Face Recognition using Self-Organizing Map and Principal Component Analysis

Dinesh Kumar
Department of Computer Science & Engineering
Guru Jambheshwar University
Hisar
Haryana, INDIA-125001
E-mail: dinesh_chutani@yahoo.com

C.S.Rai
University School of Information Technology
GGS Indraprastha University
Delhi
INDIA-110006
E-mail: csrai_ipu@yahoo.com

Shakti Kumar
Department of Electronics & Communication Engineering
Haryana Engineering College
Jagadhri
Haryana, INDIA-135001
E-mail: shakti_k@yahoo.com

***Abstract*—Face Recognition has always been a fascinating research area. It has drawn the attention of many researchers because of its various potential applications such as security systems, entertainment, criminal identification etc. Many supervised and unsupervised learning techniques have been reported so far. Principal Component Analysis (PCA) is a classical and successful method for face recognition. Self Organizing Map (SOM) has also been used for face space representation. This paper makes an attempt to integrate the two techniques for dimensionality reduction and feature extraction and to see the performance when the two are combined. Simulation results show that, though, the individual techniques SOM and PCA itself give excellent performance but the combination of these two can also be utilized for face recognition. The advantage in combining the two techniques is that the reduction in data is higher but at the cost of recognition rate.**

I. INTRODUCTION

Face Recognition has always been a fascinating research area. It has drawn the attention of many researchers because of its practical applications. Security systems, human-computer interaction, entertainment, criminal identification, video surveillance etc. are among the most common applications of it [1,2]. Many supervised and unsupervised learning techniques have been reported for face recognition. These techniques have been broadly divided into two categories namely i) Holistic or Appearance Based techniques that use the whole face as the input to recognition system, and ii) Feature Based techniques where the local features such as the eyes, and local statistics are fed to the recognition system. Turk and Pentland [4] made the first successful demonstration of machine recognition of faces. The method was known as the Eigenface method that used Principal Component Analysis (PCA) for dimensionality reduction. This approach is based on second order statistics. Linear Discriminant Analysis (LDA), also known as Fisherfaces, a supervised learning algorithm, was proposed by [5,6] and it was claimed that this method is insensitive to large variations in lighting and facial expressions. It is generally believed that algorithms based on LDA are superior to those based upon PCA. However, from some recent work [23], it was concluded that when the training data set is small, PCA can outperform LDA and also PCA is less sensitive to different training data set.

The above two methods aim to preserve the global structure where as in many real world applications, the local structure is more important. In order to preserve the intrinsic geometry of the data and the local structure, a new method was proposed by He et. al. [17] in which Locality Preserving Projections (LPP) were used for mapping the face images into the face subspace. The results showed that this algorithm is especially suitable for frontal face images.

Several other methods such as Probabilistic Subspaces [11-14], Feature Line Method [15], Evolutionary Pursuit [7], Support Vector Machines (SVM) [16] etc. have also been proposed by various researchers with their relative advantages and disadvantages. These subspace methods [5-7,11-14,15] may fail because the intra-class variation becomes zero when there is only one image per class for training. A large number of face recognition algorithms reported in literature used PCA that was based on second order statistics. Bartlett and Sejnowski [8-10] introduced method that considered the higher order statistics also. The method was based upon Independent Component Analysis (ICA).

Self-Organizing maps (SOMs) [18-19] have also been successfully used as a way of dimensionality reduction and feature selection for face space representations [20-22]. This paper makes an attempt to compare SOM, an unsupervised learning algorithm with the popular and successful classical method PCA and to see the performance of face recognition system when the two techniques are combined together. After a brief discussion of SOM and PCA in section 2, we describe the proposed method in section 3. The experiments are reported in section 4 and section 5 contains the conclusions.

0-7803-9422-4/05/$20.00 ©2005 IEEE

II. SOM AND PCA

A. Self-Organizing Maps

T. Kohonen introduced the Self-Organizing Map (SOM) [18-19]. It is an unsupervised learning process, which learns the distribution of a set of patterns without any class information. It has the property of topology preservation. There is a competition among the neurons to be activated or fired .The result is that only one neuron that wins the competition is fired and is called winner-takes all neuron. SOMs may be one-dimensional, two-dimensional or multidimensional, but the most common ones are either one-dimensional or two-dimensional maps. The number of input connections depends on the number of attributes to be used in the classification. The neuron with weights closest to the input data vector is declared the winner during the training. Then the weights of all of the neurons in the neighbourhood of the winning neuron are adjusted by an amount inversely proportional to the distance. It clusters and classifies the data set based on the set of attributes used. The algorithm is summarized as follows [25]:

1. Initialization: Choose random values for the initial weight vectors $\mathbf{w}_j(0)$, the weight vectors being different for $j = 1,2,\ldots,l$ where l is the total number of neurons
2. Sampling: Draw a sample $\mathbf{x}$ from the input space with a certain probability.
3. Similarity Matching: Find the best matching (winning) neuron $i(\mathbf{x})$ at time steps n by using the minimum distance Euclidean criterion
$$i(\mathbf{x}) = \arg\min_j \left\| \mathbf{x}(n) - \mathbf{w}_j \right\|, \quad j = 1,2,\ldots,l$$
4. Updating: Adjust the synaptic weight vector of all neurons by using the update formula
$$\mathbf{w}_{,j}(n+1) = \mathbf{w}_j(n) + \eta(n) h_{j,i(x)}(n) \left(\mathbf{x}(n) - \mathbf{w}_j(n) \right)$$
where $\eta(n)$ is the learning rate parameter, and $h_{j,i(x)}(n)$ is the neighbourhood function centered around the winning neuron $i(x)$. Both $\eta(n)$ and $h_{j,i(x)}(n)$ are varied dynamically during learning for best results.
5. Continue with step 2 until no noticeable changes in the feature map are observed.

B. Principal Component Analysis

Kirby and Sirovich [3] developed a technique based upon which the idea of eigenfaces was used. These eigenfaces represented the faces efficiently using Principal Component Analysis. An image was treated as a vector in a very high dimensional space. Only the best eigenfaces (eigenvectors of the covariance matrix of a set of images) those that had the largest eigenvalues were used to approximate the face.

Consider a set of N sample images $\{\mathbf{\Gamma}_1, \mathbf{\Gamma}_2, \cdots, \mathbf{\Gamma}_N\}$ taking values in an n -dimensional image space. Let us also consider a linear transformation mapping the original n -dimensional image space to m -dimensional feature space where $m < n$. The new feature vectors $\mathbf{Y}_k$ are defined by the following linear transformation
$$\mathbf{Y}_k = \Phi^T \mathbf{\Gamma}_k$$
where Φ is a matrix with orthonormal columns.
The covariance matrix is defined as
$$C = \sum_{k=1}^{N} (\mathbf{\Gamma}_k - \mathbf{\Psi})(\mathbf{\Gamma}_k - \mathbf{\Psi})^T$$
where Ψ is the mean image of all the samples. Only m number of n -dimensional eigenvectors $[\mathbf{V}_1, \mathbf{V}_2, \cdots, \mathbf{V}_m]$ of C is chosen that correspond to the m largest eigenvalues.

III. COMBINING SOM AND PCA

SOM is an unsupervised learning process that has the property of topology preservation. It defines a mapping from an input space onto a set of nodes in a space that has dimension much lower than that of the input space. The set of nodes is topologically ordered. An image, divided into sub blocks, is mapped to a lower dimensional space with topologically ordered set of nodes thereby providing dimensionality reduction. Further feature extraction is provided with the method known as Karhunen – Loeve (KL) transform via Principal Component Analysis (PCA). It is well known that PCA generates a set of orthogonal axes of projections known as principal components or the eigenvectors. PCA is applied to the weight matrix generated by mapping the image onto lower dimensional space using SOM. In order to further reduce the dimensionality, the eigenvectors with smaller eigenvalues are ignored and the eigenvectors corresponding to the largest eigenvalues are retained for image reconstruction.

In PCA method, the pixels of each row of the training images, taken one at a time, are concatenated vertically to form a single vector containing all the pixel values of an image thereby producing a matrix, each column of which represents an image and there are as many number of columns as the number of training images. Fifty percent of the total number of eigenvectors, corresponding to the largest eigenvalues, obtained from the covariance matrix of the training images, is retained. The test images are reconstructed after finding out its KL coefficients using the retained eigenvectors and matching is done using Euclidean norm (L2 norm) as the similarity measure.

Whereas in SOM method, the training images are mapped to lower dimension using SOM and the weight matrix of each training image is stored. At the time of recognition, the training images are reconstructed using the weight matrices and matching is done with the test image using Euclidean norm (L2 norm) as the similarity measure. The same procedure is adopted for the third technique when the two SOM & PCA are combined. We find the eigenvectors of the weight matrix, which we get through SOM training and sixty four percent of the total number of eigenvectors corresponding to largest eigenvalues and hence the KL coefficients is retained and at the time of recognition, the images are reconstructed and matching is done with the test images. In the second case, the weight matrix was obtained by using the eigenvectors and the KL coefficients and the weight vectors corresponding to first twenty neurons were considered for reconstruction of the image thereby reducing the memory space required to store the image.

TABLE 1
RECOGNITION RATE OF THE FACE RECOGNITION SYSTEM WITH VARYING NUMBER OF CLASSES

Recognition Rate (%)			
Method	Number of Classes		
	10	20	40
SOM (5 × 5)	94.06	90.72	89.92
SOM (10 × 10)	94.06	91.86	90.82
PCA	93.39	90.25	89.51
SOM+PCA (1)	77.75	72.08	62.64

TABLE 2
RECOGNITION RATE OF THE SYSTEM WITH CHANGING SUB BLOCK SIZE

Recognition Rate (%)			
Method	Size of sub block		
	(4 × 4)	(8 × 8)	(16 × 16)
SOM (5 × 5)	94.06	94.06	95.95
SOM+PCA (1)	77.75	77.17	72.83
SOM+PCA (2)	68.29	62.17	54.89

IV. EXPERIMENTATION

The ORL (Olivetti Research Laboratory)[24] face database has been used. The database contains 10 different images of 40 distinct subjects. There are variations in facial expressions and facial details. All images were taken against a dark homogeneous background with the subjects in an upright, frontal position, with tolerance for some tilting and rotation of up to about 20 degrees. The images are grayscale with a resolution of 92×112

The original image 92×112 was resized to 80×80 and it was histogram equalized prior to dividing it into sub blocks. Sub blocks of size say 4×4 were chosen. The pixels of each sub block were concatenated to form a single vector representing one sub block. We had a matrix of 16×400, having total of 400 columns, each of which representing 16 pixels corresponding to each sub block. This formed the input for the SOM. Two-dimensional SOM with say 5 nodes per dimension was trained. Finally we had a weight matrix of 25×16. PCA was applied to the transpose of the weight matrix and sixty four percent of the eigenvectors corresponding to the largest eigenvalues were retained for reconstruction of the image. Euclidean norm (L2 norm) was used as the similarity measure to see which images are most alike. As many as 5 training images and the same number of test images were used for performing the experiments. Training and test sets were used without any overlap. The experiments are as follows:

1. The number of classes was varied to see the performance of the face recognition system. As there are as many as 40 classes in the ORL face database, these were varied from 10 to 20 to 40 and recognition rate was found. Table 1 and figure 1 show the recognition rate (%) as the number of classes is varied. Each result is the average of three simulations. Two-dimensional SOM was used having 5 as the number of nodes per dimension. As is clear from the graph, the recognition rate decreases as the number of classes is increased thereby increasing the chances of similarity among the classes and hence resulting in the decrease in performance. A comparison of results was done with PCA technique. Only fifty percent eigenvectors corresponding to largest eigenvalues were retained. The results show that the recognition rate of SOM is slightly better that that of PCA technique. SOM (10 × 10) is even better than SOM (5 × 5) so far as the recognition rate is concerned. While for SOM & PCA (1) combined, sixty four percent of eigenvectors corresponding to the largest eigenvalues were retained for reconstruction of the image.

2. The second experiment was performed to see the effect of changing the sub block size on the performance of face recognition system. The experiment was performed on first 10 classes. The number of training images and test images was kept same i.e. 5. There was no overlapping between the training and test images. Table 2 and figure 2 show the recognition rates as the size of the sub block is changed. As is clear from figure 2 that there is a little change in the recognition rate for SOM and SOM & PCA combined (1) techniques whereas for SOM & PCA combined (2), it is less as compared with the first two techniques and it reduces more rapidly as the size of sub block is increased.

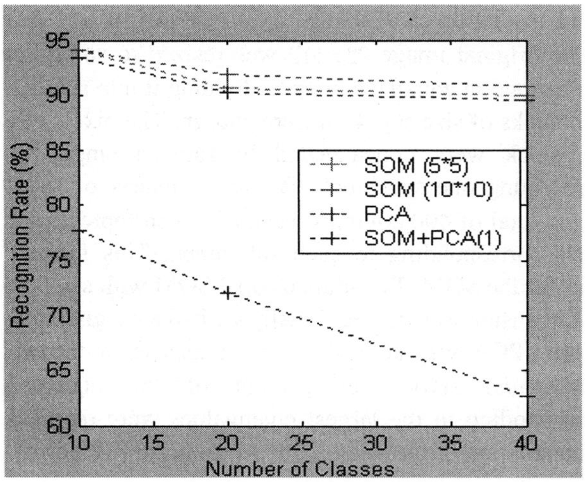

Fig. 1. Recognition Rate as a Function of Number of Classes.

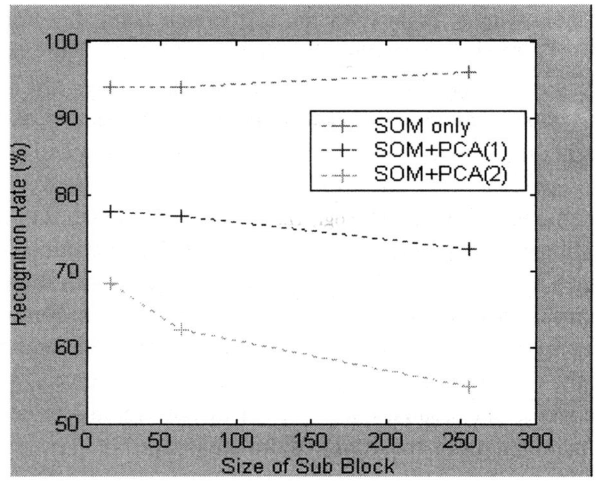

Fig. 2. Recognition Rate as a Function of Changing Sub Block Size.

V. CONCLUSIONS

In this paper a new idea was proposed in which PCA was combined with SOM for dimensionality reduction and feature extraction. The simulation results indicate that the performance of SOM (10×10) is better than SOM (5×5) because of the fact that there is more number of neurons in the earlier one. The results also indicate that the performance of face recognition system decreases as the number of classes (subjects) is increased. This is true for all the three methods i.e. SOM, PCA and SOM & PCA combined. The decrease is more in case of SOM & PCA combined as compared to other two methods. As the number of classes (subjects) increase, the chances of mismatch are more because of more similar faces. The experimental results also show that as the size of the sub block is varied, there is hardly any significant change in the performance of the recognition system for SOM and for the technique where the two SOM & PCA are combined and the sixty four percent of the eigenvectors corresponding to the largest eigenvalues are retained. But for the technique where the two are combined and lesser number of neurons of the weight matrix is used for image reconstruction, the recognition rate is less as compared to the earlier two techniques and decreases more rapidly. The reduced number of neurons and hence reduced number of features result in decrease in the recognition rate, but at the same time, it results in saving in the memory space required for storing the images.

REFERENCES

[1] W.Zhao, R. Chellapa, A. Rosenfeld, P.J Phillips,, "Face Recognition: A Literature Survey ," *ACM Computing Surveys* , Vol.35, No.4, pp. 399-458, 2003

[2] R. Chellapa, C.L.Wilson, S.Sirobey, " Human and Machine Recognition of Faces: A Survey," *Proc. IEEE*, Vol.83, No.5, pp. 705-740,1995

[3] M.Kirby and L.Sirovich, "Application of the Karhunen-Loeve Procedure for the Characterization of Human Faces," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol.12, No.1, pp.103-108, 1990

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

[5] D.L.Swets and J.J.Weng, "Using Discriminant Eigenfeatures for Image Retrieval," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol.18, No.8, pp.831-836, 1996

[6] P.N.Belhumeur, J.P.Hespanha, D.J.Kriegman, "Eigenfaces Vs Fisherfaces: Recognition using Class Specific Linear Projection," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol.19, No.7, pp.711-720, 1997

[7] C.Liu and H.Wechsler, "Evolutionary Pursuit and its Applications to Face Recognition," *IEEE Trans. on Pattern Analysis and Machine Intelligence,* Vol.22, No.6, pp.570-582, 2000

[8] M.S.Bartlett and T.J.Sejnowski, "Independent Components of Face Images: A Representation for Face Recognition," *Proc. of the 4th Annual Joint Symposium on Neural Computation, Pasadena*, 1997

[9] M.S.Bartlett, H.M.Lades, T.J.Sejnowski, "Independent Component Representations for Face Recognition," *Proc. of SPIE Symposium on Electronic Imaging: Science and Technology*; Conference on Human Vision and Electronic Imaging III, California, 1998

[10] M.S.Bartlett, "*Face Image Analysis by Unsupervised Learning and Redundancy Reduction*," Ph.D. Dissertation, University of California, San Diego, 1998

[11] B.Moghaddam, "Principal Manifolds and Bayesian Subspaces for Visual Recognition," *International Conference on Computer Vision*, Greece, pp.1131-1136, 1999

[12] B.Moghaddam and A.Pentland, "Probabilistic Visual Learning for Object Detection," *Proc. International Conference on Computer Vision*, pp. 786-793, 1995

[13] B.Moghaddam and A.Pentland, "Probabilistic Matching for Face Recognition," *IEEE Southwest Symposium on Image Analysis and Interpretation,* pp. 186-191,1998

[14] B.Moghaddam, T.Jebara, A.Pentland, "Efficient MAP/ML Similarity Matching for Visual Recognition," *Proc. of Fourteenth International Conference on Pattern recognition*, Vol. 1, pp.876-881, 1998

[15] S.Z.Li and J.Lu, "Face Recognition using Nearest Feature Line Method," *IEEE Trans. on Neural Networks*, Vol.10, No.2, pp.439-443, 1999

[16] P.J.Phillips, "Support Vector Machines Applied to Face Recognition," Technical Report, *NISTIR* 6241

[17] X.He, S.Yan, Y.Hu, P.Niyogi, H.J.Zhang, "Face Recognition using Laplacianfaces," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol.27, No.3, pp.328-340, 2005

[18] T.Kohonen, *"Self Organization and Associative Memory,"* 2nd Edition, Berlin, Germany: springer-Verlag, 1988

[19] T.Kohonen, *"Self Organizing Map,"* 2nd Edition Berlin, Germany: Springer-verlag, 1997

[20] X.Tan, S.Chen, z.H.Zhou, F.Zhang, "Recognizing Partially Occluded, Expression Variant Faces from Single Training Image per Person with SOM and Soft k-NN Ensemble,"*IEEE Trans. on Neural Networks*, Vol.16, No. 4, pp. 875-886, 2005.

[21] V.E.Neagoe, A.D.Ropot, "Concurrent Self Organizing Maps for Pattern Classification," *Proc. of First International Conference on Cognitive Informatics, ICCI'02*, pp. 304-312, 2002

[22] S.Lawrence,C.L.Giles,A.C.Tsoi, "Convolutional Neural Networks for Face Recognition , " *Proc. of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp..217-222, 1996.

[23] A.M.Martinez and A.C.Kak, "PCA versus LDA," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol.23, No.2, pp.228-233, 2001

[24] http://www.cam-orl.co.uk/facedatabase.html

[25] Simon Haykin, *"Neural Networks – A Comprehensive Foundation,"* 2nd Edition, Pearson Education, 1999