Facial Expression Recognition using Image Processing and Neural Network

Keerti Keshav Kanchi
PG Student, Department of Electronics and Communication Engineering
SDM College of Engineering and Technology
Dharwad-580002, Karnataka, India
Email-id: keertikanchi@gmail.com

Abstract— Face recognition is an important technique and has drawn the attention of many researchers due to its varying applications such as security systems, medical systems, entertainment. A face recognition system is one of the biometric information processing systems. The developed algorithm for the facial expression recognition system, which uses the two-dimensional discrete cosine transform(2D-DCT) for image compression and the self organizing map(SOM) neural network for recognition purpose, simulated in MATLAB. By using 2D-DCT we extract image vectors and these vectors become the input to neural network classifier, which uses self organizing map algorithm to recognize familiar faces (trained) and faces with variations in expression. In this paper we have developed and illustrated a recognition system for human faces using a novel self organizing map based retrieval system. SOM has good feature extracting property due to its topological ordering. The facial analytics result for the 25 images of AT &T database reflects that the face recognition rate using one the neural network algorithm SOM is 95.05% for 5 persons.

Keywords- 2D-Discrete Cosine Transform (2D-DCT), Facial Expression Recognition, Neural Network, Self Organizing Map.

I. INTRODUCTION

Face recognition is one of the most successful applications of pattern recognition and image analysis which has recently received significant attention, especially during the past several years. At least two reasons account for this trend: the first is the wide range of commercial and law enforcement applications, and the second is the availability of feasible technologies after 30 years of research. Face Recognition has gained momentum and practical vitality in the wake of increased and growing security concerns. A facial recognition and face verification system can be considered as a computer application for automatically identifying or verifying a person in a digital image in as much as the processing is carried out on digital still facial images [1]. The facial expression recognition system was introduced in 1978 by Suwa et. Al. The main issue of building a facial expression recognition system is face detection and alignment, image normalization, feature extraction, and classification. There are number of techniques which we use for recognizing the facial expression.

This paper presents a novel approach to recognize face using two dimensional discrete cosine transforms (2D-DCT) and self organize map (SOM) Neural Network as classifier. A block diagram of proposed technique for facial expression recognition using SOM Neural Network is as shown in the fig 1. In the first stage all the 25 face images are compressed for feature processing using two dimensional discrete cosine transform (2D-DCT). When the 2D DCT is applied with a mask, high-coefficients in an image are discarded [1]. Then the 2D IDCT is applied to regenerate the compressed image, which is blurred due to loss of quality and also smaller in size. In next stage uses a self-organizing map (SOM) with an unsupervised learning technique which is trained to classify vectors into groups to recognize if the subject in the input image is “present” or “not present” in the image database. After training all the 25 face images, we now take a single input image, compress it using 2D-DCT and regenerate it using IDCT. All the 25 trained images and untrained input image are simulated. The untrained input image is compared with all trained images, if the face image is classified as present, the best match image is found in the training database using minimum absolute deviation and that image is displayed otherwise if the image is not found then “image is not found in the database” is displayed.

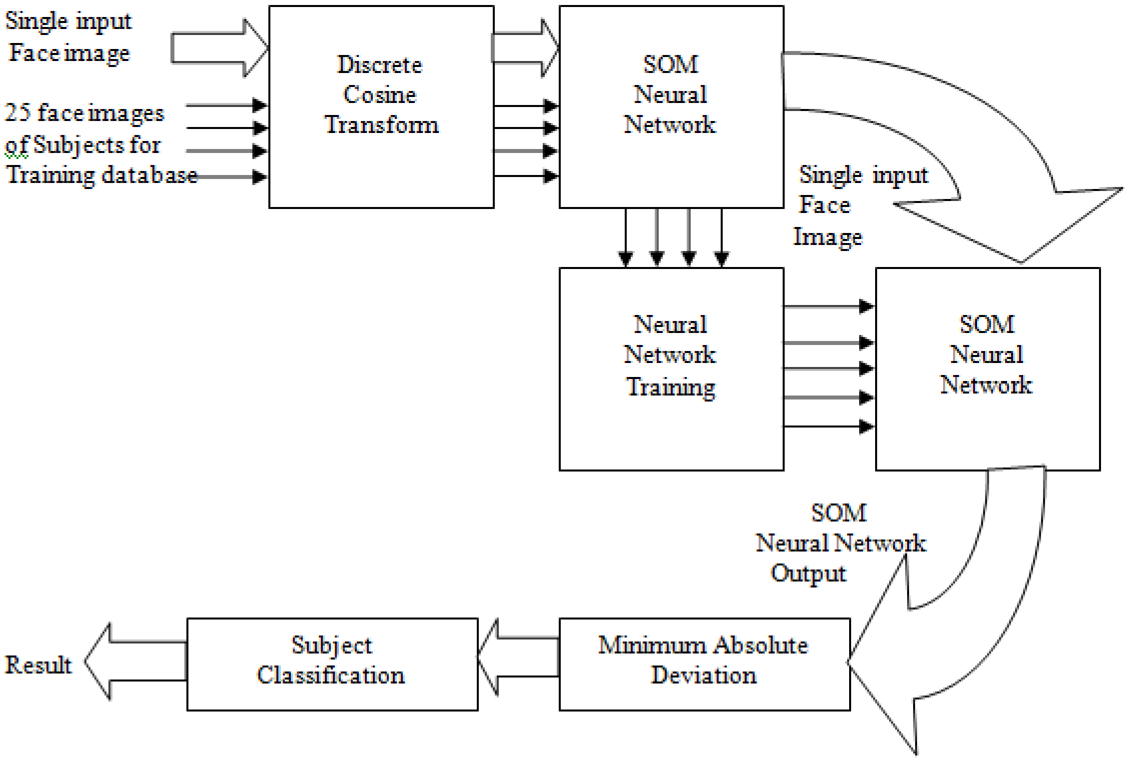

Fig. 1 Block Diagram of Facial Expression Recognition using SOM Neural Network

II. DISCRETE COSINE TRANSFORM

For feature extraction 2-D Discrete Cosine Transform (2D-DCT) is used. A discrete cosine transform (DCT) expresses a sequence of finitely many data points in terms of a sum of cosine functions oscillating at different frequencies. The use of cosine rather than sine functions is critical in these applications: for compression, it turns out that cosine functions are much more efficient, whereas for differential equations the cosines express particular choice of boundary conditions [2]. The DCT, and in particular the DCT-II, is often used in signal and image processing, especially for lossy data compression, because it has a strong "energy compaction". Most of the signal information tends to be concentrated in a few low-frequency components of the DCT.DCT is real valued and provides a better approximation of a signal with fewer coefficients.

The 2D-DCT of an M ×N matrix A is defined as follows:

$$B_{pq} = \alpha_p \alpha_q \sum_{m=0}^{M-1}\sum_{n=0}^{N-1} A_{mn} \cos\left(\frac{\pi(2m+1)p}{2M}\right)\cos\left(\frac{\pi(2n+1)q}{2N}\right), 0 \le p \le M-1, 0 \le q \le N-1 \qquad (2.1)$$

The values B_{pq} are the DCT coefficients. The DCT is an invertible transform, and the 2D-IDCT (2D Inverse-DCT) is defined as follows:

$$A_{mn} = \sum_{p=0}^{M-1}\sum_{q=0}^{N-1} \alpha_p \alpha_q B_{pq} \cos\left(\frac{\pi(2m+1)p}{2M}\right)\cos\left(\frac{\pi(2n+1)q}{2N}\right), 0 \le m \le M-1, 0 \le n \le N-1 \qquad (2.2)$$

The values α_p and α_q in equation (2.1) and (2.2) are given by:

$$\alpha_p = \begin{cases} \sqrt{\frac{1}{M}}, & p = 0 \\ \sqrt{\frac{2}{M}}, & 1 \le p \le M-1 \end{cases} \qquad \alpha_q = \begin{cases} \sqrt{\frac{1}{N}}, & q = 0 \\ \sqrt{\frac{2}{N}}, & 1 \le q \le N-1 \end{cases} \qquad (2.3)$$

The proposed technique uses the DCT transform matrix in the MATLAB Image Processing Toolbox. This technique is efficient for small square inputs such as image blocks of 8 × 8 pixels. The M × M transform matrix T is given by:

$$T_{pq} = \begin{cases} \sqrt{\frac{1}{M}}, \; p = 0, 0 \le q \le M - 1 \\ \sqrt{\frac{2}{M}} \cos \frac{\pi (2q+1)p}{2M}, 1 \le p \le M-1, 0 \le q \le M-1 \end{cases} \tag{2.4}$$

The DCT transformation matrix is implemented in MATLAB and the feature is extracted. The 2D DCT reduces the size of the data significantly by transforming the image from the spatial representation into the frequency domain. The lower frequencies are characterized by relatively larger magnitude while the higher frequencies have smaller magnitudes [3]. The higher frequency components are ignored as it does not significantly affect the accuracy of work. In short, the 2D DCT coefficient of lower frequency values captures the most relevant information of the facial expressions.

A. 2D-DCT Image Compression

The proposed design technique calculates the 2D-DCT of the image blocks of size 8 × 8 pixels using ‘8’ out of the 64 DCT coefficients for masking. The other 56 remaining coefficients are discarded (set to zero). The image is then reconstructed by computing the 2D-IDCT of each block using the DCT transform matrix computation method. Finally, the output is a set of arrays. Each array is of size 8 × 8 pixels and represents a single image [4]. Empirically, the upper left corner of each 2D-DCT matrix contains the most important values, because they correspond to low-frequency components within the processed image block.

III. SELF ORGANIZING MAP

Self organizing map also known as kohonen map is well known artificial neural network. It is unsupervised learning process in which the learning is based upon the input data which is known as unlabeled data and is independent of the desired output data. Self organizing map can also be termed as a topology preserving map [5]. There is a competition among the neurons to be activated and only one neuron that wins the competition is fired and is called the “winner”. Kohonen rule is used to learn the winner neuron and neurons within a certain neighborhood of the winning neuron. This rule allows the weight of neuron to learn an input vector so this makes it perfect for recognition. Hence in this system SOM is used as classifier. The SOM network used in this system contains N nodes ordered in two dimensional lattice architecture and each node has 2 or 3 neighbor nodes. SOM has three phases of life cycle: learning phase, training phase and testing phase.

A. Unsupervised Learning

During the learning the neurons having weight closest with the input vector declare as winner [6]. Based on winning neuron weights of all neighborhood neurons are adjusted by an amount inversely proportional to the Euclidean distance. The learning algorithm is summarized as follows:

1. Initialization: Choose random values for the initial weight vectors $w_j(0)$, the weight vector being different for j = 1, 2,...,l where l is the total number of neurons.
$$w_i = [w_{i1}, w_{i2}, ..., w_{il}]^T \in \Re^n \tag{3.1}$$
2. Sampling: Draw a sample x from the input space with a certain probability.
$$x = [x_1, x_2, ... x_l]^T \in \Re^n \tag{3.2}$$
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:
$$i(x) = \arg \min_j \|x(n) - w_j\|, \; j = 1,2.....l \tag{3.3}$$
4. Updating: Adjust the synaptic weight vector of all neurons by using the update formula:
$$w_j(n+1) = w_j(n) + \eta(n) h_{j,i(x)}(n)(x(n) - w_j(n)) \tag{3.4}$$

Where $\eta(n)$ is learning rate parameter, and $h_{j,i}x(n)$ is the neighbourhood function centred around the winning neuron. Both $\eta(n)$ and are $h_{j,i}x(n)$ varied dynamically during learning for best results.

5. Continue with step 2 until no noticeable changes in the feature map are observed.

B. Training

During the training phase feature vector are presented to the SOM one at a time. For each node net determines the output unit that is best matches for current input sample. The weight vector for the winner is adjusted with respect to the learning algorithm described in learning phase [7]. At the end of this phase each node has two values: total number of winning times for the subject present in the database, and total number of winning times for the subject not present in the database.

C. Testing

During the testing phase firstly every input vector is compared with the all the SOM nodes and then the best match is found based on minimum Euclidean distance as given in equation (3.3). After that final output result is shown.

IV. EXPERIMENTAL WORK

A. Image Database

The training database consists of 25 images, containing five subjects and each subject having 5 individual face images with 5 different facial expressions. The image databases for of 25 subjects with different facial expressions are as shown below:

(a)

(b)

Fig. 2 Training and testing image database (a) Image database for training of 25 subjects with 5 different facial expressions. (b) Untrained image for testing.

B. Validation Technique

All the face images were resized to 8 x 8 pixels and saved; the next step was to compress them by applying the 2D blocked DCT. When the 2D DCT is applied with a mask, high-coefficients in the image are discarded. Then the 2D IDCT is applied to regenerate the compressed image, which is blurred due to loss of quality and also smaller in size. For the image data to be input into the neural network, it should follow the form of only one column, despite the number of rows. Currently, all the resized and DCT compressed face images are in the form of 8 x 8 pixels. Hence the image data needed to be reshaped from an 8 x 8 matrix to a 64 x 1 array for it to be used both for the input and training database of the neural network. SOM's were found to be efficient for image data management and proved to be an accurate closest matching technique of untrained input images with trained database of images. For the design of SOM, a set of 25 image data, 5 different subjects with 5 different facial expressions for the training database was loaded into MATLAB [8]. A SOM was then created and the Parameters for the SOM network were selected to be a minimum and maximum point for each row on vector; training database. There were 64 minimum and 64 maximum points selected altogether. After the SOM neural network was created, it was trained for 2000 epochs. The figure 3 is the SOM layer weight for the 25 face images in the training database and figure 4 is the SOM weight vectors is as shown below:

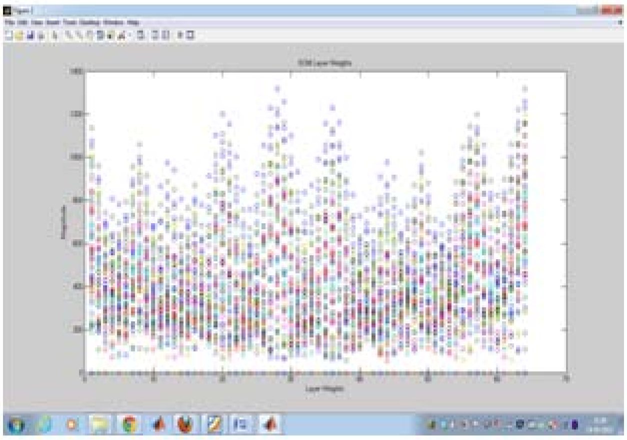

Fig. 3 SOM layer weights for the 25 face images in the training database.

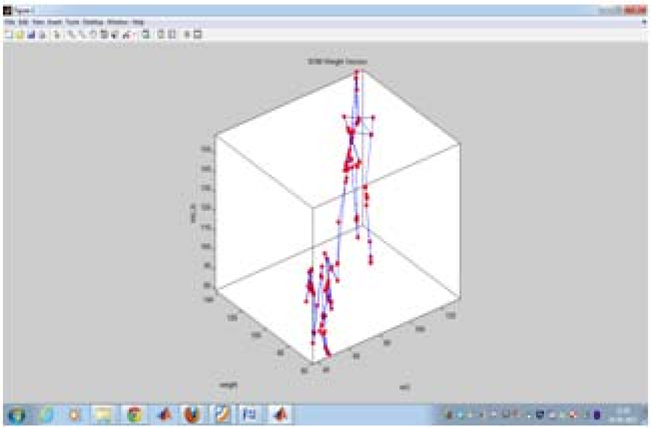

Fig.4 SOM weight vectors

After the SOM neural network was trained and simulated for the 25 images in the training database, the SOM neural network was then simulated for the single input face image. After the SOM neural network is simulated for the input face image, the image in the training database which is the closest match by the SOM neural network for the input face image is found by finding the minimum absolute deviation. After the closest matched training database images are found, they are then classified. Classification of the subject is the answer of the face recognition system.

C. Determining Optimal Number of Epochs for Training

Epochs are neural network training parameters. They are defined as one complete cycle through the neural network for all cases, which present the entire training set to the neural network. Each time a record goes through the net, it is one trial, one sweep of all records is termed as an Epoch. Less number of epochs used for training leads to less training time for the training data set. The goal is to find the optimal number of epochs for training which will produce accurate neural network results and at the same require the least amount of time for program execution. The SOM neural network was tested to determine the optimal number of epochs to be used for neural network training. This test was performed by varying the number of epochs and network training time to find the best possible recognition rate.

Table I shows the tabulated results for the test. The training database for this test consists of 25 face images of 5 subjects, each subject having 5 individual images with different facial expressions.

Table I. Comparison of number of epochs vs. network training time and Recognition rate

Number of epochs	Network Training Time (in sec)	Recognition Rate (%)
50	18.21	58.13
100	40.19	64.54
200	82.43	72.23
300	101.87	79.41
400	115.32	82.25
500	229.19	87.12
700	290.11	90.03
1000	352.45	92.45
1500	567.08	92.89
2000	766.34	95.05

Table I shows the tabulated results generated by varying the number of epochs. The most efficient number of epochs for training and with respect to fastest training time is 2000, with achieves best possible recognition rate as 95.05%.

V. CONCLUSION

This paper has presented a novel facial expression recognition technique that uses features derived from DCT coefficients, along with a self organizing map (SOM)-based classifier. The 2D-DCT and SOM neural network are the heart for the design and implementation of efficient facial expression recognition system. The system was evaluated in MATLAB using an image database of 25 face images, containing five subjects and each subject having 5 images with different facial expressions. After training for approximately 2000 epochs the system achieved a recognition rate of 95.05% for fastest network training time. The system having less computational requirement this make system well suited for low cost, real-time hardware implementation.

VI. ACKNOWLEDGMENT

It is a great pleasure to express sincere and humble gratitude to the guide Dr. Vijaya C for valuable guidance and encouragement given during the course of work. The author would also like to extend sincere thanks to the Principal, HOD and to all teaching and non-teaching staff of Electronics and Communication Engineering department of SDM College of Engineering and Technology, Dharwad, Karnataka, India for their constant support.

REFERENCES

[1] Jawad Nagi, Syed Khaleel Ahmed and Farrukh Nagi, " A MATLAB based Face Recognition System using Image Processing and Neural Networks", 4th International Colloquium on Signal Processing and its Applications, March 7-9, 2008, Kuala Lumpur, Malaysia.

[2] Dinesh Chandra Jain and V. P. Pawar," A Novel Approach For Recognition Of Human Face Automatically Using Neural Network Method", International Journal of Advanced Research in Computer Science and Software Engineering, Volume 2, Issue 1, January 2012.

[3] M.Nandini, P.Bhargavi and G.Raja Sekhar, "Face Recognition Using Neural Networks", International Journal of Scientific and Research Publications, Volume 3, Issue 3, March 2013.

[4] Digital Image Processing - R.C. Gonzalez Richards and Xiuping Jia, enlarged edition, Springer-Verlag, 1999. Woods, Addison Wesley, 1992.

[5] Simon Haykin, "Neural Network-A Comprehensive Foundation", 2nd Ed., Pearson Education, 2004

[6] Steve Lawrence, C. Lee Giles, Ah Chung Tsoi and Andrew D. Back, "Face Recognition: A Convolutional Neural Network Approach", Accepted for publication, IEEE Transactions on Neural Networks, Special Issue on Neural Networks and Pattern Recognition.

[7] Angel Noe Martinez-Gonzalez and Victor Ayala-Ramirez, "Real Time Face Detection Using Neural Networks", 2011 10th Mexican International Conference on Artificial Intelligence.

[8] Abdallah S. Abdallah, A. Lynn Abbott, and Mohamad Abou El-Nasr, "A New Face Detection Technique using 2D DCT and Self Organizing Feature Map", Proceedings Of World Academy Of Science, Engineering And Technology Volume 21 May 2007.