn the identity of previously unfamiliar faces from either dynamic sequences (10 second clip), multiple static images (5 images selected from the moving sequence, each presented for 2 seconds) or a single static image (single image presented for 10 seconds). The dynamic sequence showed the target initially facing the video camera, and then undergoing a full 360-degree rotation of the chair. At test, participants viewed a single (full-face) static image, different from any shown in the learning phase. They were asked to decide if the face shown had been present in the earlier learning phase. Results indicated that there was a significant advantage for faces learned via a coherent moving sequence. Bruce and Valentine (1988) reported a similar trend in an experiment which compared learning from video sequences of the target faces speaking, nodding etc. to sequences of single static images. Again, test images were single images taken from a different viewpoint, on a different occasion. In this experiment performance was best when the faces were learned via a moving sequence, although the difference between the moving and static condition failed to reach significance. The failure to reach significance was explained by the authors in terms of the variability of performance, for this task, across the participant population.

In later follow-up work, Lander and Bruce (2003) conducted four experiments that aimed to investigate the usefulness of rigid (head nodding, shaking) and non-rigid (talking, expressions) motion for establishing new face representations of previously unfamiliar faces. Results showed that viewing a face in motion leads to more accurate face learning, compared with viewing a single static image (Experiment 1). The advantage for viewing the face moving rigidly seemed to be due to the different angles of view contained in these sequences (Experiment 2). However, the advantage for non-rigid motion was not simply due to multiple images (Experiment 3) and was not specifically linked to forwards motion but also extended to reversed sequences (Experiment 4). Thus, although there seem to be clear beneficial effects of motion for face learning, they do not seem to be due to the specific dynamic properties of the sequences shown. Instead, the advantage for non-rigid motion may reflect increased attention to faces moving in a socially important manner.

Finally, in very recent work, Butcher *et al.* (2011) used a series of Yes/No recognition tasks, employing both same- and other-race face stimuli. Here, we aimed to further investigate the effect of motion on facial recognition and the 'other-race effect', which demonstrates that other-race faces are recognised less reliably than faces of the same race as the viewer (Malpass and Kravitz, 1969; Lindsay *et al.*, 1991). Specifically, we asked whether facial motion cues, provided by non-rigid movements, facilitate learning and recognition of other-race faces as they have been shown to do for same-race faces (see Lander and Bruce, 2003).

Importantly, the type of facial motions exhibited across different race faces has been revealed to vary (Tzou *et al.*, 2005). For instance, Tzou *et al.* (2005) found that in general Europeans exhibit larger facial movements than Asians. In particular, Europeans display significantly larger movements in the eyebrow, nose and mouth regions, although Asians were revealed to display larger movement of the eyelids. In explanation of the other race effect it has been posited that other race observers lack experience with the dimensions appropriate to code other-race faces efficiently, which leads to difficulty when differentiating between different exemplars of other-race faces (Chance *et al.*, 1982; Bothwell *et al.*, 1989). Therefore, it may be logical to reason that other race observers may lack the expertise with which to process individuating motion information exhibited by other-race faces efficiently enough for it to provide a benefit at recognition. However, familiarisation with other-race faces has been demonstrated to reduce the 'other-race effect' (McKone *et al.*, 2007). It may be that viewing moving stimuli enables the creation of more robust, descriptive face representations than static images. Accordingly, more descriptive representations, constructed with motion information, may lead to a greater sense of familiarity with target other-race faces. This increased familiarity, as a product of the availability of motion information, may lead to a moving face recognition advantage for other-race faces as is found for same-race faces (see Lander and Bruce, 2003).

So, we ask, can observers use the facial motion of other-race faces stored in facial representations to aid recognition decisions (O'Toole *et al.*, 2002; Lander and Bruce, 2003) in the same manner as they do during the recognition of same-race faces? Next we present two experiments designed to investigate this issue. In the first experiment, 60 British White participants were asked to learn British or Japanese faces shown moving or static (between-participants design). One dependent variable was measured, that of recognition accuracy with hit rates and false alarm rates considered. The design consisted of two sections, a learning phase and a recognition phase. In the learning phase participants were asked to watch short clips (2 seconds; silent) of 20 unfamiliar faces. The faces were presented either in motion or static. The faces used were selected from a bank of colour video sequences of British White and Japanese faces talking that had previously been used for facial recognition experiments (Lander *et al.*, 2007). All the sequences displayed at least the head and shoulders of the subject and were all shot from a frontal position. During the moving sequences the participant saw the target face speaking while looking directly at the camera. For the static condition, a single freeze frame was selected from the original video sequence at a stage in the video when the face was considered to be displaying a neutral expression. The same bank of video sequences was used to create the target and distracter stimuli displayed in the test phase, which were all static images. In the static learning conditions the static images used in the test phase were different to the ones used in the learning phase in order to eliminate picture-matching strategies being adopted to complete the task (e.g. Baddeley and Woodhead, 1983).

At the start of the learning phase, participants were informed that they would be shown 20 faces one at a time and that they were to watch the faces and try to remember them. In the recognition test phase, participants were shown a total of 40 faces presented sequentially one at a time as single static images, half of which were the faces shown during the learning phase and the other half were distracter faces of the same race as the learned faces. As each face was displayed the participant was asked to decide, as accurately as possible, whether they had seen the face before, whether they recognized it or not. Results from this experiment are presented in Table 11.1.

Statistical analysis of the results revealed that there was a significant main effect of race. Own-race faces were recognised significantly better than other-race faces, supporting the existence of the other-race effect (Malpass and Kravitz, 1969). The cause of the other-race effect during facial recognition has long been debated by researchers. Many theoretical explanations have been posited including the contact hypothesis (Chance *et al.*, 1982; Bothwell *et al.*, 1989) and group-processing explanations (Anthony *et al.*, 1992; Chance and Goldstein, 1996; Sporer, 2001). Despite a wealth of research into the other-race effect, firm conclusions as to why it occurs are yet to be reached, with recent accounts considering a culmination of several factors to be at the root of this effect. The results found here provide further support for previous demonstrations of the other-race effect as a robust phenomenon across various races (Brigham and Malpass, 1985; Chance and Goldstein, 1996; Meissner and Brigham, 2001; Sporer, 2001). Importantly, these results, to our knowledge, offer the first empirical evidence of an other-race effect using moving stimuli. This, in itself, is a finding that is of interest to broaden knowledge of the nature of the other-race effect.

Table 11.1 Mean hits (%) and false alarms (FAs; %) with standard deviations (SD) in each of the conditions of the experiment; 15 scores contributed to each mean.

Race of face	Same		Other	
Learning Style	Static	Moving	Static	Moving
Hits	75.0	86.3	69.3	80.0
SD	8.9	9.0	14.0	9.7
FAs	15.0	8.4	18.0	14.0
SD	12.4	7.7	13.4	12.2

Results from Experiment 1 also revealed that faces presented moving resulted in significantly better learning compared with faces presented static. Thus, we have supported our previous finding that motion information aids facial recognition at encoding when learning new faces (Pike *et al.*, 1997; Lander and Bruce, 2003; Lander and Davies, 2007). Clearly, motion information is in some manner integrated into face representations (also see Freyd, 1993) from an early stage during encoding and can be used at recognition by providing robust mental representations of dynamic faces (Pilz *et al.*, 2009). In addition, it is important to note that this beneficial effect of learning moving faces was found for both own-race and other-race faces. This result suggests that observers may be processing other-race facial motion in the same manner as they do same race faces, despite many accounts of the other-race effect proposing that a lack of experience with the dimensions of other race faces leads to the other-race effect (Cross *et al.*, 1971; Feinman and Entwistle, 1976). These results indicate that experience or contact with other race faces cannot provide the entire explanation and instead other factors, including the necessity for the observer to process in-depth information that facilitates individuation, including motion information, may play a role. Indeed, other-race faces have also been found to be processed in less detail than same-race faces (Lipp *et al.*, 2009).

In Experiment 1 motion information was only made available to the observer at learning (encoding). It is interesting to understand what extent making motion information available at recognition interacts with the advantage provided by seeing faces moving at learning. During the recognition of familiar faces the availability of motion cues has been demonstrated to increase an observer's ability to recognise a face correctly due to motion providing more individuating information (characteristic motion patterns) than is gained from seeing only a static image (Knight and Johnston, 1997; Lander *et al.*, 1999). Therefore, using exactly the same design and procedure used in the current experiment, Experiment 2 questions the following: if faces are both learned and recognised with motion information, does the motion advantage increase? Based on accounts of familiar face recognition, it is predicted that the availability of motion information at recognition will provide a cue to identity and enhance familiarity with each of the target faces providing evidence of a strong motion advantage across both same- and other-race faces when the faces are both encoded and recognised from moving stimuli.

Based on transfer-appropriate processing accounts of memory (Morris *et al.*, 1977) and study-test congruency effects, that have previously been demonstrated during the recognition of both faces and images of scenes (Lander and Davies, 2007; Buratto *et al.*, 2009), it would also be reasonable to hypothesise that the availability of motion information at both learning and recognition would lead to a larger motion advantage than is found when presentation mode is incongruent at learning and recognition.

Transfer-appropriate processing is a phenomenon concerning memory performance. It has been revealed that performance on memory tasks is not only dependent on the depth of processing (Craik and Lockhart, 1972) but also on the relationship between how the to-be-remembered information (here faces) is initially encoded and how it is later retrieved at recognition (Morris *et al.*, 1977). Memory has been found to be most accurate when the processes engaged in during encoding match those engaged in during retrieval. Thus recognition accuracy may be best when the mode of encoding (moving/static) matches the mode of presentation at recognition. So learning a face in motion and recognising it with motion information in Experiment 2 could facilitate a greater motion advantage than found in Experiment 1. Similarly, learning a face when static and recognising it from a static stimulus may be advantageous to recognition. To summarise, Experiment 2 replicates Experiment 1 except that moving images were presented at test, rather than static ones. The results are shown in Table 11.2.

Table 11.2 Mean hits (%) and false alarm (FAs; %) with standard deviations (SD) in each of the conditions of Experiment 2; 15 scores contributed to each mean.

Race of face	Same		Other	
Learning Style	Static	Moving	Static	Moving
Hits	75.7	87.3	65.0	79.3
SD	14.9	8.7	8.9	11.2
FAs	15.3	10.0	18.7	14.3
SD	12.5	8.9	11.3	10.0

As with Experiment 1, in this experiment we found that there were more correct recognitions of same-race faces than other-race faces and there were significantly more correct recognitions of faces learnt from a moving clip than when learning was carried out from a static image. Again the advantage for moving faces was found with both same- and other-race faces. Of course, one thing that is interesting to do is to compare the results from Experiments 1 and 2. Here, it is possible to investigate the combined effects of motion at face encoding (learning) and face recognition (test). When amalgamating results from Experiments 1 and 2, a recognition advantage was found for own-race faces and those learned in motion. Interestingly, there was no advantage for recognising faces in motion and no interactional effects between motion at learning and test. This finding is important to the investigation of the role of encoding and retrieval factors (and their interaction) in the recognition advantage for moving faces. For example it may be that the motion advantage for faces is stronger than the study–test congruency effect. It is important to note that our results do not entirely conflict with the results of Buratto *et al.* (2009) who investigated study–test congruency and the dynamic superiority effect. Buratto *et al.* (2009; Experiment 2) utilised an 'inclusion' and an 'exclusion' condition. In the inclusion condition participants were required to respond 'old' based solely on the content of the clip, regardless of the presentation style (static, multi-static or moving). In contrast, participants in the exclusion condition were to respond 'old' based on both the content of the clip and the presentation mode. Therefore, participants in the exclusion condition were asked to make a 'new' response for clips that had been presented previously in the study phase but in a different presentation style. As such, the inclusion condition was akin to the procedure used in our experiments presented here. For the inclusion condition Buratto *et al.* (2009) found a main effect of presentation style at encoding with recognition sensitivity highest for images learned in motion compared with multi-static or single static, as well as an interaction between encoding and recognition style. Interestingly, they also found no effect of presentation style at recognition in the inclusion condition. However, in the authors' exclusion condition, a main effect of presentation style at recognition was revealed along with the main effect of presentation style at encoding. The authors posited that recognition decisions in their inclusion condition could be based on either familiarity or recollection whereas recognition judgements in the exclusion condition were reliant upon recollection in order to correctly accept targets (old targets with same presentation style at test) and to reject pseudo-targets (old clips with changed presentation style at test), suggesting that participants were able to consciously recollect presentation style information.

Finally, when considering the role of motion for unfamiliar face learning, it is important to consider what happens as a face becomes increasingly familiar. Accordingly, Lander and Davies (2007) investigated the impact of facial motion as a previously unfamiliar face becomes known. We presented participants with a series of faces each preceded by a name, and asked participants to try to learn the names for the faces. When the participants felt they had learned the names correctly they continued onto the recognition phase in which they were presented with the same faces (same presentation method as in learning phase), and asked to name the individual. The learning phase was repeated and the participant was asked to try to learn the names of the faces again if any of the names were incorrectly recalled, after which they took the recognition test again. This procedure was replicated until the participant correctly named all 12 faces shown. In the test phase, participants were presented with 48 degraded faces; 24 as single static images and 24 moving. In the moving condition the faces were each presented for 5 seconds. Participants were informed that some would be learned faces and some would be 'new' faces, for which they had not learned names. Participants were asked to name the face or respond 'new' and to provide a response for every trial. Results suggested that facial motion learning was rapid, and as such the beneficial effect of motion was not highly dependent on the amount of time the face was seen for. Rather there was support for the idea of rapidly learned characteristic facial motion patterns, with results only revealing an advantage for recognising a face in motion (at test) when the face had been learned moving. Conversely, when the face was learned as a static image, there was no advantage for recognising moving faces compared with a static image. Indeed, it seems that participants were able to extract and encode dynamic information even when viewing very short moving clips of 5 seconds. Furthermore, the beneficial effect of motion was shown to remain stable despite prolonged viewing and learning of the face identity in Experiment 2. In this experiment, participants were assigned to one of four experimental groups. Group 1 viewed episode 1 of a TV drama before the test phase, group 2 viewed episodes 1 and 2, group 3 episodes 1 to 3 and group 4 episodes 1 to 4. Each episode was 30 minutes in length. In the test phase, participants viewed moving and static degraded images of the characters and were asked to try and identify them by character name or other unambiguous semantic information. The results revealed that, although better recognition of characters from the TV drama was seen as the number of episodes viewed increased, the relative importance of motion information did not increase with a viewer's experience with the face (O'Toole *et al.*, 2002). The size of the beneficial effect remained relatively stable across time demonstrating how rapidly motion information, through familiarisation with the to-be-recognised face, can be integrated into a face representation and utilised at recognition.

11.4 Summary

To conclude, this chapter has outlined research that has investigated the impact of motion on the recognition of already familiar faces (personally familiar or famous). A moving face recognition advantage has been found in a number of studies (for example, Knight and Johnston, 1997; Lander *et al.*, 1999). Further work has suggested that the advantage for motion is most prominent when the motion shown is natural (mixture of non-rigid and rigid rotational movement), and when the to-be-recognised face moves in a distinctive manner (Lander and Bruce, 2000; Lander and Chuang, 2005). It seems likely that as a face becomes increasingly familiar we encode the idiosyncratic aspects of its facial motion, which act as an additional cue to identity, particularly when viewing conditions are problematic. Additional work is needed to explore how patterns of motion vary according to clip and identity. For example, do some individuals consistently move in a distinctive manner? Is this demonstrated across numerous clips of the same person? Conversely, does the beneficial effect of motion simply depend on the moving clip selected – with all individuals moving in both typical and distinctive ways? An educated guess might wager that individuals do have particular ways of moving that are idiosyncratic of them, but this individuality can be more or less visually present in different clips. Thus the beneficial effect of motion may be dependent on a number of factors including familiarity, distinctiveness of motion and clip selected.

Seeing a face move when learning its identity also seems to be advantageous (Pike *et al*., 1997; Lander and Bruce, 2003). We have outlined some of our recent research that shows that this is true regardless of whether the face belongs to our own race or another race (Lander, in prep.). Further work is needed to establish the theoretical basis of the learning advantage for motion, whether this is due to additional views or increased attention to moving faces (see Lander and Bruce, 2003). Regardless of the explanation, it is clear that faces are complex, mobile surfaces which change both rigidly and non-rigidly and that the role of such motion should be considered in future face perception studies. Such research will help us to understand the way in which dynamic information from the face is represented in memory and the precise mechanism by which it facilitates the retrieval of other information about personal identity.

References

Anthony, T., Copper, C. and Mullen, B. (1992). Cross-racial facial identification: a social cognitive integration. *Personality and Social Psychology Bulletin*, **18**(3), 296–301.

Baddeley, A. D. and Woodhead, M. M. (1983). Improving face recognition ability. In S. Lloyd-Bostock and B. Clifford (Eds.), *Evaluating Witness Evidence*, pp. 125–136. Chichester: John Wiley and Sons.

Bothwell, R. K., Brigham, J. C. and Malpass, R. S. (1989). Cross-racial identification. *Personality and Social Psychology Bulletin*, **15**, 19–25.

Braje, W. L., Kersten, D., Tarr, M. J. and Troje, N. F. (1998). Illumination effects in face recognition. *Psychobiology*, **26**(4), 271–380.

Brigham, J. C. and Malpass, R. S. (1985). The role of experience and contact in the recognition of faces of own-race and other-race persons. *Journal of Social Issues*, **41**(3), 139–155.

Bruce, V., Henderson, Z., Newman, C. and Burton, A. M. (2001). Matching identities of familiar and unfamiliar faces caught on CCTV images. *Journal of Experimental Psychology: Applied*, **7**, 207–218.

Bruce, V. and Langton, S. R. H. (1994). The use of pigmentation and shading information in recognising the sex and identities of faces. *Perception*, **23**, 803–822.

Bruce, V. and Valentine, T. (1988). When a nod's as good as a wink: the role of dynamic information in face recognition. In M. M. Gruneberg, P. E. Morris and R. N. Sykes (Eds.), *Practical Aspects of Memory: Current Research and Issues, Vol. 1*, pp. 169–174. Chichester: John Wiley and Sons.

Buratto, L. G., Matthews, W. J. and Lamberts, K. (2009). When are moving images remembered better? Study-test congruency and the dynamic superiority effect. *Quarterly Journal of Experimental Psychology*, **62**(10), 1896–1903.

Butcher, N. (2009). *Investigating the dynamic advantage for same and other-race faces*. Unpublished PhD Thesis, University of Manchester.

Butcher, N., Lander, K., Fang, H. and Costen, N. (2011). The effect of motion at encoding and retrieval for same- and other-race face recognition. *British Journal of Psychology*, **102**, 931–942.

Chance, J. E. and Goldstein, A. G. (1996). The other race effect and eyewitness identification. In S. L. Sporer, R. S. Malpass and G. Koehnken (Eds.), *Psychological Issues in Eyewitness Identification*, pp. 153–176. Mahwah, NJ: Lawrence Erlbaum.

Chance, J. E., Turner, A. L. and Goldstein, A. G. (1982). Development of differential recognition for own- and other-race faces. *Journal of Psychology*, **112**, 29–37.

Christie, F. and Bruce, V. (1998). The role of dynamic information in the recognition of unfamiliar faces. *Memory and Cognition*, **26**(4), 780–790.

Craik, F. I. M. and Lockhart, R. S. (1972). Levels of processing: a framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, **11**, 671–684.

Cross, J. G., Cross, J. and Daly, J. (1971). Sex, race, age and beauty as factors in recognition of faces. *Perception and Psychophysics*, **10**, 393–396.

De Haan, E. H. F., Young, A. and Newcombe, F. (1987). Faces interfere with name classification in a prosopagnosic patient. *Cortex*, **23**, 309–316.

Feinman, S. and Entwistle, D. R. (1976). Children's ability to recognize other children's faces. *Child Development*, **47**, 506–510.

Freyd, J. J. (1993). Five hunches about perceptual processes and dynamic representations. In D. Meyer and S. Kornblum (Eds.), *Attention and Performance, X1V: Synergies in Experimental Psychology, Artificial Intelligence and Cognitive Neuroscience – A Silver Jubilee*, pp. 99–120. Cambridge, MA: MIT Press.

Haxby, J. V., Hoffman, E. A. and Gobbini, M. I. (2002). Human neural systems for face recognition and social communication. *Biological Psychiatry*, **51**, 59–67.

Hill, H., Schyns, P. G. and Akamatsu, S. (1997). Information and viewpoint dependence in face recognition. *Cognition*, **62**(2), 201–222.

Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception and Psychophysics*, **14**, 201–211.

Kemp, R. I., Pike, G. E. and Brace, N. (2001). Video-based identification procedures – combining best practice and practical requirements when designing identification systems. *Psychology Public Policy and Law*, **7**, 802–807.

Knappmeyer, B., Thornton, I. M. and Bülthoff, H. H. (2003). Facial motion can bias the perception of facial identity. *Vision Research*, **43**, 1921–1936.

Knight, B. and Johnston, A. (1997). The role of movement in face recognition. *Visual Cognition*, **4**(3), 265–273.

Lander, K. and Bruce. (2000). Recognizing famous faces: exploring the benefits of facial motion. *Ecological Psychology*, **12**(4), 259–272.

Lander, K. and Bruce, V. (2003). The role of motion in learning new faces. *Visual Cognition*, **10**(8), 897–912.

Lander, K. and Bruce, V. (2004). Repetition priming from moving faces. *Memory and Cognition*, **32**(4), 640–647.

Lander, K., Bruce, V. and Hill, H. (2001). Evaluating the effectiveness of pixelation and blurring on masking the identity of familiar faces. *Journal of Applied Cognitive Psychology*, **15**(1), 116.

Lander, K., Christie, F. and Bruce, V. (1999). The role of movement in the recognition of famous faces. *Memory and Cognition*, **27**(6), 974–985.

Lander, K. and Chuang. (2005). Why are moving faces easier to recognize? *Visual Cognition*, **12**(3), 429–442.

Lander, K., Chuang, L. and Wickham, L. (2006). Recognizing face identity from natural and morphed smiles. *Quarterly Journal of Experimental Psychology*, **59**(5), 801–808.

Lander, K. and Davies, R. (2007). Exploring the role of characteristic motion when learning new faces. *Quarterly Journal of Experimental Psychology*, **60**(4), 519–526.

Lander, K., Hill, H., Kamachi, M. and Vatikiotis-Bateson, E. (2007). It's not what you say but the way you say it: matching faces and voices. *Journal of Experimental Psychology: Human Perception and Performance*, **33**, 903–914.

Lander, K., Humphreys, G. and Bruce, V. (2004). Exploring the role of motion in prosopagnosia: recognizing, learning and matching faces. *Neurocase*, **10**, 462–470.

Light, L. L., Kayra-Stuart, F. and Hollander, S. (1979). Recognition memory for typical and unusual faces. *Journal of Experimental Psychology: Human Learning and Memory*, **5**(3), 212–228.

Lindsay, D. S., Jack, P. C. and Christian, M. A. (1991). Other-race face perception. *Journal of Applied Psychology*, **76**(4), 587–589.

Lipp, O. V., Terry, D. J., Smith, J. R. *et al.* (2009). Searching for differences in race – is there evidence for preferential detection of other race faces? *Emotion*, **9**, 350–360.

Malpass, R. S. and Kravitz, J. (1969). Recognition for faces of own and other race. *Journal of Personality and Social Psychology*, **13**(4), 330–334.

McGraw, K. O., Durm, M. W. and Durnam, M. R. (1989). The relative salience of sex, race, age, and glasses in children's social perception. *Journal of Genetic Psychology*, **150**(3), 251–267.

McKone, E., Brewer, J. L., MacPherson, S., Rhodes, G. and Hayward, W. G. (2007). Familiar other-race faces show normal holistic processing and are robust to perceptual stress. *Perception*, **36**(2), 224–248.

Meissner, C. A. and Brigham, J. (2001). *Thirty years of investigating the own-race bias in face recognition: A meta-analytic review*. Paper presented at the meetings of the American Psychology-Law Society, New Orleans, LA.

Montepare, J. M. and Zebrowitz, L. A. (1998). Person perception comes of age: the salience and significance of age in social judgments. *Advances in Experimental Social Psychology*, **30**, 93–163.

Morris, C. D., Bransford, J. D. and Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Behavior*, **16**, 519–533.

O'Toole, A. J., Roark, D. A. and Abdi, H. (2002). Recognizing moving faces: a psychological and neural synthesis. *Trends in Cognitive Sciences*, **6**(6), 261–266.

Pascalis, O. and Kelly, D. J. (2009). The origins of face processing in humans: phylogeny and ontogeny. *Perspectives on Psychological Science*, **4**(2), 200–209.

Peterson, M. F., Abbey, C. K. and Eckstein, M. P. (2009). The surprisingly high human efficiency at learning to recognize faces. *Vision Research*, **49**(3), 301–314.

Pike, G. E. Kemp, R. I., Towell, N. A. and Phillips, K. C. (1997). Recognizing moving faces: the relative contribution of motion and perspective view information. *Visual Cognition*, **4**(4), 409–438.

Pilz, S. K., Bülthoff, H. H. and Vuong, Q. C. (2009). Learning influences the encoding of static and dynamic faces and their recognition across different spatial frequencies. *Visual Cognition*, **17**(5), 716–735.

Schiff, W., Banka, L. and de Bordes Galdi, G. (1986). Recognizing people seen in events via dynamic 'mug shots'. *American Journal of Psychology*, **99**(2), 219–231.

Schultz, J. and Pilz, K. S. (2009). Natural facial motion enhances cortical responses to faces. *Experimental Brain Research*, **193**(3), 465–475.

Sporer, S. L. (2001). Recognizing faces of other ethnic groups – an integration of theories. *Psychology Public Policy and Law*, **7**(1), 36–97.

Steede, L. L., Tree, J. J. and Hole, G. J. (2007). I can't recognize your face but I can recognize its movement. *Cognitive Neuropsychology*, **24**, 451–466.

Tzou, C. H. J., Giovanoll, P., Ploner, M. and Frey, M. (2005). Are there ethnic differences of facial movements between Europeans and Asians? *British Journal of Plastic Surgery*, **58**(2), 183–195.

Valentine, T. (1991). A unified account of the effects of distinctiveness, inversion, and race in face recognition. *Quarterly Journal of Experimental Psychology Section A – Human Experimental Psychology*, **43**(2), 161–204.

Valentine, T. and Bruce, V. (1986). The effects of distinctiveness in recognising and classifying faces. *Perception*, **15**, 525–535.

Valentine, T. and Ferrara, A. (1991). Typicality in categorization, recognition and identification: evidence from face recognition. *British Journal of Psychology*, **82**, 87–102.

Vetter, V., Zielke, T. and Von Seelen, W. (1997). Integrating face recognition into security systems. In *Audio- and Video-based Biometric Person Authentication: Lecture Notes in Computer Science*, **1206**, 439–448.

Vokey, J. R. and Read, J. D. (1992). Familiarity, memorability, and the effect of typicality on the recognition of faces. *Memory and Cognition*, **20**, 291–302.

Zhang, Z. Y., Liu, A. C., Adler, D. *et al.* (2004). Robust and rapid generation of animated faces from video images: a model-based modelling approach. *International Journal of Computer Vision*, **58**, 93–119.

Chapter 12

Facial image comparison

Josh P. Davis, Tim Valentine and Caroline Wilkinson

12.1 Introduction

In this chapter, the problems associated with the individualisation of people depicted in photographic forensic evidence such as closed circuit television (CCTV) images are described. Evidence of this type may be presented in court and, even with high-quality images, human identification of unfamiliar faces has been shown to be unreliable. Therefore, facial image comparison or mapping techniques have been developed. These have been used by expert witnesses providing opinion testimony as to whether two images depict the same person or not. With *photographic video superimposition*, one image is superimposed over a second so that a series of visual tests can detect differences or similarities in facial features. With *morphological comparison analysis* facial features are classified into discrete categories, providing an indication of whether these are similar across images. Finally, with *photo-anthropometry* the proportional distances and sometimes the angles between facial landmarks are calculated and compared. Recent research using each technique is described, and the difficulties associated with their application in forensic settings evaluated. At present, no method provides certainty of identification and great care should be taken if presented in court to obtain a conviction without substantiating alternative evidence.

Government and private sector investment in crime prevention initiatives has made CCTV systems common in many urban areas. Although there are no official records, the UK probably has the highest density in the world, with at least 3 million cameras nationwide (McCahill and Norris, 2003; Norris *et al.*, 2004). There may be as many as 26 million cameras in the USA (*Washington Post*, 8 October 2005) and large-scale implementation seems inevitable elsewhere (Norris *et al.*, 2004). Widespread deployment of CCTV raises many issues. Concerns have been raised about infringement of rights to privacy (Norris and Armstrong, 1999; Introna and Wood, 2004) and crime prevention efficacy (Brown, 1995; Gill *et al.*, 2005). In this chapter, we focus on the reliability of CCTV for identification purposes.

Undoubtedly, CCTV footage can be very useful in establishing the sequence of events. When confronted with CCTV images some suspects confess to the offence. However, when identification is disputed, it can be very difficult to establish the identity of an offender captured on CCTV (Costigan, 2007; Edmond *et al.*, 2009). Research by computer scientists and by psychologists has highlighted the difficulties involved in the successful identification of unfamiliar people depicted in even the highest-quality images. In cases of disputed identification, expert witnesses may provide opinion evidence of identification in court. In this chapter some of the techniques used by experts, and the legal issues raised by this evidence, are discussed. A specific focus is on recent developments in the field, including methods that have tested identification against facial databases, and studies that have evaluated expert and public identification from CCTV footage.

Expert witnesses carry out analyses after evidence collection. However, algorithmic pattern recognition systems have been developed with the aim of identifying faces in 'real time' under two related circumstances (Brunelli and Poggio, 1993; Heisele *et al.*, 2003). One function is the verification of known individuals, for instance, to ensure access to a secure building by

Craniofacial Identification, ed. C. Wilkinson and C. Rynn. Published by Cambridge University Press.

authorised persons presenting themselves for scrutiny. A second function is for general security purposes, so that an alarm is triggered if an individual whose face is on a criminal or terrorist database enters a monitored area. In both cases, a human facial still or moving image is extracted, transformed into an abstract representation or unique individual identifier (*biometric*) for comparison against a gallery of facial images. Computer systems monitor the imagery until a 'best' image is presented and a probability for a match is generated. Facial imagery is already used as a biometric along with others such as fingerprints. One advantage for security applications is that facial biometrics can be acquired without the cooperation, consent or knowledge of the target.

Following recent terrorist attacks, the biometric industry has rapidly expanded (Introna and Wood, 2004). Norris and Armstrong (1999) predict that, when perfected, the performance of face-recognition technology will be as accurate as automatic car number plate readers. However, the three-dimensional (3D) facial surface is considerably more complex than that of a standardised number plate. The technology involved in the detection of moving facial features from a background scene is also complex, particularly if partially occluded if in a crowd, or in shadow (Hjelmas and Low, 2001), and expressions and age change facial appearance. The performance of current automatic systems is better than normal human perception under optimal conditions (i.e. controlled pose, distance, direction of lighting). Accuracy is severely impaired if views are incongruent or comparison images are recorded under different lighting or other environmental conditions (Burton *et al.*, 2001; Phillips *et al.*, 2007).

12.2 The use of CCTV images in court

Photographic evidence has been admissible in court in the UK for nearly 150 years (*R v Tolson*, 1864). CCTV footage was first used in the 1980s to provide information about theft from a retail store (*R v Fowden and White*, 1982). In a recent legal review, the Attorney General considered four situations in which it was appropriate for CCTV imagery to be used as evidence of identification (Attorney General's Reference, 2003).

12.2.1 Familiar face recognition

Individuals claiming prior familiarity with a defendant may give evidence as a witness even if the footage is no longer available. The recognition of familiar faces in CCTV images is generally robust (Bruce *et al.*, 2001; Burton *et al.*, 1999). For instance, Burton *et al.* (1999) found that university students were 90% correct when recognising lecturers from their own department in poor-quality video. A similar high level of accuracy was found in a task in which students were presented with a series of paired facial images (Bruce *et al.*, 2001). When participants were familiar with the targets, identification accuracy was extremely high. However, these images were shown in context-rich settings, such as footage from the psychology department corridors and it is less clear whether accuracy would be as high in a neutral context.

12.2.2 Unfamiliar identification by the jury

When a photographic image is 'sufficiently' clear, the jury can be asked to compare it with the defendant in the dock. In *R v Dodson and Williams* (1984), the Court concluded that: 'so long as the jury – are firmly directed that to convict they must be sure that the man in the dock is the man in the photograph, we envisage no injustice arising from this manner of evaluating evidence with the aid of what the jurors' eyes tell them is a fact which they are sure exists'. Jurors, and indeed most police officers, would be previously unfamiliar with the suspect. Identification of unfamiliar people in even the highest-quality photographs is surprisingly unreliable even with no memory demands (Bruce *et al.*, 1999, 2001; Henderson *et al.*, 2001), and when the target is present in person (Davis and Valentine, 2009). The typical positioning of CCTV cameras, often above head height with a large field of view, lessens the likelihood of obtaining clear images (Davies and Thasen, 2000). Distance from the camera to the subject (Loftus and Harley, 2004), specificity of viewpoint, expression, and environmental lighting effects all influence face matching (Hill and Bruce, 1996; Bruce *et al.*, 1987, 1999) and recognition (Bruce, 1982; Bruce *et al.*, 1987; Hill *et al.*, 1997). A mismatch of any of these factors leads to identification accuracy reductions.

12.2.3 Ad hoc expertise

A witness not previously familiar with the defendant may spend substantial time viewing and analysing evidential images, thus familiarising themselves with the accused and gaining a 'special knowledge that the court did not possess', thereby developing an 'ad hoc' expertise (*R v Clare and Peach*, 1995). Some research has been conducted on the processes involved in face familiarisation (Bonner *et al.*, 2003; Clutterbuck and Johnston, 2005). However, it is unclear how much inspection is required for identification to be as reliable as someone familiar with the culprit. Furthermore, knowing the context will be unavoidable, and context information can bias identification decisions. This has been found with the more established technique of fingerprint analyses (Dror *et al.*, 2006). International fingerprint experts at two separate time points provided assessments as to the likelihood of two fingerprints being from the same person. In the first instance, all experts gave a positive identification of the fingerprints. However, unaware that they had previously seen the fingerprints, four out of five provided different judgements when the contextual information provided suggested that a match was not expected. It is not possible to conclude that different experts would behave in the same manner. Nevertheless, it is likely that facial analytical methods would also be vulnerable to cognitive biases of this type.

12.2.4 Facial mapping or facial image comparison

Practitioners from different disciplines, including medicine, anatomy, anthropology, military surveillance, computer science and art may be invited to present opinion evidence based on professional expertise, 'of identification based on a comparison between images from the scene (whether expertly enhanced on not) and a reasonably contemporary photograph of the defendant, provided the images and the photograph are available for the jury' (Attorney General's Reference, 2003). The early use of facial image comparison experts was often by defence solicitors challenging the arrest of their clients on the evidence of police officers who claimed to recognise them as offenders in CCTV footage. These reports established innocence by demonstrating inconsistent facial structures. The majority of these cases did not reach court, as the prosecution dropped the charges. It then became inevitable that the police would utilise the same expertise to attempt to prove identification. The first Court of Appeal judgment verifying the use of expert evidence of identification in photographic images was in 1993 (*R v Stockwell*, 1993). Over the next 10 years, at least 500 expert witness facial image comparison reports were prepared annually (Bromby, 2003). This type of testimony is deemed admissible as the sole basis for a conviction, if images are good quality (*R v Hookway*, 1999; *R v Mitchell*, 2005).

In the USA following a series of court judgments (*Frye v United States*, 1923; *Daubert v Merrell Dow Pharmaceuticals Inc.*, 1993; *Kumho Tire Co v Carmichael*, 1999) all expert witness techniques are required to meet scientifically rigorous standards. In the UK, it is the prerogative of a judge to determine whether expert witnesses can provide 'information which is likely to be outside the experience/knowledge of a judge or jury' (*R v Turner*, 1975). The Association of Chief Police Officers (ACPO) specifies minimum requirements for facial analyst experts, including knowledge of facial anatomy, anthropometry, physiology and photographic image analysis techniques and that 'expertise is generally achieved through experience and is measured by the acceptance of reports presented in court' (ACPO, 2003: 8). Juries may be directed to draw their own inferences as to the credence of the expert and the evidence. However, two different experts using similar techniques can come to different conclusions (*R v Clarke*, 1995; *Church v HMA*, 1996; *R v Loveridge and others*, 2001; *R v Gray*, 2003; *R v Gardner*, 2004;). Indeed, five different facial experts were called to give evidence in the Scottish case of *Church v HMA* (1996). Three argued that the quality of crime scene CCTV images were too poor to allow analysis. In contrast, the other two experts presented evidence of reliable differences. Additional evidence in the case was provided by three eyewitnesses who positively identified the defendant in a line-up.

Some recent research suggests that experts are better than the public at facial identification from CCTV footage. One study by Wilkinson and Evans (2009, 2011) employed a CCTV system installed at the University of Manchester to record video clips of six young adult White males (targets). Sixty-one participants (30 male and 31 female) and two experts were asked to identify the target in each clip by comparison with a photographic face pool of similar males (an option of 'not present' could be chosen). The experts recorded consistently higher identification rates and lower error rates than the public, with almost double the identification rates and half the errors. The public recorded high levels of false acceptance (10%) and false rejection (49%) whether the target wore a hat or not. The experts recorded a false rejection rate of 8% and a false acceptance rate of 2% for full head identification, and a false rejection rate of 25% and false acceptance rate of 3%

when the targets wore hats. This study suggested that training and experience in facial analysis may produce more reliable facial identification. However, it does not address the fact that UK experts originate from different fields with different levels of training, or the possibility that they may have an innate 'ability' in facial recognition.

Other studies have also focused on the training of experts in relation to reliability. Lee *et al.* (2009) studied a partially trained group of postgraduate students from the University of Dundee and compared their identification ability with the public using poor-quality CCTV footage and photographic face pools. Overall, error rates were high (33%), with false acceptance rates (22%) double the false rejection rates (11%). The partially trained group was no more reliable than the public when analysing this very poor-quality footage.

12.2.4.1 Facial image comparison techniques

The focus of this chapter is on the techniques facial comparison experts may use. However, the security, storage and integrity of images must be considered. There is no digital equivalent of a photographic negative, which provides physical evidence. It might be essential to encode a digital signature or watermark within each piece of digital evidence to establish an audit trail to highlight manipulations (House of Lords, 1997/1998). Some guidelines have been published (British Standards Institute (BSI), 2005; Scientific Working Group on Imaging Technologies, 2005). However, as technology develops, additional precautions will be required.

There are three general forensic approaches to determining whether images depict the same person, often described as facial mapping or facial comparison. These are *photographic video superimposition*, *morphological comparison analysis* and *photo-anthropometry*, although they are not mutually exclusive and practitioners may combine all three. One of the primary issues when faced with facial image comparison is that a 2D image is only a representation of the 3D facial surface. Therefore ACPO (2003) recommend that images being compared should be taken from as similar a viewpoint as possible. However, even with digital images, discrepancies in source equipment can create difficulties. The optical properties of the lens, such as its focal length, can affect the relative proportion and shape of features (Harper and Latto, 2001; Edmond *et al.*, 2009). Close-up images from a wide-angled lens (e.g. in a cash machine), and a telephoto lens (used to 'zoom in' from a distance) can induce distortion.

Bramble *et al.* (2001) suggest that software filters can refine visual data to clarify and enhance edge detail. For instance, *frame averaging* techniques can be applied to multiple consecutive frames to produce one higher-quality image, clarifying static shadowed details by equalising illumination across frames. *Frame fusion* software can resolve blur caused by motion across multiple frames, producing a more stable image. However, excessive manipulations may be challenged in court.

Some image-comparison analyses are performed using optical devices such as a stereoscope. This creates an artificial 3D representation when applied across two adjacent frames, as slight movement gives an impression of depth. Proponents claim that the more experienced the practitioner, the greater the perceived enrichment of the image. However, the methodology has been criticised for being subjective in nature and for the inability to demonstrate laboratory techniques in a courtroom. Furthermore, use of a stereoscope may be inappropriate for forensic facial comparison, as when viewing the faces of different individuals in a stereoscope 'the faces blend into one in a most remarkable manner.' (From a letter written by A. L. Austin to Charles Darwin, cited by Galton, 1878.)

In the light of these issues, İşcan (1993) argues that the facial image analyst is required to 'reinvent' the methodology for every case. Part of the procedure will be an attempt to locate unique identifiers or a combination of facial features or facial measurements that can reliably distinguish the target.

Bromby (2003) recommends the use of a six-point qualitative scale to provide an assessment of a match, ranging from: 1 = Lends no support to 6 = Lends powerful support. Bromby argues that use of a scale avoids assessing feature similarity statistically against a population database. However, even if only used by an experienced facial expert, it is difficult to demonstrate objectivity. In addition, criticism has been directed at proponents for not normally providing the probability of a match of identity in court. Indeed, as a protection against miscarriages of justice, there have been calls for a national database of facial measurements so that the proportion in the population who share similar face morphology can be used to calculate the likelihood of a unique identification (*R v Gray*, 2003). Without this safeguard, the judges argued that opinions were potentially subjective, although they did not rule that evidence from facial mapping experts should be inadmissible. More recently, the same court has also ruled that knowing the likelihood of shared facial characteristics is not necessary (*R v Gardner*, 2004). The court ruled that if a technique could be shown to aid the court, an experienced practitioner using specialist equipment may present *subjective* opinion of identity in court, based on personal observations. However, professionally presented expert evidence can appear extremely convincing, making it very difficult for a jury to assess the scientific basis of the opinion.

Whatever method, in the majority of cases, a unique identification cannot be made. Even a multitude of similarities between two faces can only add support to the assertion that the two images are of the same person. In contrast, one reliable demonstrable difference that is not due to natural changes in an individual's appearance or to differences in imagery conditions will positively exclude an identity match. Images taken some time apart pose a particular issue. Ageing is accompanied by a predictable pattern of changes to the facial structure, including growth of the jaws and nose throughout childhood, altering the position and relative size of the eyes. This heart-like expansion of the head from a constrained nodal point at the junction of the brainstem and spinal cord has been described using a mathematical approximation called *cardioidal strain* (Shaw *et al.*, 1974). Other changes occur throughout adulthood and follow a predictable pattern (Gonzalez-Ulloa and Flores, 1965; Takema *et al.*, 1994; Khalil *et al.*, 1996). The skin loses elasticity due to biochemical changes in the underlying connective tissue that causes it to become less firmly attached to the underlying bone or muscles. Wrinkles form due to changes in the distribution and formation of collagenous material in the skin, a decrease in the resilience of the fibres, and a decline in the number of fibroblasts leading to dehydration. Sagging of flesh, loss of adipose tissue, blurring of iris detail, increased prominence of facial lines and hair loss also occur. An old person may appear to have sunken eyes due to resorption of adipose tissue at the orbits and more visible veins beneath the thinner orbital skin, producing dark circles below the eyes (Gonzalez-Ulloa and Flores, 1965). Nasolabial and mental creases will become more marked and deeper with increased age (Neave, 1998). Bone resorption at the alveolar processes with loss of teeth in later life will alter the jaw line and mouth significantly (Bodic *et al.*, 2005). The nose and chin will appear more prominent, the distance between the nose and the chin will decrease, with the mouth appearing to sink into the face, and there is some growth of the cartilaginous portions of the nose and the ears throughout adulthood (Neave, 1998). Although age-related changes to the skin surface follow a predictable pattern, the timing of this pattern is not predictable (Novick, 1988; Loth and İşcan, 1994; Orentreich, 1995) and changes accrue more slowly in some people so that there is a great deal of variation between individuals of the same age. Facial ageing is influenced by lifestyle and may be accelerated by external factors such as smoking, sleeping position, chronic alcohol consumption, sun damage, medication or loss of weight (Taister *et al.*, 2000). These changes are also related to genetic factors, skin type, face shape and subcutaneous fat levels. Cosmetic interventions, such as plastic surgery, mole removal and make up, can also significantly alter facial structure and theoretically, a criminal determined to evade conviction could radically change their perceived appearance. In these circumstances facial-image comparison techniques would not be useful for identification.

With *photographic video superimposition*, one image is superimposed over a second on a screen and a series of visual tests are performed for the detection of differences or similarities. Various fading mechanisms 'make one face disappear into another, with the second image eventually replacing the first' (İşcan, 1993: 63). These include visual flicker and vertical, horizontal or diagonal wiping so that a line erasing part of one image reveals part of the second. For instance, Mazumdar and Sinha (1989) developed software that allows viewing of sections of two images side-by-side. They claim that facial symmetry, or a lack of symmetry, can be highlighted, even if the target is shown in disguise. Using the technique, Sinha (1996) describes a case study by an Indian state forensics laboratory in demonstrating that two different identity photographs depicted the same individual, after a passport official questioned the resemblance.

Vanezis and Brierley (1996) report that they were asked to apply video superimposition techniques to provide opinion evidence of identity of 51 individuals in 46 UK cases. Forty were submitted by prosecuting authorities, two-thirds being robberies from banks or shops. The authors carried out frame-by-frame inspection of recordings from the crime scene, to select stills that when magnified aligned closely with suspects' photographs. They suggest that minor viewpoint disparities were not a problem, stating that 'what is acceptable depends on the experience of the examiner who should be aware of the various possible positional changes of the head' (Vanezis and Brierley, 1996: 28). The speed of video superimposition fade depends on the number of contours, such as scars in close proximity, with an increase in target features requiring a slower wipe, sometimes conducted with increased magnification. Occasionally the authors would superimpose a series of frames to highlight ill-defined features. In cases in which a positive identification was made, the ear was identified as the most useful feature, with scars and moles providing important evidence. Using this methodology, the authors claimed 11 'reliable' identifications as well as 16 'probable' and eight 'possible' identifications. They also suggest that they could exclude three of the 51 individuals due to reliable feature dissimilarities. The authors also note that they used anthropometrical indices in the examinations although these are not discussed in the paper.

Evidence from an expert witness using video superimposition was first admitted in court in the UK in the early 1990s, with the technique's status confirmed on appeal (*R v Clarke*, 1995). Nevertheless, one trial judge described it as 'really just a subjective assessment, it is not scientific; he is just a man with a magnifying glass. There are no measurements or calculations or anything of that kind' (*R v Kerrigan*, 1998). Furthermore, analysts claim to be able to 'see' details in visual images that are invisible to the untrained eye because of their 'experience and equipment' (*R v Gray*, 2003). İşcan (1993) claims that video superimposition is extremely susceptible to differences in facial viewpoint and a number of procedures such as a slow fade can increase an 'illusion' of a perfect match and provide highly persuasive evidence in court.

Morphological comparison analysis is a method by which facial features are defined and classified based on shape and size to provide an indication of whether these properties are similar across images. The technique has its scientific origins in work by Alphonse Bertillon (1853–1914) in France in the late nineteenth century. In his book *Identification Anthropometrique*, Bertillon described a classification system for use on arrested criminals using measurements of different body parts. For photographic analysis and forensic purposes, feature-by-feature classification is performed, an approach similar to fingerprinting analysis, in that it is assumed that faces have individuating characteristics. However, Mardia *et al.* (1996) note that even with distorted fingerprints the topology of shape structures are often clearly defined. In contrast, there are no highly defined connections within a face, and expression changes will alter the relative position and dimensions of the majority of facial structures.

Vanezis *et al.* (1996) examined the reliability of one morphological classification technique. Seven participants rated high-quality facial photographs of 50 males, aged 18–60 years from five different views, sub-classifying 39 feature categories into 87 different descriptors. For instance, there were three basic categories used to describe nose shape – nose tip shape, nostril visibility and nasal alae. For nose tip shape there were seven descriptors – *undecided*, *pointed*, *bilobed*, *hooked*, *rounded*, *pronounced* and *asymmetrical*, whereas there were five descriptors for nostril visibility and six for nasal alae. Fourteen categories possessed no discriminatory power or were associated with inter-assessor disagreement and were removed from further investigation. The authors suggest that the remaining categories might be appropriate for use in cases of disputed identification. However, statistical analyses to individuate different faces would have required nominal level analyses and the sample was heterogeneous in terms of age range, meaning it would be unlikely that many would be the subject of identification disputes.

Vanezis *et al.* (1996) suggest that morphological classification is most appropriate when images are of low resolution or are taken from dissimilar angles precluding the use of other facial comparison techniques. However, they note that the technique is less effective with 'average-type' people, as they tend to be classified into the same sub-categories. Furthermore, İşcan (1993) observes that features that discriminate one ethnic population from one geographical region may not adequately individuate those from another. Moreover, no large-scale databases containing exclusively morphological characteristics have been compiled to provide an indication of the likelihood of two or more individuals possessing the same features. Indeed, at least one conviction has been overturned when testimony was based on this methodology, due to the lack of the 'probability of occurrence or combinations of occurrence of particular facial characteristics' (*R v Gray*, 2003).

Finally, with *photo-anthropometry* facial landmarks are identified and the distances and sometimes the angles between them are calculated and compared across images. Measuring the face for different purposes has had a long history. Ballytyne (1984) suggests that the ancient Babylonians were probably the first proponents. According to Mardia *et al.* (1996), researchers in different disciplines have utilised various actual and photographic face measures. These include anthropologists, for the classification of faces by race or other category; surgeons, for craniofacial surgery; and orthodontists, for dentistry. However, these were mainly for the analyses of group similarities or differences and not for individuation as required in a legal context.

Absolute distances cannot easily be measured in a photograph, without knowing the exact camera distance and lens focal length (İşcan, 1993; Bramble *et al.*, 2001). Indeed, it is surprisingly complex to estimate full-body height (Bramble *et al.*, 2001; Alberink and Bolck, 2008; Lee *et al.*, 2008). Therefore, proportional analyses of the relationship between facial features in one image are compared with those in a second. In a frontal view, the *referent* distance will often be between the top of the head and the chin for vertical dimensions and the distance between the outside of the ears for horizontal dimensions. Figure 12.1 illustrates how this type of evidence can be presented in court, with the superimposition of grids over images that have been enlarged or reduced in size to visually match dimensions. If multiple images have been obtained a number of similar comparisons can be included in a report.

Details of techniques used in court have been published. Porter and Doran (2000) described methods of face measurement which proved successful in matching the identity of suspects in various identity documents and passports, resulting in 'several' successful prosecutions in Australia. Four anthropometric measurements were taken – the horizontal face width between the lower ears, the mouth width and the nose width as well as interpupillary distance, which served as the referent measurement to which proportions were expressed. Halberstein (2001) describe three cases in Florida in which between 9 and 12 anthropometric facial distances were measured and compared in crime photographs of the suspects and compared with the suspects. In two cases, facial proportions were similar and successful prosecutions were obtained. In the third case, reliable differences were identified. However, there were no tests of the uniqueness of measures against a database, and both Halberstein, and Porter and Doran carried out additional morphological comparisons. It is also not possible to determine how much weight was placed on this evidence in court.

In other studies, obtained photo-anthropometric measurements have been evaluated against databases of facial images (Catterick, 1992; Burton *et al.*, 1993; Mardia *et al.*, 1996). These studies differed substantially in database size and homogeneity, and a small database containing dissimilar faces may not provide an adequate test of a technique. Burton *et al.* (1993) examined which anthropometric measures best discriminate between genders, with hairstyle obscured and facial hair shaved (except eyebrows). They measured 18 distances between landmarks in a *frontal* view, finding that 12 proportional distances reached criteria for inclusion in being able to differentiate 85% of the 179 faces, with the highest contribution coming from

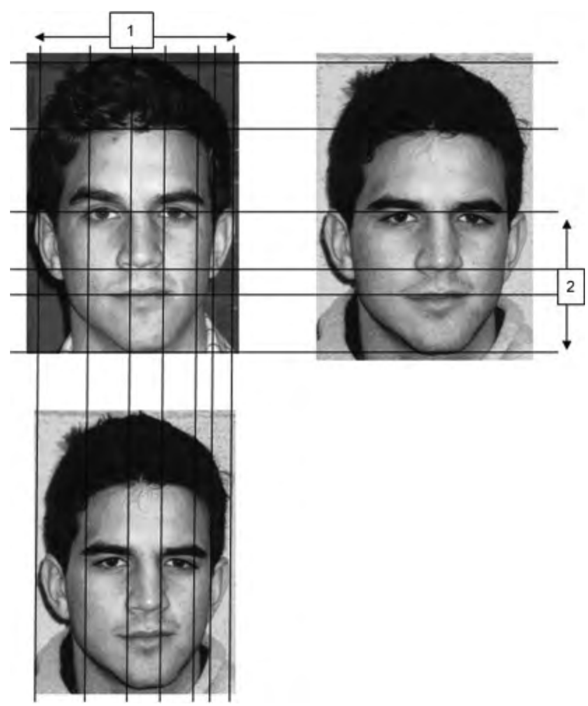


Figure 12.1 Illustration of how the results of photo-anthropometrical analyses could be presented by an expert witness in court. Accompanying the figure would be a table detailing the measurements in terms of **1**. Horizontal distances, expressed as ratios of the distance between the *superaurale* (ears), and **2**. Vertical distances, expressed as ratios of the distance from the *gnathion* (chin) to the *ectocanthians* (eyes). In this case, the same actor is depicted in the images.

eyebrow thickness, nose width at base and mouth width. The authors also conducted some analyses with additional images captured in *profile* view. They found that performance (94% accurate) equalled human ability at discriminating the gender of these faces, but only with the inclusion of 16 variables from both viewpoints.

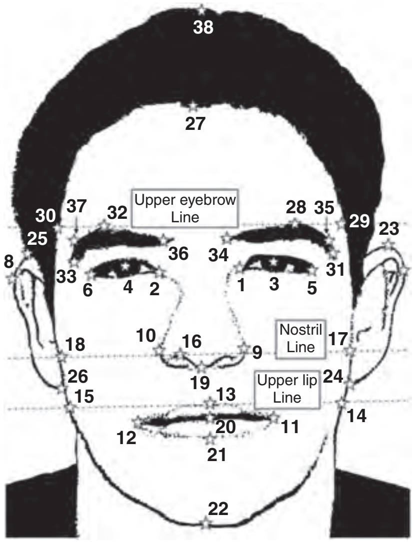
Catterick (1992) described a system of measuring distance proportions between four *frontal* landmarks using hand-held calipers against a database of passports and magazine photos, and concluded that the technique was limited in discriminating between different faces. Similarly, Kleinberg *et al.* (2007) describe a computerised measurement program in which an operator locates four facial landmarks, the *stomion* (centre of mouth), the *nasion* (the depressed area between the eyes) and the right and left *exocanthia* (outer eyes) in *frontal* photographs. The system calculates the distances and angles between the landmarks to conduct proportional analyses. The authors tested the system against a database of high-quality frontal photographs of 120 male police recruits first described by Bruce *et al.* (1999). Many of the images had proved to be difficult for participants in the original study to match by visual inspection even in ideal conditions. Kleinberg *et al.* (2007) found that it was not possible to reliably match the photograph and video still of each target using photo-anthropometry, and suggested that the 'method does not generate the consistent results necessary for use as evidence in a court of law' (Kleinberg *et al.*, 2007: 779). However, some of the video images were rotated to the left or right by up to 10%. It would perhaps be inadvisable to forensically apply this type of technique to images differing in viewpoint to this extent.

Anthropometric analyses of a database of 358 young White male faces, captured in *frontal* and *profile* views and taken in a controlled environment was conducted by Mardia *et al.* (1996). Twenty landmark (11 *frontal* and 9 *profile*) distance measurements and the angles between landmarks were collected to conduct shape analysis. There were high correlations between all measurements limiting the ability to distinguish between different faces. However, *profile* and *frontal* analyses were conducted separately and if data were combined, a more robust method of distinguishing faces may have emerged. Nevertheless, this research illustrates the difficulties involved in applying the technique even with extremely high-quality viewpoint-standardised images.

Roelofse *et al.* (2008) describe a method of combining morphological comparison and photo-anthropometric techniques with *frontal* photographs to establish the commonality of facial characteristics. Two hundred Bantu-speaking South African males aged 20 to 40, were photographed in a highly standardised environment. After removing measures that did not sufficiently vary, eight morphological features were selected for classification and sub-divided into 29 distinct categories. In addition, 12 anthropometric measurements were measured using digital calipers. These were sub-classified into discrete categories, by dividing the range of each value into three. The authors conducted separate analyses using different regions of the face to assess commonality of groups of features using both the morphological and the anthropometric categories. However, many of the faces were classified into the same categories, indicating weak individualisation. Nevertheless, inter-rater reliability was high and therefore effects of photographic distortion were small. However, dividing measurements into three was perhaps arbitrary, and some of the power of the data would have been lost.

12.2.5 Facial landmark identification

There are many unresolved issues concerning photo-anthropometric analysis. However, the technique potentially provides highly detailed, close-up measurements of facial structures, the assessment of error levels and parametric analyses, if images are of sufficient resolution and quality. Some automatic face-recognition software based on geometric feature-based algorithms use this approach and it is therefore likely to remain the focus of empirical research. Recently, *DigitalFace*, a custom software-assisted facial landmark identification system was developed by Davis and colleagues (Davis, 2007; Davis *et al.*, 2010), and has been used in legal cases. The system requires an operator to locate up to 38 specified landmark sites in *frontal* view (Figure 12.2); and 14 in *profile* view (Figure 12.3) on images displayed on a computer monitor, producing a database of 25 distance and 14 angular measurements in frontal view and 12 distance and 11 angular measurements in profile view (Figures 12.4–12.7). These extend those used in previous anthropometric (Catterick, 1992; Mardia *et al.*, 1996; Kleinberg *et al.*, 2007) and psychological studies (Burton *et al.*, 1993). *DigitalFace* operates most effectively with images from the front or side as in



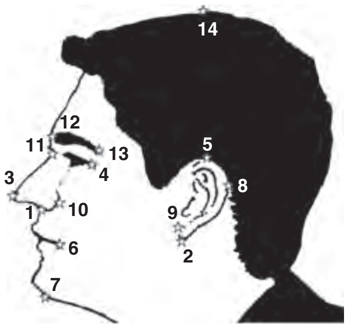
1. Right inner eye (*r endocanthian*) **2.** Left inner eye (*l endocanthian*) **3.** Right pupil centre **4.** Left pupil centre **5.** Right outer eye (*r exocanthian*) **6.** Left outer eye (*l exocanthian*) **7.** Right outer ear (*r postaurale*) **8.** Left outer ear (*l postaurale*) **9.** Right most outer point of nasal area (*r alare*) **10.** Left most outer point of nasal area (*l alare*) **11.** Right outer mouth (*r cheilion*) **12.** Left outer mouth (*l cheilion*) **13.** Top of upper left lip (*l superior labiale*) **14.** Right edge of face at upper lip line **15.** Left face edge at upper lip line **16.** Centre of left nostrum (*l supra subalare*) **17.** Right edge of face at nostril line **18.** Left face edge at nostril line **19.** Nose base (*subnasale*) **20.** Centre of mouth (*stomion*) **21.** Lower lip base (*inferior labiale*) **22.** Chin (*gnathion*) **23.** Right ear top (*r superaurale*) **24.** Right ear base (*r subaurale*) **25.** Left ear top (*l superaurale*) **26.** Left ear base (*l subaurale*) **27.** Hair line at forehead midpoint (*trichion*) **28.** Right eyebrow top (*r superciliare*) **29.** Right face edge on eyebrow top line at hair contact **30.** Left face edge on eyebrow top line at hair contact **31.** Right eyebrow base (*r orbitale superious*) **32.** Left eyebrow top (*l superciliare*) **33.** Left eyebrow base (*l orbitale superious*) **34.** Right inner eyebrow **35.** Right outer eyebrow (*r frontotemporale*) **36.** Left inner eyebrow **37.** Left outer eyebrow (*l frontotemporale*) 38. Highest point on head (*vertex*).

Figure 12.2 Locations of full-face (frontal) landmarks. Common names given in DigitalFace instructions (with anatomical definitions). Note: right and left locations are from the perspective of the viewer. Reprinted from *Forensic Science International*, **200** (1–3) Davis, J. P., Valentine, T. & Davis, R. E. Computer assisted photo-anthropometric analyses of full-face and profile facial images, 165–176. Copyright (2010), with permission from Elsevier.

police mugshot images. However, other angles are acceptable with matched viewpoints. Some *transient* measures lack medium-term permanency, such as eyebrow length or hairline, and may not be appropriate for inclusion in a forensic investigation.

Davis *et al.* (2010) describe a series of analyses, conducted to simulate 64 individual forensic investigations. Each analysis employed different sets of measures, as might be necessary dependent on visibility of facial landmarks in photographic evidence, tested against a homogeneous database of facial images of 70 individuals with a similar physical description. The aim was to examine whether novel photographs (*probes*) of eight faces taken 3 weeks previously by the same camera, would be matched with photographs of the same people (*targets*) already stored within the combined *frontal* and *profile* databases. Viewpoint in the images of the same person did not exactly match. However, all photos were posed and would meet requirements of identity documents such as passports. Indeed, unless chin supports or restraining clamps are used, it is unlikely that crime scene images would be closer in viewpoint, thus this tested the system in optimal conditions.

All measures were standardised and for each analysis, the squared Euclidian distance was computed



1. Nose base (*subnasale*) **2.** Ear base (*subaurale*) **3.** Nose tip (*pronasale*) **4.** Outer eye (*exocanthian*) **5.** Ear top (*superaurale*) **6.** Mouth corner (*cheilion*) **7.** Chin (*gnathion*) **8.** Ear rear (*postaurale*) **9.** Front of ear, point of attachment of ear lobe to cheek (*otobasion infrious*) **10.** Most lateral point of the curved part of the nose alar (*alar curvature*) **11.** Deepest landmark at the top of the nose (*sellion*) **12.** Prominent midpoint of eyebrows (*glabella*) **13.** Outer eyebrow (*frontotemporale*) **14.** Highest point on head (*vertex*)

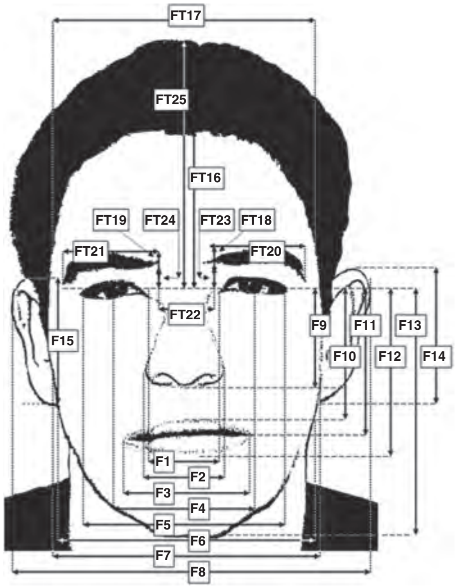
Figure 12.3 Locations of profile facial landmarks. Anatomical definitions and common names given in instructions to DigitalFace. Reprinted from *Forensic Science International*, **200**(1–3) Davis, J. P., Valentine, T. & Davis, R. E. Computer assisted photo-anthropometric analyses of full-face and profile facial images, 165–176. Copyright (2010), with permission from Elsevier.

between the measurements taken from each face in a proximity matrix. A squared Euclidean distance of zero is indicative of an exact match. A large distance indicates a high dissimilarity. The maximum value is dependent on the number of variables, cases and measurement variability. A simple decision rule was implemented. There were two criteria used to determine an identity match. The first was that the measures of two images of the *same* face (*probe* and *target*) should be closer in Euclidean space than the distance from the probe to any database *distracter*. The second, more rigorous criterion was that the distance in Euclidean space between two images of the *same* face should also be less than that between all other pairs of images of two *different* faces in the database.

With the inclusion of all frontal measures, all probes passed the primary criterion for a match to the corresponding target. However, two probes failed the secondary criterion, in that the Euclidian distance between two images of two different people was less than that between two images of the same person. A similar series of analyses was conducted in *profile* view, resulting in a similar conclusion. All probe images were correctly categorised on both the primary and secondary matching criteria, but only when all 62 *frontal and profile* measurements were included *together* in an analysis. These results show that one individual could not be reliably identified from a single image, such as that available on most single identity documents, although it should be more effective using multiple images collected from video line-ups (e.g. PROMAT, VIPER). These results support the conclusions of previous research (Mardia *et al.*, 1996; Kleinberg *et al.*, 2007; Roelofse *et al.*, 2008), illustrating that great caution should be taken when attempting to determine whether two different photographic images depict the *same* person. Some of the actors in the photographs that could not be reliably distinguished by *DigitalFace* had also been incorrectly identified as the same person in a simultaneous matching study using videos and with the actors present in person (Davis and Valentine, 2009). Therefore the investigations by Davis *et al.* (2010) simulated conditions that may occur in a forensic examination when identity is in dispute.

Expert witnesses are probably only asked to apply their techniques when images are impoverished in some manner, or if the appearance of the defendant has changed, for instance, by growing a beard. Indeed, under UK law, an expert should only be called to present evidence if a jury would be unlikely to be able to form an opinion without that assistance (*R v Turner*, 1975). With low-resolution or unclear images such as if the subject is sited some distance from the camera, features are obscured, or viewpoint is not matched, landmark identification would be more problematic, limiting the number of measurements and increasing error likelihood. Yet, cases have progressed in court with experts reporting on the use of far fewer measurements applied to images from a single viewpoint than those described by Davis *et al.* (2010).

There have been repeated calls for the establishment of large-scale databases of facial measurements in order to assess the safety of identification matching using facial mapping techniques. The database for the analyses reported by Davis *et al.* (2010) contained 70 faces from a homogeneous demographic. The results highlight the commonality of facial proportions. It could be argued that the database size was not sufficiently large for evaluation. However, the homogeneous inclusion criteria ensured that the distracter faces were highly representative of



Permanent horizontal distances			
F1	Inner eye distance (1–2)	F2	Nose width (9–10)
F3	Mouth width (11–12)	F4	Inter-pupil distance (3–4)
F5	Outer eye distance (5–6)	F6	Face width at upper lip (14–15)
F7	Face width at nostrils (17–18)	F8	Outer ear distance (7–8)
Permanent vertical distances			
F9	Eye line to nose base (1/2–19)	F10	Eye line to upper lip (1/2–13)
F11	Eye line to mouth (1/2–20)	F12	Eye line to lower lip (1/2–21)
F13	Eye line to chin (1/2–22)	F14	Right ear height (23–24)
F15	Left ear height (25–26)		
Transient distances			
FT16	Eye line to hairline (1/2–27)	FT17	Eyebrow face width (29–30)
FT18	Right eyebrow height (28–31)	FT19	Left eyebrow height (32–33)
FT20	Right eyebrow width (34–35)	FT21	Left eyebrow width (36–37)
FT22	Eyebrows distance (34–36)	FT23	Eye line to r eyebrow (1/2–31)
FT24	Eye line to 1 eyebrow (1/2–33)	FT25	Eye line to top of head (1/2–38)

Figure 12.4 Permanent and transient (marked with a T) distance measures produced by DigitalFace in full-face view. Reprinted from *Forensic Science International*, **200**(1–3) Davis, J. P., Valentine, T. & Davis, R. E. Computer assisted photo-anthropometric analyses of full-face and profile facial images, 165–176. Copyright (2010), with permission from Elsevier.

the test population. An increase in database size would probably result in more faces possessing similar facial dimensions, again increasing the potential for error in matching identity. Indeed, in an unpublished study (Clayton, 2008), the *DigitalFace* system was applied to the same set of 200 high-quality frontal facial images first described by Bruce *et al.* (1999) and used in a photo-anthropometric context by Kleinberg *et al.* (2007). Conducted in the Goldsmiths, University of London laboratory, discrimination of images of different people proved to be unreliable.

It would also be necessary to create further facial databases, if, for instance, the system was to be forensically applied to those of different ethnic backgrounds and age ranges or female targets. Bayesian statistics have recently been used to provide a measure of the likelihood that images depict the same face (Allen, 2008). However, the presentation of probability data in court is subject to potential misunderstanding (e.g. *R v Deen*, 1994; *R v Adams*, 1996). The same set of statistics can often be described in layman's terms in a variety of styles, and even minor nuances in delivery might influence the jury unduly.

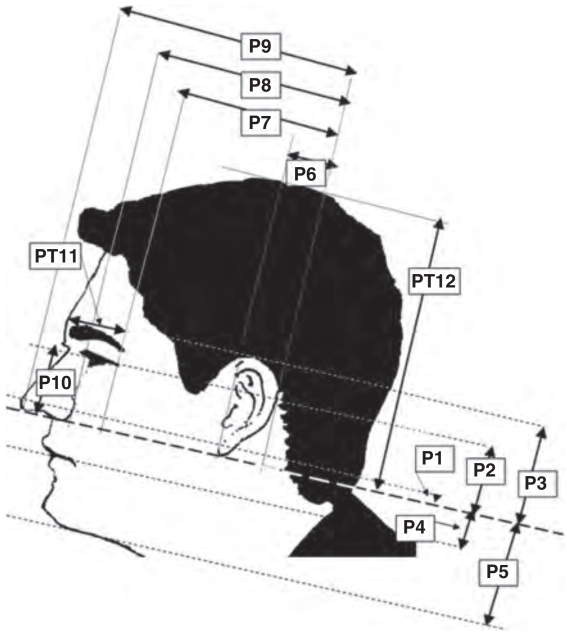


Figure 12.5 Permanent and transient (marked with a T) distance measures produced by DigitalFace in profile view. Reprinted from *Forensic Science International*, **200**(1–3) Davis, J. P., Valentine, T. & Davis, R. E. Computer assisted photo-anthropometric analyses of full-face and profile facial images, 165–176. Copyright (2010), with permission from Elsevier.

Permanent vertical distances			
P1	Nose/ear base line to nose tip (1/2–3)	P2	Nose/ear base line to eye (1/2–4)
P3	Nose/ear base line to top ear (1/2–5)	P4	Nose/ear base line to mouth (1/2–6)
P5	Nose/ear base line to chin (1/2–7)		
Permanent horizontal distances			
P6	Ear width (8–9)	P7	Rear of ear to outer eye (8–4)
P8	Rear of ear to nose rear (8–10)	P9	Rear of ear to nose tip (8–3)
P10	Nose height (3–11)		
Transient distances			
PT11	Eyebrow width (12–13)	PT12	Nose/ear base line to head (1/2–14)

12.3 Three-dimensional images

The recent development of equipment that can acquire three-dimensional (3D) images has led to suggestions that these could be used in forensic investigations in conjunction with both superimposition and photo-anthropometric techniques. For instance, Yoshino *et al.* (2000), using a 3D physiognomic range finder, demonstrated that a 2D extract can be accurately superimposed over a target image captured from a conventional camera. To ensure viewpoint equivalence, seven anthropometrical locations were marked on both images. Software automatically adjusted the 3D range finder image to match that of the 2D image by calculating the average perpendicular distance between each point. Yoshino's team (Yoshino *et al.*, 2002) calculated the reciprocal point-to-point differences against a database of 100 faces, in which novel disguised faces were entered as probes. The authors claimed a 100% identification rate, as the measured differences in two different images of the same person were always less than those of two different people. However, the faces included in the database appear to have been somewhat heterogeneous as the age range was 24–46 years. No details were given of perceived similarity, making it unclear whether any would be mistaken for another by human observers. Yoshino *et al.* (2002) suggest that 3D suspect images could be

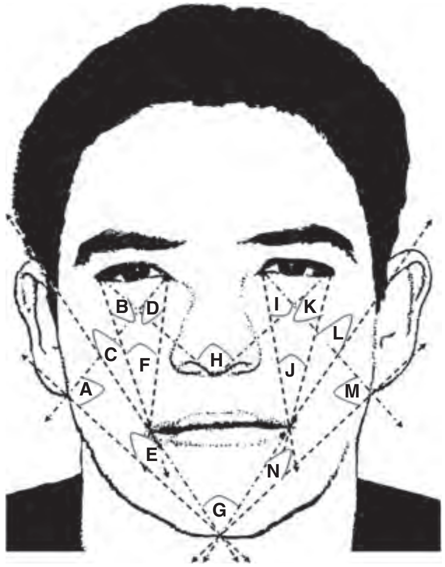


Figure 12.6 Full-face angular measurements computed by DigitalFace. Reprinted from *Forensic Science International*, **200**(1–3) Davis, J. P., Valentine, T. & Davis, R. E. Computer assisted photo-anthropometric analyses of full-face and profile facial images, 165–176. Copyright (2010), with permission from Elsevier.

Angle	Connecting lines				
A	2 (left inner eye)	-	26 (left ear base)	-	22 (chin)
B	12 (left mouth)	-	6 (left outer eye)	-	19 (nose base)
C	6 (left outer eye)	-	12 (left mouth)	-	25 (left ear top)
D	12 (left mouth)	-	2 (left inner eye)	-	26 (left ear base)
E	25 (left ear top)	-	22 (chin)	-	26 (left ear base)
F	2 (left inner eye)	-	12 (left mouth)	-	6 (left outer eye)
G	11 (right mouth)	-	22 (chin)	-	12 (left mouth)
H	5 (right outer eye)	-	19 (nose base)	-	6 (left outer eye)
I	11 (right mouth)	-	1 (right inner eye)	-	24 (right ear top)
J	1 (right inner eye)	-	11 (right mouth)	-	5 (right outer eye)
K	11 (right mouth)	-	5 (right outer eye)	-	19 (nose base)
L	5 (right outer eye)	-	11 (right mouth)	-	23 (right ear top)
M	1 (right inner eye)	-	24 (right ear base)	-	22 (chin)
N	11 (right mouth)	-	22 (chin)	-	24 (right ear base)

acquired in a similar manner to normal police mugshot photographs. The technique could then be routinely applied when security footage of an incident is obtained, by comparing the images to a 3D facial database.

Lynnerup *et al.* (2009) also studied the use of 3D laser scans for identification purposes. They recorded a 100% identification rate and a discriminatory factor of 86.7%. However, similar research conducted by Goos *et al.* (2006) using seven anthropometrical points to match a 3D laser scan to a 2D image, was less positive, being unable to categorise that a male and female volunteer were two different people. In addition, most 3D technologies suffer from image distortion from lighting anomalies and slight inadvertent body movements, as image acquisition can take several seconds (Schofield and Goodwin, 2004). Furthermore, capturing good-quality 3D images requires considerable skill, knowledge and time as well as subject cooperation. Currently available 3D scanners may be unsuitable for use in operational contexts.

12.4 Expert witnesses and the effect on jury decision-making

No published studies appear to have measured the impact of testimony from photographic comparison expert witnesses on jury decision-making. However,

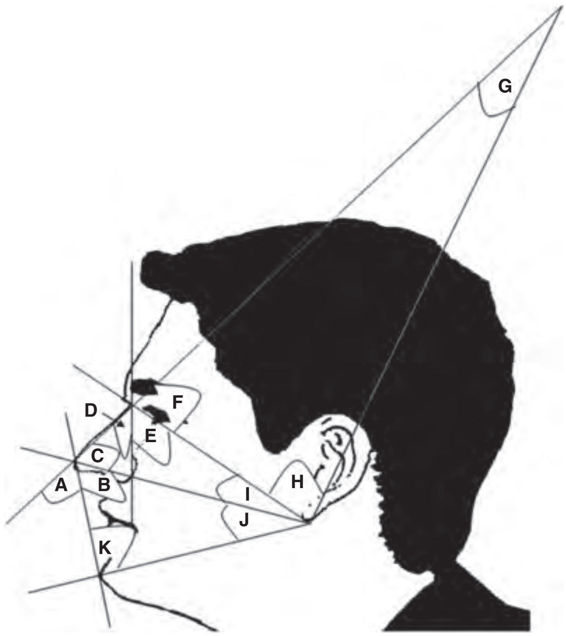


Figure 12.7 Profile angular measurements computed by DigitalFace. Reprinted from *Forensic Science International*, **200**(1–3) Davis, J. P., Valentine, T. & Davis, R. E. Computer assisted photo-anthropometric analyses of full-face and profile facial images, 165–176. Copyright (2010), with permission from Elsevier.

Angle	Connecting lines				
A	7 (chin)	-	3 (nose tip)	-	11 (top of nose)
B	2 (ear base)	-	3 (nose tip)	-	7 (chin)
C	2 (ear base)	-	3 (nose tip)	-	11 (top of nose)
D	3 (nose tip)	-	11 (top of nose)	-	10 (rear nose)
E	2 (ear base)	-	11 (top of nose)	-	6 (mouth corner)
F	2 (ear base)	-	11 (top of nose)	-	3 (nose tip)
G	3 (nose tip)	-	11 (top of nose)	-	2 (ear base)
H	4 (outer eye)	-	2 (ear base)	-	5 (top of ear)
I	3 (nose tip)	-	2 (ear base)	-	4 (outer eye)
J	3 (nose tip)	-	2 (ear base)	-	7 (chin)
K	2 (ear base)	-	7 (chin)	-	3 (nose tip)

research has been conducted on the influence of experts in eyewitness testimony (Hosch *et al.*, 1980; Cutler *et al.*, 1989). For instance, Hosch *et al.* (1980) found that participants given general information by an expert witness as to the *potential* unreliability of eyewitnesses 'lowered the importance of the eyewitness testimony' (p. 294), relative to other evidence. Although verdicts and jurors' opinions of the *credibility* of eyewitnesses were unaffected, the expert testimony caused the participants to scrutinise and discuss all evidence for longer. The authors argued that expert testimony was not a specific focus of attention during deliberations, but instead helped the participants to place appropriate weight on competing evidence. Cutler *et al.* (1989) also found that expert testimony increased the sensitivity of jurors to factors involved in eyewitness evidence without affecting belief in the accuracy of identifications.

In an unpublished study conducted at Goldsmiths College (Lacey, 2005), participants in groups played the part of jurors in assessing the guilt of a photographed 'defendant' shown simultaneously in video. Half the trials were target absent, in that someone with a close similarity of appearance to the defendant was shown in the video. For both types of trial, the belief in the guilt of the defendant after deliberation was lower when a report written by an expert witness was presented as evidence, compared with belief in guilt when no report was presented. These results are consistent with the findings of Hosch *et al.* (1980) and suggest that once participants become aware of the problems associated with making identifications they place less weight on that evidence.

12.5 Summary

The use of expert evidence to assist in the evaluation of facial identification evidence from CCTV footage can be useful in a criminal court, but the techniques require further evaluation. The reliability of any method of facial comparison involving low-quality images is questionable. Morphological classification analysis, by definition, involves grading facial features into pre-determined discrete categories, which may not be sufficiently flexible if a specific feature possesses elements of more than one category, or is on the boundary between two. Indeed, because nominal level analyses are required, it would be difficult to statistically discriminate between two different faces possessing similar characteristics (Vanezis *et al.*, 1996; Roelofse *et al.*, 2008). It has not been established that morphological analysis can distinguish reliably between unrelated people of the same age and ethnic background, especially from low-quality imagery. As such, support for a match cannot be objectively evaluated.

Experts in other identification fields have been shown to be susceptible to cognitive biases when provided with contextual information (Dror *et al.*, 2006; Dror and Rosenthal, 2008). In the UK, either the prosecution or the defence recruits experts. Although the opinions provided should be objective, it is inevitable that experts may be unconsciously influenced by the expectations of opponents in the adversarial system. Although the courts have ruled that expert evidence without objective measures should be admitted on a subjective basis only, it is hard to determine how much weight a jury may place on what might be interpreted as 'legalese'.

Currently there may be stronger safeguards in the USA than in the UK on the quality of forensic science methods (*Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 1993). The methodology used by an expert witness must have gained general acceptance in its particular academic discipline, to have been scientifically tested, and published in a peer-reviewed journal. The error rate (actual or potential) should be known (Groscup *et al.*, 2002). In the light of the work by Dror *et al.* (2006) on the influence of cognitive bias on fingerprint experts, a recent report by the National Academy of Sciences in the USA has called for investigation into the sources and rates of human error in forensic science. This report specially called for research on 'contextual bias' which occurs when the results of forensic analyses are influenced by an examiner's knowledge about the suspect's background or an investigator's knowledge of a case (National Academy of Sciences, 2009).

It is unclear at present if any facial comparison technique used in the English courts would meet the criteria in *Daubert*. A review of expert evidence in the UK has been ongoing since the autumn of 2005, mainly due to a number of medical cases in which scientific evidence was found to be questionable. It is possible that this review will recommend the adoption of similar criteria in the UK.

References

Alberink, I. and Bolck, A. (2008). Obtaining confidence intervals and likelihood ratios for body height estimations in images. *Forensic Science International*, **177**, 228–237.

Allen, R. J. (2008). Exact solutions to Bayesian and maximum likelihood problems in facial identification when population and error distributions are known. *Forensic Science International*, **179**, 211–218.

ACPO (Association of Chief Police Officers of England, Wales & Northern Ireland) (2003). National working practices in facial imagery. Available from: http://www.acpo.police.uk/asp/policies/data/garvin_facial_imaging_guidelines.doc (accessed 15 February 2006).

Attorney General's Reference No 2 of 2002 (2003) EWCA Crim 2373.

Ballytyne, J. W. (1984). The teralogical records of Chaldea. *Teratologica*, **1**, 127.

Bodic, F., Hamel, L., Lerouxel, E., Baslé, M. F. and Chappard, D. (2005). Bone loss and teeth. *Joint Bone Spine*, **72**, 215–221.

Bonner, L., Burton, A. M. and Bruce, V. (2003). Getting to know you: how we learn new faces. *Visual Cognition*, **10**, 527–536.

Bramble, S., Compton, D. and Klasen, L. (2001). Forensic image analysis: Version 2. Paper presented to 13th INTERPOL Forensic Science Symposium, Lyon, France.

British Standards Institute (BSI) (2005) Closed-circuit television (CCTV). Management and operation. Code of practice, BritBS 7958:2005. Available from: http://www.bsigroup.com/en/Shop/Publication-Detail/?pid=000000000030122793 (accessed 20 October 2006).

Bromby, M. C. (2003). At face value? *New Law Journal Expert Witness Supplement*, 28 February 2003, 302–303.

Brown, B. (1995). *CCTV in Town Centres: Three Case Studies*. Police Research Group Crime Detection and Prevention Series Paper: 68. London: HMSO.

Bruce, V. (1982). Changing faces: visual and non-visual coding processes in face recognition. *British Journal of Psychology*, **73**, 105–116.

Bruce, V., Henderson, Z., Greenwood, K. *et al.* (1999). Verification of face identities from images captured on video. *Journal of Experimental Psychology: Applied*, **5** (4), 339–360.

Bruce, V., Henderson, Z., Newman, C. and Burton, A. M. (2001). Matching identities of familiar and unfamiliar faces caught on CCTV images. *Journal of Experimental Psychology: Applied*, **7**(3), 207–218.

Bruce, V., Valentine, T. and Baddeley, A. (1987). The basis of the ¾ view advantage in face recognition. *Applied Cognitive Psychology*, **1**, 109–120.

Brunelli, R. and Poggio, T. (1993). Face recognition: features versus templates. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **15**(10), 1042–1052.

Burton, A. M., Bruce, V. and Dench, N. (1993). What's the difference between men and women: evidence from facial measurement. *Perception*, **22**(2), 153–176.

Burton, A. M., Miller, P., Bruce, V., Hancock, P. J. B. and Henderson, Z. (2001). Human and automatic face recognition: a comparison across image formats. *Vision Research*, **41**, 3185–3195.

Burton, A. M., Wilson, S., Cowan, M. and Bruce, V. (1999). Face recognition in poor-quality video: evidence from security surveillance. *Psychological Science*, **10**(3), 243–248.

Catterick, T. (1992). Facial measurements as an aid to recognition. *Forensic Science International*, **56**, 23–27.

Church v HMA (1996) SLT 383.

Clayton, L. (2008). *Matching facial identities*. Unpublished dissertation for the BSc (Hons) Psychology at Goldsmiths, University of London.

Clutterbuck, R. and Johnston, R. A. (2005). Demonstrating how unfamiliar faces become familiar using a face matching task. *European Journal of Cognitive Psychology*, **17**(1), 97–116.

Costigan, R. (2007). Identification from CCTV: the risk of injustice. *Criminal Law Review*, **August**, 591–608.

Cutler, B. L., Penrod, S. D. and Dexter, H. R. (1989). The eyewitness, the expert psychologist, and the law. *Law and Human Behavior*, **13**(3), 311–332.

Daubert v. Merrell Dow Pharmaceuticals, Inc (1993). 509 US 579 113 S Ct 2795.

Davies, G. and Thasen, S. (2000). Closed-circuit television: how effective an identification aid?. *British Journal of Psychology*, **91**, 411–426.

Davis, J. P. (2007). *The forensic identification of CCTV images of unfamiliar faces*. Unpublished thesis for the PhD Psychology at Goldsmiths, University of London.

Davis, J. P. and Valentine, T. (2009). CCTV on trial: matching video images with the defendant in the dock. *Applied Cognitive Psychology*, **23**, 482–505.

Davis, J. P., Valentine, T. and Davis, R. E. (2010). Computer assisted photo-anthropometric analysis of full-face and profile facial images. *Forensic Science International*, **200**, 165–176.

Dror, I. E., Charlton, D. and Peron, A. E. (2006). Contextual information renders experts vulnerable to making erroneous identifications. *Forensic Science International*, **156**, 74–78.

Dror, I. E. and Rosenthal, R. (2008). Meta-analytically quantifying the reliability and biasability of forensic experts. *Journal of Forensic Sciences*, **53**, 900–903.

Edmond, G., Biber, K., Kemo, R. and Porter, G. (2009). Law's looking glass: expert identification evidence derived from photographic and video images. *Current Issues in Criminal Justice*, **20**, 337–377.

Frye v United States (1923) 54 App DC 46.

Galton, F. (1978). Composite portraits. *Journal of the Anthropological Institute of Great Britain and Ireland*, **8**, 132–144.

Gill, M., Allen, J., Bryan, J. *et al.* (2005). *The impact of CCTV: Fourteen case studies*. Home Office, Research Development and Statistics Directorate Report 15/05. London: Home Office.

Gonzalez-Ulloa, M. and Flores, E. S. (1965). Senility of the face – basic study to understand its causes and effects. *Plastic and Reconstructive Surgery*, **36**, 239–246.

Goos, M. I. M., Alberink, I. B. and Ruifrok, A. C. C. (2006). 2D/3D image (facial) comparison using camera matching. *Forensic Science International*, **163**, 10–17.

Groscup, J. L., Penrod, S. D., Studebaker, C. A., Huss, M. T. and O'Neill, K. M. (2002). The effects of Daubert on the admissibility of expert testimony in state and federal cases. *Psychology Public Policy and Law*, **8**(4), 339–372.

Halberstein, R. A. (2001). The application of anthropometric indices in forensic photography: three case studies. *Journal of Forensic Science*, **46**, 1438–1441.

Harper, B. and Latto, R. (2001). Cyclopean vision, size estimation, and presence in orthostereoscopic images. *Presence*, **10**(3), 312–330.

Heisele, B., Ho, P., Wu, J. and Poggio, T. (2003). Face recognition: component-based versus global approaches. *Computer Vision and Image Understanding*, **91**(1–2), 6–21.

Henderson, Z., Bruce, V. and Burton, A. M. (2001). Matching the faces of robbers captured on video. *Applied Cognitive Psychology*, **15**, 445–464.

Hill, H. and Bruce, V. (1996). Effects of lighting on the perception of facial surfaces. *Journal of Experimental Psychology: Human Perception and Performance*, **22**(4), 986–1004.

Hill, H., Schyns, P. G. and Akamatsu, S. (1997). Information and viewpoint dependence in face recognition. *Cognition*, **62**(2), 201–222.

Hjelmas, E. and Low, B. K. (2001). Face detection: a survey. *Computer Vision and Image Understanding*, **83**(3), 236–274.

Hosch, H. M., Beck, E. L. and McIntyre, P. (1980). Influence of expert testimony regarding eyewitness accuracy on jury decisions. *Law and Human Behavior*, **4**(4), 287–296.

House of Lords: Science and Technology Committee 8th Report (1997/1998). *Digital Images as Evidence: Government Response*. Available from: http://www.publications.partliament.uk/pa/Id200102/Idinfo/reports.httm#8199798 (accessed 1 August 2009)

Introna, L. D. and Wood, D. (2004). Picturing algorithmic surveillance: the politics of facial recognition systems. *Surveillance and Society*, **2**(2/3), 177–198.

İşcan, M. Y. (1993). Introduction to techniques for photographic comparison: potentials and problems. In M. Y. İşcan and R. P. Helmer (Eds.), *Forensic Analysis of the Skull: Craniofacial Analysis, Reconstruction, and Identification*. New York: Wiley-Liss.

Khalil, A. K., Kubota, T., Tawara, A. and Inomata, H. (1996). Ultrastructural age-related changes on the posterior iris surface. A possible relationship to the pathogenesis of exfoliation. *Archives of Ophthalmology*, **114**(6), 721–725.

Kleinberg, K. F., Vanezis, P. and Burton, A. M. (2007). Failure of anthropometry as a facial identification technique using high-quality photographs. *Journal of Forensic Science*, **52**, 779–783.

Kumho Tire Co v Carmichael (1999) 526 US 137.

Lacey, A. (2005). *The influence of facial mapping expert testimony on jury decision making*. Unpublished dissertation for the BSc (Hons) Psychology at Goldsmiths, University of London.

Lee, J., Lee, E. Y., Tark, H. O., Hwang, J. W. and Yoon, D. Y. (2008). Efficient height measurement method of surveillance camera image. *Forensic Science International*, **177**, 17–23.

Lee, W. J., Wilkinson, C. M., Memon, A. and Houston, K. (2009). Matching unfamiliar faces from poor quality closed-circuit television (CCTV) footage: an evaluation of the effect of training on facial identification ability. *Axis*, **1**(1), 19–28. On-line journal of the Centre for Anatomy and Human Identification, University of Dundee.

Loftus, G. R. and Harley, E. M. (2004). How different spatial-frequency components contribute to visual information acquisition. *Journal of Experimental Psychology: Human Perception and Performance*, **30**, 104–118.

Loth, S. and İşcan, M. Y. (1994). Morphological indicators of skeletal aging: implications for paleodemography and paleogerontology. In D. E. Crews and R. M. Garruto (Eds.), *Biological Anthropology and Aging: Perspectives on Human Variation Over the Life Span*. New York: Oxford University Press.

Lynnerup, N., Clausen, M. L., Kristoffersen, A. M. and Steglich-Arnholm, H. (2009). Facial recognition and laser surface scan: a pilot study. *Forensic Science Medicine and Pathology*, **5**, 167–173.

McCahill, M. and Norris, C. (2003). Estimating the extent, sophistication and legality of CCTV in London. In M. Gill (Ed.), *CCTV*. Leicester: Perpetuity Press.

Mardia, K. V., Coombs, A., Kirkbride, J., Linney, A. and Bowie, J. L. (1996). On statistical problems with face identification from photographs. *Journal of Applied Statistics*, **23**(6), 655–675.

Mazumdar, T. and Sinha, P. (1989). Photographs of the human face and broken projective symmetry. *Journal of the Forensic Science Society*, **29**, 387–395.

National Academy of Sciences (2009). *Strengthening Forensic Science in the United States: A Path Forward*. Washington, DC: National Academies Press.

Neave, R. A. H. (1998). Age changes in the face in adulthood. In J. G. Clement and D. L. Ranson (Eds.), *Craniofacial Identification in Forensic Medicine*. Sydney: Arnold Publications.

Norris, C. and Armstrong, G. (1999). *The Maximum Surveillance Society: The Rise of CCTV*. Oxford: Berg.

Norris, C., McCahill, M. and Wood, D. (2004). Editorial. The growth of CCTV: a global perspective on the international diffusion of video surveillance in publicly accessible space. *Surveillance and Society*, **2**(2/3), 110–135.

Novick, N. L. (1988). *Super Skin*. New York: Clarkson Potter.

Orentreich, D. S. (1995). Skin. In G. L. Maddox (Ed.), *The Encyclopedia of Aging*, 2nd Edn. New York, NY: Springer.

Phillips, P. J., Scruggs, W. T., O'Toole, A. J. *et al.* (2007). *FRVT 2006 and ICE 2006: Large-scale Results*. National Institute of Standards and Technology Technical Report ISTIR 7408 available from http://iris.nist.gov (accessed 10 June 2009).

Porter, G. and Doran, G. (2000). An anatomical and photographic technique for forensic facial identification. *Forensic Science International*, **114**, 97–105.

R v Adams (1996) The Times 14 August 1996.

R v Clare and Peach (1995) 2 Cr App R 333.

R v Clarke (1995) 2 Cr App R 425.

R v Deen (1994) The Times 10 January 1994.

R v Dodson and Williams (1984) 79 Cr App R 220.

R v Fowden and White (1982) Crim LR 588.

R v Gardner (2004) EWCA Crim 1639.

R v Gray (2003) EWCA Crim 1001.

R v Hookway (1999) Crim LR 750.

R v Kerrigan (1998) Court of Appeal 11 June 1998.

R v Loveridge and others (2001) EWCA Crim 973.

R v Mitchell (2005) EWCA Crim 731.

R v Stockwell (1993) 97 Cr App R 260.

R v Tolson (1864) 4 F & F 103.

R v Turner (1975) 60 Cr App R 80 CA.

Roelofse, M. M., Steyn, M. and Becker, P. J. (2008). Photo identification: facial metrical and morphological features in South African males. *Forensic Science International*, **177**, 168–175.

Schofield, D. and Goodwin, L. (2004). *Error in 3D Capture of Facial Images*. Talk given at ICARIS London 29 October 2004.

Scientific Working Group on Imaging Technology (SWGIT) (2005). Best practices for forensic image analysis. *Forensic Science Communications*, **7**(4), 1–12.

Shaw, R. E., McIntyre, M. and Mace, W. (1974). The role of symmetry in event perception. In R. O. B. MacCleod and H. L. Pick Jr (Eds.), *Perception: Essays in Honour of James J. Gibson*. Ithaca, NY: Cornell University Press.

Sinha, P. (1996). Symmetry sensing through computer vision and a facial recognition system. *Forensic Science International*, **77**, 27–36.

Taister, M. A., Holliday, S. D. and Borrman, H. I. M. (2000). Comments on facial aging in law enforcement investigation. *Forensic Science Communications*, **2**, 2.

Takema, Y., Yorimoto, Y., Kawai, M. and Imokawa, G. (1994). Age-related changes in the elastic properties and thickness of human facial skin. *British Journal of Dermatology*, **131**(5), 641–648.

Vanezis, P. and Brierley, C. (1996). Facial image comparison of crime suspects using video superimposition. *Science and Justice*, **36**, 27–33.

Vanezis, P., Lu, D., Cockburn, J. *et al.* (1996). Morphological classification of facial features in adult Caucasian males based on an assessment of photographs. *Journal of Forensic Science*, **41**(5), 786–791.

Washington Post (8 October 2005). *Security Camera: New Witness*. Available from: http://www.washingtonpost.com/wp-yn/content/article2005/10/07/AR2005100701895 (accessed 10 October 2005).

Wilkinson, C. M. and Evans, R. (2009). Are facial image analysis experts any better than the general public at identifying individuals from CCTV images?. *Science and Justice*, **49**(3), 191–196. Corrigendum: 2011, doi: 10.1016/j.scijus.2011.07.001.

Yoshino, M., Matsuda, H., Kubota, S., Imaizumi, K. and Miyasaka, S. (2000). Computer-assisted facial image identification system using a 3D physiognomic range finder. *Forensic Science International*, **109**(3), 225–237.

Yoshino, M., Noguchi, K., Atsuchi, M. *et al.* (2002). Individual identification of disguised faces by morphometrical matching. *Forensic Science International*, **127**(1–2), 97–103.

Chapter 13

Three-dimensional facial imaging

Stephen Richmond, Alexei Zhurov and Arshed Toma

13.1 Introduction

There are many disciplines that are interested in characterising and measuring facial features and facial attractiveness. Several methods have been suggested to quantify facial shape and beauty, such as applying mathematical and geometric principles to facial morphology, for example, the golden proportion (1.618 and its reciprocal 0.618), Fibonacci series, averageness and symmetry (Ricketts, 1982; Rhodes, 2006). However, there is currently very little scientific evidence available to support the validity of these techniques. Overall facial attractiveness does not appear to depend on any single facial feature, although smiling and youthful facial appearance makes women look more attractive and facial attractiveness tends to decrease over time from ages 11 to 31 years. However, individuals tend to retain their relative levels of facial attractiveness throughout their lifespans (Tatarunaite *et al.*, 2005).

Subtle facial differences make individual faces unique. However, we are able to distinguish between individuals of different gender, age, ethnicity and race (Moss *et al.*, 1987; Kau *et al.*, 2006; Toma *et al.*, 2008; Richmond *et al.*, 2009). Facial recognition has been of prime interest in computer vision not only for security reasons but also identification of medical conditions (Zhao *et al.*, 2003). Recent studies have used static biometric methods to discriminate between normal and syndromic faces to a high level of sensitivity and specificity (e.g. Noonan syndrome) and to explore phenotype and genotype associations (Hammond *et al.*, 2005*a*, 2005*b*). Dynamic facial acquisition is a relatively new development and the relative speed of facial movement has been used to identify depression (Mergl *et al.*, 2005) and schizophrenia (Juckel *et al.*, 2008). There are a number of medical conditions that have been reported to have a significant influence on face shape, e.g. asthma (Mattar *et al.*, 2004) and Type I diabetes (El-Bialy *et al.*, 2000). Socioeconomic factors are associated with higher body mass index (Ellaway *et al.*, 1997), which in turn is related to both skeletal (Öhrn *et al.*, 2002; Sadeghianrizi *et al.*, 2005) as well as soft tissue facial shape (Ferrario *et al.*, 2004) and severity of malocclusion in adolescents (Frazao and Narvai, 2006). The growth of the craniofacial skeleton is closely related to the masticatory muscle complex with stronger biting forces associated with a smaller face height and weaker forces associated with long faces with varying thickness in the masseter muscle mass (Varrela, 1991; Varrela and Alanen, 1995; Charalampidou *et al.*, 2008). Sustained musical pursuits, such as violin/viola (Kovero *et al.*, 1997) and wind instrument playing, as well as opera singing (Brattström *et al.*, 1991), were found to have a significant effect resulting in a reduced face height. Facial width to height ratio has been reported as a predictor of aggressive behaviour (Carré and McCormick, 2008) and reduced facial expression has been associated with antisocial behaviour (Herpertz *et al.*, 2001).

13.2 Facial expressions

Facial expressions are important in non-verbal communication. Facial affect serves to communicate information about the nature of situations and the correct identification of facial affect in others is central to effective interpersonal interactions (Lancelot and Nowicki, 1997; Blair, 2003). Observing fear and/or sadness in others presents a strong cue that current behaviour is causing distress and, it is argued, inhibits antisocial behaviour (Blair, 2005). In human and primate studies distress cues inhibit aggressive behaviour (Preuschoft, 2000). It is therefore unsurprising that

Craniofacial Identification, ed. C. Wilkinson and C. Rynn. Published by Cambridge University Press.

not only do individuals with psychopathic traits demonstrate difficulty in correctly identifying facial affect, but the same is true of those who engage in antisocial behaviours generally (Walker and Leister, 1994; Woodbury-Smith *et al.*, 2005).

13.3 Genetic and environmental influences

The genetic and environmental influences on face shape are difficult to determine. However, most published studies report small convenient samples and there is a need for more objective tests, clear definition of conditions and unambiguous assessment criteria when investigating their effects on facial morphology; preferably, these criteria should be applied to a cohort of children followed from birth such as the Avon Longitudinal Study of Parents and Children (ALSPAC) (Golding *et al.*, 2001). These studies are currently being undertaken in birth cohorts exploring genotype/phenotype associations using genome-wide association studies (GWAS).

13.4 Measuring the face

The face can be measured in several ways, including (i) description of features, e.g. long/short/thin/wide; (ii) use of classification systems, e.g. Skeletal 1/2/3; (iii) visual analogue scales, e.g. 0–100 mm for facial attractiveness; (iv) direct clinical measurement, e.g. vernier calipers; (v) radiographic assessment, e.g. lateral skull radiograph; and more sophisticated techniques, such as (vi) three-dimensional (3D) imaging, e.g. surface capture and full-tissue depth capture (surface and internal tissues – computed tomography (CT), cone beam computed tomography (CBCT) and magnetic resonance imaging (MRI)) (Kau *et al.*, 2005).

Since the introduction of the cephalostat 80 years ago cephalometric analysis has become one of the main techniques in measuring soft and hard facial tissue features. Although the cephalogram provides some useful measurements it has many drawbacks in relation to the cost/maintenance of the equipment and radiation hazard, as well as projection and measurement errors (Baumrind and Frantz, 1971*a*, 1971*b*; Houston *et al.*, 1986). The image produced is a 2D projection that lacks a depth of field and the side of face furthest from the capture device has a greater magnification. The landmarks placed are usually identified as a feature of the anatomy: (i) the centre of a structure or space (e.g. Sella); (ii) juxtaposition (e.g. Nasion); (iii) deepest concavity/convexity (e.g. A and B points); (iv) extreme point – most anterior/posterior/superior/inferior (e.g. anterior nasal spine/posterior nasal spine). Unfortunately the definition of these landmarks is point based and subject to positioning/orientation of the head both during image capture and analysis, e.g. positioning of the external Frankfort plane in the cephalostat and that derived from the radiograph. There will be significant differences in point positions if a cephalogram is aligned on the Sella–Nasion line at Sella or aligned on the Frankfort plane anchored at Sella. In addition, the identified landmark point may not reflect the change seen in a discrete anatomical structure. Recent cone beam and MRI studies have the potential to overcome some of the drawbacks of landmark measures although such techniques should be restricted to major craniofacial anomalies due to the risks associated with radiation exposure. When assessing growth changes the commonly used superimposition is at Sella as the origin and the Sella–Nasion line, although we know that the posterior clinoid process moves posteriorly and Nasion anteriorly and superiorly up to the age of 18 years. There are other landmarks in the head such as the anterior cranial base, which is considered to be relatively stable from the age of 8 years. However, the anterior cranial base is not considered easy to identify and consequently not often used, although it is the most valid superimposition landmark.

If cephalometric analysis is not appropriate, alternative methods need to be sought. The use of 3D surface scanning techniques and mathematical methods (advanced methods of geometric morphometrics) are described below.

13.5 Three-dimensional imaging

Three-dimensional imaging has been used to analyse growth changes (Ferrario *et al.*, 1998; Nute and Moss, 2000), to create a database for normative populations (Yamada *et al.*, 2002; Božič *et al.*, 2009; Seager *et al.*, 2009), for assessment of clinical outcomes for orthodontic surgical procedures (McCance *et al.*, 1992; 1997*a*, 1997*b*, 1997*c*, 1997*d*; Ayoub *et al.*, 1996, 1998; Khambay *et al.*, 2002; Kau *et al.*, 2007) and for non-surgical treatments (Ismail *et al.*, 2002; Moss *et al.*, 2003). There are many 3D techniques available and these have been reported widely (Hajeer *et al.*, 2004*a*, 2004*b*; Kau *et al.*, 2005; Popat *et al.*, 2009). The main 3D techniques commonly used are surface scanning (laser and photogrammetry) and full-tissue depth (CT, CBCT and MRI).

13.6 Exploring facial characteristics

Three-dimensional imaging has the advantage of recording the facial features in the x–y, y–z and x–z planes with individual x, y and z coordinates, although the technique still has the same challenges in identifying a suitable origin and standardised reference planes (Figure 13.1). The face is aligned to the mid-sagittal and mid-intercanthal planes with the individual in natural head posture. The mid-intercanthal point would be considered as the 'origin'. The principle is that all faces should be aligned to a standardised framework enabling the analysis of facial differences with any individual at any time from any ethnic group as a part of a cross-sectional or longitudinal cohort study. To analyse facial differences 21 landmarks can be identified (Figure 13.2).

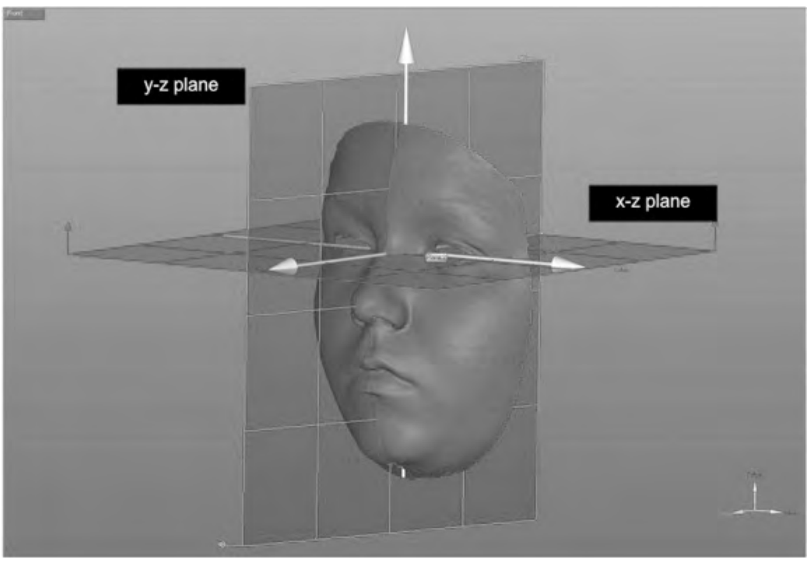
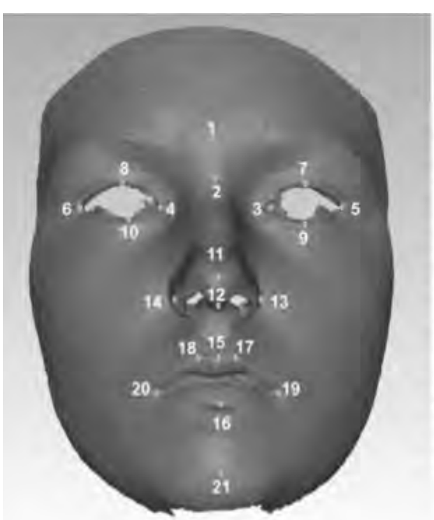


Figure 13.1 The construction of y–z and x–z planes relative to natural head posture. *Courtesy of Cardiff University.*



No	Abbrev.	Name
1	g	Glabella
2	n	Nasion
3	enL	Endocanthion: Left
4	enR	Endocanthion: Right
5	exL	Exocanthion: Left
6	exR	Exocanthion: Right
7	psL	Left palpebrale superius:
8	psR	Right palpebrale superiorus:
9	piL	Left palpebrale inferius:
10	piR	Right palpebrale inferius:
11	Prn	Pronasale:
12	Sn	Subnasale:
13	alL	Left alare:
14	alR	Right alare:
15	ls	Labiale superius:
16	li	Labiale inferius:
17	cphL	Left crista philtri
18	cphR	Right crista philtri
19	chL	Left cheilion
20	chR	Right cheilion
21	pg	Pogonion

Figure 13.2 Twenty-one facial landmarks.

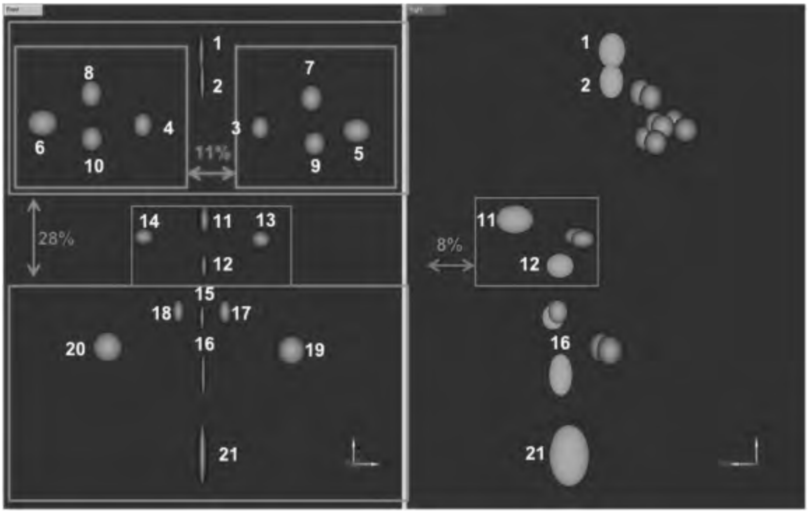


Figure 13.3a The three principal components are represented as clusters. Points 1–10 and 15–21 (28% of variance); points 3–10 (11% of the variance); and points 11–13 (8% of the variance).

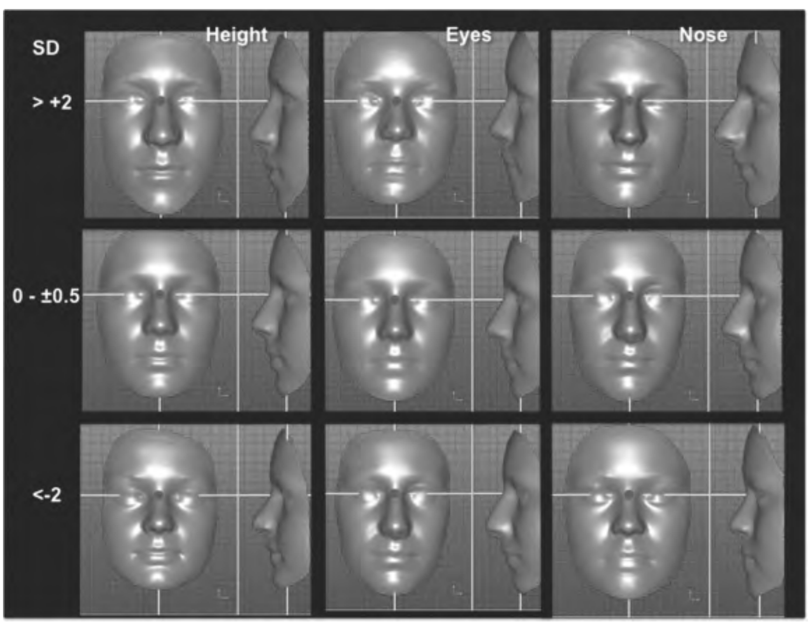


Figure 13.3b The influence of the principal components on face shape is illustrated (0 ± 0.5 and ± 2 standard deviations). First column: Long and short face; Second column: Eyes far apart and close together; Third column: Prominent nose. *Courtesy of Cardiff University.*

These landmarks were studied in a population of 350 15-year-old White children to look at the variation (Figure 13.3a). The reproducibility of these facial landmarks has been reported (Toma *et al.*, 2009). To aggregate all the facial landmarks from a population in a standardised framework it is necessary to align (translate and rotate) all the landmarks using Procrustes analysis (Bookstein, 1991). Once aligned the variation is demonstrated in 'ellipsoid-type' shapes representing 95% of the variance or two standard deviations from the mean. Different landmarks show different levels of variation in the three axes. For example, Pogonion (point 21) shows greater variation in the vertical direction compared with the horizontal. Pronasale (point 11) conversely illustrates greater variation in the horizontal plane compared with the vertical plane.

Three principal components of face shape were identified in the cohort of 15-year-old children (Richmond *et al.*, 2009) that explained 47% of the facial

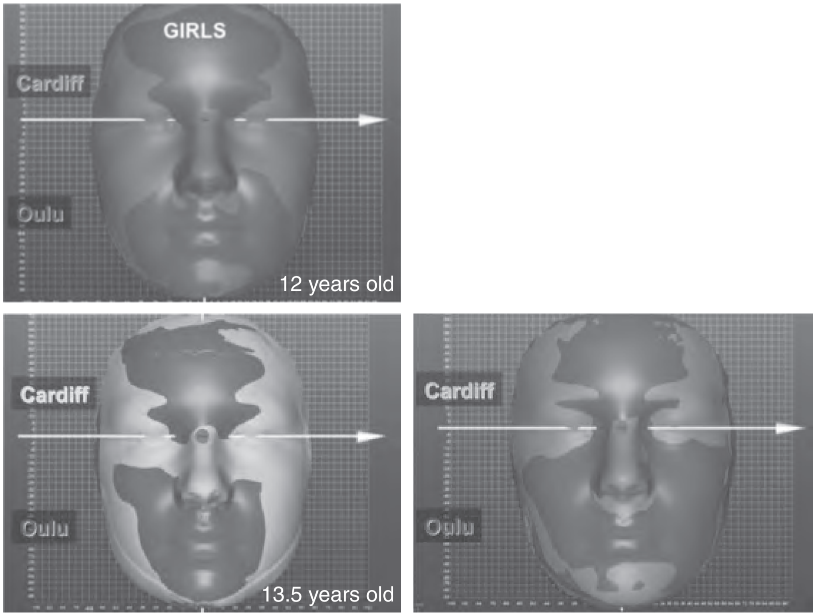


Figure 13.4 Comparison in faces for Finnish and Welsh males and females at ages 12 and 13.5 years old. Differences in the faces indicate that the Finnish faces (blue) tend to be more prominent in the upper and lower parts of the face for both age groups. *Courtesy of Cardiff University.* This figure is produced in colour in the plate section.

variance (Figure 13.3b). The first component involves the points of the face in two groups that move towards or away from each other in the Y direction. The first group (points 1–10) represents points around the eyes in addition to the nasion and glabella and the second group (points 15–21) involves points around the lips and chin. This component highlights the face height varying from short or long and explains 28% of the variance. The second component involves the distance between the eyes (points 3, 5, 7, 9 and 4, 6, 8, 10) with the effect being the eyes appearing to be closer or further apart. This component explains 11% of variance. The third component characterises the prominence of the nose (points 11–14) and explains 8% of the variance. Twenty-one landmarks can be identified but most imaging systems can record 10 000 to 50 000 facial points and these can be utilised to reconstruct the whole face (Figure 13.4) to illustrate the effect of these three components. The pattern of variation is similar for males and females for the 15-year-old age group but varies with age during the period from 12 to 16 years of age and with differing levels of sexual maturity.

13.7 Facial growth

Different ethnic groups exhibit slight variations in facial morphology (Božič *et al.*, 2009; Richmond *et al.*, 2009; Seager *et al.*, 2009). However, preliminary investigations of facial growth in different ethnic groups with similar facial morphology indicate that certain phenotypes may grow in similar ways (Richmond *et al.*, 2009). In a longitudinal study comparing the growth of facial morphologies of male and female children in Finland and Wales, the Finnish subjects demonstrated greater prominence of the face (Figure 13.5) and this prominence persisted during growth 18 months later. These studies are exploratory in nature and larger studies are required in different ethnic groups to characterise different facial morphologies at different age groups.

Two children (one Finnish and one Welsh) of the same age with similar facial morphologies were identified and followed for 3 years (Figure 13.6). It is not surprising that similar phenotypes grew in a similar way. Only the start and finish images are shown here but images were taken every 6 months and they also show similar patterns of movement throughout.

13.8 Comparison of three-dimensional facial shells

The difference in two facial shells can be compared essentially by applying a ‘best-fit’ algorithm. With this technique the computer program finds the relationship

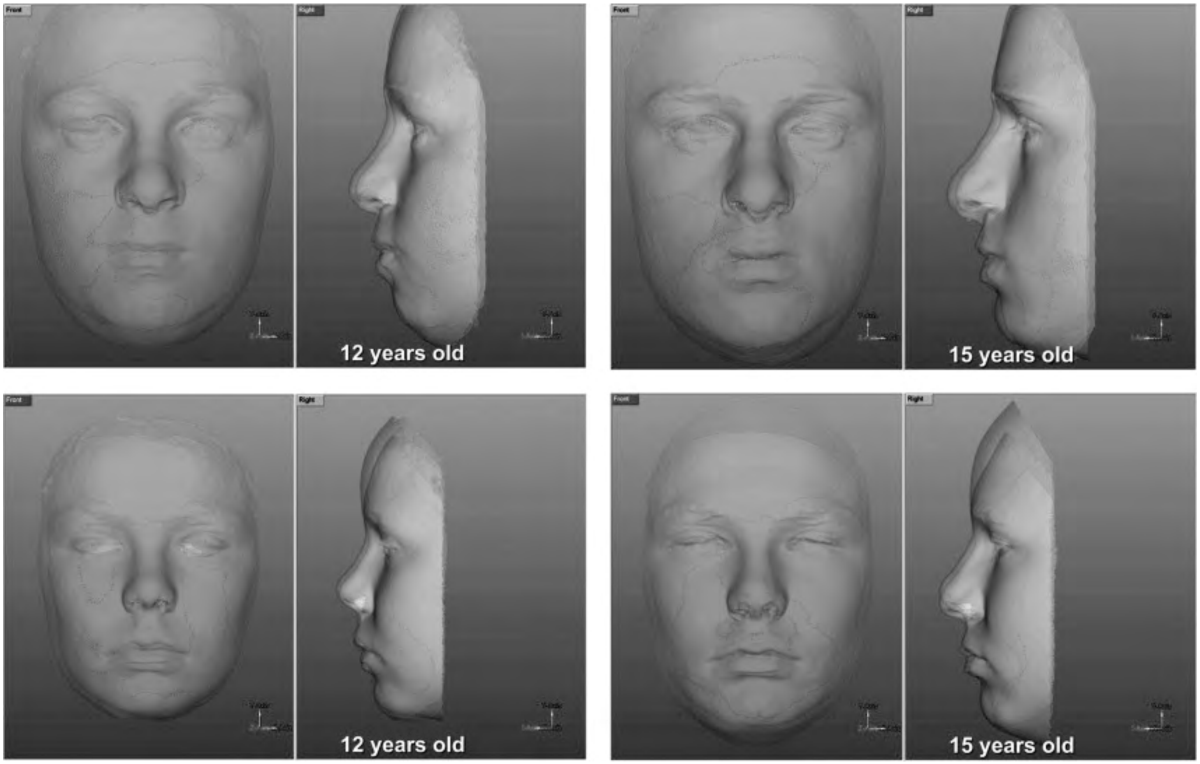


Figure 13.5 Growth of a male face between 12 and 15 years of age. Faces superimposed on a standardised framework with the mid-endocanthion as the ‘origin’ (Finnish – translucent and Welsh – white). A long face with forced lip competence is shown (top) and a short face (bottom). The different facial types (long and short face) have grown in a similar way but the long face has a different growth pattern to the short face. *Courtesy of Cardiff University.*

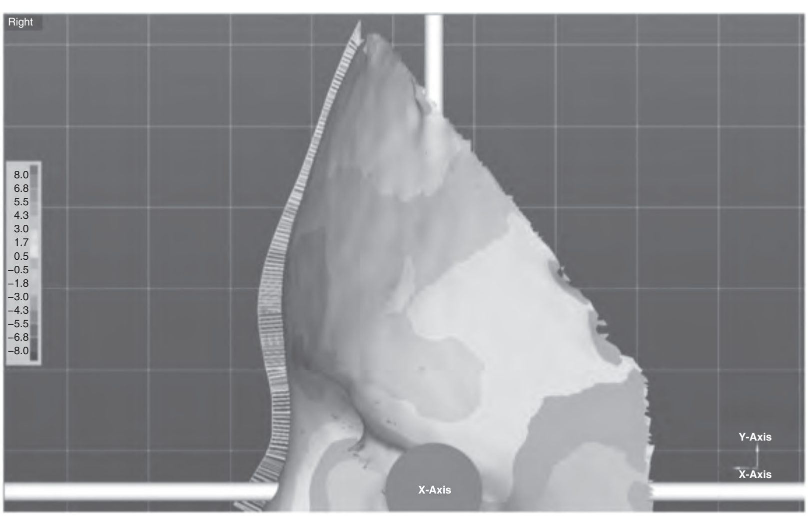


Figure 13.6 Colour deviation map with the scale on the left (–8 mm to +8 mm). The ‘warmer’ colours indicate positive changes and the ‘cooler’ colours indicate negative change. This figure is produced in colour in the plate section.

between the two faces that shows the least shell-to-shell deviation. There are many software suites available that contain these algorithms, such as Geomagic® and Rapidform®. The deviation between the shells is often displayed as a colour map. One shell is usually designated as the 'reference' shell and the shell to be compared with it the 'test' shell. Normals (lines at 90 degrees) are created from the reference shell and extended to the test shell. The distances along these normals are represented by a colour, e.g. green ± 0.5 mm and orange 3 to 4.3 mm etc. (Figure 13.7).

These colours can be used to compare any two facial shells. Colour maps are used to highlight the facial differences/changes in a cohort of male and female children at ages 12 and 17 years of age (Figure 13.8). The green represents a tolerance level of ±0.5 mm and the colour maps highlight the changes in females in which the face increases in width and height. The

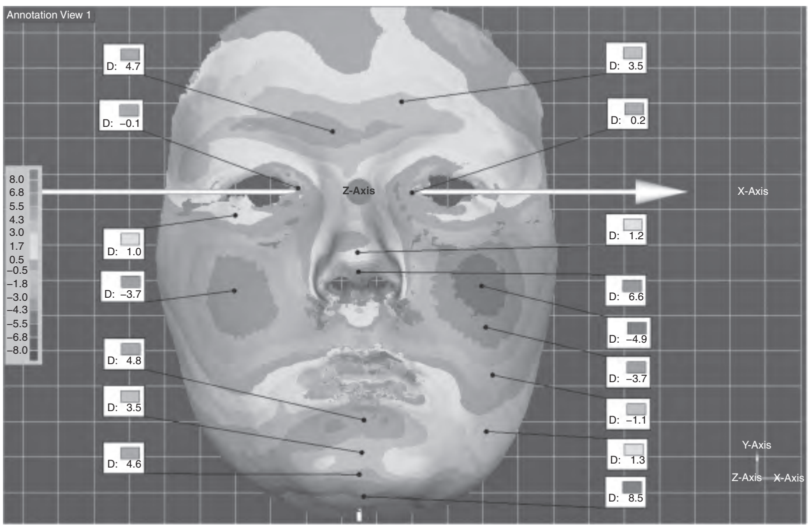


Figure 13.7 Colour map showing positive and negative changes for a male age 12 to 17 years of age (various deviation points on the face are highlighted) with the directional vectors recorded for the 21 landmarks (right). This figure is produced in colour in the plate section.

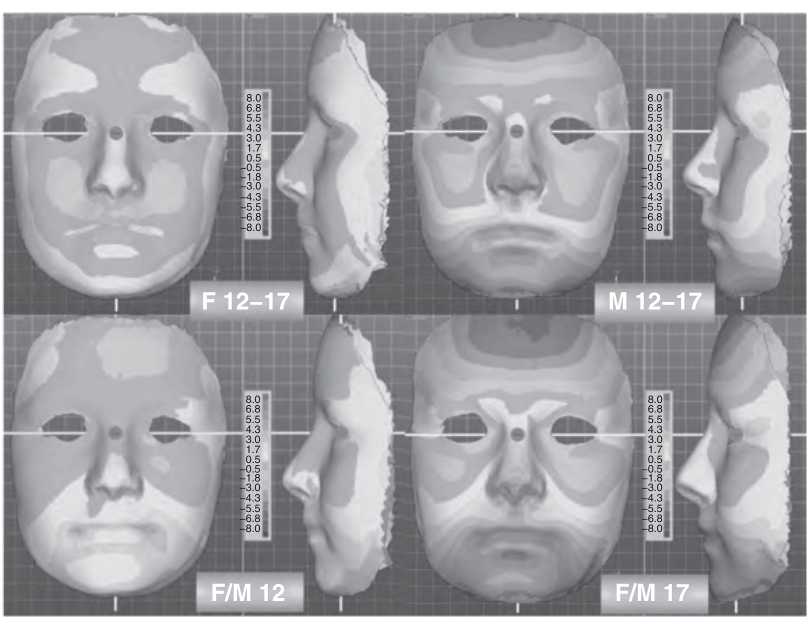


Figure 13.8 Colour maps illustrating the growth of females and males 12 to 17 years of age and differences between males and females at ages 12 and 17. *Courtesy Cardiff University.* This figure is produced in colour in the plate section.

males show greater changes with less prominence of the infra-orbital regions, greater prominence of the nose and chin. In addition there is a tendency in males to develop a backward sloping forehead. When males and females are compared there is progressive development of the nose and chin from the ages of 12 to 17 years.

13.9 Assessment of changes in facial landmarks

To compare detailed changes in landmarks for the same individual over time or different individuals a study of the directional vectors are required. The directional vector shows the change in direction and distance for a particular landmark. For example in the case illustrated (Figure 13.9), the pronasale has moved principally forward 5.3 mm and the pogonion has moved inferiorly 8.5 mm and forward 4.2 mm.

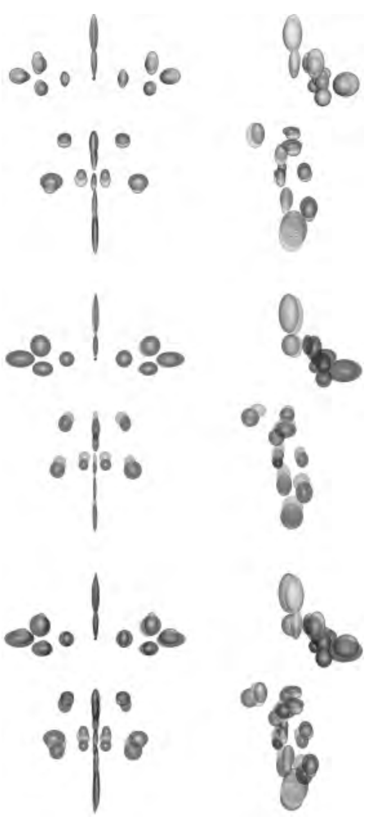


Figure 13.9 Growth of a cohort of 26 males and 23 females superimposed on a standardised framework with the mid-endocanthion point as the origin. Females (top) red – age 12 and yellow age 16. Males (middle) green – age 12 and blue – age 16. Both males and females superimposed at ages 12 and 16 (bottom). This figure is produced in colour in the plate section.

A cohort of 26 male and 23 female Welsh children were followed from the age of 12 to 16 years. The variation for the 21 landmarks is different between boys and girls and for age, highlighted by the ellipsoidal shapes in Figure 13.10. The males tend to show greater variation for all landmarks and a greater spatial change as a result of facial growth (males tend to exhibit more growth than females between the ages of 12 and 16). As the 'origin' is the mid-endocanthion there is more upward and forward growth of the nasion and glabella in boys, as well as greater forward and downward growth of the nose, mouth and chin. There is also considerable widening of the face in males, particularly of the exocanthion.

13.10 Planning facial surgery

The use of cone beam imaging is a valuable tool to evaluate the full tissue depth of an individual from the facial surface and underlying bony skeleton (Figure 13.11). Clinically these techniques are used to evaluate the relationship of the hard and soft tissues of the face to enable planning and prediction of orthognathic surgical outcomes. The individual illustrated presented with facial disharmony with a prominent chin and insufficient prominence of the mid-face. Advanced techniques have allowed the mock-up of the facial muscles and soft tissues (cartilage, fascia etc.) (Wilkinson *et al*., 2006) to enable a full biomechanical model that can be used to replicate facial movement (by activating the muscles of mastication and expressions) prior to surgery. Currently simple facial actions, for example mouth opening/closing, smiling and digust are possible. Further work will be required to develop a full range of facial actions and speech models. With this technique the outcome can be predicted to a high degree of accuracy as shown in the colour maps with the majority of the shell predictive and post-treatment falling within 0.5 mm (green).

13.11 Identifying muscle fibre orientation

Diffusion tensor magnetic resonance imaging data has been used to track nerve pathways in the brain (Kanaan *et al*., 2009). This technique can also be used to identify the fibre orientation of facial muscles (Figure 13.11). This presents the opportunity to develop a patient-specific model using the patient-derived volume and

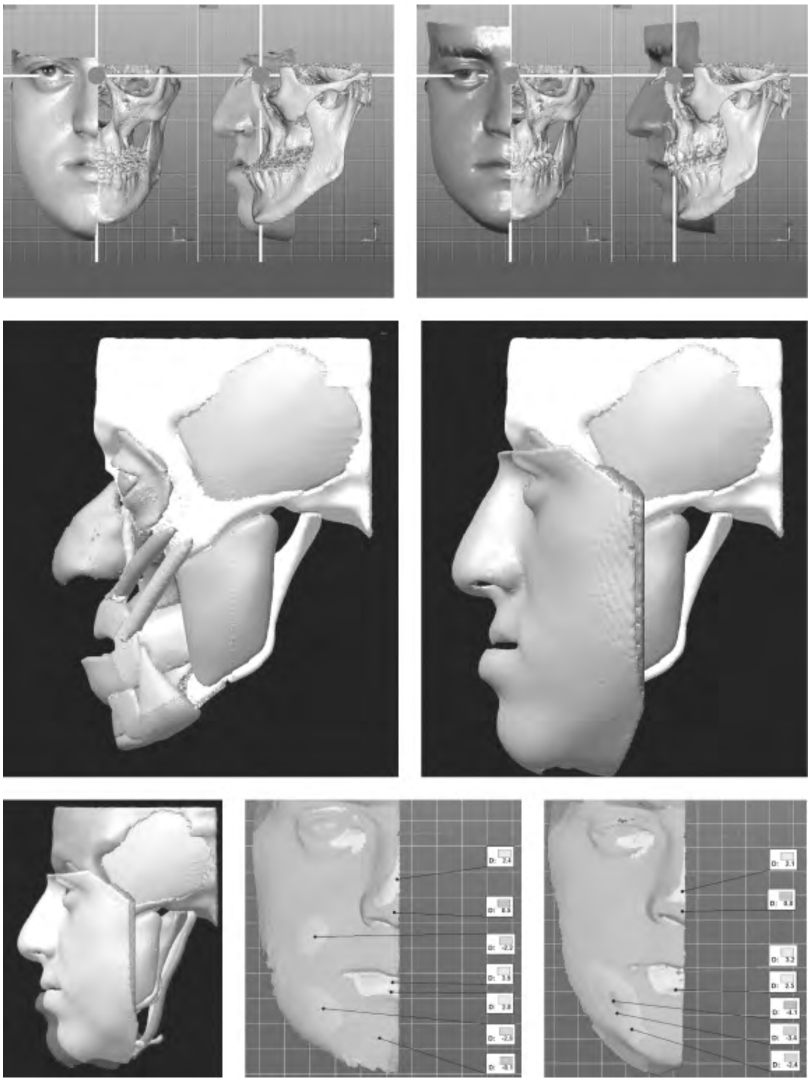


Figure 13.10 Surface laser scan superimposed on a cone beam CT. Top left (before surgery) highlights significant soft tissue in front of the maxilla and (top right) 6 months post-operative result. Muscles, cartilage, bone, subcutaneous tissues and skin are modelled enabling a functioning biomechanical head model that can currently mimic simple facial actions and expressions (middle). The difference between predicted and actual is shown (bottom left) for both 3 and 6 months (bottom middle and right) post-operatively with discrepancies at the side of the face and lower lip area (3–4 mm). *Courtesy of Cardiff University*.

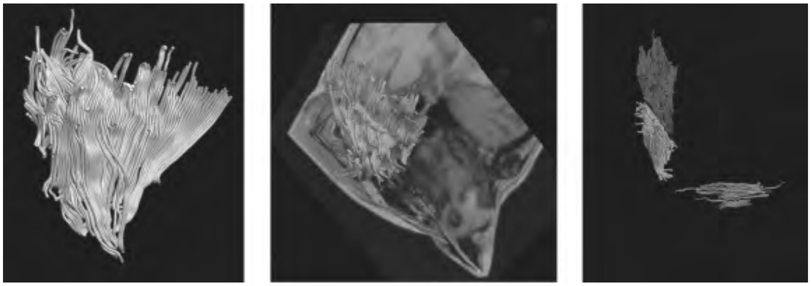


Figure 13.11 Orientation of temporalis muscle fibres isolated (left), spatial relationship of the temporalis (blue), masseter (green) and lip muscles (red) (right). *Courtesy of Cardiff University*. This figure is produced in colour in the plate section.

orientation of the facial muscles. Once all the muscles have been identified a biomechanical model can be employed to activate the muscles in the direction of the majority of the fibres to perform the various facial expressions and functions that make up daily activity.

13.12 Summary

There has been a significant advancement in 3D imaging techniques and applications. Many pilot and exploratory studies have been undertaken and the techniques developed will lead to the analysis of large cohort studies that will characterise facial morphology and determine the relative influences of genetic and environmental factors on face shape. Currently various features of the face are being characterised and their association with genes are being explored through Genome Wide Association Studies (GWAS). This will enable improved identification of individual facial features and prediction of growth/ageing. These developments will be of interest to a wide variety of disciplines and agencies.

References

Ayoub, A. F., Siebert, P., Moos, K. F. *et al.* (1998). A vision-based three-dimensional capture system for maxillofacial assessment and surgical planning. *British Journal of Oral and Maxillofacial Surgery*, **36**(5), 353–357.

Ayoub, A. F., Wray, D., Moos, K. F. *et al.* (1996). Three-dimensional modeling for modern diagnosis and planning in maxillofacial surgery. *International Journal of Adult Orthodontics and Orthognathic Surgery*, **11**(3), 225–233.

Baumrind, S. and Frantz, R. C. (1971*a*). The reliability of head film measurements. 1. Landmark identification. *American Journal of Orthodontics*, **60**(2), 111–127.

Baumrind, S. and Frantz, R. C. (1971*b*). The reliability of head film measurements. 2. Conventional angular and linear measures. *American Journal of Orthodontics*, **60**(5), 505–517.

Blair, R. J. (2003). Facial expressions, their communicatory functions and neuro-cognitive substrates. *Philosophical Transactions of the Royal Society London B Biological Sciences*, **29**(358(1431)), 561–572.

Blair, R. J. (2005). Responding to the emotions of others: dissociating forms of empathy through the study of typical and psychiatric populations. *Consciousness and Cognition*, **14**(4), 698–718.

Bookstein, F. L. (1991). *Morphometric Tools for Landmark Data Geometry and Biology*. Cambridge: Cambridge University Press.

Božič, M., Kau, C. H., Richmond, S. *et al.* (2009). Facial morphology of Slovenian and Welsh White populations using 3-Dimensional imaging. *Angle Orthodontist*, **79**(4), 640–645.

Brattström, V., Odenrick, L. and Leanderson, R. (1991). Dentofacial morphology in professional opera singers. *Acta Odontologica Scandinavica*, **49**(3), 147–151.

Carré, J. M., and McCormick, C. (2008). In your face: facial metrics predict aggressive behaviour in the laboratory and in varsity and professional hockey players. *Proceedings of the Royal Society B* doi:10.1098/rspb.2008.0873.

Charalampidou, M., Kjellberg, H., Georgiakaki, I. and Kiliaridis, S. (2008). Masseter muscle thickness and mechanical advantage in relation to vertical craniofacial morphology in children. *Acta Odontologica Scandinavica*, **66**(1), 23–30.

El-Bialy, T., Aboul-Azm, S. F. and El Sakhawy, M. (2000). Study of craniofacial morphology and skeletal maturation in juvenile diabetics (Type I). *American Journal of Orthodontics and Dentofacial Orthopedics*, **118**(2), 189–195.

Ellaway, A., Anderson, A. and MacIntyre, S. (1997). Does area of residence affect body size and shape? *International Journal of Obesity and Related Metabolic Disorders*, **21**(4), 304–308.

Ferrario, V. F., Dellavia, C., Tartaglia, G. M., Turci, M. and Sforza, C. (2004). Soft tissue facial morphology in obese adolescents: a three dimensional non-invasive assessment. *Angle Orthodontist*, **74**(1), 37–42.

Ferrario, V. F., Sforza, C., Schmitz, J. H., Miani, A. Jr and Serrao, G. (1998). A three-dimensional computerized mesh diagram analysis and its application in soft tissue facial morphometry. *American Journal of Orthodontics and Dentofacial Orthopedics*, **114**(4), 404–413.

Frazao, P. and Narvai, P. C. (2006). Socio-environmental factors associated with dental occlusion in adolescents. *American Journal of Orthodontics and Dentofacial Orthopedics*, **129**(6), 809–816.

Golding, J., Pembrey, M. and Jones, R. (2001). ALSPAC Study Team. ALSPAC – the Avon Longitudinal Study of Parents and Children. I. Study methodology. *Paediatric and Perinatal Epidemiology*, **15**(1), 74–87.

Hajeer, M. Y., Millett, D. T., Ayoub, A. F. and Siebert, J. P. (2004*a*). Applications of 3D imaging in orthodontics: part I. *Journal of Orthodontics*, **31**(1), 62–70.

Hajeer, M. Y., Millett, D. T., Ayoub, A. F. and Siebert, J. P. (2004*b*). Applications of 3D imaging in orthodontics: part II. *Journal of Orthodontics*, **31**(2), 154–162.

Hammond, P., Hutton, T. J., Allanson, J. E. *et al.* (2005*a*). Discriminating power of localized three-dimensional facial morphology. *American Journal of Human Genetics*, **77**, 999–1010.

Hammond, P., Hutton, T. J., Allanson, J. E. *et al.* (2005*b*). Discriminating power of localized three-dimensional facial morphology. *American Journal of Human Genetics*, **77**(6), 999–1010.

Herpertz, S. C., Werth, U., Lukas, G. *et al.* (2001). Emotion in criminal offenders with psychopathy and borderline personality disorder. *Archives in General Psychiatry*, **58**(8), 737–745.

Houston, W. J., Maher, R. E., McElroy, D. and Sherriff, M. (1986). Sources of error in measurements from cephalometric radiographs. *European Journal of Orthodontics*, **8**(3), 149–151.

Ismail, S. F., Moss, J. P. and Hennessy, R. (2002). Three-dimensional assessment of the effects of extraction and non-extraction orthodontic treatment on the face. *American Journal of Orthodontics and Dentofacial Orthopedics*, **121**(3), 244–256.

Juckel, G., Mergl, R., Präßl, A. *et al.* (2008). Kinematic analysis of facial behaviour in patients with schizophrenia under emotional stimulation by films with 'Mr. Bean'. *European Archives in Psychiatry and Clinical Neuroscience*, **258**, 186–191.

Kanaan, R. A. A., Borgwardt, S., McGuire, P. K. *et al.* (2009) Microstructural organization of cerebellar tracts in schizophrenia. *Biological Psychiatry*, **66**, 1067–1069.

Kau, C. H., Cronin, A. J. and Richmond, S. (2007). A three-dimensional evaluation of postoperative swelling following orthognathic surgery at 6 months. *Plastic and Reconstructive Surgery*, **119**(7), 2192–2199.

Kau, C. H., Richmond, S., Palomo, J. M. and Hans, M. G. (2005). Three-dimensional cone beam computerized tomography in orthodontics. *Journal of Orthodontics*, **32**(4), 282–293.

Kau, C. H., Zhurov, A., Richmond, S. *et al.* (2006). The 3-dimensional construction of the average 11-year-old child face: a clinical evaluation and application. *Journal of Oral and Maxillofacial Surgery*, **64**(7), 1086–1092.

Khambay, B., Nebel, J. C., Bowman, J. *et al.* (2002). 3D stereophotogrammetric image superimposition onto 3D CT scan images: the future of orthognathic surgery. A pilot study. *International Journal of Adult Orthodontics and Orthognathic Surgery*, **17**(4), 331–341.

Kovero, O., Könönen, M. and Pirinen, S. (1997). The effect of violin playing on the bony facial structures in adolescents. *European Journal of Orthodontics*, **19**(4), 369–375.

Lancelot, C. and Nowicki, S. Jr (1997). The association between receptive nonverbal processing abilities and internalizing/externalizing problems in girls and boys. *Journal of Genetics and Psychology*, **158**(3), 297–302.

Mattar, S. E., Anselmo-Lima, W. T., Valera, F. C. P. and Matsumoto, M. A. (2004). Skeletal and occlusal characteristics in mouth-breathing pre-school children. *Journal of Clinical Pediatric Dentistry*, **28**(4), 315–318.

McCance, A. M., Moss, J. P., Fright, W. R., Linney, A. D. and James, D. R. (1997*a*). Three-dimensional analysis techniques – Part 1. Three-dimensional soft-tissue analysis of 24 adult cleft palate patients following Le Fort I maxillary advancement: a preliminary report. *Cleft Palate–Craniofacial Journal*, **34**(1), 36–45.

McCance, A. M., Moss, J. P., Fright, W. R., Linney, A. D. and James, D. R. (1997*b*). Three-dimensional analysis techniques – Part 2. Laser scanning: a quantitative three-dimensional soft-tissue analysis using a color-coding system. *Cleft Palate–Craniofacial Journal*, **34**(1), 46–51.

McCance, A. M., Moss, J. P., Fright, W. R. and Linney, A. D. (1997*c*). Three-dimensional analysis techniques – Part 3. Color-coded system for three-dimensional measurement of bone and ratio of soft tissue to bone: the analysis. *Cleft Palate–Craniofacial Journal*, **34**(1), 52–57.

McCance, A. M., Moss, J. P., Fright, W. R. *et al.* (1997*d*). Three-dimensional analysis techniques – Part 4. Three-dimensional analysis of bone and soft tissue to bone ratio of movements in 24 cleft palate patients following Le Fort I osteotomy: a preliminary report. *Cleft Palate–Craniofacial Journal*, **34**(1), 58–62.

McCance, A. M., Moss, J. P. and James, D. R. (1992). Le Fort I maxillary osteotomy: is it possible to accurately produce planned pre-operative movements? *Britsh Journal of Oral Maxillofacial Surgery*, **30**(6), 369–376.

Mergl, R., Mavrogiorgou, P., Hegerl, U. and Juckel, G. (2005). Kinematical analysis of emotionally induced facial expressions: a novel tool to investigate hypomimia in patients suffering from depression. *Journal of Neurology, Neurosurgery and Psychiatry*, **76**, 138–140.

Moss, J. P., Ismail, S. F. and Hennessy, R. J. (2003). Three-dimensional assessment of treatment outcomes on the face. *Orthodontic and Craniofacial Research*, **6**(Suppl. 1), 126–131; discussion 179–182.

Moss, J. P., Linney, A. D., Grindrod, S. R., Arridge, S. R. and Clifton, J. S. (1987). Three-dimensional visualization of the face and skull using computerized tomography and laser scanning techniques. *European Journal of Orthodontics*, **9**(4), 247–253.

Nute, S. J., and Moss, J. P. (2000). Three-dimensional facial growth studied by optical surface scanning. *Journal of Orthodontics*, **27**(1), 31–38.

Öhrn, K., Al-Kahlili, B., Huggare, J. *et al.* (2002). Craniofacial morphology in obese adolescents. *Acta Odontologica Scandinavica*, **60**(4), 193–197.

Popat, H., Richmond, S., Benedikt, L., Marshall, D. and Rosin, P. L. (2009). Quantitative analysis of facial movement – a review of three-dimensional imaging techniques. *Computerized Medical Imaging and Graphics*, **33**(5), 377–383.

Preuschoft, S. (2000). Primate faces and facial expressions. *Sociological Research*, **67**, 245–271.

Rhodes, G. (2006). The evolutionary psychology of facial beauty. *Annual Review of Psychology*, **57**, 199–226.

Ricketts, R. M. (1982). The biologic significance of the divine proportion and Fibonacci series. *American Journal of Orthodontics*, **81**(5), 351–370.

Richmond, S., Toma, A. M. and Zhurov, A. I. (2009). [New perspectives on craniofacial growth.] *Orthodontie Française*, **80**(4), 359–369. [Article in French]

Sadeghianrizi, A., Forsberg, C. M., Marcus, C. and Dahllöf, G. (2005). Craniofacial development in obese adolescents. *European Journal of Orthodontics*, **27**(6), 550–555.

Seager, D. C., Kau, C. H., English, J. D. *et al.* (2009). Facial morphologies of an adult Egyptian population and an adult Houstonian white population compared using 3D imaging. *Angle Orthodontist*, **79**, 991–999.

Tatarunaite, E., Playle, R., Hood, K., Shaw, W. and Richmond, S. (2005). Facial attractiveness: a longitudinal study. *American Journal of Orthodontics and Dentofacial Orthopedics*, **127**(6), 676–682; quiz 755.

Toma, A. M., Zhurov, A., Playle, R. and Richmond, S. (2008). A three-dimensional look for facial differences between males and females in a British-Caucasian sample aged 15 1/2 years old. *Orthodontic and Craniofacial Research*, **11**(3), 180–185.

Toma, A. M., Zhurov, A., Playle, R., Ong, E. and Richmond, S. (2009). Reproducibility of facial soft tissue landmarks on 3D laser-scanned facial images. *Orthodontic and Craniofacial Research*, **12**(1), 33–42.

Varrela, J. (1991). Genetic and epigenetic regulation of craniofacial development. *Proceedings of the Finnish Dental Society*, **87**(2), 239–244.

Varrela, J. and Alanen, P. (1995). Prevention and early treatment in orthodontics: a perspective. *Journal of Dental Research*, **74**(8), 1436–1438.

Walker, D. W., and Leister, C. (1994). Recognition of facial affect cues by adolescents with emotional and behavioral disorders. *Behavioral Disorders*, **19**, 269–276.

Wilkinson, C. M., Needham, C. D., Rynn, C., Richmond, S. and Kau, C. H. (2006). A blind accuracy assessment of computer modelled facial reconstruction using computer tomography data from live subjects. *Forensic Science, Medicine and Pathology*, **2**(3), 179–187.

Woodbury-Smith, M. R., Robinson, J., Wheelwright, S. and Baron-Cohen, S. (2005). Screening adults for Asperger syndrome using the AQ: a preliminary study of its diagnostic validity in clinical practice. *Journal of Autism and Developmental Disorders*, **35**, 331–335.

Yamada, T., Mori, Y., Minami, K., Mishima, K. and Sugahara, T. (2002). Three-dimensional facial morphology, following primary cleft lip repair using the triangular flap with or without rotation advancement. *Journal of Craniomaxillofacial Surgery*, **30**(6), 337–342.

Zhao, W., Chellappa, R., Phillips, P. J. and Rosenfeld, A. (2003). Face recognition: a literature survey. *ACM Computing Surveys*, **35**(4), 399–458.

Chapter 14

Post-mortem prediction of facial appearance

Caroline Wilkinson and Amy Tillotson

14.1 Introduction

Forensic facial depiction is the interpretation of human remains in order to attempt to depict the living appearance of an individual (Gerasimov, 1971; Prag and Neave, 1997; Taylor, 2001; Wilkinson, 2004). It can be a powerful tool that significantly enhances the chances of identification of the deceased. Following major natural disasters, such as the tsunami of 26 December 2004 and Hurricane Katrina of August 2005, human remains may be extremely difficult to recognise due to decomposition or environmental effects; clothing and personal items may be lost and dental records unavailable (Thompson and Black, 2006). Even where facial preservation is sufficient for recognition by a family member to be attempted, the emotional circumstances result in many examples of false recognition (Hill, 2006). Ten per cent of victims of the tsunami and 50% of victims of the Bali bombing of 12 October 2002 were wrongly identified by facial recognition (Lain *et al.*, 2003). The social, legal and religious implications of misidentification are enormous. International investigative authorities advocate that it is vital to identify the deceased to allow remains to be returned to their families for proper recognition and religious observance, for grieving and acceptance of death and for judicial matters of estate (Lain *et al.*, 2003). With mass disasters such as these, the usual accepted methods of identification are often inappropriate and the importance of unusual and less definitive methods of identification has been recognised. The techniques of craniofacial reconstruction or post-mortem depiction can help resolve many stalemates within identification investigations (Wilkinson *et al.*, 2006).

The ultimate aim of post-mortem prediction is to recreate an in vivo countenance of an individual that sufficiently resembles the decedent to allow recognition (Prag and Neave, 1997). In forensic situations when combined with a publicity campaign, facial depiction may lead to recognition by a member of the public, and hence lead to the identification of that individual. It is currently very difficult to predict facial appearance from partially decomposed remains with any reliability, especially where analysis of the skull is not possible (where invasive assessment or radiographic analysis is inappropriate).

14.2 Post-mortem processes

Post-mortem changes to human remains are relatively well documented in forensic pathology and forensic anthropology literature, with regard to early decomposition and the accepted timeframe for these changes to occur depending upon the environmental and climatic factors involved. Differences in post-mortem changes have been observed between arid/desert like environments (Galloway *et al.*, 1989; Galloway, 1997), water/submersion environments (Haglund, 1993; Mellen *et al*, 1993; Rodriguez, 1997; Haglund and Sorg, 2002), burial environments (Berryman *et al.*, 1997; Manhein, 1997; Rodriguez, 1997; Brothwell and Gill-Robinson, 2002), cold climates (Micozzi, 1986; Komar, 1998), warm climates (Bass, 1997; Prieto *et al.*, 2004) and European and North American environments (Haglund *et al.*, 1989; Bass and Meadows, 1990; Bass, 1997; Prieto *et al.*, 2004). Although differing rates of decomposition (and some specific post-mortem changes) have been recorded in varied environments and climates, a general outline of the process of post-mortem changes can be observed. The following observations can be seen in the soft tissues in early post-mortem stages (within 2 hours of death):

Craniofacial Identification, ed. C. Wilkinson and C. Rynn. Published by Cambridge University Press.

- Skin pallor due to the cessation of the blood flow caused by the lack of cardiac activity (15 to 20 minutes).
- Flaccidity of the soft tissues due to initial post-mortem loss of muscle tone.
- Purging of stomach contents through the nose and mouth if the body is moved in the early stages after death (purging of stomach contents can also occur at a later stage due to expulsion of decomposition gases from the stomach) (Clark *et al.*, 1997).
- External and internal changes to eye colour, including: pooling of blood in the capillaries of the fundi, appearance of a thin film (minutes after death) turning to a cloudiness at the corneal surface (hours after death) and, where the eyes are open and exposed to a dry environment, the exposed eyeball will show a darkened appearance known as tache-noire de sclerotic pattern (Perper, 1980).

Following the onset of these early post-mortem changes, the body will undergo the following processes (2–4 hours after death):

- Algor mortis: the process of post-mortem body temperature decline to the ambient air temperature. As the body cools, the soft tissue fat content solidifies and will assume a shape based on external pressures, resulting in distortions relating to body position (Dix and Graham, 2000). For example, where the body lies face down, the pressure against the floor will distort the soft tissues of the face on the resting side.
- Rigor mortis: the post-mortem process of increased rigidity of muscles and joints following the initial period of flaccidity. Each muscle develops rigor at a similar rate, however smaller volume muscles, such as those of facial expression, will begin stiffening before larger volume muscles. This allows the muscles of the face, hands and feet to become rigid before the limb and trunk muscles (Perper, 1980). Rigor mortis is also responsible for the appearance of goose-flesh (anserine cutis), caused by the muscle fibres of the hair follicles (erector pilae) in the skin (Perper, 1980).
- Livor mortis (or post-mortem hypostasis): the process of pooling of blood due to gravitational forces within capillaries of the skin causing a mottled, purple and blue discolouration to the skin surface. This occurs only in areas with uncompressed blood vessels and will not appear in areas of skin in contact with the physical environment (pressure points). The resulting discolouration to the skin is known as lividity and can be useful in determining whether a body has been moved in the hours following death (Perper, 1980; Clark *et al.*, 1997).
- Bodies recovered from water may exhibit epidermal hydration and skin maceration (‘washerwoman’s hands’), manifesting as blanched, swollen and wrinkled skin.

The next stage in the post-mortem process is decomposition. This is the term describing the disintegration/dissolution of body tissues after death (Perper, 1980) and various stages have been described in the literature (Perper, 1980; Clark *et al.*, 1997; Bass, 1997; Manhein, 1997). The process of whole body decomposition, although affected by many environmental and climatic factors, tends to proceed in a predictable pattern, with any variation being limited to speed of decomposition (Clark *et al.*, 1997). The process of decomposition can be subdivided into autolysis and putrefaction.

Autolysis is the process of cellular breakdown caused by the digestive action of hydrolytic enzymes released into the cytoplasm. This process is thought to be triggered by the decrease in pH levels in the body after death (Clark *et al.*, 1997). Autolysis is responsible for a number of post-mortem changes including skin slippage, post-mortem bullae and marbling. Skin slippage results from the hydrolytic action of autolysis enzymes, which loosen the epidermal and dermal layers, causing the outer layer of the skin to slough away. In areas where the skin is separating, fluid can often accumulate as blisters or post-mortem bullae. Hair and nails may also loosen during the process of skin slippage and can become dislodged and lost (Clark *et al.*, 1997) in a process known as ‘de-gloving’. Marbling is the bluish appearance of the superficial blood vessels through the skin caused by the rupture (lysis) of red blood cells (intravascular haemolysis).

Putrefaction occurs due to the rapid overproduction of bacteria and fungi within the body following the cessation of homeostasis (Perper, 1980). Putrefaction is increased by the assimilation of the by-products of autolysis by bacterial and fungal micro-organisms, resulting in the production of malodorous gases and other by-products (Clark *et al.*, 1997). There are a number of fundamental changes observed during putrefaction, including the discolouration of tissues, dehydration of areas of the skin, swelling and bloating and destruction of the tissues.

The process of decomposition has four putrefactive and autolytic stages, which are described briefly below and in more detail in Table 14.1.

- **Putrid**: includes production of odour and distinct colour changes to the tissues, the end of rigor and fixation of lividity, initial production of gases and skin slippage (stages I–III).
- **Bloating**: includes further discolouration of tissues, marbling and further skin slippage, swelling across the whole body (stages IV–VI).
- **Destruction and saponification**: release of gases and destruction of soft tissue, production of adipocere or onset of mummification (stages VII–VIII).
- **Skeletonisation or mummification**: completion of destruction of soft tissues or completion of mummification (stages IX–X).

The process of decomposition follows the sequence seen in Table 14.1 with some deviations depending upon the environment and climate.

There is a paucity of research into the timeline for post-mortem changes to the face. In the human face the first post-mortem observations will be those associated with pallor and flaccidity then rigor, livor and algor mortis (within hours of death). Initial flaccidity of the soft tissues will cause the jaw to become slack and the mouth to hang open, due to loss of tone in the muscles surrounding the mouth fissure and those supporting the temporomandibular joint. Rigor mortis will then affect the facial muscles producing a tightened expression (Taylor, 2001) and giving a gooseflesh appearance to areas where the skin has small hairs (Clark *et al*., 1997). Rigor mortis will also affect the size of the iris by causing the muscles that control the pupil to contract, sometimes causing an artefactual difference in pupil sizes (Perper, 1980). When rigor mortis begins to wane the muscles will take on a secondary flaccidity and the jaw will again become slack (Perper, 1980). It has been observed by some practitioners (Taylor, 2001) that gravity and rigor mortis may play a part in changing the normal appearance of the eyes, with the lateral corners appearing noticeably raised.

The face will be affected by livor mortis when the body is found in a prone position and also in cases of intense congestion (i.e. congestive heart failure and

Table 14.1 Decomposition scale (adapted from Clark *et al*., 1997).

Category	Stage	Changes
Putrid	I	Early putrid odour
		Lividity fixed
		Rigor waning
		Tissues tacky
	II	Green discolouration of abdomen
		Haemolysis
		Intense livor
		No rigor
		Early skin slippage
		Drying of nose, lips and fingers
	III	Tissue gas on X-rays
		Prominent haemolysis
		Tissues soft and slick
		Skin slips easily
Bloating	IV	Early body swelling
		Discolouration of head
		No discolouration of trunk
		Gas in heart
		Marbling
		Bullae
	V	Moderate swelling
		Discolouration of head and trunk
	VI	Maximal body swelling
Destruction	VII	Release of gases
		Exhausted putrefied soft tissues
		Total destruction of blood
	VIII	Partially skeletonised
		Adipocere
		Mummification
Skeleton	IX	Skeleton with ligaments
	X	Skeleton with no soft tissues

suspension upside down) (Dix and Graham, 2000). It should also be noted that livor mortis can be difficult to observe in skin with dark pigmentation (Perper, 1980). The eyes will become darker in colour due to the settling of blood in the fundi of the eyeballs. In addition, blood will drain away from the choroidal vessels leaving an undulating dark and light background. Bodies lying in a prone position may exhibit a haemorrhagic appearance to the eyes due to gravitational blood accumulation (Perper, 1980; Dix and Graham, 2000; Curcio, 2005). Facial distortions caused by the solidification of fat during algor mortis will only be observed where the face is resting on a surface (Dix and Graham, 2000). This can be seen when the body is lying in a prone position or where the face is covered. The weight of certain covering materials, such as body bags used for recovery and other wrappings used to dump bodies, can cause distortion to the position of the nose (Wilkinson, 2004) and it is not uncommon for the print of the covering material (i.e. a heavy weave material, string patterns etc.) to be imprinted in the skin of the cheek or forehead of a victim (Berryman *et al.*, 1997). The facial tissues can also become distorted in bodies found submerged in water, due to the water pressure or action of the physical environment (Taylor, 2001).

Other early eye colour changes may also be observed. The iris colour can be altered due to changes to the cornea (a thin, clear film that turns cloudy), so that a blue iris can appear brown as a consequence of corneal clouding (Perper, 1980; Burton, 2007). Internally, retinal haemorrhage may occur frequently along with retinal oedema, leading to an increased opacity of tissues (Curcio, 2005). The sclera of the eyeball, if exposed to the air (i.e. when eyes are open) can also dehydrate producing a darkened appearance (Perper, 1980).

The decomposition process can affect the overall skin colour appearance and the following can be readily observed in the face immediately following death. Skin pallor occurs commonly 15–30 minutes after death depending upon the environment (Clark *et al.*, 1997) due to cessation of blood flow associated with cardiac arrest. A little later in the decomposition process the skin can show discolouration with light skin pigment appearing darker (Perper, 1980; Burton, 2007). Conversely, dark pigments in the skin can also be lost in certain environments (Perper, 1980; O'Brien and Kuehner, 2007). As decomposition progresses, a greenish/purplish discolouration will be observed at the face (Perper, 1980). A grey to green discolouration that becomes brown may be seen at the nose and ears in hot and dry environments (Galloway, 1997). Purging of fluids from the mouth and nose may also change skin colour at the nasal and oral area (Perper, 1980) and skin slippage will alter skin colour (initially pink and shiny followed by a matt yellow as the surface becomes dehydrated), making any post-mortem assessment of skin colour unreliable (Taylor, 2001). Skin slippage will also affect the position of the hair (Pounder, 1995; Taylor, 2001) and in hot environments, head hair may slough off as a hair mass (Mann *et al.*, 1990).

Facial hair and skin will remain intact in certain post-mortem environments, such as cases of mummification in either a dry, arid environment or in bodies preserved in water with an airtight seal, e.g. bog bodies (Slezdik and Micozzi, 1997; Brothwell and Gill-Robinson, 2002). A preserved bog body may exhibit modified hair colour due to pH changes creating lighter, reddish colours (Brothwell and Gill-Robinson, 2002) and the skin takes on a leathery appearance (Bass, 1997; Manhein, 1997; Slezdik and Micozzi, 1997). It is commonly observed in mummification that dehydration occurs in the skin around the lips (making it more difficult to observe the teeth) and at the nasal tip (Pounder, 1995; Clark *et al.*, 1997; Taylor, 2001). It is useful to note that putrefactive processes in hot and humid environments may cause the teeth to turn pink (Perper, 1980).

Two of the most influential factors in the decomposition process responsible for changes to the face are the swelling or bloating of tissues and insect activity. Swelling is observed during the earlier stages of decomposition; however, it continues throughout the process until gases are released at a later destructive phase. Bloating at the face can occur to such an extent that the living appearance will be unrecognisable (Taylor, 2001). Swelling is more noticeable at areas of the face with loose skin and includes:

- Inflation of the cheeks (Perper, 1980; Pounder, 1995; Clark *et al.*, 1997).
- Bulging of the eyeballs due to an accumulation of retrobulbar gases (Perper, 1980; Burton, 2007)
- Swelling of the eyelids causing closure (Perper, 1980; Pounder, 1995).
- Tongue protrusion from the oral cavity due to internal cavity pressure and engorgement (Perper, 1980; Pounder, 1995).

- Swelling and eversion of the lips (giving an extreme pout) (Pounder, 1995; Clark *et al*, 1997) with lip creases becoming stretched and ill-defined (Utsuno *et al.*, 2004).
- Enlargement of the ear lobes (Perper, 1980).

Facial distortion may also occur with insect activity. Insects will lay eggs in orifices and hair and eggs can be observed under the eyelids, amongst the eyelashes (Anderson and Cervenka, 2002) and inside the nostrils, mouth and ears. The eggs will distort the shape of the orifices and when the eggs hatch into maggots the activity can be strong enough to displace artificial dentures (Taylor, 2001). Once insect activity has reached the maggot stage, the face will be quickly overwhelmed and facial features will be destroyed.

Some post-mortem processes may occur less frequently. The process of saponification, which creates a substance called adipocere, can occur in certain environments. Adipocere is known as 'grave wax' or 'corpse wax' and is described as a hard, white homogeneous wax (Ruttan and Marshall, 1917), grey white caseous (Gill-King, 1997) and chalky crumbly soap (O'Brien and Kuehner, 2007;). It occurs more readily in submerged bodies, regularly affects the face at the cheeks and is caused by the initial release of fatty acids by hydrolysis of glycerol and the adoption of free sodium and potassium ions by free carboxylic acids leading to soap formation (saponification). Adipocere will retain the shape and features of a face even when advanced (Rodriguez, 1997; O'Brien and Kuehner, 2007). Saponification is more common in female, infant and obese remains (Gill-King, 1997) where the lipid content in the body is higher.

Activity by animals will also affect facial appearance. Animal carnivores are attracted to the hands (which are easy to detach from the body) and the face (due in part to the smell of the gases escaping through the facial orifices) (Mann *et al.*, 1990), and the mandible is frequently missing from exposed bodies due to animal activity.

The post-mortem processes affecting facial appearance are summarised in Tables 14.2a–f.

The majority of post-mortem research is recorded from observation of the cadaver in the course of forensic investigation, rather than from scientific experimentation. This may be a reflection of the ethics associated with the study of human remains in Europe. However, with facilities such as the Anthropological Research Facility in the University of Tennessee (Tennessee, USA), the Western Carolina University Human Identification Laboratory (North Carolina, USA), the Texas State University-San Marcos Forensic Anthropology Research Facility (Texas, USA) and the Southeast Texas Applied Forensic Science Facility (Texas, USA) it is surprising that there is such a paucity of quantifiable facial decomposition research. It is clear that facial decomposition is an area which requires considerable future research, especially with regard to human identification and disaster victim identification.

14.3 Facial depiction from post-mortem images

The police may attempt to identify bodies exhibiting peri- or post-mortem facial trauma or partial decomposition, and in these cases it may not be considered appropriate to present post-mortem facial images to the public. Frequently the police will then employ a forensic artist to produce a living facial depiction from the post-mortem image (Taylor, 2001; Gibson, 2008).

Post-mortem photography was commonplace in nineteenth century Britain and the USA (Meinwald, 1990), providing a means for memorialising dead loved ones, exhibiting executed criminals and describing crime scenes in newspapers and public publications. Memorial portraiture (or memento mori) became popular as a cheap alternative to the painted portrait and served as a keepsake to remember the deceased, especially common for children (Cameron, 1995). Often the deceased was arranged so as to appear alive, shown in repose to appear asleep or propped in a family gathering, and these portraits were often considered beautiful and sensitive (Ariés, 1985; Burns, 1990; Ruby, 1995). The same cannot be claimed for law enforcement or military images of the dead, where the aim was to display a captured and executed criminal to the public, enhancing the reputation of the police officers or military. Examples include photographs of the corpses of Jesse James in 1882 (Settle, 1996), John Dillinger in 1934 (Stewart, 2002) or Ché Guevaras in 1967 (James, 1970). These images were often brutal, stark and unarranged. Despite the current disapproval of death photography, exhibition images of criminals still infrequently appear in the media. A recent example of this is the execution image of Saddam Hussein (Singal *et al.*, 2010). Crime scene photography is currently the remit of police personnel (Blitzer and Jacobia, 2002), such as SOCOs or CSIs,

Table 14.2a Post-mortem changes to the human face.

Grouping	Sub-grouping	Post-mortem change observed	Ref. No.
The face (as a whole)	Effect on the muscles	Loss of muscle tone slackens the jaw	14
		Rigor mortis changes facial appearance from a relaxed musculature to a tensed one	18
		Rigor mortis affects the smaller muscles first, such as those relating to facial expression	10, 14
	Effect on the shape of the face	Swelling is more noticeable in areas of loose skin (giving the cheeks a distended appearance)	6, 14, 15
		Bloating at the face can hide the extent of facial fullness in life	18
		Algor mortis can cause the fat in the soft tissues to solidify and distort as a result of the physical environment	8
		Submerged bodies show a greater degree of facial distortion compared with bodies not submerged or buried	18
	Effect on areas of fat	Obese bodies are thought to lose body mass faster due to fat liquefaction	12, 15
	Effect on the colouring of the skin	Livor mortis can be more distinct in the head and neck regions in victims with intense congestion (e.g. heart failure)	8
		Grey-to-green discolouration turning to brownish shades at the nose and ears in a hot and dry environments	9
	Taphonomy	Activity by carnivores attracted to the face	12
	Adipocere	Adipocere at the cheeks in appropriate environment	13, 14, 16

Table 14.2b Post-mortem changes to the human face – skin and hair.

Grouping	Sub-grouping	Post-mortem change observed	Ref. No.
Skin and hair	Skin	Skin pallor occurs soon after death (15–30 minutes) due to cessation of blood flow	6
		Livor mortis can be hidden by darker skin pigments	14
		Light skin pigmentation may appear darker	5, 14
		Dark pigments may be lost	13, 14
	Skin colour	Decomposition causes a greenish/purple discolouration to the face	14
		A grey-to-green discolouration turning to brown at the nose and ears in hot and dry environments	9
		Purging of fluids from the mouth and nose can change skin colour at these areas	14
		Skin slippage can make skin colour determination unreliable	18

Table 14.2b (cont.)

Grouping	Sub-grouping	Post-mortem change observed	Ref. No.
	Skin texture	Rigor mortis can cause a gooseflesh appearance by affecting the erector pili muscles	14, 15
		Dehydration of the skin around the lips	18
		In specific environments mummification can cause skin to take on a leathery appearance	2, 11, 17
		Patterns from material used to cover a body can be seen imprinted on the skin	3
	Swelling of skin	Swelling is more noticeable in areas of loose skin in the face	14
	Hair position	Skin slippage due to decomposition can make the determination of hair position unreliable	15, 18
		Head hair may slough off as a hair mass in hot climates	12
		Head/facial hair and skin remains intact in bodies preserved in a water with an airtight seal (e.g. bog bodies)	4, 17
		Hair colour in bog bodies is usually modified – generally turns lighter with red tints	4

Table 14.2c Post-mortem changes to the human face – eyes and orbital area.

Grouping	Sub-grouping	Post-mortem change observed	Ref. No.
Eyes and orbital area	External appearance of the eye	Dehydration of the exposed eyeball produces a darkened area – a tache-noire de sclerotic pattern	14
		Bodies found in a prone position can have an haemorrhagic appearance at the eyes due to gravitational blood accumulation	8
	Eyeball shape/size	The eyeballs can bulge from accumulation of retrobulbar decomposition gases	5, 14
	Iris colour	Iris colour can be altered	18
		A blue iris colour can appear brown as a consequence of corneal clouding	5
	Corneal changes	A thin film appears over the cornea eventually turning cloudy	14
	Pupillary changes	Rigor mortis can alter the size of each iris and produce an artefactual difference in pupil sizes	14, 15
	Internal eyeball changes	Settling of red blood cells in the fundi of the eyeball	14
		Draining of blood from choroidal vessels leaves an undulating dark and light background	7
		Intra-retinal haemorrhage	7
		Oedema of the inner retina can opacify tissue	7

Table 14.2c (cont.)

Grouping	Sub-grouping	Post-mortem change observed	Ref. No.
	Palpebral position	The lateral (outer) corners of the eyes appear raised	18
		Insect activity (egg-laying) occurs regularly under the eyelids and amongst eyelashes	1
		Swelling of the eyelids will close the eye	14, 15

Table 14.2d Post-mortem changes to the human face – ears and hairline.

Grouping	Sub-grouping	Post-mortem change observed	Ref. No.
Ears and hairline area	Size and shape	Swelling is particularly dramatic at the ear lobes	14

Table 14.2e Post-mortem changes to the human face – nose and nasal area.

Grouping	Sub-grouping	Post-mortem change observed	Ref. No.
Nose and nasal area	Nose shape	Insect activity (egg laying) in the nostrils will cause distortion of the nose shape	18
	Nose position	Clothing/material resting on a face can distort the nasal tip (e.g. body recovery bag)	20
	Nose skin effects	Dehydration of the nasal tip and lips	6, 15

Table 14.2f Post-mortem changes to the human face – lips and mouth.

Grouping	Sub-grouping	Post-mortem change observed	Ref. No.
Lips and mouth area	Position of jaw	Loss of muscle tone has a slackening effect on the jaw – changing its position	18
	Mouth shape	Insect activity (egg laying) in the mouth will cause distortion	18
		Insect activity (egg laying) in the mouth can be strong enough to displace artificial dentures	18
	Position of tongue	The tongue swells and protrudes from the mouth	14, 15
	Lip appearance	Dehydration of the skin around the lips can make the teeth difficult to observe	6, 18
		Lip grooves are more defined in bodies with a closed mouth position	19
		Lip grooves tend to be stretched and poorly defined in bodies with an open mouth position	19
		Swelling and eversion of the lips (giving an extreme pout) can be observed	6, 15
	Colour of teeth	Putrefaction processes in a hot and humid environment can cause tooth enamel to turn pink	14

Table 14.2f (cont.)

Grouping	Sub-grouping	Post-mortem change observed	Ref. No.
	Skin colour	Decomposition causes a greenish/purple discolouration	14
		Purging of fluids from the mouth and nose can change skin colour in these areas	6, 14

Reference list for Table 14.2(a–f). See references section of this chapter for complete citations.
(1) Anderson and Cervenka, 2002; (2) Bass, 1997; (3) Berryman *et al.*, 1997; (4) Brothwell and Gill-Robinson, 2002; (5) Burton, 2007; (6) Clark *et al.*, 1997; (7) Curcio, 2005; (8) Dix and Graham, 2000; (9) Galloway, 1997; (10) Gill-King, 1997; (11) Manhein, 1997; (12) Mann *et al.*, 1990; (13) O'Brien and Kuehner, 2007; (14) Perper, 1980; (15) Pounder, 1995; (16) Rodriguez, 1997; (17) Slezdik and Micozzi, 1997; (18) Taylor, 2001; (19) Utsuno *et al.*, 2004; (20) Wilkinson, 2004.

but in the 1930s and 1940s reporters commonly photographed murder victims, accidental deaths and suicides to print in newspaper articles (Burns, 1990; Burns and Burns, 2002). The portrayal of such images has become increasingly seen as vulgar, sensationalist and taboo, a cultural shift that may be a reflection of a wider social discomfort with death.

Traditionally, the forensic artist will produce a sketch of the face or use photo-editing software to alter the post-mortem image (Taylor, 2001; Gibson, 2008). This will involve the removal of distracting detail (such as dirt, blood, maggots and detritus); the removal of trauma (such as bruising, wounds, feature distortion and inflammation); the addition of feature detail (such as open eyes, closed mouth, hair and missing parts) and the interpretation of areas affected by decomposition, based on taphonomic knowledge (such as swollen, discoloured or putrefied areas). Forensic artists who carry out this work require anatomical, taphonomic and imaging experience in order to interpret the post-mortem image. A high level of artistic skill will also be necessary to produce a realistic facial depiction.

The most important factor for post-mortem depiction is the quality of the images utilised. Ideally the forensic artist will be allowed access to the body for data collection and facial assessment, but often these images are recorded by a police photographer under less than optimal conditions.

Best practice guidelines for cadaver face photography are as follows:

- Position the entire head and hair in the image without cropping.
- Position a scale close to the face and perpendicular to the view line of the camera. The scale should neither obscure the face nor cast a shadow onto the face.
- Use soft/even lighting.
  - Outdoors – cloudy, bright light.
  - Indoors – use box lights or directional lights on either side.
- Avoid using flash, which can confuse perception of depth and facial contour. Where conditions demand the flash, replicate photographs of the face with and without flash for reference purposes.
- Record frontal and profile images of the face.
- Any close-up images of scars, tattoos or other identifying marks should include a scale.

  Frontal image:

- When the head is detached from the body, prop in an upright position and photograph from the front at eye level.
- When the head is attached to the body, position the camera using a tripod above the body. If the face is positioned looking up, position the camera in a direct vertical line above the eyes (Figure 14.1a).

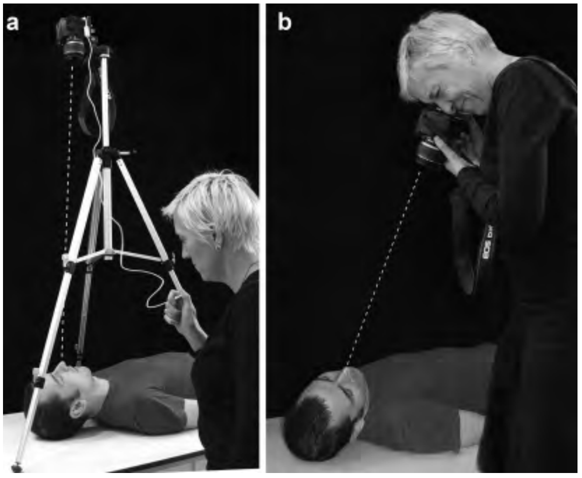


Figure 14.1 Positioning the camera to record a frontal cadaver face image.

Otherwise position the camera so that the face is 'looking' into the lens (Figure 14.1b).

- Set the tripod as tall as possible and use the zoom function in order to avoid distortion. Use a focal length between 85 and 200 mm. Do not use a wide-angle lens.
- If it is not possible to stand directly over the body, manually position the camera as described above (Figure 14.1a).

Profile image:

- When the head is detached from the body, prop in an upright position and photograph a profile view of the face, with the camera positioned level with the eyes of the cadaver.
- When the head is attached to the body and repositioning is not possible, photograph a profile view of the face, with the camera at the same height as the eyes of the cadaver (Figure 14.2).
- Position the camera at least 6 feet away and use a zoom function. Use a focal length of between 85 and 200 mm. Do not use a wide-angle lens.

Forensic artists will often produce post-mortem depictions in greyscale rather than colour, in order to minimise distraction/confusion and allow the viewer to add colour information in their imagination. Post-mortem images frequently show discoloured skin, bruising and stains, and removal of these details is easier when colour saturation is low. Where skin colour, hair colour and eye colour are uncertain, a greyscale depiction is preferable to an erroneous colour depiction.

Figure 14.2 Positioning the camera to record a profile cadaver face image.

It is important to highlight any details that may aid recognition and identification, such as dental anomalies, scars, facial creases, tattoos, clothing, headgear, jewellery or hair accessories.

Procedural guidelines for post-mortem sketches can be found in Taylor (2001) and Gibson (2008).

Where photo-editing software (such as Adobe Photoshop®) is utilised, a report should be produced detailing the process and tools employed. A database of collected frontal portrait facial images can be used as a reference for template images and the database may be divided by gender, age and ethnicity. It is preferable to employ a frontal image of the face for the depiction, as this will allow matching with the database and allow hairstyles and features to be imported. Once the image has been imported into photo-editing software (Figure 14.3a), the cadaver head can be selected using *quick mask*, *lasso* or *marquee* tools and pasted into a new layer. In this way the background can be replaced with a plain tone or neutral image and background distractions such as soil, blood, the mortuary table/trolley and the body bag can be removed (Figure 14.3b).

Imperfections to the skin surface of the cadaver face, such as discolouration, bruising, detritus and trauma, can be corrected (Figure 14.3c). The following tools are routinely employed (example tool names from Adobe Photoshop®):

Healing brush tool – corrects imperfections on the image. This tool allows the artist to paint with sampled pixels from another part of the image and matches the texture, lighting, transparency, and shading of the sampled pixels to the new area, so that the repaired pixels blend seamlessly into the rest of the image. This tool is useful for correcting small areas of skin imperfection.

Clone stamp tool – paints one part of an image over another part of the same image. The *Clone stamp* tool is useful for duplicating objects or removing a defect in an image. This tool is useful for removing larger areas of skin imperfection.

Cut and *paste* tools – selects an undamaged area of the skin (such as at the cheek) and pastes it over a large area of imperfection, such as a wound, bruise or discolouration. These tools are useful when the affected area is unilateral and the undamaged mirror facial area/feature can be copied.

Features exhibiting small distortions, such as nasal tip, ear lobe or eyebrow deformation can be corrected using distortion tools (Figure 14.3d), such as *liquify* or *transform*. These tools push, pull, rotate, reflect, pucker, or bloat any area of an image. The distortions can be subtle or drastic, but drastic distortions may cause blurring of the image.

Where features exhibit large distortions, such as those caused by physical pressure, position-related distortion, eye closure, nose deviation, slack jaw or trauma, features can be imported from the database to replace the affected area (Figure 14.3e). Features are chosen from the database to correspond to the estimated ante-mortem morphology of the cadaver feature, e.g. a similar nose from the database can be imported and scaled to fit the cadaver face. Small alterations to the shape of the imported nose can then be applied using *transform* or *liquify* tools in order to match the nose to the estimated morphology and perspective. In this way, life-like soft tissue features, teeth, hair, neck and eyes can be depicted.

The imported features will be pasted in different layers and each piece can be blended into the surrounding face using tools that soften or alter the transparency of the edge of the piece, such as *eraser*, *blend* or *burn* tools.

Clothing, shoulders and headgear can also be imported from an image database, then scaled and altered to fit the cadaver head using *transform* and *distort* tools (Figure 14.3f).

The finished head will now be made up of multiple layers containing different features and before the layers are merged together the brightness, contrast and tone balance of each can be altered to match each other and the cadaver face. Once the levels for each layer are consistent, the layers can be merged to create a final frontal image of the living face (Figure 14.3g).

Artistic post-mortem depictions are often presented to the public as an aid to recognition and it must be noted that a high degree of estimation and interpretation can be involved in the production. Nevertheless, post-mortem depictions can be very useful in forensic investigations and protect the public from exposure to disturbing and distasteful images.

Figure 14.3 Example of post-mortem depiction process using actor. a: Original post-mortem image; b: Removal of background; c: Removal of skin imperfections; d: Correction of nasal tip to remove post-mortem distortion; e: Addition of eyes, lips and chin features; f: Addition of hair, neck and clothing; g: Target image.

14.4 Post-mortem prediction using 3D statistical analysis

Since the prediction of the living face from post-mortem appearance can be difficult, research recording post-mortem changes to the face would be invaluable to the craniofacial identification field.

A recent pilot study was conducted (Tillotson, 2010) to investigate whether common facial decomposition patterns could be recorded and if these patterns could be used to predict living facial appearance. Data collection was conducted over a period of 3 months at the Anthropological Research Facility (ARF), Knoxville, University of Tennessee, USA. Six cadavers were studied, five males and one female, with a mean subject age of 68 years (age range of 55 to 82 years). The causes of death were varied, but none of the subjects presented with facial injuries or pathological conditions that would affect facial appearance.

Data collection included a visual morphological assessment (facial colour, shape and morphology and dehydration evidence at the skin), photographic records (frontal and profile views), air temperature measurements (°C and °F – using a digital thermometer at the cadaver location and average daily air temperature from a local weather station) and 3D laser scans of the faces of the cadavers (using a Polhemus Scorpion portable laser scanner). Data were collected at two intervals throughout the day, once in the morning and once in the evening, until the face became unrecognisable or when severe insect activity was encountered (enough to obscure the facial features). Photographs were recorded using a fixed tripod approximately 1–1.5 metres from the subject, with close-up images recorded for detail. Fixed registration points were employed so that the 3D face scans could be aligned to each other during analysis. The registration points needed to be unaffected by decomposition, so the left and right external auditory meatus and two bony forehead points were utilised. Figure 14.4 displays a laser scanner in use and the Figure 14.5 displays the resulting scan.

The following patterns of change in the facial morphology were observed:

The eyes and orbits:

- The soft tissue inside the orbit, including the eyeball and the eyelids, darkened in colour throughout the decomposition process for all subjects.
- Initial dehydration of the eyeball caused a decreased soft tissue volume in the orbits underlying the eyelids, producing a sunken appearance in 67% of the sample.
- Swelling of the soft tissues at the orbit occurred in all subjects, with 67% showing minor changes and 33% showing major changes. Major changes included swollen eyelids, smoothing of orbital wrinkles and a marked orbital bloating in one subject causing unrecognisable feature morphology.
- Thirty-three per cent of the sample exhibited an upturned lateral canthus.

The nose:

- Insect oviposition in the nostrils resulted in nasal swelling in all subjects and further nostril swelling followed the eggs hatching into maggots.

Figure 14.4 In situ laser scan set up for a cadaveric face.

Figure 14.5 Typical 3D laser scan of a face.

- A change in nasal skin tone was observed in 67% of subjects, with 33% exhibiting early change and 34% exhibiting change following maggot mass activity.
- Thirty-three per cent of the sample showed a colour change at the nasal tip, plus the soft tissue of the nasal tip appeared to dehydrate, with some shrinkage and skin tightening. In one subject this dehydration was severe enough to cause the morphology of the underlying cartilage to become visible.

The lips and mouth:

- Eighty-three per cent of the sample exhibited slackening of the jaw associated with initial loss of muscle tone and gravitational affects on the muscles of mastication followed by rigor mortis, resulting in soft tissue sagging and subsequent swelling in the area beneath the jaw line.
- Lip colour darkened in the entire sample and this appeared in combination with dehydration and subsequent retraction of the lips, creating a thin-lipped appearance and displaying the underlying tooth architecture more clearly.
- Thirty-three per cent of the sample displayed a downturning of the lateral corners of the mouth, resulting in a 'grimace-like' appearance.

The ears and hairline:

- A colour change caused by lividity was observed at the ears of all subjects. This manifested as an initial darkening to purple, and subsequent dark brown appearance.
- The ear lobes of 33% of the sample displayed some swelling associated with bullae formed in the cheek area of the face and subsequent sagging and swelling of the soft tissue beneath the ear.

The face as a whole:

- Eighty-three per cent of the sample exhibited an overall darkening in skin tone, which appeared to be associated with swelling and insect activity. However, one subject exhibited a consistent skin pallor, until maggot activity and skin sloughing occurred leading to a darker stained skin tone –100% of subjects exhibited sagging of the cheeks associated with jaw slackness, whilst 50% of the sample displayed an initial gaunt appearance followed by cheek swelling.
- Fifty per cent of the sample exhibited facial bloating, with one subject displaying such extensive swelling that the face became unrecognisable.

The laser scan data were initially analysed using a software program called MorphAnalyser (Tiddeman *et al.*, 1999, 2000, 2004), which is a 3D software package that enables the user to create morphs and average faces from the facial laser scan data (Ajaj *et al.*, 2007*a*, 2007*b*). MorphAnalyser created intermediate stages between scans allowing for production of an animation of the changes to the face of each subject. The animation enabled a visual timeline of decompositional changes to the subject faces to be observed, displaying some evidence of patterns between the subjects. These decompositional change patterns were then analysed utilising a piece of software called GeoMagic Qualify 10.0® (a reverse modelling software package created by GeoMagic®, North Carolina, USA). This software can be used to determine the differences between surfaces, either by using absolute mean shell deviations, standard deviations of errors during shell overlaps, maximum and minimum range maps, or histogram plots and colour maps (Kau *et al.*, 2005*a*). These methods of comparison use statistical analyses to allow detailed evaluation of two or more independent surfaces and their relationship to each other. These can then be displayed as colour maps, data tables and graphs. This type of analysis has been used for growth and twin studies (Kau *et al.*, 2005*b*). According to past research, use of laser surface scanners in combination with software such as GeoMagic Qualify 10.0® and Rapidform® (another type of reverse modelling software package created by INUS Technology Incorporated, Seoul, Korea) can produce an accurate and detailed description of the changes to the surface of a human face (Kau *et al.*, 2006), whether it is due to the growth of the facial bones (Nute and Moss, 2000) or the swelling retained from surgical procedure or abscess. Facial asymmetry in three dimensions (O'Grady and Antonyshyn, 1999) and changes in facial expression (Okada, 2001) can also be detected using this combination of scanning hardware and software analysis. This software was used to analyse the decomposition face changes in the sample by comparing each stage with the next.

Minor facial changes included:

- Little to no bloating due to decomposition gases, with the majority of the face displaying no volumetric change other than in areas of insect activity.
- Shrinkage at the orbits, specifically the eyelid regions due to dehydration of the eyeball and subsequent collapse of the orbital tissues.
- Lower lip shrinkage and the slack jaw reinforced any lower lip retraction.

Major facial changes included:

- Waning of rigor mortis, combined with positional effects on the soft tissues, causing sagging at the temporal, zygomatic and gonial regions of the face.
- Bloating of the lateral regions of the face producing a rounder overall face shape.
- Severe swelling at the soft tissue in and around the orbits resulting in bloating of the eyelids and smoothing of all facial wrinkles. Majority of bloating originating more laterally and progressing medially, resulting in orbital swelling between the lateral canthus and the bridge of the nose.
- Shrinkage at the nasal tip, glabella, upper lip and menton, consistent with sagging of the soft tissues laterally causing tightening at the midline region of the face.
- Shrinkage at the upper lip as a result of lip retraction.

MorphAnalyser then created an average face-change pattern by combining the six subject scans at three stages (Tiddeman *et al.*, 1999, 2004). This allowed assessment of a common decomposition pattern for the subjects. An example of this can be seen in Figure 14.6 for stage 1. The combined average face created using MorphAnalyser was analysed using GeoMagic Qualify 10.0® to assess whether any patterns of facial change could be visualised in the average face. The results can be seen in Figures 14.6 and 14.7.

Figure 14.6 displays the areas of volume change to the Average Face Scan between stages 1 and 2. The majority of the upper and middle face displayed little to no volume change. Shrinkage (light blue areas) was shown at the orbits, due to dehydration of the eyeballs and the subsequent collapse of the overlying eyelids, and at the mandible, due to slackening of the jaw. Bloating (yellow areas) was exhibited at the nose, side

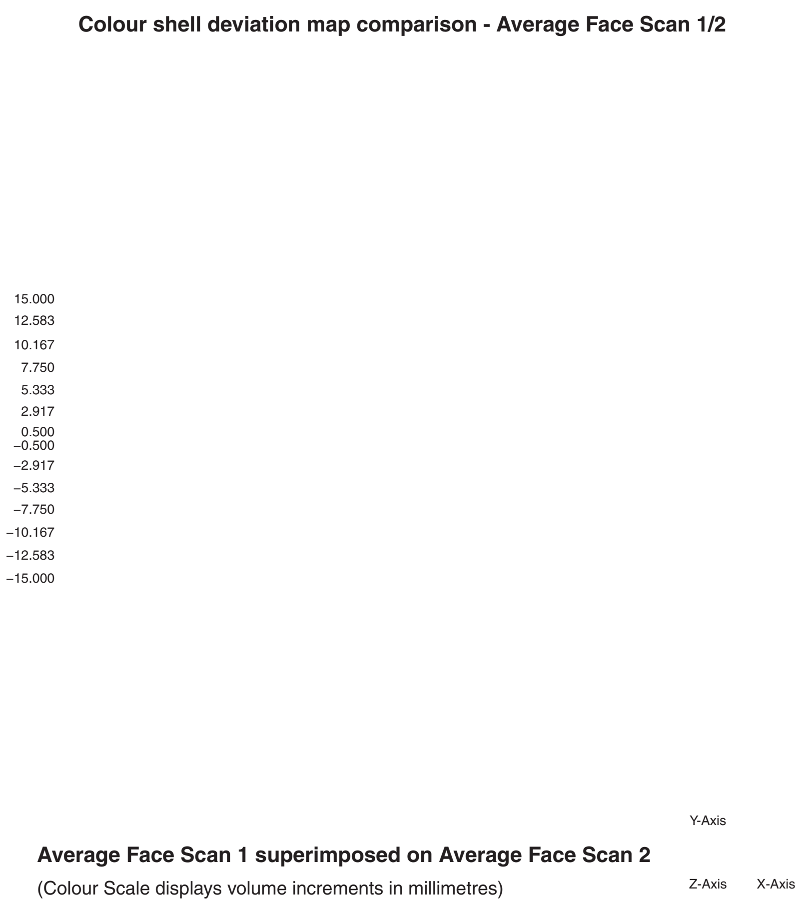


Figure 14.6 The colour shell deviation map comparison for Average Face Models 1 and 2. This figure is produced in colour in the plate section.

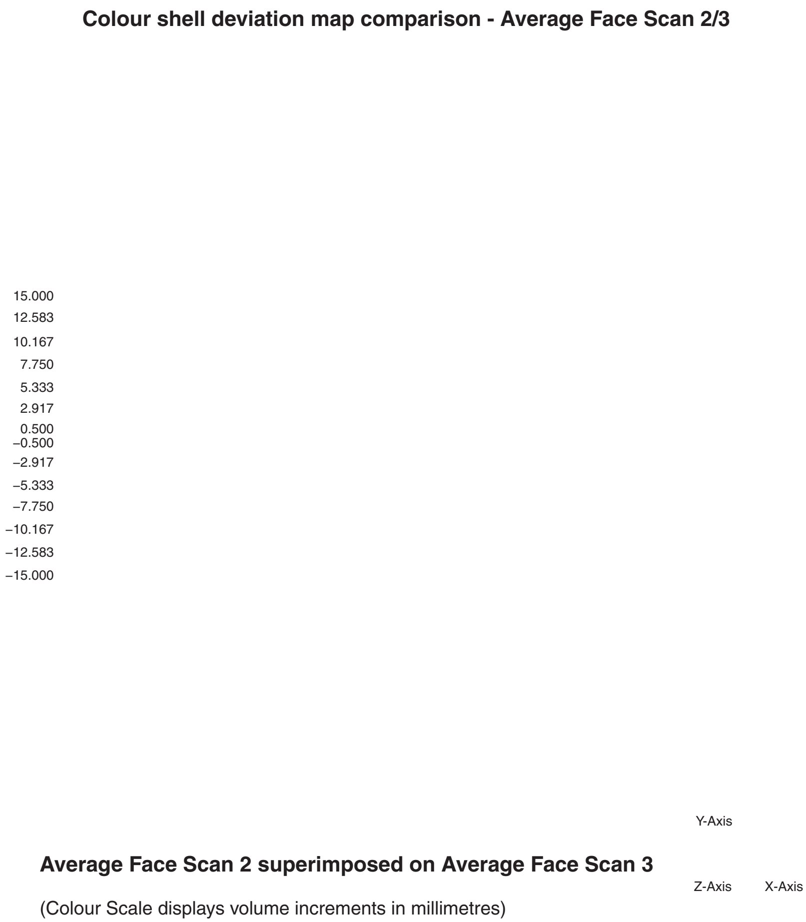


Figure 14.7 The colour shell deviation map comparison for Average Face Models 2 and 3. This figure is produced in colour in the plate section.

of the mouth and inferior cheek, due to insect activity (oviposition by flies). The large swelling seen to the right of the mouth was related to one subject with a large egg mass in this area. Bloating was also shown at the ears and may be related to a combination of the effects of gravity and the blocks used to support the head.

Figure 14.7 displays the areas of volume change to the Average Face Scan at stages 2 and 3. Shrinkage is exhibited at the orbits (similar to Figure 14.6) and is continuous with the cheek, nose and jaw areas. The shrinkage at the orbits was due to dehydration of the eyeball, whilst the volume change at the cheeks and nose was due to gravity and the effects on soft tissue following the waning of rigor mortis. Bloating is exhibited lateral to the orbits and anterior to the ears and at the gonial area. These areas of swelling were consistent with the effects of positional gravity and the accumulation of decomposition gases under the looser skin of the face. Shrinkage at the jaw and to the left side of the mouth was due to slackening of muscles of mastication. The bloating at the right side of the mouth was related to the movement of the mandible in combination with insect activity. Shrinkage of the upper lip was associated with darkening and dehydration of the tissue. The average nasal tip displays no change throughout the early decomposition process.

In conclusion, this pilot study suggests that it is possible to assess facial decomposition and establish predictable patterns for utilisation in post-mortem prediction using 3D laser scan data in combination with photographic records and morphological assessment. This study also suggests that, even with a small sample, a predictable pattern can be recorded in early to mid decomposition including:

- Initial orbital shrinkage associated with eyeball dehydration and structural tissue collapse, followed by orbital bloating associated with accumulation of decomposition gases in the eyelids and orbital tissues.
- Lateral bloating to the face at the cheeks, mouth and nose associated with the accumulation of decomposition gases in the loose skin tissue and slackness of the jaw.
- No volume change at the majority of the upper and middle face.
- Where bloating occurs with the cadaver in a supine position there will be a related compression of tissues along the midline of the face due to the gravitational pull on the swollen areas.

This study does not take into account any differences in decomposition patterns related to different face shapes, gender, ethnicity or age due to the limited sample size. Future studies might consider comparisons between different age groups, genders and ethnic groups and between brachycephalic and dolichocephalic face types, large- and small-nosed faces, and fat and thin faces.

References

Ajaj, M., Jepson, N. J., Tiddeman, B. P. and Hobson, R. S. (2007*a*). An evaluation of facial form in hypodontia using 3D imaging. *International Conference on Ectodermal Dysplasias (ED06)*, Copenhagen, Denmark.

Ajaj, M., Jepson, N. J., Gordon, P., Tiddeman, B. P. and Hobson, R. S. (2007*b*). Comparison of 3D stereophotogrammetric and cephalometric measurements in children. 85th General session and Exhibition of the International Association for Dental Research, New Orleans, 21–24 March.

Anderson, G. S. and Cervenka, V. J. (2002). Insects associated with the body: their use and analysis. In W. D. Haglund and M. H. Sorg (Eds.), *Advances in Forensic Taphonomy: Method, Theory, and Archaeological Perspectives*, pp. 173–200. Boca Raton, FL: CRC Press.

Ariés, P. (1985). *Images of Man and Death*. Translated by J. Lloyd. Cambridge, MA: Harvard University Press.

Bass, W. M. (1997). Outdoor decomposition rates in Tennessee. In W. D. Haglund and M. H. Sorg (Eds.), *Forensic Taphonomy, The Post-mortem Fate of Human Remains*, pp. 181–186. Boca Raton, FL: CRC Press.

Bass, W. M. and Meadows, L. (1990). Time since death of the human body: variables and observations in case and experimental field studies. *Journal of Forensic Sciences*, **35**, 103–111.

Berryman, H. E., Bass, W. M., Symes, S. A. and Smith, O'B. C. (1997). Recognition of cemetery remains in the forensic setting. In W. D. Haglund and M. H. Sorg (Eds.), *Forensic Taphonomy, The Post-mortem Fate of Human Remains*, pp. 165–170. Boca Raton, FL: CRC Press.

Blitzer, H. L and Jacobia, J. (2002). *Forensic Digital Imaging and Photography*. London: Academic Press.

Brothwell, D. and Gill-Robinson, H. (2002). Taphonomic and forensic aspects of bog bodies. In W. D. Haglund and M. H. Sorg (Eds.), *Advances in Forensic Taphonomy: Method, Theory, and Archaeological Perspectives*, pp. 119–132. Boca Raton, FL: CRC Press.

Burns, S. B. (1990). *Sleeping Beauty: Memorial Photography in America*. Altadena, CA: Twelve Trees Press.

Burns, S. B. and Burns, E. A. (2002). *Sleeping Beauty II: Grief, Bereavement in Memorial Photography American and European Traditions*. New York, NY: Burns Archive Press.

Burton, J. L. (2007). The external examination: an often neglected autopsy component. *Current Diagnostic Pathology*, **13**(5), 357–365.

Cameron, H. (1995). *Memento Mori: Mourning, Monuments and Memory*. Perspektive Berlin eV. http://www.ztg.tu-berlin.de/fixingid/mementomori.pdf

Clark, M. A., Worrell, M. B. and Pless, J. E. (1997). Post-mortem changes in soft tissues. In W. D. Haglund and M. H. Sorg (Eds.), *Forensic Taphonomy, The Post-mortem Fate of Human Remains*, pp. 151–164. Boca Raton, FL: CRC Press.

Curcio, C. A. (2005). Imaging maculopathy in post-mortem human eyes. *Vision Research*, **45**(28), 3496–3503.

Dix, J. and Graham, M. (2000). *Time of Death, Decomposition and Identification: An Atlas*. Boca Raton, FL: CRC Press.

Galloway, A. (1997). The process of decomposition: a model from the Arizona-Sonoran Desert. In W. D. Haglund and M. H. Sorg (Eds.), *Forensic Taphonomy, The Post-mortem Fate of Human Remains*, pp. 139–150. Boca Raton, FL: CRC Press.

Galloway, A., Birkby, W. H., Jones, A. M., Henry, T. H. and Parks, B. O. (1989). Decay rates of human remains in an arid environment. *Journal of Forensic Sciences*, **34**, 607–616.

Gerasimov, M. M. (1971). *The Face Finder*. New York, NY: Lippincott.

Gibson, L. (2008). *Forensic Art Essentials – a Manual for Law Enforcement Artists*. New York, NY: Elsevier.

Gill-King, H. (1997). Chemical and ultrastructural aspects of decomposition. In W. D. Haglund and M. H. Sorg (Eds.), *Forensic Taphonomy, The Post-mortem Fate of Human Remains*, pp. 93–108. Boca Raton, FL: CRC Press.

Haglund, W. D. (1993). Disappearance of soft tissue and the disarticulation of human remains from aqueous environments. *Journal of Forensic Sciences*, **38**(4), 806–815.

Haglund, W. D., Reay, D. T. and Swindler, D. R. (1989). Canid scavenging/disarticulation sequence of human remains in the Pacific Northwest. *Journal of Forensic Sciences*, **34**, 587–606.

Haglund, W. D. and Sorg, M. H. (2002). Human remains in water environments. In W. D. Haglund and M. H. Sorg (Eds.), *Advances in Forensic Taphonomy: Method, Theory, and Archaeological Perspectives*, pp. 201–218. Boca Raton, FL: CRC Press.

Hill, I. (2006). Physical appearance. In S. Black and T. Thompson (Eds.), *Forensic Human Identification – An Introduction*, pp. 365–378. Boca Raton, FL: CRC Press.

James, D. (1970). *Che Guevara: A Biography*. New York: Stein and Day Publishers.

Kau, C. H., Richmond, S., Savio, C. and Mallorie, C. (2006). Measuring adult facial morphology in three dimensions. *Angle Orthodontist*, **76**(5), 773–778.

Kau, C. H., Richmond, S., Zhurov, A. I. *et al*. (2005*a*). Reliability of measuring facial morphology with a three-dimensional laser scanning system. *American Journal of Orthodontics and Dentofacial Orthopedics*, **128**, 424–430.

Kau, C. H., Zhurov, A., Bibb, R., Hunter, L. and Richmond, S. (2005*b*). The investigation of the changing facial appearance of identical twins employing a three-dimensional laser imaging system. *Orthodontics and Craniofacial Research*, **8**, 85–90.

Komar, D. A. (1998). Decay rates in a cold climate region: a review of cases involving advanced decomposition from the Medical Examiner's Office in Edmonton, Alberta. *Journal of Forensic Science*, **43**(1), 57–61.

Lain, R., Griffiths, C. and Hilton, J. M. N. (2003). Forensic dental and medical response to the Bali bombing – a personal perspective. *Medical Journal of Australia*, **179** (7), 362–365.

Manhein, M. H. (1997). Decomposition rates of deliberate burials: a case study of preservation. In W. D. Haglund and M. H. Sorg (Eds.), *Forensic Taphonomy, The Post-mortem Fate of Human Remains*, pp. 469–482. Boca Raton, FL: CRC Press.

Mann, R. W., Bass, W. M. and Meadows, L. (1990). Time since death and decomposition of the human body: variables and observations in case and experimental field studies. *Journal of Forensic Sciences*, **35**(1), 103–111.

Meinwald, D. (1990). Memento mori: death and photography in 19th-century America. *CMP Bulletin*, 9 (4). http://cmp1.ucr.edu/terminals/memento_mori [accessed 18 May 2005].

Mellen, P. F. M., Lowry, M. A. and Micozzi, M. S. (1993). Experimental observations on adipocere formation. *Journal of Forensic Sciences*, **38**, 91–93.

Micozzi, M. S. (1986). Experimental study of postmortem change under field conditions: effects of freezing, thawing, and mechanical injury. *Journal of Forensic Sciences*, **31**(3), 953–961.

Nute, S. J. and Moss, J. P. (2000). Three-dimensional facial growth studied by optical surface scanning. *Journal of Orthodontics*, **27**, 31–38.

O'Brien, T. G. and Kuehner, A. C. (2007). Waxing grave about adipocere: soft tissue change in an aquatic context. *Journal of Forensic Science*, **52**(2), 294–301.

O'Grady, K. F. and Antonyshyn, O. (1999). Facial asymmetry: three-dimensional analysis using laser surface scanning. *Plastic and Reconstructive Surgery*, **104**(4), 928–937.

Okada, E. (2001). Three-dimensional facial simulations and measurements: changes of facial contour and units associated with facial expression. *Journal of Craniofacial Surgery*, **12**(2), 167–174.

Perper, J. A. (1980). Time of death and changes after death: Chapter II, Part 1, anatomical considerations. In W. U. Spitz and R. S. Fisher (Eds.), *Medicolegal Investigation of Death: Guidelines for the Application of Pathology to Crime Investigation*, 2nd Edn, pp. 14–49. Springfield, IL: Charles **C**. Thomas.

Pounder, D. (1995). *Lecture Notes, Department of Forensic Medicine, University of Dundee*. Available at: http://www.dundee.ac.uk/forensicmedicine/notes/notes.html [accessed 1 September 2009].

Prag, J. and Neave, R. A. H. (1997). *Making Faces*. London: British Museum Press.

Prieto, J. L., Magania, C. and Ubelaker, D. H. (2004). Interpretation of post-mortem changes in cadavers in Spain. *Journal of Forensic Science (technical note)*, **49**(5), 1–6.

Rodriguez III, W. C. (1997). Decomposition of buried and submerged bodies. In W. D. Haglund and M. H. Sorg (Eds.), *Forensic Taphonomy, The Post-mortem Fate of Human Remains*, pp. 459–468. Boca Raton, FL: CRC Press.

Ruby, J. (1995). *Secure the Shadow: Death and Photography in America*. Boston, MA: MIT Press.

Ruttan, R. F. and Marshall, M. J. (1917). The composition of adipocere. *Journal of Biological Chemistry*, **29**, 319–327.

Settle, W. A. (1996). *Jesse James Was His Name*. Columbia, MO: University of Missouri Press.

Singal, J., Lim, C. and Stephey, M. J. (2010). December 2006, Saddam's Last Words. 19 March 2010 *Time*. http://www.time.com/time/specials/packages/article/

0,28804,1967340_1967352_1968024,00.html [accessed 14 April 2010].

Slezdik, P. S. and Micozzi, M. S. (1997). Autopsied, embalmed, and preserved human remains: distinguishing features in forensic and historic contexts. In W. D. Haglund and M. H. Sorg (Eds.), *Forensic Taphonomy, The Post-mortem Fate of Human Remains*, pp. 483–496. Boca Raton, FL: CRC Press.

Stewart, T. (2002). *Dillinger, The Hidden Truth*. Xlibris Corporation (Self-publishers).

Taylor, K. T. (Ed.) (2001). *Forensic Art and Illustration*. Boca Raton, FL: CRC Press.

Thompson, T. and Black, S. (2006). *Forensic Human Identification – an Introduction*. Boca Raton, FL: CRC Press.

Tiddeman, B. P., Duffy, N. and Rabey, G. (2000). Construction and visualisation of three-dimensional facial statistics. *Computer Methods and Programs in Biomedicine*, **63**, 9–20.

Tiddeman, B. P., Rabey, G. and Duffy, N. (1999). Synthesis and transformation of three-dimensional facial images. In *IEEE Engineering in Medicine and Biology*, **18**(6), 64–69.

Tiddeman, B. P., Jones, B. and Perrett, D. I. (2004). Facial prototyping and transformation techniques and their applications in perception research. In H. Vuyk and P. J. F. M. Lohuis (Eds), *Facial Plastic and Reconstructive Surgery*. London: Hodder Arnold.

Tillotson, A. (2010). *Disaster victim identification using craniofacial analysis*. Unpublished PhD, University of Dundee.

Utsuno, H., Kanoh, T., Tadokoro, O. and Inoue, K. (2004). Preliminary study of post-mortem identification using lip prints. *Forensic Science International*, **149**(2), 129–132.

Wilkinson, C. M. (2004). *Forensic Facial Reconstruction*. Cambridge: Cambridge University Press.

Wilkinson, C. M., Rynn, C., Peters, H. *et al.* (2006). A blind accuracy assessment of computer-modelled forensic facial reconstruction using computed tomography data from live subjects. *Journal of Forensic Science, Medicine and Pathology*, **2**(3), 179–187.

Chapter 15

Manual forensic facial reconstruction

Ludo Vermeulen

15.1 Introduction

Manual forensic facial reconstruction is one of the most subjective, as well as one of the most controversial techniques in the field of forensic anthropology. Nevertheless, the process has had enough successful results to allow the advancement of research and methodologies. That there exists an interrelationship between the skull and face is unquestionable, but the precise nature of this relationship is complex, variable and not yet fully understood.

15.2 Controversy

Although there are several scientific branches involved, critics have said that the technique is subjective and controversial due to the presence of an artistic aspect (Stadtmuller, 1922; Suk, 1935; Stephan, 2003). It could be argued that manual reconstruction depends on the expertise of the practitioner, or even on what inspires him or her that particular day!

There have been various studies where reconstructions from several practitioners were compared, which did not suggest high reliability or reproducibility (Stadtmuller, 1922; Haglund and Reay, 1991; Helmer *et al.*, 1993; Bongartz *et al.*, 2005). Unfortunately these studies did not account for different levels of artistic/sculptural expertise or knowledge and experience in facial anatomy/anthropology, with novices/students compared with established experts. Often even the methods of facial reconstruction varied between practitioners.

Throughout the years, the manual technique has often provided spectacular results in seemingly hopeless cases, yet a large part of the scientific community remains sceptical (Suk, 1935; Stephan, 2003). This chapter aims to demonstrate that artistic input is essential when utilising the manual technique, and how the subjective aspect can be minimised.

15.3 History

Knowledge of human anatomy was greatly advanced during the Renaissance period. Artists such as Leonardo da Vinci, Michelangelo, Verrocchio and Titian developed scientific art through a deeper understanding of anatomical dissection. Their detailed illustrations of these dissections provided a thorough awareness of the skeleton as an armature for the human form. Anatomists such as Fallopio, Celalpino and Vesalius utilised the rendered images to illustrate their treatises.

An important partnership between artists and scientists began in the early days of facial reconstruction. Anatomists who required a visual method of demonstrating their data relied upon sculptors under their direction to produce three-dimensional (3D) representations (His, 1895; Kollmann and Buchly, 1898).

In 1895, anatomist Wilhelm His worked with artist Sefner to produce a reconstruction on a plaster skull cast of Johann Sebastian Bach. A facial approximation of Dante was developed in 1898 by anthropologist Kollmann with Buchly as sculptor. The reconstruction of an ancient Saxon by anatomist Merkel was assisted by Eichler in 1900 to reveal a face from ancient history (Wilkinson, 2004: 39–68). The partnership of science-based art continued in 1946, when the anthropologist Wilton Krogman worked with Mary Jane McCue as a sculptor to complete his first 3D reconstruction of a 40-year-old male of African ancestry (Taylor, 2001: 20). This association between artists and scientists continues today as more detailed studies in facial research are revealed.

Craniofacial Identification, ed. C. Wilkinson and C. Rynn. Published by Cambridge University Press.

The dream of many scientists is to produce the ideal facial reconstruction method. The technique would be rapid, reproducible and create a perfect likeness, which could be executed by almost anyone regardless of their background knowledge, capabilities or talents. Better yet, the simple push of a button would make all of this possible. Unfortunately, their dream is not a reality at present.

Even though research has evolved techniques, scientists attempt to eliminate any form of artistic input or rely on their own limited artistic abilities to develop these approximations. Yet when an important historical reconstruction receives much media attention, we often see an artist appear at the forefront (Hawass, 2005; BBC News, 2008; Braun, 2008). Was the artist consulted to make the approximation more attractive or perhaps to apply a more human aspect to the sculpture?

In these situations, a sculptor is the practitioner most often called upon for his artistic talents. Who better than an expert facial artist that specialises in 3D facial portraiture? The artist possesses three important skills. The first is the power of observation, which enables the artist to detect the small details and proportions, and measure accurately while observing his subject more often than the sculpture he is creating. The second and third skills are material knowledge and craftsmanship, which enable him to reproduce the technical details correctly and harmoniously in any number of media at hand: be it clay, plasticine or wax.

When sculpting a living model, the artist can measure and study the subject's face thoroughly from all angles to acquire the information needed to complete a 3D facial portrait. In constructing a facial approximation, the subject's face is missing and the artist can only rely on artistic experience and talent combined with scientific data applied to the skull to complete the unknown face.

Until researchers and technicians have developed a viable method of producing the 'perfect' 3D facial reconstruction, the necessity of artistic input continues to be an essential asset to the overall finished 3D facial reconstruction.

15.4 Manual technique

There are various techniques employed to reconstruct the face on a skull, including the anatomical approach (Gerasimov, 1971; Wilkinson, 2004). All anatomical elements of the human cranium are analysed and each soft tissue layer of muscle, fat and skin is modelled, using anatomical knowledge to interpret musculature from bone, until the face is complete. This technique is sometimes referred to as the Russian method, of which Mikhail Gerasimov was the pioneer. Although Gerasimov did not sculpt all the muscles of the face, he did base his technique in anatomical observation and interpretation and was himself a talented sculptor.

The anthropometric approach was primarily developed in the USA (Krogman and İşcan, 1962; Gatliff, 1984; Taylor, 2001) and is often known as the American Method. In this method pegs representing the average soft tissue depths are attached to the skull at fixed landmarks. The face is then 'fleshed out' using clay, based on the tissue depths of the markers.

The author employs a combination technique, initiated in Manchester, UK by Richard Neave (Prag and Neave, 1997) and further developed by Caroline Wilkinson and her team (Wilkinson, 2004). It includes the study of facial anatomy, expression, anthropometry, anthropology and the relationship between the soft and hard tissues of the face. This rather complex technique combines the majority of both the anatomical and anthropometrical approaches and requires a great deal of training and dexterity.

15.5 Analysis of the skull

To create a standard 3D portrait, a sculptor begins with a block of clay. To develop a forensic facial reconstruction, a skull or a replica of the skull forms the armature (Figure 15.1). In what is probably the most important phase of the reconstruction, the skull is studied extensively. Both artists and anatomists know that the skull is a unique and beautiful frame, to which the muscles and soft tissues of the face are attached. Specialists in the fields of forensic anthropology, pathology, anatomy, dentistry, orthodontics and other sciences can be consulted to offer skull assessments, providing a wealth of additional information.

Sex can usually be assigned to skeletal remains based upon the size and morphology of the cranium and mandible (Krogman and İşcan, 1986). For additional confirmation in ambiguous cases, other bony structures such as the pelvis or DNA analysis can be utilised (Brothwell, 1981).

The age is estimated mainly by examining the dentition. It is advisable to use two or three of the

Figure 15.1 An archaeological skull mounted in the Frankfurt Horizontal Plane (teeth estimated and sculpted in wax).

various techniques available as a control measure to ensure accuracy. Prior to full adulthood, the eruption of the dentition provides a relatively accurate estimate of age (Brothwell, 1981; Scheuer and Black, 2000). Subsequently, the wear and general condition of the dentition can be used to estimate age, less accurately since this varies depending on lifestyle.

Analysis of the skull also provides an indication of the ancestry of the subject (Stewart, 1948; Krogman and İşcan, 1986; Clement and Ranson, 1998). This is not always easy as geographic boundaries are difficult to define due to genetic mixing and migration. An attempt is made to categorise a skull by morphology, broadly as White, Black, Australoid or Mongoloid-type. Pathologies and traumas pertaining to the resulting sculpture are identified. The teeth and dental occlusion are thoroughly examined.

It is at this stage that some reconstruction practitioners begin to imagine the finished face. Astute reconstruction artists will concentrate on all factual details to accomplish an accurate facial approximation knowing the final face will develop throughout the entire procedure.

The skull is mounted on an adjustable stand in the Frankfort Horizontal Plane. The Frankfort Horizontal Plane is an anthropological standard position that closely approximates the natural position of the head in life. Orbitale, the lowest point on the lower margin of the orbit, is horizontally aligned with the porion, the most lateral point on the roof of the external auditory meatus (Wilkinson, 2004: 70).

15.6 The eyes

The facial reconstruction begins with the placement of artificial eyes. Great care is taken to position each eye in the centre of the socket. Once each eye has been situated, we notice immediately that there is not much room to spare surrounding the eyeball. The placement is then adjusted according to expert studies (Whitnall, 1921: 251–257; Krogman and İşcan, 1962; Stephan, 2002; Wilkinson and Mautner, 2003).

In frontal view, Krogman states, 'The apex of the cornea when viewed from norma frontalis is the juncture of two lines, one drawn from the medial edge of the orbit (maxillofrontal) to the lateral margin of the orbit (ectoconchion) and the other line bisecting the orbit between the superior and inferior margins' (Krogman and İşcan, 1962).

In lateral profile, Krogman states, 'The outer point of the cornea is approximately tangent to a centrally located line drawn from the superior and inferior margins of the orbit.' Recent studies have shown that this will result in a position that is too deep inside the eye socket. Normal protrusion should be taken as the flat plane of the iris touching a tangent from mid-supraorbital to mid-infraorbital margins (Stephan, 2002; Wilkinson and Mautner, 2003).

There is often no information relating to eye colour when reconstructing unknown skeletal remains. Ancestry of the subject may offer some clues, but to avoid risking misrecognition due to erroneous eye colour, it is safer to err on the side of caution and use a neutral tone. Artificial eyes made from plastic, glass or human prosthetic eyes made from acrylic provide a life-sized and life-like eyeball. A useful tip to circumvent the issue of incorrect eye colour is to circulate black and white images of the finished reconstruction using eyes that are neither too light nor too dark (Figure 15.2).

15.7 Tissue depth markers

To estimate the soft tissue thickness, tissue-depth markers can be fixed to the anatomical landmarks of the face. These markers are the result of many studies using different collection methods. In 1883, German anatomist Welcker used soft tissue-depth data collected from cadavers to identify the skull of the German poet, Schiller. Wilhelm His (1895) measured the tissue depths of 24 male and four female cadavers

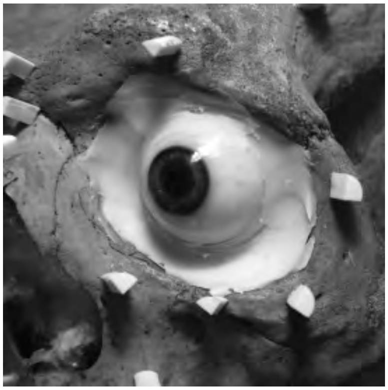


Figure 15.2 A mid-tone eyeball mounted in the orbit, to avoid misrepresentation of indeterminable eye colour.

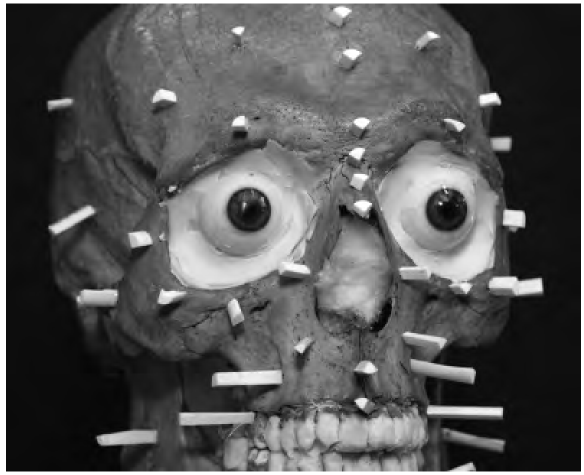


Figure 15.3 Skull with average tissue depth markers attached.

by inserting sewing needles with a rubber disc attached into the soft tissue of the faces. Having documented his efforts, the data were used to create a bust of composer Johann Sebastian Bach.

Since the days of Welcker and His, many measurements have been collected from different groups of people using various techniques. The change in tissue depth due to loss of muscle tone and shrinkage created problems with cadaveric studies. Soft tissue distortion, which occurs from drying and embalming shortly after death, and putrefaction with bloating of the face (Taylor, 2001), which may occur rapidly, even in temperate climates, prevented accurate measurements. The horizontal position of the cadaver also created false tissue-depth measurements due to the action of gravity.

Later experiments included measurements taken using radiographs, MR images and CT data, but more recent reliable measurements were performed using ultrasound, as the test subjects can be seated in an upright position (data tabulated in Wilkinson, 2004: 124–156), with the skin elasticity and muscle tone in life adding bulk to the face.

Today, the number and variety of test subjects have been extended to include data relating to sex, age and ethnic origin. Scientists and researchers have concurred on a number of fixed cranial landmarks and using calibrated and reproducible methods have determined the average thickness of the soft tissue in those areas (data tabulated in Wilkinson, 2004: 124–156). Soft tissue-depth thickness markers are fixed to these locations, indicating the mean depth of the surrounding soft tissue (Figure 15.3).

15.8 The muscles

A facial reconstruction can be thought of as a reverse dissection. An anatomist begins his process with a smooth face, then removes the soft tissue and muscles layer after layer to eventually reveal the skull. A facial reconstruction artist begins with the skull and builds layer upon layer to develop a recognisable face.

The skull displays attachment markings left by the muscles. The size and roughness of the markings and the shape of the local bone indicate the volume of the facial muscles (Wilkinson, 2010). Skull analysis is time consuming, but of the utmost importance. Each muscle placement must first be observed then meticulously developed (Figure 15.4).

15.9 The nose

The nose is the centre of the face and dominates the overall appearance of the reconstruction. Anyone who has ever been at a costume party would notice how the disguise of a different nose can completely change a face. An accurate nose is crucial to the reconstruction.

The nasal protrusion in life consists mostly of cartilage, which perishes after death. The shape and form of the nasal tip is therefore difficult to predict from skeletal remains. The anterior nasal aperture clearly indicates the position of the nose in the centre of the face, but the relationships between the skull and the width, length, inclination and form of the nose are still under research.

From the early years of anthropological facial research, there have been many studies relating to the

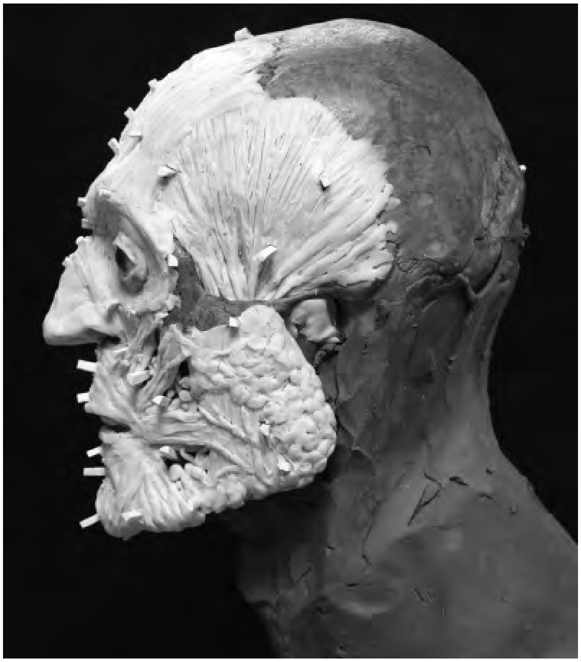


Figure 15.4 Musculature sculpted with reference to the morphology of the skull.

nasal form. A simple and commonly used method is that of Krogman (Krogman and İşcan, 1986). He stated that when reconstructing White-type skulls, the nasal aperture is measured at its widest point, and then 10 mm are added to estimate the maximum nasal width, whereas 16 mm are added to the width of the aperture for Negroid-type skulls. In the lateral view, the nose projection from subnasale to pronasale measures approximately three times the length of the anterior nasal spine. Recent studies have found this method to be inaccurate (Stephan *et al.*, 2003; Rynn and Wilkinson, 2006) plus, in practice, the anterior nasal spine is often damaged and difficult to measure.

Another method for nasal projection is provided by Gerasimov (1971) who stated that the profile of the nose is projected by two straight lines; one at a tangent to the last third of the nasal bones and the other as a continuation of the main direction of the tip of the anterior nasal spine. This is called the 'two-tangent method' and has been found to accurately predict a point on the surface of the tip of the nose (Rynn and Wilkinson, 2006), but has been found to be inaccurate at predicting pronasale: the most anterior point of the tip of the nose in the Frankfurt Horizontal Plane (Stephan *et al.*, 2003).

The aim of recent research by Rynn *et al.* (2009) was to compose a reliable and readily reproducible method of predicting nasal morphology from the bony aperture, which restricts subjectivity whilst allowing anatomical nuance to be taken into account.

Rynn *et al.* (2009) state: 'A series of simple regression equations was produced using linear distances between pairs of bony landmarks to predict nasal profile dimensions and restrict potential subjective error in Gerasimov's 'Two-Tangent' method. Maximum nasal width, the position of the alae and nostrils, and prediction of nasal asymmetry were incorporated into the resulting three-dimensional nasal prediction method.'

With this new study, we now have enough information to estimate a recognisably individualistic nose in 3D form (Figure 15.5).

15.10 The mouth

The living mouth offers a volume of information from the teeth and dental occlusion. Two forces play a major role in the appearance of our mouths. The tongue applies an outward force on our teeth and jaw and the cheeks and lips apply an inward force culminating in a kind of balance that can be read from the stance and occlusion of the teeth.

In the absence of tongue, cheeks and lips when developing a reconstruction, we get a good idea of the strength and position of the soft tissue from the teeth and jaws and their associated occlusion. The importance of a dentist or orthodontist should not be underestimated when studying the skull. They can provide vital information that defines the state and volume of the lips if the subject has an overbite, underbite or other abnormalities that would relate to the appearance of the mouth in life. In normal occlusion with the mouth at rest, the opening is 1 or 2 mm above the incisal edge of the upper central incisors (Angel, 1978).

In studies the most accurate placement for the corners of the mouth has been found to be on a line radiating from the first premolar–canine junction (Gerasimov, 1955) and/or on a vertical line, in frontal view, with the medial border of the iris (Wilkinson *et al.*, 2003) or the infraorbital foramen (Stephan and Murphy, 2008).

Gatliff (1984) estimated the vertical thickness of the lips from the gumline-to-gumline measurement, and Wilkinson *et al.* (2003) suggested that there is a

Nasal Prediction Method

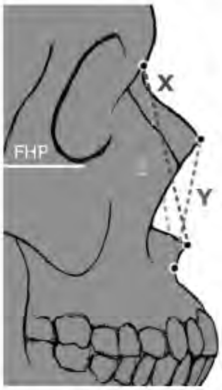


Fig 1: Craniometric Measurements

X : Nasion-Acanthion (mm)

Y : Rhinion-Subspinale (mm)

Z : Nasion-Subspinale (mm)

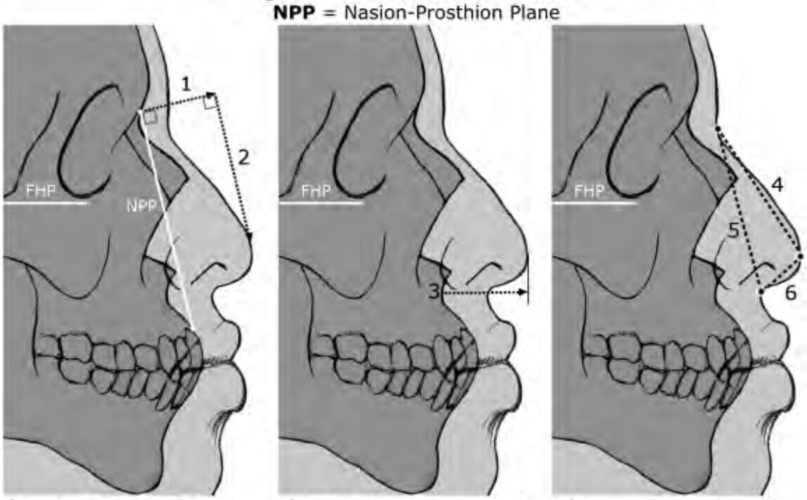
NB: The maximum aperture width (MAW) is three-fifths the maximum nasal width (MNW)

5[MAW/3] = MNW

Table 1: Profile Regression Equations

Equation number	Predicted dimension	Simplified equation	Relevant ancestry and/or sex
1.	pron ant	0.83 Y -3.5	All
2.	pron vert	0.9 X -2	All
3.	pron FHP	0.93 Y -6	All
4.	nasal length	0.74 Z +3.5	Caucasoid
5.	nasal height	0.63 Z +17	F Caucasoid
		0.78 Z +9.5	M Caucasoid
6.	nasal depth	0.5 Y +1.5	F All
		0.4 Y +5	M All

Fig 2: Soft Dimensions Predicted
NPP = Nasion-Prosthion Plane



(1) Pronasale Projection anterior from NPP **(Pron ant)**

(2) Pronasale Height in NPP **(Pron vert)**

(3) Pronasale projection from subspinale in FHP **(Pron FHP)**

(4) Nasal Length

(5) Nasal Height

(6) Nasal Depth

Figure 15.5 Nasal prediction method (Rynn *et al*., 2009). Equation numbers refer to original publication.

Figure 15.6 Estimated nose, sculpted based on Rynn *et al*. (2009)

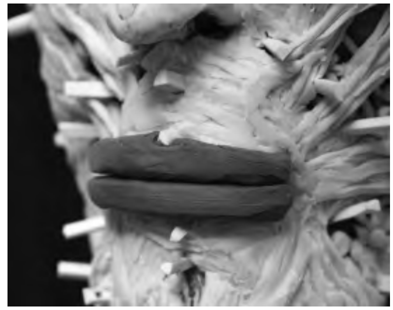


Figure 15.7 Mouth width and lip thickness based on Wilkinson *et al*. (2003) and Gerasimov (1955).

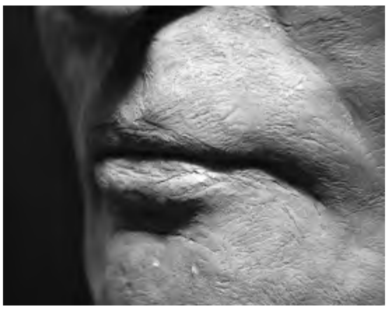


Figure 15.8 The form of the lips, influenced by dental occlusion.

positive correlation between the upper lip thickness and the maxillary teeth height; lower lip thickness and mandibular teeth height; and the total lip thickness and the total teeth height.

Science indicates the position, width and volume of the lips (Figure 15.6). With skill and experience, the artist must now apply these data to create a realistic mouth. The artist has the 'artistic freedom' to create an average mouth for the required age, gender and ethnic origin of the skeletal remains, only within the boundaries dictated by science (Figure 15.7).

15.11 The ears

Many scientists agree that information on the shape of the ear is very limited. Russian studies, by Fedosyutkin and Nainys (1993) and Gerasimov (1971), claim that the skull provides some information regarding the shape of the ears.

We do know that the external auditory meatus indicates the position of the ear, and the jaw line generally echoes the angle of the ear (Wilkinson, 2004). Commonly, the male ear is larger than the female ear and both tend to continue to grow with age. Gerasimov (1971) summarised: 'Many details of the ear's complicated relief must be intuitively reconstructed.' With so little information available, it is fortunate that the size and form of the ear usually contributes a relatively minor role in facial recognition.

15.12 Development of the eyes

The lacrimal crests and malar tubercles indicate the inner and outer canthi of the eyes (Whitnall, 1921). Beginning with the lower lid, a small strip of clay should hug the eyeball from lacrimal crest to malar tubercle and reach as high as the bottom of the iris. The upper lid should be slightly longer, more arched than the lower lid, superficial to the lower lid at the outer canthus and rest at a point halfway between the pupil and the upper margin of the iris.

Once again science dictates the data but the expertise of the artist is needed to translate the information into a human eye. The artist must consider the age of the subject as significant age-related changes occur in the area around the eyes and eyelids during life (Figure 15.8). A historical reconstruction may require expression to the face, which must also be included at this time.

Fedosyutkin and Nainys (1993) stated that the eyebrow pattern could be determined from the supraorbital bony structure. Angel (1978) stated that the eyebrows generally followed the line of the brow ridge and that they were approximately 3–5 mm above the supraorbital margin. Following this information, the artist can fashion the eyebrows in the proper position once the skin has been applied.

15.13 Skin application

A bare skull is always a strange sight, but now with muscles, nose, eyes and ears in place, the reconstruction

Figure 15.9 The ageing effect of texture around the eye.

begins to look like an actual face (Figure 15.9). The application of skin is a very important last step in completion of the reconstruction and probably the most difficult for students and scientists alike.

The manual dexterity, skill and experience of the artist are vital in placing the final layer on top of the musculature, using the markers as a guide. The practitioner must realise that the measurements are averages with a certain margin of error, but it is the artist who is able to bring proportion and harmony to the 3D portrait.

Once the skin layer is in place, finishing touches can be completed. Age-related facial details must be considered such as eye bags, sagging neck, jaw line softness, drooping eyelids and even some wrinkles. The sculptor applies these changes gently as the additional details can only be estimated and the majority of ageing data have already been incorporated when using the age-related tissue depth measurements.

Hairstyle, facial hair, scars and glasses must only be added to the reconstruction when the human remains and forensic findings dictate that they should be present. It is preferable to use an indication of hair to avoid an odd bald appearance, but beware! False information may confuse the observer and obscure recognition.

15.14 The result

The skull directly provides information crucial to a portrait artist, such as the position of the eyes, the size and position of the mouth and nose, as well as the correct measurements and proportions of the face. These vital factors determine the resemblance of a portrait, and are often missed by inexperienced artists, even in the presence of their actual model. Adherence to forensic techniques ensures that these important details are estimated as accurately as possible. If the position of any of these features is altered, the face changes immediately and the likeness will be lost. During the reconstruction, science dictates and limits the choices without the need for interpretation by the artist. In the event that there is no information available, the most probable and most frequently occurring detail of a feature is chosen using knowledge and expertise.

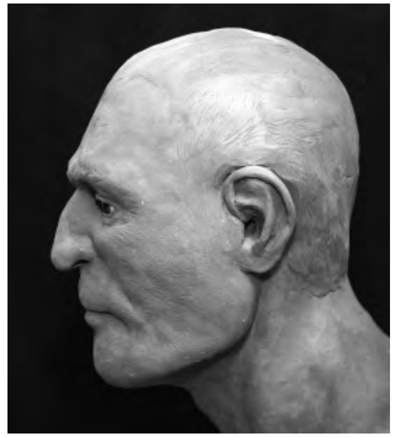


Figure 15.10 The finished facial reconstruction.

The most important contributions to forensic facial reconstruction that an artist can make are experience and manual dexterity as an artisan. The result may not be a perfect portrait of the victim (Figure 15.10), but the main goal is to create a recognisable likeness to assist an investigation of unknown human skeletal remains. We must not forget that the reconstruction is the last resort when other forensic techniques have failed. The facial approximation can lead to an extension of the list of potential victims and breathe new life into the forensic investigation.

References

Angel, J. L. (1978). Restoration of head and face for identification. *Proceedings of Meetings of the American Academy of Forensic Science*, St. Louis, MO.

BBC News (2008). Experts 'rebuild' composer's face. 29 February 2008. http://news.bbc.co.uk/1/hi/scotland/tayside_and_central/7270795.stm

Bongartz, J., Buzug, T. M., Hering, P., Tille, C. and Seitz, H. (2005). Introduction to the comparative study on facial reconstruction. *Proceedings of RSFP 2005: 2nd International Conference on Reconstruction of Soft Facial Parts*, p. 55. Remagen, 17–18 March 2005.

Braun, D. (2008). DNA-Based Neanderthal face unveiled. http://news.nationalgeographic.com/news/2008/09/080917-neanderthal-photo.html

Brothwell, D. R. (1981). *Digging up Bones*, 3rd Edn. London and Oxford: British Museum/Oxford University Press.

Clement, J. G. and Ranson, D. L. (1998). *Craniofacial Identification in Forensic Medicine*. Sydney: Arnold Publishers.

Fedosyutkin, B. A. and Nainys, J. V. (1993). The relationship of the skull morphology to facial features. In M. Y. İşcan and R. P. Helmer (Eds.), *Forensic Analysis of the Skull: Craniofacial Analysis, Reconstruction, and Identification*. New York, NY: Wiley-Liss Inc.

Gatliff, B. P. (1984). Facial sculpture on the skull for identification. *American Journal of Forensic Medicine and Pathology*, **5**(4), 327–332.

Gerasimov, M. M. (1955). *The Reconstruction of the Face from the Basic Structure of the Skull*. Translated by W. Tshernezky. Russia: Publishers unknown.

Gerasimov, M. M. (1971). *The Face Finder*. New York, NY: Lippincott.

Haglund, W. D. and Reay, D. T. (1991). Use of facial approximation in identification of Green River serial murder victims. *American Journal of Forensic Medicine and Pathology*, **12**(2), 132–142.

Hawass, Z. (2005). Press release 10 May 2005. Tutankhamun Facial Reconstruction. http://www.guardians.net/hawass/Press_Release_05–05_Tut_Reconstruction.htm

Helmer, R., Rohricht, S., Petersen, D. and Moer, F. (1993). Assessment of the reliability of facial reconstruction. In M. Y. İşcan and R. P. Helmer (Eds.), *Forensic Analysis of the Skull: Craniofacial Analysis, Reconstruction, and Identification*, pp. 229–247. New York, NY: Wiley-Liss.

His, W. (1895). Anatomische Forschungen über Johann Sebastian Bachs Gebeine und Antlitz nebst Bemerkungen über dessen Bilder. *Abhandlungen der mathematisch-physikalischen Klasse der Königlichen Sachsischen Gesellschaft der Wissenschaften*, **22**, 379–420.

Kollmann, J. and Buchly, W. (1898). *Die Persistenz der Rassen und die Reconstruction der Physiognomie prähistorischer Schädel*. Archiv für Anthropologie, 329–359.

Krogman, W. M. and İşcan, M. Y. (1962). 1st Edn. (1986). 2nd Edn. *The Human Skeleton in Forensic Medicine*. Springfield, IL: Charles C. Thomas.

Prag, J. and Neave, R. A. H. (1997). *Making Faces*. London: British Museum Press.

Rynn, C. and Wilkinson, C. M. (2006). Appraisal of traditional and recently proposed relationships between the hard and soft dimensions of the nose in profile. *American Journal of Physical Anthropology*, **130**, 364–373.

Rynn, C., Wilkinson, C. M. and Peters, H. L. (2009). Prediction of nasal morphology from the skull. *Forensic Science, Medicine, and Pathology*, **6**, 20–34.

Scheuer, L. and Black, S. (2000). *Developmental Juvenile Osteology*. London: Academic Press.

Stadtmuller, F. (1922). Zur Beurteilung der plastischen Rekonstruktionsmethode der Physiognomie auf dem Shädel. *Zeitschrift für Morphologie und Anthropologie*, **22**, 337–372.

Stephan, C. N. (2002). Facial approximation: globe projection guideline falsified by exophthalmometry literature. *Journal of Forensic Science*, **47**(4), 730–735.

Stephan, C. N. (2003). Anthropological facial ‘reconstruction’ – recognizing the fallacies, ‘unembracing’ the errors, and realizing method limits. *Science and Justice*, **43**(4), 193–200.

Stephan, C. N., Henneberg, M. and Sampson, W. (2003). Predicting nose projection and pronasale position in facial approximation: a test of published methods and proposal of new guidelines. *American Journal of Physical Anthropology*, **122**, 240–250.

Stephan, C. N. and Murphy S. J. (2008). Mouth width prediction in craniofacial identification. *Journal of Forensic Odontostomatology*, **27**(1), 2–7.

Stewart, T. D. (1948). Medicolegal aspects of the skeleton – age, sex, race and stature. *American Journal of Physical Anthropology*, **6**, 315–321.

Suk, V. (1935). Fallacies of anthropological identifications and reconstructions: a critique based on anatomical dissections. *Publications of Faculty of Science*, 270, 1–18. Brno: University of Masaryk.

Taylor, K. T. (Ed.) (2001). *Forensic Art and Illustration*. Boca Raton, FL: CRC Press.

Whitnall, S. E. (1921). *Anatomy of the Eye and Orbit*. Oxford: Oxford Medical Publications.

Wilkinson, C. M. (2004). *Forensic Facial Reconstruction*. Cambridge: Cambridge University Press.

Wilkinson, C. M. (2010). Facial reconstruction – anatomical art or artistic anatomy? *Journal of Anatomy*, **216**(2), 235–250.

Wilkinson, C. M. and Mautner, S. A. (2003). Measurement of eyeball protrusion and its application in facial reconstruction. *Journal of Forensic Science*, **48**, 4.

Wilkinson, C. M., Motwani, M. and Chiang, E. (2003). The relationship between the soft tissues and the skeletal detail of the mouth. *Journal of Forensic Science*, **48**(4), 728–732.

Chapter 16

Relationships between the skull and face

Christopher Rynn, Tatiana Balueva and Elizaveta Veselovskaya

16.1 Introduction

Craniofacial reconstruction (CFR) and approximation (CFA) are among terms commonly used to describe the procedure of predicting and recreating a likeness of an individual face based on the morphology of the skull (Gerasimov, 1955; Krogman and İşcan, 1986; Wilkinson, 2004). A variety of methods exist, most of which employ averaged tissue-depth data at various landmarks of the skull, and feature prediction guidelines to estimate the morphology of the eyes, nose, mouth and ears. Some methods also entail interpretation of general and local skull morphology to predict individual muscles of mastication and facial expression. This chapter will describe craniofacial patterns, and discuss anatomical and morphological interrelationships between the skull and the face, such as may be useful in craniofacial reconstruction.

A key principle of anatomy is that structure is inextricably related to function. Every organ, indeed every organelle in every cell in every biological organism, has been gradually evolving its way into a functional niche through the process of natural selection over innumerable generations. The human head is anatomically and architecturally fascinating because of the wide range of functions carried out by its constituent parts. Organs dedicated to four of the five special senses are housed in the craniofacial complex: the eyes (*vision*), the inner, middle and outer ears (*audition*/balance), the mouth and oropharynx (*gustation*/mastication/respiration/verbalisation) and the nasopharyngeal airway (*olfaction*/respiration), and the functionality of each is responsible in part for the structure of the head and face. The eyes of a child appear much larger and further apart than those of an adult; of course, this is only relative to the rest of the face, but the eyes must develop to a certain level, and thus a certain size, in order to function, and must be a certain distance apart for effective binocular vision, so this structural arrangement is set up early and maintained throughout development. However, the nose, mouth and lower face change shape quite drastically between infancy and adulthood, but remain functional throughout development via a series of compensatory mechanisms involving the entire craniofacial complex.

16.2 Displacement and the functional matrix

The spatial arrangement of three interconnected elements must remain within certain boundaries throughout growth and development in order to maintain functionality: (1) the brain; (2) the foramen in the cranial base through which the cranial nerves pass (generally accompanied by blood vessels); and (3) the targets of the cranial nerves (e.g. the facial sense organs and muscles of mastication and facial expression). The brain and basicranium develop earlier and faster than the facial composite (Scheuer and Black, 2000), the bones of which remain separate throughout the majority of development for good reason. The skull does not grow as a unit, but rather the bones of the face move away from one another (displacement/translation) and have both depository surfaces, upon which bone is laid down, and resorptive surfaces, from which bone is dissolved and assimilated.

Although the majority of the cranial base is internally resorptive and externally depository, surprisingly, most of the external surfaces of the facial bones are in fact resorptive (Figure 16.1) and throughout development these bones, such as the maxillae, zygomatics and mandible, are displaced apart as new bone is deposited on their internal surfaces. This explains how the lower craniofacial complex changes shape so drastically from infancy to adulthood, maintaining functionality throughout. Displacement does not

Craniofacial Identification, ed. C. Wilkinson and C. Rynn. Published by Cambridge University Press.

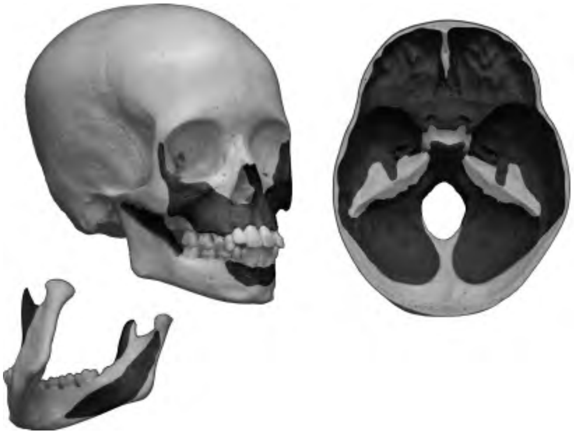


Figure 16.1 Diagram of depository (light) and resorptive (dark) areas of the developing skull (based on Enlow, D. H., Kuroda, T. and Lewis, A. B. (1971) The morphological and morphogenetic basis for craniofacial form and pattern. *Angle Orthodontist*, **41**, 161–188).

occur through the deposition on the interior surface pushing the bones apart, since this kind of pressure between the bones would inhibit deposition and eventually lead to necrosis by restricting the blood supply to the vascular osteogenetic connective tissue (Enlow and Hans, 1996). Displacement is actually driven by the functional relationships established by the soft tissues which surround and interact with a given bone. The craniofacial complex is a functional matrix (Moss and Young, 1960) and the relationship between the bone and soft tissues is reciprocal and responsive. Thus, the shape of the adult skull is based on the internal and external soft tissues of the face and head, as much as the face and head is based on the skull.

16.3 Craniofacial form

Two major factors in human evolution, bipedality and brain growth, had a particularly massive effect on the head and face. Upright posture necessitated downward rotation of the spinal column, in comparison to primates and quadrupeds. The expansion of the human forebrain is linked to the downward rotation of the face. The combined effect is as if the mandible and the soft tissues of the oropharynx were caught in a closing vice, between the descending maxilla and the spinal column. Thus, the evolution and development of the lower human face, in all its diversity, and the uniquely human chin, can be viewed as adaptation to counteract reducing function, particularly on the part of the mandible. What we think of as dental malocclusion syndromes: overjet, underbite, edge-to-edge occlusion, crossbite and overcrowding, are actually developmental solutions to this evolutionary problem, relieving the pressure on the mandible and oropharynx by remodelling the face (Enlow and Hans, 1996). In terms of shape, relative expansion of one area in the craniofacial complex appears to be balanced with relative reduction or displacement of another, both evolutionarily and developmentally.

Cranial form can generally be described as lying somewhere between the two extremes of dolichocephaly (long and narrow) and brachycephaly (short and wide), the intermediate being mesocephaly. The maximum cranial breadth divided by the maximum cranial length gives the cranial index. In dolichocephalics, the cranial index is less than 0.75; in brachycephalics, it is more than 0.8 (Hooton, 1946: 488). Dolichocephaly is linked to a leptoprosopic face (narrow, long and protrusive), and brachycephaly to a euryprosopic face (broad and flatter). In frontal view, the difference between these two extremes appears to be simply breadth, but in profile view it is clear that the whole headform is quite different, and the shape of the entire craniofacial profile and of each constituent facial feature is affected. In the White-type skull, at both extremes of the range, functional problems arise and the condition becomes pathological. Premature fusion of the sagittal suture will lead to dolichocephaly, and premature fusion of the coronal suture will lead to brachycephaly (Tessier, 1971), but on a deeper level, and in less extreme cases, the adult headform depends to a large extent on the degree of the flexion of the cranial base (Figure 16.2) i.e. the angle, in profile view, of the superior surface of the sphenoid bone as it flexes about the sella turcica (which houses the pineal gland). If the angle is more obtuse, i.e. the basicranium is relatively flat and elongated, a more dolichocephalic headform will develop. If the angle is more acute, i.e. the basicranium is short and flexed, a more brachycephalic headform will develop. In this sense, 'the basicranium is the template that establishes the shape and the perimeter of the facial growth field' (Enlow and Hans, 1996) and the craniofacial complex develops in a sequential cascade of functional adaptations based upon this template.

16.4 The mandible, mouth and nose

In White-type skulls, dolichocephaly underlies a tendency for type II dental occlusion (Enlow and

Hans, 1996). The term 'overjet' implies that the maxilla is protrusive, but this is a misconception. In fact, a retrusive maxilla is exhibited, together with an even more retrusive mandible (Mew, 1992). This leads to a restricted oropharynx, which is not functionally ideal, and so the head tends to tilt back to allow more room for the structures of the mouth and throat to function. Hence, the characteristic profile linked to dolichocephaly in White-type skulls is a sloping forehead, retrusive chin and a projecting nose, sometimes with a hump at the bridge (Mew, 1992) (Figure 16.2) as the nasal aperture becomes longer and narrower along with the entire face, as if stretched by the receding maxilla and mandible.

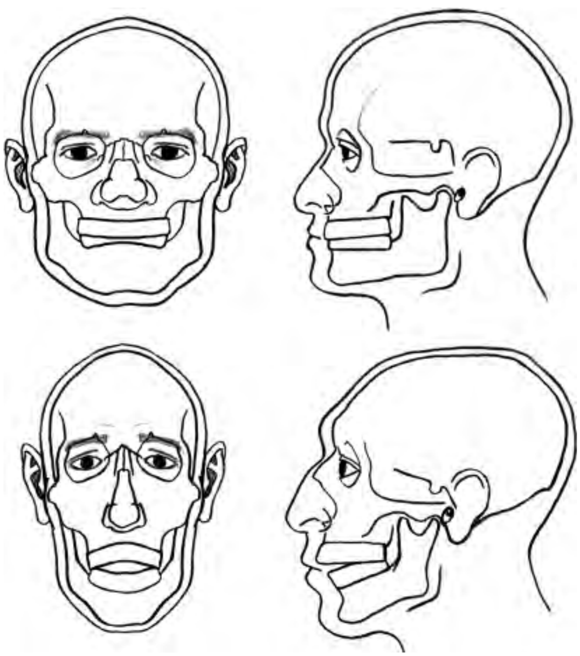


Figure 16.2 Brachycephaly (top row) and dolichocephaly (bottom row) characteristics in the White-type craniofacial complex (adapted with permission from Enlow, D. H. and Hans, M. G. (1996) *Essentials of Facial Growth*. Philadelphia, PA: W. B. Saunders).

Brachycephaly underlies a tendency for type III dental occlusion, or mandibular protrusion, which can lead to 'underbite' or 'crossbite' (Enlow and Hans, 1996). The higher degree of flexion of the basicranium means the spinal column and the maxilla are rotationally closer to one another, and maintenance of function causes the mandible to remodel accordingly. The protrusive mandible is configurationally different from the retrusive mandible, notably in the width of the ramus and the gonial angle. Previous research (Rynn, 2006) found the mean gonial angle of 25 male and 25 female North American White adults to differ by less than two degrees (Figure 16.3). Despite both ends of the male range being more acute than the female range, the red areas (showing one standard deviation either side of the mean) overlap almost completely. In the same sample, the nasoglabellar angle was significantly different between the sexes (Figure 16.3), as males tend to exhibit more projecting nasal bones and a more developed glabellar region than females. The implication being that, in White-type skulls, dolichocephaly can result in a more obtuse, and brachycephaly in a more acute, gonial angle (Figure 16.2) regardless of sex.

The width of the mouth in repose corresponds to radiating lines from the lateral borders of the canines (Fedosyutkin and Nainys, 1993), which appears to be a functional pattern relating to different types of bite and the baring of appropriate teeth, e.g. the incisor bite in contrast to the canine tearing action. Different types of dental occlusion/malocclusion are linked to different mouth and chin shapes in profile view (Figure 16.4). A recent pilot study using lateral

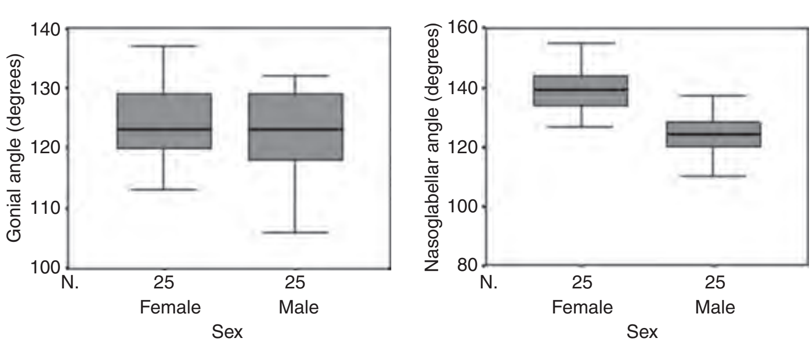


Figure 16.3 Box plots to show sexual dimorphism in gonial angle and nasoglabellar angle between 25 male and 25 female North American White adults (from Rynn, 2006).

cephalograms of dental patients (Utsuno *et al.*, 2010) found midline tissue-depth pattern to vary between type I ('normal'), type II ('overjet') and type III ('underbite') occlusion. 'Underbite' is linked to deeper tissue above the mouth and shallower tissue below, compared with type I 'normal' occlusion. 'Overjet' is linked to the opposite: shallower tissue above and deeper tissue below the mouth. Extreme 'overjet' can lead to incompetent lips, meaning the mouth is slightly open when relaxed. The muscles of the chin, particularly mentalis, become hypertonic, therefore bulkier, from the action of lip closure, which could in part explain this pattern: functional compensation on the part of the soft tissues.

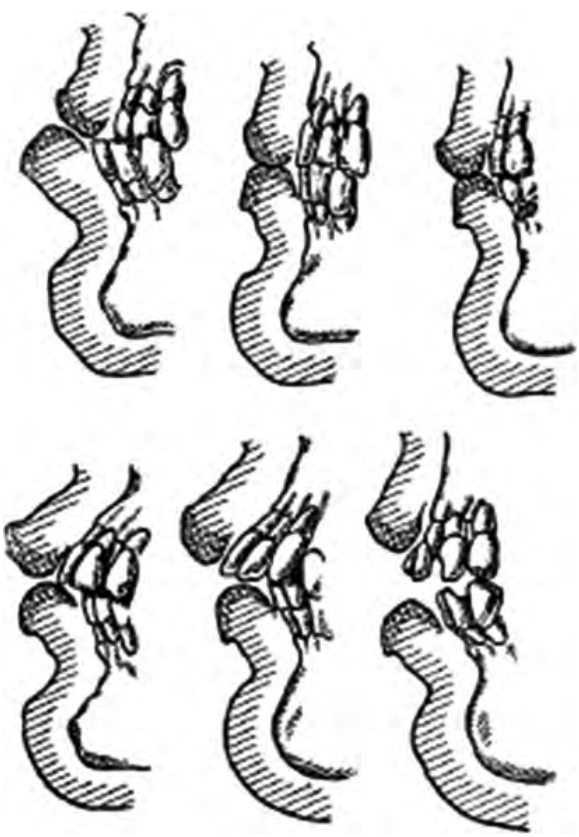


Figure 16.4 Different forms of bite and corresponding forms of the lips in profile (from Gerasimov, 1955).

The upper part of the nose correlates to the position of the ethmoid bone relative to the position and morphology of zygomatics (Figure 16.5). The cartilaginous nose appears to have evolved to maintain the sense of smell (olfaction) as the enlarging brain and rotating face resulted in a much smaller ethmoid area, orientated downwards rather than forwards, and in the superior surface of the nasal cavity rather than the posterior, relative to lower mammals.

However, the overall appearance of the nose is more related to the lower face than the upper face. Schultz (1918) postulated the level of prognathism/retrognathism to be linked to the nasal profile. He suggested that more prognathic faces appear to exhibit shorter, flatter, wider noses, and more retrognathic faces exhibit longer, more projecting, narrower noses. Schultz linked this effect directly to ancestry, because people of Sub-Saharan African ancestry are much more likely to exhibit prognathism than people of European ancestry, who tend to be more retrognathic, and nose shape varies accordingly. Dolichocephaly is common in Sub-Saharan African populations, yet is not linked to the same facial form as the White type because there is a higher degree of prognathism also common in African populations: wider teeth in a larger palate, providing extra room

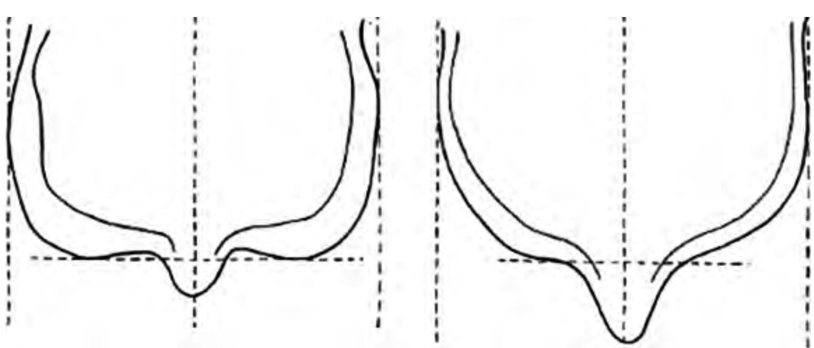


Figure 16.5 Transverse section of characteristic brachycephalic (left) and dolichocephalic (right) faces to illustrate configuration of the ethmoid and zygomatics. Tissue depth pattern over the cheekbones appears to be dependent upon the degree of undulation of the bone. In this illustration, both faces are of equal breadth, yet the dolichocephalic appears narrower (from Gerasimov, 1955).

for the structures of the oropharynx to function without necessitating the same extent of displacement and remodelling to maintain function. So the range between dolichocephaly and brachycephaly does not completely explain the shape of the face, but is merely one of myriad variables which have an effect on overall facial pattern in a holistic manner, particularly apparent in the White-type skull. If true patterns of correlation between the skull and soft tissues are to be found and employed in craniofacial reconstruction, it is more useful to examine datasets containing participants of diverse ancestry. The more ancestral diversity, the better are the chances of observing a wider range of variation, and thus extrapolating truer regression equations, which can be employed in facial reconstruction regardless of ancestry.

Recent research on the estimation of the nose (Rynn *et al.*, 2010) exploited the link between prognathism/retrognathism and nasal form, using the nasion-prosthion (NP) line as a vertical rather than the FHP as a horizontal alignment plane (Figure 16.6). It is apparent that rotation into the NP plane (NPP) (which is the nasion and prosthion line, perpendicular to the midline sagittal plane) corresponds to the postulated backwards tilt of the retrognathic dolichocephalic White-type skull (Mew, 1992). Conversely, the prognathic skull, which happens to belong to an African American (Figure 16.6) rotates downwards into the NPP from the FHP, which does not necessarily happen in terms of posture in life, but does better represent the actual projection of the nose from the face. Utilising the NPP as an axis of measurement for the nose takes into account prognathism/retrognathism and to some extent the upturned/downturned nature of the nose, so literally removes these variables from the equation. Coincidentally, on average, the soft tissue subnasale point lies level with the corresponding subspinale point when the NPP is vertical (Rynn, 2006), but not when the FHP is horizontal. In practice, the skull can be physically aligned in the FHP while nasal measurements are taken and employed in the NPP. The NP line has long been utilised in Russian methods of graphical prediction of the nasal profile (Gerasimov, 1955; Balueva *et al.*, 1988, 2004, 2009; Prokopec and Ubelaker, 2002) in that nasal profile is estimated by essentially mirroring the profile of the aperture about a line parallel to the NP line, transposed to touch the rhinion.

Glanville (1969) postulated a link between prognathism and nasal breadth. Otuyemi and Noar (1996) found people of Sub-Saharan African (specifically Nigerian) ancestry to consistently exhibit thicker dental enamel and mesiodistally wider teeth than people of European (in this case, British) ancestry, which would be echoed in the proportions of the prognathic palate. Maximum nasal width (MNW) was found to be related to maximum aperture width (MAW) on a sample which included White North American, African American and Central Asian American subjects (Rynn *et al.*, 2010) with a correlation of 0.58 ($p < 0.01$) when all ancestries were included, which dropped to 0.36 ($p < 0.01$) when the larger White North American group was isolated. The best approximation appeared to be the three-fifths rule applied across all ancestries: i.e. [MAW/3]*5 is approximately MAW. Intercanine distance, at the level of the tooth crown, was found to be linked less strongly to MNW since the correlation of 0.58 ($p < 0.01$) when all ancestries were included dropped to an insignificant 0.02 ($p > 0.05$) when the White North American group was isolated. However, when intercanine distance is taken as between the roots of the canines at the level of subspinale, a relationship strong enough to be used for MNW prediction was found in a White Central European group (Balueva *et al.*, 2009) with statistically significant correlations of 0.48 (Lithuanian males) and 0.53 (Lithuanian females), which are high considering the lower level of ancestral and sexual diversity in the samples. Regression revealed the following equations:

> Male Nasal Breadth = 18.035 + 0.44400 [Intercanine distance*]
> Female Nasal Breadth = 17.390 + 0.42393 [Intercanine distance*]
> *Intercanine distance as measured between the most prominent points on the canine alveolar bone at the level of subspinale.
> (Balueva *et al.*, 2009).

Several other features of the nose have been shown to be directly linked to the skull. The alae lie approximately 5 mm anterior and inferior to the aperture border, with the superior part of the alar groove directly linked to the inferior turbinate (christa conchalis) (Gerasimov, 1955; Balueva *et al.*, 2009; Rynn *et al.*, 2010). If employed in facial reconstruction, this pattern takes the asymmetry of the aperture into account automatically and transfers it to the alae appropriately. Conversely, lateral deviation of the bony nasal septum has been found to be linked to *contralateral* nasal deviation (Rynn *et al.*, 2010).

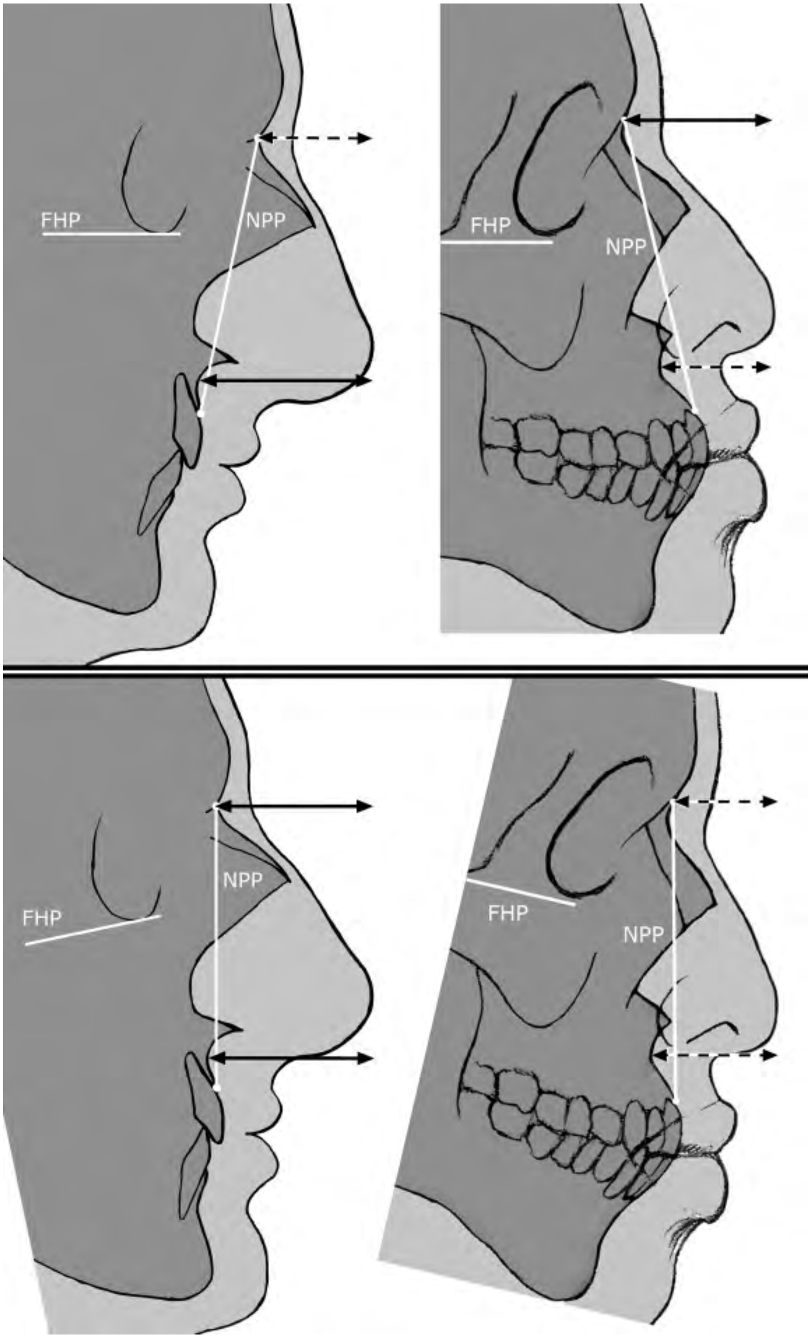


Figure 16.6 Upper row – retrognathic and prognathic facial profiles aligned in the FHP. Lower row – the same profiles aligned in the nasion-prosthion plane (NPP). Dotted lines show relatively small projection, solid lines show relatively large projection. The NPP as a vertical alignment plane enables fairer analysis of nasal projection.

This makes sense if functional patterning is postulated: if the septum deviates markedly to the left, the right aperture is larger, so the right nostril and lateral nasal cartilage will be larger, making the nose appear to deviate slightly to the right.

16.5 The eye

Eye morphology can be estimated on anatomical reconstruction rather than approximation via statistical regression. The palpebral ligaments tether the eyelids to the orbital bone, medially at the endocanthion, to a point on the 'frontal process of the maxilla situated on the medial orbital margin near its flexure, where the lower orbital wall sharply flattens' (Balueva *et al.*, 2009). The medial canthus of the eye tends to be positioned approximately 6 mm lateral to this point, with the lacrimal sac lying between, filling the lacrimal fossa. Tear drainage through the lacrimal puncta into the nasal cavity is mechanically facilitated by the lacrimal sac, which acts as a pump when squeezed and released, whether by the orbicularis oculi muscles when blinking away light tears, or more forcefully by rubbing the eyes. Laterally, Whitnall found the malar tubercle (Whitnall, 1921) to be definable, visually or by palpation, on 95% of skulls, regardless of ancestry (Wilkinson, 2004). This is the attachment point for the lateral palpebral ligament of the ectocanthion, allowing placement of the lateral canthus of the eye approximately 8 mm medial to the malar tubercle (Balueva *et al.*, 2009).

The shape of the eyelid fold echoes the shape the supraorbital border (Balueva *et al.*, 2009) (Figure 16.7). Quite a common configuration in people of Central Asian ancestry is an open orbit with a less developed margin, a relatively thick anterior lacrimal crest coupled with a flat nasal bridge, and projecting zygomatic bones, hence the term 'Mongolian' used in reference to the epicanthic eyelid fold linked to this bony relationship: 'epicanthic' because the fold curves downwards, crossing the inner canthus of the eye (Whitnall, 1921).

Similarly, with the face in repose, the eyebrows echo the form of the supraciliary arch. The lower edge of the eyebrow overlies the upper orbital edge, simulating its form (Balueva *et al.*, 2009). More developed brow ridges bulge above the eyebrow, lowering and straightening it medially, resulting in a more angular eyebrow shape. When the brow ridge is absent and the supraorbital border is rounded, the eyebrow

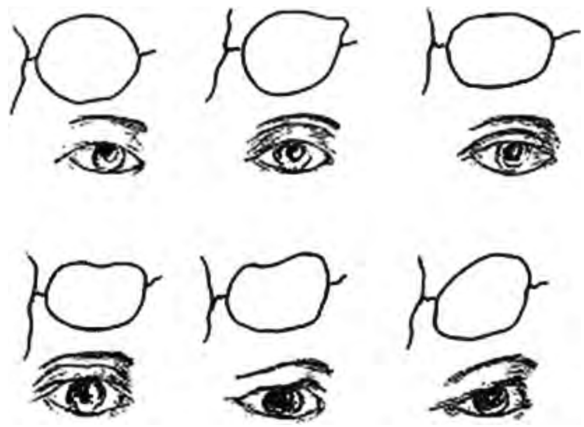


Figure 16.7 The form of the eyelid and brow related to the morphology of the orbit.

will be higher and more smoothly curved (Fedosyutkin and Nainys, 1993). Deep to the eyebrow, vertical fibres of the frontalis intermingle with horizontal fibres of the orbicularis oculi and oblique fibres of the corrugator supercilii. Corrugator supercilii is the depressor of the medial eyebrow, and is the last facial muscle to develop in utero (Gasser, 1967) deep to orbicularis oculi and frontalis. When corrugator supercilii is less developed or absent, the eyebrows remain level on lowering, as opposed to the medial part of the eyebrow being drawn down and medially into a frown (Whitnall, 1921). There is a connection between levels of testosterone, muscle mass and the degree of bony development. In this case, a more prominent bony brow ridge is linked to larger muscle mass and so, as in the cheeks over the zygomatic bones (Figure 16.6), there is considered to be a positive correlation between the degree of bony prominence of the brow ridge and increased depth of the local soft tissues (Gerasimov, 1955). This appears to be contrary to the soft tissue compensation of the mouth area in profile, indicated by the pilot study (Utsuno *et al.*, 2010) in which the soft tissue over the more prominent bone (maxilla or mandible) is shallower. However, if function is borne in mind, they have one and the same effect: increased muscle action related to function is linked to deeper soft tissues in both the brow ridge and the oral area.

The eyeball has been found to sit slightly (1–2 mm) superior and lateral of centre in the orbit (Whitnall, 1921; Stephan and Davidson, 2008) but these studies were carried out by dissection of embalmed cadavers, the soft tissues of which are likely to be somewhat distorted. Measuring clinical MRI data of living subjects, Wilkinson and Mautner (2003) found eyeball protrusion to be related to orbital depth, and stated that the iris (rather than the cornea as previously recommended) should be on a level with a straight line running from the mid-supraorbital to the mid-infraorbital border. The brachycephalic headform has the shorter anterior cranial fossa, hence characteristically shallower orbital cavities, and so a tendency towards more protrusive eyeballs (Enlow and Hans, 1996). Gerasimov (1955) found the morphology of the orbital border was also indicative of eyeball protrusion, with deep-set eyes being linked to well-developed, rugged, more 'closed' orbits, with overhanging supraorbital borders, and more 'open' orbits being linked to more protrusive eyes.

16.6 The ear

Relationships between the cartilaginous pinna and the skull are less obvious. Research (Balueva *et al.*, 2009) was carried out on samples comprising over 2000 living subjects from various ancestrally diverse Russian populations, using methods of palpation, surface marking and subsequent photogrammetry. For instance, at a certain inclination of the head at which the sternocleidomastoid muscle was relaxed, the mastoid process was palpated and marked. Subsequent photogrammetry demonstrated conformity between the inferior border of the ear lobe and the inferior border of the mastoid process.

The accepted canon in the graphic arts that ear height is correlated with the height of the nose was not confirmed. Ear height correlated more strongly with numerous anthropometric features than with either nasal height or length, but to varying extents in different Russian groups. Pinna morphology is based much more on cartilage than on bone structure. The morphological facial height from supraorbitale to gnathion was most strongly correlated with ear height in various Russian populations. In practice, the level of protrusion of the pinna is estimated in accordance with the level of development of the mastoid process of the temporal bone: a protruding ear echoing the morphology of a strongly developed mastoid process which bulges laterally (Gerasimov, 1955).

16.7 The facial surface

Averaged soft tissue-depth data employed in craniofacial reconstruction is a useful guide, providing as it does a series of points, hovering over the surface of the skull, but it is the contours between these points which make the face. Different levels of convexity, flatness or concavity between the tissue-depth markers can produce quite different appearances in faces with identical feature morphology and position, which tend to look either more or less convincing as a face, rather than like different faces. The bony surface of the cranial and orbital area provides us with fairly appropriate facial contours, but an awareness of anatomical interrelationships is necessary to predict and produce an appropriate approximation of the face in areas of deeper tissue around the cheeks and mouth region, and to some extent, the nose. The muscles of mastication, particularly temporalis and masseter, are strong and have bulk which gives shape to the face. So, if a relatively deep temporal fossa is exhibited, it stands to reason that the temporalis muscle which occupies this fossa will be relatively larger, and sculpture of the muscle with respect to the local morphology of the skull would produce a more realistic contour than would strict adherence to the average tissue depth marker, which would be lost beneath the deeper than average soft tissue. This could also apply to the mandibular ramus: if the mandibular ramus of a carnivorous quadruped is observed for exaggerated effect, it is deeply concave in harmony with the massive, bulging masseter. In the human face, gonial flaring and a relatively large lateral distance between the mandibular ramus and the zygomatic arches would logically suggest a more developed and bulkier masseter, hence, deeper than average tissue. In muscular areas such as these, if the skull surface was at the concave end of the range, it would be expected that the relevant muscles would be correspondingly more developed and have more bulk, and the facial surface would remain flat to convex. However, strict adherence to the average tissue-depth data would produce an unnatural, anatomically inappropriate concavity on the facial surface. Thus, some deviation one way or the other from the average tissue depth data may be justifiable, and it is inevitable in certain areas when a face is reconstructed anatomically.

The zygomatic major muscle defines the cheek, in that it is the underlying line which divides the anterior surface from the lateral surface of the cheek. In more dolichocephalic skulls, zygomatic major attaches more laterally to the zygomatic bone, whereas in brachycephalic skulls, it attaches more anteriorly (Gerasimov, 1955), but it is not the actual position of muscle attachment which is so different, but rather the configuration of the zygomatic bones (Figure 16.5). Zygomatic major inserts at the modiolus, along with depressor anguli oris and levator angulis oris. The modiolus is a chiasma: a knot or mesh of interwoven muscle fibres, responsible for the slight bulge of the upper lip over the lower in the area between the corner of the mouth and the nasolabial crease. Being the point of convergence for many facial muscles, the modiolus is very mobile in the animated living face, as it is pulled around into different facial expressions, changing the shape of the mouth. Depressor labii inferioris lies inferior to the lips, approximately in the areas which often remain bald even in the bearded, pulling the mouth down, as the name indicates. Depressor anguli oris is a triangular muscle, inserting at the modiolus and attaching to the mandible; it pulls down the corners of the mouth, as the name indicates. In the pentagonal area defined by the nasolabial creases and approximately by the lateral borders of depressor anguli oris, muscles lie between the facial surface and the skull, and the lack of substantial fatty tissue means logically that the facial surface relates in contour to the muscles, which can be modelled with reference to the dental occlusion and level of development of the local bone. The tissue depth of the cheeks outside of this pentagonal area is subject to the influence of the much more variable buccal fat pads, and so, average tissue-depth markers are much more useful in these areas regarding forensic facial reconstruction in the absence of any scene evidence of body mass index or facial fatness.

The facial muscles and the various ligaments which tether the face to the skull, or rather the skull to the face (at least throughout development) have relatively predictable and sometimes palpable points of attachment on the skull. Skull morphology, both local and holistic, indicates muscle morphology in some areas and attachment position in others. Since the facial muscles have such an apparent effect on the face, particularly when animated in facial expression, it makes sense to estimate and sculpt them using the skull as reference, in order to produce a more convincing, anatomically plausible face when the only evidence available is the dry skull.

References

Balueva, T. S., and Veselovskaya, E. V. (2004). New developments in facial reconstruction. *Archaeology, Ethnology and Anthropology of Eurasia*. Novosibirsk, **1**, pp. 143–152. [Russian and English].

Balueva, T., Veselovskaya, E. V. and Kobyliansky, E. (2009). Craniofacial reconstruction by applying the ultrasound method in live human populations. *International Journal of Anthropology*, **24**(2), 87–111.

Balueva, T. S., Veselovskaya, E. V., Lebedinskaya, G. V. and Pestrjakov, A. P. (1988). Anthropological types of the ancient population at the territory of USSR. In A. A. Zubov (Ed.), *Science*, Moscow. [In Russian].

Enlow, D. H., and Hans, M. G. (1996). *Essentials of Facial Growth*. Philadelphia, PA: W. B. Saunders.

Enlow, D. H., Kuroda, T. and Lewis, A. B. (1971). The morphological and morphogenetic basis for craniofacial form and pattern. *Angle Orthodontist*, **41**, 161–188.

Fedosyutkin, B. A., and Nainys, J. V. (1993). The relationship of skull morphology to facial features. In M. Y. İşcan and R. P. Helmer (Eds.), *Forensic Analysis of the Skull: Craniofacial Analysis, Reconstruction, and Identification*, pp. 199–213. New York, NY: Wiley-Liss.

Gasser, R. F. (1967). The development of the facial muscles in man. *American Journal of Anatomy*, **120**, 357–375.

Gerasimov, M. M. (1955). *The Reconstruction of the Face from the Basic Structure of the Skull*. Translated by W. Tshernezky. Russia: Publishers unknown.

Glanville, E. V. (1969). Nasal shape, prognathism and adaptation in man. *American Journal of Physical Anthropology*, **30**, 29–38.

Hooton, E. A. (1946). *Up from the Ape*. Revised Edition. New York, NY: The Macmillan Company.

Krogman, W. M., and İşcan, M. Y. (1986). *The Human Skeleton in Forensic Medicine*. Springfield, IL: Charles C. Thomas.

Mew, J. (1992). Use of the indicator line to assess maxillary position. *Cranio-View: The Journal of the Cranio Group and the Society for the Study of Craniomandibular Disorders*, **1**, 22–25.

Moss, M. L., and Young, R. W. (1960). A functional approach to craniology. *American Journal of Physical Anthropology*, **18**, 281–292.

Otuyemi, O. D., and Noar, J. H. (1996). A comparison of crown size dimensions of the permanent teeth in a Nigerian and a British population. *European Journal of Orthodontics*, **18**(1), 623–628.

Prokopec, P. and Ubelaker, D. H. (2002). Reconstructing the shape of the nose according to the skull. *Forensic Science Communications*, **4**, 1–4.

Rynn, C. (2006). Craniofacial approximation and reconstruction: tissue depth patterning and the prediction of the nose. PhD dissertation, University of Dundee.

Rynn, C., Wilkinson, C. M. and Peters, H. L. (2010). Prediction of nasal morphology from the skull. *Forensic Science, Medicine and Pathology*, **6**, 20–34.

Scheuer, L. and Black, S. (2000). *Developmental Juvenile Osteology*. London: Academic Press.

Schultz, A. H. (1918). Relation of the external nose to the bony nose and nasal cartilages in whites and negroes. *American Journal of Physical Anthropology*, **1**(3), 329–338.

Stephan, C. N., and Davidson, P. L. (2008). The placement of the human eyeball and canthi in craniofacial identification. *Journal of Forensic Sciences*, **53**, 612–619.

Tessier, P. (1971). Relationship of craniostenoses to craniofacial dysostoses, and to faciostenoses: a study with therapeutic implications. *Plastic and Reconstructive Surgery*, **48**(3), 224–237.

Utsuno, H., Kageyama, T., Uchidac, K. *et al.* (2010). Pilot study of facial soft tissue thickness differences among three skeletal classes in Japanese females. *Forensic Science International*, **195**(1–3), 165.e1–165.e5.

Whitnall, S. E. (1921). *Anatomy of the Eye and Orbit*. Oxford: Oxford Medical Publications.

Wilkinson, C. M. (2004). *Forensic Facial Reconstruction*. Cambridge: Cambridge University Press.

Wilkinson, C. M., and Mautner, S. A. (2003). Measurement of eyeball protrusion and its application in facial reconstruction. *Journal of Forensic Science*, **48**(4), 1–5.

Chapter 17

Automated facial reconstruction

Dirk Vandermeulen, Peter Claes, Sven De Greef, Guy Willems, John Clement and Paul Suetens

17.1 Introduction

Craniofacial reconstruction (CFR) can be a useful tool in the identification of a corpse that is unrecognisable due to its state of decomposition, soft-tissue mutilation or incineration, and if no other identification evidence is available. Traditional CFR methods are based on manual reconstruction by physically modelling a face on a skull replica with clay or plasticine. The progress in computer science and the improvement of medical imaging technologies during recent years have fostered the development of fast and flexible computer-based reconstruction programs. This chapter presents a comprehensive review of computerised three-dimensional (3D) CFR algorithms with particular emphasis on the more recent statistical based reconstruction methods, including a discussion of the various alternatives and problems that arise during the process of designing a CFR program.

Forensic identification of an unknown body is based on comparisons of ante- and post-mortem data, such as medical files, dental records, X-rays or DNA. Sometimes ante-mortem records are not available or incomplete; sometimes the preservation of DNA may be so poor as to make comparisons impossible. Similarly, when identifying features are completely missing or have decomposed beyond recognition but partial or complete skull fragments are available, CFR may assist the investigation. The goal of CFR is to recreate a likeness of the face of an individual immediately prior to their death. Different 2D and 3D manual or computer-aided CFR techniques have been developed for this purpose and all are based on the assumed relationship between the soft-tissue envelope and the underlying skull substrate.

Several 3D manual methods for CFR are currently being used in practice. These reconstructions consist of physically modelling a face on a skull replica (the target skull) with clay or plasticine. In the literature, such reconstructions have been commonly classified as 'Russian', 'American' or 'Combination'. (See Stephan (2006) for a discussion on this classification.) The Russian anthropologist Gerasimov (1971) was one of the first to make a manual reconstruction by modelling the complete anatomy of muscles and soft tissues covered by a thin layer, mimicking skin, onto the skull. This anatomical or 'Russian' technique, also referred to as the *morphoscopic* technique, was further refined by Lebedinskaya *et al.* (1993). During the same period, an alternative technique was developed in the USA, called the *morphometric* method (Snow *et al.*, 1970). This technique consists of building the soft-tissue layer in bulk, without much regard to the detail of the underlying anatomy, approximating tabulated average tissue depths at a sparse set of landmarks on the skull and interpolating in between. More recently, Richard Neave (Prag and Neave, 1997) used both Russian and American methods, laying the foundation for the *combined* technique, which was further developed by Caroline Wilkinson and her team at Manchester (Wilkinson, 2004). Proponents of this method claim that, since the face is reconstructed according to anatomical principles, artistic subjectivity in areas between tissue-depth landmarks is reduced (Wilkinson, 2010).

Manual reconstruction methods, however, require a high degree of anatomical and sculptural expertise and as a result remain difficult and subjective. The interpretations of two different artists can result in the creation of two substantially different faces from the same skull, as is illustrated in Figure 17.1.

Craniofacial Identification, ed. C. Wilkinson and C. Rynn. Published by Cambridge University Press.

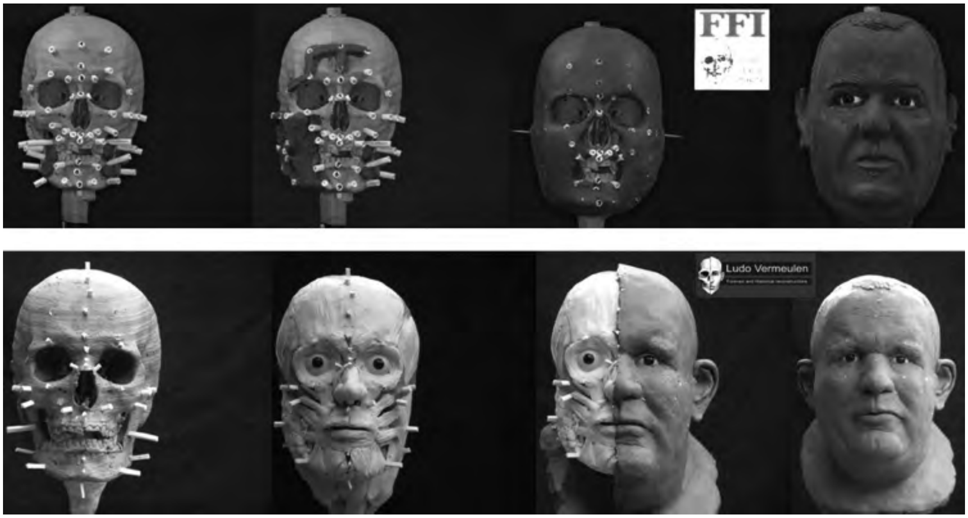


Figure 17.1 Manual reconstruction using (top) morphometric technique (courtesy of Ben Claes, FFI, Federal Police, Belgium) and (bottom) combined technique (courtesy of Ludo Vermeulen).

According to Davy and colleagues (Davy *et al.*, 2005) this point is further illustrated in the work by Haglund and Reay (1991), in which multiple facial reconstructions of several victims from the Green River serial killer were created. The results were highly variable between practitioners and little success was achieved when the reconstructions were shown to the public in an attempt to provoke recollections from people who might have known the deceased during life. Furthermore, these reconstructions are relatively time consuming, and are therefore often limited to a single reconstruction.

The progress in computer science and the improvement of medical-imaging technologies during recent years has led to the development of alternative computer-based CFR methods. A computer, compared with a human expert, is consistent and objective. Knowing all the modelling assumptions and given the same input data, a computer always generates the same output data. Furthermore, certain procedures can be automated such that the creation of multiple reconstructions from the same skull using different modelling assumptions (age, BMI, ancestry, gender, etc.) becomes possible. As a result, the CFR process becomes accessible to a wide range of people without the need for extensive expertise. An additional advantage of using computers is the ease of visualisation. The skull and the reconstructed face can be visualised simultaneously by making the face transparent, such that the face–skull relationship can be examined and, if necessary, corrected (Figure 17.2). The development of software for computerised facial reconstructions of an individual would be of benefit to various law enforcement agencies, by allowing faster, easier and more efficient generation of multiple representations of an individual.

The purpose of this chapter is to give a structured review of methods that have been proposed for computer-based CFR. A quasi-chronological list of the techniques is given in Table 17.1. We propose a general model framework for automated CFR methods and use it to analyse the published implementations and discuss the various modelling assumptions. We will also limit the discussion to computerised procedures that involve little manual input and thus exclude the computer-assisted manual sculpting procedures (Davy *et al.*, 2005) (utilising haptic feedback; Wilkinson *et al.*, 2006). This review not only updates the review presented by Wilkinson (2005) by including more recent contributions, but also focuses more on the engineering aspects.

17.2 CFR as a multivariate regression problem

Computerised techniques for CFR were discussed by Claes *et al.* (2010*b*) based on a model-based workflow, virtually mimicking manual reconstruction techniques. In this chapter, we will take an information-processing point of view by formulating craniofacial reconstruction as a multivariate data regression problem where a facial shape (the dependent output data) needs to be inferred

Table 17.1 Enlisting of 3D computerised CFR techniques in a quasi-chronological order.

Method	Reference	Craniofacial model (CFM)			Unknown Skull Representation
		Craniofacial template	Intra-subject mapping	Inter-subject mapping	
Vanezis	Vanezis (1989)	Single/Specific	Face/Tissue thicknesses	Generic/Non-uniform Scaling	Sparse
	Vanezis *et al*. (2000)	Single/Specific	Face/Tissue thicknesses	Generic/-	Sparse
Evenhouse	Evenhouse *et al*. (1992)	Single/-	Face/Tissue thicknesses	Generic/Polygon based deformations	Sparse
Evison	Evison (1996)	Single/Specific	Face/-	Generic/-	-
	Evison (2000)	Single/Specific	Face/-	Generic/-	-
Micheal	Michael and Chen (1996)	Single/Specific	Face/-	Generic/Volume Distortion Functions	-
Shahrom	Shahrom *et al*. (1996)	Single/Generic	Face/Tissue thicknesses	Generic/-	Sparse
Archer	Archer (1997); Archer *et al*. (1998)	Single/Generic	Face/Tissue thicknesses	Generic/Radial Basis Functions	Sparse
Quatrehomme	Quatrehomme (1997)	Single/Specific	Face/Skull	Generic/Radial Basis Functions	Dense/Crestlines
Nelson	Nelson and Michael (1998)	Single/Specific	Face/-	Generic/Local cylindrical coordinate	Dense/Feature Points
Attardi	Attardi *et al*. (1999)	Single/Specific	Face/Skull	Generic/Diffused Scattered Motion fields	Sparse
Bullock	Bullock (1999)	Single/Generic	Face/Tissue thicknesses	Generic/Radial Basis Functions	Sparse
Jones	Jones (2001)	Single/Specific	Face/Skull	Generic/-	Dense/Feature Points
Kahler	Kähler *et al*. (2003)	Single/Generic	Face/Muscles	Generic/-	Sparse
Claes	Claes *et al*. (2004*a*, 2004*b*, 2005*a*, 2005*b*, 2006*a*, 2006*b*)	Multiple/ Generic	Face/Tissue thicknesses	Face Specific/PCA	Sparse
	Claes (2007); Claes *et al*. (2010*a*)	Multiple/ Generic	Face/Tissue thicknesses	Face Specific/PCA	Implicit/Signed Distance Transform
Vandermeulen	Vandermeulen *et al*. (2005*a*, 2005*b*)	Multiple/ Specific	Face/Skull	Generic/Digital Cosine Transformations	Implicit/Signed Distance Transform
	Vandermeulen *et al*. (2006)	Multiple/ Specific	Face/Skull	Generic/Radial Basis Functions	Implicit/Signed Distance Transform

Table 17.1 (cont.)

Method	Reference	Craniofacial model (CFM)			Unknown Skull Representation
		Craniofacial template	Intra-subject mapping	Inter-subject mapping	
Pei	Pei *et al.* (2004)	Single/Generic	Face/Skull	Generic/Radial Basis Functions	
Andersson	Andersson and Valfridsson (2005)	Single/Generic	Face/Tissue thicknesses	Generic/-	Sparse
Berar	Berar *et al.* (2005*a*, 2005*b*, 2006)	Multiple/ Generic	Face/Skull	Face Specific/PCA	Dense/Feature Points
Davy	Davy *et al.* (2005)	Single/Generic	Face	Generic/Radial Basis Functions	Sparse
Muller	Muller *et al.* (2005); Mang *et al.* (2006)	Single/Specific	Face/Skull	Generic/Radial Basis Functions	Sparse
Subsol	Subsol and Quatrehomme (2005)	Single/Specific	Face/Skull	Generic/Radial Basis Functions	Dense/Crestlines
Tu	Tu *et al.* (2005)	Multiple/ Specific	Face/Skull	Generic/Radial Basis Functions	Dense/Range Image
	Tu *et al.* (2007)	Multiple/ Specific	Face/Skull	Generic/Radial Basis Functions	Dense/Range Image
Turner	Turner *et al.* (2005)	Multiple/ Specific	Face/-	Generic/Radial Basis Functions	Dense/Crestlines
Pei	Pei *et al.* (2008)	Single/Specific	Face/-	Generic/-	Dense/Range Image
Paysan	Paysan *et al.* (2009)	Multiple/ Generic	Face/Skull	Face Specific/PCA	Dense/Feature Points
Tilotta	Tilotta *et al.* (2010)	Multiple/ Specific	Face/Skull	Generic/Local semi-rigid	Implicit/ Extended Normal Vector Field

from skeletal information (the independent input data). Figure 17.3 illustrates this in a simplified way.

Given a particular representation of the input data, represented as a single variable on the x-axis, and a possibly different representation of the output data on the y-axis; it is possible to plot the joint representation, and thus the relation between a skull and the corresponding face, as a point in the x–y plane. Given many such points, corresponding to skull–face database exemplar pairs, it is possible to estimate a mapping from input space (the skulls on the x-axis) to the output space (the faces on the y-axis), for instance using a simple linear model as shown in Figure 17.3. This mapping can then be applied to an unknown skull to be reconstructed in a CFR scenario. Of course, this is an oversimplification of any existing CFR method. Furthermore, some of the CFR methods can only be loosely analysed in this strict mathematical framework. We will, however, use this concept to structure the discussion of the existing techniques in the rest of this chapter.

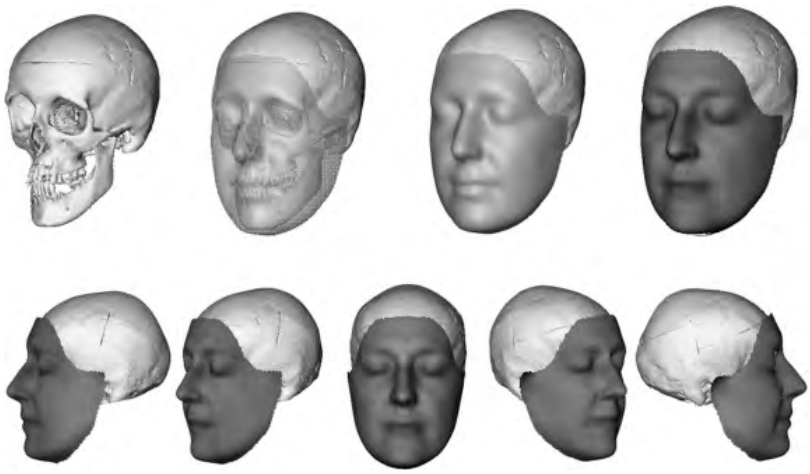


Figure 17.2 Computerised craniofacial reconstruction result shown on top of original skull with 3D rendering (different viewpoints and transparencies).

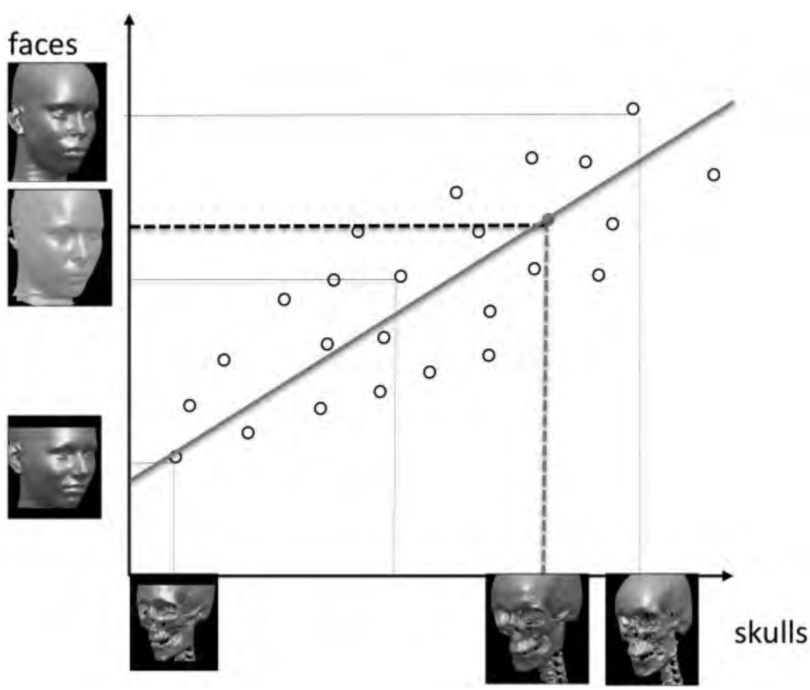


Figure 17.3 Craniofacial reconstruction as a multivariate regression problem.

Different methods of acquisition and representations for input and output data have been proposed and will be discussed in sections 17.3 and 17.4. The mapping from input to output space requires two sub-mappings. The first, intra-subject mapping, determines the (geometrical) relationship between a particular skull and the corresponding face and will be discussed in section 17.5. The second, inter-subject mapping, discussed in section 17.6, defines how to transpose such a geometrical relationship to another skull, be it the skull to be reconstructed, or a skull from a database used to define such mappings. In section 17.7 the representations and mappings are combined into a craniofacial model that is applied to reconstruct the face from a given skull. Issues involved in the graphical representation of these reconstructions are discussed in section 17.8. In section 17.9, we address the need for validation to increase the practical relevance of CFR methods in

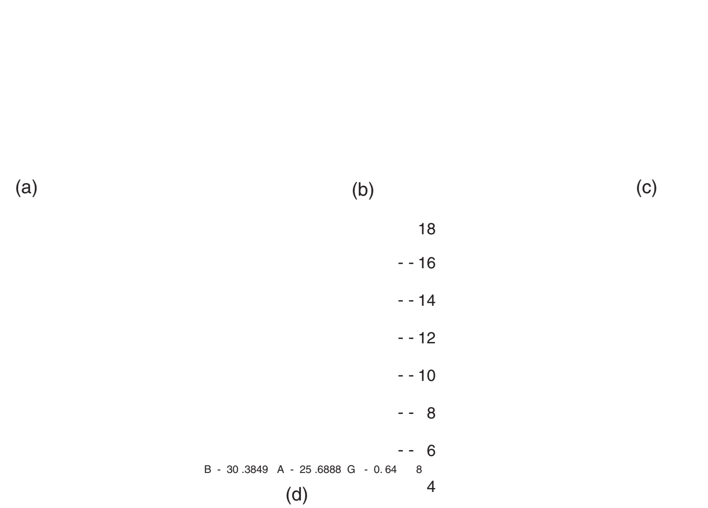


Figure 17.4 3D photogrammetric acquisition of facial shape and texture including ultrasound-based soft-tissue thickness measurement at facial landmarks (De Greef *et al.*, 2005). This figure is produced in colour in the plate section.

crime scene investigations. Finally, we conclude this review by suggesting future directions.

17.3 Data acquisition

A digitised version of the skull is required in order to code the skull shape into a machine-readable format for further processing. Moreover, for building the statistical regression model, skeletal as well as associated facial surfaces or soft-tissue thicknesses data from a matched population need to be acquired in a digital format.

Originally, skull and face surfaces were digitised using 3D laser scanning systems (Vanezis, 1989; Evison, 1996; Shahrom *et al.*, 1996; Evison, 2000; Davy *et al.*, 2005). Although laser scanning is a common and safe technique for scanning the outer-surface of 3D objects, the skull surface is very complex to reconstruct based on laser-line projections. For facial surfaces, these scanners give the possibility to scan a person in an upright lifelike position. Laser scanning or other 3D photogrammetric acquisition of database exemplars is obviously limited to the outer facial surface, thus requiring other technologies, such as ultrasound, to measure corresponding soft-tissue thicknesses (De Greef *et al.*, 2005). Figure 17.4 shows a 3D rendering, including texture as obtained by a photogrammetric procedure (ShapeCam, Eyetronics, Leuven, Belgium) and the ultrasound acquisition of soft-tissue thicknesses at the facial landmarks indicated in blue.

Thanks to recent advances in medical-imaging technology, computerised tomography (CT) scanners have become a practical alternative for acquiring a digital copy of the skull. All computerised reconstruction techniques today use CT scanners to digitise the unknown skull. Furthermore, the major advantage of also using a CT scanner to acquire database exemplars is the possibility to have both reference skull and face surface information acquired simultaneously and in the correct spatial relationship to each other (Figure 17.5).

A disadvantage however, is the level of irradiation absorbed by the subject during CT scanning, limiting the reference database to scans taken for diagnostic purposes from living patients or from deceased subjects. Furthermore, most conventional CT images are acquired of subjects in a horizontal, supine position. As a result, due to gravitational forces, facial shapes extracted from CT images will differ from the typical facial shape as viewed in an upright position (Figure 17.6), which is the normal pose for viewing, and hence recognising, faces.

Finally, CT scanners are sensitive to high-density material (e.g. amalgam) dental restorations abundant in current adult reference population subjects. These result in heavy streak artefacts in the images as shown in Figure 17.7, jeopardising the correct acquisition of either surface, without tedious or surface-modifying morphological corrections.

Recently, cone-beam CT (CBCT) scanners have been developed and are being used in medical practice. In contrast to conventional CT, CBCT scanners typically acquire images of subjects in an upright position.

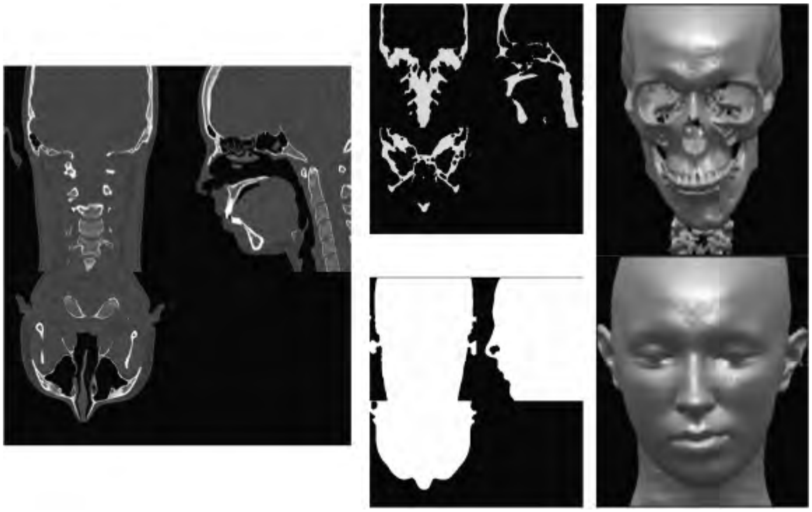


Figure 17.5 Left: cross-sectional viewing of head CT scan. Middle: Binary segmentations of head and skull volume. Right: 3D rendering.

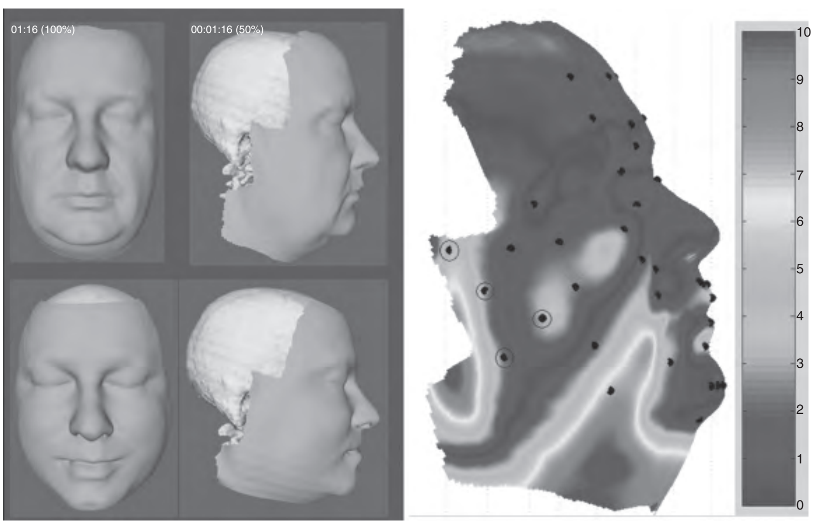


Figure 17.6 The facial geometry of a person in a horizontal supine position (bottom row) and an upright position (top row). Local surface differences visualised by means of a colour-code ranging from 0 to 10 mm (right image). This figure is produced in colour in the plate section.

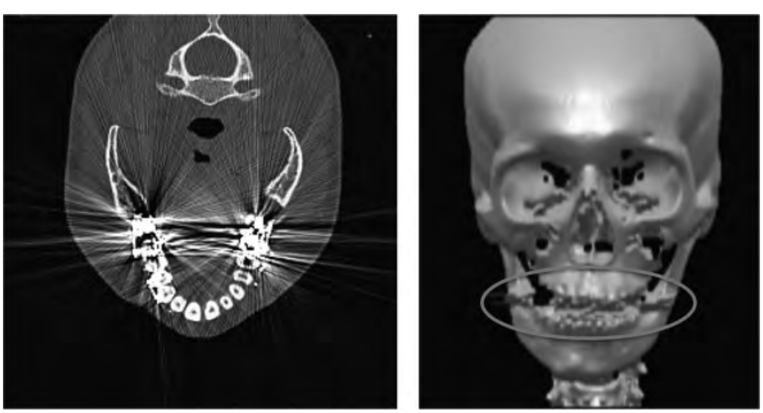


Figure 17.7 Left: axial cross-section through dental region showing artefacts caused by dental filling. Right: artefacts indicated on 3D rendering of skull surface segmentation.

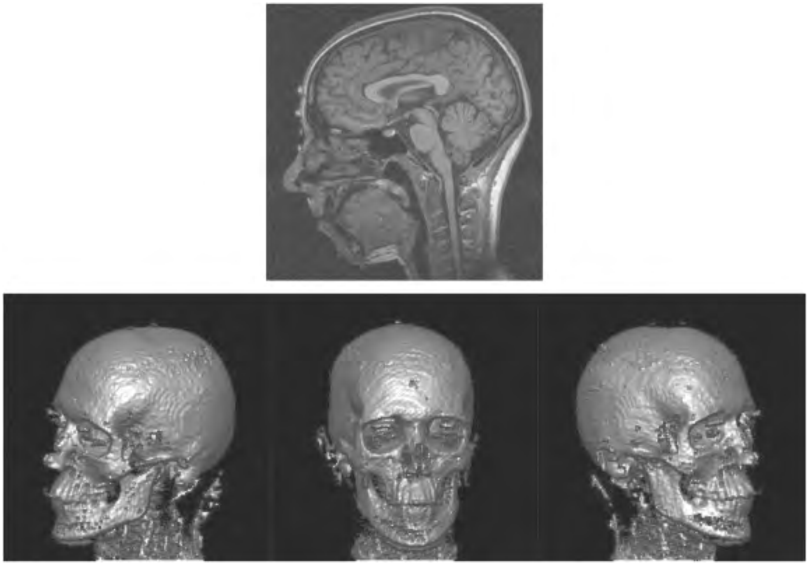


Figure 17.8 Failure of hard tissue segmentation in MR images of the maxillo-mandibular area.

Figure 17.9 Left: 3D rendering of dense point representation of skull. Middle: Sparse landmark representation. Right: Crest lines representation.

Furthermore, CBCT images can be obtained with a lower absorbed dose (Loubele *et al.*, 2009) at the cost of a slightly reduced image quality (Loubele *et al.*, 2008).

An alternative to CT scanners is the use of MRI scanners (Mang *et al.*, 2006; Paysan *et al.*, 2009) for soft-tissue imaging. In contrast to CT, MRI is considered not to be harmful. Soft-tissue depth measurements can be acquired, including differentiation between different types of soft tissue (muscle, fat). However, the same remark concerning the difference in facial shape and appearance between supine and upright position of the subject during scanning holds for MRI scanners. Furthermore, because of the poor hard-tissue visualisation in MRI (Figure 17.8) estimation of the hard tissue boundary is complicated and no robust solution for the complete facial region is yet available.

17.4 Data representation

Different data representations can be proposed for the skull and facial surfaces. Most often such surfaces are represented as a dense set of points, connected with edges, forming a triangular tessellation of the surface, which can be rendered as a 3D shaded surface as shown in Figure 17.9. In the procedures mimicking the American morphometric reconstruction procedures (Vanezis, 1989; Evenhouse *et al.*, 1992; Evison, 1996; Archer, 1997; Bullock, 1999; Vanezis *et al.*, 2000; Andersson and Valfridsson, 2005), the skull is only defined as a small set of a few tens of anatomical landmarks or feature points and associated surface normals (Figure 17.9). These points can be indicated manually or detected automatically. However, their often-ambiguous definition and their inaccurate localisation, even by automatic methods, is a major source of error (Nelson and Michael, 1998). Quatrehomme (1997) extended point sets to lines; more in particular crest lines defined as loci of local extremes of principal curvature. These lines follow the salient lines of the skull surface, such as the mandible, the orbits, the cheekbones or the temples (Subsol and Quatrehomme, 2005).

In a number of cases, surfaces are represented using a parametric model, such as a cylindrical map (Pei *et al.*, 2004) or a B-Spline representation (Archer, 1997). Surfaces can also be represented implicitly as the zero-valued iso-surface of a 3D signed distance transform (sDT) (Vandermeulen *et al.*, 2005*b*). An sDT of a closed surface is a function that associates to every point in 3D the shortest Euclidean distance to the surface, zero on the surface, positive inside and negative outside the volume encompassed by the surface. This *implicit representation* is a dense, continuous, one-to-one representation and does not only code the original surface, but simultaneously codes an infinite set of increasingly smoother surfaces inside and outside the surface away from the original surface (Figure 17.10).

Most of the facial representations used in computer-based CFR are holistic, representing a complete frontal view of the face. However, some represent the face as a collection of partial features like the nose, mouth and ears (Vanezis, 1989; Evenhouse *et al.*, 1992; Andersson and Valfridsson, 2005; Davy *et al.*, 2005; Tilotta *et al.*, 2010), which then need to be smoothly merged to improve realism.

17.5 Intra-subject (skull-to-face) mapping

The intra-subject (skull-to-face) mapping defines the geometrical relationship between a skull and the associated face of a single subject. We can make a distinction between two major approaches to model this relationship. In the first approach, the soft-tissue thicknesses are explicitly modelled as a function of the position along the skull surface. In the second approach, thicknesses are implicitly modelled by a joint representation of the Cartesian coordinates of point positions on both surfaces.

The first computer-based methods (Vanezis, 1989; Attardi *et al.*, 1999; Vanezis *et al.*, 2000) are a mere mimicking of the manual morphometric CFR methods and model the skull-to-face relationship or mapping as virtual dowels, representing soft-tissue thicknesses, at a sparse set of anatomical landmarks on the skull. The endpoints of the virtual dowels represent estimates of the associated landmarks on the face surface (Figure 17.11). In the work of Claes *et al.* (2006*b*) the virtual dowels are not defined on the skull but instead on a facial template. As a result, these dowel positions are linked to each other in a face-specific way.

In an attempt to formalise the concept of a soft-tissue 'layer' even further, some methods use a sparse soft-tissue thickness map defined on a 2D parameterisation of the skull and interpolate the soft tissues at intermediate positions. Andersson and Valfridsson (2005) use a cylindrical parameterisation of the skull in combination with bilinear interpolation of the sparse tissue thickness values, while Bullock (1999) uses a triangular tessellation of the skull and interpolates using barycentric interpolation. Instead of interpolating soft-tissue thicknesses from sparse point data, Pei *et al.* (2004) use dense tissue maps defined on a cylindrical map to represent the tissue thickness between skull surface points and face surface points measured along a radial direction perpendicular to the main cylindrical axis.

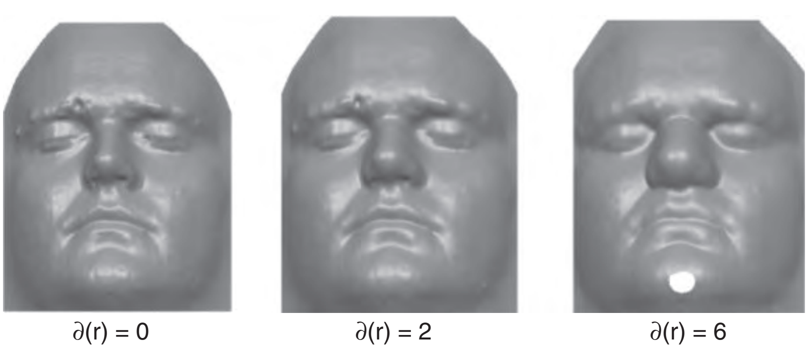


Figure 17.10 Left: zero-level isosurface of the implicit function corresponding to a facial surface. Middle and right: isosurfaces at distances of 2 and 6 mm to the facial surface.

Figure 17.11 Left: Skeletal landmarks defined on skull surface with associated normals. Right: skeletal landmarks defined via facial landmarks and inwards pointing normals to face (colour code represents soft tissue thicknesses at corresponding landmarks). This figure is produced in colour in the plate section.

In the more recent computerised approaches to CFR, the face-to-skull relationship is not modelled as a set of offsets of skull and facial surfaces via tissue thickness maps, but rather as a joint set of 3D Cartesian coordinates of points on both the skull and face surface (Nelson and Michael, 1998; Berar *et al.*, 2005*a*, 2006; Tu *et al.*, 2005; Turner *et al.*, 2005; Vandermeulen *et al.*, 2006; Paysan *et al.*, 2009; Tilotta *et al.*, 2010). The multivariate interrelationship of these coordinates is maintained via spatially continuous warping procedures.

Irrespective of the type of skull-to-face model, implicit or explicit, we clearly see a shift from sparse point-based to dense point-based, or even continuous intra-subject skull-to-face models. This can, of course, be explained historically, parallel to the availability of soft-tissue thickness data. These data were, for a long time, limited to sparse sets of craniofacial landmarks, given the difficulty in measuring them using callipers, needles or ultrasound devices. With the advent of CT imaging, in particular X-ray CT, thicknesses can now be estimated on the continuum of the skeletal base (Vandermeulen *et al.*, 2005*b*). Challenges are to measure these in a robust way given the presence of possible imaging artefacts and the complex skeletal topology of the face.

17.6 Inter-subject (skull-to-skull) deformation mapping

The inter-subject (skull-to-skull) deformation mapping describes how the skull-to-face information for one individual can be mapped to another individual. It mainly determines what type of geometric transformation is allowed or required for inter-subject matching of skeletal data and how this affects the associated skull-to-face mapping transfer. It is required for the building of statistical models for the skull-to-face information and is ultimately also applied to warp such information to the target skull.

In the original computerised CFR implementations (Vanezis, 1989; Evenhouse *et al.*, 1992) this transformation is based exclusively on the manually indicated skeletal facial landmarks that correspond to the tabulated tissue thickness values. A proper selection of tissue thickness values is set out as the length of the normal to the skull surface at the corresponding skull landmark locations. The endpoints of these virtual dowels serve as target points to which the associated facial landmark-points of a reference face need to be transformed. This transformation is extrapolated to the whole reference face using generic warping procedures based on Thin Plate Splines (TPS) (Vanezis *et al.*, 2000) or B-Splines (Archer, 1997). Pei *et al.* (2004) use generic TPS-based warping procedures to map a dense reference soft-tissue thickness map on a 2D cylindrical parameterisation of the target skull.

Alternatively, methods using the second type of skull-to-face mapping, based on the implicit coding of soft-tissue thicknesses, via an extrinsic 3D Cartesian coordinate representation of the skull and skin surface points, define the skull-to-skull transformation by deforming a complete reference skull surface to a target skull surface (Quatrehomme, 1997; Nelson and Michael,

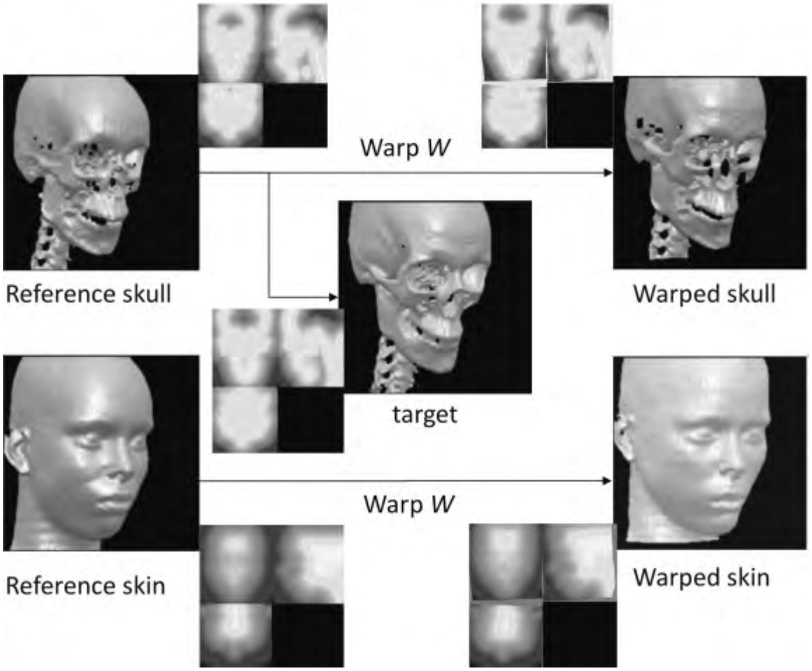


Figure 17.12 Reference skull (top left) is warped to target skull (centre) resulting in the warped reference skull (top right). This warping is applied to the reference head (bottom left) resulting in the warped reference head (bottom right). Transformations are calculated on the implicit signed distance function representations of the surfaces (small multi-slice windows). This figure is produced in colour in the plate section.

1998; Jones, 2001; Subsol and Quatrehomme, 2005; Tu *et al.*, 2005; Vandermeulen *et al.*, 2006) using, again, generic warping procedures such as TPS or B-Splines or local semi-rigid transformations (Tilotta *et al.*, 2010). The calculated skull deformation is extrapolated to the embedding 3D space and applied to the facial surface associated to the reference skull (Figure 17.12).

The use of an anatomically consistent, instead of a generic, non-rigid deformation mapping was first advocated by Claes *et al.* (2004*a*) and further elaborated in Berar *et al.* (2006) and Claes *et al.* (2006*b*, 2010*a*). Such a face-specific geometric transformation can be obtained by training or learning such mappings on a database of skulls and/or associated facial data.

In Claes *et al.* (2004*a*) the facial surface, for instance, is considered as an elastic mask, the elasticity of which is defined by the statistically allowed correlated variation of facial surfaces in a database. By changing the mask between the statistically determined boundaries, the deformation of the facial mask is restricted to take place in a face-specific way only.

Claes *et al.* (2006*b*) learn not only the statistical shape variation of facial surfaces but also the co-variation of these surfaces with the soft tissue thicknesses at a sparse set of landmarks. As a result, continuing the elastic mask analogy, elastic dowels are attached on the inside of the facial mask at those landmarks. The length and elasticity of these dowels are again based on the statistical analysis of the database and co-vary with the facial surfaces. The endpoints of these dowels correspond to craniofacial landmarks on the skull. As a result, in combination with the statistical facial shape model, they define the statistically allowed transformations between corresponding skull landmark sets. Since this method was originally developed using a database of 3D facial surfaces and ultrasound-based soft-tissue thicknesses measured at a sparse set of landmarks, the statistical model that relates the facial surface to the associated skull surface is limited to these landmarks. However, the methodology allows, given a database of almost continuous representations of skulls and faces as for instance extracted from CT, the extension to a dense statistical model of the interrelationship between skull and face surface points. This idea has been taken up by Berar *et al.* (2006) and Paysan *et al.* (2009), who model the co-variation of a dense set of points on both the skull and facial surface, extracted from a database of CT images of volunteers (Tilotta *et al.*, 2009) and combined with MRI, respectively.

The advantage of using generic deformations is that they require no training phase by the computer. However, care has to be taken when using generic deformations, because no knowledge of facial geometry and/or anatomy is incorporated. They are just

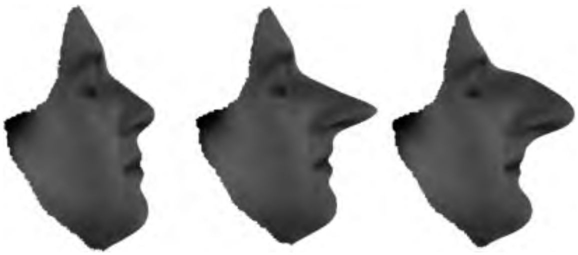


Figure 17.13 Nose tip manipulation of a face (left) with a generic TPS based deformation (middle) and a face-specific statistical-based deformation (right).

smooth (softly evolving) and as a result can deform a face into an unrealistic looking face when not used carefully (Figure 17.13).

The disadvantage of face-specific deformations, on the other hand, is that they are entirely dependent on the samples in the database. Having a small database or a database with low inter-subject variance generates a small and too restrictive a set of deformations, such that faces atypical to the database are hard to reconstruct. Furthermore, while it has not yet been clearly established how much ancestry-related information is contained in the soft-tissue structures, databases for different ethnic groups must be collected.

17.7 Craniofacial model fitting

In this section we discuss how the components described in the previous paragraphs are applied for reconstruction of a specific target skull. Every reconstruction uses a craniofacial model consisting of: (1) a computer representation of reference skulls and faces; (2) an intra-subject skull-to-face mapping model; and (3) an inter-subject skull-to-skull mapping model.

Upon application of such a craniofacial model for reconstruction of an unknown target skull, a specific craniofacial template has to be selected. This craniofacial template is a combination of a facial and skull surface representation with an intra-subject skull-to-face mapping model. This can be as simple as a set of skeletal landmark points (the skull surface representation) with associated soft tissue thicknesses (the intra-subject skull-to-face mapping) and a triangulated surface of a single individual or generic face (the face surface representation). Or it can be as complex as a joint dense skull and face point cloud representation of a single individual, where the skull-to-face mapping model is implicit in the joint representation.

The specification of the craniofacial template will strongly influence the final result. By selecting an appropriate craniofacial template the reconstruction can be tailored to the properties of the target skull, a procedure called *property normalisation*. The first step in a reconstruction, therefore, is the anthropological examination of the skull in order to determine/estimate age, gender, ancestry and stature (Reichs, 1992) or even body mass index (BMI) based on remaining soft-tissue layers on the skull or additional evidence found at the crime scene, for example the size of clothing.

Subsequently, one or more craniofacial templates are chosen/generated from the database. In the single template setup, only one reference template is used and this can either be a specific matching template (in terms of age, gender, ancestry and skull dimensions) or a representative generic template. In a multiple template setup, two approaches are also observed. The first or ‘specific’ approach generates a reconstruction per specific reference subject chosen in a database, resulting in multiple reconstructions of the unknown skull. Then the final result is generated by properly combining all the reconstructions into a single reconstruction. The second or ‘generic’ approach first integrates multiple reference craniofacial templates, which is then fitted to the target skull.

Property normalisation can be performed in a pre-reconstruction mode, where the craniofacial template is normalised to the skull properties before a CFR is made, or in a post-reconstruction mode, where only the reconstructed results are normalised to the estimated properties.

A typical pre-reconstruction normalisation approach is to select one or a set of reference heads from a database, the properties of which are similar to the ones of the skull. This is often done when using a single specific craniofacial template. However, an extended database containing enough samples for every possible sub-population that might be required is labour intensive to acquire. Alternatively, one can try to learn and model geometric facial variations originating from attribute differences between faces in a database in order to simulate age, corpulence, ancestry and gender changes. The advantage is that under-sampled sub-populations in a database can still be represented by the interpolating nature of the attribute learning process. A linear statistical interpolation of attribute differences is, for example, used in Claes *et al.* (2005*a*) and Tu *et al.* (2007) in a post-reconstruction mode and in Paysan *et al.* (2009) and Claes *et al.* (2010*a*) in a

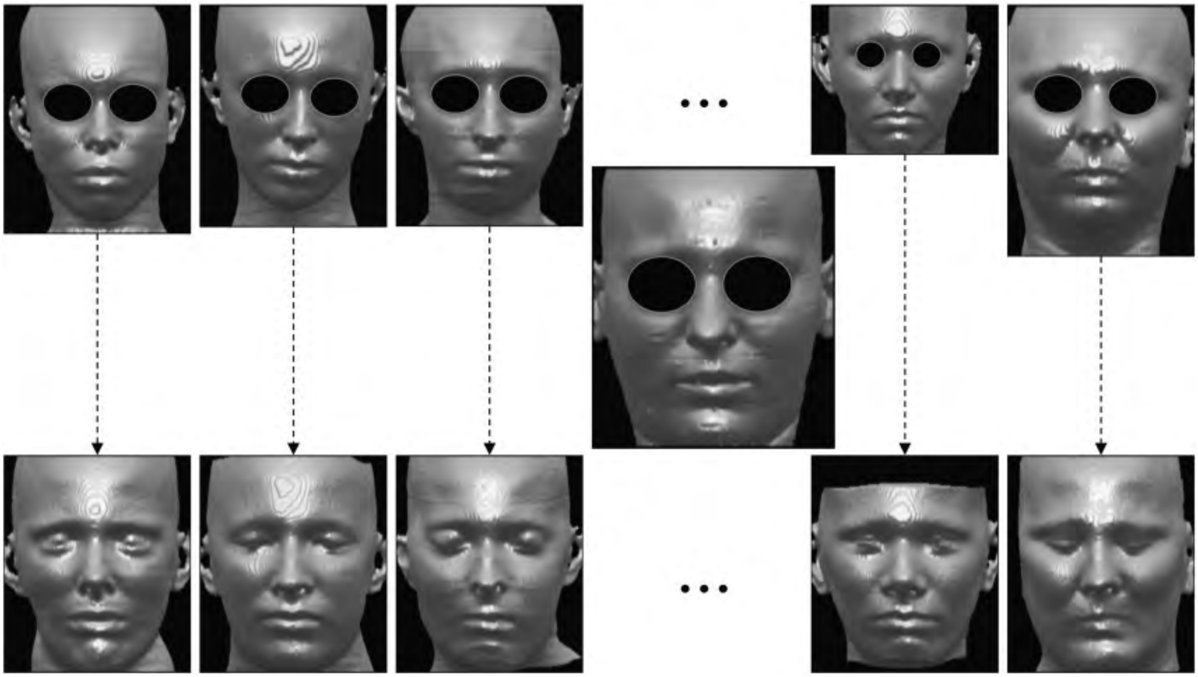


Figure 17.14 Reference bias. Top row: reference faces. Middle: target face. Bottom: Warped reference heads. Features of reference heads are still apparent after warping.

pre-reconstruction mode. A linear geometrical interpolation between two reconstructions with different skull properties is used to generate intermediate reconstruction results in Evison (2000).

If inappropriate templates are chosen, *model bias* can occur. Model bias is the leaking through of specific facial feature details, originating from the craniofacial template, into the reconstruction. Using only a single generic (for example an average face) facial template or a subject-specific facial template, the potential to produce model-biased reconstructions is high, which is illustrated in Figure 17.14. Indeed, when using a subject-specific template, based on similarity in ancestry, gender and age, unwanted facial features of the template remain visible in the final reconstruction. Using a generic face template, on the other hand, results in a reconstruction that is too smooth and non-specific. In order to reduce the model bias, it is better to work with multiple reference heads. It is still an open issue, though, how many reference heads are required for accurate reconstructions and how well these reference heads need to be attribute-matched to the target skull.

In the final step, the chosen craniofacial template is fitted to the target skull using the inter-subject skull-to-skull mapping model as detailed in section 17.6. During this stage, additional constraints may have to be imposed to deal with imperfect skull scan data, such as missing parts or acquisition artefacts (e.g. metal streak artefacts in CT). Mathematically, this can be obtained by including an additional regularisation model that further constrains the skull-to-skull mapping in the presence of noise or gross outliers. For instance, instead of perfectly interpolating the endpoints of the virtual dowels set on the target skull with a deformable facial mask, approximation of these endpoints can be imposed, resulting in smoother (generic) or more facial plausible (face-specific) deformations. Gross deviations, however, will unduly influence the deformation and will therefore have to be estimated and removed from the deformation estimation, a process called robust estimation (Claes, 2007) (Figure 17.15).

17.8 Texturing and visualisation

The information that is used to reconstruct the face from skeletal remains is exclusively based on the shape aspect of the face. Indeed, there is often little or no information from the skeletal remains about the skin texture. Generating a good approximation of the geometry of the face belonging to the unknown skull is thus the first objective of a CFR technique. 3D rendering of computer-based reconstructions of only the shape of the face results in sculpture-like representations. In order to generate a lifelike appearance, texturing the reconstruction (applying skin colour and pigmentation) might be required. 3D modelling software can be used to virtually paint the eyes and mouth onto the 3D surface of the reconstruction. This is also useful for creating particular details like scars or birthmarks when they are known to have been present. Another possibility is to use texture mapping as in Davy *et al.* (2005) and Subsol and Quatrehomme (2005), which is a process akin to applying wallpaper to a flat surface. The final textured reconstruction result can now be visualised with dedicated 3D-rendering software.

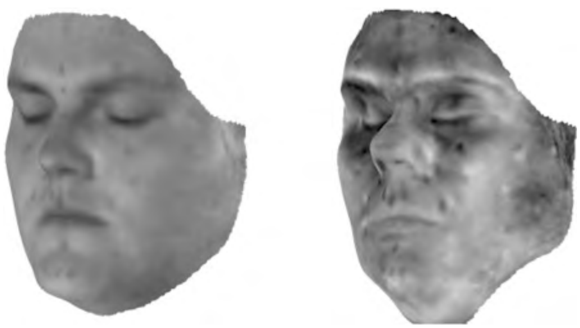


Figure 17.15 Left: A CFR example generated based on noisy skull landmarks with robust fitting of a statistical craniofacial model. Right: A CFR example generated based on the same noisy skull landmarks and statistical craniofacial model without a robust framework for model registration.

Texturing the reconstruction has to be applied cautiously, though, because it can trigger an incorrect recognition. Figure 17.16 indicates that 3D shape codes for a larger part of the information about the individuality of a face as compared with texture. Moreover, introducing a texture map from another individual introduces unwanted fine detail, decreasing the resemblance substantially. Instead, as illustrated, using an average texture map (even though not age- and gender-matched in this example) can improve the resemblance as compared with 3D shape alone.

It is necessary for average texture maps derived from different sub-populations to be created in the same way as shape models, for different sub-populations. Alternatively, Claes *et al.* (2010*a*) correlate texture information with face shape and tissue thickness information into a statistical model based on multiple faces. This gives the possibility to generate more, but not too specific-texture maps (compared with averaged-texture maps) according to the geometry of the face, because the relationship between texture and geometry is learned (Figure 17.17).

17.9 Validation

A final, but critical element in the design of a computerised CFR technique is the evaluation of the accuracy of the reconstructions obtained. Unfortunately, a rigorous validation is lacking in most of the studies

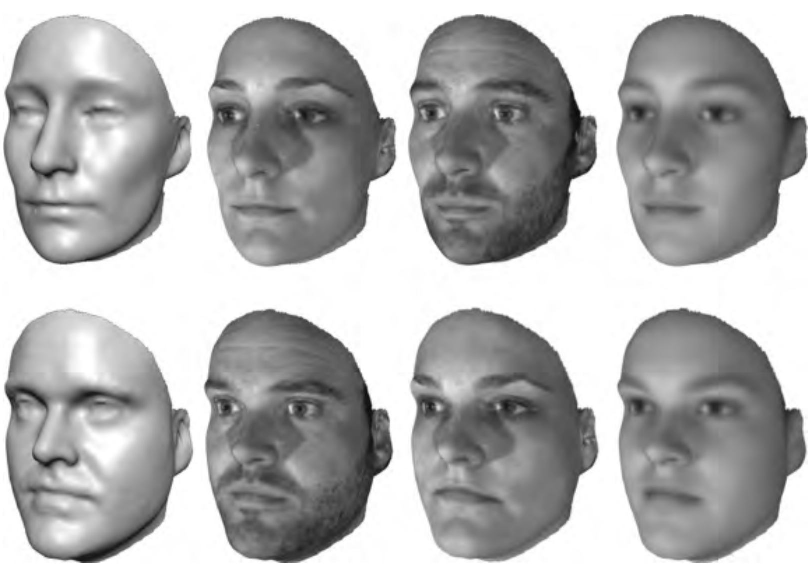


Figure 17.16 Shape versus appearance. First column: shaded surface of two individuals. Second column: corresponding textured surface. Third column: same, but texture of the two individuals switched. Last column: averaged texture from an adolescent database used in rendering. This figure is produced in colour in the plate section.

Figure 17.17 Top: 3D rendered view of test skull and ground truth 3D facial surface. Bottom: corresponding views of automated CFR result following the procedure defined by Claes and colleagues (Claes *et al.*, 2010*a*). Right: coloured representation of surface differences. This figure is produced in colour in the plate section.

5

0

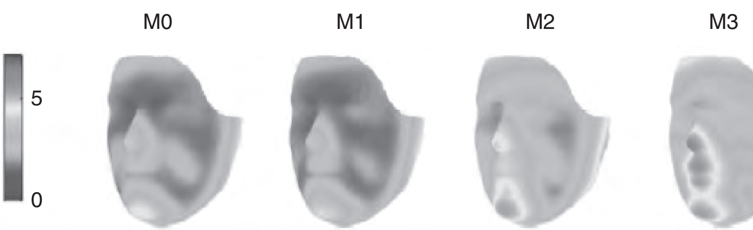


Figure 17.18 Averaged local RMSE (root mean square error) surface reconstruction results over 12 cases, depicted on the geometry of the average face from the database. M0: joint statistical model for face shape and soft tissue thicknesses. M1: separate statistical model for both. M2: generic Thin Plate Spline deformation of attribute normalised template. M3: TPS deformation of average template. This figure is produced in colour in the plate section.

involving computerised craniofacial reconstruction or is based on just a few examples of reconstructions qualified as successful based on subjective assessment of resemblance. Much of the information is anecdotal. More extensive resemblance-based validation studies, involving also unsuccessful cases, as well as quantitative studies are required if these CFR techniques are to become accepted in forensic science investigations and satisfy test of evidence in the courts (Wilkinson, 2005).

Computer-based validations are possible given a database of skulls with known facial appearance. A typical validation scenario can, for instance, be based on leave-one-out cross-validation. Here every skull in the database is removed in turn and used as a test case. The resulting facial skin surface of the reconstruction technique can then be compared with the form of the actual skin surface of the test case. The form of the actual skin surface is the golden standard or ground truth of the validation setup, enabling evaluations of the reconstruction result.

As a first instance, a quantitative reconstruction error evaluation can be performed, by observing local dimensional surface differences. Comparing the two surfaces in this quantitative way is interesting for evaluating the reconstruction performance in terms of accuracy and provides a spatial map of the difficulty of each facial region to be reconstructed (Berar *et al.*, 2006; Vandermeulen *et al.*, 2006; Claes *et al.*, 2010*a*). Figure 17.18 shows the differences as obtained by the method of Claes and colleagues (Claes *et al.*, 2010*a*), averaged over 12 test subjects for four different CFR methods, plotted on an average facial template. These results indicate that the use of face-specific mapping models (M0 and M1) outperform the generic models (M2 and M3).

However, the final goal of CFR is not reconstruction accuracy, but rather recognition or identification success. A recognition test consists of comparing the CFR result with a database of candidates including the ground truth or the actual person. The goal is to retrieve the ground truth from the database. In a computer-based recognition setup this can be accomplished automatically using facial measures of similarity (Vandermeulen *et al.*, 2006; Claes *et al.*, 2010*a*), which is currently an active research topic (Smeets *et al.*, 2010).

A more realistic, human subjective, identification process can be simulated by generating face-pool tests (Moyers, 2007). Given an image of the CFR and a set of possible candidate images extracted from the database, a human observer is asked to indicate the face from the face-pool most similar to the given CFR. An interesting future alternative, based on the work found in Hill *et al.* (2010), could be the straightforward translation of physical differences or computer-based measures of similarity into perceived human-based differences, omitting the need for cumbersome face-pool testing.

17.10 Discussion

Forensic facial reconstruction aims at estimating the facial appearance associated with an unknown skull specimen for human identification. All facial reconstruction techniques are based on the assumed relationship between the soft tissue envelope and the underlying skull substrate. Manual reconstruction methods consist of physically modelling a face on a skull replica and require a lot of anatomical and artistic expertise, resulting in subjective and time-consuming reconstructions with variable results between operators. The development of software for computerised facial reconstructions of an individual offers opportunities for faster, easier, more efficient and possibly more accurate generation of multiple representations of an individual.

This review concentrated on computerised CFR methods that require minimal manual input, where manual input is ideally restricted to the indication of a sparse set of craniofacial skeletal landmarks. However, because of this limited specific input, such (semi-) automated procedures were previously observed to generate generic approximations of the associated face (Wilkinson, 2005), compared with the more specific computer-assisted manual reconstruction procedures, which are based on careful examination of the skull and experienced anatomical knowledge (Wilkinson *et al.*, 2006).

However, over the last decade we have witnessed a clear shift in computerised CFR methods from using sparse information at a small number of landmarks to dense representations of the facial–skeletal complex. Consequently, much more information is now being used and modelled in these newer procedures, opening the possibility to recreate faces that better 'fit' the target skull. At the same time, craniofacial models can be constructed as statistical summaries of exemplars in a database. These statistical models not only code for individual differences but can also be trained to capture attribute-related trends or differences such as gender, age and ethnicity. This much richer description of statistical variation of the skeletal-facial complex therefore has large potential to substantially improve the computerised CFR results.

In order for such models to become accepted in practical forensic facial-reconstruction scenarios, some important steps still need to be taken. First and foremost, the accuracy, reliability and reproducibility of the computer-based systems need to be established. In the validation section of this review, some propositions for scientifically verifiable validation studies were listed. Ideally, a reference validation dataset, including validation scenarios should be collected and set up by the CFR community for independent and controlled assessment. Secondly, the power of these automated methods lies in the use of statistical information extracted from sufficiently large (male/female, spanning different age intervals and ethnic groups) databases of the complex interrelationship of the skull and facial surface. Such information can be extracted from conventional CT scans or, even better, as advocated in the section on data acquisition, CBCT scans, since these are taken in an upright position with less radiation dose for almost similar image quality compared to conventional CT. Again, once sufficient objective evidence is gathered from pilot studies that CFR procedures using (CB)CT-based information result in acceptable reconstructions, efforts need to be undertaken to standardise and disseminate protocols for acquisition of such data for different populations. Finally, for the 'automated' claim of these computerised methods to hold, further improvements are required in the image-processing chain to make such algorithms robust in the presence of the incomplete data available.

In conclusion, whilst CFR can still be undertaken manually, it is only a handful of sculptors who are capable of producing consistent likeness from remnant skull evidence. Many of the claims of success rates in forensic casework have been anecdotal, undocumented and difficult to verify. Situations where major losses of life had led to the need for many unknown skulls to have their faces reconstructed would overwhelm the capabilities of the handful of CFR experts who use manual methods and therefore computer-based automated methods need to be developed. Ever-cheaper and more powerful computing power has transformed medical imaging and the capability to utilise such data to construct large statistical databases from which the most plausible facial appearance from any unknown skull can be quickly inferred has become possible. These methods are still evolving but have reached the point where objective tests of performance are feasible. This is an essential development if the newer rapid, automated and objective methods are to be utilised to the full by the courts and those experts confronted with the aftermath of major disasters with many victims to be rapidly identified. This transformation will see CFR migrate from a niche art to mainstream forensic and archaeological practice in the hands of many scientists.

References

Andersson, B. and Valfridsson, M. (2005). *Digital 3D facial reconstruction based on computer tomography*. Master's Thesis, Linkopings University.

Archer, K. (1997). *Cranio-facial reconstruction using hierarchical B-spline interpolation*. Master's Thesis, The University of British Colombia.

Archer, K., Coughlan, K., Forsey, D. and Struben, S. (1998). Software tools for craniofacial growth and reconstruction. *Graphics Interface*. Vancouver, Canada.

Attardi, G., Berto, M., Forte, M. *et al*. (1999). 3D facial reconstruction and visualization of ancient Egyptian mummies using spiral CT data. *ACM SIGGRAPH, Sketches and Applications*. Los Angeles, USA.

Berar, M., Desvignes, M., Bailly, G. and Payan, Y. (2005*a*). 3D statistical facial reconstruction. *4th International Symposium on Image and Signal Processing and Analysis, ISPA*. Zagreb, Croatia.

Berar, M., Desvignes, M., Bailly, G. and Payan, Y. (2005*b*). Statistical skull models from 3D X-ray images. *2nd International Conference on Reconstruction of Soft Facial Parts, RSFP*. Remagen, Germany.

Berar, M., Desvignes, M., Bailly, G. and Payan, Y. (2006). 3D semi-landmark-based statistical face reconstruction. *Journal of Computing and Information Technology*, **14**, 31–43.

Bullock, D. (1999). *Computer assisted 3D craniofacial reconstruction*. Master's Thesis, The University of British Columbia.

Claes, P. (2007). *A robust statistical surface registration framework using implicit function representations: Application in craniofacial reconstruction*. PhD Thesis, K.U. Leuven, Belgium.

Claes, P., De Greef, S., Willems, G., Vandermeulen, D. and Suetens, P. (2004*a*). Craniofacial statistical modeling and reconstruction. *3D Modelling*. Paris, France.

Claes, P., Vandermeulen, D., De Greef, S., Willems, G. and Suetens, P. (2005*a*). Combined statistical modeling of tissue depth and 3D facial outlook for computerized facial approximation. *2nd International Conference on Reconstruction of Soft Facial Parts, RSFP*. Remagen, Germany.

Claes, P., Vandermeulen, D., De Greef, S., Willems, G. and Suetens, P. (2005*b*). Statistically deformable face models for cranio-facial reconstruction. *Proceedings 4th International Symposium on Image and Signal Processing and Analysis, ISPA*. Zagreb, Croatia.

Claes, P., Vandermeulen, D., De Greef, S., Willems, G. and Suetens, P. (2006*a*). Statistically deformable face models for cranio-facial reconstruction. *Journal of Computing and Information Technology*, **14**, 21–30.

Claes, P., Vandermeulen, D., De Greef, S., Willems, G. and Suetens, P. (2006*b*). Craniofacial reconstruction using a combined statistical model of face shape and soft tissue-depths: methodology and validation. *Forensic Science International*, **159**, 147–158.

Claes, P., Vandermeulen, D., De Greef, S. *et al*. (2010*a*). Bayesian estimation of optimal craniofacial reconstructions. *Forensic Science International*, **201**, 146–152.

Claes, P., Vandermeulen, D., De Greef, S. *et al*. (2010*b*). Computerized craniofacial reconstruction: conceptual framework and review. *Forensic Science International*, **201**, 138–145.

Claes, P., Vandermeulen, D., Suetens, P., De Greef, S. and Willems, G. (2004*b*). Computerized facial approximation using statistical models of tissue depth and 3-D facial outlook. *Conference of the International Association for Craniofacial Identification, IACI*. Dalian, China.

Davy, S., Gilbert, T., Schofield, D. and Evison, M. (2005). Forensic facial reconstruction using computer modeling software. In J. G. Clement and M. K. Marks (Eds.), *Computer-Graphic Facial Reconstruction*. Amsterdam: Elsevier.

De Greef, S., Claes, P., Mollemans, W. *et al*. (2005). Semi-automated ultrasound facial soft tissue depth registration: method and validation. *Journal of Forensic Science*, **50**, 7.

Evenhouse, R., Rasmussen, M. and Sadler, L. (1992). Computer-aided forensic facial reconstruction. *Journal of BioCommunication*, **19**, 22–28.

Evison, M. (1996). *Computerised 3D facial reconstruction* [Online]. Research School of Archaeology and Archaeological Science, University of Sheffield, UK. www.assemblage.group.shef.ac.uk/1/evison.html.

Evison, M. (2000). Modeling age, obesity, and ethnicity in a computerized 3D facial reconstruction. *Forensic Science Communications* [Online], **3**. www.fbi.gov/hq/lab/fsc/backissu/april2001/evison.htm.

Gerasimov, M. M. (1971). *The Face Finder*. New York, NY: Lippincott.

Haglund, W. and Reay, D. (1991). Use of facial approximation techniques in identification of green river serial murder victims. *American Journal of Forensic Medicine and Pathology*, **12**, 132–142.

Hill, H., Claes, P., Crocoran, M. *et al.* (2011). How different is different? Criterion and sensitivity in face-space. *Frontiers in Psychology, Perception Science*, **2**, 1–14.

Hill, H., Claes, P., Favelle, S. and Clement, J. G. (2010). Linking facial morphometrics to human perception. *Forensic Osteology and Anthropology*.

Jones, M. (2001). Facial reconstruction using volumetric data. *International Vision Modeling and Visualisation Conference*. Stuttgart.

Kähler, K., Haber, J. and Seidel, H. (2003). Reanimating the dead: reconstruction of expressive faces from skull data. *ACM Transactions on Graphics*, **22**, 554–561.

Lebedinskaya, G., Balueva, T. and Veselovskaya, E. (1993). Development of methodological principles for reconstruction of the face on the basis of skull material. In M. Y. İşcan and R. P. Helmer (Eds.), *Forensic Analysis of the Skull: Craniofacial Analysis, Reconstruction, and Identification*. New York, NY: Wiley-Liss.

Loubele, M., Bogaerts, R., Dijck, E. V. *et al.* (2009). Comparison between effective radiation dose of CBCT and MSCT scanners for dentomaxillofacial applications. *European Journal of Radiology*, **71**, 461–468.

Loubele, M., Maes, F., Jacobs, R. *et al.* (2008). Comparative study of image quality for MSCT and CBCT scanners for dentomaxillofacial radiology applications. *Radiation Protection Dosimetry*, **129**, 222–226.

Mang, A., Muller, J. and Buzug, T. (2006). A multi-modality computer-aided framework towards postmortem identification. *Journal of Computing and Information Technology*, **14**, 7–19.

Michael, S. and Chen, M. (1996). The 3D reconstruction of facial features using volume distortion. *14th Eurographics UK Conference*. London.

Moyers, D. (2007). *Validation study of ReFace* (*Reality Enhanced Facial Approximation by Computational Estimation*). Master's Thesis, University of Tennessee.

Muller, J., Mang, A. and Buzug, T. M. (2005). A template-deformation method for facial reproduction. *International Symposium on Image and Signal Processing and Analysis, ISPA*. Zagreb, Croatia.

Nelson, L. and Michael, S. (1998). The application of volume deformation to three-dimensional facial reconstruction: a comparison with previous techniques. *Forensic Science International*, **94**, 167–181.

Paysan, P., Luthi, M., Albrecht, T. *et al.* (2009). Face reconstruction from skull shapes and physical attributes. *Lecture Notes in Computer Science, Pattern Recognition*, **5748**, 232–241.

Pei, Y., Zha, H. and Yuan, Z. (2004). Tissue map based craniofacial reconstruction and facial deformation using RBF network. *Third International Conference on Image and Graphics*. Hong Kong, China.

Pei, Y., Zha, H. and Yuan, Z. (2008). The craniofacial reconstruction from the local structural diversity of skulls. *Computer Graphics Forum*, **27**, 1711–1718.

Prag, J. and Neave, R. (1997). *Making Faces using Forensic and Archeological Evidence*. London: British Museum Press.

Quatrehomme, G. (1997). A fully three-dimensional method for facial reconstruction based on deformable models. *Journal of Forensic Science*, **42**, 649–652.

Reichs, K. (1992). Forensic anthropology in the 1990s. *American Journal of Forensic Medicine and Pathology*, **13**, 146–153.

Shahrom, A., Vanezis, P., Chapman, R. *et al.* (1996). Techniques in facial identification: computer-aided facial reconstruction using laser scanner and video superimposition. *Journal of Legal Medicine*, **108**, 194–200.

Smeets, D., Claes, P., Clement, J. G. and Vandermeulen, D. (2010). Objective 3D face recognition: evolution, approaches and challenges. *Forensic Science International*, **201**, 125–132.

Snow, C., Gatliff, B. and Williams, K. M. (1970). Reconstruction of the facial features from skull: an evaluation of its usefulness in forensic anthropology. *American Journal of Physical Anthropology*, **33**, 221–228.

Stephan, C. (2006). Beyond the sphere of the English facial approximation literature: ramifications of German papers on Western method concepts. *Journal of Forensic Sciences*, **51**, 736–739.

Subsol, G. and Quatrehomme, G. (2005). Automatic 3D facial reconstruction by feature-based registration of a reference head. In J. G. Clement and M. K. Marks (Eds.), *Computer-Graphic Facial Reconstruction*, pp. 79–102. Amsterdam: Elsevier.

Tilotta, F., Glaunès, J., Richard, F. and Rozenholc, Y. (2010). A local technique based on vectorized surfaces for craniofacial reconstruction. *Forensic Science International*, **200**, 50–59.

Tilotta, F., Richard, F., Glaunès, J. *et al.* (2009). Construction and analysis of a head CT-scan database for craniofacial reconstruction. *Forensic Science International*, **191**(1–3), 112.e1–112.e12.

Tu, P., Book, R., Liu, X. *et al.* (2007). Automatic face recognition from skeletal remains. *IEEE Conference on Computer Vision and Pattern Recognition* (*CVPR*), **2007**, 1–7.

Tu, P., Hartley, R., Lorensen, W. *et al.* (2005). Face reconstruction using flesh deformation modes. In J. G. Clement and M. K. Marks (Eds.), *Computer-Graphic Facial Reconstruction*, pp. 145–162. Amsterdam: Elsevier.

Turner, W., Brown, R., Kelliher, T. *et al.* (2005). A novel method of automated skull registration for forensic facial approximation. *Forensic Science International*, **154**, 149–158.

Vandermeulen, D., Claes, P., Greef, S. D., Willems, G. and Suetens, P. (2005*a*). Volumetric deformable face models for cranio-facial reconstruction. *4th International Symposium on Image and Signal Processing and Analysis, ISPA*. Zagreb, Croatia.

Vandermeulen, D., Loubele, M., Claes, P. *et al.* (2005*b*). Low-dose CT-based soft tissue modeling for craniofacial reconstruction. *2nd International Conference on Reconstruction of Soft Facial Parts, RSFP*. Remagen, Germany.

Vandermeulen, D., Claes, P., Loeckx, D. *et al.* (2006). Computerized craniofacial reconstruction using CT-derived implicit surface representations. *Forensic Science International*, **159**, 164–174.

Vanezis, P. (1989). Application of 3-D computer graphics for facial reconstruction and comparison with sculpting techniques. *Forensic Science International*, **42**, 69–84.

Vanezis, P., Vanezis, M., Mccombe, G. and Niblet, T. (2000). Facial reconstruction using 3-D computer graphics. *Forensic Science International*, **108**, 81–95.

Wilkinson, C. (2004). *Forensic Facial Reconstruction*. Cambridge: Cambridge University Press.

Wilkinson, C. (2005). Computerized forensic facial reconstruction. *Forensic Science, Medicine, and Pathology*, **1**, 173–177.

Wilkinson, C. (2010). Facial reconstruction – anatomical art or artistic anatomy? *Journal of Anatomy*, **216**, 235–250.

Wilkinson, C., Rynn, C., Peters, H. and Taister, M. (2006). A blind accuracy assessment of computer-modeled forensic facial reconstruction using computed tomography data from live subjects. *Forensic Science, Medicine and Pathology*, **2**, 179–187.

Chapter 18

Computer-generated facial depiction

Gregory Mahoney and Caroline Wilkinson

18.1 Introduction

Numerous computerised craniofacial reconstruction (CFR) systems have developed since the early pioneering work by Moss *et al.* (1987). Many of these systems were designed to produce more time-efficient, less subjective and more reliable CFR, with the ultimate goal of complete automation. However, automated computer systems may impose a very specific set of facial characteristics and may be reliant upon facial templates (Arridge *et al.*, 1985; Evison *et al.*, 2003), average tissue depths (Vanezis *et al.*, 1989, 2000; Evenhouse *et al.*, 1992) and population-specific databases (Stratomeier *et al.*, 2005). Advances in haptic (tactile feedback interface) technology and 3D rendering software have also enabled the development of computerised systems which replicate the practitioner-led manual methods, rather than following the automated approach. These systems attempt to recreate characteristic facial morphology, and a more individualistic likeness than an approximation. Some computerised systems employ 3D animation software (Buhmann *et al.*, 2003; Eliasova *et al.*, 2003; Evison *et al.*, 2003; Kahler *et al.*, 2003; Kindermann, 2003) to model the face onto the skull, whilst other systems employ virtual sculpture software and hardware with haptic feedback (Wilkinson, 2003). The systems with haptic feedback have the advantage of allowing the practitioner to feel the surface of the skull during analysis, which provides some important skeletal detail for CFR, such as location of the malar tubercle or muscle attachment rugosity. These 3D-modelling systems have also been utilised to follow facial approximation methods (Subke and Wittke, 2005). Computerised CFR systems which require anthropological interpretation and computer modelling skills are, in general, more time consuming than the approximation systems.

There are many benefits to computer-based CFR systems. Where the skull is fragmented, some computerised systems allow skull reassembly (Subke and Wittke, 2005; Wilkinson, 2005), and this is more efficient and rapid than manual reassembly, as no physical scaffold is necessary and no damage to the specimen is possible. Computerised remodelling of missing fragments is significantly easier than manual methods, and may reduce the amount of necessary time from days to hours. Computerised systems also allow a more realistic facial appearance than the manual methods, creating familiar images consistent with photographs or film sequences (Friend, 2003; Lorenzi, 2004; BBC News, 2010).

Although some manual practitioners (D'Hollosy, 2005; Vermeulen, 2005) can produce realistic results by employing synthetic casting materials, prosthetic eyeballs, paint and real hair implantation, these are very time-consuming and are rarely appropriate in forensic scenarios. The 3D virtual-sculpture systems also allow the practitioner more flexibility, more precision and repeatability and the valuable option of variable transparency, while adhering to any physical method of forensic CFR.

A major advantage of utilising any computerised method is that there is far less risk of damage to the skull, since no matter how the 3D virtual skull model is produced (the best methods to date being an optical laser scan or CT scan) once scanned, the skull no longer needs to be handled. Of course, for fragile archaeological specimens this is extremely advantageous. The 3D skull scan data can be reproduced as a solid replica, using a rapid prototype printer (Hjalgrim *et al.*, 1995; Spoor *et al.*, 2000), to act as the base for a

Craniofacial Identification, ed. C. Wilkinson and C. Rynn. Published by Cambridge University Press.

clay sculpture; however, using the virtual skull model within the computer system provides distinct advantages and possibilities. Animation, modelling and realistic texturing, 3D displays, and their impact on recognition are a few of the areas requiring further research.

18.2 Virtual reality CFR

Virtual reality CFR has been developed utilising a hardware/software package produced by SensAble Technologies (Wilkinson, 2003, Wilkinson *et al.*, 2006). This system uses a haptic tactile feedback device, in this case the PHANTOM® desktop, in conjunction with FreeForm Modeling Plus software (SensAble Technologies, 2010*a*, 2010*b*). This method can be adapted using other systems made by SensAble, or practically any other system utilising a haptic device and 3D sculpting software. This system can be run on anything upwards of a Dual 800 MHz Pentium Processor with 1GB RAM, Windows 2000, 1024 × 768 resolution display and NVIDIA Quadro4 900XGL graphics card. The PHANTOM®/FreeForm Modeling Plus system was developed primarily for the industrial design and medical fields, and thus has a number of tools within the package that are surplus to requirements regarding CFR. Though awkward at first, with a little practice a sculptor will soon feel quite at home in the virtual environment and excited by the new possibilities, formerly unavailable to them in the physical world. The following will give a brief description of the tools and functions within the software which are ideal for CFR purposes.

FreeForm Modeling Plus essentially allows an operator to sculpt virtual 'clay' objects in 3D space (SensAble Technologies, 2010*b*). The software is structured in a way that is similar to many photo-editing software programs, in which there are multiple layers which can be grouped, and toggled on or off for viewing. Perspective can be set and turned on or off as needed for viewing. Many of the tools within the workspace would be familiar to a traditional sculptor: there are tools for carving, shaping and smoothing the clay that look and feel like tools used in the physical world. Other tools are perhaps not so familiar, but extremely valuable. The 'tug' tool allows an operator to push or pull areas of clay in any direction, and is used more than any other tool. New pieces of clay can be added in specified shapes, custom shapes, or imported from libraries. Flat planes can be aligned in any orientation, and can be sketched on using controlled lines and shapes or in a more freehand manner, plus scaled photographs can be displayed with variable transparency. These planes can also be used to slice or mirror a piece of 'clay'. The resolution of the 'clay' can be changed, and the 'clay' can be made harder or softer depending on the desired effect. Tools exist for sketching directly on the surface of an object as well as tools for measuring. There are many other tools within the software, which may be employed, but those described above are the most useful regarding CFR.

Before working in 3D space within FreeForm virtual sculpture software, CT data must be converted into a 'Certified Trust License' (.stl) file type or 'Object' (.obj) file type to be imported as a 3D model. Once imported, the skull should be aligned to the Frankfurt Horizontal Plane. Aside from placing the skull in proper orientation for accepted CFR techniques, it allows an operator to quickly switch between points of view along X, Y and Z axes.

It is also advisable to fill a working copy of the virtual skull so that no computing power is required to render interior surfaces of the skull, which are not required for CFR. Nevertheless, it is also advisable to keep an unedited skull copy in a separate layer, which can be referenced if necessary. This increases the file size, but if the layer is turned off (i.e. present but invisible) it is not taxing video memory. This can be automated to an extent, but there is an option to manually fill the interior of the skull, allowing for more control. A separate piece is created in the virtual space and is blended to fit with the exterior surface. The two pieces are then combined, which effectively fills the voids as required.

This chapter describes two virtual sculpture techniques of CFR; one that follows the traditional American method developed at the Boston Police Department (Mires *et al.*, 2003; Mahoney and Rynn, 2011) and the other that follows the traditional Manchester method developed at the University of Dundee (Wilkinson, 2004). Methods will be described for each stage of the reconstruction process, and the choice of method will depend upon the skills, preference and knowledge base of the practitioner.

18.3 Stage 1: Tissue depth peg addition

The first stage of the reconstruction process is to attach tissue depth pegs to the surface of the skull at a number of anatomical points. The number of points varies depending on the dataset employed and the dataset is chosen as the most appropriate for the sex, age and ancestry of the unknown skull. There are many different sets of data from many different geographical locations including: Australian (Domaracki and Stephan, 2006), Belgian (De Greef *et al.*, 2009), Brazilian (Tedeschi-Oliveira *et al.*, 2009), British (Wilkinson, 2002), Canadian (Vander Pluym *et al.*, 2007), Chilean (Galdames *et al.*, 2008), Chinese (Birkner, 1904), Egyptian (El-Mehallawi and Soliman, 2001), French (Tilotta *et al.*, 2009), German (His, 1895; Welcker, 1896; Helmer, 1984), Indian (Sahni *et al.*, 2002), Japanese (Suzuki, 1948; Utsuno *et al.*, 2005, 2007), Mixed Race South African (Phillips and Smuts, 1996), Namibian (Von Eggeling, 1909), Native North American (Rhine, 1983), Papuan (Fischer, 1905), Portuguese (Codinha, 2009), Swiss (Kollmann and Buchly, 1898), US White, Black and Hispanic (Rhine and Campbell, 1980; George, 1987; Manhein *et al.*, 2000), former USSR ethnic groups (Lebedinskaya and Veselovskaya, 1986) and Zulu (Aulsebrook *et al.*, 1996).

Pegs may be attached to the surface of the skull using a number of different methods. Here we outline two methods; the *planar method* (Mires *et al.*, 2003; Mahoney and Rynn, 2011) and *the imported peg method* (Wilkinson *et al.*, 2006).

18.3.1 Planar method

Frontal, posterior, lateral, superior and inferior planes are attached to the piece, creating an exterior box. These planes can be 'sketched' on with drawing tools, or have previously saved 'sketches' imported on a selected plane. The model is displayed in orthogonal projection (i.e. perspective turned off so the box sides are at right angles, in orthogonal elevation) when sketching on planes. Thus, libraries of tissue depth reference points can be created, saved and re-used. These pre-sketched tissue-depth landmarks can be moved on the planes to align them to their corresponding landmarks on the skull behind the plane. Once properly aligned to the individual skull, the sketched landmarks can be projected onto the surface of the skull piece (Figure 18.1a). Each landmark can then be selected individually, and each tissue-depth marker extruded from the surface to the appropriate depth (Figure 18.1b) and precisely perpendicular to the surface of the skull.

18.3.2 Imported peg method

A peg is first created, using the standard cylinder tool when a new file is opened, with a 10 mm length and 5 mm diameter. This peg can then be saved and imported to each anatomical point, and the length altered using the scale tool. If the peg is aligned with

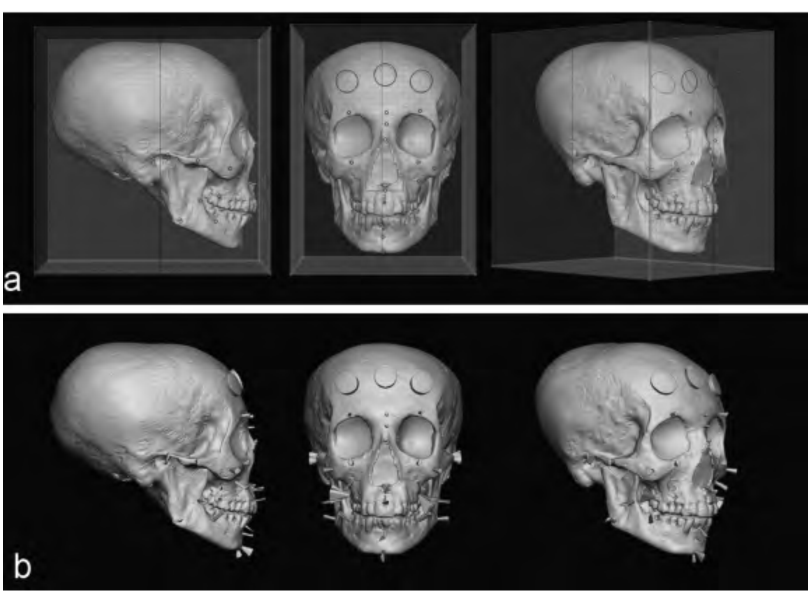


Figure 18.1 The *Planar* method of tissue depth peg addition. a: Reference markers projected onto the skull surface at the tissue depth points; b: Extrusion of the reference markers to the tissue depth at each point.

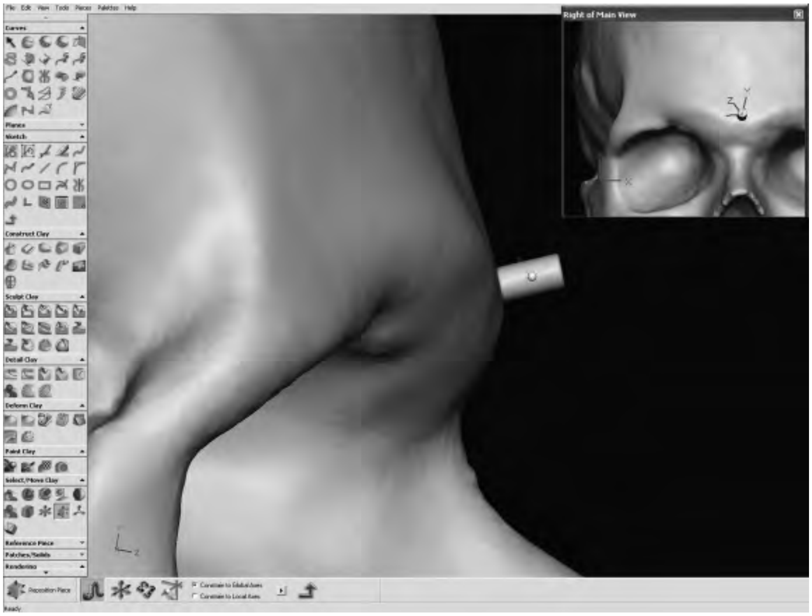


Figure 18.2 The *Imported Peg* method of tissue depth peg addition.

its length along the Y axis then the length can be altered easily as a percentage of the imported Y scale. Once the correct length, the peg can then be manually positioned at the anatomical landmark using the ‘reposition piece’ tool and fine alterations carried out at the skull surface so that the peg is aligned at right angles to the surface of the skull at the point (Figure 18.2). Bilateral pegs can be copied and pasted to the opposite side.

18.4 Stage 2: Anatomical sculpture

The American method of CFR employs a number of key anatomical structures to determine the overall face shape (Gatliff, 1984), whereas the Manchester method recommends the sculpture of the majority of facial muscles to determine detailed facial structure (Wilkinson, 2004). Therefore the American digital practitioner models anatomical structures directly onto each skull (Mires *et al.*, 2003; Mahoney and Rynn, 2011), whereas the Manchester digital method employs a database of pre-modelled muscles that are imported and deformed to fit each new skull (Wilkinson *et al.*, 2006).

18.4.1 American method

Using a copy of the filled skull, an offset piece can be created at a specific distance from the surface of the skull. Since tissue depth is most consistent and relatively shallow on the calvaria, the new piece is offset to the value associated with tissue-depth landmarks around the frontal eminence or vertex. After this piece is created, the surface of the offset piece reflects the shape of the skull exactly, but requires smoothing to more closely resemble skin, while maintaining the set distance from the surface of the skull. This approach also means that any subtleties and asymmetries of the skull are echoed in the soft tissues (Figure 18.3a).

The ‘tug’ tool within the FreeForm software essentially consists of a sphere of influence of adjustable size and strength, with the highest degree of influence at the centre of the sphere. This tool can be positioned over a tissue-depth marker and resized to a point where the radius of the sphere is about two-thirds the distance to an adjoining marker. Once positioned properly, the practitioner pulls or pushes the offset ‘skin’ piece until the furthest extent of the tissue depth marker is just visible (Figure 18.3b). This is only done with tissue depth markers along the bony eminences of the skull, as the soft tissue around these areas most closely corresponds to the surface of the skull. Areas where there are deep tissue, fat, and more robust muscles will require a different treatment.

The temporalis muscle and its overlying tissue are sculpted based upon the attachment marks evident on

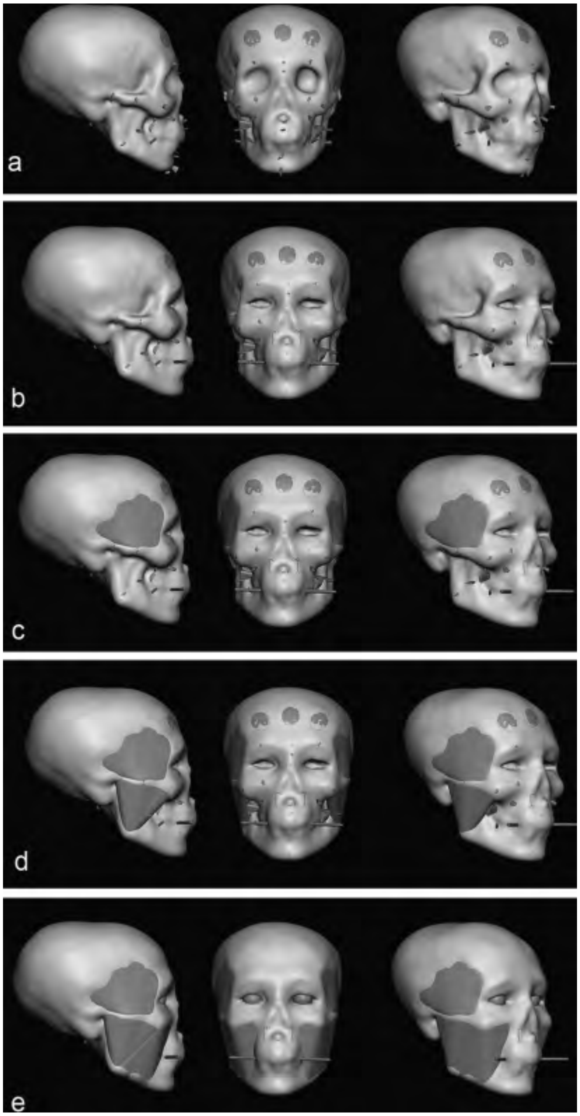


Figure 18.3 The American method of anatomical sculpture. a: An offset skin layer is created from the surface of the skull; b: Skin layer is 'tugged' to the level of the tissue depth markers over the cranium and orbits; c: Addition of the temporalis muscle and its overlying tissue; d: Addition of the 'back triangle'; e: Addition of the 'front triangle'.

the skull. One end of a column of clay is anchored to a point along the offset tissue over the zygomatic arch and zygomatic process of the temporal bone: the other end of the column is anchored to the point where the muscle would attach, along the temporal line. This line should still be evident on the surface of the skin created by the offset piece of the skull. If it is not, the operator can either turn the piece invisible or make it transparent, so the underlying skull can always be referenced. The level of transparency can be set, and the piece can still be manipulated as in the opaque view. This is a distinct advantage over applying clay on a skull, where the practitioner loses the ability to directly refer to the skull beneath the clay. When the fan shape of this muscle and tissue is complete, the columns, which are created as one piece, can be smoothed and blended into the surrounding tissue (Figure 18.3c).

In accordance with the American method of CFR, the 'back triangle' is created. This region corresponds with the plane of the lateral cheek and jaw area, superficial to the masseter muscle and parotid gland. The approach is similar to that used for the temporal region: cylinders of clay are attached along the mandibular body, from the gonial angle to the zygomatic arch between the external auditory meatus and the temporozygomatic suture. Again, this cluster of cylinders is subsequently smoothed and blended into the surrounding tissue (Figure 18.3d).

The 'front triangle' is created in a similar fashion to the 'back triangle'. This region corresponds to the soft tissue of the front plane of the face, around the nose and cheeks as well as the anterior half of the body of the mandible. Deep to the front triangle lie buccinator, zygomatic major and minor, and the levator muscles of the mouth, along with the buccal and malar fat pads. Columns are attached along the infra-orbital margin, and to the mandible between the gonial angle and approximately halfway along the mandibular body (Figure 18.3e). By this point, the face is beginning to take shape.

18.4.2 Manchester method

A database of pre-modelled muscles is employed in order to recreate the anatomical structures of the face. These muscles were previously modelled from scratch onto a template skull using the sculpting, smoothing and carving tools. The muscle models show visible surface muscle fibres and have anatomically reliable origins, insertions and morphology.

When reconstructing an unknown skull each muscle is imported and deformed to fit in relation to the proportions and size (Figure 18.4a). The deform box tool can be used to retain the basic muscle shape whilst altering the proportions and/or curvature of the surface of larger muscles, such as the masseter muscle (Figure 18.4a). The tug tool can also be employed to deform parts of the muscle in isolation (Figure 18.4b).

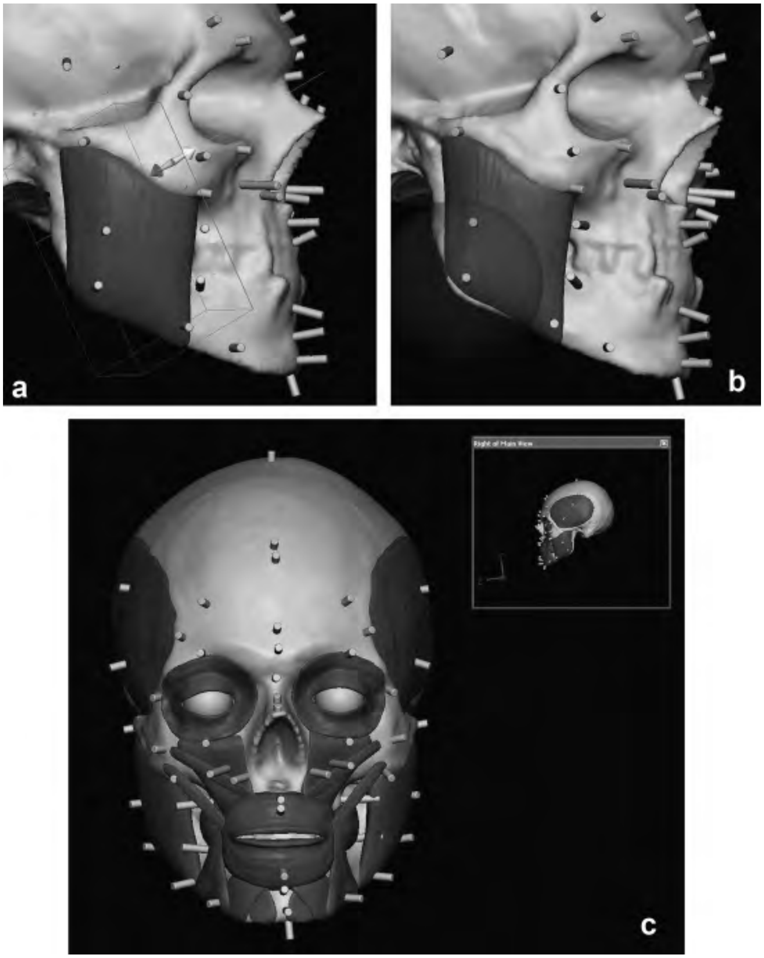


Figure 18.4 The Manchester method of anatomical sculpture. a: Muscle imported from databank to correct position and deformed proportionally using the *deform box* tool. b: Muscle deformed to fit the skull using the *deform tug* tool. c: Skull with addition of all necessary facial muscle structures.

In this way the imported muscle can be easily and quickly manipulated to fit each individual skull. Where the muscle is bilateral, the mirror tool can be utilised to recreate the second muscle, which can then be individually repositioned and manipulated to fit. All the craniofacial muscles relevant to the structure of the face are imported and deformed to build up the facial proportions, overall face shape and profile. Other anatomical structures relevant to facial appearance, such as glands (parotid gland), bones (hyoid) and vessels (superior temporal veins) are also imported and deformed to fit each skull. By this point, the face is beginning to take shape (Figure 18.4c).

18.5 Stage 3: Feature prediction

Both methods utilise libraries including certain anatomical features, such as the neck and ears. These may be pre-modelled or collected from laser scan or clinical image data. The library saves a great deal of time when compared to sculpture of these features from scratch, and where the features can be represented as templates only small modifications are necessary for each skull.

A male or female neck can be imported, scaled to correspond to the size of the skull and aligned using the mastoid processes, nuchal line, trachea position and sternoclavicular joint to determine size and relative proportions (Figure 18.5).

The library ear can also be imported, scaled, modified and positioned relative to the external auditory meatus and mastoid process. A thin column may be created from the canal to serve as a guide. A pivot point for the ear piece can be positioned at the external auditory meatus, and the ear rotated about this pivot to achieve alignment (Figure 18.6). Other features cannot be created using the library and require skeletal

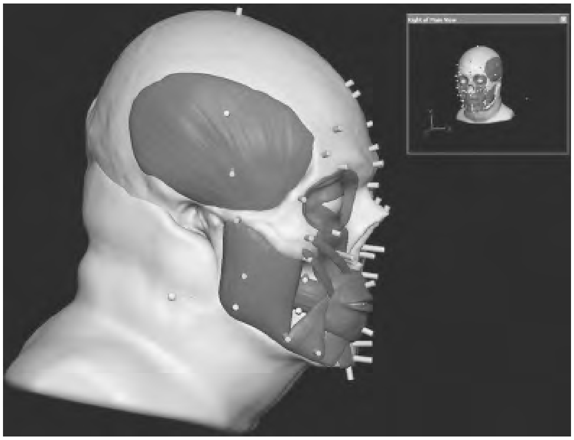


Figure 18.5 Addition of the neck. A neck model is imported from the databank and scaled, modified and positioned in relation to the skull

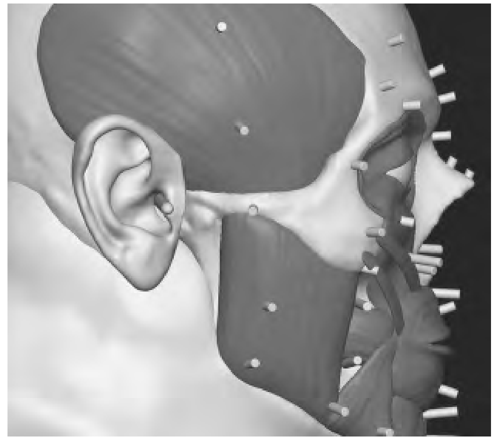


Figure 18.6 Addition of the ears. An ear model is imported from the databank and scaled, modified and positioned in relation to the skull.

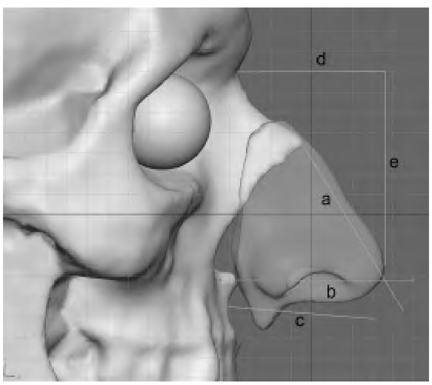


Figure 18.7 Addition of the nose using reference planes. Lines a and b relate to the two-tangent method (Gerasimov, 1955) for predicting nasal projection. Lines c, d and e relate to Rynn's soft tissue prediction measurements (Rynn and Wilkinson, 2009). a: Tangent from the last portion of the nasal bones; b: Line following the direction of the anterior nasal spine; c: Pronasale projection from subspinale in Frankfurt Horizontal Plane; d: Pronasale projection perpendicular to the nasion-prosthion plane; e: Pronasale height on a line parallel to the nasion-prosthion plane.

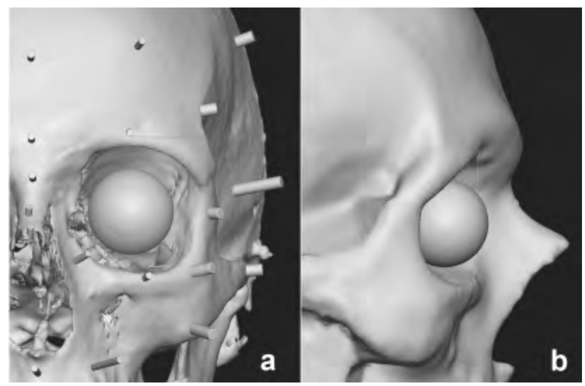


Figure 18.8 Eyeball addition using reference planes. a: Plane touching mid-supraorbital and mid infraorbital margins; b: Eyeball protruding past the plane by 3.8 mm (Wilkinson *et al*., 2003).

interpretation, utilisation of reconstruction standards and sculpture directly onto the skull.

The nose can be constructed following width and projection standards suggested by Gerasimov (1955) and nasal morphology formulae suggested by Rynn and Wilkinson (2009). Prior to constructing the nose itself, levator labii superioris is created to help with positioning of the nose. A midline sagittal plane is created which bisects the width of the nasal aperture. Reference lines can be drawn on this plane according to the standards. Using these reference lines, and an understanding of the anatomical structures, a nose is sculpted and positioned (Figure 18.7).

The eyeballs are positioned in the orbits following standards suggested by Stephan (2002) and Wilkinson *et al.* (2003), so that a tangent from the mid-supraorbital margin to the mid infraorbital margin will touch the flat plane of the iris. Globes of 25 mm diameter are created as new pieces to represent the eyeballs. Reference lines can be drawn across the orbit, from the midpoint of the supraorbital margin to the midpoint of the infraorbital margin and from the mid-medial to mid-lateral margin, creating a reference for positioning the eyeball in the orbit. These reference lines are also used to set the eyeball to the correct depth in the orbit (Figure 18.8).

The eyelids may be constructed using a variety of methods. The *extrusion method* (Mires *et al*., 2003; Mahoney and Rynn, 2011) creates an offset piece from the eyeball, the thickness of the eyelid. The shape of the eye fissure can then be sketched on this

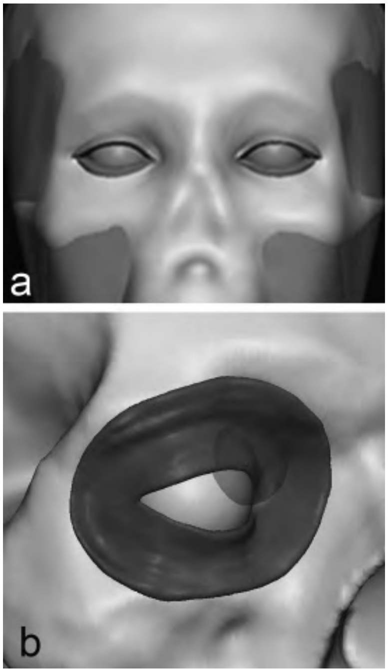


Figure 18.9 Addition of the eyelids. a: Extrusion method creating an offset piece from the eyeball; b: Addition method importing the orbicularis oculi muscle from the databank and deforming it to fit the orbital pattern.

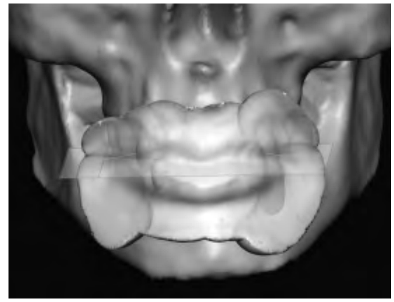


Figure 18.10 Addition of the mouth. A plane is position halfway up the maxillary incisors and lines added radiating from the lateral borders of the canine teeth (Gerasimov, 1955) to indicate the corners of the mouth. A lip model imported from the databank is deformed to fit the dental pattern and proportions of the skull.

offset piece, aligning the medial and lateral canthi to their attachment points (anterior lacrimal crest and malar tubercle respectively). The eye area can be selected and extruded back into the centre of the piece, at a depth equal to the radius of the piece, effectively creating upper and lower eyelids. Either the upper or lower lid can be manually selected, then cut and pasted in position as a separate piece. The portion of these newly created eyelid pieces that correspond to the palpebral ligaments can now be tugged to their attachment points. Be careful that the sphere of influence of the tug tool corresponds to the most medial and lateral points on the lid so the majority of the lid adheres to the shape of the eyeball. Be sure that the upper lid overhangs the lower lid at the lateral aspect. The advantage of approaching eyelid construction in this way is that the lid will automatically hug the eyeball, as in reality (Figure 18.9a).

The alternative *addition method* (Wilkinson *et al.*, 2006) utilises a pre-modelled orbicularis oculi muscle created to fit an eyeball with a 25 mm diameter. This is imported and positioned so that the palpebral part touches the eyeball at the upper and lower lid (the imported orbicularis oris muscle was originally modelled over an eyeball of the same diameter, so that it will always hug the eyeball as in life). Since skulls exhibit different sizes, proportions and structure the orbital part of orbicularis oculi will not necessarily be in the correct position and this can be repositioned using the 'deform' tools to allow the orbital part to touch the bones around the orbital margin whilst leaving the palpebral part in proper alignment touching the eyeball. To complete the eyes, the soft tissue between the supra- and infraorbital margins and the corresponding upper and lower lids are tugged towards the lids themselves. Be careful to pay attention to the eyelid patterns, which are dictated by the morphology of the eye socket, as well as taking into account the subcutaneous fat (Figure 18.9b).

The mouth is constructed following standards suggested by Gerasimov (1955), Subtelny (1959), Angel (1978), Koch *et al.* (1979), Holdaway (1983), Talass *et al.* (1987), Wilkinson *et al.* (2003), Stephan and Henneberg (2003) and Stephan and Murphy (2008). Reference lines can be drawn onto a plane and projected onto the tissues for placement of the corners of the mouth and thickness of the upper and lower lips. Two thin columns can also be created, radiating from the junctions between the maxillary canine and first premolar teeth to position the corners of the mouth. Using these references the oral fissure is then carved, paying attention to the dental occlusion and morphology. It is also possible to import lips from a library of mouths, whether sculpted or laser scanned, that can be chosen for similarity and distorted to fit each new skull (Figure 18.10).

18.6 Stage 4: Skin layer

The American method has already constructed a skin layer that can be combined with all other features and smoothed at the seams. The Manchester method requires the addition of a skin layer over the muscle structure using the tissue depth pegs as guides. This can be created by two different methods.

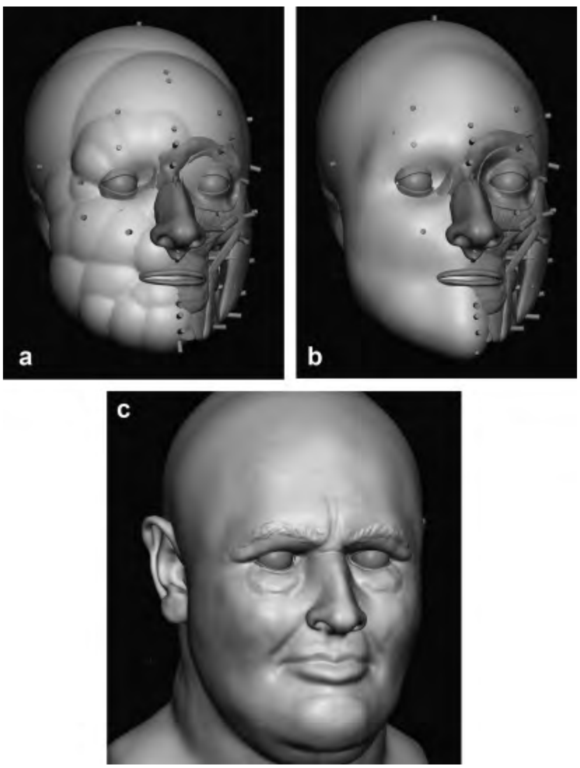


Figure 18.11 Addition of the skin layer using the *addition method*. a: Balls of clay are added over the muscle structure using the tissue depth pegs as guides; b: The skin layer is smoothed and pulled to the level of the tissue depth pegs; c: Addition of fine surface detail.

The *addition method* involves the creation of a skin layer as low-resolution balls of clay. The balls of clay vary in size depending on the area with very large spheres employed at the cranium to encompass the calvaria, moderate spheres added at the cheeks and jaw to follow the shape of the zygomatic bones and mandible, and small balls around the mouth and nose where the contours are more complicated. These balls can then be repeatedly smoothed and modified until a skin layer has been constructed with contours that follow both the structure of the muscles and the tissue depth guides. The resolution of this layer can then be increased as more fine detail is sculpted and as it is combined with other feature layers (Figure 18.11).

The *offset method* involves the creation of a piece at a specific distance from the surface of a copy of the combined muscle and skull layers. The new piece is offset to approximately 4 mm (close to the tissue depth at the vertex). The surface of the offset piece reflects the shape of the skull and the muscle structure, but requires smoothing to resemble skin. This approach also ensures that any subtleties and asymmetries of the skull and muscle structure are echoed in the soft tissues. The 'tug' tools are then employed to further shape the skin surface at areas that require greater tissue depths and finer detail is added in a similar way to the first method (Figure 18.12).

Any age-related changes such as sagging skin, deep wrinkles, or other fine-tuning can now be modelled. At this point, detail can be added by increasing the resolution and hardness of the combined piece. Fine lines, wrinkles and the oral fissure can all be dealt with in more detail.

18.7 Addition of textures

Models created within FreeForm are monochromatic, though the operator can set the colour of the working piece. FreeForm Modeling also allows the operator to

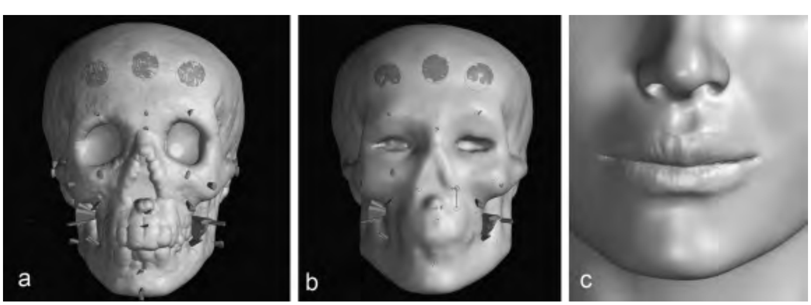


Figure 18.12 Addition of the skin layer using the *offset method*. a: An offset piece is created from the muscle structure; b: the offset piece is smoothed and pulled to the level of the tissue depth pegs; c: Addition of fine surface detail.

add colour using traditional paint tools from alternative software paint packages. There is even a FreeForm capability that allows model rendering using a Mental Ray® Rendering Engine to create different textures and surfaces. However, these FreeForm Modeling tools are not robust when compared with other available third-party applications. This aspect is better dealt with by exporting the model into another 3D modelling/animation program for texturing and rendering. This would depend on budget and time constraints, but allows for the creation of potentially photo-realistic results.

Some authors (Macris, 1987) suggest that a lower level of realism may be beneficial to recognition in a forensic scenario, although it is possible that increased exposure to photo-realistic computer depictions will affect our ability to recognise such faces as our relationship with technology evolves. A similar response has been seen in the film and animation field, where rapid technical obsolescence has led to the adjustment of public expectation and experience. It may be just a matter of time before the capabilities of image production render current representations inadequate.

We can take advantage of the current capabilities of computer software to create more engaging representations of the face, not only with regard to texture, but with facial expression. It is therefore possible to create faces as moving, speaking or interactive models.

Models created within FreeForm can be quite dense, often consisting of upwards of a 3-million triangle mesh. Exporting the model may be difficult where there is such high resolution and often 3D modelling/animation software packages require a mesh of quadrilateral polygons rather than triangular polygons. Where the mesh requires conversion it is important not to lose any model detail that might affect recognition and identification. The challenge is to lighten the FreeForm model resolution while optimising detail loss during the conversion process. FreeForm Modeling allows selective decimation of a model to considerably lower the number of triangles, while the morphology of the model remains true. However, a model can be decimated too far and become unrecognisable.

Finally, where facial animation is necessary the topology will require redefinition so that edge flow corresponds to the natural flow of the face. The triangle mesh within FreeForm is not created with this purpose in mind, so the model will need to be processed before importing into modelling/animation software. Once a new mesh has been created with quadrilateral polygons, then the operator will have a more appropriate model for texturing and rendering.

While there are third-party software solutions to address the problem of model detail compromise, FreeForm Modeling Plus also has the tools to address this issue. The operator can sketch curves directly onto the model within FreeForm Modeling Plus and can also set the node frequency or sensitivity of the lines being sketched onto the surface. The operator can also dictate that the curves being sketched snap to the surface of the underlying model. With this approach, a directed mesh can be created whilst respecting the fidelity of the original model. Once completed each quadrilateral section now consists of NURBS (Non-Uniform Rational B-Spline) curves on each side, which can then be selected to form a NURBS patch. Each NURBS patch reflects a vector-based representation of the corresponding surface of the original model. The NURBS patch can also be subdivided into segments dictated by the operator, allowing total control for the number of polygons desired. The only real stipulation here is that all polygons must have the same number of subdivisions to allow for the patches to be joined together seamlessly. After all the patches have been created they can be joined together to create a copy of the model with a topography recreated entirely of NURBS surfaces.

With this approach one can create an optimal mesh with appropriate edge flow for animation and modelling, while remaining as true to the original model as possible. While the morphology of the original and the NURBS model appear identical, the accuracy of this technique has yet to be tested, but it is anticipated that any differences between the two models will be negligible. The model can then be exported into a number of 3D file formats appropriate for third-party software for texturing and animation.

18.8 Skull reassembly and remodelling

There have been a number of studies to evaluate skull reassembly using computerised systems. One study (Mackenzie, 2007) utilised the FreeForm Modeling Plus software to reassemble five fragmented skulls. The skull fragments were scanned using a handheld Polhemus Fastscan Scorpion laser scanner and the reassembled skulls were compared with the intact skulls using surface-to-surface anthropometry (Geomagic Qualify) software. The intact skulls were recorded prior to fragmentation as laser scan models. Mackenzie found that all the mandibles, and three out of five reassembled crania showed less than 2.0 mm of error at more than 60% of the skull surface. The reassembled mandibles showed higher accuracy than the crania, with one mandible showing 70–79% and two showing 80–89% with < 2.0 mm error, whereas two crania showed only 50–59%, two 60–69% and one 70–79% with < 2.0 mm error. In this study the scanning technology appeared to be the limiting factor rather than practitioner error.

There have been a couple of studies evaluating the estimation of missing skull parts using skull remodelling (Colledge, 1996; Ismail, 2008). One study (Ismail, 2008) utilised the FreeForm Modeling Plus software to remodel five caucasoid-type skull models with missing parts. The original skulls were scanned using a hand held Polhemus Fastscan Scorpion laser scanner and a different bone was removed from each skull model (zygomatic, frontal, maxilla, nasal and mandible) and remodelled in a blind study. The remodelled skulls were compared with the original skulls using surface-to-surface anthropometry (Geomagic Qualify) software. Ismail found that the two most inaccurately remodelled bones were the mandible (average 0.36 mm error) and the zygomatics (average ±0.05 mm error). However, when the inaccurate skull models were employed for CFR and compared to ante-mortem images, there was only a slight decrease in the assessed likeness of the faces (using panel resemblance rating and facial superimposition assessments), suggesting that any errors in skull remodelling would not have a detrimental effect on CFR.

18.9 Three-dimensional models from two-dimensional data

Where access to the original specimen is not possible either directly or via 3D clinical imaging, methods of 3D model production from 2D data (such as radiographs, photographs and craniometrics) may be utilised.

Radiographs and/or photographs are used as templates and multiple views are aligned using cranial points as registration marks. Computer modelling software is then employed to create a 3D model (Wilkinson, 2005) or distort a template mesh (Davy *et al.*, 2005) to reproduce the skull morphology. Extrapolation of surface morphology between the views is inevitable and the more views that are available the more accurate the resulting 3D model of the skull. Since these methods include a certain degree of estimation and loss of detail on the surface of the bone, any resulting CFR/approximation would require photographic records of the skull and/or an appreciation of the decrease in accuracy of any resulting face. Care must be taken when establishing anatomical landmarks.

In 2004 analysis of the face of St Nicolas was carried out (Sunday Times, 2004). St Nicolas is the canonical and most popular name for Nikolaos of Myra (270 – 6 December 346), a saint and Greek Bishop of Myra (Demre, in Lycia, part of modern-day Turkey). In 1087, his relics were furtively transferred to Bari, in southeastern Italy, where they are housed today in a tomb at the cathedral. Access to the skeleton of St Nicolas is not possible as the remains are sealed in a tomb. However, a thorough anthropological assessment was carried out in 1953 by an Italian anthropologist, Martino, and he produced a large number of craniometric measurements, scaled drawings, radiographs and high quality black and white photographs of the skull (Martino, 1987). Utilising the photographs and radiographs as templates a 3D model of the skull of St Nicolas was produced using 3D modelling software and the accuracy of the proportions and scale of the skull model was checked using the numerous craniometrics measurements (Figure 18.13a). A 3D computerised facial reconstruction was then produced following the usual method (Wilkinson *et al.*, 2006) and realistic skin, eye and hair textures were applied to the head model to create a realistic representation of the face of St Nicolas (Figure 18.13b) (BBC2, 2004).

18.10 The accuracy of virtual computer CFR

It is important to assess the accuracy and reliability of any computer system for CFR, especially where this system may be employed in forensic investigation. There have been a couple of published studies (Wilkinson *et al.*, 2006; Lee *et al.*, in preparation) utilising the FreeForm Modeling Plus and PHANTOM® system. The first study used two skulls from CT data of living individuals from the USA (one male and one female), two CFR practitioners

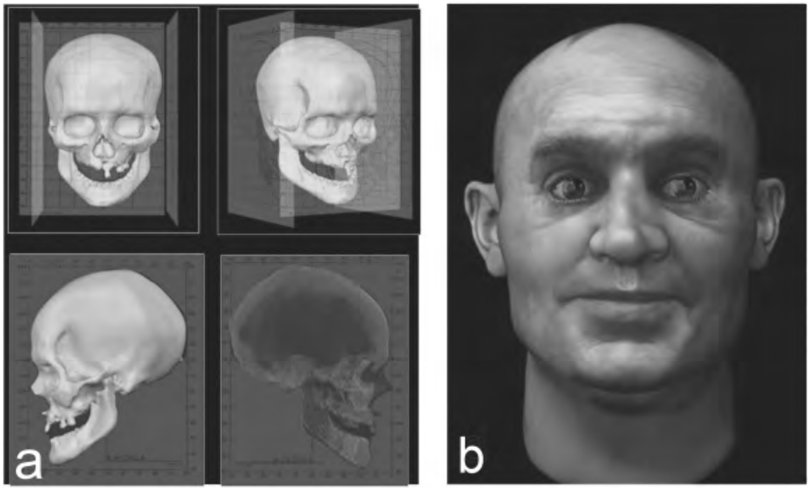


Figure 18.13 Facial analysis of St Nicolas. a: 3D model of the skull was produced utilising photographs, craniometrics, scaled drawings and radiographs as templates; b: Facial reconstruction of St Nicolas from the skull model. Texture courtesy of Image Foundry Studios.

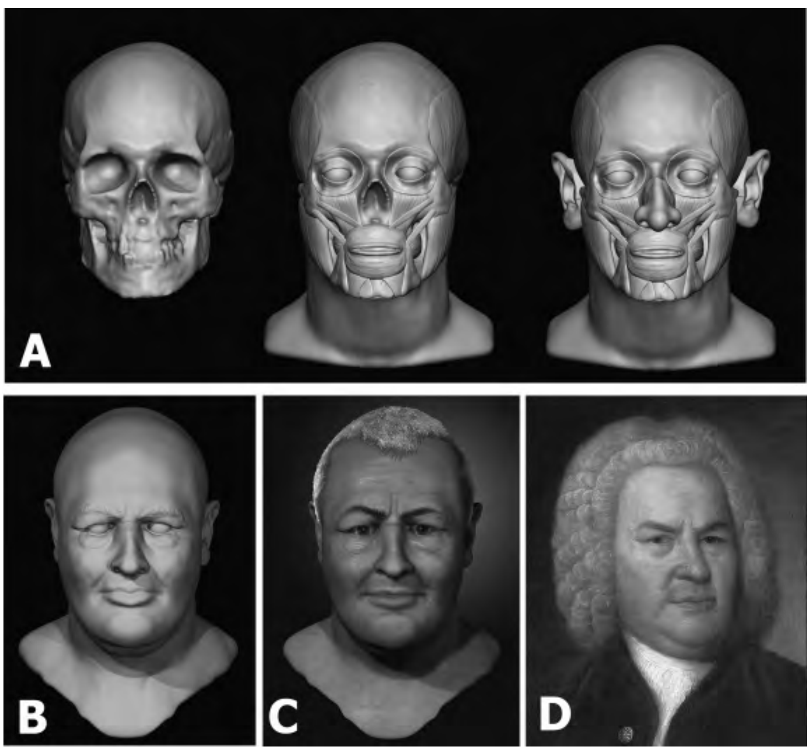


Figure 18.14 Facial depiction of J. S. Bach. The computerised facial reconstruction (A) utilised a 3D model of the skull from laser scan data. Texture was added to the resulting face (B) using the Haussmann 1746 portrait (D) as reference material regarding the degree of fatness, eye condition, skin colour, eye colour, hair colour and skin textures (C). Courtesy of the Bachhaus, Eisenach and Dr Caroline Wilkinson and Janice Aitken, University of Dundee.

and face-pool and surface-to-surface anthropometry assessment. Both reconstructions in this study received majority hit rates (>70%) in the face-pool assessment and showed less than 2.5 mm of error at more than 60% of the facial surface, when compared with the 3D model of the living face from CT data. Errors between the facial surfaces were, in part, a reflection of the difference in position between the reconstructed upright face and the prone subject in the CT scanner.

Lee *et al.* (2011) studied three skulls from CT data of living individuals from Korea (one female and two male) using one CFR practitioner and surface-to-surface anthropometry assessment. Reconstructions in this study showed less than 2 mm of error at 54% and less than 2.5 mm of error at 65% of the facial surface, when compared with the 3D model of the living face from CT data. This study employed cone CT scan data where the subject was in an upright position making the face-to-face comparison more relevant.

18.11 Examples of 3D computer CFR

Recent analysis of the face of Johann Sebastian Bach (1685–1750), the famous German composer and organist, utilised 3D CFR from the skull and texture addition from portrait detail (Hansen, 2008). The 3D computerised CFR was produced using a laser scan of the bronze copy of the skull of Bach provided by the Bachhaus, Eisenach (Figure 18.14). The skull suggested a middle-aged man with a large nose, strong chin and under-bite. Tissue-depth data from contemporary German men in their sixties were employed for the reconstruction (Helmer, 1984). Bach was known to have sat for the painter Elias Gottlob Haussmann in 1746, a few years before his death (Müller, 1935). Haussmann (1695–1774) served as court painter at Dresden, and from 1720 as the official portraitist at Leipzig. Written records of Bach's life suggest that he had eye problems that left him with swollen, sore eyelids and the portrait also depicted this affliction (David and Mendel, 1945; Zegers, 2005). This portrait was, therefore, utilised to determine facial fatness and eye detail, eye colour, skin colour and facial ageing details (Figure 18.14). Skin colour, eye colour and skin texture details were taken directly from the portrait during the texture addition process (Wilkinson and Aitken, 2009). The combination of skeletal assessment and portrait evaluation presented a facial depiction of Bach that can be considered as accurate as is possible with the material available (Hansen, 2008).

References

Angel, J. L. (1978). Restoration of head and face for identification. *Proceedings of Meetings of American Academy of Forensic Science*, St. Louis, MO.

Arridge, S. R., Moss, J. P., Linney, A. D. and James, D. R. (1985). Three-dimensional digitisation of the face and skull. *Journal of Maxillofacial Surgery*, **13**, 136–143.

Aulsebrook, W. A., Becker, P. J. and İşcan, M. Y. (1996). Facial soft-tissue thickness in the adult male Zulu. *Forensic Science International*, **79**, 83–102.

BBC News (2010). Dundee scientist brings Shakespeare's face to life. 13 September 2010. http://www.bbc.co.uk/news/uk-scotland-tayside-central-11281581 [accessed: 15 September 2010].

BBC2 (2004). The Real Face of Santa. 18 December http://news.bbc.co.uk/cbbcnews/hi/newsid_4090000/newsid_4091400/4091447.stm

Birkner, F. (1904). Beitrage zur rassenanatomie der gesichtsweichteile. *Corr Bl Anthrop Ges Jhg*, **34**, 163–165.

Buhmann, D., Bellman, D., Kahler, K. *et al.* (2003). Computer-aided soft tissue reconstruction on the skeletonised skull. *Proceedings of the 1st International Conference on Reconstruction of Soft Facial Parts* (*RSFP*), pp. 37–39. Potsdam, Germany.

Codinha, S. (2009). Facial soft tissue thicknesses for the Portuguese adult population. *Forensic Science International*, **184**, 1–80.

Colledge, H. (1996). Loss a face? The effect on the outcome of cranio CFR when part of the skull is missing. Thesis submitted to the University of Manchester for the degree of MSc in Medical Art in the Faculty of Medicine.

David, H. T., and Mendel, A. (1945). *The Bach Reader: A Life of Johann Sebastian Bach in Letters and Documents*. New York, NY: WW Norton and Co.

Davy, S. L., Schofield, D. and Evison, M. P. (2005). Creating a 3D skull model from two-dimensional images: problems and practicalities in computerised CFR. *Proceedings of the 2nd International Conference on Reconstruction of Soft Facial Parts*. Remagen, Germany.

De Greef, S., Vandermeulen, D., Claes, P., Suetens, P. and Willems, G. (2009). The influence of sex, age and body mass index on facial soft tissue depths. *Forensic Science, Medicine and Pathology*, **5**, 60–65.

D'Hollosy, M. (2005). http://www.skullpting.nl/english/index1.html.

Domaracki, M. and Stephan, C. N. (2006). Facial soft tissue thicknesses in Australian adult cadavers. *Journal of Forensic Science*, **51**(1), 5–10.

Eliasova, H., Dvorak, D. and Prochazka, I. O. (2003). Facial three-dimensional reconstruction. *Proceedings of the 1st International Conference on Reconstruction of Soft Facial Parts* (*RSFP*), pp. 45–48. Potsdam, Germany.

El-Mehallawi, I. H., and Soliman, E. M. (2001). Ultrasonic assessment of facial soft tissue thickness in adult Egyptians. *Forensic Science International*, **117**, 99–107.

Evenhouse, R. M., Rasmussen, M. and Sadler, L. (1992). Computer-aided forensic facial reconstruction. *Journal of Biological Chemistry*, **19**, 22–28.

Evison, M. P., Davy, S. L., March, J. and Schofield, D. (2003). Computational forensic facial reconstruction. *Proceedings of the 1st International Conference on Reconstruction of Soft Facial Parts* (*RSFP*), pp. 29–34. Potsdam, Germany.

Fischer, E. (1905). Anatomische Untersuchungen an den Kopfweichteilen zweier Papua. *Corr Bl Anthrop Ges Jhg*, **36**, 118–122.

Friend, T. (2003) Could this be the profile of a queen? *USA Today*. http://www.usatoday.com/news/science/2003–08–12-nefertiti-usat_x.htm [accessed 8 December 2003].

Galdames, I. C. S., Alarcón, G. E. S. and Matamala, D. A. Z. (2008). Best parameters for sexual dimorphism in the facial thickness tissue with ultrasonic assessment. *Saúde, Ética and Justiça*, **13**(2), 60–64.

Gatliff, B. P. (1984). Facial sculpture on the skull for identification. *American Journal of Forensic Medicine and Pathology*, **5**(4), 327–332.

George, R. M. (1987). The lateral craniographic method of facial reconstruction. *Journal of Forensic Sciences*, **32**(5), 1305–1330.

Gerasimov, M. M. (1955). *The Reconstruction of the Face from the Basic Structure of the Skull*. Translated by W. Tshernezky. Russia: Publishers unknown.

Hansen, J. (2008). *Bach Through the Mirror of Medicine. Exhibition Catalogue*. Bachhaus Eisenach.

Helmer, R. (1984). *Schädelidentifizierung durch elektronische Bildmischung*. Heidelberg: Kriminalistik-Verlag.

His, W. (1895). Anatomische Forschungen uber Johann Sebastian Bach's Gebeine und Antlitz nebst Bemerkungen uber dessen Bilder. *Abh MathPhysikal KI Kgl Sachs Ges Wiss*, **22**, 379–420.

Hjalgrim, H., Lynnerup, N., Liversage, M. and Rosenklint, A. (1995). Stereolithography: potential applications in anthropological studies. *American Journal of Physical Anthropology*, **97**, 329–333.

Holdaway, R. A. (1983). A soft tissue cephalometric analysis and its use in orthodontic treatment planning. *Part I. American Journal of Orthodontics*, **84**, 1–28.

Ismail, N. (2008). *Computerised 3D craniofacial remodelling of incomplete skulls: the effect on facial reconstruction*. MSc Human Identification Dissertation, University of Dundee.

Kahler, K., Haber, J. and Seidel, H. P. (2003). Reanimating the dead: reconstruction of expressive faces from skull data. *ACM/SIGGRAPH Computer Graphics Proceedings*, **22**(3), 554–567.

Kindermann, K. (2003). Innovative approaches to facial reconstruction using digital technology. *Proceedings of the 1st International Conference on Reconstruction of Soft Facial Parts* (*RSFP*), p. 127–132. Potsdam, Germany.

Koch, R., Gonzales, A. and Witt, E. (1979). Profile and soft tissue changes during and after orthodontic treatment. *European Journal of Orthodontics*, **1**, 193–199.

Kollmann, J. and Buchly, W. (1898). Die Persistenz der Rassen und die Reconstruction der Physiognomie prahistorischer Schadel. *Archives fur Anthropologie*, **25**, 329–359.

Lebedinskaya, G. V., and E. V. Veselovskaya (1986). Ultrasonic measurements of the thickness of soft facial tissue among the Bashkirs. *Annales Academiae Scientiarium Fennicae Ser. A. 5 Medica*, **175**, 91–95.

Lee, W. J., Wilkinson, C. M. and Hwang, H. S. (2011). An accuracy assessment of forensic computerised facial reconstruction employing cone-beam computed tomography from live subjects *Journal of Forensic Sciences*.

Lorenzi, R. (2004). Santa Claus' face reconstructed. *Discovery News*. http://dsc.discovery.com/news/briefs/20041213/santa.html [accessed 17 December 2004].

Mackenzie, S. (2007). *An evaluation of computer assisted skull re-assembly using hand held and automated laser scanners*. MSc Human Identification dissertation, University of Dundee.

Macris, T. F. (1987). Composite art, general principles for man and machine. *Identification News*, August 1987.

Mahoney, G. and Rynn, C. (2011). Facial reconstruction: new approaches. In M. Thali, M. D. Viner and B. G. Brogdon (Eds.), *Brogdon's Forensic Radiology*, 2nd edn, pp. 479–484. , Boca Raton, FL: CRC Press.

Manhein, M. H., Listi, G. A., Barsley, R. E. *et al.* (2000). In vivo facial tissue depth measurements for children and adults. *Journal of Forensic Sciences*, **45**(1),48–60.

Martino, L. (1987). Le Reliquie di S.Nicola – Studio anatomo-anthropologico dei resti scheletrici rinvenuti nella sua tomba in Bari (alla ricognizione canonica del 5 maggio 1953). *Biblioteca di San Nicola*, Centro Studi Nicolaiani, Bari.

Mires, A. M. W., Mahoney, G., Manhein, M. H. and Barrows, E. B. (2003). Forensic facial reconstruction using FreeForm software, poster presentation. *Proceedings of the American Academy of Forensic Sciences*, **9**.

Moss, J. P., Linney, A. D., Grindrod, S. R., Arridge, S. R. and Clifton, J. S. (1987). Three-dimensional visualisation of the face and skull using computerised tomography and laser scanning techniques. *European Journal of Orthodontics*, **9**, 247–253.

Müller, G. (1935). *Bach Portraits. The Musical Quarterly*. Oxford: Oxford University Press.

Phillips, V. M., and Smuts, N. A. (1996). Facial reconstruction: utilization of computerized tomography to measure facial tissue thickness in a mixed racial population. *Forensic Science International*, **83**(1), 51–59.

Rhine, J. S., and Campbell, H. R. (1980). Thickness of facial tissues in American Blacks. *Journal of Forensic Sciences*, **25**(4), 847–858.

Rhine, S. (1983). *Tissue thickness for South-western Indians*. PhD thesis. Physical Anthropology Laboratories. Maxwell Museum, University of New Mexico.

Rynn, C. and Wilkinson, C. M. (2009). Prediction of nasal morphology from the skull. Abstract from IACI 2008 conference. *Journal of Forensic Science, Medicine and Pathology*, **5**(1), 50–51.

Sahni, D., Jit, I., Gupta, M., Singh, P. and Suri, S. (2002). Preliminary study on facial soft tissue thickness by magnetic resonance imaging in Northwest Indians. *Forensic Science Communications*, **4**(1).

SensAble Technologies (2010*a*). http://www.sensable.com/haptic-phantom-desktop.htm [accessed 15 September 2010].

SensAble Technologies (2010*b*). http://www.sensable.com/products-freeform-systems.htm [accessed 15 September 2010].

Spoor, F., Jeffery, N. and Zonneveld, F. (2000). Imaging skeletal growth and evolution. In P. O'Higgins and M. Cohn (Eds.), *Development, Growth and Evolution: Implications for the Study of the Hominid Skeleton*, pp. 124–161. London: Academic Press.

Stephan, C. (2002). Facial approximation: globe projection guideline falsified by exophthalmometry literature. *Journal of Forensic Science*, **47**(4), 730–735.

Stephan, C. N., and Henneberg, M. (2003). Predicting mouth width from inter-canine width – a 75% rule. *Journal of Forensic Science*, **48**(4), 725–727;

Stephan, C. N., and Murphy, S. J. (2008). Mouth width prediction in craniofacial identification: cadaver tests of four recent methods, including two techniques for edentulous skulls. *Journal of Forensic Odontostomatology*, **27**(1), 2–7.

Stratomeier, H., Spee, J., Wittwer-Backofen, U. and Bakker R. (2005). Methods of forensic facial reconstruction. *Presented at the 2nd International Conference on Reconstruction of Soft Facial Parts* (*RSFP*), Remagen, Germany.

Subke, J. and Wittke, M. (2005). CAD Enhanced soft-tissue reconstruction in forensics with Phantom® 3D Touch – an electronic modelling tool with haptic feedback. *Presented at the 2nd International Conference on Reconstruction of Soft Facial Parts* (*RSFP*), Remagen, Germany.

Subtelny, J. D. (1959). A longitudinal study of soft tissue facial structures and their profile characteristics, defined in relation to underlying skeletal structures. *American Journal of Orthodontics and Dentofacial Orthopedics*, **45**(7), 481–507.

Sunday Times (2004). Now do you believe in Santa Claus? *Sunday Times Magazine*, 12 December, pp. 19–24.

Suzuki, H. (1948). On the thickness of the soft parts of the Japanese face. *Journal of the Anthropological Society of Nippon*, **60**, 7–11.

Talass, M. F., Talass, L. and Baker, R. C. (1987). Soft-tissue profile changes resulting from retraction of maxillary incisors. *American Journal of Orthodontics and Dentofacial Orthopedics*, **5**, 385–394.

Tedeschi-Oliveira, S. V., Melani, R. F. H., de Almeida, N. H. and de Paiva, L. A. S. (2009). Facial soft tissue thickness of Brazilian adults. *Forensic Science International*, **193**, 1–127.

Tilotta, F., Richard, F., Glaune, J. *et al.* (2009). Construction and analysis of a head CT-scan database for craniofacial reconstruction. *Forensic Science International*, **191**, 112–124.

Utsuno, H., Kageyama, T., Deguchi, T. *et al.* (2005). Facial soft tissue thickness in Japanese female children. *Forensic Science International*, **152**, 101–107.

Utsuno, H., Kageyama, T., Deguchi, T. *et al.* (2007). Facial soft tissue thickness in skeletal type I Japanese children. *Forensic Science International*, **172**, 137–143.

Vander Pluym, J., Shan, W. W., Taher, Z. *et al.* (2007). Use of magnetic resonance imaging to measure facial soft tissue depth. *Cleft Palate–Craniofacial Journal*, **44**, 52–57.

Vanezis, P., Blowes, R. W., Linney, A. D. *et al.* (1989). Application of three-dimensional computer graphics for facial reconstruction and comparison with sculpting techniques. *Forensic Science International*, **42**, 69–84.

Vanezis, P., Vanezis, M., McCombe, G. and Niblett, T. (2000). Facial reconstruction using 3-D computer graphics. *Forensic Science International*, **108**(2), 81–95.

Vermeulen, L. (2005). http://www.ludovermeulen.be/index.php?pag=home

Von Eggeling, H. (1909). Anatomische untersuchungen an den Kopfen con cier Hereros, einem Herero- und einem Hottentottenkind. *Forschungsreise im westrichen und zentraien Sudafrika*, pp. 323–348. Schultze. Jena, Denkschriften.

Wilkinson, C. M. (2002). In vivo facial tissue depth measurements for White British children. *Journal of Forensic Science*, **47**(3), 459–465.

Wilkinson, C. M. (2003). 'Virtual' sculpture as a method of computerised facial reconstruction. *Proceedings of the*

1st International Conference on Reconstruction of Soft Facial Parts (*RSFP*), pp. 59–63. Potsdam, Germany.

Wilkinson, C. M. (2004). *Forensic Facial Reconstruction*. Cambridge: Cambridge University Press.

Wilkinson, C. M. (2005). Computerized forensic facial reconstruction – an accuracy study. *Proceedings of the 2nd International Conference on Reconstruction of Soft Facial Parts* (*RSFP*). Remagen, Germany

Wilkinson, C. M., and Aitken, J. (2009). The face of Bach. *Abstract from IACI 2008 conference. Journal of Forensic Science, Medicine and Pathology*, **5**(1), 54.

Wilkinson, C. M., Motwani, M. and Chiang, E. (2003). The relationship between the soft tissues and the skeletal detail of the mouth. *Journal of Forensic Science*, **48**(4), 1–5.

Wilkinson, C. M., Rynn, C., Peters, H. *et al*. (2006). A blind assessment of computer-modelled forensic facial reconstruction using computed tomography data from live subjects. *Journal of Forensic Science, Medicine and Pathology*, **2**(3), 179–187.

Zegers, R. H. C. (2005). The eyes of Johann Sebastian Bach. *Archives of Ophthalmology*, **123**, 1427–1430.

Chapter 19

Craniofacial superimposition

Mineo Yoshino

19.1 Introduction

The identification of skeletal remains has been achieved by several methods: matching of ante-mortem dental records with teeth of the skull, matching of ante-mortem radiographs with various skeletal parts, matching of facial photographs with the skull by craniofacial superimposition and DNA analysis of bones or teeth and reference samples. Ante-mortem dental records and radiographs have been generally accepted to suggest positive identification of skeletal remains. However, the availability of these ante-mortem records is relatively rare in criminal cases. Molecular approaches such as mitochondrial DNA (mtDNA) analysis have been applied to bone identification and have improved an assessment for personal identification, but such analyses are expensive and time-consuming. Craniofacial superimposition is the most popular technique for identifying unknown skulls because facial photographs of the presumed person can be easily obtained from the victim's family.

In the earliest stage of craniofacial superimposition, the skulls of historically important individuals such as Immanuel Kant, Johann Sebastian Bach and Josef Haydn were compared with relevant portraits, busts and death masks using skull morphology, osteometry and physiognomy (Stewart, 1979; Krogman and İşcan, 1986; Gruner, 1993; Taylor and Brown, 1998; Glassman, 2001). Although these comparisons were performed for academic purposes, their methods were basically applied to subsequent research in the field of forensic sciences.

From the forensic standpoint, the most instructive example of the application of craniofacial superimposition technique is provided by the two skulls from the Ruxton murder case of 1935 in England (Glaister and Brash, 1937). In this case, there were two female skulls, listed as skull no. 1 and skull no. 2. It was assumed that one was that of Mrs Ruxton and the other belonged to the housemaid Mary Rogerson. Comparison using a superimposition technique demonstrated that skull no. 1 could not possibly be the skull of Mrs Ruxton, and that skull no. 2 might be the skull of Mrs Ruxton. It was also demonstrated that skull no. 2 could not possibly be the skull of Miss Rogerson, but that skull no. 1 might be the skull of Miss Rogerson. In this case, the craniofacial superimposition evidence was not accepted as a means for confirming positive identification of two remains, but was allowed as supporting evidence in the identity of the two skulls by the court. With methodological viability established in the forensic community, craniofacial superimposition has been applied to many cases, the most famous of which are the Dobkin case (Simpson, 1943), the Wolkersdorfer case (Gordon and Drennan, 1948), the Howick Falls murder case (Prinsloo, 1953) and the identification of the remains of Nazi war criminal, Josef Mengele (Teixeira, 1985; Helmer, 1987). Other case reports also indicated the effectiveness of craniofacial superimposition for personal identification (Sen, 1962; Basauri, 1967; Vogel, 1968; Sekharan, 1971; Reddy, 1973; Gejvall, 1974; Sivaram and Wadhera, 1977; Janssens *et al.*, 1978; Brown, 1982; McKenna *et al.*, 1984; Bastiaan *et al.*, 1986; Webster *et al.*, 1986; Loh and Chao, 1989; Yoshino *et al.*, 1989; Ubelaker *et al.*, 1992; Shahrom *et al.*, 1996; Solla and İşcan, 2001; Bilge *et al.*, 2003; Ghosh and Shinha, 2005; Fenton *et al.*, 2008).

As a general rule, it can be stated that craniofacial superimposition is of greater value in ruling out a match between the skull and the facial photograph.

Craniofacial Identification, ed. C. Wilkinson and C. Rynn. Published by Cambridge University Press.

In some cases, a facial photograph of another person may well be consistent with the skull in question. Therefore, forensic examiners must be well versed in the anatomy of the skull and the face for the effective utilization of craniofacial superimposition technique. When evaluating anatomical consistency between these parts, special attention should be paid to their outline, soft-tissue thickness at various anthropometric landmarks of the skull and positional relationships between the skull and the face.

This chapter comprises an overview of the technical development of craniofacial superimposition, criteria for assessing a match between the skull and the face, and the reliability of this technique in personal identification.

19.2 Superimposition techniques

From the aspect of technical development, craniofacial superimposition can be seen to have passed through three phases: the photograph, video and computer-assisted superimposition techniques (Aulsebrook *et al.*, 1995). The still photographic superimposition technique was developed in the middle of the 1930s. The video superimposition technique has been widely used since the latter half of the 1970s. In the latter half of the 1980s, the computer-assisted superimposition was introduced.

19.2.1 Still photographic superimposition

When performing still photographic superimposition, examiners must enlarge the comparison photograph to the size of the skull in question, and then position the skull in the same orientation as the facial photographs.

19.2.1.1 Enlargement of a facial photograph

The facial photograph is enlarged to scale using measurable objects (Glaister and Brash, 1937; Gordon and Drennan, 1948; Sekharan, 1971; Janssens *et al.*, 1978), anatomical landmarks in the skull and the face (Gruner and Reinhard, 1959; Dorion, 1983), anthropometric measurements and partial sizes of the skull and the face (Simpson, 1943; Reddy, 1973; McKenna *et al.*, 1984; Webster *et al.*, 1986; Loh and Chao, 1989; Yoshino and Seta, 1989; Jayaprakash *et al.*, 2001). Prinsloo (1953) enlarged the head of portrait to a size convenient for work, and then the print of the skull was made to a size slightly smaller than that of the enlargement of the head, to allow for the thickness of the covering soft tissues.

In the Ruxton case (Glaister and Brash, 1937), the life-sized enlargement of Mrs Ruxton's portrait was made based on the size of a tiara on her head. For Miss Rogerson's portrait, a gate was used as an enlargement scale. Gordon and Drennan (1948) used a tie, Sekharan (1971) a chair and a saree, and Janssens *et al.* (1978) used the button of a sweater as an enlargement scale.

Gruner and Reinhard (1959) used the appropriate cephalometric points and the corresponding craniometric points for enlargement of a portrait. Dorion (1983) enlarged the facial photograph to a 1:1 ratio using the skull as reference.

Simpson (1943) used the interpupillary distance and the distance between the two central points of the orbits. Reddy (1973) followed the criterion established by Simpson (1943): the distance between the two central points of the orbit of the skull equates to the interpupillary distance of the person to whom the skull belonged. Sekharan (1971) described that an approximate life-size enlargement of the face is obtained from a snapshot, keeping approximately the interpupillary distance as 6 cm. However, McKenna *et al.* (1984) suggested that the interpupillary distance varied widely in their research sample, and therefore, rejected the use of an average figure as the basis for photographic enlargement. Loh and Chao (1989) proposed a new approach, using the true interpupillary distance of a second person as the scale for the enlargement of the presumed person, when two people were shown in the same angle as the ante-mortem photograph. Yoshino and Seta (1989) recommended the use of craniometric measurements of the skull and the published data of soft-tissue thickness at anthropometric points for making a positive half-size transparency. The frontal view facial photograph is enlarged from the original photograph based on the bizygomatic breadth of the skull and the soft-tissue thickness at the zygion. In the oblique or lateral view, the facial photograph enlargement is based on the glabella–gnathion distance and the soft-tissue thickness at the gnathion. Jayaprakash *et al.* (2001) used the inter-Whitnall's malar tuberclar (Stewart, 1983) distance measured from the skull as the reference scale for bringing out the biorbital breadth in the facial photograph for life-size enlargement. They mentioned that the biorbital breadth provides the greatest horizontal distance in the skull and facial photograph thereby minimizing error. Maat (1989) noted that the larger the separation of the two relevant landmarks, the more accurate is magnification of the photograph. McKenna *et al.* (1984) and Webster *et al.* (1986) used measurements of the anterior teeth of an unidentified skull to prepare the life-size transparency when the anterior teeth are found in a facial photograph.

19.2.1.2 Orientation of the skull to facial photographs

Correct orientation of the skull to the facial photograph is another important problem in the superimposition technique. Sekharan (1993) described that the range of movements of the head includes the following factors: flexion or extension (forward to backward tilt), lateral flexion right or left (sideways bend) and rotation. According to him, flexion or extension is the most important for successful application of the superimposition technique.

A craniophore used for anthropometry and a pivoting head tripod are most commonly used. Glaister and Brash (1937) used a specially designed craniophore within a cubic metal frame. Gruner and Reinhard (1959) devised an optical bench that consisted of a skull rest, a frame with criteria lines, a thin plastic plate and a camera. They determined the position of the skull to correspond to the facial photograph with the help of outline drawings, making points and criteria lines. Gruner and Schulz (1969) recommended that the soft-tissue thickness be marked at the anthropometric points with a corresponding layer of plasticine or wax to improve precision.

To date, a number of devices have been designed to position the skull so that it will adjust to the orientation of the still or video image of the face. McKenna *et al.* (1984) mounted the skull on a rest adapted from a 'phantom-head' holder (used by dental students), which permits the skull to be oriented in three spatial planes. McKenna (1988) also devised a goniometer with a video camera to be able to document and reproduce the position of the skull relative to the attitude of the face in the photograph. Bastiaan *et al.* (1986) developed an adjustable support for the skull with protractor scales, which allows controlled movement in three dimensions and its position to be quantitatively shown. Brocklebank and Holmgren (1989) designed an armature to hold the skull and a pan-and-tilt device, incorporating calibrated measurement scales, enabling independent movements in each of the Cartesian coordinates. However, these devices are awkward and time-consuming because the examiner must manually rotate the skull and view the transparent facial photograph simultaneously. Some motor-driven skull positioning devices have been developed (Yoshino and Seta, 1989; Kumari and Sekharan, 1992; Lan, 1992; Sekharan, 1993; Seta and Yoshino, 1993; Yoshino *et al.*, 1997; Taylor and Brown, 1998) which provide multi-positional movement of the skull by remote control. In 1987, Taylor and Brown developed a system that would simulate the natural movement of the head and permitted controlled, reproducible movement of the skull in six directions (Taylor and Brown, 1998). They described that precision adjustments could be made in 1 mm steps and 1 degree segments. Kumari and Sekharan (1992) designed a skull-positioning device comprising a tripod stand and a pan-and-tilt head unit equipped with a craniophore. Lan (1992) devised a skull position adjusting mechanism capable of manoeuvering the skull freely in six directions. Yoshino *et al.* (1997) used a pulse motor-driven mechanism for controlling the movement of the skull instead of the chain-gear motor-driven mechanism used in the previous system (Yoshino and Seta, 1989; Seta and Yoshino, 1993). The pulse motor-driven mechanism was superior to the chain-gear one in several respects such as speed, precision and fine control. The skull rest can be moved by remote control levers in six directions: upward and downward (± 20 mm), left and right (± 40 mm), forward and backward (± 40 mm), rotation around the axis (± 100°), extension and flexion (± 30°), and tilt to the left and right (± 15°). The combination of these movements adequately establishes the orientation of the skull to multi-positional facial photographs.

19.2.1.3 Procedure for still photographic superimposition

'Viewing through' techniques, in which the orientation of the skull to the facial photograph is achieved by looking through a transparent film of the face, are optimal for still photographic superimposition.

In the Ruxton case (Glaister and Brash, 1937), the orientation of the skull to match the portrait was based on the probable angle of rotation of the head in the portrait, then the skull position was adjusted to conform as closely as possible to the head position by placing a full-size transparent outline of the portrait on the viewing screen of the camera. To compare the outlines of the skull and portrait, the salient features of both full-size photographs were outlined with Indian ink, and transferred to tracing papers. Two anthropometric points, the nasion and the prosthion, were marked on the skull and portrait outlines. The transparent outline of the portrait was then superimposed over the outline of the skull by means of these points and the lines between them. Four registration marks were placed on the corner of the superimposed tracing of outlines. In order to demonstrate the direct photographic superimposition, the tracing paper with the superimposed outlines was then placed in turn on the positive print of the portrait and a paper negative of the skull, and the registration marks transferred to each by pricking through their central points. The positive portrait and the negative skull were re-photographed on X-ray film, thus producing a transparent negative of the portrait and a transparent positive of the skull, each with the registration marks photographed. Finally, the two films were superimposed by piercing the central points of the registration marks, and were then re-photographed on X-ray film by transmitted light. This method was modified and used in several cases (Sen, 1962; Basauri, 1967; Sekharan, 1971; Reddy, 1973; Sivaram and Wadhera, 1977). According to Reddy (1973), an enlarged positive print of the face was taken from the negative supplied and then reproduced on quarter-size cut film. The negative thus obtained was placed under the ground glass of the camera, and the outline of the face was drawn using very fine pen and Indian ink. A 1-inch scale was placed on the forehead of the skull, and the skull placed on an appropriate skull rest that enabled it to be arranged in the same perspective as that of the portrait. The skull was then photographed on quarter-size cut film. The negative was enlarged to natural size using the 1-inch scale, then printed on a bromide paper. A positive transparency was then taken on 24 × 30 cm X-ray film. Finally, the natural-size transparency of the skull and the life-size transparency of the portrait were superimposed and photographed on quarter-size cut film using transmitted light.

As described in the preceding section, the optical bench was specially designed for the 'viewing through' technique (Gruner and Reinhard, 1959). The optical bench was 3 m in length, the skull was fixed at one end and a camera fixed at the other end. The appropriate anthropometric points (e.g. nasion, gnathion, etc.) were marked on an enlarged transparent photograph of the head and then three guidelines were drawn based on these points with Indian ink. Three criteria lines within the frame were adjusted to the guidelines on the transparent photograph. The outline of the head and the guidelines were drawn on the thin plastic plate. When the criteria lines of the frame coincided with the guidelines on the thin plastic plate, the skull position was set, taking soft-tissue thickness into account. The enlarged transparent photograph of the head without the marker points and the guidelines was set on to the thin plastic plate, and then photographed. Finally, removing the frame and thin plastic plate, a double exposure of the positioned skull was taken. In the optical bench devised by Brocklebank and Holmgren (1989), a camera was attached to two parallel rails allowing the distance between camera and skull to be varied.

Dorion (1983) devised a special apparatus for craniofacial superimposition. A Graflex 4×5 camera with a 135 mm lens at f-4.5 was mounted directly in front of the skull at a distance of 95.25 cm. The beam splitter and front surface mirror used for the 'viewing through' device were positioned so as to reflect the image of the subject's photographic transparency. By varying the intensity of the quartz lights mounted in front of the skull, a superimposed image of the skull and facial photographic transparency was obtained when viewing through the camera. The skull was mounted on a pivoting head tripod and moved until its orientation coincided with that of the facial photographic transparency.

In 1988 a combined apparatus for photographic and video superimposition was designed (Yoshino and Seta, 1989; Seta and Yoshino, 1993). This apparatus reproduced the skull image on ground glass at half the original size, via four plates of reflection mirror and a projection lens (Fujinon 600 mm, f-11.5). The facial photographic transparency was made from the original photograph, to correspond to half of the skull size, with enough space to account for the thickness of soft tissue at anthropometric points. The image of the skull on the ground glass was then photographed using a cabinet-size film. The skull and face films were finally superimposed on the light box based on anthropometric landmarks. In this apparatus, the distance between the skull and the camera was set at about 1.7 m to reduce perspective distortion of the skull image. Maat (1989) mentioned that the principle of central projection and a minimum photographic distance of 1.5 m are important preconditions. Lan and Cai (1993) described that the optimal lens distance from the skull for superimposition is 1 m. Eliasova and Krsek (2007) presented a mathematical model of the standard projective camera and the mathematical description of perspective to eliminate the distortion of the skull image in the superimposition process.

Jayaprakash *et al.* (2001) concede that the application of mensural parameters is intrinsically inadequate for bringing out the life-size enlargement or for positioning the skull, and achieving the final correspondence between the skull and face photograph continues to remain an 'expert' job.

19.2.2 Video superimposition

Since Helmer and Gruner (1977) introduced the video superimposition technique to skull identification, this technique has been popularly used in the fields of forensic anthropology and odontology. The video superimposition presents enormous advantage over the conventional photographic one. This technique eliminated the time-consuming enlargement of the facial photograph to match the skull, and allowed control over the level of transparency of either the skull or the facial image (fade-in and fade-out) on the TV monitor for overall assessment of how well the two images match. This also allowed the various sectioning images of the skull and facial image for evaluating the positional relationships between the skull and the face.

19.2.2.1 Equipment

Helmer and Gruner (1977) used two video cameras for taking images of a skull and a facial photograph, an electronic mixing unit and two TV monitors, and recorded the process of superimposition. They mentioned that this method was simple to use and offered ease of recognisability of relationships between bone and soft tissue. Koelmeyer (1982) applied this technique to his comparison of a skull, facial photographs and a radiograph. He utilised three video cameras to reproduce the images of the facial photograph, the skull X-ray and the partially reassembled skull. Using the switcher joystick to fade-out or fade-in and sweep, he progressively superimposed the X-ray picture and the skull image on the facial photograph on both vertical and horizontal axes. Koelmeyer's assessment of the technique was that the fade-out or fade-in produced a more pleasing image, while the sweep capability permitted better comparison of bony landmarks. Brown (1983) used two video cameras and a special-effects generator to display simultaneous single-screen images from a skull and a facial photograph. The special-effects generator provided a variety of combinations of the superimposed images on the monitor. Dorion (1983) and Loh and Chao (1989) employed one video camera for superimposition. They used a transparency of the photograph of the face set just in front of the skull, and displayed the skull and transparency on a single monitor. Krogman and İşcan (1986) performed a number of superimpositions using two three-plumbic commercial studio cameras, a switcher for special effects mixing and dissolving, an ordinary 19-inch monitor, and 3/4 inch videotapes. Iten (1987) recommended the following equipment for video superimposition: two video tubes, electronic and mixer units, and three monitors. The skull was photographed using tube 1 and the image reproduced on monitor 1. The facial photograph was then reproduced with tube 2 on monitor 2. The mixed picture and horizontal and vertical sections were appraised on monitor 3 by means of the electronic mixer units. Delfino *et al.* (1993) utilised one video camera equipped with an optical device that is a plane-parallel semi-reflective glass in a rectangle for taking both the skull and facial images. Solla and İşcan (2001) utilised equipment consisting of one video camera with tripod, a digital video mixer unit, a video recorder, a monitor of high-resolution image, a character generator, a skull-positioning stand and a photographic stand.

In general, the video superimposition system consists of two video cameras, a skull-positioning rest, a photographic stand for the facial photograph, a video-image mixing device, a TV monitor and a videotape recorder (Bastiaan *et al.*, 1986; Yoshino and Seta, 1989; Lan, 1992; Seta and Yoshino, 1993; Austin-Smith and Maples, 1994; Shahrom *et al.*, 1996; Taylor and Brown, 1998; Jayaprakash *et al*, 2001; Fenton *et al.*, 2008). Some research groups adapted the motor-driven skull rest to video superimposition to quickly adjust the position of the skull to the orientation of the facial image (Yoshino and Seta, 1989; Lan, 1992; Seta and Yoshino, 1993; Taylor and Brown, 1998).

19.2.2.2 Procedure for video superimposition

To make the comparison, examiners must strive to establish the orientation of the skull–facial photograph as firmly in the video superimposition as it is in the still photographic superimposition. The size of the skull image is easily adjusted to that of the facial image on the monitor using the zoom mechanism of the video camera, taking soft-tissue thickness into account. Tissue thickness markers are placed on the skull at the appropriate anatomical landmarks (Austin-Smith and Maples, 1994; Fenton *et al.*, 2008). Krogman and İşcan (1986) noted that the duration of the entire video process was approximately 1 h and can be accomplished by two examiners, one to watch the monitor and the other to adjust the camera orientation.

In the video-transparency superimposition technique (Dorion, 1983; Loh and Chao, 1989; Shahrom *et al.*, 1996; Solla and İşcan, 2001), the facial photograph is enlarged approximately to life-size and photocopied onto a transparency. According to Shahrom *et al.* (1996), the skull is placed on a skull holder and adjusted to the same position as that of the facial photograph. It is viewed through a video camera and the image can be seen on a TV monitor. The transparency is then superimposed over the skull image on the TV monitor. The size of the skull image is increased or reduced by adjusting the focal length and the focusing of the video camera in such a way that the skull image fits the facial image based on the anatomical landmarks. The landmarks used are the lateral aspect of the orbits to the ectocanthia, the anterior nasal spine to the uppermost part of the philtrum and the occlusal (bite) line of the teeth to the lip line of a closed mouth, respectively. This technique was slightly modified by Solla and İşcan (2001), who took the facial image with the video camera. The facial image was digitised and stored within the digital video mixer unit, and then a transparent plastic sheet was taped to the monitor. Key anatomical landmarks of the facial image were traced onto this sheet. After removing the facial image from the monitor by the digital mixer controls, the skull image was taken with the same video camera and then the orientation of the skull was adjusted based on key anatomical landmarks taking soft-tissue thickness into account. Finally, the skull was manipulated manually until the position approximated that of the facial image.

When two video cameras were used for superimposition, the facial photograph was firstly taken with one of the two cameras so the image fills most of the TV monitor, and then the skull is taken with the other camera. The skull image is sized so that it can be superimposed on the facial image on the TV monitor. After the skull and facial images are adjusted for size and basic or gross orientation, the exact orientation of the skull image is accomplished by finely adjusting the outlines and the anatomical landmarks between both images. The fade-out and wipe image modes help this adjusting process. Iten (1987) used the ratio of the distance between the eyes, and the distance between the eye and auditory canal as axes for adjusting the correct rotation and inclination of the skull, respectively. Gross positioning of the skull in relation to the orientation of the face in the photograph was achieved using some calculating methods (Lan and Cai, 1993). Shahrom *et al.* (1996) applied a different method to determine the skull position. Since there is some difficulty in fine positioning of the skull, the facial photograph is rotated in the coronal plane to match the skull angulation along that plane. Manipulation of the skull position or angulation using the skull holder is done only in two planes, i.e. the sagittal and horizontal planes. Jayaprakash *et al.* (2001) found it expedient to orient the skull in relation to the face in the photograph by aligning the nasal perspective in the skull and the facial photograph using the horizontal sweep facility in the video superimposition device. Fenton *et al.* (2008) used the 'dynamic orientation process' to arrive at the best fit possible in the alignment of the skull with the facial photograph. The first step in the dynamic orientation process is to align the skull and facial images at the porion. In the second step, the left and right Whitnall's malar tubercles are aligned with the left and right ectocanthion points of the face. These steps are critical in setting the skull at the correct angle of inclination and declination so that it approximates the angle of the facial image. Subsequently, the subnasale and gnathion points of the skull are adjusted to align with the subnasale and gnathion on the facial image, respectively. Even in the video superimposition, the exact positioning of the skull is determined by trial-and-error manipulation of the skull (Austin-Smith and Maples, 1994).

The superimposed image on the monitor can be photographed directly with the camera and the entire process can be recorded on videotape.

19.2.3 Computer-assisted superimposition

As computer technology advances, computer-assisted superimposition has become a popular method of identifying an unknown skull. Computer-assisted superimposition is divided roughly into two categories from the point of view of identification strategy. The first method is to digitise the skull and facial photograph using a video-computer with appropriate software, and then to compare the two images morphologically by image processing (Chai *et al.*, 1989; Lan, 1992; Ubelaker *et al.*, 1992; Cai and Lan, 1993; Lan and Cai, 1993; Yoshino *et al.*, 1997; Bilge

Figure 19.1 Computer-assisted skull identification system using video superimposition. A: Skull-positioning box; B: Video superimposition unit; C: Computer-assisted skull identification unit; D: Photo-stand for facial photograph. The skull position is moved in any direction by remote control lever using the pulse motor-driven mechanism.

et al., 2003; Al-Amad *et al.*, 2006). Some researchers have attempted to automatically superimpose the skull image on the facial images using genetic algorithms (Nickerson *et al.*, 1991; Ibanez *et al.*, 2008, 2009) and neural networks (Ghosh and Shinha, 2001). Ibanez *et al.* (2009) classified these methods as computer-based craniofacial superimposition. According to them, the works of Ubelaker *et al.* (1992) and Yoshino *et al.* (1997) are typical examples of computer-assisted superimposition in the sense that a digital infrastructure supports the process, but its potential utility is not used to automate it. The second is to evaluate the fit between the skull and facial image by morphometric examination (Delfino *et al.*, 1986, 1993; Bajnoczky and Kiralyfalvi, 1995; Yoshino *et al.*, 1997; Ricci *et al.*, 2006).

19.2.3.1 Morphological examination using digitised images

Manual superimposition

Ubelaker *et al.* (1992) designed the simplest system, consisting of a collection of proprietary software and associated hardware such as a personal computer, a data tablet, a colour display monitor and a video camera. This system allowed the necessary images of the skull and facial photograph to be taken with only one camera and to directly compare both the digitised images, including the opportunity to remove the soft tissue to view the underlying skeletal structure on the monitor. In this system, the skull was manually manipulated to adjust the marked landmarks on the facial images. Lan and his group (Chai *et al.*, 1989; Lan, 1992; Cai and Lan, 1993; Lan and Cai, 1993) developed the image-superimposition processing system. They described that the relation between the proportion and distance between superimposed landmark lines can be checked by moving the cross cursor in accordance with the determining lines, interrelated relationships, soft-tissue thickness and profile morphological curves of Chinese adults stored in the program, and the value is measured on the monitor in real time and analysed by the computer.

Yoshino's skull identification system (Yoshino *et al.*, 1997) consists of two main units, namely a video superimposition unit and a computer-assisted skull identification unit (Figure 19.1). After the determination of the orientation and size of the skull to those of the facial photograph using the video superimposition system, the skull and facial photograph images are digitised and stored within the computer. The scaling of the digitised skull image is performed by converting the actual measurement between the landmarks (e.g. zygion–zygion, nasion–gnathion, etc.) into the number of pixels on the monitor. The comparison of positional relationships between the digitised skull and face is performed by fading out and in, or wiping in the vertical or horizontal plane with the mouse in operation. Furthermore, the distance between the anatomical landmarks and the thickness of soft tissue at the landmarks are measured semi-automatically (Figure 19.2).

Recently, commercial software such as Adobe Photoshop has been used (Bilge *et al.*, 2003; Al-Amad *et al.*, 2006). The scales of the skull and facial images

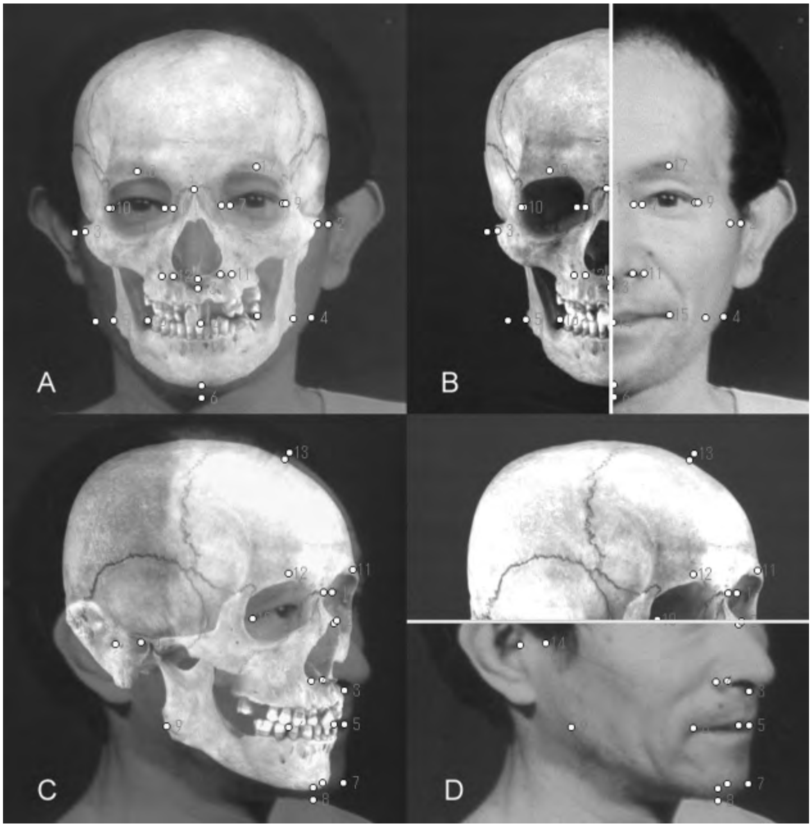


Figure 19.2 Superimposition of digitised skull and facial images. A and B: Total mixing and vertical wipe images of the frontal view; C and D: Five total mixing and horizontal wipe images of the oblique view. The distance between the anatomical landmarks and the thickness of soft tissue at the landmarks are measured by means of pair-points with mouse operation.

are adjusted using the 'free transforming' command. The facial image is transformed to transparent form by 40–50% using the 'set transparent' command, and then superimposed on the skull image.

Automatic superimposition

Nickerson *et al*. (1991) developed a methodology for comparing the skull and the facial image using the near-optimal fit between the three-dimensional (3D) skull surface mesh and the two-dimensional (2D) digitised facial photograph. In this method, the digitised 3D skull image is reconstructed from the original skull using a laser range finder. A genetic algorithm (binary-coded genetic algorithm) is used to find the optimal parameters of similarity and perspective transformation that overlays the 3D skull model on the 2D digitised facial image. Using four anatomical landmarks, the computer visually depicts the superimposed results. This technique could reduce or eliminate the previous problems with the skull/facial image scaling and orientation. The final image is reviewed by forensic specialists to ensure the sizing, exact positioning, and alignment of landmarks on the skull image fit the facial image. Ibanez *et al*. (2008) improved upon Nickerson's approach. They used a real-coded genetic algorithm and fuzzy location of cephalometric landmarks. According to them, genetic algorithms may be the most representative evolutionary algorithms. In their algorithm, 12 registration transformation parameters in a real-value array were encoded. The use of fuzzy landmarks aims to address the problem of location uncertainty in the facial image. Every landmark is given by a rectangular array of pixels defined by diagonally opposite corners. Furthermore, Ibanez *et al*. (2009) designed three different evolutionary algorithms, that is, two variants of a real-coded genetic algorithm and a covariance matrix adaptation evolution strategy (CAM-ES) in order to find a good fit between the 3D skull model and the 2D facial image. They evaluated the actual performance of these algorithms including a binary-coded genetic algorithm in three forensic identification cases, and then concluded that CAM-ES had a good performance, achieving high-quality solutions in all the cases and showing a high robustness. They mentioned that this method could be used now as a tool for automatically obtaining a good quality superimposition to be manually refined by the forensic expert in a quick way.

Ghosh and Shinha (2001) designed a craniofacial superimposition system, an extended symmetry perceiving adaptive neuronet (ESPAN), from a special automated version of the photo/video superimposition technique. Two 2D segment images of the skull, that is, the front skull plane and the upper skull contour are obtained, and then the final cranial image including both images is reconstructed. In this system, two networks needed to be separately trained for the front skull plane and the upper skull contour. ESPAN has a capability to take care of ambiguities due to soft-tissue thickness during facial feature selection in the nearly front facial photograph. By virtue of its robust self-deterministic decision-making capability, ESPAN searches out the specific features from input-ambiguous information by a reasonably organised defuzzification procedure (Ghosh and Shinha, 2001). The facial components such as the eyebrows, eyes, nose and lips are compared with the front skull plane, and the upper facial contour such as the zygomatic arch and the forehead with the upper skull contour using piece-wise combinations. ESPAN is an interactive program and does not implement a full automation, i.e. it should be complemented by human expertise.

19.2.3.2 Morphometric examination using digitised images

Delfino *et al.* (1986, 1993) developed an original Shape Analytic Morphometry software package yielding analytical descriptions by polynomial function, Fourier harmonic analysis and Janus procedure to assess the fit between the outline of the skull image and the facial image. The polynomial function is used for smoothing the curve representing the investigated contour. The square root of the mean of square error was used to calculate the distance between polynomial function curves. The Fourier analysis is a mathematical technique that was developed to describe complex waveforms by reducing them to a series of simple sinusoidal waves of different frequencies. In the Fourier analysis, lower-order harmonics represent the rough outline of the shape and higher-order ones correspond to increased details in the contour. The Janus procedure is used to evaluate the symmetry differences between two profiles. Yoshino *et al.* (1997) introduced the polynomial function and Fourier harmonic analysis that was proposed by Delfino *et al.* (1986, 1993) as the morphometric examination. For evaluating the fit between the outline of the skull image and the facial image, the corresponding segments of both outlines are extracted (Figure 19.3). The polynomial function curves obtained from Figure 19.3 are shown in Figure 19.4. In actual cases, these morphometric analyses cannot be always applied to compare the outline between the skull image and the facial image because the appropriate outlines of both images are extracted from only the lateral or oblique image. If quantitative data based on morphometric analyses are obtained in addition to the anatomical and anthropometric evaluation, the judgement of the skull identification will be more objective and reliable.

Bajnoczky and Kiralyfalvi (1995) mathematically evaluated the coordinate values of the pair-anatomical landmarks in the skull and the face using the normal frontal and lateral views. Eight to twelve pair-points were recorded and expressed as pixel units, and then the final matrix was established by computer-aided processing. They noted that although the result obtained from this method is objective and easily interpreted for lay people, it should be used only in combination with classic video-superimposition and could be regarded as an independent check.

Ricci *et al.* (2006) formulated an algorithm able to compare the facial image with the skull radiograph in the anterior position. In total, 14 anatomical points were manually marked with a cross on each facial image and corresponding skull radiograph, and then both images were superimposed by the operator. They mentioned that the algorithm is able to calculate the mean value of the total distance in crosses of the facial image and the radiograph for assessing the similarity, and to compensate for up to 10° of head rotation with head inclination being corrected during the pre-elaboration image movement phase.

19.3 Criteria for assessing anatomical consistency between the skull and the face

In order to identify someone through the craniofacial superimposition, examiners must investigate the outline, soft-tissue thickness at various anthropometric landmarks and positional relationships of the skull to

Figure 19.3 Digitised skull and facial images (A) and the selected segments (the forehead line between arrows) for assessing the fit between the outline of both images (B). Thin-line images (B) are automatically extracted from A, and then the selected segments are manually determined.

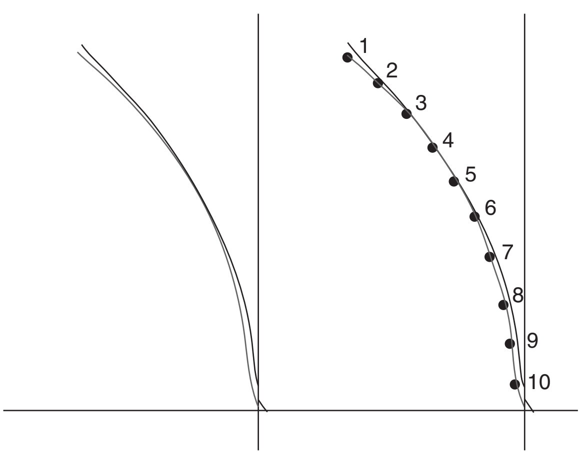


Figure 19.4 Polynomial function curves obtained from the selected segments in Figure 19.3 showing a good match (left) and 10 measurement points for calculating the distance between both curves (right). In this case, the sum of the horizontal distance between two curves is 7.2 mm.

face parts based on anatomical data after the size and orientation of the skull to those of the facial image are achieved. These criteria have been summarised (Krogman and İşcan, 1986; Yoshino and Seta, 1989, 2000, Taylor and Brown, 1998). Taylor and Brown (1998) pointed out that average soft-tissue thickness in specific locations might be used as guides but never as definitive measures. They also described that examiners must have an adequate knowledge and experience of the anatomy of the skull, the morphology of the face and the interrelationships of both. Readers can find detailed data and descriptions about the thickness of soft tissue at the anthropometric landmarks and positional relationships between the skull and the face in Chapter 18.

Austin-Smith and Maples (1994) showed morphological requirements for establishing a consistent fit between the skull and the face. They recommended that these morphological features in both the frontal and lateral views should correspond in the superimposition. Yoshino *et al.* (1997) used 18 examination criteria – four for the outline, seven for soft-tissue thickness and seven for positional relationships – for evaluating anatomical consistency between the skull and the face. Jayaprakash *et al.* (2001) listed the criteria for assessing the fit of the skull with the facial photograph during superimposition. A standardised protocol of the number of points or specific areas to be

Table 19.1 Criteria for assessing anatomical consistency between the skull and the face in frontal view.

Outline		Soft-tissue thickness		Positional relationships	
Skull	**Face**	**Skull**	**Face**	**Skull**	**Face**
Temporal line	Forehead	Zygion	Zygion	Supraorbital margin	Eyebrow
Lateral line of zygomatic bone	Cheek outline	Gonion	Gonion	Medial orbital margin	Endocanthion
Mandibular line	Lower jaw outline	Gnathion	Gnathion	Lateral orbital margin (Whitnall's malar tubercle)	Ectocanthion
				Orbit	Eye-slit
				Lateral margin of piriform aperture	Alare
				Cutting edge of upper central incisor	Stomion
				Teeth (premolar)	Cheilion
				Occlusal line	Oral-slit

Table 19.2 Criteria for assessing anatomical consistency between the skull and the face in lateral or oblique view.

Outline		Soft-tissue thickness		Positional relationships	
Skull	**Face**	**Skull**	**Face**	**Skull**	**Face**
Frontal bone contour	Forehead outline	Trichion	Trichion	Supraorbital margin	Eyebrow
Outline from nasion to rhinion	Nasal dorsum line	Glabella	Glabella	Lateral orbital margin (Whitnall's malar tubercle level)	Ectocanthion
		Nasion	Nasion		
Mental outline	Chin outline	Rhinion	Rhinion	Nasion	Higher than nasal root
Gonial outline	Jaw angle outline	Slight inferior to nasospinale	Subnasale	Lateral margin of piriform aperture	Alare
		Pogonion	Pogonion	Lower margin of piriform aperture	Subnasale
		Gnathion	Gnathion		
				Incisor	Stomion
				Teeth (canine, premolar)	Cheilion

compared in superimposed images has not been adopted. Table 19.1 and Table 19.2 show criteria generally used in actual cases. The temporal and occipital regions, including the vertex and opisthocranion, cannot be examined in most cases because these regions are commonly covered with hair.

By assessing the concordance between identifiable landmarks and morphological features, judgements can be made about inclusion or exclusion of the skull with the facial photograph. The determination of identity in craniofacial superimposition comparison cannot be automatically made at present. In future, some appropriate algorithms may give an automated judgement to assist the forensic examiner to make the final expert opinion.

19.4 Reliability of craniofacial superimposition

It is generally accepted that the unknown skull without the mandible cannot be positively identified as the presumed person, even if a good match is seen in the craniofacial superimposition image (Koelmeyer, 1982). Koelmeyer insisted that the complete skull is required to obtain any degree of accuracy.

Reddy (1973) noted that craniofacial superimposition is of a more negative value because it can be definitely stated that the skull and the facial photograph are not those of the same person, but if they tally, it can only be stated that the skull could be that of the person in the photograph, because of the possibility that another skull of that size and contour may tally with the photograph. In general, it is considered that craniofacial superimposition is more valuable for exclusion purposes. Dorion (1983) also stated that the craniofacial superimposition should not serve as the sole basis for positive identification.

Experimental studies have been carried out for evaluating the reliability of craniofacial superimposition. Helmer *et al.* (1989) quantified the individuality of human skulls by using craniometric measurements and their probability distribution on 52 European skulls. They concluded that skulls were comparable in their individuality to fingerprints. According to Ubelaker *et al.* (1992), credibility of comparison by craniofacial superimposition was enhanced through study of 52 similar skulls from collections of human skeletons of the Smithsonian Institute. In this experimental study, the unknown skull appeared to match the facial photograph, while the skulls of the four individuals in the collection that were found to be most similar to the unknown skull showed distinct differences when compared with the facial photograph. Austin-Smith and Maples (1994) compared three skulls of known identity to 97 lateral-view and 98 frontal-view photographs of non-related individuals, yielding a total of 585 superimpositions. Lateral-view and frontal-view superimpositions were identified incorrectly in 9.6% and 8.5% of the sample, respectively. The incidence of false matches was reduced to 0.6% of the sample when a frontal-view and lateral-view photograph of the same individual were both compared to one skull. These data led us to request multiple photographs from widely varying angles when it was necessary to prove or disprove identity by craniofacial superimposition (Austin-Smith and Maples, 1994; Austin, 1999). Yoshino *et al.* (1995) investigated the anatomical consistency of craniofacial superimposition images for evaluating the validity in personal identification using 52 skulls in forensic cases that were corroborated by other scientific and circumstantial evidence. This study suggested that the outline from the trichion to the gonion in the lateral or oblique view is the preferable portion for personal identification, and the craniofacial superimposition is reliable when two or more facial photographs taken from different angles are used in the examination.

When the anterior teeth can be observed in the facial photograph, they can be used as reliable data. The coincidence of dentition between the skull and facial photograph could lead to positive identification (McKenna *et al.*, 1984; Webster *et al.*, 1986; Solla and İşcan, 2001). As suggested by the forensic work done in the Howick Falls murder case (Prinsloo, 1953), pathological findings or congenital abnormalities in the skull and the facial photograph, in addition to points of consistency between the skull and the facial photograph, certify they belong to the same person.

Glassman (2001) gave a ranked score between Grades I and IV. The ranking is based on qualitative assessment of the number and closeness of anatomically matched areas. According to him, Grade I represents a close match with strong concordance in all anatomical areas available for comparison, and no area dictates exclusion. Grade II reflects a somewhat less convincing comparison and is described as a reasonable match with strong concordance in most anatomical areas and no area dictates exclusion. Grade III is used in those cases where the comparison cannot be used to definitively exclude a match, but is judged unlikely due to a number of anatomical areas that exhibit poor concordance. Grade IV is assigned when comparison of one or more areas indicates definite exclusion. From our practical examination (Yoshino *et al.*, 1995), it is stated that a skull could be positively identified as the presumed person if > 13 examination criteria are anatomically satisfied in the superimposition images. Fifteen out of 52 unknown skulls were examined using only the frontal view image of the missing person and exhibited < 12 matching criteria, giving a probable identification.

In order to obtain reliable results, the examiner should select a good-quality facial photograph that was taken in recent years. When craniofacial superimposition comparison is applied to identify juvenile skeletal remains, the most recent photograph of a presumed person should be used because the size and proportion of juvenile skulls constantly change throughout development (Yoshino *et al.*, 1989). To reduce an optical distortion of the facial image, a photograph showing the presumed person in a position near to the central area of it should be used.

Bilge *et al.* (2003) proposed that multidisciplinary techniques such as anthropology, odontology and DNA profiling should be employed to identify a decapitated human head. In our institute, mtDNA analysis has been used together with craniofacial superimposition for personal identification of unknown skulls without the mandible since 1999, and short tandem repeats (STR) typing has been applied since 2003 (Imaizumi *et al.*, 2005). Skeletal remains exposed to severe environmental conditions for long periods contain insufficient amount of nuclear DNA for analysis, but ample amounts of mtDNA. Therefore, mtDNA analysis is applicable to almost criminal cases. Unfortunately, the discriminatory power of mtDNA analysis is not so high because mtDNA is maternally inherited and is not unique to an individual (Holland and Parsons, 1999). Results obtained from mtDNA analysis should be used as supporting data for craniofacial superimposition comparison. According to Imaizumi *et al.* (2005), in some cases in which amplifiable nuclear DNA is extracted from bone or teeth, STR typing is a potentially powerful tool for identifying unknown skeletal remains. Combined use of craniofacial superimposition and STR typing would give a positive identification.

19.5 Conclusions

Craniofacial superimposition techniques have been developed, and its accuracy and reliability have been improved. Even so, examiners must understand the limitations of this technique. Craniofacial superimposition is generally considered to be effective for excluding the presumed person when comparing the unknown skull with the presumed person. The judgement of exclusion can be made based on one or more obvious anatomical discrepancies between the skull and the facial photograph. Thus, the possibility of false exclusion in this technique would be low. However, how many anatomical consistencies between the skull and the facial photograph are needed for positive identification? The author described that the skull could be positively identified as the presumed person when 13 or more examination criteria were anatomically matched in the superimposition images. However, this is only personal opinion induced from a small-sample examination. In order to avoid false inclusion, examiners must check as many as possible matching criteria using two or more facial photographs taken from different angles.

Recently, forensic scientists are gradually inclining to utilise DNA technology for identifying unknown skeletal remains. When the skull without the mandible or one facial photograph of frontal view is provided as evidential samples, combined use of the craniofacial superimposition technique and DNA analysis is recommended and would lead to more reliable identification.

References

Al-Amad, S., McCullough, M., Graham, J., Clement, J. and Hill, A. (2006). Craniofacial identification by computer-mediated superimposition. *Journal of Forensic Odonto-stomatology*, **24**, 47–52.

Aulsebrook, W. A., İşcan, M. Y., Slabbert, J. H. and Becher, P. (1995). Superimposition and reconstruction in forensic facial identification: a survey. *Forensic Science International*, **75**, 101–120.

Austin, D. (1999). Video superimposition at the C. A. Pound Laboratory 1987 to 1992. *Journal of Forensic Sciences*, **44**, 695–699.

Austin-Smith, D. and Maples, W. R. (1994). The reliability of skull/photograph superimposition in individual identification. *Journal of Forensic Sciences*, **39**, 446–455.

Bajnoczky, I. and Kiralyfalvi, L. (1995). A new approach to computer-aided comparison of skull and photograph. *International Journal of Legal Medicine*, **108**, 157–161.

Basauri, C. (1967). A body identified by forensic odontology and superimposed photographs. *International Criminal Police Review*, **205**, 37–43.

Bastiaan, R. J., Dalitz, G. K. and Woodward, C. (1986). Video superimposition of skulls and photographic portraits: a new aid to identification. *Journal of Forensic Sciences*, **31**, 1373–1379.

Bilge, Y., Kedici, P. S., Alakoc, Y. D., Ulkuer, K. U. and Ilkyaz, Y. Y. (2003). The identification of a dismembered human body: a multidisciplinary approach. *Forensic Science International*, **137**, 141–146.

Brocklebank, L. M., and Holmgren, C. J. (1989). Development of equipment for the standardization of skull photographs in personal identification by photographic superimposition. *Journal of Forensic Sciences*, **34**, 1214–1221.

Brown, K. A. (1982). The identification of Linda Agostini. *American Journal of Forensic Medicine and Pathology*, **3**, 131–141.

Brown, K. A. (1983). Developments in cranio-facial superimposition for identification. *Journal of Forensic Odonto-stomatology*, **1**, 57–64.

Cai, D. and Lan Y. (1993). Standards for skull-to-photo superimposition. In M. Y. İşcan and R. P. Helmer (Eds.), *Forensic Analysis of the Skull: Craniofacial Analysis, Reconstruction, and Identification*, pp. 171–181. New York, NY: Wiley-Liss.

Chai, D., Lan, Y., Tao, C. *et al.* (1989). A study on the standard for forensic anthropologic identification of the skull-image superimposition. *Journal of Forensic Sciences*, **34**, 1343–1356.

Delfino, V. P., Colonna, M., Vacca, E., Potente, F. and Introna, F. (1986). Computer-aided skull/face superimposition. *American Journal of Forensic Medicine and Pathology*, **7**, 201–212.

Delfino, V. P., Vacca, E., Potente, F., Lettini, T. and Colonna, M. (1993). Shape analysis morphometry in computer-aided skull identification via videosuperimposition. In M. Y. İşcan and R. P. Helmer (Eds.), *Forensic Analysis of the Skull: Craniofacial Analysis, Reconstruction, and Identification*, pp. 131–159. New York, NY: Wiley-Liss.

Dorion, R. B. J. (1983). Photographic superimposition. *Journal of Forensic Sciences*, **28**, 724–734.

Eliasova, H. and Krsek, P. (2007). Superimposition and projective transformation of 3D object. *Forensic Science International*, **167**, 146–153.

Fenton, T. W., Heard, A. N. and Sauer, N. J. (2008). Skull photo superimposition and border death: identification through exclusion and the failure to exclude. *Journal of Forensic Sciences*, **53**, 34–40.

Gejvall, N. G. (1974). Superimposition plus SEM comparison of hair cuticle for identification purposes. *OSSA: International Journal of Skeletal Research*, **1**, 99–103.

Ghosh, A. K., and Shinha, P. (2001). An economised craniofacial identification system. *Forensic Science International*, **117**, 109–119.

Ghosh, A. K., and Shinha, P. (2005). An unusual case of cranial image recognition. *Forensic Science International*, **148**, 93–100.

Glaister, J. and Brash, J. C. (1937). *The Medico-Legal Aspects of the Buck Ruxton Case*, pp. 144–170. Edinburgh: E. and S. Livingstone.

Glassman, D. M. (2001). Methods of superimposition. In Taylor, K. T. (Ed.), *Forensic Art and Illustration*, pp. 477–498. Boca Raton, FL: CRC Press.

Gordon, I. and Drennan, M. R. (1948). Medico-legal aspects of the Woldersdorfer case. *South African Medical Journal*, **22**, 543–549.

Gruner, O. (1993). Identification of skulls: a historical review and practical applications. In M. Y. İşcan and R. P. Helmer (Eds.), *Forensic Analysis of the Skull: Craniofacial Analysis, Reconstruction, and Identification*, pp. 29–45. New York, NY: Wiley-Liss.

Gruner, O. and Reinhard, R. (1959). Ein photografisches Verfahren zur Schadelidentifizierung. *Deutsche Zeitshrift fur Gesamte Gerichtliche Medizin*, **47**, 247–256.

Gruner, O. and Schulz, G. (1969). Uber eine Vereinfachung der photografischen Schadelidentifizierung. *Beitrage zur Gerichtliche Medizin*, **26**, 132–137.

Helmer, R. P. (1987). Identification of the cadaver remains of Josef Mengele. *Journal of Forensic Sciences*, **32**, 1622–1644.

Helmer, R. P., and Gruner, O. (1977). Vereinfachte Schadelidentifizierung nach dem Superprojektionsverfahren mit Hilfe einer Video-Anlage. *Zeitschrift fur Rechtsmedizin*, **80**, 183–187.

Helmer, R. P., Schimmler, J. B. and Rieger, J. (1989). Zun Beweiswert der Schadelidentifizierung mit Hilfe der Video-Bildmischtechnik unter Berucksichtigung der Kraniometrischen Individualitat menschlicher Schadel. *Zeitschrift fur Rechtsmedizin*, **102**, 451–459.

Holland, M. M., and Parsons, T. J. (1999). Mitochondrial DNA sequence analysis: validation and use for forensic casework. *Forensic Science Review*, **11**, 21–50.

Ibanez, O., Ballerini, L., Cordon, O., Damas, S. and Santamaria, J. (2009). An experimental study on the applicability of evolutionary algorithms to craniofacial superimposition in forensic identification. *Information Science*, **179**, 3998–4028.

Ibanez, O., Cordon, O., Damas, S. and Santamaria, J. (2008). Craniofacial superimposition based on genetic algorithms and fuzzy location of cephalometric landmarks. *Hybrid Artificial Intelligence System, LNAI*, **5271**, 599–607.

Imaizumi, K., Noguchi, K., Shiraishi, T. *et al.* (2005). DNA typing of bone specimens: the potential use of the profiler test as a tool for bone identification. *Legal Medicine*, **7**, 31–41.

Iten, P. X. (1987). Identification of skulls by video superimposition. *Journal of Forensic Sciences*, **32**, 173–188.

Janssens, P., Hansch, C. and Voorhamme, L. (1978). Identity determination by superimposition with anthropological cranium adjustment. *OSSA: International Journal of Skeletal Research*, **5**, 109–122.

Jayaprakash, P. T., Srinivasan, G. J. and Amravaneswaran, M. G. (2001). Cranio-facial morphoanalysis: a new method for enhancing reliability while identifying skulls by photo superimposition. *Forensic Science International*, **117**, 121–143.

Koelmeyer, T. D. (1982). Videocamera superimposition and facial reconstruction as an aid to identification. *American Journal of Forensic Medicine and Pathology*, **3**, 45–48.

Krogman, W. M. and İşcan, M. Y. (1986). Restoration of physiognomy. In *The Human Skeleton in Forensic Medicine*, 2nd edn, pp. 413–457. Springfield, IL: Charles C. Thomas.

Kumari, T. R. and Sekharan, P. C. (1992). Remote control skull positioning device for superimposition studies. *Forensic Science International*, **54**, 127–133.

Lan, Y. (1992). Development and current status of skull-image superimposition: methodology and instrumentation. *Forensic Science Review*, **4**, 125–136.

Lan, Y. and Cai, D. (1993). Technical advances in skull-to-photo superimposition. In M. Y. İşcan and R. P. Helmer (Eds.), *Forensic Analysis of the Skull: Craniofacial Analysis, Reconstruction, and Identification*, pp. 119–129. New York, NY: Wiley-Liss.

Loh, F. C. and Chao T. C. (1989). Skull and photographic superimposition: a new approach using a second party's interpupil distance to extrapolate the magnification factor. *Journal of Forensic Sciences*, **34**, 708–713.

Maat, G. J. R. (1989). The positioning and magnification of faces and skulls for photographic superimposition. *Forensic Science International*, **41**, 225–235.

McKenna, J. J. I., Jablonski, N. G. and Fearnhead, R. W. (1984). A method of matching skulls with photographic portraits using landmarks and measurements of dentition. *Journal of Forensic Sciences*, **29**, 787–797.

McKenna, J. J. I. (1988). A method of orientation of skull and camera for use in forensic photographic investigation. *Journal of Forensic Sciences*, **33**, 751–755.

Nickerson, B. A., Fitzhorn, P. A., Koch, S. K. and Charney M. (1991). A methodology for near-optimal computational superimposition of two-dimensional digital facial photographs and three-dimensional cranial surface meshes. *Journal of Forensic Sciences*, **36**, 480–500.

Prinsloo, I. (1953). The identification of skeletal remains. *Journal of Forensic Medicine*, **1**, 11–17.

Reddy, K. S. N. (1973). Identification of dismembered parts: the medicolegal aspects of the Nagaraju case. *Forensic Science*, **2**, 351–374.

Ricci, A., Marella, G. L. and Apostol, M. A. (2006). A new experimental approach to computer-aided face/skull identification in forensic anthropology. *American Journal of Forensic Medicine and Pathology*, **27**, 46–49.

Sekharan, P. C. (1971). A revised superimposition technique for identification of the individual from the skull and photograph. *Journal of Criminal Law, Criminology and Police Science*, **62**, 107–113.

Sekharan, P. C. (1993). Positioning the skull for superimposition. In M. Y. İşcan and R. P. Helmer (Eds.), *Forensic Analysis of the Skull: Craniofacial Analysis, Reconstruction, and Identification*, pp. 105–118. New York, NY: Wiley-Liss.

Sen, N. K. (1962). Identification by superimposed photographs. *International Criminal Police Review*, **162**, 284–286.

Seta, S. and Yoshino, M. (1993). A combined apparatus for photographic and video superimposition. In M. Y. İşcan and R. P. Helmer (Eds.), *Forensic Analysis of the Skull: Craniofacial Analysis, Reconstruction, and Identification*, pp. 161–169. New York, NY: Wiley-Liss.

Shahrom, A. W., Vanezis, P., Chapman, R. C. *et al*. (1996). Techniques in facial identification: computer-aided facial reconstruction using a laser scanner and video superimposition. *International Journal of Legal Medicine*, **108**, 194–200.

Simpson, K. (1943). Rex versus Dobkin: The Baptist church cellar murder. *Medico-Legal Review*, **11**, 132–145.

Sivaram, S. and Wadhera C. K. (1977). Identity from skeleton: a case study. *International Criminal Police Review*, **308**, 158–160.

Solla, H. E., and İşcan, M. Y. (2001). Skeletal remains of Dr. Eugenio Antonio Berrios Sagredo. *Forensic Science International*, **116**, 201–211.

Stewart, T. D. (1979). Reconstruction of facial soft parts. In *Essentials of Forensic Anthropology*, pp. 255–274. Springfield, IL: Charles C. Thomas.

Stewart, T. D. (1983). The points of attachment of the palpebral ligaments: their use in facial reconstructions on the skull. *Journal of Forensic Sciences*, **28**, 858–863.

Taylor, J. A., and Brown, K. A. (1998). Superimposition techniques. In J. G. Clement and D. L. Ranson (Eds.), *Craniofacial Identification in Forensic Medicine*, pp. 151–164. London: Arnold.

Teixeira, W. R. G. (1985). The Mengele report. *American Journal of Forensic Medicine and Pathology*, **5**, 279–283.

Ubelaker, D. H., Bubniak, E. and O'Donnell, G. (1992). Computer-assisted photographic superimposition. *Journal of Forensic Sciences*, **37**, 750–762.

Vogel, G. (1968). Zur Identifizierung unbekannter Toter: Unter besonderer Berucksichtigung von Zahnstatus, Rontgen-Vergleichsuntersuchung und Fotografischer Identifizierungsmethode. *Kriminalistik*, **22**, 630–634.

Webster, W. P., Murray, W. K., Brinkhous, W. and Hudson, P. (1986). Identification of human remains using photographic reconstruction. In *Forensic Osteology*, pp. 256–289. Springfield, IL: Charles C. Thomas.

Yoshino, M., Imaizumi, K., Miyasaka, S. and Seta, S. (1995). Evaluation of anatomical consistency in cranio-facial superimposition images. *Forensic Science International*, **74**, 125–134.

Yoshino, M., Matsuda, H., Kubota, S. *et al.* (1997). Computer-assisted skull identification system using video superimposition. *Forensic Science International*, **90**, 231–244.

Yoshino, M., Miyasaka, S., Sato, H., Miyake, B. and Seta, S. (1989). Identification of infant skeletal remains: case report. *Japanese Journal of Legal Medicine*, **43**, 511–516.

Yoshino, M. and Seta, S. (1989). Personal identification of the human skull: superimposition and radiographic techniques. *Forensic Science Review*, **1**, 23–42.

Yoshino, M. and Seta, S. (2000). Skull–photo superimposition. In *Encyclopedia of Forensic Sciences*, *Vol. 2*, pp. 807–815. London: Academic Press.

Chapter 20

Juvenile facial reconstruction

Caroline Wilkinson

20.1 Introduction

The reconstruction of the face of a child is different from adult facial reconstruction. There are many difficulties associated with juvenile remains, including less accurate sex and ancestry assignment, the more emotive and sensitive nature of an investigation into the death of an unknown child, and the less-defined skeletal details associated with underdeveloped skulls. There also may be advantages associated with juvenile remains, such as increased public awareness, increased media attention and more accurate age estimation. Historically this subject has not been separated from adult facial reconstruction, although the differences between adult and juvenile skulls are significant.

20.2 Facial growth

The development of the skull throughout childhood produces extreme changes in facial appearance and these may be so drastic that an infant might become unrecognisable after only a few months (Y'Edynak and Işcan, 1993). There are three principal regions of craniofacial development: the brain and basicranium, the facial and pharyngeal airway, and the oral complex. Each of these regions has its own timetable of development, but all are inseparably linked as an interrelated whole (Enlow and Hans, 1996).

The infant brain and basicranium are precocious in development relative to the craniofacial complex and body, so that the head constitutes one-quarter of total body length, with the proportionate head size decreasing into adulthood, to closer one-eighth of total body length. The size of the neonatal cranium is limited by the maternal pelvis, and the face is as small as possible to allow effective birth passage. The dental and nasal regions are therefore late developing, causing the ramus of the mandible to be short, leading to a vertically short and wide (brachycephalic) facial appearance (Enlow, 1982). Once born the infant will need to dramatically increase body size, and related lung enlargement provides oxygen to the increased body volume, whilst dental development enables more efficient food intake. These changes in turn lead to nasal passage augmentation, through an increase in nose size, to enhance respiration and mandible reconfiguration to provide the skeletal foundation for tooth eruption. The face and body grow rapidly into adolescence, whilst the cranium shows limited further maturation. In effect the rest of the body catches up with the early development of the brain. The basicranium prescribes a template that establishes the growth fields within which both the mandible and nasomaxillary complex develop. Structures grow proportionally more and for a longer period of time the further they are from the neurocranium; therefore the growth of the mandible begins later and continues for longer than the midfacial or orbital development (Enlow and Hans, 1996).

The infant head has a bulbous forehead, prominent cheekbones and a flat face with large, wide-set, bulging eyes, a dainty jaw, puffy cheeks, low-placed ears and a small mouth.

It is necessary for a newborn baby to suckle in order to sustain life; therefore, the neonatal nose will not be prominent and will exhibit an up-turned columella, low nasal bridge and concave profile. This nasal shape allows the infant to breathe whilst suckling and prevents suffocation at the mother's breast. Therefore infant nasal shape is unrelated to the eventual adult nasal appearance, and even people with the largest, most prominent noses will have had small, uncharacteristic noses as infants. At 1 year of age the nasal dimensions are only 50–73% of their adult size, and to achieve maturity nasal tip protrusion has to increase by 96%, nasal tip inclination decrease by

Craniofacial Identification, ed. C. Wilkinson and C. Rynn. Published by Cambridge University Press.

9.8 degrees and columella inclination decrease by 17.8 degrees (Bjork, 1947; Tanner, 1952; Farkas and Hreczko, 1994; Feik and Glover, 1998). As the child matures into adolescence the nose will develop and at approximately 8 years of age the nose should be recognisable in its adult morphology and proportionate size. Farkas and Posnick (1992) state that full nasal maturity occurs at age 10 years in girls, and 14 years in boys; with the adolescent growth spurt leading to faster nose growth in boys, creating larger, more prominent noses in men than women.

The superior to inferior orbital margin line of an infant will exhibit vertical alignment or posterior inclination due to the diminutive nasal and oral regions and precocious cranial development. As the nasal and oral regions develop into adolescence this line will become more anteriorly inclined and the forehead will become less bulbous and upright. Anthropometry studies (Bjork, 1947; Tanner, 1952; Farkas and Hreczko, 1994; Feik and Glover, 1998) found that forehead inclination decreased by 13.7 degrees and nasofrontal angle increased by 7.9 degrees with an age increase.

Infant cheeks have large buccal fat pads to protect the delicate facial bones and rapidly developing nasal and oral regions; this produces a smooth-skinned, round-cheeked appearance to the infant face. As the child matures the buccal fat pads recede producing a leaner, more angular adult face. The buccal fat is more persistent in girls than boys, and tissue depth studies show plumper cheeks in girls than boys well into adolescence (Manhein *et al.*, 2000; Wilkinson, 2002).

The eyes of a full-term newborn are approximately 65–75% of adult size (Whitnall, 1932; Bjork, 1947; Tanner, 1952; Farkas and Hreczko, 1994; Feik and Glover, 1998); they have colour vision and very myopic visual acuity (20/400). The entire eyeball increases a little less than three-fold between birth and maturity (Scammon and Wilmer, 1950) and this precocious development may be due to the importance of sight for communication and social interaction between humans. It is established in psychology circles that facial perception and the interpretation of facial expression are vital skills for the development and survival of humans, with particular importance in the processing of information relating to eyes and eye-gaze direction (Farroni *et al.*, 2002). A neonate can see a parent's face when being held, and will often gaze intently at that face. Studies showing preferential attention to perceived faces with direct gaze provide compelling evidence that the human neonate is born prepared to detect socially relevant information (Baron-Cohen, 1995). Since the eyes are mature at a very young age, they will become progressively smaller relative to the face as the child grows. The eyes will also appear to become closer together with increased age, even though the eyeballs are actually similarly spaced in adults and children; it is the enlargement of other facial features and lower face proportions that produces this phenomenon. Anthropometric studies show that intercanthal and interocular distances increase minimally (7–24%) during childhood (Bjork, 1947; Tanner, 1952; Farkas and Hreczko, 1994; Feik and Glover, 1998) in comparison with other features. Ronneburger *et al.* (2006) measured the iris and cornea diameter of children between 1 month of age and adolescence, and also carried out a longitudinal study of 13 individuals from infancy to adolescence. They found a small amount of growth, 0.3 mm, over a mean age difference of 8.3 years, and concluded that, after birth, the fastest growth of the cornea and iris occurs during the first few months of life.

Infant ears appear to be low-placed and, as the face elongates, they appear to rise. In fact the ears move downward during development, but the face enlarges even further downwards, so that the relative position of the ears rises. The facial feature dimension that reaches maturity earliest is ear width; with females at 6 years and males at 7 years (Bjork, 1947; Tanner, 1952; Farkas and Hreczko, 1994; Feik and Glover, 1998).

The emergence of the teeth leads to enlargement and reconfiguration of the mandible and maxillae, with squaring of the chin and jaw, enlargement of the ramus of the mandible, expansion of the masticatory muscles and the flaring of the gonial angles, so that the whole lower face takes on a more U-shaped appearance. Farkas and Posnick (1992) stated that the adolescent growth spurt leads to a faster mandible growth in boys than girls, producing larger, more projecting jaws in men than women. In both sexes to achieve adult size mandible height has to increase by 50% and mouth width by 52% of size at 1 year of age. Anthropometry studies (Bjork, 1947; Tanner, 1952; Farkas and Hreczko, 1994; Feik and Glover, 1998) showed that the intergonial distance reaches its mature size between ages 5 and 15 years, and upper lip incline decreases by 11.9 degrees and lower lip inclination increases by 12.6 degrees into adulthood. Other 3D studies of facial growth (Nute and Moss, 2000) showed that lower face width increases by approximately 1–3 mm a year, and the intergonial width increases by approximately 3–5 mm a year between the ages of 5 and 10 years. They also found that the chin prominence varied little until after the age of 8 years.

Farkas and Posnick (1992) studied facial growth in American children, and found that the face reaches maturity by about 10 to 13 years in girls, and 12 to 15 years in boys. Farkas and Posnick (1992) stated that full maturity in head length, breadth and circumference occurs at age 10 years in girls, and 14 years in boys.

20.3 Facial recognition of children

Since an infant's face is immature and the features underdeveloped and diminutive, infants' faces will resemble each other more than they resemble adult faces. This suggests that unfamiliar juvenile faces will be more difficult to distinguish and recognise than adult faces. It is well established that we have more difficulty determining sex from juvenile faces than adults (Bruce *et al.*, 1993; Wild *et al.*, 2000), but unfamiliar recognition studies have mostly focused on adult faces. Some studies show that children are better than adults at recognising facial images of other children (Anastasi and Rhodes, 2005), although worse than adults at recognising facial images of adults. These studies, along with others relating primarily to adults (Bäckman, 1991; Fulton and Bartlett, 1991; Perfect and Harris, 2003), suggest that there is an own-age bias to facial recognition, similar to the well-established own-race bias (see Meissner and Brigham, 2001, for a review). Other studies confirmed an own-age bias for adults, but did not show a significant own-age bias for children (Chung, 1997). However, Wild *et al.* (2000) found that faces of children and adults were equally recognisable (by both children and adults).

There is some evidence to suggest that recognition of familiar juvenile faces is more reliable than unfamiliar juvenile face recognition (Porter *et al.*, 1984; Kaitz *et al.*, 1988). Porter *et al.* (1984) found that within 33 hours post-partum, mothers recognised photographs of their own offspring when presented with those of unrelated neonates. Furthermore, adult subjects were able to match photographs of unfamiliar mothers and their infants, and establish the sex of neonates, at a greater than chance level of accuracy. However, there are no published studies relating to the recognition of familiar children's faces by non-parents.

The recognition of missing children does not have a high success rate. Abducted children are not frequently recognised by the general public, even when intense media attention is given to the investigation (MST, 2007). The NCMEC (National Center for Missing and Exploited Children), who release thousands of images of missing children a week in the USA, reports that only one in six missing children (17%) is recovered because someone viewed a child's picture and called authorities (NCMEC, 2004).

These research results and forensic statistics have significant implications for the recognition of reconstructions of juvenile faces. We do not know whether facial depictions of children are likely to be recognised as readily as those of adults and the role of juvenile facial reconstruction in forensic investigation is uncertain. However, forensic investigations utilising juvenile facial reconstruction report a good level of success: juvenile depictions show a higher success rate than those of adults (100% success in juvenile cases compared with 65% in adult cases; Wilkinson, 2010), but the number of cases is much smaller than those of adults (less than 10% of adult cases). It is also unclear whether success is due to the increased media attention and heightened emotional response of the public to unidentified juvenile remains, rather than any resemblance to the target child.

20.4 Adaptation of reconstruction standards for juvenile skulls

Whilst juvenile facial-tissue research is common due to the introduction of ultrasonic measurement, the majority of facial anthropology research relating to the prediction of facial features from skeletal analysis has been restricted to adults. This is, in part, due to the ethical considerations relating to clinical images of children (SCoR, 2005, 2008) and data protection issues restricting access to clinical juvenile data (ICO, 2006). The health risks involved in CT and radiograph collection limit the use for children (SCoR, 2005), and the expense of MRI acquisition and problems in the visualisation of bone in MR images has excluded its widespread use in this field. Even where clinical images are necessary for medical diagnosis and treatment, acquisition will be limited in relation to craniofacial exposure (Feigal, 2001). This, in turn, limits research relating to juvenile craniofacial analysis.

Therefore it is common for adult facial prediction standards to be employed in juvenile facial depiction. Practitioners have suggested variations to some facial prediction standards, where there is anatomical evidence for juvenile differences, such as eyeball size

(Scammon and Armstrong, 1925; Whitnall, 1932), head position (Gerasimov, 1955; Wilkinson, 2004), buccal fat size and position (Xiao *et al.*, 1999; Zhang *et al.*, 2002) and ear pattern (Adamson *et al.*, 1965; Niemitz *et al.*, 2007).

20.4.1 Eyeball size

Weiss (1897) recorded a mean neonatal eyeball diameter of 16 mm, which is 67% of mean adult diameter (24 mm). Other studies have published tables relating to juvenile eyeball size (Krause, 1832; Engel, 1850; Scammon and Armstrong, 1925) and it is established that the adult eyeball size is achieved by 8 years of age with the fastest growth in the first year of life. Therefore juvenile (2–8 years) reconstructions should use a mean diameter of 22 mm.

20.4.2 Head position

Adult head position is described by the Frankfurt Horizontal Plane (FHP), and has been used since 1884 as one of the acceptable ways to approximate the physiological horizontal (Moorrees and Kean, 1958). However, Madsen *et al.* (2008) studied 12–18-year-olds and suggested that the palatal plane or Krogman–Walker (KW) plane are more accurate representations of natural head position. The KW line is described as occipitale to maxillon (Rothstein and Yoon-Tarlie, 2000). In addition, since adults tend to view the faces of children from above, it is considered preferable to tilt the skull of a child so that the FHP is higher anteriorly than posteriorly.

20.4.3 Buccal fat size and position

It is well recognised that children have proportionally larger buccal fat pads than adults (Xiao *et al.*, 1999; Zhang *et al.*, 2002). Zhang and colleagues (2002) dissected three juvenile and eight adult cadavers, and found that the anterior and intermediate lobes are larger and richer in fat, the posterior lobe is smaller and the pterygoid and buccal processes are larger in children than adults. Therefore the buccal fat pad should be included in the facial reconstruction of juvenile skulls as this will have a significant effect upon the facial proportions and appearance of a child.

20.4.4 Ear pattern

It is established that children have smaller ears than adults, with a 3-year-old having ears 85% of adult size (Adamson *et al.*, 1965; Carsten *et al.*, 2007) and ear height continuing to grow into late adulthood. It is also established that adult ear prominence and width are achieved by 10 years of age (Adamson *et al.*, 1965; Carsten *et al.*, 2007). Lin and Furnas (2009) recorded the mean length of the ear (superaurale to subaurale) in boys as 55 mm when aged 6 years, 60 mm when aged 12 years, and 62 mm when aged 18 years. In girls, the values were 54 mm, 58 mm and 58 mm at ages 6, 12 and 18 years, respectively. The mean width of the ear (preaurale to postaurale) was 34 mm, 35 mm and 36 mm in boys, and 33 mm, 34 mm and 34 mm in girls at ages 6, 12 and 18 years, respectively (Lin and Furnas, 2009). In addition, Da Silva Freitas *et al.* (2008) recorded an average cephaloauricular angle of 47.7 degrees. Prominent ears are known to be linked to prominent mastoid processes, which tend to push the concha forward, causing auricular prominence (Lin and Furnas, 2009). Pohl (2006) states that 15% of the auricle should be above the horizontal eye line (inner to outer canthus line) and the average angle between the vertical axis (the line perpendicular to the horizontal eye line) and the longitudinal axis of the ear (superior aspect of the outer helix to the inferior border of the earlobe) is between 10 and 30 degrees. In addition, Pohl (2006) states that the length of an ear can be roughly estimated by measuring the distance between the arch of the eyebrow and the base of the nasal alae.

20.4.5 Mouth shape

The morphology and pattern of the mouth and lips is estimated using orthodontic and anthropology standards. There have been a number of recent studies that have improved mouth prediction (Stephan and Henneberg, 2003; Wilkinson *et al.*, 2003; Stephan and Murphy, 2008), but all these studies involved adult subjects, often elderly adults, and these standards have never been tested on children. In addition, children often display mixed dentition (deciduous and permanent) and since it is well established that dental eruption and pattern have a direct effect upon lip shape (Subtelny, 1959; Rudee, 1964; Roos, 1977; Koch *et al.*, 1979; Waldman, 1982; Holdaway, 1983; Denis and Speidel, 1987; Talass *et al.*, 1987) we cannot assume that the adult standards will be applicable to juvenile subjects. Certainly the standards relating tooth enamel height to lip thickness (Wilkinson *et al.*, 2003) cannot be applied to children where partial enamel exposure, dental eruption and tooth loss are present. It is established that dental, eye and lip development follow different schedules (Bjork, 1947; Tanner, 1952; Farkas and Hreczko, 1994; Enlow and Hans, 1996; Feik and Glover, 1998), so the standards relating intercanine width (Stephan and Henneberg, 2003) and infraorbital foramina position (Stephan and Murphy, 2008) to the cheilion points are not applicable. On the other hand there are no published juvenile standards, and it is clear that a great deal more research is required into the prediction of juvenile lip shape from dental pattern. In practice, mouth shape can be estimated using the adult standards as *cautious guides*.

20.4.6 Juvenile tissue depths

In 2000, Manhein and her colleagues studied 515 White, Black and Hispanic North American children, between the ages of 3 and 18 years, using ultrasonic imaging, to create the first useful database of children's facial tissue depths. Wilkinson (2002) followed with a study of 200 White British children using ultrasound measurement, and since then there have been a small number of other published juvenile databases on Japanese (Utsuno *et al.*, 2005, 2007, 2010) and North American (Williamson *et al.*, 2002) children.

References

Adamson, J. K., Rome E., Hortox, C. E. and Crawford, H. H. (1965). The growth pattern of the external ear. *Plastic and Reconstructive Surgery*, **36**(4), 466–470.

Anastasi, J. S. and Rhodes, M. G. (2005). An own-age bias in face recognition for children and older adults. *Psychonomic Bulletin and Review*, **12**, 1043–1047.

Bäckman, L. (1991). Recognition memory across the adult life span: the role of prior knowledge. *Memory and Cognition*, **19**, 63–71.

Baron-Cohen, S. (1995) The eye direction detector (EDD) and the shared attention mechanism (SAM): two cases for evolutionary psychology. In C. Moore and P. J. Dunham (Eds.), *Joint Attention: Its Origins and Role in Development*, pp. 41–59. Mahwah, NJ: Lawrence Erlbaum.

Bjork, A. (1947). The face in profile: an anthropometrical x-ray investigation on Swedish children and conscripts. *Svensk Tandlak-T*, **49** (Suppl.).

Bruce, V., Burton, A. M. and Hanna, E. (1993). Sex discrimination: how do we tell the difference between male and female faces? *Perception*, **22**, 131–152.

Carsten, N., Nibbrig, M. and Zacher, V. (2007). Human ears continue to grow throughout the entire lifetime according to complicated and sexually dimorphic patterns; conclusions from a cross-sectional analysis. *Anthropologischer Anzeliger*, **65** (4), 391–413.

Chung, M. S. (1997). Face recognition: effects of age of subjects and age of stimulus faces. *Korean Journal of Developmental Psychology*, **10**, 167–176.

Da Silva Freitas, R., Sanchez, M. E., Manzotti, M. S. *et al.* (2008). Comparing cephaloauricular and scaphaconchal angles in prominent ear patients and control subjects. *Aesthetic Plastic Surgery*, **32**(4): 620–623.

Denis, L. K. and Speidel, T. M. (1987). Comparison of three methods of profile change prediction in the adult orthodontic patient. *American Journal of Orthodontics and Dentofacial Orthopedics*, **92**, 396–402.

Engel, J. (1850). Ein Beitrag ZUT Pliysik des Auges. *Vierteljahreschr. f. prakt.* Heilk., Bd. 7, S. 167–208.

Enlow, D. H. (1982). *A Handbook of Facial Growth*, 2nd edn. Philadelphia, PA: W. B. Saunders.

Enlow, D. H. and Hans, M. G. (1996). *Essentials of Facial Growth*. Philadelphia, PA: WB Saunders.

Farkas, L. G and Hreczko, T. A. (1994). Age-related changes in selected linear and angular measurements of the craniofacial complex in healthy North American Caucasians. In L. G. Farkas (Ed.), *Anthropometry of the Head and Face*, 2nd edn. New York, NY: Raven Press.

Farkas, L. G. and Posnick, J. C. (1992). Growth and development of the head. *Cleft-Palate Craniofacial Journal*, **29**, 301–329.

Farroni, T., Csibra, G., Simion, F. and Johnson, M. H. (2002). Eye contact detection in humans from birth. *Proceedings of the National Academy of Sciences USA*, **99**(14), 9602–9605.

Feigal, D. W. (2001). *FDA Public Health Notification: Reducing Radiation Risk from Computed Tomography for Pediatric and Small Adult Patients*. CDRH/FDA. Available at: www.fda.gov/cdrh/safety/110201-ct.html2.

Feik, S. A. and Glover, J. E. (1998). Growth of children's faces. In J. G. Clement and D. L. Ranson (Eds.), *Craniofacial Identification in Forensic Medicine*. Sydney: Arnold Publishers.

Fulton, A. and Bartlett, J. C. (1991). Young and old faces in young and old heads: the factor of age in face recognition. *Psychology and Aging*, **6**, 623–630.

Gerasimov, M. M. (1955). *The Reconstruction of the Face from the Basic Structure of the Skull*. Translated by W. Tshernezky. Russia: Publishers unknown.

Holdaway, R. A. (1983). A soft tissue cephalometric analysis and its use in orthodontic treatment planning. Part I. *American Journal of Orthodontics*, **84**, 1–28.

ICO (2006). *Information Comissioner's Office – Protecting Children's Personal Information*. November 2006.

http://www.ico.gov.uk/Global/data_protection_research_archive.aspx

Kaitz, M., Good, A., Rokem, A. M. and Eidelman, A. (1988). Mothers' and fathers' recognition of their newborns' photographs during the postpartum period. *Journal of Developmental and Behavioral Pediatrics*, **9**, 223–226.

Koch, R., Gonzales, A. and Witt, E. (1979). Profile and soft tissue changes during and after orthodontic treatment. *European Journal of Orthodontics*, **1**, 193–199.

Krause, C. (1832). Einige Bemerkung iiber den Bau und die Dimensionen des menschlichen Auges. *Arch. f. Anat. u. Physiol., Bd.*, **6**, S. 86–120.

Lin, S. J. and Furnas, D. (2009). Ear, Prominent Ear. eMedicine Specialties; Plastic Surgery; Head/Neck. http://www.emedicine.medscape.com/article/1290275-overview [accessed 28 September 2010].

Madsen, D. P., Sampson, W. and Townsend, G. C. (2008). Craniofacial reference plane variation and natural head position. *European Journal of Orthodontics*, **30**(5), 532–540.

Manhein, M. H., Barsley, R. E., Listi, G. A. *et al.* (2000). In vivo facial tissue depth measurements for children and adults. *Journal of Forensic Science*, **45**(1), 48–60.

Meissner, C. A. and Brigham, J. C. (2001). Thirty years of investigating the own-race bias in memory for faces: a meta-analytic review. *Psychology, Public Policy, and Law*, **7**, 3–35.

Moorrees, C. F. A. and Kean, M. (1958). Natural head position, a basic consideration in the interpretation of cephalometric radiographs. *American Journal of Physical Anthropology*, **16**(2), 213–234.

MST (2007) What impact has media coverage of Madeleine McCann had? Media Standards Trust, 05/06/2007. http://www.mediastandardstrust.org/projects/debatestandards/debatedetails.aspx?sid=3863 [accessed 21 September 2010].

NCMEC (2004) National Centre for Missing and Exploited Children Press Release – 4 October 2004. http://www.missingkids.com/missingkids/servlet/NewsEventServlet?LanguageCountry=en_USandPageId=1750 [Accessed 21 September 2010].

Niemitz C., Nibbrig M. and Zacher V. (2007). Human ears grow throughout the entire lifetime according to complicated and sexually dimorphic patterns – conclusions from a cross-sectional analysis. *Anthropologischer Anzeiger*, **65**(4), 391–413.

Nute, S. J. and Moss, J. P. (2000). Three-dimensional facial growth studied by optical surface scanning. *Journal of Orthodontics*, **27**, 31–38.

Perfect, T. J. and Harris, L. J. (2003). Adult age differences in unconscious transference: source confusion or identity blending? *Memory and Cognition*, **31**, 570–580.

Pohl, C. A. (2006). Abnormalities in ear shape and position. In E. K. Chung, J. A. Boom, G. A. Datto and P. S. Matz (Eds.), *Visual Diagnosis in Pediatrics, 1st edn*. Philadelphia, PA: Lippincott Williams and Wilkins.

Porter, R. H., Cernoch, J. M. and Balogh, R. D. (1984). Recognition of neonates by facial-visual characteristics. *Pediatrics*, **74**(4), 501–504.

Ronneburger, A., Basarab, J. and Howland, H. C. (2006). Growth of the cornea from infancy to adolescence. *Ophthalmic and Physiological Optics*, **26** (1), 80–87.

Roos, N. (1977). Soft tissue profile changes in Class II treatment. *American Journal of Orthodontics*, **72**, 165–175.

Rothstein, T. and Yoon-Tarlie, C. (2000). Dental and facial skeletal characteristics and growth of males and females with Class II, division 1 malocclusion between the ages of 10 and 14 (revisited) – part I: characteristics of size, form, and position. *American Journal of Orthodontics and Dentofacial Orthopedics*, **117**, 320–332.

Rudee, D. A. (1964). Proportional profile changes concurrent with orthodontic therapy. *American Journal of Orthodontics*, **50**, 421–434.

Scammon, R. E. and Armstrong, E. L. (1925). On the growth of the human eyeball and optic nerve. *Journal of Comparative Neurology*, **38**(2), 165–219.

Scammon, R. E. and Wilmer, H. A. (1950). Growth of the components of the human eyeball: II. Comparison of the calculated volumes of the eyes of the newborn and of adults, and their components. *Archives of Ophthalmology*, **43**(4), 620–637.

SCoR (2008). *Society of Radiographers – Code of Conduct and Ethics*. Document Library. http://doc-lib.sor.org.

SCoR (2005). *Society of Radiographers – The Child and The Law: The Roles and Responsibilities of the Radiographer*. London: SCoR.

Stephan, C. N. and Henneberg, M. (2003). Predicting mouth width from inter-canine width – a 75% rule. *Journal of Forensic Science*, **48**, 725–727.

Stephan, C. N. and Murphy, S. J. (2008). Mouth width prediction in craniofacial identification: cadaver tests of four recent methods, including two techniques for edentulous skulls. *Journal of Forensic Odontostomatology*, **27**(1), 2–7.

Subtelny, J. D. (1959). A longitudinal study of soft tissue facial structures and their profile characteristics, defined in relation to underlying skeletal structures. *American Journal of Orthodontics and Dentofacial Orthopedics*, **45**(7), 481–507.

Talass, M. F., Talass, L. and Baker, R. C. (1987). Soft-tissue profile changes resulting from retraction of maxillary incisors. *American Journal of Orthodontics and Dentofacial Orthopedics*, **5**, 385–394.

Tanner, J. M. (1952). The assessment of growth and development in children. *Archives of Diseases in Childhood*, **27**, 10–33.

Utsuno, H., Kageyama, T., Deguchi, T. *et al.* (2005). Facial soft tissue thickness in Japanese female children. *Forensic Science International*, **152**(2–3), 101–107.

Utsuno, H., Kageyama, T., Deguchi, T. *et al.* (2007). Facial soft tissue thickness in skeletal type I Japanese children. *Forensic Science International*, **172**(2), 137–143.

Utsuno, H., Kageyama, T., Uchidac, K. *et al.* (2010). Facial soft tissue thickness in Japanese children. *Forensic Science International*, **199**(1), 109e1–109.e6.

Waldman, B. H. (1982). Change in lip contour with maxillary incisor retraction. *Angle Orthodontist*, **52**, 129–134.

Weiss, L. (1897). Uber das Wachstums des menschlichen Auges und uber die Verlnclerung der Muskelinsertionen am wachsenden Auge. *Anat. Hefte, Bd.*, **8**, S. 193–248.

Williamson, M. A., Nawrocki, S. P. and Rathburn, T. A. (2002). Variation in midfacial tissue thickness of African-American children. *Journal of Forensic Sciences*, **47**(1), 25–31.

Whitnall, S. E. (1932). The anatomy of the human orbit and accessory organs of vision. Oxford: Humphrey Milford/ Oxford University Press.

Wild, H. A., Barrett, S. E., Spence, M. J. *et al.* (2000). Recognition and sex categorization of adults' and children's faces: examining performance in the absence of sex-stereotyped cues. *Journal of Experimental Child Psychology*, **77**(4), 269–291.

Wilkinson, C. M. (2002). In vivo facial tissue depth measurements for White British children. *Journal of Forensic Science*, **47**(3), 459–465.

Wilkinson, C. M. (2004). *Forensic Facial Reconstruction*. Cambridge: Cambridge University Press.

Wilkinson, C. M. (2010). *University of Dundee statistics for facial reconstruction success in forensic investigation*. 28 September 2010. Available from author.

Wilkinson, C. M., Motwani, M. and Chiang, E. (2003). The relationship between the soft tissues and the skeletal detail of the mouth. *Journal of Forensic Sciences*, **48**, 728–732.

Xiao, H., Bayramicli, M. and Jackson, L. T. (1999). Volumetric analysis of the buccal fat pad. *European Journal of Plastic Surgery*, **22**, 177–181.

Y'Edynak, G. J. and Işcan, M. Y. (1993). Craniofacial evolution and growth. In M. Y. İşcan and R. P. Helmer (Eds.), *Forensic Analysis of the Skull: Craniofacial Analysis, Reconstruction, and Identification*, pp. 11–27. New York, NY: Wiley-Liss.

Zhang, H. M., Yan, Y. P., Qi, K. M., Wang, J. Q. and Liu, Z. F. (2002). Anatomical structure of the buccal fat pad and its clinical adaptions. *Plastic and Reconstructive Surgery*, **109**, 2509–2518.

Index

Figure 5.1 An example of landmarking a face image. A set of landmarks are specified by hand on an image. The original image is shown on the left and the landmarked image is in the centre. The image is then warped, to place the landmarks into the same positions as those of a reference shape, as shown on the right.

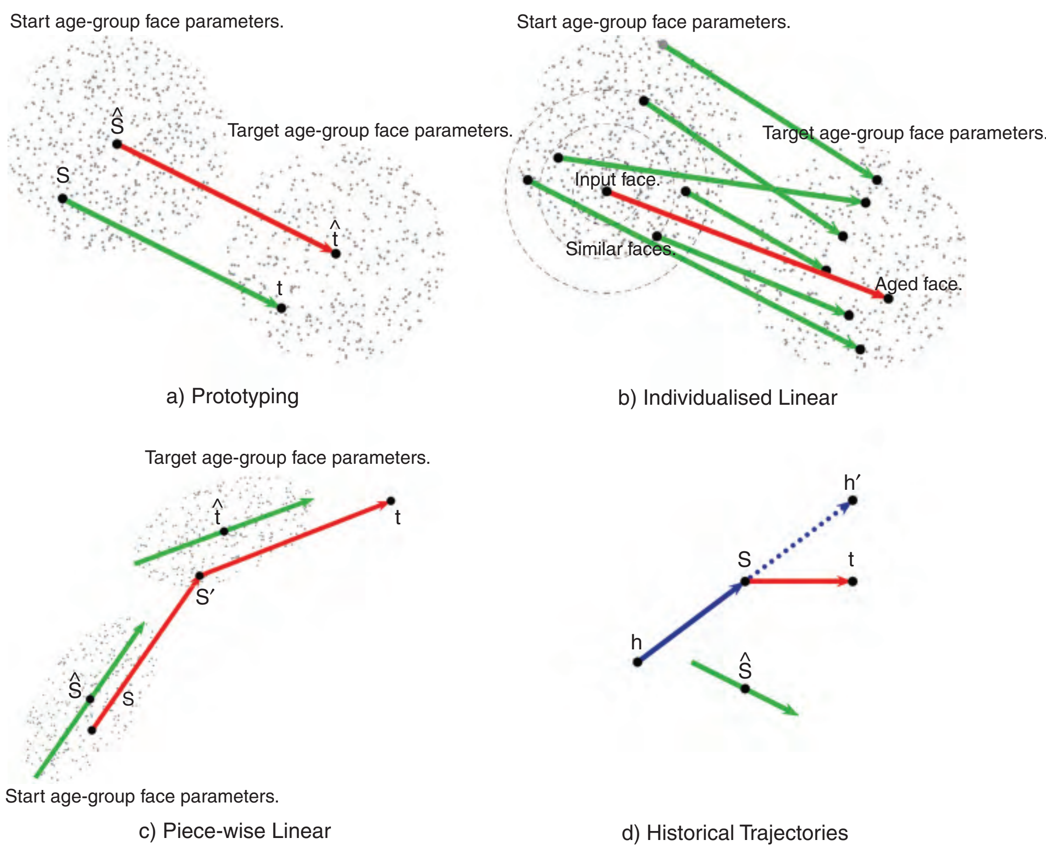


Figure 5.4 A comparison of two ageing functions. The Prototyping function, (a), describes a transform of a input face (**s**) to (**t**). **ŝ** and **t̂** represent the source and target Prototypes respectively. The individualised function (shown in its linear incarnation for simplicity) in (b), shows how a new ageing trajectory is created from the training set. Those transforms that originate 'near' the unseen face have the strongest influence on its final trajectory. Diagram (c) shows the piece-wise linear ageing used by Scandrett *et al.* (2006). A trajectory for each age group is calculated as an age-weighted average (shown in green). A face can be aged by stepping from group to group along the trajectories defined for each group. Diagram (d) shows how an historical ageing trajectory works (Scandrett *et al.*, 2006). Given two faces one (**s**) the starting point of the ageing, and one (**h**) of the same person at a younger age. The trajectory predicts a face at **h′**. This trajectory is modified by the trajectory from an overall age group **ŝ** to create a target at **t**.

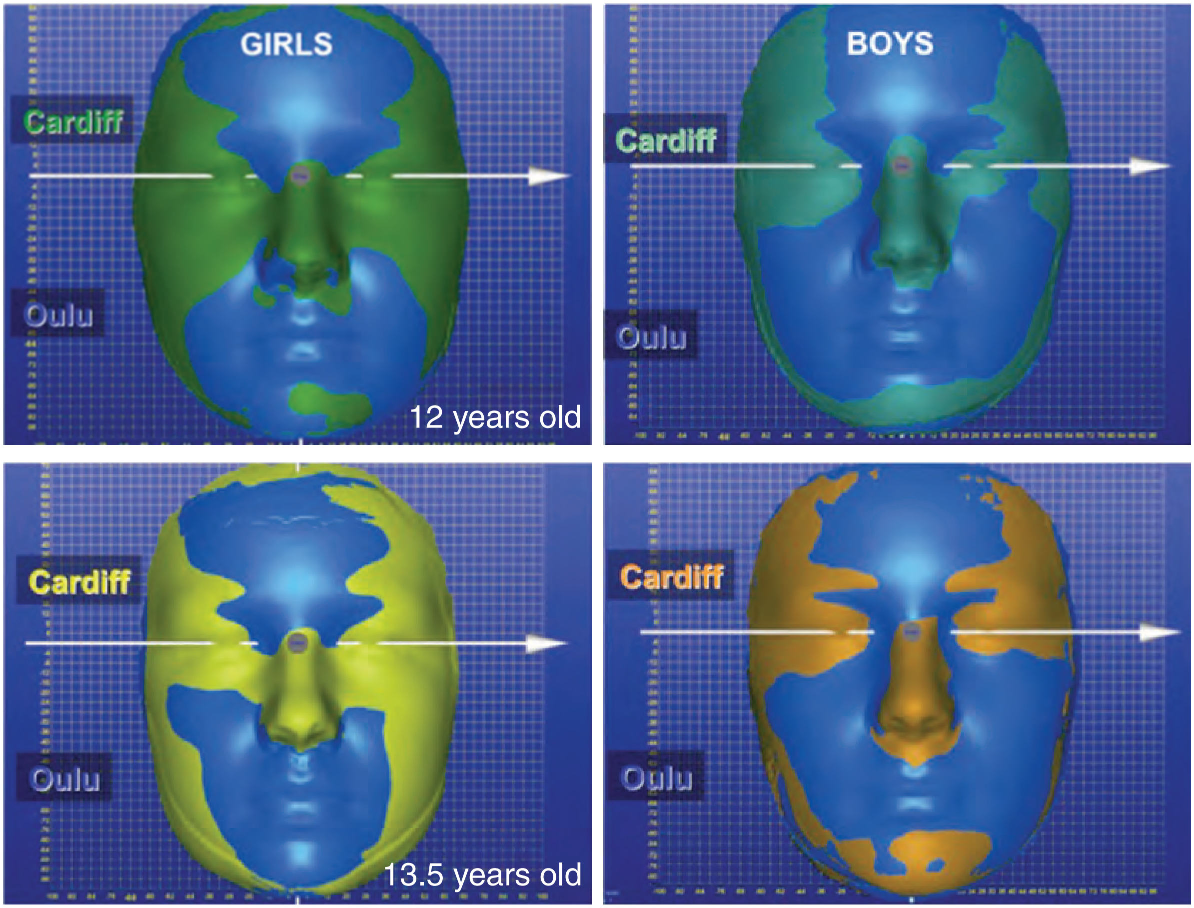


Figure 13.4 Comparison in faces for Finnish and Welsh males and females at ages 12 and 13.5 years old. Differences in the faces indicate that the Finnish faces (blue) tend to be more prominent in the upper and lower parts of the face for both age groups. *Courtesy of Cardiff University.*

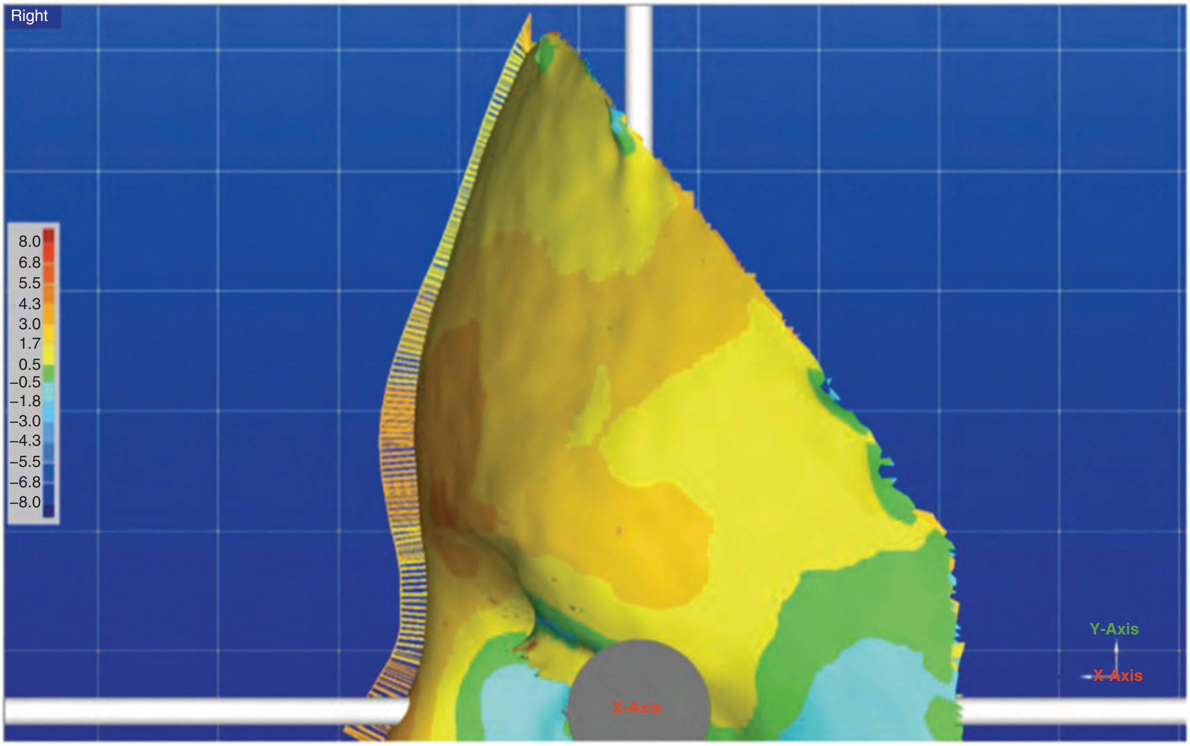


Figure 13.6 Colour deviation map with the scale on the left (–8 mm to +8 mm). The 'warmer' colours indicate positive changes and the 'cooler' colours indicate negative change.

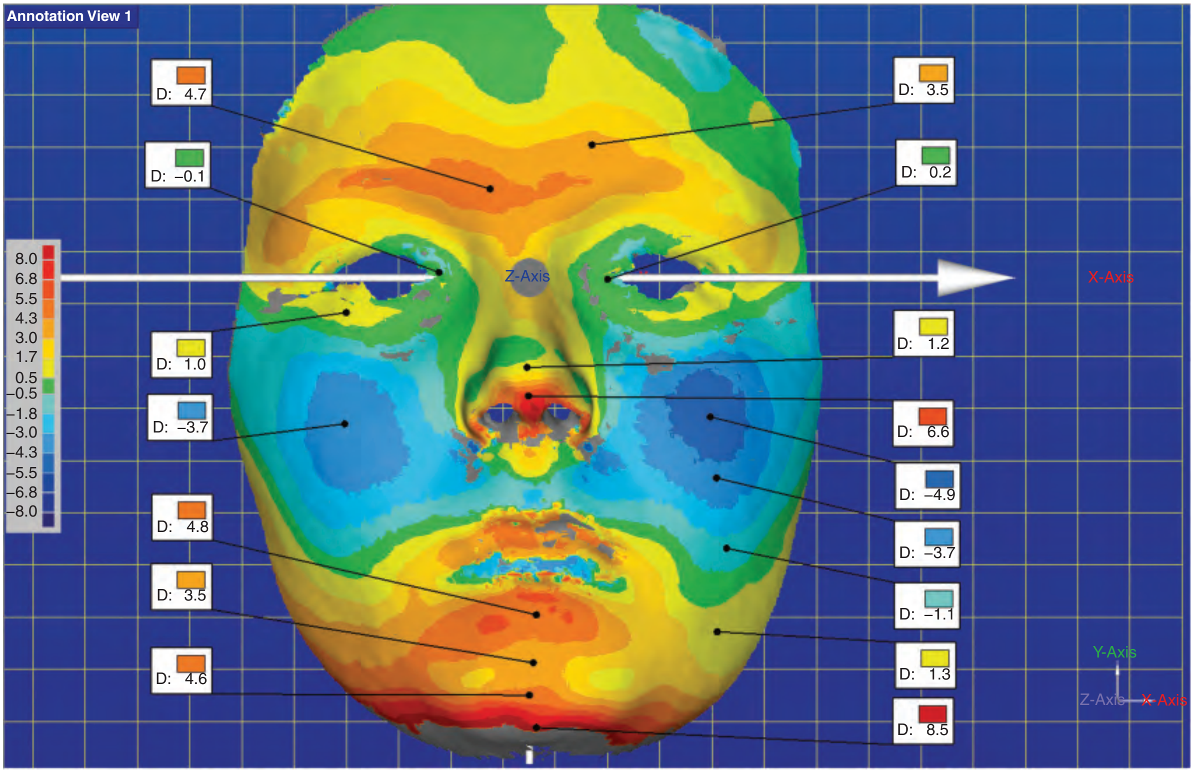


Figure 13.7 Colour map showing positive and negative changes for a male age 12 to 17 years of age (various deviation points on the face are highlighted) with the directional vectors recorded for the 21 landmarks (right).

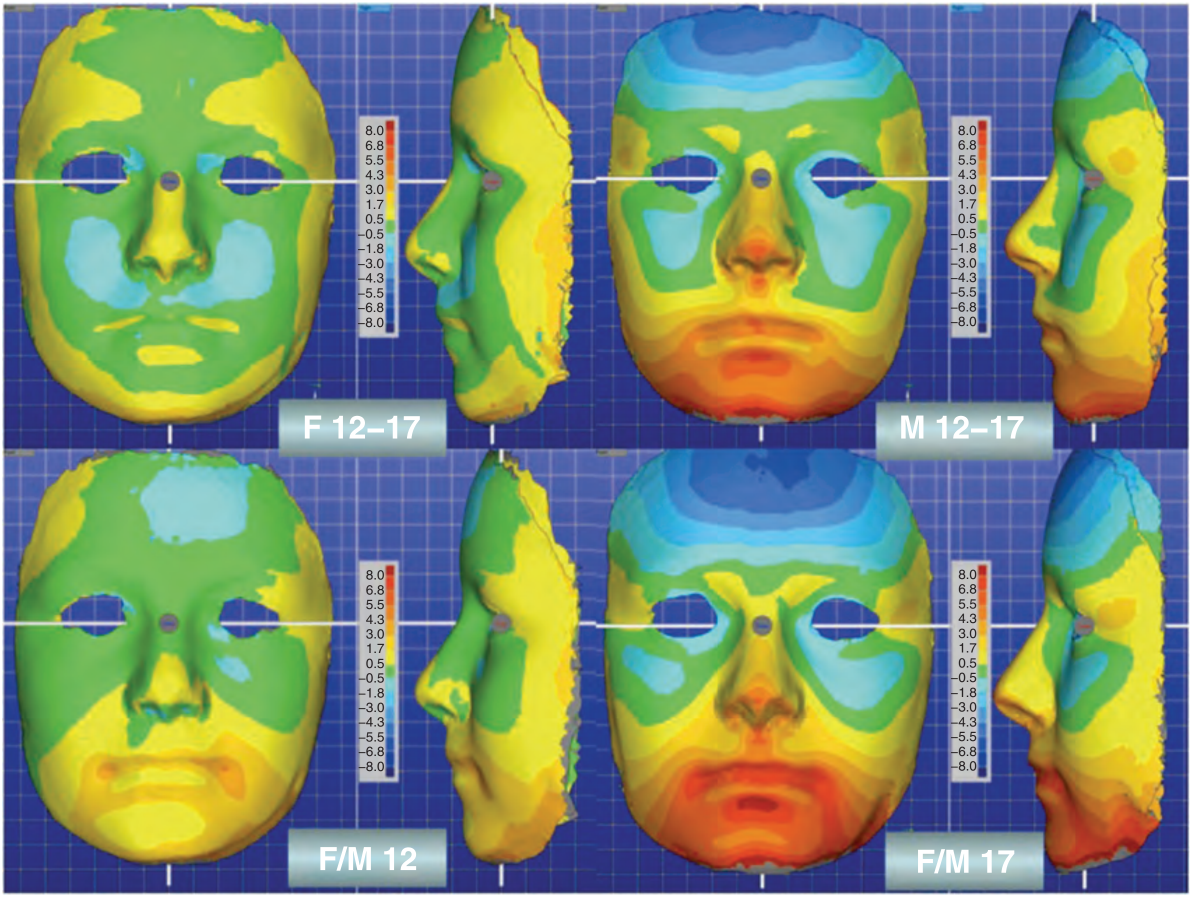


Figure 13.8 Colour maps illustrating the growth of females and males 12 to 17 years of age and differences between males and females at ages 12 and 17. *Courtesy Cardiff University.*

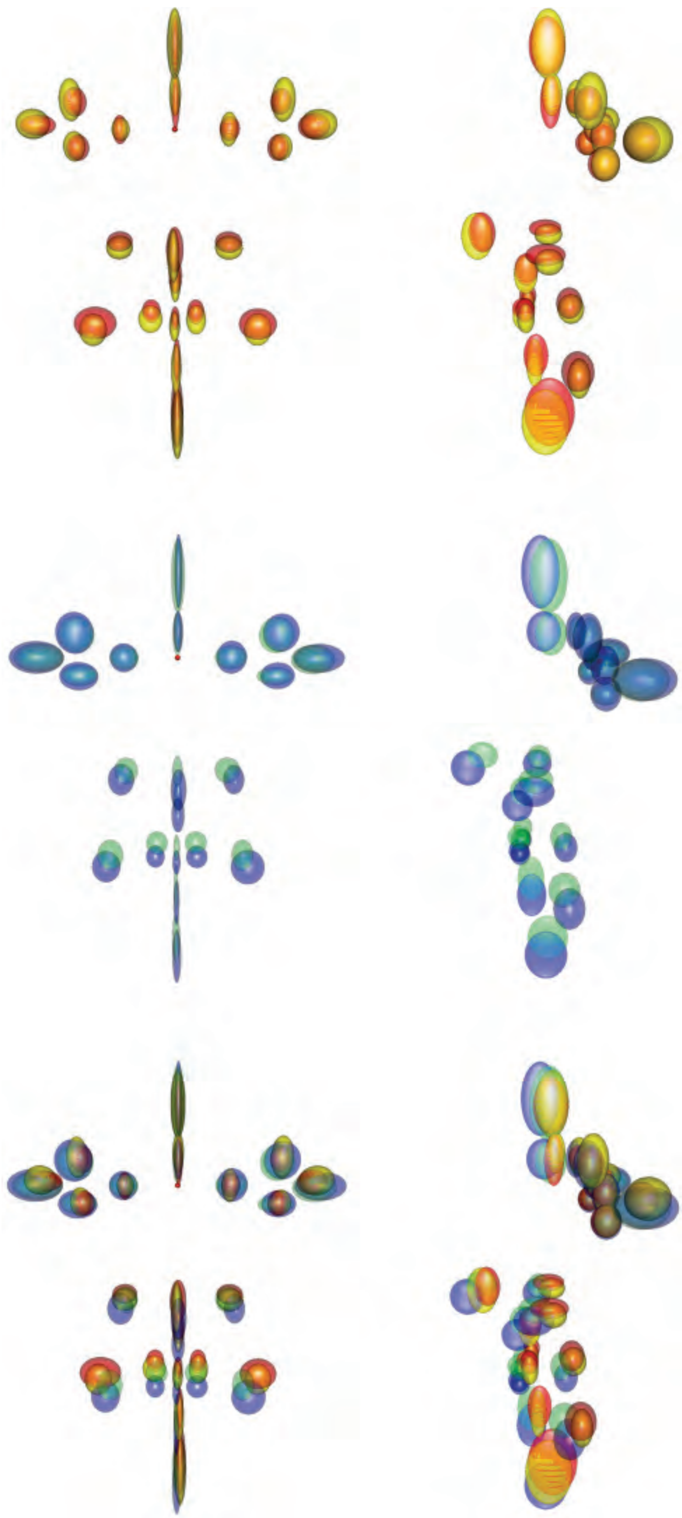


Figure 13.9 Growth of a cohort of 26 males and 23 females superimposed on a standardised framework with the mid-endocanthion point as the origin. Females (top) red – age 12 and yellow age 16. Males (middle) green – age 12 and blue – age 16. Both males and females superimposed at ages 12 and 16 (bottom).

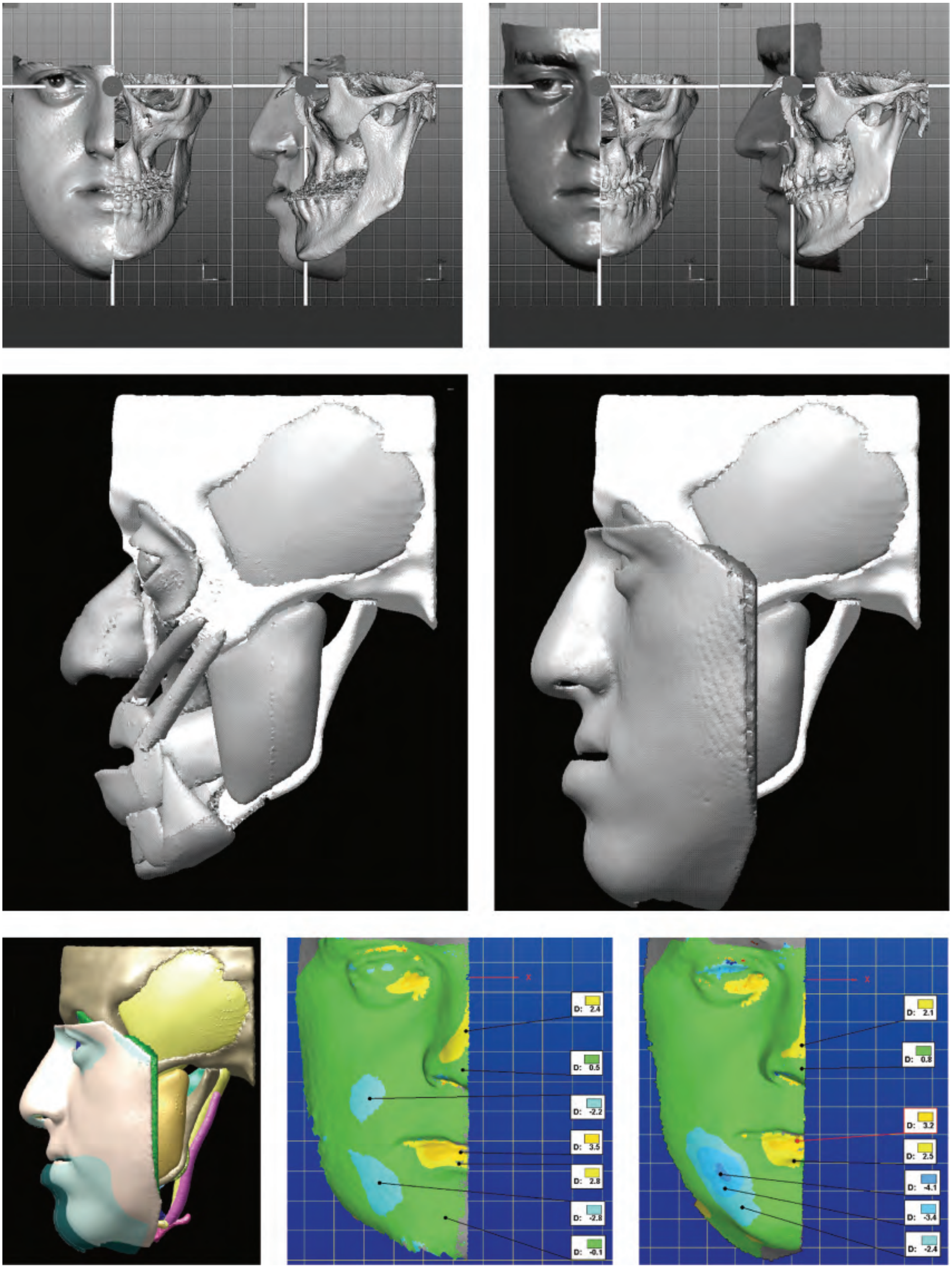


Figure 13.10 Surface laser scan superimposed on a cone beam CT. Top left – before surgery highlights significant soft tissue in front of the maxilla and 6 months post-operative result (top right). Muscles, cartilage, bone, subcutaneous tissues and skin are modelled enabling a functioning biomechanical head model that can currently mimic simple facial actions and expressions (middle). The difference between predicted and actual is shown (bottom left) for both 3 and 6 months (bottom middle and right) post-operatively with discrepancies at the side of the face and lower lip area (3–4 mm). *Courtesy of Cardiff University.*

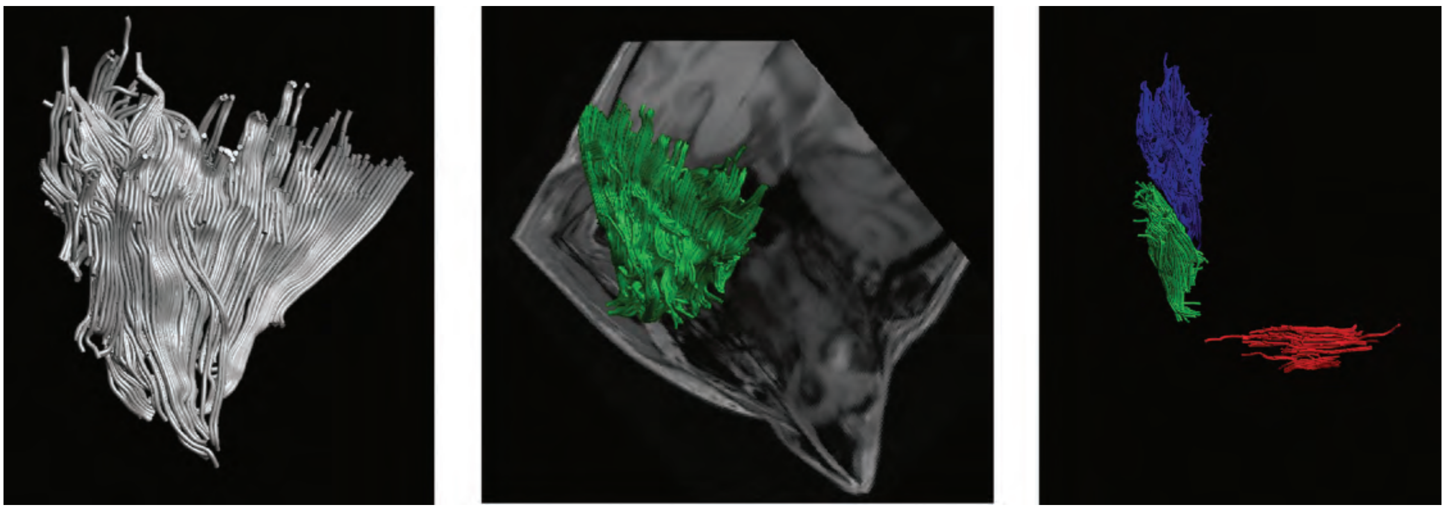


Figure 13.11 Orientation of temporalis muscle fibres isolated (left), spatial relationship of the temporalis (blue), masseter (green) and lip muscles (red) (right). *Courtesy of Cardiff University*.

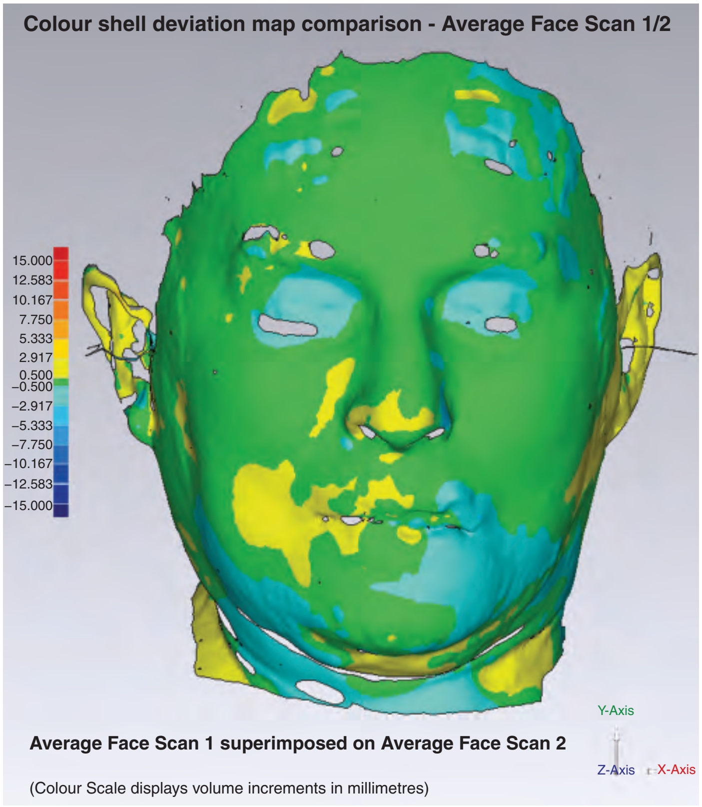


Figure 14.6 The colour shell deviation map comparison for Average Face Models 1 and 2.

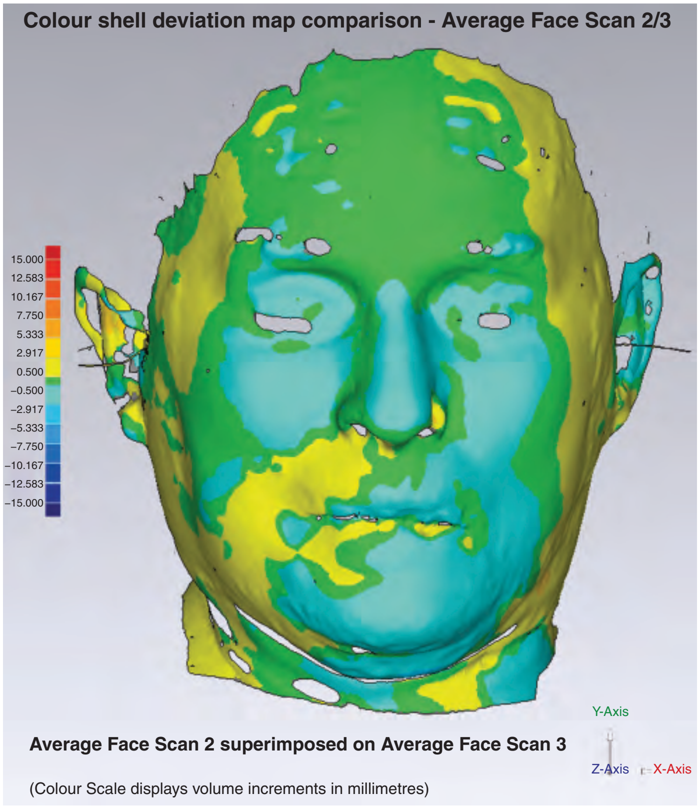


Figure 14.7 The colour shell deviation map comparison for Average Face Models 2 and 3.

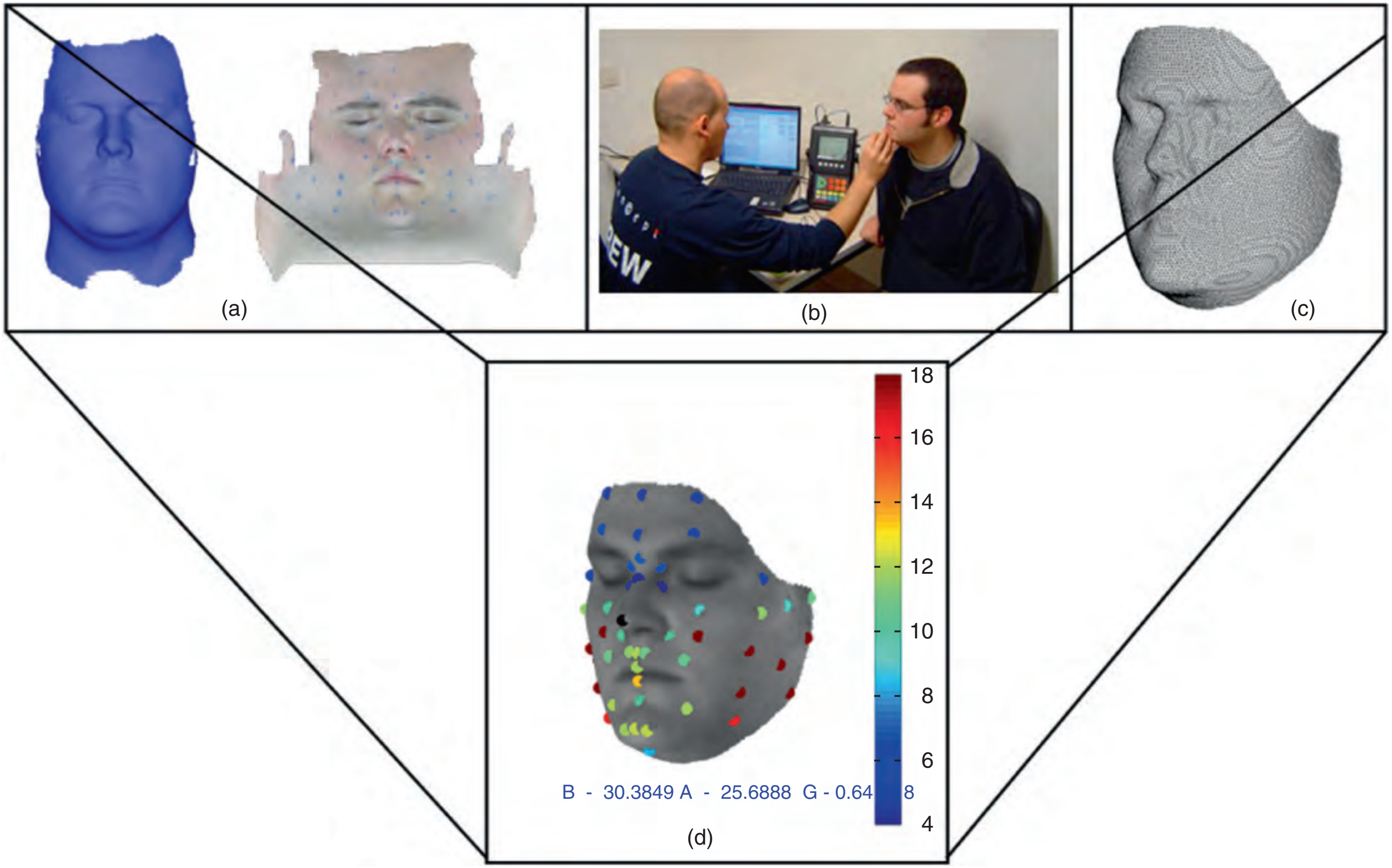


Figure 17.4 3D photogrammetric acquisition of facial shape and texture including ultrasound-based soft-tissue thickness measurement at facial landmarks (De Greef *et al.*, 2005).

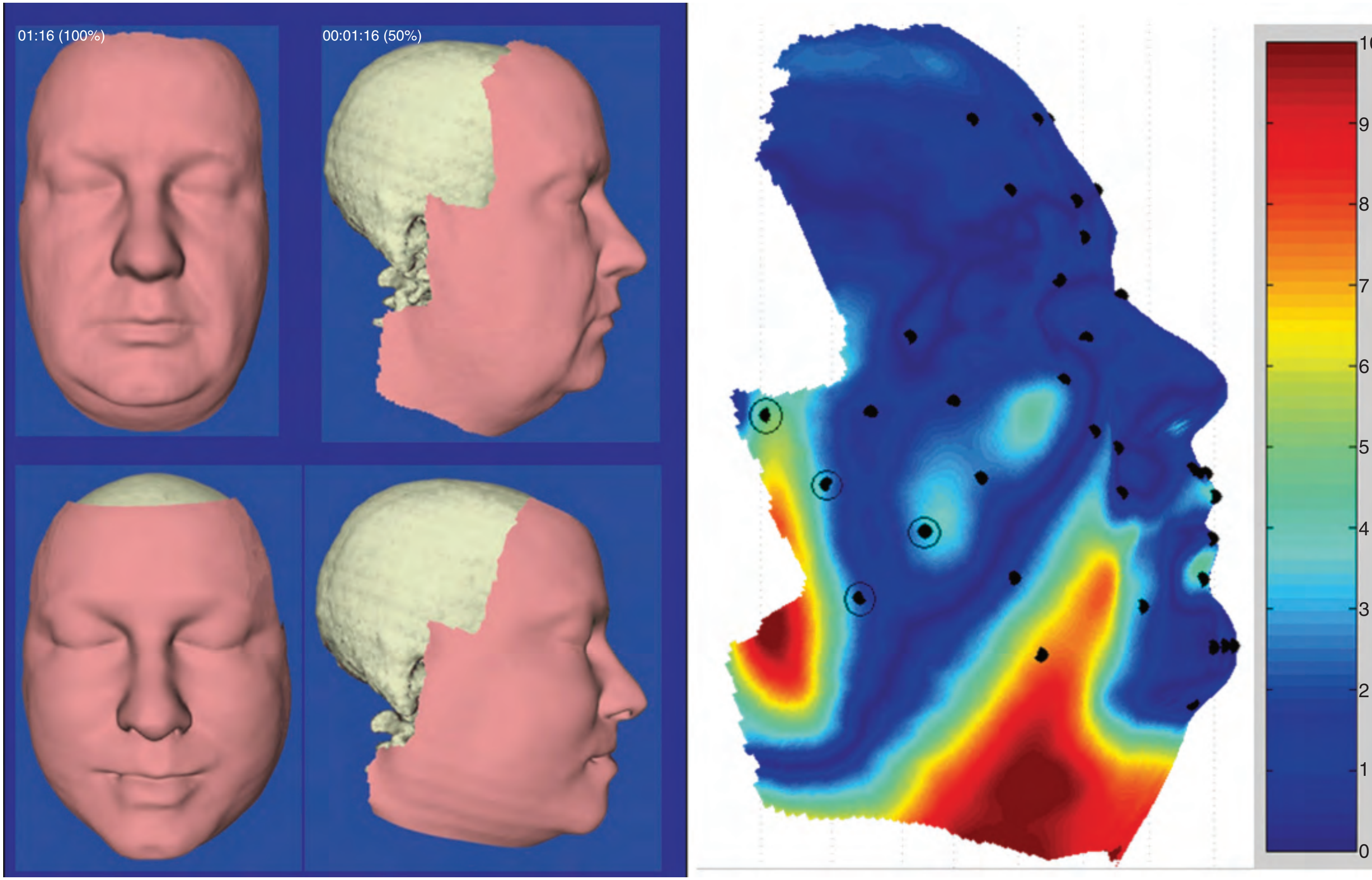


Figure 17.6 The facial geometry of a person in a horizontal supine position (bottom row) and an upright position (top row). Local surface differences visualised by means of a colour-code ranging from 0 to 10 mm (right image).

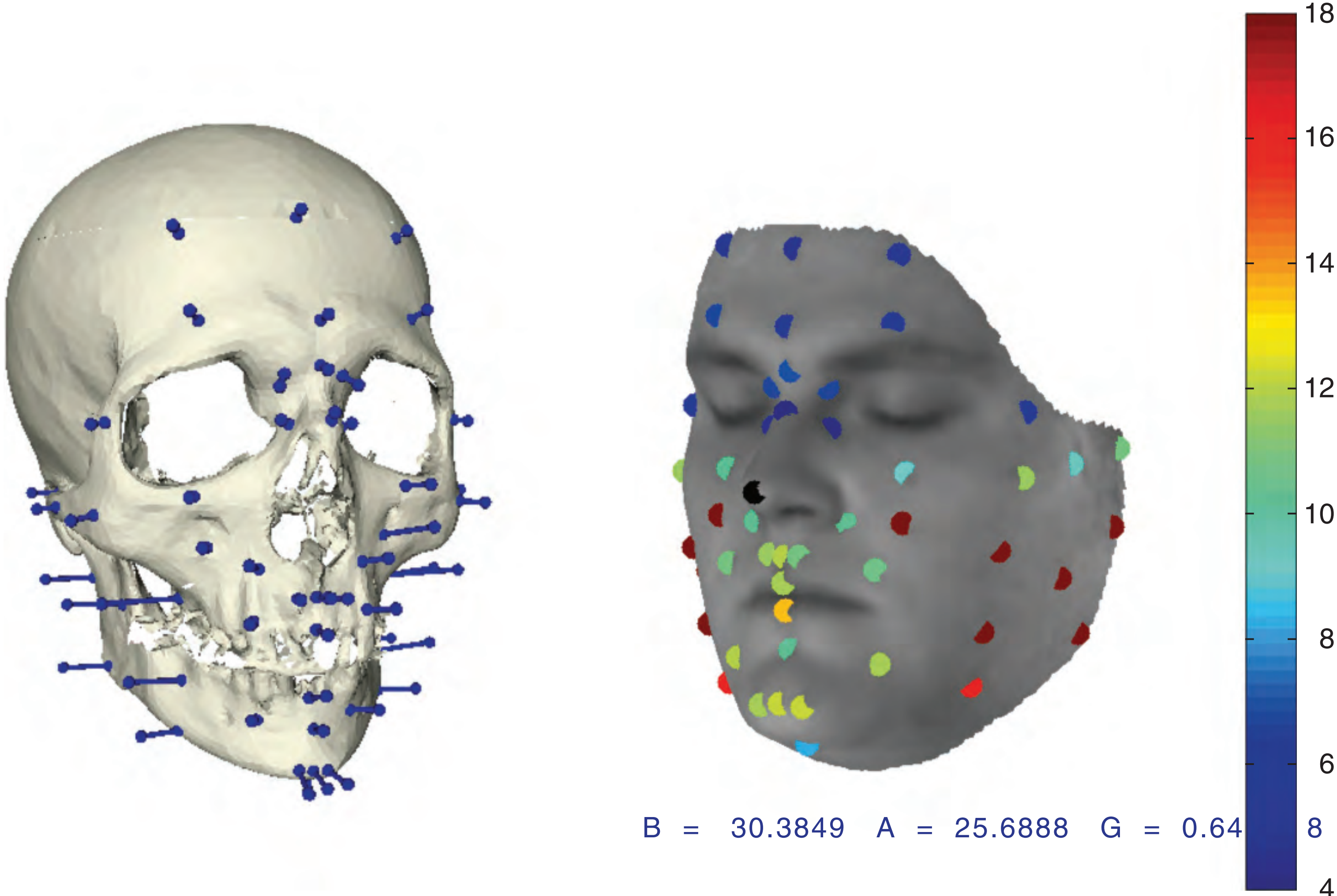


Figure 17.11 Left: Skeletal landmarks defined on skull surface with associated normals. Right: skeletal landmarks defined via facial landmarks and inwards pointing normals to face (colour code represents soft tissue thicknesses at corresponding landmarks).

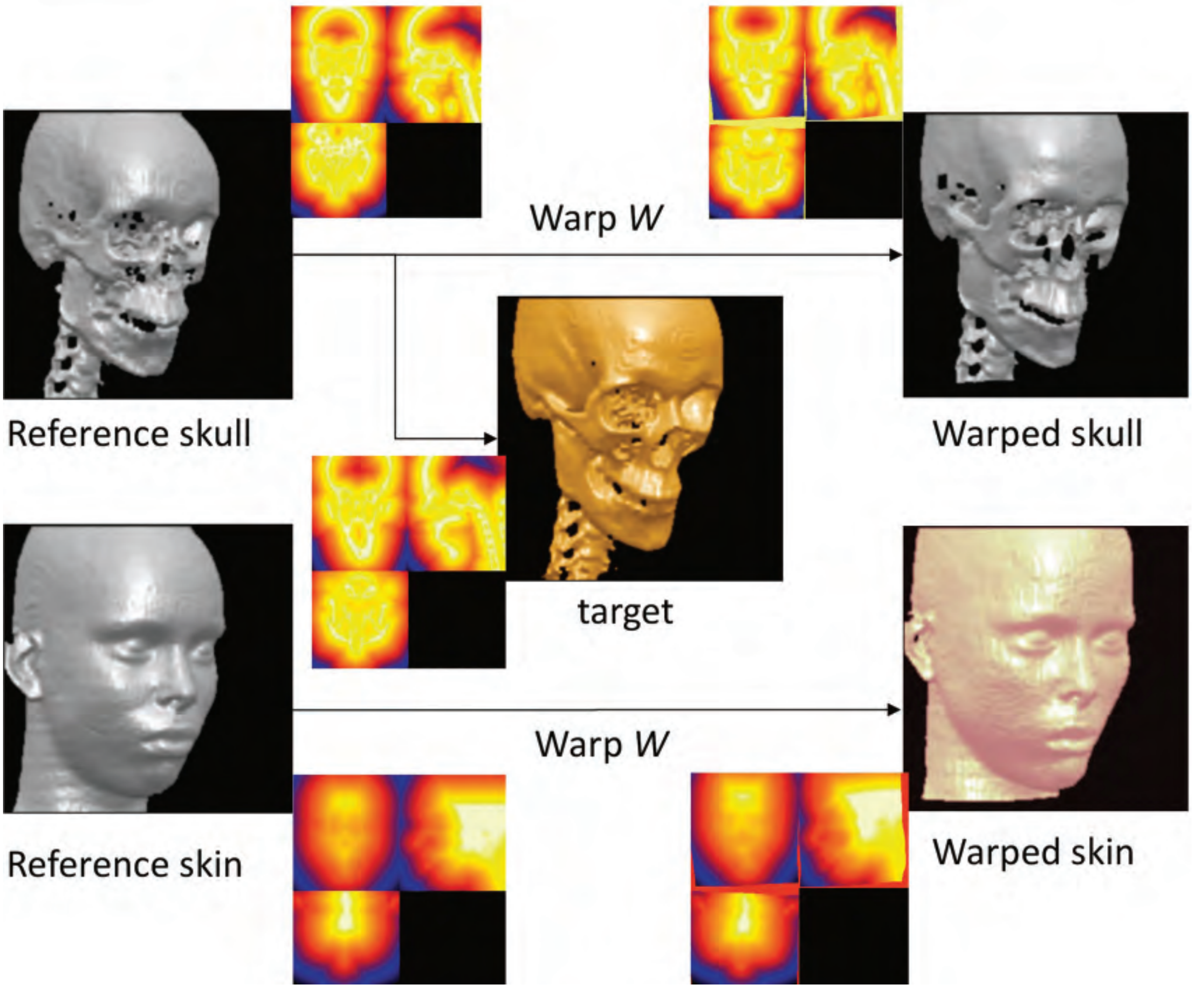


Figure 17.12 Reference skull (top left) is warped to target skull (centre) resulting in the warped reference skull (top right). This warping is applied to the reference head (bottom left) resulting in the warped reference head (bottom right). Transformations are calculated on the implicit signed distance function representations of the surfaces (small multi-slice windows).

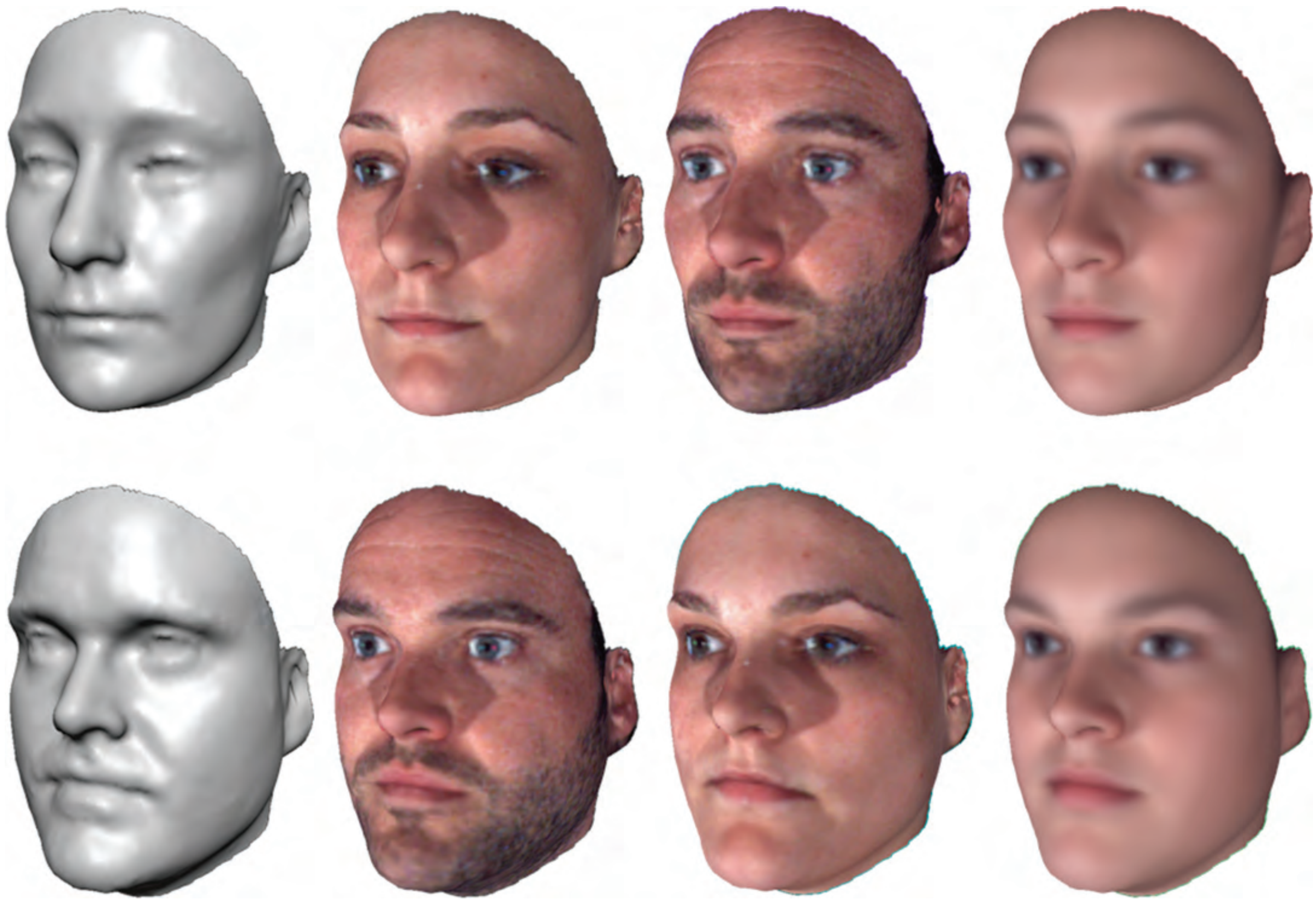


Figure 17.16 Shape versus appearance. First column: shaded surface of two individuals. Second column: corresponding textured surface. Third column: same, but texture of the two individuals switched. Last column: averaged texture from an adolescent database used in rendering.

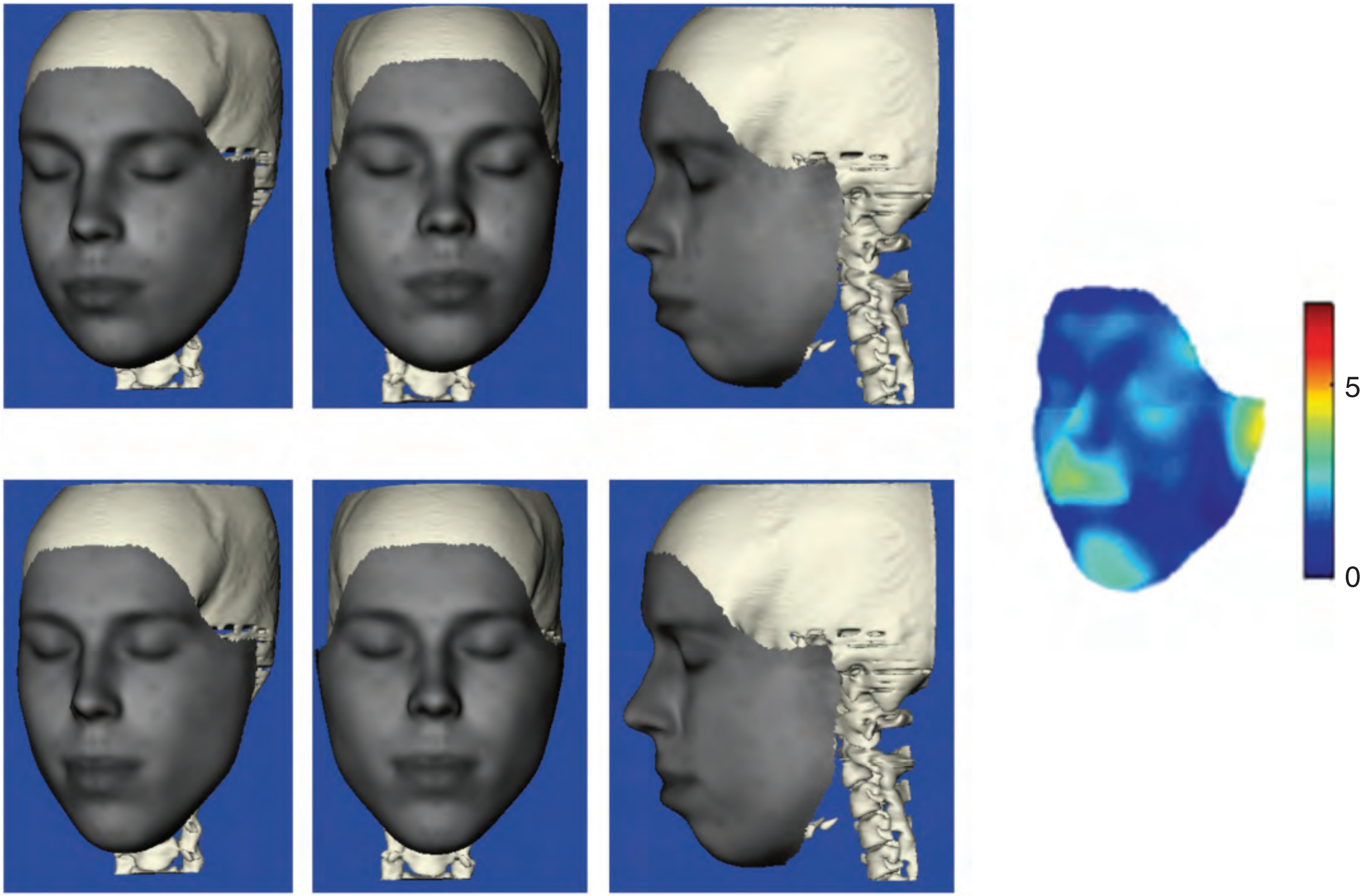


Figure 17.17 Top: 3D rendered view of test skull and ground truth 3D facial surface. Bottom: corresponding views of automated CFR result following the procedure defined by Claes and colleagues (Claes *et al*., 2010*a*). Right: coloured representation of surface differences.

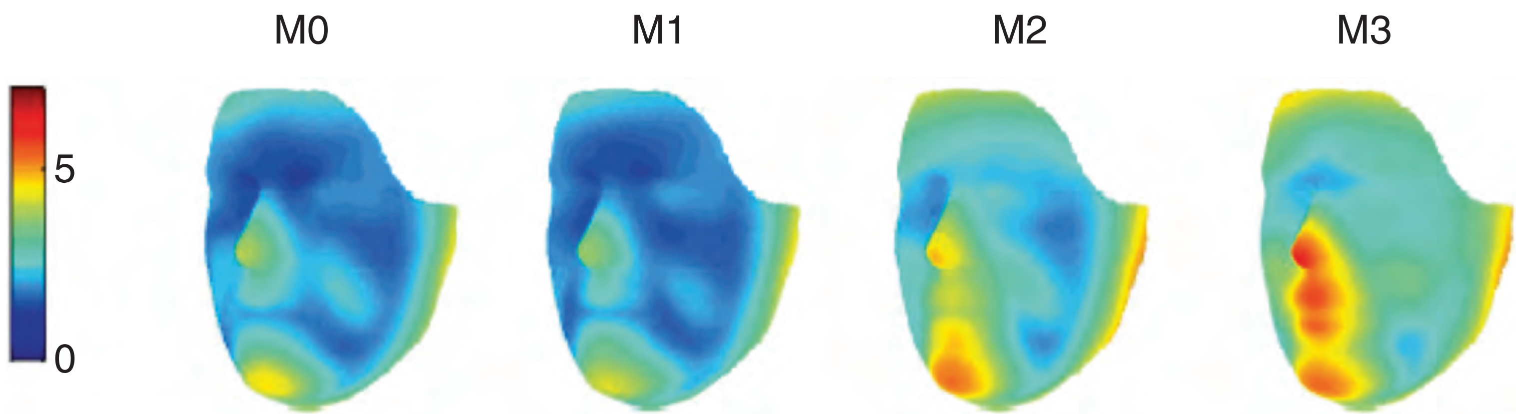


Figure 17.18 Averaged local RMSE (root mean square error) surface reconstruction results over 12 cases, depicted on the geometry of the average face from the database. M0: joint statistical model for face shape and soft tissue thicknesses. M1: separate statistical model for both. M2: generic Thin Plate Spline deformation of attribute normalised template. M3: TPS deformation of average template.