Facial Detection based on PCA and Adaptive Resonance Theory 2A Neural Network

M. P. Mahyabadi, H. Soltanizadeh, Sh. B. Shokouhi
College of Electrical Engineering
Iran University of Science and Technology
Pmahyabadi@iust.ac.ir, H_Soltanizadeh@iust.ac.ir, Bshokouhi@iust.ac.ir

Abstract

The prominent feature of Adaptive Resonance Theory neural network is its ability to cluster arbitrary number of input patterns. In this paper, we investigate the performance of ART-2A acting as a classifier in a face recognition system. The Olivetti-Oracle Research Lab (ORL) database of 400 facial images of 40 subjects is used for training and evaluation, and the performance of this network is compared to the MLP and RBF counterparts. The Principal Components of facial images have been used as inputs for the proposed networks. The experimental results reveal that the ART-2A network offers better recognition accuracy even when the illumination of the images is varied or they are corrupted by additive random noise. Moreover the training procedure of the ART-2A is much less time consuming than the other neural networks and its adaptation is fast while introducing a new sample.

Introduction

Face recognition is of substantial interest for a variety of applications, such as expert identification, witness face recognition, bank/store/court and security department, etc. It refers to the process of labeling an image as belonging to a certain image class.

In pattern recognition in general and in face recognition in particular, the most popular and necessary problem is dimensionality reduction. Instead of using N intensity values for an N pixel image, it is generally feasible to specify an image by a set of M features, where $M \ll N$. The selected features must be able to uniquely represent the right class for their corresponding facial images.

Principal component analysis (PCA), one of the most thoroughly investigated approaches to face recognition, is a powerful statistical technique to extract specific facial components, with dimensionality reduction. It decreases the vector's dimension with assurance that the data loss would be the lowest [1-4]. It is also known as Karhunen-Loejve expansion, eigenpicture, eigenvector, and eigenface. It can estimate the coordinates of subspace to represent images as a few components as possible by following the maximization of second order statistics.

The other point in pattern recognition is classifier's type; a reliable, robust and fast system is more desired. The attractiveness of using neural network could be due to its nonlinearity in the network. In the field of competitive neural networks, a key problem is that they don't always form stable clusters. If the number of input patterns is not too large, or if the input patterns do not form too many clusters, then the learning eventually stabilizes. However, the standard competitive networks do not have stable learning in response to arbitrary input patterns. The learning instability occurs because of network's adaptability (or plasticity) which causes prior learning to be eroded by more recent learning. This problem has been solved by Adaptive Resonance Theory (ART), a family of self organizing [5-7]. It is capable to cluster arbitrary sequence of input patterns into stable reorganization codes. Some extensions of ART neural network are ART1, ART2, ART3, ARTMAP, FUZZY ART. ART1 algorithm is exclusively designed for binary input patterns, while the others accept continuous valid vectors. ART3 introduced a more sophisticated biological model for the reset required for ART. The ARTMAP architecture consists of two ART modules that are connected by an inter-ART associative memory. This net has been modified to incorporate FUZZY logic, FUZZY ART. However, the specific ART2 neural networks, such as ART 2A, ART 2A-C and ART 2A-E are also extended [5].

This work represents a facial recognition system consisting of a PCA stage which feeds the principal components of facial image to the networks, ART 2A, MLP [7-9] and RBF [4], [8-11]. The main concern is to compare clustering performance and robustness to illumination variations and noise of ART with the other networks.

The paper describes

- The design and underlying principles of ART Network;
- ART-2A algorithm;
- The experimental results of alternative neural classifiers (MLP, RBF & ART 2A);
- Comparing the performance of networks when the illumination of the images is varied or they are corrupted by additive random noise.

The ART-2A Network

The ART networks are designed to allow the user to control the degree of similarity of patterns placed on the same cluster. The resulting of the number of clusters then depends on the distance between all input patterns, presented to the network during training periods. The similarity parameter, vigilance ρ is typically limited to the range [0, 1]. If the degree of similarity between current input pattern and best fitting prototype J is at least as high as vigilance ρ, this prototype is chosen to represent the cluster containing the input. In contrast, if the degree of similarity between current input pattern and best fitting prototype does not fit into the vigilance interval [ρ, 1], a new cluster has to be installed, where the current input is most commonly used as the first prototype or cluster center.

These nets cluster inputs by using unsupervised learning which occurs in a set of feedback connections from layer 2 to layer 1. Input patterns may be presented in any order. Each time a pattern is presented, an appropriate cluster unit is chosen and that cluster's weights are adjusted to let the cluster unit learn the pattern. Some excellent features of ART, besides its clustering capabilities, are performance, economic usage of memory and temporal of stored knowledge.

Basic processing module of ART networks is an extended competitive learning network, as shown in Fig.1 [5], [7]. The m neurons of an input layer F_1 register values of an input pattern $\boldsymbol{I} = (\boldsymbol{i}_1, \boldsymbol{i}_2, ..., \boldsymbol{i}_m)$. Every neuron of output layer F_2 receives a bottom-up net activity t_j, built from all F_1-outputs S = I. The vector elements of $\boldsymbol{T} = (\boldsymbol{t}_1, \boldsymbol{t}_2, ..., \boldsymbol{t}_n)$ can be perceived as the results of comparison between input pattern I and prototypes $W_1 = (\mathrm{w}_{11}, \ldots, \mathrm{w}_{1\mathrm{m}})$, …,

$W_n = (w_{n1}, \dots, w_{nm})$. These prototypes are stored in the synaptic weights of the connections between F_1 - and F_2 -neurons.

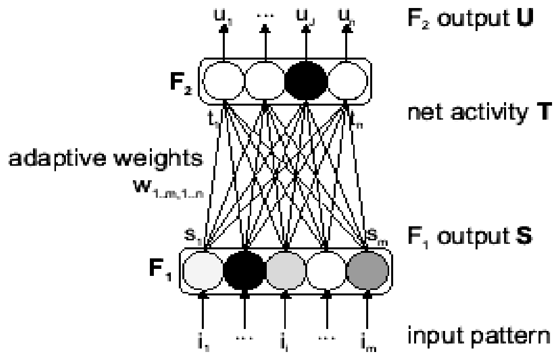


Figure 1: Basic structure of the ART network

The ART-2A network can be characterized by its preprocessing, choice, match and adaptation where choice and match define the search circuit for a fitting prototype [5]. The initial weights for the ART2 net are as follows;
Initial top-down weights must be small to ensure that no reset will occur for the first pattern placed on a cluster unit. The norm of the initial bottom-up weights must be less than the norm of the weights after training, to prevent the possibility of a new winner being chosen during resonance as the weights change. Larger values of bottom-up weights encourage the net to form more clusters.
The central functions of the ART 2A- algorithm are [5]:

- Preprocessing

No negative input values are allowed and all uncoded input vectors A are normalized to unit Euclidean length, denoted by function symbol $\boldsymbol{N}$ as follows,

$$\boldsymbol{I} = \boldsymbol{N}(\boldsymbol{A}) = \frac{\boldsymbol{A}}{\sqrt{\sum_{\boldsymbol{i}=1}^{\boldsymbol{m}} \boldsymbol{a}_i^2}} = \frac{\boldsymbol{A}}{\|\boldsymbol{A}\|} \tag{1}$$

$$\boldsymbol{a}_i \geq 0 \ \forall \boldsymbol{i}, \quad \|\boldsymbol{A}\| > 0$$

Carpenter and Grossberg [4], [6] suggested an additional method of noise suppression to contrast enhancement by setting all input values to zero, which do not exceed a certain bias θ as defined by,

$$\boldsymbol{I} = \boldsymbol{N}\left(\boldsymbol{F}\left(\boldsymbol{N}\left(\boldsymbol{A}\right)\right)\right)$$

$$\boldsymbol{F}\left(\boldsymbol{X}\right)_i = \begin{cases} \boldsymbol{x}_i & \boldsymbol{if} \ \ \boldsymbol{x}_i > \theta \\ 0 & \boldsymbol{else} \end{cases} \quad , \quad \left\|\boldsymbol{F}\left(\boldsymbol{N}\left(\boldsymbol{A}\right)\right)\right\| > 0 \tag{2}$$

This kind of contrast enhancement does only make sense if characteristic features of input patterns, leading to a distribution on different clusters, are coded exclusively in their highest values. With θ bounded by

$$0 \leq \theta \leq \frac{1}{\sqrt{m}} \tag{3}$$

The upper limit will lead to complete suppression of all patterns having the same constant value for all elements.

• Choice

Bottom-up net activities, leading to the choice of a prototype, are determined by

$$t_j = \begin{cases} I.W_j & \text{if } j \text{ indicates a} \\ & \text{committed prototype} \\ \alpha.\sum_{i=1}^{m} i_i & \text{other wise} \end{cases} \tag{4}$$

$$0 \le \alpha \le \frac{1}{\sqrt{m}} \tag{5}$$

Bottom-up net activities are determined differently for previously committed and uncommitted prototypes. The choice parameter $\alpha \ge 0$ again defines the maximum depth of search for a fitting cluster. With $\alpha = 0$, all committed prototypes are checked before an uncommitted prototype is chosen as winner. The simulations in this paper apply $\alpha = 0$.

• Match

Resonance and adaptation occurs either if J is the index of an uncommitted prototype or if J is a committed prototype and

$$\rho \le IW_J = t_J \tag{6}$$

• Adaptation

Adaptation of the final winning prototype requires a shift towards the current input pattern,

$$W_J^{(new)} = N\left(\eta I + (1-\eta) W_J^{(old)}\right) \qquad 0 \le \eta \le 1 \tag{7}$$

ART 2A-type networks always use fast-commit slow-recode mode. Therefore the learning rate is set to $\eta = 1$ if J is an uncommitted prototype and to lower values for further adaptation.

Since match and choice do not evaluate the values of un- committed prototypes, there is no need to initialize them with specific values. ART 2A-related networks should not be used in fast learning mode with $\eta \cong 1$, because proto-types then begin to `jump' between all patterns assigned to their cluster, instead of converging towards their mean.

Experiment Results

The used data base for training and evaluation contains of a set of 400 facial images of 40 subjects which is constructed under various depth and plane rotation. The images are acquired in different emotional expression. The images are 93*112 pixels in size with 256 gray levels per pixel. Some examples of the images with different levels of illumination of are depicted in Fig.2b and Fig.2c. They are also corrupted by different percentage of additive random noise (Fig2.d).

Each experiment consists of four steps: Preprocessing, generation of eigenfaces, training and testing the classifiers. The first step consists of size equalization, noise suppression, putting more emphasis on the desired feature of the image. Olivetti face dataset is composed of 10 face images per a person, so in the second step, the training set consists of seven randomly selected images of each individual. The other three images are defined as testing sets in turn. Then the PCA are used both to decrease the vector's dimension and to extract the eigenfaces. In the third step, the input vectors, after being projected onto the face space, are presented to all classifiers that should be

trained. In the eventual step, the performances of the classifiers are evaluated. Each test image is projected onto eigenfaces, obtained in the second step, and fed to each classifier.
In addition, the images, corrupted by different percentage of additive salt-pepper noise, are presented to each classifier. The noise percentage is defined as the ratio of the number of noisy pixels to the total number of image pixels.

$$Noise\ Ratio = \frac{the\ number\ of\ \ noisy\ \ pixels}{the\ number\ of\ \ image\ pixels} \tag{8}$$

Furthermore the illumination of the images is varied and the robustness to illumination variations of ART is compared to the other networks. The illumination of image is varied by,

$$g_{new} = g_{old} \pm \Delta I \times 256 \tag{9}$$

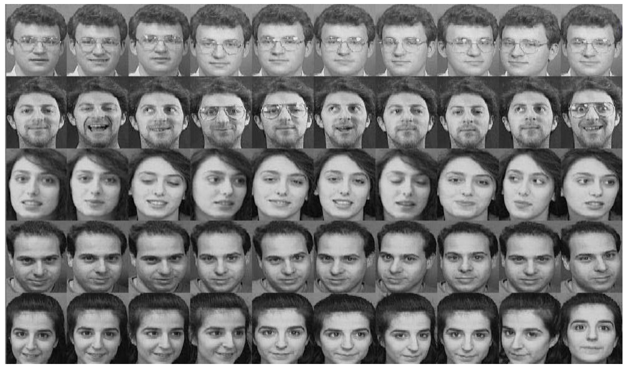
Where g is the gray level of pixels, and ΔI is the percentage of illumination variation and thus, the image is darkened or brightened.
The results of experiments are summarized in table1. It is found that there is considerably difference in recognition rate accuracy when using three networks in the recognition. In addition, the ART-2A is surprisingly robust to high level additional noise, so the noise suppression is considered in the ART-2A algorithm (Eq.2) and, it is the one of the outstanding features of ART-2A besides its high accuracy and clustering capabilities. Also the MLP and RBF networks are very sensitive to illumination and darkening variations of the images, whereas the ART-2A is surprisingly robust to the illumination variations with high level.
In addition, it has become evident that there are obvious differences in time consumption of the training procedure when using three networks. The time consumption of the ART-2A is about few seconds which is an order of magnitude less than that of the other networks (the approximate value of time consumption is several minuets in these networks). Also the training time of the MLP and RBF neural networks will be increased considerably with the sample adding. Further more, ART-2A will respond to the new introduced sample immediately, while the previously mentioned networks will exhibit less. Hence ART-2A architectures keep stable in all inspected environments and suit for particular applications can be selected.

Conclusion

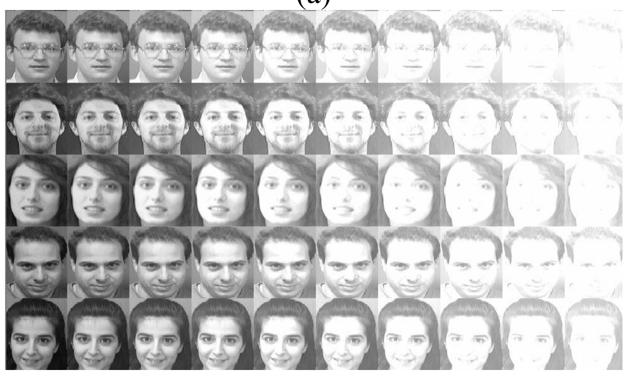
The performance of a face recognition system using PCA for features extraction and ART-2A network as classifier is compared with the performance of the MLP and RBF counterparts. From the results of the recognition experiments, it is certified that the ART-2A is the best classifier when the images are corrupted by additive random noise and the illumination of images are varied. Moreover, the training procedure of the ART-2A is much less time consuming than the other neural networks and its adaptation is fast while introducing a new sample is added.

Table 1: Recognition accuracy and the recognizable additional noise percentage

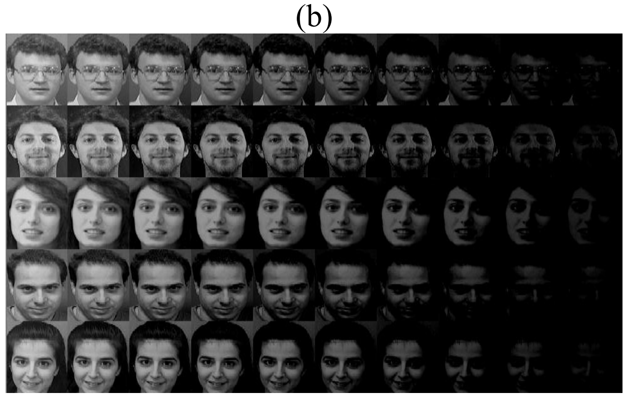
Network Type	MLP	RBF	ART 2A
Recognition Rate (%)	87.6	92.3	97.8
Robustness Accuracy to Additional Noise (%)	10.3	12.1	44.5
Robustness Accuracy to illumination Variations (%)	10.8	14.5	32.7



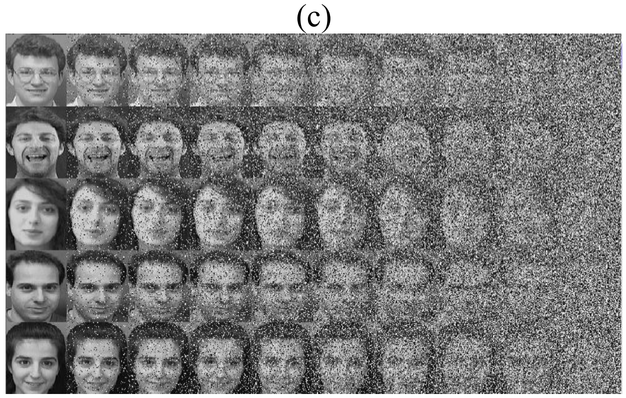
(a)



(b)



(c)



(d)

Figure 2 (a): Some examples of the ORL dataset. (b), (c) the images with different levels of illumination. (d) The images are corrupted with additive noise.

References

[1] C. Conde, A. Ruiz and E. Cabello1, PCA vs. Low Resolution Images in Face Verification, Proceedings of the 12th International Conference on Image Analysis and (ICIAP'03) IEEE, 2003.

[2] S. Romdhani, Face Recognition Using Principal Components Analysis, Ph.D. Thesis, 1997.

[3] J. B. Roseborough and H. Murase, Partial Eigenvalue Decomposition for large Image Sets with Run-Length Coding, Pattern Recognition, Vol. 28, No.3, pp. 421-430, 1995.

[4] C. E. Thomaz, R.Q. Feitosa, A. Veiga, Design of Radial Basis Function Network as Classifier in Face Recognition Using Eigenfaces, Neural Networks, Proceedings. Vth Brazilian Symposium on, pp. 118 – 123, 9-11 Dec. 1998.

[5] T. Frank, K. F. Kraiss, T. Kuhlen, Comparative Analysis of Fuzzy ART and ART_2A Network Clustering Performance, IEEE, Transaction on Neural Networks, Vol. 9, pp. 544 -559, No. 3, 1998.

[6] X. H. Liu, Z. Z. Yu, J. Dunan, L. B. Zhang, M. Lio, Y. C. Liang and C. G. Zhou, Face Recognition Using Adaptive Resonance Theory, Proc. The Second International Conference on Machine Learning and Cybernetics, Xian, 2-5 Nov, 2003.

[7] Laurene Fausett, Fundamentals of Neural Networks, Arechitecture, Algorithms and Applications. 2nd ed, Prentice-Hall International Inc., 1994.

[8] Martin. T Hagan, Howard B.Demuth, Mark Beale., Neural Network Design, Boston: PWS Pub., 1996.

[9] S. Haykin, Neural Networks, Macmillan College P.C, Inc., 1996.

[10] O. F. Romero, A. A. Betanzos, Adaptive Pattern Recognition in the Analysis of Cardiotocographic Records, IEEE Transaction on Neural Networks, Vol. 12, No. 5, September 2001.

[11] A. P. Timoszczuk and E. F. Cabral Jr, RBF Neural Networks and MTI for Text Independent Speaker Identification, Proceedings on Neural Networks, December, 9-11, Brazil, 1998.

Biographies

M. P. Mahyabadi is PhD student at Iran university of Science and Technology. She earned her B.S. (Electronics Engineering, 1996) degree from Bahonar University and M.S. (Electronics Engineering, 1999) degree from Guilan University. Her research is in Image Processing, signal processing, and optoelectronics.

H. Soltanizadeh is PhD student at Iran university of Science and Technology. He earned his B.S. (Electronics Engineering, 1998) degree from Mazandaran University and M.S. (Electronics Engineering, 2002) degree from Iran University of Science and Technology. Her research is in Image Processing, Signal processing, and pattern recognition.

Sh. B. Shokouhi is Associate Professor of Electronics Engineering, He earned his B.S. (Electronics Engineering, 1986), M.S. (Electronics Engineering, 1989) degrees from Iran University of Science and Technology and Ph.D. (Electronics Engineering, 1999) degree from University of Bath. Dr. Shokouhi teaches a variety of courses to postgraduate engineering students. His research is in Image Processing, Machine Vision, Digital and Analog Electronics.