Recognition of unfamiliar faces

Peter J.B. Hancock, Vicki Bruce
Department of Psychology, University of Stirling

A. Mike Burton
Department of Psychology, University of Glasgow

pjbh1@stir.ac.uk
+44 1786 467675
Fax +44 1786 467641

For Trends in Cognitive Sciences

Summary

People are excellent at identifying faces familiar to them, even from very low quality images, but are bad at recognising, or even matching, faces that are unfamiliar. In this review we shall consider some of the factors which affect our abilities to match unfamiliar faces. Major differences in orientation (e.g. inversion) or greyscale information (e.g. negation) affect face processing dramatically, and such effects are informative about the nature of the representations derived from unfamiliar faces, suggesting that these are based on relatively low-level image descriptions. Consistent with this, even relatively minor differences in lighting and viewpoint create problems for human face matching, leading to potentially important problems over the use of images from security video images. The relationships between different parts of the face (its "configuration") are as important to the impression created of an upright face as local features themselves, suggesting further constraints on the representations derived from faces. The review then turns to consider what computer face recognition systems may contribute to understanding both the theory and the practical problems of face identification. Computer systems can be used as an aid to person identification, but also in an attempt to model human perceptual processes. There are many approaches to computer recognition of faces, including ones based on low-level image analysis of whole face images, which have potential as models of human performance. Some systems show significant correlations with human perceptions of the same faces, for example recognising distinctive faces more easily. In some circumstances, some systems may exceed human abilities on unfamiliar faces. Finally, we look to the future of work in this area, that will incorporate motion and three-dimensional shape information.

Keywords
Face recognition; face identification; face perception; face representations; face configurations; video evidence; principal component analysis

Human face recognition

It is commonly held that humans are rather good at recognising faces. We do generally identify familiar faces with little effort, despite possibly large variations of lighting, viewpoint, expressions and "disguises" such as beards, spectacles and hats. Moreover, familiarity with a face permits identification even from very low quality images[1]. Some early experiments appeared to show a similar facility with unfamiliar faces, with participants able correctly to identify large numbers of target images from distracters. However, it has become clear that such abilities have more to do with picture recognition than face recognition and that our ability to remember, or even to match, unfamiliar faces is rather poor.

It is this much harder task of unfamiliar face processing which is the main subject of this review, though comparisons between familiar and unfamiliar face processing will also be included where these may throw light on unfamiliar face processing. Recognising and comparing unfamiliar faces are tasks of obvious importance to the practical problems of identifying people glimpsed or filmed at the scene of a crime, and understanding such processes is important for facilitating automatic methods of identification to supplement human vision. Problems of eye-witness memory for briefly viewed faces have been noted for some 25 years or so, but with the increased use of surveillance cameras for recording events in public places, it is timely now to reflect upon factors which may affect human abilities to compare faces visually as well as affecting memory. The first part of this article provides an overview of what is known about our abilities to recognise and match unfamiliar faces. We consider how these abilities are affected by variations in imaging conditions (viewpoint, lighting) and by intrinsic properties of individual faces (e.g. their distinctiveness). We then describe some computer systems which have been developed to recognise faces, and consider how well their performance compares with humans. Few of these systems, however, were developed explicitly with the aim of modelling human recognition processes. In the final part of the article we consider what additionally can be learned about human face recognition abilities from attempts to model it.

We may usefully divide the problems associated with face recognition into two types: imaging problems, such as variations in viewpoint and lighting; and those inherent in the structure of faces, such as their configuration or distinctiveness. The distinction is, however, not entirely clear-cut, for example there is evidence that face-specific processing is used to solve the problem of recognition across different viewpoints. We first consider the effects of viewpoint and lighting, before turning to more face-specific issues such as configuration and distinctiveness.

Viewpoint and orientation

Much of the work on face perception has focussed on front views of faces, but a face is a complex 3D object, which must be recognised across different directions of lighting and despite rigid changes in viewpoint and non-rigid variations in expression.

There seems to be some advantage in recognition and memory for an angled view - often set at 20, 30 or 45 degrees rotation from frontal (the "3/4" viewpoint). In various studies this has shown up as an advantage for study, test, or both, of 3/4 views compared with frontal views or profiles (which usually produce poorest performance). This is partly because such views lie between frontal and profile and therefore there is less possible change in orientation between learning and test when changed views are used. However, that is not the whole story, as the recognition rates when there is no change of orientation are highest at 3/4 view, confirmed again by O'Toole et al.[2] Hill et al.[3] found that while generalisation from one profile to the other was poor, generalisation from one 3/4 view to the other was very good. Performance at profile may be bad because it obscures the configural information that underlies normal face processing.

Inverting a face causes particular problems with identification, thought to be caused by disruption of the processing of the configuration of the face, discussed further below. The fall-off in recognition with increasing change of viewpoint between learning and test is larger for inverted faces than for upright ones,[4] suggesting that face-specific processes are normally used to help solve the viewpoint problem. O'Toole et al.[5] looked in detail at the pattern of responses for different faces and concluded that there are two separable components that contribute to recognition performance at differing viewpoints. One correlates with performance between frontal and 3/4 views, the other with performance between 3/4 and profile views. The facial characteristics underlying these components are not clear, but might be something to do with global and local distinctive features, respectively.

Lighting and negation

Photographic negatives invert the pattern of brightness across the image, and are extremely difficult to recognise, even harder than inverted images.[6] This has implications for the mechanisms underlying human face identification. In principle it should be possible for a human observer to derive information about the size and position of facial features from a negative: edges will still show as luminance gradients. Thus the difficulties experienced with negative images suggest that visual representations of faces are closely tied to properties of the original images of the faces, rather than based on derived measurements.

Line drawings of faces are equally easy (or difficult) to recognise whether black on white or white on black, suggesting that it is information about shading that is disrupted by negation. This was tested by Hayes et al.[7] who showed that two-tone images could be recognised equally well in negative if first filtered by a high pass spatial filter, which leaves only edges in the image. Lower spatial frequencies contain the information about shading, shown above to be beneficial for identification. If these are negated, perception of the underlying shape may be impaired.

The effects of photographic negation on an image resemble those of lighting from below. Such lighting also disrupts identification, as shown by Johnston et al.[8] They also showed that bottom lighting ameliorated the effects of negation or inversion on identification rates. So an inverted or negated familiar face is easier to recognise if it has also been bottom-lit, implying that inversion has some effect on perception of shape as well as disrupting processing of the configuration.

Hill and Bruce [9] investigated the effects of lighting direction on a face-matching task using 3D surface images of heads. Pairs of images were shown and participants had to decide whether the same person was shown in each. Matching the images, even with the same viewpoint, was much harder when the direction of lighting differed. Again, this suggests that representations preserve image properties, since similar contours could be derived from heads lit from different directions. Matching across different viewpoints was easier when both were lit from above than when both were lit from below.

While negating the luminance of an image disrupts face processing, Kemp et al.[10] demonstrated that changing image hues does not. They produced face images in which either the luminance, the hue or both were negated. Recognition of familiar faces was impaired by the change in luminance, but not by hue. However, recognition memory for pictures of unfamiliar faces was affected by changes in hue, suggesting that initial picture memory is hue dependent, but that this variability gets excluded in the process that accompanies familiarisation with a face.

Configuration

A number of strands of evidence indicate that the relationships between different face features (their "configuration") are as important for the overall impression given as the components themselves. We seem particularly sensitive to the configuration of upright faces - inversion removes this sensitivity. The most dramatic example of this is probably the “Thatcher illusion” [11] illustrated in Figure 1.

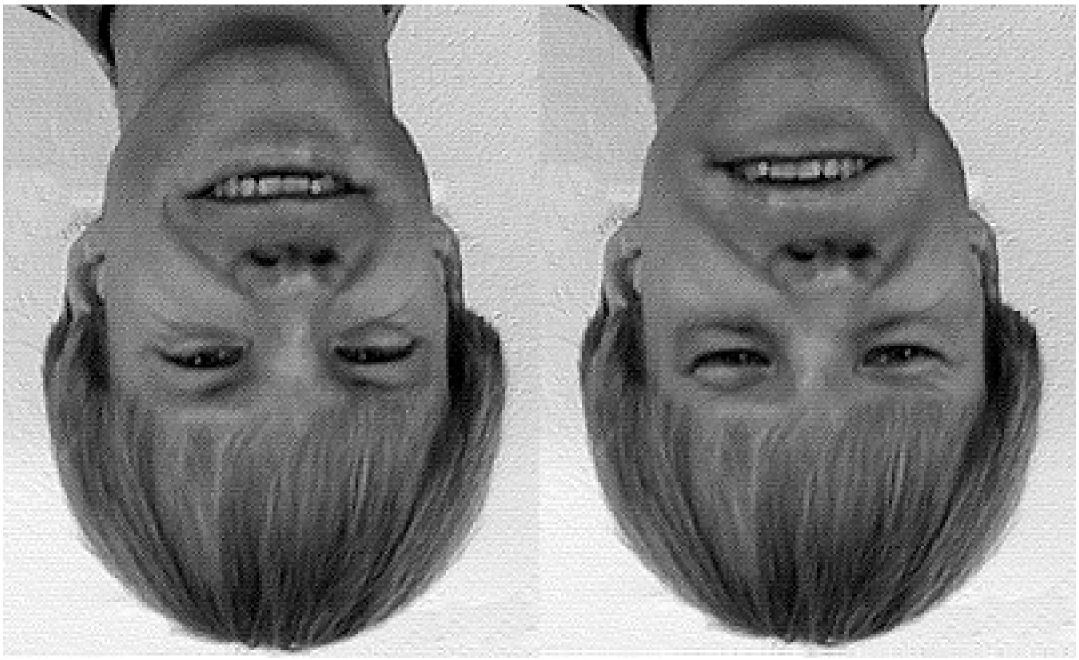


Figure 1. The "Thatcher" illusion so-named because a picture of Margaret Thatcher's face was originally used by Thompson[11], is here illustrated using Peter Hancock's face. Rectangles of a face image containing the eyes and the mouth are cut out and inverted. Viewed upside down, it is difficult to see anything amiss, but when seen the correct way up, the altered face is hideous. Inversion of the whole face disrupts our ability to perceive the illegal alignment of features, and the effect is seen with unfamiliar faces as well as familiar ones.

Inverting a picture of anything that has a correct way up makes it harder to recognise, but the effect on faces is disproportionate[12]. Recently Farah et al.[13] have shown that the disproportionate effect of inversion on faces extends to the perceptual level. Using a rapid presentation same/different task, they showed that the performance with faces deteriorated more with inversion than with words or images of houses.

An elegant demonstration of our configural perception of faces, and the ability of inversion to disrupt it, was given by Young et al.[14] They combined the top half of one face with the bottom of another. When correctly aligned, it was hard to recognise the individual identities of the two halves. If misaligned, identification was much easier. When the faces were inverted, this difference disappeared. It is apparently easier to concentrate on the features of the relevant half of the face when interference through configural processing from the more distant parts of the composite face is removed following inversion.

What do we mean by "configuration"? Work by Bartlett and Searcy[15,16] indicates it is the ability to perceive spatial relationships between features which is disrupted by inversion. They modified images of faces, moving features about to produce configurations which, when viewed normally, are clearly unreasonable, with, for example, much too big a gap between mouth and nose.

Figure 2. Which of the two side images differs from the one in the middle? Now turn the journal upside down. Spatial displacements of features (left, this way up) are readily seen in upright faces but become virtually invisible with inverted ones. In contrast, local changes such as thickening of the eyebrows remain readily detected in inverted faces. These manipulations have been made to a picture of Mike Burton's face.

When the images were inverted, the “grotesqueness” ratings were much lower. However, when the faces were modified by making a change to a particular feature, such as adding vampire teeth, inversion had relatively little effect. Leder and Bruce[17] showed that similar modifications also affected memorability of unfamiliar faces. Faces rendered more distinctive either by modification to a local feature or by changing the relationships between features were both more memorable than the unaltered originals when viewed upright, but only the locally altered faces were better remembered when inverted (see Figure 2). Other researchers (e.g. Tanaka and Farah[18]) have suggested that configural means "holistic" - the upright face pattern may not be decomposed into parts or features at all. A related suggestion, more compatible with the key role of spatial relationships, might be that the features important for face perception include a number of global features or "metric" *variations* [19] as well as "simple" local features such as eyes or nose.

It is difficult to explore these issues empirically, since any change to a local feature tends also to change spatial relationships, the "holistic" pattern of the face, and any more global features we might propose. Leder and Bruce[20] provide some evidence supporting spatial relationships over "holistic" processing, but the jury is still out.

Recent advances in fMRI have enabled the effects of inversion on brain activity to be studied. It has been found that when faces are inverted, there is a relatively larger response from the cortical areas involved in normal object perception.[21,22] The response from the fusiform "face area" was undiminished, suggesting that while face-specific processes continue to operate for inverted faces, there is a greater contribution from normal object areas.

Distinctiveness and face space

The way we perceive a face affects the likelihood that it will be remembered. The most obvious characteristic is that of distinctiveness: unusual faces are better remembered than typical ones. Numerous studies have shown that faces rated as distinctive (e.g. "how easy would it be to pick this person out in a crowd?") are subsequently better recognised on test.[23] Faces rated as typical are more likely to give rise to a false positive recognition. Distinctiveness also affects how quickly a known face is identified.[24] Thus someone with a striking face such as Arnold Schwarzenegger would be more rapidly identified than someone who is more typical, such as Leonardo diCaprio. Where the task is to identify whether an item is a face or a non-face (with non-faces constructed such that features are rearranged), distinctive faces are at a disadvantage, whether familiar or unfamiliar.

A caricature, produced either by an artist or by a computer, emphasises the differences between a face and the average. Line drawings that are rather poorly recognised if veridical are better identified when caricatured.[25] A caricatured photograph of a famous person may be rated as a better likeness than the original.[26] One interpretation of the effect of caricaturing is that moving a face away from the relatively crowded centre of face space (see below) increases the average distance to other known faces, effectively increasing distinctiveness and facilitating recognition.[27] Caricature effects are most often observed in already familiar faces, though occasional effects have been reported with unfamiliar ones too.[28]

A popular way to conceptualise our internal representation of faces is that of a face space.[29] There are hypothesised to be a number of dimensions upon which a face can be measured, so that a given face can be described by a location in a multi-dimensional, usually Euclidean, space. This type of face space model has been used to offer accounts of effects such as distinctiveness and caricature on recognition e.g.[30]. These accounts rely on ideas such as the relative density of different parts of face space, and on the distance between particular representations within the space. Recent work provides insights about the possible nature of the space.[31,32]

There are a number of candidates for what the underlying dimensions might be. One way to tackle this is to perform multi-dimensional scaling (MDS) on human similarity ratings. An example is given by Busey[33] who found that six identifiable dimensions accounted for three quarters of the variance in a set of images of bald men: age; race; facial adiposity (plumpness); facial hair; aspect ratio (short-fat/long-thin); and facial hair colour. When target faces are not bald, hair colour and style is an important dimension. An alternative approach is to perform a statistical analysis, such as PCA (see Box 1), on a set of face images, and to establish whether there are correspondences between the derived components and human performance. A number of studies have found such correlations[34-36] but it is often hard to give labels to the dimensions. PCA may offer a feel for the dimensionality of the space: e.g. Scheuchenpflug found that 13 components best described the human data. However, such computer-derived measures still fare relatively poorly in accounting for human data. Dailey et al.[37] attempted to model human recognition performance using descriptions given by PCA, Gabor wavelets and human MDS data (from Busey[33]). The human MDS data gave a much better account, especially when paired with a kernel density estimation model that allocated bigger kernels (regions of face space) to more distinctive faces.

Failures of recognition

Early work on face recognition in the laboratory suggested that recognition rates for unfamiliar faces could be high. However, early attempts to test recognition in the field[38] indicated that real-life performance was much worse. Participants were asked to spot a "suspect" in a shopping area and met with almost complete failure. A number of cases of misidentification leading to criminal conviction have also highlighted the difficulty of recognition of unfamiliar faces. A current concern stems from the rise of closed circuit video surveillance of public areas, with most major town centres in the UK now covered with cameras. Increasingly, suspects are caught in the act on video and there follows a requirement to identify the individual concerned. While presentation of the video evidence may elicit a guilty plea, there are increasing numbers of cases where a misidentification has been claimed.

Recent work in our group has been investigating the ease with which video images can be matched with photographic stills. If the viewer is familiar with the target, performance is extremely good, even with very poor quality video.[1] When the same video clips are shown to observers unfamiliar with the targets, performance is much worse. Even with high quality video, matching to photographs is surprisingly error-prone. Bruce et al.[39] used stills from a VHS video, recorded on the same day as Photo-CD quality still images. A video still was presented along with an array of 10 photographs of men chosen to be of similar general appearance to the target. The participants' task was to decide whether the target was present in the array. Even when viewpoint and expression matched, the average performance was wrong 30% of the time. Changes in viewpoint or expression caused even more errors.

Poor performance is found even when the task is made simpler as illustrated in Box 2. Here, participants are told (correctly) that the target person is present, and they simply have to choose which one is the match. Even in this case, error rate is about 20% on average, with some faces giving rise to very high error rates. These are very surprising results. There is no memory load on subjects, and no requirement to make a decision quickly. So, even when given free-viewing conditions and simply asked to match a face presented at the same time as a line up of 10 photos, subjects perform surprisingly poorly. These studies have shown that there was little effect of using colour rather than greyscale images, and perhaps more surprisingly, the use of animated video clips rather than stills does not improve performance in these matching tasks. This suggests that the visual comparisons of such unfamiliar faces are based on relatively low level image comparisons, which are not helped by more abstract 3D information that might be extracted as a result of movement. More work is required to identify what aspect of the matching is causing the high failure rate. Possibilities include different non-linearities in the brightness between video and still photograph, and different perspectives from the two lenses. Whatever the cause, it is clear that human matching of unfamiliar faces is far from infallible even under seemingly ideal conditions. Not surprising then, that there are moves towards using computer based matching systems in an attempt to help out. To be useful, an automatic solution to this problem will have to out-perform humans, clearly an ambitious aim.

Computer recognition

Automated face recognition has obvious commercial and security applications and there are a wide variety of approaches.[40] The problem for automatic systems is the same as for humans: at an image level, the variability in the appearance of a given face under variations of lighting, pose and expression far exceeds the variation between two different faces under the same conditions. A useful system has to find a way through the extraneous variability to reach identity. Even the task of finding a face in an image, normally effortless for humans, is computationally complex and difficult to achieve.

A complete overview of the range of computational approaches is beyond the scope of this article, see Chellappa et al.[40] and Grudin[41] for recent comprehensive reviews. Instead we shall look at a small number of systems that have been studied more intensively and for which comparisons exist. Comparing the results from different papers is difficult because of the variability of training and test sets. Unless the faces used have been collected very carefully, it is quite possible for extraneous variables, such as background, clothing or subtle changes in lighting, to provide cues to identity. This is particularly a problem with self-organising models, which will adapt themselves to whatever cues are present. More constrained models, with some notion of the structure of a face built in, are less likely to be fooled.

In order to allow proper comparisons, the same face set needs to be used with different systems. A controlled series of comparisons has been carried out using the FERET database.[42] The database is in two parts, a public set that may be used during development of a system and a hidden test set used to conduct comparisons. We shall look briefly at the technology underlying some of the more successful systems in the FERET competition.

Computer-based approaches to face recognition usually have two main stages: an image processing stage, where the face image is converted to an internal representation, followed by a matching stage, where the target is compared to the stored representations of known faces. Although the evidence is that humans use a rather holistic representation for faces, a number of computer-based approaches continue to use purely feature-based representations.[43] Brunelli and Poggio[44] compared such a template-based feature matching system with one that used geometrical features such as the shape and position of mouth and nose and found the former to be more successful. At the other extreme are purely holistic representations, such as PCA on whole face images. Falling in-between are systems that use local feature representations, but which then code the relationships between them. Note that the local features used may not correspond with our idea of features, for example, the output of a localised Gabor filter, rather than a nose.

The simplest kind of matching is to compute a distance, often Euclidean, between the target and each of the stored representations. The closest is taken to be the match, usually subject to some threshold in order to reduce the chances of false identifications. This form of matching was used by the early PCA-based system.[45] However, it ignores potentially useful knowledge given by the presence of multiple views of each person in the database. One approach to this is to use Linear Discriminant Analysis, which produces an optimal linear classification between faces belonging to one person and all others.[46] This system performed well in the FERET test. Meanwhile, Moghaddam et al added a non-linear classification function to the PCA-based system.[47] They begin by performing PCA on intra-personal variation as well as the usual inter-personal variation. They can then compute the probability that each type of variation accounts for the difference between a target image and an item in the database. This produces a marked increase in recognition rates, amongst the best in the FERET competition.

An example of a system that does much more than a simple distance match is the graph matching system of Wiskott et al.[48] The face images are initially processed by banks of Gabor filters (or jets) that reflect human early visual processing. Internal representations of faces are held as labelled graphs, which have to be distorted somewhat to match the incoming face description. The model that suffers least strain is declared the winner, provided the match reaches some threshold. This system is also highly competitive in the FERET tests.

While some of the systems now being developed may represent good engineering approaches to face identification, they may bear little resemblance to human perception. Comparisons between human and computer performance are therefore of interest, for although agreement does not imply that the computer system is working in the same way that we are, disagreement can help to rule out candidates, and differences may shed light on what happens as we become familiar with a face. We have compared the PCA approach with the graph matching system and found that the latter provided a better overall account of human face recognition, while PCA (at least with unprocessed images) was more sensitive to image variations and thus more resembled our abilities with unfamiliar faces.[49] There have recently been some other comparisons of these two systems, or variants thereof, with human performance.[36,37,50] Box 2 explores how some current systems fare at the same task of identity verification from video images that we have used with human participants.

Integration of perceptual and cognitive models

In this review we have concentrated on the *perception* of faces, focussing on the visual analysis of relatively unfamiliar faces. However, there is a very large parallel literature concerned with *cognitive* or memorial aspects of face recognition. This literature is concerned mainly with delineating the different information-processing routes involved in extracting different types of meaning from the face, such as its emotional expression versus its identity, and with the stages of information processing involved in identifying familiar people - for example distinguishing a stage where the familiarity only is known, from a stage where knowledge of a person's occupation or nationality can be retrieved, from a stage where the person's name can be retrieved. The modal model of cognitive aspects of face processing was developed by Bruce and Young[51] and this remains a much-used framework today. The face identification component of that model was implemented by Burton et al.[52] in a simple connectionist architecture. The original and implemented versions of the model have been used to provide accounts for a number of phenomena, such as priming (the facilitation of recognition by an earlier encounter with an identical or related item), everyday recognition failure, the particular difficulty of recalling names even after the face has been identified, and a variety of impairments to face identification which may arise as a result of brain damage.

The perceptual and cognitive literatures have tended to run in parallel. So, models of memory for familiar people have tended to remain agnostic about the visual processes which deliver their input, developing accounts based upon hypothetical rather than real face patterns. Similarly, most perceptual models have not been constrained by what is known about the different kinds of information in memory which can be accessed via the face. This separation is unproblematic for many effects, particularly those that have a clear locus. However, there are a number of effects which require an account of both perceptual and memory processes. We have mentioned above the importance of familiarity with a face for recognition and matching. Similarly, some effects of distinctiveness and priming differ according to whether faces are known or not.

We have recently offered an account which ties together perceptual and cognitive aspects of face recognition.[53] This model uses a front-end image processing system based on PCA (See box 1). Information from this analysis is passed on to a connectionist system[52] developed to account for the later stages of cognitive identification. This integrated model is able to capture a number of aspects of human face recognition, for example, a significant negative correlation between recognition time and human rated distinctiveness of the faces.

It is our belief that future models of face processing must incorporate both perceptual and cognitive aspects if they are to tackle some of the most important and interesting aspects of the problem. Moreover, since cognitive models have emphasised processes involved in the recognition of already familiar faces, and research on face perception has often focused on the processing of unfamiliar ones, the marriage of these approaches will be necessary also if we are to understand the ways in which representations change as faces become familiar. Although our own model may very well turn out not be the most useful, we believe it makes a start in bringing together these areas of study which traditionally remain separate.

Current research: adding dimensions

Most work to date has used static, 2-dimensional images of faces. More powerful computers have started to lift these restrictions, e.g. Blanz and Vetter[54] describe a very sophisticated 3-D model of faces that can produce variations in appearance to order. These techniques are finding their way into psychological testing with some unexpected results, for example, it was found that caricatures in 3D look older, rather than more recognisable.[55] The effects of motion on recognition are also beginning to be tested. Benefits of movement for the task of unfamiliar face recognition are unclear (e.g. Christie & Bruce[56] and Bruce et al.[39] report no benefits, while Pike et al.[57] find advantages for study of moving items in recognition memory), but there have recently been some intriguing demonstrations that motion can aid the identification of familiar faces seen in images which are otherwise difficult to recognise.[58,59]. This suggests that as faces become familiar, their characteristic patterns of movement may become represented in memory as well as the static form.

Outstanding questions

What internal changes accompany the change in performance as we become familiar with a face, to go from the situation where it is hard even to match photographs to one where a highly impoverished stimulus is adequate for recognition?
What are the dimensions that we use to describe faces internally?
How important is motion in face identification?
How important is 3D information?
Will it be possible to produce computer systems that are consistently more reliable than humans at identifying unfamiliar faces?

References

1. Burton, A. M. et al. (1999) Face recognition in poor-quality video: Evidence from security surveillance. *Psychological Science* 10, 243-248.

2. O'Toole, A. J. et al. (1998) Stimulus-specific effects in face recognition over changes in viewpoint. *Vision Research* 38, 2351-2363.

3. Hill, H. et al. (1997) Information and viewpoint dependence in face recognition. *Cognition* 62, 201-222.

4. Moses, Y. et al. (1996) Generalization to novel images in upright and inverted faces. *Perception* 25(4), 443-461.

5. O'Toole, A. J. et al. (1998) The perception of face gender: The role of stimulus structure in recognition and classification. *Memory and Cognition* 26, 146-160.

6. Bruce, V. and Langton, S. (1994) The use of pigmentation and shading information in recognising the sex and identities of faces. *Perception* 23, 803-822.

7. Hayes, T. et al. (1986) Recognition of positive and negative bandpass-filtered images. *Perception* 15, 595-602.

8. Johnston, A. et al. (1992) Recognising faces: Effects of lighting direction, inversion, and brightness reversal. *Perception* 21, 365-375.

9. Hill, H. and Bruce, V. (1996) Effects of lighting on the perception of facial surfaces. *Journal of experimental psychology-human perception and performance* 22(4), 986-1004.

10. Kemp, R. et al. (1996) Perception and recognition of normal and negative faces: the role of shape from shading and pigmentation cues *Perception* 25, 37-52.

11. Thompson, P. (1980) Margaret Thatcher: A new illusion. *Perception* 9, 483-484.

12. Valentine, T. (1988) Upside-down faces: A review of the effect of inversion upon face recognition. *British Journal of Psychology* 79, 471-491.

13. Farah, M. et al. (1998) What is "special" about face perception? *Psychological review* 105(3), 482-498.

14. Young, A. W. et al. (1987) Configurational information in face perception. *Perception* 16, 747-759.

15. Bartlett, J. C. and Searcy, J. (1993) Inversion and configuration of faces. *Cognitive Psychology* 25, 281-316.

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

17. Leder, H. and Bruce, V. (1998) Local and relational aspects of face distinctiveness. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology* 51(3), 449-473.

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

19. Cooper, E. E. and Wojan, T. J. (1996) Differences in the coding of spatial relations in faces and objects *Investigative Opthalmology and Visual Science* 37, 842.

20. Leder, H. and Bruce, V. (In press) When inverted faces are recognised: the role of configural information in face recognition *Quarterly Journal of Experimental Psychology*.

21. Aguirre, G. K. et al. (1999) Stimulus inversion and the responses of face and object-sensitive cortical areas *Neuroreport* 10, 189-194.

22. Haxby, J. et al. (1999) The effect of face inversion on activity in human neural systems for face and object perception *Neuron* 22, 189-199.

23. Light, L. L. et al. (1979) Recognition memory for typical and unusual faces. *Journal of Experimental Psychology: Human Learning and Memory* 5, 212-228.

24. Valentine, T. and Bruce, V. (1986) Recognizing familiar faces: The role of distinctiveness and familiarity. *Canadian Journal of Psychology* 40, 300-305.

25. Rhodes, G. et al. (1987) Identification and ratings of caricatures: Implications for mental representations of faces. *Cognitive Psychology* 19, 473-497.

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

27. Lewis, M. B. and Johnston, R. A. (1999) A unified account of the effects of caricaturing faces *Visual Cognition* 6, 1-42.

28. G. Rhodes, Superportraits: Caricatures and recognition. (Psychology Press, Hove, ed. 1996).

29. Valentine, T. (1991) A unified account of the effects of distinctiveness, inversion, and race in face recognition. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology* 43A, 161-204.

30. Byatt, G. and Rhodes, G. (1998) Recognition of own-race and other-race caricatures: Implications for models of face recognition. *Vision Research* 38, 2455-2468.

31. Burton, A. M. and Vokey, J. R. (1998) The face-space typicality paradox: Understanding the face-space metaphor. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology* 51A, 475-483.

32. Wickham, L. H. V. et al. (2000) Facial distinctiveness: its measurement, distribution and influence on immediate and delayed recognition. *British Journal of Psychology* 91, 99-123.

33. Busey, T. (1998) Exemplar-based models of face recognition. *Journal of Mathematical Psychology* 42(4), 494.

34. O'Toole, A. J. et al. (1994) Structural aspects of face recognition and the other-race effect. *Memory and Cognition* 22, 208-224.

35. Hancock, P. J. B. et al. (1996) Face processing: Human perception and principal components analysis. *Memory and Cognition* 24, 26-40.

36. Scheuchenpflug, R. (1999) Predicting face similarity judgements with a computational model of face space. *Acta Psychologica* 100, 229-242.

37. Dailey, M. N. et al. (1999) Facial memory is kernel density estimation (almost). *Advances in Neural Information Processing 11*.

38. Logie, R. H. et al. (1987) Face recognition, pose and ecological validity. *Applied Cognitive Psychology* 1, 53-69.

39. Bruce, V. et al. (In Press) Verification of face identities from images captured on video *Journal of Experimental Psychology: Applied*.

40. Chellappa, R. et al. (1995) Human and machine recognition of faces - a survey. *Proceedings of the IEEE* 83(5), 705-740.

41. Grudin, M. A. (2000) On Internal Representations in Face Recognition Systems *Pattern Recognition* 33, 1161-1177.

42. Phillips, P. et al. (1998) The Feret database and evaluation procedure for face-recognition algorithms. *Image and vision computing* 16(5), 295-306.

43. Doi, M. et al. (1998) A robust face identification against lighting fluctuation for lock control. *Proceedings of the Third International Conference on Automatic Face and Gesture Recognition*. 42-47. Los Alamitos, California, IEEE Computer Society.

44. Brunelli, R. and Poggio, T. (1993) Face recognition - features versus templates. *IEEE transactions on pattern analysis and machine intelligence* 15(10), 1042-1052.

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

46. Zhao, W. et al. (1998) Discriminant analysis of principal components for face recognition. *Proceedings of the Third International Conference on Automatic Face and Gesture Recognition*. 336-341. Los Alamitos, California, IEEE Computer Society.

47. Moghaddam, B. et al. (1998) Beyond eigenfaces: probabilistic matching for face recognition. *Proceedings of the Third International Conference on Automatic Face and Gesture Recognition*. 30-35. Los Alamitos, California, IEEE Computer Society.

48. Wiskott, L. et al. (1997) Face recognition by elastic bunch graph matching. *IEEE transactions*

on pattern analysis and machine intelligence 19(7), 775-779.

49. Hancock, P. J. B. et al. (1998) A comparison of two computer-based face identification systems with human perceptions of faces. *Vision Research* 38, 2277-2288.

50. Kalocsai, P. et al. (1998) Face similarity space as perceived by humans and artificial systems. *Proceedings of the Third International Conference on Automatic Face and Gesture Recognition*. 177-180. Los Alamitos, California, IEEE Computer Society.

51. Bruce, V. and Young, A. (1986) Understanding face recognition. *British Journal of Psychology* 77, 305-327.

52. Burton, A. M. et al. (1990) Understanding face recognition with an interactive activation model. *British Journal of Psychology* 81, 361-380.

53. Burton, A. M. et al. (1999) From pixels to people: A model of familiar face recognition. *Cognitive Science* 23, 1-31.

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

55. O'Toole, A. et al. (1997) Three-dimensional caricatures of human heads: distinctiveness and the perception of facial age. *Perception* 26(6), 719-732.

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

57. Pike, G. E. et al. (1997) Recognizing moving faces: The relative contribution of motion and perspective view information. *Visual Cognition* 4, 409-437.

58. Knight, B. and Johnston, A. (1997) The role of movement in face recognition. *Visual Cognition* 4, 265-273.

59. Lander, K. et al. (1999) The role of movement in the recognition of famous faces *Memory and Cognition* 27, 974-985.

Box 1 Principal components analysis of faces.

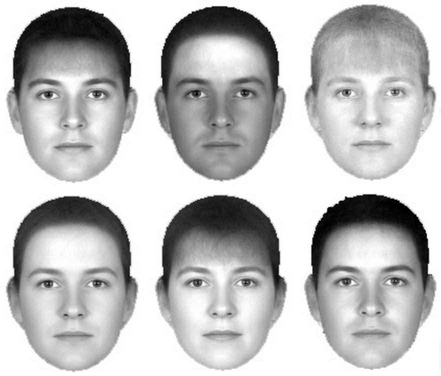


Figure I From left to right, the first three image components, illustrated by subtracting (top row) and adding the component to the average face (though note that the sign of the change is arbitrary). These early components are largely dominated by lighting and hair effects, but note that the second strongly codes face sex.

Principal components analysis is a standard statistical technique for reducing large volumes of data, which has been proposed as a model of face recognition.[a] The analysis yields a basis set for coding faces in a compact fashion, perhaps 20-50 real numbers, which can be matched easily with those in a database. Recognition performance can be improved by first standardising the shape of each face, by marking key features such as eyes and nose and then morphing the face to the average positions.[b] This leaves the image PCA with a more linear space, since the features all appear in the same location on each face. The resultant *eigenfaces* may best be illustrated by adding and subtracting them from the average face, see Figure I above. This set came from an analysis of 50 male and 50 female faces. The second component codes strongly for sex, as was found by O'Toole et al.[c], but clearer here because of the initial shape averaging.

The shape vectors that result from defining the location of key features may also be subjected to PCA to reveal the major modes of variation of the configurations of the features, *eigenshapes*. The effects of these can also be shown by adding and subtracting them from the average face shape,[d] see Figure II. Animations of these changes may be seen at http://www.stir.ac.uk/psychology/Staff/pjbh1. The codes for the image and shape components of each face form the input to the integrated model of face recognition described in the main text.

PCA is able to match some aspects of human performance, such as the other race effect.[e] If PCA is performed on a set of faces of one race, then other faces from that race are coded better than faces from a different race. The other race varies along different dimensions, which are not well captured by those derived from the first set. However, as the figure above indicates, PCA on raw images is very sensitive to variations in image properties such as lighting. As is shown in Box 2, human recognition of unfamiliar faces is also affected by changes of the image.

Figure II From left to right, the first two and the ninth shape components, illustrated by subtracting (top row) and adding the component to the shape of the average shape-free face. The first codes head size, along with an element of face sex (women in this set have smaller heads, even after normalising for pupil centres). The ninth is included because it clearly captures another aspect of sex differences.

References

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

b. Craw, I. and Cameron, P. (1991) Parameterising images for recognition and reconstruction. Mowforth, P. *Proceedings of the British Machine Vision Conference*. pp 367-370. London, Springer-Verlag.

c. O'Toole, A. J. et al. (1991) Classifying faces by race and sex using an autoassociative memory trained for recognition. Hammond, K. J. and Gentner, D. *Proceedings of the Thirteenth Annual Conference of the Cognitive Science Society*. pp 847-851. Hillsdale, NJ, Erlbaum.

d. Lanitis, A. et al. (1995) Automatic Face Identification System Using Flexible Appearance Models *Image and Vision Computing* 13, 393-401.

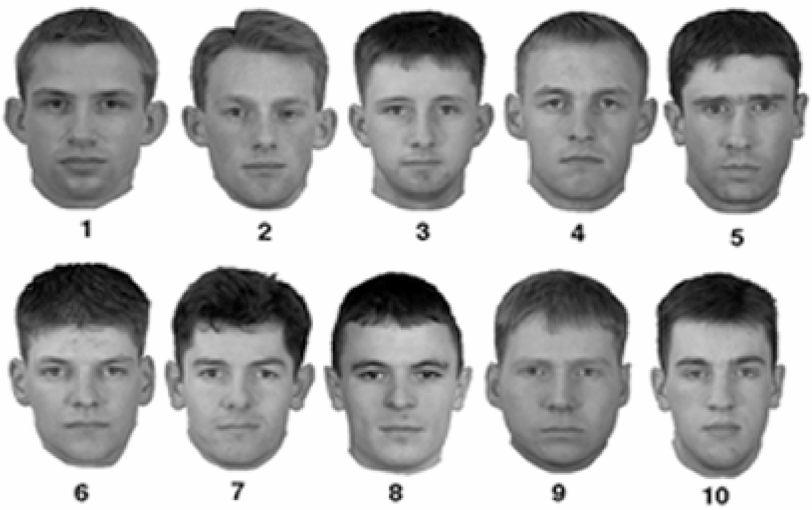
e. O'Toole, A. J. et al. (1994) Structural aspects of face recognition and the other-race effect. *Memory and Cognition* 22, 208-224.

Box 2: Matching unfamiliar faces by people and computers

Compare the face at the top with the ten shown below. Which one is the same as the person shown at the top? This is an example of the kind of task reported by Bruce et al.[a] Participants were shown a single target video-image presented with 10 photographs. Their task was to match the target against the photographic arrays. In the first of this series of experiments, the target was present in half the arrays and participants were asked to determine whether or not he was present, and which one if present. In later experiments all arrays contained a target, and participants were asked to find the closest matching face from the array. The experiments produced surprising results. Despite the high quality images, a large number of errors arose even when viewpoints were coincident as illustrated here. On average, errors were made on about 30% of trials when the target was present on only half the trials, and on 20% of trials when the target was always present. Performance dropped further when there was a change in viewpoint or expression between the faces to be compared. Performance in these studies was not improved significantly if participants were shown short video clips rather than still images for comparison. Nor did it make any difference whether the target, array faces, or both were shown in colour rather than grey-scale. We also showed that matching performance was dominated by the external features of the faces to be matched. If only the internal features of the target (top) image were shown, performance was incorrect 51% of the time.

These results are important, as they suggest that image quality is not the only factor which limits abilities to identify unknown faces shown on video images. Moreover, it underlines the need to examine automatic methods to improve face verification.

In parallel studies, we (Miller et al. unpublished manuscript) have examined how well computer image analysis systems can do the same task, using the same face sets as humans. We have compared two automated image analysis systems. The automated systems are Principal Components Analysis (PCA, see box 1) and Gabor Wavelet analysis.[b] Neither of the systems we examined was very successful at the tasks given, and in general, neither was as good as human vision. On the task illustrated in this box, using full-face neutral expression images, human vision was 78% correct; the graph-matching system managed 70% correct, but PCA based upon the first 40 components managed only 59% correct. However, a modification of the PCA-based system to use Mahalanobis rather than Euclidean distance[c] - in some respects a means of compensating for the differences in lighting between different photographic media - did yield performance which could exceed human perception (86% using 40 components), though only when viewpoints of the faces to be matched were closely coincident. When there was variation in the viewpoints to be compared, human performance by far surpassed any of the PCA-based systems.



References

a. Bruce, V. et al. (In Press) Verification of face identities from images captured on video. *Journal of Experimental Psychology: Applied*.

b. Wiskott, L. et al. (1997) Face recognition by elastic bunch graph matching. *IEEE transactions on pattern analysis and machine intelligence* 19(7), 775-779.

c. Craw, I. et al. (1999) How should we represent faces for automatic recognition? *IEEE Transactions on Pattern Analysis and Machine Intelligence* 21(8), 725-736.