Face Recognition using SOM Neural Network with Different Facial Feature Extraction Techniques

Nisha Soni
JEC Kukas
Jaipur, India

Mahendra Kumar
Mewar University Gangrar
Chittorgarh, India

Garima Mathur
JEC Kukas
Jaipur, India

ABSTRACT

This paper deals with 3 different techniques for feature extraction of image. Face detection is a necessary first-step in face recognition systems, with the purpose of localizing and extracting the face region from the background. The Self-Organizing Map (SOM) Neural Network has been used for training of database and simulation of FR system. The developed algorithm for the face recognition system formulates an image-based approach, using discrete wavelet transform (DWT), discrete cosine transform (DCT) and Sobel edge detection, simulated in MATLAB. Simulation results are very promising.

General Terms

Face Recognition (FR); Neural Network**;** image processing.

Keywords

Face Recognition (FR); Discrete Cosine Transform (DCT); Discrete Wavelet Transform (DWT); Sobel Edge detection (SED); SOM Neural Network**.**

1. INTRODUCTION

Automatic recognition of people is a challenging problem which has received much attention during the recent years due to its many applications in different fields. Face recognition is one of those challenging problems and up to date, there is no technique that provides a robust solution to all situations and different applications that face recognition may encounter. Face recognition has several characteristics that are advantageous for consumer applications. In addition to, the need for an automatic face recognition system especially at the border control, like airports is becoming very important to strengthen the security. Generally, feature extraction and classification criterion are the two basic operations of any face recognition system. As a result, to improve the recognition performance of such systems one has to enhance these operations. Prior to classification, the features should be extracted from the human face images. Feature extraction in the sense of some linear or nonlinear transforms of the face images with subsequent feature selection is commonly used for reducing the dimensionality of these images so that the extracted features are as representative as possible. However, the problem of extracting features from a human face remains a barrier to apply the practical applications, since it is influenced by the lighting condition, illumination changes, various backgrounds and individual variations. In recent years, several methods for feature extraction have been proposed [1-5]. Of these reported techniques, the principle component analysis (PCA)) [1] and linear discriminant analysis (LDA) [2] are regarded to be the well known feature extraction methods. PCA linearly projects the image space along the direction of maximal variance into a low-dimensional subspace. These methods are statistical linear projection methods which largely rely on the representation of the training samples. The major drawbacks with these techniques are the performance drop of face recognition whenever face appearances are subject to variations by factors such as illumination, expression, pose, accessories and aging. Moreover, they require intensive computations in determining the basis space for a large number of images and when adding new images to a facial image database. On the other hand, multi-resolution techniques have been used for feature extraction in many face recognition systems [5–8]. Among them, the most popular are the discrete wavelet transform (DWT). DWT has been used in various face recognition systems in order to extract multiple sub-band face images. These sub-band images contain coarse approximations of the face as well as horizontal, vertical and diagonal details of faces at various scales. These wavelet-based methods focus on the sub-bands that contain the most relevant information to better represent the face image.

Further we use the discrete cosine transform (DCT) has been employed in face recognition [9]–[12]. The DCT has several advantages over the PCA. First, the DCT is data independent. Second, the DCT can be implemented using a fast algorithm. The discrete cosine transform (DCT) represents an image as a sum of sinusoids of varying magnitudes and frequencies. The DCT has the property that, for a typical image, most of the visually significant information about the image is concentrated in just a few coefficients of the DCT. For this reason, the DCT is often used in image compression applications. For example, the DCT is at the heart of the international standard lossy image compression algorithm known as JPEG. The other one technique is Sobel Edge detection is the process of localizing pixel intensity transitions. The edge detection have been used by object recognition, target tracking, segmentation, and etc. Therefore, the edge detection is one of the most important parts of image processing. The remainder of the paper is organized as follows: In Section 2, Three proposed face recognition method is briefly explained. The SOM neural network system is described in Section 3. Experimental results and conclusions are presented in Sections 4 and 5, respectively.

2. FACE RECOGNITION METHOD

A block diagram of the proposed face recognition system is shown in fig. 1. In this figure the three key steps of face recognition, namely, face detection, facial feature extraction and face recognition are identified. Detailed descriptions of these steps are addressed in the following sub-sections.
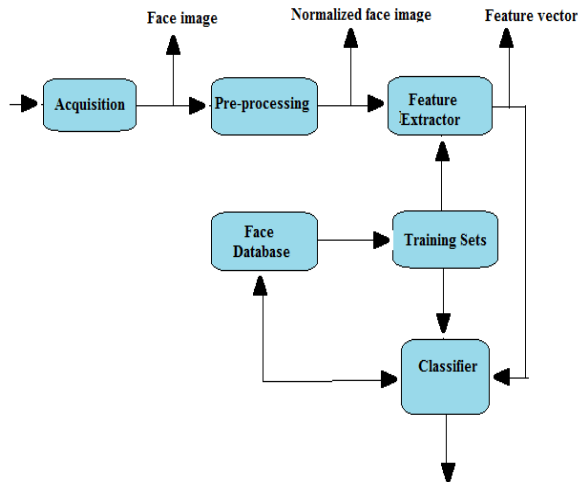
Fig 1: Block diagram of the face recognition system [15]

2.1 Preprocessing

Face detection is the first step in any face recognition system. Many face detection techniques have been proposed in the past decade . They can be classified into geometry-based face detectors and color-based face detectors. Among the geometry-based face detectors, a method examines the triangle-relationship between eye and mouth regions to identify the face region. In addition to, the traditional eye detection methods can be simply and efficiently implemented for frontal face images but can be difficult for complex images. Moreover, skin color has been proven to be an effective image feature for face detection. The automatic face detectors based on skin color are relatively fast and therefore are beneficial for consumer applications such as real-time face recognition embedded in a smart home environment . In the preprocessing stage of the proposed system, a facial region based on skin color detection is cropped from an input image. The obtained facial region is then resized into an 8×8 pixel image to make the face recognition system scale invariant. After then, histogram equalization is applied to enhance the image brightness and contrast.

2.2 Facial Feature Vector Extraction

In search result there are so many technique available for facial extraction .We are dealing with feature extraction methods based which is based on the transforms such as the DWT, DCT and Sobel edge detection also commonly represent the face images with a large set of features. The features of every image stored in our data based are extracted and then stored into the feature vector. Once the feature vectors for all existing images are developed the new database consist of all feature vectors is formed and then stored inside our storage device. To retrieve all images that are similar to the target image, we must extract the features of the target image and compare it with all features vectors. Discrete cosine transform (DCT) is a powerful transform to extract proper features for face recognition. After applying DCT to the entire face images, some of the coefficients are selected to construct feature vectors.

3. SELF-ORGANIZING MAP (SOM) NEURAL NETWORK

3.1 Overview

The self-organizing map also known as a Kohonen Map is a well-known artificial neural network. 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 the "winner"[15]. A SOM network identifies a winning neuron using the same procedure as employed by a competitive layer. However, instead of updating only the winning neuron, all neurons within a certain neighborhood of the winning neuron are updated using the Kohonen Rule. The Kohonen rule allows the weights of a neuron to learn an input vector, and because of this it is useful in recognition applications. Hence, in this system, a SOM is employed to classify Sobel, DWT and DCT-based vectors into groups to identify if the subject in the input image is "present" or "not present" in the image database [13].

3.2 Network Architecture

SOMs can be one-dimensional, two-dimensional or multi-dimensional maps. The number of input connections in a SOM network depends on the number of attributes to be used in the classification [15].
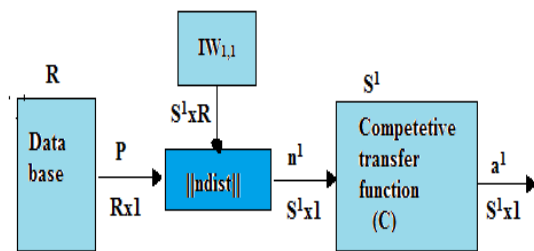

Fig 2: Architecture of a simple SOM Neural Network [15]

The input vector p shown in fig. 2 is the row of pixels of the input compressed image. The ||dist|| box accepts the input vector p and the input weight matrix $IW^{1,1}$, which produces a vector having S^1 elements. The elements are the negative of the distances between the input vector and vectors $_iIW^{1,1}$ formed from the rows of the input weight matrix. The ||dist|| box computes the net input n^1 of a competitive layer by finding the Euclidean distance between input vector p and the weight vectors. The competitive transfer function C accepts a net input vector for a layer and returns neuron outputs of 0 for all neurons except for the winner, the neuron associated with the most positive element of net input n^1. The winner's output is 1. The neuron whose weight vector is closest to the input vector has the least negative net input and, therefore, wins the competition to output a 1. Thus the competitive transfer function C produces a 1 for output element a^1_i corresponding to i^*, the "winner". All other output elements in a^1 are 0[16].

$$n^1 = - \left\| IW_{11} - p \right\| \qquad (3.1)$$

$$a^1 = compet(n^1) \qquad (3.2)$$

Thus, when a vector p is presented, the weights of the winning neuron and its close neighbours move toward p. Consequently, after many presentations, neighbouring neurons learn vectors similar to each other[16]. Hence, the SOM network learns to categorize the input vectors it sees.

The SOM network used here contains N nodes ordered in a two-dimensional lattice structure. In these cases, each node has 2 or 4 neighboring nodes, respectively. Typically, a SOM has a life cycle of three phases: the learning phase, the training phase and the testing phase.

3.3 Unsupervised Learning

During the learning phase, the neuron with weights closest to the input data vector is declared as the winner. Then weights of all of the neurons in the neighborhood of the winning neuron are adjusted by an amount inversely proportional to the Euclidean distance. It clusters and classifies the data set based on the set of attributes used. The learning algorithm is summarized as follows [15]:

1. Initialization: Choose random values for the initial weight vectors $w_j(0)$, the weight vectors being different for $j = 1,2,...,l$ where l is the total number of neurons.

$$w_i = [w_{i1}, w_{i2}, ..., w_{il}]^T \in \Re^n \qquad (3.3)$$

2. Sampling: Draw a sample x from the input space with a certain probability.

$$x = [x_1, x_2, ..., x_l]^T \in \Re^n \qquad (3.4)$$

3. Similarity Matching: Find the best matching (winning) neuron i(x) at time t, $0 < t \le n$ by using the minimum distance Euclidean criterion:

$$i(x) = \arg \min_{j} \left\| x(n) - w_j \right\| \qquad j = 1\ 2...,l \qquad (3.5)$$

4. Updating: Adjust the synaptic weight vector of all neurons by using the update formula:

$$w_j(n+1) = w_j(n) + \eta(n) h_{j,i(x)}(n)(x(n) - w_j(n)) \qquad (3.6)$$

where $\eta(n)$ is the learning rate parameter, and $h_{j,i}x(n)$ is the neighborhood 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.

Training images are mapped into a lower dimension using the SOM network and the weight matrix of each image stored in the training database. During recognition trained images are reconstructed using weight matrices and recognition is through untrained test images using Euclidean distance as the similarity measure. Training and testing for our system was performed using the MATLAB Neural Network Toolbox.

3.4 Training

During the training phase, labeled Sobel-vector, DCT vectors, DWT vectors are presented to the SOM one at a time. For each node, the number of "wins" is recorded along with the label of the input sample. The weight vectors for the nodes are updated as described in the learning phase. By the end of this stage, each node of the SOM has two recorded values: the total number of winning times for subject present in image database, and the total number of winning times for subject not present in image database [17].

3.5 Testing

During the testing phase, each input vector is compared with all nodes of the SOM, and the best match is found based on minimum Euclidean distance, as given in (3.5)[15]. The final output of the system based on its recognition, displays if the test image is "present" or "not present" in the image database.

4. EXPERIMENTAL RESULT

4.1 Image Database

A face image database was created for the purpose of benchmarking the face recognition system. The image database is divided into two subsets, for separate training and testing purposes. During SOM training, 30 images were used, containing six subjects and each subject having 5 images with different facial expressions. fig 3 shows the training and testing image database constructed.
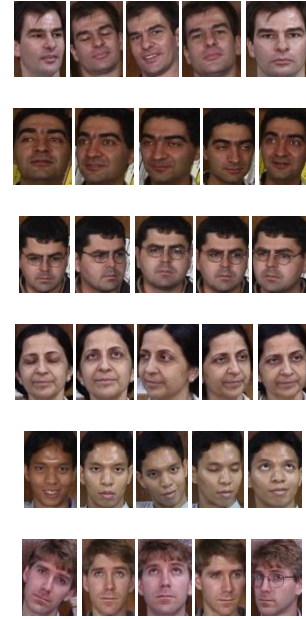

Fig 3: Image database for training

The face recognition system presented in this paper was developed, trained, and tested using MATLAB™ 7.5. The computer was a **Windows** 8 with a 2.40GHz Intel(R) core(TM) i3-3110M processor and 2 GB of RAM.

4.2 Validation of Technique

The pre-processed grayscale images of size 8×8 pixels are reshaped in MATLAB to form a 64×1 array with 64 rows and 1 column for each image. This technique is performed on all 5 test images to form the input data for testing the recognition system. Similarly, the image database for training uses 30 images and forms a matrix of 64×30 with 64 rows and 30

columns. The input vectors defined for the SOM are distributed over a 2D-input space varying over [0 255], which represents intensity levels of the gray scale pixels. These are used to train the SOM with dimensions [64 2], where 64 minimum and 64 maximum values of the pixel intensities are represented for each image sample. The resulting SOM created with these parameters is a single-layer feed forward SOM map with 128 weights and a competitive transfer function. The weight function of this network is the negative of the Euclidean distance [13]. As many as 5 test images are used with the image database for performing the experiments. Training and testing sets were used without any overlapping Fig. 4 shows the result of training and testing simulated in MATLAB using the image database and test input image shown in Fig 3.
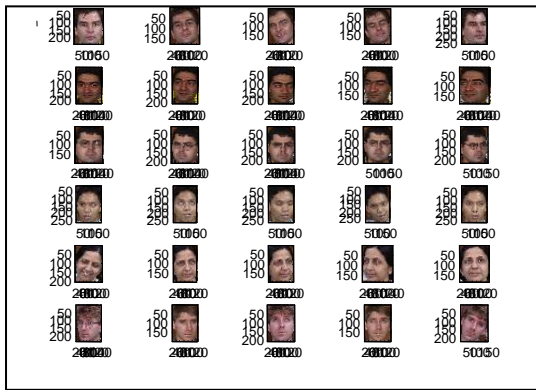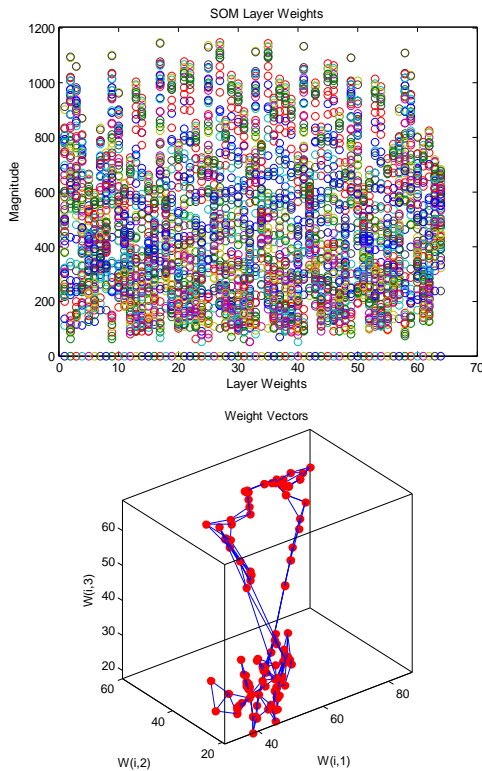

Fig 4: Training & Testing Image Database

Fig 5: (a) SOM Layer Vector (b) Weight Vector

Table 1. Comparison table of face recognition time b/w three methods at different epoch

Image	Epoch	Method ⇨ Time ⇩	DWT	DCT	SED
S01_01	100	Training time	4.6851	5.9940	4.6029
		Execution time	21.9060	115.7445	14.0523
S01_01	500	Training time	24.7147	24.8952	24.4737
		Execution time	44.8569	43.3684	43.3329
S01_01	1000	Training time	46.0579	46.2868	45.6684
		Execution time	58.5632	59.6358	68.7433

Fig 6: Best match image of subject found in training database (a) Face recognition using DWT (b) Face recognition using DCT (c) Face recognition using SOBEL edge detection.

5. CONCLUSION

This paper present's a novel face recognition technique that uses features derived from DCT, DWT, Sobel coefficients, along with a SOM-based classifier. The system was evaluated in MATLAB using an image database of 30 face images, containing six subjects and each subject having 5 images with different facial expressions. After training for approximately 1000 epochs the system achieved a recognition rate of 100% for 5 consecutive trials. A reduced feature space, described for experiment, dramatically reduces the computational requirements of the methods. DWT feature extraction method gives better results compared to DCT and SED methods. This makes our system well suited for high speed, low-cost, real-time hardware implementation.

Table 2. Comparison at 1000 epoch between SED, DCT and DWT

Image	Method	Epoch	Training time	Execution time
S01_01	SED	1000	47.3178	62.1368
S01_02			47.5833	59.7113
S01_03			45.2376	54.5345
S01_04			45.2924	54.2835
S01_05			45.5312	54.6159
S01_01	DCT	1000	47.1712	68.8538
S01_02			49.1798	64.5863
S01_03			50.8347	55.6666
S01_04			49.9102	58.6306
S01_05			49.3563	67.9594
S01_01	1D-DWT	1000	48.7260	95.2298
S01_02			56.5270	74.7419
S01_03			47.1116	71.5741
S01_04			47.0752	59.7149
S01_05			46.3630	77.4150
S01_01	2D-DWT	1000	46.4098	60.6708
S01_02			46.8866	62.0113
S01_03			46.7219	60.4715
S01_04			46.1503	64.3105
S01_05			46.6969	53.0163

6. REFERENCES

[1] M. A. Turk and A. P. Pentland, "Face recognition using eigenfaces," Proc. of IEEE Conference on Computer Vision and Pattern Recognition, pp. 586-591, June 1991.

[2] X. He, S. Yan, Y. Hu, P. Niyogi, and H. Zhang, "Face recognition using Laplacian faces," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 27, no. 3, pp. 328-340, 2005.

[3] Y. Bai, L. Lianwen Jin, and Q. Huang, "Novel face recognition approach based on steerable pyramid feature," IEEE International Conference on ICIP2009, Cairo, Egypt, pp. 4165-4168, 2009.

[4] Z. M. Hafed and M. D. Levine, "Face recognition using the discrete cosine transform," International Journal of Computer Vision, vol. 43, no. 3, pp. 167-188, 2001.

[5] H. K. Ekenel and B. Sankur, "Mult-resolution face recognition," Image and Vision Computing, vol. 23, pp. 469–477, 2005.

[6] C. Garcia, G. Zikos, and G. Tziritas, "Wavelet packet analysis for face recognition," Image and Vision Computing, vol. 18, no. 4, pp. 289–297, 2000.

[7] L. Shen and L. Bai, "A review on Gabor wavelets for face recognition," Pattern Analysis and Applications, vol. 9, no. 2, pp. 273-292, 2006.

[8] D. Kim, I. Jeon, S. Y. Lee, P.K. Rhee, and D. J. Chung, "Embedded face recognition based on fast genetic algorithm for intelligent digital photography," IEEE Transactions on Consumer Electronics, vol. 52, no. 3, August 2006.

[9] D. Koller, and M. Sahami, "Towards optimal feature selection," In ICML1996, Bari, Italy, pp. 87–95, 1996.

[10] M. Raymer, W. Punch, E. Goodman, L. Kuhn, and A. Jain, "Dimensionality reduction using genetic algorithms," IEEE Transactions on Evolutionary Computation, vol. 4, no 2, pp. 164-171, 2000.

[11] C. Liu and H. Wechsler, "Evolutionary pursuit and its application to face recognition," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 22, no. 6, pp. 570-582, 2000.

[12] K. H. Tan and M. Ghanbari "Layered image coding using the DCT pyramid," IEEE Trans. on Image Processing, vol. 4, no. 4, pp. 512-516, April 1995.

[13] AYBAR, E., "Topolojik Kenar _slecleri", Anadolu Üniversitesi, Fen Bilimleri Enstitüsü, Ph.D. thesis, 2003.

[14] Image Toolbox (for use with Matlab) User's Guide, The MathWorks Inc., 2000.

[15] J. Nagi, "Design of an Efficient High-speed Face Recognition System",Department of Electrical and Electronics Engineering, College of Engineering, Universiti Tenaga Nasional, March 2007.

[16] A. Abdallah, M. Abou El-Nasr, and A. Lynn Abbott, "A New Face Detection Technique using 2D DCT and Self Organizing Feature Map" in Proc. of World Academy of Science, Engineering and Technology,Vol. 21, May 2007, pp. 15-19.

[17] Y. Zi Lu and Z. You Wei, "Facial Expression Recognition Based on Wavelet Transform and MLP Neural Network", in Proc. 7th International Conference on Signal Processing, ICSP 2004, Vol. 2, Aug 2004, pp. 1340-1343.