Face Recognition Rate Using Different Classifier Methods Based on PCA

Dr. Eyad I. Abbas
Department of Electrical Engineering,
University of Technology
Baghdad, Iraq
30013@uotechnolog.edu.iq

Mohammed E. Safi- MIEEE
Department of Electrical Engineering,
University of Technology
Baghdad, Iraq
30165@uotechnolog.edu.iq

Dr. Khalida S. Rijab
Department of Electrical Engineering,
University of Technology
Baghdad, Iraq
30049@uotechnolog.edu.iq

***Abstract*— this paper describes the different classifier methods with minimum means of clusters to achieve face recognition rate of humans from the feature extracted of training face image data for many sets of images as a data base. Principal Component Analysis (PCA) is a robust method used as feature extraction techniques for face recognition but the recognition decreases with the variation of person's actions. The features extracted for face images are light insensitive, individual, hidden, and activity effective to biometric recognition. The face recognition treats as two dimensions recognition problems, the fact is to take the advantage of these human faces are straight pose in general may be represented as a small set of two dimension characteristics view. The training and testing face images are selected from Research Laboratory for Olivetti and Oracle (ORL) face database, which have minimum pose variation. Three classifier methods are used to obtain the distance of recognition. These classifiers are: the Euclidian distance method, the Squared Euclidian Distance method, and the City-Block Distance method. By Clustering the difference of training image with images set for each person and determined the mean to it, the minimum mean is representing the recognition of the person. The cluster method with Squared Euclidian Distance method produces higher a recognition rate 100% near the Euclidian Distance method which gives a human face recognition rate 98% higher than the City-Block Distance method which gives a recognition rate 95%.**

Keywords—:Human Face Recognition; Distance Methods; PCA

I. INTRODUCTION

The face recognition for human is one of the most important issues inbiometric sciences based on personalidentification [1]. In the last few years, many different successful applications of image analysis, has received significant development [2]. Biometrics are a recognition methods based on person behavioral characteristics or unique physical. The big advantage in human face recognition using in biometric authentication, has arisen from many reasons. The first reason is in the concern of public security like facilities and places such as airports, parks, and sport fields. The second reason is theinternet network challenges like shopping based on electronic commerce. And the third reason is using of the human face recognition in biometric identification is one of the high accuracy level by using least intrusive methods. The Principal component analysis (PCA) is a one of the powerful statistical technique used to analyze the data in classes. The basic form of the PCA is to minimize the dimensionality of the data set that consist of a large number of the correlated variables, while maintain as possible of the variation present in the set of data. The mathematical formula of PCA is clinging in the standard deviation, and the eigenvectors, and the eigenvalues [3]. The final process of facial recognition is distance measures. Different methods used to estimate the distance between the test image and data base of images. The classification satisfied with the minimum distance between the training faces and the test image. Many of approaches are introduced to optimize the distance of image recognition.

J. Jia, Y. Xu, S. Zhang, and X. Xue, 2016[4] proposed a new framework for facial expression classification. Improved PCA to extract the features and uses the random forest algorithm as the classifier. It is possible to distinguish more expressions, such as smiling and laughing in happy, amazing and shocking in surprise.

E. I. Abbas, 2015[5] This paper presented an approach to study the effect of the different eigenfaces levels on the faces recognition rate using principal component analysis. The experiment was conducted on 50 images from the database of faces (ORL), using 40 images for the training set and 15 images for the test group. The results proved that the proposed method is effective and successful in obtaining recognition rate up to 100% in the third level when using ten eigenfaces.

A. Alkandari, S. J. Aljaber, 2015[6], introduced to recognize, identify the facial image without human intervention. PCA is a useful used for reduce dimension vector to better recognize images. The outhers recommended PCA for image recognition without mention recognition rate that on the results.

M. P. Gawande, and D. G. Agrawal, 2014[7] introduce face recognition system using Principal Component Analysis with

978-1-5386-2955-0/17/$31.00 ©2017 IEEE

II. THE PCAM ETHOLOGY

The PCA algorithm isreduce effectively the dimensions of the human face images, also keep its status of identifying information [8]. The mathematicalform of PCA takes three primary steps.The first step is creating a transformation matrix using the training of the human face images. The second step, the training face images are arrangedinto the matrix vectors. Ultimately, the test face image is recognize by introducing it in the subspace of eigenfaces and compared with the trained face images in the subspace of eigenfacedomain.The PCA technique can be shown in the following aspects:

a- Let each facial image define by $X(x,y)$ be a two dimension (m×n) is interpreted as a one dimension (mn). Then the training set of face imagesis collected as $\{X_1, X_2, \dots\dots\dots, X_N\}$. The average ofthis training set is calculated by:

$$\overline{X} = \frac{1}{N}\sum_{i=1}^{N} X_i \ \dots\dots\dots (1)$$

b- Form the convariance matrix as:

$$C = \frac{1}{N}\sum_{i=1}^{N}(X_i - \overline{X})(X_i - \overline{X})^T \dots\dots\dots (2)$$

c- Compute the eigenvectors corresponding to eigenvalues by using:

$$CV = \lambda V \dots\dots\dots (3)$$

Where: V is the set of the eigenvectors, and λ is the eigenvalue.

d- Descending sort of the order eigenvectors corresponding to the eigenvalues.

e- Each of mean values is centering image project into eigenspace using:

$$W_i = V_i^T(X_i - \overline{X}) \dots\dots (4)$$

f- The test image should be mean centered as the training phase into the same eigenspace.

g- The test of face image is now compared withthe data baseof the face images in eigen domain.

h- Estimate the similarity measures using one of the distance methods as classifier. The recognition will be if the training face image with the minimum distanceobtain to the test face image can used to identify.

III. THE DISTANCE CLASSIFIER METHODS:

The distance measure is the final and most important step of face recognition. Images of human faces are converted into an eigen domain and sorted as matrices. The distance between the matrices of two face images is determined how similar the test face image with trained image. Three methods of distance are used in this workforClassification. These methods are:

1- Euclidean Distance method: is the widly method used in distance. The Euclidean distance equation is:

$$d(x,y) = \sqrt[2]{\sum_{i=1}^{l}(x_i - y_i)^2} \dots\dots\dots\dots (5)$$

Where x, y in the data set X and x_i, y_i are the i^{th} coordinates of xand y, respectively.

2- The Squared Euclidean Distance method: The Squared Euclidean Distance is a like Euclidean Distance Classifier without the square root:

$$d(x,y) = \sum_{i=1}^{l}(x_i - y_i)^2 \ \dots\dots\dots\dots (6)$$

3- City-Block Distance method: also called, rectilinear distance. It interprets the distance between spots on grid space and take absolute to the result as follows:

$$d(x,y) = \sum_{i=1}^{l}|x_i - y_i| \ \dots\dots\dots\dots (7)$$

IV. THE ORL FACEDATABASE

The directory of ORL contains a set of faces for 40 distinct persons of 10 different images. The some of persons appear in face image taken at varying lighting slightly, different time, facial expression in eyes, simling, oraintation. As well all images are taken in dark homogeneous background. The image files are stored as PGM format with size of each image is (112×92), 8 bits, and gray levels[9]. Fig.(1) shows example for 10 individuals 10 images per personal from the ORL database for human face images using in training and testing.

V. THE SYSTEM SOFTWARE ALGORITHM

The system algorithm for the face recognition in the spatial domain using PCA is shown in the flowchart Fig. (2). The training images set are converted into a set of eigenfaces, then the weights of trained images are determined and stored in the set W. For an unknown image X, the weightsare predicted from that image and stored in the vector Wx. the two vectors W and Wx are compared using distance classifications methods.

Fig. 1. The ORL database for training and testing

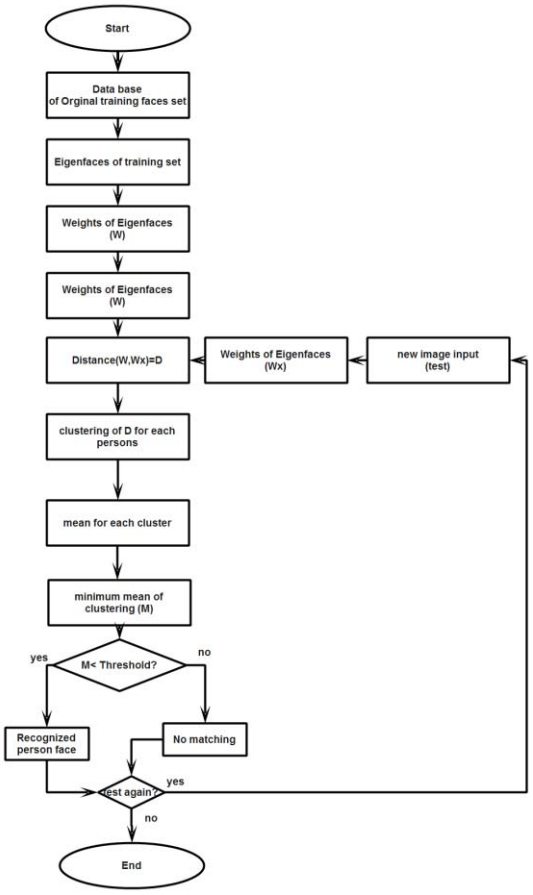

Fig. 2. Flowchart of SystemAlgorithm

VI. EXPERIMENTAL RESULTS:

The training set and new, unknown face images are picked out carefully from the ORL database in order to reducevariation in poses. This database consists of 10 classes, each class represents one person,each class, consist of 10 face images as shown in Fig.(1). All the face images enter into training and testing. The image format has been transformed to Joint Photographic Experts Group (JPEG) format because it's easy to process. The normalize of the training face images is used to minmize errors caused bylighting conditions and background. In Fig.(3) shows the eigenfaces for training faces sorting as descending form of eigenvalues. The distanceshows the similarity of the input image to the images on our training set and the maximum for three methods. Based on the averge of persons clusterswith these methods the decision of recognition is configure. Fig (4) illustrates the samples of weight for three test images and the minimum values for each distance method, while Table(1) content a part of minimum and maximum and difference of values obtains from experiments to show the recognition rate for each method. The table (2) is show illustrate the overall result of three distance methods applied to get the recognition rate for each case. It is clear the Euclidean Distance Classifier is powerful and effective with recognition rate 100%, follow it second The Squared Euclidean Distance Classifier with recognition rate98%, And finally become City-BlockDistance Classifier with recognition rate 95%.

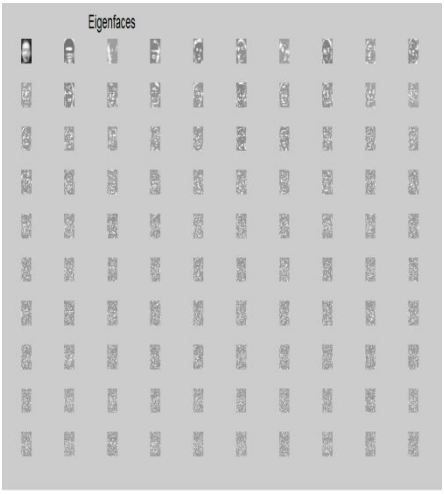

Fig. 3. Eigenface ranked

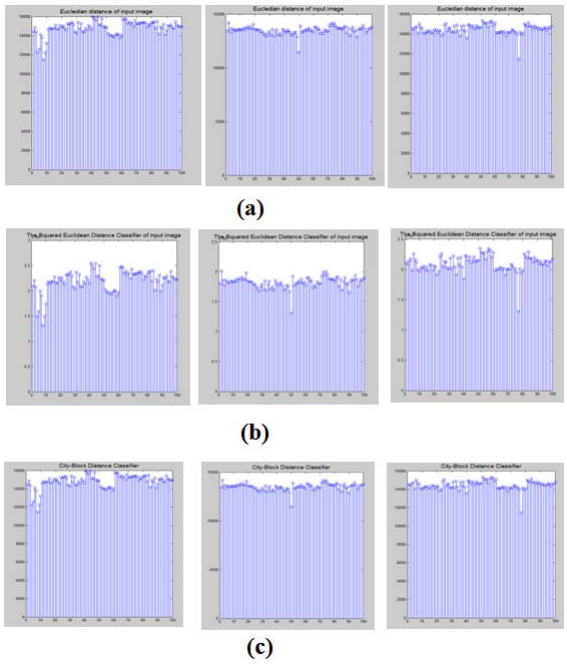

Fig. 4. Samples of weights for three test images: a-Enclidean b- Squared. c-City-Block

Table I. Minimum and Maximum and difference values for three distance methods

	A	B	C	D	E	F	G	H	I	J	K	L
1	euclidean distance				square					city		
2	max	min	difference		max	min	difference			max	min	difference
3	14430.13	11471.33	2958.801		2.08E+08	1.31E+08	76628700			14422.86	11462.76	2960.099
4	16539.27	11458.93	5080.345		2.74E+08	1.31E+08	142261716			16538.55	11456.94	5081.603
5	14832.86	11469.5	3363.355		2.2E+08	1.31E+08	88812454			14836.07	11458.47	3377.596
6	17005.4	11452.85	5552.547		2.89E+08	1.31E+08	158023216			17005.79	11453.12	5552.671
7	16709.17	11458.6	5250.567		2.79E+08	1.31E+08	147840304			16707.48	11458.6	5248.884
8	15907.41	11460.74	4446.662		2.53E+08	1.31E+08	121538712			15898.79	11456.48	4442.315
9	15547.48	11462.97	4084.507		2.41E+08	1.31E+08	110132103			15537.53	11457.87	4079.657
10	15975.78	11457.69	4518.09		2.55E+08	1.31E+08	123931339			15975.5	11458.15	4517.35
11	16951.82	11460.94	5490.885		2.88E+08	1.31E+08	156200406			16956.22	11459.2	5497.024
12	15672.46	11455.78	4216.681		2.46E+08	1.31E+08	114564487			15676.94	11454.35	4222.59
13	14466.38	11463.96	3002.426		2.09E+08	1.31E+08	77761929			14457.99	11457.37	3000.613
14	14208.88	11461.37	2747.51		2.02E+08	1.31E+08	70670635			14210.17	11456.8	2753.371
15	14386.03	11461.21	2924.825		2.07E+08	1.31E+08	75444085			14375.83	11455.15	2920.685
16	14285.52	11461.5	2824.018		2.04E+08	1.31E+08	72765605			14283.44	11456.48	2826.958
17	14503.97	11461.05	3042.916		2.1E+08	1.31E+08	79126114			14505.79	11457.91	3047.875
18	14671.28	11460.25	3211.028		2.15E+08	1.31E+08	83978792			14671.24	11457.14	3214.094
19	15196.86	11457.38	3739.478		2.31E+08	1.31E+08	99654397			15193.92	11454.28	3739.634
20	14625.05	11459.31	3165.734		2.14E+08	1.31E+08	82644573			14623.76	11454.68	3169.08

TABLE II. RECOGNITION RATE FOR THREE DISTANCE METHODS

Methods with clustring	Recognition rate
Squared Euclidean Distance	100%
Euclidean Distance	98%
City-Block Distance	95%

VII. CONCLUSIONS

The modal biometric system proposed in this work using PCA with different methods for classification using distance difference integrated with minimum mean of difference for clustering images. The PCA algorthimfor Face Recognition process is simple, fast and works well under constrained environment. It is robust feature extraction for face recognition but it must optimize with clustering for person recognition.The face recognition based on PCA was applayed on three distance techniqeswith the ORL database. The test results show that PCA with mean of clustering gave best results with Squared Euclidian distance method with recognition rate 100% than the Euclidian distance method and the City Block distance method, while, the recognition rate using the Euclidian distance is approximately close to Squared Euclidian distance classifier rate, and both techniques is greater than the recognition rate for City-Block distance.

REFERENCES

[1] [1] L. Hong and A. Jain " Integrating faces and fingerprints for personal identification", IEEE Trans. PAMI, Vol 20, pp. 1295-1307, 1998.

[2] [2] X. Xie, and K. Lam, "An efficient illumination normalization method for face recognition". Pattern Recognition Letters 27,609-617. c ELSEVIER (2006).

[3] [3] A. Khan and H. Farooq "Principal Component Analysis-Linear Discriminant Analysis Feature Extractor for Pattern Recognition", IJCSI International Journal of Computer Science Issues, Vol. 8, Issue 6, No 2, November 2011.

[4] [4] J. Jia, Y. Xu, S. Zhang, and X. Xue," The facial expression recognition method of random forest based on improved PCA extracting feature ",International Conference on Signal Processing, Communications and Computing, DOI: 10.1109/ICSPCC.2016.7753643, IEEE, 2016.

[5] [5] E. I. Abbas "Effect of Eigenfaces Level on The Face Recognition Rate Using Principal Component Analysis", Eng. & Tech, Journal, Vol. 33, No. 3, 2015.

[6] [6] A. Alkandari,and S. J. Aljaber, "Principle Component Analysis Algorithm (PCA) for Image Recognition", Second International Conference on Computing Technology and Information Management, DOI: 10.1109/ICCTIM.2015.7224596, IEEE, 2015.

[7] [7] M. P. Gawande, and D. G. Agrawal, " Face recognition using PCA and different distance classifiers", Journal of Electronics and Communication Engineering (IOSR-JECE), ISSN: 2278-8735, Vol. 9, Issue 1, 2014.

[8] [8] C. Han, "Modular PCA Face Recognition Based on Weighted Average". Modern Applied Science, Vol. 3, No. 11, November 2009.

[9] [9] The ORL database, Available: http://www.cam-orl.co.uk/face