AN EYE-GAZE TRACKING AND HUMAN COMPUTER INTERFACE SYSTEM FOR PEOPLE WITH ALS AND OTHER LOCKED-IN DISEASES

Shuo Samuel Liu, Andrew Rawicz, Teng Ma, Cheng Zhang,
Kyle Lin, Siavash Rezaei, Eion Wu
Simon Fraser University

ABSTRACT

Eye tracking is one of the most important ways for people with ALS and other locked-in diseases to communicate. The majority of current eye tracking computer input systems are too expensive and not very user-friendly. This paper proposes an infrared sensor/emitter based eye tracking system called EyeLive, which does not use the common camera based approach. The hardware, eye tracking algorithm, and user interface are described, discussed, and compared with camera based eye tracking systems. The performance of the system is presented with experimental data. The advantages of the EyeLive system such as low cost, user friendliness, portability, and eye strain reduction are also discussed.

KEYWORDS

Eye gaze tracking, human-computer interface, communication, ALS

INTRODUCTION

Amyotrophic lateral sclerosis (ALS) is a neurodegenerative disease that leads to progressive paralysis of voluntary muscles. Patients eventually lose their ability to move or speak, but retain the ability to move their eyes [1]. Therefore, eye movement is the most natural way for late-stage ALS patients to communicate.

The eye-gaze tracking systems on the market and under research today are very expensive, costing \$5,000 to \$20,000. For example, the ERICA system costs USD\$7,000-8,000 [2], [3]. Most of these systems use a video camera to capture images of the eyes and the face, and thus tracking the movement of the pupil. Although cameras are relatively cheap today, the complex algorithms and the necessity of frequent calibration render the systems expensive and not very user-friendly.

The EyeLive system presented in this paper aims to significantly reduce the cost, increase the portability, and improve the user interface of an eye-gaze tracking and human computer interface system. Instead of the camera-based approach used by most of the literature and commercial systems, EyeLive uses infrared sensors installed on a pair of glasses to detect the direction of the eye gaze. This lightweight approach requires minimal hardware and computation, increases the mobility, and allows easy integration with other devices.

SYSTEM OVERVIEW

As shown in Figure 1, the EyeLive system consists of three major components: the infrared sensors/ emitters, the microcontroller circuit, and the computer software. The infrared sensors measure the eye reflections and send analog signals to the microcontroller, which does the A/D conversion and transfers the data to the computer through USB. The computer software uses algorithms to determine the direction that the eye is looking at, and controls the graphical user interface.

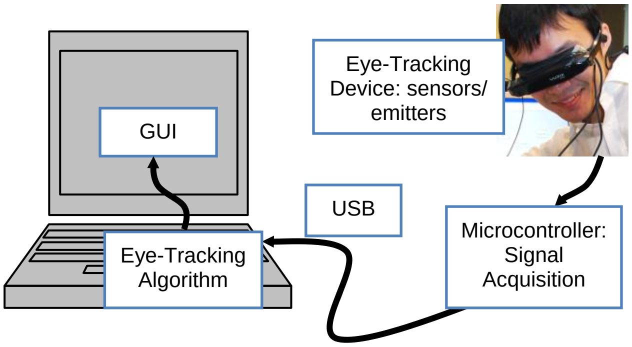

Figure 1: Overall design of the system

EYE TRACKING HARDWARE AND ALGORITHM

Hardware

Two other systems use similar designs to the EyeLive system. The EyeTouch system [4] in Figure 2-b uses 4 IrDA (e.g. remote control) sensor/emitter pairs mounted onto the glasses frame for both eyes. The Owl system [5] in Figure 2-c has 8 infrared phototransistor/LED pairs mounted around an eye. The EyeLive system (Figure 2-a), on the other hand, consists of 4 infrared LEDs (white) and 4 infrared phototransistors (black) mounted in front of the glasses for one eye. The phototransistors and LEDs are arranged with equal and symmetric spacing between each other to minimize the interference of an LED on adjacent phototransistors, and to maximize the scattering of light from an LED.

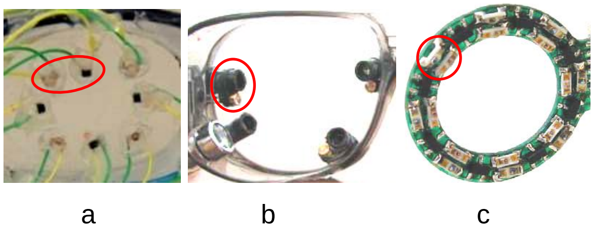

Figure 2: Comparison of discrete sensor/emitter based eye tracker designs

EyeLive reduces the number of sensors from 8 in the Owl system to only 4, because 4 sensors are sufficient to very accurately detect 4 directions and eye blink, and quite accurately detect 8 directions. The sensor/emitter hardware in front of the glasses is less bulky than both the EyeTouch and the Owl, which allows the user to see more portion of the screen and is more aesthetically pleasing.

Data acquisition and processing

The 4 LEDs are turned on one at a time, and the intensity of the reflected light is measured by the 4 phototransistors and converted to a 10-bit number. Therefore, each measurement has 16 data points acquired at 4608 Hz. After 12 measurements are collected, they pass through a mean filter to remove high frequency noise. Then a 16D vector representing the averaged measurement is sent to the computer via USB at 24 Hz.

Calibration

Before normal usage, the user must first look at 5 positions on the screen (Figure 3) and close the eye. The system collects information about the user's unique eye profile. For each direction, many measurements are taken during about 2 seconds.

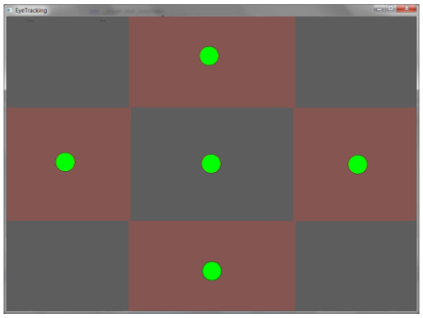

Figure 3: Calibration points

Then the collected measurements are passed through a combination of medium and mean filters, resulting in a 16D vector to be used as the identifier for that direction.

Algorithm

During normal usage each incoming measurement received by the computer is compared with the identifiers. A number of algorithms can be used to determine the direction, such as mean absolute error (MAE), mean square error (MSE), and principal component analysis (PCA). MAE is currently used because it is the simplest method to implement, it requires the least computation, and the result is as good as the other algorithms.

With some initial training, all the people who tested the system have been able to use it accurately without any difficulty. To numerically verify the accuracy of the algorithm, the user looks at 20 additional yellow spots on the screen, as shown in Figure 4.

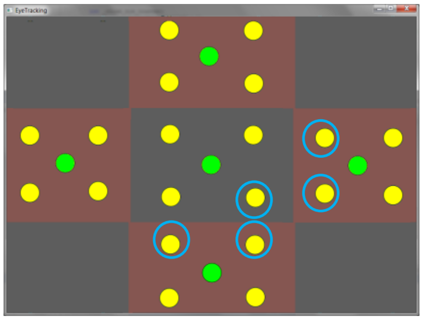

Figure 4: Additional points for verification

The yellow spots define the boundary within which the user is most likely to look in each direction. The measurement for each yellow spot is then classified using the MAE algorithm to see if the classification matches the intended direction.

40 trials were conducted on 4 individuals. The result shows 0% error rate in classifying eye close. The algorithm shows only 4% error rate in classifying the 5 directions (middle/up/down/left/right), with 0% error rate in half of the trails. The spots where the errors occur (Figure 4, in blue circles) are the transition spots between middle and one of the four other directions. One reason is that the user may not focus precisely onto the spot during the experiment, thus resulting in an error.

GRAPHICAL USER INTERFACE

EyeLive's graphical user interface contains a unique and innovative keyboard, as shown in Figure 5. The user selection method is similar to a Morse code.

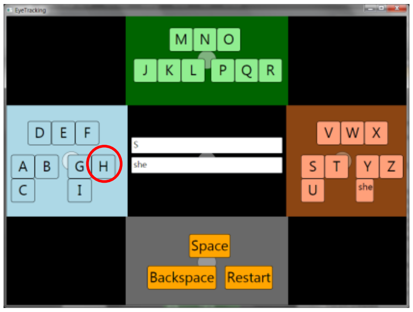

Figure 5: Keyboard, first step

The 26 alphabet letters are divided into groups. For example, if the user wants to type the letter "H", s/he first looks to the left, and select the *Left* grid by either closing or dwelling the eye for 1 second. Then the user repeats the process by selecting the *Right* grid (Figure 6), then the *Up* grid (Figure 7).

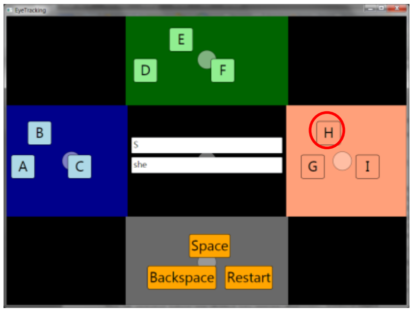

Figure 6: Keyboard, second step

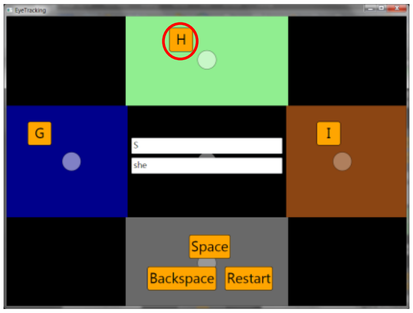

Figure 7: Keyboard, third step

Comparing with the camera-based systems with a full keyboard user interface, EyeLive's user interface has many advantages. Human eyes are much better at scanning around than focusing on a particular spot. Therefore, the eyes require a lot of effort to focus on a particular button on a full keyboard, but require very little effort to roughly look at one of 4 directions. As a result, EyeLive's keyboard reduces eye strain. With training, the eyes can build muscle memory and move to one of 4 directions very quickly. So although 3 selections are required to enter a letter on EyeLive instead of 1 selection per letter on a full keyboard, faster selection can partially offset this disadvantage.

EyeLive also has a word prediction feature to speed up typing, and a text-to-speech feature to make the user interface more interactive.

FURTHER IMPROVEMENT

Preliminary results show that EyeLive can detect up to 8 directions, but the error rate is higher than detecting 4 directions. Thus, further investigation can be done to develop more sophisticated variations of MAE, MSE, PCA and other algorithms. As a result, the user interface usually needs only two selections per letter instead of three, which increases the typing speed. Artificial intelligence techniques could also be incorporated to profile a user during long term in order to increase accuracy.

CONCLUSION

In this paper, an infrared sensor/emitter based eye tracking and computer interface system is described and implemented. The system's experimental performance in terms of accuracy is presented. The EyeLive system has comparable (although slower) typing speed than a camera based system, but it costs much less, reduces eye strain, increases accuracy, enhances mobility, and improves user-friendliness. Therefore, the EyeLive system is very competitive compared to both camera-based and other similar optical eye tracking devices.

ACKNOWLEDGEMENTS

We would like to thank other team members at Simon Fraser University who have contributed this project: Siyong Zhu implemented part of the software, and Priyanka Deshmukh implemented part of the hardware.

REFERENCES

[1] The ALS Association, Oct 2008, "What is ALS," http://www.alsa.org/als/what.cfm (Mar 4, 2010).

[2] Axistive, Jun 23, 2007, "ERICA Eye Tracking," http://www.axistive.com/erica-eye-tracking.html (Mar 4, 2010).

[3] Eye Response Technologies (now DynaVox Mayer-Johnson), 21 July 2009, "Eye Tracker Erica," http://www.cogain.org/wiki/Eye_Tracker_Erica (Mar 4, 2010).

[4] Cihan Topal, Ömer Nezih Gerek and Atakan Doğan, "A head-mounted sensor-based eye tracking device: eye touch system," 2008.

[5] Dale Grover, "Progress on an eye tracking system using multiple near-infrared emitter/detector pairs with special application to efficient eye-gaze communication," *Proceedings of COGAIN 2006: Gazing into the Future*, pp. 21-25, 2006.