Buletinul Ştiinţific al Universităţii "Politehnica" din Timişoara

Seria ELECTRONICĂ şi TELECOMUNICAŢII
TRANSACTIONS on ELECTRONICS and COMMUNICATIONS

Tom 54(68), Fascicola 1, 2009

LDA versus PCA Followed by a Neural Classifier for Facial Image Recognition

Alexandru Cristian Mugioiu[1]

Abstract – We present a comparison of two feature selection methods for face recognition: Linear Discriminant Analysis (LDA) and Principal Component Analysis (PCA). We assume that the classifier is a neural system called Concurrent Self-Organizing Map (CSOM). The ORL Database of Faces is used for testing the above mentioned cascade. The experimental results are given.

Keywords: pattern recognition, facial recognition, CSOM, PCA, LDA

I. INTRODUCTION

In modern computerized society there is a growing need for identity verification of a person. One non-intrusive technique witch has good results for identity verification is the face recognition. A picture of a person can be easily obtained even without the consent of that person. Things like makeup and sunglasses are not welcome in such a system and, of course, frontal views come along with better results against side-views.

Face recognition systems have a large variety of actual or potential applications like: access control, virtual access control, electronic transactions, identity authentication, hidden surveillance, low-enforcements and a lot more.

A face recognition algorithm is constructed from two parts: feature selection and classification (Fig. 1). Face recognition is a vast domain in witch there are a lot of different approaches for either feature extraction or classification.

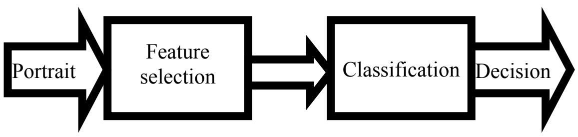

Fig. 1 – Face recognition process

This paper experiments and evaluates a cascade of Linear Discriminant Analysis (LDA) or Principal Component Analysis (PCA) for feature selection followed by Concurrent Self-Organizing Maps (CSOM) [1], [2] for classification.

II. Feature Selection

Feature selection is a key problem in pattern recognition, heaving a strong influence to the performance of statistical classification. Feature selection is a process of linear or nonlinear transformation of the initial space. This transformation is useful because it reduces significantly the workspace, also, some algorithms may have better results in the transformed space than in the initial space.

A. *Principal Component Analysis*

PCA is a widely used technique for dimensionality reduction and pattern recognition. Trough this transformation, images are projected into a subspace where the first orthogonal dimension captures the greatest amount of variance among the images [7]. The main goal of the PCA is dimensionality reduction, therefore the eigenvectors of the covariance matrix should be found in order for a valid solution. The eigenvectors correspond to the directions of principal components of the original data and their statistical significance is given by their corresponding eigenvalues. The eigenvectors constitute the transformation matrix *K*. In order to obtain dimensionality reduction there are kept only a small part of the eigenvectors.

PCA is a simple dimensionality reduction statistically method, witch tends to became a standard method for face recognition.

B. *Linear Discriminant Analysis*

Linear Discriminant Analysis (LDA) is also a widely used method for feature extraction and dimensionality reduction in pattern recognition [5]. It tries to find the best projection in witch the training samples belonging to different classes are best separated. It seeks transformation that maximizes the ratio of the determinant of the between-class scatter matrix and the determinant of the within-class scatter matrix of

[1] Facultatea de Electronică şi Telecomunicaţii, Universitatea Politehnica Bucureşti
Catedra de Electronică Aplicată

the projected samples or finding an optimal projection such as:

$$W_{LDA} = \arg\max \frac{\left|\underline{W}^T \underline{S}_B \underline{W}\right|}{\left|\underline{W}^T \underline{S}_W \underline{W}\right|} = \left[\underline{W}_1 \quad \underline{W}_2 \quad \ldots \quad \underline{W}_m\right]$$

There is one statement witch deserves attention for this algorithm and that is $\det\{\underline{S}_W\} = 0$. This problem appears whenever there is a small number of training images in comparison with the dimension of the training data. To avoid it, first is applied PCA, hence the Fisherfaces algorithm [5]. All this considering, the overall transformation matrix will have the following formula:

$$\underline{W}_{opt} = \underline{W}_{LDA} \cdot \underline{W}_{PCA}$$

LDA is a feature selection method witch is sensitive to background, scale, rotation and translation variations, while it behaves pretty well for lighting variations.

III. Classification with Concurrent Self-Organizing Maps

The CSOM system was firstly proposed and developed by V. Neagoe in [1] and [2]. It was tested for pattern recognition: grayscale face recognition using PCA [2], [4], satellite image classification [2], color facial recognition [3], speech recognition [4]. CSOM is a collection of Self-Organizing Map modules used for classification issues. Each module is dedicated in learning the images of only one pattern. The number of SOM modules is equal with the number of classes. The training set is divided in subsets corresponding to classes (fig. 2). After training, the input pattern is presented to the set of trained modules, the closest weight vector decides the winning neuron and the index (class) of the module in witch resides that neuron gives the classification result (fig. 3).

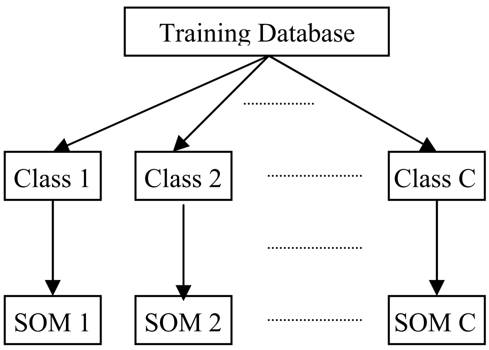

Fig. 2 – CSOM training

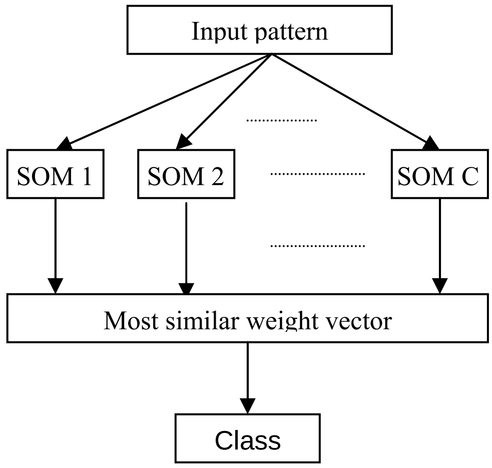

Fig. 3 – CSOM classification

The whole CSOM network is a supervised one but each SOM module is trained in an unsupervised manner. If another class is needed after the training of the CSOM network, then another module is simply added. CSOM combines the attributes of the SOM network with the possibility of adding another class without the time consuming of the retraining process of the whole network. The time spent to train a CSOM network is a lot smaller than the needed time for the training of a SOM network of similar size by the number of the modules.

IV. Experimental Results

The algorithm is tested using the ORL database of faces. This database contains the images of 40 subjects. Each subject has ten images of 112x92 pixels with 256 gray levels. The images were reduced by mediation to 56x46 pixels with 256 gray levels.

Five portraits of each subject are used for training the system and the other five are used for testing it. For rough testing the algorithm, we choose two different partitions for training and testing. In the first partition of training we chose the most similar portrait of each subject. The remaining images are used for testing. The resulted testing set is harder to recognize due to the lack of variation in the training set. The second partition was constructed the other way around: the most different images are chosen for training, while the remaining images are used in testing. Naturally, the testing images are easier to recognize.

Both PCA and LDA algorithm are used for feature selection. By applying PCA the number of features is reduced to 100. A CSOM network classifies the PCA feature vectors. The number of neurons for each SOM module of the CSOM classifier is kept constant. LDA is used with the vectors resulted after applying PCA, keeping the same number of features. A similar CSOM classifier is used for classification of the LDA features.

Tables one through six compare the two methods of feature extraction methods, while the tables seven to ten compare the three architectures of the SOM modules in the CSOM classifier and the eleventh table compares three neural classifiers

Table 1 – Correct recognition rate for CSOM using SOM modules with rectangular architecture for the difficult partition [%]

Feature selection	Number of neurons per SOM module						
	1	3	4	6	8	9	10
PCA	71	73.5	**75**	72.5	74	74	73.5
LDA	72	72.5	72	73	72.5	**75**	73.5

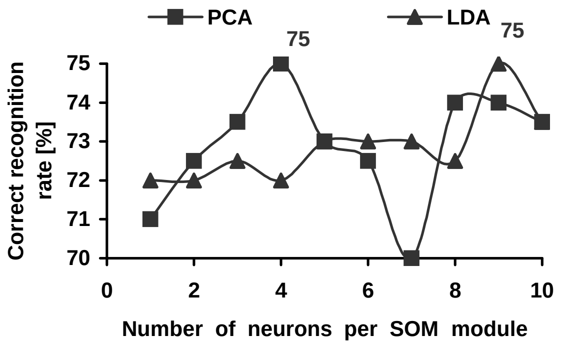

Fig 4 - Correct recognition rate for the difficult partition, CSOM using SOM modules with rectangular architecture

Table 2 – Correct recognition rate for CSOM using SOM modules with rectangular architecture for the simple partition [%]

Feature selection	Number of neurons per SOM module						
	1	3	4	6	8	9	10
PCA	99.5	**100**	99.5	99	98.5	99	97
LDA	**100**	**100**	**100**	99.5	99.5	99	99

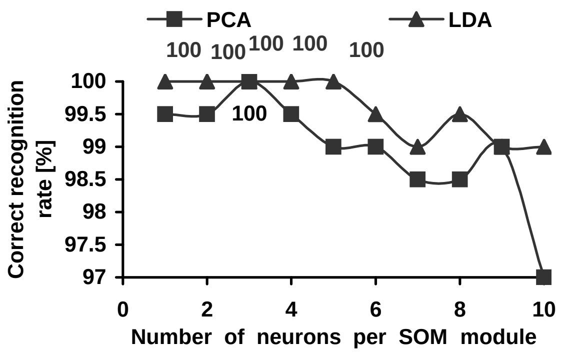

Fig 5 - Correct recognition rate for the simple partition, CSOM using SOM modules with rectangular architecture

Table 3 – Correct recognition rate for CSOM using SOM modules with cylindrical architecture for the difficult partition [%]

Feature selection	Number of neurons per SOM module						
	1	3	4	6	8	9	10
PCA	71	72.5	**74.5**	73	74	73	73.5
LDA	72	73	72	73	72.5	**74.5**	73

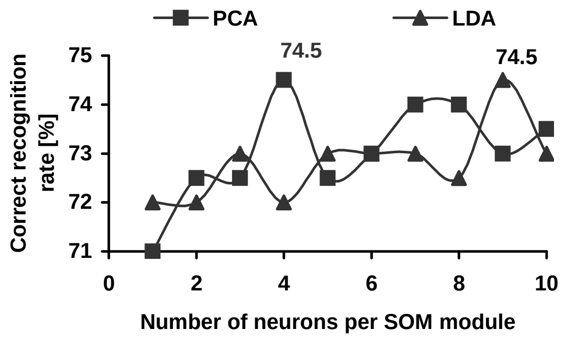

Fig 6 - Correct recognition rate for the difficult partition, CSOM using SOM modules with cylindrical architecture

Table 4 – Correct recognition rate for CSOM using SOM modules with cylindrical architecture for the simple partition [%]

Feature selection	Number of neurons per SOM module						
	1	3	4	6	8	9	10
PCA	99.5	**100**	99.5	98.5	98.5	98	97.5
LDA	**100**	**100**	**100**	99.5	99.5	99.5	99

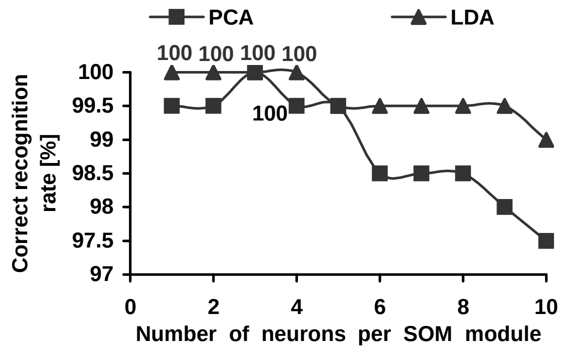

Fig 7 - Correct recognition rate for the simple partition, CSOM using SOM modules with cylindrical architecture

Table 5 – Correct recognition rate for CSOM using SOM modules with toroidal architecture for the difficult partition [%]

Feature selection	Number of neurons per SOM module						
	1	3	4	6	8	9	10
PCA	71	73	72.5	**75.5**	74.5	75	**75.5**
LDA	72	72	72	72	72	73.5	72.5

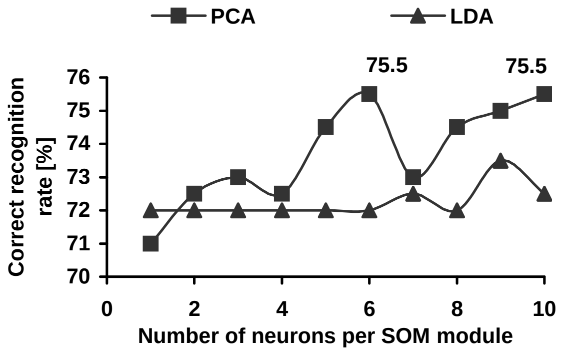

Fig 8 - Correct recognition rate for the difficult partition, CSOM using SOM modules with toroidal architecture

Table 6 – Correct recognition rate for CSOM using SOM modules with toroidal architecture for the simple partition [%]

Feature selection	Number of neurons per SOM module						
	1	3	4	6	8	9	10
PCA	99.5	**100**	99.5	99.5	99.5	99.5	99.5
LDA	**100**	**100**	**100**	**100**	**100**	**100**	99

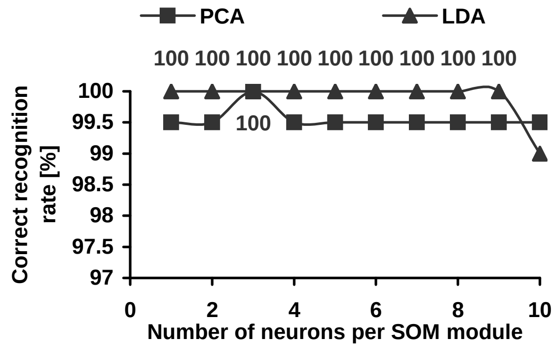

Fig 9 - Correct recognition rate for the simple partition, CSOM using SOM modules with cylindrical architecture

Table 7 – Correct recognition rate, CSOM – PCA for the difficult partition [%]

Architecture of SOM modules	Number of neurons per SOM module						
	1	3	4	6	8	9	10
Rectangular	71	73.5	75	72.5	74	74	73.5
Cylindrical	71	72.5	74.5	73	74	73	73.5
Toroidal	71	73	72.5	**75.5**	74.5	75	**75.5**

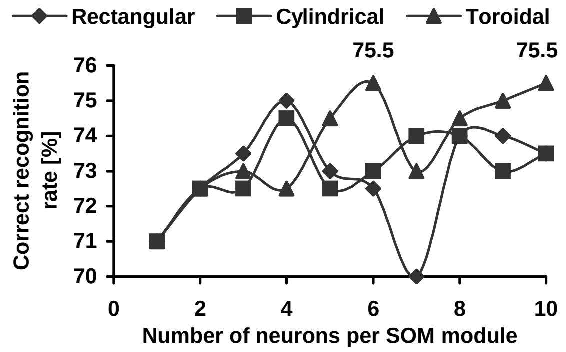

Fig 10 - Correct recognition rate for the difficult partition, CSOM – PCA

Table 8 – Correct recognition rate, CSOM – PCA for the simple partition [%]

Architecture of SOM modules	Number of neurons per SOM module						
	1	3	4	6	8	9	10
Rectangular	99.5	**100**	99.5	99	98.5	99	97
Cylindrical	99.5	**100**	99.5	98.5	98.5	98	97.5
Toroidal	99.5	**100**	99.5	99.5	99.5	99.5	99.5

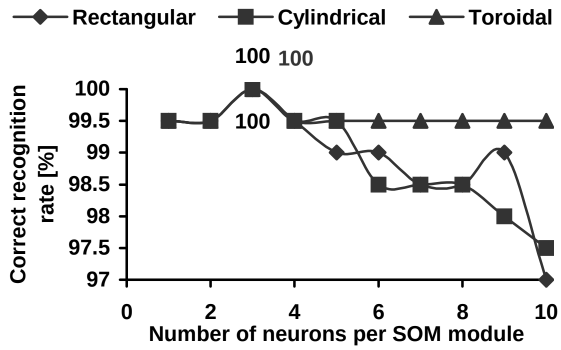

Fig 11 - Correct recognition rate for the simple partition, CSOM – PCA

Table 9 – Correct recognition rate, CSOM – LDA for the difficult partition [%]

Architecture of SOM modules	Number of neurons per SOM module						
	1	3	5	6	8	9	10
Rectangular	72	72.5	73	73	72.5	**75**	73.5
Cylindrical	72	73	73	73	72.5	74.5	73
Toroidal	72	72	72	72	72	73.5	72.5

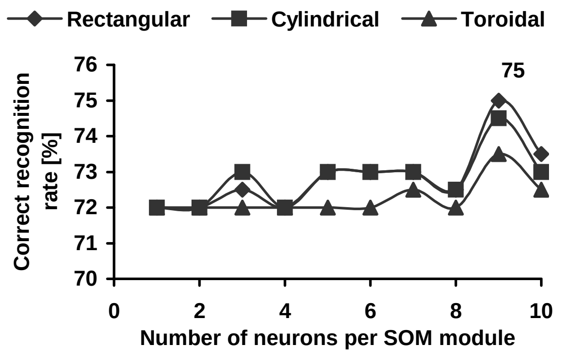

Fig 12 - Correct recognition rate for the difficult partition, CSOM – LDA

Table 10 – Correct recognition rate, CSOM – PCA for the simple partition [%]

Architecture of SOM modules	Number of neurons per SOM module						
	1	3	4	6	8	9	10
Rectangular	**100**	**100**	**100**	99.5	99.5	99	99
Cylindrical	**100**	**100**	**100**	99.5	99.5	99.5	99
Toroidal	**100**	**100**	**100**	**100**	**100**	**100**	99

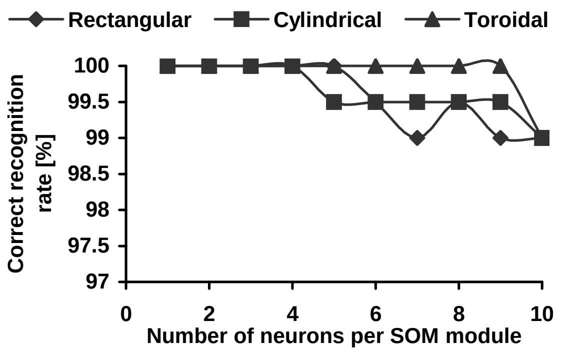

Fig 13 - Correct recognition rate for the simple partition, CSOM – LDA

We compared the best results of the CSOM classifier with two other standard neural classifiers, Multi-Layer Perceptron (MLP) network and Radial Basis Function (RBF) network, using the same feature selection method. In the MLP case, the Bayesian Regulation Backpropagation is the algorithm used for training. The number of neurons in the intermediate layer is 24 neurons and the number of epochs is 500. The RBF network has a number of neurons in the intermediate layer equal with the number of training vectors (200); the network itself establishes the number of neurons.

Table 11 – Correct recognition rate, neural classification for the difficult partition [%]

Neural classifier	Feature selection	
	PCA	LDA
CSOM (top results)	**75.5**	**75**
MLP	73.5	73.5
RBF	61.5	56

Both MLP and RBF are implemented by Matlab® toolbox. The CSOM implementation is also made in Matlab and it is based on Helsinki SOM Toolbox [8].

V. Concluding Remarks

1. The cascade LDA – CSOM is an interesting approach for face recognition, combining the discriminative power of the LDA with the good classification performance of the CSOM neural classifier.
2. For the considered experiments, the LDA leads to slightly better results than PCA in the case of the simple partition, while for the difficult partition PCA is slightly better than LDA.
3. For the simple partition, we obtained 100% correct recognition rate for both PCA and LDA cascaded with CSOM, with more maximum results for LDA. Moreover, for LDA, we obtained maximum recognition rate for one neuron only per SOM module (for any architecture).
4. The top result of the recognition rate for the difficult partition is 75.5%, obtained for the cascade PCA – CSOM, while the LDA – CSOM leads to 75%.
5. In three cases out of four the toroidal architecture of the SOM modules behaves best, while the rectangular architecture does best in one case.
6. By comparing the CSOM face recognition performances to those obtained using other standard neural networks as MLP and RBF, one obtains best results for CSOM in both cases of PCA and LDA feature selection. This is why we have chosen CSOM as a classifier in our cascade.
7. The training time of the CSOM is radically smaller than the training time of the MLP network.
8. These concluding remarks are valid for the two partitions obtained from ORL face database. The performance comparison problem of PCA vs LDA still remains for other applications.

REFERENCES

[1] V. Neagoe, "Concurrent Self-Organizing Maps for Automatic Face Recognition" In: *Proc. of the 29th International Conference of the Romanian Technical Military Academy*, November 15-16, 2001, Bucharest, Romania. Section 9 (Communications). ISBN: 973-8290-27-9. (2001) pp. 35-40.
[2] V. Neagoe and A. Ropot, "Concurrent Self-Organizing Maps for Pattern Classification", *Proc. of First IEEE International Conference on Cognitive Informatics, ICCI 2002*, 19-20 August 2002, Calgary, Alberta, Canada, ISBN 0-7695-1726-9, pp. 304-312.
[3] V. Neagoe and A. Mugioiu, "A fully neural approach to color facial image recognition", *Proc. of the World Automation Congress, 2008 (WAC 2008), International Symposium on Soft Computing in Industry (ISSCI'08),*Sept. 28–Oct. 2, 2008, Hawaii, USA, ISBN:978-1-889335-38-4, IEEE Catalog No. 08EX2476.
[4] V. Neagoe and A. Ropot, "Concurrent Self-Organizing Maps – a Powerful Artificial Neural Tool for Biometric Technology", *Proc. of World Automation Congress WAC '04*, Seville, 3, (2004), ISBN 1-889335-20-7, IEEE Catalog 04EX832C
[5] W. Zhao, R. Chellappa and A. Krishnaswamy, "Discriminant Analysis of Principal Components for Face Recognition," pp.336, Third IEEE International Conference on Automatic Face and Gesture Recognition (FG'98), 1998.
[6] Fisher R. A. (1936). "The Use of Multiple Measurements in Taxonomic Problems". *Annals of Eugenics* 7: 179–188.
[7] Yambor W.S., "Analysis of PCA-Based and Fisher Discriminant-Based Image Recognition Algorithms", Technical Report CS-00-103, Computer Science Department, Colorado State University, July 2000.
[8] http://www.cis.hut.fi/projects/somtoolbox/