Face Recognition Sub-division under Health Monitoring of Pc Users

Sanjana Vijaykumar Kalyanappagol
Department of Electronics and Communication
BMS College of Engineering, Bengaluru, India

Srirama N
Department of Electronics and Communication
BMS College of Engineering, Bengaluru, India

Ravilla Nithya Sree
Department of Electronics and Communication
BMS College of Engineering, Bengaluru, India

Abstract — This project mainly deals with the health of PC users and gamers who works on PC for long hours. The goal of this sub project is mainly to evaluate various face detection and face recognition methods and to provide a complete solution for recognizing the faces with high accuracy and very good response rate.

Keywords- Face detection, Face recognition, Viola Jones Algorithm, Hidden Markov Models, 2D-DCT.

I. Introduction

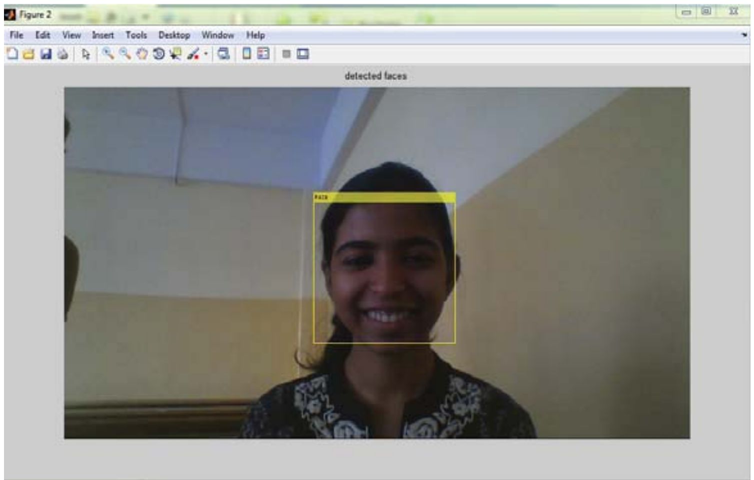
Now-a-days so many PC users are facing the problems regarding musculoskeletal affections due to sedentary and repetitive works. This problem has been detected by World Health Organization also. Our project aims at solving this problem to some extent.

The sub-topic is Face recognition and detection of faces in an image. Thanks to the advancement in image processing, which made our work easier. A depth camera such as webcam or laptop camera with sensors is used to track the facial points and position, and orientation data.

Face recognition in this main project is essential to maintain a database that keeps track of health of the various PC users in an institution or organization. Maintaining a database is necessary as we may need it in future for analysis. This is a substitute to the traditional manual logging system i.e., signing in record books. This project can be divided into two parts- Face detection and Face recognition. First, let us analyze the algorithms and techniques of face detection.

II. FACE DETECTION

In digital camera technology ‘Face Detection’ is also termed as ‘face priority AF (auto focus). It’s a function of camera that detects faces so that the camera can set the focus and appropriate exposure for the shot automatically [Fig. 1].



[Fig. 1] : Face Detection

Paul Viola and Michael Jones presented a fast and robust method for face detection [Fig 8]. This algorithm was 15 times quicker than any technique at that time of its introduction.

The facial feature extraction algorithm [2][3] is a good example of feature searching, achieving 82% accuracy with invariance to gray and color information, failing to detect faces with glasses and hair covering the forehead. A similar system proposed by Jeng Et Al reported an 86% detection rate [4]

Probabilistic face models based on multiple face appearance have also been used in many systems including Yow and Cipollas model reporting a 92% detection rate [5]. But Viola Jones algorithm tops this list giving an accuracy rate of 93.9% and the detection rate with the least minimal false detections. [Fig 6]

This algorithm provides a frontal face detection system which achieves detection and false positive rates which are equivalent to the best published results. [1]

This face detection system is most clearly distinguished from Sung and Poggio approach [6] [9] in its ability to detect faces extremely rapidly. Operating on 384 by 288 pixel images, faces are detected at 15 frames per second on a conventional 700 MHz Intel Pentium III.

Viola Jones algorithm consists of the fallowing components

1) Integral Image representation allows fast feature computation.
2) Adaboost based classifier training procedure.
3) Classifier cascade allows fast rejection of non-face image.

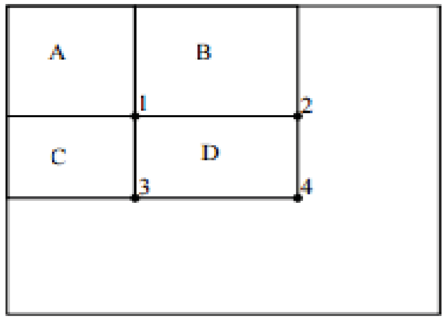
A. INTEGRAL IMAGES

Now let us introduce a new representation of an image called Integral image which facilitates for the features detected by the detector to be computed very easily.

In order to compute the features very rapidly and fast at many scales, the introduction of the new representation of the image becomes very essential. The Integral image can be computed from an Image using few operations per pixel.

The integral images facilitates for the fast feature evaluation. We can compute rectangular features by using integral images. The integral image at location (x, y) contains the sum of pixels and to the left of (x, y) is shown in equation I

$$Ii(x, y) = \sum I(x', y') \longrightarrow (I)$$



[Fig. 3]: Pixels in the rectangle box

The sum of the pixels within the rectangle can be computed with four array references.

Value at Integral Image location 1 is sum of pixels in rectangle A [Fig 3]. Similarly, value at location 2, 3, 4 is

i) A+B

ii) A+C

iii) A+B+C+D

Sum within D can be computed as 4 +1-(2+3)

$$s(x, y) = s(x, y-1) + i(x, y) \longrightarrow (II)$$

$$ii(x, y) = ii(x-1, y) + s(x, y) \longrightarrow (III)$$

Where ii(x, y) is the integral image and i(x, y) is the original image in equation (II) and (III)

In Viola–Jones algorithm a slightly modified version of Adaboost algorithm is used to select a small set of features and train the classifier.

B. FEATURE DISCUSSION

We use a set of features which are the derivatives of the Haar Basis functions. Haar like functions can be computed at any scale or location in constant time. Rectangular features are primitive compared to the steerable filter- (which provides texture analysis, Image compression), but rectangular features provide Rich image representation which provides effective learning. There are over 180,000 rectangle features associated with every image. These numbers are very large compared to other numbers. But our aim is to combine very small features to form an efficient classifier. In order to support this we have to design a weak learning algorithm. For every feature the weak classifier determines the optimum threshold classifier function.

$$H\,j(x) = \begin{cases} 1 \text{ if } p_j f_j(x) < p_j \theta_j & \longrightarrow (IV) \\ 0 \text{ otherwise} \end{cases}$$

Weak classifier $h_j(x)$, feature f_j and threshold θ_j and parity p_j indicating the direction for inequality as shown in equation IV. One has to make sure that features are selected in the early boosting stage which provide error rates in the range 0.1-0.3. Later boosting stages yield higher error in the range 0.4-0.5.

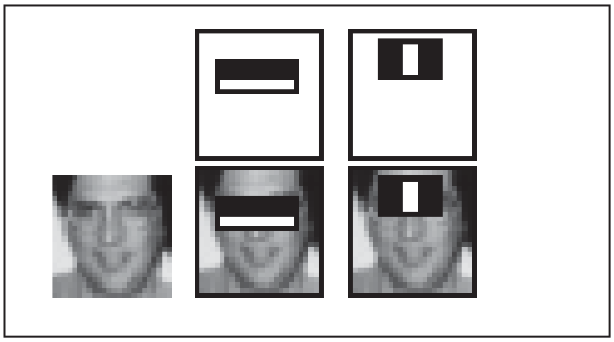
C. LEARNING DISCUSSION

Our appliaction needs an effective approach which is used for future use. For this Papageorgeiou et al[8] proposed a scheme which selects 37 features out of 1734 features.

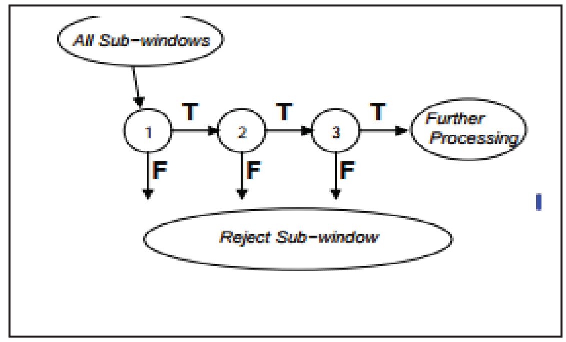
Others include:
1) Roth Et Al
2) Winnow learning process etc.

For performing the task of face detection the initial features detected by the Adaboost algorithm are easily interpretable and are also meaningful [Fig 4]

First selected feature focuses on the property that region of eye is darker than region of nose and cheeks. The second property is that- eyes are darker than the bridge of the nose.



[Fig. 4]: Face detection performance



[Fig. 5]: Cases classifiers

D. THE ADDITIONAL CASCADE

Algorithm for constructing a cascade of classifiers which drastically increases detection performance and also reducing the computation time. [Fig. 6] Boosted and small classifiers can be constructed which reject the negative sub windows and detecting all the positive instances. Overall form of detection process depends on a degenerate decision tree which is referred to as a cascade When we get a positive result from the first classifier this triggers the evaluation of the 2nd classifier, A positive outcome from both first and second classifiers triggers the third. A negative result leads to immediate rejection of the sub-window.

The stages in cascade are formulated by training the classifiers using Adaboost[7], then we have to adjust the threshold to minimize the false negatives. Cascading training process involves two types of trade offs .

In majority of the cases classifiers with higher features will predominantly achieve higher detection rates and lower false positive rates [Fig 5] These however require more time to compute.

We could define optimization frame work
- i) specify no. of classifier stages
- ii) No. of features in each stage
- iii) Threshold of each stage

This is however a complex task. In practice a very simple frame work is used to produce an effective classifier .Stages are added until the Overall target for false positive and detection rate is met.

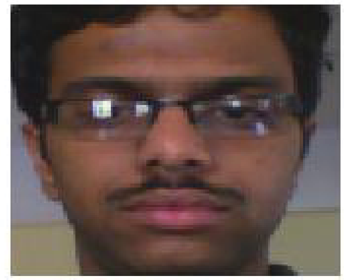
E. SPEED

Complete face detector has 38 stages with about 6000 features which results in fast average detection times.

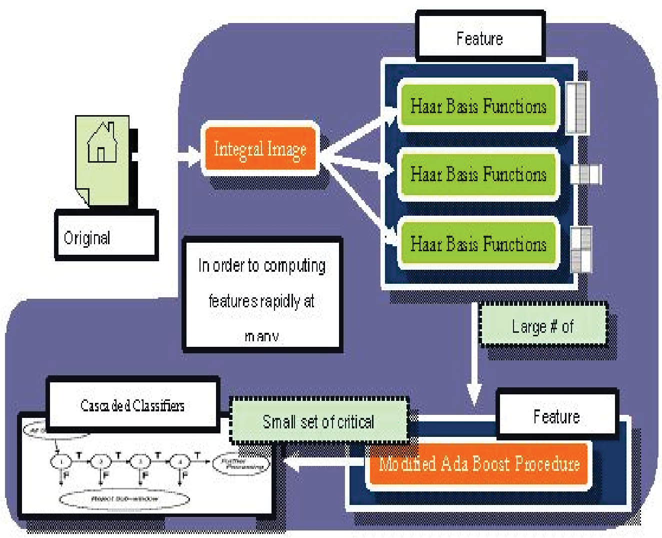
The cost of evaluating our detector at every scale and location is much faster than finding and grouping edges throughout the image. The speed of cascaded detector is directly related to the number of features evaluated per scanned sub-window. The final detector is scanned across the image at multiple scales and locations. Scaling is obtained by scaling the detector rather than scaling the image.

Detector \ False detections	10	31	50	65	78	95	167
Viola-Jones	76.1%	88.4%	91.4%	92.0%	92.1%	92.9%	93.9%
Viola-Jones (voting)	81.1%	89.7%	92.1%	93.1%	93.1%	93.2 %	93.7%
Rowley-Baluja-Kanade	83.2%	86.0%	-	-	-	89.2%	90.1%
Schneiderman-Kanade	-	-	-	94.4%	-	-	-
Roth-Yang-Ahuja	-	-	-	-	(94.8%)	-	-

[Fig. 6]: Detection rates for various numbers of false positives on the MIT+CMU test set containing 130 images and 507 faces.



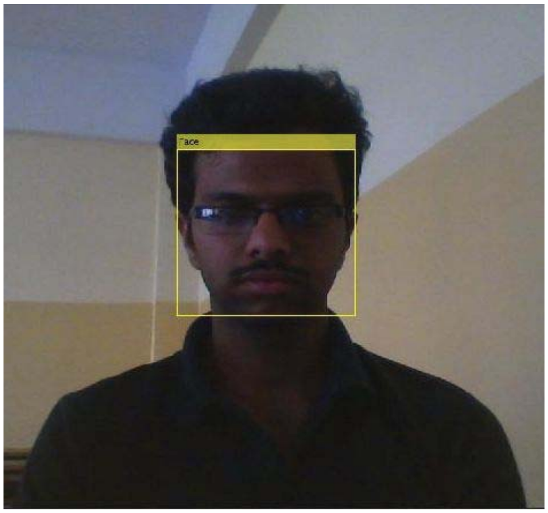
[Fig. 7] : Output of a viola jones based face detection



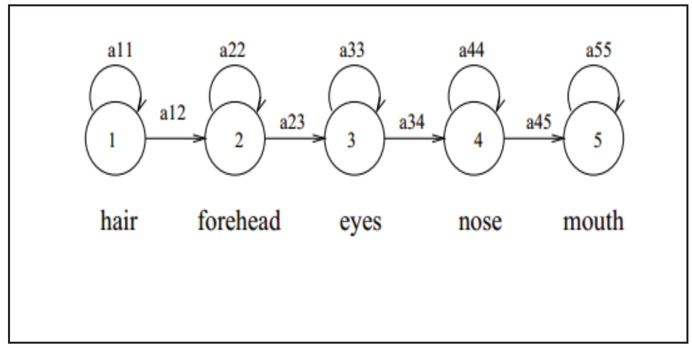
[Fig. 8]: Viola Jones Algorithm

III. FACE RECOGNITION

Face recognition (Facial recognition) [Fig 9] is a type of biometric software that can identify the face of that specific user. This is an integral part of the health monitoring of long time PC users. It is essential to maintain the database of computer users, duration of use and their health status. Face recognition is used to know the identity of the user. Most of the current face methods works with numeric codes called face prints. These face prints works quickly and can identify the individuals under favorable conditions.



[Fig. 9]: Face Recognition



[Fig. 10]: Left to Right HMM for Face Recognition

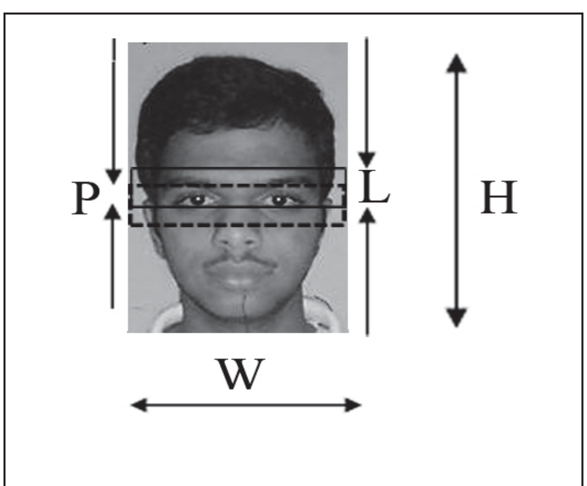
Markovian mixture with discriminative face alignment is an algorithm used for face detection. This algorithm fails to solve the problems of illumination and facial orientation[10] .We are using the HIDDEN MARKOV MODELS for face recognition, with feature extraction accomplished using 2D-DCT, that overcomes the drawback. Given the success of HMM in speech and character recognition and the work of Samaria [12], we have used HMM in face recognition. Hidden Markov Models (HMM) are a set of statistical models used to characterize the statistical properties of a signal [13]. HMM consist of two interrelated processes: (1) an underlying, unobservable Markov chain with a finite number of states, a state transition probability matrix and an initial state probability distribution and (2) a set of probability density functions associated with each state.

E. IMPLEMENTION OF HMM

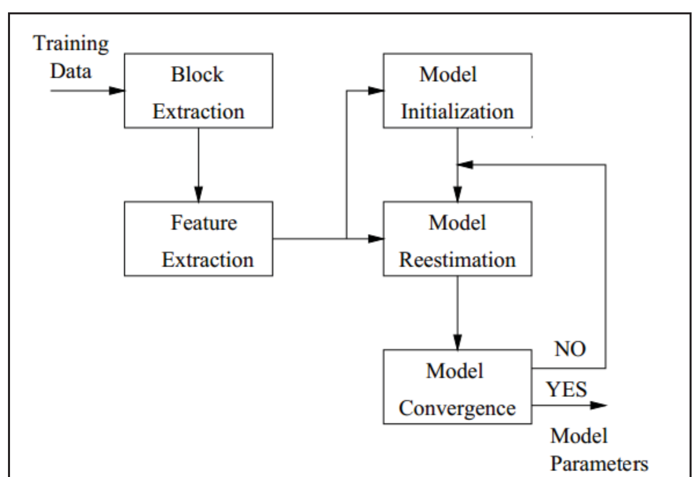
For frontal face images, the significant facial regions (hair, forehead, eyes, nose and mouth) come in a natural order from top to bottom, even if the images are taken under small rotations in the image plane and/or rotations in the plane perpendicular to the image plane. Each of these facial regions is assigned to a state in a left to right 1D continuous HMM. [9] The state structure of the face model and the non-zero transition probabilities a_{ij} are as shown in [Fig. 10].

F. FEATURE EXTRACTION

In face recognition using HMM models [Fig 13], extraction of the main face features i.e., hairs, forehead, eyes, nose and mouth is of significant importance. Irrespective of the methodology adopted in feature extraction, the frame of the face is divided into numerous sub blocks.



[Fig. 11]: Face image parameterization and blocks



[Fig. 13]: Training Mode extraction

As shown in [Fig 11], each face image of width W and height H is divided in to overlapping blocks of height T and width W. The amount of overlap between consecutive blocks is L. The number of blocks extracted from each face image equals the number of observation vectors T and is given by: T = (H - L /L – P) + 1. The system recognition rate significantly depends on the parameters P and L. However, the system recognition rate is not very sensitive to variations in L, as long as P is large (P <= L-1) [10]

In [6] the observation vectors consist of all the pixel values from each of the blocks, and therefore the dimension of the observation vector is L * W (L = 10 and W = 92).

The use of the pixel values as observation vectors has two important disadvantages: First, pixel values are very sensitive to image noise, image rotation or face tilt, shift, and change in illumination and second, the large dimension of the observation vector leads to increase in computational complexity and processing time of the training and recognition systems, making the process cumbersome. This can be a major problem for face recognition over large databases or when the recognition system is used for real time applications.

Using 2D-Cepstrum methods for feature extraction [8] has drawbacks. A cepstrum is the result of taking the Inverse Fourier transform (IFT) of the logarithm of the estimated spectrum of an image signal. However, approximating a spectral envelope via regularized discrete cepstrum coefficients has two problems-the selection of the spectral peaks and frequency axis is an elaborate procedure. Cepstrum features are extracted using only the local information of a given time series without making use of the training data.

Principal Component Analysis (PCA) is a statistical procedure that uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables called principal components. PCA utilizes the training data to find the linear transformation and approximates the subspace in which most of the variance of the training data lies. Although this is an advantage of PCA over Cepstrum, a linear transformation is not necessarily a good method of feature extraction from the data.

Symbolic Dynamic Filtering (SDF) is yet another method of feature extraction which is complex and not optimized for feature extraction of a human face.

Hence we use 2-Dimensional–Discrete Cosine Transform (2D-DCT).The observation vectors consist of a set of 2D-DCT coefficients that are extracted from each block. The DCT compression properties for natural images make the use of this transform a suitable feature extraction technique for the face recognition system. The features used in this approach, contain the 2D-DCT coefficients inside a rectangular window of size 13*3 over the lowest frequencies in the DCT domain. The window size has been chosen to cover the most significant coefficients, i.e. the coefficient that contains most of the signal energy.

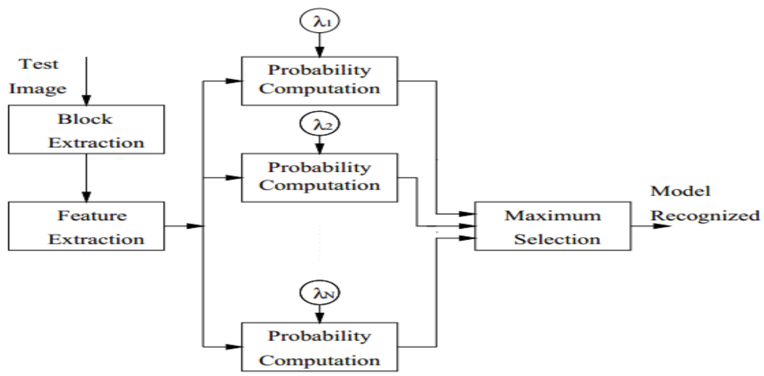
G. TRAINING THE FACE MODELS

While implementing face recognition in real time applications, the system must have a high recognition rate of over 90%, which includes recognition with different facial expressions, eye wear and hair style. Hence we adopt training of face models in our software.

Each individual in the database is represented by a HMM face model. A set of five images representing different instances of the same face are used to train each HMM. Following the block extraction, a set of 39 2D-DCT coefficients obtained from each block are used to form the observation vectors. The observation vectors are effectively used in the training of each HMM.

First, the HMM $\lambda = (A : B : \pi)$ is initialized. The training data is uniformly segmented from top to bottom in N = 5 states and the observation vectors associated with each state are used to obtain initial estimates of the observation probability matrix B. The initial values for A and are set given the left to right structure of the face model.
In the next steps the model parameters are re-estimated using the E-M procedure [11] to maximize $P(O|\lambda)$ [Fig 12]. The iterations stop, after model convergence is achieved i.e., the divergence between model probability at consecutive iterations (k and k + 1) is smaller than a threshold C, as shown in equation V.

$$| P (O|\lambda^{(k+1)}) - P (O | \lambda^{(k)}) | < C. \longrightarrow (V)$$



[Fig 12]: Block Diagram of Face Recognition

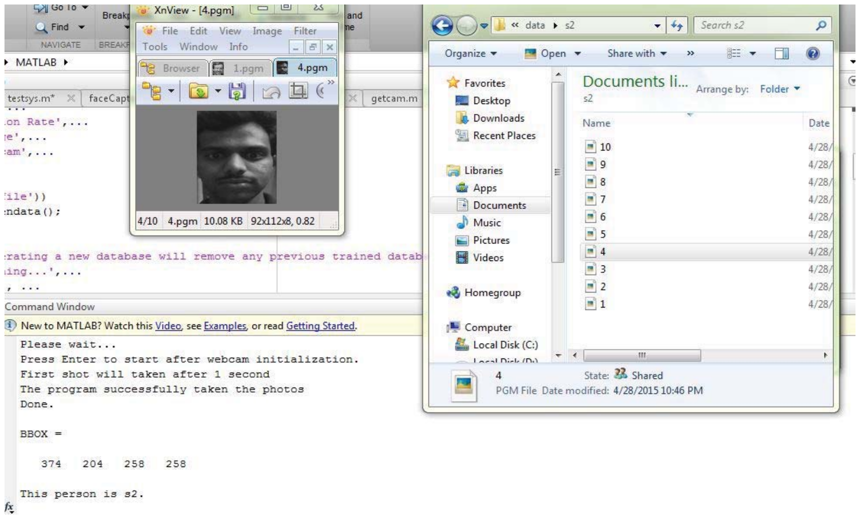
IV.CONCLUSION

We have proposed a novel method for face recognition which includes face detection and face verification/recognition as its subunits. The face detection uses Viola-Jones algorithm to identify the faces. This algorithm is inbuilt in the MATLAB software, hence it is easy to implement the algorithm. This algorithm detects up to 45 degrees of face rotation, which satisfies our requirement. The algorithm is efficient and provides required accuracy for face detection. It has fast feature computation and efficient feature selection. It can detect faces of any size.

The face recognition uses Hidden Markov Model algorithm. This algorithm can be used for real time applications. The HMM algorithm is more efficient as compared to the other face recognition algorithms such as Eigen face method achieved a recognition rate of 84% with L = 10 and P = 9. On the same database the recognition rate of the Eigen face method is 73% and the recognition rate of the HMM based approach presented in [6] is 84% over a fraction of the same database. The approach presented in this paper performs at a recognition rate better than the Eigen face method and decreases significantly the recognition time of the HMM based method in [6] while preserving the same recognition rate. The processing time required to compute the likelihood of one test image given a face model is decreased from 25 seconds reported in [6] (C language processing time) to 2.5 seconds in the present work (Matlab processing time).

Using lower dimensional feature vector, the computational complexity and recognition time is reduced. This algorithm is used for face recognition for wide range of facial expressions and orientations. The further improvement of this algorithm includes study of pseudo 2D HMM algorithm. This module is very essential part of the project as it is used for maintaining the database of the PC/laptop users. Using this module in the health monitoring of PC users will help us keep a track about the person using PC undergoing Health Monitoring. It helps us know the identity of the person who is not maintaining appropriate posture or/and not blinking his eyes required number of times, which may cause the user certain health problems.

V. RESULTS



VI. ACKNOWLEDGEMENT

We would like to sincerely thank Mrs.D.R.Ambika, for guiding us in this project. We are very grateful for her constant support throughout this project.

REFERENCES

[1] International Journal of Computer Vision 57(2), 2004(first published in CVPR ‘01) PAUL VIOLA, Microsoft Research MIKE JONES, Mitsubishi Energy Research Lab (MERL).

[2] A Neural Network Based Face Detection Approach Anjana Mall, Mrs. Shusmita Ghosh, Dept. Of Electronics & Telecommunication, CSVTU, Bhilai Chhattisgarh, Bhilai – India.

[3] De Silva, L.C., Aizawa, K., and Hatori, M.(1995). Detection and tracking of facial features by using a facial feature model and deformable circular template. IEICE Trans. Inform Systems. E78-D(9), 1995-1207.

[4] Jeng, S.H.Liao, H.Y.M., HAN, C.C., Chern, M.Y., and Liu, Y.T.(1998). Detection using geometrical face model: An efficient approach. Pattern recognition, 31.

[5] Yow, K.C. and Cipolla. R (1997). Feature-Based Human face detection, Image Vision Computing, 15(9).

[6] Sung and Poggio, 1998; Rowley Et Al., 1998; Osuna Et Al., 1997A; Schneiderman and Kanade, 2000; Roth Et Al., 2000.

[7] Robert E Schapire. AT&T Labs ,Shannon laboratory ,Proceedings of the Sixteenth International Joint Conference on Artificial Intelligece ,1999.

[8] Papageorgiou, C., Oren, M., and Poggio, T. 1998. A general framework for object detection. In International Conference on Computer Vision.

[9] Sung, K. and Poggio, T. 1998. Example-based learning for viewbased face detection. IEEE Patt. Anal. Mach. Intel., 20:39–51.

[10] Markovian Mixture Face Recognition with Discriminative Face Alignment, Ming Zhao, Google Inc. and Tat-Seng Chua, National University of Singapore.

[11] Image Based face detection and recognition, Fiazan Ahmad, Beijing University, China, Aaima Najam, COMSATS Institute of Information Technology, Pakistan and Zeeshan Ahmed, Education University, Pakistan.

[12] F. Samaria and S.Young, “HMM based architecture for face identification,” Image and Computer Vision, vol. 12, pp. 537-583, October 1994.

[13] L. Rabiner and B. Huang, Fundamentals of Speech Recognition, Englewood Cliffs, NJ: Prentice-Hall, 1993.

[14] F. Samaria and A. Harter, “Parametrisation of stochastic model for human face identification,” in Proceedings of the Second IEEE workshop on Application of Computer Vision, 1994.

[15] C. Wu, “On the convergence properties of the EM algorithm.” Annuals of Statistics, Vol 11, no. 1, pp. 95-103, 1983.