Facial Expression Detection Using Artificial Neural Network

Bhavesh V.Vaghasiya[1] Prof. M.D.Amipara[2]

[2]Principle

[1]Noble engineering college (junagadh) [2]BIET (junagadh)

Abstract— automatic recognition of people is a challenging problem which has received much attention during recent years due to its many applications in different fields. Human Sate Recognition as a facial expression recognition is one of those challenging problems and up to date, there is no technique that provides a robust solution to all situations. In this paper all five universally recognized basic emotions namely angry, disgust, happy, sad and neutral. This paper presents a new technique for facial expression recognition. This technique uses an image-based approach towards artificial intelligence by removing redundant data from face images through image compression using the two-dimensional discrete cosine transform (2D-DCT). The DCT extracts features from face images based on skin color. Feature vectors are constructed by computing DCT coefficients. A self-organizing map (SOM) using an unsupervised learning technique is used to classify DCT-based feature vectors into groups to identify if the subject in the input image is "present" or "not present" in the image database. Facial expression recognition with SOM is carried out by classifying intensity values of grayscale pixels into different groups. Evaluation was performed in MATLAB using an image database of 25 face images, containing five subjects and each subject having 5 images with different basic facial expressions.

Keywords: Facial expression recognition, discrete cosine transform, self-organizing map, neural network, artificial intelligence.

I. INTRODUCTION

During the past several years, human emotional state recognition has attracted a significant interest in the scientific community, and it plays a vital role in human centered interfaces. Many applications, such as virtual reality, video-conferencing, user profiling, and customer satisfaction studies for broadcast and web services, require efficient facial expression recognition in order to achieve the desired results. Therefore, the impact of facial expression recognition on the above-mentioned application areas is constantly growing.

Mehrabian [1] indicates that the verbal part (i.e. spoken words) of a message contributes only for a 7% of the effect of the message, the vocal part (i.e. voice information) contributes for 38% while facial expressions of the speaker contributes for 55% of the effect of the spoken message. Hence in order to develop "Active Human Interface" that realizes heart to heart communication between intelligent machine and human beings.

There are many techniques for facial expression recognition. In this paper technique discussed for expression recognitions Self- Organizing Map technique, which is based on DCT and Neural Network. The remainder of this paper is organized as follows. Section II, describes the different methods used for facial expression recognition. Section III discusses DCT computation on face images. Section IV describes the design and architecture of the SOM neural network. Section V shows experimental results of the system. Section VI gives the conclusion drawn from the experiment performed.

II. DIFFERENT TYPES OF METHOD FOR FACIAL EXPRESSION RECOGNITION

Different techniques have been proposed to classify facial expressions, such as Neural Network [2-4], Support Vector Machine (SVM) [5], Bayesian Network (BN) [6] and rule-based classifiers [7-9]. In Lyons et al.' work [10], the principle components of the feature vectors from training images were analyzed by LDA to form discriminate vectors, and facial image classification was performed by projecting the input vector of a testing image along the discriminate vectors. Cohen et al. compared different Bayes classifiers [11], and Gaussian Tree-Augmented-Naive (TAN) Bayes classifiers performed best. Bartlett et al. [5] performed systematic comparison of different techniques including AdaBoost, SVM and LDA for facial expression recognition.

III. DISCRETE COSINE TRANSFORM

The discrete cosine transform is an algorithm widely used in different applications. The most popular use of the DCT is for data compression, as it forms the basis for the international standard loss image compression algorithm known as JPEG [12]. 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. Extracted DCT coefficients can be used as a type of signature that is useful for recognition tasks.

Face images have high correlation and redundant information which causes computational burden in terms of processing speed and memory utilization. The DCT transforms images from the spatial domain to the frequency domain. Since lower frequencies are more visually significant in an image than higher frequencies, the DCT discards high-frequency coefficients and quantizes the remaining coefficients. This reduces data volume without sacrificing too much image quality [13]. The 2D-DCT of an M × N matrix A is defined as follows:

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

(3.1)

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

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

(3.2)

The values αp and αq in (3.1) and (3.2) are given by:

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

(3.3)

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:

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

(3.4)

Nearest-neighbor interpolation is performed using the MATLAB Image Processing Toolbox to resize preprocessed images from size 512 × 512 pixels to image blocks of size 8 × 8 pixels. 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.

IV. SELF-ORGANIZING MAPS

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" [14]. 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 DCT-based vectors into groups to identify if the subject in the input image is "present" or "not present" in the image database [13]. SOMs can be one-dimensional, two-dimensional or multidimensional maps. The number of input connections in a SOM network depends on the number of attributes to be used in the classification [13].

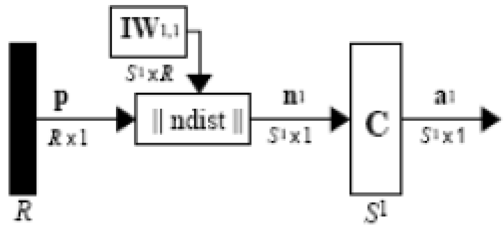

Fig. 1: Architecture of a simple SOM

The input vector p shown in Fig. 6 is the row of pixels of the DCT compressed image. The ||dist|| box accepts the input vector p and the input weight matrix IW1, 1, which produces a vector having S1 elements. The elements are the negative of the distances between the input vector and vectors iIW1, 1 formed from the rows of the input weight matrix. The ||dist|| box computes the net input n1 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 n1. 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 a1 i corresponding to i*, the "winner". All other output elements in a1 are 0[14].

$$n^1 = -\left\| IW_{1,1} - p \right\| \quad (4.1)$$

$$a^1 = compet(n^1) \quad (4.2)$$

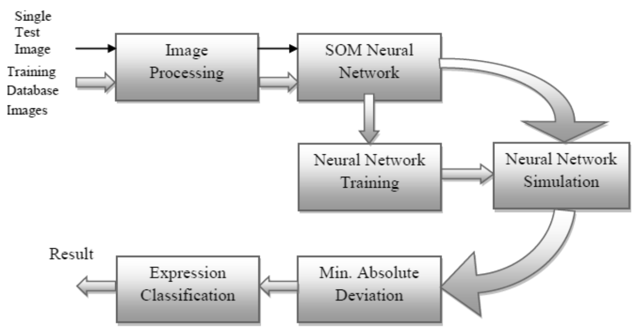

Fig. 2: Purposed Method for SOM

Thus, when a vector p is presented, the weights of the winning neuron and its close neighbour move toward p. Consequently, after many presentations, neighbouring neurons learn vectors similar to each other [8]. 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.

Fig. 2 shows purposed method of SOM based on DCT.

V. EXPERIMENTAL RESULT

This algorithm is implemented on the MATLAB software. For the expression recognition process there two types of data base are used, which are training database and testing database, In Training database several person's photographs are taken with different five expressions.

Output/ Desired	Happy	Sad	Angry	Disgust	Neutral
Happy	100	0	0	0	0
Sad	0	100	0	0	0
Angry	0	0	100	0	0
Disgust	0	0	0	100	0
Neutral	0	0	0	0	100

Table. 1: Confusion matrix for training database using SOM

Output/ Desired	Happy	Sad	Angry	Disgust	Neutral
Happy	97	03	0	0	0
Sad	0	98	0	01	01
Angry	0	00	99	01	0
Disgust	0	02	01	97	0
Neutral	0	01	0	0	99

Table. 2: Confusion matrix for testing database using SOM.

VI. CONCLUSION

This paper has presented a novel face recognition technique that uses features derived from DCT coefficients, along with a SOM-based classifier. 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 1000 epochs the system achieved a recognition rate of 98% for 10 consecutive trials.

REFERENCES

[1] A.Mehrabian, Communication without words. Psychology today, vol.2, no.4, pp.53-56, 1968.
[2] Z. Zhang, M.J. Lyons, M. Schuster, S. Akamatsu, Comparison between geometry-based and Gabor-wavelets-based facial expression recognition using multi-layer perceptron, in: IEEE International Conference on Automatic Face & Gesture Recognition (FG), 1998.
[3] Y. Tian, Evaluation of face resolution for expression analysis, in: CVPR Workshop on Face Processing in Video, 2004.
[4] C. Padgett, G. Cottrell, Representing face images for emotion classification, in: Advances in Neural Information Processing Systems (NIPS), 1997.
[5] M.S. Bartlett, G. Littlewort, M. Frank, C. Lainscsek, I. Fasel, and J. Movellan, Recognizing facial expression: machine learning and application to spontaneous behaviour, in: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2005.
[6] I. Cohen, N. Sebe, A. Garg, L. Chen, T.S. Huang, Facial expression recognition from video sequences: temporal and static modeling, Computer Vision and Image Understanding 91 (2003) 160–187.
[7] M. Pantic, L. Rothkrantz, Expert system for automatic analysis of facial expression, Image and Vision Computing 18 (11) (2000) 881–905.
[8] M. Pantic, L.J.M. Rothkrantz, Facial action recognition for facial expression analysis from static face images, IEEE Transactions on Systems, Man, and Cybernetics 34 (3) (2004) 1449–1461.
[9] M. Pantic, I. Patras, Dynamics of facial expression: recognition of facial actions and their temporal segments from face profile image sequences, IEEE Transactions on Systems, Man, and Cybernetics 36 (2) (2006) 433–449.
[10] M.J. Lyons, J. Budynek, S. Akamatsu, Automatic classification of single facial images, IEEE Transactions on Pattern Analysis and Machine Intelligence 21 (12) (1999) 1357–1362.
[11] I. Cohen, N. Sebe, A. Garg, L. Chen, T.S. Huang, Facial expression recognition from video sequences: temporal and static modelling, Computer Vision and Image Understanding 91 (2003) 160–187.
[12] E. Ifeachor and B. Jervis, Digital Signal Processing: A Practical Approach. Prentice Hall, 2001.
[13] J. Nagi, "Design of an Efficient High-speed Face Recognition System", Department of Electrical and Electronics Engineering, College of Engineering, University Tenaga Nasional, March 2007.
[14] 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.

