ELSEVIER

June 1994

Pattern Recognition Letters 15 (1994) 551-558

Pattern Recognition Letters

Analysing front view face profiles for face recognition via the Walsh transform

Xiaoguang Jia *, Mark S. Nixon **

Department of Electronics and Computer Science, University of Southampton, Southampton SO9 5NH, United Kingdom

Received 19 March 1992

Abstract

This paper describes the extraction of the facial profile from a frontal view to provide a measure for automatic face recognition. The profile is derived from the intensity projection of the face image and is described using the Walsh power spectrum. Results have demonstrated its potential as a measure for face recognition whilst it remains relatively immune to head movement and to small changes in illumination.

Key words: Feature extraction; Face recognition; Walsh transform

1. Introduction

The development of automatic face recognition has evolved via two main themes, via systems using neural networks and systems based on detailed face features (Jia and Nixon, 1992). Neural networks can appear to provide satisfactory recognition through training and sometimes through developing features (Phillips and Smith, 1989) and several neural network face recognition systems have been established (Hancock, 1990; Aleksander et al., 1992). It appears that neural networks can identify a face from a small population such as required in security applications, for example in access control. However strategies based on detailed face feature analysis appear more suited to identifying a face from a large population, such as might be required by the police. Systems based on detailed face feature measurements have, so far, employed two views of the face (Nixon et al., 1992), the front view (Jia and Nixon, 1992; Wong et al., 1989; Buhr, 1986), and the side view (Harmon et al., 1981; Wu and Huang, 1990) but only as separate strategies. Systems using a side view of a face have excluded features available from a frontal view, which restricts their feature lists. Systems based on a frontal view have included more features, generally by including more geometrical measurements.

Basing a recognition procedure on a frontal view alone rather than on a side view appears a wise choice, since a frontal view has major advantages that a much richer feature set can be available and that most documents record face images only as frontal views, such as passport photographs. In order to satisfy recognition requirements within large populations a rich feature set will be required. This cannot be achieved by just taking more geometrical measurements. It is manifest that face recognition can be based on a side view alone, but the feature set does not benefit from the richer feature set available in a front view. In order, therefore, to include the advantages of recogni-

* Current address: School of Astronautics, Harbin Institute of Technology, Harbin 150006, People's Republic of China.

** Corresponding author.

SSDI 0167-8655(93)E0068-Y

tion based on the side view of a face within a strategy based on a front view alone, it is possible to determine a version of the profile by intensity projection of a frontal view of a face.

2. Extraction and description of the profile from a frontal view

2.1. Extraction of profiles from frontal views

Human vision can normally infer 3-D shape correctly from a 2-D grey-level image by interpreting intensity information in the image and with some prior knowledge of the object. However, automatic inference of the 3-D shape has proven to be remarkably complex. Intensity distribution on a single 2-D monochrome image does contain the 3-D shape information of the object and some methods have been developed for otaining depth information of the object surface (Horn, 1986). Photometric stereo recovers the orientation of surface patches from a number of images taken under different lighting conditions. The techniques of photogrammetry and stereo interpret 3-D surface from a stereo pair of images by recovering the transformation between the two camera coordinate systems. Shape-from-shading analyzes the radiometry of image formation by measuring parameters calculated from the image of the illuminated object to determine the depth of surface. Laser-based systems have been used to obtain range data which contains explicit shape information. In these techniques either a stereo imaging system is required or the exact relative position of illumination sources to the object, such as distance and angles, must be known or controllable, and intensive computation is necessary.

In face recognition stereo imaging or precise prior control of lighting is a demanding requirement in application. Also, we seek to analyse a face for purposes of recognition and do not require to be able to reconstruct a face from a feature description. We therefore aimed to describe the profile feature of a face by a simpler method and without any special lighting requirements.

Intensity reveals the 3-D shape of the object on a 2-D image and the profile feature of a face is potentially available from a frontal view photograph of a face within a vertical central region of the face image. Accordingly a vertical rectangular region from the top of the forehead to the bottom of the chin along the middle of the face is defined as the face profile feature region. This region is well marked in an image since the position of the eyes can be established with good resolution. The intensity variation in this region is employed to represent the profile feature of a face.

The intensity distribution of an image can be represented by an intensity projection. The intensity projection of an image $f(x, y)$ along the direction w on the line z is defined as $p_w(z)$ (Jia and Qian, 1985)

$$p_w(z) = \int_z f(x, y)\, \mathrm{d}w\,. \tag{1}$$

The profile has been detected by the projection taken along the direction perpendicular to the vertical axis of the profile feature region. Example faces and their profile projections are shown in Fig. 1. This projection reflects the relation between the peaks and valleys of intensity with the face profile feature and dis-

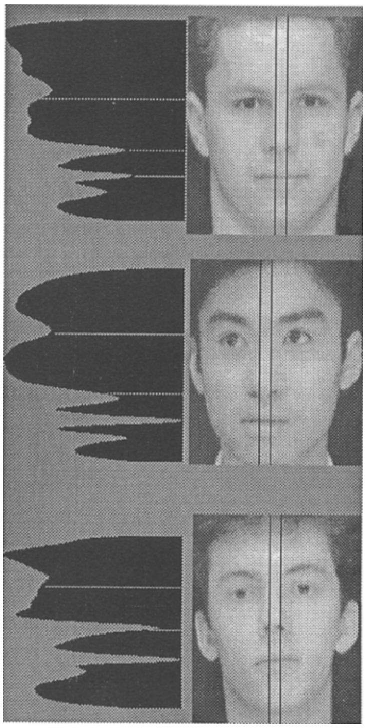

Fig. 1. Example faces and their profile projections.

tances between fiducial points, although it is not exactly the same as the profile of a face.

This projection then needs to be resampled into the length of N points, where $N=2^m$ and m is an integer, for further fast transformation. The image resolution is 256×256 points each with 256 grey-levels and since the profile feature is contained within this, $m=7$ and hence $N=128$ are a suitable choice.

2.2. Description of the profile projections

In order to achieve an efficient description of the profile projection several series of feature data were derived from the same resampled projection for further comparison. The Fourier transform and autocorrelation are the methods most commonly used to analyse characteristics of data series. The Walsh transform, and related dyadic autocorrelation, also have a wide application in signal processing and transform spectroscopy (Beauchamp, 1975). The Walsh transform has found extensive use in communications since it is based on sequency components which is an appropriate mode for decomposition and study of digital signals. It was included in this study to determine whether such as basis had any advantages in this application. Seven different description have therefore been explored and these are

(i) the resampled projection itself, expressed as $p(i)$;

(ii) the autocorrelation function of $p(i)$, $ACP(i)$;

(iii) the dyadic autocorrelation function of $p(i)$, $DACP(i)$;

(iv) the Fourier transform of $p(i)$, $FTP(i)$;

(v) the Walsh transform of $p(i)$, $WTP(i)$;

(vi) the Fourier power spectrum of $p(i)$, $FPSP(i)$; and

(vii) the Walsh power spectrum of $p(i)$, $WPSP(i)$.

The autocorrelation function of $p(i)$, $ACP(i)$, is given by

$$ACP(i)=\frac{1}{N}\sum_{k=0}^{N-1}p(k)p[(i+k)\bmod N]\,. \tag{2}$$

The Fourier power spectrum of $p(i)$, $FPSP(i)$, is the Fourier transform of $ACP(i)$ where

$$FPSP(i)=F[ACP(i)]=F[p(i)]F[p(i)]^* \tag{3}$$

where $F[\ ]$ is the Fourier transform.

A major difference between Walsh and Fourier spectral analysis is that Walsh analysis is expressed in terms of sequency as opposed to frequency components. Because Walsh spectral decomposition is circular time shift, it is not possible to achieve the power spectrum via the autocorrelation function directly. An indirect method (Beauchamp, 1975), which is equivalent to the Wiener–Khintchine method, derives the Walsh power spectrum via the dyadic autocorrelation of $p(i)$, $DACP(i)$, where:

$$DACP(i)=\frac{1}{N}\sum_{k=0}^{N-1}p(k)p(i\oplus k) \tag{4}$$

where $\oplus$ is a modulo 2 addition. The Walsh transform of the dyadic autocorrelation is the Walsh power spectrum $WPSP(i)$, and satisfies

$$WPSP(i)=W[DACP(i)]=W[p(i)]W[p(i)] \tag{5}$$

where $W[\ ]$ is the Walsh transform. This is defined as the dyadic power spectra by Walsh transform which corresponds to the circular power spectra obtained using the Fourier transform.

To reduce the effect of truncation on the autocorrelation function and Fourier transform, $p(i)$ is put through a Hamming Window filter, $W_h(n)$, where

$$W_h(n)=\alpha+(1-\alpha)\cos\left[\frac{\pi(2n-N)}{N}\right],\quad 0\leqslant n\leqslant N-1 \tag{6}$$

before calculating the autocorrelation function and the Fourier transform, where $\alpha=0.54$.

2.3. Comparison of different descriptions of profile projection

A good feature description should be able to be used to distinguish an object from a selection of similar objects of the same class. The seven descriptions of profile projection are compared by study of the standard deviation, σ_d, of the differences between the profile features of faces for all descriptions. Three groups of data were used for the comparison. First, seven descriptions were derived from each of three profile projections. Hence three 7×128-dimensional feature matrices were formed, in which each row represents one description of the profile projection. The relative difference between every two matrix elements was then calculated for the seven rows individually. Given

$$d_{ij} = \frac{|x_{ij} - y_{ij}|}{\sqrt{x_{ij} y_{ij}}}, \quad i = 1, 2, ..., 7; \quad j = 1, 2, ..., 128 \quad (7)$$

where x_{ij} and y_{ij} are the elements of matrices X and Y (the jth feature element of the ith profile description of face image X and Y, respectively). Three difference matrices were formed indicating feature differences of every two profiles for each of the seven descriptions. Given the standard deviation

$$\sigma_d = \left[\frac{1}{127} \sum_{j=1}^{128} (d_{ij} - \bar{d}_i)^2 \right]^{1/2} \quad (8)$$

where

$$\bar{d}_i = \frac{1}{128} \sum_{j=1}^{128} d_{ij}$$

for every two faces seven values of standard deviation were calculated individually corresponding to seven profile descriptions. Three groups of standard deviations of differences were obtained and shown in Fig. 2 with the horizontal axis indicating seven different profile descriptions in the order of (i) to (vii) as in Section 2.2., and the vertical axis indicating the standard deviation σ_d. A larger value of the standard deviation of the difference indicates a better description of the profile feature to distinguish between different faces. In the three curves in Fig. 2 both the autocorrelation, (ii), and the Fourier power spectrum, (vi), show their poor discriminatory ability by an even poorer performance than that achieved with the original data. This similarity in performance is to be expected from the relation between the autocorrelation and the Fourier power spectrum. However, the performance of the Fourier transform, (iv), as opposed to the power spectrum, appears to be the second largest value in two of three curves. The description of (vii), the Walsh power spectrum *WPSP(i)*, which demonstrates the largest difference in three groups and gives a clear improvement to that achieved with the original data is, therefore, illustrating the best presentation of profile features from different faces and was hence chosen as the most appropriate mode for analysing the profile feature.

2.4. Walsh power spectrum analysis

Though Fourier spectral analysis is a well-known and powerful method for signal analysis, the Walsh power spectrum has two advantages in this application which enable the Walsh power spectrum to give a better description of the profile feature. The first advantage is that it is possible for the power spectrum to be sequence-limited although the corre-

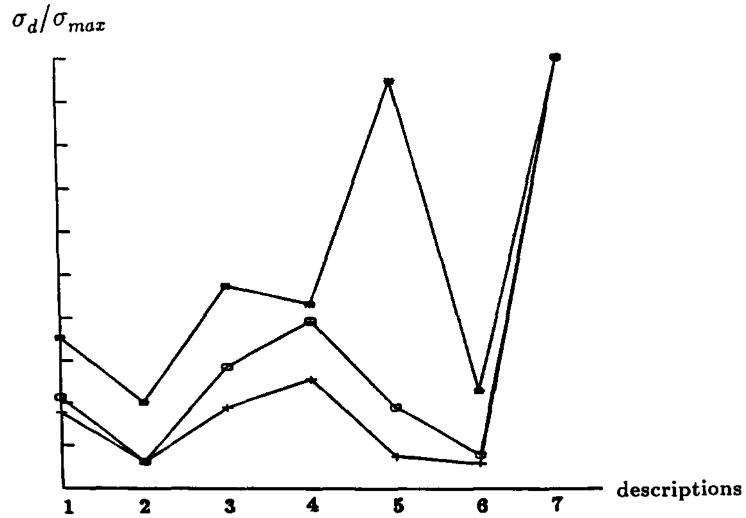

Fig. 2. Feature differences on three objects by seven descriptions: (i) *p*; (ii) *ACP*; (iii) *DACP*; (iv) *FTP*; (v) *WTP*; (vi) *FPSP*; (vii) *WPSP*.

sponding time function is time-limited (Beauchamp, 1975). This is in contrast to the behaviour of the Fourier transform in a power spectrum definition where a time-limited function cannot have a frequency-limited power spectrum. This may be the reason why a better description is achieved with the Walsh power spectrum than with the Fourier power spectrum even though a Hamming window was used before the Fourier transform of the projection data to reduce the variance in the power spectral density estimates. The same series length was kept for both transforms.

The second advantage is that the Walsh power spectrum is more phase sensitive than the Fourier power spectrum. In the Fourier power spectrum all coefficient items are defined by

$$R(n)=X(n)X(n)^{*}=|X(n)|^{2} \tag{9}$$

where $X(n)$ are the Fourier coefficients of the data series and $R(n)$ their Fourier power spectrum, in which phase shift of any coefficient does not emerge. Though the basic shape of a profile projection rarely varies from face to face, given that profiles are all brighter on the forehead and darker under the nose tip and between the lips, the undulation of peaks and valleys and the variation in their relative position, which may appear as phase shift in the Fourier transform while causing a global change in the Walsh transform, reflects the individual features.

3. Results

3.1. The Walsh power spectrum of profiles of different faces

The Walsh power spectrum of a profile extracted from a frontal view has been applied to a number of face images for assessment. By assessing the relative difference, d, between every two profile features where

$$d=\sum_{i=1}^{128}\frac{|x_i-y_i|}{\sqrt{x_i y_i}} \tag{10}$$

and where x_i and y_i are the feature elements of face image X and Y, these features have shown to be sufficiently reliable to discriminate between different persons' faces and to match different pictures of the same person.

Results of the relative difference evaluated for profile features are illustrated in Fig. 3. In this figure the horizontal axis is the face number and the vertical axis the inverse of the difference (i.e., the match). Exemplar match data has been evaluated for matching face No. 7, the top image in Fig. 1, with the 39 other faces. Face 8 is a different image of the same subject of face No. 7 which is reflected by a clear match between the two faces. A slightly poorer match is found between face No. 31 and face No. 7 (face No. 31 is the middle image in Fig. 1). The projections of image No. 7 and No. 31 have almost the same number of peaks and valleys (which are of similar width too), which corresponds to basic sequency components, but different shapes of the peaks, which should be related to phases and higher-order sequencies. The face which appears to differ most from face No. 7 by Walsh transform analysis is face No. 28, which is the bottom image in Fig. 1. The obvious difference between the projections of image No. 28 and No. 7 confirms the poor match by their Walsh power spectra.

3.2. The Walsh power spectrum of profiles of different images of same person

The algorithm has also been assessed for tolerance to face rotation. A group of frontal view images of the same person were taken with the face rotated from side to side and up and down, as shown in Fig. 4 which has these 8 face views together with the original view of the face (on the top row). The match for the profile extracted from the face images in Fig. 4, together with other faces, is shown in Fig. 5 in which Nos. 1–8 are profiles from Fig. 4 (these except the top one) which are the same person as No. 9 (top image in Fig. 4), but with face rotation. In Fig. 5 Nos. 1–4, in the second row of Fig. 4, show side-side movement; Nos. 5–8, in the bottom row of Fig. 4, show up-and-down movement; and the remainder of the faces Nos. 11–50 are of subjects different from No. 9. A good match to No. 9 appears for profiles Nos. 1–5; the match between No. 9 and No. 6 and No. 7 is still better than that between No. 9 and most other face profiles. However, face profile No. 8 fails to match No. 9 which was when the face was inclined significantly towards the camera. These results demonstrate that though face images are rarely taken with the camera

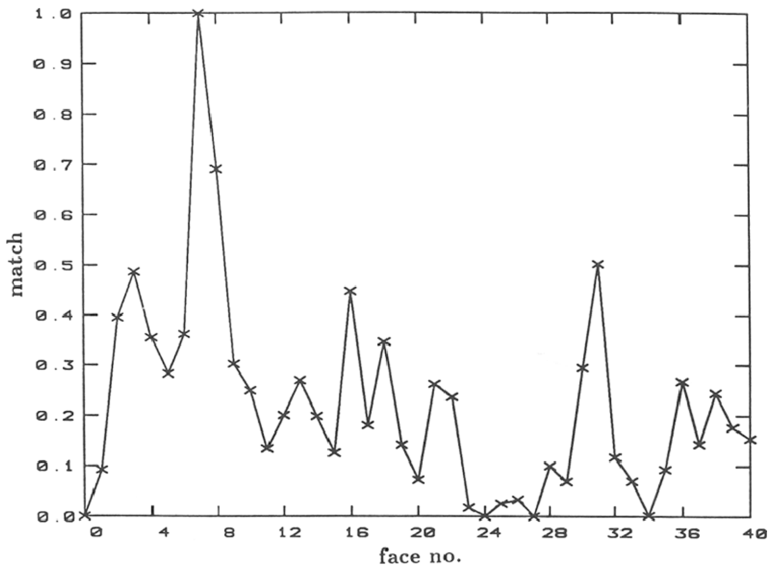

Fig. 3. Match between faces on the Walsh power spectrum of the profile.

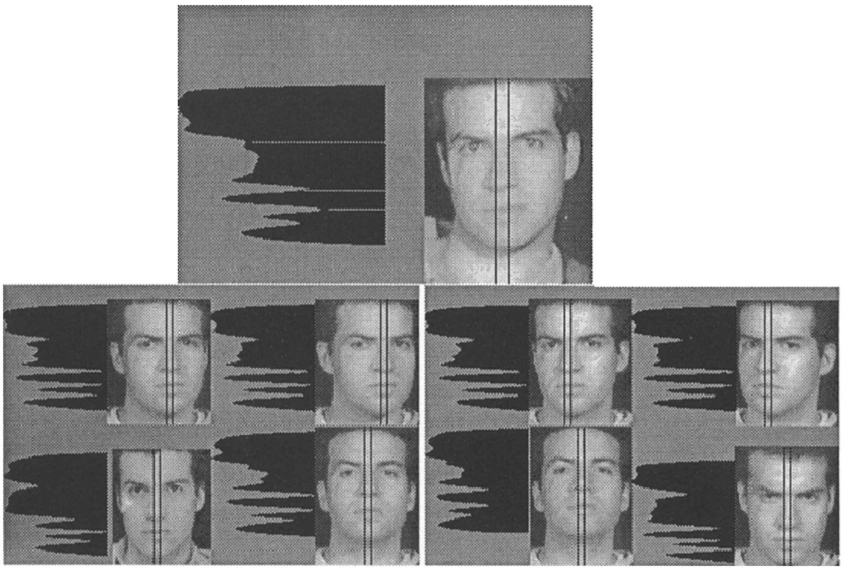

Fig. 4. Exemplar faces of a same person with different rotation and their profile projections.

exactly normal to a face, the profile feature described by the Walsh power spectrum of profile projection still reflects the profile feature of the face when compared with other faces and they also indicate particularly that the profile feature is more tolerant to side-to-side rotation of the face, than to up-and-down movement. This is because the profile measure, the projection, is taken vertically along the face, and in consequence inclining the face severly towards or from the camera causes a significant variation to the measure of the profile. This accords with the observed phase sensitivity of the Walsh transform described earlier and is consistent with expectations of its performance.

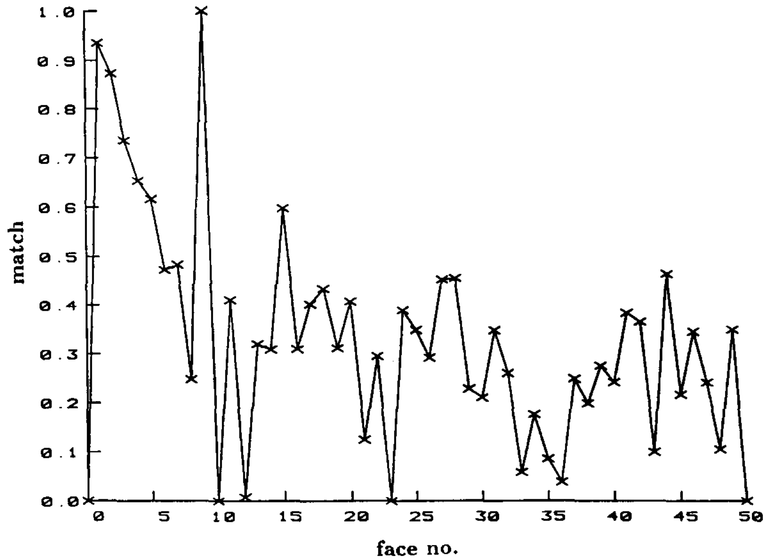

Fig. 5. Match between faces on the profile feature.

3.3. The Walsh power spectrum of profiles of the same face under different lighting

The effects of lighting on profile feature have also been assessed. When lighting changes slightly, the profile feature is naturally still contained within the projection. Movement of the faces are not only interpreted to represent the effect of variation in rotation but also to cause variation in illumination by changing the position of shaded areas on a face. The result of the previous section can therefore be assumed to reflect tolerance not just to side-to-side movement but also to illumination as well. However, if the lighting varies greatly this may have disastrous consequences since it will lead to a great difference in the grey-level face images, particularly shadows, thus affecting the projection of profiles. This is evident in the shadow caused by up-and-down face movement. It is therefore clear that extreme lighting can cause the technique to fail. These studies have been based on a laboratory environment but without precisely controlled illumination. The relative positions of light sources to a face have not been fixed – indeed this is difficult to accomplish given the variation in size of the subjects. On a day-to-day basis the illumination therefore changes from image to image. This demonstrates that it is tolerant to ambient illumination though it is clear that extreme illumination conditions would present serious problems – as they would in any vision exercise on feature extraction and description.

4. Conclusions

This paper has presented a novel method for detecting the profile feature of a face from a frontal view for purposes of automatic face recognition. This profile feature is extracted by an intensity projection from the central vertical rectangular region of a face and described by the Walsh power spectrum which was chosen as the feature descriptor from six other descriptors including the Fourier transform according to its ability to distinguish the differences of the profiles of different faces.

The technique has been applied to face images of different subjects and to different images of the same person but where the face was rotated or where the lighting varied slightly. The results of these studies have shown that this profile feature represented by the Walsh power spectrum appears to be reliable to reflect the identity of a person and the difference between faces. The assessment of the profile via the Walsh power spectrum was carried out not only on frontal views where the face was normal to the camera but also on images of faces which were rotated slightly from side-to-side and up-and-down. The results showed a good tolerance to these effects. The effects of lighting have also been considered and only extreme conditions will cause the technique to fail.

This strategy of extracting a profile feature from a frontal view for automatic face recognition is therefore feasible, and the methods of obtaining this measure by intensity projection and by describing it with Walsh power spectrum are sufficiently reliable to warrant further investigation. Possible further studies include texture discrimination to reduce the effects of hair on the forehead and analysis of shadows in the face to improve discrimination abilities.

References

Aleksander, I., et al. (1992). *Proceedings of IEE Colloquium on Machine Storage and Recognition of Faces.* London, UK, Jan. 1992.

Beauchamp, K.G. (1975). *Walsh Functions and Their Applications.* Academic Press, London.

Buhr, R. (1986). Analyze and Klassifikation von Gesichtsbildern. *ntzArchiv* 8 (10), 245–256.

Hancock, P.J.B. (1990). Gannet: Design of a neural net for face recognition by genetic algorithm. *IEEE Genetic Algorithms, Neural Networks and Simulated Annealing,* Glasgow, UK. IEEE UK and RI section Joint Chapter, 1–10, May 1990.

Harmon, L.D., M.K. Khan, R. Lasch and P.F. Raming (1981). Machine identification of human faces. *Pattern Recognition* 13 (2), 97–110.

Horn, B.K.P. (1986). *Robot Vision.* MIT Press, Cambridge, MA; McGraw-Hill, New York.

Jia, X. and M.S. Nixon (1992). Extending the feature set for automatic face recognition. *IEE 4th Internat. Conf. on Image Processing and Its Applications,* Maastricht, The Netherlands, Apr. 1992.

Jia, X. and G.H. Qian (1985). Object shape feature extraction and recognition. *J. Harbin Institute of Technology,* Special Issue. HIT Press, May 1985.

Nixon, M.S., J.N. Carter and X. Jia (1992). On biometric systems. *IEE Colloquium on Systems, Aspects of Machine Perception and Vision,* London, Jan. 1992.

Phillips, W.A. and L.S. Smith (1989). Conventional and connectionist approaches for face processing by computer. In: A.W. Young and H.D. Ellis, Eds., *Handbook of Research on Face Processing.* North-Holland, Amsterdam, 513–518.

Wong, K.H., H.M. Law and P.W.M. Tsang (1989). A system for recognizing human faces. *IEEE ICASSP,* 1638–1642.

Wu, C.J. and J.S. Huang (1990). Human face profile recognition by computer. *Pattern Recognition* 23 (3/4), 255–259.