An Eye-Gaze Tracking and Human Computer Interface System for People with ALS and Other Locked-in Diseases

Shuo Samuel Liu*    Andrew Rawicz    Siavash Rezaei    Teng Ma    Cheng Zhang
Kyle Lin    Eion Wu

School of Engineering Science, Simon Fraser University, Burnaby, BC V5A IS6, Canada

Received 3 Sep 2010; Accepted 17 Mar 2011; doi: 10.5405/jmbe.836

Abstract

Eye tracking is one of the most natural ways for people with amyotrophic lateral sclerosis and other locked-in and paralysis diseases to communicate. The majority of existing eye-tracking computer input systems use cameras to capture images of the user’s eyes to track pupil movements. Most camera-based systems are expensive and not user-friendly. This paper proposes an eye-tracking system called EyeLive that uses discrete infrared sensors and emitters as the input device. The eye-tracking and calibration algorithms classify eye positions into six categories, namely looking up, looking down, looking left, looking right, looking straight ahead (i.e. middle direction), and eyes closed. A graphical user interface optimized for EyeLive is also developed. It divides the screen into a nine-cell grid and uses a hierarchical selection approach for text input. EyeLive’s hardware, eye-tracking algorithm, calibration, and user interface are compared to those of existing eye-tracking systems. The advantages of the EyeLive system, such as low cost, user friendliness, and eye strain reduction, are discussed. The performance of the system in classifying eye positions is experimentally tested with eight healthy individuals. The results show a 5.6% error rate in classifying eye positions in five directions and a 0% error rate in classifying closed eyes. Additional experiments show that the average typing speed is 1.95 words/min for a novice user and 2.91 words/min for an experienced user. The tradeoffs of lower typing speed versus other advantages comparing to other systems are explained.

Keywords: Eye-gaze tracking, User interface, Communication, Amyotrophic lateral sclerosis (ALS), Locked in disease

1. 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.

Most existing eye-gaze tracking systems use a video camera to capture images of the eyes to track pupil movement. The systems are either desktop-mounted (i.e., the camera is mounted on a table or monitor), such as Tobii [2-6] and ERICA [7,8], or head-mounted, such as EyeSeeCam [9-11]. Although many of these systems provide high-precision eye tracking, they are very expensive due to high hardware and data processing requirements. In addition, the user interfaces [2-13] are not particularly user-friendly for ALS patients.

* Corresponding author: Shuo Samuel Liu
Tel: +1-778-316-9484
E-mail: samerlite@ieee.org

The present study proposed the EyeLive system, which uses infrared emitters and sensors installed on a pair of glasses to detect the direction of eye gazes. The EyeLive system uses simple eye-tracking algorithms and has a graphical user interface (GUI) that increases accessibility and reduces eye strain.

2. System description

2.1 System overview

As shown in Fig. 1, the EyeLive system consists of three major components: infrared sensors/emitters, a microcontroller circuit, and computer software. The infrared sensors measure eye reflections and send analog signals to the microcontroller. The microcontroller converts the analog signals into digital data, and transfers the data to the computer through a USB cable. Finally, the computer software determines the direction of eye gazes and controls the GUI.
A/D convertors are sent to the mean measurements (e.g. mean filter). The resulting data processing (e.g. mean filter) is used as the identifier for the given direction. The calibration time for each direction is 2 seconds, making it more aesthetically pleasing.

Two temples and a stretchable strip on the top of the head prevent slippage without causing discomfort. Since EyeLive only detects eye position in 5 directions, a slippage of less than 2 mm is acceptable.

The system consists of low-cost hardware and the infrared sensor/emitter unit. The whole system can be priced below $1000 after considering R&D, marketing, support, and other costs.

2.3 Eye-tracking and calibration algorithms

The 4 LEDs on the EyeLive headset are turned on one at a time. When an LED is on, the intensity of the reflected light is measured by the four phototransistors and converted to four-dimensional vector (e.g. mean filter) to determine user’s intention. If in calibration mode, the 18 outliers with respect to the largest mean measurements (\(m_{avg,1}\)) of 12 measurements to remove high-frequency interference such as high-frequency low-amplitude eye trembles. 24 sets of mean measurements (\(m_{avg,1,2}, a_{avg,1,2}, \ldots, a_{avg,1,24}\), i = 1 to 24) are sent to the computer each second via USB (thus 16*12*24=4608). A set size of 12 was chosen as the optimal value. An excessively small set size provides an insufficient number of measurements to smooth out interference, and wastes USB bandwidth by sending mean measurements too frequently. An excessively large set size filters out some useful movements such as blinking and low-frequency eye jerks. In addition, the microcontroller has enough memory to store and process only 12 measurements.

Before normal usage, the user must first look at 5 positions on the screen (Fig. 3) and close his or her eyes. The system collects information about the user’s unique eye profile. For each position, 48 sets of mean measurements (\(m_{avg,2}\)) are acquired in 2 seconds. For each of the 16 dimensions, the median of the 20 remaining measurements is calculated. Then, the 18 outliers with respect to the largest absolute distance from the median are discarded, and the mean of the 48 measurements is calculated. The resulting 16-dimensional vector \(p\) is used as the identifier for the given direction. The calibration time for each direction is 2 seconds,
as determined from experiments (1 second is too short for a gaze to stabilize, and after 2 seconds the eyes start to blink and the gaze destabilizes due to eye strain).

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 position, such as the mean absolute error (MAE), the mean square error (MSE), and principal component analysis (PCA). The MAE is used because it has the lowest error rate (see results section). The MAE also requires the least computation and is the simplest to implement.

### 2.4 User Interface

EyeLive’s GUI presents a keyboard, as shown in Fig. 4(a). To select a letter, a hierarchical selection approach similar to Morse code is used. The 26 letters of the alphabet are divided into groups. For example, if the user wants to type the letter “H”, he or she first looks to the left, and selects the Left cell by either closing or dwelling the eyes for 1 second. Then the user repeats the process by selecting the Right cell (Fig. 4(b)), and then the Up cell (Fig. 4(c)). EyeLive also has a word prediction feature to speed up typing, and a text-to-speech feature to make the user interface more interactive. In addition, the letters can be replaced by functions like “watch TV” or “light on” in an alternative keyboard design, so that the user can control his/her environment.

Compared to camera-based systems with a full QWERTY keyboard [12,13], EyeLive’s GUI leads to lower eye strain and has higher accessibility for patients with decreased oculomotor functions. Based on Fitts’s law [18-20], the time required to reach a target decreases with increasing target size. Experiments have verified that button size has a significantly positive effect on selection speed for an eye-tracking system [21]. Therefore, with 9 very large buttons, the EyeLive GUI has a higher button selection speed compared to that for a full QWERTY keyboard. When focusing on a particular object, the eye muscles bring the two eyes to convergence so that the images seen by the two eyes can be fused. When the eyes focus on an item for too long, muscular stress and strain can occur [22]. The users all reported that they experienced less eye strain with the EyeLive GUI even if it takes three selections to choose a letter instead of one selection on a QWERTY keyboard. Thus, by reducing the time required for a button selection, the EyeLive interface reduces eye strain. Furthermore, with reduced oculomotor function due to antisaccade and remembered saccade [23], it is very difficult for ALS patients to accurately focus on a small button, making the EyeLive keyboard more suitable. With training, the user can recall the sequence of three eye position to choose a particular letter and perform the sequence very quickly, further increasing the selection speed.

The EyeLive keyboard is slower to use than a full QWERTY keyboard because three selections are required to enter a single letter. The average typing speed is 1.95 words/min for a novice user and 2.91 words/min for an experienced user.

After consulting with many ALS patients and their caregivers, it was found that writing short messages and controlling the environment are more important than writing large amounts of text. Therefore, functionality outweighs pure typing speed. The EyeLive system is designed to provide easy access to frequently used functions. For example, the letters on buttons can be replaced by light switches, TV controls, and music player controls.

Three special function buttons are shown in the bottom cell: “backspace”, “space”, and “restart”. Only two selections are required to activate these frequently used buttons. The user first looks down, and then left, up, or right to select “backspace”, “space”, or “restart”, respectively. When a space is inserted, the system reads the word out loud. The “restart” button takes the user back to the first selection (Fig. 4(a)).
The pEYEwrite system [24-29] uses a hierarchical selection approach similar to that used in the proposed GUI. pEYEwrite groups the letters into 6 equally divided regions of a pie. The system was designed primarily for cursor-based control, and is harder to navigate using 5 eye positions. EyeLive’s 9-cell grid is optimized for navigation with 5 eye positions by fully utilizing the entire screen. The four cells for typing (up/down/left/right) use a very large font for users with poor vision. The middle cell displays the entered words. The four corner cells can display information such as time and a calendar.

3. Results

Nine individuals (eight healthy and one with ALS) were recruited for experiments with the EyeLive system. The individual with ALS did not finish the test. The procedures followed were in accordance with the ethical standards of the responsible committee on human experimentation (institutional and national) and with the Helsinki Declaration of 1975, as revised in 2004. The primary safety concern is the intensity of the infrared light emitted from the LEDs. The light intensity from the LEDs in EyeLive is below 1 mW/cm², which is only 10% of the exposure limit in the ICNIRP (1997) guidelines for infrared radiation [30], and similar to the average corneal exposure from infrared radiation in sunlight [31]. The safety concerns were explained to the test subjects and written consent was obtained.

After the test subjects were introduced to the operation procedures of the system and allowed to practice for ten minutes, they were able to use the system accurately without any difficulty. To numerically verify the accuracy of the algorithm, the 8 healthy individuals with normal eye sight each performed 10 trials. Each subject first calibrated the system by looking at 5 dark gray spots on the screen, as shown in Fig. 5. Then, the subject looked at the 5 dark gray spots and 20 additional light gray spots on the screen (thus a total of 25 eye movements). The light gray spots define the boundary within which the user is most likely to look in each direction. The measurement for each light gray spot was then classified using the MAE, MSE, and PCA algorithms to determine whether the classification matched the intended direction.

![Diagram](image)

Figure 5. Additional points for verification.

The error rates for classification in the 5 directions obtained using the MAE, MSE, and PCA algorithms are shown in Tables 1, 2, and 3, respectively. Table 4 shows that the overall error rates obtained using the MAE, MSE, and PCA algorithms are 5.6%, 6.6%, and 6.3%, respectively. The spots where the errors occurred (indicated by arrows in Fig. 5) are the transition spots between the middle cell and a cell in one of the other four directions. A possible reason for these errors is that the user may not have focused his or her gaze precisely onto the spot during the experiment. In practice, the user can easily avoid such errors by moving the eye farther in the intended direction. In addition, EyeLive is perfect in classifying “eyes closed” (error rate of 0%).

<table>
<thead>
<tr>
<th>Subject #</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trial 1</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Trial 2</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Trial 3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Trial 4</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Trial 5</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Trial 6</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Trial 7</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Trial 8</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Trial 9</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Trial 10</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>1</td>
</tr>
</tbody>
</table>

Mean number of errors/trial: 1.4, 1.7, 1.3, 1.4, 1.3, 1.5, 1.2
Mean error rate: 5.6%, 6.8%, 5.2%, 5.6%, 5.2%, 5.2%, 6.0%, 4.8%

<table>
<thead>
<tr>
<th>Subject #</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trial 1</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Trial 2</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Trial 3</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Trial 4</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Trial 5</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Trial 6</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Trial 7</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Trial 8</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Trial 9</td>
<td>0</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Trial 10</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>1</td>
</tr>
</tbody>
</table>

Mean number of errors/trial: 1.7, 1.9, 1.5, 1.6, 1.5, 1.6, 1.8, 1.6
Mean error rate: 6.8%, 7.6%, 6.0%, 6.4%, 6.0%, 6.4%, 7.2%, 6.4%

<table>
<thead>
<tr>
<th>Subject #</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trial 1</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Trial 2</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Trial 3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Trial 4</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Trial 5</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Trial 6</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Trial 7</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Trial 8</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Trial 9</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Trial 10</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>1</td>
</tr>
</tbody>
</table>

Mean number of errors/trial: 1.6, 1.8, 1.5, 1.6, 1.5, 1.6, 1.4
Mean error rate: 6.4%, 7.2%, 6.0%, 6.4%, 6.0%, 6.4%, 6.4%, 5.6%
4. Discussion

Preliminary experiments show that EyeLive can detect up to 9 eye positions (i.e., all 9 cells in the grid in Fig. 7), but not in a consistent manner. Further investigation can be done to develop more sophisticated variations and combinations of the MAE, MSE, PCA, and other algorithms. If the system can detect 9 eye positions, the number of selections per letter could be reduced to two, increasing typing speed. Artificial neural network techniques may also help increase the accuracy in detecting 9 eye positions by finding patterns from collected user data.

Compensation for device slippage greater than 2 mm will be a topic of further research. Continuous recalculation of the identifier and fast recalibration during run-time are possible solutions.

One male individual with ALS conducted 10 trials with his consent under the supervision of his wife. Unfortunately, neither the EyeLive system nor other high-precision proven eye-tracking systems like Tobii could correctly track his eye movements. In future research, more ALS patients will be recruited to verify EyeLive’s eye tracking accuracy and usability.

The typing speed may have been even higher if the test subjects were able to use the system on a regular basis for a month, but such experiment were been performed because the test subjects could not commit to using it regularly and only one device was made due to budget and time limitations. In future research, more prototypes will be made and ALS patients will be allowed to use the system for an extended period of time.

5. Conclusion

An infrared sensor/emitter-based eye-tracking and computer interface system was proposed and implemented. Compared to camera-based systems with a full QWERTY keyboard, the proposed EyeLive system has a lower price point, higher accessibility, and reduced eye strain at the cost of slower typing speed. Eight healthy individuals performed ten trials each with EyeLive, with an error rate of only 5.6%.

Acknowledgements

The authors would like to thank Siyong Zhu for implementing part of the software and Priyanka Deshmukh for implementing part of the hardware. We would also like to thank the ALS Society of BC and the GF Strong Rehabilitation Center for providing their expert advice on assistive technologies and the needs of ALS patients. The Wighton Engineering Development Fund, the Simon Fraser University Engineering Society, and the Kaiser Foundation are acknowledged for their generous financial support of the project.

References