Face Liveness Detection

A BACHELOR'S MINI PROJECT

submitted in partial fulfillment

*of the requirements for the completion of the **5th** semester*

of the

UNDER GRADUATE PROGRAM

in

INFORMATION TECHNOLOGY
(B.Tech in IT)

Submitted By:

1. SHIVAM AGGARWAL (IIT2011158)

2. SHUBHAM AGARWAL (IIT2011165)

3. SHANTANU AGRAWAL (IIT2011172)

Under the Guidance of:

Prof. G. C. Nandi
Head of Information Technology Division, and Dean (Academic)
IIIT-Allahabad

INDIAN INSTITUTE OF INFORMATION TECHNOLOGY
ALLAHABAD – 211 012 (INDIA)

November, 2013

CANDIDATE'S DECLARATION

We hereby declare that the work presented in this project report entitled "**FACE LIVENESS DETECTION"**, submitted towards the Mid-Semester evaluation of mini project in Fifth semester of B. Tech. (Information Technology) at Indian Institute of Information Technology, Allahabad, is an authenticated record of our original work carried out from July 2013 to Nov 2013 under the guidance of **Prof. G. C. Nandi**. Due acknowledgements has been made in the text to all other materials used. The project was done in full compliance with the requirements and constraints of the prescribed curriculum.

Place: Allahabad
Date:

SHIVAM AGGARWAL
IIT2011158
SHUBHAM AGARWAL
IIT2011165
SHANTANU AGRAWAL
IIT2011172

CERTIFICATE FROM SUPERVISOR

We do hereby recommend that the mini project report prepared under our supervision by PROF. G. C. NANDI titled HEAD OF INFORMATION TECHNOLOGY DIVISION and DEAN(ACADEMICS) be **accepted** in the partial fulfillment of the requirements of the completion of 5th semester of Bachelor of Technology in Information Technology **for Examination**

Date:
Place: Allahabad

Prof. G. C. Nandi
Head of Information Technology Division and Dean(Academics), IIITA

Committee for Evaluation of the Project

_____ _____

_____ _____

ACKNOWLEDGEMENT

The author would like to express his sincere gratitude to Prof. G. C. Nandi for his constant support and effort towards the completion of my project.

Place: Allahabad
Date:

Shivam Aggarwal (IIT2011158)
Shubham Agarwal (IIT2011165)
Shantanu Agrawal (IIT2011172)

B Tech 3rd Year, IIITA

ABSTRACT

Face authentication is commonly offered as an alternative to passwords for device unlock. However, available face authentication systems are vulnerable to simple spoofing attacks. To defend against these vulnerabilities, we propose a face authentication system that includes a secrecy challenge. During authentication, the user must not only show his face but also gaze at a pattern that moves across the screen randomly. Using a novel method for estimating the noise level in the gaze tracking data, our system adapts the system's parameters to enable secure, hands-free authentication.

Existing face liveness detection algorithms adopt behavioral challenge-response methods that require user cooperation. To be verified live, users are expected to obey some user unfriendly requirement. We present a multispectral face liveness detection method, which is user cooperation free. Moreover, the system is adaptive to various user-system distances. After grossly locating a face, we first find the areas which left and right eyes lies in. According their probabilities, the precise eye positions are decided. Then Hough transformation is applied to detect cornea so as to track the eye movements. We then proposed a synchronization mechanism which established a synchronization with the challenge given with the response given by the system.

Table of Contents

1. INTRODUCTION

Faces can also be used as a way to authenticate a person to a device because of the unique features that the face of each and every person possesses, like eyes, lips, nose etc.

Facial recognition for dynamic images, in security, is a way to identify a person liveness through the different features present on a face of that person. There is a demand of user-friendly applications which can secure our assets and protect our privacy. These applications should be fast enough and reliable and should not be vulnerable towards breaches. Developing this kind of applications leads us to an emerging concept called biometrics.

Biometrics is an emerging area of bioengineering, it is the automated method of recognizing person based on a physiological or behavioral characteristic. Biometrical face recognition, also known as Automatic Face Recognition (AFR), is one of the most attractive approach as it focuses on the identification of a person as a normal person do.

The problem of face recognition can be divided into two parts:-

(a) Face recognition
(b) Face verification

The first stage is a stage to identify the person's face in an image. While the second stage is a stage for extracting relevant feature extraction or important information for the discrimination.

But using only the above techniques, faces cannot be used as secure way of authentication as it can be breached using public knowledge of faces[1]. Our main focus is to detect liveness on the face recognized to overcome the problem.

In this project, we present a real-time method of distinguishing a live person from images, videos and 3-D model. We are capturing the movement of eyes and lips to a randomly generated pattern for real-time authentication.

1.1 MOTIVATION

Many applications for face recognition have been envisaged. Commercial applications have so far only scratched the surface of the potential. Installations are so far are limited in their ability to handle pose, age, & lighting variations, but as technologies to handle this effects are developed, huge opportunities for deployment exist in many domains. Installation are so far are limited in their ability to handle pose, age, & lighting variations, but as technologies to handle this effects are developed, huge opportunities for deployment exist in many domains :

1. Access Controls: Face verification, matching a face against a single enrolled example, is well within the capabilities of current personal computer hardware. Since PC cameras have become widespread, their use for face based Pc logon has become feasible, though take -up seems to be very limited. Increased ease of use over password protection is hard to argue with today's somewhat unreliable & unpredictable systems, & for few domains is their motivation to progress beyond the combinations of password & physical security that protect most enterprise computers.

2. Identification Systems: This is an identification task where any new applicant being enrolled must be compared against the entire database of previously enrolled claimants, to ensure that they are not claiming under more than one identity. Unfortunately face recognition is not currently able to reliably identify one person among the millions enrolled in a single state database, so demographics are used to

narrow the search. Here a more accurate system such as fingerprint or iris based person recognition is more technologically appropriate, but face recognition is chosen because it is more acceptable & less intrusive.

3. Surveillance: The application domain where most interest in face recognition is being shown is probably surveillance. Video is the medium of choice for surveillance because of the richness & type of information that it contains naturally, for application that require identification, face recognition is the best biometric for video data.

2. PROBLEM DEFINITION AND SCOPE

Face authentication available nowadays can easily be spoofed using commonly available technologies like

- Photographs
- Digital image
- Videos

In the social media era, users' faces are often available online. Highly contrasted photographs of user's faces such as pictures on Facebook, Twitter, and LinkedIn etc. can be taken even without the knowledge of the user.

Faces should be considered public knowledge, which means face authentication lacks a secrecy component. Secrecy is the strength behind passwords and PINs; to be secure, it may be something that face authentication must incorporate.

To defend against these vulnerabilities we propose a face authentication system that includes user friendly challenges.

By finding a robust solution for the above problem we can integrate this with the existing techniques of authentication to introduce a new level of security.

3. LITERATURE SURVEY

Photos taken in conditions similar to users' enrollment conditions most likely result in successful logins [2]. However, existing work fails to examine one important question: how "good" do these images need to be? How many pixels does an attacker need to conduct a successful attack? Is it possible to break a face authentication system with the low quality photographs that are widely posted online?

According to a survey conducted at University of California, Berkeley, Nokia Research and Oxford University: On testing four commercially available face authentication systems: Dell Fast Access (version 2.4.95) [3]; HP Face Recognition for HP ProtectTools (version 2.0.1.651); Lenovo Veriface (version 3.0) [4]; and Toshiba Face Recognition (version 3.1.18) [5]. The result obtained is shown as

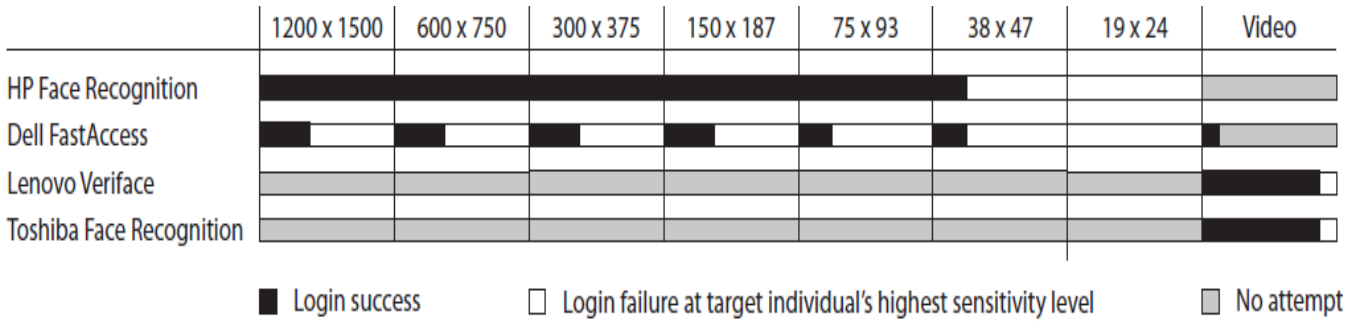

Figure 1: Summary of photo and video spoofing attempts at the systems' highest sensitivity levels. Note that we were able to reliably authenticate with images at the systems' default sensitivity levels.

The results show that photographs with an extremely low resolution are capable of spoofing both the Dell and HP systems. The HP system was much easier to reliably

spoof, perhaps because they could not increase the sensitivity level over the default. At higher sensitivity levels, the Dell system was sensitive to changes in lighting and position. They were able to reliably spoof the Dell system if they decreased the sensitivity level back to its default. A video was also able to spoof the Dell system. Turning off liveness detection on the Lenovo system also allowed them to authenticate using still images.

There are many ways to eliminate such threats by tracking the movement of different features on the face. We used proportions of the face to get an estimate of eye region. One such method is eye localization.

Zhu [6] and Haro [7] propose to perform real time eye tracking based on combing its appearance, the bright pupil effect, and motion characteristics in video stream. However, this technique strongly depends on the lighting conditions and size of the pupils. [8], [9] and [10] use facial structural knowledge such as Hough transform, symmetry detector, projection analysis etc. to detect eyes. In these methods, the physical properties of the eyes are not taken into account. The appearance-based methods [11] [12] detect eyes based on the intensity distribution of the objects. They collect a large amount of training data under different conditions, and rely on techniques from statistical analysis and machine learning to find the relevant characteristics of eyes and non-eye samples. But in most of these methods only eye patch was considered. As a matter of fact, eyebrows or thick spectacle frames sometimes look so similar to a closed eye that the classifier often makes a wrong decision. So both the Eye patch and eye neighborhood should be considered.

4. PROPOSED APPROACH

To solve the above mentioned problem, we have first detected the nearest (one at a time) face using the Haar classifier of face from the video so obtained by the web camera, then we had crop the eye region of the face using estimated proportion, then using the estimated proportion again we were able to detect the left and the right eye of the detected face. We then used Hough transformation for detection the cornea of both the eyes separately and drawing a circle around the cornea from the canter coordinates and the radius so obtained from the transformation.

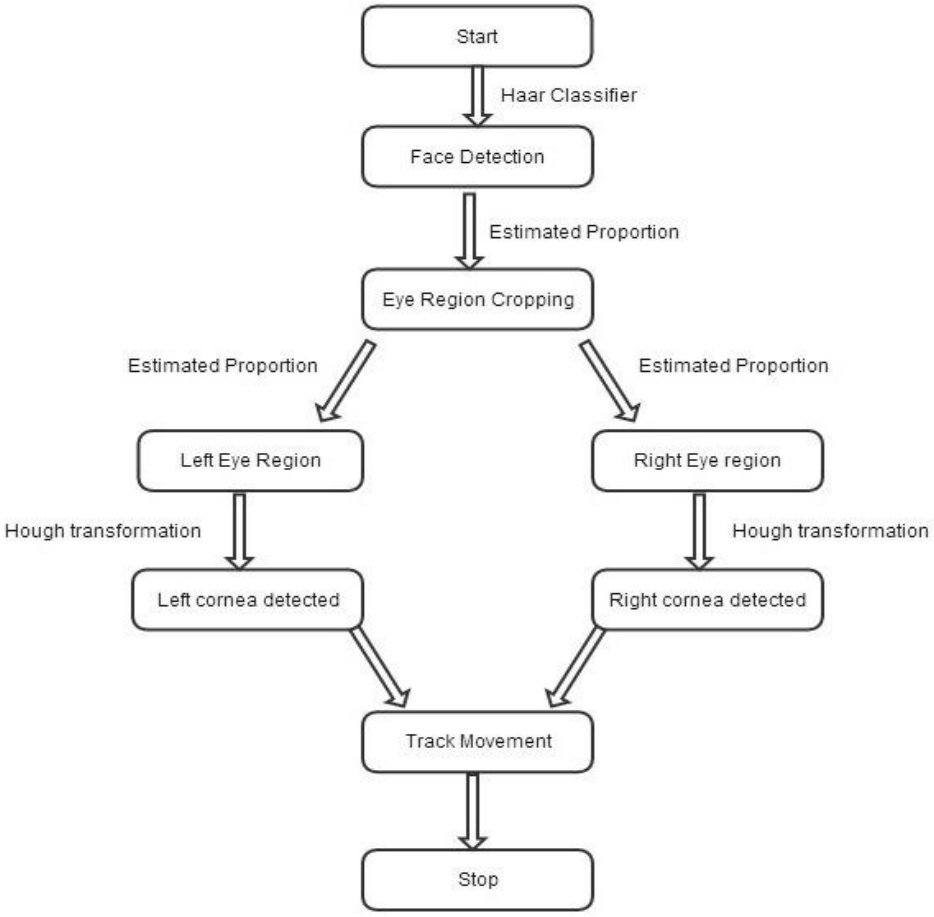

Figure 2: Flow chart describing our algorithm

1. **Haar Classifier for face:** The core basis for Haar classifier object detection is the Haar-like features. These features, rather than using the intensity values of a pixel, use the change in contrast values between adjacent rectangular groups of pixels. 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-like features, as shown in Figure 1 are used to detect an image. Haar features can easily be scaled by increasing or decreasing the size of the pixel group being examined. This allows features to be used to detect objects of various sizes.

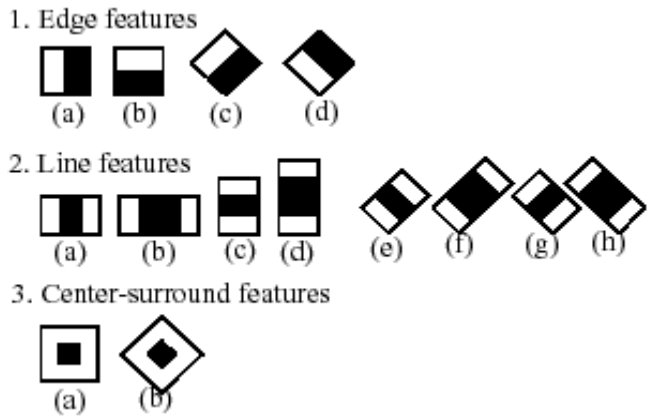

Figure 3: Haar Features

Classifiers Cascaded

Although calculating a feature is extremely efficient and fast, calculating all 180,000 features contained within a 24 × 24 sub-image is impractical [Viola

2001, Wilson 2005]. Fortunately, only a tiny fraction of those features are needed to determine if a sub-image potentially contains the desired object.

In order to eliminate as many sub-images as possible, only a few of the features that define an object are used when analyzing sub-images. The goal is to eliminate a substantial amount, around 50%, of the sub-images that do not contain the object. This process continues, increasing the number of features used to analyze the sub-image at each stage. The cascading of the classifiers allows only the sub-images with the highest probability to be analyzed for all Haar-features that distinguish an object. It also allows one to vary the accuracy of a classifier. The inverse of this is also true. Viola and Jones were able to achieve a 95% accuracy rate for the detection of a human face using only 200 simple features.

2. **Estimated Proportion for calculating eye region:** According to various surveys conducted over proportionate features of face, Fabian Timm and Erhardt Barth[13] collected some data and used it to judge our left and right eye regions in the face area. for left eye region, it will be 13% away from the face and 31% below the face rectangle but will not exceed 37% in width and 18% in height for right eye region it will be 63% await from the face and 31% below the face rectangle and will not exceed 7% in width and 18% in height

3. **Hough Transformation:** The Hough transform is a well-established family Of algorithms for locating and describing geometric figures in an image. However, the computational complexity of the algorithm used to calculate the transform is high

The classical Hough transform maps an image into an abstract parameter space via a voting process. Each significant pixel in the input image is examined to find the set of target objects of which it could form part. Each pixel can therefore be considered to generate a "vote" for each of the Plausible target objects. Target objects that are actually present in the image will be consistent with a relatively large number of image pixels, so will accumulate more votes

In practice, votes are accumulated in an abstract parameter space known as the Hough space, which has dimension equal to the number of parameters required to uniquely describe the figure in question. For example, a Hough transform intended to find arbitrary circles will have a three-dimensional Hough space, with two parameters corresponding to the coordinates of the Centre of the circle and a third to describe the circle's radius.

This problem is exacerbated as the complexity of the target objects increases because the dimension of the Hough space increases commensurately.
There are several modifications that can be made to the simple algorithm to reduce the time and/or storage demands of the Hough transform.

- The first approach allows significant peaks in Hough space to be generated without accumulating votes from every image pixel. The speed increase is roughly proportional to the sparseness of the sampling.

- The second approach is to generate fewer spurious votes in Hough space by extracting more information from the input image than is available in a single pixel. Examples of this approach include using the image gradient to constrain vote location.

For doing so, we propose a real time challenge response system. This system would be restrict spoofing but would require the cooperation of the user who is trying to authenticate it. If a user is able to respond correctly, the he will be authenticated else he would not be able to authenticate the system.

Challenge Response Mechanism

- Mechanism in which one party presents a question ("challenge") and another party must provide a valid answer ("response") to be authenticated.
- The users are required to look into a certain direction, which is randomly chosen by the system.
- By estimating the eye movements, the system verifies the user's response to the direction challenge. The movement estimation is based on detection and subsequent tracking of suitable features.

Steps involved in the process would include:

1) Recognize the face.
2) Detect the position of the eyeballs.
3) Throw the randomly generated challenge prepared as shown in the figures.
4) Record the response of the user for that challenge by detecting the coordinates of the new eyeball position.
5) Get the relative difference in previous and new positions of eyeballas.
6) If the response is as expected, authentication is successful.
7) Otherwise, it is identified that user is not live. Hence, the system cannot be authenticated

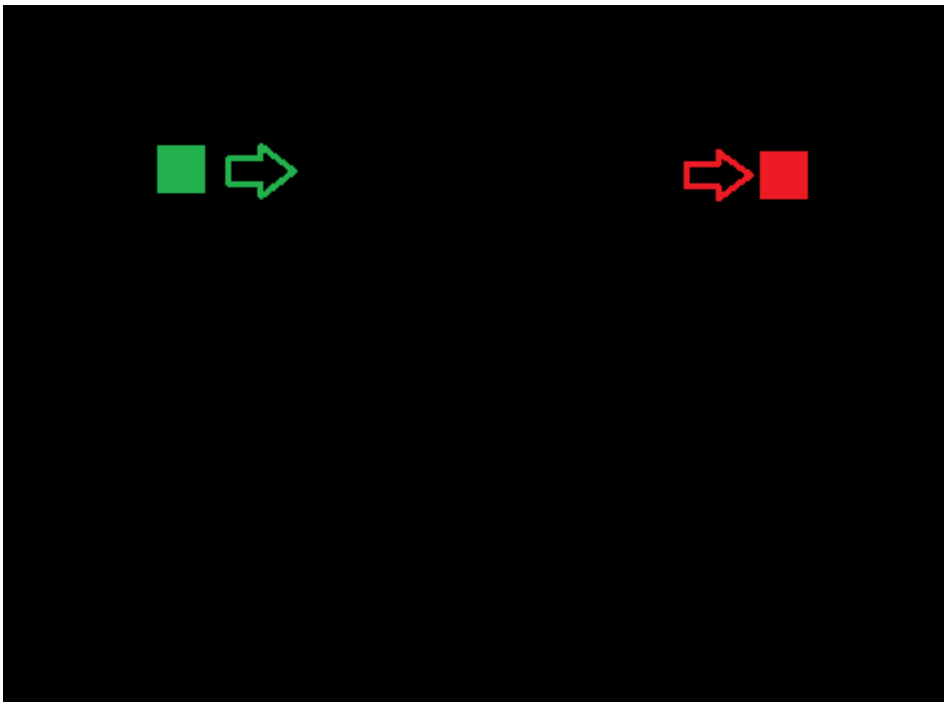

Figure : 4 one of the random pattern generated , the user has to see from the green dot to the right dot for authentication.

5.1 HARDWARE REQUIREMENTS

The hardware requirements for our project are listed as follows:

- A processor capable of performing calculations at a medium speed, so that it could serve as a part of a Real-Time System.
- A Webcam, to take input as the video or picture of users. It might be the one build inside the laptop or may be an external one.

We shall use:

- An Intel i3 processor.
- Webcam inbuilt in the laptop.

5.2 SOFTWARE REQUIREMENTS

The Software requirements for our project are listed as follows:

- A cross-platform IDE (Integrated development Environment) with Runtime Environment.
- MingW or G++ compiler for C++ programming language.
- OpenCV library integrated with the IDE and compiler.

We shall use:

- Code Blocks, (The open source, cross-platform IDE)
- MingW Compiler
- OpenCV library integrated into Code Blocks

6. ACTIVITY TIME CHART

Task	Start Date	End Date	Percentage of Task Completed
Literature Survey	July 25, 2013	November 1, 2013	100%
Learning basic Open CV and integrating with IDE	August 12, 2013	October 10, 2013	100%
Study of Haar cascade classifier and implementation	September 1, 2013	September 8, 2013	95%
Study of Hough Transformation and its implementation	September 9, 2013	September 26, 2013	95%
Study of proportionate division of face	September 26, 2013	October 15, 2013	100%
Synchronization of challenge with eye movement	October 5, 2013	November 22, 2013	95%

7. Face Liveness Detection

The basic work done includes face and eye detection along with detection of cornea which would be used for capturing the direction of movement of eyes.

7.1 Face and Eye Detection

For face detection we have used Haar classifier of face detection. The key advantage of a Haar-like feature over most other features is its calculation speed.

Steps for face detection using Haar cascade classifier:

1. Capture the frame from web camera.
2. Gray the frame and equalize the Hist
3. Detect faces from the frame using Haar cascade classifier
4. For each detected face
5. Detect eyes using Haar eye cascade classifier
6. For the eyes detected draw the circle

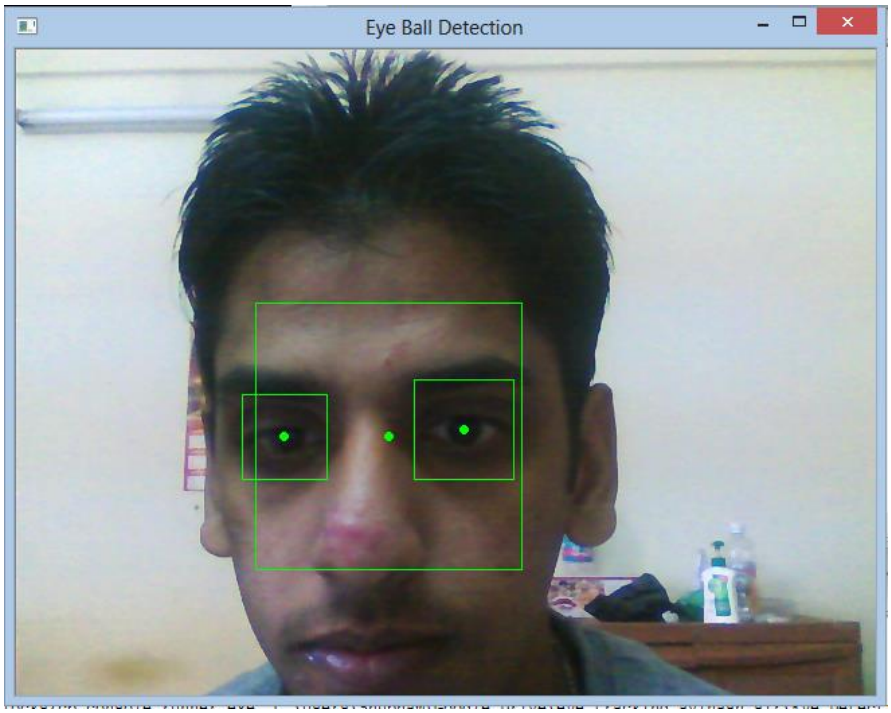

Figure 5: Showing misclassification by using haar eye classifier

But the accuracy obtained using Haar classifier for eye is not good enough because several misclassifications were found while using it. So we changed the approach for detecting eyes by cutting the probable portion of eyes by the estimated proportions for eyes.

Based on the position of the detected face and anthropometric relations, we extract rough eye regions relative to the size of the detected face.

By using the Hough transformation, the cornea position is calculated which would be used for tracking the movement of eyes.

Here we have not used the inbuilt Hough transformation but we have implemented it as follows.

Determine the probabilistic range of eye feature in the image (xmin to xmax, ymin to ymax and rmax)

search the whole image pixel by pixel
    search in range of eye feature
        if pixel value around 0
            calculate radius of the possible circle
            if radius < rmax
                Then that pixel coordinate with the radius value is given a vote

Count the pixel value and radius with maximum votes (x, y, r)

Plot a circle with center(x, y) and radius r

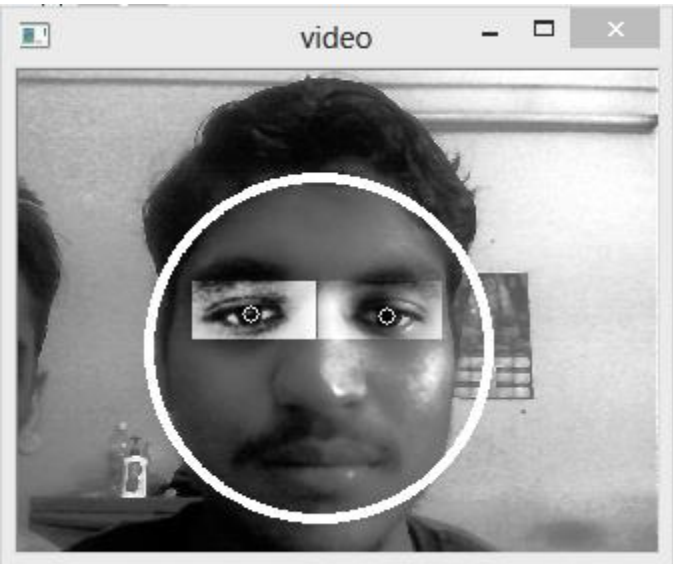

Figure 6: Face Detection

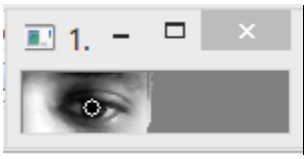

Figure 7: right eye

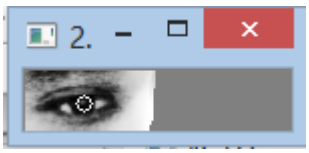

Figure 8: left eye

Figure 5, 6 and 7 are the snapshots of the algorithm that we had implemented. There is much accuracy as compared to figure 4 when we have used Hough transformation instead of Haar classifier for eye detection. By implementing Hough there is no misclassification and the cornea points are accurately detected.

7.2 Movement Synchronization

Till now, we have not actually synchronized the eye movements with the pattern generated on the screen, for doing so, we first tried to find out the eye corners and then tried to find out where the person is looking. For this we used that – if the distance between the eye balls is closer to left corner, then he would be seeing left and if the person's eye ball is closer than the right corner, then he would be seeing to the right.

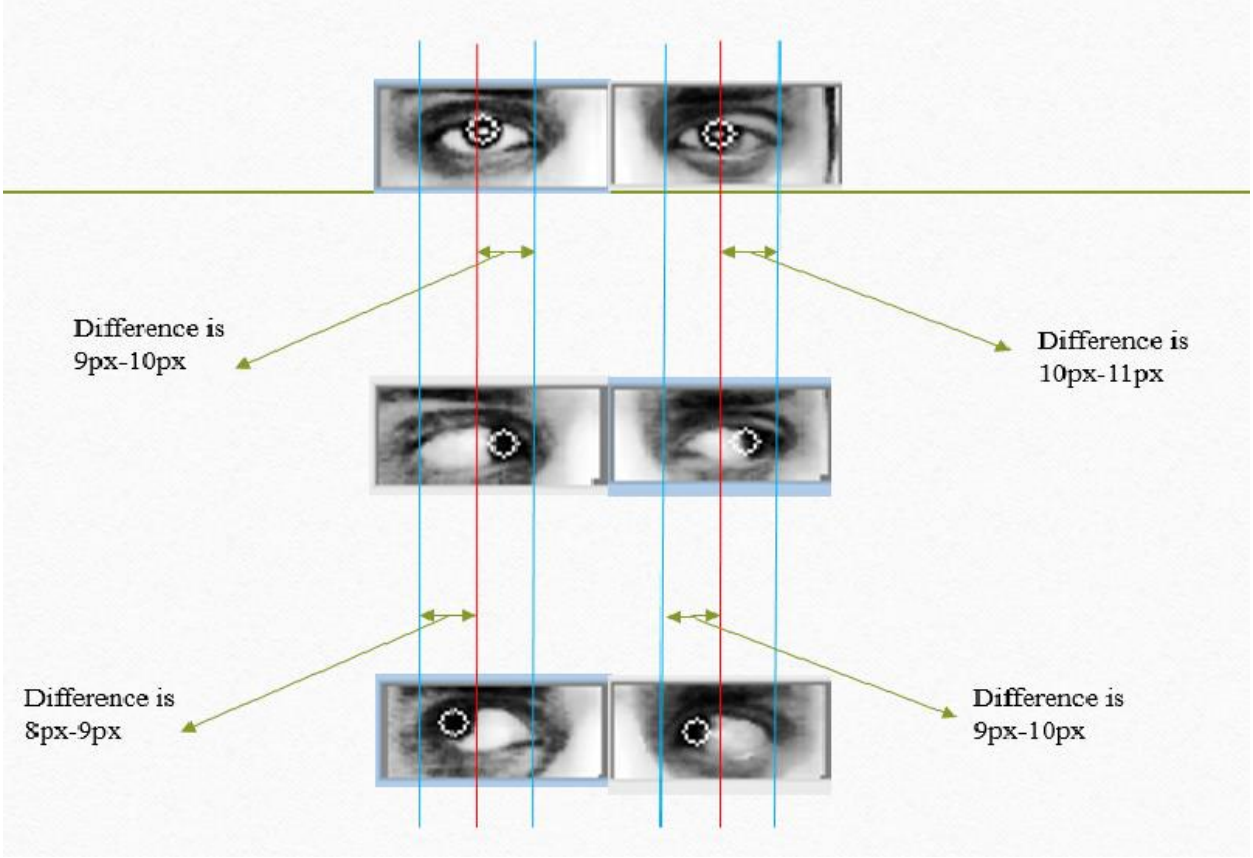

Figure 9: finding the left , right direction : 9(a) showing the right eye movement, 9(b) showing the centre postion of eye balls, 9(c) showing the left eye movement

Another way to detect the eye movements is mapping of screen with the eye movements that is what is used by the eye gazing methods. The extreme coordinates of screen is scaled with respect to the eye motion. After detecting the eye ball, mapping of screen size with the left and right extreme position is done. First, the user is required to tell the position of his/her right and left extreme points.

This is done by:

1) The person see to its top left most point on screen and presses a key(space bar) to indicate the program about the left extreme value at that distance.
2) The person then see to its bottom right most point on screen and presses a key to indicate the program about the right extreme value at that distance.
3) After that two point would be generated randomly on the screen – one of green color and another of red color. Person has to look from the green point generated on the screen to the red one. If system could verify the motion, he has passed this step test. Otherwise either person have to try again or he could not be authenticated.
4) The step four is repeated random number of time.

Then the system records these points and make further decision on the basis of these coordinates. Then the system classify the points in x direction into two parts – Left part and Right Part. Due to fluctuations in the real time eyeball detected coordinates, we wanted to take means of the coordinated generated. For this we used a data structure queue, which after every 1 mili second, check for the number of points classified as left, right and center in case of queue for X direction tracking or up and down for the Y direction tracking. The count having the maximum value at that time would be classified as the position of the direction of motion of eye balls.

Steps involved in this classification include:

1) Get the bottom - right and top - left most extreme values for that person
2) Scale the screen along the x direction according to the difference so obtained from the two extreme values
3) Scale the screen along y direction according to the difference so obtained from the two extreme values
4) For each eye ball coordinate obtained
    - Classify the coordinates according to the respective position
    - Put them in a queue
    - Count the frequency of the classifications present in the queue

- Match them with the point that we had generated on the screens
- If the matching is correct
    - The user is determined to be a live user, so he can access the system
- Else
    - The user is not a live user, hence he cannot access the system.

The point generation is random, so there is not any possibility of spoofing by videos or from any other means.

Snapshots indicating the motion of the randomly generated points as well as the challenges being respond by the system.

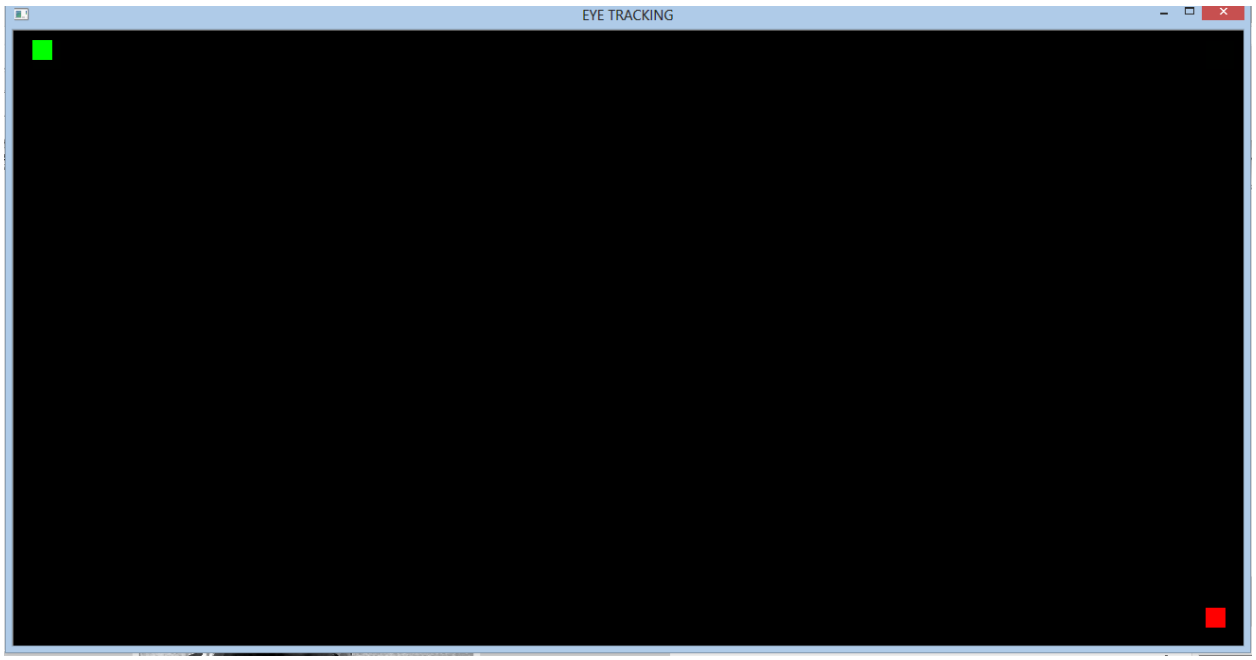

Figure 10: The person has to view the green point and starting from green point he has to see the red point.

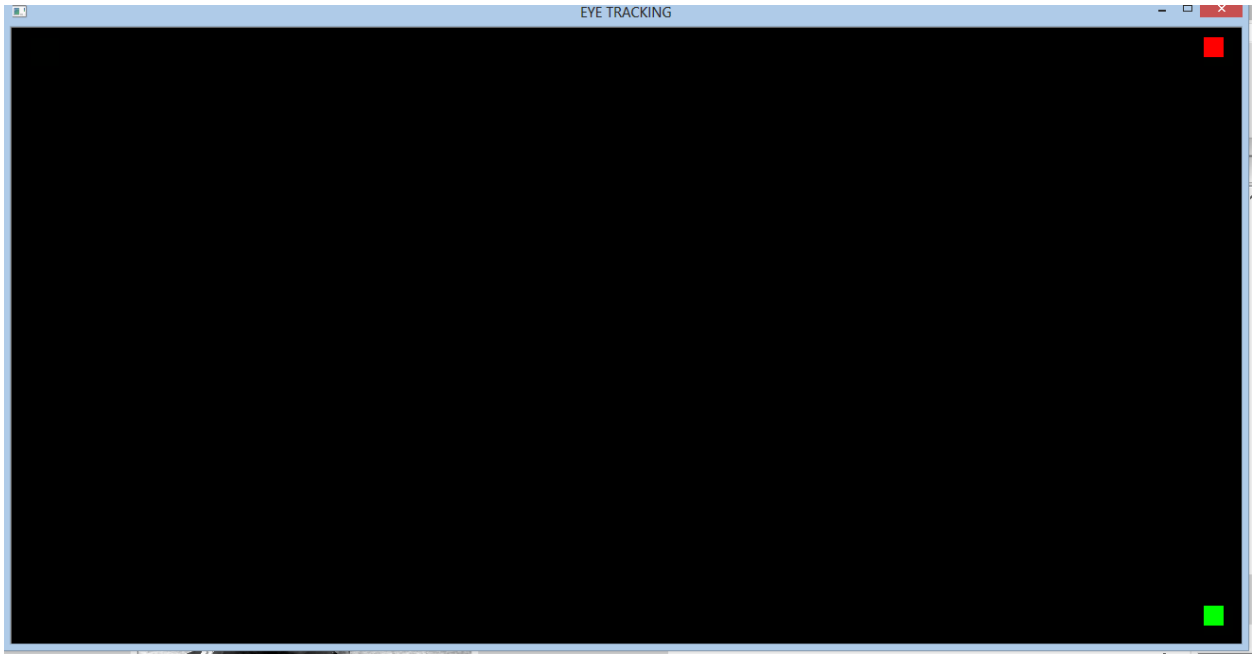

Figure11: as the person looks at the previous red point becomes green and a new red point is generated on the screen.

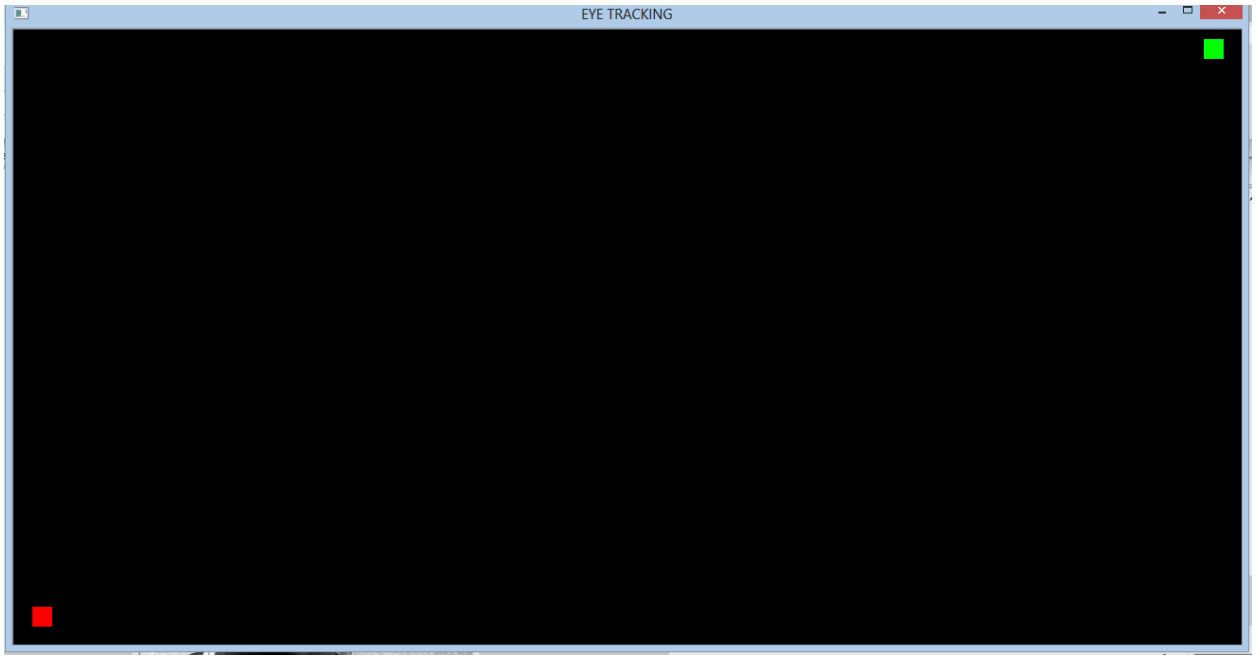

Figure11: similarly the person goes on looking till then the system verifies the user as genuine.

7.3 Timeout Mechanism

Time out is defined as a specified period of time that will be allowed to elapse in a system before a specified event is to take place, unless another specified event occurs first; in either case, the period is terminated when either event takes place.

A timeout mechanism is then used to restrict the system from verifying the videos. Our time out period is 2s, in this very time person must see from the green point to the red point.

Difference between the time when the person was looking at the green point and when the person was looking at the red point is recorded.

If he/she is not able to complete the task in these 2 second interval then the person is not live, otherwise he is said to be live person.

The process of generation of point is repeated and if the person passes all these live test correctly then he can access the system. Otherwise he cannot access the system. Doing so would add the robustness to the system.

Conclusion and Results

We have successfully determined the liveness of the user without the use of any special high-tech gadgets. By using a challenge response mechanism, the system differentiates the live person from the photos, videos. Using eye as determining feature, we need only normal web camera which is commonly available nowadays.

Although there are certain limitations like

1) The person must be seated 30 cm from the screen,
2) The person must not move his head while verification,
3) The person should be able to track the red point within 2 seconds.

According to our analysis, optimum distance from the camera is 30cm. Maximum accuracy obtained at optimum distance is around 85-90%. As we go away from the camera, the accuracy of the program decreases. Moving away 10cm decreased the accuracy to about 15-20%. Also, if the head movement is large during the process of authentication, chances of misclassification rises greatly.

Accuracy of the program also depends on the participation of the user as misclassifications increase if user is not able to respond to the challenge within timeout period.

REFRENCES

[1] SAFE: Secure Authentication with Face and Eyes Arman Boehm, Dongqu Chen, Mario Frank, Ling Huang, Cynthia Kuo, Tihomir Lolic, Ivan Martinovic and Dawn Song

[2] A. F. Abate, M. Nappi, D. Riccio, and G. Sabatino. 2d and 3d face recognition: A survey. Pattern Recognition Letters, 28(14):1885–1906, 2007.

[3]Dell FastAccess. http://www.sensiblevision.com/en-us/support/dellsupport.aspx.

[4]LenovoVeriface. http://support.lenovo.com/enUS/detail.page?LegacyDocID=MIGR-72561.

[5]Toshiba Face Recognition. http://us.toshiba.com/computers/research-center/technology-guides/face-recognition.

[6] Z. Zhu, Q. Ji. "Real-time eye detection and tracking under various light conditions." ACM papers.2001

[7] A. Haro, M. F. "Detecting and tracking eyes by using their physiological properties, dynamics, and appearance." Proc. Of IEEE Conf. on CVPR, 2000

[8] T.Kawaguchi, D.Hikada, and M.Rizon, "Detection of the eyes from human faces by hough transform and separability filter," Proc. of ICIP, pp.49-52, 2000

[9] D.Reisfeld, H.Wolfson, Y.Yeshurun, "Context free attentional operators: the generalized symmetry transform", Int.Journal of Computer Vision, 1995.

[10] S.Baskan, M. M. B., V.Atalay, "Projection based method for segmentation of human face and its evaluation". Pattern Recognition Letters 23, 1623-1629, 2002

[11] J. Huang, X.H. Shao, H. Wechsler. “Pose Discrimination and Eye Detection Using Support Vector Machines”. Proceeding of NATO-ASI on Face Recognition: From Theory to Applications, 1998

[12] S.Lucey, S.Sridharan, V.Chandran, “Improved facial feature detection for AVSP via unsupervised clustering and dicriminant analysis”, EURASIP Journal on Applied SignalProcessing, vol 3, pp.264-275, 2003

[13] Fabian Timm and Erhardt Barth “Institute for Neuro- and Bioinformatics”, University of Lubeck, Ratzeburger Allee 160, D-23538 L ¨ ubeck, Germany ¨ pattern Recognition Company GmbH, Innovations Campus Lubeck, Maria-Goeppert-Strasse 1, D-23562 L ¨ ubeck, Germany ¨ ftimm, barthg@inb.uni-luebeck.de

SUGGESTIONS OF BOARD MEMBERS