

Available online at www.sciencedirect.com



Pattern Recognition Letters

Pattern Recognition Letters 28 (2007) 335-341

www.elsevier.com/locate/patrec

Face recognition using multiple facial features

A.N. Rajagopalan *, K. Srinivasa Rao, Y. Anoop Kumar

Image Processing and Computer Vision Laboratory, Department of Electrical Engineering, Indian Institute of Technology Madras, Chennai 600 036, India

Available online 19 May 2006

Abstract

We propose a face recognition method that fuses information acquired from global and local features of the face for improving performance. Principle components analysis followed by Fisher analysis is used for dimensionality reduction and construction of individual feature spaces. Recognition is done by probabilistically fusing the confidence weights derived from each feature space. The performance of the method is validated on FERET and AR databases.

© 2006 Elsevier B.V. All rights reserved.

Keywords: Face recognition; Principal components analysis; Fisher's linear discriminant; Block histogram modification; Edginess image; Distance in feature space; Fusion

1. Introduction

Enabling computers to recognize faces is guite a challenging problem. With society becoming more and more electronically connected, the capability to automatically establish identity of individuals using face as a biometric is becoming increasingly important and essential. In the literature, various algorithms have been proposed for the face recognition problem. Among the works based on facial features is that of Wu and Huang (1990) who used fiducial marks extracted from face profiles for recognition. Brunelli and Poggio (1993) advocate the idea of template matching for face recognition. A popular approach for face recognition is based on a compact representation of faces using the Karhunen-Loeve transform. This is also called principal components analysis (PCA) or the eigenface method (Turk and Pentland, 1991). Researchers have also proposed a combination of PCA and Fisher's linear discriminant (FLD) for face recognition (Belhumer et al., 1997). While PCA is optimal for representation of faces, the FLD uses class-specific information for higher discriminability. Yet another popular approach is elastic graph bunch matching (Wiskott et al., 1997) which uses a complex graph matching algorithm for recognition. Recently, multi-modal identification techniques have been receiving a lot of attention. The idea is to integrate different biometric cues such as face, fingerprint, speech, and gait (Hong and Jain, 1998; Kittler et al., 1998; Shakhnarovich and Darrell, 2002; Yacoub et al., 1999) for robustness.

We propose a novel method for face recognition that combines information from multiple facial features for improving accuracy and robustness. The work described here is an extension of the work reported in Srinivasa et al. (2004). The facial features that we consider are the grayscale image of the face, the edginess image of the face, and the eyes. The edginess image is robust to variations in illumination while the eyes are robust to facial expressions and occlusions. The idea is to use complimentary information for improving overall recognition performance. We assume that the eye locations are given so that the facial features can be extracted. Before feature extraction, a block histogram modification technique is applied to compensate for local changes in illumination. PCA in conjunction with FLD is then used to encode the facial features in a lowerdimensional space. The distance in feature space (DIFS) values are calculated for all the training images in each of the feature spaces and these values are used to compute the distributions of the DIFS values. The distributions of the DIFS values are very useful in characterizing the differences

^{*} Corresponding author. Tel.: +91 44 2257 4433; fax: +91 44 2257 8360. *E-mail address:* raju@ee.iitm.ac.in (A.N. Rajagopalan).

between imposters and true persons. In the recognition phase, given a test image, the three facial features are extracted and their DIFS values are computed in each feature space. Each feature provides an opinion on the claim in terms of a confidence value which is measured by integrating the DIFS distributions of each feature space with respect to the DIFS value computed in that feature space. The confidence values of all the three features are then probabilistically fused for final recognition. The proposed fusion method works quite well and yields a significant improvement in recognition over that achievable with any single feature.

In Section 2, we discuss the eigenface technique. Section 3 describes the global and local facial features used in this work. Section 4 deals with the construction of individual feature spaces and the issue of illumination compensation. A novel approach to fuse information from multiple facial features is discussed in Section 5. Experimental results are given in Section 6. The paper concludes with Section 7.

2. Principal components analysis

A face image of size $N \times N$ pixels can be viewed as a vector of dimension N^2 . However, such a high-dimensional representation is too detailed. Since facial features are similar in overall configuration across individuals, it is possible to describe faces quite compactly using PCA. Let a face image f(x, y) from the training set be a 2-D $N \times N$ array of intensity values. Let $f_{i,m}$ be the N^2 dimensional vector representing the *m*th training image of the *i*th person. If in the dataset there are I number of people, each having M number of images, then we have a total of $K = I \cdot M$ training images. The average image for the entire dataset is given by $\psi = \frac{1}{K} \sum_{i=1}^{I} \sum_{m=1}^{M} f_{i,m}$. An estimate of the covariance matrix C of the face dataset can be obtained as

$$C = \frac{1}{K} \sum_{i=1}^{I} \sum_{m=1}^{M} \phi_{i,m} \phi_{i,m}^{\mathrm{T}}$$
(1)

where $\phi_{i,m}$ is the mean subtracted image of $f_{i,m}$ and is given by $\phi_{i,m} = f_{i,m} - \psi$. The weight vector in the PCA (eigenface) space corresponding to the *m*th training image $\phi_{i,m}$ of the *i*th person can be derived as $w_{i,m} = E_{pca}^{T} \phi_{i,m}$. Here, $E_{pca} = [e_1, e_2, \dots, e_{K'}]$ consists of only the first K' significant eigenvectors of *C*. These are also called eigenfaces (see Fig. 1). Since $K' \ll N^2$, the weight vector which represents the face image in the PCA space is of a low-dimension. The average weight vector w_i in the PCA space for the *i*th person (or class) is given by



Fig. 1. Some typical eigenfaces.

$$w_{i} = \frac{1}{M} \sum_{m=1}^{M} w_{i,m}$$
(2)

When a new test image pattern T is presented to the system, its weight vector is computed as $w = E_{pea}^{T}(T - \psi)$. The pattern is declared to belong to that face class in the training set for which $||w - w_i||$ is the smallest over all j.

PCA yields projection directions that maximize the total scatter across all faces. It retains unwanted variations due to lighting and expressions which limits its performance. For FERET database, the maximum accuracy of the base-line PCA (Phillips et al., 2000) is only 80% and with variations in illumination it drops to 22%.

3. Global and local features

Face recognition approaches that consider only the entire face as a feature do not take into account just what other aspects of the face stimuli are important for recognition (Penev and Atick, 1996). We propose to use global as well as local features. The motivation for incorporating local features into a recognition system stems from the fact that it is possible for humans to recognize a face from only parts of it. In addition to the entire face image, we consider two other features; namely, the edginess image of the face, and the eyes. The edginess image is a global facial feature that is reasonably robust to illumination. It is a measure of the change in intensity from one pixel to the next. The eves are quite robust to facial expressions and occlusions. Eyes are essentially unaffected by beards and mustaches, making them invaluable for the face recognition task. Since each feature contributes differently, the idea is to utilize the complementary nature of these contributions to get improved accuracy. Of course, one must have a novel way of fusing information.

To extract the edginess image, we employ 1-D processing (Venkatesh et al., 2002) along two orthogonal directions as follows. To detect the horizontal component of edginess, a discrete approximation of the 1-D Gaussian filter is first used to smooth the image horizontally to reduce the effect of noise. A discrete approximation of the first-order derivative of the 1-D Gaussian function is next used in the orthogonal direction (i.e., vertically) to find the horizontal component of edginess. The vertical component of edginess is computed in a similar manner by carrying out the above steps in the orthogonal direction. The final edginess image is a gray-valued image and is obtained by taking the absolute sum of the horizontal and the vertical components. The edginess at a pixel gives the magnitude of the gradient of the intensity at that pixel location in the image.

It must be mentioned here that in order to extract the facial features, one requires automatic and accurate face detection in cluttered backgrounds. This is a challenging research problem in itself and various efforts have been proposed to tackle this (Rajagopalan et al., 2000). Following other works (Phillips et al., 2000), we assume that the face has been cropped out of a scene. Usually, a face is

cropped with respect to the eye locations since the separation between the eyes is independent of facial expressions, up and down movements of the face, etc.

4. Construction of feature spaces

The three facial features described in the previous section are quite high-dimensional and cannot be used directly. For dimensionality reduction, we apply the PCA technique to each feature. To provide class-specific discrimination for achieving better recognition accuracy, we employ Fisher's linear discriminant (FLD). A face recognition system must recognize a face from its novel image despite variations in illumination. For this purpose, we propose a block histogram method which modifies the histogram of a test image in accordance with that of a uniformly illuminated reference image.

4.1. Linear discriminant analysis

The Fisher's linear discriminant (FLD) is a classical technique in pattern recognition first developed by Fisher (1936) for taxonomic classification. FLD is an example of a class-specific method in the sense that it tries to shape the scatter in order to make it more reliable for classification. It involves eigenanalysis of a product of two matrices, one of which is inverted. To ensure that the matrix to be inverted is not rank-deficient, we first perform dimensionality reduction using the PCA technique discussed in Section 2. The PCA spaces are derived independently for each feature. The weight vectors in the PCA space are given as input to FLD. The FLD analysis that follows is equally applicable to all the three facial features.

For an *I* class problem, the between-class scatter matrix for a given feature is defined as $S_b = \sum_{i=1}^{I} (w_i - \bar{w})(w_i - \bar{w})^T$ where w_i is the average weight vector in the PCA space for the *i*th class. The quantity \bar{w} is the average weight vector of all the classes in the PCA space of that feature and is given by $\bar{w} = \frac{1}{I} \sum_{i=1}^{I} w_i$. The within-class scatter matrix is defined as $S_w = \sum_{i=1}^{I} (\frac{1}{M} \sum_{m=1}^{M} (w_{i,m} - w_i)(w_{i,m} - w_i)^T)$ where $w_{i,m}$ is as defined in Section 2. Mathematically, FLD selects the projection matrix E_{fld} so as to maximize the ratio of the determinant of the between-class scatter matrix to the within-class scatter matrix of the projected samples, i.e., the projection matrix E_{fld} is chosen such that

$$E_{\rm fld} = {}^{\rm arg} F^{\rm max} \frac{|F^{\rm T} S_{\rm b} F|}{|F^{\rm T} S_{\rm w} F|} = [f_1, f_2, \dots, f_k]$$
(3)

Eigenvector f_i can be determined by solving the generalized eigenvalue problem (Belhumer et al., 1997) $S_{\rm b}f_i = \lambda_i S_{\rm w}f_i$, i = 1, 2, ..., k. Since there are at most I - 1 nonzero generalized eigenvalues, an upper bound on k is I - 1.

The projection matrix E_{opt} which is a combination of the eigen and Fisher projections is given by $E_{opt}^{T} = E_{fld}^{T} E_{pca}^{T}$. Thus, the final feature vector corresponding to the *i*th person with *M* training images is computed as

$$w'_{i} = \frac{1}{M} \sum_{m=1}^{M} E_{\text{opt}}^{\text{T}} \phi_{i,m} = \frac{1}{M} \sum_{m=1}^{M} E_{\text{fld}}^{\text{T}} E_{\text{pca}}^{\text{T}} \phi_{i,m}$$
(4)

where $\phi_{i,m}$ is the mean subtracted training image. These steps are repeated independently for each of the three facial features to construct the three feature spaces.

4.2. Block histogram modification

Various approaches have been proposed in the literature to alleviate the effect of illumination variations for face recognition (Adnin et al., 1997). However, till to date, a revolutionary solution remains elusive. Among the existing methods, the illumination cones technique (Georghiades et al., 2001) outperforms most other methods. But it is computationally intensive and requires at least seven images per person. We propose here a simple but effective block histogram modification (BHM) technique for illumination compensation.

Assume that a reference image Y taken under well-controlled lighting conditions is available. The goal is to bring the local illumination levels of an input image X to those of the reference image Y. Both the images are assumed to be of the same size ($N \times N$ pixels). Consider a block image B_I from X with pixel locations ranging from 1 to M and also a block image $B_{\rm R}$ from Y at the corresponding pixel locations. Let $p_x(x)$ and $F_x(x)$ be the probability density function and the distribution function, respectively, of the intensity variable $x \ (\geq 0)$ belonging to the input block image B_I . Also, let $p_v(y)$ and $F_v(y)$ be the probability density function and the distribution function, respectively, of the intensity variable $y \ (\geq 0)$ belonging to the reference block image $B_{\mathbf{R}}$. Note that $F_x(x) = \int_0^x p_x(u) du$ while $F_{v}(y) = \int_{0}^{y} p_{v}(u) du$. The final output block image B_{O} with pixel intensity value $z \ge 0$ will have the desired density $p_{v}(y)$ and distribution $F_{v}(y)$ provided $z = F_{v}^{-1}[F_{x}(x)]$.

If histogram modification is done blockwise independently, it can lead to artifacts. To smooth out intensity changes across adjacent blocks, we weight and overlap the blocks. This prevents edges and patches from appearing in the illumination compensated image. The histogram modified block image intensity values are scaled with a window to yield

$$B_{\mathcal{O}}(n,m) = B_{\mathcal{O}}(n,m)U(n,m), \quad 1 \le n, m \le M$$
(5)

where the windowing filter $U(n,m) = \frac{4nm}{M^2}$ for $1 \le n, m \le \frac{M}{2}$; $\frac{4m(M-n+1)}{M^2}$ for $\frac{M}{2} < n \le M$, $1 \le m \le \frac{M}{2}$; $\frac{4n(M-m+1)}{M^2}$ for $1 \le n \le \frac{M}{2}$, $\frac{M}{2} < m \le M$; and $\frac{4(M-n+1)(M-m+1)}{M^2}$ for $\frac{M}{2} < n, m \le M$. Note that the window has been chosen such that the sum of the weights in the overlapping region is 1. By simultaneously shifting the blocks in both the horizontal and the vertical directions in steps of $\frac{M}{2} + 1$ pixel locations, and adding pixel intensity values in overlapping regions, we arrive at the final image Z.

In Fig. 2, we give examples of images taken under different illumination directions and the corresponding intensity



Fig. 2. Images before and after illumination compensation using BHM.

normalized images using our method. The reference image was the same. As can be seen, the method works quite well.

5. The fusion approach

Once the feature spaces are constructed, an important quantity of interest is the distance in feature space (DIFS). Consider (say) the face feature space. Given a face image γ , its weight vector in the PCA–FLD feature space can be derived as $w' = E_{opt}^{T}(\gamma - \psi)$ where all the terms are as described earlier. The DIFS value which measures how well γ matches with the *i*th class is given by $||w - w'_i||$. The concept of DIFS is applicable to all three feature spaces. We now propose a fusion strategy to integrate information coming from face, edginess image of the face, and eyes. The methodology involves computing the distributions of the DIFS values at different ranks in each feature space. The distributions are computed empirically from the training data as explained below.

Let us consider a single feature (say) face. As discussed in Section 4.1, the face training set is used to construct the PCA-FLD space for faces. Let ϵ^{α} (a random variable with density function $f^{\alpha}(\epsilon^{\alpha})$ represent the DIFS value of all individuals' training images with their own classes, i.e., $f^{\alpha}(\epsilon^{\alpha})$ describes the distribution of the DIFS value for the genuine case. It is also possible to project every image in the training set onto the PCA-FLD face space which would yield DIFS values for that image with respect to all the individuals in the database. Without loss of generality, let us assume that $\epsilon_1, \epsilon_2, \ldots, \epsilon_I$ are the DIFS values arranged in an increasing rank order and that these are statistically independent. Let ϵ_i (a random variable with density function $f_i(\epsilon_i)$ denote the DIFS value at rank *i*. Note that both ϵ^{α} and ϵ_i are based on the entire training set and are not individual-specific. The relative difference in the DIFS value at rank *i* is defined as

$$\Delta_i = \epsilon_i - \epsilon^{\alpha}, \quad 1 \leqslant i \leqslant I \tag{6}$$

where *I* is the total number of individuals/identities/classes in the database. The distribution $f_i(\Delta_i)$ at rank *i* is then given by

$$f_i(\Delta_i) = \int_{-\infty}^{\infty} z_i(\epsilon_i, \Delta_i) \mathrm{d}\epsilon_i$$

where

$$z_i(\epsilon_i, \Delta_i) = \frac{I!}{(I-i)!(i-1)!} F(\epsilon_i)^{i-1} [1 - F(\epsilon^{\alpha})]^{I-i} f(\epsilon_i) f(\epsilon^{\alpha})$$
(7)

Here, $z_i(\epsilon_i, \Delta_i)$ describes the joint distribution of the *i*th rank DIFS value ϵ_i and its distance from the true DIFS value ϵ^{α} .

The term $f(\epsilon_i)$ is the probability density function of the absolute DIFS value at rank *i* and $F(\epsilon_i)$ is the corresponding distribution function. More details on the distribution of distances can be found in Gumbel (1958). Since closed form expressions for $f(\epsilon_i)$, $f(\epsilon^{\alpha})$, $F(\epsilon_i)$, and $F(\epsilon^{\alpha})$ are not available for the problem on hand, we derive $f_i(\Delta_i)$ and $f(\epsilon^{\alpha})$ empirically by using the DIFS values for ϵ_i and ϵ^{α} from the images in the training set. Since there are *I* individuals each having *M* training images, we have $I \cdot M$ DIFS values or sample realizations for the random variable ϵ^{α} and so also for ϵ_i for each rank *i*. Note that the mean value of the DIFS distributions $f_i(\Delta_i)$ will increase with rank *i*.

When a new test image γ arrives, its DIFS values $\epsilon'_1, \epsilon'_2, \ldots, \epsilon'_i$ with respect to all individual classes in the PCA-FLD face space are arranged in an increasing order. Let $\epsilon^{\alpha'}$ denote the DIFS value for the top rank. The relative DIFS values Δ'_i can be computed using Eq. (6) as $\Delta'_i = \epsilon'_i - \epsilon^{\alpha'}$. For the given image, the confidence weight assigned for the hypothesis that the image belongs to that of the identity/class at rank *i* is computed as

$$P_{\text{face}}(i) = P_i(\Delta'_i) \cdot P_{\text{order}}(i), \quad 1 \le i \le I$$
(8)

In the above equation, $P_i(\Delta'_i)$ describes how close the *i*th rank DIFS value is to the top rank and is given by $P_i(\Delta'_i) = \int_{\Delta'_i}^{\infty} f_i(\Delta_i) d\Delta_i$. The term $P_{order}(i)$ assigns an appropriate weight depending on the rank position and the top rank DIFS value. If the top rank DIFS value for the given image γ is very small, then that image will most likely correspond to the actual identity. Hence, $P_{order}(i)$ should fall very sharply as rank *i* increases. On the other hand, if the top rank DIFS value is large, then the top person may not be the actual identity, and hence the confidence weight should fall gradually to accommodate even individuals at lower ranks (i.e., higher values of *i*). Thus, depending on the top rank DIFS value $\epsilon^{\alpha'}$, we give relative weightage to the person at rank *i* using the gamma distribution. The gamma distribution is given by

$$P_{\text{order}}(i) = \frac{\beta}{\gamma(\lambda)} \left(\frac{i-1}{\theta}\right)^{(\lambda\beta-1)} e^{(-\frac{i-1}{\theta})}, \quad \theta = \frac{1}{p}, \ 1 \leqslant i \leqslant I \quad (9)$$

For our problem, $p = \int_{\epsilon^{a'}}^{\infty} f^{\alpha}(\epsilon^{\alpha}) d\epsilon^{\alpha}$. We define $P_{\text{order}}(i)$ with $\lambda = 1$ and $\beta = 1$. If p = 1, the curve given by the above equation falls very sharply. However, when the top rank identity is not a genuine one, the combined effect of $P_i(\Delta'_i)$ and $P_{\text{order}}(i)$ is to allow to accommodate identities at lower ranks also.

In an exactly similar manner, we compute $P_{edge}(i)$ and $P_{eye}(i)$ for the other two features. For mathematical convenience, the three facial features (face, edginess, and eyes) are assumed to be independent. If I_1, I_2, \ldots, I_I are the identity indicators of the individuals in the database, then the final confidence weight of an identity I_i is obtained by multiplying the confidence weights contributed from each feature space of that identity, i.e.,

$$P(I_i) = P_{\text{face}}(I_i) \cdot P_{\text{edge}}(I_i) \cdot P_{\text{eye}}(I_i)$$
(10)

The identity $ID(\gamma)$ for a given image γ is determined by the following criterion:

$$ID(\gamma) = \begin{cases} I_k & \text{if } P(I_k) > \tau\\ Imposter & \text{otherwise} \end{cases}$$

where

$$P(I_k) = \max\{P(I_1), P(I_2), \dots, P(I_I)\}$$
(11)

Threshold τ is chosen such that an untrained person will not be recognized at all.

6. Experimental results

We demonstrate the performance of the proposed method on the standard FERET and AR face databases. The required facial features were cropped with reference to the eye locations provided along with the database. Fig. 3 shows the extracted facial features for an individual in the dataset. The eye locations were used to account for rotation and scaling, when necessary. Before constructing the feature spaces, all the images were intensity normalized using the BHM technique described in Section 4.2.

The FERET database contains 14,126 images comprising of 1199 individuals. Since the images are acquired during different photo sessions, this dataset contains significant



Fig. 3. (a) Face image, (b) edginess image, and (c) the cropped eyes.

variations in pose, illumination and facial expressions. The FA images (regular frontal faces of persons) were used as the gallery set (1196 images), whereas four categories of probe sets were used for comparison against the gallery set. The first probe category was the FB probe set (1195 images). This indicates an alternative frontal image, taken seconds after the corresponding FA images. The second probe category contained all duplicate frontal images and is referred to as the Duplicate I probe set (722 images). The third category of probe set is the FC set (194 images) which contains images taken on the same day but with different camera and illumination. The fourth category of probe set is called the Duplicate II set (234 images). These images are duplicates of FA images but taken at least one year between the acquisition of the gallery images (FA) and the probe images.

A commonly used performance measure for face recognition is the cumulative match score (CMS), i.e., the recognition accuracy in the top *n* ranks. The CMS plots for the top *n* ranks are shown in Fig. 4 for all the four probe categories and for different features viz., face, edginess, eyes, and the fusion case. For n = 1 the plots clearly reveal the advantage of using fusion. The fusion method performs consistently well for all the probe categories and yields an average improvement of about 5% as compared to the best accuracy achievable with any individual feature. In Table 1, we have also compared the performance of our method with the partially automatic face recognition algorithms (Phillips et al., 2000). The recognition accuracy for these algorithms was deciphered from the plots therein. From the table, we observe that the performance of the proposed method is comparable to the best reported results. For the FB, Duplicate I and Duplicate II probe sets, our method



Fig. 4. CMS plots for (a) FB, (b) Duplicate I, (c) FC, and (d) Duplicate II probe set.

 Table 1

 Recognition accuracy for different algorithms on the FERET database

Probe set	% Recognition accuracy at rank 1			
	Fusion	UMD 97	USC	MIT 96
FB	98.3	96.5	95	94.8
Duplicate I	68	46	58	57
FC	59	59	82	32
Duplicate II	54	21	46	34

has better accuracy compared to others. On the FC probe set, we come second.

The performance of our system was also tested on untrained people to check how well it rejects unknown persons. This was done using false acceptance rate (FAR) and false rejection rate (FRR) curves. The relation between the two rates is controlled by the acceptance threshold of the system. The value of FRR and FAR at the point where the plots cross is called the equal error rate (ERR). For good recognition, the ERR value should be as small as possible. The system was trained with 482 out of 1199 individuals. A total of 1446 images were used for training, three images per subject. For the FAR plot, 1440 images were used as probe from the remaining 717 untrained individuals. For the FRR plot, 304 images were used as probe from 482 trained individuals. The FAR and FRR plots for face, edginess, eyes, and the fusion method are shown in Fig. 5(a)-(d), respectively. For face, edginess, and eves, the acceptance threshold value is in terms of the top rank DIFS value, whereas for the fusion method it is in terms of the final confidence value. Note that the ERR values for face, edginess, and eye are quite high (31%, 30%, and

34%, respectively). In contrast, the ERR value for the fusion method is only 12%.

We also conducted experiments on the AR face database (Martinez and Benavente, 1998) which contains rich variations in expressions and many facially occluded images. During training, we selected 630 images comprising of 126 individuals with five images per subject. The recognition performance was tested using a probe set containing 1759 images of all the trained individuals. The images in the test set were different from the training set. The CMS plots are shown in Fig. 6 for different features viz., face, edginess and eyes individually, as well as for the fusion case. We observe that for n = 1 there is an improvement of almost 6% with fusion over other individual features.

The FAR and FRR plots for face, edginess, eyes, and fusion are shown in Fig. 7(a)–(d), respectively. While computing FAR plots, the system was trained with 90 persons out of the 126 members in the database. A total of 380 probe images of the remaining untrained people were used for test-



Fig. 6. CMS curves for the AR database.



Fig. 5. FAR and FRR plots for FERET database: (a) face only, (b) edginess only, (c) eyes only, and (d) fusion.



Fig. 7. FAR and FRR plots for the AR database using (a) face only, (b) edginess only, (c) eyes only, and (d) fusion.

ing. While computing the FRR values, the system was again trained with the complete dataset and a probe set of 1209 images comprising of all 126 persons was used. We observe that the ERR values for face, edginess, and eyes are high (25%, 29.5%, and 25%, respectively). In contrast, the ERR value for the fusion method is much lower and is only 10%.

7. Conclusions

We have proposed a new methodology that combines information gathered from multiple facial features, namely, the face, the edginess image of the face, and the eyes for robust and accurate face recognition. The distributions of the DIFS values in each feature space yield confidence weights for performing recognition. To compensate for changes in illumination, a block histogram modification method was advocated. When the fusion method was tested on FERET and AR databases, its recognition accuracy as well as its ability to reject imposters was found to be quite good.

References

- Adnin, Y., Moses, Y., Ulman, S., 1997. Face recognition: The problem of compensating for changes in illumination direction. IEEE Trans. Pattern Anal. Machine Intell. 19, 712–732.
- Belhumer, P.N., Hespanha, J.P., Kriegman, D.J., 1997. Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. IEEE Trans. Pattern Anal. Machine Intell. 19, 711–720.
- Brunelli, R., Poggio, T., 1993. Face recognition: Features versus templates. IEEE Trans. Pattern Anal. Machine Intell. 15, 1042–1052.
- Fisher, R., 1936. The use of multiple measures in taxonomic problems. Ann. Eugenic. 7, 179–188.
- Georghiades, A.S., Belhumeur, P.N., Kriegman, D.J., 2001. From few to many: Illumination cone models for face recognition under differing

pose and lighting. IEEE Trans. Pattern Anal. Machine Intell. 23, 643-660.

- Gumbel, E.J., 1958. Statistics of Extremes. Columbia University Press.
- Hong, L., Jain, A., 1998. Integrating face and fingerprints for personal identification. IEEE Trans. Pattern Anal. Machine Intell. 20, 1295–1307.
- Kittler, J., Hatef, M., Duin, R.P.W., Matas, J., 1998. On combining classifiers. IEEE Trans. Pattern Anal. Machine Intell. 20, 226–239.
- Martinez, A.M., Benavente, R., 1998. The AR face database. CVC Technical Report #24, June 1998.
- Penev, P.S., Atick, J.J., 1996. Local feature analysis: A general statistical theory for object representation. Network: Comput. Neural Systems 7, 477–500.
- Phillips, P.J., Moon, H., Rizvi, S.A., Rauss, P.J., 2000. The FERET evaluation methodology for face-recognition algorithms. IEEE Trans. Pattern Anal. Machine Intell. 22, 1090–1103.
- Rajagopalan, A.N., Kumar, K.S., Karlekar, J., Manivasakan, R., Patil, M.M., Desai, U.B., Poonacha, P.G., Chaudhuri, S., 2000. Locating human faces in a cluttered scene. Graph. Models Image Process. 5, 323–342.
- Shakhnarovich, G., Darrell, T., 2002. On probabilistic combination of face and gait cues for identification. In: Proc. Fifth IEEE Internat. Conf. on Automatic Face and Gesture Recognition.
- Srinivasa, R.K., Ghadi, L.S., Anoop, K.Y., Rajagopalan, A.N., 2004. Robust face recognition by fusion. In: Proc. Indian Conf. on Computer Vision, Graphics and Image Processing, pp. 669–674.
- Turk, M., Pentland, A., 1991. Eigenfaces for recognition. J. Cogn. Neurosci. 3, 76–96.
- Venkatesh, B.S., Palanivel, S., Yegnanarayana, B., 2002. Face detection and recognition in an image sequence using eigenedginess. In: Indian Conference on Computer Vision, Graphics and Image Processing, pp. 97–101.
- Wiskott, L., Fellous, J.M., Kruger, N., Malsburg, C., 1997. Face recognition by elastic bunch graph matching. IEEE Trans. Pattern Anal. Machine Intell. 19, 775–779.
- Wu, C.J., Huang, J.S., 1990. Human face profile recognition by computer. Pattern Recognition 23, 255–259.
- Yacoub, S.B., Abdeljaoued, Y., Mayoraz, E., 1999. Fusion of face and speech data for person identity verification. IEEE Trans. Neural Networks 10, 1065–1074.