APPEARANCE BASED FACE RECOGNITION BY PCA AND LDA

S. K. Hese, M. R. Banwaskar

Department of Electronics & Telecommunication, MGM’s College of Engineering Nanded
Near Airport, Nanded, Maharashtra, India

ABSTRACT

In the recent years, Face Recognition has become one of the most challenging tasks in pattern recognition field. Human face is most obvious human identifier. The face is the most visible part of human anatomy and act as the first distinguishing factor of a human being. Each individual has his own uniqueness and this could be one of the most transparent and unique feature to differentiate a person from one to another. Images can be analyzed and faces can then be identified, before they can be recognized. There are different methods of face recognition which involve a series of steps that serve to capturing, analyzing and comparing a face to a database of stored images. Different algorithms have been designed for face recognition. The aim of this paper is to present independent and comparative study of most popular appearance based linear algorithms of face recognition based on projection methods such as PCA and LDA. In this work, we have applied image preprocessing methods such as gray scale conversion and modified histogram equalization on database prior to PCA and LDA, in order to enhance the quality of image and overall recognition performance. It is generally said that the algorithms based on LDA are better than the algorithms based on PCA. But in present research work we have found that sometimes PCA also outperforms LDA with respect to dimensionality.

Keywords- Face recognition, grayscale conversion, image preprocessing, LDA, modified histogram equalization pattern recognition, PCA.

I. INTRODUCTION

Face recognition is one of the important challenges in appearance based recognition field. It has got significant attention for decades because of its numerous potential applications. It includes identification and verification of the person which is helpful in maintaining national security, law enforcement, and public safety field [1, 2]. Face
recognition has also proven useful in applications such as human-computer interaction, virtual reality, database retrieval, multimedia, and computer entertainment. Research in automatic face recognition started in the 1960’s [1, 3]. Kirby and Sirovich [4] were among the first to apply principal component analysis (PCA). Turk and Pentland popularized the use of PCA for face recognition [5]. They used PCA to compute a set of subspace basis vectors (which they called “eigenfaces”) [5] for a database of face images, and projected the images in the database into the compressed subspace. One alternative is Fisher’s linear discriminant analysis [6]. For any N-class classification problem, the goal of LDA is to find the N-1 basis vectors that maximize the interclass distances while minimizing the intraclass distances. At one level, PCA and LDA are very different: LDA is a supervised learning technique that relies on class labels, whereas PCA is an unsupervised technique. LDA has been compared to PCA in several studies [7, 8, 9]. One characteristic of both PCA and LDA is that they produce spatially global feature vectors. In other words, the basis vectors produced by PCA and LDA are non-zero for almost all dimensions, implying this change to a single input pixel, will alter every dimension of its subspace projection.

As we know that any image or face has size n x m pixels which require n.m dimensional space. This space is too large and needs to be reduced for better recognition which is achieved by dimensionality reduction techniques [9]. We have two most popular techniques for these purposes that are principal component analysis (PCA) and linear discriminant analysis (LDA) which is also known as Fisher discriminant analysis [1, 5, 6]. For better performance we have implemented these two algorithms with several preprocessing factors such as gray scale conversion and modified histogram equalization before recognition algorithms. Image preprocessing techniques represent an essential part of face recognition systems, which has a great impact on the performance and robustness of the recognition procedure. Amongst the number of techniques already presented in the literature, histogram equalization has emerged as the dominant preprocessing technique and is regularly used for the task of face recognition. The aim of this paper is to study the PCA and LDA with respect to face recognition rate and dimensionality. The experiments are based on face94 database. The organization of this paper is done in six sections. Section II describes the preprocessing methods performed on facial images. Section III provides introduction to PCA and its mathematical derivation. In section IV we discussed LDA and the related mathematical analysis. Results and conclusion are presented in section V & VI respectively.

II. IMAGE PREPROCESSING

The following preprocessing techniques have been applied in order to enhance the quality of input images which also helped to improve the recognition performance of the algorithms.

Grayscale Conversion

In order to retain as much as information of images, the color images are converted into grayscale images. This is the first step of experiment. As color images (RGB images) are composed of 3 channels to present red, green and blue components in RGB space. Pixels in grayscale images are stored 8–bit integer to represent color into black and white [10, 11].
Histogram Equalization

An image histogram is a graphical representation of the tonal distribution in a digital image. It plots the number of pixels for each tonal value. The horizontal axis of the graph represents the tonal variations, while the vertical axes represent the number of pixels in that particular tone. Particularly, histogram of a gray-scale image is in the horizontal axes at 256 brightness levels, and in the vertical axes, the number of times this level appears in the image. In Histogram Equalization the global contrast enhancement is obtained using the cumulative density function of the image as a transfer function [12, 13]. The result is a histogram approximately constant for all gray values. While Local Histogram Equalization enhances details of the image, global Histogram Equalization enhances the contrast of whole image. The problem of the first approach is that the output is not always realistic, but in this case the representation of the face image must be invariant to lighting variations and not a realistic image [13]. To equalize the image histogram the cumulative distribution function (cdf) has been computed. The cdf of each gray level is the sum of its recurrence and the recurrence of the previous gray level in an image. The histogram equalization formula is:

\[
h(v) = \text{round} \left( \frac{\text{cdf}(v) - \text{cdf}_{\text{min}}}{(M \times N) - \text{cdf}_{\text{min}}} \times (L - 1) \right)
\]

Where cdf_{min} is the minimum value of the cumulative distribution function, M and N are the width and the height of image, and L is the number of gray levels used. The result is an equalized and normalized image.

III. PRINCIPAL COMPONENT ANALYSIS

Principle Component Analysis (PCA) is a dimensionality reduction technique that is used for image recognition and compression. It is also known as Karhunen-Loeve transformation (KLT) [1, 2, 4, 5] or eigenspace projection. The major advantage of PCA is that the eigenface approach helps reducing the size of the database required for recognition of a test image. The trained images are not stored as raw images; rather they are stored as their weights which are found out by projecting each and every trained image to the set of eigenfaces obtained. In the language of information theory, the relevant information in a face needs to be extracted, encoded efficiently and one face encoding is compared with the similarly encoded database. The trick behind extracting such kind of information is to capture as many variations as possible from the set of training images.

Mathematically, the principal components of the distribution of faces are found out using the eigenface approach [1, 5, 11]. First the eigenvectors of the covariance matrix of the set of face images is found out and then they are sorted according to their corresponding eigenvalues. Then threshold eigenvalue is taken into account and eigenvectors with eigenvalues less than that threshold values are discarded. So ultimately the eigenvectors having the most significant eigenvalues are selected. Then the set of face images are projected into the significant eigenvectors to obtain a set called eigenfaces. Every face has a contribution to the eigenfaces obtained. The best M eigenfaces from an M dimensional subspace is called “face space”[5]. Each individual face can be represented exactly as the linear combination of “eigenfaces” or each face can also be approximated using those significant eigenfaces obtained using the most significant eigen values. Now the test image
subjected to recognition is also projected to the face space and then the weights corresponding to each eigenface are found out. Also the weights of all the training images are found out and stored. Now the weights of the test image are compared to the set of weights of the training images and the best possible match is found out. The comparison is done using the “Euclidean distance” measurement. Minimum the Euclidean distance maximum will be the match.

Calculating Eigenfaces

Consider an N x N face image \( \Gamma(x, y) \) as a vector of dimension \( N^2 \), so the image can be thought as a point in \( N^2 \) dimensional space. Acquired database of M images can therefore be mapped to a collection of points in high dimensional “facespace” as \( \Gamma_1, \Gamma_2, \Gamma_3, \ldots, \Gamma_M \).

To compute Eigenfaces, first an average of all training images needs to be computed. The average image \( \Psi \) is computed by using following equation.

\[
\Psi = \frac{1}{M} \sum_{n=1}^{M} \Gamma_n
\]

Each image \( \Gamma_1 \) differs from the average image \( \Psi \) by the vector \( \Phi_i = \Gamma_i - \Psi \) where \( i = 1, 2, \ldots, M \). The covariance matrix \( C \) of the data set is defined by Equation given below.

\[
C = \frac{1}{M} \sum_{n=1}^{M} \Phi_n \Phi_n^T = AA^T
\]

Where the matrix \( A = [\Phi_1, \Phi_2, \ldots, \Phi_M] \). The matrix \( C \) has a dimension of \( N^2 \times N^2 \) eigenvectors and eigenvalues, and, for typical image sizes, this size would be a very high value. Therefore, we need a computationally feasible method to determine these eigenvectors. If the number of data points in the image space is less than the dimension of space \( M < N^2 \), there will be only \( (M - 1) \) meaningful eigenvectors, and the rest of the eigenvectors will have eigenvalues of zeros.

Consider eigenvectors \( \nu_i \) of \( A^T A \) such that

\[
A^T A \nu_i = \mu_i \nu_i
\]

Multiplying both sides by \( A \), we have

\[
A A^T A \nu_i = \mu_i A \nu_i
\]

From above equation, we can see that \( A \nu_i \) are the eigenvectors of \( C = AA^T \). Then we can construct an \( M \times M \) matrix as follows.

\[
L = A^T A
\]
Afterwards we can find the M eigenvectors $u_i$ of L. These vectors determine linear combinations of the M training set face images to form Eigenfaces $u_i$.

$$u_i = \sum_{k=1}^{M} u_{ik}\Phi_k$$

Eigenvectors $u_i$ are in fact images, and they are called eigenfaces, and the eigenvectors with the highest eigenvalues are more useful than the rest. Therefore those $M^I < M$ eigenfaces that are most significant are used for constructing the “face subspace” for projections that are used in identifying or classifying images.

**Eigenfaces for Projection and Classification**

In order to identify an input image, first of all, that image needs to be transformed into the face space by taking inner products with eigenfaces. These inner products compose a vector, which is the representation of input image in face space. Then we compare this vector with those of known images using Euclidean distance. The input test image is then classified as the closest match.

A new face image $\Gamma$ is transformed into its eigenface components (i.e., projected into “face space”) using the following operation.

$$\omega_k = u_k^T(\Gamma - \Psi) \quad k = 1, 2, ..., M^I$$

Where $\omega_k$ is the $k$-th co-ordinate of the $\Phi$ in the new “face space”. The above equation describes point-by-point image multiplication and summation, resulting in the scalar value (dimension 1x1) defined as a weight that describes each face. The weights form a feature vector as given by below Equation.

$$\Omega^T = [\omega_1, \omega_2, ..., \omega_M]$$

Vector $\Omega^T$ in above equation describes the contribution of each eigenface in representing the input test face image, where the eigenfaces represent the basis set for face images. This vector is used in finding the class that the input image belongs to, if there are more than one image describing a person, otherwise it is used to find which single image is most similar to the test image. This decision is reached by computing the Euclidean distance between input image and training images in the face database (i.e., class projections), as given in Equation

$$\varepsilon_k = ||\Omega - \Omega_k||^2$$

Where $\Omega_k$ describes $k$-th face class, which is the average of the eigenface representation of the face images for each individual in the training set. A face will be classified as belonging to some class if the minimum $\varepsilon_k$ is below some specified threshold. Otherwise, that image will be classified as unknown face.
IV. LINEAR DISCRIMINANT ANALYSIS

Linear Discriminant is a “classical” technique in pattern recognition where it is used to find a linear combination of features which characterize or separate two or more classes of objects or events [6, 7]. The resulting combination may be used as a linear classifier or more commonly, for dimensionality reduction before it can be classified. The LDA methods or Fisher discriminant, group images of the same class and separate images of different classes. Images are projected from N-dimensional space (where N is the number of pixels in the image) to \( \frac{C}{C-1} \) dimensional space (where C is the number of classes of images). For example, consider two sets of points in 2-dimensional space that are projected onto a single line. Depending on the direction of the line, the points can either be mixed together (Fig. 1(a) or separated (Fig. 2(b)). In computerized face recognition, each face is represented by a large number of pixel values. Linear discriminant analysis is primarily used here to reduce the number of features to a more manageable number before classification. Each of the new dimensions is a linear combination of pixel values, which form a template. The linear combinations obtained using Fisher's linear discriminant are called Fisher faces.

Fisher discriminant finds the line that best separates the points. To identify a test image, the projected test image is compared to each projected training image, and the test image is identified as the closest training image. As with eigenspace projection, training images are projected into a subspace. The test images are projected into the same subspace and identified using a similarity measure. What differs is how the subspace is calculated. The following sections describe how to find the Fisher discriminant for a set of images.

The within-class scatter matrix, also called intra-personal, represents variations in appearance of the same individual due to different lighting and face expression, while the between-class scatter matrix, also called the extra-personal, represents variations in appearance due to a difference in identity [6, 7, 9, 12]. By applying this method, we find the projection directions that on one hand maximize the Euclidean distance between the face images of different classes and on the other hand minimize the distance between the face images of the same class. The ratio is maximized when the column vectors of the projection matrix are the eigenvalues of \( s_w^{-1}s_b \) [17, 18].

![Fig.1. (a) Points mixed when projected onto line.](image1.png)

![Fig.1. (b) Points separated when projected onto line.](image2.png)
Calculating Fisher Faces

Linear Discriminant Analysis (LDA) method tries to find the subspace that best discriminates different face classes as shown in Fig. 2. It is achieved by maximizing the between-class scatter matrix $S_b$, while minimizing the within-class scatter matrix $S_w$ in the projective subspace. $S_w$ and $S_b$ are defined as

$$
S_w = \sum_{j=1}^{c} \sum_{i=1}^{N_j} (X_i^j - \mu_j) (X_i^j - \mu_j)^T
$$

Where $X_i^j$ is the $i$th sample of class $j$, $c$ is the number of classes, and $N_j$ is the number of samples in class $j$.

$$
S_b = \sum_{j=1}^{c} (\mu_j - \mu)(\mu_j - \mu)^T
$$

Where $\mu$ represents the mean of all classes. The subspace for LDA is spanned by a set of vectors $W= [W_1, W_2... W_d]$ satisfying

$$
W = \text{argmax} \left| \frac{W^T S_b W}{W^T S_w W} \right|
$$

The within class scatter matrix represents how face images are distributed closely within classes and between class scatter matrix describes how classes are separated from each other. When face images are projected into the Discriminant vectors $W$, face images should be distributed closely within classes and should be separated between classes, as much as possible. In other words, these Discriminant vectors minimize the denominator and maximize the numerator in above equation and therefore can be constructed by the eigenvectors of $S_w^{-1} S_b$ [17, 18].
V. EXPERIMENTAL RESULTS

In this work, we have used face94 database which contains 3000 images having size 180x200. We have used MATLAB.10 for implementation. First of all the image preprocessing steps are carried out on this database and then PCA and LDA are implemented simultaneously as discussed above. Following figures shows the results of PCA and LDA based algorithms.

After preprocessing, eigenfaces were obtained by PCA as shown in Fig. 3. Similarly fisherfaces were obtained by LDA based algorithm as displayed in Fig. 4. Finally the recognition process gave the most matching recognized face from training database as shown in Fig. 5.

Fig. 3. Eigenfaces

Fig. 4. Fisherfaces

Fig. 5. Recognized face by PCA and LDA based algorithms
Then the performance curve of PCA and LDA was plotted as given in Fig. 6. The graph (Fig. 6) clearly indicates that it is not LDA always outperforms PCA, but sometimes PCA also outperforms LDA with increasing dimensionality.

![Performance curves of PCA and LDA](image)

**Fig. 6.** Performance curves of PCA and LDA

In this study, we have applied two most popular appearance based face recognition methods i.e. PCA and LDA along with image preprocessing factors on face94 database containing 3000 images. The Euclidean Distance based classifiers were used for both methods of face recognition systems. The results obtained shows that LDA always may not outperform PCA also result shows that sometimes PCA also gives better recognition rate than LDA with increasing dimensionality. As shown in Fig. 6, therefore, it may be concluded that for some value of dimensionality, PCA gives higher recognition rate than LDA. In future we want to extend our work by using another approach of face recognition that is Independent component analysis and compare all algorithms (PCA, LDA, and ICA) for their performances. The work further may be extended for occluded image databases.

**REFERENCES**


