A Facial Expression Recognition Model using Support Vector Machines

Sivaiah Bellamkonda [a*], N.P. Gopalan [b]

[a,b]*Department of Computer Applications, National Institute of Technology, Tiruchirappalli - 620015, India*

Abstract

Facial Expression Recognition (FER) has gained interest among researchers due to its inevitable role in the human computer interaction. In this paper, an FER model is proposed using principal component analysis (PCA) as the dimensionality reduction technique, Gabor wavelets and Local binary pattern (LBP) as the feature extraction techniques and support vector machine (SVM) as the classification technique. The experimentation was done on Cohn-Kanade, JAFFE, MMI Facial Expression datasets and real time facial expressions using a webcam. The proposed methods outperform the existing methods surveyed.

Index Terms: Facial Expression Recognition, Principal Component Analysis, Support vector Machine, Gabor Wavelets, Local Binary Pattern, Machine Vision, Human Computer Interaction.

1. Introduction

Facial expression [1] or emotion is an important nonverbal cue in communication. For the development of intelligent robotics, critical health care, student satisfaction and other application areas, emotional interaction between machine and human became the fundamental basis. So, researchers are intensely trying to improve the accuracy [15-18] of the FER systems.

Ekman and Friesen [2] introduced 6 basic facial expressions (emotions) as shown in Fig. 1, which are Angry, Disgust, Happy, Fear, Sad, Surprise. According to Mehrabian [3], 55% communicative cues can be judged by facial expression, which implies that the facial expression is the major modality in human communication.

* Corresponding author.
E-mail address: sivaiah.bk@gmail.com

Fig.1. Six Basic Expressions from JAFFE Database

Expressions are formed when we stretch or enlarge facial muscles on the face, but in Facial Action Coding System (FACS) [5], each muscle stretch is considered as an Action Unit (AU) where these AUs will form various expressions. It is very challenging to detect AUs and hence facial expressions with these AUs. So, researchers are concentrating either on active patches or whole face, which we call as a holistic approach. There are numerous feature extraction techniques and classification techniques developed. SVM is popular classifier for FER systems where as some researchers are using Neural Networks [15], Hidden Markov Models [19, 20] and even KNNs [21, 22].

2. Related Work

Muzammil Abdulrahman [9] proposed a facial expression recognition model using Gabor wavelet transform along with PCA and LBP. In this paper, Gabor wavelets are used to extract features from images. PCA and LBP algorithms are used as dimensionality reduction techniques. The experiments were carried on JAFFE database.

Yanpeng Liu [1] proposed a FER model using LBP as feature extraction techniques where features are extracted from active facial patches. PCA is used as a dimensionality reduction technique in this model. Softmax regression classifier is used for classification purpose where experimentation was done on the Cohn Kanade facial expression dataset.

Parth Patel [4] has proposed a FER model using PCA and Discrete Wavelet Transform (DWT). These two techniques are used for extracting features from the Face. SVMs are used for classification purpose. Experimentation was done on JAFFE database.

Muzammil Abdulrahman [6] has proposed a model for FER using SVMs. PCA and LBP algorithms were used for extracting features. SVMs are used for classification purpose. Experiments were conducted on both JAFFE and Mevlana University Facial Expression (MUFE) databases.

Mahesh Kumbhar [8] proposed a model for FER using Gabor Wavelet. 2-D Gabor Function is used for extracting the features and PCA is used for dimensionality reduction. Feed-forward neural networks with 32 input, two hidden layers, 40 to 60 hidden neurons and four output neurons have been used for classification purpose. Experiments were conducted on JAFFE database.

This paper is organized into six sections. The part discusses about the importance and motivation of Facial Expression Recognition. The second section deals with literature review. The third section is the proposed methodology for FER. The fourth part will be the implementation of the proposed method. The fifth section is about results and discussion. The sixth section which is the last section concludes the paper.

3. Methodology

In this paper, we introduced a model with the combination of the three methods such as Gabor wavelets, PCA and LBA for FER task. Here, PCA is used for dimensionality reduction then Gabor wavelets and LBP are

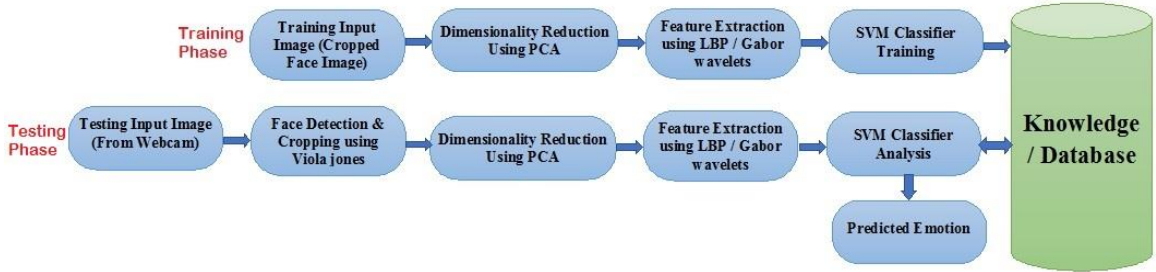used for feature extraction. Fig. 2 shows the proposed method in detail.

Fig.2. System Architecture of the Proposed Method

In the Training phase, the system is to be trained with training samples which are still images. The training input images are already pre-processed such as the face detection and cropping for ease of training purpose. The dimension of the input face image is reduced using PCA. Then Gabor wavelets or Local binary pattern technique is used for extracting the distinct features from the image. The image database is constructed using these features where this database is used for training the SVM classifier. Once the classifier is trained and acquired with the knowledge about the classification pattern, the system is ready to test the real-world expression inputs from the webcam.

In the testing phase, the input image is captured from the webcam for the real-time testing sample. As the image is obtained from the webcam, there would be unwanted detail in the picture such as other objects, background, etc. To remove such noise from the image, we detect the face region from the image and then crop the detected face image. This cropped face image is then processed. The dimension of the input test image is reduced using PCA as we have done in the training phase. Then the features are extracted from the image using Gabor wavelet or Local binary pattern. Then the obtained features are used with the SVM classifier to find the label of the expression.

3.1. Principal Component Analysis

PCA [10-12] can be used as both dimensionality reduction technique and feature extraction technique. It identifies patterns in data which are similarities and differences. Once we find patterns in data, we can compress those data, which act as a dimensionality reduction technique. Assume that there are N images {x1, x2, x3……xN} which can be constructed as a vector of size t. The PCA performs mapping of the original t-sized feature vector onto an f-sized feature sub vector such that f is always smaller than t. The obtained feature vector Yi € Rf is by equation 1:

$$Y_i = W_{PCA}^T X_i (1,2,3....,N) \tag{1}$$

Here W_{PCA} is the transformation matrix and *i* is the number of images. The columns of W_{PCA} are the *f* eigenvectors with the *f* largest Eigen values of the scatter matrix S_r, where S_r can be defined by equation 2:

$$S_r = \sum_{i=1}^{N} (X_i - \mu)(X_i - \mu)^T \tag{2}$$

Where $\mu \in R^t$ which is the mean image of all the images.

3.2. Gabor Wavelets

The specialty of Gabor Wavelets [11-14] is that they extract local features of an image even at different orientations in spatial as well as frequency domains. The GWs finds essential features in an image such as frequency selectivity, orientation selectivity, spatial localization, and quadrature phase relationship. The GW kernel especially in spatial domain uses a Gaussian function. The GWs kernel can be defined by equation 3:

$$\psi(x, y, \varpi, \theta) = \frac{1}{2\pi\sigma^2} e^{-(\frac{x'^2 + y'^2}{2\sigma^2})} e^{i\varpi x'} \tag{3}$$

Where (x, y) is the pixel position, ϖ is the central frequency, θ is the orientation and σ is the standard deviation. The parameters x' and y' can be defined by equation 4 & 5:

$$x' = x\cos\theta + y\sin\theta \tag{4}$$

$$y' = -y\sin\theta + y\cos\theta \tag{5}$$

For example, Fig. 3 shows the magnitude of the Gabor at four scales and the Real Gabor filter bank (GFB) with four different scales and six different orientations.

(a)   (b)
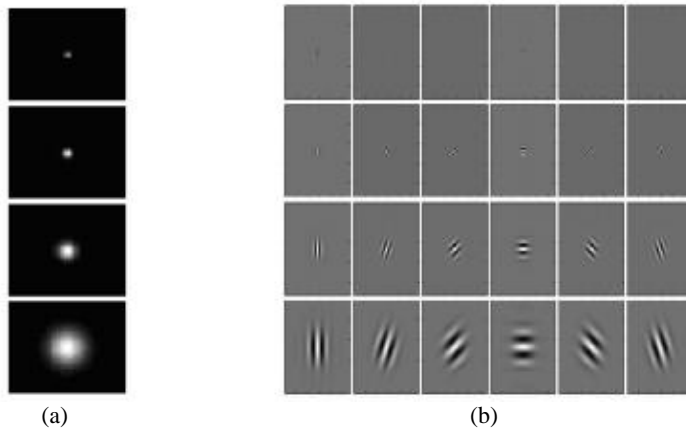
Fig.3. (a) Magnitude of the GFB at four scales; (b) Real part of the Gabor Kernels at Four Scales and Six Orientations.

The GW feature illustration $\varphi(m, n(x, y))$, is obtained by convolution of the GFB $\psi(x, y, \varpi, \theta)$ with input image as given by equation 6:

$$\varphi_{m,n}(x, y) = I(x, y) * \psi(x, y, \varpi, \theta) \tag{6}$$

3.3. Local Binary Patterns

LBP [5, 16] has its roots in 2D texture analysis. LBP will summarize the local pattern of an image by

comparing each pixel with its neighboring pixels. A threshold value is considered to examine all the adjacent pixels of a center pixel to make 1 if the intensity of a pixel is greater than or equal to the threshold or 0. Then the binary pattern is considered as the local image descriptor. This operator was initially considered for 3×3 pixel matrix, resulting 8 bit codes based on the 8 pixels around the central pixel. Fig. 4 gives an example of the basic LBP operator.
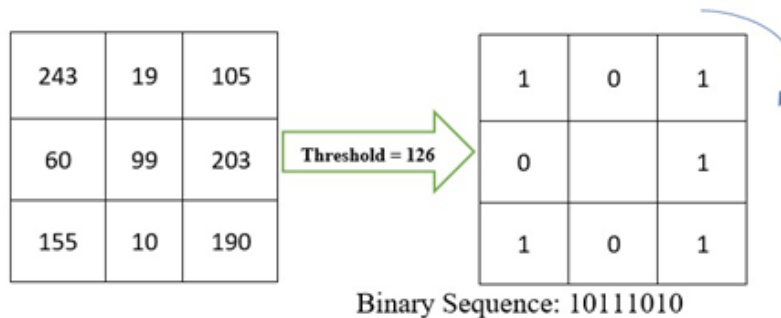

Fig.4. LBP Operator Working Example

3.4. Support Vector Machines

SVM [7] is a supervised learning technique with associated learning algorithms that analyze data for classification or regression. The SVM training algorithm builds a model for a given training data, which will assign to one category or another for new data, making it a non-probabilistic binary linear classifier. When SVM was introduced, it used for two category classification only. A multi-class classification is a combination of two or more two-category classifications.

With a support vector machine, the gap between classes will be maximized as well as the accuracy of classification is also improved. SVM can solve the problems of an inadequate sample of FER and large variance of capacity between different expressions.

FER comes under nonlinear classification problem. SVM may use linear algebra and geometry to separate input data into a high dimensional feature space through a selected nonlinear mapping function. This nonlinear mapping function is nothing but kernel function and a learning algorithm is formed to use the kernel functions. Kernel functions include linear, polynomial, RBF and sigmoid. In this paper, RBF kernel function is used.

It uses two threads which are one-to-many and one-to-one. SVM generates n different classifiers for n different classes. One-to-one thread selects two different classes as one SVM classifier. Then it will generate n × (n − 1)/2 SVM sub-classifiers. FER comes to multi-class classification issue.

4. Experimentation

The proposed model was implemented using MATLAB environment. The testing was conducted on Cohn Kanade, JAFFE and MMI Facial expression data sets. Cohn Kanade database is a benchmark dataset for FER. There are 97 subjects in this database there; all of them are university students ranged in age between18 to 30 years. 65% are female students, 15% are African-American, and 3% are Asian or Latino. Images are saved with pixel sizes of 640x480 of 8-bits for grayscale values, which have been resized to 200x200 pixels, and used for training.

The JAFFE database is also used in the experiment, which contains 210 images of 10 people, having 7 expressions. Each image is saved with a resolution of 256x256. In our experimentation, the original images are used without being altered.

MMI Database contains images of 20 subjects of 31 different expressions for each subject. Participants followed the FACS coding system while capturing expressions. Images in the dataset include all six basic emotions. The images were captured at a resolution of 1200x1900 pixels.

5. Results and Discussion

The accuracy of the system is measured in percentage of correctly classified expression. We can also construct a confusion matrix with the list of emotions that are correctly classified or wrongly classified. Accuracy can be defined as shown in equation 7:

$$Accuracy = \frac{Number-of-Correctly-classified-samples}{Total-number-of-testing-samples} \times 100 \tag{7}$$

The experiments were conducted with bench mark datasets such as Cohn Kanade, JAFFE, and MMI Facial expression datasets. The training images were pre-arranged in respective folders for convenience of reading and constructing database. The testing images are different images than the training images.

Table 1 is the confusion matrix for PCA dimensionality reduction combined with Gabor feature extraction on JAFFE database. The average accuracy of this method is 84.17%.

Table 1. PCA with Gabor Wavelet on JAFFE Dataset

	Anger	Disgust	Fear	Happy	Sad	Surprise
Anger	82	4	3	5	4	2
Disgust	5	85	5	1	2	2
Fear	5	2	79	4	7	3
Happy	3	2	6	85	1	3
Sad	2	3	3	4	86	2
Surprise	3	4	4	1	0	88

Table 2 depicts the confusion matrix of PCA dimensionality reduction combined with Gabor feature extraction on MMI database. The average accuracy of this method is 85.83%.

Table 2. PCA with Gabor Wavelet on MMI Dataset

	Anger	Disgust	Fear	Happy	Sad	Surprise
Anger	88	2	3	2	3	2
Disgust	3	86	3	3	3	2
Fear	3	2	82	4	6	3
Happy	2	3	6	85	1	3
Sad	2	3	3	4	86	2
Surprise	2	4	3	2	1	88

Table 3 is the confusion matrix of PCA dimensionality reduction combined with Gabor feature extraction on Cohn-Kanade database. The average accuracy of this method is 93%.

Table 3. PCA with Gabor Wavelet on Cohn-Kanade Dataset

	Anger	Disgust	Fear	Happy	Sad	Surprise
Anger	94	2	1	2	1	2
Disgust	1	92	1	1	1	2
Fear	1	2	94	1	2	0
Happy	2	1	2	92	2	1
Sad	1	1	1	2	93	2
Surprise	1	2	1	2	1	93

Table 4 lists the confusion matrix for PCA dimensionality reduction combined with LBP extraction on JAFFE database. The average accuracy of this method is 86%.

Table 4. PCA with LBP on JAFFE Dataset

	Anger	Disgust	Fear	Happy	Sad	Surprise
Anger	86	3	3	4	2	2
Disgust	4	86	5	1	2	2
Fear	4	2	80	4	7	3
Happy	3	2	5	87	1	2
Sad	1	3	3	3	88	2
Surprise	2	4	4	1	0	89

Table 5 is the confusion matrix for PCA dimensionality reduction combined with LBP extraction on MMI database. The average accuracy of this method is 88.16%.

Table 5. PCA with LBP on MMI Dataset

	Anger	Disgust	Fear	Happy	Sad	Surprise
Anger	88	3	3	2	3	1
Disgust	2	88	3	3	3	1
Fear	4	2	84	4	3	2
Happy	3	2	5	87	1	2
Sad	1	2	3	3	90	2
Surprise	2	3	2	1	0	92

Table 6 is the confusion matrix for PCA dimensionality reduction combined with LBP extraction on Cohn-Kanade database. The average accuracy of this method is 96.83%.

Table 6. PCA with LBP on Cohn-Kanade Dataset

	Anger	Disgust	Fear	Happy	Sad	Surprise
Anger	97	1	1	1	0	0
Disgust	0	98	0	0	1	1
Fear	1	0	96	2	1	0
Happy	1	1	1	95	1	1
Sad	1	0	1	1	97	0
Surprise	0	0	1	1	0	98

The summary of the average performances is listed in Table 7. The obtained results were compared with the existing methods which are reviewed in the literature. The combination of PCA with LBP outperforms well compared to other methods. The PCA dimensionality reduction with LBP feature extraction method works very well with the Cohn-Kanade dataset.

Table 7. Accuracies Obtained in Percentages

Reference	Dataset		
	JAFFE	Cohn-Kanade	MMI
Yanpeng Liu [1]	96.30	-	-
Parth Patel [4]	96.67	-	-
Muzammil [6]	87	-	77
Mahesh [8]	72.50	-	-
Proposed PCA+Gabor	84.17	93.00	85.83
Proposed PCA+LBP	88.00	96.83	88.16

Fig. 5 shows the comparative analysis in a graphical representation of the accuracies obtained by the proposed model and the existing models. The graph clearly shows that the proposed model has improved the performance than the existing methods, particularly for Cohn Kanade dataset, the Local Binary Pattern out performs on all the methods.
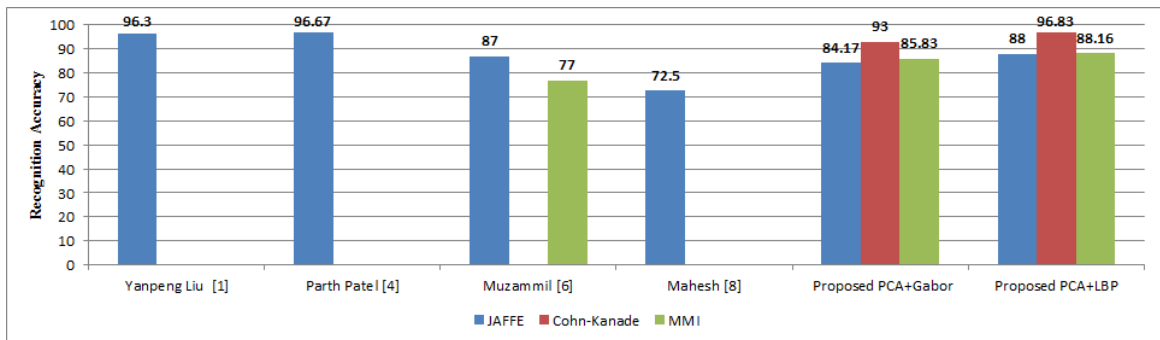

Fig.5. Comparison of Obtained Accuracy with Existing Models

6. Conclusion

In this paper, PCA is used for dimensionality reduction, which drastically improves the processing speed of the overall system. Then Gabor wavelets and Local Binary Patterns were used as feature extraction techniques, which improved the accuracy of the proposed system with the most distinguishable features of the expressions. SVM is used as the classifier which gives a better classification over other classifier algorithms. The overall setup of the proposed model improved the accuracy when compared with existing models. The proposed model gives an average accuracy of 84.17% for JAFFE using Gabor wavelets, 93.00% for MMI using Gabor wavelets, 85.83% for Cohn Kanade using Gabor wavelets, 88.00% for JAFFE using LBF, 88.16% for MMI using LBF and 96.83% for Cohn Kanade using LBF feature extraction techniques, where all these methods commonly used PCA as dimensionality reduction and SVM as a classifier.

References

[1]   Yanpeng Liu, Yuwen Cao, Yibin Li, Ming Liu, Rui Song, "Facial Expression Recognition with PCA and LBP Features Extracting from Active Facial Patches", IEEE International Conference on Real-time Computing and Robotics June 6-9, 2016, Angkor Wat, Cambodia, pp. 368-341, 2016.
[2]   P. Ekman, and W. Friesen, "Facial Action Coding System: A Technique for the Measurement of Facial Movements", Consulting Psychologists Press, California, 1978.
[3]   Mehrabian.A, "Communication without Words", Psychology Today, Vo1.2, No.4, pp. 53-561968.
[4]   Parth Patel, Khushali Raval, "Facial Expression Recognition Using DWT-PCA with SVM Classifier", International Journal for Scientific Research & Development, Vol. 3, Issue 03, pp. 1531-1537, 2015.
[5]   Yuan Luo, Cai-ming Wu, Yi Zhang, "Facial expression recognition based on fusion feature of PCA and LBP with SVM", Optik - International Journal for Light and Electron Optics Volume 124, Issue 17, pp. 2767-2770, 2013.
[6]   Muzammil Abdurrahman, Alaa Eleyan, "Facial expression recognition using Support Vector Machines", 23rd IEEE Conference on Signal Processing and Communications Applications Conference (SIU), 16-19 May 2015, Malatya, Turkey, 2015.
[7]   Anushree Basu, Aurobinda Routray, Suprosanna Shit, Alok Kanti Deb, "Human emotion recognition from facial thermal image based on fused statistical feature and multi-class SVM", 2015 Annual IEEE India Conference (INDICON), 17-20 Dec. 2015, New Delhi, India, 2015.
[8]   Mahesh Kumbhar, Manasi Patil, Ashish Jadhav, "Facial Expression Recognition using Gabor Wavelet", International Journal of Computer Applications, Volume 68, No.23, PP. 13-18, 2013.
[9]   Muzammil Abdulrahman, Tajuddeen R. Gwadabe, Fahad J. Abdu, Alaa Eleyan, "Gabor Wavelet Transform Based Facial Expression Recognition Using PCA and LBP", IEEE 22nd Signal Processing and Communications Applications Conference (SIU 2014), Trabzon, Turkey, 23-25 April 2014, pp. 2265 - 2268, 2014.
[10]   Shilpa Sharma, Kumud Sachdeva, "Face Recognition using PCA and SVM with Surf Technique", International Journal of Computer Applications, Volume 129, No.4, pp. 41-47, 2015.
[11]   Vinay A., Vinay S. Shekhar, K. N. Balasubramanya Murthy, S. Natarajan, "Face Recognition using Gabor Wavelet Features with PCA and KPCA - A Comparative Study, 3rd International Conference on Recent Trends in Computing 2015 (ICRTC-2015), Procedia Computer Science 57, pp. 650–659, 2015.
[12]   Anurag De, Ashim Sahaa, Dr. M.C Pal, "A Human Facial Expression Recognition Model based on Eigen Face Approach", International Conference on Advanced Computing Technologies and Applications (ICACTA-2015), pp. 282-289, 2015.
[13]   Liangke Gui, Tadas Baltrusaitis, and Louis-Philippe Morency, "Curriculum Learning for Facial Expression Recognition", 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017), Washington, DC, USA, 2017.

[14] Zhiming Su, Jingying Chen, Haiqing Chen, "Dynamic facial expression recognition using autoregressive models", 7th International Congress on Image and Signal Processing (CISP), 14-16 Oct. 2014, Dalian, China, 2014.

[15] Mao Xu, Wei Cheng, Qian Zhao, Li Ma, Fang Xu, "Facial Expression Recognition based on Transfer Learning from Deep Convolutional Networks", 11th IEEE International Conference on Natural Computation (ICNC), 15-17 Aug. 2015, Zhangjiajie, China, 2015.

[16] Pan Z., Polceanu M., Lisetti C., "On Constrained Local Model Feature Normalization for Facial Expression Recognition", In: Traum D., Swartout W., Khooshabeh P., Kopp S., Scherer S., Leuski A. (eds) Intelligent Virtual Agents. IVA 2016, Lecture Notes in Computer Science, Vol. 10011. Springer, Cham, 2016.

[17] Shaoping Zhu, "Pain Expression Recognition Based on pLSA Model", The Scientific World Journal, Volume 2014, 2014.

[18] Myunghoon Suk and Balakrishnan Prabhakaran, "Real-time Mobile Facial Expression Recognition System – A Case Study", 2014 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 23-28 June 2014, Columbus, OH, USA, 2014.

[19] Zineb Elgarrai, Othmane El Meslouhi, Mustapha Kardouchi, Hakim Allali, "Robust facial expression recognition system based on hidden Markov models", International Journal of Multimedia Information Retrieval, Volume 5, Issue 4, pp. 229–236, 2016.

[20] Arnaud Ahouandjinou, Eug`ene Ezin, Kokou Assogba, Cina Motamed, Mikael Mousse, Bethel Atohoun, "Robust Facial Expression Recognition Using Evidential Hidden Markov Model", https://hal.archives-ouvertes.fr/hal-01448729, 2017.

[21] Guo Y., Zhao G., Pietikäinen M., "Dynamic Facial Expression Recognition Using Longitudinal Facial Expression Atlases", In Fitzgibbon A., Lazebnik S., Perona P., Sato Y., Schmid C. (eds) Computer Vision – ECCV 2012, Lecture Notes in Computer Science, vol. 7573. Springer, Berlin, Heidelberg, 2012.

[22] Nikunj Bajaj, S L Happy, Aurobinda Routray, "Dynamic Model of Facial Expression Recognition based on Eigen-face Approach", Proceedings of Green Energy and Systems Conference, Long Beach, CA, USA, 25 November 2013.

Authors' Profiles

Dr. N.P. Gopalan is Professor at Department of Computer Applications, National Institute of Technology, Tiruchirappalli, Tamil Nadu, India. He obtained his Ph.D. from Indian Institute of Science, Bangalore, India. His research interests are in Data Mining, Distributed Computing, Cellular Automata, Theoretical Computer Science, Image Processing and Machine Intelligence.

B. Sivaiah is pursuing Ph.D. at Department of Computer Applications, National Institute of Technology, Tiruchirappalli. He obtained his B. Tech. and M. Tech in Computer Science & Engineering from Jawaharlal Nehru Technological University, Hyderabad, Andhra Pradesh, India in 2007 and 2010 respectively. His current research interests include image processing, neural networks, machine learning, and intelligent systems.