Construction of Facial Composites from Eyewitness Memory

Tredoux, Colin Getty; Frowd, Charlie; Vredeveldt, Annelies; Scott, Kyra

published in
Biomedical Visualisation
2023

DOI (link to publisher)
10.1007/978-3-031-13021-2_8

document version
Publisher's PDF, also known as Version of record

citation for published version (APA)
Tredoux, C. G., Frowd, C., Vredeveldt, A., & Scott, K. (2023). Construction of Facial Composites from Eyewitness Memory. In L. Shapiro, & P. M. Rea (Eds.), Biomedical Visualisation: Volume 13 – The Art, Philosophy and Science of Observation and Imaging (Vol. 13, pp. 149-190). (Advances in Experimental Medicine and Biology; Vol. 1392). Springer. https://doi.org/10.1007/978-3-031-13021-2_8

Construction of Facial Composites from Eyewitness Memory

8

Colin Getty Tredoux, Charlie Frowd, Annelies Vredeveldt, and Kyra Scott

Abstract

Law enforcement agencies often rely on practical technologies to help witnesses and victims of crimes construct likenesses of faces from memory. These 'face composites' are typically circulated to law enforcement officers and made accessible to the public in the hope that someone familiar with the depicted person will recognise their likeness and thus provide the police with a suspect. We will review methods for constructing such likenesses from memory dating back to the portrait parlé of Alphonse Bertillon (Signaletic instructions including the theory and practice of anthropometrical identification. Werner Company, 1896) and the composite images of Francis Galton (Nature 18:97-100, 1878). We will also review more modern methods, ranging from the overlay techniques of Identi-Kit (McDonald, c 1959) and Photo-Fit (Penry J. The Police Journal 43:307, 1970) to feature-based computerised composite systems such as Identi-Kit 2000, FACES, and ProMat. Most early systems were based on the common-sense notion that sectioning a face is invertible: just as a face can be sectioned into components, so it can be recreated by arrangements of sections. This assumption appears to be unwarranted. The underlying problem with earlier face systems may have been the absence of a representational or computational theory. This led in the late 1990s to the development of the so-called third-generation holistic composite systems, which are based on underlying statistical and mathematical models of face images (e.g. ID [Tredoux et al. South African Computer Journal 2006:90–97, 2006], EvoFIT [Frowd CD, Hancock PJB, & Carson D. ACM Transactions on Applied Psychology (TAP) 1:1-21, 2004a], E-FIT [Gibson et al., International Conference on Visualisation, 146–151, 2003]). A special focus of the chapter will be on these newer technologies and other recent technological innovations. Our approach will be to review (i) the methods of operation, (ii) the techniques identified by psychologists and other researchers for improving the quality of information obtained from memory, and (iii) the empirical data on the effectiveness of

C. G. Tredoux (✉) · K. Scott
Department of Psychology, University of Cape Town, Cape Town, South Africa
e-mail: colin.tredoux@uct.ac.za; SCTKYR001@myuct.ac.za

C. Frowd
School of Psychology and Computer Science, University of Central Lancashire, Preston, England, UK
e-mail: cfrowd1@uclan.ac.uk

A. Vredeveldt
Department of Criminal Law and Criminology, Vrije Universiteit Amsterdam, Amsterdam, Netherlands
e-mail: a.vredeveldt@vu.nl

L. Shapiro, P. M. Rea (eds.), *Biomedical Visualisation*, Advances in Experimental Medicine and Biology 1392, https://doi.org/10.1007/978-3-031-13021-2_8

these systems at representing faces from memory. We will consider related issues, too, including the question of whether face composites damage witness memory, and the ethics of face composition.

Keywords

Memory · Eyewitness · Face composite · Face likeness · Synthetic face

8.1 Introduction

Eyewitnesses to crimes are frequently required by law enforcement agencies to assist in the construction of visual likenesses of criminal offenders. These 'face composites' can be created with the assistance of a portrait artist or various mechanical and automated methods, such as Identi-Kit and Photo-Fit. In recent times, probably because of the large-scale impact of computing technologies, composite systems have become much more sophisticated and more widely used.

Police forces in many countries rely on the creation of face composites as part of their investigative practice. Composites are distributed to newspapers and are posted on information boards in shopping centres, in schools, on the internet, and on television. They are a way of obtaining identifications of existing suspects, but they are also used to generate leads for potential new suspects. A representative survey of US law enforcement agencies found that 35.5% (112) of agencies use composites and on average each created 3 composites in the year under survey, 2010 (Police Executive Research Forum 2013). In South Africa, Schmidt and Tredoux (2006) found that over 500 composites were produced annually in a city of 4.5 million people.

When a suspect is discovered that resembles the composite picture, eyewitnesses to the original event may be required to attempt an identification from a line-up. However, if an innocent person who happens to look like the composite is arrested, there could be grave consequences. This appears to have happened to Kirk Bloodsworth, a US Army veteran, who resembled a composite released to the media by police and could not provide a convincing alibi for his whereabouts at the time of a young girl's murder. Two witnesses identified him in a line-up. He was sentenced to death for the murder at trial and to two life sentences at retrial (Junkin 2005) but was exonerated by DNA testing after 9 years in prison. In this case and some others like it, an innocent person's coincidental resemblance to a composite was sufficient not only to get him arrested in the first place but also to get him identified by eyewitnesses and convicted for a crime he did not commit. Despite this problem, composites continue to be widely used. When there are no other clues to a perpetrator's identity than a witness's memory, turning the memory into a composite sketch to assist with the search for a perpetrator seems likely to continue as a police practice.

A particularly well-known case involving the generation of composite images is that of the 'Yorkshire Ripper', a murderer who dismembered young women in the 1970s in Yorkshire. It took the English police several years to apprehend, try, and convict Peter Sutcliffe for these crimes, and in the process of searching for the perpetrator, many composite likenesses were constructed and published. These are shown in Fig. 8.1, and many do not resemble each other very well at all.

The production of a face composite, which is the construction of a face from memory, is without doubt a challenging task. Can it be done with any degree of accuracy and with any degree of consistency? Does it help the police apprehend guilty suspects, and does it lead to the solution of crimes? Might the act of viewing or creating composites actually harm memory? These are important questions, not least because composite creation is a costly exercise for police forces but also for our understanding of face memory processes. Our goal in this chapter is to review historic attempts to visualise faces from memory, especially in police work. We review early

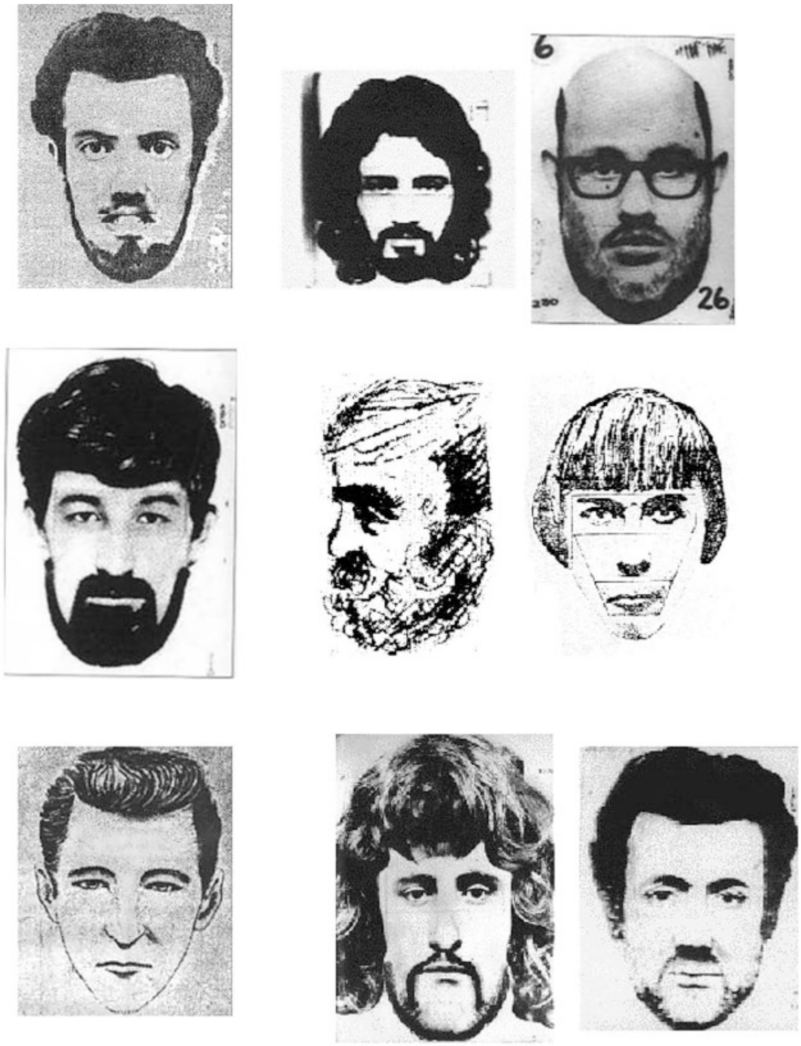

Fig. 8.1 Various face composite constructions of the 'Yorkshire Ripper' from interviews with witnesses

attempts from just after the rise of modern policing in the nineteenth century, all the way through to modern cutting-edge techniques that depend on artificial intelligence systems. We also review research on the efficacy of such systems and their possible adverse consequences.

8.2 History of Face Composition

8.2.1 Portrait Artists

Portrait artists appear to have been used in policing since at least the 1880s but may well have been used much earlier in other contexts to draw faces from memory. Several high-profile cases in the late nineteenth century captured the public imagination, and sketches from portrait artists in those cases were said to have helped police identify and detain the suspects. These cases include that of Percy Lefroy Mapleton, who was convicted for a murder that took place on a train in London, England, and of Dr. Crippen, who murdered his wife and was caught crossing to the USA on a ship, with the assistance of an artist's sketch of his imagined current appearance (see Taylor 2000, and Davies and Valentine 2007,

Fig. 8.2 Artist's sketch of Percy Mapleton, wanted for murder, circa 1881. Source: Wikimedia.org

for accounts). These early sketches were often highly stylised and caricature-like (see Fig. 8.2).

Portrait artists have remained in use in many police forces in 140 years since the Mapleton case. This is understandable, since their training is precisely to render face sketches in ways that assist human perception. Indeed, some research suggests that portrait artists are better than average on face perception tasks (e.g. perceptual discrimination, face matching; cf. Devue and Barsics 2016; Hsiao et al. 2021), and Hsiao et al. speculate that portrait artists may also be better able to access invariant information about faces.

Police sketch artists are usually trained portrait artists who interview eyewitnesses and draw the perpetrator's face through close interaction and conversation. Some artists have developed specific techniques for this purpose, sometimes to the point of developing a 'translation language' that facilitates the translation of the witness's memory into a facial composition. For example, the witness may be asked to think of a face feature as if it were similar to a common object and then to indicate the width or curvature of the face by thinking of the appearance of the object. We refer later in the chapter to studies that have tested the ability of sketch artists to construct accurate likenesses based on eyewitness memory.

It is not common now for police forces to hire sketch artists because of the high cost of hiring them, but some police forces do hire sketch artists and remain convinced of their special skills (see Boylan 2001).

8.2.2 Bertillon's Portrait Parlé and Anthropometry

Sociologists have argued that the rise of sketch artistry within police services in the nineteenth century is part of a sea change in the conceptualisation and response to crime. One element driving this change was a new concern with the identification of individuals within nation states, especially criminals. In an attempt to quell an apparent 'epidemic' of crime, which jurists and bureaucrats thought could be traced to repeat offenders (recidivists), a new demand was created for the identification of criminals. If recidivists were destined to repeat their crimes, there was an urgent need to identify them (for a much more extensive discussion along these lines, see Cole 2001). Several innovations for gathering information about identity sprung up at this time, which were initially concentrated on the task of identifying criminals but which would later become used to identify the citizenry in general. One consisted of recording detailed physical and behavioural descriptions in logbooks, and systems of index cards, typically arranged by name, and containing detailed descriptions, which could hopefully lead to later identification and prosecution. Early systems were easily defeated by criminals who used aliases but were later improved with extensive collections of photographs of faces of criminals ('galleries of rogues'), although it soon became clear that there were limitations in using photographs to verify identity, as appearance could change naturally over time or with some effort on the part of criminals.

These attempts to cast a net that ensnared all criminal recidivists went hand in hand with a forensic anthropology that tried to find scientific and taxonomic bases for identifying criminals. These were often biologically misguided and deterministic, for instance, proposing that variations in skull physiognomy were associated with variations in criminality (so-called phrenology). Perhaps the most successful approach in the late nineteenth century and early twentieth century was the anthropometric classification system devised by Bertillon (1896). Criminals (or indeed any person of interest) could be measured on a number of physical dimensions (e.g. head length, head breadth, foot size, index finger length), facial dimensions (forehead, ear, nose measurements), and many other additional characteristics (e.g. age, eye colour, height, weight). These were entered on an index card, along with frontal and profile photographs of the subject, and could be used for identification purposes. Bertillon considered the reduction of physical appearance to a set of codes to be a key component of his system, and he called the method of doing so 'signaletics'. Bertillon did not ascribe to the view that criminality was the consequence of innate character or behavioural types but saw his system clearly as a way of precisely describing individual differences, allowing the accurate construction of appearance from descriptions. Indeed, police officers were for a time trained in the use of the Bertillon signaletics to describe faces they had seen, on the assumption that the signaletic representation of a face was a truer, more invariant coding of a face than a photograph, which could change considerably if taken under different conditions and at different times. Figure 8.3 is an image of the Bertillon index card (or *portrait parlé,* 'spoken portrait', as it was known in France at the time), as applied to its progenitor, Alphonse Bertillon.

Bertillon's system was used to capture the details of hundreds of thousands of people in continental Europe, the UK, the USA, and many other countries. Indeed, a 'Bertillon cabinet' was considered an essential part of a well-equipped police station, and in Paris, the system took up much of an entire building. Bertillon's method of indexing individuals (known widely then as 'Bertillonage') was also very useful for organising and searching the hundreds of thousands of photographic images police services built up (over 100,000 per decade in Paris alone in the late nineteenth century). Indeed, he applied the then novel idea that statistical distributions can be used to order and retrieve information: he used the binomial distribution and the idea of binomial search extensively and cleverly. Bertillon achieved some spectacular successes with his system. A ship, the Drummond Castle, sunk en route from Cape Town to London, near Molène in France, with the loss of 242 crew and passengers. Bertillon photographed and applied his 'signaletics' to 27 of the recovered bodies, and 10 were identified from his descriptions (Higgs 2011). On the other hand, Bertillon also appeared as an expert witness for the prosecution in the infamous case against Captain Alfred Dreyfus (Paleologue 1957), where he attempted to apply ideas similar to 'signaletics' to handwriting, which helped build the case against the innocent Dreyfus but was quickly debunked.

The enormous cost of the Bertillon system ultimately made it impractical for most police forces, and it fell into disuse when it became clear that fingerprints, which would often be left as traces at crime scenes, were a more easily collected form of unique identification of individuals. Fingerprinting technology became part of police practice in the late nineteenth century, and 'Bertillonage' and 'signaletics' became obsolete soon afterwards.

Nevertheless, it was an important step in the systemisation of the idea that criminals and other wanted people could be discovered by issuing information about their physical appearance, even if police did not have a visual likeness available. In particular, the key idea in 'signaletics', that careful measurement of an individual could reduce an individual's appearance to a small set of index numbers, or coefficients, allowing highly efficient and rapid transmission of information about physical appearance of people, lives on today in the Identi-Kit, Photo-Fit, and FACES software composite systems. Modern composite systems differ from Bertillonage in that the visual likeness is the key artefact, rather than a signaletic encoding.

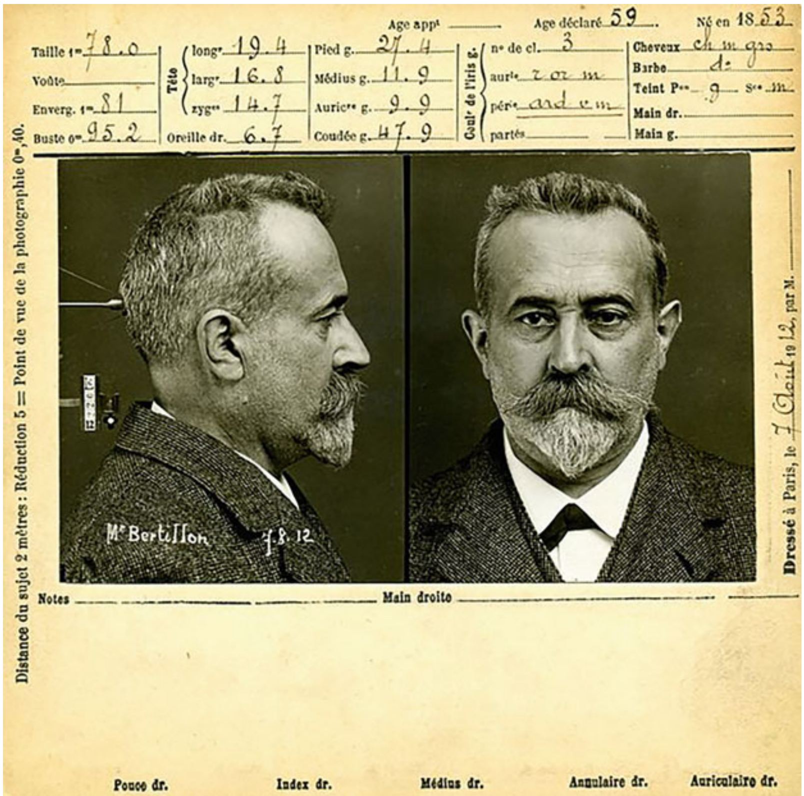

Fig. 8.3 Alphonse Bertillon's entry in his eponymous Bertillon anthropometric indexing system. Source: Wikimedia.org

8.2.3 Galton's Composite Images

One interesting attempt that has a striking resemblance to the modern 'eigenface' synthesis methods is that devised by the Victorian polymath and early differential psychologist Francis Galton. Galton experimented with multiple exposures of photographic negatives of human faces on light-sensitive material, combining the images of different people into a single composite (for an overview, see Chap. 12 in 1924). Although he thought at first that he might be able to characterise the physical appearance of criminal types in this way (see Fig. 8.4), he realised that this was unlikely to succeed and instead turned his attention to the more general use of 'analytic photography'. For instance, he developed a theory of so-called transformers, functions which could 'convert' one photograph to another, e.g. an unsmiling subject to a smiling subject, a St. George's Cross to a St Andrew's cross, and so on. We can see in this the lineaments of the idea that summing and weighting basis image vectors (e.g. 'eigenfaces') could create images of faces and that we can transform one face into another by changing the weights. Also of considerable interest is his 'automated profile portraiture'. Taking measures at standard points on many hundreds of profiles of faces, he segmented the sketches into features, determining an 'average' feature. Individual faces could be indexed according to the deviation of their individual features from the averages, which allowed for extraordinary efficiency of representation (four coordinates made up the index). The index could in principle be used to construct an image of the person by resizing features and adding angles between features and connecting points. What is missing from this method, though, is the ability to create faces that are not in the original database, but it would be an easy generalisation to do so (sampling coordinates from an appropriate distribution of the coordinates, for instance). Although Galton's ideas were not implemented

Fig. 8.4 Composite images of convicted criminals by Francis Galton. Source: Galton.org (scan of page from Pearson, K. *The life, letters and labours of Francis Galton*. Cambridge University Press, 1914, 1924, 1930)

in a practical way in the service of policing, it seems clear that they foreshadowed the development 50 years later of the featural face composite kits and 100 years later of the eigenface software composite systems.

8.3 Solutions from the Twentieth Century

8.3.1 Manual Systems

8.3.1.1 Identi-Kit

Hugh McDonald and Harry Rogers, detectives working in Los Angeles in the 1950s, recognised the need for a unified system capable of reproducing human faces without the help of trained police artists. They developed a face recall kit, which McDonald named 'Identi-Kit', marketed through Smith and Wesson. The kit contained several hundred line drawings of facial features reproduced on acetate sheets. The police would begin the interview of an eyewitness by obtaining a general description of the perpetrator, asking the witness to describe each feature in turn. An initial composition could be created from these descriptions by selecting acetate sheets corresponding to features in the description, overlaying them to form a composite face. This initial composition could be improved by adding or changing features until witnesses were satisfied that they had achieved a good likeness or could do no better. Some aspects of the composition could be modified using manual artistic methods and tools, emphasising details, for instance, with a wax crayon. The elements of the Identi-Kit system were numbered so that information about the portrait could be quickly transmitted, allowing it to be constructed in part from a distance. Identi-Kit was one of the first facial composite systems and is still widely used in the USA and elsewhere. The current version is Identi-Kit 7, which is a software reinvention of the original Identi-Kit. Figure 8.5 shows a composite produced by an experienced police Identi-Kit operator from a witness interview and Fig. 8.6 a sketch construction by an experienced portrait artist, for comparison.

8.3.1.2 Photo-Fit

Identi-Kit was adopted soon after production in many countries, including the UK. However, it was judged in the UK to suffer from practical problems, chief among these being that to produce a satisfactory Identi-Kit composition took many hours. Other problems included the unavailability of UK hairstyles and accessories (the kits were US-made). UK police forces sought more locally appropriate alternatives. Jacques Penry (a man of questionable credentials, real

Fig. 8.5 Identi-Kit 2000 image created by portrait composite operator, from witness memory (target was the same person as in the right panel)

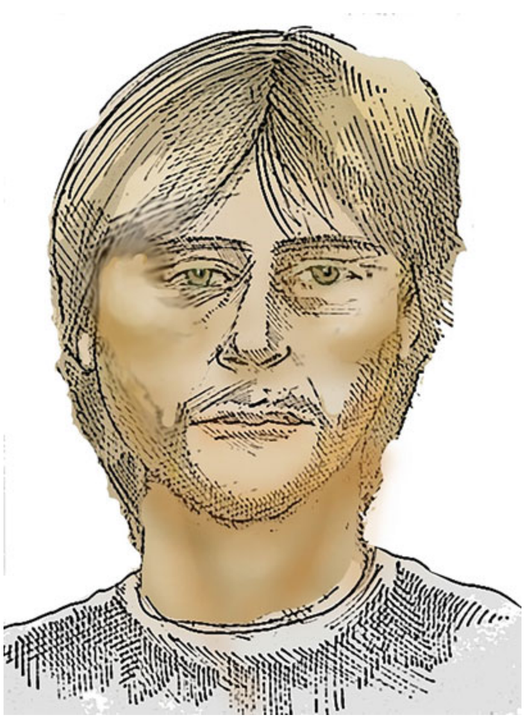

Fig. 8.6 Sketch image created by portrait artist, from witness memory (target was the same person as in the left panel)

name Bill Ryan), who called himself a 'facial topographer', proposed a planned face composite system to the Home Office Police Research Branch, to be composed of snippets of face portrait photographs, and would allow the construction of photo-realistic portraits of perpetrators from witness memory. Penry was given access to the photographic 'mugshot' books of several UK police forces to gather enough facial features to construct a working kit. Two prototypes were tested, and Penry negotiated a deal with a board game company that it would produce the kit under licence if adopted by the Home Office. The full kit, when released as Photo-Fit, had a large collection of hairstyles and foreheads, eyes, noses, mouths, and chins. There were accessories such as eyeglasses, hat wear, facial hair components, and face wrinkles/lines. The White male kit had a total of 855 elements. There were also kits for White females, African, Asian, and Middle Eastern male faces. The new system was rapidly adopted by police forces in the UK and, around the world, in 23 countries, including Brazil, the USA, Brazil, Nigeria, and Australia. Seemingly immediately successful, Photo-Fit was in use for decades. Figure 8.7 is an example construction of a face with Photo-Fit, with the face in full view. For an extensive, if critical, review of the development of Photo-Fit, and its inventor (an apparent 'pseudoscientist'), see Lawrence (2020).

8.3.2 Computerised Feature-Based Systems

Computerisation of the composite 'kits' led to several benefits. One important benefit is that composite systems could provide a much greater range of facial features for which a face could be constructed. Simply providing many more examples of hair, eyes, brows, noses, and so forth allows for a more accurate composite to be created. In practice, composite systems store many examples of each facial feature. These examples are classified into their physical characteristics to allow specific types to be identified. Eyebrows, for example, could be

Fig. 8.7 Photo-Fit construction using the 1970s kit. Source: Revell and Wiley (2007), with permission

classified by their shape, thickness, and colour. In fact, the ability to store a virtually unlimited number of items in a computer database allows faces to be constructed of both sexes and for many races. Indeed, feature databases can include scars, marks and wrinkles, facial hair, and 'accessories' worn by perpetrators: hats, hoods, glasses, jewellery, and even tattoos. Provision of clothing is also possible.

For mechanical systems, facial features were printed onto transparencies or stiff card, causing obvious limitations for the production of a desired face. Computerisation greatly improved the way that facial features could be sized and positioned on the face, further expanding the number of possible faces that could be constructed. Computer graphics technology also allowed features to be blended together more acceptably. This improvement itself overcame an issue with the original Photo-Fit system, namely, that the presence of a dark outline between facial features—a result of features being slotted together in a template—interfered with recognition of the face (Ellis et al. 1978).

A third important development relates to the way in which an eyewitness would select facial features. Previously, facial components tended to be shown in isolation to a complete face. A witness would select from pages of eyes, noses, mouths, etc. However, there is considerable evidence that this is not the best method to select facial features (e.g. Davies and Christie 1982; Tanaka and Farah 1993; Tanaka and Sengco 1997). We perceive features in the context of other features. In fact, the presence of one feature can affect the appearance of another (e.g. Wells and Hryciw 1984; Yasuda 2005; Young, Hellawell, and Hay 1987), and face recognition seems to be based more on the whole face—that is, it is holistic in nature (e.g. Young and Bruce

2011). As a consequence, composite systems were developed to allow facial features to be selected in the context of an intact face: a more effective procedure (e.g. Davies and Christie 1982; Skelton et al. 2015; Tanaka and Farah 1993).

The result of these improvements led to the construction of more accurate faces, as discussed later in the chapter, but curiously led to a practical problem: There were now far too many example features available in a feature database that could be presented to an eyewitness in a whole-face context. In one of its databases, for example, PRO-fit has 173 sets of eyes, 169 eyebrows, and 176 noses. To overcome this issue, an eyewitness was asked to provide a description of the perpetrator's face, the result of which was used to identify suitable examples from the composite's database, to present to the eyewitness.

There are various methods that a description of a face could be obtained from an eyewitness, but arguably the best was to make use of a cognitive interview. The cognitive interview (CI) was developed in the late 1980s principally by US Psychologists Ronald Fisher and Edward Geiselman as an alternative to the question-and-answer format that was then common practice in policing (e.g. Fisher et al. 1989; Geiselman, Fisher, MacKinnon, and Holland 1985). The CI was designed to enable a person to recall as much accurate information as possible and, given the importance to police and other investigations, has been the subject of considerable research and revision (e.g. Wells and Hasel 2007). In essence, the CI contains a set of techniques, or *mnemonics*, designed to facilitate recall. Witnesses are invited to think back to the time of an incident and recall it freely, in as much detail as possible, without guessing. They may be asked to repeat recall, as taking a further extensive route through memory may trigger information not recalled at the first attempt, or to describe the event from another person's perspective or in a different temporal order (e.g. starting from the most recent part of the event). Applied to composite construction, eyewitnesses are asked to describe the face in such an uninterrupted format (Frowd 2012). It is also a usual practice to read back a witness's recall and ask whether further information can be remembered, as part of *cued* recall. For example, 'You mentioned the eyes were small and light in colour. Can you remember anything further?'

The process of composite construction now involves a two-stage process, recall of the perpetrator's face and construction of a composite using a composite system. Such a procedure should be carried out by a practitioner trained and experienced in cognitive interviewing as well as in the relevant composite system. We note that additional training is usually required as composites may require artistic work carried out on the face. For example, hair selected by a witness may require changes to its length or style or a logo added to a cap or hood. Curiously, one of the aims of creating composite kits was to allow police personnel without formal training in drawing techniques to create composite images with witnesses, but it is clear that such skills, as well as skills in interviewing techniques, are necessary and important for effective construction of composites (e.g. Davies et al. 1983; Fisher et al. 1989; Frowd 2012; Memon et al. 1997; Wells et al. 2007).

8.3.2.1 Mac-a-Mug pro

Mac-a-Mug Pro is a composite software, first introduced for the Apple Macintosh II computer in 1986 (Shaharazam 1986). The software programme contained an extensive feature set for that time, with hundreds of hairlines, eyebrows, noses, mouths, and accessories. These were all low-resolution raster images of sketch-like features. The most significant advance was that the user could edit and resize features, using some simple graphic editing tools, thus creating a synthetic face that was composed of different elements (see Fig. 8.8), but gave a convincing impression of an individual's face. Mac-a-Mug software proved popular with face recognition researchers, but it is not clear whether it was ever used in policing (a survey conducted in the late 1990s suggested not; see McQuiston-Surrett et al. 2006).

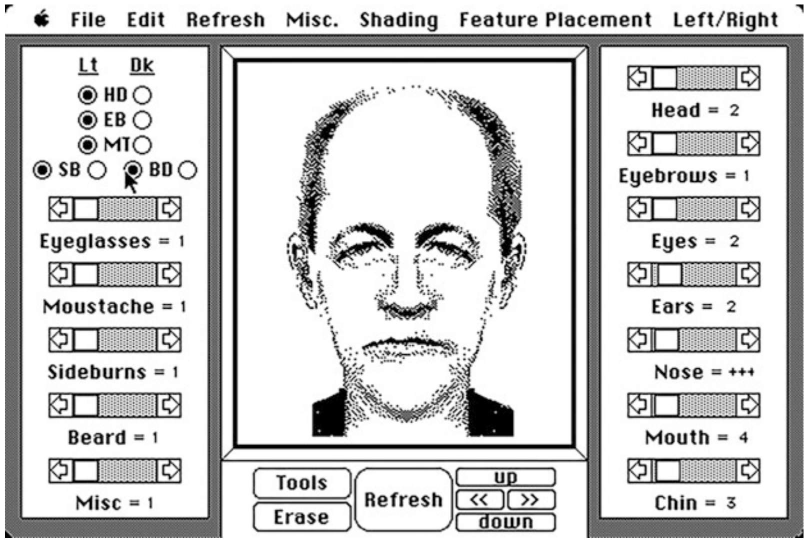

Fig. 8.8 Mac-a-Mug Pro software interface, circa 1986. See https://mac-a-mug.simplykiwi.com/mac-a-mug.html for an html emulator of the original software

8.3.2.2 Faces

The featural composition software FACES was released on Windows platforms in 1998 (Côté 1998) and became popular with police forces and researchers with version 3 (see McQuiston-Surrett et al. 2006, who report it being widely used in the USA in the late 1990s and early 2000s). Its creator, Pierre Côté, created it in a very similar manner to the way that Jacques Penry created Photo-Fit, from photographic images of face features but with the benefit of digital manipulation. The most recent version of the software, FACES 4.0, has an easy-to-understand graphical user interface, which makes it possible in principle for a novice to construct a face composite on their own after a relatively brief training period, but, as suggested above, better results are likely to be achieved with training and experience. The programme takes a featural approach to face construction, the operator choosing and assembling different facial features, in sequence, to construct a likeness of the perpetrator. FACES 4.0 includes hundreds of hairstyles, head shapes, sets of eyebrows, pairs of eyes, noses, lips, other face features, and accessories. All facial features are assembled on the right-hand side of the interface and are added to a whole-face composite on the left by clicking on them (see Fig. 8.9). Every feature can be manipulated in size, position, and colour. An important concept with this system is that, although an operator works with segmented features on the right-hand side of the screen, selecting a feature automatically places it in a whole-face context, as mentioned above.

8.3.2.3 E-FIT and PRO-Fit

E-FIT and PRO-fit are applications developed by different software companies from a common DOS programme, CD-FIT. Both systems have been involved in police investigations in the UK, Europe and the USA. As described below, research (Frowd et al. 2005a, b, 2007a) indicates that these two systems are equivalent in terms of their effectiveness, and so we consider them together here. Similar to the FACES system, they have a large database of photographic quality facial features classified and stored in a computer database. As standard practice, individual features are selected in the context of a complete face and usually blend acceptably on the face. Both make use of an artwork package to allow enhancements to be carried out under the direction of an eyewitness, a package that is integrated into PRO-fit but is a separate commercial application for E-FIT. As part of standard police practice in the UK (e.g. NPIA 2009), comprehensive training courses have been available since the 1990s, including training on photo editing and cognitive interviewing techniques.

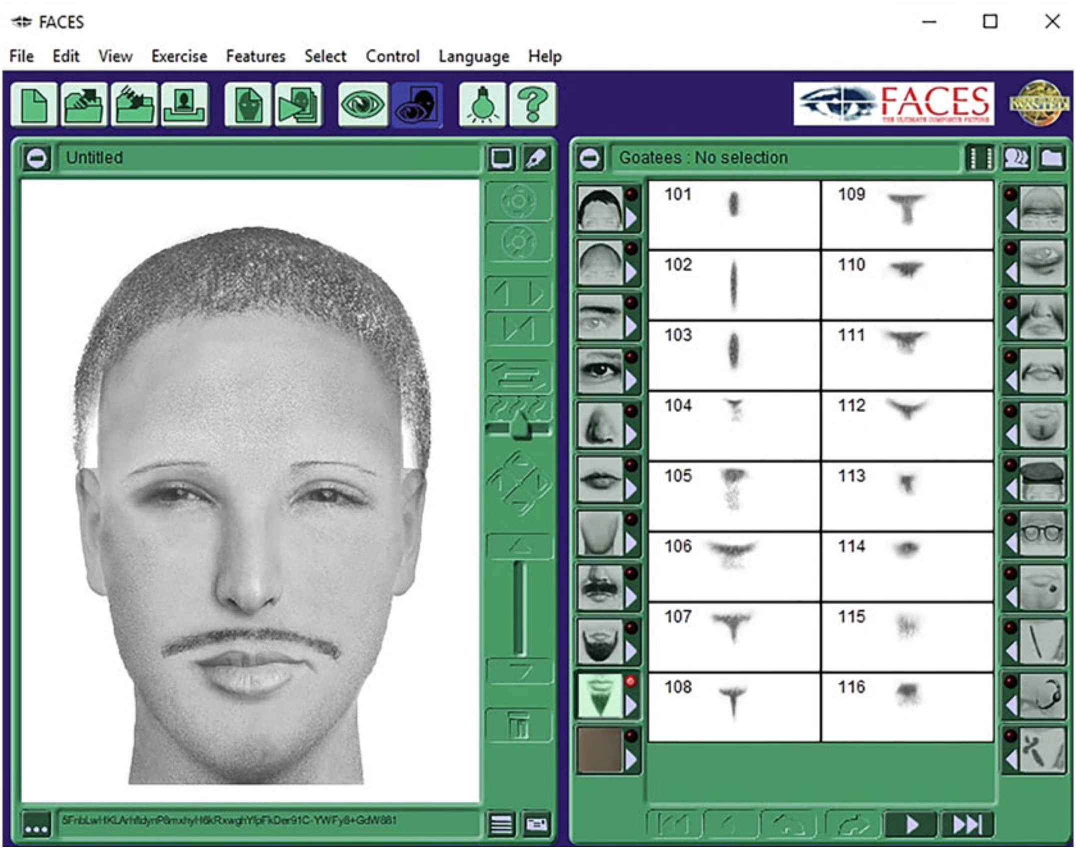

Fig. 8.9 FACES software graphical user interface (GUI), introduced in 1998. See https://facialcomposites.com/pages/faces-4-0-software-cloud-edition-demo for an html5 usable interactive version of the programme

8.3.3 Research Evaluation of Computerised Systems and Sketch

Assessing the effectiveness of composite systems, so that results are applicable to real-world use of these systems by police practitioners, requires considerable care and attention. There has been an impressive amount of excellent research conducted on mechanical systems (see Ellis and Shepherd 1996, for a review). The measures that were used assessed composite accuracy by asking participants to match composites to target pictures or rate composite likeness for accuracy. While these tasks indicate effectiveness to some extent, they do not capture the way that composites are deployed in the real world, where police and members of the public attempt to *name* composites. As such, important elements of the police procedure were developed into a gold standard research protocol (Frowd et al. 2005a, b).

This gold standard required two stages. In the first, ‘witnesses’ were recruited into a study and shown the face of a target identity for a fixed time (60 seconds in the original assessments). Importantly, as is usually the case in a police investigation, this face would be chosen to be unfamiliar to the witness. Having studied the face, a witness would later be interviewed by a trained interviewer. The interviewer would ask each witness to recall the face using a cognitive interview, as described above, and then guide the person through the process of constructing a composite.

Witnesses would be encouraged to create the best likeness possible in an open-ended session, and an artwork package was made available to make any necessary adjustments to the face. The resulting composites would then be shown to another group of participants, people who were selected to be familiar with the relevant identities, to name. Finally, after attempting to name composites, participants would be asked to name a photograph of the targets, as a check that they were familiar with the relevant identities—and so should be able to correctly name the composites. Statistical tests such as analysis of variance (ANOVA) would then assess differences between the systems on the average (mean) number of correct names given.

In the first assessment (Frowd et al. 2005a), ten photographs of celebrity faces were used as targets. These images were high-quality and should enable each person to form a very good memory of the face. Careful selection was carried out so that participants were paired randomly with a face that was unfamiliar to them and constructed a single composite of this face using one of five production systems. Systems were the mechanical Photo-Fit; the computerised E-FIT and PRO-fit systems; Sketch, with a highly experienced forensic artist drawing out the face by hand under the direction of the witness; and a recognition system called EvoFIT that was in development at the time, described later in the chapter. All interviews were conducted between 3 and 4 hours after a target face had been seen. Interviewers did not see any of the target faces shown to witnesses, followed the same basic procedure to administer a cognitive interview (including *free* and *cued* recall), and worked with witnesses to construct the best likeness possible in an open-ended face construction session.

The study revealed that people's ability to correctly name composites varied markedly by system. Composites from E-FIT were correctly named at 19.0%, PRO-fit at 18.0%, Sketch at 9.2%, and Photo-Fit at 6.2%; results from EvoFIT are discussed in the next section. The study thus indicates superiority for the computerised feature systems, when compared against the archaic Photo-Fit. Also, half of the targets had been selected to be distinctive in appearance, and results indicated that these identities led to composites that were correctly named more effectively overall than faces that were more average in appearance, a finding that clearly supports superiority of distinctive faces (e.g. Shapiro and Penrod 1986; Valentine 1991).

A subsequent assessment attempted to replicate results. The same basic design was followed except that the delay between a person ('witness') encoding a target and both describing this face and creating a composite of it was increased to 2 days, a more typical interval found in criminal investigations. A different set of target images was selected, again enabling interviewers to remain blind to their identities, and the FACES 3.0 computerised feature system was used in place of Photo-Fit. Average (mean) correct naming of composites was now woeful: E-FIT at 0%, PRO-fit at 1.3%, Sketch at 8.1%, and FACES at 3.2%. Sketch clearly emerged best, even though these composites were still not named effectively. Using a long retention interval of 1 or 2 days, other research has similarly found low levels of correct naming for these systems, typically in the region of 5%, as well as from the Identi-Kit 2000 computerised feature system (Fodarella et al. 2021; Frowd et al. 2007a, 2010a, 2015). Further assessments also indicate the general superiority of Sketch compared with computerised feature systems (Frowd et al. 2015).

8.3.3.1 Improving Computerised Systems and Sketch

Research has attempted to understand the reason why composites from computerised systems, in particular those constructed by witnesses after a long interval from seeing a target face, cannot be readily named. One issue concerns memory, in that total recall of facial features declines with time, measurable even after an hour, more so after a day and a week (Ellis, Shepherd, & Davies 1980). In fact, Frowd and Goodfellow (2017) demonstrate that the description used to locate facial features in the composite system is less effective after a delay, as is the resulting composite. The major impact, though, is that construction of the internal part of the face, the part that encompasses the region around the eyes, eyebrows, nose, and mouth, is simply less accurate (e.g. Frowd et al. 2011a, b). It has been known for a long time that we rely on these internal features when recognising a face with which we are familiar (e.g. Campbell et al. 1999; Ellis et al. 1979; Young et al. 1985). In contrast, for faces that are unfamiliar, we rely more on the exterior, external features—in particular hair and face shapes (e.g. Frowd et al. 2011a, b). Therefore, an eyewitness who has only seen a perpetrator's face once is likely to rely on the exterior features, while someone attempting to recognise the composite will tend to focus on the internal features.

Frowd and colleagues proposed that a sensible way to improve identifiability of a composite would be to improve construction of the internal features, so that recognition of the resulting image would be better. In fact, this approach has been remarkably successful, and we outline some methods in this section that have successfully improved identification of composites from computerised systems as well as from artists' sketches.

8.3.3.1.1 Combining Memories

One approach has explored the potential benefit of combining memories from different eyewitnesses. In situations where there is more than one witness to a crime, good practice is for policing to interview witnesses independently (e.g. PACE Code D 2017). It is well-known that human memory is subject to bias and error (e.g. Loftus 2003), but a careful comparison of multiple independent eyewitness accounts can provide confidence in the veracity of information that is consistent. Bruce et al. (2002) considered the situation where observers each viewed a target identity and independently constructed a single composite of the face. The researchers combined composites of the same identity together into a 'morphed' composite, so called because of the image manipulation technique used to create an average representation of the face. Assessment of the resulting composites revealed that the morphed composite was more effective than individual composites on average. In fact, the morphed image tended to be as effective as the best in the set and was sometimes better. Bruce et al. proposed that the superiority of the morphed image emerged as errors present in the individual composites were not related to each other and, through averaging, tended to cancel each other out, the result of which is a more accurate representation. The findings prompted a change in police guidelines in the UK: best practice is to interview eyewitnesses independently, with each person creating a separate composite that is combined to create a morphed composite, for publication in the media (NPIA 2009). Other research has found similar results (e.g. Davis et al. 2015; Frowd et al. 2006a, 2012b; Wells and Hasel 2007).

While this approach is sometimes used in a police investigation, it does of course require more than one observer (although see Ness 2001, for how this process could be adapted for a single observer). In the subsequent sections, we outline some of the techniques that have improved the effectiveness of composites when only a single observer has seen a perpetrator.

8.3.3.1.2 Recalling the Environmental Context

Considerable research has highlighted the importance of the environment (e.g. the room) in which encoding of information takes place. In a classic study, for example, recalling a list of words in the same physical environment as where they were seen, underwater or on the surface, led to more effective recall relative to when encoding and recall environment conditions were incongruent (Godden and Baddeley 1975). For unfamiliar faces, presenting cues that have been present at encoding facilitates recognition (Shapiro and Penrod 1986). For construction of faces, asking a person to recall the (physical and psychological) environment from memory enables a person to create a more accurate face using Photo-Fit (Davies and Milne 1985), while returning to the place in which a target had been encoded is also effective, albeit to a lesser extent. Recent research has used such a guided memory procedure for recalling environmental context using contemporary interviewing procedures and systems.

Fodarella et al. (2021) investigated the PRO-fit featural system and EvoFIT, an eigenface system described later in the chapter. Systems were assessed using the gold standard procedure including the use of a cognitive interview to obtain a free description of the face. Inviting a person to recall the physical (room) and psychological (internal) environment prior to face recall led to composites that were correctly named more often than either when recall of the environment was not requested or when the witness returned to the room in which the target identity had been seen.

Detailed memory recall of the context (DMRC) is an effective procedure to facilitate facial composites. This simple technique has also been applied to artists' sketches, improving their correct naming (Kuivaniemi-Smith et al. 2018). Fodarella et al. (2021) demonstrated that DMRC leads to greater recall of information about the face, a benefit that carries over to the ensuing construction of the face. It is perhaps worth mentioning that asking a witness to recall (write down) the appearance of a face between encoding and the face construction interview itself (compared to not recalling the face in the intervening period) is particularly effective at facilitating composites from featural, sketch, and eigenface systems (Brown et al. n.d.; Portch et al. 2017a).

In the following sections, we assess two further mnemonics that have also been successful in facilitating face construction.

8.3.3.1.3 Facilitating Holistic Face Processing

Face construction principally involves two cognitive processes, face recall and face recognition. Techniques of the cognitive interview, including the DMRC mnemonic, improve the effectiveness of composites by facilitating the former. Face recognition, though, is a holistic procedure and tends not to be augmented by techniques that improve recall (e.g. Vredeveldt et al. 2015; Wells and Hryciw 1984). In fact, such techniques may even hinder recognition (e.g. Alogna et al. 2014; Frowd and Fields 2011; Schooler and Engstler-Schooler 1990). Instead, facilitating a witness's face recognition should help him or her to make more accurate selection of facial features. This was achieved by asking witnesses to focus on the character or personality of a face, implemented as a series of 'holistic' mnemonics to the cognitive interview.

In Frowd et al. (2008a), witnesses were asked to think back to the time when they had seen an unfamiliar target and then to freely describe it. They then were asked to think silently about the character or personality of this face. After 1 minute, witnesses were given seven characteristics to rate for the target face on a three-point scale, *low*, *medium*, and *high*. Characteristics included intelligence, friendliness, and arrogance. As illustrated in Fig. 8.10, composites constructed using the PRO-fit feature system were correctly named at a higher rate following the use of these holistic mnemonics. Their advantage to feature-based construction has been replicated many times (e.g. Frowd et al. 2015; Skelton et al. 2019) and found to extend to face construction using an eigenface system (e.g. Frowd et al. 2012b; Frowd et al. 2013). Pilot work has also indicated how these mnemonics can be used to facilitate production of forensic sketches (Frowd et al. 2015).

8.3.3.1.4 Increasing the Focus on Internal Facial Features

Skelton et al. (2019) argued that de-emphasising, or even masking, external features should facilitate construction of internal features. This approach was inspired by its effectiveness for the EvoFIT eigenface system, as mentioned later. It turned out, however, that worse composites ensued when the approach was applied to a computerised featural system. The reason seems to be that recalling the appearance of the face activates memories of facial features, internal and external, but viewing only the internal region for feature selection produces an incongruent context (as external features were not shown on this face to create a composite). However, Skelton et al. found that composites were greatly facilitated when holistic recall mnemonics were used, as these led to a focus of attention on the internal features. In the project, once internal

Fig. 8.10 Composites created in Frowd et al. (2008b) of EastEnders's character Billy Mitchell (played by actor Perry Fenwick) following a standard type of face recall cognitive interview (left) and following this interview plus mnemonics for holistic recall (right). In the study, two witnesses watched a video of the character and, after the relevant cognitive interview, constructed a single composite of the actor using a computerised feature system

features had been constructed, witnesses viewed and then constructed external features. Also, demonstrating how to produce a more effective featural composite, Skelton et al. highlight the importance of consistent processing of information from one stage of face construction to the next, an important principle that we return to later.

8.3.3.1.5 Overcoming Changes of Appearance

Perpetrators attempt to evade detection in various ways. One way is to change their appearance, perhaps by wearing a hat, glasses, and /or hooded top, but equally they may wear a balaclava and may even change their hair or facial hair. There is good evidence that changes in appearance interfere with face recognition (e.g. Henderson et al. 2001; Righi, Peissig, and Tarr 2012; Sporer 1993). Brown et al. (2018) considered the impact of a change of hair along with selective concealment of a composite, the latter based on an advantage previously observed when adding sunglasses to a composite (McIntyre et al. 2010). As predicted, they found that composites were markedly less identifiable when a target face had been seen—unknown to the witness—with a change of hair. Brown et al. tested a new 'array' format for recognition, one that included the original composite along with three depictions presenting the composite wearing a hat, sunglasses, or both of these adornments. When a witness had seen a target with a change of hair, the array format produced equivalent naming to a composite constructed from a target with veridical hair. Therefore, concealing the part of a composite that had changed was advantageous.

A further result is worth mentioning. Brown et al. (2018) tested a novel procedure for naming, one in which a single composite was presented first and then the array format. When a target had been encoded with changed hair, the array improved identification, as mentioned above. However, for veridical hair, correct naming increased significantly between the first and the second presentation. Therefore, the research demonstrated a practical technique for improving recognition of a finished (completed) composite that is not only advantageous when a perpetrator has changed his or her hair but also in general for composites.

In the following section, we outline other techniques that have attempted to improve recognition of finished composites, several of which are in current police use.

8.3.3.1.6 Facilitating Recognition of Finished Composites

Composite images from all systems may be inaccurate in terms of the shape of facial features, the colouring of features, the appearance of the skin, and the spacing between features (so-called relational information, e.g. Diamond and Carey 1986). Helping police and members of the public to recognise these 'finished' composites in the presence of error is of course advantageous for criminal investigations. One successful approach was mentioned earlier: combining individual composites to create a 'morphed' (average) face. When only one composite is available, another approach has been to present a composite that dynamically transforms by facial caricature (Frowd et al. 2007c, 2012b). It is well known that artists capture facial characteristics, sometimes using only a few lines on the page, by exaggerating distinctive aspects of the face. Here, researchers used a computer technique that progressively changed the shape information in the composite face, exaggerating and then de-emphasising distinctive aspects of the face. This 'dynamic' caricature technique is illustrated in Fig. 8.11. It is very effective at triggering recognition, since a different frame in the sequence serves as a preferable cue to identity than the original composite. The technique has been shown to be effective for all types of composite, featural, sketch, and eigenface. Several composite systems have caricaturing now built in, to allow production of a dynamic caricature. The technique has been used in criminal investigations (see Frowd et al. 2012a, and *Case Study* in the section later on *EvoFIT*).

A second approach has been to present a composite as a linear stretch. Such a representation is thought to make it more difficult to perceive facial features (as they are stretched), and when a person normalises the face, the result is a less error-prone, more identifiable representation; as an alternative, turning a piece of paper containing a printed composite to the side, a 'perceptual' stretch, so that the face appears long and thin, similarly enhances recognition, as does presenting the face against a perceptual 'backdrop', as illustrated in Fig. 8.12 (Frowd et al. 2013; Frowd et al. 2014; Skelton et al. 2019).

Several other techniques in addition to those mentioned above (Brown et al. 2018; McIntyre et al. 2010) have similarly facilitated recognition by (i) horizontally misaligning the face (McIntyre et al. 2016), (ii) reducing the level of a composite's texture (Frowd et al. 2008b), and (iii) adding characteristic (person-specific) motion to the face (Lander et al. 2017).

8.3.3.1.7 Combining Techniques

Before considering more modern approaches to composite construction, it is perhaps worth mentioning that several projects have considered combining some of the aforementioned techniques, to promote a more effective composite. For example, averaging a face to create a 'morphed' composite and then presenting it as a dynamic caricature are more effective than either averaging or caricaturing alone (Frowd et al. 2012a). Skelton et al. (2019) assessed PRO-fit composites constructed following holistic recall mnemonics and masking of internal features; subsequent composites were named with the face front-on or using the perceptual stretch technique. Independent benefits were observed for each technique, with very accurate correct naming emerging when all three techniques were used in conjunction with each other. The research indicates that very effective composites can be produced so long as witnesses have a good memory of the face. The research also indicates that combining techniques is a fruitful pursuit, one which we apply to the production of eigenface systems discussed in the next section.

8.4 Solutions in the Twenty-First Century

8.4.1 Eigenface Theory

We have reviewed research in this chapter that found featural composite systems to produce poor-quality representations of faces, especially

Fig. 8.11 Example frames in a dynamic caricature sequence. In this example, a composite from the EvoFIT eigenface system has been caricatured at −60%, −30%, 0%, +30%, and + 60%. While positive values exaggerate distinctive shape aspects of the face, negative values de-emphasise them. The face in the middle (0%) is the original, constructed composite. The sequence is usually seen to change smoothly as an animated GIF over a 21-frame sequence that repeats every 6 seconds

Fig. 8.12 An example composite presented in the perceptual backdrop image format

when created from long-term memory but often even when created while the face was in full view of research participants. Tredoux (2002) speculated that the underlying problem with existing face composite systems at that time may have been the lack of a representational or computational theory driving the reconstructions. As mentioned above, Frowd et al. (2007b) suggested that the problem related to focus on the exterior, external features. Valentine 1991 had advocated a multidimensional model of perceptual face representation for several years prior to that but had not provided a computational basis for underpinning his model. However, engineers had at roughly the same time been using a computational basis for representing faces (without a perceptual theory), and this computational basis could be used to create a powerful face construction system, as was later shown by Tredoux et al. (1999, 2006), Hancock (2000), Frowd et al. (2003), and Gibson et al. (2003). The mathematics underlying the computational basis is fully described by Sirovich and Kirby (1987), and an important extension is the so-called 'active appearance model' (Cootes et al. 1998).

A brief, non-mathematical summary is as follows: A set of face images is digitised and standardised in terms of position, average light intensity, and size. Certain landmarks are placed on each face (points marking head shape, edges of the eyes, nose, and other facial areas) allowing separation of the face shape from its texture information. The normalised data for shape and texture are then subjected to a principal component analysis (PCA), separately. PCA produces a set of basis vectors, or eigenfaces, for both shape and texture. These are in turn melded into an appearance model, but texture and shape remain separately manipulable in the model. The appearance model forms a representational basis for the original set of faces. This basis has many useful properties. For example, the Euclidean distance between two faces in the space formed by the eigenfaces can be used as an index of their physical and perceptual similarity (e.g. Hancock, Burton, and Bruce Bruce 1996; Tredoux 2002). However, the most important consequence of the computational basis is that faces that do not appear in the original set, but come from a population of similar faces, can be reconstructed with a high degree of accuracy. A linear combination of the eigenfaces, using the coordinates as weights, produces the reconstruction. The population of faces used to generate the eigenfaces obviously imposes limits on the faces that can be successfully projected into the 'face space'.

As Figs. 8.13 and 8.14 show, the eigenface composition of faces is a function of the coefficients that create weighted combinations of eigenfaces. The problem in creating faces from memory in interaction with witnesses thus becomes one of finding a suitable set of coefficients. The research groups that have created composite systems with eigenface technology have usually relied on the so-called genetic search algorithms (see Johnston and Franklin 1993, for an early example) or on hybrid genetic search and competitive learning algorithms (Tredoux et al. 1999) or bespoke algorithms (e.g. MChoice; Tredoux et al. 2006). All the algorithms we know of are essentially examples of optimisation search. A typical user-controlled search in an eigenface system proceeds as follows: The system generates a random sample of faces, for example, 12, and these are shown to the user in a 4 x 3 array (an example is shown as Fig. 8.15). The user selects one or more faces that resemble the target face in their memory. The system aggregates the coefficients selected in the previous step and uses them to generate a new array of faces, to be shown to the user. This process continues until the user is satisfied that the resemblance is good enough or is not improving. The number of generations required to obtain a satisfactory result appears to depend on many factors, including the sensitivity of the user's search criterion and the strength of the memory representation. Tredoux et al. (2006) ran a simulation where they replaced the user input at each generation with an automatic random selection of faces among those with the lowest 50% distance to the target face and showed that the system could produce a good likeness within 15 to 25 generations when using their bespoke algorithm, MChoice.

There have been a number of technical improvements to eigenface systems in the past two decades. Perhaps the most widely used

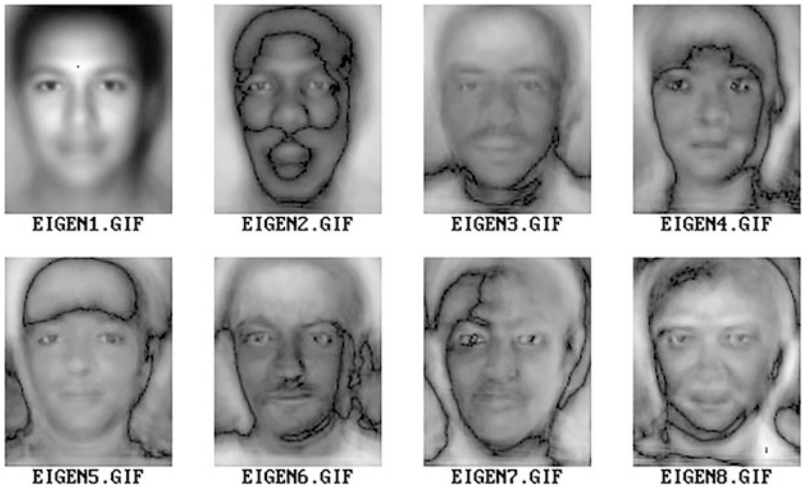

Fig. 8.13 The average face (Eigen1.gif) and a representation of the first seven eigenfaces from analysis of a set of 520 South African faces (see Tredoux 2002)

Fig. 8.14 A synthetic face created through eigenface composition: 0.1*Eigen2 + 0.3*Eigen3+ 0.3*Eigen6+ 0.3*Eigen1

improvement or alternative is the so-called 3D Morphable Face Models of a group at Basel University (Blanz and Vetter 2003).

8.4.2 Systems

8.4.2.1 Id

Tredoux et al. (1999, 2006) developed ID, an eigenface-based system, and have made it publicly available since then. There are presently ten face models for use with ID: (A) models of White, Black, coloured, and Indian South African faces, both male and female, of varying ages, in which the base data for each model are about 300 to 350 images; (B) models of Egyptian, Maghreban, and French European populations, which are smaller, with about 100 base images each; and (C) a small-scale model of Dutch Europeans, which allows manipulation of emotional expressions.

Settings in ID allow configuration of parameters such as size of face thumbnails, number of faces in each generation of synthetic images, and several relating to the search algorithms. ID is capable of generating entire faces or of modifying particular facial features on a previously generated face, using local eigenfaces and local coefficient optimisation. The main interface used by the witness is shown in Fig. 8.15, and an example of a good quality synthetic face made with ID is shown in Fig. 8.16. ID also provides some tools to manipulate the current generation, including increasing or decreasing similarity of faces in the current generation and searching in the space for holistic properties of faces, illustrated in Figs. 8.17 and 8.18.

8.4.2.2 EvoFIT

EvoFIT has been developed for over 20 years (Frowd et al. 2004a, b). However, the first version that was sufficiently effective for use in police investigations, producing composites with correct

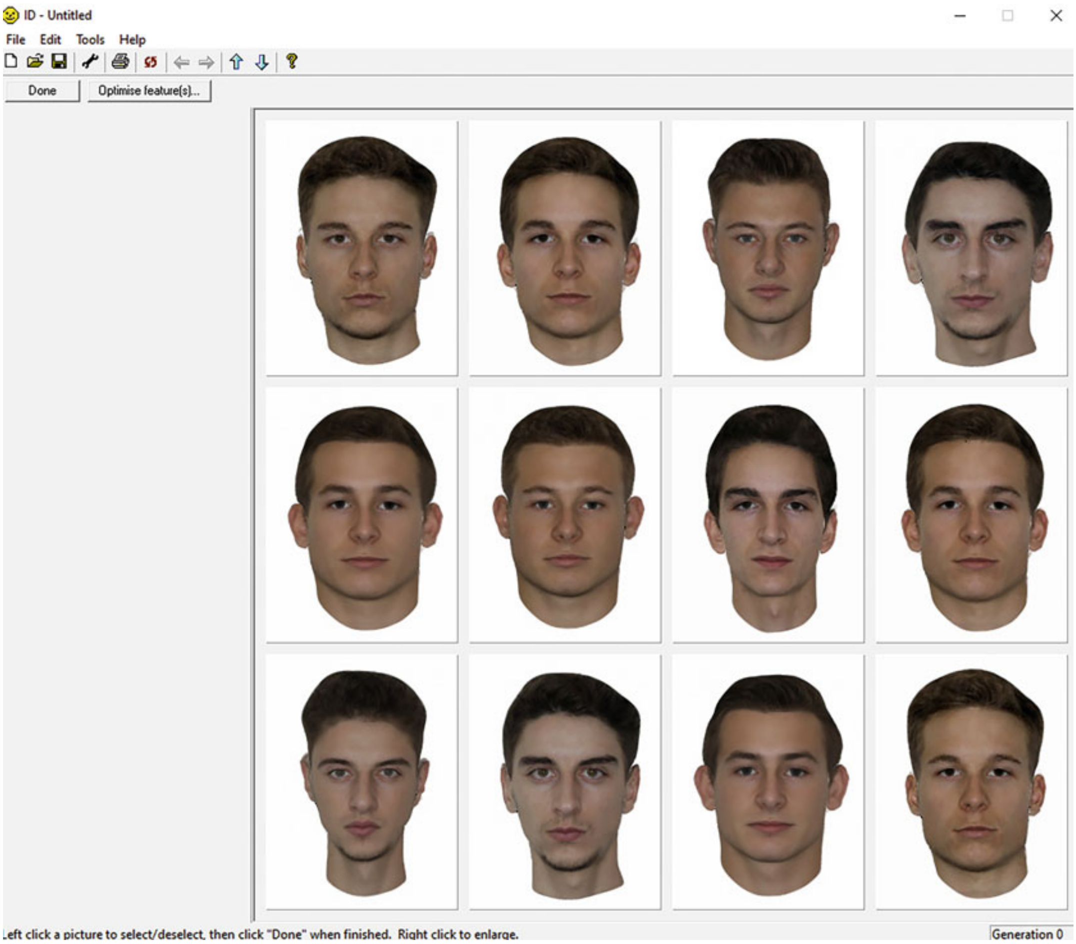

Fig. 8.15 GUI controlling face synthesis in ID

naming of 25% after a long retention interval, emerged in 2007 (Frowd et al. 2008b, 2010a). The procedure to evolve a face is detailed (for a technical review, see Frowd et al. 2004a, b), and we present a summary here of this basic version (system enhancements are discussed subsequently).

Similar to faces in ID, faces in EvoFIT are synthesised from a separate principal component analysis (PCA) of facial shape and texture information from training sets of around 70 reference faces classified by age, race, and gender. This system deliberately presents facial images in greyscale based on evidence that colour is not valuable for familiar face recognition under normal circumstances (Frowd et al. 2006c). Initially, suitably scaled random weights were used to combine eigenfaces in the PCA shape model, to create a random shape vector, and eigenfaces in the PCA texture model, to create a random texture vector. Information from the shape vector was then applied to a smooth (averaged) face to give a single face that changed by feature shape and placement; information from the texture vector gave a plausible appearance for shade of eyes, eyebrows, skin tone, etc. At the start, eyewitnesses were asked to select suitably matching hair and other external features that were combined with the randomly generated facial texture. See Fig. 8.19, for example, images used to synthesise a face. This process was repeated to give a screen of 18 smooth faces and 18 textured faces to present to eyewitnesses.

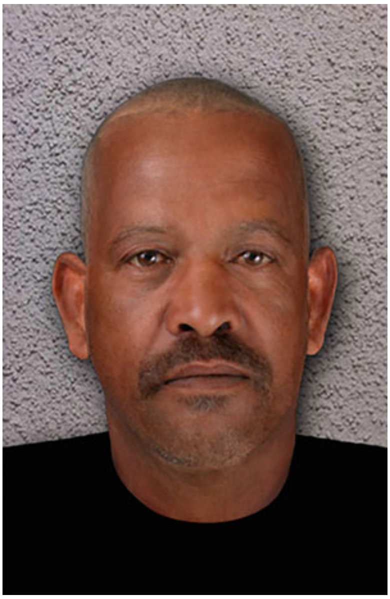

Fig. 8.16 Example of good quality synthetic face generation in ID

In the first pass through the system, Generation 1, eyewitnesses selected faces from a total of 12 such screens. First, eyewitnesses were shown four screens of smooth faces, selecting the two best-matching items on the first three screens, to a total of six items, and making any alternative choices on the fourth. The six selected smooth faces were presented and the eyewitness nominated the best overall match. Next, four screens of facial texture were superimposed on this best-matching shape, and eyewitnesses selected in the same way as before (for smooth faces). The result was six smooth and six textured faces. Combinations of these items were then presented for an eyewitness to select the best overall match.

Faces were presented again to eyewitnesses to select by smooth, textured, and a combination of both. This time, faces were synthesised not by using random values, but by combining parameters from faces that had been selected in the previous generation. This process involved randomly selecting a pair of faces, combining their coefficients to create another set, with half selected randomly from one face and half from the other. In this procedure, the 'best' face was carried forward unaltered into the next generation as part of an *elitist* approach; it was also given twice the number of breeding opportunities, the aim of which was to facilitate conversion around this presumably more effective face. To maintain population variance, a small percentage of coefficients was subject to mutation, one which replaced a coefficient with a suitably scaled random value.

At the end of Generation 2, software tools were used to improve the match of this 'evolved' face. This involved a set of holistic tools (Frowd et al. 2006b) that changed age, weight, health, and seven other such global scales. Each time, an eyewitness identified the best-matching change or left the face unaltered; any change made was carried forward to the next scale. Then, the face was enhanced by altering the greyscale colour of the eyes, eyebrows, mouth, etc. before enhancing the face further in a 'shape' tool which altered the shape and placement of individual features on the face. An artwork package and a 'warp' tool (especially for enhancing hair) were made available to make further adjustments to the face.

This procedure was found to be effective only some of the time. However, it was apparent that eyewitnesses were distracted by the presence of hair and other external features, and so their

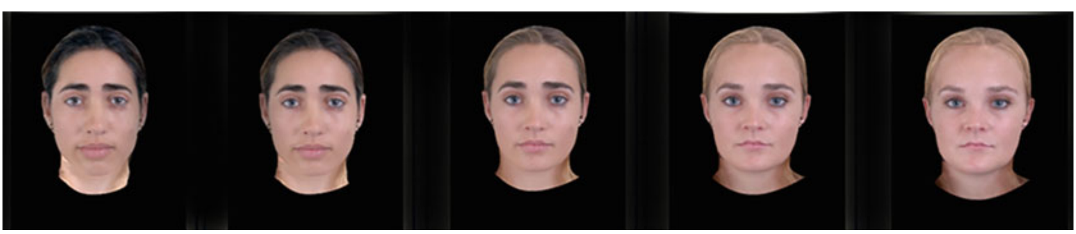

Fig. 8.17 Manipulation of perceptual dimensions in ID (left to right, Southern to Northern European)

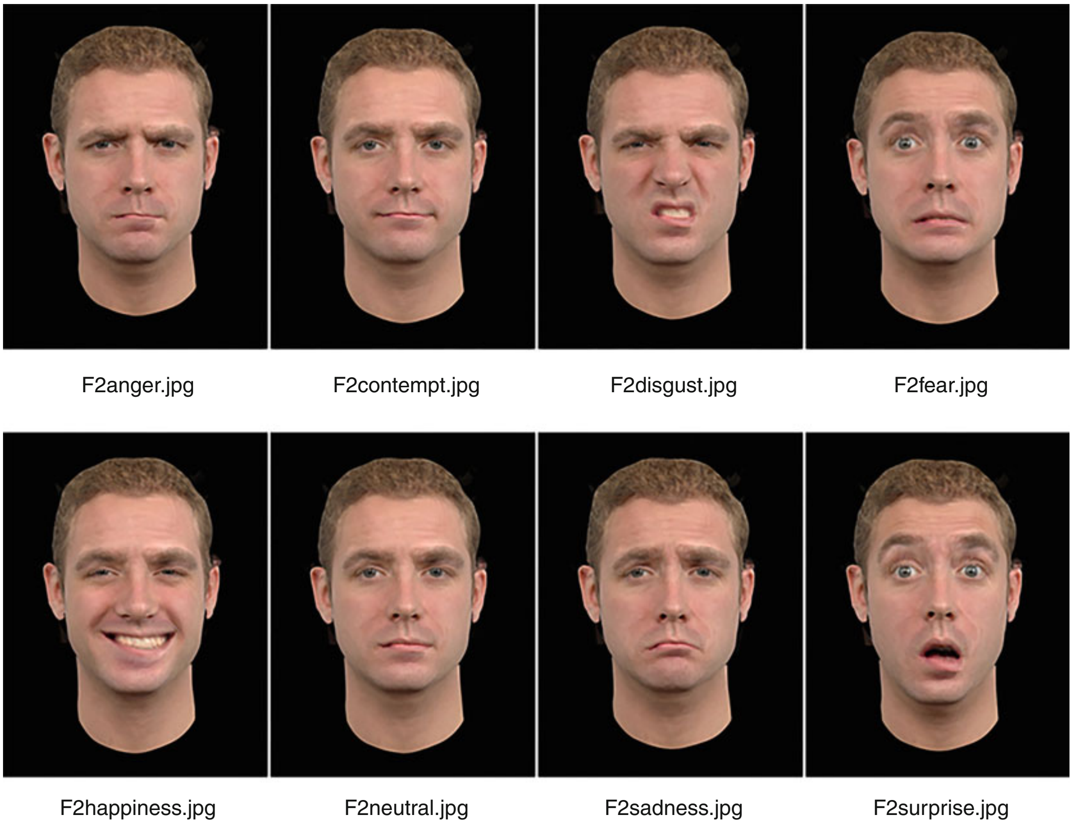

Fig. 8.18 Eigenemotions: generation of emotional expressions in ID

presence was de-emphasised by blurring them using a Gaussian filter. Gaussian blur was applied when witnesses selected hair and external features at the start and disabled at the end of the process prior to saving the face to disk. An example array using this de-emphasising procedure was shown previously in Fig. 8.19. Using the gold standard procedure, this version of EvoFIT produced composites that were named by other people on average at 25% correct (Frowd et al. 2010b), as detailed below.

8.4.2.2.1 Police Assessment

It is estimated that EvoFIT has been deployed in over 5000 criminal investigations, mainly in criminal investigations of serious crime (e.g. rape, aggravated burglary) in the last 10 years (Frowd et al. 2019; REF 2014). There have been several assessments of performance of this system, with the main measure usually whether a composite produced by a witness or victim directly led to identification of a suspect (e.g. Frowd et al. 2010a, 2011c, 2012a). These assessments indicate that identification has increased substantially over time, reflecting subsequent development (see section on *enhancements to eigenface technologies*). The latest assessment indicates an identification rate of 60%; of these cases, 29% led to a guilty verdict (i.e. provided an outcome based on suitably supportive additional evidence).

8.4.2.2.2 Case Study

There have been many cases where an EvoFIT composite produced by a witness to (or a victim of) crime has directly led to the identification of a perpetrator (e.g. see Frowd et al. 2010a, 2011c, 2012a). We describe one such case, as part of Operation Hatton, that has been reported widely in the media. In this example, Greater Manchester

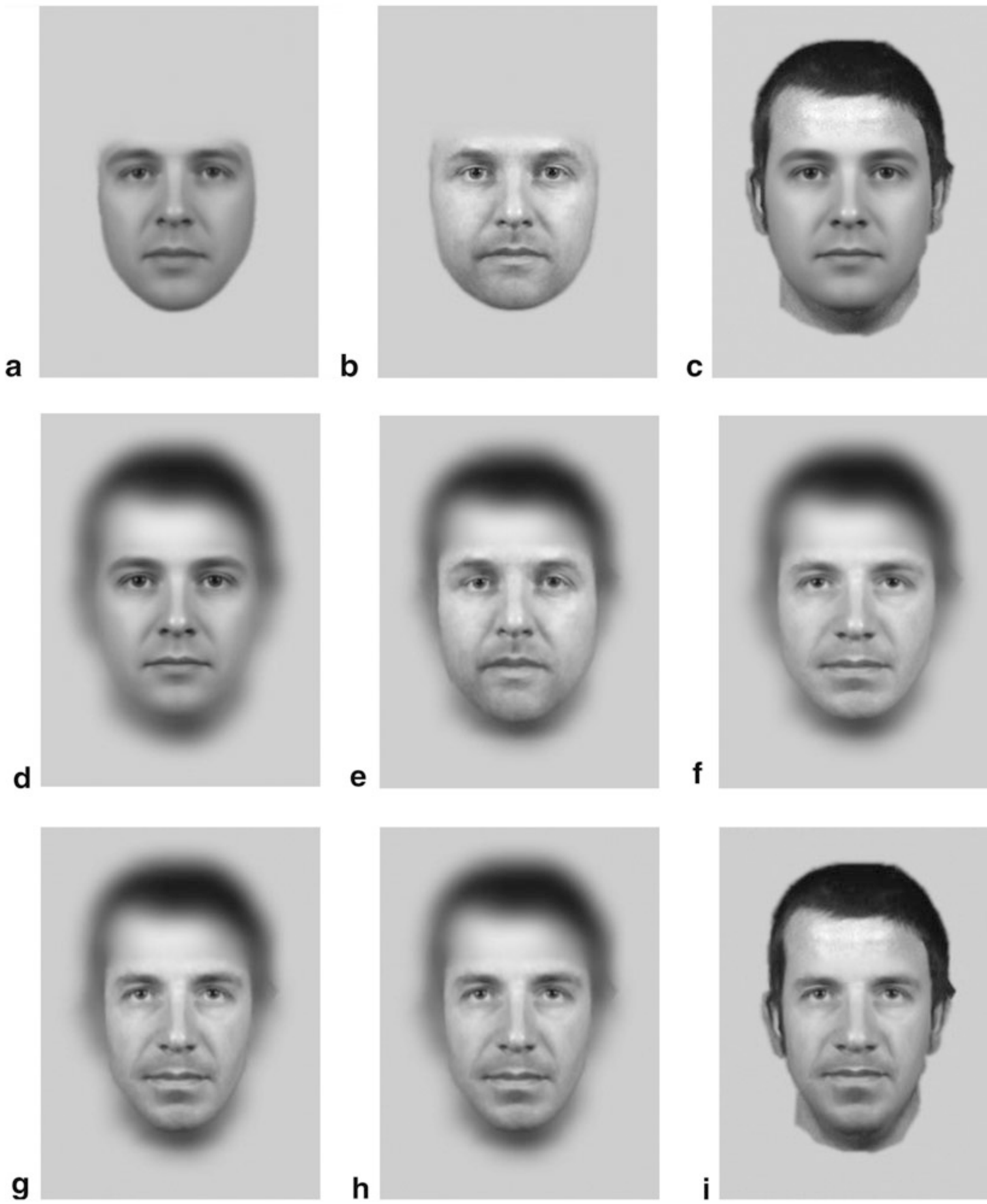

Fig. 8.19 Stages in the creation of an EvoFIT facial composite. Image (**a**) example smooth face, (**b**) example textured face, (**c**) selected external features, (**d**) selected external features with Gaussian blur applied, (**e**) combining smooth and textured face with blurred external features, (**f**) 'evolved' face, (**g**) face after holistic tool use, (**h**) face after shape changes, and (**i**) final image

Police were attempting to locate the whereabouts of a serial rapist (Frowd et al. 2012a). The crimes took place around 2009 and 2010. The young male adult in question would stalk his female victims, attacking them in the south part of Manchester in the early hours of the morning. His behaviour was callous, in one case using a weapon in the attack and in another phoning the victim's father to say that his daughter was safe and that he was bringing her home.

In the most recently known crime, an EvoFIT composite was constructed of the offender 7 days after the event. The victim described the face to a forensic practitioner and proceeded through the various screens to construct an EvoFIT composite of her attacker. The procedure used to construct the face is as described above except that faces were presented as internal features, rather than with blurred external features, to reflect the improved protocol that had recently been developed (see Fig. 8.20 and *Enhancements to Eigenface Technologies*).

The constructed composite (Fig. 8.21) was released into the media as part of a public appeal for information. The dynamic caricature technique was also used with this composite (http://tiny.cc/animated-composite-1). One name was consistently given to police, Asim Javed. He was a local person who worked in a fast food outlet in the Chorlton area of Manchester, presumably known by many people in the community. He was arrested as a suspect in the investigation and interviewed. During the interview, he confessed to the attacks, stating that he would have continued offending had he not been stopped. He received an indefinite prison sentence.

8.4.2.3 EFIT-v

EFIT-V, previously called EigenFIT, follows a broadly similar procedure to the eigenface systems to construct a face. There are several detailed descriptions for how this system operates (e.g. Gibson et al. 2003; Gibson et al. 2009). Here, we outline the procedure to construct a face that is detailed in Valentine et al. (2010), as this procedure provides a measured level of performance. In this case, individual composites were constructed from immediate memory and were named by other people, on average, at 20% correct.

The starting point for composite construction is selection of a facial database and hairstyle. The eyewitness is then presented with a screen of nine synthesised (randomly generated) faces each containing the selected hairstyle. Similar to the ID eigenface system, individual faces are presented that change by both shape and texture information (cf. EvoFIT), although for E-FIT-V faces only render texture as faint information. As is normal for this type of system, none of the faces usually bear a strong resemblance to the intended person. The eyewitness is asked to select the best match, and a genetic algorithm is used to create another screen of nine faces. This process is repeated, with faces on successive screens becoming more similar to each other and (ideally) becoming more similar to the intended face.

The system has various options that allow a flexible approach to creating a face. For example, two faces on the screen can be blended together; an array of faces can be rejected, allowing a new set of faces to be shown; and facial features can be 'frozen', that is, keeping a selected feature

Fig. 8.20 Example screenshot of smooth faces (left) and textured faces (right) with external features presented with a moderate level of Gaussian blur

Fig. 8.21 The EvoFIT composite created by a victim 7 days after the attack in Operation Hotton. For copyright reasons, we are not able to reproduce a photograph of the person convicted in this case, but a search of the internet (e.g. using terms *Operation Hatton Asim Javed EvoFIT*) should reveal news stories that include the composite shown here along with a photograph of the person convicted

unaltered on successive screens. There are also software tools that allow adjustment of the face similar to more traditional feature systems. Here, individual facial features can, for example, be manipulated by height and width. The age of the face can also be adjusted and the face caricatured.

There is a more recent version of the system, E-FIT-6, which we understand allows an even more flexible approach to face construction.

8.4.3 Research Evaluations of Eigenface Systems

There has been a series of formal assessments of eigenface systems that have involved the gold standard procedure. Frowd et al. (2010a) assessed the version of EvoFIT described above compared to a standard computerised featural system, PRO-fit. Eyewitnesses encoded a single face for 1 minute and, 2 days later, described the face using a cognitive interview and constructed a single composite face using EvoFIT or PRO-fit, as per random assignment. The resulting composites were named by other people at 24.5% correct for EvoFIT and 4.2% correct for PRO-fit. The study also revealed that EvoFIT composites created without blurring hair and external features were poorly named, demonstrating the advantage of this technique for face selection, and that correct naming was facilitated following the use of the holistic scales (e.g. to adjust age, weight, health, and other global properties of the face). Police were also given this version of EvoFIT to assess. Lancashire and Derbyshire constabularies reported that its composites created by witnesses and victims directly led to an arrest in 20% of cases (Frowd et al. 2011b), comparable with the formal assessment in the laboratory.

EFIT-V has been assessed by Valentine et al. (2010). This assessment largely followed the gold standard procedure except that each eyewitness constructed four faces with a target 'in view' during construction and four faces from memory. With respect to the latter, the procedure did not involve a long retention interval, with the four individual composites constructed immediately after a video of a target had been seen by the eyewitnesses. Construction of the faces was guided by the experimenter and used recommended operating procedures to evolve the face. This included, where appropriate, blending faces together, rejecting an entire array, 'freezing' (fixing the appearance of) individual facial features, manipulating the width or height of individual features, and adjusting the apparent age of the face.

Valentine et al. (2010) found that average naming of individual composites was measured at 20% correct. The researchers also found that correct naming increased when participants viewed a 'morphed' composite—that is, an image averaged together from four individual composites—replicating the advantage of a four-item morph found by Bruce et al. (2002). (Note that Valentine et al. also found that morphed composites were more effective when combined from different eyewitnesses than from one eyewitness, as one might anticipate.) While these results seem positive, system performance would appear to be comparable to computerised feature systems assessed after a short retention interval (Frowd et al. 2005a): presumably correct naming would be much lower following a longer forensically relevant delay. Such a suggestion is consistent with subsequent research indicating that rated likeness of these eigenface composites reduces with increasing retention interval (Davis et al. 2011).

Evaluations of the ID composite programme are reported in Tredoux et al. (1999), Prag (2005), Tredoux et al. (2006), Sullivan (2007), and Schmidt (2010). Since the ID programme is used mostly for purposes of synthesising faces for use in face and person recognition experiments, there is less formative research on its effectiveness than for EvoFIT or E-FIT. In the earliest evaluation, of a prototype ID programme built with MATLAB, simulated witnesses were able to build composites of three target faces in full view that were recognised from seven-person line-ups by independent judges 51% of the time (but up to 70% of the time for one face in particular). When building a composite of a celebrity face from memory, simulated witnesses were able to construct composites that were recognised 58% of the time from line-ups, but when the face was of a person previously unknown, recognition dropped to 17%. A later version of the ID programme was evaluated by Prag (2005; it was known then as ‘E-Face’) and was again tested in ‘in-view’ and ‘from-memory’ conditions, intended to simulate strong versus weak memory encoding. In-view constructions were rated for similarity to the target, and four foils, by independent judges. When these were thresholded, the ratings translated to an identification rate of approximately 51%. A different ‘mugshot reduction’ task suggested a lower identification rate of approximately 30%, but the task was quite different in nature, being intended to simulate the reduction in search space for a composite.

Tredoux et al. (2006), drawing on Prag’s study, asked witnesses to a simulated cell phone theft to construct faces with either FACES 4 or ID. When made with the target face in view, FACES 4 composites were better recognised than ID composites, but when made from memory, ID composites were better recognised than FACES 4 composites (this latter difference was evident when using a standard interview method but disappeared when using a cognitive interview). Sullivan (2007) exposed witnesses to a simulated fraudulent encounter with a lengthy 20-minute encoding and a 5-week delay. They compared composites created with FACES 4 and ID, under different interviewing conditions. Witnesses were asked to construct composites of the target face after a guided memory interview (Malpass and Devine 1981) or a standard police interview. There were also some additional comparison conditions: Composites made by different witnesses of the same target were morphed, and an expert operator tried to reconstruct the target faces while in view, as did additional participants who had not seen the target. Subsequent accuracy on a six-person line-up was 68% for the in-view expert operator, 53% for the in-view non-expert participants, 25% for the 5-week delayed guided memory constructions, 21% for the 5-week delayed standard interview condition, and 21% for the morphing condition. Recognition of composites created with FACES and ID showed complex interactions with perpetrator identity, with ID composites resulting in better recognition than FACES composites for some perpetrators, and vice versa for other perpetrators. Generally, the ‘from-memory’ composites were more recognisable when using the ID system, and the ‘in-view’ composites were more recognisable when using the FACES system.

In contrast, Schmidt (2010) found slight advantages for the composite system Identi-Kit 2000 over ID in a likeness rating task but equivalent performance in a line-up identification task. However, ID appeared to promote fewer correct choices in target-present line-ups and more correct choices in target-absent line-ups.

8.5 The Cutting Edge

8.5.1 Enhancements to Eigenface Technologies

There have been several notable improvements to eigenface systems. Based on published research, the focus in this section is on the EvoFIT system, but there does not seem to be any real reason why these findings should not also apply to other eigenface technologies. Indeed, some of the enhancements to the ID system have been described above, but since the development of ID has been less thoroughly evaluated empirically than the development of EvoFIT, we leave them as described above. Note that this section considers EvoFIT as described in the previous section as the basic system, here outlining important developments that have been assessed using the gold standard procedure. The same as for the computerised feature systems, the approach has been to assess techniques that attempt to (i) improve the accuracy with which eyewitnesses construct a composite's internal facial features and (ii) render a composite's internal features more identifiable. Several of the successful techniques for these computerised featural systems—in particular, the holistic cognitive interview, the focus on internal features, and the use of dynamic caricature—also apply to EvoFIT (and presumably apply to other eigenface systems) and have been the result of co-development.

8.5.1.1 Facilitating Holistic Face Processing

One of the most important procedures to improve identification of EvoFIT composites has been application of the holistic mnemonics for the face recall cognitive interview. In a study in which these mnemonics were administered by practitioners to eyewitnesses in the same way as for feature systems, Frowd et al. (2012a) observed an increase in naming accuracy from 21% without these mnemonics to 39% when they were administered.

8.5.1.2 Increasing the Focus on Internal Facial Features

Research in parallel with this work was assessing the appropriate level of Gaussian blur to apply to the external features in face arrays, for eyewitnesses to select best matches. The original design included a moderate level of blur, based on the rationale that face selection would be valuable in the context of some degree of external features. However, Frowd et al. (2012a) found that best performance was observed when infinite blur was applied—that is, when external features were not seen at all. Indeed, Frowd et al. found that any presence of the external features was a distraction, and so a more optimal procedure is to present internal features only in face arrays, with external features selected at the end, after holistic and shape tool use (see Fig. 8.22a, b for example screens). Naming increased from 22.7% with the previous level of external features blur to 46.6% for evolving the internal features only (and selecting external features at the end). This method of face construction has been replicated (Frowd, Skelton et al. 2013) and is current police practice.

8.5.1.3 Focus on the Eye Region

A second important development has been to shift the focus of witness attention to the eye region. As mentioned earlier, the internal features of the face are particularly important when recognising familiar faces (e.g. Ellis et al. 1979), but, therein, arguably the region around the eyes is the most important. This is presumably because we pay most attention to this area due, in part, to its multiple roles for communication (e.g. Henderson et al. 2005). Fodarella et al. (2017) found that correct naming increased from 19.8% when eyewitnesses were asked to select for overall match in face arrays (the norm at that time) to 42.9% when asked to focus on the eye region. In connected research, Portch et al. (2017b) found that the holistic mnemonics no longer enhanced composite naming when the focus of attention at construction was on the eye region. Skelton et al. (2019) confirmed that the

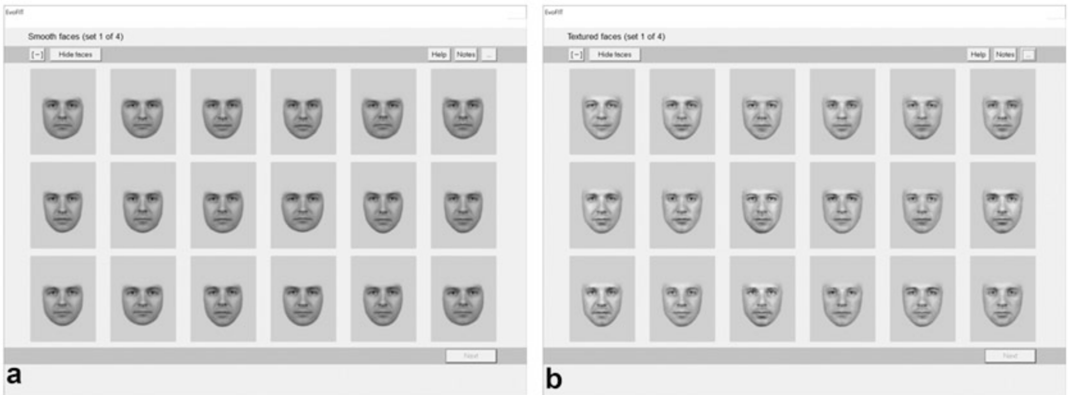

Fig. 8.22 (**a**) Example screenshot of smooth (left) and (**b**) textured internal features (right) for eyewitness selection with internal features only. External features are selected and shown on the face towards the end of the process

holistic mnemonics no longer enhanced EvoFIT composites unless a further mnemonic was administered last, one which requested eyewitnesses to make whole-face judgements for the region around the eyes. As mentioned earlier for computerised feature systems, the research highlights the importance of a consistent focus of attention from cognitive interview to face construction.

8.5.1.4 Population Size

Current research is assessing the importance of population size: the number of faces presented to eyewitnesses for selection. A larger population size should increase the chance of evolving an identifiable likeness, but because eyewitnesses may become fatigued, this assumption may not be entirely correct. As described above, 4 screens of 18 faces were presented to eyewitnesses, a population size of 72 smooth faces and 72 textured faces. However, with a focus on the eye region, reducing the number of screens has been shown to promote more effective composites (Frowd 2021). The research suggests that evolving a not-so-good likeness, one with fewer screens, is allowing an eyewitness to make more effective enhancements to the face using holistic and shape tools. Current police practice involves a population size of 36 faces, and the most recent assessment suggests that this could be halved.

8.5.1.5 Replication and Combining Techniques

The same as for previous systems, projects have assessed techniques in combination with each other. Usually though, once a technique has been found to be effective, further research will tend to incorporate the technique as standard, as was the case in Skelton et al.'s (2019) Experiment 3, which assessed holistic mnemonics using a version of EvoFIT that presented internal features only (cf. blurred external features) in face arrays. When considering the potential positive impact of combining techniques, research has involved the use of enhancement techniques that apply to finished EvoFIT composites, such as naming composites by perceptual stretch (e.g. Frowd et al. 2013, 2014) or by dynamic caricature (Frowd et al. 2008b, 2012a).

In one project, Frowd et al. (2013) found independent effects of holistic (cf. no holistic) mnemonics, system presentation of internal features (cf. blurred external features), and naming with the face side-on (cf. front-on). They demonstrated that these techniques could be combined to improve correct naming: in fact, they found that when all three techniques were used together, naming emerged that was very high, at an astonishing 74% correct, and was greater than all three techniques applied separately.

Fig. 8.23 Example composites constructed of well-known faces from forensic training, to illustrate the effectiveness of eigenface systems. Can you guess the identities? The answers are listed at the end of the chapter

As a final, amusing demonstration of EvoFIT's capabilities, please see Fig. 8.23, which shows example composite constructions of eight well-known celebrities.

8.5.2 Generative Adversarial Networks

A very recent technological addition to composite software is the so-called GAN (Generative Adversarial Network), which appears to allow for the creation of synthetic face images that are difficult to detect as being 'fake'. The concept of generative modelling is not novel within the field of machine learning, dating back to the energy-based models of the 1980s (Bond-Taylor et al. 2022). However, only within recent years, with advances in computing power alongside access to large-scale open-source datasets, have generative models been viable to implement (Bond-Taylor et al. 2022). The main goal of generative

Fig. 8.24 A popular technique for generating deepfakes is to perform face-swapping. Here, the original portrait of Mozart (shown bottom left) is swapped with the face of Abraham Lincoln (shown top left) while applying the same painting style to produce a highly realistic fake portrait painting

modelling is to generate 'fake' data by learning the underlying distribution of training data and then sampling from this estimated distribution (Creswell et al. 2018). In theory, this allows generative models to mimic any underlying data distribution to produce synthetic data in a wide variety of formats, including audio, video, and text (Turhan and Bilge 2018). Deep generative models (DGMs), like other generative models, utilise deep neural networks, composed of many hidden layers and parameters, to learn approximations of complex, high-dimensional distributions of training data in an unsupervised manner (Creswell et al. 2018; Ou 2018). Some examples of data generated by DGMs include well-known 'deepfakes' relating to false video footage, images, or voices of famous celebrities or politicians acting in absurd ways (Tolosana et al. 2020; see Fig. 8.24). In this section, the use of DGMs to generate synthetic facial images will be discussed by focusing on Generative Adversarial Networks.

Generative Adversarial Networks (GANs) were initially proposed by Goodfellow et al. (2014) and rest on the key defining notion of 'adversarial training'. Adversarial training pits two neural networks against each other, namely, a generator model and a discriminator model (Goodfellow 2016). The generator estimates the underlying distribution of the training data and generates fake data. This is 'fed' to the discriminator in the hopes of 'fooling' it, whereas the discriminator aims to correctly distinguish between fake data it receives from the generator and real data it is fed from the training data. The generator has no exposure to the real training data, and the only way it learns is via feedback from the discriminator, which is exposed to both real training data and fake generated data. The discriminator learns by receiving ground-truth

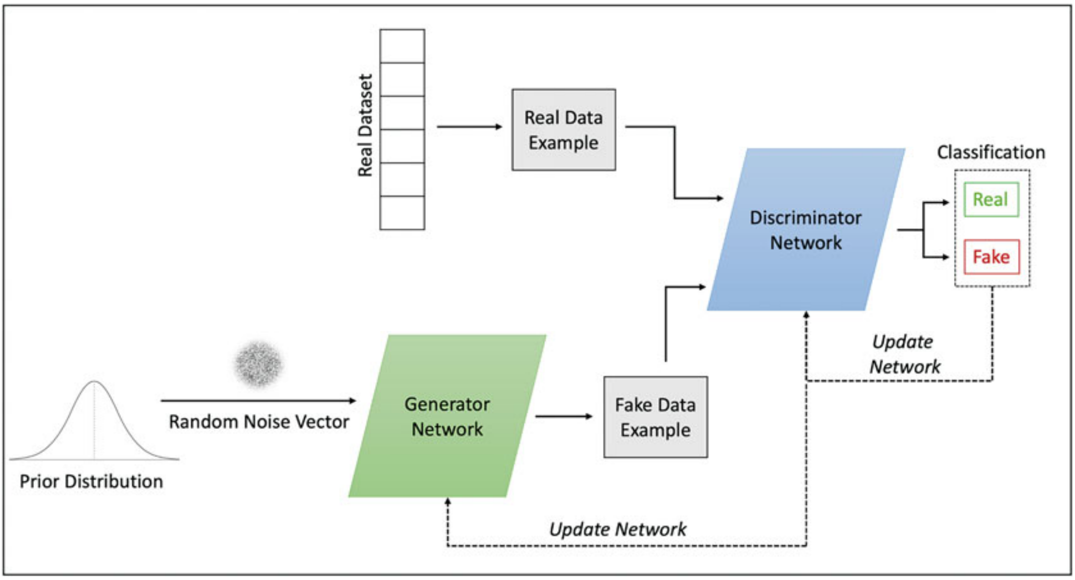

Fig. 8.25 The generator network takes in random noise sampled from a prior distribution to generate fake data. The fake data, alongside sampled real data, is fed into the discriminator network, which must decide on whether the data are real or fake. Both networks are then updated appropriately upon the discriminator network's decision

feedback on whether the data was real, allowing for it to better distinguish between real and fake data (see Fig. 8.25). This process is often put as an analogy: the generator can be thought of as an art forger creating forged paintings, whereas the discriminator is a detective tasked with identifying whether paintings are authentic or forged. For the art forger to keep forging, the forgeries must be continuously improved in order to dupe the detective. On the other hand, if the detective is to capture the forger, there must be continuous improvement at distinguishing authentic paintings from forgeries. This adversarial process allows for the simultaneous training of both the discriminator and the generator model: continuing refinement of the generator occurs via interaction with the discriminator, allowing the generator to update and improve its generation process sufficiently over time to finally 'outsmart' the discriminator. The training is complete once the discriminator cannot discern between the real or fake data at higher than chance probability (Goodfellow et al. 2014). GANs have facilitated significant improvements in artificial image generation, and the key idea seems to have been the incorporation of both a discriminator model and a generator model within a single architecture. This allows for robust and realistic modelling of very complex data, such as natural face images.

Typical GAN architecture for image generation in addition utilises deep Convolutional Neural Networks (CNNs), providing a hierarchical approach to image synthesis. CNNs utilise convolutional layers that systematically increase in size to gradually construct a synthetic image (Bouvrie 2006), mimicking neurons in the mammalian visual cortex (Gu et al. 2018). Due to the layered structure of a deep CNN, both semantic information and texture can be separated and independently extracted (Gatys et al. 2016). In terms of face image generation, this refers to the preservation of the defining arrangement of facial features that make a face a face (i.e. semantic information) while also allowing independent extraction of texture (e.g. skin tone, eye colour, age, etc.) that allows for variation between faces. This method was adopted in Nvidia's StyleGAN (Karras et al. 2019), which utilises deep CNNs to influence texture (referred to as 'style'), through careful input control. This control over 'style' and

Fig. 8.26 Synthetic 'fake' faces of individuals who do not exist, as generated by StyleGAN. See https://thispersondoesnotexist.com/ for more examples

disentanglement of features has resulted in the StyleGAN becoming the leading GAN for synthetic face image generation presently.

StyleGAN maps an original latent constant from a learnt distribution into an intermediate latent space via a feedforward network, applying a non-linear mapping to 'output' a vector, which is further transformed into styles, that in turn serve as input to various layers of the CNN making up the generator model (Karras et al. 2019). This provides control over the generator model's face generation with style input into earlier layers, providing coarse-grain variations (such as changes to pose, face shape, general hairstyle), while style input into later layers provide fine-grain variations (such as changes to colour schemes and facial microstructures). A random component is added to the generation process via injection of Gaussian noise into the various layers of the CNN of the generator model. This 'stochastic variation' allows for arbitrary changes of details to a face such as pore refinement, reflection of light in eyes, facial shadows, and how individual hair strands fall, while global aspects remain unchanged. Refinements to the StyleGAN architecture have occurred with the subsequent releases of StyleGAN2 (Karras et al. 2020) and StyleGAN3 (Karras et al. 2021) to achieve even better image quality of synthetic faces (see Fig. 8.26).

GANs have emerged as a novel technique for composite generation by applying control over image synthesis and allowing for the editing of highly realistic synthetic face images. Transparent latent-space GAN (TL-GAN) provides an approach for directly controlling the generation of synthetic faces based on broad attributes such as sex, age, skin tone, nose, mouth, and eyes to directly alter synthetic faces via human interaction (Guan 2018). This concept was further expanded with the introduction of a composite generating GAN (CG-GAN), which introduces an interactive evolutionary GAN-based approach for facial composite generation (Zaltron et al. 2020). The system allows for direct feature manipulation over GAN-generated faces using TL-GAN's techniques while also applying an interactive evolutionary search algorithm to present users with synthetic faces that can be selected and morphed together in an iterative process of generations until a synthetic face of adequate likeliness is produced. CG-GAN allows for the locking of certain facial features to ensure that specific facial features remain unchanged between generations. Additionally, individual synthetic faces can be manually altered by the user with 'sliding scale' controls. The system also allows for unweighted and weighted morphing of multiple individual GAN composites if more than one witness produces a composite.

An additional possible use of GANs, currently being investigated by the fourth and first authors of the present chapter, is the use of GANs to generate synthetic images of faces from verbal descriptions (Scott et al. 2020). This is a return, with twenty-first-century technology, to the idea that Bertillon introduced to the world in the nineteenth century.

8.5.2.1 Limitations of GANs

Although GANs can produce powerful, realistic results, there are several pitfalls to using them. First, very large and diverse face image databases of adequate resolution are required to train the network (Pavan Kumar and Jayagopal 2021). For instance, StyleGAN used a publicly available face database, known as Flickr-Faces-HQ (FFHQ), which is comprised of 70,000 high-resolution images of human faces (https://github.com/NVlabs/ffhq-dataset). Additionally, due to the complexity of adversarial training, GANs can be very difficult to train (Bond-Taylor et al. 2022). Training convergence may not be achieved due to the generator model and discriminative model not reaching an equilibrium during training, such as when the discriminator model is unable to detect fakes produced by the generator model at levels above chance guessing (Barnett 2018). Mode collapse can also happen, resulting in poor diversity among generated synthetic data as only a small representation of the distribution has been learnt due to the generator producing synthetic data that easily fooled the discriminator early on in training (Saxena and Cao 2021). Additionally, instability can occur due to unsuitable network architecture and poor initiation decisions (Roth et al. 2017). Finally, training of GANs can be costly as it requires efficient computer hardware consisting of enough good quality graphical processing units (GPUs) and ample memory storage (Bhattacharya 2021).

8.6 Do Composites Damage Memory?

We have focused in earlier sections on how to produce the best quality composites. What we have not yet considered, however, is the effect that composites may have on memory. Some researchers have speculated that the process of constructing or even just viewing a facial composite may damage the eyewitness's original memory of the perpetrator's face. In fact, such claims have been presented as scientific fact in the news media (e.g. Munger 2006; Roth 2007) and in legal and police guidelines (e.g. McNamara 2009; Stoughton Police Department 2019). In this section, we review the scientific evidence on the effects of viewing and constructing a composite, respectively.

When an eyewitness sees a composite made by another eyewitness, this may constitute a form of post-event information (see Loftus 2005, for a review of post-event information research), and the composite may presumably alter the original memory of the face or make the original memory harder to access than the competing memory of the facial composite. This is in line with source-monitoring theory (Johnson et al. 1993), which states that people often do not store the source of their memory correctly—in the present example, a memory representation of a face formed while viewing the perpetrator's face may be confused with a representation formed by viewing another eyewitness's composite. If the composite closely resembles the perpetrator, this can be beneficial, strengthening the memory for the original target. If, on the other hand, the composite image does not resemble the perpetrator, this may be detrimental. Sporer et al.'s (2020) narrative analysis of studies on the effect of viewing another person's face composite on subsequent recall and recognition of the target's face showed no support for the idea that viewing an accurate composite might enhance subsequent face recall. In contrast, their analysis did show that viewing another person's composite, especially if it is misleading, may adversely affect both the recall of specific features and the accuracy of identification for a perpetrator. However, they also conclude that the negative effects are limited to certain circumstances, of short duration and are relatively easily reversed. In practical terms, there seems little real danger.

When eyewitnesses construct a composite, this can also influence their original memory for the perpetrator. Creating a composite might help the eyewitness to visually rehearse the perpetrator's face, thereby reinforcing the memory of the face (cf. Meissner and Brigham 2001). In contrast, it might also interfere with the original memory of the face, perhaps even more so than merely viewing a composite made by another eyewitness (cf. Wells et al. 2005). Conflicting results have been reported in the literature, including beneficial effects, no effects, and harmful effects of composite construction on later identification accuracy. It is therefore not surprising that a recent meta-analysis (Tredoux et al. 2021) found no significant overall effects of composite construction on identification performance. Further, they concluded that the current literature does not provide an adequate basis to test potential moderators, even though we might expect that the effects of constructing a composite might depend on the quality of the encoding, the quality of the composite, and the presence or absence of the perpetrator in the line-up, among other things.

In conclusion, the findings of the narrative analysis (Sporer et al. 2020) and meta-analysis (Tredoux et al. 2021) are not in line with the strong claims published in news outlets and practitioner guidelines, which suggest that viewing or creating facial composites impairs memory. Based on the current body of scientific evidence, we are not yet able to formulate sound and empirically based policy recommendations about the ways in which viewing or creating a facial composite may impact memory.

8.7 Are Face Composites Better than Descriptions of Faces?

Assessing the quality of face composites is a practical issue of considerable importance. We outlined one method earlier in the chapter, as developed by Frowd and colleagues (the so-called gold standard, in which composites are created by a witness unfamiliar with the face and then tested for recognition by someone familiar with the face). Another way of thinking about the effectiveness or recognisability of composites though is to ask whether composites are more effective than alternative methods of achieving the same goal (i.e. directing police to the offender). Verbal and written descriptions of offenders are easier (and less expensive) to gather from witnesses than composites, and are arguably more robust than visual likenesses or composites (since they are more likely to describe invariant properties of faces), and—most importantly—only present the information about offender identity that witnesses are confident about. This idea was anticipated by Bertillon, as discussed earlier in the chapter, especially in his idea of 'signaletics'. Face composites, being likenesses of faces, are obliged to present information about all of the face, even if the witness is not confident about all details. A small number of studies have compared the effectiveness of composites versus simple verbal descriptions in securing facial recognition of targets. Christie and Ellis (1981) reported that descriptions of faces were more accurate than Photo-Fit composites when using recognition and classification tasks to assess accuracy. McQuiston-Surrett and Topp (2008) replicated this finding, showing in their study that descriptions were more likely to result in accurate identifications of perpetrators than composites were. Lech and Johnston (2011) found that descriptions appeared to be more accurate even than composites made with the new-generation E-FIT-V synthesis system.

These results are surprising. Christie and Ellis (1981) suggested that the task of achieving visual similarity of faces in face composites is hindered by the interference of specific patterns between the visual images stored in memory and the visual representations to be generated, which is a problem that is less evident for verbal descriptions. Christie and Ellis reported findings from an additional manipulation, in which raters were given *both* descriptions and face composites, but this did not result in better performance. Ness (2003), on the other hand, reported that independent raters who were given both descriptions and composite items performed better than those who were given only descriptions or composite items. This scant line of research definitely needs to be revisited in future studies of composite effectiveness.

8.8 The Ethics of Face Composition

Sociologists and historians have pointed to an intermeshed history of face science and the social control asserted by nascent nation states in the nineteenth century over both recidivist criminals and, to a lesser extent, ordinary citizens (see Cole 2001; Davie 2003; Ellenbogen and Langmead 2020; Higgs 2011; Pavlich 2009). Some of this work was also clearly connected to what are now seen as dubious and biologically misguided theories of physiognomic determinism (e.g. phrenology, Lombroso's ideas about atavism and the origin of crime) and the search for 'a distinctive criminal type', which characterised British criminology then. There is little doubt that the technology of composing faces that we have reviewed in this chapter grew out of those concerns (see Higgs 2011) and out of those technologies. This does not mean that the ideas behind 'signaletics' and 'composite photography' are invalid, but the potential for such technologies to be used for state and social control remains an important issue and arguably much more problematic ethically now than in the nineteenth century. Composing faces is closely connected to the task of searching for faces, and the technologies of face recognition and the uses to which they are put have become particularly controversial in recent times. Van Noorden (2020) lists a number of unethical uses of face recognition technology: there are several instances of Chinese state-funded research reporting training of face recognition software to distinguish faces of Uyghur people from Tibetans and Koreans (ethically problematic since it is now evident that the Chinese state is responsible for ongoing human rights abuses against Uyghur Muslims), building face recognition models from faces of people without their consent, many of which have been used in military and surveillance software but also in commercial applications, and misrecognising faces of ethnic minorities because models have been constructed without a sufficient representation of minority people in the database—among other issues. The creation of 'deepfakes' by sophisticated GANs has also become controversial, and there are known instances of the creation of non-consensual pornography, political disinformation, and financial fraud (Johnson and Diakopoulos 2021). It seems inevitable that laws regulating the creation of synthetic faces and people will follow, and it seems appropriate that face composite researchers and developers consider a code of ethics specific to their practices as soon as possible.

8.9 Discussion and Conclusion

The problem of finding out who committed a crime is as old as crime itself. Throughout the centuries, police have tried to find new ways to locate perpetrators and bring them to justice. In the nineteenth century, the first known facial composites were published. These were drawn by portrait artists and were, at least in some cases, successful in bringing perpetrators to justice. Shortly afterwards, Bertillon's *portrait parlé* became widely used. This involved keeping careful records of apprehended criminals' physical characteristics (e.g. head length, height, eye colour), so that if they reoffended, it would be easy to find and identify them again. Of course, finding out who committed a crime became only more difficult as the population increased and people started travelling further afield. In the second half of the twentieth century, manual systems were developed that allowed police officers to compose portraits by piecing together facial features from a 'kit' of parts (e.g. Identi-Kit, Photo-Fit). With the advent of the computer, these featural methods were computerised (e.g. Mac-a-Mug Pro, FACES, E-FIT, Identi-Kit 2000, and PRO-fit), which greatly improved the number of choices in features, the ease of use, and the quality of the composites. Nonetheless, evaluation of these methods using the 'gold standard' protocol (Frowd et al. 2005a, b) with a realistic delay showed that people were generally unable to recognise the person whom the composite was intended to portray (around 5% accuracy). Naming accuracy can be improved with various tweaks that facilitate either witness memory (e.g. through mnemonics) or later recognition of composites (e.g. through morphing or caricaturing) but tends to remain relatively low unless witnesses have a good memory of the perpetrator's face.

In the twenty-first century, a wholly new approach to composite construction was developed, moving away from featural systems and towards holistic systems. Systems based on eigenface theory, such as ID, EvoFIT, and E-FIT-V, let witnesses select the faces from an array that best approximate the target, after which the computer generates more faces similar to the selected faces, and this process repeats itself until the eyewitness achieves the best likeness possible. These systems are more in line with how people process faces—as a whole rather than as separate features—and evaluations have shown that people are better at naming the resulting composites (around 20% accuracy). Recent enhancements, such as removing distracting regions of the face and focusing attention to the eye region, have improved naming accuracy even further. Although facial composites created with holistic systems tend to look more realistic than those created with featural systems, people are still able to distinguish reasonably well between fake and real faces. In very recent years, however, Generative Adversarial Networks have been able to create fake faces that cannot be distinguished from real faces.

As has become clear in this chapter, facial composite technology has advanced greatly in recent years. At the time of writing, composites can be created that are indistinguishable from real faces. These developments are not uniformly good news, however. First, legislation to prevent potential misuses of the technology is still lagging behind the technological advancements. This means that face recognition tools can be used by governments to spy on their citizens and that realistic fake faces can be used to spread misinformation and mislead unsuspecting people (e.g. through scams on the internet). Second, many of the facial composite systems have been developed with a relative disregard for minority groups, resulting in poorer matching composites and a higher potential for mistaken arrests, false identifications, and wrongful convictions of suspects from minority groups. Conviction on the basis of overwhelming evidence, some of which must *not* be based on eyewitness identification, is crucial. Finally, being exposed to facial composites, especially if they look like real faces, could under some circumstances interfere with eyewitnesses' original memory for the perpetrator, though results are mixed on this issue.

Ultimately, we need to ask ourselves whether the use of facial composites is a more effective method of finding out who committed a crime than alternative methods, such as a simple verbal description of the perpetrator. Limited research to date suggests that a verbal description may actually be more effective in describing the perpetrator than even a high-quality facial composite. Perhaps, Bertillon had it right all along with his *portrait parlé*. If a verbal description is really more effective than creating a composite, this would present a simpler and less risky method of finding suspects. It is therefore crucial that this line of research is expanded. At the same time, researchers should continue to investigate potential improvements to facial composite construction. Now that we have technologically advanced to the point that we can artificially create highly realistic faces, perhaps it is time to shift our focus to important psychological aspects of facial composites. Future research should investigate how to make composites resemble the real perpetrator even better while at the same time considering safeguards to prevent impairments to the eyewitness's original memory.

Answers. Composites of well-known identities in this book chapter have been used to illustrate the effectiveness of production systems. Some composites are easier to recognise than others. Figure 8.23, from left to right, TV personality and entrepreneur Simon Cowell, US President Joe Biden, musician will.i.am, Former US President Donald Trump, TV presenter Anthony (Ant) McPartlin, singer Nicole Scherzinger, Irish comedian and television presenter Dara Ó Briain, and actress Whoopi Goldberg.

References

Alogna VK et al (2014) Registered replication report: schooler and Engstler-Schooler (1990). Perspect Psychol Sci 9:556–578

Barnett SA (2018) Convergence problems with generative adversarial networks (gans). arXiv preprint arXiv:1806.11382

Bertillon A, McClaughry RW (1896) Signaletic instructions including the theory and practice of anthropometrical identification. Werner Company

Bhattacharya G (2021) From DNNs to GANs: review of efficient hardware architectures for deep learning. arXiv preprint arXiv:2107.00092

Blanz V, Vetter T (2003) Face recognition based on fitting a 3d morphable model. IEEE Trans Pattern Anal Mach Intell 25(9):1063–1074

Bond-Taylor S, Leach A, Long Y, Willcocks CG (2022) Deep generative modelling: a comparative review of VAEs, GANs, normalizing flows, energy-based and autoregressive models. IEEE Trans Pattern Anal Mach Intell 44:7327–7347

Bouvrie J (2006) Notes on convolutional neural networks

Boylan J (2001) Portraits of guilt: the woman who profiles the faces of America's deadliest criminals. Pocket Star

Brown C, Portch E, Fodarella C, Tran L, Wood E, et al. (n.d.) The impact of retention interval: improving facial-composite construction using an early free recall mnemonic of the cognitive interview

Brown C, Portch E, Skelton FC, Fodarella C, Kuivaniemi-Smith H, Herold K, Hancock PJB, Frowd CD (2018) The impact of external facial features on the construction of facial composites. Ergonomics 62:575–592

Bruce V, Ness H, Hancock PJB, Newman C, Rarity J (2002) Four heads are better than one. Combining face composites yields improvements in face likeness. J Appl Psychol 87:894–902

Campbell R, Coleman M, Walker J, Benson PJ, Wallace S, Michelotti J, Baron-Cohen S (1999) When does the inner-face advantage in familiar face recognition arise and why? Vis Cogn 6:197–216

Christie DF, Ellis HD (1981) Photofit constructions versus verbal descriptions of faces. J Appl Psychol 66(3):358. https://doi.org/10.1037/0021-9010.66.3.358

Cole SA (2001) Suspect identities: a history of criminal identification and fingerprinting. Harvard University Press, Cambridge, MA and London

Côté P (1998) FACES 3.0 ultimate composite picture. https://facialcomposites.com/

Cootes TF, Edwards GJ, Taylor CJ (1998) Active appearance models. In: European conference on computer vision. Springer, Berlin, Heidelberg, pp 484–498

Creswell A, White T, Dumoulin V, Arulkumaran K, Sengupta B, Bharath AA (2018) Generative adversarial networks: an overview. IEEE Signal Process Mag 35(1):53–65

Davie N (2003) Criminal man revisited? Continuity and change in British criminology, c. 1865–1918. J Vic Cult 8(1):1–32

Davies GM, Christie D (1982) Face recall: an examination of some factors limiting composite production accuracy. J Appl Psychol 67:103–109

Davies GM, Milne A (1985) Eyewitness composite production: a function of physical reinstatement of context. Crim Justice Behav 12:209–220

Davies GM, Milne A, Shepherd JW (1983) Searching for operator skills in face composite reproduction. J Police Sci Adm 11:405–409

Davies G, Valentine T (2007) Facial composites: forensic utility and psychological research. In: The handbook of eyewitness psychology: volume II. Psychology Press, pp 73–98

Davis JP, Sulley L, Solomon C, Gibson S (2011) A comparison of individual and morphed facial composites created using different systems. In: Howells G, Sirlantzis K, Stoica A, Huntsberger T, Arslan AT (eds) 2010 IEEE international conference on emerging security technologies, pp 56–60

Davis JP, Simmons S, Sulley L, Solomon C, Gibson S (2015) An evaluation of post-production facial composite enhancement techniques. J Forensic Pract 17 (4):307–318. https://doi.org/10.1108/JFP-08-2015-0042

Devue C, Barsics C (2016) Outlining face processing skills of portrait artists: perceptual experience with faces predicts performance. Vis Res 127:92–103

Diamond R, Carey S (1986) Why faces are and are not special: an effect of expertise. J Exp Psychol Gen 11: 107–117

Ellenbogen J, Langmead A (2020) Forms of equivalence: bertillonnage and the history of information management. Technol Cult 61(1):207–238

Ellis HD, Davies GM, Shepherd JW (1978) Remembering pictures of real and ‘unreal’ faces: some practical and theoretical considerations. Br J Psychol 69:467–474

Ellis HD, Shepherd JW, Davies GM (1979) Identification of familiar and unfamiliar faces from internal and external features: some implications for theories of face recognition. Perception 8:431–439

Ellis HD, Shepherd JW (1996) Face recall methods and problems. In: Sporer SL, Malpass RS, Koehnken G (eds) Psychological issues in eyewitness identification. Hillsdale, NJ, Lawrence Erlbaum. Ch, p 5

Fisher RP, Geiselman RE, Amador M (1989) Field test of the cognitive interview: enhancing the recollection of actual victims and witnesses of crime. J Appl Psychol 74:722–727

Fodarella C, Frowd CD, Warwick K, Hepton G, Stone K, Date L, Heard P (2017) Adjusting the focus of attention: helping witnesses to evolve a more identifiable composite. Forensic Research & Criminology International 5. https://doi.org/10.15406/frcij.2017.05.00143

Fodarella C, Marsh S, Chu P, Athwal-Kooner H, Jones H, Skelton E, Wood E, Jackson E, Frowd CD (2021) The importance of detailed context reinstatement for the production of identifiable composite faces from memory. Vis Cogn 29:180–200

Frowd, C. (2012). Facial recall and computer composites. Craniofac Identif 42– 56. https://doi.org/10.1017/cbo9781139049566.004

Frowd CD, Hancock PJB, Carson D (2004a) EvoFIT: a holistic, evolutionary facial imaging technique for creating composites. ACM Transactions on Applied Psychology (TAP) 1:1–21

Frowd CD, Hancock PJ, Carson D (2004b) EvoFIT: a holistic, evolutionary facial imaging technique for creating composites. ACM Transactions on applied perception (TAP) 1(1):19–39

Frowd CD, Carson D, Ness H, Richardson J, Morrison L, McLanaghan S, Hancock PJB (2005a) A forensically valid comparison of facial composite systems. Psychol Crime Law 11:33–52

Frowd CD, Carson D, Ness H, McQuiston D, Richardson J, Baldwin H, Hancock PJB (2005b) Contemporary composite techniques: the impact of a forensically-relevant target delay. Leg Criminol Psychol 10:63–81

Frowd CD, Bruce V, Plenderleith Y, Hancock PJB (2006a) Improving target identification using pairs of composite faces constructed by the same person. IEE conference on crime and security. IET, London, pp 386–395

Frowd CD, Bruce V, McIntyre A, Ross D, Fields S, Plenderleith Y, Hancock PJB (2006b) Implementing holistic dimensions for a facial composite system. J Multimed 1:42–51

Frowd CD, Bruce V, Storås K, Spick P, Hancock PJB (2006c) An evaluation of morphed composites constructed in a criminal investigation. Proceedings of the 16th conference of the European Association of Psychology and law. IP-PA Publishing, London, pp 59–66

Frowd CD, McQuiston-Surrett D, Anandaciva S, Ireland CE, Hancock PJB (2007a) An evaluation of US systems for facial composite production. Ergonomics 50:1987–1998

Frowd CD, Bruce V, McIntyre A, Hancock PJB (2007b) The relative importance of external and internal features of facial composites. Br J Psychol 98:61–77

Frowd CD, Bruce V, Ross D, McIntyre A, Hancock PJB (2007c) An application of caricature: how to improve the recognition of facial composites. Vis Cogn 15:1–31

Frowd CD, Bruce V, Smith AJ, Peter JB, Hancock PJ (2008a) Improving the quality of facial composites using a holistic cognitive interview. J Exp Psychol Appl 14(3):276–287. https://doi.org/10.1037/1076-898X.14.3.276

Frowd CD, Park J, McIntyre A, Bruce V, Pitchford M, Fields S, Kenirons M, Hancock PJB (2008b) Effecting an improvement to the fitness function. How to evolve a more identifiable face. In: Stoica A, Arslan T, Howard D, Higuchi T, El-Rayis A (eds) ECSIS symposium on bio-inspired, learning, and intelligent Systems for Security. CPS. (Edinburgh), NJ, pp 3–10

Frowd CD, Pitchford M, Bruce V, Jackson S, Hepton G, Greenall M, McIntyre A, Hancock PJB (2010a) The psychology of face construction: giving evolution a helping hand. Appl Cogn Psychol 25:195–203

Frowd CD, Hancock PJB, Bruce V, McIntyre A, Pitchford M, Atkins R, Webster A, Pollard J, Hunt B, Price E, Morgan S, Stoika A, Dughulia R, Maftei S, Sendrea G (2010b) Giving crime the 'evo': catching criminals using EvoFIT facial composites. In: Howells G, Sirlantzis K, Stoica A, Huntsberger T, Arslan AT (eds) 2010 IEEE International Conference on Emerging Security Technologies, pp 36–43

Frowd CD, Hancock PJB, Bruce V, Skelton FC, Atherton C, Nelson L, McIntyre A, Pitchford M, Atkins R, Webster A, Pollard J, Hunt B, Price E, Morgan S, Greening R, Stoika A, Dughulia R, Maftei S, Sendrea G (2011a) Catching more offenders with EvoFIT facial composites: lab research and police field trials. Global Journal of Human Social Science 11:46–58

Frowd CD, Skelton F, Butt N, Hassan A, Fields S (2011b) Familiarity effects in the construction of facial-composite images using modern software systems. Ergonomics 54:1147–1158

Frowd CD, Hancock PJ, Bruce V, Skelton FC, Atherton CJ, Nelson L et al (2011c) Catching more offenders with Evo FIT facial composites: lab research and police field trials. Glob J Hum-Soc Sci 11(3):35–46

Frowd CD, Pitchford M, Skelton F, Petkovic A, Prosser C, Coates B (2012a) Catching even more offenders with EvoFIT facial composites. In: Stoica A, Zarzhitsky D, Howells G, Frowd C, McDonald-Maier K, Erdogan A, Arslan T (eds) IEEE *proceedings of 2012 third international conference on emerging security technologies*, pp 20–26

Frowd CD, Skelton FC, Atherton C, Pitchford M, Bruce V, Atkins R, Gannon C, Ross D, Young F, Nelson L, Hepton G, McIntyre AH, Hancock PJB (2012b) Understanding the multi-frame caricature advantage for recognising facial composites. Vis Cogn 20:1215–1241

Frowd CD, Fields S (2011) Verbalisation effects in facial composite production. Psychol Crime Law 17:731–744. https://doi.org/10.1080/10683161003623264

Frowd CD, Skelton F, Hepton G, Holden L, Minahil S, Pitchford M, McIntyre A, Brown C, Hancock PJB (2013) Wholeface procedures for recovering facial images from memory. Sci Justice 53(2):89-97 S1355030612001517. https://doi.org/10.1016/j.scijus.2012.12.004

Frowd CD, Jones S, Fodarella C, Skelton FC, Fields S, Williams A, Marsh J, Thorley R, Nelson L, Greenwood L, Date L, Kearley K, McIntyre A, Hancock PJB (2014) Configural and featural information in facial-composite images. Sci Justice 54:215–227

Frowd CD, Erickson WB, Lampinen JL, Skelton FC, McIntyre AH, Hancock PJB (2015) A decade of evolving composite techniques: regression- and meta-analysis. Journal of Forensic Practice 17:319–334

Frowd CD, Goodfellow Y (2017) An investigation into whether poor quality composites are a result of diminished facial recall. In: Bindemann M, Megreya A (eds) Face processing: systems. Disorders and Cultural Differences, Nova Science

Frowd CD, Portch E, Killeen A, Mullen L, Martin AJ, Hancock PJB (2019) EvoFIT facial composite images: a detailed assessment of impact on forensic practitioners, police investigators, victims, witnesses, offenders and the media. In: Stoica A, Howells G, McDonald-Maier K, Erdogan A, Arslan T (eds) Proceedings of IEEE International Conference on Emerging Security Technologies. University of Essex, Colchester, UK. https://doi.org/10.1109/EST.2019.8806211

Galton FJ (1878) Composite portraits. Nature 18:97–100

Gatys LA, Ecker AS, Bethge M (2016) Image style transfer using convolutional neural networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 2414–2423

Geiselman RE, Fisher RP, MacKinnon DP, Holland HL (1985) Eyewitness memory enhancement in the police interview: Cognitive retrieval mnemonics versus hypnosis. J Appl Psychol 70(2):401–412. https://doi.org/10.1037/0021-9010.70.2.401

Gibson S, Solomon C, Maylin M, Pallares-Bejarano A (2003) Facial attribute manipulation for composite systems and computer graphics applications. In: 3rd International Conference on Visualisation, Imaging and Imaging Processing, pp 146–151

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

Godden DR, Baddeley AD (1975) Context-dependent memory in two natural environments: on land and underwater. Br J Psychol 66:325–331

Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S et al (2014) Generative adversarial nets. Adv Neural Inf Proces Syst 27

Goodfellow I (2016) Nips 2016 tutorial: generative adversarial networks. arXiv preprint arXiv:1701.00160

Gu J, Wang Z, Kuen J, Ma L, Shahroudy A, Shuai B et al (2018) Recent advances in convolutional neural networks. Pattern Recogn 77:354–377

Guan S (2018) Generating custom photo-realistic faces using AI. Medium. Retrieved 27 February 2022, from https://blog.insightdatascience.com/generating-custom-photo-realistic-faces-using-ai-d170b1b59255

Hancock, PJB., Burton, AM, Bruce V (1996). Face processing: Human perception and principal components analysis. Mem Cogn, 24(1), 26–40

Hancock PJB (2000) Evolving faces from principal components. Behav Res Methods Instrum Comput 32(2):327–333

Henderson JM, Williams CC, Falk RJ (2005) Eye movements are functional during face learning. Mem Cogn 33:98–106

Henderson Z, Bruce V, Burton AM (2001) Matching the faces of robbers captured on video. Appl Cogn Psychol 15:445–464

Higgs E (2011) Identifying the English: a history of personal identification 1500 to the present. Bloomsbury Publishing

Hsiao JH, An J, Zheng Y, Chan AB (2021) Do portrait artists have enhanced face processing abilities? Evidence from hidden Markov modeling of eye movements. Cognition 211:104616

Johnson MK, Hashtroudi S, Lindsay DS (1993) Source monitoring. Psychol Bull 114(1):3

Johnson DG, Diakopoulos N (2021) What to do about deepfakes. Commun ACM 64(3):33–35

Johnston VS, Franklin M (1993) Is beauty in the eye of the beholder? Ethol Sociobiol 14(3):183–199

Junkin T (2005) Bloodsworth. Algonquin Books

Karras T, Laine S, Aila T (2019) A style-based generator architecture for generative adversarial networks. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp 4401–4410

Karras T, Laine S, Aittala M, Hellsten J, Lehtinen J, Aila T (2020) Analyzing and improving the image quality of stylegan. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp 8110–8119

Karras T, Aittala M, Laine S, Härkönen E, Hellsten J, Lehtinen J, Aila T (2021) Alias-free generative adversarial networks. Adv Neural Inf Proces Syst 34

Kuivaniemi-Smith HJ, Richardson B, Nash R, Uther M, Frowd CD (2018) The impact of mental reinstatement of context on sketch composites. European Association of Psychology and Law (EAPL), Turku, Finland

Lander K, Portch E, Frowd CD, Cootes T (2017) Can animating face composites make them more recognisable? European Association of Psychology and Law (EAPL), Belgium

Lawrence P (2020) Policing, 'science', and the curious case of photo-fit. Hist J 63(4):1007–1031

Lech AM, Johnston R (2011) The relative utility of verbal descriptions and facial composites in facial identifications. International Journal of Bio-Science and Bio-Technology 3:1–16

Loftus EF (2003) Our changeable memories: legal and practical implications. Nat Rev Neurosci 4:231–234. https://doi.org/10.1038/nrn1054

Loftus EF (2005) Planting misinformation in the human mind: a 30-year investigation of the malleability of memory. Learn Mem 12(4):361–366

Malpass RS, Devine PG (1981) Guided memory in eyewitness identification. J Appl Psychol 66:343

McIntyre A, Frowd CD, Bruce V, Hancock PJB (2010) Looking at people who are wearing glasses: the impact of glasses on matching photographs of unfamiliar faces and recognizing familiar composite faces. In: Howells G, Sirlantzis K, Stoica A, Huntsberger T, Arslan AT (eds) IEEE international conference on emerging security technologies, pp 36–43

McIntyre A, Hancock PJB, Langton SRL, Frowd CD (2016) Holistic face processing can inhibit recognition of forensic facial composites. Law Hum Behav 40: 128–135

McNamara JM (2009) Sketchy eyewitness-identification procedures: a proposal to draw up legal guidelines for the use of facial composites in criminal investigations. Wis Law Rev:763–800
McQuiston-Surrett D, Topp LD, Malpass RS (2006) Use of facial composite systems in US law enforcement agencies. Psychol Crime Law 12(5):505–517
McQuiston-Surrett D, Topp LD (2008) Externalizing visual images: examining the accuracy of facial descriptions vs. composites as a function of the own-race bias. Exp Psychol 55(3):195–202. https://doi.org/10.1027/1618-3169.55.3.195
Meissner CA, Brigham JC (2001) A meta-analysis of the verbal overshadowing effect in face identification. Appl Cognitive Psychol 15(6):603–616
Memon A, Wark L, Holley A, Bull R, Koehnken G (1997) Eyewitness performance in cognitive and structured interviews. Memory 5:639–655
Munger D (2006) When crime-fighting tools go bad: problems with the face-composite system. Cognitive daily. http://scienceblogs.com/cognitivedaily/2006/09/27/when-crimefighting-tools-go-ba/
National Policing Improvement Agency NPIA (2009) Facial Identification Guidance. Unpublished
Ness H (2003) Improving composite images of faces produced by eyewitnesses [PhD Thesis, University of Stirling]. Storre. http://hdl.handle.net/1893/12555
Ou Z (2018) A review of learning with deep generative models from perspective of graphical modeling. arXiv preprint arXiv:1808.01630
Paleologue M (1957) My secret diary of the Dreyfus case. Secker and Warburg
Pavan Kumar MR, Jayagopal P (2021) Generative adversarial networks: a survey on applications and challenges. Inter J Multimed Inf Retr 10(1):1–24. https://doi.org/10.1007/s13735-020-00196-w
Pavlich G (2009) The subjects of criminal identification. Punishment Soc 11(2):171–190
Pearson K (1924) *Life and letters of Francis Galton* (1914–1930), vol 2. Cambridge University Press, Cambridge, England
Penry J (1970) Photo-fit. The Police Journal 43(7): 307–316
Police Executive Research Forum (2013) A national survey of eyewitness identification procedures in law enforcement agencies. National Institute of Justice, Washington, DC
Portch E, Frowd CD, Brown C (2017a) Tell me again about the face: using repeated interviewing techniques to improve feature-based facial composite technologies. In: Howells G et al (eds) Proceedings of IEEE 2017 seventh international conference on emerging security technologies. University of Kent, Canterbury
Portch E, Logan K, Frowd CD (2017b) Interviewing and visualisation techniques: attempting to further improve EvoFIT facial composites. In: Howells G et al (eds) Proceedings of IEEE 2017 seventh international conference on emerging security technologies. University of Kent, Canterbury, UK
Prag B (2005) Testing featural and holistic composite technologies under different methods of face recall. Unpublished M.Soc.Sci. thesis. University of Cape Town
REF (2014) Providing software, training and support to the police to allow them to identify criminal suspects using facial-composite images. Impact case study (University of Central Lancashire), Research Excellence Framework (REF). https://impact.ref.ac.uk/casestudies/CaseStudy.aspx?Id=1761
Revell G, Wiley M (2007) Photofit: self-p[ortraits. http://www.gilesrevell.com/files/photofit2.pdf
Righi G, Peissig JJ, Tarr MJ (2012) Recognizing disguised faces. Vis Cogn 20(2):143-169 3. https://doi.org/10.1080/13506285.2012.654624
Roth M (2007) Why police composites don't always hit mark. Pittsburgh Post-Gazette. http://www.post-gazette.com/news/nation/2007/03/25/Why-police-composites-don-t-always-hit-mark/stories/200703250152
Roth K, Lucchi A, Nowozin S, Hofmann T (2017) Stabilizing training of generative adversarial networks through regularization. Advances in neural information processing systems, 30
Schmidt HC (2010) Determinants of estimated face composite quality. Unpublished PhD dissertation. University of Cape Town. https://open.uct.ac.za/handle/11427/11722?show=full
Saxena D, Cao J (2021) Generative adversarial networks (GANs) challenges, solutions, and future directions. ACM Computing Surveys (CSUR) 54(3):1–42
Scott K, Tredoux CG, Nortje A (2020) The relationship between format of facial-identification information and identification accuracy [Conference session]. The British Psychological Society. https://www.bps.org.uk/
Schmidt HC, Tredoux CG (2006) Utilisation and usefulness of face composites in the south African police service-an evaluation study. South African Journal of Criminal Justice 19(3):303–314
Schooler JW, Engstler-Schooler TY (1990) Verbal overshadowing of visual memories: some things are better left unsaid. Cogn Psychol 22:36–71
Shapiro PN, Penrod SD (1986) Meta-analysis of facial identification rates. Psychol Bull 100:139–156
Shaharazam I (1986) The mac-a-mug pro manual. Shaharazam, Milwaukee, WI
Sirovich L, Kirby M (1987) Low-dimensional procedure for the characterization of human faces. J Opt Soc Am A 4(3):519–524. https://doi.org/10.1364/JOSAA.4.000519
Skelton FC, Frowd CD, Speers K (2015) The benefit of context for facial-composite construction. J Forensic Prac 17:281–290
Skelton FC, Frowd CD, Hancock PJB, Jones HS, Jones BC, Fodarella C, Battersby K, Logan K (2019) Constructing identifiable composite faces: the importance of cognitive alignment of interview and construction procedure. J Exp Psychol Appl 26. https://doi.org/10.1037/xap0000257

Sporer SL (1993) Clothing as a contextual cue in face recognition. Ger J Psychol 17:183–199

Sporer SL, Tredoux CG, Vredeveldt A, Kempen K, Nortje A (2020) Does exposure to facial composites damage eyewitness memory? A comprehensive review. Appl Cognitive Psychol 34(5):1166–1179

Stoughton Police Department (2019) Policy and procedure: eyewitness identification. https://www.ci.stoughton.wi.us/vertical/sites/%7B801AC7AB-1155-4D50-B8C6-60A370EC007F%7D/uploads/06_14_EYEWITNESS_IDENTIFICATION.pdf

Sullivan T (2007) Comparing featural and holistic composite systems with the aid of Guided Memory techniques. Unpublished MA thesis. University of Cape Town

Tanaka JW, Farah MJ (1993) Parts and Wholes in Face Recognition. Q J Exp Psychol Sec A 46(2):225–245. https://doi.org/10.1080/14640749308401045

Tanaka JW, Sengco JA (1997) Features and their configuration in face recognition. Mem Cogn 25(5):583–592. https://doi.org/10.3758/BF03211301

Taylor KT (2000) Forensic art and illustration. CRC press

Tolosana R, Vera-Rodriguez R, Fierrez J, Morales A, Ortega-Garcia J (2020) Deepfakes and beyond: a survey of face manipulation and fake detection. Information Fusion 64:131–148

Tredoux C (2002) A direct measure of facial similarity and its relation to human similarity perceptions. J Exp Psychol Appl 8(3):180

Tredoux C, Rosenthal J, Da Costa L, Nunez D (1999) Reconstructing faces with an eigenface composite system. SARMAC III, Boulder, Colorado

Tredoux C, Nunez D, Oxtoby O, Prag B (2006) An evaluation of ID: an eigenface based construction system. South African Computer Journal 2006(37):90–97

Tredoux CG, Sporer SL, Vredeveldt A, Kempen K, Nortje A (2021) Does constructing a facial composite affect eyewitness memory? A research synthesis and meta-analysis. J Exp Criminol 17(4):713–741

Turhan CG, Bilge HS (2018) Recent trends in deep generative models: a review. In: 2018 3rd International Conference on Computer Science and Engineering (UBMK). IEEE, pp 574–579

Valentine T (1991) A unified account of the effects of distinctiveness, inversion and race in face recognition. Q J Exp Psychol 43A:161–204

Valentine T, Davis JP, Thorner K, Solomon C, Gibson S (2010) Evolving and combining facial composites: Between-witness and within-witness morphs compared. J Exp Psychol Appl 16(1):72–86. https://doi.org/10.1037/a0018801

Van Noorden R (2020) The ethical questions that haunt facial-recognition research. Nature 587(7834):354–359

Vredeveldt A, Tredoux CG, Kempen K, Nortje A (2015) Eye remember what happened: eye-closure improves recall of events but not face recognition. Appl Cogn Psychol 29(2):169–180. https://doi.org/10.1002/acp.3092

Wells GL, Hryciw B (1984) Memory for faces: encoding and retrieval operations. Mem Cogn 12:338–344

Wells GL, Hasel LE (2007) Facial Composite Production by Eyewitnesses. Curr Dir Psychol Sci 16(1):6–10. https://doi.org/10.1111/j.1467-8721.2007.00465.x

Wells GL, Memon A, Penrod SD (2007) Eyewitness evidence: improving its probative value. Psychological sciences in the public interest 7:45–75

Wells GL, Charman SD, Olson EA (2005) Building face composites can harm lineup identification performance. J Exp Psychol Appl 11(3):147

Yasuda T (2005) Effects of configural information on face shape perception. Jpn J Psychol 76(2):131–138. https://doi.org/10.4992/jjpsy.76.131

Young AW, Bruce V (2011) Understanding person perception. Br J Psychol 102(4):959–974. https://doi.org/10.1111/j.2044-8295.2011.02045.x

Young AW, Hellawell D, Hay DC (1987) Configurational information in face perception. Perception 16(6):747–759. https://doi.org/10.1068/p160747

Young AW, Hay DC, McWeeny KH, Flude BM, Ellis AW (1985) Matching familiar and unfamiliar faces on internal and external features. Perception 14:737–746

Zaltron N, Zurlo L, Risi S (2020) Cg-gan: an interactive evolutionary GAN-based approach for facial composite generation. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 34, No. 03, pp. 2544–2551)