Webcam-Based Eye Interactive System for Locked-In Syndrome Patients*

[1]E. Effah and [1]J. A. Addo
[1]University of Mines and Technology, P.O. Box 237, Tarkwa, Ghana

Effah E. and Addo J. A., (2016), "Webcam-Based Eye Interactive System for Locked-In Syndrome Patients", *Ghana Journal of Technology,* Vol. 1, No. 1, pp. 32 - 39.

Abstract

Locked-In Syndrome (LIS) patients cannot move nor talk but in most cases have the ability to blink and move their eyes. This project harnesses the blinking abilities of LIS patients from a live video stream as a way of interacting with a computer system. Farneback's optical flow image processing algorithm is applied to the frames of the video stream to read blinks. Face detection technique is employed to increase the probability of detecting the eyes in a frame. The various input gestures (short and long blinks) are used to trigger corresponding events to execute a particular action. The resulting system can be used in the home and in any favorable environment where there is enough light. The system will help LIS patients to communicate, improve the quality of life of LIS patients, and immensely widen the area of Human-Computer Interaction (HCI). The mammoth novelty of this paper is the fact that our system was deployed using the ubiquitous Webcam accessory of the Computer.

Keywords: Locked-In Syndrome, Live Video Stream, Farneback's Algorithm, Face Detection, Blink

1 Introduction

National Institute of Neurological Disorders and Stroke (NINDS) in Ghana defines locked-in syndrome (LIS) as a neurological disorder characterized by complete paralysis of voluntary muscles in all parts of the body except for those that control eye movement. It may result from traumatic brain injury (normally through accidents), diseases of the circulatory system (strokes), diseases of the nervous system (Amyotrophic Lateral Sclerosis), or medication overdose. Individuals with LIS are conscious and can think and reason, but are unable to speak or move. The disorder leaves individuals completely mute and paralyzed. LIS patients are deprived of speech, writing and the basic body movements and hence cannot communicate.

Mortality is high in the early stages of LIS (acute LIS) at 87 % within the first 4 months for LIS of vascular origin, early rehabilitation and more effective nursing care have been reported to reduce mortality (Patterson and Grabois, 1986). A retrospective study of in-patients with stroke (a cause of LIS in Ghana) admitted to the Komfo Anokye Teaching Hospital (KATH), Kumasi, from January 2006 to December 2007 showed that stroke constitutes 9.1 % of total medical adult admissions and 13.2 % of all medical adult deaths within the period under review. The stroke case fatality rate is 5.7 % at 24 hours, 32.7 % at 7 days, and 43.2 % at 28 days (Agyemang *et al.*, 2012). The percentage of mortality rate of stroke patients can be reduced (Patterson and Grabois, 1986) with LIS interactive systems but the cost of acquisition, installation and running discourage clients.

This paper presents a design and implementation of an affordable webcam-based interactive system that helps LIS patients to conveniently communicate. Because this system uses webcam, it is deployable domestically to help improve the quality of life of LIS patients. Application of this system is envisaged to reduce mortality rate of LIS patients in Ghana. However, the implementation of this system is limited to conscious and aware patients; patients with blinking capabilities; patients without eye defects (no spectacles); and patients who can read (English).

1.1 Communication and LIS Patients

Communication is the act or process of using words, sounds, signs, or behaviors to express or exchange information or to express ideas, thoughts, feelings, etc., to someone else. Verbal or Oral languages, written words and sign languages are by far the standard ways of communication. Computers now bridge the communication between two parties at different geographical locations and with language differences. Even, people with disabilities now communicate with the help of computers.

LIS patients lack all the abilities to engage in the standard ways of communication. As such, computer-based communication systems are built to help them communicate and control their environment, use a word processor coupled to a

*Manuscript received August 23, 2016
Revised version accepted September 16, 2016

speech synthesizer, and access the worldwide net (Anon., 2007b). Some research works in Information Communication Technology (ICT) and the medical field have significantly improved the quality of lives of LIS patients. Existing eye-controlled, computer-based communication technology currently allows the patient to control his environment, use a word processor-coupled to a speech synthesizer, and access the worldwide net.

1.2 ICT, Human-Computer Interaction (HCI) and Health

ICT is an enabling technology which provides various solutions in the healthcare sector, ranging from electronic patient records and health information networks to intelligent prosthetics and robotized surgery (Butter *et al.*, 2008). Artificial Intelligence (AI), HCI and Robotics for healthcare is an emerging field, which is expected to grow in the face of demographic change, expected shortages of healthcare personnel, calls for improving quality of life for the elderly and disabled, and the need for even higher quality care, for example high precision surgery (Butter *et al.*, 2008). All these factors stimulate innovation in the domain of ICT for Healthcare.

Ageing and its concomitant chronic diseases is a great challenge for national development and the public healthcare systems (Anon., 2010). Among the areas of research interest for the aged is the HCI. HCI has expanded rapidly and steadily for three decades, attracting professionals from many other disciplines and incorporating diverse concepts and approaches. To a considerable extent, HCI now aggregates a collection of semi-autonomous fields of research and practice in human-centered informatics. However, the continuing synthesis of disparate conceptions and approaches to science and practice in HCI has produced a dramatic example of how different paradigms can be reconciled and integrated in a vibrant and productive intellectual project (Soegaard and Dam, 2002). HCI is not limited to developing user interfaces but also developing effective and natural ways of interacting with computers.

1.3 Literature Evaluation of LIS Systems

1.3.1 Dasher Keyboard

Dasher, now managed by the Inference group, was invented by David J. C. Mackay in 1997 and developed by David Ward in 1998. Dasher is an information-efficient text-entry interface, driven by natural continuous pointing gestures. Dasher is a competitive text-entry system wherever a full-size keyboard cannot be used. Typical instances are when operating a computer one-handed, by joystick, touchscreen, trackball, or mouse; when operating a computer with zero hands (i.e., by head-mouse or by eye tracker).

The eye tracking version of Dasher allows an experienced user to write text as fast as normal handwriting - 29 words per minute; using a mouse; experienced users can write at 39 words per minute. Dasher used with an Eye Tracker is a very efficient tool for communication for LIS patients. Fig. 1 demonstrates how after typing 'techn' the program makes it far easier to end the word with 'ology'.

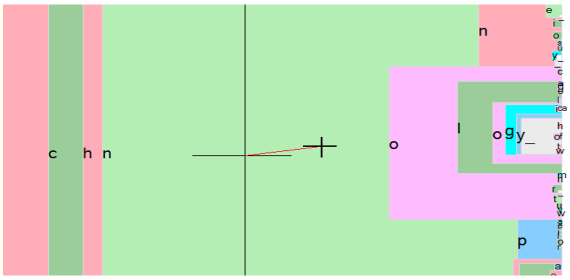

Fig. 1 Dasher Keyboard Interface (Source: Anon, 2007a)

1.3.2 Optikey

Optikey, developed by Julius Sweetland is an assistive on-screen keyboard designed to bring mouse, keyboard, and speech control to LIS patients. It is designed to be used with an eye-tracking device that is low cost in order to bring mouse control, keyboard control, and speech. Optikey works with an eye-tracker device. The software works right out of the box once a person has their eye-tracking device installed.

Optikey can replace your mouse and allow you to click, scroll and drag with exactitude anywhere on your screen. The program's basic mouse actions include basic mouse actions such as left, middle, right click or scrolling. Optikey allows input into applications like Microsoft (MS) Office Word. Optikey can be used without an eye-tracker. Example is shown in Fig. 2.

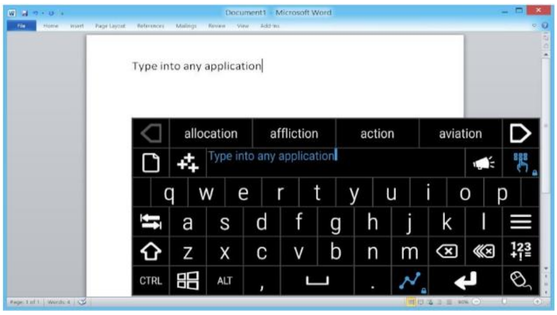

Fig. 2 Optikey as Input to MS Word (Source: Weiss, 2015)

1.3.3 Eye Tracker Device (ETD)

An eye-tracker Fig. 3 is a device for measuring eye positions and eye movement. Eye trackers are used in research on the visual system, in medicine, marketing and in product design. Eye trackers are powerful enough to track the pupil. Eye trackers used with Dasher and Optikey are a very powerful tool for communication for the disabled.

Fig. 3 Head Mounted Eye Tracker Device (Source: Anon., 2005)

1.3.4 LIS Systems Review Conclusion

The technologies discussed above are very expensive to come by. The cheapest among them is Optikey, which is free but is used with extremely expensive hardware like Eye Tracker Device. This paper in its novelty developed a very cheap system, domestically deployable and also based on the Webcams on personal computers. This system is affordable to implement because it uses inexpensive tools and also requires cheap accessories to use. A few more patients could live a better life with this system.

1.4 Image Processing

Image processing is using mathematical operations to process any form of signal for which the input is an image (Gonzalez, 2008). The output of image processing may be either an image or a set of characteristics or parameters related to the image. Most image-processing techniques involve treating the image as a two-dimensional signal and applying standard signal-processing techniques to it (Miller, 2012). Images are also processed as three-dimensional signals where the third-dimension, mostly being time or on a few occasions the z-axis.
Digital image processing is the use of computer algorithms to perform image processing on digital images (Sarfraz, 2013). Digital image processing allows the use of much more complex algorithms, and hence, can offer both more sophisticated performance at simple tasks, and the implementation of methods, which would be impossible by analog means (Sarfraz, 2013). In particular, digital image processing is the only practical technology for classification, feature extraction, pattern recognition and projection.

Some techniques which are used in digital image processing include: pixilation; linear filtering; image editing; image restoration; principal component analysis; independent component analysis; hidden Markov models; anisotropic diffusion; partial differential equations; self-organizing maps; neural networks and wavelets (Tabassum, 2012).

1.4.1 OpenCV

OpenCV is an open source image processing library written in C++, and designed for computational efficiency and real-time applications. Written in optimized C/C++, the library takes advantage of multi-core processing and the hardware acceleration of the underlying heterogeneous compute platform.

OpenCV's application areas include: 2D and 3D feature toolkits; Egomotion estimation; Facial recognition system; Gesture recognition; HCI; Mobile robotics; Motion understanding; Object identification; Segmentation and recognition; Stereopsis stereo vision: depth perception from 2 cameras; Structure from motion (SFM); Motion tracking; Augmented reality.

OpenCV contains a statistical machine learning library that equips its applications with these tools: Boosting; Decision tree learning; Gradient boosting trees; Expectation-maximization algorithm; k-nearest neighbor algorithm; Naive Bayes classifier; Artificial neural networks; Random forest; Support vector machine (SVM)

1.4.2 Matrix (Mat) type

Mat is the primary data structure in OpenCV. It stores images and their components. It is a two-dimensional array of pixels. A pixel is the smallest unit of an image. A pixel is represented in OpenCV as channels which is an array. A 1-channeled pixel for grayscale images, 3 for colored RGB images and 4 for Colored RGBa images. The channels simply represent array size. A Mat type is therefore a two-dimensional array of arrays of integer or float type. Fig. 4 represents an abstract memory structure of a 4-channeled image.

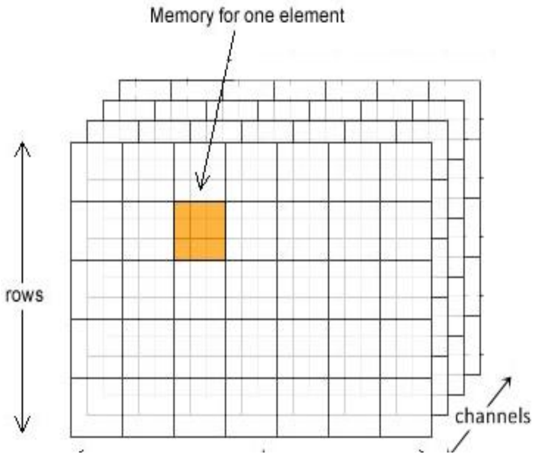

Fig. 4 Structure of a (5 rows × 7 columns) x 4 Channels Mat type.

1.4.3 Cascade Classifier for Object Detection

A classifier is a trained file with a few hundred sample views of a particular object (i.e., a face or a car), called positive examples, that are scaled to the same size (say, 20×20), and negative examples - arbitrary images of the same size (Anon., 2014). Classifiers are applied to a region of interest (of the same size as used during the training) in an input image. The classifier outputs a "1" if the region is likely to show the object and "0" otherwise. To search for the object in the whole image, the search window can be moved across the image to check every location using the classifier. The classifier is designed so that it can be easily "resized" in order to be able to find the objects of interest at different sizes, which is more efficient than resizing the image itself. So, to find an object of an unknown size in the image the scan procedure should be done several times at different scales. The basic classifiers are XML decision-tree classifiers with at least 2 leaves.

1.4.4 Optical Flow

Optical flow or optic flow is the pattern of apparent motion of objects, surfaces, and edges in a visual scene caused by the relative motion between an observer (an eye or a camera) and the scene. It is 2D vector field where each vector is a displacement vector showing the movement of points from first frame to second. OpenCV provides two methods for calculating optical flow: Lucas-Kanade's algorithm for tracking corner-like features and tracking them to compute the optical flow; Farneback algorithm for calculating dense optical flow of the whole image. Optical flow has application areas like video compression and stabilization, motion detection and kinematic analysis.

2 Resources and Methods Used

2.1 OpenCV Environment

This system is developed on the MS Windows operating system and as such, OpenCV binaries, built with Visual C++ compilers, is used to implement the project. The environment variable for OpenCV, is correctly setup and configured appropriately.

2.2 System Requirements

This system is a desktop application that serves as a communication tool for LIS patients. The application takes inputs such as eye blinks using a webcam. The application presents a list of requests and statements in order of Maslow's hierarchy of needs. A user scrolls through the list and selects one at a time. The system skips involuntary blinks. Naturally, involuntary blinks are very short therefore the short input blink should hold a little longer.

2.3 Use Case

From the user's point of view (Fig. 5), we have one user of the system: LIS patient. There are two use cases and are listed below:

(i) User Scroll through a list of request or statement using short blinks. The system beeps once to alert user of the end of a short blink;

(ii) User selects an item from list using long blinks. The system beeps twice for the second time to indicate a long blink.

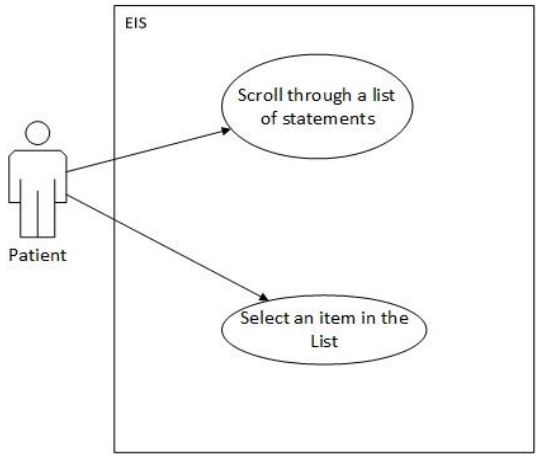

Fig. 5 Use Case Diagram

2.4 Functional Components

The system has two main components: Blink detection component, responsible for detecting eye blinks, and the Requests component for managing requests.

2.4.1 Blink Detection Component

The blink detection component is the core of the system. It is the part of the system that reads and processes image data captured from webcam to detect blinks and the type of blink (short or long). It takes advantage of OpenCV's optical flow algorithm and Farneback's algorithm to detect eye movement (blinks). This algorithm has lower time complexities even in the worst case. It includes other sub-components as the face and eye detectors. It also has information about the state of the eye (closed or opened state). It also starts a timer whenever the eye is closed and stops when opened. The timer classifies the blink as short or long. Figs. 6 and 7 present the flowcharts for blink detection and request components respectively whilst Fig. 8 shows a detailed flow chart of the our system as proposed.

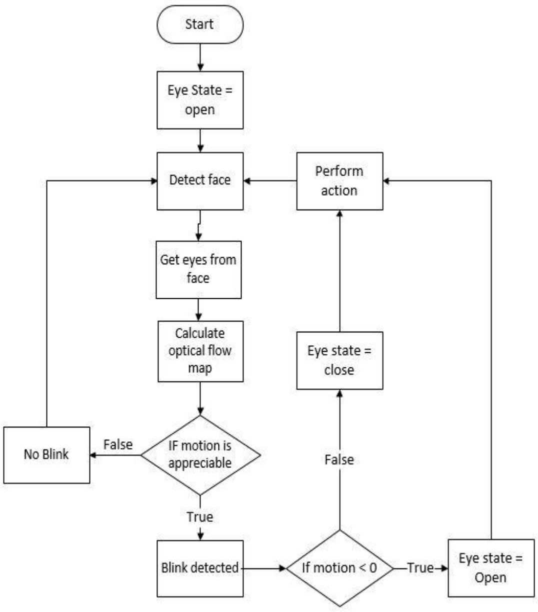

Fig. 6 Blink Detection Component

2.4.2 Requests Component

The requests component stores the list of predefined statements. It also holds the index of the selected request in the list. The requests component knows nothing about the blink detector component. It only receives index changes from the blink detector. It is also responsible for drawing request buttons on the frame. The "ActiveRequest" value is received from the blink detector component. Fig. 8 shows a flow chart of the request component. The list is organized in order of Maslow's theory of needs in order to ensure easier navigation through.

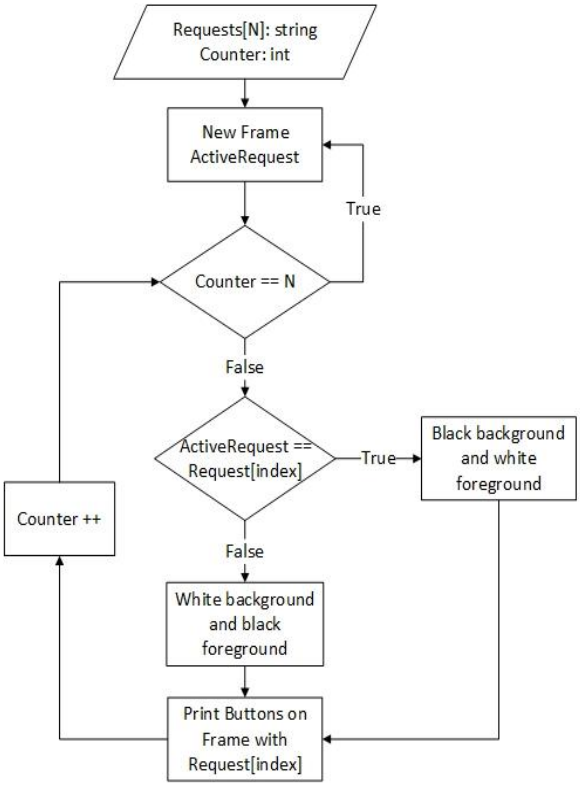

Fig. 7 Request Component

3 Results and Discussion

3.1 The System and its Operations

The system has only three basic request for testing purposes. The items in the list are: Hungry; Thirsty; Washroom. The items are organized in order of Maslow's hierarchy of needs. When the application is started, the webcam and screen of computer is adjusted such that the user's face is captured in the center of the application window. The user should also be able to view the face and the list on the screen clearly.

3.1.1 Scrolling Through the List

The user closes the eyes and waits for a single beep sound at time t = 0.60 s. The timer starts immediately the eyes are closed. The user should then open the eyes before t = 2.00 s. The selection then changes to the next on the list. Figs. 9 and 10 show the user's action to scroll from "Hungry" to "Thirsty". Fig. 10 clearly shows that "Thirsty" is now highlighted. The red circles in the top corners in Fig. 9 indicates that the eyes are closed. The circles are green if the eyes are open as in Fig. 10.

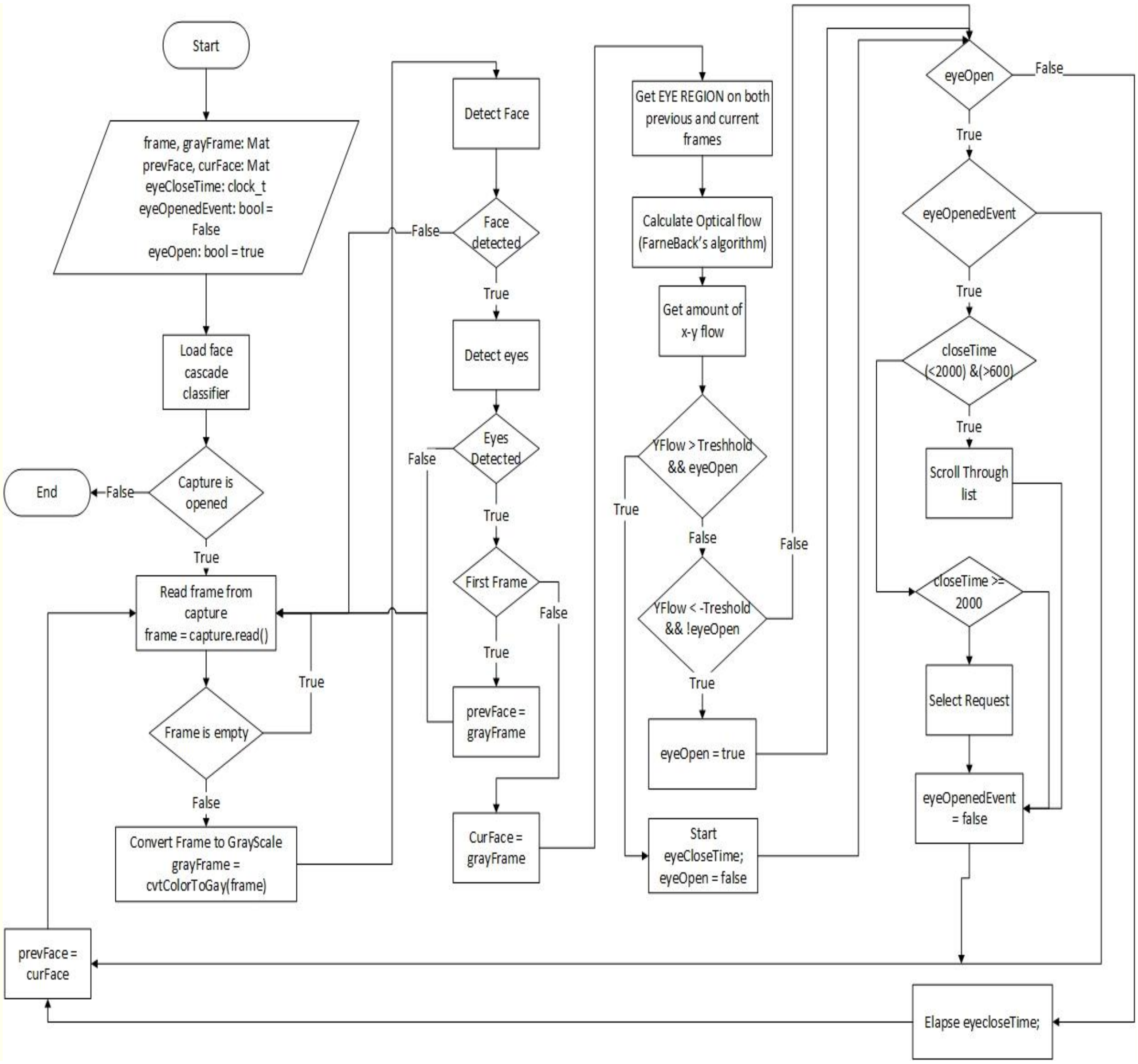

Fig. 8 Detailed System Flow Chart

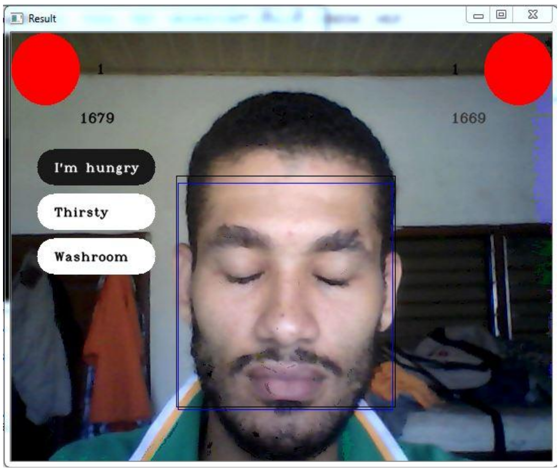

Fig. 9 User Closes Eyes for 0.6s < t < 2.0s

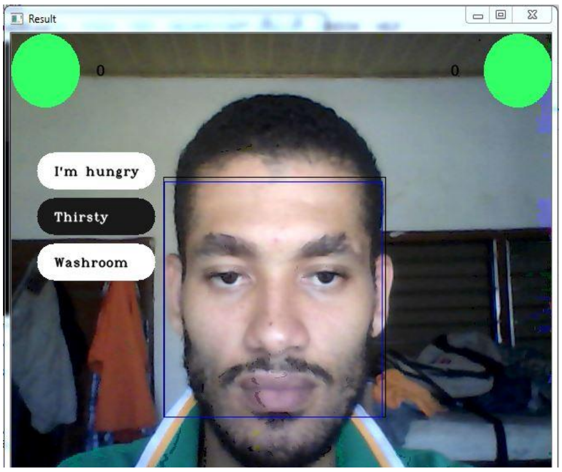

Fig. 10 User Opens Eyes Between 0.6s< t < 2.0s

3.1.2 Selecting an Item

The user keeps the eyes closed after the first beep at t = 0.60s till a second double beep is made at t = 2.00s. Once the eyes are opened after t = 2.00s the highlighted item is selected and printed on the console. Fig. 11 shows the output of the selection.

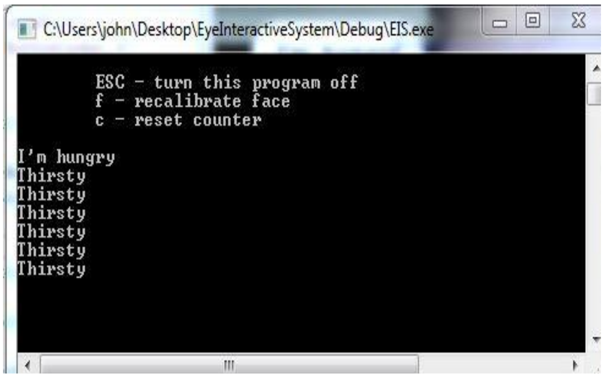

Fig. 11 The Output to the Console

4 Conclusions and Recommendations

In conclusion, this system will help LIS patients to communicate. The quality of life of LIS patients can be improved. The life expectancy of LIS patients can also be increased. The present system was deployed on 4 gigabytes INTEL processor which processes graphics slowly. This system is highly recommended to be implemented on Nvidia cuda graphic processors for faster processing since Nvidia processors are faster even with algorithms with higher time complexities. Further studies can be done to add more features and input gestures to the system. It would be much easy to use if requests are categorized than a tall list of requests displayed.

References

Agyemang, C., Attah-Adjepong, G., Owusu-Dabo, E., De-Graft A. A., Addo, J., Edusei, A. K., Nkum, B. C., Ogedegbe, G., (2012), "Stroke in Ashanti region of Ghana", *Ghana Medical Journal*, Vol. 46, pp. 12 -17.

Anon., (2005), "Eye tracking on the ISS", *https://en.wikipedia.org/wiki/Eye_tracking_on_the_ISS*, Assessed: March 22, 2016.

Anon., (2007a), "Dasher", Better Living through Technology, *http://www.bltt.org/software/dasher/*, Assessed: March 20, 2016.

Anon., (2007b), "Locked-In Syndrome", *http://www.ninds.nih.gov/disorders/lockedinsyndrome/lockedinsyndrome.htm*, Assessed: February 18, 2016.

Anon., (2010), "Why ICT for Health", *http://www.ictforhealth.net/index.php?option=comcontent&view=article&id=45&Itemid=27*, Assessed: March 17, 2016.

Anon. (2014), "Haar Feature-based Cascade Classification for Object Detection", *http://www.docs.opencv.org/2.4/modules/objdetect/doc/cascade_classifier.html*, Assessed: March 22, 2016.

Baumstarck, K., Alessandrini, M., Blandin, V., Villemeur, T. B. and Auquier, P., (2015), "Quality of life in patients with locked-in syndrome: Evolution over a 6-year period", *Orphanet Journal of Rare Diseases*, pp. 1-8.

Butter, M., Rensma, A., Boxsel, J. V., Kalisingh, S., Schoone, M., Leis, M., Gelderblom, G.J., Cremers, G., Wilt, M., Kortekaas, W., Thielmaan, A., Cuhls, K., Sachinopoulou, A., Korhonen, I., (2008), "Robotics for healthcare", Innovation for Life, Vol. 5, pp. 4 – 5.

Gonzalez, R. C. and Woods, R. E., (2008) "Digital Image Processing", *Prentice Hall*, pp. 1-3.

Miller, F. P., Vandome, A. F., McBrewster J., (2010), "Image Processing", *VDM Publishing*, 72 pp.

Lienhart, R. and Maydt, J., (2002), "An Extended Set of Haar-like Features for Rapid Object Detection", *Proceedings of the IEEE Transactions on ICIP*, Vol. 1, pp. 900-903.

Merriam-Webster Dictionary, "Communication" *http://www.merriamwebster.com/dictionary/communication*, Assessed: May 30, 2016.

Patterson, J. and R., Grabois, M. (1986), "Locked-in syndrome: a review of 139 cases", *American Heart Association*, Vol. 17, No. 4, pp. 758 – 763.

Sarfraz, M., (2013), "Intelligent Computer Vision and Image Processing", *Information Science Reference*, 701 E. Chocolate avenue, Hershey PA 17033, p. XV.

Soegaard, M., Dam and R. F., (2002), "The Encyclopaedia of Human-Computer Interaction", *https://www.interactiondesign.org/-literature/book/the-encyclopedia-of-human-computer-interaction-2nd-ed*, Assessed: March 10, 2016.

Tabassum S. N., (2012), "Digital Computing Image Processing Network Technology", *International Journal of Scientific and Engineering Research*, Vol. 3, pp. 3 – 4.

Weiss, T. C., (2015), "OptiKey is an assistive on-screen keyboard designed to bring mouse, keyboard, and speech control to people with speech and motor limitations", *http://www.-disabled-world.com/assistive devices/computer/optikey.php*, Assessed: 22nd March, 2016.

Authors

Emmanuel Effah is a Lecturer in the Computer Science and Engineering Department of the University of Mines and Technology. He holds the degrees of MSc (Information Technology) from the Open University of Malaysia and BSc. (Electrical and Electronic Engineering) from University of Mines and Technology (UMaT) Tarkwa, Ghana. His current research areas include Smart Systems, Data Mining and Digital Signal Processing.

John Afetsi Addo is currently a National Service Personnel in the University of Mines and Technology. He holds a Bachelor degree (BSc) in Computer Science and Engineering from the University of Mines and Technology (UMaT), Tarkwa, Ghana. His research interest is in Artificial Intelligence.