IJESRT

INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY

Analyzing Face Recognition Using Pca and Comparison between Different Distance Classifier

Bhaskar Gupta[*1], Anil Kumar Singh[2]

[*1,2] Associate Professor, Electronics Engineering Department, BBDIT, Ghaziabad, UP, India

bhaskar.g@rediffmail.com

Abstract

Principal component analysis (PCA) is a widely used technique that is quite efficient and reliable when used for face recognition. Face is the most dominant feature that strongly communicates identity of the person. It proves to be very useful and secure if used for biometric identification. Principal component analysis uses Eigen Face approach for extracting effective features (Eigen vectors) from a given face database and these features corresponds to the dissimilarities among the faces. Every face in the database can be represented as a linear combination of these eigenvectors with appropriate weight associated with each of the eigenvector. This paper presents a methodology for face recognition using Principal Component Analysis algorithm and compares two different distance classifiers i.e. Euclidean distance and Manhattan distance on the basis of recognition rate obtained by varying the number of Eigen faces.

Keywords : Face recognition, Eigen Face, Principal Component Analysis, Euclidean Distance, Manhattan Distance.

Introduction

In the recent years, human face recognition, a technique that detects and identifies human faces is gaining importance in the field of biometrics. The human face is a highly intricate and dynamic structure with characteristics that can adversely change with time but it is also the feature that best distinguishes a person. Humans can recognize thousands of faces learned throughout their lifetime and identify familiar faces at a glance even after years of separation [2]. This skill is quite robust, despite large changes in the visual stimulus due to viewing conditions, expression, aging, and distractions such as glasses or changes in hairstyle. Computational models of face recognition, in particular, are interesting because a computer that can recognize faces could contribute to a wide variety of problems, including criminal identification, security systems, image and film processing, and human computer interaction. Thus these models add not only to theoretical insights but also to practical applications. Developing such a computational model is a quite difficult task.

The Eigen face based PCA approach used in this paper does not focus on detecting individual features like eyes, mouth, nose and head outline and developing a face model by determining the position, size, and relationship among these features (Feature Based Method) but it uses entire raw face image as an input and extracts features that are different among these faces (Holistic method).2. P

PCA Algorithm

Principal Component Analysis (PCA) is a dimensionality reduction technique in which multidimensional data is reduced by extracting the desired number of principal components. Without the need of intense effort PCA provides us with an outline that describes how to reduce a complex data set to a lower dimension to reveal certain unknown useful and simple information. This is the case when there is a strong correlation between observed variables. This strong correlation is present in human faces because every face has various features in common like two eyes, one nose, one mouth etc. at defined positions. By capturing the variation in a set of faces, independent of any judgment of features it is possible to extract the relevant information in a face image which is encoded as efficiently as possible, and compared with a database of encodings developed similarly for every face [1]. PCA approach transforms face images into a small set of relevant features images called Eigen faces which are ghostly images represented by eigenvectors that are the principle components of the faces and characterize the variation between them. Each individual face can be represented exactly in terms of a linear combination of the Eigen faces. Also not all but only few relevant eigenvectors with largest eigenvalues are sufficient to represent each face image. This method thus reduces the dimensionality of data space by projecting data from M-dimensional space to P-dimensional space, where P<<M [5].

After the dimensional reduction of the face space, the distance is measured between two images for recognition and if the distance is less than some threshold value, then it is considered as a known face, else it is an unknown face [6].

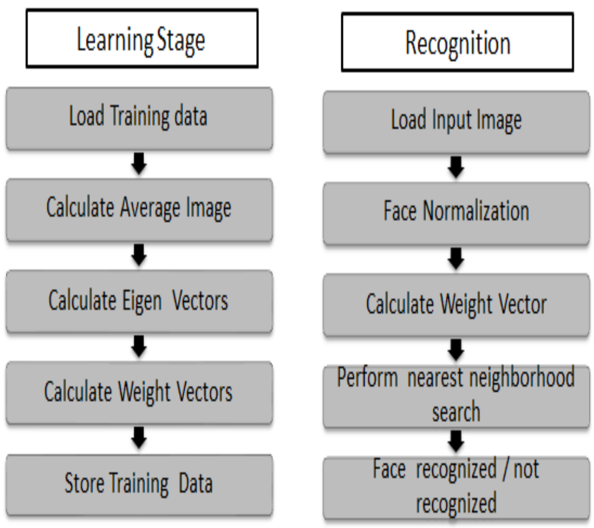

Fig 1: brief steps for Eigen face approach

Mathematical Procedure fOR PCA

1. The first step is to obtain a set S with M face images. We have taken M = 45(fig.2). Each image is transformed into a vector $\Gamma$and placed into the set.

$$S = \{\Gamma_1, \Gamma_2, \Gamma_3, \dots\dots, \Gamma_M\} (1)$$

2. Calculate mean image $\boldsymbol{\Psi}$(fig 3)

$$\psi = \frac{1}{M}\sum_{i=1}^{M}\Gamma_{\mathrm{i}} \ (2)$$

3. Then we have to find the difference $\boldsymbol{\Phi}$ between the input image and the mean image

$$\phi_i = \Gamma_{\mathrm{i}} - \psi (3)$$

4. Next we seek a set of M orthonormal vectors, $\boldsymbol{u_n}$,which best describes the distribution of the data. The kth vector,$\boldsymbol{u_k}$**,** is chosen such that

$$\lambda_k = \frac{1}{M}\sum_{n=1}^{M}(u_k^T\phi_n)^2 \ (4)$$

is maximum, subject to

$$u_l u_k = \delta l_k = \begin{cases} 1, if\ l = k \\ 0, otherwise \end{cases} (5)$$

Note: u_k and$\lambda_k$are the eigenvectors and eigenvalues of the covariance matrix **C**

5. We obtain the covariance matrix **C** in the following manner

$$C = \frac{1}{M}\sum \phi_n\phi_n^T$$

$$= AA^T (6)$$

Where A = $[\phi_1\phi_2\phi_3 \dots . \phi_M]$.

6. L = A^TA, where

$$L_{mn} = \phi_m^T\phi_n (7)$$

7. Once we have found the eigenvectors (fig.4),$\boldsymbol{v_l}$,$\boldsymbol{u_l}$

$$u_l = \sum_{k=1}^{M} V_{lk}\phi_k; \quad \text{l=1,…..,M} (8)$$

Recognition Procedure

1. A new face is transformed into its Eigen face components. First we compare our input image with our mean image and multiply their difference with each eigenvector of the L matrix. Each value would represent a weight and would be saved on a vector Ω.

$$\omega = u_k^T(\Gamma - \psi) (9)$$

for k = 1,…..,M'.

$$\Omega^T = [\omega_1, \omega_2, \dots, \omega_{M'}] (10)$$

2. We now determine which face class provides the best description for the input image. Here we have used two different distance classifiers i.e. Euclidean distance and Manhattan distance classifier.

Euclidean distance:

$$d(x,y) = \|x - y\|^2 = \sum_{i=1}^{k}(x_i - y_i)^2 \qquad (11)$$

Manhattan Distance:

$$d(x,y) = |x - y| = \sum_{i=1}^{k}|x_i - y_i| \ (12)$$

3. The input face is considered to belong to a class if $\varepsilon_k$is below an established thresholdθ_ε, then the face image is considered to be a known face. If the difference is above the given threshold, but below a second threshold, the image can be determined as unknown face. If the input image is above these two thresholds, the image is determined NOT to be a face.

 ISSN: 2277-9655

4. If the image is found to be an unknown face, then we can add that image to our training set for future recognitions. We have to repeat steps 1 through 7 to incorporate this new face image.

Experimental Results

Training Database

While there are many databases in use currently, the database in this paper is taken from *Face Recognition Data, University of Essex,UK* [4]. There are 45 faces in the database. There are faces of 15 people with 3 images of each person with different expressions.

Test database

In the test database there are 30 images in which some are of the known individuals in the database but with different expression and others are of unknown individuals and few are images that do not contain a face at all.

Threshold

A threshold is chosen such that if

$e_r < \theta$; Known face

$e_r > \theta$; Unknown face

Threshold value plays an important role in selecting a known face. If there will be no threshold then it might be possible that a test image with no face at all could get recognized as a known face (its distance value can be minimum from an image in database).

Therefore we have to set a threshold heuristically for our image databases [3].

We have chosen the threshold for our database empirically. By using trial and error method the threshold is chosen which depends on various factors like number of images in database, number of images per person, features of database etc. In this paper the threshold value chosen is 5.5×10^3 for Euclidean distance and $8.5 \times 10^3$for Manhattan distance classifier.

Results

To study the effect of different distance classifiers (like Euclidean distance and Manhattan distance) and the effect of changing the number of Eigen faces on face recognition we have performed experiments on the database using MATLAB.

Database contains 45 images (3 images of each of the 15 individuals) and test database contains 30 images of known and unknown individuals. Dimension of each image of the database is 200x180x3 and in vector form the size is 200x200.

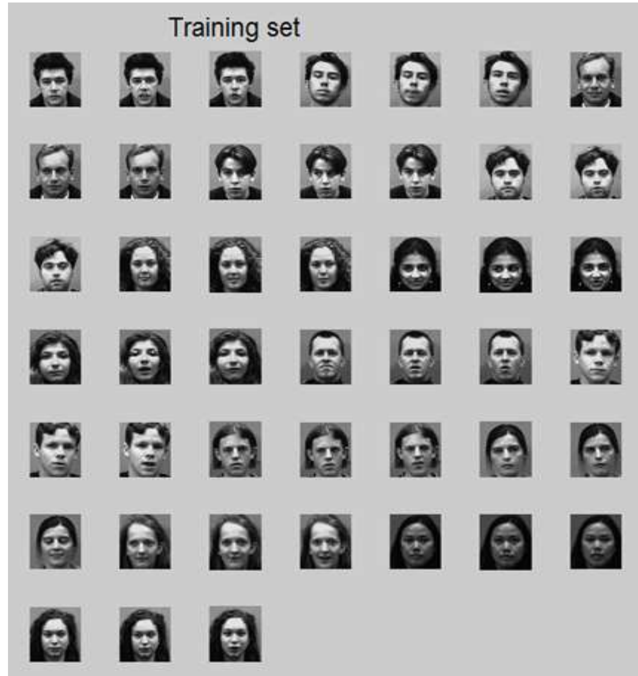

Fig. 2 Training database

Fig. 3 Average Fac

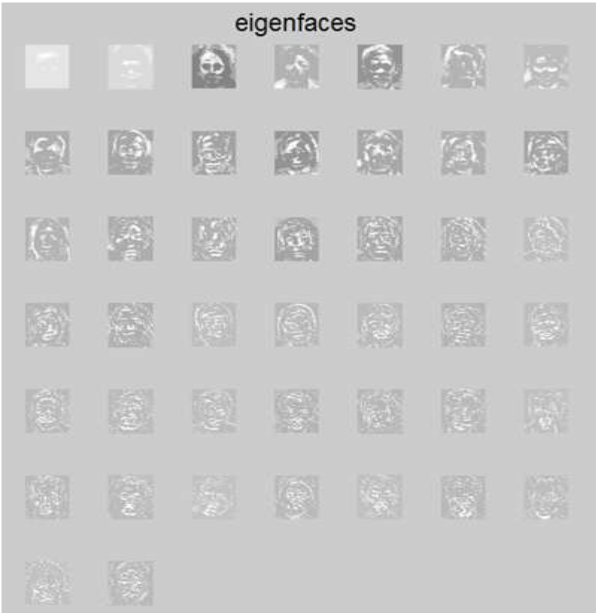

Fig. 4 Eigen faces

Parameters considered for comparison between Euclidean distance classifier and Manhattan distance classifier are recognition rate and number of Eigen faces required for maximum recognition rate.
Recognition rate reflects the percentage of faces recognized correctly as known or unknown when test database faces are evaluated. It is desirable to have maximum recognition rate by using less number of Eigen faces because it clearly makes the procedure simple and fast.

TABLE I.RESULT OF FACE RECOGNITION

Number of principle components(Eigen face)	Recognition Rate	
	Euclidean Distance	Manhattan Distance
5	66.66%	73.33%
15	87%	94%
20	94%	94%
30	94%	94%

Conclusion

Face recognition method using PCA is proposed and comparison between Euclidean distance and Manhattan distance classifier is done. We have used the database containing 45 images of 15 different persons (3 images of each person with different expressions).

From the results, it can be concluded that, a maximum recognition rate of 94% can be achieved using either Euclidean or Manhattan distance. For recognition using Euclidean distance classifier, it is required to take 40% of Eigen faces with highest eigenvalues but for Manhattan distance classifier around 30% of the Eigen faces are sufficient. It is observed that using Manhattan distance classifier less number of Eigen faces are required for achieving the same recognition rate as compared to Euclidean distance classifier.

References

[1] Matthew A. Turk and Alex P. Pentland. "*Face recognition using eigenfaces*". Proc.Of IEEE Conf. on Computer Vision and Pattern Recognition, pp. 586-591, June 1991

[2] SezinKaymak"*Face Detection,Recognition and Reconstruction using Eigenfaces''*,Technical Report, Eastern Mediterranean University,Turkey, spring 2003

[3] ShubhenduTrivedi"*Face Recognition using Eigen faces and Distance Classifiers:ATutorial"*,Feb 2009, http://onionesquereality.wordpress.com/2009/02/11/face-recognition-using-eigenfaces-and-distance-classifiers-a-tutorial/

[4] Face Recognition Data,University of Essex,UK(http://cswww.essex.ac.uk/mv/allfaces/index.html)

[5] MarijetaSlavkovic, DubravkaJevtic*"Face Recognition Using Eigenface Approach"*,Serbian Journal of Electronics Engineering, Vol 9,No 1,pp 121-130, Feb 2012.

[6] Sheifali Gupta, O.P. Sahoo, Ajay Goel, Rupesh*Gupta "A New Optimized Approach to Face Recognition Using EigenFaces"*, Global Journal of Computer Science and Technology, Vol 10,April 2010.

[7] Eigenface Tutorial,http://www.pages.drexel.edu