Fuzzy Logic Based Eye-Brain Controlled Web Access System

Suvir P. Mirchandani¹

Correspondence: Suvir Mirchandani, Fox Chapel Area Senior High School, Pittsburgh, PA 15238, USA. E-mail: smirchan@gmail.com

Received: July 18, 2015 Accepted: August 9, 2015 Online Published: November 6, 2015

Abstract

Accessing the Web is crucially important in today's society because of communication, education, and entertainment opportunities. Paralyzed or paretic individuals are unable to capitalize on these opportunities using traditional human-computer interaction methods. Thus, I develop a low-cost web browsing system for such individuals, integrating eye and brain control in a novel fashion to relay and interpret navigation commands. The system combines gaze position estimates obtained from a new image processing algorithm with brain concentration levels sensed and transmitted by an electroencephalogram headset. Since user intent may itself be uncertain, the system incorporates a novel fuzzy logic algorithm for combining brainwave and eye position inputs to determine the user's targeted hyperlink. The algorithm adopts an approach based on exponential smoothing to efficiently keep a record of historical signals. Experimental evaluation established that the first attempt success rate of the system lies between 87% and 95% with 95% confidence. Error recovery accuracy is 98.4%, resulting in a second attempt success rate of 99.1%. This assistive technology, using off-the-shelf components, costs less than \$100 excluding the tablet computer, and has the potential to affordably open up the Web to millions of individuals.

Keywords: accessible computing, electroencephalography, eye tracking, fuzzy logic, human-computer interaction, hybrid brain-computer interface

1. Introduction

Individuals with paralysis and paresis resulting from disease, injuries, or old age lack the ability to control their limbs and may be speech impaired. For example, "locked-in" individuals numbering half a million worldwide have complete loss of voluntary movement and speech (Randolph & Moore Jackson, 2010). These individuals find it difficult to communicate directly with friends and family. Indirect communication using computers is a possibility; however, existing systems that work without mouse, touch, or speech input can be expensive, cumbersome, and require extensive training. Consequently, these individuals can experience social isolation which in turn can cause serious behavioral health issues (J. Cacioppo & S. Cacioppo, 2014).

The primary goal of this work is to develop and evaluate a system that can improve the quality of life of such individuals by allowing them to easily communicate with society. One of the modern interfaces that can be used as a gateway into society is the World Wide Web. The proposed approach is a multimodal mechanism for allowing disabled individuals to browse the web, giving them access to the communication, education, and entertainment opportunities available to others. The duality of brain and eye controls makes this system fast, intuitive, and accurate. Moreover, the system is inexpensive; the cost is less than one hundred dollars.

We typically use multiple "senses" to interact with a computer. For example, we use vision to read the contents on a monitor and then use touch or typing to input the desired command. Specifically, when browsing the web, we scan the currently displayed webpage and then click (using touch or a mouse) on a link to go to the next page. Instead of using our fingers to tap or click, we may be able to use voice activation in some cases. However, individuals who can neither control their limbs nor speak are unable to use the World Wide Web affordably. As a result, they miss out on societal interaction – something that others can take for granted and frequently do through, for example, social media websites. This system offers a potential solution for such disabled individuals.

In the proposed system, users still use vision to scan the monitor. However, instead of touch, keyboard, or voice input for relaying a command about where to navigate, this system uses an innovative integration of eye and brain control. This is the first contribution of this work. The second contribution is the development of an image

¹ Fox Chapel Area Senior High School, Pittsburgh, PA, USA

processing algorithm for eye control. Since information from such processing can be imprecise because of both uncertain user intent when interacting with a computer (particularly when browsing the web) and inaccuracy in measuring the eye movements, this study uses fuzzy logic to analyze the input in a novel fashion, processing not just the most recent signal, but a combination of previously recorded signals. This is the third main contribution. The fourth contribution is the implementation of this system in a user-friendly application for the iPad. Finally, system evaluation demonstrates that its first attempt accuracy is 90.9% and the accuracy with a second attempt is 99.1%. Since this assistive system has the potential to create new opportunities for locked-in individuals, I use the acronym FLYAWAY (Fuzzy Logic based eYe-brAin controlled Web Access sYstem).

This paper is organized as follows. Section 2 provides a literature review and hardware assessment. Section 3 provides an overview of this system. Section 4 supplies the technical details of the various elements of FLYAWAY, and discusses the algorithmic development and hardware-software integration. Section 5 describes the data collection and analysis. Finally, Section 6 concludes the paper and provides some future research directions.

2. Literature Review and Hardware Assessment

2.1 Literature Review

Multimodal human-computer interaction systems use more than one modality (for example, speech, vision, facial expressions, gestures, pen, and touch) (Oviatt, 2012). Multimodal systems are becoming increasingly prevalent as they can improve or even replace traditional input methods such as keyboard and mouse (Thiran et al., 2010). By using multiple modalities, we can increase the reliability of the inference about user intent, even if the same information is conveyed using more than one method. As an example, lip movements from a video signal can compensate for loss of an audio feed. Multimodal systems can also incorporate complementary information from the incoming signals. The famous demonstration "Put That There" by the Architecture Machine Group at MIT augments voice commands with gestures to manipulate objects on a screen (Bolt, 1980). Because a greater amount of data needs to be collected and analyzed in multimodal systems, the required computational power increases. Even though modern computing systems are becoming more powerful, the "curse of dimensionality" manifests itself. This drawback is described later in this paper.

FLYAWAY uses eye and brain data in a multimodal form. The use of these two types of input signals in tandem is recent and limited, and has typically relied on expensive equipment for both input modalities (McMullen et al., 2014; Zander et al., 2010). Even though eye tracking has been used for the last couple of decades, brain control is still in its infancy. Only recently has affordable and easy-to-use brain activity data acquisition hardware started to appear. FLYAWAY integrates eye and brain inputs using fuzzy logic.

2.1.1 Eye Tracking and Gaze Estimation

Eye tracking focuses on localization of the eye in a face image and gaze estimation tracks eye gaze paths. In an excellent survey on models for eye tracking, Hansen & Ji (2010) point out that despite active and significant progress in the area, "eye detection and tracking remains challenging." Challenges that researchers have identified include ambient lighting, head pose, occlusion of the eye by eyelids, and characteristics unique to the individual: size, color, and reflectivity of the iris.

Geometric and photometric characteristics of the eyes are important for eye detection. There are generally three different approaches that are used for eye detection. These are (i) *shape-based* approaches which geometrically model parts of the eye, (ii) *appearance-based methods* which detect eyes based on color distribution or filter response of a facial image, and (iii) *feature-based* methods which identify inherent characteristics of the eye, such as the limbus, pupil, and cornea reflections.

Some simple shape models, such an elliptical model, are not capable of capturing variations of eye features such as eyelids and eyebrows. Deformable shape models consist of scalable and deformable templates to allow for a more detailed fit of the eye shape, but these methods are computationally demanding and may fail when the eye changes shape (e.g., when the eyes are wide open or occluded by eyelashes) (Hansen & Ji, 2010). A specific type of appearance-based methods that is based on filter response of an image includes the analysis of Haar-like features. Haar-like features have high computational efficiency (Castrillón et al., 2007). Feature-based methods tend to be less sensitive to variations in viewing angle and illumination. One particular method (Yang et al., 1998) locates pupils by searching for dark regions that satisfy certain constraints using a threshold algorithm. The accuracy of this method depends on dark regions such as eyebrows and dark circles around the eyes. If a common threshold is used for both eyes, however, this method can fail due to variable light conditions across the face.

After locating the eye, the point on the screen at which the user is looking needs to be determined. This is called gaze position estimation. In order to determine gaze position, we have to analyze both the head pose and eye orientation. Some low-cost approaches require head gear to be worn on the user (EyeWriter, 2010; Li & Parkhurst, 2006), but this can be cumbersome for the user and a caregiver; FLYAWAY instead assumes that head position is constant with respect to the camera and uses a built-in tablet camera for the facial video feed (see Section 2.2.2).

The approach for eye tracking in this paper uses Haar feature analysis and a dynamic adaptive threshold, which results in a reliable system for eye detection despite barriers presented by the use of low-resolution camera input. Hansen and Ji (2010) pointed out that future research for eye tracking should include low-cost approaches, high degree of tolerance, and the ability to interpret gaze behavior. The method developed in this study is low-cost, and since interpreting a gaze requires a high degree of tolerance, this method uses fuzzy logic.

2.1.2 Brain-Computer Interfaces

Computer input devices traditionally use secondary sensors; i.e., their signals are a consequence of cognitive change or muscle movement (Hinckley & Wigdor, 2002). In contrast, primary sensing directly measures brain activity, possibly through electroencephalography (EEG), or muscle movement, for example through electromyography (EMG). Electroencephalography measures the electrical voltage fluctuations created by neurons in the brain. EEG technology is primarily used in medical environments and scientific data gathering. EEG-based systems are also being developed for computer input control. A more detailed description of EEG technology is provided in Section 4.2.

A brain-computer interface (BCI) is a "non-muscular channel for sending messages and commands to the external world" providing a direct connection between the brain and a computer (Wolpaw et al., 2002). Müller-Putz et al. (2011) note that BCIs, often using EEG signals for input, are a suitable options for people with motor disabilities to give them a method of control without the need for movement. Allison et al. (2012) provide a recent survey of brain control research. Current brain-machine interface technology has limitations, and researchers have overcome these limitations by using it in multimodal settings and in a semi-autonomous framework (McMullen et al., 2014; Müller-Putz et al., 2011; Zander et al., 2010). Such applications are known as *hybrid BCIs* (Pfurtscheller, et al., 2010). In a study combining EMG and EEG signal input, Leeb et al. (2010) propose augmenting brain signals with muscular signals available in a disabled person to utilize any remaining functionalities for control. In the same vein, FLYAWAY combines brain signals and eye input to take advantage of two critical functionalities in a locked-in individual.

FLYAWAY is distinctive in the way it uses brain activity to sense user intent in that it does not require expensive equipment (see Section 2.2.3). The use of brain imaging remains in its infancy (Hinckley & Wigdor, 2002), and "the development of efficient BCIs and their implementation in hybrid systems that combine well-established methods in [human-computer interaction] and brain control will not only transform the way we perform everyday tasks, but also improve the quality of life for individuals with physical disabilities" (Papadelis et al., 2013).

2.1.3 Fuzzy Logic and Theory

Lotfi Zadeh (Zadeh, 1965) proposed the theory of *fuzzy sets* in 1965. Fuzzy sets differ from classical sets in that every element of a universal set \mathcal{U} has a *degree* of membership that belongs to the closed interval [0, 1], as opposed to the clear membership or nonmembership of classical sets. Classical sets are also referred to as *crisp sets* because the membership states of elements are unambiguously defined. Fuzzy sets have an increasing number of applications in natural language processing, risk analysis, and engineering (e.g., Cambria et al. 2014; Ross, 2009; Zimmerman, 2010). A more detailed explanation along with how FLYAWAY uses fuzzy logic is included in Section 4.3.

2.2 Hardware Assessment

This subsection discusses hardware used for running the mobile application and for detecting eye and brain signals.

2.2.1 Tablet Computer

The application is written for Apple iPad to ensure mobility and ease-of-use while having access to sufficient computational power. The application has been tested on an iPad (third generation) with 16GB of storage and an Apple A5X dual-core chip.

2.2.2 Eye detection and Tracking Equipment

FLYAWAY uses the built-in iPad frontal camera, thereby eliminating the requirement for an external camera and

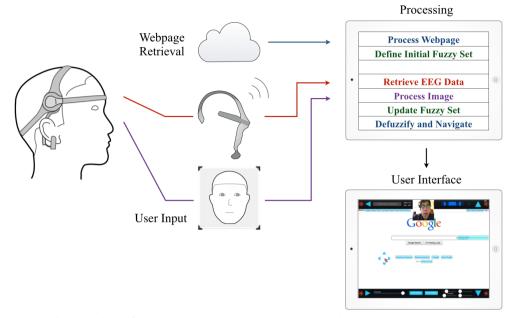


Figure 1. Schematic Overview of the System. FLYAWAY combines brain and eye inputs to select the appropriate hyperlink, presented in a tailored web browser

an additional camera expense. As discussed above, approaches for eye tracking which are low-cost often require head gear to be worn on the user (EyeWriter, 2010; Li & Parkhurst, 2006). Some commercial eye tracking systems use specialized cameras, and can cost more than \$4,000 (Tobii Technology, 2013).

2.2.3 Electroencephalography Equipment

The cost of EEG equipment can range from \$2,000 to \$100,000 for medical diagnosis or scientific research, and \$100 to \$500 on the consumer market (Goldberg, 2012). To align with the goal of creating a low-cost system, this work used the NeuroSky MindWave Mobile, the least expensive EEG headset at the time of this publication, to measure the user's attention level. The headset connected to the iPad via Bluetooth. Even though this hardware costs less than \$100, integrating the input from the brain control with the input from eye control resulted in an accurate system.

3. System Configuration and Schematic Overview

FLYAWAY allows a user to utilize eye tracking and brain control to navigate the web. In existing eye tracking interfaces, the primary form of interaction with the computer is "clicking" on a button by simply looking at it for a certain amount of time (called *dwell* time). Since cognitive ability often remains in individuals who are motor-impaired (Wilson et al., 2011), FLYAWAY utilizes the concentration level of the user as a second method of control. This bimodal control, using the eyes and the brain, utilizes economical and easily available hardware. Figure gives a schematic of the system. While more details of the implementation are given in Section 4, an overview is given here. When a user opens a webpage, it is downloaded onto the user's mobile device. FLYAWAY processes the webpage to find clickable links and highlight them on the web browser, which I design and implemented as an iPad app. As the user browses the web page, the gaze is tracked using image processing on camera input. This information helps FLYAWAY define and update a fuzzy set. The elements of this fuzzy set are the various hyperlinks and their membership degrees reflect the user's intent. The membership degrees are constantly updated using the most recent gaze estimation point. Simultaneously, the electroencephalogram headset is used to measure the user's attention level and transmit this information to the iPad. Together, EEG attention level signals and gaze estimation are used to determine the user's intended link.

4. System Development

4.1 Eye Control: Image Processing

The frontal camera of the iPad captures the facial image of the user. This image is processed using the OpenCV software library (Bradski & Kaehler, 2014) to find a general region around the eyes and subsequently approximate

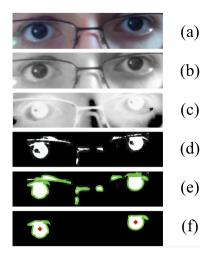


Figure 2. Eye Detection and Tracking Sequence. This figure visually depicts the sequence of the image processing algorithm

the centroid of each iris. More details are provided below.

4.1.1 Eye Detection

The red channel in an RGB eye image is especially useful for distinguishing the pupil from other parts of the eye (Fukuda et al., 2011). This system extracts the red channel from the image inputted by the camera, and uses the Haar Cascade Classifier two_eyes_large (Castrillón et al., 2007) to detect a basic Region of Interest (ROI) around both eyes (Figures (a) and (b)).

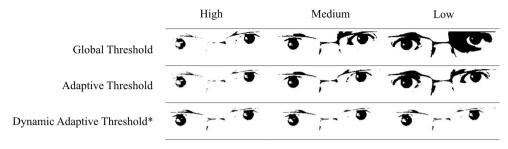
4.1.2 Eye Tracking

Locating the Iris. The frame is inverted (Figure (c)) by taking the complement of the grayscale pixel values. As will be shown, this inversion helps in the calculation of the centroid of the iris. By observation, the iris and pupil tend to be the darkest parts of the eye area (and thus the lightest parts on the inverted grayscale image). Therefore, binarization is used to distinguish the iris. As noted previously, a global binarization threshold would not be sufficient due to variations in lighting across the face. This study addresses that problem by applying an adaptive binary threshold on the inverted grayscale image (Figure (d)). Let binaryImage(x, y) denote the value of the binary image at coordinate (x, y), grayFrame(x, y) denote the pixel value at coordinate (x, y), x denote the mean of the pixel values in a neighborhood region around (x, y), and x denote a constant. The function binaryImage(x, y) is computed as follows:

binaryImage
$$(x, y) = \begin{cases} 255 & \text{if } \operatorname{grayFrame}(x, y) > \tau(x, y) - C \\ 0 & \text{otherwise.} \end{cases}$$
 (1)

In the adaptive binary threshold formula, C is a preset constant. However, in testing the algorithm, it was determined that the contrast of the eye and surrounding areas affect the optimal value of C. To ensure that the iris shape approximation is accurate for all situations and is robust to shade and lighting conditions, FLYAWAY *dynamically* adjusts C based on the binary image resulting from the adaptive binary threshold. Experimentation helped conclude that the optimal iris contour's area is about 3% of the eye region detected. Thus, if the iris contour is greater than the optimal area, C is decreased (so the pixels are more selectively assigned to 255). Likewise, if the iris contour's area is lower than desired, C is increased.

Figure displays a comparison of the three different thresholding techniques. The first row presents an example of global thresholding with high, medium, and low values for the threshold. The adaptive binary threshold adjusts to, for example, the differential in lighting on the two sides of the face. However, the high, medium, and low values of C still affect the resulting binary image (second row). In the last row, C dynamically responds to the current binary image output (and thus, the different thresholding levels do not apply). Adjusting C results in the most consistent detection of the iris. FLYAWAY uses OpenCV's findContours() function to locate the bounding areas of white regions in the binary image (Figure (e)). The contours are filtered by area to remove noise.



* Since the Dynamic Adaptive Threshold adjusts to the optimal value of *C*, the "High," "Medium," and "Low" thresholds do not apply.

Figure 3. Thresholding Techniques. FLYAWAY implements dynamic adaptive thresholding to improve the accuracy of iris contours

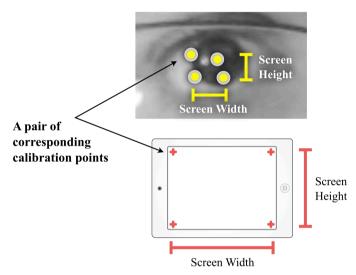


Figure 4. Calibration. A four-point calibration procedure improves the accuracy of gaze estimation

Centroid Calculation. FLYAWAY calculates moments on each filtered contour to compute the iris centroid (see the two dots in the iris contours in Figure (f)). The moment $M_{mn}(A)$ of an array A with x- order m and y- order n is defined by

$$M_{mn}(A) = \sum_{x} \sum_{y} (A(x, y) \cdot x^{m} y^{n}), \qquad (2)$$

where A(x, y) is the value of the pixel at location (x, y) in array A. The x-value (y-value) of the centroid of a finite set of points in 2-D space is the mean of all x-values (y-values). Recall that the grayscale image was inverted (Figure (c)). Pixels within the iris approximation contour have a value of 1 (white). Thus, $M_{00}(A)$ yields the area of the contour A, and the coordinates of the centroid of array A can be expressed as

Centroid(A) =
$$\left(\frac{M_{10}(A)}{M_{00}(A)}, \frac{M_{01}(A)}{M_{00}(A)}\right)$$
. (3)

Depending on the side (right or left) of the centroid, each centroid is added to an array.

4.1.3 Gaze Estimation

To perform gaze estimation, the iris centroid approximation must be mapped to a point on the screen. To accomplish this, a four-point calibration method is implemented, during which the user looks sequentially at each of four visual stimuli on the screen. Upon selection of each stimulus, the eye image is captured, and the centroid approximations are used as a calibration point. The calibration points are stored in an array for further processing. The upper part of Figure provides examples of the centroid approximations overlaid on each other for comparison and the lower part shows the four visual stimuli. The calibration point associated with each stimulus is used for

mapping the eye image coordinates onto a point on the screen. Let (X_{Gaze}, Y_{Gaze}) be the mapped gaze estimation point, (X_C, Y_C) be the current centroid position, (X_{TR}, Y_{TR}) and (X_{TL}, Y_{TL}) be the coordinates of the top right and top left calibration points, (X_{BR}, Y_{BR}) and (X_{BL}, Y_{BR}) be the coordinates of the bottom right and bottom left calibration points, and (ScreenRes_x, ScreenRes_y) be the resolution of the screen in pixels. The following formula is used for estimating gaze position:

$$(X_{\text{Gaze}}, Y_{\text{Gaze}}) = \left((X_C - X_{TL}) \cdot \frac{\text{ScreenRes}_x}{\frac{X_{TR} - X_{TL} + X_{BR} - X_{BL}}{2}}, (Y_C - Y_{TL}) \cdot \frac{\text{ScreenRes}_y}{\frac{Y_{TR} - Y_{BR} + Y_{TL} - Y_{BL}}{2}} \right). \tag{4}$$

A cursor image is drawn centered at the gaze estimation point. It is assumed that the head position of a user with paralysis could be kept stationary; it would also be easy to keep the device stable, for example, by attaching a bracket to a wheelchair or bed.

4.2 Brain Control: EEG Processing

Electroencephalography relies on electrical impulses created by neurons in the brain. When neurons communicate, current is created. To measure brain activity, voltage fluctuations are measured from the scalp. Spikes in electrical potential are known as neural oscillations. Different frequencies in these oscillations can be attributed to different actions and stages of consciousness (Berger, 1929).

Delta waves have frequency up to 4 Hz. They normally occur when one is in deep sleep. Theta waves which range from 4 to 8 Hz are linked to daydreaming and emotional stress. Alpha waves which have frequency ranging from 8 to 13 Hz are associated with relaxation. Mu waves which are found in the same frequency range are associated with the intention of motor movement. Beta waves have frequency in the range 12 to 30 Hz. They are associated with focused concentration, attention, and arousal. Gamma waves which range from frequency 30 Hz and up are associated with perception and cognition (Georgieva et al., 2014).

EEG equipment has traditionally been used in the medical field, but is now becoming more readily available. EEG is the most common approach for brain signal acquisition without surgically implanted devices (Randolph & Moore Jackson, 2010). The EEG control augments the eye tracking interface to serve as another level of control. As the user focuses their attention on a hyperlink on the screen, the program determines whether he/she wants to select it.

After a one-time pairing procedure, the headset automatically connects to the iPad using the built-in Bluetooth communication capability. This application uses the ThinkGear SDK for iOS for communication with the headset. The eSenseAttention value provided by the headset primarily uses Beta wave values to measure concentration, focus, and directed mental activity of the user.

4.3 Integrated Eye and Brain Control Using Fuzzy Logic

Fuzzy logic is used to handle the potential uncertainty in user intention about where to navigate next. Since hyperlinks are the primary interactive item on a webpage, they are regarded as locations that the user could wish to select. For every gaze estimation point, a fuzzy set is created containing elements representing all of the hyperlinks. To illustrate, some definitions are first necessary.

4.3.1 Fuzzy Set Theory

A set is defined as a collection of distinct elements. Every set has a characteristic function, representing an element's membership in the set. In classical, or *crisp*, sets, an element's membership or nonmembership is well defined.

Let \mathcal{U} denote a universal set and \emptyset denote a null set. Let $\operatorname{char}_{\mathcal{A}}(x)$ denote the characteristic function of a set $\mathcal{A} \subseteq \mathcal{U}$ and $x \in \mathcal{A}$.

- \mathcal{A} is a *crisp set* if and only if $\forall x \in \mathcal{A}$, $\operatorname{char}_{\mathcal{A}}(x) \in \{0, 1\}$.
- \mathcal{A} is a *fuzzy set* if and only if $\forall x \in \mathcal{A}$, char $_{\mathcal{A}}(x) \in [0, 1]$.

Thus, fuzzy sets generalize crisp sets. Henceforth, fuzzy sets are denoted by a tilde over an uppercase script letter. Figure shows an example when the characteristic functions have an inverse relationship with the values of x.

Fuzzy sets are useful when information is imprecise. For example, we are often imprecise in our assessments or descriptions of situations. We use words such as "virtually," "practically," "typically," and "roughly" which denote

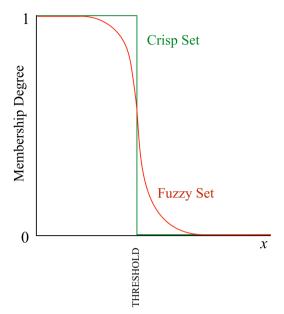


Figure 5. Examples of Characteristic Functions. In a fuzzy set, membership degree of an element can take up any value between 0 and 1

different degrees of certainty. Since these words might mean different things to different people, they cannot be analyzed using classical set theory. Fuzzy sets can model situations of this type, and have been used in several engineering applications. However, in the situation that this system models, where user intentions may be unclear and there may be measurement errors, the novel application of fuzzy sets is immensely useful.

Fuzzy set theory follows all but two of the axioms of classical set theory. The union of fuzzy sets \mathcal{A} and \mathcal{B} is defined by the characteristic function $\operatorname{char}_{\tilde{\mathcal{A}} \cup \tilde{\mathcal{B}}}(x) = \max(\operatorname{char}_{\tilde{\mathcal{A}}}(x), \operatorname{char}_{\tilde{\mathcal{B}}}(x))$ and the intersection is defined by the characteristic function $\operatorname{char}_{\tilde{\mathcal{A}} \cap \tilde{\mathcal{B}}}(x) = \min(\operatorname{char}_{\tilde{\mathcal{A}}}(x), \operatorname{char}_{\tilde{\mathcal{B}}}(x))$. Although De Morgan's principles apply to fuzzy sets, $\tilde{\mathcal{A}} \cup \tilde{\mathcal{A}}^C \neq \mathcal{U} \land \tilde{\mathcal{A}} \cap \tilde{\mathcal{A}}^C \neq \emptyset$.

4.3.2 Determining Membership Degrees

Every hyperlink on the webpage is assigned a membership degree that reflects the degree of desirability of the link to the user. The membership degree of each link depends on the location of each link and the gaze estimation point.

Location of Hyperlinks. In order to find the location of hyperlinks, this program uses the HTML source and JavaScript. Upon loading of a webpage, it obtains and processes the HTML to add an id attribute to every hyperlink element. The HTML attribute href indicates a hyperhyperlink, so a regular expression is initiated to locate all hyperlinks and append each occurrence with id="Link%u" where %u specifies an integer relating the position of the hyperlink occurrence within the HTML. The JavaScript functions getElementById() and hypergetClientRects() are used to find the spatial coordinates of every hyperlink rendered within the browser. **Membership of Hyperlinks.** The location of a hyperlink is used as one of the factors that determine its membership degree. Let \mathcal{L} be the set of links L_i and $\widetilde{\mathcal{F}}_t$ be the fuzzy set for the gaze position at time t. The distance d_i between link L_i with coordinates $L_{i,x}$ and $L_{i,y}$ and the cursor C (drawn at the gaze estimation point, with coordinates C_x and C_y) is defined by the Euclidean metric $d_i = \sqrt{(C_x - L_{i,x})^2 + (C_y - L_{i,y})^2}$ (see Figure). The relative distance of link L_i yields values between 0 and 1, where relative distance to the cursor is its distance to the cursor divided by the sum of the distances of all links to the cursor. The greater the distances of the cursor to the other links, the closer the relative distance of L_i is to 0. However, since the end result requires links closer to the cursor to have higher values, the program takes the difference of 1 and the relative distance.

This function, $\delta_t(L_i)$ at measurement t for link L_i is represented as

$$\delta_t(L_i) = \frac{\sum\limits_{j \in \mathcal{L} \setminus \{L_i\}} d_j}{\sum\limits_{j \in \mathcal{L}} d_j} \qquad \forall i, \forall t.$$
 (5)

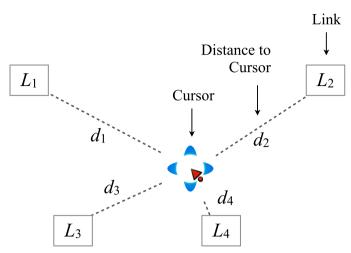


Figure 6. Cursor and Hyperlinks. The membership degree of a hyperlink in the fuzzy set depends on its relative distance from the cursor

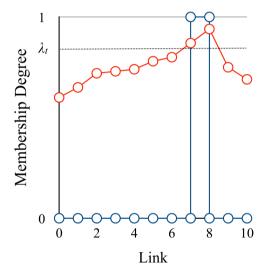


Figure 7. Illustration of λ -Cut. FLYAWAY defuzzifies the fuzzy set using a modified λ -cut in which the value of lambda is dynamic

4.3.3 Updating Membership Degree Using Exponential Smoothing

Every new gaze estimation point creates a new fuzzy set. Some weight should be given to previous gaze estimations to monitor and verify user intent. Considering only the most recent gaze estimation was not a desirable option because of, for example, inaccurate measurement and because users may simply be scanning all available hyperlinks without intending to click one. The option existed to use the union or intersection of the most recent fuzzy sets, but neither alternative was viable. The union would have the maximum membership degree among all the measured fuzzy sets for each element, which would lead to undesired membership degrees. Likewise, the intersection would take the minimum, and the minimum values are not important to determining the user intent. Instead, exponential smoothing is used to give a weight ω to the most recent measurement, but still some positive weight $1-\omega$ to the previous membership degree. Let $\mu_{\widetilde{\mathcal{F}}_i}(L_i)$ be the membership degree of link L_i in the fuzzy set $\widetilde{\mathcal{F}}_i$. Then,

$$\mu_{\tilde{\mathcal{F}}_{t}}(L_{i}) = (\omega)\delta_{t}(L_{i}) + (1 - \omega)\mu_{\tilde{\mathcal{F}}_{t-1}}(L_{i}) \qquad \forall i, \forall t.$$
 (6)

Using exponential smoothing allows FLYAWAY to store only the current measurement and the membership degree for the previous reading. However, this membership degree has embedded in it *all* of the previous membership degrees. In this fashion, the memory requirements are minimized and the "curse of dimensionality" is mitigated.

4.3.4 Defuzzification Using λ -Cuts

Although the system uses fuzzy sets to model impreciseness, the user's intended hyperlink in the end has to be determined and clicked. Thus, the fuzzy sets are *defuzzified*. To determine if the user intends to select a hyperlink, a λ -cut for defuzzification is used (for example, see Ross, 2009). This work develops and implements a *conditional* λ -cut. Before discussing the conditional λ -cut, consider the regular λ -cut. Let the defuzzified set \mathcal{F}_t^{λ} (which is crisp) be the λ -cut set of $\tilde{\mathcal{F}}_t$. \mathcal{F}_t^{λ} is defined as $\mathcal{F}_t^{\lambda} = \{L \in \mathcal{L} : \mu_{\tilde{\mathcal{F}}_t}(L) \geq \lambda\}$. In other words, any element $L \in \tilde{\mathcal{F}}_t$ with a degree of membership that is greater than or equal to λ belongs to \mathcal{F}_t^{λ} .

Figure graphically illustrates a λ -cut. This system uses λ -cuts to separate the hyperlinks with a high enough membership to be clicked from the others. When the attention level exceeds a threshold θ , a λ -cut is performed to extract the hyperlinks with the top values. However, to prevent $|\mathcal{F}_t^{\lambda}|$ from becoming unmanageable, this system defines a conditional λ -cut that changes the value of λ depending on the situation. This value at time t is referred to as λ_t . Let κ and Δ be constants. Let μ_{max} be the maximum membership value in $\tilde{\mathcal{F}}_t$. The value of λ_t depends on μ_{max} :

$$\lambda_t = \begin{cases} \mu_{max} - \Delta & \text{if } \mu_{max} \ge \kappa \\ \kappa & \text{otherwise.} \end{cases}$$
 (7)

In Figure , links 7 and 8 have membership degrees exceeding the cutoff λ_t and become members of the defuzzified set.

4.3.5 Link Selection and Target Expansion

The program determines which hyperlink to click using a conditional λ -cut, as described in Section 4.3.4. If $|\mathcal{F}_t^{\lambda}| = 0$, then no link is clicked. If $|\mathcal{F}_t^{\lambda}| = 1$, the JavaScript functions getElementById() and click() are used to simulate the click of the appropriate hyperlink. Note that the use of λ -cuts allows for a click to be made as long as the membership degree is high enough and does not require the cursor to go exactly over the target, similar to (Grossman, 2005). When $|\mathcal{F}_t^{\lambda}| > 1$, there are multiple links that are close together and it is difficult to determine the one in which the user is interested. In this case, following previous research (Hwang et al., 2013), a rectangle around the identified links is magnified. Membership degrees of all links are reset. This strategy allows a user to have more fine-grained and precise control over their choice. Target expansion uses the Objective-C UIScrollView method zoomToRect.

4.3.6 Error Recovery Mechanism

In case of a mistake, the system has an error recovery mechanism. Two options were considered: (i) the user closes his/her eyes for a certain period of time, e.g. five seconds; and (ii) a back arrow. Since a quadriplegic may not have full voluntary control of his/her eyelids, the system implements a back button presented in the top left corner of the interface (Figure). Also, since user intent would be clear in the case of error recovery, membership values are overridden for gaze estimations within a fixed distance of the back button.

5. Experimental Design and Results

This section describes the experimental evaluation of FLYAWAY. The experiment investigated how accurately a cross-section of individuals could navigate a webpage using this system. Following previous researchers investigating the performance of non-traditional assistive technologies (for example, Randolph & Moore Jackson, 2010), this research studied individuals' initial use of FLYAWAY with just a short introduction. Training effects or learning through experience due to long-term use of the system were not incorporated in this study. Approval of the study was obtained from the Intel International Science and Engineering Fair (ISEF) Scientific Review Committee and affiliated Institutional Review Board.

5.1 Interface Design

Figure shows the interface of this application with two example webpages loaded, and major elements labeled.

5.2 Participants

Participants in this study were at least 18 years old and included students, school teachers, librarians, university professors, working professionals, retired individuals, and homemakers. Their educational qualifications ranged from high school graduate to PhD. No compensation was offered for participation in the system testing. The sample consisted of 31 individuals to evaluate the system. The sample characteristics are as follows. All participants were non-disabled; while it would have been preferable to conduct an evaluation with quadriplegic individuals, a representative sample could not be obtained. 68% of the individuals were women and 32% were men. 87% of the sample had used computers for more than 10 years, but only 29% of the sample had used a mobile device

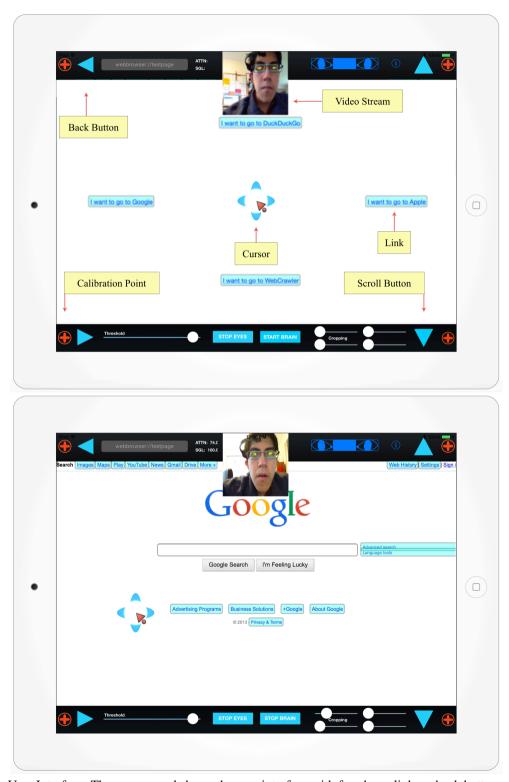


Figure 8. User Interface. The upper panel shows the user interface with four hyperlinks, a back button, and scroll buttons. The lower panel shows the interface after the link to the Google homepage is clicked



Figure 9. Data Collection. Two participants from the study using FLYAWAY

(smartphone or tablet) for more than four years. 19% reported using the Internet on mobile devices for more than one hour per day, in comparison to 68% who reported using the Internet on computers for more than one hour per day. Finally, 94% reported using keyboard and mouse/trackpad as the primary input method while using computers, and 81% reported using touch as the primary input method while using mobile devices. Two studies were conducted using this sample.

5.3 Effectiveness of First-Attempt

The first study evaluated the first-attempt accuracy in the system; i.e., how accurate the system is in allowing the user to navigate to the intended webpage or to scroll on the first try.

5.3.1 Stimuli and Procedures

Each respondent was first told the motivation for developing FLYAWAY and was then asked to sign a consent form. He/she was then asked to complete a brief questionnaire requesting demographic and computer/internet usage information. The consent form and questionnaire are available upon request. Subsequently, the EEG headset was positioned on the participant, and the Bluetooth connection and headset measurements were established and verified by asking the respondent to concentrate and to relax.

Next, the participant was asked to rest his/her chin on a chin rest to keep the head steady. Figure shows the setup for two study participants using FLYAWAY. After calibration, the participant was asked to select any one of the four links on a test webpage, shown in the upper panel of Figure . This process was repeated four times. Next, each participant was asked to look at the up and down scrolling buttons, in any order. This resulted in six observations per individual, for a total of 186 observations.

Before presenting the results, it would be useful to note that the accuracy of the system can depend on the complexity of the webpage. While the system was tested on a static webpage as described above, some websites include dynamic content in the same spatial location of the webpage. In such cases, it is important to click at just the right moment to navigate to the desired webpage. Such complex webpages were not evaluated using FLYAWAY in this study, but that can be a topic for future research.

5.3.2 Results and Analysis

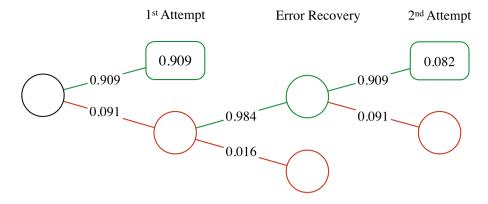
Let p denote the proportion of times for which this system accurately navigates to the intended page on the first attempt. In 169 out of the 186 cases during the data collection, the system navigated to the intended link. The point estimate, \hat{p} of the first attempt success rate is therefore $\hat{p} = 169/186 = 0.909$.

Next, a confidence interval was developed for the success proportion of the first attempt. A sample size check showed that the study has a large enough sample size (see, for example, Keller, 2014):

$$n \cdot \hat{p} = 186 \cdot 0.909 = 169 > 9$$
 and $n \cdot (1 - \hat{p}) = 186 \cdot (1 - 0.909) = 17 > 9$.

The standard error is 0.021 and the lower and upper limits of the 95% confidence interval for p are 0.87 and 0.95, respectively.

Therefore, the data provides evidence with 95% confidence that the first attempt success rate lies between 87% and 95%.



Second attempt success rate = 0.909 + 0.082 = 0.991

Figure 10. Decision Tree. Analysis of success rate with error recovery and second attempt

A reasonable hypothesis for this situation is $H_0: p \le 0.70$; $H_a: p > 0.70$. The reason for this hypothesis stems from (Kübler et al., 2009; see also Randolph & Moore Jackson, 2010). The data in the sample provides sufficient evidence to reject the null hypothesis with an alpha of 0.01 (p-value < 0.001).

5.4 Effectiveness of Error Recovery

The second study investigated the effectiveness of the error recovery mechanism of this system.

5.4.1 Stimuli and Procedures

This study was conducted immediately following Study 1 for each respondent. The respondent was told that the objective of this part of the data collection was to check the effectiveness of error recovery in case of incorrect navigation. The respondent was asked to click the back button four times, resulting in a total of 124 observations.

5.4.2 Results and Analysis

Let p_e denote the proportion and \hat{p}_e denote the sample proportion of correct error recovery selection. Since 122 out of the 124 observations resulted in correct error recovery, the point estimate of p_e is $\hat{p}_e = 122/124 = 0.984$.

5.5 Success Rate with Second Attempt

The error recovery success rate was used in conjunction with the first attempt rate to calculate the success proportion with a second attempt. Figure graphically shows how this computation was done. Since selection and error recovery are independent events, the probability of failure on the first attempt, success in error recovery and success in the second attempt were multiplied to obtain 0.082, which is the joint probability of these three events. Consequently, the success rate with a second attempt is 0.909 + 0.082, which equals 0.991.

5.5.1 Determining the Confidence Interval for Success Rate with Second Attempt

In Section 5.3, determining the confidence interval was straightforward. However, determining the confidence interval for the success rate with the second attempt is complicated because it involves three random quantities. Since no closed form expression could be located, a Monte-Carlo simulation was performed with 100 trials to construct the confidence interval. For each trial, a random number for the first attempt success rate was generated from a normal distribution with a mean of 0.909 and a standard deviation of 0.021. Denote this value as X_1 . Since the error recovery percentage is already very high (0.986), this value was not simulated. Denote this value as Y. Finally, a second stage success rate was simulated assuming that the first stage resulted in an incorrect choice. To do this, a random number X_2 was generated from the same distribution as that of X_1 . The second attempt success rate for this trial was computed as $X_1 + (1 - X_1) * Y * X_2$. Next, the mean and the standard deviation of the 100 such values were computed, which were used to determine the 95% confidence interval for the success rate with a second attempt. Based on the simulation, it was concluded that the upper and lower limits of the 95% confidence interval for the success rate with a second attempt is are 0.9845 and 0.9966, respectively.

6. Conclusion and Future Directions

This paper describes the development of a low-cost and accurate web browsing system targeted to help individuals

with motor impairments. The system utilizes multimodal control by capturing eye movements and brain attention levels. These signals are used as inputs to initiate and continuously update the estimated desirability of links on a webpage. A novel fuzzy logic algorithm helps accomplish this task. While the system meets the current objectives, there are several directions that can be explored to improve the system.

Three criteria that are important in designing a system of this type are speed, accuracy, and cost. For a given cost, a trade-off has to be made between speed and accuracy. Although the iPad (third generation) camera has a high resolution capability, a low resolution input is used to improve image processing efficiency. It would be important to see if these algorithms can be fine-tuned so that they can process higher-quality images in the same amount of time. Of course, improvements in the processing power of mobile devices provide opportunities to extend the frontier at which this trade-off needs to be made. Potential algorithmic improvements to explore include (i) determining a function for κ based on number of links $|\mathcal{L}|$, (ii) implementing alternate defuzzification mechanisms, and (iii) further fine-tuning the attention level threshold θ , exponential smoothing constant ω , and conditional λ -cut constants κ and Δ .

Alternate methods of user control that hold potential for specific abilities include eye-mouth control, cheek-eyebrow control, and closing eyes for error recovery. A dynamic interaction interface centered on *ability* rather than disability, termed as *ability-based design* (Wobbrock et al., 2011), would automatically adapt to a wide variety of users.

While the system is currently implemented for the iPad, the image processing algorithm and fuzzy logic algorithm can be ported to other mobile tablet operating systems. Developing a cross-platform system could increase its versatility.

FLYAWAY could also be expanded to a computer-wide control system. On a longer-term basis, this approach could be used to exploit the "Internet of Things." Those facing motor-impairments could use multimodal control to use home appliances in customized settings for further improving their quality of life.

Acknowledgments

The author would like to acknowledge participants in this study as well as Professor Dennis Galletta, Ph.D., of the University of Pittsburgh for his support and comments on a previous version of this paper.

References

- Allison, B. Z., Leeb, R., Brunner, C., Müller-Putz, G. R., Bauernfeind, G., Kelly, J. W., & Neuper, C. (2012). Toward smarter BCIs: extending BCIs through hybridization and intelligent control. *Journal of Neural Engineering*, 9(1), 013001. http://dx.doi.org/10.1088/1741-2560/9/1/013001
- Berger, H. ber das Elektroenzephalogramm des Menschen I-XIV. Arch. Psychiat. Nervenkr. (1929-1938), (87-108).
- Bolt, R. A. (1980). "Put-that-there": Voice and gesture at the graphics interface. ACM SIGGRAPH Computer Graphics, 14(3), 262-270. http://dx.doi.org/10.1145/965105.807503
- Cacioppo, J. T., & Cacioppo, S. (2014). Social relationships and health: The toxic effects of perceived social isolation. *Social and personality psychology compass*, 8(2), 58-72. http://dx.doi.org/10.1111/spc3.12087
- Cambria, E., & White, B. (2014). Jumping NLP curves: a review of natural language processing research [review article]. *IEEE Computational Intelligence Magazine*, 9(2), 48-57. http://dx.doi.org/10.1109/mci.2014.2307227
- Castrillón, M., Dniz, O., Guerra, C., & Hernández, M. (2007). ENCARA2: Real-time detection of multiple faces at different resolutions in video streams. *Journal of Visual Communication and Image Representation*, 18(2), 130-140. http://dx.doi.org/10.1016/j.jvcir.2006.11.004
- EyeWriter. (2010). The EyeWriter. Retrieved August 10, 2015, from http://www.eyewriter.org/.
- Fukuda, T., Morimoto, K., & Yamana, H. (2011). Model-based eye-tracking method for low-resolution eye-images. In 2nd Workshop on Eye Gaze in Intelligent Human Machine Interaction.
- Georgieva, P., Silva, F., Milanova, M., & Kasabov, N. (2014). EEG Signal Processing for Brain-Computer Interfaces. In *Springer Handbook of Bio-/Neuroinformatics* (pp. 797-812). Springer Berlin Heidelberg. http://dx.doi.org/10.1007/978-3-642-30574-0_46
- Grossman, T., & Balakrishnan, R. (2005). The bubble cursor: enhancing target acquisition by dynamic resizing of the cursor's activation area. In Proceedings of the SIGCHI conference on Human factors in computing

- systems (pp. 281-290). ACM. http://dx.doi.org/10.1145/1054972.1055012
- Goldberg, A. (2012). The Feasibility of Consumer Grade EEG Devices as Computer Input. *Proceedings of The National Conference On Undergraduate Research (NCUR) 2012*.
- 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. http://dx.doi.org/10.1109/tpami.2009.30
- Hinckley, K., & Wigdor, D. (2002). Input technologies and techniques. *The human-computer interaction hand-book: fundamentals, evolving technologies and emerging applications*, 151-168. http://dx.doi.org/10.1201/b10368-12
- Hwang, F., Hollinworth, N., & Williams, N. (2013). Effects of target expansion on selection performance in older computer users. *ACM Transactions on Accessible Computing*, 5(1), 1. http://dx.doi.org/10.1145/2531922.2514848
- Kaehler, A., & Bradski, G. (2014). Learning OpenCV (2nd ed.). O'Reilly Media.
- Keller, G. (2014). Statistics for Management and Economics (10th ed.). Cengage Learning.
- Kübler, A., Furdea, A., Halder, S., Hammer, E. M., Nijboer, F., & Kotchoubey, B. (2009). A Brain-Computer Interface Controlled Auditory Event-Related Potential (P300) Spelling System for Locked-In Patients. *Annals of the New York Academy of Sciences*, 1157(1), 90-100. http://dx.doi.org/10.1111/j.1749-6632.2008.04122.x
- Leeb, R., Sagha, H., Chavarriaga, R., & del R Millan, J. (2010). Multimodal Fusion of Muscle and Brain Signals for a Hybrid-BCI. 2010 Annual International Conference of the IEEE Engineering in Medicine and Biology. http://dx.doi.org/10.1109/iembs.2010.5626233
- Li, D., & Parkhurst, D. J. (2006). openEyes: an open-hardware open-source system for low-cost eye tracking. *Journal of Modern Optics*, 53(9), 1295-1311. http://dx.doi.org/10.1080/09500340600619072
- McMullen, D. P., Hotson, G., Katyal, K. D., Wester, B. A., Fifer, M. S., McGee, T. G., ... & Crone, N. E. (2014). Demonstration of a Semi-Autonomous Hybrid Brain-Machine Interface Using Human Intracranial EEG, Eye Tracking, and Computer Vision to Control a Robotic Upper Limb Prosthetic. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 22(4), 784-796. http://dx.doi.org/10.1109/tnsre.2013.2294685
- Müller-Putz, G. R. Breitwieser, C., Cincotti, F., Leeb, R., Schreuder, M., Leotta, F., ... & Milln, J. D. R. (2011). Tools for brain-computer interaction: a general concept for a hybrid BCI. *Frontiers in Neuroinformatics*, 5. http://dx.doi.org/10.3389/fninf.2011.00030
- Oviatt, S. (2012). Multimodal interfaces. *The human-computer interaction handbook: Fundamentals, evolving technologies and emerging applications, 14*, 405-430.
- Papadelis, C., Braun, C., Pantazis, D., Soekadar, S. R., & Bamidis, P. (2013). Using Brain Waves to Control Computers and Machines. *Advances in Human-Computer Interaction* 2013, 1-2. http://dx.doi.org/10.1155/2013/802063
- Pfurtscheller, G., Allison, B. Z., Brunner, C., Bauernfeind, G., Solis-Escalante, T., Scherer, R., ... & Birbaumer, N. (2010). The hybrid BCI. *Frontiers in Neuroscience*, 4(42). http://dx.doi.org/10.3389/fnpro.2010.00003
- Randolph, A. B., & Moore Jackson, M. M. (2010). Assessing fit of nontraditional assistive technologies. *ACM Transactions on Accessible Computing*, 2(4), 16:1-16:31. http://dx.doi.org/10.1145/1786774.1786777
- Ross, T. J. (2010). Fuzzy logic with engineering applications. http://dx.doi.org/10.1002/9781119994374
- Thiran, J.-P., Marqus, F., & Bourlard, Hervé (2009). *Multimodal Signal Processing: Theory and applications for human-computer interaction*. Academic Press.
- Tobii Technology. (2013). Access your computer through eye gaze. Retrieved February 21, 2014, from http://www.tobii.com/assistive-technology/north-america/products/hardware/pceye/
- Wilson, B. A., Hinchcliffe, A., Okines, T., Florschutz, G., & Fish, J. (2011). A case study of Locked-In-Syndrome: Psychological and personal perspectives. *Brain Injury*, 25(5), 526-538. http://dx.doi.org/10.3109/02699052.2011.568034
- Wobbrock, J. O., Kane, S. K., Gajos, K. Z., Harada, S., & Froehlich, J. (2011). Ability-based design: Concept, principles and examples. *ACM Transactions on Accessible Computing*, 3(3), 9.

- http://dx.doi.org/10.1145/1952383.1952384
- Wolpaw, J. R., Birbaumer, N., McFarland, D. J., Pfurtscheller, G., & Vaughan, T. M. (2002). Brain-computer interfaces for communication and control. *Clinical Neurophysiology*, 113(6), 767-791. http://dx.doi.org/10.1016/s1388-2457(02)00057-3
- Yang, J., Stiefelhagen, R., Meier, U., & Waibel, A. (1998). Real-time face and facial feature tracking and applications. In AVSP'98 International Conference on Auditory-Visual Speech Processing.
- Zadeh, L. A. (1965). Fuzzy sets. *Information and Control*, 8(3), 338-353. http://dx.doi.org/10.1016/s0019-9958(65)90241-x
- Zander, T. O., Gaertner, M., Kothe, C., & Vilimek, R. (2010). Combining eye gaze input with a brain-computer interface for touchless human-computer interaction. *International Journal of Human-Computer Interaction*, 27(1), 38-51. http://dx.doi.org/10.1080/10447318.2011.535752
- Zimmermann, H. J. (2010). Fuzzy set theory. *Wiley Interdisciplinary Reviews: Computational Statistics*, 2(3), 317-332. http://dx.doi.org/10.1002/wics.82

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/3.0/).