Computer Vision Based Employee Activities Analysis

Md. Zahangir Alom, Nabil Tahmidul Karim, Shovon Paulinus Rozario*, Md. Rezwanul Hoque, Md. Rumman Bin Ashraf, Saikat Lawrence Rozario
School of Engineering and Computer Science, BRAC University, Dhaka, Bangladesh
*Email: shovon990 {at} gmail.com

Abstract-**This paper presents a real time human face recognition based automated system for monitoring the work schedule and attendance of the employees in offices. Keeping track of the attendance records of the employees and monitoring their working hours is a time consuming task for the management of any company. Though the existing facial and card based attendance systems keeps the record of the attendance, measuring the actual time an employee works in office is still a challenge. This experimental system comes up with a solution to mark the attendance along with monitoring the time an employee spends in his/her workstation for work through real time face recognition. Haar cascade classifier is used for human face detection followed by implementation of Principal Component Analysis algorithm for recognition of the faces measuring the Euclidean distance of the eigenvalues. The system allows the company administrator to monitor the report of the employee working hours through a website or any android smart-phone providing those companies the scope of increasing management efficiency.**

Keywords- office management solution, real time face recognition, principal component analysis, haar cascade classifier.

I. INTRODUCTION

In today's world, managing employees is becoming crucial than ever in order to keep pace with the competing business world. Along with recording attendance for a general measurement of individual working hours of the employees, it is job of the administration of the companies to make sure that the employees are working while they are in their office campus. Traditionally, besides recording handwritten attendance, many of the companies use fingerprint, facial and card based attendance system for automatically recording the office hours of the employees. However, it is practically very critical and time consuming process for the administration to monitor every employee individually to ensure their presence in the workstations. In solution to this problem, this paper presents a system which not only records the employee attendance, but also keeps track of the hours an individual employee spends in his workstation. Since a significant portion of the employees in every company generally performs desk job where everyone has a fixed workstation, this system can significantly improve the efficiency of the management through reducing the hassle of manually monitoring the presence of their employees monitored directly from the system using their preferred device (laptop or android device). The system performs human face detection and recognition algorithms on video feeds received from the workstation cameras in order to record whether individual employees are present in their workstation or not ensuring the transparency of data served to the management.

II. RELATED WORKS

Even in a technically advanced world vast majority of the companies still follow the manual attendance system to mark the attendance where employees register their entry time and departure time manually in attendance papers. This creates the burden of manually counting the number of working hours for each day for every individual employee, which is not only a time consuming process but also often lacks the actual reflection of the time that the employees work. Some of the companies use fingerprint recognition based biometric attendance systems [1]. RFID technology [2] is also used in many cases for attendance management, which requires the users to waive the RFID card near a reader for identification of the person logging in and/or out. However, since attendance is marked using the card, anyone can give attendance of anyone as long as they have the cards. M. Smaili. Et. al [3] presented a system that authenticates users recognizing the iris of the individuals for tracking attendance. The above mentioned systems can be effective in case of marking general attendance of the employees, but they do not measure the actual time an employee spends working in his workstation. J. G. Roshan Tharanga and co. represented an automated system where real time face recognition technology was used for smart attendance [5]. However, his system focused mainly on the leave requests of the employees rather than concentrating on their total work-schedule and at the same time their attendance. On the other hand, this proposed system uses face recognition for monitoring the work schedule and attendance of each employee and demonstrates the result over two platforms. Thus, the system ameliorates the attendance system for a company and exhibits the ideal result in terms of validity and precision.

III. PROPOSED APPROACH

The proposed system uses face recognition technique for identification of employee presence in a workstation. The system consists of three parts-

- A local server application
- A server-side web application
- A mobile application

There is a CCTV camera at the entrance of the office campus which sends the feed to the local server for recording the arrival and departure time of the employees. The webcams in individual workstations capture the real-time video of the corresponding desks. The local server at the office campus runs the application for processing all the video feeds sent from the workstation webcams in order to identify whether an employee is present in his desk or not [5]. The data is then sent to the remote server which stores all the records of how much time everyone spent on their own desk along with the timestamp of their arrival at office and departure at the end of the day. The statistical analysis of the employee activities in the office can be viewed through employee profiles using the server-side web application or android mobile application.

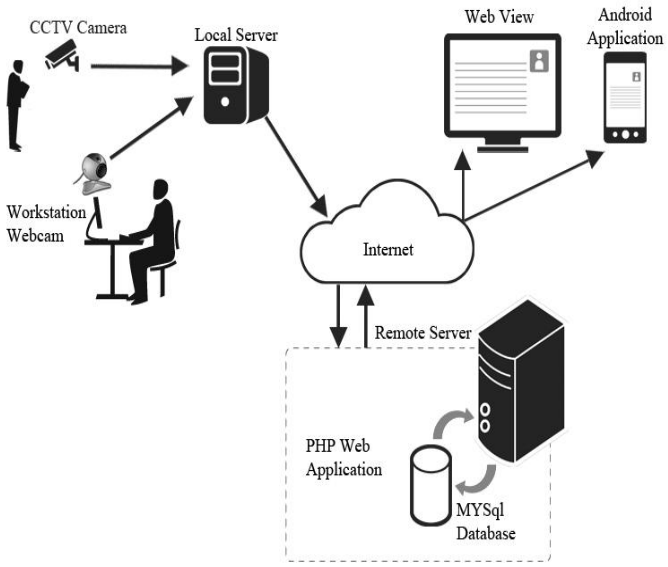

Figure 1: Overview of the system

IV. Technical Implementation

A. Local Server Application

The image processing for employee face detection and recognition is performed in the local server using the [4] OpenCV image processing library in .NET application. The employees need to be registered in the system providing their personal information and pre-captured images along with the assigned camera id numbers against their corresponding workstations. The registration process requires an admin authorization which enables the employees to be eligible for accessing their own records through the web application as well as the android mobile application.

When an employee arrives at the main entrance of the office, a CCTV camera sends the feed to the local server and a timestamp is recorded as entry time for the employee. Considering that it takes a maximum of five minutes to reach the desk from the main entrance, the local server application uses the workstation camera of the corresponding desk for checking whether the employee is present in his/her desk or not within the allocated time gap. If not, then it starts to record the time as offline period for the employee. The Haar Cascade Classifier [6] is run for detection of human face in the desk followed by implementation of PCA algorithm for recognition of the employees showed in Figure-6. Since each and every employee is assigned a fixed workstation containing a camera with a unique "CameraDeskID", the application periodically (every 12 minutes) checks the presence of the employees in their desk against the assigned camera id number. If at a given time, the detection algorithm does not find any face for 15 times at a stretch, the system will record the timestamp and start measuring the offline period of the employee. The departure time is marked using the CCTV camera in the main exit at the time of leaving the office.

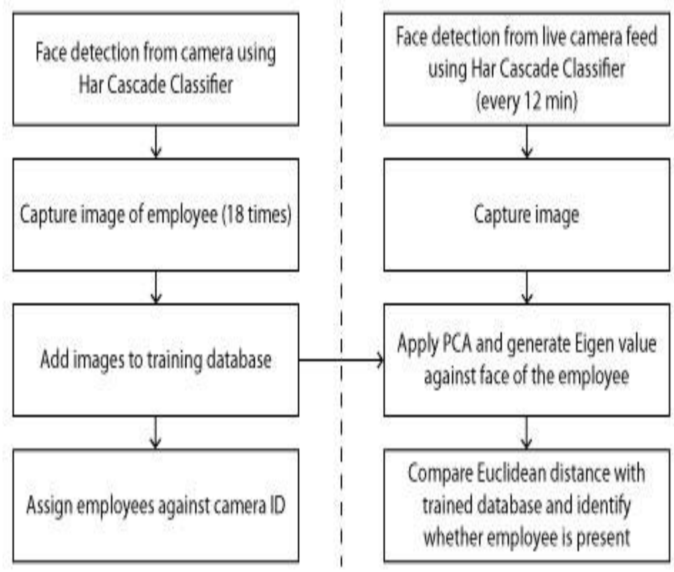

Figure 2: Process flow of local server application

B. Detection Process

The local server application runs the detection process on the video feeds received both from the entrance CCTV camera and individual workstation camera to detect facial features for checking existence of employees in their desk, followed by the recognition algorithm run later on to identify the actual person against the assigned workstation id number. Haar Classifier proposed by Paul Viola and Michael Jones [8] is used for detection of the human faces. Initially, the images collected from live video feed are converted from RGB to Grayscale image format in order to get clear edges for detecting the key points. The Haar classifier object detection uses Haar-like features showed in Figure -3[8], where the change in contrast values between adjacent rectangular groups of pixels are used instead of using the intensity values of a pixel. The contrast variances between the pixel groups are used to determine relative light and dark areas. Two or three adjacent groups with a relative contrast variance form a Haar-like feature. Haar features can easily be scaled by increasing or decreasing the size of the pixel group being examined which allows features to be used for detecting objects of various sizes [10].

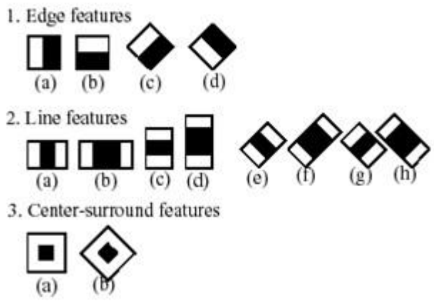

Figure 3: Common Haar features [8]

The simple rectangular features of an image are calculated using an intermediate representation of an image, called the integral image [8] ~ [9] which is an array containing the sums of the pixels' intensity values located directly to the left of a pixel and directly above the pixel at location (*x*, *y*) inclusive. So if *A*[*x*,*y*] is the original image and *AI*[*x*,*y*] is the integral image then the integral image is computed as shown in the following equation and illustrated in the Figure - 4.

$$AI[x,y] = \sum A(x',y');\, where\ x' \leq x\ and\ y' \leq y \quad (1)$$

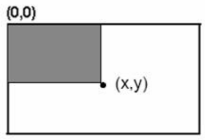

Figure 4: Summed area of integral image [8]

The rotated integral image is calculated by finding the sum of the pixels' intensity values that are located at a forty five degree angle to the left and above for the x value and below for the y value. So if A[x, y] is the original image and AR[x, y] is the rotated integral image then the integral image is computed as shown in the following equation.

$$AR[x,y] = \sum A(x',y');\text{where } x' \leq x, x' \leq x - |y - y'| \quad (2)$$

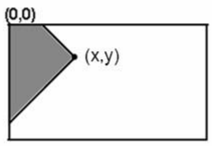

Figure 5: Summed are rotated integral image [8]

Using the appropriate integral image and taking the difference between six to eight array elements forming two or three connected rectangles, a feature of any scale can be computed.

Although calculating a feature is extremely efficient and fast, calculating all the 180,000 features contained within a 24 × 24 sub-image is impractical. In order to eliminate as many sub-images as possible, only a few of the features that define a face are used when analyzing sub-images [4]. The cascading of the classifiers allows only the sub-images with the highest probability to be analyzed for all Haar-features. In order to train the classifiers, gentle AdaBoost algorithm and Haar feature algorithms are implemented [6]. It requires two sets of images with one set not containing the object, in this case a facial feature, which is going to be detected. This set of images is referred to as the negative images. The other set of images, the positive images, contain one or more instances of the face**.** The location of the objects within the positive images is specified by: image name, the upper left pixel and the height, and width of the object. The positive set of images passes all the stages of cascade classifiers and thus results in a face region [8].

In the proposed system, the faces of the employees are detected at their workstations by identifying the Haar like features in order to evaluate the presence of the employees. If a face is discovered and thus detected, a red square box is drawn around the face to represent that a face has been detected showed in Figure-6 and 7.

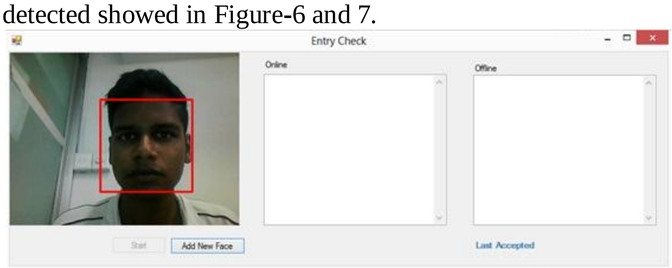

Figure 6: Face detection

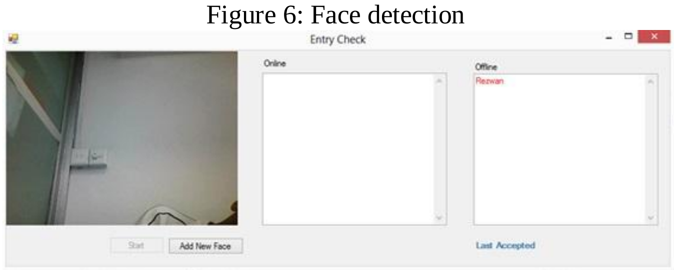

Figure 7: Testing detection with absence of face

C. Recognition Process

Principal Component Analysis (PCA) algorithm is implemented for identification of the detected faces from the gray images [7]. A 1-D vector can represent a 2-D facial image by concatenating each row (or column) into a long thin vector. Suppose there are *M* vectors of size *N* (= rows of image *x* columns of image) representing a set of sampled images. p_js' represent the pixel values.

$$x_i = [p_1 \ldots \ldots p_N]^T, i = 1, \ldots \ldots, M \quad (3)$$

The images are mean centered by subtracting the mean image from each image vector. Let *m* represent the mean image.

$$\mathrm{m} = \frac{1}{\mathrm{M}} \sum_{\mathrm{i}=1}^{\mathrm{M}} \mathrm{x_i} \quad (4)$$

And mean centered image is defined as w_i,

$$\mathrm{w_i} = \mathrm{x_i} - \mathrm{m} \quad (5)$$

The intention is to find a set of e_i's which has the largest possible projection onto each of the $\mathrm{w_i}$'s and a set of *M* orthonormal vectors e_i for which the quantity is,

$$\lambda_\mathrm{i} = \frac{1}{\mathrm{M}} \sum_{\mathrm{n}=1}^{\mathrm{M}} \left(\mathrm{e_i^T w_n}\right)^2 \quad (6)$$

maximized with the orthonormality constraint,

$$e_l^T e_k = \delta_{lk} \tag{7}$$

It has been shown that the eigenvectors and eigenvalues of the covariance matrix generates the e_i's and λ_i's

$$C = WW^T \tag{8}$$

where W is a matrix composed of the column vectors w_i placed side by side. The size of C is N *x* N which could be enormous, like in the case of 64 *x* 64 images, which creates the covariance matrix of size 4096 *x* 4096. Thus it is not practical to solve for the eigenvectors of C directly in such cases**[7].** A common theorem in linear algebra states that the vectors e_i and scalars λ_i can be obtained by solving for the eigenvectors and eigenvalues of the M *x* M matrix W^TW. Let $d_i$and μ_i be the eigenvectors and eigenvalues of W^TW, respectively.

$$W^TWd_i = \mu_i d_i \tag{9}$$

By multiplying both sides of (9) by W,

$$W^TW(Wd_i) = \mu_i(Wd_i) \tag{10}$$

which means that the first M - 1 eigenvectors e_i and eigenvalues λ_i of $WW^T$are given by $Wd_i$and μ_i respectively. Normalization should be done toWd_iin order for it to be equal to e_i. Since the system sums up a finite number of image vectors, M, the rank of the covariance matrix cannot exceed M – 1(The **-**1 come from the subtraction of the mean vector m) [12].

The eigenvectors corresponding to nonzero eigenvalues of the covariance matrix produce an orthonormal basis for the subspace within which most image data can be represented with a small amount of error [13]. The eigenvectors are sorted from high to low according to their corresponding eigenvalues. The eigenvector associated with the largest eigenvalue is one that reflects the greatest variance in the image. That is, the smallest eigenvalue is associated with the eigenvector that finds the least variance [12]. They decrease in exponential fashion, meaning that the roughly 90% of the total variance is contained in the first 5% to 10% of the dimensions. A facial image can be projected onto M' (<<M) dimensions by computing

$$\Omega = [v_1 v_2 \ldots . v_{M'}]^T \tag{11}$$

Where v_i=$e_i^T w_i$. v_i is the i^{th} coordinate of the facial image in the new space, which came to be the principal component. The vectors e_i are called eigenimages or eigenfaces[13].

So, Ω_k describes the contribution of each eigenface in representing the facial image by treating the eigenfaces as a basis set for facial images [14]. The face class that provides the best description of an input facial image is the simplest method for determining the face class k that minimizes the Euclidean distance.

$$\varepsilon_k = ||(\Omega - \Omega_k)|| \tag{12}$$

Where $\Omega_k$is a vector describing the k^{th}face class. If $\varepsilon_k$is less than some predefined threshold θ_ε, a face is classified as belonging to the class k[12].

Once the eigenfaces have been computed, the recognition of a person is processed where it must be decided if the individual has already been seen.

The eigenvectors, computed by PCA are in the direction of the largest variance of the training vectors, called as eigenfaces. Each eigenface can be considered as a feature [12]. When a particular face is projected onto the face space, its vector describes the importance of each of those features in the face. The face is expressed in the face space by its eigenface coefficients (or weights) [7]. A large input vector, facial image can be handled by taking its small weight vector in the face space. Since the dimensionality of the image space is much larger than that of face space this reconstructs the original face with some errors.

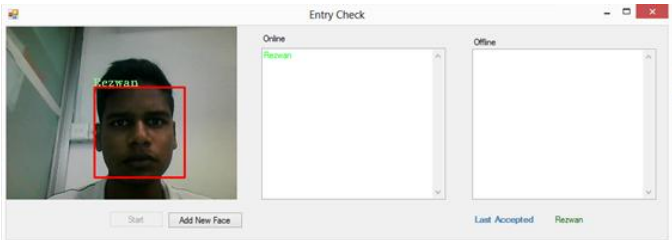

Figure 8: Face recognition

Each face in the training set is transformed into the face space and its components are stored in memory. The population of this face space should be done with these known faces. An input face is given to the system, and then it is projected onto the face space. The system computes its distance from all the stored faces [7].

Thus the distance determines whether the face is recognized or not. In Figure-8, in the proposed system, after detecting an employee's face, PCA is run to recognize his face at the workstation to record his timestamp and therefore, calculates the total time he spent at the workstation.

D. WEB AND MOBILE APPLICATION

The local server pushes the timestamp and status of the employees to the remote server database. The server-side PHP web application and the android mobile application depicts the statistics of individuals making it accessible for the employees to check their status through their profile showed in both Figure-9 and Figure-10. Therefore, the status of each individual is represented in a graphical bar chart in the web application showed in Figure-9, where Y-Axis represents time shown in every hour interval and X-Axis representing days of the month in single day interval. Moreover, any employee can observe the total summary of his working hours along with the illustrated time periods spent in and away from his desk. The management holds the full authorization for accessing all the detailed analysis reports of the current year or previous years of all the employees with the ability to manage their records directly from the web application.

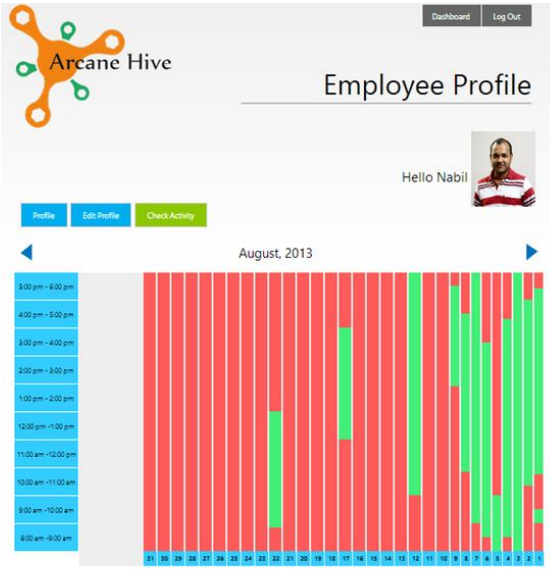

Figure 9: Employee activity statistics in the web app

In Figure- 9, the green bar shows how long the user was working on his desk and the red bar shows how long the user was away from his desk.

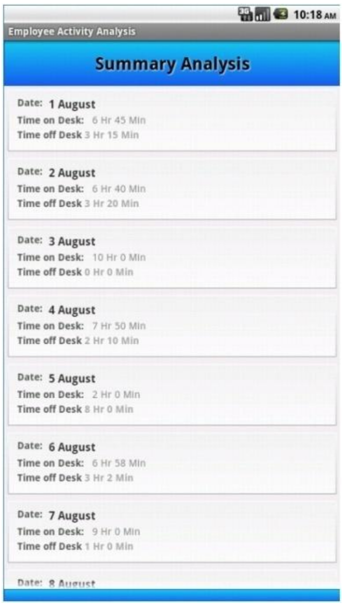

Figure 10: Mobile application preview

V. Discussion

The designed system is able to detect human presence at the designated desks accurately. It also works well in case of face recognition, though there are two concerns regarding this issue. The first issue is the impact of environment in the process of face recognition. The recognition of an employee from a captured image depends on the image quality where outside environment directly affects the outcome. For example, if the room is too dark or if the camera is set on such a position that it directly faces the light source of the room, the system fails to generate accurate results. To measure the accuracy, the system is run a number of times in different intensities of lights, where the result of the recognition is True Recognition (TR) if it identifies a person accurately, and False Recognition (FR) in case of wrong identification of the person. The equation used for calculating accuracy is,

$$\text{Accuracy} = TR / (TR + FR) \qquad (13)$$

The system generates the result of recognition with an accuracy rate of 84.58% for 240 test cases.

The second issue concerning the efficiency of the system is the amount of time that the system takes for processing of images from real-time video feeds for all the employees. The proposed system processes (detection and identification) the captured feeds sequentially, which will increase the time gap between checking for each desk as the number of employees increases. Considering an office for 20 employees, it takes about 12 minutes for the developed system to come back to the same workstation again in order to check the status of the employee. Therefore, it can be a big concern for a company that has hundreds of employees in its office. This time difference can be decreased using parallel data processing.

VI. Conclusion and Future Works

There are not many works aimed towards supervision of employee activity analysis using facial recognition in the office campus. The developed system thus fulfills its purpose of increasing efficiency for the office management in order to monitor the employees without any hassle. The accuracy of the experimental results also demonstrates the effectiveness of the approach for the developed system. The next study of this research is to overcome the processing limitations for decreasing time consumption using parallel processing methods. Parallel processing has shown remarkable results in improving the efficiency of various systems. The developed system's performance will be augmented with the help of this method.

References

[1] O. Shoewu, O.A. Idowu "Development of Attendance Management System using Biometrics", The Pacific Journal of Science and Technology, Vol. 13, May 2012.

[2] Arulogun O. T., Olatunbosun, A., Fakolujo O. A., and Olaniyi, O. M. "RFID-Based Students Attendance Management System", International Journal of Scientific & Engineering Research, Volume 4, Issue 2, February 2013.

[3] Kadry, S. and M. Smaili. "Wireless Attendance Management System Based on Iris Recognition", Scientific Research and Essays, Vol. 5(12), pp. 1428-1435, 18 June, 2010.

[4] "Face Detection using Haar Cascades", OpenCV 3.0.0- dev documentation,

http://docs.opencv.org/trunk/doc/py_tutorials/py_objdetect/py_face_detection/py_face_detection.html

[5] J. G.. Roshan Tharanga, S. M. S. C. Samarakoon, T. A. P. Karunarathne, K. L. P. M. Liyanage, M. P. A. W. Gamage, D. Perera, "SMART ATTENDANCE USING REAL TIME FACE RECOGNITION (SMART - FR)", Department of Electronic and Computer Engineering, Sri Lanka Institute of Information Technology (SLIIT), Malabe, Sri Lanka.

[6] M. Gopi Krishna, A. Srinivasulu, Prof (Dr.) T.K.Basak, "Face Detection System on Adaboost Algorithm Using Haar Classifiers", International Journal of Modern Engineering Research (IJMER), Vol.2, Issue.6, Nov-Dec. 2012.

[7] Kyungnam Kim, "Face Recognition using Principle Component Analysis", Department of Computer Science, University of Maryland, College Park, MD 20742, USA.

[8] Phillip Ian Wilson, and Dr. John Fernandez, "FACIAL FEATURE DETECTION USING HAAR CLASSIFIERS", Texas A&M University.

[9] Adolf F., "How-to build a cascade of boosted classifiers based on Haar-like features".

[10] Yi-Qing Wang, "An Analysis of the Viola-Jones Face Detection Algorithm", Image Processing Online, 2014, pp. 128–148. http://dx.doi.org/10.5201/ipol.2014.104

[11] The Code Project, "EMGU Multiple Face Recognition using PCA and Parallel Optimisation", 05 October 2011.

[12] Vytautas PERLIBAKAS, "Face Recognition Using Principal Component Analysis and Wavelet Packet Decomposition", Image Processing and Analysis Laboratory Kaunas University of Technology, October 2003.

[13] Abhishek Singh, Saurabh Kumar, "Face Recognition Using PCA and Eigen Face Approach", Department of Computer Science and Engineering National Institute of Technology Rourkela Rourkela – 769008, India.

[14] Matthew Turk, Alex Pentland, "Eigenfaces for Recognition", Vision and Modeling Group, The Media Laboratory, Massachusetts Institute of Technology