Facial Expression Classification System with Emotional Back Propagation Neural Network

* V.Ramachandran[1] B. Sivaiah[2] Dr. E. Srinivasa Reddy[3]
[1]Research Scholar, Dept. of CSE, Acharya Nagarjuna University, Nambur, Guntur, AP, India.
[2] Assistant Professor, Dept. of CSE, Vignan's University, Vadlamudi, Guntur, AP, India.
[3]Professor & Principal, Acharya Nagarjuna University, Nambur, Guntur, AP, India.

Abstract: Facial expression recognition is also gaining interest among the researchers because of its inevitable advantages in image retrieval which can be extended many fields like medicine, artificial intelligence, robotics and neural networks. So it is one of the hot topic for researchers. Existing methods such as PCA, LDA, LPP etc. with Euclidian distance classifier are popular. Neural network classifiers are also used along with above for classification. In this paper, the pattern averaging and PCA are used for feature extraction. In this work, feed forward neural network with added emotional coefficients (EBPNN) for facial expression classification is being proposed. The network is trained with back propagation algorithm. The results are compared with normal feed forward neural network with back propagation. The proposed algorithm is producing better results over the existing.

Keywords: Face Recognition, Biometrics, Thermal facial signatures, Eigenfaces.

1. Introduction

Any automatically measurable, robust and distinctive physical characteristic or personal trait that can be used to identify an individual or verify the claimed identity of an individual. Measurable means that the characteristic or trait can be easily presented to a sensor, located by it, and converted into a quantifiable, digital format. The robustness of a biometric refers to the extent to which the characteristic or trait is subject to significant changes over time. Distinctiveness is a measure of the variations or differences in the biometric pattern among the general population. Biometrics is used for human recognition which consists of identification and verification. Facial recognition records the spatial geometry of distinguishing features Of the face. Different vendors use different methods of facial recognition, however, all focus on measures of key features of the face. Because a person's face can be captured by a camera from some distance away, facial recognition has a clandestine or covert capability (i.e. the subject does not necessarily know he has been observed).

For this reason, facial recognition has been used in projects to identify card counters or other undesirables in casinos, shoplifters in stores, criminals and terrorists in urban areas. Facial Recognition Also Provides a Surveillance Capability Desire to Locate Specific Individuals like Criminals, Terrorists, Missing children.

Researchers are studying the role of emotions in artificial intelligence (AI) from a variety of viewpoints: to develop agents and robots that interact more gracefully with humans, to develop systems that use the analog of emotions to aid their own reasoning, or to create agents or robots that more closely model human emotional interactions and learning. In human-computer or human-human interaction systems, emotion recognition systems could provide users with improved services by being adaptive to their emotions. In virtual worlds, emotion recognition could help simulate more realistic avatar interaction

Advantages of Facial Recognition Surveillance:

- Uses faces, which are public
- Involves non-intrusive, contact-free process
- Uses legacy databases
- Integrates with existing surveillance systems

In the proposed system, there are three phases: i) Face detection and cropping from image database, ii) feature extraction from the database, iii) classification of facial expressions using neural network classifier. The proposed system was tested on the Cohn Kanade facial expression database (Web Resource). Six expressions of 10 subjects and 10 samples in each expression are used for training and testing. The basic expressions like, surprise, fear, happy, sad, anger, and disgust are considered in the system.

2. Proposed Facial Emotion Classification method

The proposed architecture in this work contains the following stages: preprocessing of input images, feature extraction, training, classification, and database. Preprocessing 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. 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.

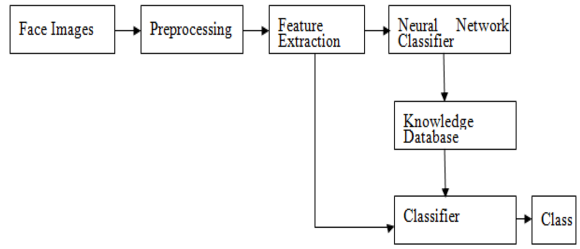


Figure 1 Architecture of the proposed method

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.

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.

Figure 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.

Each image is considered as a bunch of matrices of a fixed size and taken average of each matrix in the image, so that the image is reduced to lower dimension.

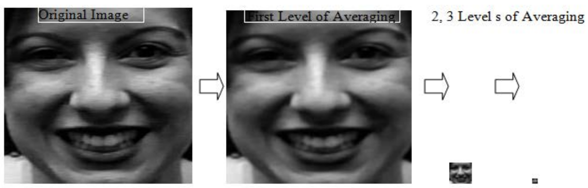


Figure 3 Pattern averaging

Facial Component analysis:

A 2-D facial image can be represented as 1-D vector by concatenating each row (or column) into a long thin vector. Let's suppose we have M vectors of size N (= rows of image * columns of image) representing a set of sampled images. pj's represent the pixel values. The images are mean centered by subtracting the mean image from each image vector. Let m represent the mean image.

Our goal is to find a set of ei's which have the largest possible projection onto each of the wi's. We wish to find a set of M orthonormal vectors ei for which the quantityis maximized with the ortho normality constraint

It has been shown that the ei's and¸ i's are given by the eigenvectors and eigenvalues of the covariance matrix C = WWT where W is a matrix composed of the column vectors wi placed side by side. The size of C is N * N which could be enormous. A common theorem in linear algebra states that the vectors ei and scalars¸ i can be obtained by solving for the eigenvectors and eigenvalues of the M *M matrix WTW. Let di and µi be the eigenvectors and eigenvalues of WTW, respectively.

By multiplying left to both sides by W, which means that the first (M – 1) eigenvectors ei and eigenvalues ꭓi of WWT are given by Wdi and µi, respectively. Wdi needs to be normalized in order to be equal to ei. Since we only sum up a finite number of image vectors, M, the rank of the covariance matrix cannot exceed M -1 (The -1 come from the subtraction of the mean vector m).

The eigenvectors corresponding to nonzero eigen values of the covariance matrix produce an orthonormal basis for the subspace within which most image data can be represented with a small amount of error. The eigenvectors are sorted from high to low according to their corresponding eigenvalues. The eigenvector associated with the largest eigenvalue is one that reflects the greatest variance in the image.

That is, the smallest eigenvalue is associated with the eigenvector that finds the least variance. They decrease in exponential fashion, meaning that the roughly 90% of the total variance is contained in the first 5% to 10% of the dimensions. A facial image can be projected onto M′ (<< M) dimensions by computing where vi = eTi wi. vi is the ith coordinate of the facial image in the new space, which came to be the principal component. The vectors ei are also images, so called, eigenimages, or eigenfaces in our case. They can be viewed as images and indeed look like faces. So, Ω describes the contribution of each eigenface in representing the facial image by treating the eigenfaces as a basis set for facial images. The simplest method for determining which face class provides the best description of an input facial image is to find the face class k that minimizes the Euclidean distance Where Ωk is a vector describing the kth face class. If Ɛk is less than some predefined threshold θe, a face is classified as belonging to the class k.

Once the eigenfaces have been computed, several types of decision can be made depending on the application. What we call face recognition is a broad term which may be further specified to one of following tasks:

- Identification where the labels of individuals must be obtained,
- Recognition of a person, it must be decided if the individual has already been seen,
- Categorization where the face must be assigned to a certain class.

Since a face is well represented by the face space, its reconstruction should be similar to the original; hence the reconstruction error will be small. Non-face images will have a large reconstruction error which is larger than some threshold θr. The distance Єk determines whether the input face is near a known face.

Figure 4 Original Face Images of Cohn Kanade Database

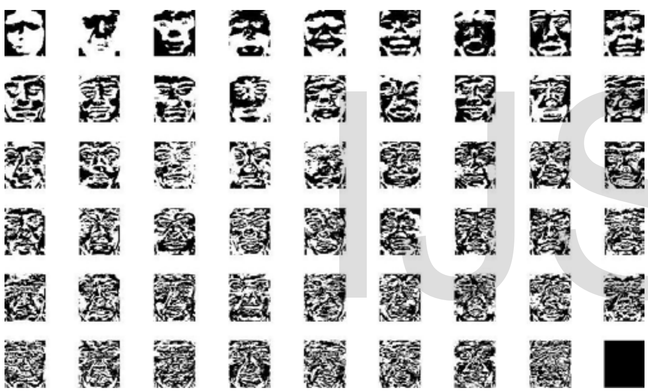


Figure 5 Eigen Faces of the Original Images

Feed Forward Neural Network:

This architecture specifies the classification with neural network using pattern averaging of input images. The training images were taken and applied the pattern averaging. The remaining features are input to the feed forward neural network to train the network. 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.6.

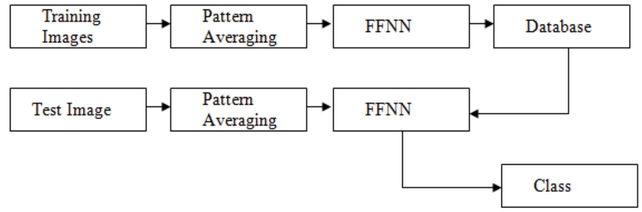


Figure 6 Architecture of existing FFNN classification with Pattern Averaging

Pattern Averaging with Emotional Back Propagation Neural Network:

This architecture proposes the classification with neural network using pattern averaging of input images. The training images were taken and applied the pattern averaging. The remaining features are input to the emotional back propagation neural network to train the network. 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 7.

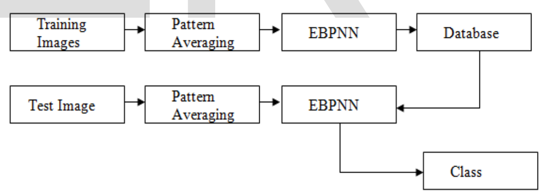


Figure 7 Proposed architecture of EBPNN classification with pattern averaging

PCA with Emotional Back Propagation Neural Network:

This architecture specifies the classification with neural network using PCA applied on input images. The training images were taken and applied the PCA. The feature vectors collected from the PCA are input to the emotional back propagation neural network to train the network.

The neural network will produce the knowledge database. In the process of testing, the test input image will be

ISSN 2229-5518

applied PCA and the gained 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 8.

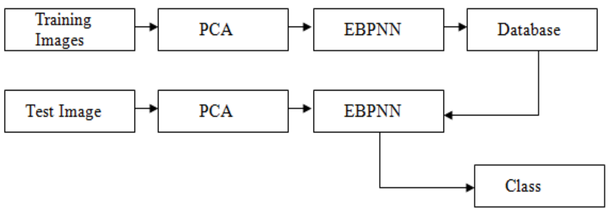


Figure 8 Proposed architecture of EBPNN classification with PCA

The proposed three architectures uses the emotional back propagation neural network architecture. The generalized architecture of the proposed system is shown in Fig 9.

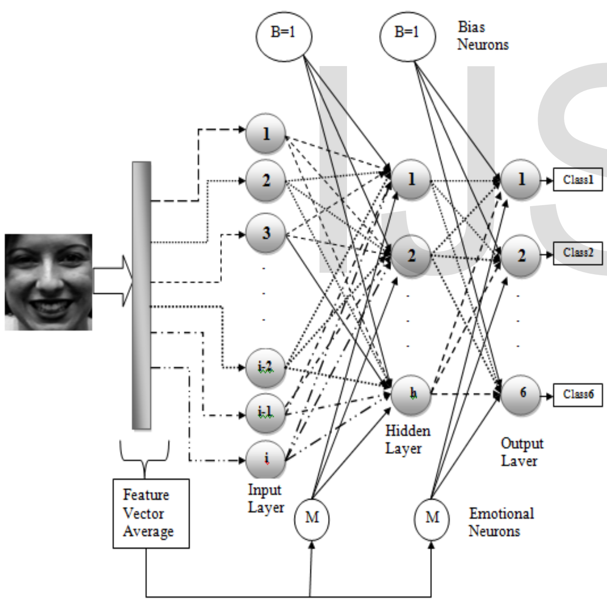


Figure 9 Generalized EmBP-based neural network

3. 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. Sample images from the Cohn Kanade database are shown in Fig.4.

It is observed that PCA with back propagation neural network is giving good results than feature averaging (FeaAVG) with back propagation neural network; where as PCA with Emotional back propagation neural network is giving better performance than FeaAVG with emotional back propagation or simple back propagation neural networks, either with PCA or FeaAVG. The tests performed are person dependent tests. The variation in the images of testing and training is more in the database.

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

TRAIN SAMPLES	TEST SAMPLES	FeaAVG FFNN	PCA FFNN	FeaAVG EBPNN	PCA EBPNN
2	8	75	79	77	77
3	7	85	82	74	78
4	6	80	82	91	84
5	5	92	74	86	80
6	4	84	87	93	91
7	3	96	89	84	90
8	2	98	95	98	98

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 PCA with emotional back propagation neural network is 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.

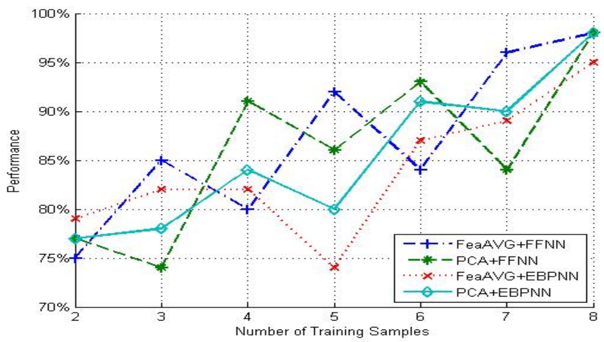


Figure 10 Plot against various results

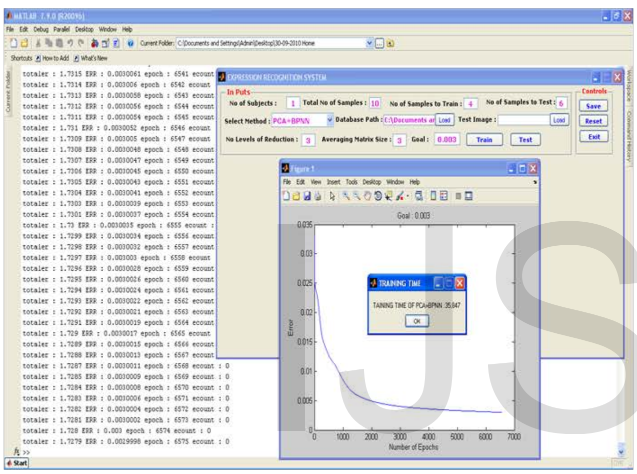
The confusion matrix is created for each of the test. The test is performed on five subjects.

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

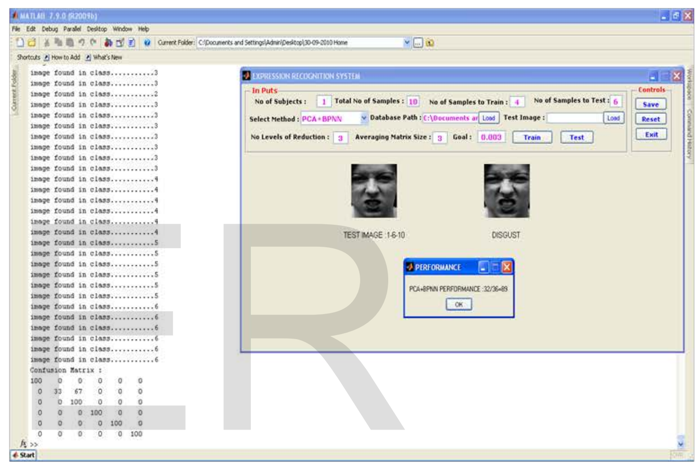
Figure 11 FeaAVG+FFNN Confusion Matrix

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

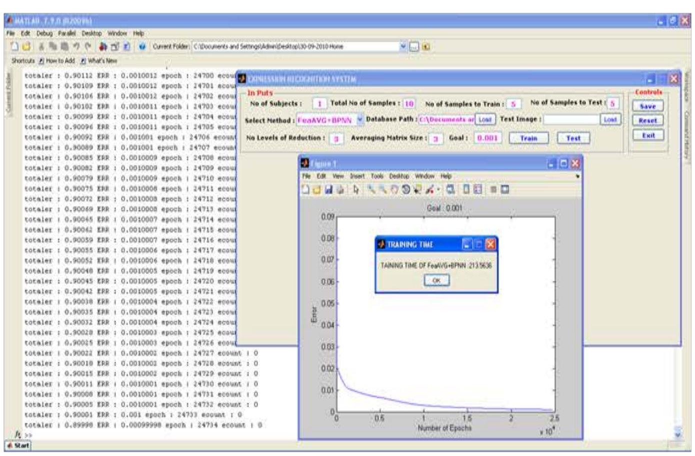
Training of PCA with Emotional BPNN and Error minimization plot:



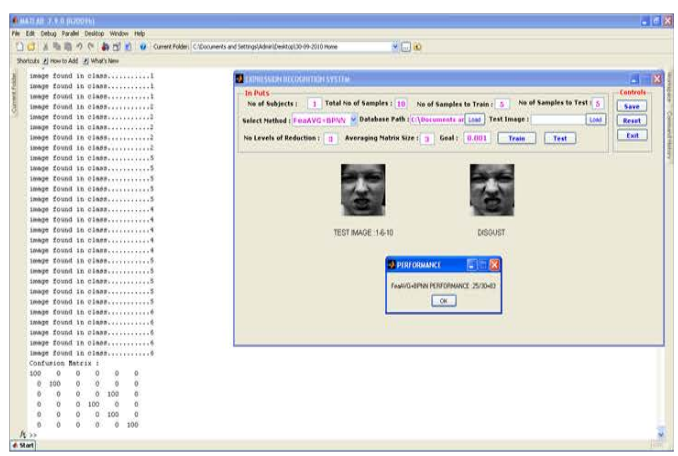
Testing of PCA with Emotional BPNN and Performance:



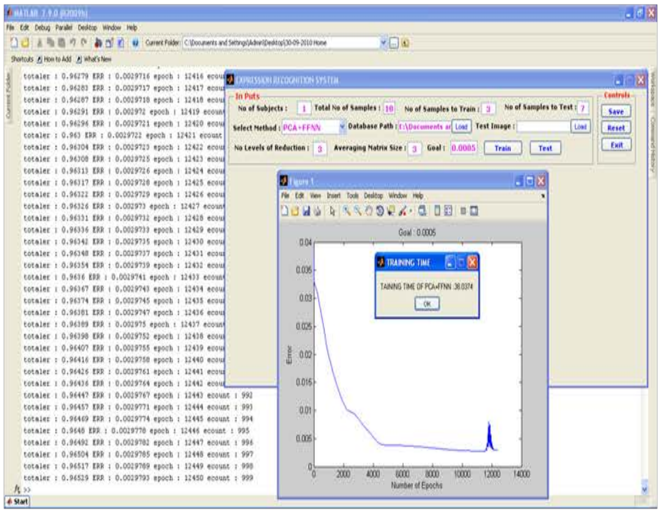
Training of FeaAVG with EBNN and Error minimization plot:



Testing of FeaAVG with EBPNN and Performance:



Training of PCA with BPNN and Error minimization plot:



Testing of PCA with BPNN and Performance:

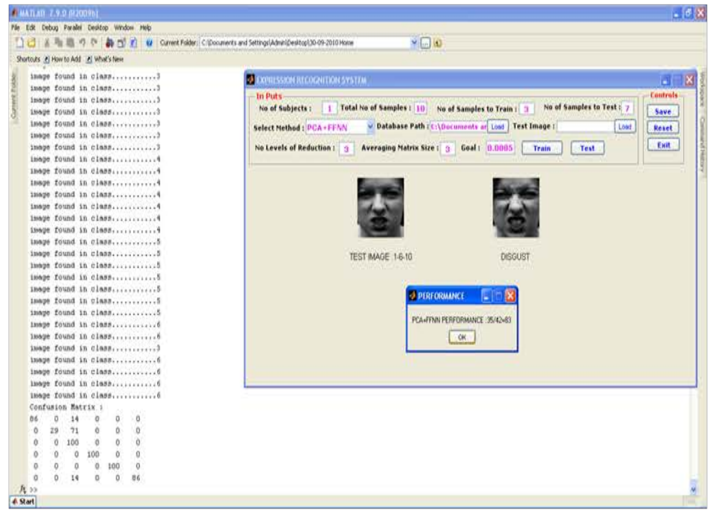


Figure 12 Experimental Results

All other elements are the mis-classifications. The test and training results of various facial emotional classification methods is shown in Fig.12. Experimental results show that the proposed architecture improves the performance of the facial expressions. Based on the results we can conclude that the proposed emotional back propagation neural network with principal component analysis is best in both cases of minimization of training time of neural network and performance as well. 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.

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.

Acknowledgements

The authors would like to thank the anonymous reviewers for their comments which were very helpful in improving the quality and presentation of this paper.

References:

[1] B.Toth, “Biometric Liveness Detection”, Information Security Bulletin, October 2005.

[2] S. De Geef, P. Claes, D. Vandermeulen, W. Mollemans, and P.G. Willems, “Large-Scale In-Vivo

Caucasian Facial Soft Tissue Thickness Database for Craniofacial Reconstruction," Forensic Science, vol. 159, no. 1, pp. S126-S146, May 2006.

[3] S. Baker and S.K. Nayar, "Pattern Rejection," Proc. IEEE Conf. Computer Vision and Pattern Recognition, 1996, pp. 544-549.

[4] S.G. Kong, J. Heo, B.R. Abidi, J. Paik, M.A. Abidi, Recent advances in visual and infrared face recognition—a review, Comput. Vision Image Understanding 97 (2005) 103–135.

[5] R.C. Gonzalez and R.E. Woods, Digital Image Processing, 3rd Edition, Prentice Hall, 2002

[6] M. Turk and A. Pentland, "Face recognition using eigenfaces", Proc. IEEE Conf. on Computer Vision and Pattern Recognition, pp. 586-591, 1991.

[7] M. Turk and A. Pentland, "Eigenfaces for Recognition," J. Cognitive Neuroscience, vol. 3, no. 1, 1991.

[8] M. Turk and A. Pentland, "Face Recognition Using Eigenfaces," Proc. IEEE Conf. on Computer Vision and Pattern Recognition, 1991, pp. 586-591.

[9] P.N. Belhumeur, J.P. Hespanha, and D.J. Kriegman, "Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection," European Conf. Computer Vision, 1996, pp. 45-58

[10] Y. Cui, D. Swets, and J. Weng, "Learning-Based Hand Sign Recognition Using SHOSLIF-M," Int'l Conf. on Computer Vision, 1995, pp. 631-636.

Authors Profile:

V.Ramachandran, is a research scholar at Acharya Nagarjuna University, Nambur. He got his 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 retrieval , human vision & pattern recognition. He did several projects in image processing .

B. Sivaiah have received the B.Tech. degree in Computer Science & Engineering from Jawaharlal Nehru Technological University, Hyderabad, Andhra Pradesh, India in 2007 and M.Tech. degree in Computer Science & Engineering from Jawaharlal Nehru Technological University, Kakinada, Andhra Pradesh, India in 2010. Presently he is working as assistant professor in Department of Computer Science & Engineering, Vignan's University, Andhra Pradesh, India. His current research interests include image processing, neural networks, machine learning, intelligent systems, pattern recognition and artificial intelligence.

Dr. E. Srinivasa Reddy, Professor in Computer Science Department of Acharya Nagarjuna University. He did Master of Technology computer science engineering from sir Mokhsagundam Visweswariah University. He did M.S from BITS Pilani. He received his Ph.D. from CSE Department, ANU, Guntur. His research areas of interest are in various domains including Digital Image Processing.