

NEURAL NETWORK BASED SUPERVISED SELF ORGANIZING MAPS FOR FACE RECOGNITION

A.S.Raja¹ and V. JosephRaj²

¹Research Scholar, Sathyabama University, Jeppiar Nagar, Chennai, Tamil Nadu, India
csehod@capeitech.org,

²Professor, Kamaraj College, Thoothukudi, Tamil Nadu, India
v.jose08@gmail.com

ABSTRACT

The word biometrics refers to the use of physiological or biological characteristics of human to recognize and verify the identity of an individual. Face is one of the human biometrics for passive identification with uniqueness and stability. In this manuscript we present a new face based biometric system based on neural networks supervised self organizing maps (SOM). We name our method named SOM-F. We show that the proposed SOM-F method improves the performance and robustness of recognition. We apply the proposed method to a variety of datasets and show the results.

KEYWORDS

Biometrics, Face, Supervised Self Organizing Maps (SOM).

1. INTRODUCTION

The increase of terrorism and other kinds of criminal actions, such as fraud in e-commerce, increased the interest for more powerful and reliable ways to recognize the identity of a person [1, 2]. To this end, the use of behavioural or physiological characteristics, called biometrics, is proposed. Biometrics is best defined as measurable physiological and or behavioural characteristics that can be utilized to verify the identity of an individual [1]. Many physiological characteristics of humans, i.e., biometrics, are typically invariant over time, easy to acquire, and unique to each individual. Therefore the biometrics traits are increasingly adopted for civilian applications and no longer confined for forensic identification.

The recognition of individuals without their full co-operation is in high demand by security and intelligence agencies requiring a robust person identification system. Many face recognition algorithms have been proposed so far [3, 4, 5, 6, 7, 8]. Algorithms related to recognition of face, hand geometry, iris, voice recognition have also been proposed (See Handbook of Biometrics [9]). It is estimated that 5% of the population does not have legible fingerprints [1], a voice could be altered by a cold and face recognition systems are susceptible to changes in ambient light and the pose of the subject.

A typical biometric system usually consists of that specific biometric detection scheme followed by an extraction methodology (which shrinks the dimensionality of useful information) and then a classifier to make the appropriate decision. There are linear approaches such as principal component analysis (PCA) or Eigen faces method [10], independent component analysis (ICA)

[11,12], and linear discriminant analysis (LDA) [13,14] available in literature. Besides there are several nonlinear manifold analysis approaches that used the Kernel method, such as kernel principal component analysis (KPCA) and kernel linear discriminant analysis [15,16, 17]. A survey of methods and the comparative results from literature are presented below in Table 1.

Previous works	Feature extraction method	Database	Correct recognition rate
[18]	FSS or Eigenfaces method	Yale face database and constructed a particular database	FSS is varying from 77.8 to 95.3% Eigenface is varying from 65.6 to 76%
[19]	Eigenfaces method, ICA, or LDA	FERET data set	Varying from 64.94 to 83.85%
[20]	Edge information + traditional PCA + ICA	Indian face database and Asian face database	Varying from 74 to 95%
[21]	Kernel discriminant analysis	FERET, ORL and GT databases	Error recognition rate varying from 43 to 1.5%
[22]	Eigenfaces or fisherfaces	FERET dataset	Error recognition rates ranging from 35 to 1.5%
[23]	Eigenfaces, MLP as a feature extractor, Or SOM network	MIT face database	Ranging from 72.4 to 83.07%
[Present Paper]	Neural Network Based Supervised Self Organizing Maps For Face Recognition	IIT Delhi Database	Ranging from 88.25% to 98.3%

In this paper, we provide a new approach for appearance-based face based recognition that does not require a subject's cooperation called as SOM-F method. We use neural network based supervised self organizing maps for dimensionality reduction as a part of SOM-F method. We show that our proposed SOM-F method improves the performance and robustness of recognition when compared to methods proposed in literature. SOM-F stands for Supervised self organizing maps (SOM) for Face.

The remainder of this paper is organized as follows: In section 2 we discuss an object identification technique suitable for face. In section 3 we explain how SOM is used with Face for recognition. In section 4 we discuss the results obtained using this SOM-F method. Paper concludes with conclusion and future direction.

2. OBJECT DETECTION

We extract the regions of interest using a Haar like features based object detector provided by the open source project OpenCV library [24]. This form of detection system is based on the detection of features that display information about a certain object class to be detected. Haar like features encode the oriented regions whenever they are found, they are calculated similarly to the coefficients in Haar wavelet transformations. These features can be used to detect objects, in this case the human face. The Haar like object detector was originally proposed by Viola and Jones [25] and later extended by Lienhart and Maydt [26].

We used the dataset related to face retrieved from [27] for our experiments. To create the face detector we also used 1000 positive samples and 2500 negative images. The positive images were scaled to a size of 25x25 to reflect the rectangular dimensions of the face. The face detector

worked well with a few falsely detected faces, the problem was overcome by selecting the larger detected object.



Fig. 1.1 Example raw images from the dataset of IIT Delhi [27]



Fig. 1.2 Example images from the dataset of IIT Delhi [27] after face detection.

3. SOM & METHODOLOGY USED IN THIS PAPER

Self Organizing Map (SOM) is a special kind of unsupervised computational neural network [28] that combines both data projection (reduction of the number of attributes or dimensions of the data vectors) and quantization or clustering (reduction of the number of input vectors) of the input space without loss of useful information and the preservation of topological relationships in the output space.

A few concepts are useful to understand the workings of the technique. The input space (also called signal) is the set of input data we employ to feed the algorithm; the set of input data in our case refers to the set of images that we use for training; typically, the observations are multidimensional and are thus expressed by using a vector for each of them. In our case the observations refer to the pixels present in each image (in our case the dimension is $25 \times 25 = 625$; the vector size of each image is 625). On the contrary, the output space (trained network, network or SOM) refers to the low-dimensional universe in which the algorithm represents the input data. It usually has two-dimensions, and is composed of a set of elements called neurons (or nodes)

which are interconnected, hence the network. What the algorithm does is to represent the input space onto the output space, keeping all the relevant information and ordering observations in a way such that topological closeness in the output space implies statistical similarity in the input space.

The input space is composed by n -dimensional vectors we want to visualize/cluster in a low-dimensional environment. We can express the input vector t as:

$$x=[\xi_1(t),\xi_2(t),\dots,\xi_n(t)]^T \in \mathbb{R}^n,$$

where $\xi_i(t)$ represents the value for each dimension.

The output space is an array of x by y neurons (nodes) topologically connected following a kind of geometrical rule (the most common topologies being circles, squares and hexagons). In our case $x = 11$ and $y = 11$. Each of the nodes is assigned a parametric real vector of initially random values that we call model, and express as:

$$m_i=[\mu_{i1},\mu_{i2},\dots,\mu_{in}]^T \in \mathbb{R}^n$$

Last, we may also define as $d(x, m_i)$ any distance metric between two vectors x and m_i . The most widely used is the Euclidean distance, although other specifications are also valid.

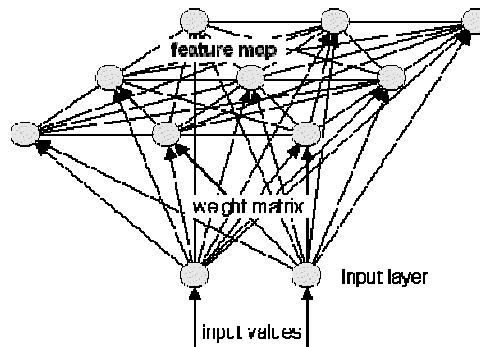


Fig. 2 Self Organizing Map (SOM)

What we are looking for is a topologically-ordered representation of the signal space into the network. That is done by the SOM in an iterative process called training, in which each signal vector is sequentially presented to the output space. The best matching unit (b.m.u.) for x is defined as the neuron minimizing the distance to x . When this is found, the b.m.u. is activated and an adaptive process starts by which such neuron and its topological neighbours are modified by the following scheme:

$$m_i(t+1)=m_i(t)+h_{ci}(t)[x(t)-m_i(t)],$$

where t and $t + 1$ represent, respectively, the initial and the final state after the signal has activated the neuron; $h_{ci}(t)$ is called neighbourhood function and expresses how the b.m.u. and its neighbours are modified when activated by a signal; usually, the linear or Gaussian versions are used. This process is repeated over many cycles before the training is finished. The neighbourhood function depends on several parameters relevant for this stage: the distance between the b.m.u. and the modified neuron (so the further away the neuron is, the smaller the

adjustment); a learning rate $\alpha(t)$ that defines the magnitude of the adjustment, and gradually decreases as the training cycles advance; and the neighbourhood radius, which decides which of the surrounding neurons of the b.m.u. are also modified, and also decreases over the training stage and the self arranging (organization) of the input observations.

This procedure may be used as a visualization tool for multidimensional datasets as well as a clustering method. In the first case, we would want to see how the different observations are mapped into the SOM to discover (dis)similarities, making use of the topological preservation of the statistical characteristics, and study how the different dimensions are distributed; in the second one, the network would have a relatively small number of neurons (as many as clusters we would want to obtain) and we would focus on analysing which observations are grouped with which. In our case, images which have similar face characteristics gets grouped together within the respective nodes/maps.

4. RESULTS

We applied the SOM-F method to around 113 image datasets obtained from IIT Delhi [27] and made lot of studies. The results are shown in diagrammatic form as below:

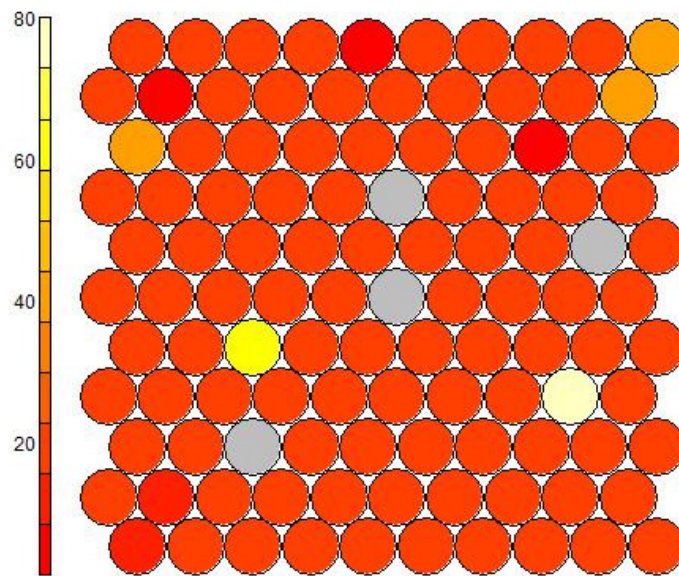


Fig. 3 Counts plot of the map obtained from the face dataset. Empty units are depicted in gray. The colour in each cell represents the number of face images which went into that cell.

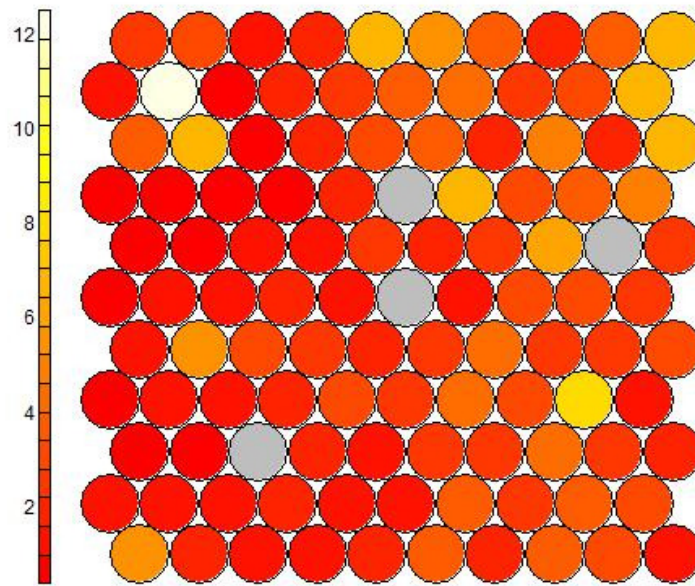


Fig 4. shows the quality of the mapping;

In Fig. 3, the background colour of a unit corresponds to the number of samples mapped to that particular unit; they are reasonably spread out over the map. Only one of the unit is empty: no samples have been mapped to that unit. Fig. 4 shows the mean distance of objects, mapped to a particular unit, to the vector of that unit. A good mapping should show small distances everywhere in the map.

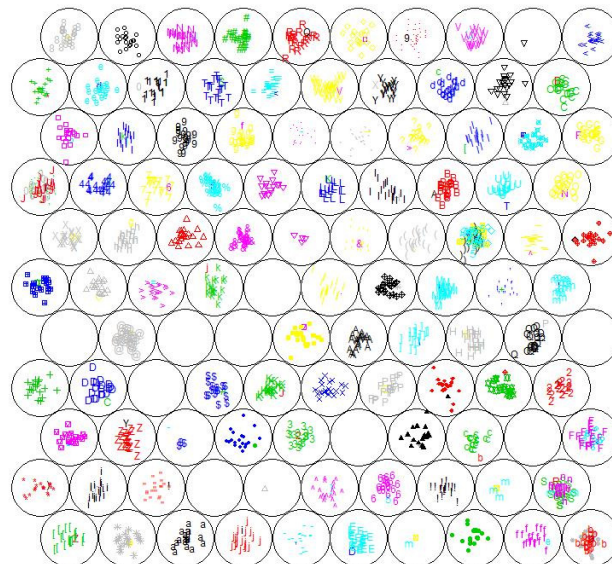


Fig. 5 Mapping of the 113 face images in a eleven-by-eleven SOM

Fig. 5 shows the mapping of images related to SOM. From the dataset, one shall infer that each subject has 20 face images related to him. Figure 5 reveals this out clearly.

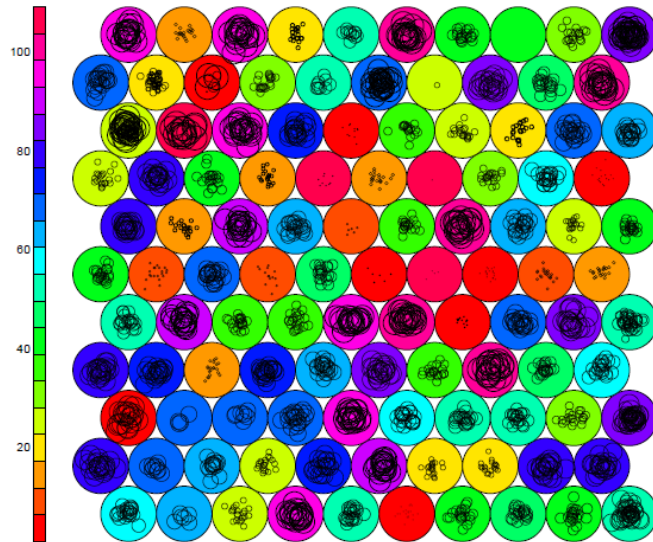


Fig. 6 SOM for the Face data.

Fig.6 shows the mapping of images when the SOM map has been formed; the background colour indicate the person number (from 1 to 113 as explained in the colour palette); the radius of the circles drawn inside a node also reemphasizes this point (face image corresponding to person numbered 1 shall have a smaller radius when compared to the face image corresponding to a person numbered 113);

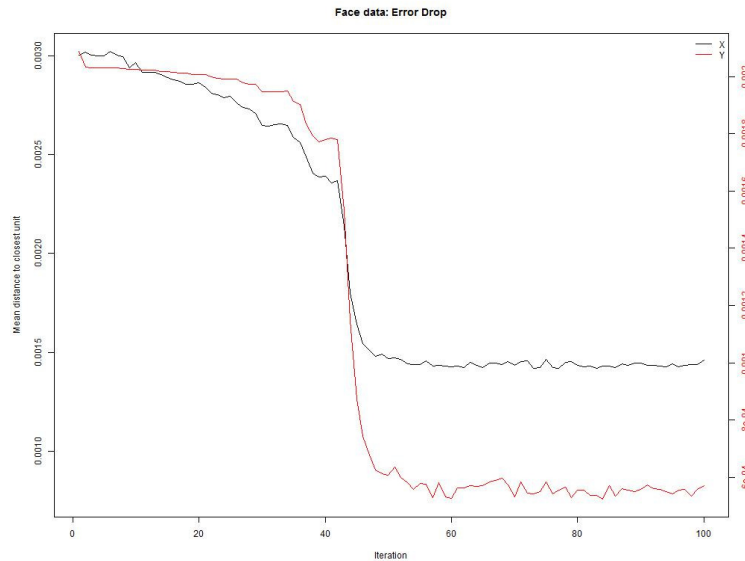


Fig. 7 Training progress, as measured by the average distance of an object with the closest codebook vector unit for face dataset.

During training, the nodes vectors become more and more similar to the closest objects in the dataset. We optimize the parameters by visualisation of this process as shown in Fig. 7. The learning took approximately 9 seconds and the recognition took less than a second.

5. CONCLUSIONS

A new feature extraction metric (named SOM-F) has been proposed. We have shown that our proposed SOM-F method improves the performance and robustness of recognition when compared to methods proposed in literature.

REFERENCES

- [1] Jain AK, Bolle R, Pankanti S (eds) (1999) *Biometrics: Personal identification in networked society*. Kluwer Academic Publishers, Boston/Dordrecht/London
- [2] Wechsler H, Phillips JP, Bruce V, Folgeman Soulie F, Huang TS (eds) (1997) *Face recognition – From theory to applications*. ASI NATO series, vol 163, Springer, Berlin Heidelberg New York.
- [3] Zhao WY, Chellappa R, Rosenfeld A, Philips PJ (2000) *Face recognition: A literature survey*. UMD CfAR technical report CAR-TR-948
- [4] Turk M, Pentland A (1991) Eigenfaces for face recognition. *J Cognitive Neurosci* 3(1):71–86
- [5] Jain A, Flynn P, Ross AA, *Handbook of Biometrics*, Springer, Heidelberg (2008)
- [6] Wiskott L, Fellous JM, Krüger N, von der Malsburg C (1997) Face recognition by elastic bunch graph matching. *IEEE T Pattern Anal* 19(7):775–779
- [7] Kotropoulos CL, Tefas A, Pitas I (2000) Morphological elastic graph matching applied to frontal face authentication under well-controlled and real conditions. *Pattern Recogn* 33(12):1935–1947
- [8] Penev PS, Atick J (1996) Local feature analysis: A general statistical theory for object representation. *Network–Comp Neural* 7(3):477–500
- [9] de Vel O, Aeberhard S (1999) Line-based face recognition under varying pose. *IEEE T Pattern Anal* 21(10):1081–1088
- [10] Alvarado GJ, Pedrycz W, Reformat M, Kwak KC (2006) Deterioration of visual information in face classification using eigenfaces and fisherfaces. *Int J Mach Vis Appl* 17(1):68–82
- [11] Karande KJ, Talbar SN (2009) Independent component analysis of edge information for face recognition. *Int J Image Proc* 3(3): 120–130
- [12] Liu C, Wechsler H (1999) Comparative assessment of independent component analysis (ICA) for face recognition. In: *Proceedings of the 2nd international conference on audio- and video-based biometric person authentication*, pp 211–216
- [13] Li J, Zhao B, Zhang H (2009) Face recognition based on PCA and LDA combination feature extraction. In: *Proceedings of the 1st IEEE international conference on information science and engineering*, pp 1240–1243
- [14] Zhang H, Deng W, Guo J, Yang J (2010) Locality preserving and global discriminant projection with prior information. *Int J Mach Vis Appl* 21:577–585
- [15] Jarillo G, Pedrycz W, Reformat M (2008) Aggregation of classifiers based on image transformations in biometric face recognition. *Int J Mach Vis Appl* 19:125–140
- [16] Lu J, Plataniotis KN, Venetsanopoulos AN (2003) Face recognition using kernel direct discriminant analysis algorithms. *IEEE Trans Neural Netw* 14(1):117–126
- [17] Bach FR, Jordan MI (2002) Kernel independent component analysis. *Int J Mach Learn Res* 3:41–48.
- [18] Shan S, Gao W, Zhao D (2003) Face recognition based on facespecific subspace. *Int J Imag Syst Technol* 13:23–32
- [19] Delac K, Grgic M, Grgic S (2005) Independent comparative study of PCA, ICA, and LDA on the FERET data set. *Int J Imag Syst Technol* 15:252–260
- [20] Karande KJ, Talbar SN (2009) Independent component analysis of edge information for face recognition. *Int J Image Proc* 3(3): 120–130
- [21] Jiang X, Mandal B, Kot A (2009) Complete Discriminant evaluation and feature extraction in kernel space for face recognition. *Int J Mach Vis Appl* 20:35–46
- [22] Alvarado GJ, Pedrycz W, Reformat M, Kwak K (2006) Deterioration of visual information in face classification using eigenfaces and fisherfaces. *Int J Mach Vis Appl* 17(1):68–82

- [23] Oravec M, Pavlovic J (2007) Face Recognition methods based on feedforward neural networks, principal component analysis and self-organizing map. Radio Eng 16(1):51–57
- [24] <http://sourceforge.net/projects/opencvlibrary/>
- [25] Viola, P., Jones, M.: Rapid object detection using boosted cascade of simple features. In: Proceedings of IEEE Computer Vision and Pattern Recognition (2001)
- [26] Lienhart, R., Maydt, J.: An extended set of Haar-like features for rapid object detection. In: Proceedings of IEEE International Conference on ISOMeE Processing, pp. 900–903 (2002)
- [27] Face Database Source Link, <http://www4.comp.polyu.edu.hk/~csajaykr/>
- [28] Fischer, M. M. (2001). Computational neural networks: Tools for spatial data analysis. In M. M. Fischer & Y. Leung (Eds.), Geocomputational modelling: Techniques and applications (pp. 79–102). Heidelberg: Springer.

Authors

Raja. A. S is currently a Research Scholar in Sathyabama University, Jeppiar Nagar, Chennai, Tamil Nadu, India. His recent research interests include Biometrics, Management, Graph Theory etc. He has also been awarded the 'Young IT Professional Award' for the year 2011 by Computer Society of India.



V. Joseph Raj received his PhD degree from M.S. University, Tirunelveli, India and P.G in Anna University, Chennai, India. He is presently working as a Professor in Kamaraj College, Thoothukudi, India. He has also worked as a Professor in European University of Lefke, North Cyprus, Turkey. He is guiding PhD scholars of various Indian and Foreign universities. He has a vast teaching experience of about 20 years. His research interests include neural network and biometrics.

