JCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.-7, Special Issue, 5, March 2019 E-ISSN: 2347-2693

Face Recognition System using Modular Principal Component Analysis

S.P. Sundarsingh[1*], C.D. Daniel Dharamaraj[2]

[1]Research Centre in Computer Science,V.H.N.S.N College (Autonomous), Virudhunagar,India
[2] Dept. of Computer Science, Research Centre in Computer Science,V.H.N.S.N College (Autonomous), Virudhunagar, India

Corresponding Author: sumisundar.ps@gmail.com

DOI: https://doi.org/10.26438/ijcse/v7si5.140145 | Available online at: www.ijcseonline.org

Abstract- This paper aims to present face recognition based on Principal Component Analysis (PCA) and Modular Principal Component Analysis (MPCA) approach. The PCA based face recognition method is not very effective under the conditions of varying poses and expressions rather than the proposed MPCA method. In the MPCA method the original face image was partitioned into tiny sub-images and then PCA technique is applied for each sub-image. Since a few of the normal facial features of an individual do not differ even when the pose and expression may differ, the proposed method manages these variations and takes only a few numbers of principal components for matching the faces for similarity. The proposed method improves the recognition rates with less number of principal components when compared with the conventional PCA method. This present system is tested with two standard face databases and results are presented.

Keywords: Eigen faces; Euclidean Distance; Face Recognition; MPCA; Principal Component Analysis.

I. INTRODUCTION

Face Recognition (FR) is one of the divisions of biometric recognition. Human face recognition, as one of the most successful applications of image analysis and understanding, has received significant attention in the last decade [1]. For humans, it is an instinctive and innate ability human, for computers, it is viewed as a complicated problem. It is easy for a human to decide where an object ends and when the next starts, but to a computer, a

digital image is just a matrix of pixels. The human face is full of information but working with all the information is time consuming and less efficient. The process of computer based facial recognition determines identity using facial features as elements of distinction and discards other useless information. The importance of research on FR is determined by both its wide range of potential applications and scientific challenges [2]. The automated methods of facial recognition, even though work very well, do not recognize subjects in the same manner as a human brain [3, 4]. Our goal is to develop a face recognition system that is fast, reasonably simple, and accurate in constrained environments.

FR is viewed as high-dimensional pattern recognition problem, even though low resolution face images generate huge dimensional feature spaces. In addition to the problems of large computational complexity and memory storage, this high dimensionality makes very difficult to obtain statistical models of the input space using well defined parametric models [5, 6]. However, the intrinsic dimensionality of the image space is much lower than the dimensionality of the face space. Many techniques can be used for face recognition but Principal Component Analysis (PCA) is one of the most victorious techniques that have been used in face recognition .Sirovich et al. [7, 8] first used PCA to efficiently represent pictures of human faces. PCA extracts efficient features from high-dimensional vectors of input data and encodes the relevant information in a face image as efficiently as possible to reduce the data. This encoded data will be use as the features of a collection of faces. The aim of the present work is to extract these features, which are the variations in the face images, to be able to compare them.

Since, the traditional PCA method only utilizes the global information of face images, its recognition accuracy deeply affected by the variation in head pose, lighting condition and facial expression in the image [9-12]. But some of local facial features of an individual do not change when the facial expression and pose change. Recent improvement is the Modular Principal Component Analysis [13]. In this technique, the face images are divided into smaller regions and the PCA approach is applied for each one of these sub images. It keeps the relation between regions as well as the global information of the face.

In this paper, the state of the art in face recognition is achieved over the different facial expressions and poses for different persons is discussed by using the most promising face recognition algorithm on PCA based on eigenface technique and MPCA. It describes how these techniques can be generalized to handle different facial expressions and poses to find the correct similarity on persons faces accurately and compute the recognition rate accuracy.

The rest of the paper is organized as follows: In section 2 the proposed MPCA technique is formulated for selecting number of principal components for classification. Design and implementation are explained in section 3. Section 4 presents the experimental results and followed by the discussion and conclusion is presented in section 5.

II. MODULAR PRINCIPAL COMPONENT ANALYSIS

Most of the automatic face recognition algorithms evaluate faces as one unit which leads to problems due to variations in expression, illumination and pose [14-17]. This neglects the important fact that few facial features are expression invariant and others are more susceptible to the expressions. MPCA is one such improvement over PCA. This MPCA approach the entire face image is divided into smaller regions and the feature extraction is computed for each one of these regions. For each sub region of training set, average sub-region computation, covariance matrix, eigenvectors and the weight set calculation as in PCA. The weights will be more representative of the local information of the face. The weights of the face images will closely match the weights in the same region of an individual face image under normal conditions. The given face is classified as the face belong to that class which is at the nearest Euclidean distance in the face space. Since some variations on face images do not affect the entire information of the faces, this approach takes advantage of the unaffected regions of the face to increase its accuracy rate. Therefore it is expected that improved recognition rates can be obtained by using the following MPCA approach. Furthermore it also recognizes the exact matched face image from selecting few numbers of principal components rather than PCA.

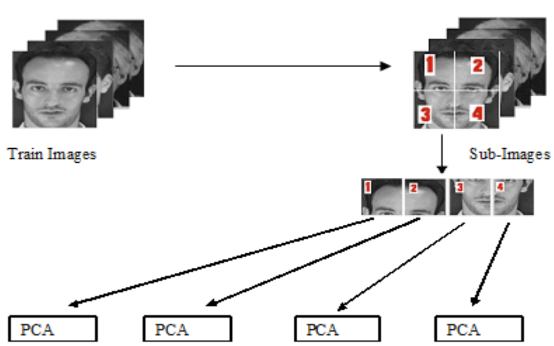


Fig.1 Dividing the whole image into sub images and applying PCA technique

In this method, each image in the training set is divided into N smaller regions. Hence the size of each sub-image will be r ×c/N.These sub-images can be represented mathematically as

$$I_{ij}(m, n) = I_i\left(\frac{r}{\sqrt{N}}(j-1) + m, \frac{c}{\sqrt{N}}(j-1) + n\right)\ \forall_{i,j} \quad (1)$$

where i varies from 1 to M, M being the total number of images in the training set, j varies from 1 to N, N being the number of sub-images and m and n vary from 1 to $r/\sqrt{N}$ and $c/\sqrt{N}$.

The next step is to find mean image of all the training sub-images is computed as

$$\Psi = \frac{1}{M.N} \sum_{i=1}^{M} \sum_{j=1}^{N} I_{ij} \quad (2)$$

Subtracting the image matrix with its mean,

$$Y_{ij} = I_{ij} - \Psi \quad (3)$$

The covariance matrix is computed from these sub-images as follows,

$$C = \frac{1}{M.N} \sum_{i=1}^{M} \sum_{j=1}^{N} Y_{ij} . Y_{ij}^{T} \quad (4)$$

From the covariance matrix, Eigen values and Eigenvectors are computed. The outcome of the covariance matrix is symmetric. The image weights are computed by multiplying the eigenvectors and generating the principal components as follows,

$$PC = (I_{ij} - \Psi) . \nu \quad \forall_{ij} \quad (5)$$

where ν are eigenvectors.

Next classifying the test image and the weights are also computed for the test sub-images using eigenvectors as shown in the following equation:

$$\Omega w = PC^{T} . (\Gamma_{test\ j} - \Psi); \ w=1, 2.......M, \ \forall j \quad (6)$$

The Euclidean distance between two weight (feature) vectors thus provides a measure of similarity between the corresponding images. Mathematically, recognition is finding the minimum Euclidean distance ε_k, between a test point and a training point given in the following.

$$\varepsilon_k = \sqrt{\| \Omega_w - \Omega_k \|^2}; \ k=1,......M \quad (7)$$

For the extraction of significant information from the sample images of two databases, we have to choose an optimal number of PCs. We can achieve this by calculating variance

$$\text{Variance} = \frac{\lambda_i}{\sum_{i=1}^{M} \lambda_i} \times 100 \quad (8)$$

III. DESIGN OF THE PROPOSED SYSTEM

A user friendly Graphical User Interface (GUI) program is developed in the present work. The source code of the face recognition system is written in MATLAB. The process explained in section 3 is designed in Matlab 7.10 (R2010a) to read the databases, process them and to display the resultant images. Once sample images (UMIST and ATT database) are loaded into the GUI, evrytime the recognition process is done in these loaded images. The minimum distances of sample face images and recognition is calculated and displayed in the GUI window. The quality of the resultant images are assessed and significant information is retrieved from the software system. These activities are performed by clicking the menu items in the GUI. The Fig. 2 shows the flow diagram of the proposed system.

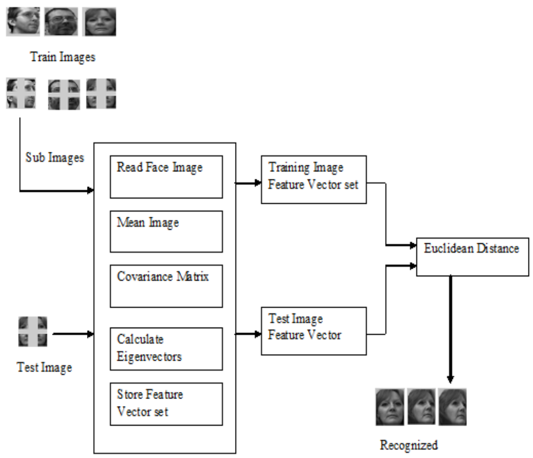


Fig.2 System Using proposed Technique

IV. RESULTS AND DISCUSSION

In this section, the performances of traditional PCA-based algorithm and the MPCA-based algorithm are evaluated with two standard face image databases. The ATT database consists of face images with varying expressions and the UMIST database consists of face images with varying poses. Table.1 shows the details of image databases, utilized for the present study. A class consists of face images belonging to either training or testing.

Table 1:Image databases used in the present work for training and testing.

Phase	Database	Number of classes	Number of images in each class	Total number of images	Image size
Training	UMIST ATT	20 20	20 9	400 180	112×112 112×112
Testing	UMIST ATT	20 20	1 1	20 20	112×112 112×112

During the first study, twenty sample images per class from UMIST are considered for training and one sample image per class are taken for testing. The images of different poses of each person ranging from profile to frontal views are included in the samples. The sample images used for training and those for testing in the present work are displayed in Fig.3 and Fig.4 respectively.

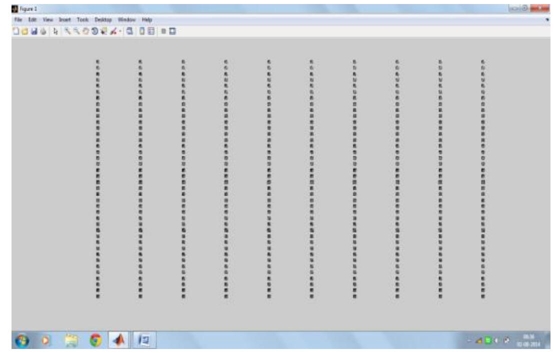


Fig.3. 400 sample images taken from UMIST for training

Fig.4 20 images taken from UMIST for testing

Sample images are arranged into single column vectors. The average vector and the covariance matrix are computed. Subsequently, as described in section 2 the eigenvalues and their eigenvectors are then calculated from the covariance matrix. Finally the first few largest eigenvalues and the corresponding eigenvetors are selected. Thus, the images that are having more information is found and displayed in Fig.5 and their values are presented in Table 2. Such images called eigenfaces that is principal components.

Table 2: Recognition rates of PCA technique of face recognition

Eigenvectors	Recognition Rate
3	30%
5	40%
10	45%
19	50%

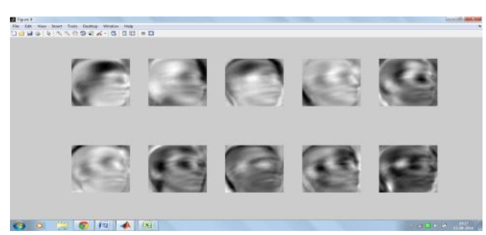


Fig.5 First ten principal components (Eigenfaces)

In order to increase the recognition rates more number of eigenvectors must be chosen. The following GUI window (Fig. 6) shows the results of PCA technique.

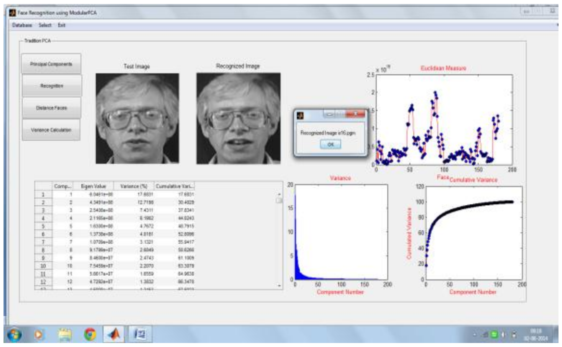
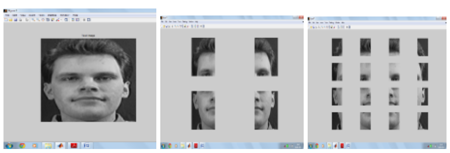


Fig.6 Face Recognition system using PCA Technique

In PCA technique, the whole image is used for calculating eigenvectors. In MPCA technique, the image is divided into smaller regions and the PCA technique is applied into each region. The methodology had already been explained in section 2. Each sample image (Fig.7a) is divided into four regions (Fig.7b) or sixteen regions (Fig.7c).



7a 7b 7c

Fig.7 Graphical representation of image subdivision based on MPCA technique (7a. Whole Image, 7b four regions and 7c sixteen regions)

In the present work, we observe increased recognition rates with few numbers of eigenvectors, taken for processing. The recognition rates of an image divided into 4 regions using MPCA technique are presented in Table 3.

Table 3: Recognition rates of present MPCA technique.

Eigenvectors	Recognition Rate
3	75%
5	80%
10	90%
19	95%

In the present work, 400 sample images have been taken from UMIST database and each sample image is divided into 4 regions (400×4=1600) or 16 regions (400×16=6400) in training set. The system is tested with 4 regions, increased recognition rates achieved by selecting the minimum number of eigenvectors. Furthermore it also recognizes the exact matching face image for the test image. Fig.8 shows the results obtained from 10 eigenvectors out of 1600 regions (400 sample images) and 90% of recognition accuracy is achieved.

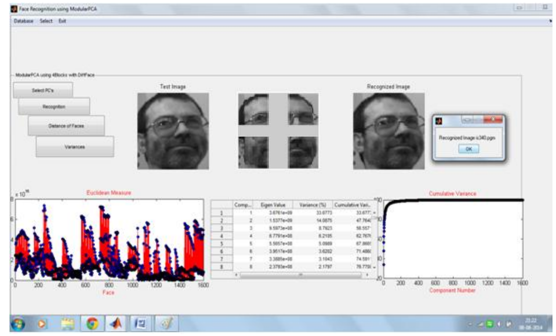


Fig.8 Face Recognition using 4 regions

The following Fig.9 and Fig.10 displays the variance table for PCs. In this table column 5 shows the cumulative variance used for selecting the required number of PCs for testing.

table

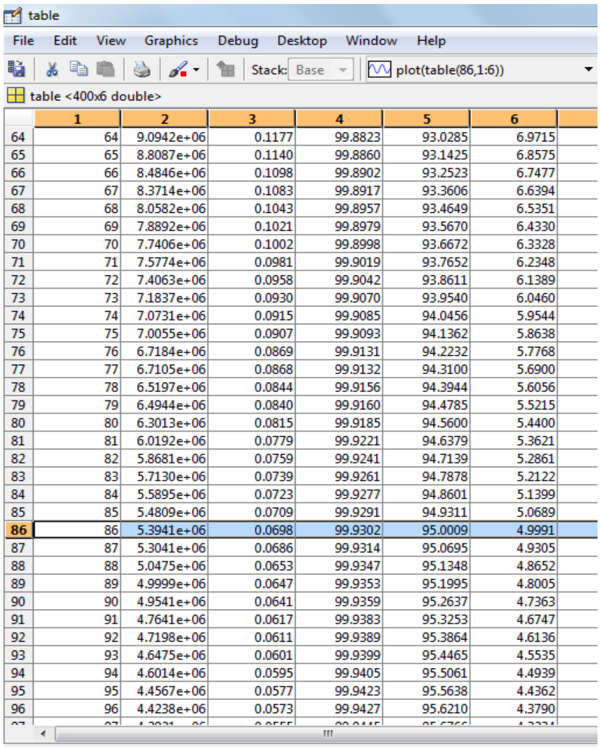
File Edit View Graphics Debug Desktop Window Help

Stack: Base plot(table(19,1:6))

table <6400x6 double>

	1	2	3	4	5	6	7
1	1	5.3918e+09	46.9107	53.0893	46.9107	53.0893	
2	2	1.8838e+09	16.3894	83.6106	63.3001	36.6999	
3	3	1.3661e+09	11.8852	88.1148	75.1854	24.8146	
4	4	5.0794e+08	4.4193	95.5807	79.6046	20.3954	
5	5	4.2766e+08	3.7208	96.2792	83.3254	16.6746	
6	6	3.0912e+08	2.6895	97.3105	86.0149	13.9851	
7	7	1.7823e+08	1.5507	98.4493	87.5656	12.4344	
8	8	1.6308e+08	1.4188	98.5812	88.9844	11.0156	
9	9	1.2857e+08	1.1186	98.8814	90.1030	9.8970	
10	10	1.0399e+08	0.9048	99.0952	91.0077	8.9923	
11	11	9.1874e+07	0.7993	99.2007	91.8071	8.1929	
12	12	7.7414e+07	0.6735	99.3265	92.4806	7.5194	
13	13	5.8824e+07	0.5118	99.4882	92.9924	7.0076	
14	14	5.4346e+07	0.4728	99.5272	93.4652	6.5348	
15	15	4.6804e+07	0.4072	99.5928	93.8724	6.1276	
16	16	4.6198e+07	0.4019	99.5981	94.2744	5.7256	
17	17	3.9267e+07	0.3416	99.6584	94.6160	5.3840	
18	18	3.1311e+07	0.2724	99.7276	94.8884	5.1116	
19	19	3.1015e+07	0.2698	99.7302	95.1583	4.8417	
20	20	2.7165e+07	0.2363	99.7637	95.3946	4.6054	
21	21	2.4033e+07	0.2091	99.7909	95.6037	4.3963	
22	22	2.2078e+07	0.1921	99.8079	95.7958	4.2042	
23	23	2.0922e+07	0.1820	99.8180	95.9778	4.0222	
24	24	1.9051e+07	0.1658	99.8342	96.1436	3.8564	
25	25	1.8191e+07	0.1583	99.8417	96.3019	3.6981	
26	26	1.6657e+07	0.1449	99.8551	96.4468	3.5532	
27	27	1.5998e+07	0.1392	99.8608	96.5860	3.4140	
28	28	1.5014e+07	0.1306	99.8694	96.7166	3.2834	
29	29	1.2668e+07	0.1102	99.8898	96.8268	3.1732	
30	30	1.1965e+07	0.1041	99.8959	96.9309	3.0691	
31	31	1.1710e+07	0.1019	99.8981	97.0328	2.9672	
32	32	1.1017e+07	0.0959	99.9041	97.1286	2.8714	
33	33	1.0009e+07	0.0871	99.9129	97.2157	2.7843	

Fig.9



	1	2	3	4	5	6
64	64	9.0942e+06	0.1177	99.8823	93.0285	6.9715
65	65	8.8087e+06	0.1140	99.8860	93.1425	6.8575
66	66	8.4846e+06	0.1098	99.8902	93.2523	6.7477
67	67	8.3714e+06	0.1083	99.8917	93.3606	6.6394
68	68	8.0582e+06	0.1043	99.8957	93.4649	6.5351
69	69	7.8892e+06	0.1021	99.8979	93.5670	6.4330
70	70	7.7406e+06	0.1002	99.8998	93.6672	6.3328
71	71	7.5774e+06	0.0981	99.9019	93.7652	6.2348
72	72	7.4063e+06	0.0958	99.9042	93.8611	6.1389
73	73	7.1837e+06	0.0930	99.9070	93.9540	6.0460
74	74	7.0731e+06	0.0915	99.9085	94.0456	5.9544
75	75	7.0055e+06	0.0907	99.9093	94.1362	5.8638
76	76	6.7184e+06	0.0869	99.9131	94.2232	5.7768
77	77	6.7105e+06	0.0868	99.9132	94.3100	5.6900
78	78	6.5197e+06	0.0844	99.9156	94.3944	5.6056
79	79	6.4944e+06	0.0840	99.9160	94.4785	5.5215
80	80	6.3013e+06	0.0815	99.9185	94.5600	5.4400
81	81	6.0192e+06	0.0779	99.9221	94.6379	5.3621
82	82	5.8681e+06	0.0759	99.9241	94.7139	5.2861
83	83	5.7130e+06	0.0739	99.9261	94.7878	5.2122
84	84	5.5895e+06	0.0723	99.9277	94.8601	5.1399
85	85	5.4809e+06	0.0709	99.9291	94.9311	5.0689
86	86	5.3941e+06	0.0698	99.9302	95.0009	4.9991
87	87	5.3041e+06	0.0686	99.9314	95.0695	4.9305
88	88	5.0475e+06	0.0653	99.9347	95.1348	4.8652
89	89	4.9999e+06	0.0647	99.9353	95.1995	4.8005
90	90	4.9541e+06	0.0641	99.9359	95.2637	4.7363
91	91	4.7641e+06	0.0617	99.9383	95.3253	4.6747
92	92	4.7198e+06	0.0611	99.9389	95.3864	4.6136
93	93	4.6475e+06	0.0601	99.9399	95.4465	4.5535
94	94	4.6014e+06	0.0595	99.9405	95.5061	4.4939
95	95	4.4567e+06	0.0577	99.9423	95.5638	4.4362
96	96	4.4238e+06	0.0573	99.9427	95.6210	4.3790

Fig.10 Variance table in MATLAB editor window

Next for the same database, the system is tested with 16 regions (400×16=6400) and is shown in Fig 11 by achieving 95% of recognition accuracy. Exact matching is done only by using 19 eigenvectors instead of 6400.

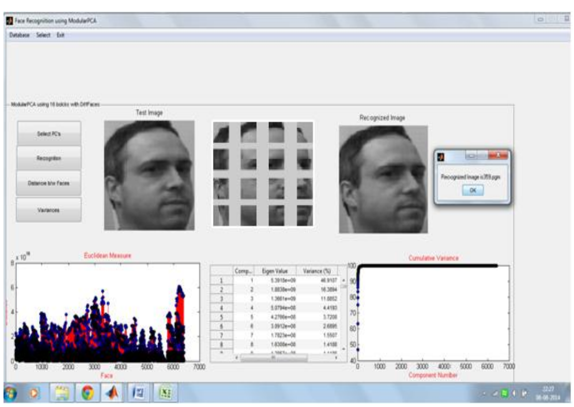


Fig .11 Face Recognition using 16 regions

The Euclidean distances for each region are computed for the two instances (M=1600 and M=6400). For each of the region, the minimum distance and its corresponding region are observed. The face is thus recognized from that region is classified correctly. With the same pose of images all the 4 regions are exactly matched in 4 regions case and 16 regions are matched in 16 regions case, plotted in Fig 12 and 13 respectively. The system is tested with the face images with different poses.

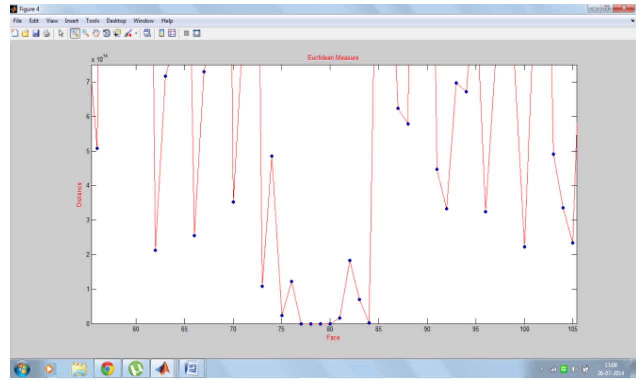


Fig.12 Euclidean distance graph of 4 regions

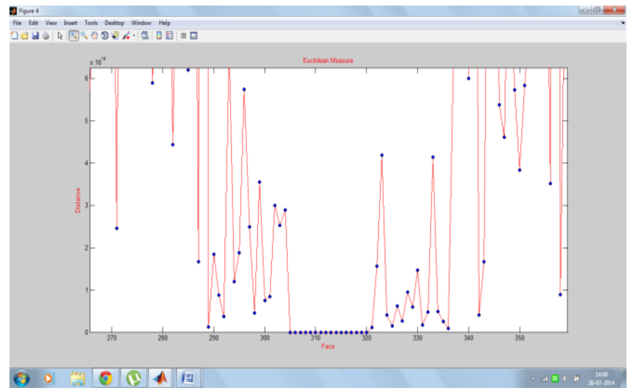


Fig.13 Euclidean distance graph of 16 regions

The recognition rates are computed and plotted as shown in Fig.14 This study has been performed with varied number of regions and principal components. A better classification results has been achieved with 19 principal components, selected from 6400 principal components using UMIST dataset and 10 principal components selected from 2880 using ATT database. A recognition accuracy of 95% is obtained for UMIST dataset and 95% is obtained for ATT dataset.

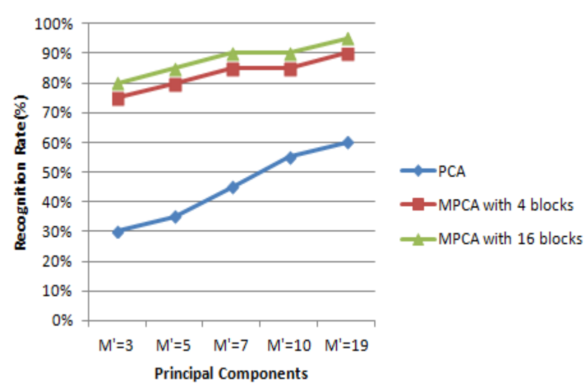


Fig.14 Recognition accuracy of MPCA technique for pose (UMIST) Database

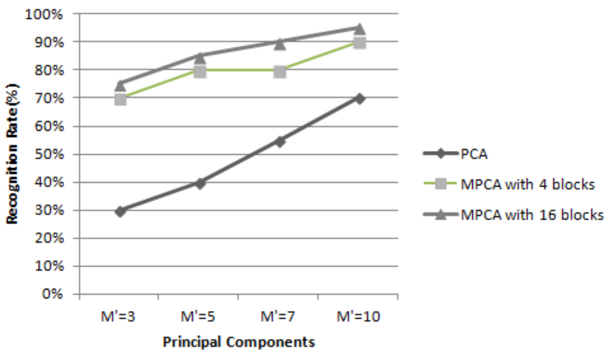


Fig.15. Recognition accuracy of MPCA technique for Expression (ATT) Database

The following Table.4 shows the overall results of recognition rate when choosing the number of principal components using MPCA technique. This technique uses the number of regions of images and this is very helpful for matching the similarity faces from the test image and database images by choosing only a minimum number principal components. Hence, the table shows the recognition rate, number of regions, database, total number of PC's and selected PC's.

Table 4: Recognition accuracy in ATT and UMIST datasets

Dataset	Method	Regions	High dimensional space (M=Total number of PC's) M	Recognition Accuracy (%) Selected number of PCs(M') M'=3	M'=5	M'=7	M'=10	M'=19
ATT	PCA	Entire Image	180	30	40	55	70	70
	MPCA	4	720	70	80	80	90	90
		16	2880	75	85	90	95	95
UMIST	PCA	Entire image	400	30	35	45	55	60
	MPCA	4	1600	75	80	85	85	90
		16	6400	80	85	90	90	95

V.CONCLUSION

In this work, the main advantage of face recognition method based on MPCA has been discussed. The ability to extract the features of sub-images, which better reflects the differences between the face images and is presented in a graphical user interface. The smaller the number of images, the process becomes much simpler and faster. An improved face recognition accuracy is achieved by selecting only a few numbers of PC'S based on MPCA technique. The recognition system is tested with UMIST and ATT face databases and produces better recognition rates. A recognition accuracy of 95% is obtained for UMIST dataset and 95% is obtained for ATT dataset.

REFERENCES

[1] Muge Carikci, Figen Ozen, "A face recognition method based on eigenface method", Prcedia Techonology 1 (2012) 118-123.
[2] H. Wang, S. Chen, Z. Hu, B. Luo , "Probabilistic two-dimensional principal component analysis and its mixture model for face recognition", Neural Comput & Applic (2008) 17: 541-547
[3] M. Turk, A. Pentland, "Eigenfaces for recognition", J. Cognitive Neurosci. 3 (1) (1991) 71-85
[4] M. Turk, A. Pentland, "Face recognition using eigenfaces", Proceedings of IEEE Conference on Computer Vision and Pattern Recognition 1991, pp. 586-591
[5] P.N Belhumeur, J. P. Hespanha, and 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, Jul. 1997.
[6] R.O.Duda, P. E. Hart, and D. G. Stork," Pattern classification", 2nd ed.. USA: John Wiley & Sons, 2001.
[7] L. Sirovich , M. Kirby, "Low-Dimensional Procedure for Characterization of Human Faces," J. Optical Soc. Am., vol. 4, pp. 519-524, 1987.
[8] M. Kirby, L. Sirovich, "Application of the KL Procedure for the Characterization of Human Faces", IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 12, no. 1, pp. 103-108, Jan. 1990.
[9] S.Chen, Y.Zhu, "Subpattern-based principal component analysis", Pattern Recognition 37 (2004) 1081-1083.
[10] K. Tan, S.Chen, "Adaptively weighted sub-pattern PCA for face recognition", Neurocomputing 64 (2005) 505-511.
[11] Hussein Rady, "Face recognition using Principle Component Analysis with different distance classifiers", International Journal of Computer Science and Network Security, Volume 11, No.10, October 2011.
[12] Minkook Cho, Hyeyoung Park," A feature analysis for dimension reduction based on a data generation model with class factors and environment factors", Computer Vision and Image Understanding, 113 (2009) 1005-1016.
[13] R. Gottumukkal, V. K. Asari," An improved face recognition technique based on modular PCA approach", Pattern Recognition Lett. 25 (4) (2004) 429-436.
[14] Rajesh Kumar Gupta, Umesh Kumar Sahu," Real time face recognition under different conditions", International Journal of Advance Research in Computer Science and Software Engineeirng, Volume 3,Issue 1, January 2013.
[15] Jie Wang, K. N. Plataniotis, A. N. "Venetsanopoluos, Select discriminant for face recognition", Pattern Recognition Letters 26 (2005) 1470-1482.
[16] Xiaoyang Tan, Bill Triggs, "Enhanced Local Texture Feature Sets for Face Recognition Under Difficult Lighting Conditions", IEEE Transactions on Image Processing, Vol. 19, No .6, June 2010.
[17] Javier Ruiz-del-Solar, Pablo Navarrete, "Eigenspace – Based Face Recognition: A Comparative Study of Different Approches", IEEE Transcations On Systems, Man, And Cybernetics – PART C: Applications and Reviews, Vol., 35, No. 3, August 2005.
[18]http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html
[19]http://see.xidian.edu.cn/vipsl/database_Face.html#UMIST

