

# Face Detection using Neural Network & Gabor Wavelet Transform

Avinash Kaushal<sup>1</sup>, J P S Raina<sup>2</sup>

<sup>1</sup>GCET, Greater Noida, U.P., India; <sup>2</sup>BBSBEC, Fatehgarh Sahib, Punjab, India

E-mail : kaushalpinkie@rediffmail.com

**Abstract :** This paper is based on classifications of the features of a face detected using Gabor filter feature extraction techniques in image processing. The feature vector based on Gabor filters used as the input of the classifier, which is a Feed Forward Neural Network (FFNN) on a reduced feature subspace learned by an approach simpler than Principal Component Analysis (PCA). The effectiveness of the proposed method is demonstrated by the experimental results on testing large number of images and comparisons with state of the art method. Face detection and recognition has many applications in a variety of fields such as security systems, video conferencing and identification.

**Keywords :** Face detection, Gabor wavelet, Feed Forward Neural Network (FFNN) classifier, Multilayer perceptron

## I. Introduction

Human face detection and recognition is an active area of research spanning several disciplines such as image processing, pattern recognition and computer vision. Face detection and recognition are preliminary steps to a wide range of applications such as personal identity verification, video-surveillance, lip tracking, facial expression extraction, gender classification, advanced human and computer interaction. Most methods are based on neural network approaches, feature extraction, Markov chain, skin color, and others are based on template matching [1]. Pattern localization and classification is the step, which is used to classify face and non-face patterns. Many systems dealing with object classification are based on skin color. In this paper we are interested by the design of an ANN algorithm in order to achieve image classification. This paper is organized as follows: In section II, we give an overview over classification for face detection. Description of our model is discussed in Section III. Section IV deals with the training method.

## II. Classification for face detection

While numerous methods have been proposed to detect face in a single image of intensity or color images. A related and important problem is how to evaluate the performance of the proposed detection methods [1]. Many recent face detection papers compare the performance of several methods, usually in terms of detection and false alarm rates. It is also worth noticing that many metrics have been adopted to evaluate algorithms, such as learning time, execution time, the number of samples required in training, and the ratio between detection rates and false alarms. In general, detectors can make two types of errors: false negatives in which faces are missed resulting in low detection rates and false positives in which an image is declared to be face.

False negative =  $\frac{\text{Number of Missed Faces}}{\text{Total Number of Actual faces}}$

False Positive =  $\frac{\text{Number of Incorrect Detected Faces}}{\text{Total Number of Actual faces}}$

Face detection can be viewed as two-class Recognition problem in which an image region is classified as being a "Face" or "non Face". Consequently, face detection is one of the few attempts to recognize from images a class of objects for which there is a great deal of within-class variability. Face detection also provide interesting challenges to the underlying pattern classification and learning techniques. The class of face and no face image are decidedly characterized by multimodal distribution function and effective decision boundaries are likely to be nonlinear in the image space. Pattern localization and classification are CPU time intensive being normally implemented in software, however with lower performance than custom implementations. Custom implementation in hardware allows real-time processing, having higher cost and time-to-market than software implementation. Some works [2, 3, 4] uses ANN for classification, and the system is implemented in software, resulting in a good performance (10 sec for localization and classification). A similar work is presented in [5], aiming to object localization and classification.

We are interested in the implementation of an ANN algorithm & design of a Gabor filter in order to provide better image classification. The MLP (Multi-layer Perceptron) algorithm is used to classify face and non-face patterns before the recognition step.

## III. Multi-layers Perceptron

The MLP neural network [1] has feed forward architecture within input layer, a hidden layer, and an output layer. The input layer of this network has N units for an N dimensional input vector. The input units are fully connected to the I hidden layer units, which are in turn, connected to the J output layers units, where J is the number of output classes. A Multi-Layers Perceptron (MLP) is a particular of artificial neural network [6]. We will assume that we have access to a training dataset of I pairs  $(x_i, y_i)$  where  $x_i$  is a vector containing the pattern, while  $y_i$  is the class of the corresponding pattern. In our case a 2-class task,  $y_i$  can be coded 1 and -1.

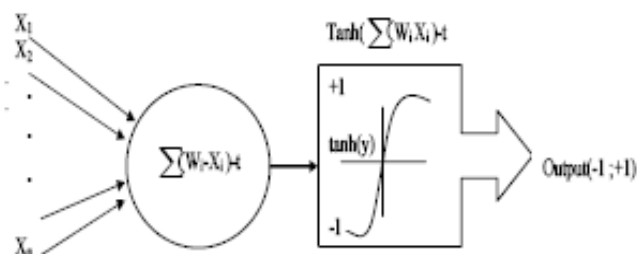


Fig.1 The neuron of supervised training [8]

We considered a MLP (Multi-Layers Perceptron) with a 3 layers, the input layer is a vector constituted by  $n^2$  units of neurons ( $n \times n$  pixel input images). The hidden layer has  $n$  neurons, and the output layer is a single neuron which is active to 1 if the face is presented and to otherwise.

The activity of a particular neuron  $j$  in the hidden layer is written by

$$S_j = \sum_{i \in \text{input}} w_{ji} x_i, x_i = f(s_j) \quad (1), f$$

a sigmoid function.

Where  $W_{ji}$  is the set of weights of neuron  $i$ ,  $b_1(i)$  is the threshold and  $x_i$  is an input of the neuron.

Similarly the output layer activity is:

$$S_j = \sum w_{ji} x_i$$

In our system, the dimension of the retina is  $27 \times 18$  pixels represent human faces and non-face, the input vector is constituted by 2160 neurons, the hidden layer has 100 neurons.

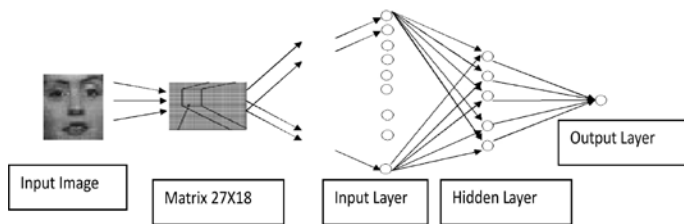


Fig.2 (a) Architecture of proposed System

We are designing a feed forward neural network with one hundred neurons in the hidden layer and one neuron in the output layer which prepares images for training phase. All data form both “face” and “non-face” folders will be gathered in a large cell array. Each column will represent the features of an image, which could be a face, or not.

Rows are as follows:

- Row 1: File name
- Row 2: Desired output of the network corresponded to the feature vector.
- Row 3: Prepared vector for the training phase

We will adjust the histogram of the image for better contrast. Then the image will be convolved with Gabor filters by multiplying the image by Gabor filters in frequency domain. To save time they have been saved in frequency domain before Features is a cell array contains the result of the convolution of the image with each of the forty Gabor filters. These matrices have been concatenated to form a big  $135 \times 144$  matrix of complex numbers. We only need the magnitude of the result. That is why “abs” is used.  $135 \times 144$  has 10,400 pixels. It means that the input vector of the network would have 19,400 values, which means a large amount of computation. So we have reduced the matrix size to one-third of its original size by deleting some rows and columns. Deleting is not the best way but it save more time compare to other methods like PCA. We should optimize this function as possible as we can.

First training the neural network and then it will return the trained network. The examples were taken from the Internet database. The MLP will be trained on 500 face and 200 non-face examples.

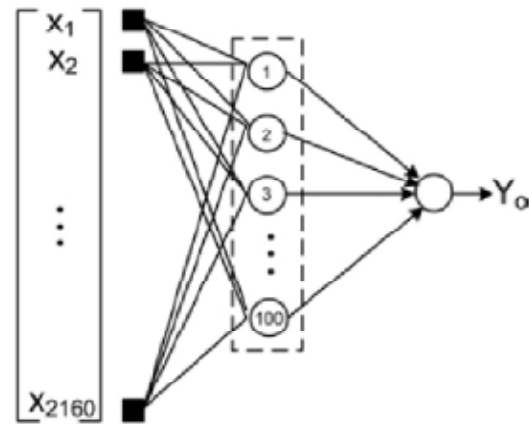


Fig. 2(b) Architecture of proposed system

**IV. Training Methodology**

The MLP with the training algorithm of feed propagation is universal mappers, which can in theory, approximate any continuous decision region arbitrarily well. Yet the convergence of feed forward algorithms is still an open problem. It is well known that the time cost of feed forward training often exhibits a remarkable variability. It has been demonstrated that, in most cases, rapid restart method can prominently suppress the heavy-tailed nature of training instances and improve efficiency of computation.

Multi-Layer Perceptron (MLP) with a feed forward learning algorithms was chosen for the proposed system because of its simplicity and its capability in supervised pattern matching. It has been successfully applied to many pattern classification problems [9]. Our problem has been considered to be suitable with the supervised rule since the pairs of input-output are available. For training the network, we used the classical feed forward algorithm. An example is picked from the training set, the output is computed.

**V. Algorithm Development & Results**

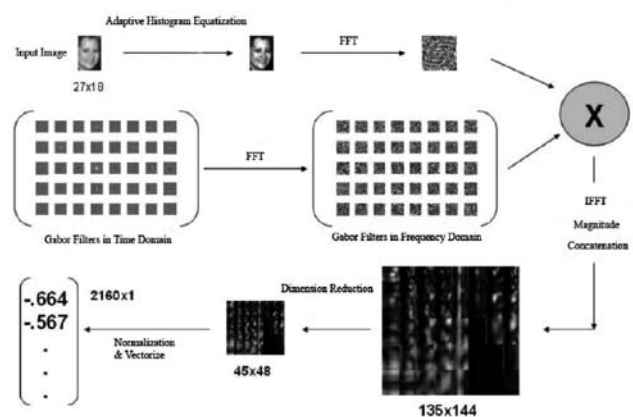


Fig. 4 Steps involved in Feature extraction

**VI. 2D Gabor Wavelet Representation of Faces**

Since face recognition is not a difficult task for human beings, selection of biologically motivated Gabor filters is well suited to this problem. Gabor filters, modeling the responses of simple cells in the primary visual cortex, are simply plane waves restricted by a Gaussian envelope function [18].

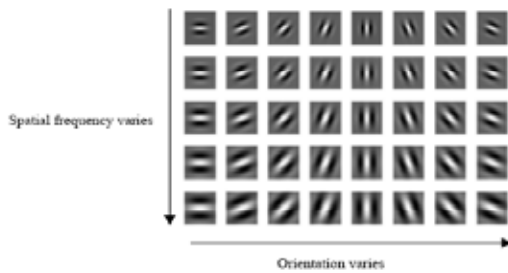


Fig. 5: Gabor filters correspond to 5 spatial frequencies and 8 Orientation (Gabor filters in Time domain)

An image can be represented by the Gabor wavelet transform allowing the description of both the spatial frequency structure and spatial relations. Convolving the image with complex Gabor filters with 5 spatial frequency ( $\nu = 0, \dots, 4$ ) and 8 orientation ( $\mu = 0, \dots, 7$ ) captures the whole frequency spectrum, both amplitude and phase (Fig. 5). In Fig. 6, an input face image and the amplitude of the Gabor filter responses are shown below.

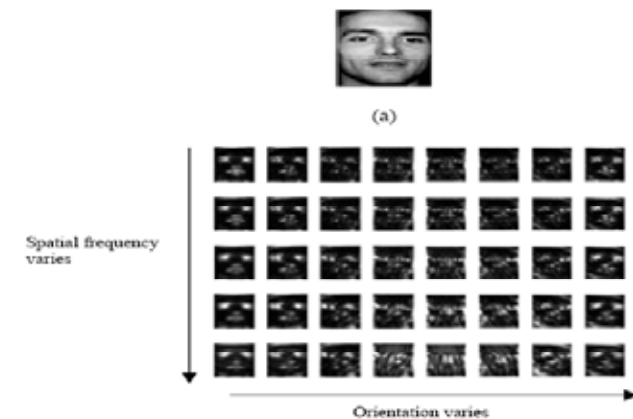


Fig. 6 (a & b): Example of a facial image response to above Gabor filters, 6(a) original

face image (from internet database), and b) filter responses. One of the techniques used in the literature for Gabor based face recognition is based on using the response of a grid representing the facial topography for coding the face. [19, 20, 21, 22]. Instead of using the graph nodes, high-energized points can be used in comparisons which forms the basis of this work. This approach not only reduces computational complexity, but also improves the performance in the presence of occlusions.

**A. Feature extraction**

Feature extraction algorithm for the proposed method has two main steps in (Fig. 8):

- (1) Feature point localization
- (2) Feature vector computation.

**1. Feature point localization**

In this step, feature vectors are extracted from points with high information content on the face image. In most feature-based methods, facial features are assumed to be the eyes, nose and mouth [14,16]. However, we do not fix the locations and also the number of feature points in this work. The number of feature vectors and their locations can vary in order to better represent diverse facial characteristics of different faces, such as dimples, moles, etc., which are also the features that people might use for recognizing faces (Fig. 7).



Fig. 7: Facial feature points found as the high-energized points of Gabor wavelet responses

From the responses of the face image to Gabor filters, peaks are found by searching the locations in a window  $W_0$  of size  $W \times W$  by the following procedure:

A feature point is located at  $(x_0, y_0)$ , if

$$R_j(x_0, y_0) = \max_{(x,y) \in W_0} (R_j(x, y))$$

$$R_j(x_0, y_0) > \frac{1}{N_1 N_2} \sum_{x=1}^{N_1} \sum_{y=1}^{N_2} R_j(x, y),$$

(3-4)  
 $j=1, \dots, 40$

where,  $R_j$  is the response of the face image to the  $j$ th Gabor filter  $N_1 N_2$  is the size of face image, the center of the window,  $W_0$  is at  $(x_0, y_0)$ . Window size  $W$  is one of the important parameters of proposed algorithm, and it must be chosen small enough to capture the important features and large enough to avoid redundancy. Equation (4) is applied in order not to get stuck on a local maximum, instead of finding the peaks of the responses. In our experiments a  $9 \times 9$  window is used to search feature points on Gabor filter responses. A feature map is constructed for the face by applying above process to each of 40 Gabor filters.

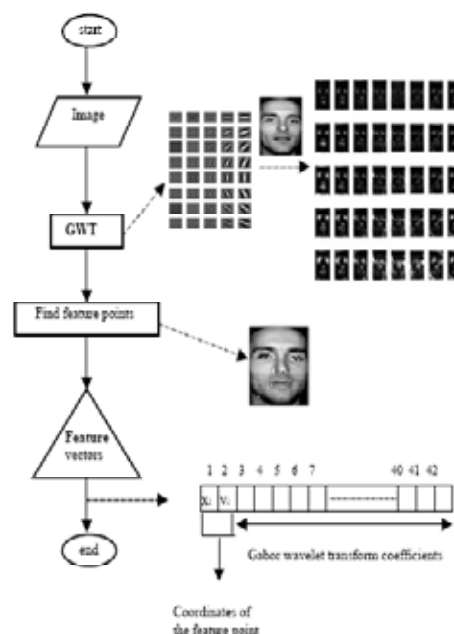


Fig. 8: Flowchart of the feature extraction stage of the facial images

**2. Feature vector generation**

Feature vectors are generated at the feature points as a composition of Gabor wavelet transform coefficients.  $k$ th

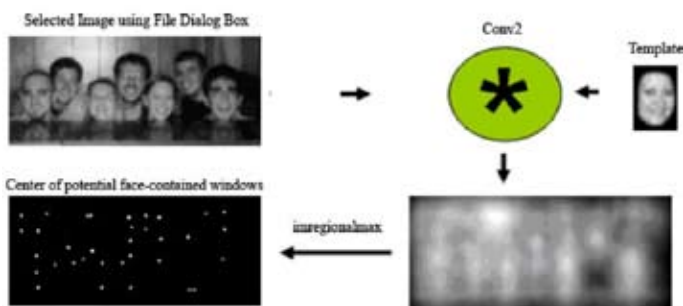
feature vector of  $i$ th reference face is defined as,

$$v_{i,k} = \{x_k, y_k, R_{i,j}(x_k, y_k) \mid j = 1, \dots, 40\}$$

While there are 40 Gabor filters, feature vectors have 42 components. The first two components represent the location of that feature point by storing  $(x, y)$  coordinates. Since we have no other information about the locations of the feature vectors, the first two components of feature vectors are very important during matching (comparison) process. The remaining 40 components are the samples of the Gabor filter responses at that point. Although one may use some edge information for feature point selection, here it is important to construct feature vectors as the coefficients of Gabor wavelet transform. Feature vectors, as the samples of Gabor wavelet transform at feature points, allow representing both the spatial frequency structure and spatial relations of the local image region around the corresponding feature point.

**VII. Results**

**A. First Section**



**B. Second Section**

In this section the algorithm will check all potential face-contained windows and the windows around them using neural network. The result will be the output of the neural network for checked regions.



Fig. 10.1 Cell .net

**C. Third Section**



Fig. 10.2

**1. Filtering above pattern for values above threshold**

**(xy\_)**



Fig10.3

**2. Dilating pattern with a disk structure (xy\_)**



Fig10.4

**3. Finding the center of each region**



Fig. 10.5

**4. Draw a rectangle for each point**

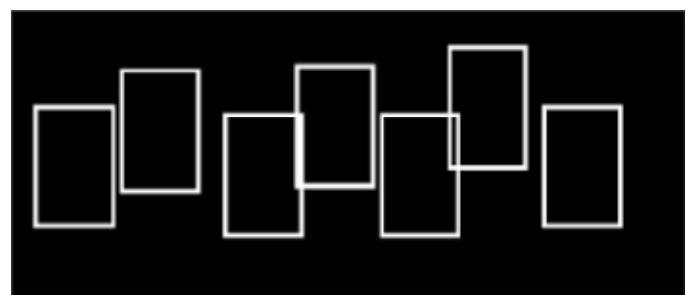


Fig10.6

**5. Final Result will be like this-**

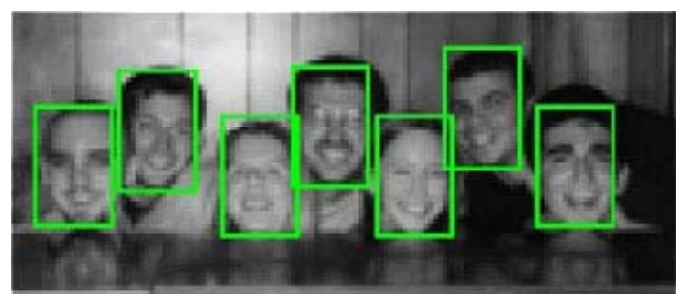


Fig 10.7



This architecture will be implemented using Matlab in a graphical environment allowing face detection in a database. It will be evaluated using the test data of 500 images containing faces, on this test set we obtained a good detection.

### VIII. Conclusion

Face detection has been an attractive field of research for both neuroscientists and computer vision scientists. Since 1988, many algorithms have been proposed as a solution to automatic face detection. Although none of them could reach the human detection performance, currently two biologically inspired methods, namely eigenfaces and elastic graph matching methods, have reached relatively high detection rates. Eigenfaces algorithm has some shortcomings due to the use of image pixel gray values. As a result system becomes sensitive to illumination changes, scaling, etc. and needs a beforehand pre-processing step [7,8]. Satisfactory recognition performances could be reached by successfully aligned face images. When a new face attends to the database system needs to run from the beginning, unless a universal database exists. Unlike the eigenfaces method, elastic graph matching method is more robust to illumination changes, since Gabor wavelet transforms of images is being used, instead of directly using pixel gray values. Although detection performance of elastic graph matching method is reported higher than the eigenfaces method, due to its computational complexity and execution time, the elastic graph matching approach is less attractive for commercial systems [21]. Although using 2-D Gabor wavelet transform seems to be well suited to the problem, graph matching makes algorithm bulky. Moreover, as the local information is extracted from the nodes of a predefined graph, some details on a face, which are the special characteristics of that face and could be very useful in detection task, might be lost [10,11,12,17]. In this paper, a new approach to face detection with Gabor wavelets & feed forward neural network is presented. The method uses Gabor wavelet transform & feed forward neural network for both finding feature points and extracting feature vectors. From the experimental results, it is seen that proposed method achieves better results compared to the graph matching and eigenfaces methods, which are known to be the most successive algorithms.

Although the proposed method shows some resemblance to graph matching algorithm, in our approach, the location of feature points also contains information about the face. Feature points are obtained from the special characteristics of each individual face automatically, instead of fitting a graph that is constructed from the general face idea. In the proposed algorithm, since the facial features are compared locally, instead of using a general structure, it allows us to make a decision from the parts of the face. For example, when there are sunglasses, the algorithm compares faces in terms of mouth, nose and any other features rather than eyes. Moreover, having a simple matching procedure and low computational cost proposed method is faster than elastic graph matching methods [15,18]. Proposed method is also robust to illumination changes as a property of Gabor wavelets, which is the main problem with the eigenfaces approaches. A new facial image can also be simply added by attaching new feature vectors to reference gallery while such an operation might be quite time consuming for systems that need training. Feature points, found from Gabor responses of the face image, can give small deviations between different conditions (expression, illumination, having glasses

or not, rotation, etc.), for the same individual. Therefore, an exact measurement of corresponding distances is not possible unlike the geometrical feature based methods [13]. Moreover, due to automatic feature detection, features represented by those points are not explicitly known, whether they belong to an eye or a mouth, etc. Giving information about the match of the overall facial structure, the locations of feature points are very important. However using such a topology cost amplifies the small deviations of the locations of feature points that are not a measure of match. Gabor wavelet transform of a face image takes 1.1 seconds, feature extraction step of a single face image takes 0.2 seconds and matching an input image with a single gallery image takes 0.12 seconds on a Pentium IV PC. Note that above execution times are measured without code optimization.

### References

- [1] Yang M.H., Kriegman D.J., and Ahuja N., "Detecting Faces in Images: A Survey", IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol.24, No.1, January 2002.
- [2] Rowley H. A., Baluja S. , Kanade T., "Neural Network-Based Face Detection", IEEE Trans. On Pattern Analysis and Machine Intelligence, vol.20, No. 1, Page(s). 39-51, 1998.
- [3] ZhenQiu Z., Long Z., Li S.Z., Jiang Z.H., "Real-time multi-view face detection", Proceeding of the Fifth IEEE International Conference on automatic Face and Gesture Recognition, pp: 142-147, 20-21 May 2002.
- [4] Gavrilu, D.M; Philomin, V., "Real-Time Object detection for Smart Vehicles". International Conference on Computer Vision (ICCV99). Vol. 1. Corfu, Greece, 20-25 September 1999.
- [5] Molz R.F., Engel P.M., Moraes F.G., Torres L., Robert M., "System Prototyping dedicated to Neural Network Real-Time Image Processing", ACM/SIGDA ninth international Symposium On Field Programmable Gate Arrays (FPGA 2001).
- [6] Theocharides T., Link G., Narayanan V., Irwin M.J., "Embedded Hardware Face Detection", 17th International Conference on VLSI Design, Mumbai, India, January 5-9, 2004.
- [7] Yang F. and Paindavoine M., "Prefiltering for pattern Recognition Using Wavelet Transform and Face Recognition", The 16th International Conference on Microelectronics, Tunisia 2004.
- [8] Yang F. and Paindavoine M., "Prefiltering for pattern Recognition Using Wavelet Transform and Neural Networks", Advances in imaging and Electron physics, Vol. 127, 2003.
- [9] Li X. and Areibi S., "A Hardware/Software co-design approach for Face Recognition", The 16th International Conference on Microelectronics, Tunisia 2004.
- [10] Gajski D., Dutt v, Wu A., "High-Level Synthesis: Introduction to Chip and System Design", Kluwer Academic Publishers, Boston, 1992.
- [11] Kanade T., "Picture processing by computer complex and recognition of human faces", Technical report, Kyoto University, Dept. of Information Science, 1973.
- [12] Lin S., Kung S. , and Lin L. , "Face Recognition / Detection by Probabilistic Decision-Based Neural Network", IEEE Transactions on Neural Networks, Vol.8, pp.114-132, 1997.
- [13] Brunelli R., Poggio T., "Face Recognition: Features vs. Templates", IEEE Trans. on PAMI, Vol. 12, No. 1, Jan.

- 1990.
- [14] Yang M. H., Ahuja N. , and Kriegman D., “A survey on face detection methods”, IEEE Transactions on Pattern Analysis and Machine Intelligence, to appear 2001.
- [15] Ranganath S. and Arun K., “Face Recognition Using Transform Features and Neural Network”, Pattern Recognition, Vol. 30, pp. 1615-1622, 1997.
- [16] Lawrence S. , Giles C., Tsoi A. and Back A., “Face Recognition: A Convolutional Neural Network Approach”, IEEE Trans. on Neural Networks, vol. 8, pp. 98-113, 1997.
- [17] Turk M. and Pentland A. . “Eigenfaces for recognition”, Journal of Cognitive Science, pp.71-86, 1991.
- [18] Perkins D. , “A definition of caricature and recognition”, Studies in the Anthropology of Visual Communication, Vol. 2, pp. 1-24, 1975.
- [19] Wiskott L., Fellous J. M. , Krüger N. and Malsburg C.V.D., “Face Recognition by Elastic Graph Matching”, In Intelligent Biometric 110 Techniques in fingerprint and Face Recognition, CRC Press, Chapter 11, pp. 355-396, 1999.
- [20] Lyons M. J. , Budynek J. and Akamatsu S., “Automatic Classification of Single Facial Images”, IEEE Trans. on PAMI, Vol. 21, No.12, December 1999.
- [21] Krüger N., Pöttsch M. and Malsburg C.V.D., “Determining of Face Position and Pose with a learned Representation Based on Labeled Graphs”, Image and Vision Computing, Vol. 15, pp.665-673, 1997.
- [22] Duc B. , Fisher S. , and Bigün J. , “Face Authentication with Gabor Information on Deformable Graphs”, IEEE Transactions on Image Proc., vol.8, no.4, pp.504-515, 1999.



Avinash Kaushal was born on 3rd November, 1981. She has received her B.Tech. in Electronics and Instrumentation Engineering and M.Sc. in Information Technology. She is Pursuing her M.Tech.. Currently, she is working as a Lecturer in Galgotias College of Engg. & Technology, Greater Noida, U.P., India. She has a Teaching Experience of 6 years.