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. As 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 5 million cameras nationwide (McCahill & 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 & Armstrong, 1999; Introna & 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 & 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 & 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 & 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 & 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 & Thasen, 2000). Distance from the camera to the subject (Loftus & Harley, 2004), specificity of viewpoint, expression, and environmental lighting effects all influence face matching (Hill & 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 & 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, 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 judgement 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 judgements (*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) 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 were consistently better 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 (54%) whether the target wore a hat or not. The experts recorded a false rejection rate of 8% and a false acceptance rate of 3% for full head identification, and a false rejection rate of 25% and false acceptance rate of 2% when the targets wore hats. This study suggested that training and experience in facial analysis produces 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 & 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 & 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 & 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 & 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 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 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 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 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. Currently the most common application of this technique is probably for the identification of human remains. 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 & 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. Bromby (2003) describes 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 (Figure 12.1). 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 in crime photographs 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

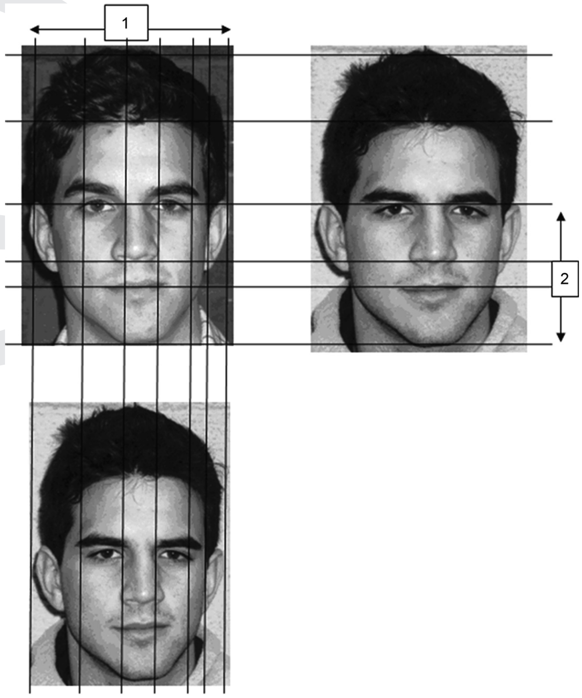


Figure 12.1 Illustration of how the results of photo-anthropometrical analyses could be presented by an expert witness in court. Accompanying the figure would be a table detailing the measurements in terms of **1**. Horizontal distances, expressed as ratios of the distance between the *superaurale* (ears), 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.*, submitted), 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

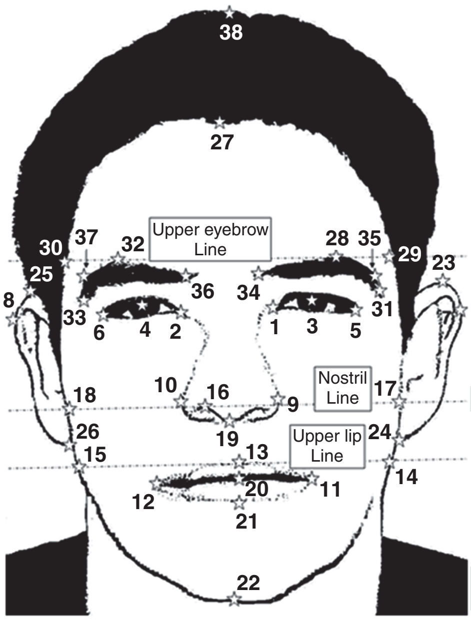
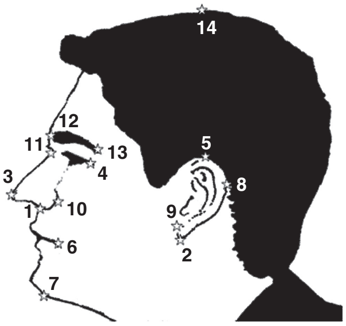


Figure 12.2 Locations of full-face (frontal) landmarks. Common names given in DigitalFace instructions (with anatomical definitions). Note: right and left locations are from the perspective of the viewer.

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*).

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.



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.

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 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 were 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 & 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

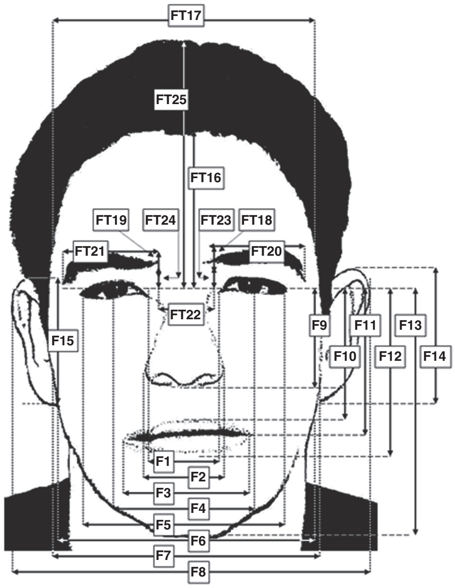


Figure 12.4 Permanent and transient (marked with a T) distance measures produced by DigitalFace in full-face view.

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)

distracter faces were highly representative of 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.

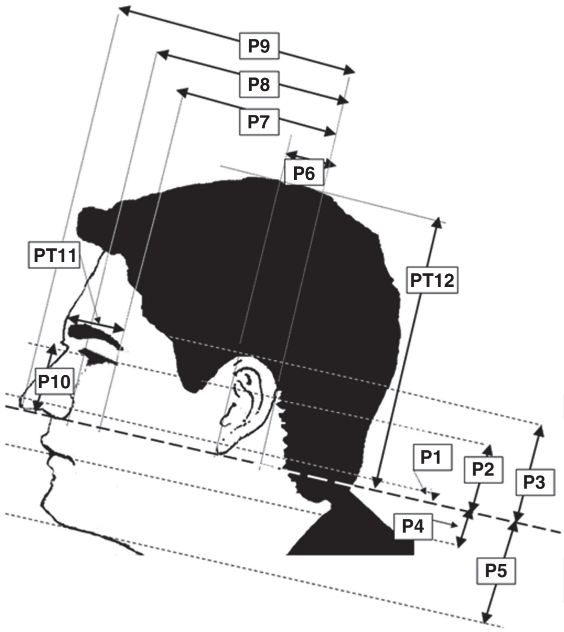


Figure 12.5 Permanent and transient (marked with a T) distance measures produced by DigitalFace in profile view.

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

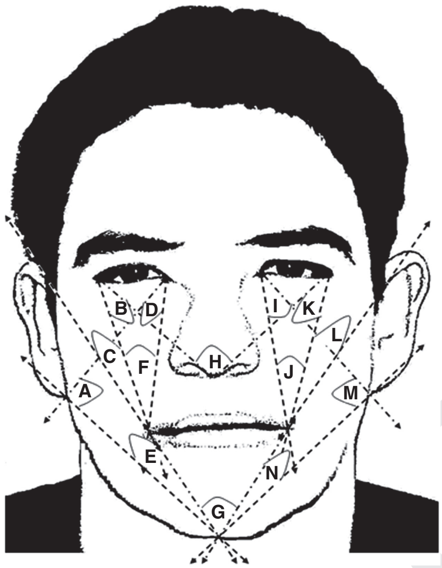


Figure 12.6 Full-face angular measurements computed by DigitalFace.

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 & 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,

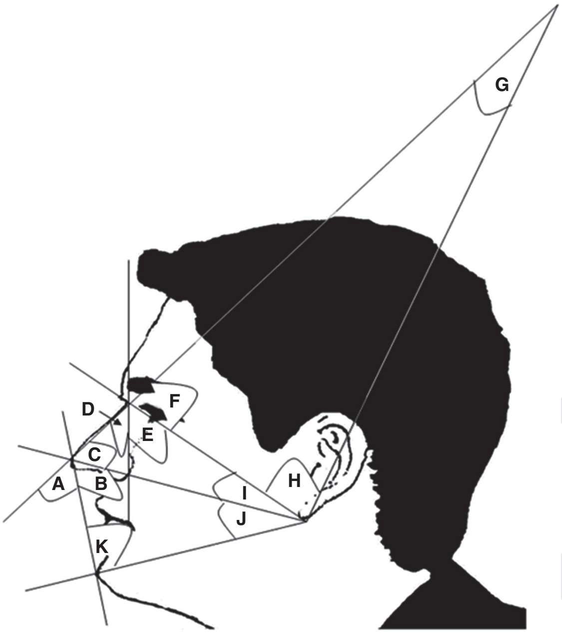


Figure 12.7 Profile angular measurements computed by DigitalFace.

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 a pilot 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 & 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 are 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.

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.