Scientific Journal Impact Factor (SJIF): 1.711
International Journal of Modern Trends in Engineering and Research
ISSN (ONLINE) : 2349 - 9745
ISSN (PRINT) : 2393 - 8161

Face Recognition Using Singular Value Decomposition and Hidden Markov Model

Prof. Dr. P.D.Bhamre, Miss. Swati.B.Memane

[1,2]*Department of E&TC, K.K.Wagh college of Engineering, Nashik, Maharashtra, India.*

Abstract—Biometrics is an automatic method of recognizing a person based on his physiological characteristic. Face recognition system (FRS) is an emerging field in biometrics. FRS can be used for identifying a person in highly secured areas as well as where person authentication is required; it is one of the best methods because face structure cannot be copied like a password. In most of the face recognition system front pose face images are considered when the database is created. But when the query image is not its front face image but side view of the face or any other pose or facial expression other than the image in the database, the system fails to identify the person. In this paper a method to solve this problem is given. The database contains the facial images of different persons with different facial expressions and poses. The features are extracted using Singular value decomposition method (SVD). A HMM (Hidden Markov Model) is used to train these features of the images from the given set of data. These resultant HMM's then compared with the HMM of the available face image to identify the person. This method gives good results compared to other 2D face recognition systems which are affected by pose variations and intensity of the 2D image. The system is tested using ORL standard database and the algorithm for this system is simulated using MATLAB software.

Keywords—Face recognition system, Hidden markov model, Singular value decomposition, ORL Database, Yale database.

I. INTRODUCTION

Biometrics is a science and technology of measuring and analyzing biological data of human body. In computer security, biometrics refers to measurable characteristics of human body that can be used for identification of a person. There are several types of biometric identification schemes such as face, fingerprints, hand geometry, signature, voice etc. A matcher compares the normalized signature with the set of normalized signatures in the system's data base and provides a similarity score that compares the individual's normalized signature with each signature in the database set [1]. Face as a biometric feature has number of advantages over the other biometric features such as, it is the main biometric feature that humans use to recognize one another. Face recognition has the advantage of being universal over other major biometric features because a face image can be easily captured without the persons knowledge such as using a surveillance camera, whereas fingerprints or other biometric features such as iris image, palm prints are captured with much more difficulty and that cannot be captured with a good quality without the persons co-operation. In forensics applications the images of the criminals available are not proper or with different angles and expressions called as Mug shots. These images are most of the time captured by a surveillance camera and are not proper. It then becomes difficult to identify the person. To recognize faces across poses, 3-D approaches are used as given in [2], the most well-known 3D approach for face recognition is the 3-D morphable model given in [2]. 3D methods are considered as an effective method for identifying persons with different poses but 3D images increases the computational complexity of the system. Reliability is an important factor in highly secured areas like defence; hence such model can be used in these areas in spite of increased complexity but in general purpose use such high complexity system is not affordable, so a system with accuracy equal to 3D system and less complex than the 3D system has been developed using 2D images.

II. ORGANIZATION OF PAPER

The rest of this paper is organized as follows: section 3 gives the literature review of different types of systems. Section 4 explains the details of the proposed method of face recognition system. Section 5 covers the results obtained. Section 6 provides a conclusion and gives recommendations for future study.

III. LITERATURE SURVEY

Face recognition is an important research topic, this is because face recognition has numerous practical applications such as bankcard identification, access control, Mug shots searching, security monitoring, and surveillance system [3]. Face recognition system can be formulated as: given an image, identify or verify one or more persons in the image by comparing with faces stored in a database [3]. Some image processing techniques extract feature points such as eyes, nose, and mouth and use it as input data for the application. Various approaches have been proposed to extract these facial points from the images. This section gives an overview of various methods of face recognition.

A. Eigen Faces

Eigen face is one of the most widely investigated approaches to face recognition. In reference [4] principal component analysis is used to efficiently represent pictures of faces. It is mentioned that any face images can be approximately reconstructed by a small collection of weights for each face and a standard face picture (Eigen picture). The weights describing each face are obtained by projecting the face image onto the Eigen picture. The method proposed in reference [5] uses Eigen faces, which was motivated by the technique of Kirby and Sirovich given in [4], for face detection and identification. In mathematical terms, Eigen faces are the principal components of the distribution of faces, or the eigenvectors of the covariance matrix of the set of face images. The authors reported 96 percent, 85 percent, and 64 percent correct classifications averaged over lighting, orientation and size variations, respectively. The authors explained the robust performance of the system under different lighting conditions by significant correlation between images with changes in illumination. However, [6] showed that the correlation between images of the whole faces is not sufficient for satisfactory recognition performance. Reference [7] extended their early work on eigenface to eigenfeatures corresponding to face components, such as eyes, nose, and mouth. They used a modular Eigen space which was composed of the above Eigen features like Eigen eyes, Eigen nose, and Eigen mouth. This method was less sensitive to appearance changes than the standard eigenface method. The system achieved a recognition rate of 95 percent on the FERET database of 7,562 images of approximately 3,000 individuals. Eigen face appears as a fast, simple and practical method. However, in general, it does not provide invariance over changes in scale and lighting conditions.

B. Geometry based technique

Geometrical feature matching techniques are based on the computation of a set of geometrical features from the face image. The total geometrical configuration is described by a vector representing the position and size of the main facial features, such as eyes and eyebrows, nose, mouth and the shape of face outline [8]. In geometry based technique features are extracted using the size and the relative position of important components of images. First the direction and edges of important component is detected and then feature vectors from these edges and direction are prepared. In some methods the grayscales difference of unimportant components and important components are found, the important components of a face are eyes and eyebrows, nose and mouth. By using this data, Haar-like feature block is used in Adboost method [9] to change the grayscales distribution into the feature, these features are then further used for comparison. Geometrical feature matching technique used by Kanade in [10] is based on the extraction of a set of geometrical features forming the picture of a face. Their system achieved 75 percent recognition rate on a database of twenty persons, using two images per person; one for training and the other for test. Mark Nixon presented a geometric measurement for eye spacing with the Hough transform technique to detect the instance of a circular shape and of an ellipsoidal shape. The result of this paper illustrate that it is possible to derive a measurement of the spacing by detection of the position of both the iris [11]. All these techniques require image to be properly captured. Hence the image used for these methods must have the cooperation of the person whose image is to be captured, which in many cases is not possible. These techniques require a threshold for comparison, which may adversely affect the achieved performance.

C. Template based Matching

In template matching based techniques the test image is represented as a two-dimensional array of intensity values and is compared using a suitable metric, such as the Euclidean distance, with a single template representing the whole face. There are many other ways of template matching on face recognition. More than one face template from different viewpoints can also be used to represent an individual's face. A face from a single viewpoint can also be represented by a set of multiple distinctive smaller templates [12]. Another technique described will extract facial feature based on the previously designed templates using appropriate energy function and the best match of template in facial image yield the minimum energy. Methods have been proposed by Yuille in [13], detecting and describing features of faces using deformable templates. In deformable templates the feature of interest, for example an eye, is described by a parameterized template. These parameterized templates enable a priori knowledge about the expected shape of the features to guide the detection process. An energy function is defined to link peaks, edges and valleys in the image intensity with corresponding properties of the template. After that the template matching is done with the image, by altering its parameter values to minimize the energy function, thereby deforming itself to find the best fit. For the descriptor purpose final parameter value is used. In the template based eye and mouth detection, first an eye template is used to detect the eye from image. Then a correlation is found out between the eye templates with various overlapping regions of the face image. Eye region has a maximum correlation with the template. The Yuille, Fischler and Elschlager in [13], measured features automatically and described a linear embedding algorithm that used local feature template matching and a global measure of fit to find and measure facial features. These algorithms require a priori template modeling, in addition to their computational costs, which clearly affect their performance. Genetic algorithms have been proposed for more efficient searching times in template matching.

D. Appearance based Technique

This approach processes the image as two dimensional patterns. The concept of feature in this approach is different from simple facial features such as eyes and mouth. Any extracted characteristic from the image is referred to a feature. This method found the best performance in facial feature extraction because it keep the important information of image and reject the redundant information. Method such as principal component analysis (PCA), independent component analysis (IDA), DCT or Singular Value Decomposition (SVD) is used to extract the feature vector. The main purpose of doing this is to reduce the large dimensionality of observed variable to the smaller dimensionality of independent variable without losing much information. This technique would be later the foundation of the proposal of many new face recognition algorithms [14]. In PCA analysis high order dependencies exist and this is the disadvantage of this method because much information may contain in the high order relationship. While other method ICA uses technique independent component analysis which not only uses second-order statistics but also uses high order statistics. In DCT the set of coefficients is very large and hence it becomes complex to analyze and store such data. Whereas compared to all these methods SVD has less number of important coefficients and gives best result with low complexity. Some problems with IDA method is that it requires image matrices to be transformed into vectors, which are usually of very high dimensionality and this causes high computational cost and complexity. This approach uses skin color to isolate the face area from the non face area in an image. Any non-skin color region within the face is viewed as a candidate for eyes or mouth [14]. The performance of such techniques on facial image databases is rather limited [15].

IV. PROPOSED METHODOLOGY

In the previous section we have discussed about various methods of face recognition, this section gives us the detail information about the Singular value decomposition and Hidden markov model which is used in the proposed method of face recognition.

A. Singular value decomposition

Singular value decomposition (SVD) can be looked at from three compatible points of view. On one hand, one can see it as a method for transforming correlated variables into a set of uncorrelated ones that better expose the various relationships among the original data items. At the same time, SVD is a method for identifying and ordering the dimensions along which data points exhibit the most variation. The third way of viewing SVD is that once we have identified where the most variation is, it is possible to find the best approximation of the original data points using fewer dimensions. Hence, SVD can be seen as a method for data reduction. The basic idea behind SVD is taking a high dimensional, highly variable set of data points and reducing it to a lower dimensional space that exposes the substructure of the original data more clearly and orders it from most variation to the least. SVD is based on a theorem from linear algebra which says that a rectangular matrix 'A' can be broken down into the product of three matrices- an orthogonal matrix 'U', a diagonal matrix' S', and the transpose of an orthogonal matrix 'V' . The theorem is usually presented as given in equation.

$$A_{mxn}=U_{mxm}\,S_{mxn}\,V^{T}_{nxn} \qquad (4.1)$$

Where, $U^TU=I$ & $V^TV=I$
The columns of U are orthonormal eigenvectors of AA^T, columns of V are orthonormal eigenvectors of A^TA and S is a diagonal matrix containing the square roots of eigen values from U or V in descending order.
An image with 64x64 resolutions has an S matrix with same dimensions, which means 64 distinct singular values. It has been shown that the energy and information of a signal is mainly conveyed by a few big singular values and their related vectors. To select some of these coefficients as feature, a large number of combinations of the singular values and other elements of matrices U & V were evaluated in [3]. It was found that SVD contains three matrixes (U, S and V) use of two first coefficients of matrix S and first coefficient of matrix U as three features $U_{11}S_{22}S_{11}$ associating each block have the best classification rate[3].

B. Hidden Markov Model

Hidden Markov Models (HMMs) are widely used in pattern recognition applications, most notably for speech recognition but can be extended to face recognition systems. A process in the HMM class can be described as a finite-state Markov Chain with a memory less output process which produces symbols in a finite alphabet. However, from the perspective of an observer who knows the parameters of some representation of the process and is able to observe the output symbols but not the internal states, things look different. For some processes there are infinitely many distinct states of such an observer's knowledge about the status of the process. This knowledge is defined in terms of conditional distributions on future symbols. This is the sense in which there can be infinitely many states. These states are more relevant than the original finite set of states. Hence to study the process, these states are selected since they allow for optimal prediction [4]. Hidden Markov Model is a Markov Chain with an associated output mechanism which takes either states or transitions between states to either symbols or distributions on symbols. Hidden Markov Models are useful in modeling one dimensional data in face finding, object recognition and face recognition. HMM is associated with non-observable hidden states and an observable sequence generated by the hidden states individually. The elements of a HMM are:

N= S is the number of states in the model, where S= $\{s_1, s_2,.., s_N\}$ is the set of all possible states. M= V is the number of the different observation symbols, where V= $\{v_1, v_2,...., v_M\}$is the set of all possible observation symbols. Each observation vector is a vector of observation symbols of length T. T is defined by user. A= $\{a_{ij}\}$ is the state transition probability matrix, where:

B = $b_j(k)$ is the observation symbol probability matrix, where;

$$b_j(k)=P[o_t=v_k \,|q_t=s_j],\ 1\leq j\leq N,\ 1\leq k\leq M \qquad (4.2)$$

$\pi = \{\pi_1, \pi_2, \pi_3, ...,\pi_N\}$ is the initial state distribution, where

$$\pi_i = P[q_1 = s_i] \quad 1\leq N \qquad (4.3)$$

Using shorthand notation HMM is defined as $\lambda = (A,B, \Pi)$ HMMs generally work on sequences of symbols called observation vectors, while an image usually is represented by a simple 2D matrix. In the case of using a one dimensional HMM in face recognition problems, the recognition process is based on a frontal face view where the facial regions like hair forehead, eyes, nose and mouth come in a natural order from top to bottom. In this project we divided image faces into seven regions and each is assigned to a state in a left to right one dimensional HMM is as shown in fig 4.1 [4]. The main advantage of the above model is its simple structure and flexibility for adjustment of parameters. Hence as seen before singular value decomposition is a method for transforming correlated variables into a set of uncorrelated ones that better expose the various relationships among the original data items. It is a method for identifying and ordering the dimensions along which data points exhibit the most variation. This shows that once identified where the most variation is, it's possible to find the best approximation of the original data points using fewer dimensions. These are the basic ideas behind SVD: taking a high dimensional, highly variable set of data points and reducing it to a lower dimensional space that exposes the substructure of the original data more clearly and orders it from highest variation to the least. Hence, it shows that SVD is a good method for data reduction. Therefore the sequence of tokens generated by an HMM gives some information about the sequence of states. The image divided into blocks acts as the states for the HMM as shown in fig 4.2 [3]

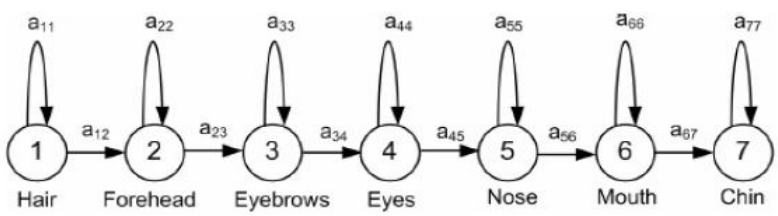

Fig.4.1. 7 state HMM Model

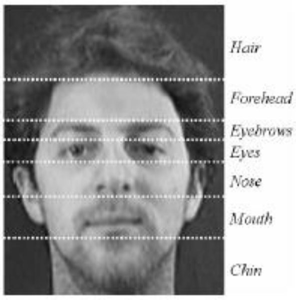

Fig 4.2 Division of face image into 7 states

C. Algorithm of Proposed Method

1. First, the image from the database is obtained and converted to gray scale.
2. Secondly, the image is resized to around 50 percent of its original size. So as to reduce the computational complexity.
3. In order to compensate the flash effect and reduce the salt noise, a nonlinear minimum order static filter is used. The filter has a smoothing role and reduces the image information.
4. After all the preprocessing on the image, the image is divided into blocks having an overlap of around 75 percent. The number of blocks extracted from each image is given by [3]:

$$T= \{H-L/ L-P\} + 1 \qquad (4.3)$$

Where, H=height of the image
P=L-1; L= overlap size.
5. SVD for each block obtained in the previous step is calculated and the value for $U_{11,}$ S_{22}, S_{11} are found.
6. Each U, S & V value is then given to the HMM model for training and an HMM model for a person is generated.
7. Similarly HMM model for each person is generated and saved in the database for further comparison.
8. While testing, all the steps are repeated and an HMM model for the query image is found. This model is then compared with the available database to find the probable match.

As given in the algorithm Fig.4.3 shows the block diagram of the proposed method.

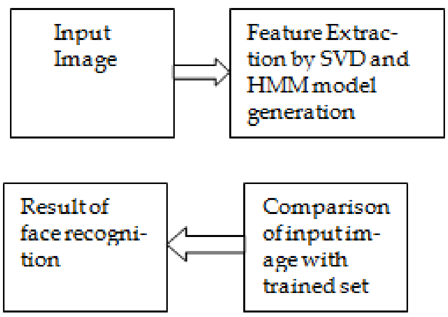

Fig 4.3 Face recognition system block diagram

D. Database

ORL database.

Volume 02, Issue 10, [October – 2015] ISSN (Online):2349–9745 ; ISSN (Print):2393-8161

The ORL (Olivetti Research laboratory) database now-a-days called as 'The database of faces' consists a set of face images taken between April 1992 and 1994. There are 10 different images of each of 40 distinct subjects, named from $s1$ to s_{40}, s_{xy} is the x^{th} image of person y. The images were taken at different times, varying the lightning, facial expressions and facial details like glasses. All the images were taken against a dark h mogeneous background [16]. The images are in PGM format. The size of each image is 112x92 pixels with gray scale. Different images of a single person in the database is as shown in fig.4.4

Fig 4.4 ORL database

V. RESULTS

The image of a person is recognized by its recognition coefficients. The recognition coefficients are found by comparing the features of test image with the HMM model of each person.The matching occurs where the recognition coefficient is higher. The Fig 5.1 below shows the plot of recognition coefficient for each person when the test image is of person $s1$. $s11$, $s22$, $s33$ are the corresponding images of person 1, 2 and 3 respectively. It can be concluded from the plot that when the test image is compared to all other images the recognition coefficient for correct recognition has to be the highest. It can be summarized that for every correct recognition the range of recognition coefficient is roughly around e^-90 to e^-150 and for a wrong matching the coefficients has a difference of e^-100 or greater than that of the correct matching coefficient.

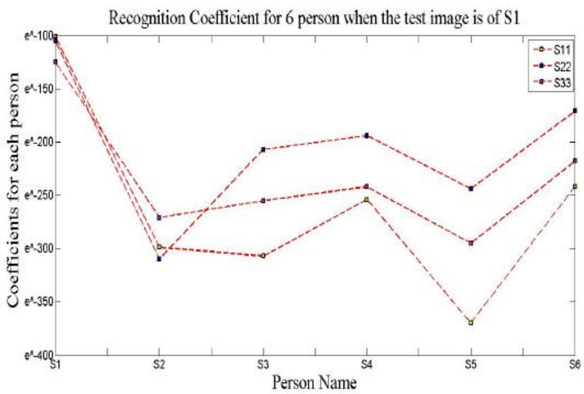

Fig 5.1 Recognition coefficient for 6 person for test image S_1

The system is tested by varying the following parameters and finding the best solutions for highest recognition rate and less computational time.

1. Image size
2. Different Block height.

Image size

Image size is an important parameter, as the image size is greater the time required for computation of parameters will increase. So we have to find an image size where the computation time required is minimum as well as the recognition rate is higher.
It is found that the recognition rate decreases gradually as the image size reduces as shown in fig 5.2. This is because as we go on reducing the image size the image information gets lost hence we get lower recognition rate. The optimal solution can be found out from this graph for a recognition rate above 96 percent can be obtained when the image size is 75X85 or 64X64 with less computational time.

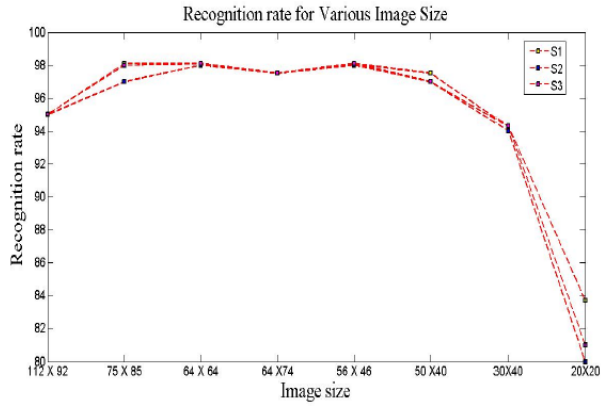

Fig 5.2 Recognition rate for various Image size

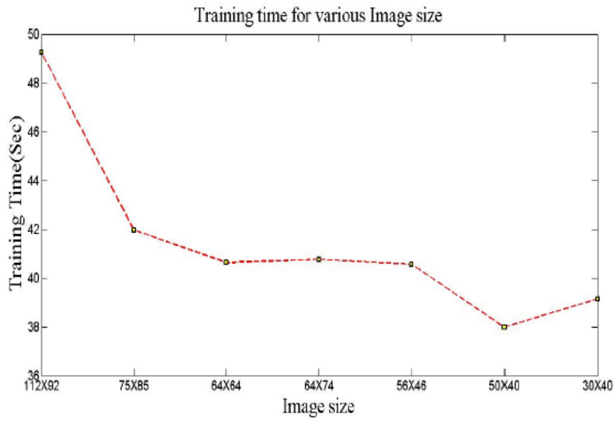

Fig 5.3 Time for various Image size

Block Height
Block height of image and its overlap can also affect the recognition rate we took the results for various block height of the image. The results are taken for fixed image size 75X85 and for a particular training set as shown in fig 5.3 and fig 5.4. From the results it can be concluded that the block height 5 is the best solution when recognition rate and time is considered.

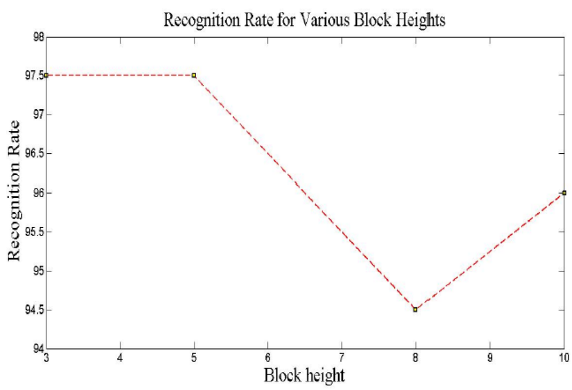

Fig 5.4 Block height vs Recognition rate

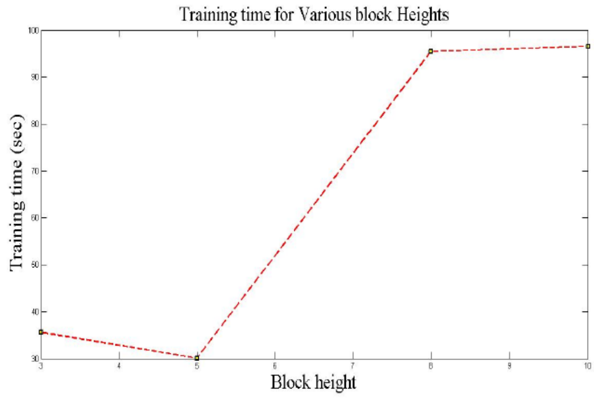

Fig 5.5 Block height vs computational time

Thus the system is analyzed with various parameters such as image size & block height to find the best probable solution. The choice of parameters depends upon the application of the system. If it is used for a security purpose then the parameters which give the highest recognition rate in spite of all the other parameters must be selected but where time and complexity are considered we have to find the optimal solution for it. Considering all the parameters to its optimum value from the study we see that the best solution for ORL database is given by following set of parameters.

Image size	Block height	Recognition Rate (%)	Computational time(sec)
75X85	5	97.5	37.06
75X85	5	98.12	42.16

VI. CONCLUSION AND FUTURE SCOPE

Face recognition applications such as information security, access management, biometrics etc. A fast and efficient system is presented in this paper. Images of each face are converted into a sequence of blocks. Each block is featured by a few number of its SVD parameters. Every person is associated with a hidden markov model. The previous systems designed such as given in [2] attains a recognition rate of 98 % using 2D HMM which makes the system bit complex, also given in [4] the system have recognition rate of 99 percent for Yale database. The proposed system is tested on ORL which has images with different poses. The best recognition rate encountered is 98.1 % percent for ORL. This is obtained with optimized system parameters chosen after a deep study of their influence on recognition rate and training time. Better results can be obtained with higher number of training images and original size of the image. But keeping the original size of the image makes the system complex and slow. Future work can be extended to the use of larger dataset with varying poses as well as different expressions. For such complicated database the same method will not give such efficiency hence the feature extraction method must be worked to be stronger for such database. The HMM can be extended to 2D HMM to improve the results but with increased complexity, a midway has to be found between the increased complexity, time required for training and the recognition rate as the future scope of this project.

REFERENCES

[1] A. S. Tolba, A.H. El-Baz, and A.A. El-Harby, "Face Recognition: A Literature Review", International Journal of Signal Processing 2; Feb 2006.

[2] V.Blanz T.Vetter, "Face Recognition on fitting a 3D morphable Model", IEEE Transactions, Pattern Analysis, Machine Intelligence, Vol. PAMI-25, pp 1063-1074, Sep 2003.

[3] H.Miar-Naimi P.Davari, ”A New Fast And efficient HMM bassed Face Recognition System using a 7 state HMM along with SVD coefficient.”, International Journal of Electronics and Electrical Engg. Vol 4, Jan 2008.
[4] M. Kirby and L. Sirovich, ”Application of the Karhunen- Love procedure for the characterisation of human faces,” IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 12, pp. 831-835, Dec. 1990.
[5] M. Turk and A. Pentland, ”Eigenfaces for recognition,” J. Cognitive Neuroscience, vol. 3, pp. 71-86, 1991.
[6] M.A. Grudin, ”A compact multi-level model for the recognition of facial images,” Ph.D. thesis, Liverpool John Moores Univ., 1997.
[7] A. Pentland, B. Moghaddam, and T. Starner, ”View-Based and modular eigenspaces for face recognition,” Proc. IEEE CS Conf. Computer Vision and Pattern Recognition,pp. 84-91, 1994.
[8] S. Tamura, H. Kawa, and H. Mitsumoto, ”Male/Female identification from 8 X 6 very low resolution face images by neural network,” Pattern Recognition, vol. 29, pp. 331-335, 1996.
[9] M. Jones and P. Viola,”Face Recognition Using Boosted Local Features”, IEEE International Conference on Computer Vision, 2003.
[10] M. Nixon,” Eye spacing measurement for facial recognition”,. Proceedings of the Society of Photo-Optical Instrument Engineers, SPIE, 575(37):279-285, August 1985.
[11] Yuille, A. L., Cohen, D. S., and Hallinan, P. W., ”Feature extraction from faces using deformable templates”, Proc. of CVPR, (1989).
[12] R. Bruneli and T. Poggio, ”Face recognition: features versus templates,” IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 15, pp. 1042-1052, 1993.
[13] Bhoi, M. N. Mohanty,” Template Matching based Eye Detection in Facial Image”, International Journal of Computer Applications (0975 - 8887) Volume 12- No.5, December 2010.
[14] M. Turk and A. Pentland, ”Eigenfaces for recognition” Journal of Cognitive Neuroscience, 3(1):71-86, 1991
[15] J.Lim, Y. Kim J. Paik” Comparative Analysis of Wavelet- Based Scale-Invariant Feature Extraction Using Different Wavelet Bases” International Journal of Signal Processing, Image Processing and Pattern Recognition Vol. 2, No. 4, December, 2009.
[16] http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase