Face Recognition using PCA Algorithm

Khushboo Dagar[1] Dr. Sat Pal Singh[2]
[1,2]Baba Mastnath University, Haryana, India

Abstract— Principal Components Analysis (PCA) is a basic method widely used in face feature extraction and recognition. In this paper, A Face recognition method using eigenface is proposed and also we have attempted a study of principal component analysis using Euclidean Distance. Experimental results with FAR and FRR are shown that the performance of PCA is good.

Key words: PCA, Face Recognition

I. Introduction

Face recognition has been a fast growing, challenging and interesting area in real-time applications. Lots of achievements have been obtained in face recognition .The human face is complex and dynamic structure with characteristics that can quickly change with time. Face recognition is applied in many important areas such as security systems, identification of criminals, and verification of credit cards and so on. Many face features make development of facial recognition systems difficult [6]. The task of recognition of human faces is sometimes complex. The human face is full of information but working with all the information is time consuming and less efficient. It is better to get unique and important information and discards other useless information to make system efficient. [5]. the term face recognition refers to identifying, by computational algorithms. This operation can be done by means of comparisons between the unknown face and faces stored in a database [3]. Principal component Analysis is usually used because of its acceptable performance [3].

Principal Components Analysis is a basic method widely used in face feature extraction and recognition. A set of eigenface can be generated by performing a mathematical process called principal component analysis (PCA) on a large set of images depicting different human faces. The purpose of PCA is to reduce the large dimensionality of the data space to the smaller dimensionality of feature space.

A. Face Recognition

Face is the index of mind. It is a complex structure and needs a good computing technique for recognition. Face recognition system are easily confused by changes in illumination, variation in poses and change in angles of faces. Many techniques are being used for security and authentication purposes which includes areas in detective agencies and military purpose[8].

PCA is designed to transform a high-dimensional image into a low-dimensional one. It is based on statistical features, and eigenvectors of covariance matrix from a set of face image samples are used to represent the whole face features approximately [7].The Eigen faces approach is now largely superseded in the face recognition literature. A set of eigenface can be generated by performing a mathematical process called principal component analysis (PCA) on a large set of images depicting different human faces. Eigen faces can be considered a set of "standardized face ingredients", derived from statistical analysis of many pictures of faces. Any human face can be considered to be a combination of these standards faces[6].

There are two phases in the identification phase: training and testing. In the training phase the eigenvalues and eigenvectors of the training set are extracted and the eigenvectors are chosen based on the top eigenvalues. Training set is a set of clean images without any duplicates. In the testing phase the algorithm is provided a set of known faces and a set of unknown faces as the probe set. The algorithm matches each probe to its possibly identity in the gallery.

II. PCA in Face Recognition

PCA was invented in 1901 by Karl Pearson. Principal component analysis (PCA) is a classic technique used for compressing higher dimensional data sets to lower dimensional ones for data analysis, visualization, feature extraction, or data compression. The Principal Component Analysis (PCA) is one of the most successful techniques that have been used in image recognition and compression. PCA reduce the large dimensionality of the data space to the smaller intrinsic dimensionality of feature space which is needed to describe the data economically. The PCA is a technique that allows, from extracting the eigenvectors and eigenvalues of the covariance matrix of images, to create a space of reduced images, which contains the “main components” of the images for subsequent recognition. Although PCA is widespread adopted in the recognition of face images[5].

PCA is the concept of Eigenvalues and Eigen faces and is a method of transforming a number of correlated variables into a smaller number of uncorrelated variables. . PCA is applied to the task of face recognition by converting the pixels of an image into a number of eigenface feature vectors, which can then be compared to measure the similarity of two face images. PCA involves the calculation of the eigenvalue decomposition of a data covariance matrix or singular value decomposition of a data matrix, usually after mean centering the data for each attribute[2]. After feature extraction, the image is represented as a feature vector. The Euclidean Distance is used similarity measures to measure the similarity of test and training images feature.

A. The PCA Algorithm

Suppose Γ is an N^2x1 vector, corresponding to an NxN face image I [1].

1) Step 1: obtain face images I1, I2, ..., IM (training faces)
2) Step 2: represent every image Ii as a vector Γi
3) Step 3: compute the average face vector Ψ:

$$\Psi = \frac{1}{M}\sum_{i=1}^{M}\Gamma_i$$

4) Step 4: subtract the mean face:

$$\Phi i = \Gamma i - \Psi$$

5) Step 5: compute the covariance matrix C:

$$C = \frac{1}{M}\sum_{n=1}^{M}\Phi_n\Phi_n^T = AA^T \quad (N^2 \text{x} N^2 \text{ matrix})$$

where A = [Φ1 Φ2 . . . Φ*M*] (N^2x*M* matrix)

6) Step 6: compute the eigenvectors ui of AA^T

(The matrix AA^T is very large)

6.1) Step 6.1: consider the matrix A^TA (MxM matrix)

6.2) Step 6.2: compute the eigenvectors vi of A^TA

$$A^TA\, v_i = \mu_i v_i$$

$$\Rightarrow AA^T A v_i = \mu_i A v_i$$

$$\Rightarrow CAv_i = \mu_i Av_i \quad \text{or} \quad Cu_i = \mu_i u_i \quad , \quad \text{where } u_i = Av_i$$

Thus, $AA^T$and A^TA have the same eigenvalues and their eigenvectors are related as follows: ui = Avi .

6.3) Step 6.3: compute the M best eigenvectors of AA^T : ui = Avi

7) Step 7: keep only K eigenvectors (corresponding to the K largest eigenvalues)

Each face (minus the mean) Φi in the training set can be represented as a linear combination of the best K eigenvectors:

$$\hat{\Phi}_i - mean = \sum_{j=1}^{K} w_j u_j, \quad (w_j = u_j^T\Phi_i)$$

Each normalized training face Φi is represented in this basis by a vector:

$$\Omega_i = \begin{bmatrix} w_1^i \\ w_2^i \\ \dots \\ w_K^i \end{bmatrix}, \quad i = 1, 2, \dots, M$$

Image recognition using Euclidean distance in PCA:

Given an unknown face image Γ , follow these steps:

1) Step 1: normalize Γ: Φ = Γ −Ψ

2) Step 2: project on the eigenspace:

$$\hat{\Phi} = \sum_{i=1}^{K} w_i u_i \quad (w_i = u_i^T\Phi)$$

3) Step 3: represent □ as:

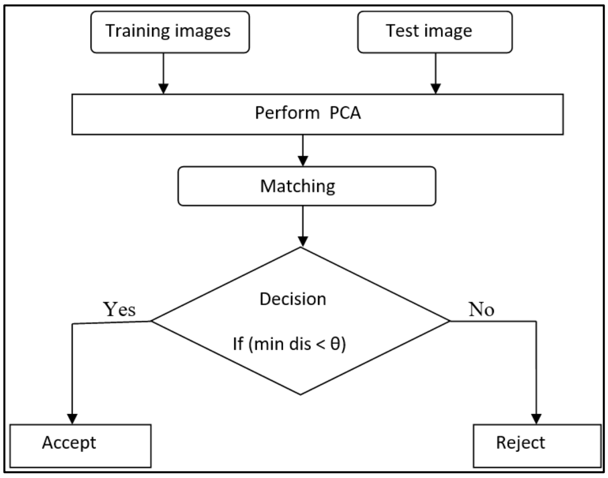
$$\Omega_i = \begin{bmatrix} w_1^i \\ w_2^i \\ \dots \\ w_K^i \end{bmatrix}, \quad i = 1, 2, \dots, M$$

4) Step 4: find $e_r = min^l \,||\Omega - \Omega^l||$

5) Step 5: if $e_r < Tr$, then Γ is recognized as face *l* from the training set

The distance e_r is called distance within the face space. We use the common Euclidean distance to compute e_r. If the minimum distance between test face and training faces is higher than a threshold θ, the test face is considered to unknown, otherwise it is known. It is necessary to set a threshold value that will allow us to determine whether a person is in the database or not.

B. Flowchart of Proposed System



III. EXPERIMENT / RESULTS

The Frontal face dataset which is Collected by Markus Weber at California Institute of Technology, 450 face images with resolution 896 x 592 pixels in the Jpeg format, Belongings to 27 or so unique people under with different lighting/ expressions/ backgrounds. In our experimental result, we have used 5 individuals with different 5 face images.

The training database contains 120 images of 15 persons (8 images per each person), a test database has 40 images of different individuals (30 known and 10 unknown). By using PCA, we have reduced the dimensions from (896 x 592) to (100 x 100). First largest 30 eigenvalues are selected for 30 Eigen faces instead of 120 eigenface. The Euclidean distance is used to find out the distance between the test image and training images. The training image which has the minimum distance with the test image is supposed to be an individual with perfect match. The proposed MATLAB program works in jpeg format images.

The average face of 190 training images is shown in Fig1.

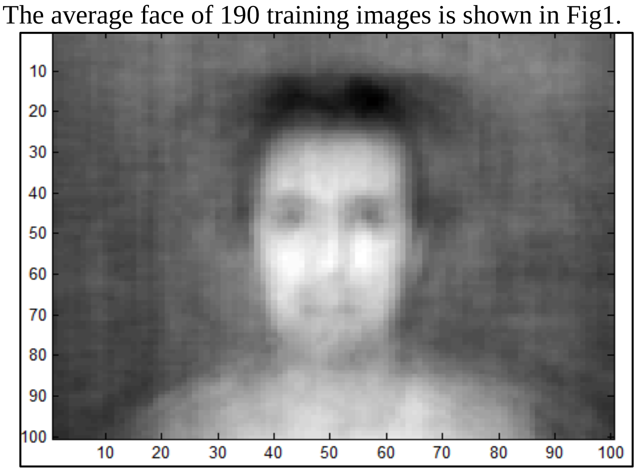


Fig. 1: Shows the Average Face of Training Database

Each eigenvalue corresponds to a single eigenvector. We do not have to consider eigenvectors that correspond to small eigenvalues because they do not carry important information about the image [6].Therefore, we considered only 30 eigenvalues instead of 120. We discard the small eigenvalues.

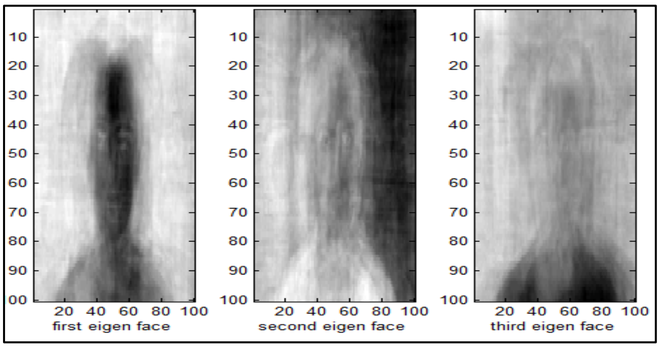


Fig. 2: Shows The First Three Eigen faces With High Values

The first three high eigenvalues of eigenvectors are

1) First eigenvalue 148.150043009616
2) Second eigenvalue 52.3848641925491
3) Third eigenvalue 40.9971295622560

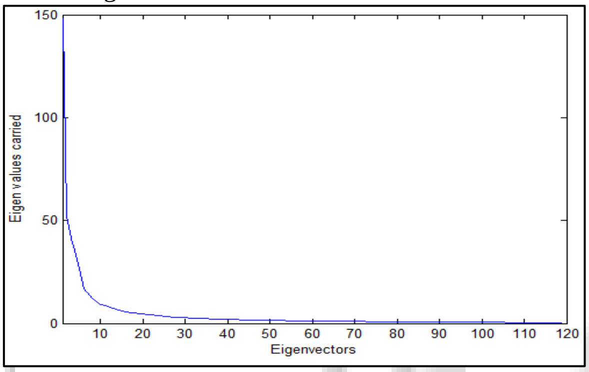


Fig. 3: Shows the Eigenvalues Corresponding the Eigenvectors

The figure clearly shows that 1 eigenvector contains 30 percent of basic features and 10 eigenvectors contain 60 percent. We have chosen 30 eigenvalues for our work because 30 eigenvectors contains 80 percent of basic features that is more enough.

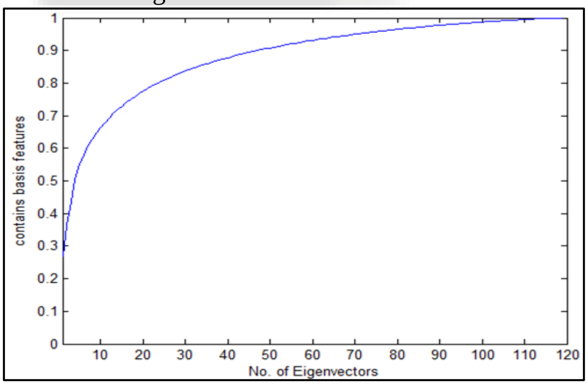


Fig. 4: Shows The Eigenvectors Contains Basic Features.

To recognize the face, we calculate the distance of test image from each image from the training database. It is obvious that if the minimum distance between the test image and other images is zero, the test image entirely matches the image from the training database. If the distance is greater than zero but less than a certain threshold, it is a known person with other facial expression, otherwise it is an unknown person [6].

The test was done for the person who was present in the database but has a different facial expression. The test image has a minimum Euclidean distance which is less than the threshold (θ) (Fig5), so it is a known face. Minimum distance shows us which image from the database matches the test image best. Fig6 shows the Euclidean distance between the test image and all 120 images from database.

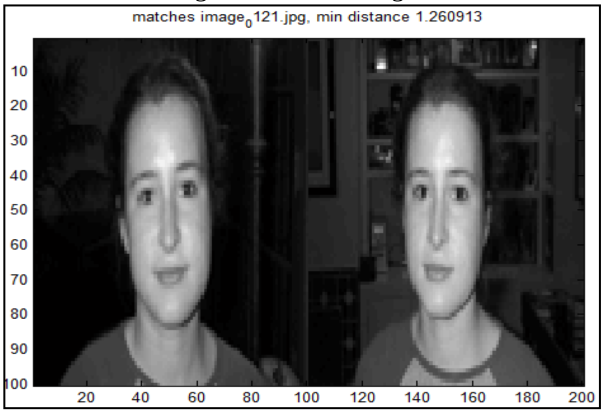


Fig. 5: Test Image & Recognized Image from the Training Base

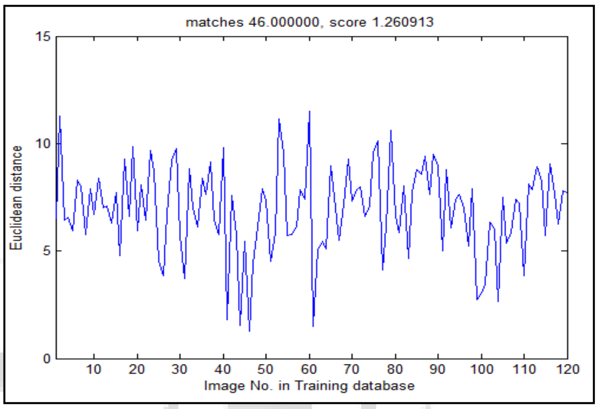


Fig. 6: Euclidean Distance between the Test Image & Images From The Database

During face recognition, a threshold value was used to control the accuracy and was accepted from the user as input value. In order to evaluate the overall performance of the biometric system, the False Acceptance Rate (FAR), False Recognition Rate (FRR), Accuracy and Recognition Time were used. From these two important measures, the accuracy of the system can be calculated using the formula [9].

$$\text{Accuracy} = 100 - (\text{FAR} + \text{FRR})/2)$$

$$\text{FAR} = \frac{\textbf{number of false Acceptance}}{\textbf{total number of Imposter Attempts}}$$

$$\text{FRR} = \frac{\textbf{number of false Rejection}}{\textbf{total number of Genuine Attempts}}$$

Table 1. Shows the result of FAR and FRR obtained for a threshold value specified.

Method	FAR	FRR	Accuracy (%)
PCA	0.30	0.26	99.72

Table 1: Accuracy Analysis

IV. CONCLUSION

In this paper, Face recognition method using Eigen faces is proposed. We used database of face images which contains 120 images of 15 different persons (8 images per person). From the results, it can be concluded that for accurate recognition, it is sufficient to take about 25% Eigen faces with the highest eigenvalues. It is also clear that the recognition rate increases with the number of training images per person. If the minimum distance between the test image and other images is zero, the test image entirely matches the image from the training database. If the distance is greater than zero but less than a certain threshold, it is a known person with other facial expression, otherwise it is an unknown person.

REFERENCES

[1] M. Turk and A. Pentland, "Eigenfaces for Recognition", Journal of Cognitive Neuroscience, vol.3, no. 1, pp. 71-86, 1991, hard copy.
[2] “Eigenfaces face recognition (MATLAB)”, 2 December 2010 based on M. Turk and A. Pentland, seminal work.
[3] Dr.H.B.Kekre, Sudeep D. Thepade, Akshay Maloo, “Eigenvectors of Covariance Matrix using Row Mean and Column Mean Sequences for Face Recognition”, International Journal of Biometrics and Bioinformatics (IJBB), Volume (4): Issue (2).
[4] Turk, M.A. and A.L. Pentland, 1991, “Face recognition using Eigen faces”, Proc. IEEE computer society Conf Computer Vision and pattern recognition, pp: 586-591.
[5] Taranpreet Singh Ruprah ,”Face Recognition Based on PCA Algorithm”, Special Issue of International Journal of Computer Science & Informatics (IJCSI), ISSN (PRINT) : 2231–5292, Vol.- II, Issue-1, 2.
[6] Marijeta Slavkovic , Dubravka Jevtic, “Face Recognition Using Eigenface Approach”, SERBIAN JOURNAL OF ELECTRICAL ENGINEERING, Vol. 9, No. 1, February 2012, 121-130.
[7] Jun-ying Gan, Dang-pei Zhou, Chun-zhi Li, “A Method for Improved PCA in Face Recognition”, International Journal of Information Technology, Vol. 11 No. 11 2005.
[8] V.Vijayakumari, “Face Recognition Techniques: A Survey”, World Journal of Computer Application and Technology 1(2): 41-50, 2013.
[9] Dr. V. Radha et al., “Comparison of PCA Based and 2DPCA Based Face Recognition” , Journal of Engineering Science and Technology Vol. 2 (12), 2010, 7177-7182