# Performance Evaluation of an Eye Tracking System under Varying Conditions

**Arpita Ray Sarkar†, Goutam Sanyal† and Somajyoti Majumder††,**

†National Institute of Technology, M.G. Avenue, Durgapur, WB 713209, India
††CSIR-Central Mechanical Engineering Research Institute, Durgapur, WB 713209, India

**Summary**
Presently, eye tracking is a one of the most important domain of research of computer vision dealing with the detection of eyes and associated parts, especially the corners, centres etc. to identify the point-of-interest of any person. This methodology has huge potential for application for the patients who have lost their communication abilities to communicate with the outside world using eye tracking. A large number of eye trackers are available commercially, but they failed to become popular due to higher cost and complexity. As a result, researchers are engaged around the world to develop low-cost eye tracking systems using most available components in the market. But, none of these systems are evaluated against different difficult and critical situations which they may come across. Such evaluation of the basic underlying methods may save time, effort and cost towards successful technological development. The present work evaluates the performance of an eye tracking methodology to be used to develop a rehabilitation system for stroke patients under different difficult and critical situations, such as different illuminations during day and night, low light during night, users with and without spectacles, variable distances between the camera and users etc. The optimum range of operation, minimum value and type of illumination, orientation of the light sources has been established through experimentation.
***Key words:***
*Face detection; Eye tracking; Haar-like features; Hough circle; Stroke; Rehabilitation*

## 1. Introduction

Eye tracking is a methodology to identify the point-of-interest where the person is looking at from his or her eye movements. Continuous tracking of eyes can reveal the behaviour, needs, emotional states, desires, cognitive processes etc. of a person [1]. As a result eye tracking has huge potential for application in fatigue assessment of drivers, surveillance, controlling robots and prosthetics [2], advertising etc. Over the last few years, substantial research has been carried out in this domain to develop robust and low-cost eye tracking systems for human-computer-interactions. Some of them are passive in nature and monitor only the movements of eyes to generate a conclusion for final decision, such as determination of behaviour of the driver, fatigue or drowsiness of the driver. In some other cases, active interactions can be performed by using eye tracking as a method for input to the computer. Controlling the mouse cursor, typing texts, using for rehabilitation supports are a few examples where the eye tracking system is directly used for controlling other devices. This has led to development of commercial eye trackers twenty years ago. But these commercial systems have failed to become popular, due to the higher price (even the minimum cost is around \$1500 [3]), often inability to operate in outdoor environments (where infrared lights do not work properly) and sometimes due to the discomfort caused due to wearing them on the body.

So, several researchers/ students are engaged in the development of low-cost, yet robust eye tracking systems using the easily available as well as cheaper components and open-source softwares. After development of these systems, none of them are tested scientifically and vigorously to define the performances. Rarely, some laboratory trials are performed to prove their performances. They are not evaluated in simulated environments closely replicating the actual environments. This present paper describes a methodology for evaluating the performance of a low-cost eye tracking system developed for the purpose of 24×7 support and rehabilitation of stroke and other locked-in patients. Different lighting conditions (illumination and position), different aged people (30 – 80 years), users with and without spectacles, variable in-between distances of 3 Ft to 9 Ft have been set as the parameters for evaluation.

## 2. Related Works

A substantial amount of research has been carried out by different researchers around the world for the development of eye tracking methodology and system, but only a few work related to the evaluation of such systems have been performed as reported in the literature. The evaluation process can be performed before final deployment either by following standard procedures as described in different International/National standards for application in a specific domain or by assessment against the critical parameters in a simulated environment similar to that of

Manuscript received April 5, 2017
Manuscript revised April 20, 2017

actual environment where it is supposed to work. A few of such evaluation works for eye tracking systems as reported in the literatures are discussed below.

Zhang & MacKenzie [4] have followed the procedure described in ISO 9241 – Part 9 which presents the ergonomic requirements for office work with visual display terminals and requirements for non-keyboard input devices. Performance testing was limited to compare the throughput in bits/s for four point-and-select tasks, three involving eye tracking and one using a standard mouse. Similar approach has also been followed by MacKenzie [5] for evaluating two commercial eye trackers for computer input. However, they did not perform the evaluation for longitudinal study and text entry applications. Also the measurement of biomechanical load has not been performed due to unavailability of sophisticated equipment.

Majority of the evaluation works have been performed in simulated environments through execution of actual tasks supposed to be performed by the systems. Some of them have used their own developed systems from commercially available components and some used commercially available costlier eye tracking systems. The ITU Gaze Tracker, developed by San Agustin et. Al. [6] has been evaluated in an eye-typing task using two different typing applications, one for low speed (3.56 words per min) and other for higher speed (6.78 words per min.). The system has been assessed for usability by a user with severe motor impairments in the home for typing texts through a wall-projected interface using his eye movements. Sibert & Jacob [7] have evaluated their eye gaze interaction technique through two experiments, one with circle experiment for determination of raw performance and other with letter experiment, where users first decides which object to select and then finds it. It has been found that the proposed gaze interaction technique is faster in selecting than selection with a mouse. Evaluation has also been performed [8] by determination of reliability and precision of the eye tracking system in two different virtual reality applications using the head-mounted device. It has been noted that the precision is found to be sufficient to identify the body part where the user is looking at and identify the behavioural statistics. Pfeiffer et. Al. have [9] carried out an evaluation to test accuracy, precision and application performance of two algorithms in combination with two available eye trackers. The main objective of the study was to find the combination of software and hardware suitable for 3D gaze-based interaction in Virtual Environments. But there are limitations depending on the applied VR technology in terms of insufficienct channel separation of the applied stereoscopy method and limited interaction space of the desktop based VR platform. Comparisons have been made by Nevalainen & Sajaniemi [10] for performance evaluation of three commercially available eye trackers (Tobii 1750, ASL 504 Pan/Tilt Optics and ASL 501 Head Mounted Optics) for psychology of programming research. Ahlstrom & Dukic [11] have compared availability, accuracy and precision between two commercial systems, one uses single camera and the other uses three cameras for monitoring the state of the drivers. The study has revealed that one-camera system is thus mostly suitable for in-vehicle applications.

The success of all the above methodologies or techniques depends up on the successful eye tracking which comes first. If different features of an eye and parts of the face (if required), such as the eye corners, eye balls, eye centres, nose, lip lines etc. are detected properly, gaze tracking can become easier. So, it is essential to correctly detect those required features first and then only the methodologies or techniques can be successful. As a result, the performance evaluation of the initial steps in critical conditions is necessary to save time and effort, but no such approach has still been reported in the literature. The present work describes the evaluation of an eye tracking methodology, already developed with intended application for continuous support and rehabilitation of stroke or other locked-in patients.

## 3. Methodology

Accurate eye tracking requires few essential steps to be followed. Most of the times, researchers use webcams close to the eyes to capture the image of the eyes. This causes inconvenience to the users. But here as the camera is placed at a distance from the user, it is utmost necessary to follow various image processing techniques to obtain best quality image for final tracking.

Initially, the face region is detected from the captured image amongst the cluttered background and then the eye region is detected from the previously detected face region. This approach provides better result than the direct detection of eye regions as the most of the cameras (webcam) are usually of low resolutions. As a result, the captured images are blurred and very difficult to work on. This approach substantially reduces the portion of the image on which image processing is to be done and this helps to reduce the computational load as well as processing time. The details of the associated steps have been described below.

### 3.1 Face and eye detection

Camera images containing clutter background should be processed first to detect the face region correctly. Then face detection is performed followed by detection of eye region. The details of the face and eye detection techniques have been described in the following paragraphs.

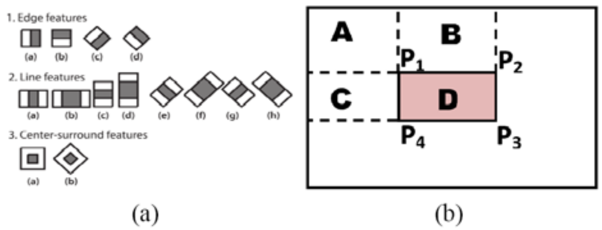



Fig. 1 (a) Haar-like Features [12] (b) Computation of sum of the pixels within rectangle D using integral image

The present research uses Haar-like feature-based cascade classifiers proposed by Paul Viola and Michael Jones [12] in 2001 to detect faces and eye regions from any image. Such Haar-like features have been presented in Fig. 1(a). In this method, a window of targeted size is moved over the input image and the Haar-like features are calculated for each sub-section of the image using integral image. Integral image at location $(x, y)$ is defined as the sum of all pixels within the rectangle ranging from the top left corner at location $(0,0)$ to bottom right corner at location $(x, y)$:

$$I(x, y) = \sum_{x' \le x, y' \le y} i(x, y) \tag{1}$$

where $i(x, y)$ is the grey level of input image. Rectangle features can be calculated after the completion of the calculation of all integral images. For example, as shown in Fig. 1(b) to calculate the sum of the pixels within region D, the integral images at $P_1$, $P_2$, $P_3$, $P_4$ are used. It can be written as:

$$D = A + B + C + D - (A + B) - (A + C) + A \tag{2}$$

or $D = I(P_4) - I(P_2) - I(P_3) + I(P_1)$ (3)

AdaBoost (Adaptive Boosting) is a machine learning technique proposed by Yoav Freund and Robert Schapire [13] and it can be used in conjunction with other learning algorithms. Viola and Jones introduced a highly efficient real-time face detection methodlogy combining fast features calculation with the AdaBoost algorithm and the cascade technique. All possible windows need to be checked for different positions and scales for detection of objects (faces or eyes) in an image. Use of a cascade of classifiers is preferred for real-time face or eye detection. Each stage of cascade, trained by the AdaBoost, will accept almost 100% of the positive images (contains objects) and reject 20 – 50% of the negative (contains non-objects) images. Every sub-window that is rejected at k-stage will be concluded that it doesn't contain a face or eye and ignored in the later stages. Linking n stages, the object in image can be detected with high rate.

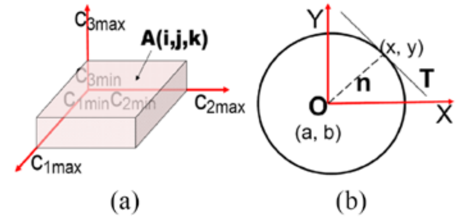



Fig. 2 (a) Representation of accumulator in 3D space (b) Relationship between (x,y,θ) and (a,b)

## 3.2 Eye centre detection

The next step is to detect the eye centres from the extracted eye regions. There are several methods for detection of eye centres from eye regions, such as application of colour based filter or shape based filter or use of Hough circles [14][15] etc. Another novel method is to train classifiers for detection of eyeballs like that of faces and eye regions. However, here Improved Hough circle based approach proposed in [15] has been used to detect eye centres.

Hough transform is an approach to detect circle in an image due to its applicability to any function in the form $g(v, c) = 0$, where $v$ and $c$ are the co-ordinate and co-efficient vector respectively. The co-ordinates of point on the circle satisfy Eq. 4.

$$(x - c_1)^2 + (y - c_2)^2 = c_3{}^2 \tag{4}$$

where $(c_1, c_2)$ is the co-ordinate of centre of the circle and $c_3$ is the radius of the circle. These three parameters $(c_1, c_2, c_3)$ can be mapped to a three dimensional rectangular co-ordinate system, so each accumulator $A(i, j, k)$ is equivalent to a cube structure graphic as shown in Fig. 2(a). Then the process of detection of points on the circle is evolved into refining and gradually increasing the values of $c_1$ and $c_2$ in $c_1c_2c_3$ plane and obtaining the value of $c_3$ which satisfies Eq. 4. The value stored in the accumulator $A(i, j, k)$ will become the number of many conical intersections in three dimensional parametric space.

Theoretically Hough transform can detect circles in images easily, but it is seldom applied for circle detection in digital image processing due to large number of parameters of round function, more computational complexity and poor detection efficiency. This has led to improvement of the Hough transform for application to two dimensional parametric space. Followings are the two associated steps for circle detection using improved Hough transform.

Step - I: First the circle centre is to be determined in two dimensional Hough transform space

Step - II: Then the circle radius is computed by statistic histogram

As shown in Fig. 2(b), the tangent $T$ to the circle will touch the circle at the point $(x, y)$ and the normal $n$ to the tangent will pass through this point and the centre of the circle. This principle can be used to find the centre of the circle. This normal component can be obtained using grey-scale edge detection operator. Then the three dimensional parameters $(x, y, \theta)$ can be mapped to two dimensional space $(a, b)$, where $(x, y)$ is the co-ordinate of the normal vector and $\theta$ is the angle of the normal vector with the x-axis. It is obvious from Fig. 2(b) that

$$\tan\theta = \frac{(y-b)}{(x-a)} \tag{5}$$

or $b = a\tan\theta + (y - x\tan\theta)$ (6)

The radius of the circle can be calculated from Eq. 7.

$$r = (x-a_0)^2 + (y-b_0)^2 \tag{7}$$

where $(a_0, b_0)$ is the co-ordinate of the centre of the circle and it can be calculated from Step - I. Traversing through all the points $(x_k, y_k)$ a statistical histogram of the related radius can be plotted and the peak value will denote the radius. The algorithm for the Improved Hough transform can be written as:

- First find edges and then for each edge point perform the following steps.
- Draw a circle with centre on the edge point with radius r and increment all co-ordinates that the perimeter of the circle passes through in the accumulator
- Identify one or several maxima in the accumulator
- Map the obtained parameters $(r, a_0, b_0)$ corresponding to the maxima back to the original image

Finally the detected eye centres need to be transformed to a global co-ordinate so as to create a map between the eye centres and co-ordinates on a display monitor or screen placed at a distance from the user.

## 4. Experiments

The main objective of the work is to evaluate the performance of an eye tracking system to be used for rehabilitation of stroke patients or other locked-in patients. Here, the daily activities of the patients can be supported by such a rehabilitation system and necessary support can be provided based up on the eye tracking on some pre-defined basic need icons. As the system is supposed to provide support to the patients day and night, it is essential to evaluate the performance of the system in various lighting conditions inside a room in day and night with users wearing spectacles and without spectacles. Finally, the essential requirement is to identify the optimum location/distance of the camera from the patient for best performance at varying illuminations during day and night.

The hardware components of the system include Intel NUC Kit (with 2.4 GHz Intel Celeron Processor N2820, DDR3 RAM and Intel HD Graphics), a USB camera (Logitech HD Webcam C525 with capability for auto-focus and auto-light correction) and a monitor. The Intel NUC kit is operated with 12V power. The USB camera has been selected based on its performance and price as mentioned in [16]. USB camera can be directly connected to the Intel NUC based minicomputer. For display standard monitors are used. One can also use low cost and portable projector and screen to replace the monitor.

Microsoft Visual Studio 2013 with Open CV libraries has been used here as the programming platform. Advantage of using Open CV is that it comes with a trainer as well as a detector. It contains many pre-trained classifiers for faces, eyes, smiles etc. developed by different researchers over the time. One can use either these pre-built classifiers or can develop own classifiers using positive and negative images. For the current work, pre-built libraries and classifiers have been used.

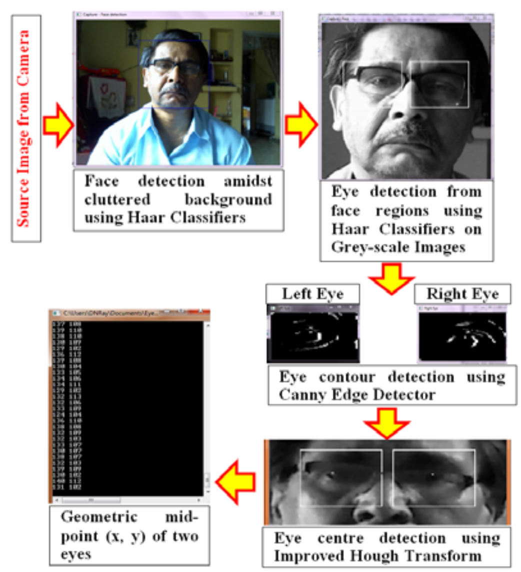
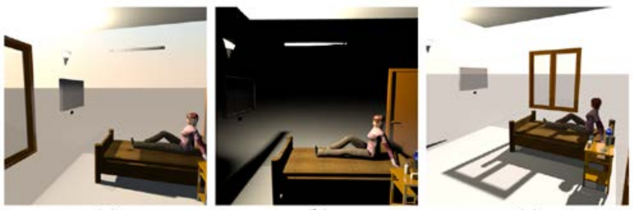



Fig. 3 Different associated steps of image processing for eye centre detection from a video using Improved Hough Transform

At the first step, the video is captured by the camera and loaded frame-by-frame. Then the required xml files (i.e. classifier files) are loaded and applied to each of these frames. A group of three consequent detections are merged as one. Pruning is necessary to exclude regions of these captured frames where it is unlikely to contain a face. The smallest size of face to detect can be set accordingly depending up on the needs. Similar steps are followed to detect the eyes. The detected regions are marked by a rectangular box. Thereafter the eye centres are detected from the eye regions (region of interest) and marked with '+' signs. The midpoint of those two eyes are also calculated and marked in the picture with black diamond '◊' sign. The different windows associated with the detection are presented in Fig. 3.



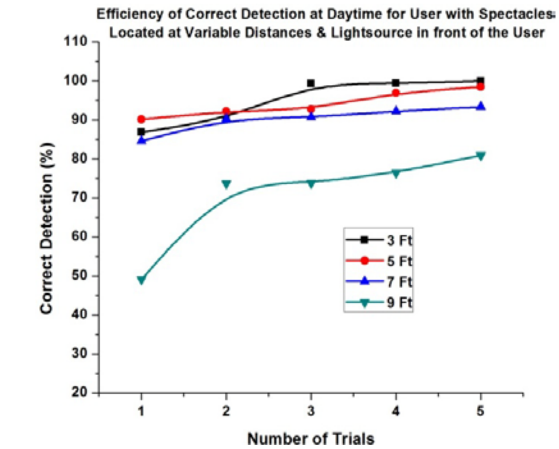
(a) (b) (c)

Fig. 4: View of the experimentation room (a) during daytime with light source in front of user (b) during night with light source in front of user (c) during daytime with light source on one side of the user

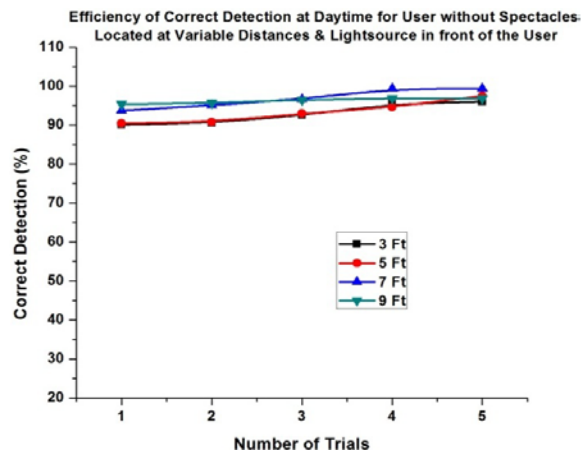
As the proposed system has to work for patients living in normal rooms of houses or hospitals or physiotherapy clinics, the experiments have been carried out in a room with dimension 12 ft by 12 ft similar to that in [3]. The views of the experimentation room under different illuminations and conditions are shown in Fig. 4(a) - (c). The user rests on a bed in the room in an inclined position. Here as the system is not wearable one, it is very much necessary to find out the optimum location of the camera. SO, the camera is placed at different distances from the user to find out the best performance. The distances between the camera and the eye of the user have been set as 3 Ft, 5 Ft, 7 Ft and 9 Ft. Position of the light source with reference to the patient is another important factor that may affect the performance of the eye tracking system. For this the light source has been placed mostly in front of the users, but performance has also been evaluated placing the light source on one side of the user. The monitor is mounted on the wall facing the patient directly. The video data has been stored on the computer and later processed for further analysis. A number of users have been engaged to gather the data for different illuminations (during day and night) and parameters/ conditions. The minimum value of the illumination was kept close to 4 Lux as specified in IS: 4347 – 1967 (Reaffirmed 2005) using low-power lights [17]. Some of the users used spectacles and some users were without spectacles. The online video data for the window with the detection signs are stored on the computer using third-party software.

## 5. Results and Discussions

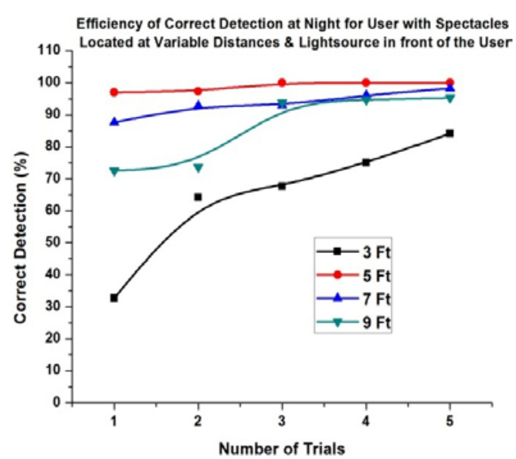
The data stored on the computer were later analyzed to evaluate the performance of the eye tracking system. The average number of frames for the video data set is 370 for different lighting conditions (day and night) with users wearing spectacles and without spectacles.



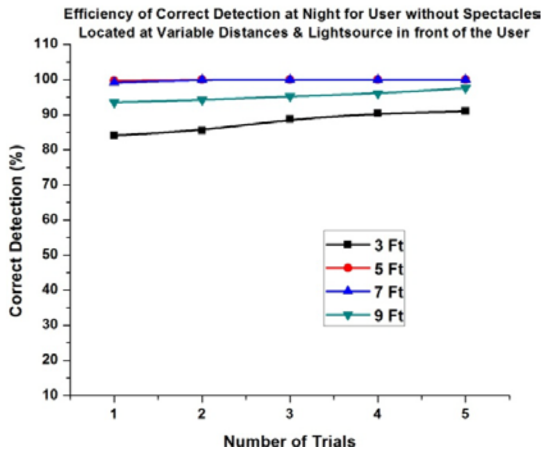

(a)




(b)



(c)




(d)

Fig. 5: Efficiency of detection with light source in front during (a) daytime for users with spectacles (b) daytime for users without spectacles (c) night for users with spectacles (d) night for users without spectacles

The average illuminations during each set of experiments have been found to be different depending upon the distance between the user and location as well as intensity of the light source. In daytime the illumination is obviously higher with natural light source. During night it is lower. Though experiments during the night were started with very low illumination, but later normal illumination as found in day-to-day life was maintained. The average values of illuminations during the experiments for different positions and types of the light sources have been presented in Table - 1.

Table 1: Average Values of Illuminations for Different In-between Distances and Location of Light source during the Experiments

| ***Sl.*** | ***Experimental Conditions*** | ***Dist.-3Ft*** | ***Dist.-5Ft*** | ***Dist.-7Ft*** | ***Dist.-9Ft*** |
|---|---|---|---|---|---|
| 1. | (a) Time of the day: Daytime (b) Light source: In front of users | 380 Lux | 350 Lux | 230 Lux | 160 Lux |
| 2. | (a) Time of the day: Night (b) Light source: In front of users | 5 Lux | 60 Lux | 36 Lux | 18 Lux |
| 3. | (a) Time of the day: Daytime (b) Light source: On left side of users | 78 Lux | 82 Lux | 128 Lux | 155 Lux |

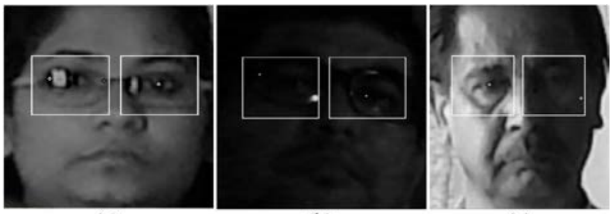
Here, the main objective is to detect both the eye centres correctly and the correct detection will be considered when both the eye centres have been detected in a frame. Percentage of correct detection for any video data set is the total number of such correctly detected frames against the total number of frames captured. Percentage of correct detection at daytime and night for users with and without spectacles and light source in front of the patient have been calculated and presented in the graphs in Fig. 5(a) to (d). Simultaneously, the effect of placement of the light source to one side of the user has also been investigated especially during the daytime and has been discussed here.

Fig. 5(a) depicts the percentage of correct detection during daytime for users (with spectacles) located at different distances from the camera and the light source placed in front of the user. The average illumination for experiments during daytime with light source placed in front of the user is presented at sl. No. 1 of table 1. The average percentages of correct detection in this case for the in-between distance of 3 Ft, 5 Ft, 7 Ft and 9 Ft are 95, 94, 90 and 70 respectively. It is very normal to note that the lowest efficiency of detection is obtained for largest in-between distance of 9 Ft. The standard deviations for the first three cases lie within 6 and for the farthest case, it is around 13. The efficiency of detection for in-between distance of 9 Ft is lower due to presence of glaring on the glasses. The grey scale image with glaring has been presented in Fig. 6(a).

Similarly the percentages of correct detection for users without spectacles at daytime (light source placed in front) have been presented in Fig. 5(b). In all the cases the average percentages of detection lie above 90. The lowest one for the in-between distance of 3 Ft is approximately 93% and highest one for in-between distance of 7 Ft is around 97%. Here the range of standard deviations is lower than the earlier case and lies within 3. It is worthy to note that the minimum one is for the in-between distance of 7 Ft and it is 0.7 only. This reveals that the data are very concentrated around the mean value which in turn refers the robustness of the system.

The most critical situation for correct detection may occur in the night when the illumination is low and user wears spectacles. Glaring, reflection and availability of sufficient lights are the major hindrances for such situation. It has been found from Fig. 5(c) that the average percentages of correct detection in night for users with spectacles and located at a distance of 3 Ft, 5 Ft, 7 Ft and 9 Ft are 65, 99, 94 and 86 respectively. Except the first and last cases, the standard deviations are very close (within 4). For in-between distance of 3 Ft standard deviation is around 19.5 and that is 12 for the distance of 9 Ft.

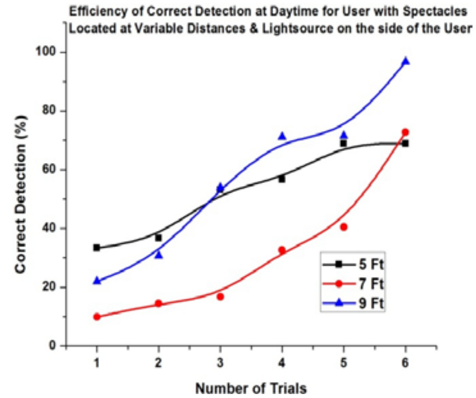
However, the average efficiency of detection for cases when users are not wearing spectacles at night is higher than expected. Absence of glasses helps the eye tracking system to detect the eye centres correctly. The average percentages of correct detection are 88, 99.9, 99.8 and 95 for the in-between distances of 3 Ft, 5 Ft, 7 Ft and 9 Ft respectively. The average standard deviation is found to be the lowest for this case. The maximum and minimum standard deviations are approximately 3 and 0.15 for in-between distances of 3 Ft and 5 Ft respectively.



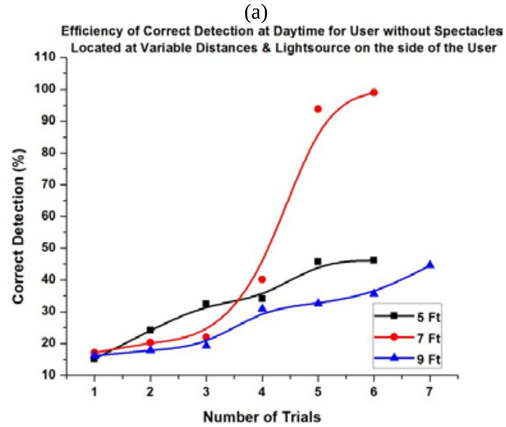
(a) (b) (c)

Fig. 6 The grey scale image of the faces for (a) users (9 Ft away from the camera) wearing spectacles during daytime with light source in front of the user showing presence of glaring on the glasses (b) users (3 Ft away from camera) wearing spectacles during night with very low illumination showing glaring on the glasses (c) user at daytime with light source on one side showing shadow on the half of the face

Performance evaluation under low illuminations is also an essential requirement for such rehabilitation systems. So, different illuminations at different in-between distances have been carried out here during the night. Initially for the nearest position i.e. for in-between distance of 3 Ft the illumination was kept very low i.e. around 5 Lux as per IS: 4347 – 1967 using low power light (10 Watt night-lamp) and this has resulted in lower efficiency of 65% for users with spectacles. Low light and glaring are the two main difficulties for detection in this case. The grey scale image with detection marks in low illumination is shown in Fig. 6(b). But for users without spectacles, the efficiency of 88% is quite appreciable for such low illumination. Later, normal lights (18 Watt LED light) often used in homes/ hospitals during night has been used to generate a variety of different illuminations. The average values of illumination are kept at 60, 36 and 18 Lux for in-between distances of 5 Ft, 7 Ft and 9 Ft respectively. In all the cases, the light source is placed 12 Ft away in front of the users at a height of 10 Ft from the floor as often done in normal homes. As discussed above, the range of detection is better and acceptable at normal lighting conditions for living in night due to presence of uniform brightness on the face.



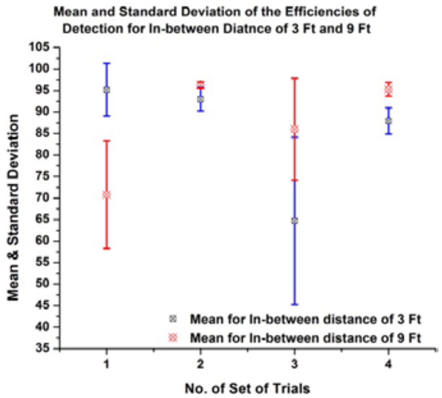

(a)



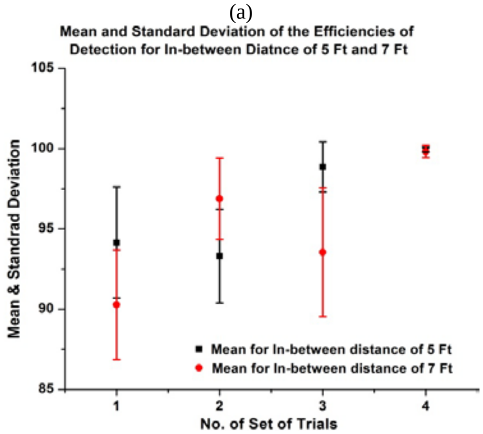

(b)

Fig. 7: Efficiency of detection when light source on the left side during (a) daytime for users with spectacles (b) daytime for users without spectacles

A distinguishable deterioration in the performance is noticed when the (natural) light source is located on the side of the user. Here, the user was so placed that the open window (light source) lies on one side of the user. A typical grey scale image of the user for such scenario has been presented in Fig. 6(c). The average percentage of efficiency of detection ranges from 33 to 65 only. This can be noticed from Fig. 6(c) that there is shadow/ darkness on the farthest side of the face of the user due to placement of the light source on the other side. As a result the face is not uniformly enlightened and only one eye centre is found to be detected all the times. The average efficiencies of detection for users with spectacles and without spectacles are 52 and 41 respectively as shown in Fig. 7(a) and (b). Also the data are widely spread with a standard deviation of around 20. Such lower efficiency with higher standard deviation may not be acceptable for any eye tracking system to be used for the purpose of rehabilitation.



(a)



(b)

Fig. 8 Mean and Standard Deviation of the Efficiency values of Correct Detection for Different Users (with and without spectacles) during Day and Night for In-between Distances of (a) 3 Ft and 9 Ft (b) 5 Ft and 7 Ft

However the present eye tracking system reveals a linear behaviour irrespective of illumination, appearance of users (with or without spectacles) and increased distance. The mean values of detection for users with spectacles during daytime, users without spectacles during daytime, users with spectacles at night and users without spectacles at night are approximately 88%, 95%, 86% and 96% respectively. These average values include highest values of correct detection as 100% as well as lowest value as 32%. So, it reveals that the present eye tracking system has the capability to detect eye centres 100% correctly if proper illumination and camera with capability of detection in lower illumination are used. The primary requirement is that illumination should be sufficient and uniform on the face as well as the position of the light and camera should be such to reduce the glaring on the glasses. In the present research work the minimum acceptable value of illumination has been found to be 18 Lux at night. This has resulted in average percentages of detection as 86 and 95 for users with and without spectacles respectively for a distance of 9 Ft between the camera and the user.

Though an average percentage of correct detection of 87 has been achieved for different users with in-between distance of 9 Ft and the light source placed in front of the user, but the standard deviation is higher i.e. more than 6 as shown in Fig. 8 (a). Similar situation also arises in case of the in-between distance of 3 Ft. The average percentage of correct detection and the standard deviation are 85 and 7.8 respectively. However, the average percentages of detection for in-between distances of 5 Ft and 7 Ft are found to be 97 and 95 respectively. The standard deviations are also lower and in the range of 2 and 2.6 for in-between distances of 5 Ft and 7 Ft respectively as shown in Fig. 8(b). So, the optimum range of operation of the present eye tracking system is found to be between 5 Ft to 7 Ft and minimum acceptable illumination is 18 Lux (at night) for uniformly enlightened face with the light source placed in front of the users.

A detailed comparison of the present work with other existing evaluation works has been presented in the table-2 below. The above discussion reveals that the present method can provide 100% efficiency of detection, when proper and uniform illumination is used. In none of the evaluations, discussed in the table, such performance has been achieved. The maximum efficiency has been obtained as 80% by Zhang & MacKenzie [4], MacKenzie [5] at a distance of 2 Ft. The maximum in-between distance tested is only 3 Ft (i.e. 90 cm) resulting in average efficiency of 54% for selection of circle and 76.5% for selection of letter. It is also to note that performance in lower illumination has not been performed though it is essential for system to be used in real-time. The present work has been tested successfully under lower illumination also.

Table 2: Comparison with Related Works

| ***Paper Reference*** | Zhang & MacKenzie, MacKenzie | Agustin et. Al. | Sibert & Jacob | Ciger et. Al. | Present Method |
|---|---|---|---|---|---|
| ***Objectives/ Experiments*** | Pointing & selection of an on-screen target using | Eye-typing test to produce texts | Selection of targets using dwell time | Evaluated Reliability & Precision Test | Eye tracking for selection of icons |

| | | | | | |
|---|---|---|---|---|---|
| ***Distance*** | 60 cm (2 Ft) | 60 cm (2 Ft) | 90 cm (3 Ft) | Not applicable (Head mounted) | 91 cm (3 Ft) to 274 cm (9 Ft) |
| ***Performance*** | 3.78 bits/s w.r.t to 4.68 bits/s for mouse (80%) | Typing of 3.56 to 6.78 words per min. | 54 % (Mean = 503.7 w.r.t. Mean for Mouse = 931.9) for circle selection & 76.5% (Mean = 1103 w.r.t. Mean for Mouse = 1441) for Letter | Maximum 79% Reliability | Average 91% (Maximum 100% & minimum 65%) |
| ***Low-light Experiments*** | Not performed | Not performed | Not performed | Not performed | Yes (18 Lux) |

## 6. Conclusion

Proper assessment of the essential and initial steps of any method or technology can save a lot of time, effort and sometimes cost towards the successful development. Majority of the works, as reported in the literature, evaluate the performance at the end. But, pre-evaluation of the underlying steps can provide corrective measures beforehand so as to avoid delays and failures in the final development. The present work evaluates the performance of an eye tracking technique that is a major and most essential step towards the development of a vision-based rehabilitation system. This technique uses Haar-feature based Cascade classifier and Improved Hough transformation for detection of the eyes and eye centres respectively. Performance of this hybrid methodology is then evaluated in a simulated environment similar to that of actual. The illumination, different positions and types of light sources, distance between the camera and the user, different users (with and without spectacles) are varied to check the robustness, acceptability and efficiency of the detection.

It has been found that the developed eye tracking system can easily achieve an average overall efficiency of correct detection as 91% (the maximum is 100% and the minimum is 65%) for a range of distances of 3 Ft to 9 Ft. Though the optimum range has been found to be between 5 Ft to 7 Ft, performance for a distance of 9 Ft is quite better and acceptable. The average percentage of correct detection for such case during day and night for users with and without spectacles has been achieved as 87, but the standard deviation is little higher i.e. 6.66. In the most preferred cases, this standard deviation has been around 2.3 with average value as 96%. From table-2 it is obvious that none of the works reported in the literature has achieved such higher efficiency. Furthermore, correct detection at low light during the night is another challenging issue for rehabilitation systems. The situation becomes more critical if the user wears spectacles. The detection of eye centres becomes difficult due to glaring of glasses. The present system has been tested against a minimum illumination of 5 Lux. However, the minimum effective illumination has been found to be 18 Lux which results in an average percentage of correct detection as 86 at night for user with spectacles for in-between distance of 9 Ft. It is clear from the literatures that such higher efficiency has not been achieved by any other eye tracking systems for similar situations. So, it can be concluded that this methodology has the potential to be used for detection of correct and accurate gazes towards further development of an efficient rehabilitation and support system. Work is in progress to develop the final gaze detection system and implement it in real time for rehabilitation of stroke and other locked-in patients.

**Acknowledgments**

Authors are thankful to the Head and other faculty members of Department of CSE, NIT-Durgapur and SR Lab, CSIR – CMERI Durgapur for their continuous help and support. Authors are also grateful to Dr. D. N. Ray for his continuous suggestions and advices.

## References

[1] D. W. Hansen, Q. Ji, “In the Eye of the Beholder: A Survey of Models for Eyes and Gaze”, IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 32 , No. 3; March 2010, pp. 478 – 500

[2] C. Castellini, G. Sandini, “Gaze Tracking for Robotic Control in Intelligent Teleoperation and Prosthetics”, The 2nd Conf. on Communication by Gaze Interaction (COGAIN 2006) : Gazing into the Future; Turin, Italy; Sept. 2006, pp. 73 - 77

[3] A. Ray Sarkar, G. Sanyal, S. Majumder, “Methodology for a Low-Cost Vision-Based Rehabilitation System for Stroke Patients”, Advancements of Medical Electronics: Lecture Notes in Bioengineering 2015, Springer, Jan. 2015, pp 365 – 377

[4] Xuan Zhang, I. Scott MacKenzie; Evaluating Eye Tracking with ISO 9241 - part 9; In Proc. of the 12th international conference on Human-computer interaction: intelligent multimodal interaction environments; Beijing, China; July 22 - 27, 2007; Pages 779-788

[5] I. S. MacKenzie; Evaluating eye tracking systems for computer input; In Gaze interaction and applications of eye

tracking: Advances in assistive technologies, P. Majaranta et. Al.; Hershey, PA: IGI Global; 2012; Pages 205-225

[6] Javier San Agustin, Henrik Skovsgaard, Emilie Mollenbach, Maria Barret, Martin Tall, Dan Witzner Hansen, John Paulin Hansen; Evaluation of a Low-Cost Open-Source Gaze Tracker; In Proc. Of the 2010 Symposium on Eye-Tracking Research & Applications; Austin, Texas; March 22 - 24, 2010; Pages 77-80

[7] Linda E. Sibert, Robert J. K. Jacob; Evaluation of eye gaze interaction; In Proc. of the SIGCHI conference on Human Factors in Computing Systems; The Hague, The Netherlands; April 01 - 06, 2000; Pages 281-288

[8] Jan Ciger, Bruno Herbelin and Daniel Thalmann; Evaluation of Gaze Tracking Technology for Social Interaction in Virtual Environments; In Proc. of the 2nd Workshop on Modeling and Motion Capture Techniques for Virtual Environments (CAPTECH'04); 2004; Pages 1-6

[9] Thies Pfeiffer, Marc E. Latoschik, Ipke Wachsmuth; Evaluation of Binocular Eye Trackers and Algorithms for 3D Gaze Interaction in Virtual Reality Environments; JVRB - Journal of Virtual Reality and Broadcasting; Vol. 5, No. 16; 2008

[10] Seppo Nevalainen and Jorma Sajaniemi; Comparison of Three Eye Tracking Devices in Psychology of Programming Research; In Proc. 16th Workshop of the Psychology of Programming Interest Group; Carlow, Ireland; April 2004; Pages 1-8

[11] Christer Ahlstrom, Tania Dukic; Comparison of eye tracking systems with one and three cameras; Proceedings of Measuring Behavior 2010; Eindhoven, Netherlands; August 24 - 27, 2010; Pages 196-199

[12] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features", In Proc. of IEEE Conference Computer Vision and Pattern Recognition, 2001

[13] Y. Freund and R. E. Schapire, “A decision-theoretic generalization of on-line learning and an application to boosting”, J. of Computer and System Sciences, Vol. 55, No. 1, August 1997, pp. 119 – 139

[14] M. Soltany, S. T. Zadeh and H.-R. Pourreza, “Fast and Accurate Pupil Positioning Algorithm using Circular Hough Transform and Gray Projection”, In Proc. of 2011 Int. Conf. on Computer Communication and Management, Vol. 5, Singapore, pp. 556 - 561

[15] D. Chen, J. Bai, Z. Qu, "Research on Pupil Center Location Based on Improved Hough Transform and Edge Gradient Algorithm", In Proc. of National Conf. on Information Technology and Computer Science (CITCS); Lanzhou, China; Nov. 2012, pp. 47 – 51

[16] A. Raysarkar, G. Sanyal and S. Majumder, “Performance Comparison of Different Low Cost Cameras for Development of an Eye-Tracking System” [unpublished]

[17] Bureau of Indian Standards, Indian Standard: Code of Practice for Hospital Lighting, IS: 4347 – 1967 (Reaffirmed 2005)

**Arpita Ray Sarkar** received the B. Tech. degree in IT from Kalyani University in 2004 and M. Tech. degree in Information Security from National Institute of Technology, Durgapur, India in 2012. Currently she is pursuing Ph. D degree from NIT, Durgapur in the field of computer vision. Her field of interest includes image processing, computer vision, rehabilitation systems etc.

**Prof. Goutam Sanyal** is currently working as Professor and Head of Dept. of Computer Sc. & Engg in National Institute of Technology, Durgapur. He has received the Ph.D degree from Jadavpur University in 2000 and achieved B.E. and M.Tech degree from R E College Durgapur, West Bengal, India. in 1977 and 1990. He has More than 30 years of experience in Research and teaching. His research interests are Computer Vision/Image Processing, Machine Learning, NLP, Computer Network, Information Security and Cloud Computing.

**Dr. Somajyoti Majumder** is a doctorate in sensor data fusion techniques from University of Sydney, Australia. He has obtained his Masters degree from Indian Institute of Technology, Kharagpur and Bachelor’s degree from University of Calcutta. He is working in the field of robotics and automation system, data fusion, Artificial Intelligence, Image processing for last 30 years. He is currently holding the post of senior scientist at CSIR-CMERI, Durgapur.