Comparisons between human and computer recognition of faces

Bruce, Vicki; Hancock, Peter J.B.; Burton, A. Mike

Published in:
Proceedings - 3rd IEEE International Conference on Automatic Face and Gesture Recognition, FG 1998

DOI:
10.1109/AFGR.1998.670983

Licence:
Other

Comparisons between Human and Computer Recognition of Faces

Vicki Bruce and Peter J.B. Hancock
Department of Psychology
University of Stirling
Stirling FK9 4LA
Scotland, UK

A. Mike Burton
Department of Psychology
University of Glasgow
Glasgow G12 8QB
Scotland, UK

Abstract

This paper reviews characteristics of human face recognition that should be reflected in any psychologically plausible computational model of face recognition. We then summarise recent results which compare aspects of human face perception and memory with the performance of two computer models which each claim some degree of biological plausibility. We show how the performance of each is correlated with human performance on the same images, but that each explains rather different aspects of human performance with these faces. We conclude with a discussion of the coding of image sequences by humans and computers.

1: Introduction

There is no necessary link between techniques developed by engineers to automate face recognition, and natural mechanisms used by the human visual system to achieve the same end. For example, recognition of individual iris patterns is currently attracting considerable attention as a reliable way of recognising individuals automatically [1], but such patterns play little or no part in natural face recognition. Nonetheless it is often instructive to compare natural and engineered face recognition processes. Dramatic claims are sometimes made about the biological plausibility of particular face recognition algorithms, but usually in the absence of careful examination of the extent to which human and machine performance coincide when each is given a similar task to perform with the same set of images. Conversely, psychologists and neuroscientists may draw unsophisticated theoretical conclusions from their observations if they work in ignorance of computational possibilities. In this paper we first review some of the characteristics of human face recognition that must be reflected in any computational model of the process, and then go on to describe recent results which compare aspects of human face perception and memory with the performance of two computer models each claiming some degree of biological plausibility. Finally, we suggest how future models will need to pay serious attention to coding of multiple images of each encountered face.

2: Characteristics of human face recognition

Space does not permit a full review of characteristics of human face recognition (for a more extensive treatment comparing human and computer recognition see [2]. For more extensive general reviews of the field see edited collections [3] and [4]). Psychological research over the past 20 years or so has shown rather convincingly that human vision processes upright faces in a ‘holistic’ or configural way, rather than as a set of independent facial features (for examples see [5, 6]). Importantly, human face matching of unfamiliar face images is badly affected by variations in this holistic image created by changes in viewpoint or illumination. For example, Bruce [7] found recognition memory for unfamiliar faces dropped substantially when there was a change in viewpoint and/or expression between study and test (see [8] for some recent data on viewpoint dependency). Such effects are important because they are relevant to the question of whether human face recognition operates by coding 2D intensity images in a fairly low-level way (as occurs in many successful computer models of face recognition), or whether human vision derives a 3D model from these 2D intensity images.

Adini, Moses and Ullman [9] showed that changes in lighting, like those of viewpoint, can produce greater changes in a face image than a change in identity. Moreover, they showed that supposedly lighting-independent representations such as edge representations and Gabor convolutions were much more badly disrupted by changes in lighting than human observers were in their experiments.

However, other studies have shown that people can also find face matching very difficult when there are changes in lighting direction between the images to be compared. Johnston, Hill and Carmen [10] showed that familiar faces were more difficult to recognise when lit from below than when lit from above, an effect which they suggested could contribute to the well-documented effects of face inversion, since upside-down faces will tend to be lit from below. Consistent with this, they found that effects of inversion were less for bottom-lit faces (which appeared top-lit when inverted) than for top-lit faces (which appeared bottom-lit when inverted), though this could not account entirely for the effects of inversion. Hill and Bruce [11] showed that matching face surfaces for identity was made much more difficult when the surfaces were lit from different directions. In their experiments subjects had to determine whether two face images were of the same or different people, with viewpoint (three-quarter view or profile) and lighting direction (top or bottom) either matched or mis-matched across these images. They found that a change in lighting made matching more difficult even when matching viewpoints were presented in profile, where it might have been assumed that the occluding contour could provide a lighting-invariant cue. Moreover, matching across different viewpoints was very much better when the surfaces were lit from above rather than below, and a series of control experiments using faces and non-face objects in different orientations revealed this to be a genuine benefit of lighting from above, rather than an artefact of the different features visible under different lighting conditions. One explanation of these findings is that they reveal the use of shape-from-shading processes in the construction of face representations which incorporate the assumption that lighting comes from above [12].

The effects of lighting change on face identification and matching suggest that representations for face recognition are crucially affected by changes in low-level image features. This conclusion is reinforced by the dramatic effects of photographic negation on face recognition (e.g. [13]). Bruce and Langton [14] showed that negation had an even greater detrimental effect than inversion on the identification of familiar faces. They went on to examine whether this was because negation affected patterns of shading, and hence the derivation of 3D shape from shading, by examining how negation affected the recognition and matching of surface images derived from laser scanning which lacked any pigmented or textured features. Bruce and Langton found that negation had no significant effect on the recognition and accuracy of matching of these surface images and this led them to attribute the negation effect to the alteration of brightness information about pigmented areas. A negative image of a dark-haired Caucasian, for example, will appear to be a blonde with dark skin. However, Kemp et al [15] showed that the hue values of these pigmented regions do not themselves matter for face identification. Familiar faces presented in "hue negated" versions (where, for example, red areas are replaced with green) but with luminance values preserved, were recognised as well as those with original hue values maintained, though there was a decrement in recognition memory for pictures of faces when hue was altered in this way. This suggests that episodic memory for pictures of unfamiliar faces can be sensitive to hue, though representations of familiar faces seems not to be. This distinction between memory for *pictures* and for *faces* is an important one to which we return. Kemp et al suggested that their results favoured an explanation of the negation effect in terms of the disruption of shape from shading processes.

Kemp et al's demonstrations that hue values are unimportant for face identification are among a number of studies showing that the image features mediating face recognition appear to be gross rather than precise. Areas of light and dark need to be preserved for good identification of faces, and line drawings which lack these features are very much poorer representations for identity than those which preserve them [16, 17]. However, these features need only be preserved at relatively coarse scale, and face identification is possible when much of the finer scale information is removed by spatial filtering. For example, Bachmann [18] followed up early work by Harmon and others [19] on pixellation and found that participants were quite well able to identify one of a small number of target faces provided that more than 15 pixels horizontally were used.

In summary, effects of viewpoint, lighting, negation and the overall importance of relatively coarse scale information about patterns of light and dark in face images are consistent with the use of relatively low-level, coarse scale image features in the identification of faces. Some of these findings, however, additionally point to the possible use of patterns of shading for the derivation of 3D models for face recognition. However, if 3D models are derived in face recognition, it is difficult to understand why face recognition appears so viewpoint dependent. This underlines the importance of distinguishing between different uses made of facial information when considering representational possibilities. A 3D model of a face may be used to help parse or normalise an image, and certainly will have to be made explicit to mediate physical actions made to faces (e.g. to kiss or stroke a face), but this does not necessarily mean that 3D models are useful for identification. Recognition of surface images devoid of pigmentation and texture is very poor [20], and classical sculptors used to paint their busts, additional hints that pigmented features carry important information for face individuation.

3: Comparisons with computer models

The weight of evidence briefly reviewed above suggests that representations for human face recognition are based upon the analysis of relatively low level image-features from the whole facial pattern, rather than on more abstract derived measurements, though this is not to deny a possible role for an abstract model of a face which may be used in the alignment of face images. Given this, it is interesting that in the recent "FERET" competition funded by the US Army Research Laboratory to find the most successful artificial face recognition method [21], two of the most successful systems were those of Pentland et al. which is based upon Principal Components Analysis (PCA) of image pixel values [22, 23] and of von der Malsburg and colleagues based upon graph matching of Gabor wavelets [24, 25].

Each of these systems has also claimed some psychological and/or neurobiological plausibility, but without comparing each of these models with human vision directly on the same tasks, and using the same images, it is impossible to tell how reasonable these claims are, and whether any similarity to human vision arises because of the specific or more general (e.g. image-based) nature of the systems. Unfortunately for these purposes the FERET test uses image sets too large for comparative performance with people to be obtained. However, in recent research in Stirling and Glasgow, we have compared the performance of human perception and memory for faces with that of these two different computer-based systems.

Our investigations of the PCA system have been the more extensive. Analysis of the correlations between pixel intensity values of a set of face images yields a set of "eigenfaces" [22], cf. Sirovich and Kirby [26], which can be used to describe and code new faces. The codes for new faces can be compared with those stored in order to recognise faces as those of specific known individuals. However, the success of this technique for compact coding and recognition of faces depends crucially on the alignment of the images. Typically, image sets are approximately standardised by alignment of the eyes of each face. In our work we have followed the suggestion of Craw and Cameron [27] and aligned faces carefully by morphing them all to a common "shape-free" shape. (Currently the morphing relies on the manual location of a set of key points on each face image, but various techniques including optical flow are potentially available to automate this stage). Thus each face provides a shape vector (which describes how it departs from the shape-free norm), and a "texture" vector, which describes the 2D array of intensities in the "shape-free" version of the face. The texture vector also contains some information about shape, of course. PCA is done separately on the texture and shape vectors associated with each face. Using this technique, we [28] confirmed earlier results by O'Toole and her colleagues [29, 30] in showing that there was a strong correlation between human and PCA recognition memory performance when both were tested with the same set of face images. Moreover, we showed that this correlation was improved by the operation of separately coding for shape and texture (in the shape-free images).

However, these investigations were limited in that they were effectively exploring human and PCA picture memory rather than face recognition, since the study and test images used in the human and computational experiments were the same. More recently, we [31] have extended our work to investigate what happens when people and the computer systems are asked to identify the same person shown in varying images, and also explored in more detail how closely human and computer estimates of similarities between different people compare. In these investigations we have directly compared the PCA-based system with that developed by von der Malsburg and colleagues at Bochum (see [24], and [25] for more recent developments). In this system, faces are coded by families of Gabor-type wavelets, at several scales and orientations, located at a number of places around the face. These locations are found automatically for a new face by comparison with stored models. The face locations form a labelled graph of activity vectors (known as "jets") of the wavelets attached to every vertex on the graph. A graph is stored of each known face, and the derived graphs of test faces are distorted to form the best possible match against each of the stored graphs in turn. The stored face which yields the least distortion is taken as the match.

We obtained a set of images of the faces of fifty different young men, each in neutral plus one or more different facial expressions. The neutral (N) faces were coded by the PCA (making use of the shape-free transformation described above) and by the graph-matching systems. We then used the set of varying expression (E) images as test faces for each of these systems. Each performed extremely well at recognising the changed individuals. Of more interest is to compare this performance (specifically, the confidence of each of these matches) with human memory performance using the same faces. We compared confidence measures for each of the NE comparisons in the computer models with a number of different memory measures obtained from participants who were asked to try to recognise which were old or new items when tested with the same (N followed by N) or different (N followed by E) images. Interestingly, the graph-matching system gave similar, and significant, correlations between its confidence doing NE matches, and human performance in both the NN and the NE task (correlation coefficients of 0.33 and 0.32 respectively). However, the PCA system gave a greater correlation with human performance obtained when matching identical images in the NN condition (correlation with hit rate of 0.41), but a much smaller, and non-significant correlation with the NE data (correlation with hit rate of only 0.17). Thus the PCA system performance when matching different images of the same people's faces co-varies with human matching of identical images of these faces. These data suggest that PCA does a better job of accounting for the similarity that people see between specific images of faces, while graph-matching may do a better job at accounting for similarity between faces.

Similar conclusions were reached when we turned to examine how well each system accounted for the similarities seen between different people's faces. To do this, we compared human judgements of the similarities seen between each of the 50 faces in the set with judgements of these similarities obtained from each of the computer systems by examining their rank ordering of the goodness of match to each of the non-targets as each target image was matched. The human judgements of similarity were obtained by the simple method of asking observers to sort the faces into piles of similar appearance, and counting the number of times each pair was sorted together as a measure of similarity. Forty observers sorted the face images with hair visible, a further forty sorted the images with hair removed, in order to get measures of similarity less dominated by hairstyle. While this method of obtaining similarity ratings is simple, it is also rather crude, and there are many ties in the human data obtained in this way. Nonetheless, each computer system yielded significant (though numerically small) correlations with human similarity data. However, again the pattern of correlations was rather different between the two systems. The graph matching system produced similar correlations to the human ratings to faces shown both with and without hair, but the PCA system gave much higher correlations to the ratings obtained with hair.

How can we interpret these findings? It seems that each system is predicting (some of the) variance in human performance but in slightly different ways. The graph-matching system gives a better account of how people recognise faces when images vary, while the PCA system provides rather a good account of the coding of specific images of individual faces. In other words, PCA may provide a better model for human picture memory, and graph-matching a better model of human face memory. Both these processes (which Bruce and Young [32] referred to as "pictorial" and "structural" coding) are important in human face recognition, with the recognition of relatively unfamiliar faces dominated by pictorial processes and that of more familiar faces dominated by structural coding. It seems that each of the computer systems captures important, but different, aspects of human face coding.

4: Multiple images of faces

The comparisons reported above were between human and computer analysis of faces where individual face exemplars were coded and remembered. This is not a very ecologically valid situation. While human ability to remember and compare individual face images is an activity which is of interest in the modern world, where we use individual images to access identities and as identifiers, it is not a natural activity. Human brains evolved means of encoding individuating descriptions from dynamic face sequences. How do people build up stable representations of faces from varying facial images of them? The answers to this may provide further clues which will be useful to engineers in the future.

There is some evidence that variations in facial appearance may lead to the storage of a "prototype" abstracted from these varying images. Posner and Keele [33] showed people dot patterns which varied around a central but unseen prototype pattern. Following training on patterns which were moderate distortions from this prototype, the prototype itself was subsequently classified as accurately as any of the actually studied patterns. Solso and McCarthy [34] demonstrated a similar effect using faces composed from varying facial features using Identikit II. Prototype faces were constructed from particular features, and participants studied variants of these which differed in one or more of the features. At test, the previously unseen prototype which was composed of the most commonly presented features was falsely recognised as an old face more confidently than any of the faces which had actually been studied.

The problem with experiments such as Solso and McCarthy's is that it is not clear whether it is tapping what happens when variations of a single face are seen (e.g. the variations of a specific individual's face over fast changes in expression or slower ones such as weight or age changes), or whether it is tapping the representation of several different faces in memory (with different eyes, noses, etc.) in a way which makes more typical ones seem falsely familiar. The use of varying exemplars which are comprised of different facial features confounds these two possibilities.

Bruce, Doyle, Dench and Burton [35] examined prototype effects using a number of distinctly different face identities constructed from a computer-based kit of face features. Each different face was itself varied in terms of the placement of its facial features, by moving the internal features of the face up and down the image by specific numbers of pixels (e.g. moving the eyes, nose and mouth upwards or downwards within the face frame by 2, 4 or more pixels). Such variations, at least when minor, resemble very approximately changes in face shape during ageing and so are plausible as variations which could occur to the appearance of an individual face. When participants were shown extreme variants of each of these identities in an incidental memory task, they later found the unseen prototype images as familiar as any exemplars that they had actually studied. For example, if participants were only shown variants of a particular face with its features displaced up or down by 10 pixels, they found the original, zero displacement face as familiar as the ones they had earlier studied. Bruce [36] and Cabeza, Bruce, Kato and Oda [37] found similar effects using images of real faces rather than schematic ones.

Bruce [36] and Cabeza et al [37] went on to explore the limits of this prototype effect, and found that while variations of faces within the same viewpoint gave rise to prototype effects, variations in head angle did not. So, for example, if participants studied faces whose head angle was shown at plus or minus 30 degrees from an unstudied prototype angle, they did not later tend to find the prototype angle as familiar as the studied ones. These findings suggest that different exemplars of the same person's face may somehow be amalgamated in memory (averaging is one possible form of amalgamation) but in a viewpoint-specific way, such that representations are established separately at a (probably small) number of discrete angles.

5: Moving images of faces.

While discussion of the effects of multiple exemplars on face encoding brings us a step closer to natural face recognition, we must also consider the possible additional effects of facial motion on representations for face recognition. A dynamic face image sequence of, say, 25 frames per second, potentially contains useful information both by virtue of the multiple images presented in such a sequence, with their variations in viewpoint, expression etc., and by virtue of dynamic information itself, which might, for example, yield a better representation of the face in 3D than static images alone.

Recently, Knight and Johnston [38] showed that the identification of famous faces shown in photographic negative could be significantly enhanced when the faces were shown moving rather than static. In follow up research in Stirling, Karen Lander has extended this finding to other conditions where identification is made difficult, e.g. by thresholding the images or showing them in blurred or pixellated formats. We have also checked that the effect is not due to the additional static information in multiple views, since a benefit for seeing animated sequences is found even after attempts are made to equate the static information in moving and static conditions. Although we must be cautious in generalising results obtained in the recognition of degraded, famous faces to those relevant to recognition of faces seen in more natural conditions, these data certainly suggest that information is contained in patterns of natural animation that can prove useful for recognition in at least some circumstances.

In our own work we have been less successful in demonstrating any advantage for animated sequences in the recognition of previously unfamiliar faces. Christie and Bruce [39] showed no advantage for the recognition of novel views of previously unfamiliar faces when the faces had been studied in animated compared with non-animated image sequences. However, some advantages for animated sequences were found in studies reported by Hill et al [8] and by Pike et al [40], so it may be that such effects depend rather critically on the range and extent of motion shown. Certainly, for movement to benefit the identification of famous faces it must enter into face representations at some stage, and our current work is pursuing these questions.

Such questions about the role of multiple and animated image sequences for human recognition of faces become important in the context of increased reliance on security video surveillance systems for capturing and establishing the identities of criminals. We do not yet know what consequences there may be of, for example, sparse storage of discrete samples of such videos, nor in what ways such footage should be shown to people to maximise the chance of accurate identification of the people shown. At a theoretical level, more psychological models of face storage and recognition should be extended to address explicitly the issue of how representations are derived and accessed from such sequences.

Some computational models are already exploring the use of image sequences to derive invariant information about individual face identities, as well as information specifying pose, e.g. [41, 42]. As far as we are aware such studies have not yet incorporated direct comparisons with human performance in similar circumstances. The further extension of both human experimental studies and computational modelling to image sequences is a promising and important one for future studies.

References

[1] J. Daugman. Phenotypic vs. genotypic approaches to face recognition. In Wechsler, H. et al (Eds). *Face recognition: From theory to applications*. Springer, 1998.

[2] V. Bruce, P.J.B. Hancock and A.M. Burton. Human face perception and identification. In Wechsler, H. et al (Eds). *Face recognition: From theory to applications*. Springer, 1998.

[3] A.W. Young and H.D. Ellis (Eds). *Handbook of research on face processing*. Amsterdam, North Holland, 1989.

[4] V. Bruce, A. Cowey, A.W. Ellis and D.I. Perrett (Eds) *Processing the facial image*, Oxford University Press, 1992.

[5] A.W. Young, D.J. Hellawell and D.C. Hay. Configurational information in face perception. *Perception*, 16, 747-59, 1987.

[6] J.W. Tanaka and M.J. Farah. Parts and wholes in face recognition. *Quarterly Journal of Experimental Psychology*, 46A, 225-46, 1993.

[7] V. Bruce. Changing faces: Visual and non-visual coding processes in face recognition. *British Journal of Psychology,* 73, 105-116, 1982.
[8] H. Hill, P.G. Schyns and S. Akamatsu. Information and viewpoint dependence in face recognition, *Cognition,* 62, 201-222, 1997.
[9] Y. Adini, Y. Moses, and S. Ullman. Face recognition: The problem of compensating for changes in illumination direction. *IEEE Transactions on pattern analysis and machine intelligence,* 19, 721-732, 1997.
[10] A. Johnston, H. Hill, and N. Carman. Recognising faces: effects of lighting direction, inversion and brightness reversal. *Perception,* 21, 365-375, 1992.
[11] H. Hill, and V. Bruce. Effects of lighting on matching facial surfaces. *Journal of Experimental Psychology: Human Perception and Performance,* 22, 986-1004, 1996.
[12] V. Ramachandran. Perception of shape from shading. *Nature,* 331, 163-166, 1988.
[13] R.E. Galper. Recognition of faces in photographic negative. *Psychonomic Science,* 19, 207-208, 1970.
[14] V. Bruce, and S. Langton. The use of pigmentation and shading information in recognising the sex and identities of faces. *Perception* , 23, 803-822, 1994.
[15] R. Kemp, G. Pike, P. White and A. Musselman, A. Perception and recognition of normal and negative faces - the role of shape from shading and pigmentation cues. *Perception,* 25, 37-52, 1996.
[16] V. Bruce, E. Hanna, N. Dench, P. Healy and A.M. Burton. The importance of "mass" in line drawings of faces. *Applied Cognitive Psychology,* 6, 619-628, 1992.
[17] G.M. Davies, H.D. Ellis and J.W. Shepherd. Face recognition accuracy as a function of mode of representation. *Journal of Applied Psychology,* 63, 180-187, 1978.
[18] T. Bachmann. Identification of spatially quantised tachistoscopic images of faces: How many pixels does it take to carry identity? *European Journal of Cognitive Psychology*, 3, 87-103, 1991.
[19] L.D. Harmon and B. Julesz. Masking in visual recognition: Effects of two-dimensional filtered noise. Science, 180, 1194-1197, 1973.
[20] V. Bruce, P. Healey, A.M. Burton, T. Doyle, A. Coombes, and A. Linney. Recognising facial surfaces. *Perception,* 20, 755-769, 1991.
[21] P.J. Phillips. Foundations of face recognition. In Wechsler, H. et al (Eds). *Face recognition: From theory to applications*. Springer, 1998.
[22] M. Turk and A. Pentland. Eigenfaces for recognition. *Journal of Cognitive Neuroscience,* 3, 71-86, 1991.
[23] A. Pentland, B. Moghaddam and T. Starber. View-based and modular eigenspaces for face recognition. *Proceedings of IEEE Computer Society conference on computer vision and pattern recognition,* 84-91, 1994.
[24] M. Lades, J.C.Vorbruggen, J. Buhmann, J. Lage, C. von der Malsburg, R.P. Wurtz, and W. Konen. Distortion invariant object recognition in the dynamic link architecture. *IEEE Transactions on Computers,* 42, 300-311, 1994.
[25] L. Wiskott, J.M. Fellous, N. Kruger, and C. von der Malsburg. Face recognition by elastic bunch graph matching *IEEE Transactions on pattern analysis and machine intelligence,* 19, 775-779, 1997.
[26] L. Sirovich and M. Kirby. Low dimensional procedure for the characterisation of human faces. *Journal of the Optical Society of America,* 4, 519-524, 1987.
[27] I. Craw and P. Cameron. Parameterised images for recognition and reconstruction. In P. Mowforth (Ed.), Proceedings of the British Machine Vision Conference. Berlin: Springer-Verlag. 1991.
[28] P.J.B. Hancock, A.M. Burton and V. Bruce. Face processing: human perception and principal components analysis. *Memory & Cognition*, 24, 26-40, 1996.
[29] A.J. O'Toole, H. Abdi, K.A. Deffenbacher, D. Valentin, D. and J.C. Bartlett. Simulating the "other-race effect" as a problem in perceptual learning. *Connection Science,* 3, 163-178, 1991.
[30] A.J. O'Toole, K.A. Deffenbacher, D. Valentin and H. Abdi. Structural aspects of face recognition and the other-race effect. *Memory & Cognition,* 22, 208-224, 1994.
[31] P.J.B. Hancock, V. Bruce and A.M. Burton. A comparison of two computer-based face identification systems with human perceptions of faces. *Vision Research,* 1998.
[32] V. Bruce and A.W. Young. Understanding face recognition. British Journal of Psychology, 77, 1986.
[33] M. Posner and S.W. Keele. On the genesis of abstract ideas. *Journal of Experimental Psychology*, 77, 353-363, 1968.
[34] R.L. Solso and J.E. McCarthy. Prototype formatio of faces: A case of pseudo-memory. *British Journal of Psychology*, 72, 499-503, 1981.
[35] V. Bruce, T. Doyle, N. Dench and M. Burton. Remembering facial configurations. *Cognition,* 38, 109-144, 1991.
[36] V. Bruce. Stability from variation: the case of face recognition. *Quarterly Journal of Experimental Psychology,* 47A, 5-28, 1994.
[37] R. Cabeza, V. Bruce, T. Kato and M. Oda. The prototype effect in face recognition: Extension and limits. *Memory and Cognition.* 1998.
[38] B. Knight and A. Johnston. The role of movement in face recognition. *Visual Cognition*, 4, 265-274, 1997.
[39] F. Christie and V. Bruce. The role of movement in the recognition of unfamiliar faces. *Memory and Cognition*, 1998.
[40] G.E. Pike, R.I. Kemp, N.A. Towell and K.C. Phillips. Recognizing moving faces: the relative contribution of motion and perspective information. *Visual Cognition*, 4, 409-437, 1997.
[41] A. Psarrou, S. Gong and H. Buxton. Modelling spatio-temporal trajectories and face signatures on partially recurrent networks. Proceedings International Conference on Neural Networks, ICNN'95, Ch 620, 2226-31, 1995.
[42] D.B. Graham and N.M. Allison. Characterising virtual eigensignatures for general purpose face recognition. In Wechsler, H. et al (Eds). *Face recognition: From theory to applications*. Springer, 1998.