ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 4, Issue 2, February 2017

Comparative Study of PCA and LDA Algorithms for Automated Attendance System Using Face Recognition

[1] Ms. Sarika Ashok Sovitkar [2] Dr. Seema S. Kawathekar
[1] Ph.D., Student, [2] Associate Professor
[1][2]Department of CS&IT, Dr, Babasaheb Marathwada University, Aurangbad.

***Abstract*- now a day as lots of research was done in computer vision and image processing from last few decades. It is possible to develop a system for automated attendance using face recognition technique. In this paper we implement the PCA and LDA algorithm for face recognition technique and compare result. We found that in PCA as the number of samples per persons in training dataset does on increase the accuracy also increase for recognition. To store the huge sample dataset is the main drawback of this algorithm and also we found that in real time there are such requirement the will have to recognize a person from only one sample dataset as training dataset. In LDA we extract the components in training phase and compare them with the extracted components of dataset of testing image in testing phase. As we succeed to decrease the number of samples to only two samples for training and the results found are encouraging. The experiments are performed using different variations in lighting, illumination, facial expression, partial occlusion and imprecise localization of face area.**

Keywords: Principle Components Analysis(PCA), Linear Discriminates Analysis (LDA), eigen face, eigen vector. Scatter Matrix.

I. INTRODUCTION

based on effective ID information extracted by computers from facial pictures. With the development of technology, performance of computers is becoming more & more powerful. Also the development of digitalized information requires secure authentication more than ever. Compare with biometric authentication methods like figure prints, iris recognition, palm print recognition, face recognition system is more friendly, convenient & easier for users to accept. This makes face recognition technology a research hotspot.

Automated real time attendance using face recognition marking system can help schools and higher education in many ways. There is no doubt that an attendance management system will help save time and money by eliminating a great deal of manual processes involved in attendance and leave entry and calculating hours attended. With automatic class attendance system, teachers can more accurately and quickly track student's time on the classroom.

Face recognition is a hard problem because there exist many image variations in real-life such as lighting, illumination, facial expression, partial occlusion and imprecise localization of face area. The typical approach in handling these variations is to use large and representative training sample sets. Also there are some difficulties from environment like distance, size, illumination etc. Limitations on computer to process huge data in a short period.

II. A BRIEF REVIEW OF PCA ALGORITHM

The PCA treat an image that has N X M pixels as an vector with a length of NM, which also can be treat as a point in an M X M dimensional space. PCA usually have three functionalities in processing original data 1.Reduce dimension, 2. Remove relevance and 3. Probability estimation

2.1 Feature extraction method

PCA method has been widely used in applications such as face recognition and image compression. PCA is a common technique for finding patterns in data, and expressing the data as eigenvector to highlight the similarities and differences between different data [2]. The following steps summarize the PCA process[11].

[1] Get image data in form of matrix as A1, A2….Am.
[2] Calculate the mean Ø as S Ai /M. i=1
[3] Subtract mean from original data (image data) A1- Ø.

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 4, Issue 2, February 2017

[4] Calculate covariance matrix C= AT A.
[5] Calculate eigen values of the covariance matrix.
[6] Choose components with highest information and form a feature vector.
[7] Derive new data set and project the eigen face image.

2.2 Nearest- Neighbor classification

Recog_wt= P` * A2 | A2 is created from testing data projection testing image to eigen-space.
euDis(i,j) = sqrt ((recog_wt(I,j) – train_wt(I,j)) ^2)

Find the Euclidian distance from each testing face to each known faces. These projections (feature vectors) are then used as classification features in face recognition. [8] In general features extracted from PCA are subjected to distance classifiers. The distance between the features of probe image and features of trained images is calculated. If the distance is less than the threshold (Θ) then the probe image is recognized.

$$e_r = min\|\omega - \omega_i\|$$

where : er is Euclidean distance,
ω is image vector,
and i is number of trained image.

III. A BRIEF REVIEW ON LDA ALGORITHM

LDA-based methods have been proved to be an effective appearance-based approach for developing practical face recognition systems. The objective of LDA algorithm is to find an optimal projection W*, from original sample feature space to a transformed lower dimensional feature space, such that the ratio of the between-class scatter Sb and the within-class scatter Sw is maximum, i.e.

$$W^*: R^d \to R^m, d > m,\ W^* = \arg\max_w \frac{\det(W^T S_b W)}{\det(W^T S_w W)} \quad (1)$$

This problem is equivalent to solve the following eigenvalue problem:

$$S_w^{-1} S_b W = W\Lambda \quad (2)$$

However, LDA-based algorithms suffer from a small sample size problem in which the with-in class scatter matrix is singular. When there is only one training sample available, the within-class scatter matrix Sw cannot be calculated. In turn LDA method can't be applied directly. This is called one training sample problem for LDA.

3.1 Mathematics of LDA

[1] A = X – [M_e M_e ……. M_e] | *The same as PCA*.
[2] $S_b = S_b + ClsMean_i * ClsMean_i$ | *Computer N_2* N_2 between class scatter matrix.*
[3] $S_w = S_w + (X(j) – ClsMean_i) * (X(j) – ClsMean_i)`$ | *Compute N_2 * N_2 within class scatter matrix*.
[4] S_{bb} = P1` * S_b * P1 |*Using the eigenspace calculated in PCA*
S_{ww} = P1` * S_w * P1 | *project S_b, S_w to Eigenspace*.
[5] [V,D] = eig(S_{bb}, Sww) |*Generalized Eigenvalue* S_{bb} * V = S_{ww} * V * D. *decomposition solved*
[6] eigvals =diag(D)
λ = eigvals(1: Mp)
P = P1 * V(: 1 :Mp)
Train_wt = P1 * A | Store training weights for classifier use.

The testing phase of Linear Discriminant Analysis is as shown in below figure 1.

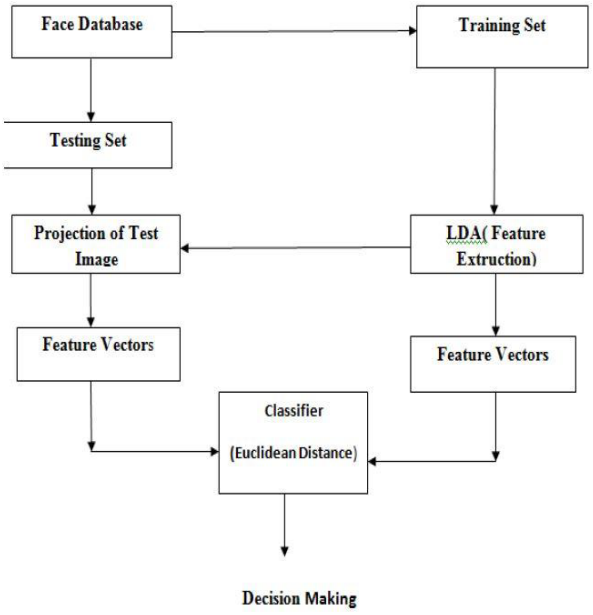

Figure 1. Steps involved in LDA

3.2 Application Of LDA

1. LDA techniques used in statistics, pattern recognition and machine learning to find a linear combination of features characterized or separated two or more classes of objects or events.
2. Dimensionality reduction.
3. Linear classifier.

This paper combines the concept of component-based approach and LDA method, and develops a

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 4, Issue 2, February 2017

component based LDA method for face recognition with one training sample. The rest of this paper is organized as follows. We will give a brief review on LDA algorithm in Section 3. The database details are represented in section 4. Our proposed method and experimental results are presented in.

IV. DATABASE USED

4.1 Training Datasets I

For the experiment we used AT & T standard database of faces which is also known as "The ORL Database of faces." There are ten different images of each of 40vdistinct subjects. For some subjects, the images were taken at different times, varying the lighting, facial expression (open or closed eyes, smiling or not smiling) & facial details such as with glasses or without glasses.

All the image were taken against a dark homogeneous background with the subjects in an upright, frontal position. All the files are in PGM format, which are converted into JPEG for experiment with 256 grey levels per pixel. The size of each image is 92 x 112 pixels.

Figure 2. Training database of AT & T

4.2 Training Datasets II

We use the frontal camera to capture the faces for the experiment with 720x1280 resolution of HP Truevision HD. The captured images are in colored format which are further converted into grayscale of size 92x112 pixels with JPEG file format to be used for experiment. We use 20 different subjects vary in the range of age i.e. first age group is between 5 to 10 second age group between 15 to 25 and third age group is between 30 to 40 with varying in expression, pose. Also the images are captured from 1 meter distance in different lighting condition. Per subject minimum 10 images are stored into the database as a result the database contains 214 images of faces as shown in below figure3.

Figure 3. Database of faces using frontal camera of HP laptop.

4.3 Test Database

Here we take one image for each subject contains 40 subjects of AT & T as well as 20

Figure 4. Samples from test database.

V. EXPERIMENTS AND RESULTS

5.1 experiment 1: For PCA algorithm

- In this experiment, we select 2 images of 4 persons, total 28 & one image is for testing in test

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 4, Issue 2, February 2017

folder. The resolution of the normalized image is 92 x112. Calculate the principle component of each person.

- As we are going to increase number of samples per person from selected 2 to 10 images, we found that as the number of samples is increased the recognition rate also increased from 50 % to 98 % as shown in table 1.

SR.NO.	PERSON	NUMBER OF SAMPLES	% OF ACCURACY
1.	1	2	50 %
2	2	2	49 %
3	3	2	52 %
4	4	2	54%
5	1	3	60%
6	2	3	55%
7	3	3	54%
8	4	3	52%
9	1	4	63%
10	2	4	64%
11	3	4	65%
12	4	4	56%
13	1	5	70%
14	2	5	69%
15	3	5	66%
16	4	5	68%
17	1	6	75%
18	2	6	74%
19	3	6	73%
20	4	6	76%
21	1	7	80%
22	2	7	96%
23	3	7	97%
24	4	7	90%

Table 1: Result for PCA Using FERET Database

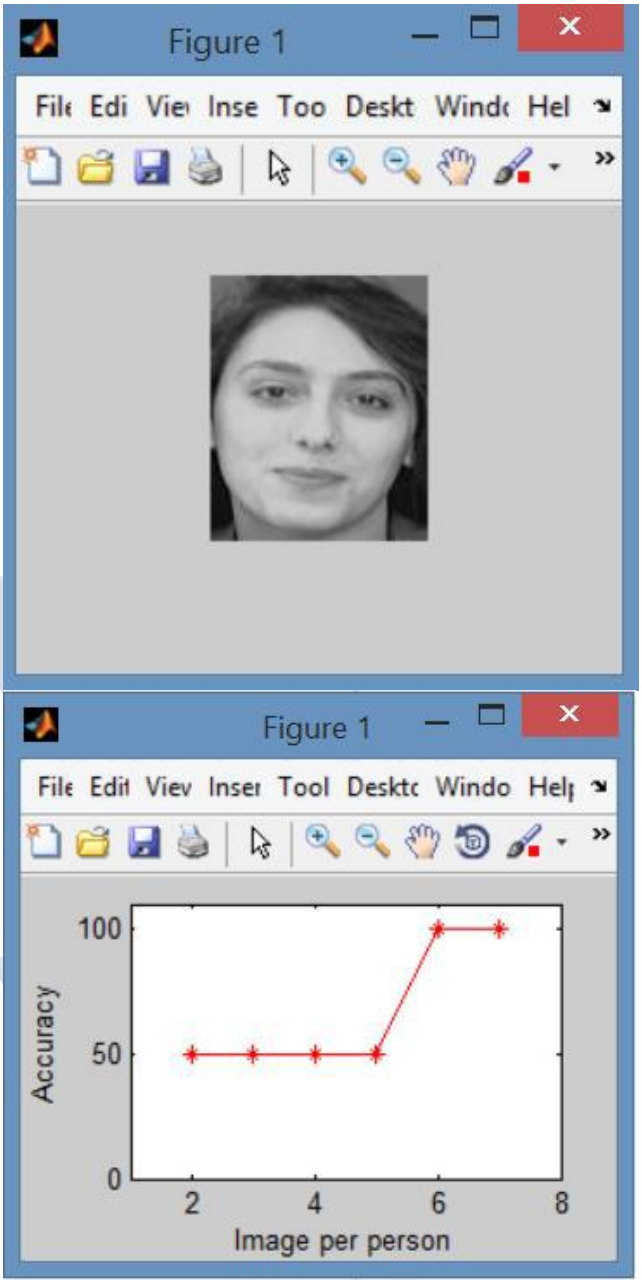

Figure 5. Screen shots for PCA algorithm.

5.2 Experiment 2

In this experiment we take the real time database for faces with different variations such as lighting, expressions & partial occlusion. For this experiment using PCA algorithm we get the following result as shown in table 2. As we calculated the false acceptance ratio and false rejection ratio as 80.8% and 25.50% respectively.

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 4, Issue 2, February 2017

SR.NO.	PERSON	NUMBER OF SAMPLES	% OF ACCURACY
1.	1	8	93.73 %
2	2	7	75.89 %
3	3	12	98.00 %
4	4	4	45.44%
5	1	10	92.89%
6	2	8	78.89%
7	3	10	92.9%
8	4	5	52.00%
9	1	6	62.00%
10	2	4	46.51%
11	3	9	88.89%
12	4	12	98.99%

Table 2: Result For Pca Using Real Time Database

5.3 Experiment 3

In this experiment we implement LDA algorithm for feature extraction for the solution of large samples per person problem. We take 2 samples as training set from FEREST dataset. We observed that as there is only 2 samples but the accuracy rate is 100% for LDA. Same experiment was developed for real time dataset which contains variations in lighting, pose, expression etc. got the 98.63% result.

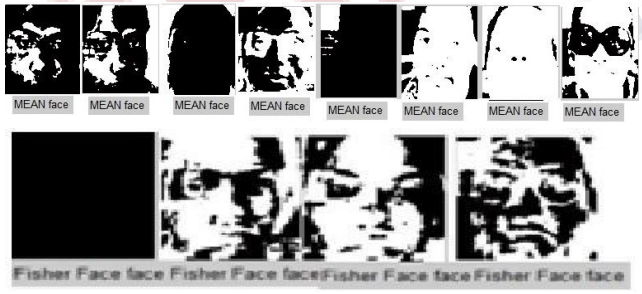

Figure 6.Some mean Faces & Figure faces obtained from LDA algorithm

Figure 7. Result showing the match found successfully.

VI. CONCLUSION

In this paper we successfully shown the experimental result obtained from PCA & LDA implementation. As we can conclude that to reduce the no. of sample problem the LDA gives the better result than PCA. Experimental results show that our component-based LDA method is more robust than that of PCA algorithm. FERET face database is used for evaluation and the results are encouraging.

Acknowledgment

The authors would like to thank for the US Army Research Laboratory for contribution of the FERET database. S. A. Sovitkar sincerely thanks to all the colleagues, students and other members for their valuable support, suggestion and encouragement.

REFERENCES

[1] W. Zhao, R. Chellappa, A. Rosenfeld, and J. Phillips, "Face Recognition: A Literature Survey", Technical Report, CFAR-TR00-948, Univ. of Maryland, 2000. (Revised 2002).

[2] M Turk and A Pentland, "Eigenfaces for recognition", Journal of Cognitive Neuroscience, Vol. 3, No.1, pp. 71-86, 1991.

[3] P. N. Belhumeur, J.P. Hespanha, D.J. Kriegman, "Eigenfaces vs. Fisherfaces: recognition using class specific linear projection", IEEE Trans. Pattern Anal. Mach. Intell., Vol. 19, No. 7, pp. 711-720, 1997.

[4] A M Martinez and A. C. Kak, "PCA versus LDA", IEEE Trans. Pattern Anal. Mach. Intell., Vol. 23, 2001, pp. 228-233, 2001.

[5] H. Yu and J. Yang, "A direct LDA algorithm for high-dimensional data --- with application to face recognition", Pattern Recognition, Vol. 34, pp. 2067- 2070, 2001.

[6] P C Yuen, W S Chen and J Huang, "A novel subspace LDA algorithm for face recognition with small number of training images", Submitted to IEEE Transactions on Circuits and Systems for Video Technology, Special issue on Image- and Video based Biometrics, 2002.

[7] R Huang, Q Liu, H Lu and S Ma, " Solving small sample size problem in LDA", Proceeding of International Conference in Pattern Recognition (ICPR 2002), Vol. 3, Aug 2002.

[8] A M Martinez, "Recognizing imprecisely localized partially occluded and expression variant faces from a single sample per class", IEEE Trans. Pattern Anal. Mach. Intell., Vol. 24, N0. 6, pp. 748-763, 2002.

[8] Jianxin Wu and Zhi-Hua Zhou, "Face Recognition with One Training Image per Person", Pattern Recognition Letters, vol.23, no.14, pp.1711-1719, 2002.

[9] L Wiskott, J Fellous, and C von der Malsburg, "Face Recognition by Elastic Bunch Graph Matching", IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 19, pp. 775-779, 1997.

[10] J Huang, P C Yuen and J H Lai, Combining local and global features for face recognition, Proceedings of ACCV 2002. No. 23-25, 2002.

[11] S.A. Sovitkar, S.S. Kawthekar, "A Review On-Automated Attendance Management System Using Face Recognition", Golden Research Thoughts (GRT), Vol. 5, no. 16, 2016.