International Journal of Applied Engineering Research ISSN 0973-4562 Volume 12, Number 9 (2017) pp. 1819-1825

Performance Comparison of Different Low Cost Cameras for Development of an Eye-Tracking System

Arpita Ray Sarkar*, Goutam Sanyal and Somajyoti Majumder
National Institute of Technology, Durgapur, India.
CSIR-Central Mechanical Engineering Research Institute, India.
ORCID ID: 0000-0001-6629-2059

Abstract

Eye tracking is very useful to monitor and assess the behaviors, needs, emotional states, and desires of human beings, such as monitoring the drowsiness of a driver, or providing rehabilitation support to stroke and locked-in patients. The commercial eye trackers, developed 20 years ago, have failed to become popular for several reasons. As a result, development of low-cost eye-tracking systems using easily available and affordable cameras has become a major domain of research. The camera's performance plays a vital role in efficient eye tracking in spite of different illuminations, users, and backgrounds. The present research evaluates the performance of six popular, low-cost, easily available cameras for development of an efficient eye-tracking system to be used for rehabilitation purposes of stroke and other locked-in patients. Different illuminations at day and night, different aged users, and users with and without spectacles have been considered during the evaluation.

Keywords: Performance Evaluation, Webcams, Digital Cameras, Face Detection, Image Processing, Patient Rehabilitation

INTRODUCTION

Eye tracking is a methodology to identify the point-of-interest a person is looking at from his or her eye movements. Continuous tracking of eyes can reveal a person's behaviors, needs, emotional states, desires, and cognitive processes (Hansen & Ji, 2010). As a result, eye tracking has huge potential for application in such areas as fatigue assessment of drivers, surveillance, controlling robots and prosthetics (Castellini & Sandini, 2006), and advertising etc. Based on these demands, substantial research in this domain to develop robust eye -tracking systems for human-computer-interactions led to the development of commercial eye trackers 20 years ago. But these commercial systems have failed to become popular, due to high cost (even the minimum cost is around $1,500; Ray Sarkar, Sanyal, & Majumder, 2015), frequent inability to operate in outdoor environments (where infrared lights do not work properly), and discomfort caused due to wearing them on the body.

Today, several researchers and students are engaged in the development of low-cost, yet robust eye-tracking systems using easily available and affordable components as well as open-source software.

The camera plays one of the major roles in the development of such an eye-tracking system. Good visibility in low-light conditions often poses a challenge for these systems, as they are expected to be used for 24/7 support and for rehabilitation of stroke and other locked-in patients in hospitals, clinics, rehabilitation centers, and homes. Thus, the cost of the system should be kept as low as possible. Therefore, it is necessary to compare the performance of different available low-cost (within 100 US$) cameras and to identify the most suitable, low-cost one for developing an eye-tracking system. This article describes a comparative study of the performance of different low-cost cameras towards selection of the best one for development of an eye-tracking system. Different lighting conditions, different aged people, and users with and without spectacles from a distance of 5 feet have been used for the evaluation.

RELATED WORKS

The efficiency of eye-tracking systems depends on the targeted users, different parameters (illumination, position of camera and light source, distance between camera and user, occlusion, etc.), and hardware. A single hardware may not address all the applications simultaneously. For example, the eye-tracking hardware for indoor application may not be suitable for application outdoors, as it may use infrared light (which may not work properly in an outdoor environment) in the detection of the eye center. Therefore, it is essential to evaluate the hardware beforehand for a particular application.

Most of the hardware for today's eye-tracking systems can be categorized into two different classes based on the constituent hardware components. One uses the commercially available eye trackers (e.g., S2 of Mirametrix, TM4 and VT2 of EyeTech DS, EyeX of Tobii, EyeFollower and EAS Binocular of LC Technologies Inc., GP3 of Gazepoint. etc.), and others use different available cameras for eye tracking. Commercial eye trackers are compact and have proven their capabilities mostly in indoor environments. But they are costlier for common people and hence have not been popular even 20 years after development and commercialization. Several works have been reported in the literature to compare the performance of different eye trackers (Ahlstrom & Dukic, 2010; Cheng & Vertegaal, 2004; Funke et al., 2016; Janthanasub & Meesad, 2015; Nevalainen & Sajaniemi, 2004).

The other category uses either webcams or analog or digital cameras for the purpose of eye tracking. Natural light and infrared light have been used for illuminating the environment.

Most of these cameras are replaceable due to their lower cost and easy availability. The camera is the first and one of the most essential components for eye-tracking systems. The performance of an eye-tracking system is highly influenced by the quality of the camera, even if the same algorithm with the same set-ups is utilized. Every effort should be made to identify the best camera at lowest cost for a particular application in a typical environment.

However, the majority of the works available in the literature are for performance evaluation of the complete eye-tracking system rather than for the camera only in a simulated environment or in an environment that is a close-approximation to the real environment. Zhang and MacKenzie (2007) have followed the procedures described in ISO 9241 - Part 9, which presents the ergonomic requirements for office work with visual display terminals and the 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. The ITU Gaze Tracker, developed by Agustin et. al. (2010), has been evaluated in an eye-typing task using two different typing applications, one for low speed (3.56 words per minute) and other for higher speed (6.78 words per minute). Sibert and Jacob (2000) have evaluated their eye-gaze interaction technique through two experiments, one with a circle experiment for determination of raw performance and the other with a letter experiment, where users first decide which object to select and then find it. Evaluation has also been performed (Ciger, Herbelin, & Thalmann, 2004) by determination of reliability and precision of the eye-tracking system in two different virtual reality applications using the head-mounted device. Pfeiffer, Latoschik, and Wachsmuth, (2008) have made 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 none of the works describe or evaluate the reason for the use of a particular camera model for eye tracking. However, Ferhat (2012) has described the basis of the use of a particular camera for his work. He has identified high frame rate, high resolution, and good image quality as their requirements. A number of different low-cost cameras, such as Logitech C615 (1080p), Logitech (720p), Microsoft HD-3000 (720p), and several other less expensive models (from brands such as Trust, NIUM, and Hercules) have been tested, with Logitech C615 selected as the most suitable. But detailed reasons for this choice are not available anywhere in the work.

The present work describes the experimental evaluation of different low-cost cameras (within US$100), including Logitech C525 and C920, for developing a low-cost eye-tracking system for day-and-night rehabilitation support to stroke and other locked-in patients. Extreme difficult situations, such as very low illumination and users with spectacles (occlusion and glaring are a very common phenomenon in such cases) have been imposed for the evaluation without any special lighting or special attachments. Such experimental evaluation in a real-time scenario has not been performed elsewhere as revealed from the literature.

METHODOLOGY

Accurate eye tracking requires a few essential steps to be followed. Most of the time, researchers use webcams close to the eye to capture the image of the eyes (often referred as Do-It-Yourself; Mantiuk, Kowalik, Nowosielski, & Bazyluk, 2012). This causes inconvenience to the users, but as the camera is placed at a distance from the user, it is necessary to follow various image-processing techniques to obtain the best quality images for final tracking.

First, the face region is detected from the captured image among 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 most of the cameras (webcam) are usually of low resolution. 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, which helps to decrease the computational load as well as the processing time.

Haar feature-based cascade classifiers have been used here to detect the face and eye. Viola and Jones (2001) proposed an effective way of object detection known as the Haar feature-based cascade classifier. It is a highly efficient, real-time face detection system that combines fast-features calculation with the AdaBoost algorithm (Freund & Schapire, 1997) and the cascade technique. For detection of objects (faces or eyes) in an image, all possible windows need to be checked for different positions and scales. Use of a cascade of classifiers is suggested 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 images (contains non-objects). For every sub-window that is rejected at k-stage, it will be concluded that it doesn't contain a face and is ignored in the later stages. Linking n stages, the object in the image, can be detected at a high rate.

The next step is to detect the eye centers from the extracted eye regions. For eye center detection, several methods are available, such as application of color-based filtering, shape-based filtering, or detection using the circle Hough Transform (Chen, Bai, & Qu, 2012; Soltany, Zadeh, & Pourreza, 2011). However, in the present research, the Improved Hough Transform circle-based approach proposed in Chen et al. (2012) has been used to detect eye centers.

Although the Hough Transform can detect circles in images theoretically, it is hardly applied for circle detection in digital image processing due to large numbers of parameters of round function, more computational complexity, and poor detection efficiency. As a result, this has led to improvement of the Hough Transform for application to two-dimensional parametric space. The following are the two associated steps for circle detection using The Improved Hough Transform:-

- Step - I: Find the circle center in the two-dimensional Hough Transform space.
- Step - II: Obtain the circle radius by statistic histogram.

Finally, the detected eye centers must be transformed to a global co-ordinate so as to create a map between the eye centers and the co-ordinates on a display monitor or screen placed at a distance from the user. The summary of these steps is presented in Figure 1.

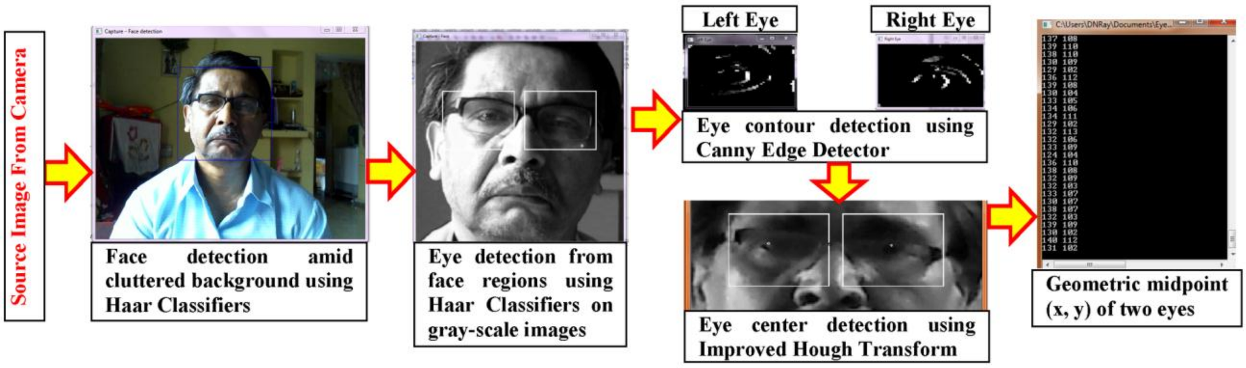


Figure 1. Different associated steps of image processing for eye tracking from a video

EXPERIMENT

The main objective of the work is to evaluate the performance of different low-cost (within 100$) and easily available USB Web cameras, wireless cameras, and night-vision cameras for development of an eye-tracking system to be used for rehabilitation of stroke patients or other locked-in patients. Here, the patients' daily activities can be supported by such a system, and necessary support can be provided based 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 compare the performance of the cameras in various lighting conditions of the room during the day and at night with users with and without spectacles from a standard distance.

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 or analog) camera, and a monitor. The Intel NUC Kit is operated with 12V power. The USB camera can be directly connected to the mini computer. The night-vision cameras are analog in nature, and they can only be connected to the computer through a frame grabber or TV tuner card. The six different cameras as mentioned in Table 1 are used here for performance evaluation. A standard monitor has been used for display. One can also use a low-cost and portable projector and screen to replace the monitor.

Table 1. Specifications and approximate price of different low-cost cameras

Sl.	Camera Make and Model	Approximate Price (US$)	Type	Resolution (Pixels)	Photo Quality	Special Features
Camera1	Logitech HD Webcam C525	58	USB, HD	1280 × 720	8 mega pixels	• Autofocus • Autolight correction
Camera2	Logitech HD Webcam Pro C920	65	USB, HD	1920 × 1080	15 mega pixels	• Autofocus • Low-light correction
Camera3	Frontech JIL 2243 Webcam	12	USB, CMOS	300K pixels	20 mega pixels	• White balance • Manual adjusting focus
Camera4 (wireless) and Camera5 (wired)	JMK IR Wireless Night Vision CCTV Security System Camera	44	2.4 GHz wireless and wired	PAL: 628 × 582 NTSC: 510 × 492	380 TV lines	Minimum illumination: 3LUX
Camera6	JMK 007 Mini Pinhole Wireless Color Camera	30	1.2 GHz wireless CMOS	380 TV lines (horizontal)	-	Minimum illumination: 3LUX

Microsoft Visual Studio 2013 with Open CV libraries has been used here as the programming platform. The advantage of using Open CV is that it comes with a trainer and a detector. It contains many pre-trained classifiers for faces, eyes, smiles, and similar which have been developed by different researchers over the time. These pre-built classifiers can be used, or classifiers using positive and negative images can be developed. For the current work, pre-built libraries and classifiers have been used.

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 the 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 to detect can be set accordingly, depending on needs. Similar steps are followed to detect the eyes. A rectangular box marks the detected regions. Thereafter, the eye centers are detected from the eye regions (region of interest) and marked with white diamond '◊' signs. The geometric midpoint (x, y co-ordinates w.r.t fixed camera position) of those two eyes are also calculated and presented in a different command window.

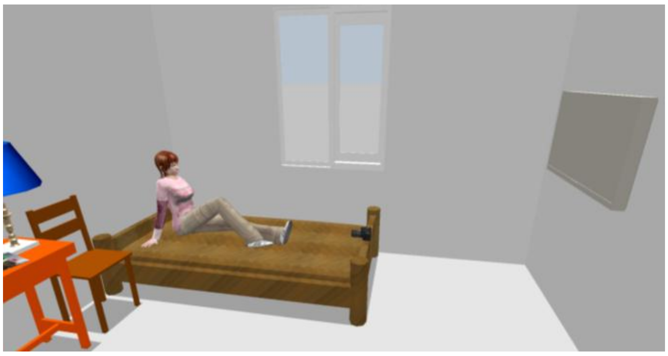


Figure 2. Experimental set-up showing the patient laying half-way on the bed, camera placed 5 feet away at a convenient location, and the display screen placed on the wall facing the patient

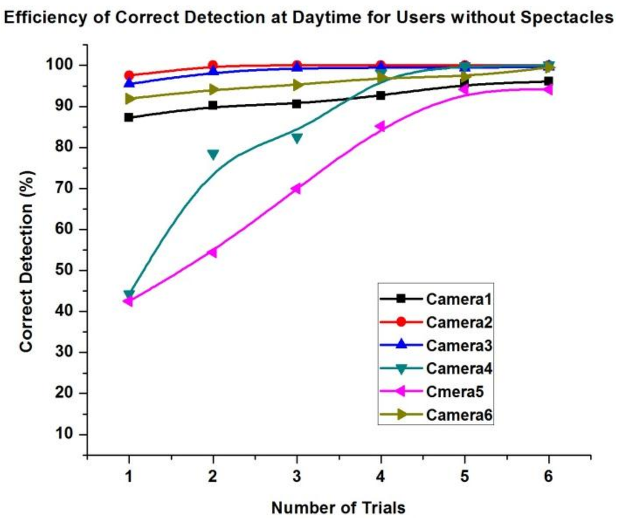
The main application of the system will be for providing support and rehabilitation to locked-in patients who have undergone spinal surgery or have been affected by stroke or spinal injury. The proposed system has to work for patients living at home or in hospitals or physiotherapy clinics. The experiments have been carried out in a 12-foot by 12-feet room similar to that in Ray Sarkar et al. (2015). The view of the experimentation room is shown in Figure 2. The user rests on a bed in the middle of the room in an inclined position. The camera is fixed 5 feet away from the patient. The monitor is mounted on the wall facing the patient directly. The video data can be stored on the computer and later processed for further analysis. Several users as mentioned in Table 2 have been engaged to gather the data for different illuminations (day and night) and conditions. The minimum value of the illumination was kept close to 4 Lux as specified in IS: 4347 – 1967 (Bureau of Indian Standards, 2005) using low-power lights. Some of the users wore spectacles and some did not. The online video data for the window with the detection signs are stored on the computer using third-party software.

Table 2. Profile of users engaged in the experiment

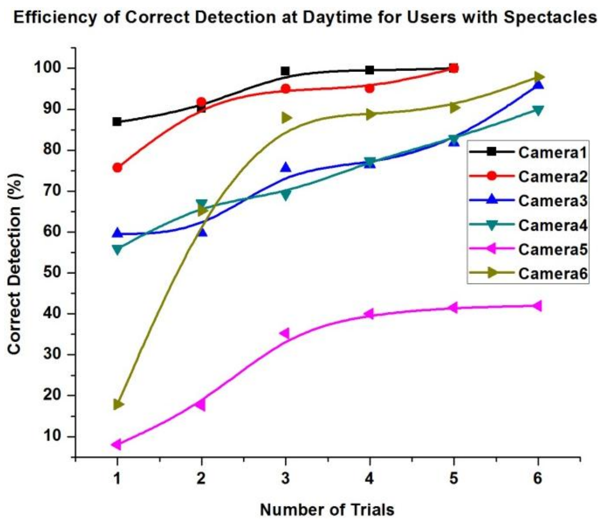
Age Group	30 - 40	40 - 50	50 - 60	60 - 70	70 - 80
No. of Users	2	-	2	2	1
Disabilities	No	-	No	No	Yes

RESULTS AND DISCUSSION

The data stored on the computer were analyzed to evaluate the performance of different low-cost cameras. The total number of frames for different cameras for different lighting conditions (day and night) with users with and without spectacles has been recorded for this purpose. Approximately 1, 400 numbers of frames have been analyzed from the video data for each and every test condition.

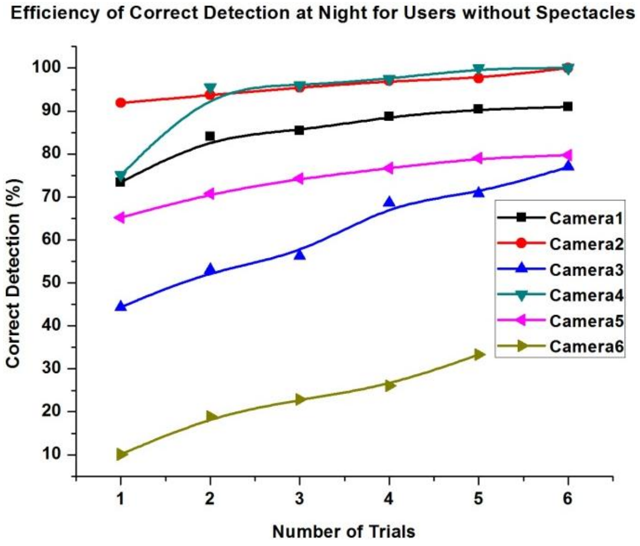


(a)

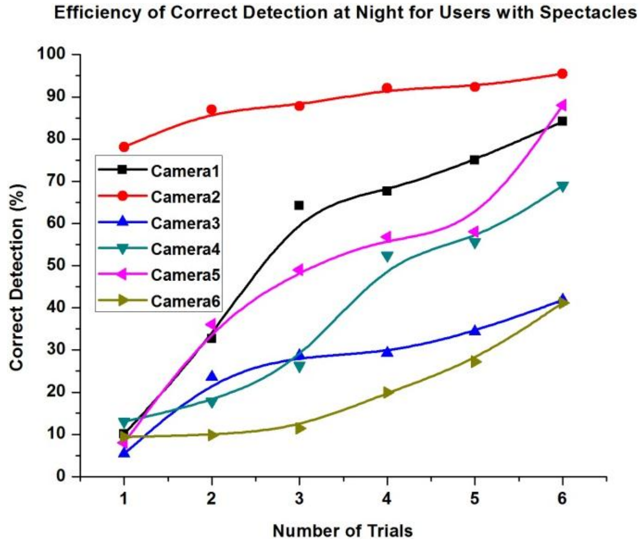


(b)

International Journal of Applied Engineering Research ISSN 0973-4562 Volume 12, Number 9 (2017) pp. 1819-1825



(c)



(d)

Figure 3. Efficiency of detection for different cameras at (a) daytime for users without spectacles, (b) daytime for users with spectacles, (c) night for users without spectacles, and (d) night for users with spectacles

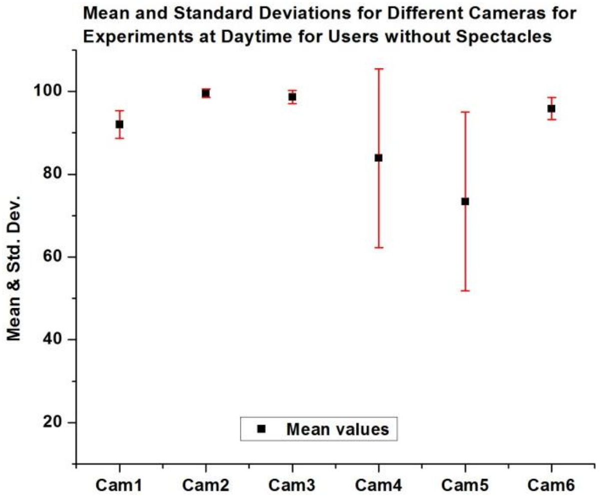
The average illumination during the daytime was measured to be approximately 78 Lux and 5 Lux at night. Detection of both eye centers rightly has been considered as the correct detection. The percentage of correct detection is the total number of such correctly detected frames against the total number of frames captured. The percentage of correct detection for different cameras during the day and at night for users with and without spectacles has been calculated and presented in the graphs in Figure 3(a) to (d).

Figure 3(a) depicts the percentage of correct detection during the day for users without spectacles using different cameras. Four cameras (i.e., camera1, camera2, camera3, and camera6) have a very close range of detection for such condition, and all are above 90%. The other two cameras (i.e., camera 4 and camera5) have the average percentage of detection as 84 and 73, respectively. Also, the standard deviations for these two cameras are very high at around 20. This is presented in Figure 4(a).

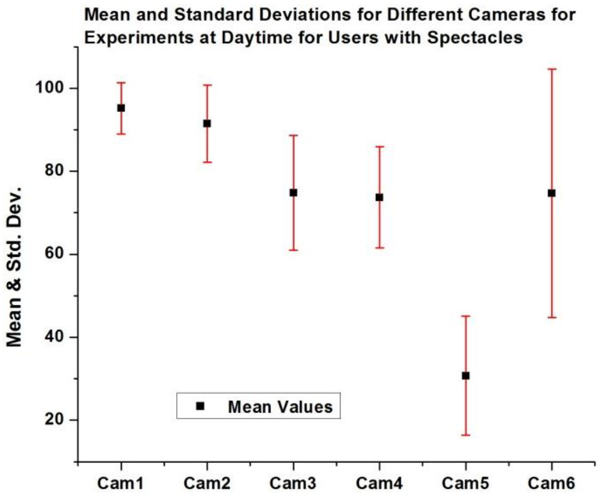
Similarly, the percentages of detection for users with spectacles during the day are presented in Figure 3(b). Two cameras (i.e., camera1 and camera2) have mean values of detection above 90%, and the other three cameras (except camera5) have an average percentage of detection above 73. Camera5 has an average percentage of detection as approximately 31. Here the range of standard deviations is also higher than the earlier one, as shown in Figure 4(b). Except for the first two cameras, the values of standard deviation lie between 12 to 30. For the first two cameras, these values are low and below 10.

The average percentages of detection for the cases at night for users without spectacles as presented in Figure 3(c) and are widely spread from 20 to 95. Camera2 and camera4 have average values of detection around 95%. At about 22%, camera6 has the lowest mean value. As shown in Figure 4(c), the standard deviations for all the cameras, except camera3, lie below 10. For camera3, the standard deviation is around 12.

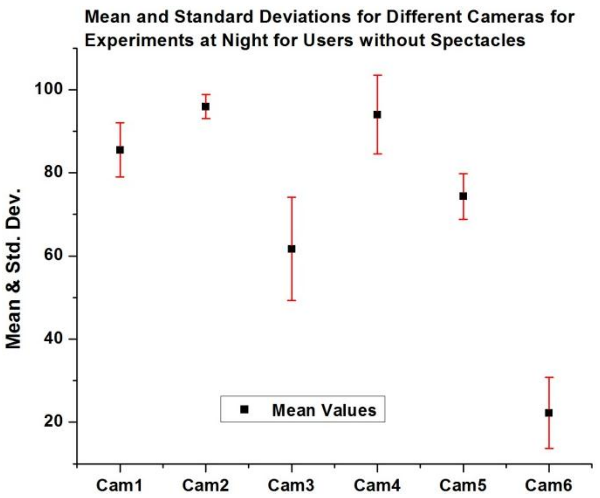
The most critical situation for detection of eye centers is at night, with very low illumination of around 5 Lux for users with spectacles. It can be noted from Figure 3(d) that, except for camera2, the overall range of detection is very low, around 50%. But for camera2, it is nearly 88%. Obviously, the standard deviations for this case are widely spread from a range of approximately 12 to 28. However, for camera2, it is around 6 due to its higher resolution and low-light compatibility. Camera3 shows poor performance at night, although it has higher resolution than camera2. This is because camera3 uses white LED lights in low light, which creates reflection on the spectacles, making it difficult for detection. Camera6 also shows poor performance for the experiments at night.



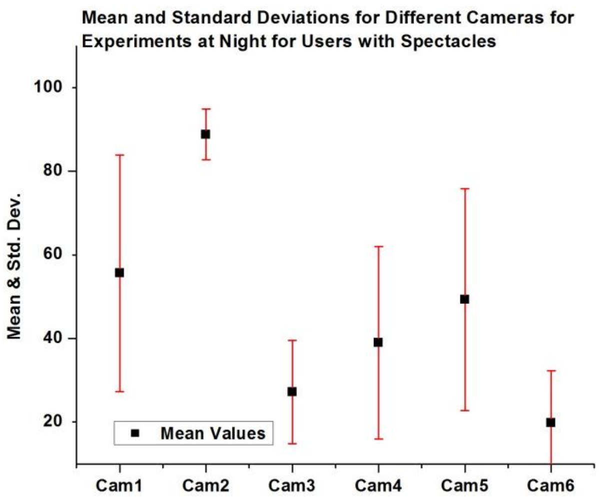
(a)



(b)



(c)



(d)

Figure 4. Mean and standard deviation of the efficiency values for correct detection for different cameras at (a) daytime for users without spectacles, (b) daytime for users with spectacles, (c) night for users without spectacles, and (d) night for users with spectacles

However, the present study reveals that the mean values of detection for users with and without spectacles during the day, and for users with and without spectacles at night are approximately 91%, 73%, 72%, and 47%, respectively. These average values include the highest values of correct detection as 100% and the lowest value as 10%. This reveals that the eye-tracking system has the capability to detect eye centers 100% correctly if the proper lighting and camera are used. This means that the bottleneck is the environmental illumination and the quality of the image or camera. The toughest situation is at night for users with spectacles. The lowest mean value of correct detection for a good camera at night for users with spectacles is around 89%. The average values of standard deviation, except for the case of users with spectacles at night, also lie within 15. For cases with spectacles at night, standard deviation is above 15. However, for the overall situation, the values are very closely repetitive in nature and lie within a narrow band.

CONCLUSION

Successful eye tracking has huge potential for application in different domains of human civilization. Although commercial eye trackers have been available in the market for the last 20 years, they failed to become popular due to higher cost, operational complexity, and difficulty of obtaining spares. As a result, research on the development of low-cost eye-tracking systems has become popular using easily available and affordable webcams and, analog and digital cameras.

This research found that selection of a good camera with capability of operation in very low illumination would make eye tracking easier and more efficient. Use of normal lighting is always preferred to special infrared or mono-chromatic lighting.

The present study evaluates the performance of six low-cost cameras (with prices below US$100) in different normal illuminations (daytime: 78 Lux; night: 5 Lux) with users with and without spectacles. Although an earlier work used the Logitech C615 camera for eye tracking, here the performance of Logitech C920 has been found superior in comparison with other models. The efficiency of detection for this camera in extremely difficult situations (i.e., at night with users with spectacles) is approximately 90% (mean value: 88.78%), and at normal illumination during daytime, it is 100% (mean value: 99.59%). This camera can now be used to develop a vision-based, low-cost rehabilitation system for 24/7 support to stroke and other locked-in patients. Work is in progress to define the mapping function between the eye center detected by the camera and the screen placed in front of the patient displaying useful icons of daily basic needs.

REFERENCES

[1] Agustin, J. S., Skovsgaard, H., Mollenbach, E., Barret, M., Tall, M., Hansen, D. W., & Hansen, J. P. (2010, March). Evaluation of a low-cost open-source gaze tracker. In *Proceeding of the 2010 Symposium on Eye-Tracking Research & Applications* (pp. 77 80). Austin, Texas.

[2] Ahlstrom, C., & Dukic, T. (2010, August). Comparison of eye tracking systems with one and three cameras. In *Proceedings of Measuring Behavior* (pp. 196- 199), Eindhoven, Netherlands.

[3] Bureau of Indian Standards. (2005). *Indian standard: Code of practice for hospital lighting (IS: 4347–1967)* New Delhi, India.

[4] Castellini, C., & Sandini, G. (2006, September). Gaze tracking for robotic control in intelligent teleoperation and prosthetics. In *Proceedings of the 2nd Conference on Communication by Gaze Interaction: Gazing Into the Future* (pp. 73–77), Turin, Italy.

[5] Chen, D., Bai, J., & Qu, Z. (2012, November). Research on pupil center location based on improved Hough transform and edge gradient algorithm. In *Proceedings of the National Conference on Information Technology and Computer Science* (pp. 47 -51), Lanzhou, China.

[6] Cheng, D., & Vertegaal, R. (2004, March). An eye for an eye: A performance evaluation comparison of the LC technologies and Tobii eye trackers. In *Proceedings of the 2004 Symposium on Eye Tracking Research & Applications* (pp. 61),. San Antonio, Texas.

[7] Ciger, J., Herbelin, B., & Thalmann, D. (2004, December). Evaluation of gaze tracking technology for social interaction in virtual environments. In *Proceedings of the 2nd Workshop on Modeling and Motion Capture Techniques for Virtual Environments* (pp. 1 6). Zermatt, Switzerland.

[8] Ferhat, O. (2012). *Eye-tracking with webcam-based setups: Implementation of a real-time system and an analysis of factors affecting performance* (Master's thesis). Universitat Autònoma de Bercelona, Spain.

[9] Freund, Y., & Schapire, R. E. (1997). A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences*, *55*(1),. 119 – 139.

[10] Funke, G., Greenlee, E., Carter, M., Dukes, A., Brown, R., & Menke, L. (2016, September). Which eye tracker is right for your research? Performance evaluation of several cost variant eye trackers. In *Proceedings of Human Factors and Ergonomics Society 2016 Annual Meeting* (pp. 1239–1243), Washington DC.

[11] Hansen, D. W., & Ji, Q. (2010). In the eye of the beholder: A survey of models for eyes and gaze. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *32*(3), 478 - 500.

[12] Janthanasub, V., & Meesad, P. (2015). Evaluation of a low-cost eye tracking system for computer input. *KMUTNB International Journal of Applied Science and Technology*, *8*(3), 1 – 12.

[13] Mantiuk, R., Kowalik, M., Nowosielski, A., & Bazyluk, B. (2012, January). Do-it-yourself eye tracker: Low-cost pupil-based eye tracker for computer graphics applications. In *Proceedings of the 18th International Conference on Multimedia Modeling* (pp. 115 - 125), Klagenfurt, Austria.

[14] Nevalainen, S., & Sajaniemi, J. (2004, April). Comparison of three eye tracking devices in psychology of programming research. In *Proceedings of the 16th Workshop of the Psychology of Programming Interest Group* (pp. 1 -8). Carlow, Ireland.

[15] Pfeiffer, T., Latoschik, M. E., & Wachsmuth, I. (2008). An evaluation of binocular eye trackers and algorithms for 3D gaze interaction in virtual reality environments. *Journal of Virtual Reality and Broadcasting*, *5*(16), 1-14.

[16] Ray Sarkar, A., Sanyal, G., & Majumder, S. (2015). Methodology for a low-cost vision-based rehabilitation system for stroke patients. In S. Gupta, S. Bag, K. Ganguly, I. Sarkar, & P. Biswas (Eds.), *Advancements of medical electronics: Lecture notes in bioengineering* (pp. 365- 377). Kalyani, India. Springer

[17] Sibert, L. E., & Jacob, R. J. K. (2000, April). Evaluation of eye gaze interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 281- 288)., The Hague, Netherlands.

[18] Soltany, M., Zadeh, S. T., & Pourreza, H.-R. (2011, April). Fast and accurate pupil positioning algorithm using circular Hough transform and gray projection. In *Proceedings of the 2011 International Conference on Computer Communication and Management, Vol. 5* (pp. 556- 561), Singapore.

[19] Viola, P., & Jones, M. (2001, December). Rapid object detection using a boosted cascade of simple features. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Vol. I* (pp. 511 -518),. Kauai, Hawaii.

[20] Zhang, X., & MacKenzie, I. S. (2007, July). Evaluating eye tracking with ISO 9241: Part 9. In *Proceedings of the 12th International Conference on Human-Computer Interaction: Intelligent Multimodal Interaction Environments* (pp. 779- 788)., Beijing, China.