

# Face Recognition System based on Subspace Linear Discriminant Analysis

Sandeep Mishra

M. Tech Scholar, Digital Electronics  
Dept of E&Tc Engg, LNCT Jabalpur

Anupam Dubey

Assistant Professor,  
Dept of E&Tc Engg, LNCT Jabalpur

Nisha Bhatt

Professor & Head,  
Dept of E&Tc Engg, LNCT Jabalpur

**Abstract-** This paper is based on subspace linear discriminant analysis in which facial features are extracted by using Principal Component Analysis followed by Linear Discriminant Analysis based dimension reduction techniques. On the basis of literature review methodology is designed which includes sequence as preprocessing, dimension reduction by PCA, feature extraction for class separability by LDA and finally classification. ORL face database is used whose experimental results is extracted as recognition rate of 96.35% and hence giving efficiency of the proposed method over traditional methods of automatic face recognition systems.

**Keywords-** Biometrics, PCA, LDA, Recognition Rate

## I. INTRODUCTION

Face recognition technologies have been widely used in commercial, law enforcement, and military applications. An automatic system performs face detection, face verification and finally face recognition in various high alert areas such as airport security, traffic supervision and monitoring, HCI and environments like cars, mobiles. But challenges associated with this technology of biometrics are deeper under variations in illumination conditions, occlusion, poses, expression, aging, and disguises. Researchers had proposed methods till now, and hence found that appearance-based approaches yield good outcome, which is operated directly on images or appearances of face, where complete face image is operated and processed as two dimensional matrix. The result face image features are then used to represent the class in which it is lying. Features are used such that they have high separability when processed further and lower separability is discarded [1][2].

To decrease computational time, database, generated after various transformations, are symbolize in a lower dimensional space. Efficient appearance-based methods like Principal Component Analysis, Linear Discriminant Analysis and Independent Component Analysis, make use of mathematical techniques like eigenfaces, fischerfaces and independent components respectively. PCA gives class representations which are in an orthogonal linear space. However LDA generates class discriminatory information in a linear separable space which is not necessarily orthogonal [3]. These techniques give good results under varying condition. In these techniques dimensions are first reduced by

eigen faces and then applying fischer space for class separability.

## II. RELATED WORK AND BACKGROUND

In feature based approach key information of facial features such as eyes, nose, mouth, and chin, are gathered with the help of deformable templates and extensive mathematics and converted into a feature vector. Yuille et. al. [4] proposed deformable templates techniques, where face and its features are determined by interactions with the face image. In Image based approach information theory concepts like class information, class separability, independent components, energy contents of whole face image are utilised. As whole face image is used so this method is also termed as holistic method. Turk et. al. [5] developed Principal Component Analysis based face recognition using eigenface techniques. The term eigenface is used because mathematical algorithms using eigenvectors represent the primary components of the face. Weights are used to represent the eigenface features so a comparison of these weights permits identification of individual faces from a database. Zhao t. al. [6] uses linear discriminant analysis to maximize the scatter between different classes and minimize the scatter between the input data in the same class. While PCA tries to generalize the input data to extract the features, LDA tries to discriminate the input data by dimension reduction. Utilizing both this techniques face recognition is performed by subspace LDA.

## III. FEATURE EXTRACTION ALGORITHM AND CLASSIFICATION TECHNIQUES

In this section PCA and LDA techniques are explained mathematically. Principal Component Analysis is used to reduce dimension of dataset while Linear Discriminant Analysis is used to increase differences between different class so as to increase accuracy and reduce computational cost.

### A. Principal Component Analysis

Principal Component Analysis method is dimension reduction technique that maximizes the scatter of all projected samples. It is achieved by utilizing eigenspace.

In PCA, one dimensional feature vectors of each image in training dataset are extracted and hence training dataset

image space is generated. PCA categorize images according to distance between feature vectors of image space [7]

For  $N$  number of sample images  $(x_1, x_2, x_3 \dots x_N)$  with an  $n$ -dimensional image space, if each image belong to one of  $C$  classes  $\{X_1, X_2, \dots, X_C\}$ , then in PCA, linear transformation of the original  $n$  dimensional image space into an  $m$ -dimensional feature space will be performed, where  $m < n$ .

The new feature vectors  $y_k$  will be then given by

$$y_k = \mathbf{W}^T x_k \quad (k = 1, 2, 3, \dots, N) \quad (1)$$

Where  $y_k \in \mathbb{R}^m$  is a linear transformation matrix &  $\mathbf{W} \in \mathbb{R}^{n \times m}$  is a matrix with orthonormal columns.

From here the total scatter matrix,

$$S_T = \sum_{k=1}^N (x_k - \mu)(x_k - \mu)^T \quad (2)$$

where  $\mu \in \mathbb{R}^m$  is the mean image training database, after applying the linear transformation  $\mathbf{W}^T$ , and transformed feature vectors  $(y_1, y_2, y_3 \dots y_N)$  will be given by  $\mathbf{W}^T S_T \mathbf{W}$ .

In PCA, the optimal projection matrix  $\mathbf{W}_{opt}$  is selected to maximize the determinant of the total scatter matrix of the projected samples, i.e.

$$\mathbf{W}_{opt} = \arg \max_{\mathbf{W}} |\mathbf{W}^T S_T \mathbf{W}| \quad (3)$$

$$\mathbf{W}_{opt} = [\mathbf{W}_1 \ \mathbf{W}_2 \ \dots \ \mathbf{W}_m] \quad (4)$$

where  $\mathbf{W}$  is the is the set of  $n$ -dimensional eigenvectors of total scatter matrix,  $S_T$ , corresponding to the  $m$  largest eigenvalues [8].

#### B. Linear Discriminant Analysis

Linear Discriminant Analysis make use of fischer space method to achieve maximum discrimination of classes and also to achieve dimensionality reduction. In LDA within-class and between-class scatter matrices are defined [6].

For a set of  $N$  sample images  $(x_1, x_2, x_3 \dots x_N)$ , taking values in an  $n$ -dimensional image space, and suppose that each image belongs to any one of  $C$  classes  $\{X_1, X_2, \dots, X_C\}$ .

Hence Within class scatter matrix is given by

$$S_W = \sum_{i=1}^C \sum_{x_k} (x_k - \mu_i)(x_k - \mu_i)^T \quad (5)$$

And Between class scatter matrix is given by

$$S_B = \sum_{i=1}^C N_i (\mu_i - \mu)(\mu_i - \mu)^T \quad (6)$$

where  $\mu_i$  is the mean image of class  $X_i$ , and  $N_i$  is the number of samples in class  $X_i$ .

In LDA, the optimal projection  $\mathbf{W}_{opt}$  is selected to maximizes the ratio of the determinant of the between-class scatter matrix of the projected samples to the determinant of the within-class scatter matrix of the projected samples.

$$\mathbf{W}_{opt} = \arg \max_{\mathbf{W}} \frac{|\mathbf{W}^T S_B \mathbf{W}|}{|\mathbf{W}^T S_W \mathbf{W}|} \quad (7)$$

$$\mathbf{W}_{opt} = [\mathbf{W}_1 \ \mathbf{W}_2 \ \dots \ \mathbf{W}_m] \quad (8)$$

where  $\mathbf{W}$  is the set of generalized eigenvectors of  $S_B$  and  $S_W$  corresponding to the  $m$  largest generalized eigenvalues  $\lambda$ , such that

$$S_B \mathbf{W}_i = \lambda_i S_W \mathbf{W}_i \quad (i = 1, 2, \dots, m) \quad (9)$$

Also, there should be eigenvectors corresponding to at most  $(C-1)$  largest eigenvalues [9] [10].

#### C. Classification Techniques

In order to recognize any image from dataset, either distance between the images in  $N$ -dimensional space is used or similarity between them is adopted. For best match either distance measure should be minimum or similarity measure should be maximum for high level of similarity [11]. Some technique are discussed below:

1)  $L_1$  norm: Also known as the city block norm or the sum norm. It sums up the absolute difference between pixels. The  $L_1$  norm of an image  $X$  and an image  $Y$  is:

$$L_1(X, Y) = \sum_{i=1}^N |X_i - Y_i| \quad (10)$$

2)  $L_2$  norm: Also known as the Euclidean norm or the Euclidean distance when its square root is calculated. It sums up the squared difference between pixels. The  $L_2$  norm of an image  $X$  and an image  $Y$  is:

$$L_2(X, Y) = \sum_{i=1}^N (X_i - Y_i)^2 \quad (11)$$

3) Covariance: Calculates the angle between two normalized vectors. Taking the dot product of the normalized vectors. The covariance between images  $A$  and  $B$  is:

$$\text{cov}(A, B) = \frac{A}{\|A\|} \cdot \frac{B}{\|B\|} \quad (12)$$

The similarity measurement technique, covariance is also known as the angle measure.

#### IV. PROPOSED FACE RECOGNITION ALGORITHM

PCA performs dimension reduction by projecting the data onto the eigenface space upon while LDA performs class separability by classifying the eigenface space projected data.

The method consists of four stages:-

- Preprocessing
- PCA
- LDA
- Classification.

Flowchart of the subspace linear discriminant analysis is shown below.

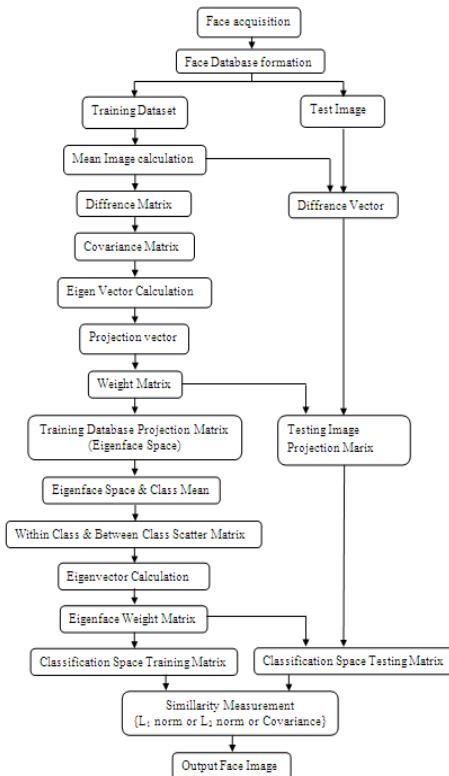


Fig. 1 Flowchart for Face Recognition by Subspace LDA method

## V. EXPERIMENTAL RESULTS

In this work, Olivetti and Oracle Research Laboratory (ORL) face database is used which provided 400 face images of 40 individuals in JPG format. Out of 10 face images of each individual, 8 face images of each person are taken for training and 2 images are taken for testing dataset. These images are of size 92x112 pixels with 256 gray levels. For some subjects, the images were taken at different times, varying the lighting, facial expressions (open / closed eyes, smiling / not smiling) and facial details (glasses, moustaches, and beard). All the images were taken same background as shown in fig. 4.1.



Fig. 4.1 Preview of individual's face image from ORL Face database

Original images are resized to 64x64 pixels so that the input space has the dimension of 4096 followed by histogram equalization. After vectorization column wise, mean image is calculated as shown in fig. 4.2.



Fig 2 Mean Image

Some highest Eigen face of training dataset are represented in shown in fig. 4.3.

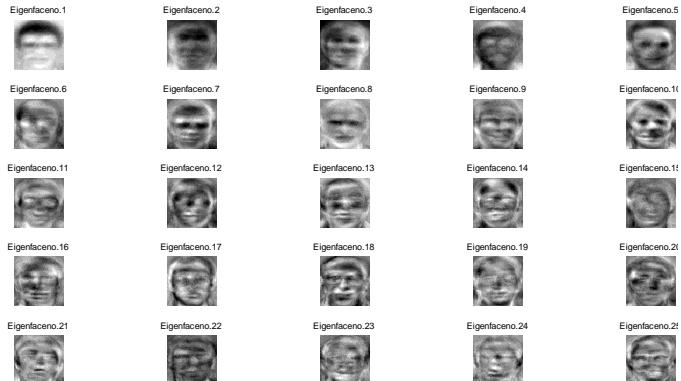


Fig 3. 16 Eigen Faces

These eigenfeatures from eigenface space projection matrix is derived and over it LDA is applied. LDA further reduces dimension of projection matrix and increases recognition rate and reduces computation time. Classification space projection matrix, derived after LDA, is classified with various techniques. Result after classification is shown in fig. 4.4



Fig 4 Recognized Face Image after Subspace LDA

PCA based face recognition is performed by extracting 70, 100, 250 & 300 eigen faces and compared their recognition rates by  $L_1$ ,  $L_2$  and covariance.

TABLE I RECOGNITION RATE OF PCA BASED FACE RECOGNITION WITH VARYING PCA FEATURES

S. No	No. of PCA features	Classification Techniques		
		$L_1$ Norm	$L_2$ Norm	Covariance
1	70	95.00%	93.75%	95.00%
2	100	96.25%	96.25%	93.75%
3	250	96.25%	97.50%	93.75%

4	300	97.50%	97.50%	92.50%
---	-----	--------	--------	--------

LDA followed by PCA and hence Subspace LDA based face recognition is performed by extracting 40, 70 & 100 PCA features and corresponding 15, 25 & 39 LDA features and classified by City Block Distance Calculation. Various recognition rates are then compared.

TABLE II RECOGNITION RATE OF SUBSPACE LDA BASED FACE RECOGNITION WITH VARYING PCA & LDA FEATURES

S.No.	PCA Features	LDA Features		
		15	25	39
1	<b>40</b>	77.50%	86.25%	<b>96.35%</b>
2	<b>70</b>	72.50%	82.50%	90%
3	<b>100</b>	66.50%	65%	77.50%

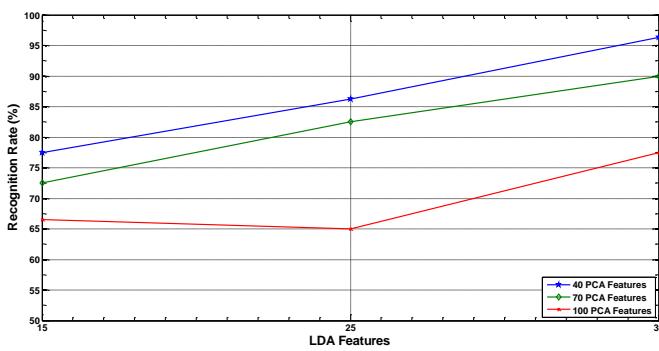


Fig 5 Graphical Representation of Recognition Rate with varying LDA Features

TABLE III COMPARISON OF RECOGNITION RATE OF PROPOSED ALGORITHM WITH THAT OF OTHER METHODS

Method	Recognition Rate
PCA with L <sub>2</sub> Norm [5]	84.00%
DCT Based Face Recognition [14]	84.50%
K-means [15]	86.75%
PCA & LDA with L <sub>2</sub> Norm [4][6]	94.80%
Fuzzy Ant with fuzzy C-means [15]	94.82%
Proposed Algorithm	<b>96.35%</b>

## VI. CONCLUSION

This work is subspace LDA which is PCA followed by LDA hence classification is performed by City Block Distance Calculation. ORL Face Database is used for experiment, which provided face images under varying facial condition [16]. Experimental results are giving recognition rate equal to 96.35 % at 40 numbers of PCA features. Hence it can be concluded that even if PCA features is less LDA is giving better result.

## REFERENCES

- Zhao W., Chellappa R., Phillips P. J. and Rosenfeld A., "Face Recognition: A Literature Survey", ACM Computing Surveys, Vol. 35, No. 4, pp. 399–458, 2003,
- Patil A.M., Kolhe S. R. and Patil P.M, 2D Face Recognition Techniques: A Survey, International Journal of Machine Intelligence, ISSN: 0975–2927, Volume 2, Issue 1, pp-74-8, 2010
- Belhumeur P.N., Hespanha J.P., and Kriegman D. J., "Eigen faces vs. Fisher faces: Recognition using class specific linear projection", IEEE Trans. Pattern Anal. Machine Intel., vol. 19, PP. 711-720, 1997.
- Yuille A. L., Cohen D. S., and Hallinan P. W., "Feature extraction from faces using deformable templates", Proc. of CVPR, 1989
- Turk M., Pentland A., "Eigen faces for face recognition", Journal cognitive neuroscience, Vol. 3, No.1, 1991.
- Zhao W., Chellappa R., Krishnaswamy A., "Discriminant analysis of principal component for face recognition", .IEEE Trans. Pattern Anal. Machine Intel., Vol 8, 1997
- Turk M. and Pentland A., Eigenfaces for recognition. J. Cogn. Neurosci. 3, 72–86, 1991
- Lee S. J., Yung S. B., Kwon J. W., and Hong S. H., "Face Detection and Recognition Using PCA", pp. 84-87, IEEE TENCON, 1999.
- Swets D.L. and Weng J.J., "Using Discriminant Eigen features for image retrieval", IEEE Trans. Pattern Anal. Machine Intel, vol. 18, PP. 831-836, 1996.
- Etemad K. and Chellappa R., "Face Recognition Using Discriminant Eigenvectors", IEEE Transaction for Pattern Recognition, pp. 2148-2151,1996.
- Phillips P. J., and Moon H., "Comparison of Projection-Based Face Recognition Algorithms", , IEEE Transaction for Pattern Recognition, pp. 4057-4062, 1997.
- Sirovich L. and Kirby, M., "Low-dimensional procedure for the characterization of human faces", International Journal of Optical Society4, 3,pp. 519-524, 1987
- Pang S., Ozawa S., Kasabov N., "Incremental linear discriminant analysis for classification of data streams", IEEE Trans. on Systems, Man and Cybernetics, vol. 35, no. 5, pp. 905-914, 2005.
- Zhao S., Grigat R. R., "Multi block fusion scheme for face recognition", Int. Con. on Pattern recognition (ICPR), Vol. 1, pp. 309-312, 2004.
- Makdee S., Kimpan C., Pansang S. , "Invariant range image multi – pose face recognition using Fuzzy ant algorithm and membership matching score" Proceedings of 2007 IEEE International Symposium on Signal Processing and Information Technology , pp. 252-256, 2007
- Olivetti & Oracle Research Laboratory, The Olivetti & Oracle Research Laboratory Face database of Faces, <http://www.camorl.co.uk/facedatabase.html>