Facial Expression Classification with Enhanced Feature Extraction using Particle Swarm Optimization and Emotional Back Propagation Neural Network

Vedantham Ramachandran, and Dr. E.Srinivasa Reddy

***Abstract*—**Human face-to-face communication plays an important role in human communication and interaction. In recent years, several different approaches have been proposed for developing methods of automatic facial expression analysis. In this paper we have proposed a novel facial expression recognition system which chooses the optimized features using particle swarm optimization (PSO) from the features calculated from principle component analysis (PCA) of input face images. These optimized features are then used to train the emotional backpropagation neural network (EBPNN). Using this neural network classifier, emotions are classified. The proposed architecture yields good results when PSO features compared with normal PCA features.

***Index Terms*—**Back propagation neural network, Emotion classification, Facial expression, Human Computer Interaction, Particle swarm optimization.

I. Introduction

Development of a robust FER system is still a challenging issue, largely because of various unpredictable facial variations and complicated exterior environmental conditions [1]. These variations make it difficult to pre-locate facial regions and perform robust and accurate feature extraction. Many efforts have been made to overcoming these variations, especially in pose and illumination condition. Although these studies have achieved promising results, to be best of our knowledge, no approach has been reported on handling face localization errors (e.g. changes in face scale and location), and relatively little attention on overcoming facial occlusions. The significant impact of these two types of variations on FER performance has been specifically highlighted in recent works [2].

Particle swarm optimization (PSO) is a population based stochastic optimization technique developed by Dr. Eberhart and Dr. Kennedy in 1995, inspired by social behavior of bird flocking or fish schooling [3].

PSO shares many similarities with evolutionary computation techniques such as Genetic Algorithms (GA). The system is initialized with a population of random solutions and searches for optima by updating generations. However, unlike GA, PSO has no evolution operators such as crossover and mutation. In PSO, the potential solutions, called particles, fly through the problem space by following the current optimum particles. The detailed information will be given in following sections.

Compared to GA, the advantages of PSO are that PSO is easy to implement and there are few parameters to adjust. PSO has been successfully applied in many areas: function optimization, artificial neural network training, fuzzy system control, and other areas where GA can be applied [4].

PSO is used it to solve the optimization problems. In PSO, each single solution is a "bird" in the search space. We call it "particle". All of particles have fitness values which are evaluated by the fitness function to be optimized, and have velocities which direct the flying of the particles. The particles fly through the problem space by following the current optimum particles.

PSO is initialized with a group of random particles (solutions) and then searches for optima by updating generations. In every iteration, each particle is updated by following two "best" values. The first one is the best solution (fitness) it has achieved so far. (The fitness value is also stored.) This value is called pbest. Another "best" value that is tracked by the particle swarm optimizer is the best value, obtained so far by any particle in the population. This best value is a global best and called gbest. When a particle takes part of the population as its topological neighbors, the best value is a local best and is called lbest.

After finding the two best values, the particle updates its velocity and positions with following equation (a) and (b).

```
v[] =
v[]+c1*rand()*(pbest[]-present[])+c2*rand()*(g
best[]-resent[])…(a)
present[] = persent[] + []………………………(b)
```

v[] is the particle velocity, persent[] is the current particle (solution). pbest[] and gbest[] are defined as stated before. rand () is a random number between (0,1). c1, c2 are learning factors. usually c1 = c2 = 2.

Vedantham Ramachandran, Research Scholar, Dept. of CSE, AcharyaNagarjuna University, Nambur, Guntur, AP, India

Dr. E.Srinivasa Reddy, Prof. & Dean, University College of Engg. and Tech., Acharya Nagarjuna University , Nambur, Guntur, AP, India

The pseudo code of the procedure is as follows

```
For each particle
    Initialize particle
END
Do
    For each particle
        Calculate fitness value
        If the fitness value is better than the
best fitness value (pBest) in history
            set current value as the new pBest
    End

    Choose the particle with the best fitness
value of all the particles as the gBest
    For each particle
        Calculate particle velocity according
equation (a)
        Update particle position according
equation (b)
    End
While maximum iterations or minimum error
criteria is not attained
```

Particles' velocities on each dimension are clamped to a maximum velocity Vmax. If the sum of accelerations would cause the velocity on that dimension to exceed Vmax, which is a parameter specified by the user. Then the velocity on that dimension is limited to Vmax.

Figure 1 shows the framework of a general FER research and development system. It contains six main steps: face pre-processing (face detection, tracking, and normalization), feature extraction, feature selection, emotion classification, emotion representation, and performance evaluation [5]. The input data can be static images or video sequences. The classified facial expressions are then represented in different methods and the performance is evaluated using different measurements[18].

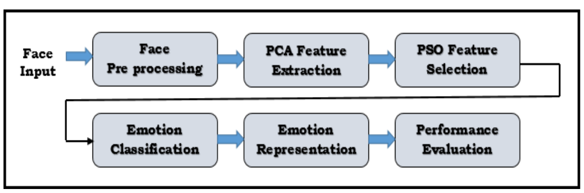

Fig 1: An Expression Recognition system

II. Proposed Facial Recognition System

The proposed architecture in this work contains the following stages: pre-processing of input images, feature extraction, training, classification, and database [6]. This work proposes a new solution to the facial expression recognition problem, describing a facial recognition system that can be used in application of Human computer interface [14]. Pre-processing of input images includes, face detection and cropping. Feature extraction is the process of deriving unique features from the data and can be accomplished by specific algorithms like Feature averaging, principal component analysis etc [7]. Training of neural network will be done by giving the extracted features as input to the neural network with specified network parameters.

Classification will be done by the neural network according to the specified targets in the network. The following figure shows the proposed architecture of the Facial recognition system. Pre-processing is the next stage after entering the data into the facial expression recognition system. The important data that is needed for most facial expression recognition methods is face position. In pre-processing module images are resized from 256 x 256 pixel value to 280 x 180 pixel values. The Sobel method has been used to identify the face edges.

Face images are taken from Cohn Kanade database of facial expressions. The original image contains time and camera model also. For better performance, face is detected and cropped and saved as separate image. The cropped image is then used to extract features. These features are given as input to the neural network and will be trained to gain knowledge.

III. Preprocessing

The testing image will also be preprocessed and features will be extracted and input to the neural network. The classifier of the neural network will classify the expression of the input test image.

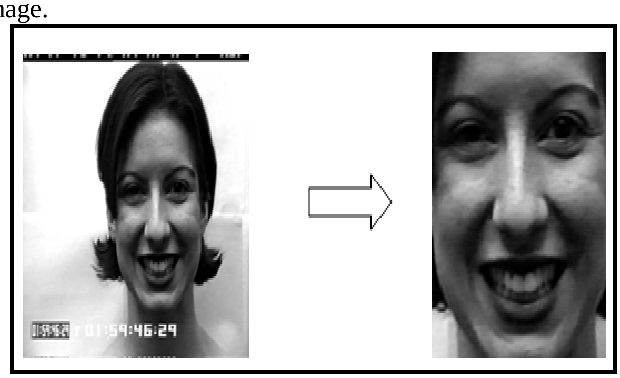

Fig. 2 Face detection and cropping

In order to perform data reduction, the first step is to take the required data from an image. So the face is detected and cropped from original image as shown in Fig.2.

- In this paper two architectures are proposed. The first architecture consists of Feed forward neural network where the input layer neurons are the features selected by PSO from Eigen vectors of the principle component analysis [6]. Figure 3 shows the first architecture being proposed.

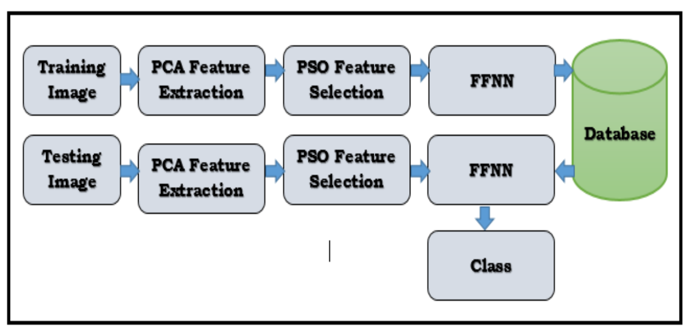

Fig. 3 Architecture of FFNN classification with PCA features optimized by PSO.

IV. PSO Coefficients With Emotional Back Propagation Neural Network

This is the second architecture proposes the classification with neural network using PSO Coefficients selected from PCA features of input images. The training images were taken and applied the PCA and then PSO is used to select optimized features. The remaining features are input to the emotional back propagation neural network to train the network[15]. The neural network will produce the knowledge database.

In the process of testing, the test input image will be applied pattern averaging and the remaining features will be used to classify through the neural network classifier and using the knowledge database gained from training. The architecture is shown in the Fig 4.

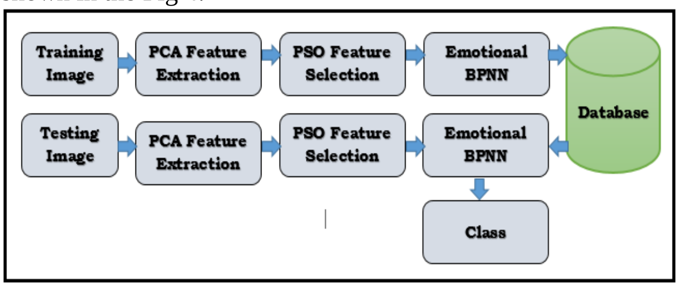

Fig. 4 Proposed architecture of EBPNN classification with PSO

The proposed architectures use the feed forward and emotional back propagation neural network [3] architectures. The generalized architecture of the proposed system is shown in Fig 5. And an adhoc cmparision was done between these two kinds of neural network appraches over the facial expression recognition. A brief functioning of artificial neural networks along with its implemntation was described in the following paragraph.

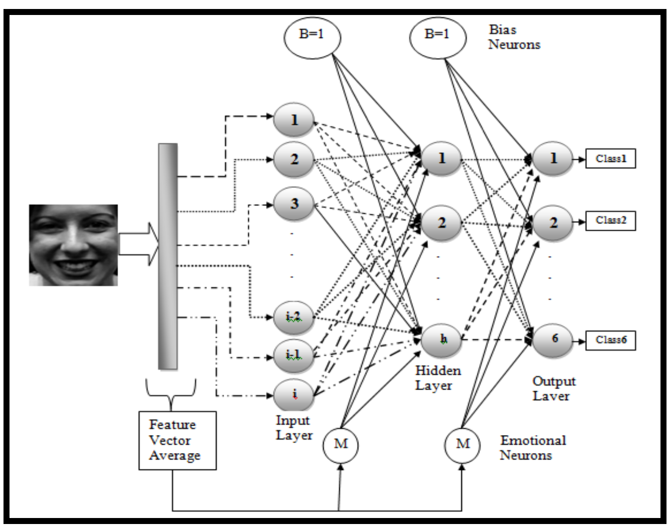

Fig. 5 Generalized EBP-based neural networks

V. Tests, Results & Conclusions

The implementation of neural network consists of training and testing. The training and testing is performed on Cohn Kanade facial expression database. The database consists of 2000 images of 200 subjects. About 600 images were used in this work for the training and testing process.

The performance of the system is measured by varying the number of images of each expression in training and testing. Following table shows the performance of the proposed method along with the other methods.

TABLE 1
COMPARISON OF RESULTS ON COHN KANADE DATABASE

Training samples	Testing samples	Training Time		Recognition Rate	
		PSO +FFNN	PSO +BPNN	PSO +FFNN	PSO +BPNN
8	2	140.81	132.36	97	100
7	3	124.19	110.32	97	100
6	4	54.50	52.43	96	98
5	5	27.45	22.69	94	97
4	6	12.86	10.08	90	94
3	7	5.36	4.148	86	90
2	8	2.45	2.12	82	88

The recognition performance increases as the number of training samples increases. The lower the number of training samples the lesser the recognition rate. It is found that the PSO features with emotional back propagation neural network are yielding the better results even the training samples are less. The performance plot was shown against various algorithms, number of training images and their performances.

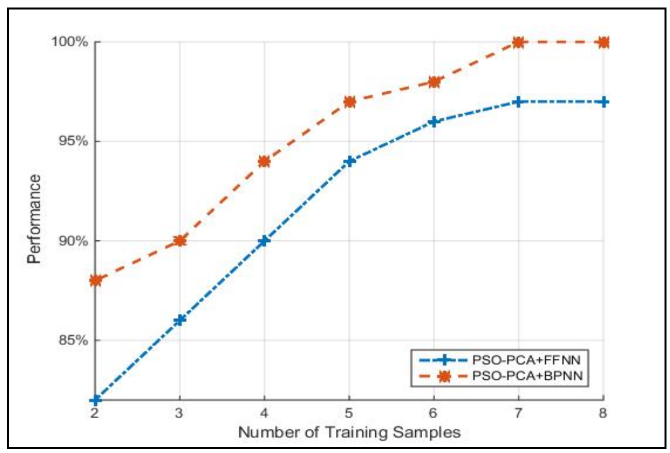

Fig. 6 Plot against FFNN & BPNN Recognition Rate

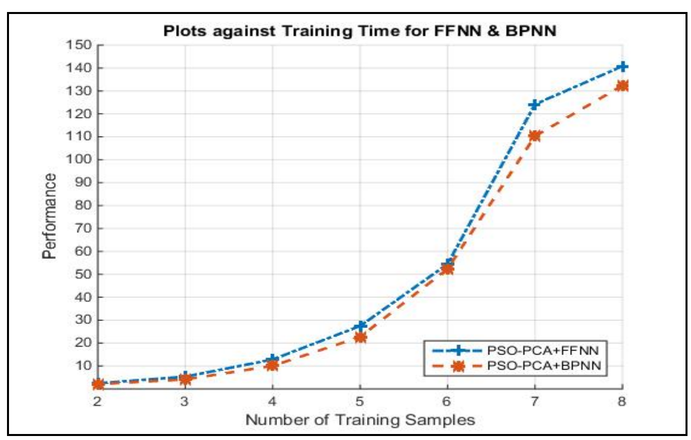

Fig. 7 Plot against FFNN & BPNN Training Times

The confusion matrix is created for each of the test. The test is performed on five subjects.

TABLE 2
PSO+PCA+BPNN CONFUSION MATRIX

Found → Actual ↓	Surprise	Fear	Happy	Sad	Anger	Disgust
Surprise	100	0	0	0	0	0
Fear	0	100	0	0	0	0
Happy	0	0	60	0	0	40
Sad	0	0	0	100	0	0
Anger	0	0	0	60	40	0
Disgust	0	0	0	0	0	100

The confusion matrix shows the percentage of correct classifications and mis-classifications also. Diagonal elements show the correct classification results.

All other elements are the misclassifications. The test and training results of various facial emotional classification methods is shown in Table 2. Experimental results show that the proposed architecture improves the performance of the facial expressions [8]. Based on the results we can conclude that the proposed emotional back propagation neural network with DCT is best in both cases of minimization of training time of neural network and performance as well[13]. Since the emotional parameters were introduced, the training time for the single iteration may be little more but the overall training time is reduced in achieving the minimization of error.

The performance and training time of the neural network depends on the parameters selected like learning coefficient and momentum factor. The number of hidden neurons is also affecting the performance of the neural network. Experiments were carried out by altering the learning coefficient and number of hidden neurons and the types of sigmoid functions[17].

The optimal value for learning rate is 0.02, which produces the best performance for facial expression recognition. The number of hidden neurons is same as the number of input neurons. Sigmoid action function is used in both hidden layer and output layer for activating the neurons. In the classification part of the emotional back propagation neural network, the time very less when compared to other neural networks.

The work can be extended to clustering techniques like segmentation for the lower training times and higher performance. Since the training data is still images, there is more dependency on the image data like lighting, illumination conditions, poses of the faces, variations in expression and gender of the person also [18].

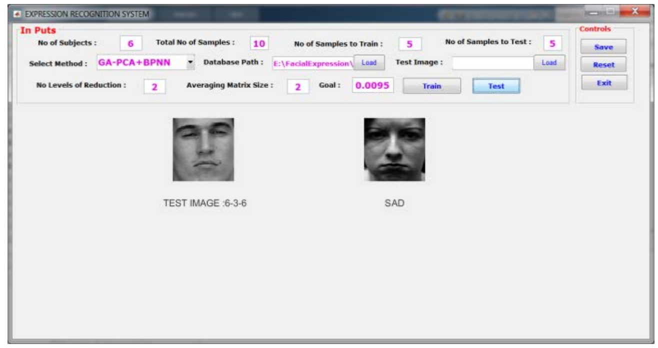

Fig. 8. GUI Result for the Developed System

VI. CONCLUSIONS AND FUTURE WORK

We present an approach for facial expression estimation that combines state-of-the-art techniques for model-based image interpretation and sequence labelling. Learned objective functions ensure robust model-fitting to extract accurate model parameters. The classification algorithm is explicitly designed to consider sequences of data and therefore considers the temporal dynamics of facial expressions. Future work aims at presenting the classifier training data that is obtained from various publicly available databases to reflect a broader. Variety of facial expressions. Furthermore, our approach will be tuned towards applicability in real-time. It is planned to create a working demonstrator from this approach.

REFERENCES

[1] Vedantham Ramachandran et.al., “A novel Facial Expression Classification System using Emotional Back Propagation Artificial Neural Network and Genetic Algorithm” International Journal of Applied Engineering Research (IJAER), Volume-10, Issue-11, November 2015

[2] Kennedy, James. "Particle swarm optimization." Encyclopedia of machine learning. Springer US, 2011. 760-766.

[3] Vedantham Ramachandran et.al., “Facial Expression Classification System with Emotional Back Propagation Neural Network”, International Journal of Scientific & Engineering Research(IJSER), Volume-04, Issue-10, September 2013

[4] P Phaneendra et.al., “Human Face Detection and Recognition using PCA and DCT in HMM”, International Journal of Scientific & Engineering & Technology Research(IJSETR), Volume-04, Issue-35, August 2015

[5] Vemulakonda Srinivasa Rao et.al., “ANN based facial clustering technique using PCA and DCT approaches “,International Journal of Advanced Research & Innovative Research(IJARIR), Volume-7, Issue-18,December 2015

[6] V.Ramachandran et.al., “Image Retrieval System with User Relevance Feedback “,International Journal of International Journal of Computer & Communication Technology (IJCCT), Volume-3, Issue-5, July 2004

[7] Adnan Khashman , “A modified back propagation learning algorithm with added emotional coefficients”, IEEE transactions on neural networks, vol. 19, no. 11, November 2008 http://dx.doi.org/10.1109/TNN.2008.2002913

[8] Vedantham Ramachandran et.al., “An Enhanced Facial Expression Classification System Using Emotional Back Propagation Artificial Neural Network with DCT Approach”, International Journal of Computer Science Engineering and Information Technology Research(IJCSEITR), Volume-04, Issue-05, August 2015

[9] D. Rumelhart, G. E. Hinton, and R. J.Williams, “Learning internal representations by error propagation,” in Parallel Distributed Processing, D. Rumelhart and J. McClelland, Eds. Cambridge, MA: MIT Press, 1986, vol. 1.

[10] H. Murase and S. K. Nayar, “Visual learning and recognition of 3-D objects from appearance,” Int. J. Comput. Vis., vol. 14, pp. 5–24, 1995. http://dx.doi.org/10.1007/BF01421486

[11] P. M. Chavan, Manan C .Jadhav, Jinal B. Mashruwala, Aditi K. Nehete, Pooja A. Panjari, “Real Time Emotion Recognition through Facial

Expressions for Desktop Devices", International Journal of Emerging Science and Engineering (IJESE) Volume-1, Issue-7, May 2013

[12] Ajit P. Gosavi, S. R. Khot, "Facial Expression Recognition Using Principal Component Analysis", International Journal of Soft Computing and Engineering (IJSCE), Volume-3, Issue-4, September 2013

[13] V.Ramachandran, et.al "Facial Expression Classification Systems with Emotional Back Propagation Neural Network " International Journal of Scientific and Engineering Research (IJSER) - (ISSN 2229-5518), Volume 4, Issue 9.during Aug 2013

[14] Lt. Dr. S. Santhosh Baboo, Lt. Dr. S. Santhosh Baboo," Face Emotion Analysis Using Gabor Features In Image Database for Crime Investigation", International Journal of Data Engineering (IJDE), Volume (2) : Issue (2) : 2011

[15] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection," IEEE Trans. Pattern Anal. Mach. Intell., vol. 19, no. 7, pp. 711–720, Jul. 1997. http://dx.doi.org/10.1109/34.598228

[16] S. Z. Li, X. W. Hou, H. J. Zhang, and Q. S. Cheng, "Learning Spatially localized, parts-based representation," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2001, pp. 207–212. http://dx.doi.org/10.1109/cvpr.2001.990477

[17] L. L. Huang and A. Shimizu, "Combining classifiers for robust face detection," in Lecture Notes in Computer Science. Berlin, Germany: Springer-Verlag, 2006, vol. 3972, pp. 116–121.

[18] C. Park, M. Ki, J. Namkung, and J. K. Paik, "Multimodal priority verification of face and speech using momentum back-propagation neural network," in Lecture Notes in Computer Science. Berlin, Germany: Springer-Verlag, 2006, vol. 3972, pp. 140–149.

Vedantham Ramachandran, is a research scholar at Acharya Nagarjuna University, Nambur and an associate professor at VasireddyVenkatadri Institute of Technology. He did B. Tech Computer Science & Systems engineering Degree from Andhra University and M. Tech in Computer Science Engineering from JNTU, Kakinada. He is very much interested in image processing, medical image retrieval, human vision & pattern recognition. He did several projects in image processing.

Dr. E. Srinivasa Reddy, Professor & Dean R&D in University Engineering College of Acharya Nagarjuna University. He did Master of Technology in Computer Science Engineering from sir Mokhsagundam Visweswariah University, Bangalore. He also did Mater of Science from Birla Institute of Technological Sciences, Pilani. He did Philosophical Doctorate in Computer Science Engineering from Acharya Nagarjuna University, Guntur. Prof.E.S.Reddy, guided several research scholars of PhD to success. His research areas of interest are in various domains including Digital Image Processing.