Assistive System Using Eye Gaze Estimation for Amyotrophic Lateral Sclerosis Patients

Cheng-Chieh Chiang
Department of Information Technology
Takming University of Science and Technology
Taipei, Taiwan

Abstract: Amyotrophic lateral sclerosis (ALS) patients cannot control their muscle except eyes in the later stage of the disease progress. This paper aims to develop an eye-based assistive system that is controlled by the eye gaze to help ALS patients improve their life quality. Two main functions are proposed in this paper. The first one is called HelpCall that can detect the users' eye gaze to active the corresponding events. ALS patients can "talk" with other people more easily by looking at specific buttons in the HelpCall system. The second one is an eye-control browser that allows the users browsing web pages in Internet. We design an interface that embeds the IE browser into several buttons controlled by the user eye gaze. ALS patients can visit the Internet only using their eyes in our proposed system. This paper discusses our ideas for the assistive system and then describes the design and implementation of our proposed system in details.

Keywords: Amyotrophic Lateral Sclerosis; eye-based assistive system; eye gaze; HelpCall; eye-control browser.

1. INTRODUCTION

Amyotrophic lateral sclerosis (ALS) is a neurodegenerative disease that patients gradually lose the control for their muscle in the whole body. Eventually, patients cannot move their any part of bodies, except eyes, in the end of the disease progress. In the last stage of their life, ALS patients are locked in the frozen body and lose the ability to communicate with other people. The rough description about the ALS disease can be found in the Wikipedia site [1].

Eye is the only one part that most of ALS patients can control in the last stage of the ALS. It has much potential to improve the life quality if ALS patients can "do something" using their eyes. Hence, an assistive system that can be controlled by the eye gaze should be helpful for ALS patients. Now there have been some products called Eye Tracker [2][7][9][10] that can estimate what target the user looks at. In current many eye tracker system have been released for different goals that are introduced in Section 2.

This paper aims to design an eye-based assistive system for ALS patients to build a communication bridge. What kinds of applications do ALS patients need? The assistive system does not need complex or rich of functions. Instead, we try to design a simple but helpful user interface for patients. In this paper two applications to help ALS patients in their life are provided. First, we design an assistive function called HelpCall, which allows the patients can easily "talk" with other people by looking at specific buttons in screen. For example, a caregiver may ask whether the patient wants to watch TV or not, and then the patient only need to stare at a "Yes" button to answer it. Second, an eye-control browser is proposed to allow the patients browsing webpages in Internet. Using this eye-control browser, the patients can determine by themselves what pages they want to view.

Figure 1 draws the brief flowchart of our system. When a video frame is captured, we first employ a face detection method to detect the face region in image. In general, a human face contains a fixed layout consisting of eyes, nose, and mouth. Hence, a method for eye detection is performed on a rough eye-area in face to determine the eye regions. Moreover, we design the horizontal and vertical projection for

pupil localization due to the high contrast between the pupil and the surrounding area. The pupil localization can define the coordinate of pupil in eye, and then we construct a mixture of Gaussian Models to estimate the corresponding eye gaze. Finally, the eye gaze can drive the mouse control in the two applications, HelpCall and the eye-control browser, proposed in this paper.
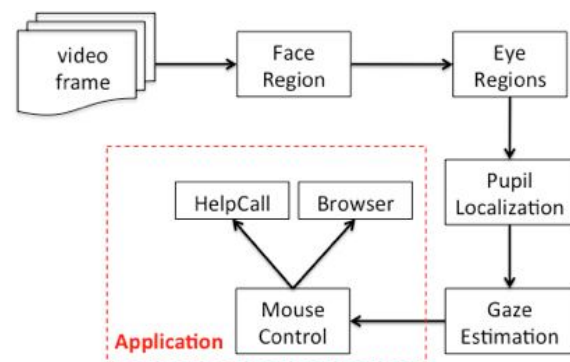

Figure 1. Example of an image with acceptable resolution

The rest of this paper is organized as the follows. Section 2 reviews previous works related to this paper. Section 3 describes how to localize the pupil position in eye, and then Section 4 presents our approach to estimate the user eye gaze. The design and implementation of the proposed assistive system are shown in Section 5. Finally, Section 6 draws our conclusions and discusses potential extensions for this work.

2. RELATED WORKS

Many eye tracking systems have been released in the market for different goals, such as EyeLink [2], MangoldVision [7], SMI iView [9], and Tobii [10]. Most of eye tracking systems request to fix the head position in order to simplify the pupil localization. For example, both EyeLink and SMI iView need to construct a "head holder" in Figure 2(a) and 2(b) to keep the accurate estimation for eye gaze estimation. This case of eye tracking system is not appropriate for ALS patients who need to lie on a bed.
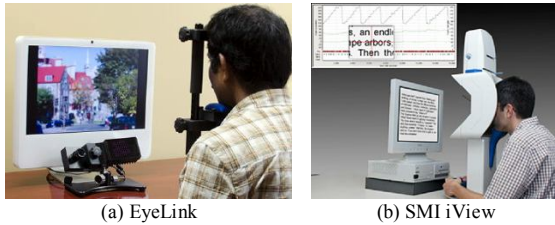
(a) EyeLink   (b) SMI iView

Figure 2. Both EyeLink in (a) and SMI iView in (b) need to construct a "head holder" to keep the head position.

Another type of eye tracking systems is to embed a camera into a glasses to track the pupil position, such as MangoldVision [7] and Tobii [10] shown in Figire 3. These two systems are wearable, hence ALS patients can wear the camera-glasses on bed to control the computer system. In fact, Tobii is somewhat widely used for ALS patients. Unfortunately, the prices of Tobii systems are also a little high. Our work in this paper tries to employ a common camera to reduce the cost of the whole system. Moreover, we design two specific functions to help ALS patients to improve their inconvenient life.
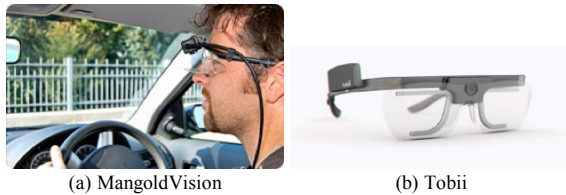

(a) MangoldVision   (b) Tobii

Figure 3. MangoldVision in (a) and Tobii in (b) designs wearable glasses that appends a camera to estimate the pupil position.

3. PUPIL LOCALIZATION

Pupil is one of the most significant features in eye. Given a video frame, this section presents our procedures to locate the pupil position in eye, including the face detection to separate the face region, eye region detection on the face region, and determining the pupil coordinate in eye.

3.1 Face Detection

Face detection is a key technology to locate or identify human in image. Many state-of-the-art approaches have been proposed to treat the face detection in either image or video under different conditions of the real world [5][6].

This work adopted the adaboosting approach [11], which was first proposed by Viola and Jones in 1999, to automatically detect face regions in a video frame. This approach collects a large number of Haar-like features that are fast computed by the integral image. Then, an adaboosting approach [3] is employed to select most significant Haar-like features of face regions. In general, many Haar-like features may be contained to construct an efficient classifier in the adaboosting approach, but that should also need more computational time for face detection. In order to achieve the real-time requirement, Viola and Jones designed a cascading structure, the basic concept is shown in Figure 4, to fast filter most of trivial non-face regions out using fewer Haar-like features in the several starting stages and then accurately to detect face regions using more features in the ending stages.
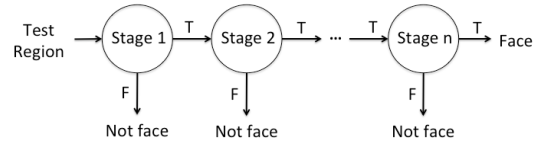

Figure 4. The cascade structure to speed up the face detection

The adaboosting approach of face detection has been implemented in the openCV library [8] and was widely used in many applications. Although this approach may be failed if the orientation of face regions is skew, it is still appropriated for our framework due to skewed human faces can be simply ignored in eye-related applications such as the eye state monitoring for car drivers.

3.2 Eye Region

The eye region in face contains a strong characteristic that there is high contrast between the eyes and the surrounding areas. Thus, the gray channel is more appropriate than color channels in representing eye regions. Hence, the first step to detect the eye region is to convert the face image from color to gray. We adopt the simplest averaged method for the gray conversion: gray = (R+G+B)/3, where R, G, and B are the pixel values in the red, green, and blue channels, respectively.

Next, we employ the adaptive thresholding method [4] to make a binary image from the gray image of face. In the traditional thresholding method for binarization, it is difficult to choose a proper threshold globally for a whole image. Instead a global threshold, the adaptive thresholding method computes the average of pixel values in a slide window to be a local threshold value associated with the center pixel of the slide window. The adaptive thresholding method can make a better binarization by adapting the intensity various locally. The details of the adaptive thresholding method can be found in [4].

Now a binary image of face can be available from a video frame. Our next step is to determine the accurate position of the pupil in eye. Since our goal of this work is to design a practical system for ALS patients, the computation in our methods must be simple such that the system can have fast responses in work. Thus, we employ the horizontal and vertical projection method [12] to fast but efficiently capture the eye regions. Assume that $I(x, y)$ is the pixel value at (x, y) in a candidate of eye region with size m by n pixels. Because the image is converted to binary, the pixel values have to be either 0 or 1. The horizontal (H_{porj}) and vertical (V_{proj}) projection on the binary image can be defined as the follow,

$$H_{proj}(x) = \sum_{x=1}^{n} I(x, y), \text{ where } x = 1 \text{ to } m$$

$$V_{proj}(y) = \sum_{x=1}^{m} I(x, y), \text{ where } y = 1 \text{ to } n. \tag{1}$$

Figure 5 presents an illustration that performs the horizontal and the vertical projections in a binary eye image. Due to the low intensity of the pupil, the maximal peaks in the projections can indicate the center in the two coordinates. That is to say, the center coordinate of the pupil can be located according to the two maximal peaks.
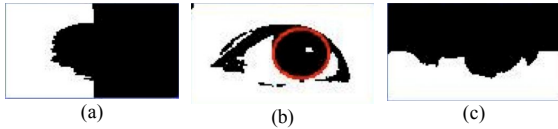
Figure 5. An illustration for eye region decision. (a) and (c) are the results of the horizontal and the vertical projections in the binary eye image (b), respectively. Thus, the intersection of maximal peaks in the two projections can locate the eye region drawn in the red circle of (b).

4. GAZE ESTIMATION

In order to understand what target the user is looking at, we have to construct the mapping between the coordinate in screen and the pupil position in eye. Unfortunately, the coordinate scales of screen and eye region are very different. Assume that an image is 640x480 captured from a common webcam, and hence we cannot expect to get an eye region with high resolution. In our implementation, the width of an eye region is often less than 100 pixels in our system. In contrast to pupil, the resolution is often very high in the current screen. Hence, the coordinate mapping from pupil to screen could be very sensitive to the accuracy of the pupil localization.

The whole screen are divided into 3x3=9 grids instead of high-resolution coordinates to reduce the sensitivity of the gaze estimation. Let $L = \{L_i| i=1, …, 9\}$ be the nine areas in the screen, and then our task is to convert the coordinate of pupil position in eye to the nine grid. To this end, we construct nine Gaussian distributions, denoted G_i with mean μ_i and variance σ_i^2, associated with the nine grids $L = \{L_i| i=1, …, 9\}$. Given a pupil position (x, y), the gaze can be assigned to the screen grid L_{i*}, where

$$i^* = \underset{i}{\arg\max}\, G_i(x, y) \qquad (2)$$

The mean μ_i and variance σ_i^2 of Gaussian distribution G_i can be computed by the training images that are captured by asking for the patient looking at the grid L_i. When the proposed assistive system starts, the screen first shows a circular point in a grid and then asks for the patient watching this point. Repeating nine times with changing grids, the system can capture a lot of training images that the patient looks at each of grids. Therefore, the mean and variance of the pupil coordinates associated with a grid can be computed to determine the Gaussian distribution.

5. DESIGN AND IMPLEMENTATION

This section presents our design of the assistive system that is controlled by the user gaze. Two systems are proposed in this paper. The first one is called HelpCall that can help ALS patients communicate with people. The second one is an eye-control browser to help the patient does not need a mouse but use the eye gaze to control the browser in Internet.

5.1 HelpCall

The idea of HelpCall system is intuitive: one button presents one meaning. The users can look at a specific button to present their opinions without any talks. Figure 6 shows the interface of this system that contains four dedicated and five user-defined functions. The four dedicated functions are:

- TV: ask for watching TV
- CALL: trigger an alarm to call the caregiver.
- YES/NO: a simple answer in talk.

This system preserves five functions that patients can define them according to their needs. All of nine functions can be edited by the users if necessary.

Figure 6. The user interface of HelpCall system, containing four dedicated (blue) and five user-defined functions.

5.2 Eye-Controlled Browser

ALS patients have been imprisoned in body, hence our idea is to design an eye-controlled browser that can allow them contact rich contents of Internet. The most challenging issue of this system is to design an appropriate interface that cannot only provide a simple way browsing a web page for ALS patients but also cover the complex functions of a browser.
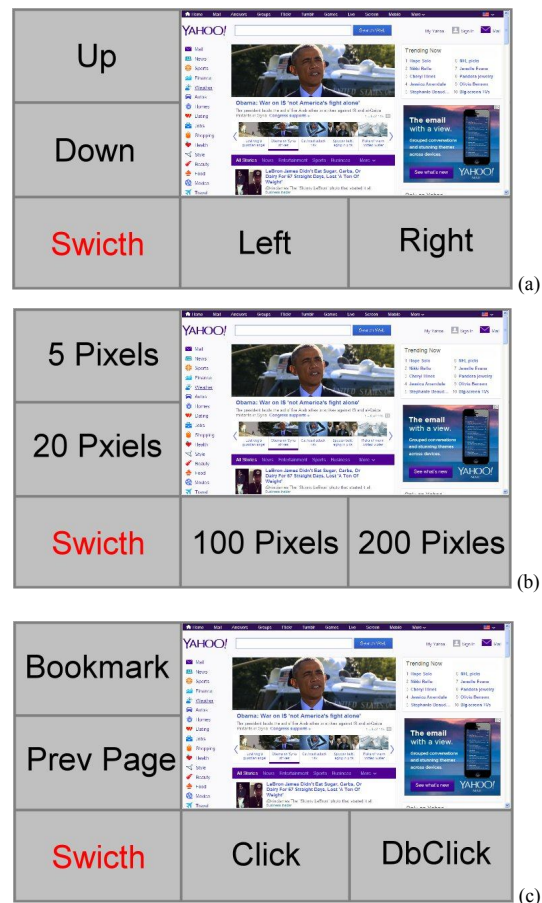

Figure 7. The user interface of our proposed eye-controlled browser system. Buttons in three pages can help ALS patients to control the IE browser without mouse.

The interface of the proposed system is shown in Figure 7 that embeds several buttons in order to control the cursor by eye gaze instead of moving mouse. The interface is also based on the layout of 9x9 grids, where the browser covers four right-upper grids and functional buttons are on the other five grids. The "Switch" button can change functions such that 4x3=12 functional buttons are provided to control the browser. In the first page of Figure 7(a), buttons are used for moving the cursor with the directions. The buttons of Figure 7(b) can determine the step size of cursor moving. Figure 7(c) provides four important functions, containing "Click" and "DbClcik" for mouse click and double click, respectively, "Bookmark" for opening the bookmark page, and "Pre Page" for returning the previous page.

ALS patients can completely control the browser using our proposed interface by their eye gaze. First, the users can look at "Switch" button to change the page of Figure 7(a), and then enable the four directional buttons to move the cursor in the browser. If the users want to change the moving step of cursor, they just need to switch to the second page in Figure 7(b). When the cursor stays at a link of interest in the web page, the users can switch the page again to Figure 7(c) and enable the "Click" function to enter the hyperlink.

6. CONCLUSION AND FUTURE WORK

ALS patients should be difficult to control their body muscle in the late stage of life. This paper proposes an assistive system with two functions, HelpCall and an eye-controlled browser, to help ALS patients based on the computer vision technologies. The HelpCall system provides nine buttons of grids to present nine responses for ALS patients by using eye gaze. The eye-controlled browser embeds additional buttons to allow ALS patients moving the cursor in the browser using their eyes. These two assistive functions can definitely improve the quality in life for ALS patients.

Two main extensions are included in our future works. Using eye gaze to control the interface is not an easy way for users in a practical system. Hence, we plan to develop the blink detection for the proposed assistive system and then to perform a user study of this system controlled by the eye gaze and blink. The second potential extension for this work is to design an eye-controlled keyboard, either eye gaze or blink, that users can enter words in the browser such that the functions in the proposed assistive system can be more complete.

7. REFERENCES

[1] Amyotrophic lateral sclerosis,
http://en.wikipedia.org/wiki/Amyotrophic_lateral_sclerosis

[2] EyeLink, http://www.sr-research.com/

[3] Freund, Y., Schapire, R., and Abe, N., A short introduction to boosting. Journal-Japanese Society For Artificial Intelligence, 1999. 14: p. 771-780.

[4] Gonzalez, R. C. and Woods, R. E., Digital Image Processing, 3rd Edition. Prentice Hall, 2007.

[5] Hsu R.-L., Abdel-Mottaleb M., and Jain A. K., Face detection in color images. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002. 24(5): p. 696-706.

[6] Liao S., Jain A. K., Li S. Z., Partial face recognition: alignment-free approach. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013. 35(5): p. 1193-1205.

[7] MangoldVision,
http://www.mangold-international.com/home.html

[8] OpenCV, http://opencv.org/

[9] SMI, http://www.smivision.com/en.html

[10] Tobii, http://www.tobii.com/

[11] Viola P. and Jones M. J., Robust real-time face detection. International journal of computer vision, 2004. 57(2): p. 137-154.

[12] Zhou Z.-H. and Geng X., Projection functions for eye detection. Pattern recognition, 2004. 37(5): p. 1049-1056.