



Facial Expression Recognition by Multiple Feature Extraction under Highly Corrupted Noisy Environment

Renu Nagpal*

Lecturer: C.S.E dept.

Rayat Institute of Engineering and Information Technology

Ropar, India

er.renunagpal@gmail.com

Sumeet Kaur

Assistant Professor: C.E dept.

Yadavindra College of Engineering

Talwandi Sabo, India

Purba_sumeet@yahoo.com

Pooja Nagpal

Lecturer: C.S.E dept.

Rayat Institute of Engineering and Information Technology

Ropar, India

Poojanagpal48@gmail.com

Abstract: A person's face is considered as the mirror of the mind. It is an important biometric feature for personal identification. Facial expressions and the changes in facial expressions provide important information about effective state of the person, his temperament and personality, psychopathological diagnostic information, related to stress levels, truthfulness etc. Body language and facial expressions are the best ways to know the personality of a person and the response of a person in various situations. The facial expressions tell us about concealed emotions which can be used to verify if the information provided verbally is true. Facial recognition and expression analysis is rapidly becoming an area of interest in computer science and in the design communities of human computer interaction. It plays an essential role in communications and in social interactions with other human beings which deliver rich information about their emotions. Facial expression plays an important role in interpersonal relations. Automatic recognition of facial expressions can be an important component of natural human-machine interface. Human being possesses an ability of communication through facial emotions in day to day interaction with others. Recognizing human facial expression and emotion by computer is an interesting and challenging problem. It provides a key mechanism for understanding and conveying emotions. This study aims at developing intelligent computers or robots that are mind implemented. This document demonstrates how multiple features are extracted and emotions are detected from noisy images. Facial expression is the best channel for emotion recognition and making machine intelligent. An automatic system for the recognition of facial expressions is based on a representation of the expression, learned from a training set of pre-selected meaningful features. However, in reality the noises that may embed into an image document will affect the performance of face recognition algorithms. As a first we investigate the emotionally intelligent computers which can perceive human emotions. In this research paper five emotions namely angry, fear, happy, sad along with neutral is tested from database in noisy environment of salt and pepper noise. Very high recognition rate has been achieved for all emotions along with neutral on the training dataset as well as user defined dataset. The proposed method uses Neural Networks by which emotions are classified.

Keywords: biometrics; facial expression; neural network; histogram equalization; multiple features

I. INTRODUCTION

Biometric technologies are automated methods of verifying or identifying the identity of a living person based on a physiological or behavioral characteristics. Biometrics has also been used to refer to the emerging field of technology devoted to identification of individuals using biological traits. The face plays a crucial role in interpersonal communication. By seeing a face we can recognize a person's identity, expression, age, gender. This information is irreplaceable for the normal conduct of human communications. If machines could recognize such information from a human face, humans and machines might thereby communicate more smoothly, robustly and harmoniously. In human interaction, the articulation and perception of facial expressions form a communication channel, that is additional to voice and that carries crucial information about the mental, emotional and even physical states of the conversation. In the simplest form, facial expressions can indicate whether a person is happy or

angry. More subtly, expressions can provide either conscious or subconscious feedback from listener to speaker to indicate understanding or doubts toward what the speaker is saying. Facial expression analysis deals with analysis of different facial motion changes by extraction of facial parameters. A typical system extracts number of facial parameters from an image, and classifies the image into the set of defined expressions.

Humans are capable of producing thousands of facial actions during communication that vary in complexity, intensity and meaning. Emotion or intention is often communicated by subtle changes in one or several discrete features. The addition or absence of one or more facial actions may alter its interpretation. Also, some facial expressions may have a similar gross morphology but indicate varied meaning for different expression intensities. Hence in order to capture the subtlety of facial expression in nonverbal communication, it is proposed to develop a computer vision system with a user interface that automatically extract features and their motion

information, discriminate subtly different facial expressions and estimate expression intensity.

Expression recognition relied heavily on the information such as geometry, texture, shape, models of muscular motion, feature correspondences and complex modeling schemes. While this information characterizes the underlying physical processes well, it is not available in all cases and is often difficult to compute from images or videos alone. The muscles that contract to produce the facial expressions of anger, fear, sadness, disgust, and enjoyment are the same the world over, regardless of sex or culture. There may also be universal expressions for surprise, contempt, and embarrassment as it is for anger, fear, sadness, disgust and enjoyment, but the evidence is not as complete. Facial recognition systems are computer programs that are used for automatically identifying a person. This technology works by using several facial features in a person's image and comparing these with existing images in the database. Facial recognitions systems are used as an additional and mass security measure and are comparable to the other biometric security systems available today such as retina scanners, fingerprint scanners, etc.

During the transmission of images over the network, some random usually unwanted variation in brightness or color information may be added as noise. Image noise can originate in film grain, or in electronic noise in the input device such as scanner digital camera, sensor and circuitry, or in the unavoidable shot noise of an ideal photon detector. Slow shutter speed and in low light having high exposure of the camera lens are also some of the reasons that noise gets added to the image. Noise causes a wrong conclusion in the identification of images in authentication and also in pattern recognition process. The noise should be removed prior to performing image analysis processes. The identification of the nature of the noise is an important part in determining the type of filtering that is needed for rectifying the noisy image.

Noise in imaging systems is usually either additive or multiplicative. In practice these basic types can be further classified into various forms such as amplifier noise or Gaussian noise, Impulsive noise or salt and pepper noise, quantization noise, shot noise, film grain noise and non-isotropic noise. However, in our experiments, we have considered only salt and pepper impulsive noise. The difficulty in removing salt/pepper noise from binary image is due to the fact that image data as well as the noise share the same small set of values (either 0 or 1) which complicates the process of detecting and removing the noise. This is different from grey images where salt/pepper noise could be distinguished as pixels having big difference in grey level values compared with their neighborhood. Many algorithms have been developed to remove salt/pepper noise in document images with different performance in removing noise and retaining fine details of the image. Most methods can easily remove isolated pixels while leaving some noise attached to graphical elements. Other methods may remove attached noise with less ability in retaining thin graphical elements. These conditions in turn stress the importance of the design of robust face recognition algorithms that retain recognition rates even under noisy environments. In general all the face recognition algorithms use any one or the combinations of the features namely shape, texture, colors or intensity to represent the facial image structure. It has been seen from previous works that the appearance based representations that uses the intensity or pixel values produces the better result compared

with other techniques. But the intensity features are very vulnerable to image noises that may add with the original image during transmission or during the capturing processes itself. In reality, most of the face recognition algorithms that uses appearance based representations are considered only for the noiseless environments and are not dealing with different type of noises occurred in the image. In this work, experiments have been conducted to reveal the robustness of our proposed technique under salt and pepper type of noise.

A. Organization of the Paper

The rest of the paper is organized as follows: Section 2 describes various types of applications for facial expression recognition. The implementation overview of proposed technique is given in section 3. In section 4, the experimental results have been discussed with respect to percentage of correct recognition considering JAFFE facial image database under salt and pepper noisy environments. The paper is concluded with some closing remarks and future scope in section 5.

II. APPLICATIONS

Automatic face expression recognition systems find applications in several interesting areas. With the recent advances in robotics, especially humanoid robots, the urgency in the requirement of a robust expression recognition system is evident. As robots begin to interact more and more with humans and start becoming a part of our living spaces and work spaces, they need to become more intelligent in terms of understanding the human's moods and emotions. Expression recognition systems will help in creating this intelligent visual interface between the man and the machine. Humans communicate effectively and are responsive to each other's emotional states. Computers must also gain this ability. This is precisely what the Human-Computer Interaction research community is focusing on: namely, Affective Computing. Expression recognition plays a significant role in recognizing one's affect and in turn helps in building meaningful and responsive HCI interfaces. Apart from the two main applications, namely robotics and affect sensitive HCI, expression recognition systems find uses in a host of other domains like Telecommunications, Behavioral Science, Video Games, Animations, Psychiatry, Automobile Safety, Affect sensitive music juke boxes and televisions, Educational Software, etc. As expression recognition systems become more real-time and robust, we will be seeing many other innovative applications and uses.

III. IMPLEMENTATION OVERVIEW

The design and implementation of the Facial Expression Recognition System can be subdivided into three main parts. The first part is image processing, second part is a recognition technique which includes training of the images and the third part is testing. The image processing part consists of image acquisition of noisy image as input through scanning or from JAFFE database by putting salt and pepper noise in to the image. Filtering, Feature Extraction, Region of Interest clipping, Quality enhancement of image. The second part consists of the artificial intelligence which is composed by Back Propagation Neural Network and Radial Basis Neural network. First training of the neurons is there and then testing is done. The First Part consists of several image processing

techniques. First, noisy face’s image acquisition is achieved by scanner or from JAFFE database then adaptive median filter is used to remove noise from the image and finally features are extracted. The region of interest is eye and lips, eyes or lips are clipped. These extracted features of image are then fed into Back-Propagation and Radial Basis Neural Network for training. In the Second Part, Back-propagation and RBF algorithms are used. It consists of two layers. At input to hidden layer Back-propagation Neural Network is used which consists of feed forward and feed backward layers and at hidden to output layer, RBF Neural Network is used for classification of expressions. A special advantage of the technique is that the expression is recognized even there is more noise density in the image up to 0.9. The range of density is taken between 0.1 to 0.9. Second by taking three extra inputs of statistical features like mean, median and standard deviation the classification is easy and there will be more correctness to recognize the facial expression. Third, dual enhancement of image is there, first at the removal of noise and second by using histogram equalization.

As the recognition machine of the system, a three layer neural network has been used that is trained with several times on various input ideal and noisy images forced the network to learn how to deal with noise. The window size is of 9. We can increase the size of window with that computation time is also increased. The variation in the density of noise is taken from 0.2 to 0.9. The adaptive filter removes up to 90% of noise from the image with more accuracy and the learning ranges from 0.1 to 0.9. Main accuracy or goal is 0.01, for that it takes more computation time. For single image total of 10 iterations are needed for zero error but for 21 images the error is zero in total of 15000 epochs because there are different images and different types of motions so more iterations are there. The outline of Facial Recognition System is shown in figure 1.

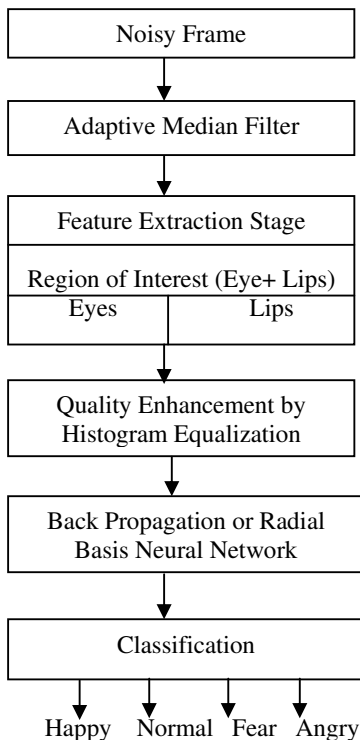


Figure 1. Outline of Facial Expression Recognition System

A. Input Noisy Image

This forms the first state for the face recognition module. To this module a noisy face image of salt and pepper noise is passed as an input for the system. The input image is randomly picked up from the data base used for training and evaluated for the recognition accuracy.

B. Adaptive Median Filter

A median filter is an example of a non-linear filter it is very good at preserving image detail. To run a median filter: consider each pixel in the image, sort the neighboring pixels into order based upon their intensities and replace the original value of the pixel with the median value from the list. Median filter is good at removing salt and pepper noise from an image, and also cause relatively little blurring of edges, and hence are often used in computer vision applications.

C. Preprocessing and Feature Extraction

In this phase of operation the face image is transformed to operational compatible format, where the face image is resized to uniform dimension, the data type of the image sample is transformed to double precision and passed for feature extraction. As the first step in image processing, the region of interest (ROI) of a lip and an eye or only lips region or only eyes region have been selected in the acquired images. The ROI image is converted into grayscale image.

D. Histogram Equalization

A histogram equalization method has been applied before obtaining the filtered grayscale image. Histogram equalization improves the contrast in the grayscale and its goal is to obtain a uniform histogram. The histogram equalization method also helps the image to reorganize the intensity distributions. New intensities are not introduced into the image. Existing intensity values will be mapped to new values but the actual number of intensity pixels in the resulting image will be equal or less than the original number. In the image sequence, the histogram equalized image is filtered using average and median filters in order to make the image smoother. Hence, the histogram equalized image is split into of lip ROI and eye ROI regions and then the regions are cropped from the full image. This has solved the problem of light intensity variations.

E. Training using Neural Network

An artificial neural network is a non linear and adaptive mathematical module inspired by the working of a human brain. It consists of simple neuron elements operating in parallel and communicating with each other through weighted interconnections. The typical components of a neural network are

- One input layer of neurons
- One or more hidden layers of neurons
- One output layer

The way in which these components of a neural network are arranged defines the architecture of that neural network. Figure 2 shows a typical architecture of a neural network.

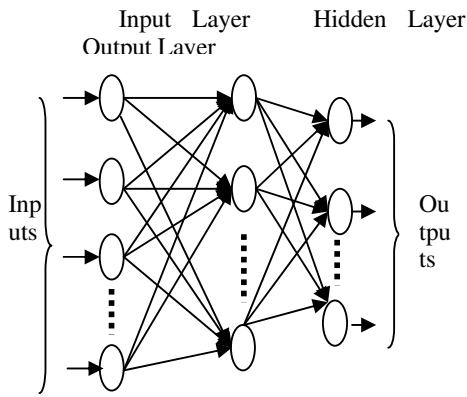


Figure 2. Typical neural network architecture

In training total of 625 neurons are taken, hidden neurons are 25 and output neurons are 4. Total of 13 pairs are trained with the different emotions of happy, angry, fear and neutral. In this phase epochs and errors are calculated of particular face region. Two parameters are used total numbers of epochs and errors are calculated and in neural network. For more accuracy more computation time is needed. The main accuracy or goal is 0.01 then it takes more computation time. The Classification of Neural Network includes two types of neural networks that were trained based on the input parameters extracted. Back Propagation Neural Network and Radial Basis Neural Network.

a) *Back Propagation Neural Network*: The most widely used neural network is the Back Propagation algorithm. This is due to its relatively simplicity, together with its universal approximation capacity. The parameters used in Back Propagation Neural Network are shown in table 1.

Table I. Parameters for Back Propagation Neural Network

Input Neurons	625 { 25 x25}
Hidden Neurons	75
Output Neurons	4
Network Size	625-75-4
Training Pairs	21
Learning Rate	0.5
Maximum Epochs	5000
Activation Functions	Sigmoid, Sigmoid
Error Goal	0.001

b) *Radial Basis Functions*: Radial Basis Functions (RBF) has attracted a great deal of interest due to their rapid training, generality and simplicity. The RBF is basically composed of three different layers: the input layer, which basically distributes the input data; one hidden layer, with a radially symmetric activation function, hence the network's name; and one output layer, with linear activation function. The parameters used in Radial Basis Neural Network are shown in table 2.

Table II. Parameters for Radial Basis Neural Network

Input Neurons	625 { 25 x25}
Hidden Neurons	75
Output Neurons	4
Network Size	625-75-4

Training Pairs	21
Learning Rate	0.5
Maximum Epochs	5000
Activation Functions	Sigmoid, Radbas
Error Goal	0.001

F. Classification

To demonstrate the capability and the accuracy of the recognition stage, selected databases of 21 faces of 4 classes are considered. The faces presented are the inputs into the training stage where a representative set of facial features were determined. After training, new images are processed and entered into the recognition stage for identification.

IV. EXPERIMENTAL RESULTS

For our experiments, the facial images from the facial image database JAFFE are used. The database contains 213 images of 7 facial expressions (6 basic facial expressions + 1 neutral) posed by 10 Japanese female models. Each image has been rated on 6 emotion adjectives by 60 Japanese subjects. Figure 3 shows the sample images used in our experiments collected from JAFFE face database. In our experiments, we have taken 3 persons images as 21 images for a single person from that 13 pairs are trained in the training phase means a total of 63 images are taken from that 39 are trained with different emotions we have used common type of noise namely, salt and pepper impulsive noise that affect the biometric image processing applications. In order to show the robustness of our face recognition method, these noises are introduced in the JAFFE database face images. Figure 3 shows the sample of image database.

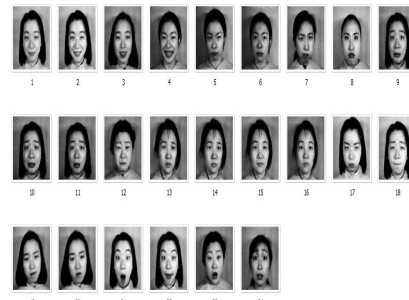


Figure 3. Sample images from JAFFE database

When the noise level increases, the face images get affected more and sometimes are not visible. Hence in our experiments, we have considered mean and variance varying from 0.1 to 0.9. To start with, the probing image set is formed by applying the salt and pepper noise with mean and variance equal to 0.2 on all the images of the JAFFE face database. All the images in the JAFFE database without adding any noise are taken as the prototype image set. Hence we got all images in prototype set. The feature set is formed by applying our technique on both the sets. In the experimental phase, we take the first image of the first subject from the prototype image set as the query image and the top matching ten images are found from a set of all probe images. If the top matching images lie in the same row (subject) of the prototype query image, then it is treated as a correct recognition. The number of correct recognized images for each query image in the prototype image set is calculated and the results are shown in the following figures for salt and pepper noise of variance 0.6

The results consists of three sections first is the preprocessing results which is shown in figure 4,5 and 6 .Figure 4 shows the eye feature to recognize facial expression, figure 5 shows the lip feature and figure 6 shows the mouth(eyes & lips) feature to recognize facial expression. Statistical features for all of the features are shown in tables 4 , 5and 6 Second section consists of the results of training of neural network starting with first there is a table of total frames in table1,then 2,3,and 4 tables consists of the training of eyes, lips and mouth features. The graphs for all of three features is shown in figure 7,8, and 9.Third section consists of testing of frames as there is total number of frames in table ,table , and consists of frames for eye, lip and mouth and the there is result of classification.

A. Preprocessing Results

1) Preprocessing Results for eye feature

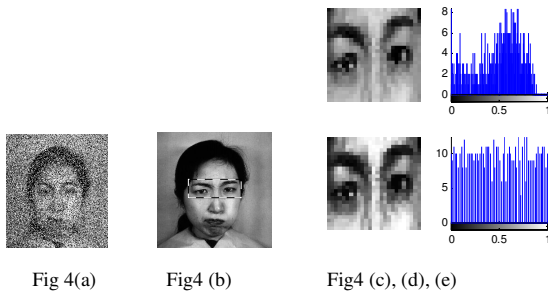


Figure 4 (a) Noisy Image with noise variance 0.6
 (b) Cropped Eyes Region in restored image
 (c) Cropped Eye region and Histogram
 (d) Enhancement of cropped Eyes by Histogram Equalization
 (e) Histogram after enhancement

Table III. Statistical Features of Eyes

Statistical Features	Noisy Frame	Restored Frame	Cropped Frame	Enhanced Frame
Mean	0.4838	0.4703	0.5021	0.5001
Median	0.5177	0.5221	0.5708	0.4921
Standard Deviation	0.3901	0.2266	0.2646	0.2938

2) Preprocessing Results for lip feature

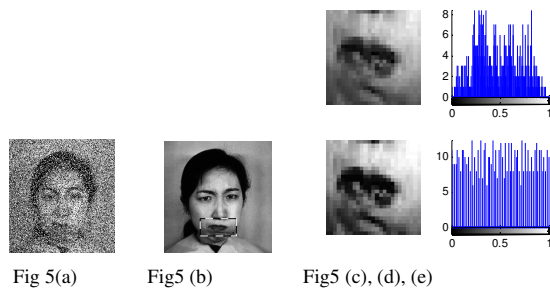


Figure 5 (a) Noisy Image with noise variance 0.6
 (b) Cropped Lip Region in restored image
 (c) Cropped Lip region and Histogram
 (d) Enhancement of cropped Lip by Histogram Equalization
 (e) Histogram after enhancement

Table IV. Statistical Features of Lips

Statistical Features	Noisy Frame	Restored Frame	Cropped Frame	Enhanced Frame
Mean	0.4838	0.4699	0.5157	0.4999
Median	0.5177	0.5177	0.5177	0.4921
Standard Deviation	0.3901	0.2268	0.2283	0.2935

3) Preprocessing Results for mouth feature

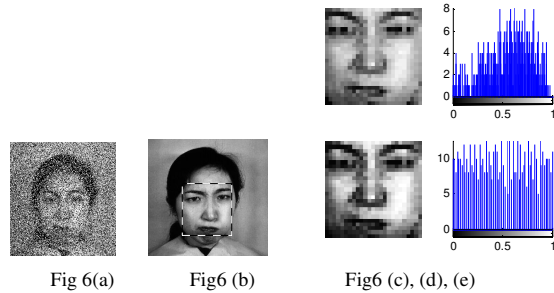


Figure 6 (a) Noisy Image with noise variance 0.6
 (b) Cropped mouth Region in restored image
 (c) Cropped mouth region and Histogram
 (d) Enhancement of cropped mouth by Histogram Equalization
 (e) Histogram after enhancement

Table V. Statistical Features of mouth (eye+lip)

Statistical Features	Noisy Frame	Restored Frame	Cropped Frame	Enhanced Frame
Mean	0.4838	0.4702	0.5260	0.5005
Median	0.5177	0.5221	0.5487	0.4921
Standard Deviation	0.3901	0.2270	0.2446	0.2936

B. Training of Neural Network Results

Table VI. Total Training Frames

Frame	Total Frames	Classes			
		Angry	Fear	Happy	Normal
1	21	3	4	4	2

Table VII. Training of Lips

Max Epochs	Maximum Error	Minimum Error
5000	2.4	0.5
10000	1.2	0.4
15000	0.9	0.1

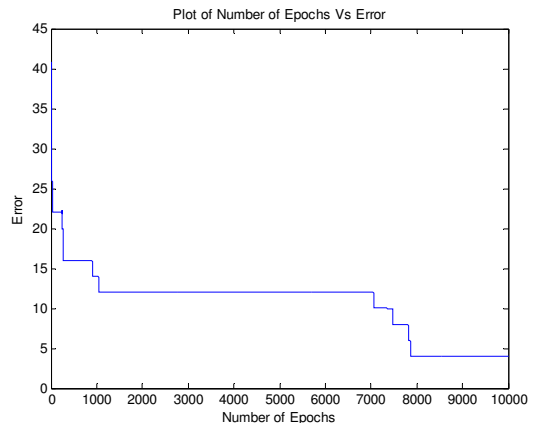


Figure7. Graph of Epochs versus Error for Lips

Table VIII. Training of Eyes

Max Epochs	Maximum Error	Minimum Error
5000	3.7	0.6
10000	1.2	0.9
150000	1.6	1.0

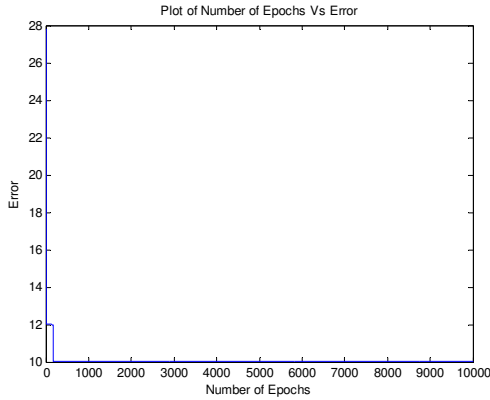


Figure 8 Graph of Epochs versus Error for Eyes

Table IX. Training of Eyes and Lips

Max Epochs	Maximum Error	Minimum Error
5000	2.3	0.8
10000	1.4	0.4
150000	0.8	0.1

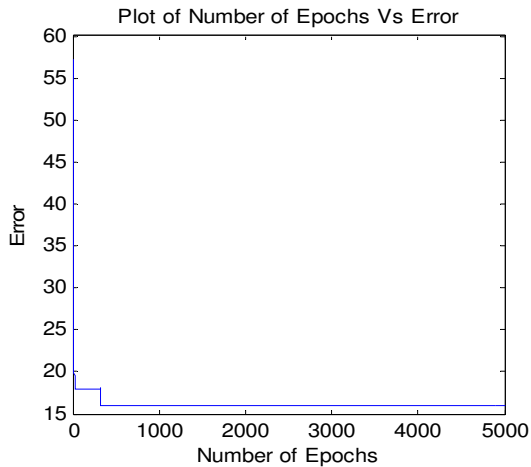


Figure 9 Graph of Epochs versus Error for Lips and Eyes

C. Testing of Frames

Table X. Testing of Lips

Tested Frames	Correct Classification	Wrong Classification	Performance
13	12	1	90

Table XI. Testing of Eyes

Tested Frames	Correct Classification	Wrong Classification	Performance
13	12	1	90

Table XII. Testing of Eyes And lip

Tested Frames	Correct Classification	Wrong Classification	Performance
13	13	0	100

V. RESULTS OF CLASSIFICATION



Angray
Fig10 (a)



Happy
Fig10 (b)



Fear
Fig10(c)



Normal
Fig10 (d)

Figure 10 (a) Result of Classification of Angry Expression
(b) Result of Classification of Happy Expression
(c) Result of Classification of Fear Expression
(d) Result of Classification of Normal or Neutral Expression

VI. DISCUSSION,CONCLUSION ANSD FUTURE WORK

This paper has covered automatic expression recognition. Similar architectures and processing techniques are often used for facial expression recognition and face recognition, despite the duality that exists between these recognition tasks. 2-D monochrome facial image sequences are the most popular type of pictures used for automatic expression recognition. Although a variety of face detection techniques have been developed, robust detection and location of faces or their constituents is difficult to attain in many cases. Features for automatic expression recognition aim to capture static or dynamic facial information specific to individual expressions. Geometric, kinetic, and statistical- or spectral-transform-based features are often used as alternative representation of the facial expression prior to classification.

A wide range of classifiers, covering parametric as well as non-parametric techniques, has been applied to automatic expression recognition. Generally speaking, automatic expression recognition is a difficult task, which is afflicted by the usual difficulties faced in pattern recognition and computer vision research circles, coupled with face specific problems. As such, research into automatic expression recognition has been characterized by partial successes, achieved at the expense of constraining the imaging conditions, in many cases.

Unresolved research issues are encapsulated in the challenge of achieving optimal pre-processing, feature extraction or selection and classification, under conditions of data variability. Sensitivity of automatic expression recognition to data variability is one of the key factors that have curtailed the spread of expression recognizers in the real world.

However, few studies have systematically investigated robustness of automatic expression recognition under adverse

VII. CONCLUSION

Multiframe Radial Basis Neural Network is used for emotion detection. In this work, a multiple feature options such as face, eyes and lips are used for emotion detection. This technique can be used as robust face emotion detection algorithm. Another advantage of this technique is enhancement of cropped image such as face, eyes, and lips are made by histogram equalization. The average efficiency of this algorithm is 90% in case of large database emotion detection. Classification of emotions such as Fear, Angry, Normal and Happy are considered.

VIII. FUTURE WORK

Future work of this technique is: The computation bottleneck of algorithms can be reduced by training of Radial Basis Neural Network by PSO or by ordered hybrid optimization techniques. A number of emotions should be considered more to make algorithm universal. The same algorithm should be developed by using GUI.

IX. REFERENCES

- [1] Tingfan Wu, Marian S. Bartlett and Javier R. Movellan, "Facial Expression Recognition Using Gabor Motion Energy Filters" *IEEE CVPR workshop on Computer Vision and Pattern Recognition for Human Communicative Behavior Analysis.in*, (2010)
- [2] G.Sofia ,M. Mohamed Sathik , "An Iterative approach For Mout Extraction In Facial Expression Recognition", *Proceedings of the Int. Conf. on Information Science and Applications ICISA 2010*, 6 February 2010, Chennai, India.
- [3] Ruba Soundar Kathavarayan and Murugesan Karuppusamy, "Preserving Global and Local Features for Robust Face Recognition under Various Noisy

conditions.

- Environments", *International Journal of Image Processing (IJIP) 2010*, Volume (3), Issue (6)
- [4] Jagdish Lal Raheja, Umesh Kumar, "Human Facial Expression Detection from detected in Captured image using Back Propagation Neural Network", *International Journal of Computer Science and Information Technology (IJCSIT)*, Vol 2, No.1, February 2010.
- [5] C.R Vimal chand, "Face and gender Recognition Using Genetic Algorithm and Hopfield Neural Network", *Global Journal of Computer Science and Technology*, Vol. 10 Issue 1 (Ver. 1.0), 2010
- [6] Ramesha K, K B Raja, Venugopal K R and L M Patnaik, "Feature Extraction based Face Recognition, Gender and Age Classification", *International Journal on Computer Science and Engineering*, Vol. 02, No.01S, 2010, 14-23
- [7] Sarawak Anam, Md. Shohidul Islam, M, A, Kashem, M.N.Islam, M.R.Islam and M.S.Islam, "Face Recognition using Genetic Algorithm and Back Propagation Neural Network", *Proceedings of the International Multiconference of Engineers and Computer Scientist 2009*.Vol.1.IMCSE 2009.
- [8] P.Y. Simard, H.S. Malvar, "An efficient binary image activity detector based on connected components", *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing*, 229–232, 2004.
- [9] L. Beaurepaire, K.Chehdi, B.Vozel, "Identification of the nature of the noise and estimation of its statistical parameters by analysis of local histograms", *Proceedings of ICASSP-97*, Munich, 1997.
- [10] Noise Models, http://homepages.inf.ed.ac.uk/rbf/CVonline/LOCAL_COPIES/VELDHUIZEN/node11.html
- [11] Image Noise, http://en.wikipedia.org/wiki/Image_noise
- [12] H.S.M. Al-Khaffaf , A.Z. Talib, R. Abdul Salam, "A Study on the effects of noise level, cleaning method, and vectorization software on the quality of vector data", *Lecture Notes in Computer Science* 299-309.