Face Recognition: the Problem of Compensating for Changes in Illumination Direction

Yael Moses*, Yael Adini, and Shimon Ullman

Department of Applied Mathematics and Computer Science,
The Weizmann Institute of Science, Rehovot 76100, Israel.
* Current address: Robotics Group, Dept Engineering Science,
Parks Rd, Oxford University, OX1 3PJ, UK
e-mail: yael@robots.ox.ac.uk

Abstract. Recognizing faces is a difficult problem due to the generally similar shape of faces combined with the considerable variability in images of the same face under different viewing conditions. In this paper we consider image variation due mainly to illumination conditions. We study several image representations that are often considered insensitive to changes of illumination conditions, such as edge maps, derivatives of the grey-level image, and the image convolved with Gabor filters. For each of these representations, we compare the differences between images of the same face under different imaging conditions, with differences between images of distinct faces. The comparison is performed using a controlled database of faces, in which each of the imaging parameters (illumination, viewing position, and expression) is controlled separately. The main result of these studies is that the variations between the images of the same face due to illumination and viewing directions are almost always larger than image variations due to a change in face identity. For illumination changes, this reversal is almost complete except for representations that emphasized the horizontal features. However, even for these representations, systems based only on comparing such representations will fail to recognize up to 30% of the faces in our database. We conclude that these representations are insufficient by themselves to overcome the variation between images due to changes in illumination direction as well as changes due to viewing position and expression.

1 Introduction

Recognizing faces is a difficult problem due to the generally similar shape of faces together with the considerable variations between images of the same face. The image of a face changes with facial expression, age, viewing position, illumination conditions, noise, etc. The task of a face recognition system is to recognize a face in a manner that is as independent as possible of these image variations. Psychophysical experiments show that the human visual system can identify faces from their novel images despite considerable variations due to illumination direction (Moses et al., 1993) and viewpoint changes (Patterson and Baddeley, 1977; Davies et al., 1978; Bruce, 1982; Moses et al., 1993).

Lecture Notes in Computer Science, Vol. 800
Jan-Olof Eklundh (Ed.)
Computer Vision – ECCV '94

The question then is how a face recognition system can identify a face despite the variation between different images of the face. In this paper we focus mainly on variations due to illumination changes, and briefly consider variations due to changes in viewing position and expression. Two main approaches for dealing with the variation between images due to illumination changes have been used in the past. The first uses the grey-level information to extract the three dimensional shape of the object, namely, a shape from shading approach (e.g. Horn and Brooks,1989). The second approach constructs a representation of the image and the model that is relatively insensitive to illumination using a local operator. An example of such a representation is edge maps extracted from the grey-level images (e.g. Marr and Hildreth, 1980; Canny, 1986).

Extracting shape from shading from an image is an ill-posed problem. All proposed solutions to this problem make assumptions about either the object shape and reflectance properties, or the illumination condition. Such assumptions may be too strict for the face recognition task. We therefore consider here representations that are often used by face recognition systems, and considered insensitive to illumination changes. The edges of the image grey-level function are often considered the basic image representation for general object recognition and in particular for face recognition (Kanade, 1977; Wong et al., 1989; Govindaraju et al., 1989; Brunelli and Poggio, 1991). For biological visual systems, there is both physiological and psychophysical evidence that incoming images are first passed through local multiple, parallel channels that are both orientation and spatial frequency specific (Wilson and Bergen, 1979; Watson and Robson, 1981; Daugman, 1987). It was suggested that the response profile of a simple cell could be described by Gabor-like or DOG functions (Daugman, 1984; De-Valios and De-Valios, 1988; Mercelja, 1980; Marr and Hildreth, 1980; Pollen and Ronner, 1983). Gabor functions are also used by several artificial face recognition systems that filter the grey-level image before they attempt to recognize the faces in the image (Brunelli and Poggio, 1991; Buhmann et al., 1993; Manjunath et al., 1992). Other image representations used by face recognition systems to reduce the effect of changes of illumination conditions on face images include the derivatives of the grey level distribution (Brunelli and Poggio, 1991; Edelman et al., 1992), and in addition to the above transformations, a non linear transformation such as logarithmic transformation of the image intensities (Reisfeld and Yeshurun, 1992).

Ideally, an image representation used for recognition should in particular be invariant to illumination changes. It has been shown theoretically that for the general case, a function invariant to illumination does not exist (Moses and Ullman, 1992). Similar results regarding variations due to changes of viewing position was shown by (Burns et al., 1992; Clemens and Jacobs, 1991; Moses and Ullman, 1992). The objects considered by Moses and Ullman were unconstrained 3D objects, consisting of n independent patches in space. For recognition systems that are limited to a certain classes of objects, this limitation does not necessarily apply. We therefore consider image and model representations that are relatively insensitive to illumination for specific classes of objects, e.g. faces. This includes the following image representations: the original grey-level image, the edge map of the image, the image filtered with Gabor functions, and the first and second derivatives of the grey-level image. In addition, several of these representations were also further processed by a log function of the intensity.

The question we address is whether these image representations are sufficient by themselves to overcome the variation between images due to illumination direction. Often, when these representations are used by recognition systems, their evaluation is determined by demonstrating the performance of the whole system on a relatively small database of faces. These databases are not guaranteed to have variations between images due to each of the separate imaging parameters (e.g. illumination condition), and it is therefore impossible to evaluate a single component of the system (e.g. the component that deals with variation due to illumination changes). When recognition systems are developed, it is important to analyze how well they actually cope with variations due to a given imaging parameter. When no theoretical analysis exists, it is important to have a well controlled database to examine the systems performance.

To study the performance of common recognition techniques systematically, we constructed a special database of faces, in which the face identity and each of the following imaging parameters was controlled separately: illumination, viewing position, and expression. The actual distances between pairs of images (or image representations) that vary due to change in face identity were computed and compared with the variation between pairs that vary due to a change in each of the other imaging parameters.

In the rest of the paper, a description of the empirical study of the image variation (Section 2) and its results (Section 3) are given. Finally a discussion of these results is presented (Section 4). The full details of the methods and results can be found in (Adini et al., 1993).

2 Methods

Face database: Five images (size 256×176 pixels) of each of 26 different faces were used. All faces were of males with no glasses, beards, or any other distinctive features (see example of four faces in Figure 1). All images were taken under tightly controlled illumination and viewing position conditions. The camera was attached to a robot arm which controlled the camera locations to frontal and 17^o from the frontal view on the horizontal axis (Figure **2a,b**). Normalized frontal views for all faces were obtained by fixing the location of the face symmetry axis, the external corners of the eyes, and the bottom of the nose, before the pictures were taken. Two different illuminations were used by turning on and off fixed light sources (see Figure 2a,e). The subjects were asked to bear a neutral expression, a smile expression, and a drastic expression, and to remain still (see Figure **2a,c,d** respectively). [1]

Masks: To avoid background interference, we extracted and considered only the face part of the image. There is some psychological evidence suggesting that

[1] The images were selected from a larger database, the Weizmann Facebase, that consists of 66 images for a face.

Fig. 1. Four faces out of the 26 faces: frontal view with left illumination.

different face parts make different contribution to face recognition (see review Shepherd *et al.* 1981). We therefore considered several regions of the face separately: the entire face without the hair, the eyes area, and the lower part of the face (see Figure 2e,f,g).

Distance functions: The distance between pairs of images was computed using simple distance measures that are often used to measure distances between general 2-D distributions. These measures can be considered as an indication of an objective distance between the images when no knowledge about the image formation function is used. This includes: *Pointwise distance* which is the average difference between the grey-level values of all pairs of corresponding pixels (i.e., two pixels in the same location); *Affine-GL* distance which is the minimum Euclidean distance between the grey-level values of one of the image and any affine transformations of the grey-level values of the other image. Note that the Affine-GL distance compensates for uniform affine transformation of the grey-level values; *Regional distance* which is the average of the minimum difference between the grey-level value of a pixel, and the grey-level value of each pixel in the (9 x 9) neighbourhood of the corresponding pixel. Note that the regional distance compensates for a displacement of up to 3 pixels of the image in the plane.

Procedure: Three distance functions, defined above, were applied to pairs of images, where the images in each of the pairs differ due to one of the following sources: face identity, illumination, viewing position, smile expression, and drastic expression. These distances were measured between the original image and each of the following image representations:

Gabor Filters: A Gabor function is a multiplication of sinusoidal (or cosine) grid with a Gaussian. Formally, a Cosine and a Sine Gabor functions are given by: $\cos\left[(\frac{2\pi}{\lambda})(x\cos(\theta)+ysin(\theta))\right] e^{\frac{-(r^2)}{2\sigma^2}}$; $\sin\left[(\frac{2\pi}{\lambda})(x\cos(\theta)+ysin(\theta))\right] e^{\frac{-(r^2)}{2\sigma^2}}$, where λ and θ are the grid wave length and orientation, σ is the standard deviation (scale) of the Gaussian, and $r^2 = x^2 + y^2$.

The effects of several parameters of the Gabor function were studied:

- Symmetric *vs.* non-symmetric function (i.e. sine or cosine Gabors).
- The orientation of the grid, denoted by θ, with $\theta = \{0^0, 45^0, 90^o, 135^o\}$.
- The scale of the Gaussian (standard deviation), denoted by σ, with $\sqrt{2}\sigma = \{5, 8, 12, 19\}$ pixels, and $\lambda = \sqrt{2}\sigma/2$.
- The ratio between the scale and the grid wave length, denoted by σ/λ, with

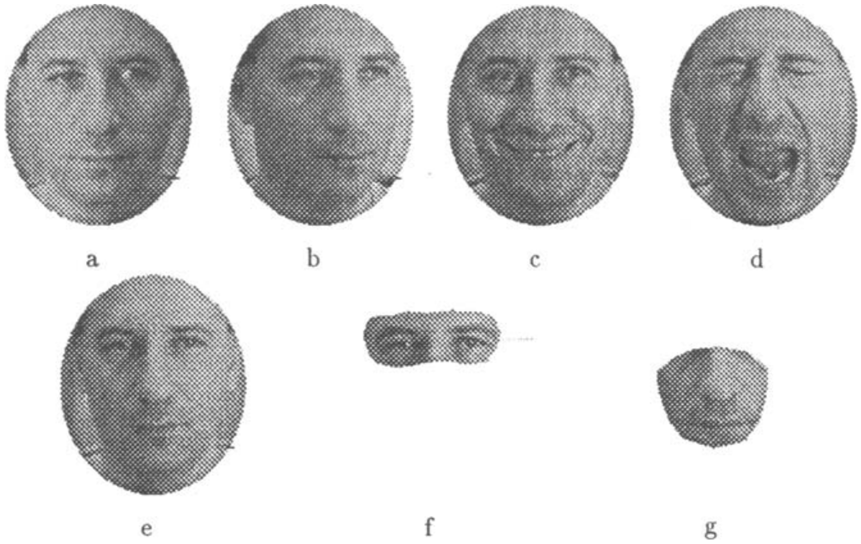

Fig. 2. Five images of the same face: (a) frontal view and left illumination; (b) 17° away from the frontal view on the horizontal axis and left illumination; (c) A smile expression taken from frontal view with left illumination; (d) A drastic expression taken from frontal view with left illumination; (e) frontal view and right illumination The three masks used; (e) The entire face (without the hair); (f) the eyes area; and (g) the lower part of the face.

$\lambda = \{3, 6, 12, 24\}$ for a constant scale ($\sqrt{2}\sigma = 6$), and a constant orientation ($\theta = 90^o$).

Edge representation: There are quite a few edge detectors in the literature. We considered only one of them, the edge representation computed by the DRF edge detector (Shen and Castan, 1987). The edge map of the images were first smoothed by a Gaussian function to obtain a grey-level image, and the same distance functions as in the other representations were then applied. We expect that any simple method that measures the distance between edge maps will give similar qualitative results to the one suggested here.
Laplacian of a Gaussian filter: The Laplacian of a Gaussian filter computes the second derivative of an image that was first blurred by a Gaussian function (Marr and Hildreth, 1980). The Laplacian of a Gaussian filter is given by:
$\nabla^2 G(r) = \frac{-1}{\pi\sigma^4}(1 - \frac{r^2}{2\sigma^2})\, e^{\frac{-r^2}{2\sigma^2}}$. The representation was computed by convolving an image with the Laplacian of a Gaussian (zero-crossings were not computed). The scales of the Gaussian considered were $\sigma = \{2, 4\}$.
Grey-level derivatives: To compute the derivatives of a smoothed grey-level image, the grey-level image was convolved with the derivative of a Gaussian. Three filters were considered: Symmetric derivative, $\frac{r}{\sigma^2}\, e^{\frac{-r^2}{2\sigma^2}}$, with $\sqrt{2}\sigma = \{9, 12\}$.; Horizontal derivative, $\frac{x}{\sigma^2}\, e^{\frac{-x^2}{2\sigma^2}}$, with ($\sqrt{2}\sigma = \{6, 9, 12\}$); Vertical derivative, $\frac{y}{\sigma^2}\, e^{\frac{-y^2}{2\sigma^2}}$, with $\sqrt{2}\sigma = \{6, 9, 12\}$.

Log representation: Several of the above representations were further processed by a log function in order to generate a new representation. These representations include the images convolved with Sine Gabor functions and the first and second derivatives of a Gaussian.

3 Results

In this section we present the results of comparing the variation between images due to face identity with the variation between images of the same face due to all other imaging parameters. Altogether, 17,589 different pairs of images were considered, with three different face masks.

Ideally, we would like the image variation due to face identity always to be larger than the image variation between images of the same face. Since none of the representations considered here has this property, we propose a measure to evaluate how well each of the representations can overcome the variability between images due to a given imaging parameter. The measure is based on the performance of a simple recognition system. The system compares distances between a target image and a model image after processing to some intermediate representation (for example the images are filtered with a Gabor function). The matched model is chosen as the one with a minimal distance to the target image. The database of this system contains one image of each face.

Let a *missed-face* be a face that the the system fails to recognize for a given image representation and a given imaging parameter. We classify a given face as a missed-face if the distance between the two images of this face taken with different values of the imaging parameter is larger than the distance between the target image and one other image in the database (after the appropriate representation is applied). The percentage of missed faces from the set of faces in our database, which we denote by *miss-percentage*, is used to evaluate a given image representation with respect to a given imaging parameter. Zero and 100% *miss-percentages* correspond to perfect recognition and total failure of the system, respectively. The system may fail to recognize a given face because it confuses it with one or more faces from its database. Let the *miss-degree* be the average number of faces that the system confuses for each missed face. If the miss-degree is high, then it is unlikely that the system will misidentify a face due to accidentally similar pairs of faces in the database. (Note that a different possible database for each target image is used to compute the miss-percentage and miss-degree.)

We now present the results of evaluating how well each of the image representations overcome the variation between images of the same face due to a given imaging parameter, using the miss-percentage and miss-degree measures. **Illumination:** for the original images, the miss-percentage was 100%, that is, the recognition system would fail to recognize all faces in the database when the direction of the illumination changes from left to right. Furthermore, the miss-degree for the original images was above 96%. This implies that the system may confuse each face with almost all other faces in its database.

For most image representations considered, the miss-percentage was above 80%. The representations for which the miss-percentage were below 80% were the ones that were sensitive to horizontal features, such as a Sine Gabor filter at orientation of 90^o (horizontal), derivative of a Gaussian in the vertical direction, and some of the edge maps of the images. However, the lowest miss-percentage was still very high (61%). Furthermore, even for the representation with relatively low value of miss-percentage, the miss-degree was above 30%. The lowest miss-percentage (30%, was obtained by further processing of some of these representations by a log function, However, the miss-degree in this case was high (above 53%).

Viewing position: for the original images as well as for all other representations, the miss-percentage was above 84% for each of the face masks. It follows that the recognition system would fail to recognize almost all faces in its database if the target images were taken from a different viewing positions than those of the images in its database (17^o apart on the horizontal axis). Furthermore, the miss-degree for the original images was above 30%, which implies again that it is not accidental that the system cannot identify a face, but that many faces interfere with the identification of each missed face.

Expression: when the variation due to a smile was considered for the original images, the miss-percentage was 0% for the full face mask. However, if the system considers only part of the face, such as the eyes area or the lower part, then the miss-percentage increased to 30% and 60%, respectively. The recognizability of the faces were impaired after several image representations were considered. For example, for the image representation computed by convolving the image with a horizontal Sine Gabor function, the miss-percentage increased to 34% for the entire face. When variation due to a drastic expression was considered, the results were different. For the original images the miss-percentage was 60% for the full face mask and above 80% for the eyes area. As in the smile case, other image representations often increase the miss-percentage.

The full list of results and examples of the images with different representations can be found in (Adini et al., 1993). Similar results were obtained for the Affine-GL and the Regional distance functions, and for the three different face masks.

4 Summary and Discussion

In this work we studied several widely used image representations, computed by local operators, and often considered relatively insensitive to illumination changes. The question we have addressed was whether these image representations were sufficient by themselves to overcome the variation between images due to illumination direction. To study the performance of common recognition techniques systematically, we constructed a special database of faces, in which each of the following imaging parameters was controlled separately: illumination, viewing position, expression, and face identity. The actual distances between pairs of images (or image representations) that vary due to change in face identity were computed and compared with the variation between pairs that vary due to change in each of the other imaging parameters.

We conclude that all the image representations considered in this paper are insufficient by themselves to overcome the variation due to changes of light source location. The same result holds when variations between images of the same face due to other imaging parameters such as viewing position and expression are considered. Image representations that are sensitive to horizontal features can reduce the distance between images due to illumination relative to the distance between images due to identity. Further non-linear processing, the log of these representation, can decrease even further the sensitivity of these representations to changes of light source location. However, these representations are still insufficient for recognizing even one third of the faces in the database.

We would like to remark that the database used in this paper was suitable for showing the negative results that were presented. However, in order to evaluate a face recognition system, it is important to evaluate it on additional imaging parameters such as hair style, location of the face in the image, etc. Moreover, it is essential to consider a combination of several imaging parameters such as images that vary due to both illumination direction and viewing position. We have used simple comparison methods to evaluate the representations, since more complicated methods may include additional processing that compensate for the image variations due to a given imaging parameter. If such processing exits, it should be presented as such and evaluated separately.

An efficient face recognition system should deal explicitly with the variation due to changes of illumination direction. It is beyond the scope of this paper to suggest such a process, but we briefly consider here several basic possible approaches.

A straightforward multiple image approach: where a face model consists of a large set of images of the same face. The recognition process consists of computing and comparing the distances between a given image and every image of the model set. The crucial question, then, is how many images should a model contain. When other imaging parameters, such as the viewing position and expression, are considered as well, the number of images for a given model should be the product of the sample size for each of the spaces. Such schemes also have limited generalization capacity beyond the parameter values that had been sampled and stored. Furthermore, a psychophysical study of face identification (Moses et al., 1993) shows that the human visual system can identify faces in novel images after learning a single view of a face, which does not agree with such an approach.

Model-based approaches: usually address the problem of compensating for the viewing position parameter. In this case, several images of the same face can be considered to extract (either directly or indirectly) information about the three dimensional shape of the face. In principle it may be possible to extend some of these systems to cooperate with variation due to illumination direction as well. An example of the extension of the Linear Combination approach (Ullman and Basri, 1991) to deal with variations due to illumination can be found in

Moses (1933). However, such approaches are again inconsistent with the ability of the human visual system to generalize to novel views from a single image.

Class-based approaches: in which the knowledge about the general shape of a face can be used in order to compensate for the differences between images due to a given imaging parameter. An example of such processing is extracting special facial features in a manner that is independent of the illumination direction. This can be performed by choosing, for example, the stable edges from an edge map of a face. Another class-based approach is to define a representation that is invariant to illumination and viewing position. For example, representations invariant to illumination of bilaterally symmetric objects were shown to exist theoretically (Moses, 1993). Faces in particular belongs to the class of bilaterally symmetric objects, and therefore such a representation may be used. It may also be possible to incorporate knowledge about facial shape to the shape from shading approach. In this case, the 3D shape of the object may be extracted independently of the viewing position and illumination direction.

Finally, it is interesting to compare the sensitivity in face identification in novel images of the methods considered here to those of the human visual system. A psychophysical study of face identification in novel upright and inverted images taken with new illumination and viewing locations was performed (Moses et al., 1993). The results were that the identification of faces in upright novel images were above 97% correct. For inverted images, the performance of humans subjects reduced to 89% and 85% for new viewing position and illumination directions, respectively.

None of the representations considered in this paper is sufficient to achieve a performance comparable with that of the human visual system on upright faces. We therefore compare our results to those of the human visual system on inverted images. The human visual system can better generalize to changes in the light source location for inverted faces than to variations in viewing position. Our results regarding the original images were that the image variations due to illumination change are larger than image variation due to changes of viewing position. However, when image representations that are sensitive to horizontal features were used, the results agreed better with the sensitivity of the human visual system to image variations. That is, the image variations due to illumination change were smaller than image variation due to changes of viewing position. This may indicate that although the representations considered are insufficient by themselves to overcome the variation between images of the same face, they may still take part in the early stages of processing by the human visual system. However, to achieve the performance of face identification exhibited by the human visual system in upright images, additional processing must take place.

The variations between images of the same face due to changes of expression were not considered in the psychophysical experiment described above. However, intuitively we can say that the variations due to expression are often harder to compensate for than those due to changes of illumination and viewing position. The results of our experiments were that some representations make the variations due to expression harder to compensate for than for the original grey-level images. This result may suggest again that some of these image processing may indeed take place in the human visual system, and as a result overcoming the variations due to expression becomes more difficult than we would expect.

To conclude, overcoming image variations due to illumination direction is a basic problem in face recognition. Existing approaches to this problem use mainly universal representations i.e. representations that are not specific to faces. In this paper we have shown that many of these representations are insufficient to overcome the variations due to illumination. We suggest, therefore, that further effort should be made to overcome image variations due to changes in illumination conditions. Using knowledge regarding the specific domain of faces may help to find a representation that is insensitive to light source location.

Acknowledgements: We are grateful to Eli Okon and Oded Smikt for technical assistance in setting up the face image acquisition system. We also thank the people who were patient enough to have long series of their snapshots taken. We thank Ronen Basri, Nicola Ferrier, and Larry Shapiro for their comments on a draft of this paper. S.U. was supported in part by NSF Grant IRI 8900267.

References

Adini, Y., Moses, Y., , and Ullman, S. (1993). Face recognition: the problem of compensating for illumination changes. Technical report cc93-21, The Weizmann Institute of Science.

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

Brunelli, R. and Poggio, T. (1991). HyperBF networks for real object recognition. In *Proceedings IJCAI*, pages 1278–1284, Sydney, Australia.

Buhmann, J., Lades, M., and Eeckman, F. (1993). A silicon retina for face recognition. Technical Report No. 8596-CS, Institut für Informatik, Universität Bonn.

Burns, J., Weiss, R., and Riseman, E. (1992). The non-existence of general-case view-invariants. In Mundy, J. and Zisserman, A., editors, *Geometrical Invariance in Computer Vision*. M.I.T. press.

Canny, J. F. (1986). A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8:679–698.

Clemens, D. and Jacobs, D. (1991). Model-group indexing for recognition. In *Proc. CVPR-91, and IEEE Trand PAMI vol 13, No 10 p.1007-1017.*

Daugman, J. G. (1984). Spatial visual channels in the fourier plane. *Vision Res.*, 24(9):891–910.

Daugman, J. G. (1987). Image analysis and compact coding by oriented 2d gabor primitives. *SPIE processing*, 758:19–30.

Davies, G. M., Ellis, H., and Shephered, J. W. (1978). Face recognition accuracy as a function of mode of representation. *Journal of Applied Psychology*, 92:507–523.

Davis, L. S. (1975). A survey of edge detection techniques. *Computer Graphics and Image Processing*, 4:248–270.

De-Valios, R. and De-Valios, K. (1988). *Spatial Vision*. Oxford Univ. Press, New York.

Deriche, R. (1987). Optimal edge detection using recursive filtering. In *proc. ICCV, London*, pages 501–504.

Edelman, S., Reisfeld, D., and Yeshurun, Y. (1992). A system for face recognition that learns from examples. In Sandini, S., editor, *Proc. European Conference on Computer Vision*, pages 787–791. Springer-Verlag.

Govindaraju, V., Sher, D. B., Srihari, R., and Srihari, S. N. (1989). Locating human faces in newspaper photographs. In *Proc. CVPR 89*, pages 549–554.

Haralick, R. M. (1984). Digital step edges from zero crossings of second directional derivatives. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6:58–68.

Horn, B. K. P. and Brooks, M. (1989). *Seeing shape from shading.* MIT Press, Cambridge, Mass.

Kanade, T. (1977). *Computer recognition of human faces.* Birkhauser Verlag. Basel ans Stuttgart.

Manjunath, B., Shekhar, C., Chellappa, R., and von der Malsburg, C. (1992). A robust method for detecting image features with application to face recognition and motion correspondence. In *ICPR B*, pages 208–212.

Marr, D. and Hildreth, E. (1980). Theory of edge detection. *Proc. R. Soc. Lond. B*, 207:187–217.

Mercelja, S. (1980). Mathematical description of the responses of simple cortical cells. *J. Opical Soc.*, 70:1297–1300.

Moses, Y. (1993). *Face recognition: generalization to novel images.* PhD thesis, Weizmann Institute of Science.

Moses, Y., Edelman, S., and Ullman, S. (1993). Generalization to novel images in upright and inverted faces. Technical report cc93-14, The Weizmann Institute of Science.

Moses, Y. and Ullman, S. (1992). Limitation of non-model-based recognition schemes. In Sandini, G., editor, *Proc. ECCV-92*, pages 820–828. Springer-Verlag.

Patterson, K. and Baddeley, A. (1977). When face recognition fails. *Journal of Experimental Psychology: Human Learning and Memory*, 3:406–417.

Pollen, D. and Ronner, S. (1983). Visual cortical neurons as localized spatial frequency filters. *IEEE Transactions on System, Man and Cybernetics, SMC-13*, pages 907–916.

Reisfeld, D. and Yeshurun, Y. (1992). Robust detection of facial features by generalized symmetry. In *ICPR A*, pages 117–120.

Shen, J. and Castan, S. (1987). An optimal linear operator for edge detection. In *Proc. SPIE-87.*

Shepherd, J., Davies, G., and Ellis, H. (1981). Studies of cue saliency. In Davies, G. M., Ellis, H., and Shepherd, J., editors, *Perceiving and Recognizing Faces*, pages 104–131. Academic Press.

Torre, V. and Poggio, T. (1986). On edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8:147–163.

Ullman, S. and Basri, R. (1991). Recognition by linear combinations of models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13:992–1005.

Watson, A. and Robson, J. (1981). Discrimination at threshold: Labeled detectors in human vision. *Vision Research*, 19:19–32.

Wilson, H. R. and Bergen, J. R. (1979). A four mechanism model for threshold spatial vision. *Vision Research*, 19:19–32.

Wong, K., Law, H., and Tsang, P. (1989). A system for recognising human faces. In *Proc. ICASSP*, pages 1638–1642.