IEEE, International Conference on Information, Communication, Instrumentation and Control (ICICIC – 2017)
Paper Id: 347

Facial Expression Recognition using Gabor filter and Multi-layer Artificial Neural Network

Kunika Verma[1], Ajay Khunteta[2]
Department of Electronics Engineering
Rajasthan Technical University
Kota, India.
kunika.verma.77@gamil.com[1], akhunteta@rtu.ac.in[2]

***Abstract* — This paper deciphers the Facial Expression Recognition (FER) using Gabor Filter and Artificial Neural Network (ANN), extracts the facial expression using Gabor filter and then classify the facial expressions using the multi-layer artificial neural network. Recognizing facial expressions of human beings in an image processing by a computer is an interesting and challenging research work. This paper is based on detection and classification of facial emotion expressions. At first, we design an algorithm to detect the face reason image in the whole image using Viola-Jones detection algorithm, then by using Gabor filter extracts the facial features in the spatial domain. The Gabor filter is used to capture the whole frequency spectrum in all directions. And then extract meaningful facial features using Gabor Filter. Finally, they have been successfully classified the facial expressions using the extracted Gabor features of face image used as an input to the Artificial Neural Network classifier. The experimental results on database images of JAFFE show the robustness and better recognition rates of the proposed approach.**

Keywords*— *Facial Expression Recognition (FER); Viola-Jones algorithm; Gabor Filter; Artificial Neural Network (ANN).

I. INTRODUCTION

Research on Facial Expressions and Physiognomy dated back to the early Aristotelian era. Facial expressions play an essential role in human being, non-verbal interpersonal communications. Facial emotions are a state of feeling like thoughts, expressions, the variety of modes and psychological changes. The applications of facial expression recognition have the wide range of areas such as image retrieval, psychological area, video conferencing and defense services, etc. The main purport of facial expression recognition is to introduce a general way of communication in human and computer machine interaction. So, FER has become an active research area in pattern recognition, computer vision, medical applications and artificial intelligent robotics, etc.

The facial expressions are commonly classified in the six basic facial expressions such as Anger, Disgust, Fear, Happiness, Sadness and Surprise and in this paper, we also deliberate 'Neutral'. This paper represents the technique which is used to detect and identify the human facial expressions from digital images from the digital dataset. FER approach can be divided into three main steps so that first is face detection in an image is known as pre-processing and detection, second is facial features extraction which is the method used to represent the facial expressions, and the last one is classification which is a method that classify the seven basic expressions using extracted features.

Li [1] holds that facial expression recognition using Gabor filter and GLCM under partial occlusion and used the K-NN method for classification. Azmi [2] proposed the local Gabor binary pattern histogram sequence (LGBPHS) into the facial expression recognition field and proven that the combined algorithm as higher recognition rate than each single one. Wang [3] introduced the Viola-Jones face detection algorithm this paper describes the whole algorithms of the Viola-Jones method as an integral image, Ada-boost and cascade classifier. Ilonen [4] describes the efficient computation of Gabor features and implementation details of multi-resolution Gabor filtering in both spatial and frequency domains. Kaushal [5] emphasized the face detection technique using neural network and Gabor wavelet transform and described the multi-layers perceptron neural network. The study of various types of facial expression recognition techniques using different and many types of detection, extraction, and classification processes are described in [6, 7]. Stevens [8] described the artificial neural network for the purpose of data pattern recognition. Nazil [9] calculated the facial expression classification using statistical and spatial features like entropy, moment, energy, mean, variance and standard deviation and neural network also. The experiments on Japanese Female Facial Expression (JAFFE) [10] shows the high robustness improve the recognition rates of the proposed method. Many researchers had also used the different-different techniques for facial expression recognition [11, 12, 13, 14 and 15].

II. FACIAL EXPRESSION RECOGNITION (FER)

FER has three steps as follows (Fig. 1):

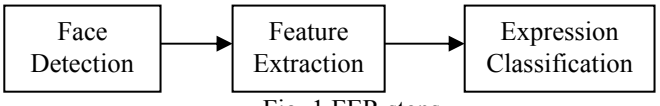


Fig. 1 FER steps

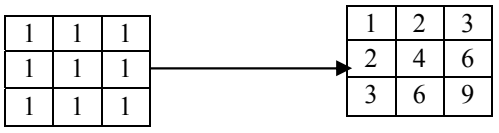
A. Face Detection using Viola-Jones Algorithm

Preprocessing is the first step of face detection. Mostly, at the time of taking a photograph, the face area is not the only area taken but also the background. So, the region of interest (ROI) is only face image on the whole image such as it reduce the processing time and increases the reliability by removing background as unnecessary information. The background and

the hair of the images in JAFFE database do not contribute information in recognizing the expressions [6, 7].

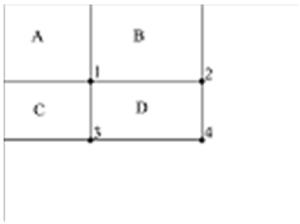
In this paper, use the Viola-Jones face detection algorithm. Viola-Jones projected a face detection system that allows fast feature extraction. The Viola-Jones face detection method has three steps as follows [3]:

1) Integral Image: In Viola-Jones face detection algorithm the input image turns into an integral image (Fig. 2). The integral image is the summation of all pixels above and to the left of the concerned pixels. (Fig. 3).



Input image Integral image

Fig. 2 Integral image construction



Sum within D = (4+1)-(2+3) = (A+B+C+D+A) - (A+B+A+C)

= D

Fig. 3 Sum calculation of the pixels

For feature extraction, the Haar – like features are the rectangular type that is obtained by an integral image. This is calculated by subtracting the sum of white rectangles from the sum of black rectangles (Fig. 4).

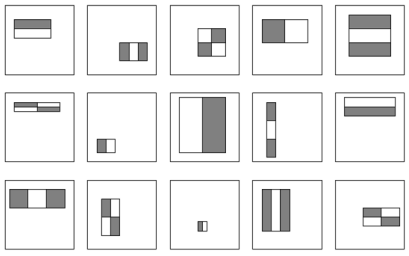


Fig. 4 Haar - like features

2) Adaboost: Features are extracted from subwindows of a sample image. The base size of subwindow is 24 X 24 pixels. Each feature types are scaled and shifted across all possible combinations. In a 24 X 24 pixel subwindow, there are ~160,000 possible features to be calculated. So, we use AdaBoost (Adaptive Boosting), that is machine learning algorithm which helps to find only the best features among all these 160,000+ features. These features are also called weak classifier.

3) Cascade Classifier: The linear combination of 'weak' classifier is used to construct a 'strong' classifier using cascading. The basic principle of the Viola-Jones face detection algorithm is to scan many times through the same image, each time with a new size. If an image contains one or more faces it is obvious that evaluated sub-window should still be non–faces. So the algorithm should concentrate on discarding non-faces quickly and spend more time on probable face regions (Fig. 5).

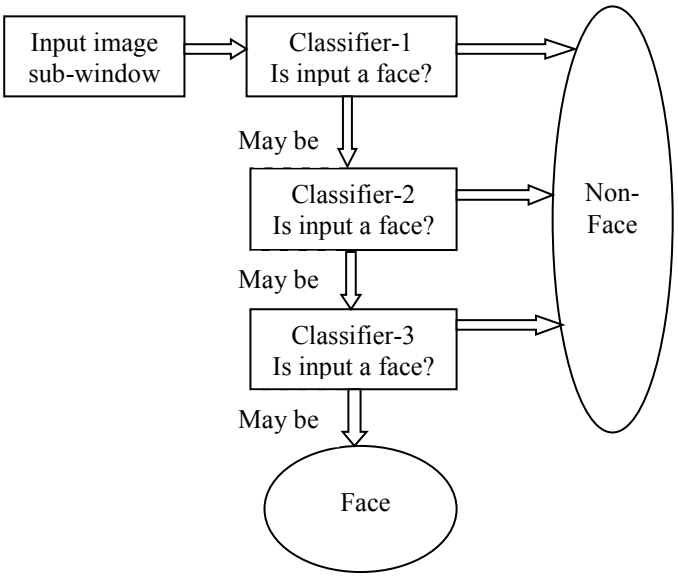


Fig. 5 Cascade classifier

B. Feature Extraction using Gabor Filter

Gabor filter could extract both the time (spatial) and frequency domain. The Gabor filter used in many applications, such as target detection, fingerprint recognition, image analysis and compression, edge detection, document analysis, character recognition, texture segmentation and classification, face recognition and iris detection, etc. [1, 2, 4, 5].

Gabor filters have been proved to be successfully applied to FER. Gabor filters are used via multi–resolution structures means filter consisting of multi frequencies and multi orientations. The multi-resolution structure of Gabor filter relates the Gabor filters to wavelets [4]. The 2-D Gabor filter is defined in (1) as follows –

$$\psi_{\omega,\theta}(x,y)=\frac{1}{2\pi\sigma_x\sigma_y}\exp\left[-\frac{1}{2}\left(\frac{x'^2}{\sigma_x^2}+\frac{y'^2}{\sigma_y^2}\right)\right]\exp\left[j\omega x'\right] \quad (1)$$

$$x'=x\cos\theta+y\sin\theta \quad (2)$$

$$y'=-x\sin\theta+y\cos\theta \quad (3)$$

Where (x, y) is the pixel position in the spatial domain, θ is the orientation of Gabor filter, ω is the radial center frequency, and σ is the standard deviation of the Gaussian function along

the x-axis and y-axis. The main parameters of the Gabor filter are its frequency and orientation. For feature extraction of facial image, the image are convolved with the Gabor filter as follows (4) –

$$G_{\omega,\theta}(x,y) = I(x,y) * \psi_{\omega,\theta}(x,y) \quad (4)$$

By using different–different orientations and frequencies the Gabor features are extracted. In general, Gabor filter bank uses five frequencies and eight orientations to extract the Gabor features for facial feature recognition.

The selection of discrete rotation any be θ_i and calculated the orientations must be spaced uniformly as follows (5) –

$$\text{(Orientation)}\ \theta_i = \frac{(i-1)\pi}{O}, \qquad i = 1, 2, ..., O. \quad (5)$$

For reduction of complexity and computation to half we use responses on angles [π, 2π[since they are complex conjugate of responses on [0, π[in a case of real-valued input. And the frequency calculated as follows (6) –

$$\text{(Frequency)}\ f_i = \frac{f_{\max}}{\left(\sqrt{2}\right)^{i-1}}, \qquad i = 1, 2, ..., S \quad (6)$$

Where O is the total number of orientations and S is that of scales, k=2 for an octave and k=√2 for half octave spacing [4]. So, total 40 Gabor feature is used to disintegrate and obtain the multi-frequency (scale) and multi-orientation Gabor features. We prefer the real Gabor amplitude information to depict the human facial expressions for its smoothness and stability. So, 40 Gabor magnitude pictures (GMP) as Gabor features [1, 2]. Fig. 6 show the Gabor filter bank corresponds to five spatial frequencies and eight orientations.

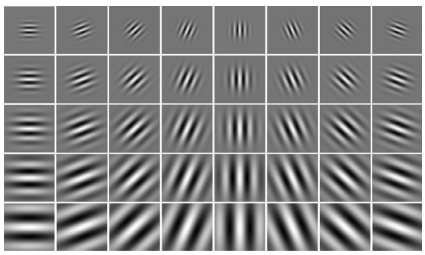


Fig. 6 40 Gabor features.

C. *Classification using Artificial Neural Network (ANN)*

The artificial neural network is an information processing paradigm that is inspired by the way biological nervous systems, such as the brain, process information. Fig. 7 shows the biological neural network which consists of a number of nerve cells (neurons) which are connected with each other via synapses. A neuron has a single output called dendrites. Axons transmit their nerve impulses via the neuromuscular junction with the help of synapses (neurotransmitter) situated in the axon terminal.

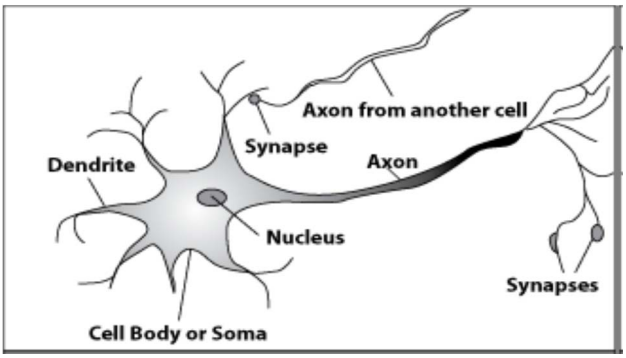


Fig. 7 A biological neural network [8]

ANNs are the most simplified model of biological neural networks. It uses the multi-layered feed-forward network which is an extension of single layered networks. Feed-forward ANN allows the signal to travel only one way from input to output. There is no feedback (loops). Feed-forward ANN tends to be straight forward networks that associate inputs with outputs. They are extensively used in pattern recognition. This type of organization is also referred to as bottom-up or top-down (Fig. 8.).

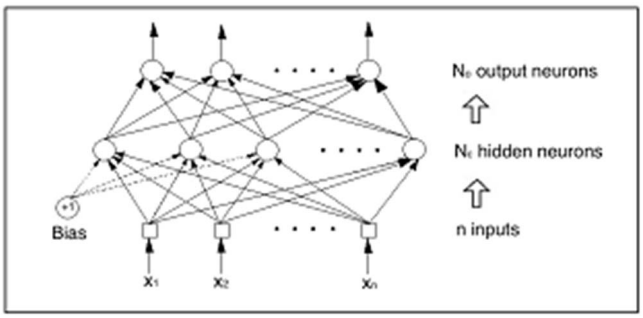


Fig. 8 Multi-layered network [8]

A network of this type contains an input layer, an output layer, and one and more intermediate layer. These intermediate layers are called "hidden" layers because their activity is not accessible from outside the network like input and output layers are. The number of neurons in the hidden layers can be difficult to determine as there are no specific rules and theorems. By using NN successful approaches have been achieved in approximation, categorization, and prediction and pattern recognition fields [8].

Pattern recognition is the most important application of neural network. In this paper, used multilayer feed-forward NN back propagation algorithm. It involves the data as training and testing. During the training process, the network is stimulated with a predefined set of data, called training set. Neural Network process [10] as follows:

- Training – There are generally 4 steps in the training process –

a. Assemble the training data
b. Create the network object
c. Train the network

d. Simulate the network response to new inputs.

- Target – Targets are user defined for the seven expressions. We used [1000000] for Angry, [0100000] for Disgust, [0010000] for Fear, [0001000] for Happy, [0000100] for Neutral, [0000010] for Sad and [0000001] for Surprise.
- Epoch – Epoch is the step in training process. In our dataset, numbers of epoch are 100 and number of hidden layer neurons are 50.
- Test – This process is for testing the image. In the testing phase, feed the images as input to test whether it classify the face into the respective class labels, or not.

III. DATABASE

One of the most important aspects of developing any new recognition and detection system is the choice of the database that will be used for testing the new system. The facial expression analysis database is grouped into two categories – first is Ekman and Friesen called the six basic emotions – happiness, sadness, fear, disgust, surprise and anger. And second is facial action coding system (FACS), which was design to describe subtle changes in the set of facial muscles [10].

A. *Japanese Female Facial Expression (JAFFE) Database*

TABLE I. JAFFE DATABASE [10]

No. of Subjects	Expressions	Image Resolution
10	7	256 X 256

The JAFFE database contains 213 images of 10 Japanese female models. Each subject (emotion) was recorded to three or four times while displaying the seven emotions. Fig. 9 shows example images of JAFFE database.

KL.AN3.169 KL.DI1.170 KL.FE1.174 KL.HA2.159 KL.NE2.156 KL.SA1.161 KL.SU2.165

Fig. 9 Example images of JAFFE database of KL [10]

IV. PROPOSED METHOD

This paper describes the proposed system of FER. This proposed methodologies as being explained in fig. 10. This whole process is as follows:

Step 1: Pre-processing for background removal and detection of only the face region from JAFFE database image using Viola-Jones face detection algorithm.

Step 2: Using Gabor filter, extracted the facial features. By using 5x8 Gabor filter bank total obtained 40 GMPs.

Step 3: Then, convolve the 40 GMP Gabor filter with the detected face image and the extract of the facial features.

Step 4: Next, calculated the mean of 40 GMPs of facial features.

Step 5: Then the whole process is repeated with all the images on JAFFE database images and create the database of mean values for input of NN.

Step 6: Using NN, classified the facial expression with the input of mean values of Gabor facial features.

V. SIMULATION AND RESULT

In this paper used the input image of JAFFE database is the front face images of the Japanese woman. Firstly, they detect the facial image of the whole image using Viola-Jones face detection algorithm. Then this detected face is the input of the feature extraction using Gabor filter. The Gabor filter extracts the facial features as 40 GMPs. Then calculate the mean of 40 GMPs. The mean value of all the facial expressions is the input of the NN classifier. So, finally, the NN classifier classifies the seven facial expressions using multi-layer feed-forward back propagation algorithm.

Fig. 10 shows the experiment of the facial expression recognition, results of this proposed method. Fig. 11 shows the final confusion matrix of 10 different model and 7 different expressions, total 70 facial images of JAFFE database. This confusion matrix shows that the recognition ratio is 85.7% using Gabor filter with the neural network. Table II shows the comparative literature study of different-different facial expression recognition methods using JAFFE database and the proposed method shows the high accuracy recognition rate.

TABLE II. COMPARATIVE LITERATURE METHODS FOR FER USING JAFFE DATABASE

Sr. No.	**Techniques**	**Successive Recognition Rate (%)**
1.	Statistics, Spatial Features, and NN [9]	70%
2.	LBP, Haar-like Feature, and ANN [12]	75%
3.	Viola-Jones face detector, Gabor Filter, and NN **[Proposed method]**	85.7%

VI. CONCLUSION AND FUTURE SCOPE

The paper represents a face recognition approach by using Gabor filter and neural network techniques. The theory neural networks are the adequate tool for control pattern recognition. The proposed method is robust for JAFFE database images. This proposed method is going to have an impact on our day to day life by enhancing the way we interact with computers, our surrounding living, and work spaces.

The differences do exist in facial features and facial expressions between cultures (for example Europeans and Asians) and age groups (adult and children), is the other major challenge. Facial expression recognition systems must become robust against such changes. The FER is on-going and interesting area for research. Hence there are the different path to explore it, as they can use the different features which can

improve its accuracy from 85.7% to high. We can try some other techniques for FER on some other databases to increase its accuracy.

Acknowledgment

I thank to my guide and everyone who gives me support to convey out this paper.

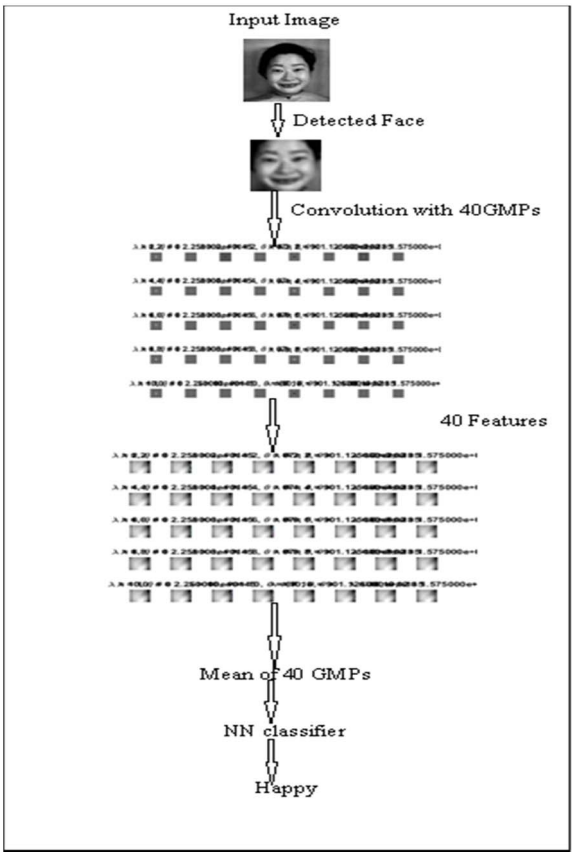


Fig. 10 Flow chart of proposed method and results of FER

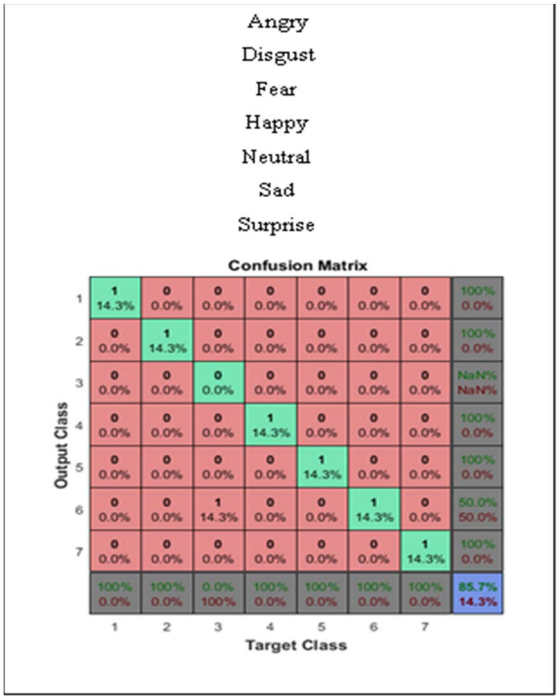


Fig. 11 Confusion Matrix show 85.7% recognition rate.

References

[1] Li R., Liu P. and Jia K., Wu Q., "Facial expression recognition under partial occlusion based on Gabor filter and gray-level co-occurrence matrix ," Intrenational Conference on Computational Intelligence andCommunication networks,2015.

[2] R. Azmi, S. Yegane, "Facial expression reconition in the presence of occlusion using local Gabor binary patterns,". 20th Iranian Conference on Electrical Engineering. Current Associates, Inc.:Tehran,pp.742-747,2012.

[3] Wang Yi-Qing, "An analysis of the Viola-Kones face detection algoritm," Image Processing On Line, 4 (2014),pp. 128-148, 2014.

[4] Ilonen J., Kamarainen J.-K., and Kalvianen H., "Efficient computation of Gabor features," Lappeenranta Universitty of Technology, 2005.

[5] Kaushal Avinash, Raina J. P. S., "Face detection using neural network and Gabor wavelet transform," International Journal of Computer Science and Technology, Vol. 1, Issue 1, September 2010.

[6] Hemalatha G., Sumathi C. P., "A study of techniques for facial detection and expression classification," International Journal of Computer Science & Engineering Survey (IJCSES) Vol.5, No.2, April 2014.

[7] Kolap Ashlesha D., Shrikhande S.V., and Jagtap Nitin K., "Review on various face recognition techniques," International Journal of Innovative Research in computer and Communication Engineering (An ISO 3297:2007 Certified Organization), Vol. 3, Issue 3, March 2015.

[8] Stevens T., Lovric D., "Artificial neural network (ANN)- in data pattern recognition for monitoring purpose," Bachelor thesis, Lunds Universitet, Helsingborg, June 2011.

[9] Nazil Perveen, Gupta Shubhrata and Verma Keshri, "Facial expression classification using statistical, spatial features and neural network," International Journal of Advances in Engineering & Technology, ISSN: 2231-1963, Vol. 4, Issue 1, pp.424-435, July 2012.

[10] Gross Ralph, "Face databases," In S. Li and A. Jain, Handbook of face recognition. Springer – Verlag,2005.

[11] Thomas Nisha, Mathew Mercy, "Facial expression recognition system using neural network and Matlab,".

[12] Sim Sang-Kwon, Cho jun-dong, " Expression recognition based on the artificial neural network using error backpropagation learning algorithm," International Journal of Applied Engineering Research, ISSN 0973-4562, pp. 820-823, Volume 11, Number 2, 2016.

[13] Yadollahi Mohammadreza, Prochazka Ales, "Artificial neural networks in pattern recognition," Institute of chemical Technology, Department of computing and Control Engineering.

[14] Saudagare Pushpaja V., Chaudhari D.S., "Facial expression recognition using neural network – an overview," International Journal of Computing and Engineering (IJSCE), ISSN: 2231-2307, Volume-2, Issue-1, March 2012.

[15] Linge Ganesh, Pawar Meenakshi, "Neural network based face recognition using PCA," International Journal of Computer Science and Information Technologies, Vol. 5(3), 4011-4014,2014.