Volume 2 No. 2, February 2012 ISSN 2223-4985
International Journal of Information and Communication Technology Research

http://www.esjournals.org

Improved Fisher Face Approach for Human Recognition System using Facial Biometrics

[1]Vijayshree More, [2]Abhay Wagh
[1]E&Tc Engineering department
JNEC, Aurangabad, India.
Deputy Secretary,
[2]Govt. of Maha. Higher & Tech Edu., Mumbai, India.

ABSTRACT

In this paper, a person is identified with face as a biometric feature using modified fisher face and fuzzy fisher face. The premise of this paper is to introduce modified fisher face, fuzzy fisher face and include gradual level of assignment to class being regarded as a membership grade which helps to improve recognition results. Performance of the said system is compared with traditional fisher face methods using Fisher's linear discriminant analysis. For all these methods of face recognition an assumption was made about the same level of relevance of each face to the corresponding category. The comprehensive experiments completed on ORL, and Yale face databases show improved recognition rates and reduced sensitivity to variations between face images caused by changes in illumination and viewing directions.

Keywords—*linear discriminant analysis, fisher face, fuzzy fisher face, face recognition*

I. INTRODUCTION

Humans can easily recognize faces, spoken words, handwritten or printed digits, images and many other things in everyday life. If there are a limited number of categories or classes then recognition performance may be improved; however, the same might not be very efficient if several categories are present. Face recognition is largely motivated by the need for surveillance and security, telecommunication and digital libraries, human-computer intelligent interaction, and smart environments [1 - 4]. In practice, face recognition is a very difficult problem due to a substantial variation in light direction, different face poses, and diversified facial expressions. A good face recognition methodology should consider representation as well as classification issues, and a good representation method should require minimum manual annotations. Face recognition depends heavily on the particular choice of features used by the classifier [5]. One usually starts with a given set of features and then attempts to derive an optimal subset of features leading to high recognition performance with the expectation that similar performance can also be displayed on future trials using unseen test data. The most well known classification techniques used for face recognition are those of eigenface [6] and fisher face [7]. The eigenface method relies on a transformation of feature vectors by utilizing principal components and is referred to as principal component analysis (PCA) used to derive a starting set of features; the other naming used there is the Karhunen–Loeve (KL) expansion. In essence, the PCA dwells on a linear projection of a high-dimensional face image space into a new low-dimensional feature space. The major problem coming with the use of the eigenface technique is that it can be affected by variations in illumination conditions and different facial expressions. It is also worth stressing that the PCA is oriented toward the representation in low-dimensional spaces but not necessarily optimal in terms of face classification. Applying PCA technique to face recognition, Turk and Pentland developed a well-known Eigenface method. The Eigenface method, however, does not consider the classification aspect, as it is based on the optimal representation criterion in the sense of mean-square error.

To improve the PCA standalone classification performance, one needs to combine further this optimal representation criterion with some discrimination criterion. One widely used discrimination criterion in the face recognition community is the Fisher linear discriminant (FLD), or linear discriminant analysis (LDA) [9], a well-known technique for dimensionality reduction. The second well-known approach coming under the name of fisher face is insensitive to large variation in the conditions already enumerated above. It uses both PCA and FLD. It is worth stressing that by maximizing the ratio of between-scatter matrix and within-scatter matrix, FLD produces well separated classes in a low-dimensional subspace, even under severe variation in lighting and facial expressions. There are a various enhancements made to FLD direct LDA, uncorrelated discriminant transformation [8], most discriminating feature [9]. In LDA, the dimensional embeddings are reduced in such a way that the orientation of the projected data of classes on an arbitrary line or space is well-separated from each other. There are some limitations in applying LDA directly viz. within class scatter matrix can become singular due to high dimensionality of original feature vectors in comparison with low number of training vectors available. To overcome this limitation, a number of authors have proposed the use of class-independent PCA prior to LDA in the feature extraction stage. Swets et al. [11] showed two stage PCA plus LDA method where PCA is first used for dimension reduction so as to make within class scatter matrix non-singular before the application of LDA especially when training samples are scarce. Belhumeur et al. [12] proposed a projection method which is based on LDA and PCA techniques for face recognition. In their technique class-independent PCA is first reduce the original space and then LDA is applied to reduce the dimension. Zhao et al. [13,14] demonstrated a technique based on the combination of LDA and PCA. A complete Kernel Fisher discriminant (KFD) was introduced to implement kernel PCA plus LDA strategy by Yang et al. [15] after KFD implementation by Mika et al. [16]. Wu et al. [17] presented a direct LDA method that is applicable to small sample size problems. Jian et al. [18] suggested subspace algorithm for determining the optimal projection for LDA that addressed two LDA problems viz. “small sample size” and “illumination and pose variations”. Xiaogang et al. [19] then presented a unified framework using PCA, LDA and Bayes techniques for face recognition. Ye et al. [20] showed generalized optimization criteria based on pseudoinverse for discriminant analysis to address undersample problems.

Organization of this paper is described as below: Section II provides a well-known technique of fisher face. Section III describes a proposed modified fisher face and fuzzy fisher face approach. Section IV reports on comprehensive simulation results completed for several commonly used face databases such as ORL, and Yale. Finally, concluding comments and references are included.

II. CONVENTIONAL METHOD OF FACE RECOGNITION USING FISHER FACE APPROACH

FLD is a popular discriminant criterion that measures the between- class scatter normalized by the within-class scatter.

Let $\omega_1, \omega_2, ..., \omega_L$ and $\mathrm{N}_1, \mathrm{N}_2, ..., \mathrm{N}_L$, denote the number of classes and the number of images within each class, respectively. Let $\mathrm{M}_1, \mathrm{M}_2, ..., \mathrm{M}_L$ and M be the means of the classes and the grand mean. The within- and between-class scatter matrices, S_ω , S_b are defined as follows:

$$S_\omega = \sum_{i=1}^{L} p(\omega_i)\varepsilon\{(Y^{(p)} - M_i)(Y^{(p)} - M_i)^T \mid_{\omega i}\} \quad (1)$$

$$S_b = \sum_{i=1}^{L} p(\omega_i)\varepsilon(M_i - M)(M_i - M)^T \quad (2)$$

Where, $p(\omega_i)$ is *a priori* probability, $S_\omega, S_b \in mR$, and L denotes the number of classes. FLD derives a projection matrix that maximizes the ratio $\frac{|P^T S_b P|}{|P^T S_\omega P|}$. This ratio is maximized when P consists of the eigenvectors of the covariance matrix A

$$S_\omega^{-1} S_b \psi = \psi \Delta \quad (3)$$

where ψ , $\Delta \in R^{m \times m}$ are the eigenvector and Eigen value matrices of $S_\omega^{-1} S_b$ respectively.

III. PROPOSE ALGORITHM

a) Modified Fisher Linear Discriminant approach

The Modified Fisher linear discriminant Model (MFLD) improves the generalization capability of FLD by decomposing the FLD procedure into a simultaneous diagonalization of the two within- and between-class scatter matrices. The simultaneous diagonalization is step wisely equivalent to two operations whitening the within class scatter matrix and applying PCA on the between-class scatter matrix using the transformed data.

To achieve the enhanced performance, MFLD preserves a proper balance between the need that the selected eigen values (corresponding to the principal components for the original image space) account for most of the spectral energy of the raw data, i.e., representational adequacy, and the requirement that the eigenvalues of the within class scatter matrix (in the reduced PCA space) are not too small, i.e., better generalization. The choice of the range of principal components (m) for dimensionality reduction takes into account both the spectral energy and the magnitude requirements. The eigenvalue spectrum of the covariance matrix provides a good indicator for meeting the energy criterion; one needs then to derive the eigenvalue spectrum of the within-class scatter matrix in the reduced PCA space to facilitate the choice of the range of principal components so that the magnitude requirement is met.

The stepwise FLD procedure derives the eigenvalue and eigenvectors of $S_\omega^{-1} S_b$ as the result of the simultaneous diagonalization of S_ω and S_b.

$$S_\omega v \Xi = \Xi \Gamma \text{ and } \Xi^T \Xi = 1 \quad (4)$$

$$\Gamma^{-1/2} \Xi^T S_\omega \Xi \Gamma^{-1/2} = 1 \quad (5)$$

where,

$\Xi, \Gamma \in R^{m \times m}$ are the eigenvector and the diagonal eigenvalue matrices of S_ω respectively

Different spectra are obtained corresponding to different number of principal components utilized. Tread

Volume 2 No. 2, February 2012 ISSN 2223-4985
International Journal of Information and Communication Technology Research

http://www.esjournals.org

off is to be find to optimize the behavior of the trailing eigenvalues in the reduced PCA space with the energy criteria for original image space. The eigen vectors derived are as shown in Figure 1.

Figure 1. Eigen vector images.

After the feature vector is derived, MFLD first diagonalizes the within-class scatter matrix S_ω using (4) and (5). Ξ and Γ are the eigenvector and the eigenvalue matrices corresponding to the feature vector $\Upsilon^{(p)}$. MFLD proceeds then to compute the between-class scatter matrix as follows:

$$\Gamma^{-1/2}\Xi^T S_b \Xi \Gamma^{-1/2} = K_\phi \qquad (6)$$

Diagonlize now the new between-class scatter matrix κ_b

$$\kappa_b \Theta = \Theta\gamma \quad \text{and} \quad \Theta^T \Theta = I \qquad (7)$$

where,

$\Theta, \gamma \in R^{m\times m}$ eigenvector and diagonal eigenvalue matrices of κ_b respectively

The overall transformation matrix of MFLD is now defined as follows

$$T = \Xi \Gamma^{-1/2} \Theta \qquad (8)$$

b) Fuzzy Fisher Linear Discriminant (FFLD) approach

The fisher face presented in the previous section has exhibited a substantial advantage over the PCA as far as classification aspects are concerned. But in order to improve the same technique, an alternative can be used which sophisticate the use of class assignment of patterns (faces). Further to refine classification results so that they could affect the within-class and between-class scatter matrices and enhance the performance of the classifier. Having this in mind, an obvious choice is to look at the fundamental results available in the setting of fuzzy nearest neighbor classifiers.

Given a set of feature vectors transformed by the PCA, $X = \{x_1, x_2, \ldots x_N\}$, a fuzzy c- class partition of these vectors specifies the degrees of membership of each vector to the classes. The partition matrix denoted by μ_{ij} for $i = 1, 2, \ldots, c$ and $j = 1, 2, \ldots, N$ satisfies two obvious properties

$$\sum_{i=1}^{c} \mu_{ij} = 1 \qquad (9)$$

$$0 < \sum_{j=1}^{N} \mu_{ij} < N \qquad (10)$$

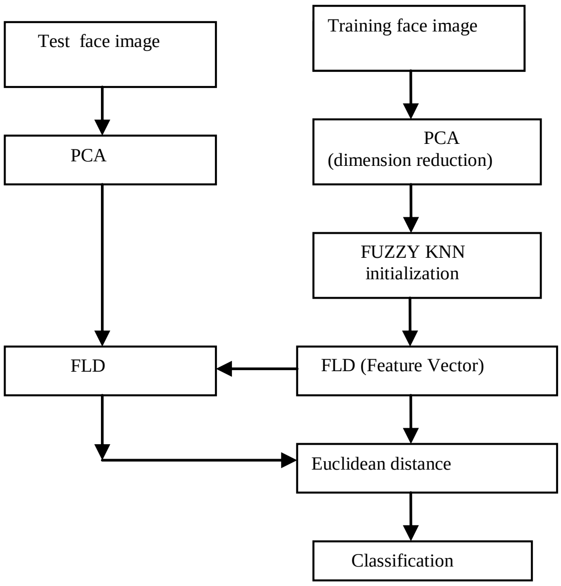

Figure 2. A general flow of computing for the fuzzy fisher face method.

The first condition helps us assure sound mathematical tractability. The computations of the membership degrees are realized through a sequence of steps:

***Step* 1:** Compute the Euclidean distance matrix between pairs of feature vectors in the training.

***Step* 2:** Set diagonal elements of this matrix to infinity (practically place large numeric values there).

***Step* 3:** Sort the distance matrix (treat each of its columns separately) in an ascending order. Collect the class labels of the patterns located in the closest neighborhood of the pattern under consideration (as we are concerned with "*k*" neighbors, this returns a list of "*k*" integers).

***Step* 4:** Compute the membership grade to class i for j^{th} pattern

$$\mu_{ij} = \begin{cases} 0.51 + 0.49(n_{ij}/k) & \text{if } i = \text{the same as the lable of the } j^{th} \text{ pattern} \\ 0.49(n_{ij}/k) & \text{if } i \neq \text{the same as the lable of the } j^{th} \text{ pattern} \end{cases} \tag{11}$$

In the above expression, n_{ij} stands for the number of the neighbors of the j^{th} data (pattern) that belong to the i^{th} class. After the examination of the membership location formula we conclude that the method attempts to "fuzzify" or refine the membership grades of the labeled patterns.

Figure 3. Fuzzy fisher faces

IV. RESULTS AND DISCUSSIONS

The algorithms are tested on Yale, and ORL database to compute recognition rate. The recognition rate for above proposed algorithms is carried out in two parts as in-database and out-database. In database means training images and testing images are the same whereas for out-database training images and testing images are different. The result obtained for in-database are listed in table I and result obtained for out-database are listed in table II

Table I. In-Database Recognition Rate

Face database	Recognition rate		
	FLD	MFLD	FFLD
ORL	95%	100%	100%
Yale	94%	100%	100%

MFLD and FFLD have 100% identification rate. In contrast, FLD has only 94% for Yale database and 95% for ORL database. It means that FLD doesn't learn well when we train the input images. In fact, FLD only maximizes the distance of each subject. As a result, these projected coefficients of different subject could be distributed in the same area and FLD system cannot identify these faces.

Table II. Out-Database Recognition Rate

Face database	Recognition rate		
	FLD	MFLD	FFLD
ORL	75.5	90.7	95.5
Yale	83.45	91.4	94.8

The MFLD and FFLD technique outperforms in identifying a person with face as a biometric feature as compared with traditional FLD. The experimental results show improved recognition rates and reduced sensitivity to variations between face images caused by changes in illumination and viewing directions.

V. CONCLUSIONS

We have proposed a generalized version of the fisher face method for face recognition by including refined information about class membership of the binary labeled faces (patterns). By doing this we were able to reduce sensitivity of the method to substantial variations between face images caused by varying illumination, viewing conditions, and facial expression. Experimental results showed a consistently better classification rates in comparison to other "standard" methods such as Eigenface and fisher face when applied to ORL, Yale.

ACKNOWLEDGEMENT

Authors are thankful to Dept. of Computer Science, Dr BAMU, Aurangabad for providing necessary facilities to carry out this research work.

REFERENCES

[1] R. Chellappa, C. L.Wilson, and S. Sirohey, "Human and machine recognition of faces: A survey," *Proc. IEEE*, vol. 83, pp. 705–740, May 1995.

[2] J. Daugman, "Face and gesture recognition: Overview," IEEE Transaction Pattern Analysis and Machine Intelligence, vol. 19, no. 7, pp. 675–676, 1997.

[3] A. Pentland, "Looking at people: Sensing for ubiquitous and wearable computing," IEEE Transaction Pattern Analysis and Machine Intelligence, vol. 22, pp. 107–119, Jan. 2000.

[4] A. Samal and P. A. Iyengar, "Automatic recognition and analysis of human faces and facial expression: A survey," *Pattern Recognition*, vol. 25, no. 1, pp. 65–77, 1992.

[5] C. Liu and H.Wechsler, "Evolutionary pursuit and its application to face recognition," IEEE Transaction Pattern Analysis and Machine Intelligence, vol. 22, pp. 570–582, June 2000.

[6] M. Turk, A. Pentland, "Face recognition using eigenfaces," Proceeding of IEEE Conference on Computer Vision and Pattern Recognition, Maui, HI, USA, June 1991, pp. 586–591. (second Paper M. Turk and A. Pentland, "Eigenfaces for recognition," *J. Cogn. Neurosci.*, vol. 13, no. 1, pp. 71–86, 1991.

[7] P.N. Belhumeur, J.P. Hespanha, D.J. Kriegman, "Eigenfaces vs. fisherfaces: recognition using class specific linear projection," IEEE Transaction Pattern Analysis and Machine Intelligence, 19 (7) (1997) 711–720.

[8] M. Kirby and L. Sirovich, "Application of the Karhunen-Loeve procedure for the characterization of human faces," IEEE Transaction Pattern Analysis and Machine Intelligence, vol. 12, pp. 103–108, Jan. 1990.

[9] R. A. Fisher, "The use of multiple measures in taxonomic problems," *Ann. Eugenics*, vol. 7, pp. 179–188, 1936.

[10] R.O. Duda, P.E. Hart, "Pattern Classification and Scene Analysis," Wiley, New York, 1973.

[11] D.L. Swets, J. Weng, "Using discriminative eigenfeatures for image retrieval," IEEE Transaction Pattern Analysis and Machine Intelligence, 18 (8) (1996) 831–836.

[12] P.N. Belhumeur, J.P. Hespanha, D.J. Kriegman, Eigenfaces vs. fisherfaces: recognition using class specific linear projection, IEEE Transaction Pattern Analysis and Machine Intelligence, 19 (7) (1997) 711–720.

[13] W. Zhao, R. Chellappa, P.J. Phillips, "Subspace linear discriminant analysis for face recognition," Technical Report CAR-TR-914, Center for Automation Research, University of Maryland, College Park, 1999.

[14] W. Zhao, R. Chellappa, N. Nandhakumar, "Empirical performance analysis of linear discriminant classifiers," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 1998, pp. 164–169.

[15] J. Yang, A.F. Frangi, J.-y. Yang, D. Zhang, J. Zhong, "KPCA plus LDA: a complete kernel Fisher discriminant framework for feature extraction and recognition," IEEE Transaction Pattern Analysis and Machine Intelligence, 27 (2) (2005) 230–244.

[16] S. Mika, G. Rätsch, J. Weston, B. Schölkopf, A. Smola, K.-R. Müller, "Constructing descriptive and discriminative nonlinear features: Rayleigh coefficients in kernel feature spaces," IEEE Transaction Pattern Analysis and Machine Intelligence, 25 (5) (2003) 623–628.

[17] X.-J. Wu, J. Kittler, J.-Y. Yang, K. Messerm S. Wang, "A new direct LDA (D-LDA) algorithm for feature extraction in face recognition," Proceedings of the International Conference on Pattern Recognition vol. 4, 2004, pp. 545–548.

[18] H. Jian, P.C. Yuen, C. Wen-Sheng, "A novel subspace LDA algorithms for recognition of face images with illumination and pose variations," Proceedings of the International Conference on Machine Learning and Cybernetics, vol. 6, 2004, pp. 3589–3594.

[19] W. Xiaogang, T. Xiaoou, "A unified framework for subspace face recognition," IEEE Transaction Pattern Analysis and Machine Intelligence, 26 (9) (2004) 1222–1228.

[20] J. Ye, R. Janardan, H.P. Cheong, P. Haesun, "An optimization criterion for generalized discriminant analysis on undersampled problems," IEEE Transaction Pattern Analysis and Machine Intelligence, 26 (8) (2004) 982–994.