

Integrated electromyogram and eye-gaze tracking cursor control system for computer users with motor disabilities

Craig A. Chin, PhD;^{1*} Armando Barreto, PhD;¹ J. Gualberto Cremades, EdD, PhD, CC-AASP;² Malek Adjouadi, PhD¹

¹Department of Electrical and Computer Engineering, College of Engineering and Computing, Florida International University, Miami, FL; ²Department of Sport and Exercise Sciences, Barry University, Miami Shores, FL

Abstract—This research pursued the conceptualization, implementation, and testing of a system that allows for computer cursor control without requiring hand movement. The target user group for this system are individuals who are unable to use their hands because of spinal dysfunction or other afflictions. The system inputs consisted of electromyogram (EMG) signals from muscles in the face and point-of-gaze coordinates produced by an eye-gaze tracking (EGT) system. Each input was processed by an algorithm that produced its own cursor update information. These algorithm outputs were fused to produce an effective and efficient cursor control. Experiments were conducted to compare the performance of EMG/EGT, EGT-only, and mouse cursor controls. The experiments revealed that, although EMG/EGT control was slower than EGT-only and mouse control, it effectively controlled the cursor without a spatial accuracy limitation and also facilitated a reliable click operation.

Key words: assistive technology, cursor control, electromyogram, eye-gaze tracking, mean power frequency, motor disabilities, multimodal cursor control, point of gaze, spectral analysis, universal access.

INTRODUCTION

Typically, nondisabled individuals communicate with a computer using standard input devices such as a mouse, trackball, touch pad, or keyboard. An estimated 250,000 to 400,000 individuals in the United States and more than 2 million people worldwide live with spinal cord injury or spinal dysfunction [1–2], and many of these individuals

are unable to use standard input devices. Given the increasing pervasiveness of computer-based systems in most of our daily activities and the increasing levels of communication and social participation that occur over the Internet, clearly, facilitating access of these individuals to graphical user interface (GUI)-driven computer systems is an important technical goal.

With today's GUI-based personal computer software, most of the human-to-computer interaction is based on selection operations, which consist of two steps:

- Pointing: Positioning the cursor at the desired location on the screen, over the appropriate area or icon.
- Clicking: Executing the mouse down-up function that is interpreted by the computer's operating system as an indication to complete the selection of the item associated with the icon at the location of the screen cursor.

Abbreviations: 2-D = two-dimensional, ANOVA = analysis of variance, EEG = electroencephalogram, EGT = eye-gaze tracking, EMG = electromyogram, FMMNN = fuzzy min-max neural network, GUI = graphical user interface, MPF = mean power frequency, POG = point of gaze, PSD = power spectral density, SD = standard deviation.

*Address all correspondence to Craig A. Chin, PhD; Department of Electrical and Computer Engineering, College of Engineering and Computing, Florida International University, 10555 West Flagler St, Miami, FL 33174; 305-348-3711. Email: cr_chin@hotmail.com
DOI: 10.1682/JRRD.2007.03.0050

Considering the previous paragraphs, our group sought to create a hands-free cursor control system that would empower individuals who cannot use standard input devices to perform point-and-click operations. This system would enable a user to perform these cursor control operations by identifying patterns in electromyogram (EMG) signals, which are associated with predefined facial movements of the user, and in eye-gaze tracking (EGT) paths, associated with the user's eye movements.

Electromyography is the study of muscle function through the monitoring of electrical signals emitted by muscles [3]. When a surface electrode is placed on the skin above a superficial muscle while it is contracting, it will receive electrical signals emanating from several muscle fibers associated with different motor units. The spatiotemporal summation of these electrical signals results in an EMG signal. Therefore, the EMG signal effectively monitors muscle activity.

EMG-based cursor control systems typically monitor EMG signals from a targeted set of superficial muscles, which are associated with a group of movements that the user can still perform. A number of algorithms can be used to recognize the EMG patterns associated with each movement so as to produce the associated cursor action. Some examples are given in the following paragraphs.

Chang et al. designed a real-time EMG discrimination system in which five distinct motions of the neck and shoulders were used to produce five commands [4]. They accomplished real-time discrimination by using the cepstral coefficients of the input EMG signals as feature inputs to a modified maximum likelihood distance classifier. A 95 percent recognition rate and a response time of less than 0.17 s were achieved for the six subjects tested.

Barreto et al. created a real-time system that used EMG signals from cranial muscles and electroencephalogram (EEG) biosignals from the cerebrum's occipital lobe to control the two-dimensional (2-D) movement of the cursor, perform left-clicks, and switch the cursor control function on and off [5]. The system performed periodogram estimations of the power spectral density (PSD) of the EMG signals over discrete windows. They classified these spectral data by considering amplitude thresholds to determine the onset of a contraction and then using spectral power summations aggregated over specific frequency bands between 8 and 500 Hz to determine which muscle had contracted. The results of point-and-click tests revealed that, although this form of EMG control was effective, its average task completion time was slow (16.3 s) compared with that of a mouse (1–2 s).

Kim et al. introduced an EMG system for cursor control that interpreted six predefined wrist motions into the cursor actions: left, right, up, down, click, and rest. A fuzzy min-max (minimum-maximum) neural network (FMMNN) was used as a classifier [6]. Difference absolute mean values were extracted from the EMG signals and used as training features in the FMMNN. The recognition rate obtained was 97 percent for the 10 people they used to test the system.

The development of the system reported here is based on modifying and augmenting the EMG cursor control system that Barreto et al. created [5]. The advantages of EMG-based cursor control are that it provides the user with the ability to perform small discrete cursor movements, and it possesses a robust and stable "clicking" procedure. However, as mentioned previously, this system performs slowly compared with a mouse-operated system in point-and-click tests and could potentially become tiresome if the user is required to make large excursions across the screen.

An alternative form of cursor control is EGT. It uses the user's gaze direction to determine the position of the cursor. To better understand this technology and its limitations, one needs to consider the physiological aspects of eye movements. The general mechanism used by the eyes to examine a visual scene consists of two types of eye movements: the saccade and the fixation. A saccade is a ballistic motion that moves the eye from one area of focus of the visual scene to another. After a saccade, a period of relative stability follows. This period is called a fixation, and it allows the eye to focus light on an area of the retina called the fovea. During a fixation, the eyes still exhibit small, jittery motions, usually less than 1° visual angle in size [7].

EGT techniques determine the user's visual line of gaze by taking video images of the eye to establish a relationship between the geometric orientation of specific features in the eye image and the line of gaze. The most popular EGT technique today uses the relative position of the bright eye (pupil) center and the center of the glint (corneal reflection) to determine the line of gaze [7–11]. Once the line of gaze is determined, the point of gaze (POG) is found by allowing the line of gaze to intersect with the plane of the scene being viewed (typically the computer screen). The mapping between screen coordinates and eye-gaze direction is determined by a calibration procedure, performed in advance of the use of the EGT system. The "raw" POG coordinates produced by the EGT system are generally processed further by a fixation identification algorithm, which will extract fixation coordinates from the POG coordinates.

These fixation coordinates are then used to update the cursor position. Selections or left-clicks may be implemented with a time threshold placed on the eye-gaze dwelling within a small area of the screen, which will result in a left-click operation being issued if this threshold is exceeded. Eye blinks may also be used to produce left-click operations. Dwell time is more natural to the user [7] and thus is more prevalent in its usage.

A seminal work in the field of EGT-based control of the cursor is that of Ware and Mikaelian [8]. The EGT technique that they used required a dwell time of 400 ms. In their article, they presented experiments to investigate the viability of EGT as a pointing technique. The results showed task times of less than 1 s and that task time and error rate increased significantly for target sizes less than 1.5° visual angle.

Hutchinson et al. described an Eye-Gaze Response Interface Computer Aid in their article [9]. The EGT-based cursor control system used a 2 to 3 s dwell time as a selection criterion, and the testing of their system produced some notable observations: The bright eye effect, used to define the center of the pupil reflection, was not observable in 5 to 10 percent of the candidates; the head must remain fairly stationary for the eye image to be captured; and the accuracy of the system was limited.

The fixation identification algorithm used by Robert J. K. Jacob in his eye-tracking cursor control technique used a 100 ms temporal threshold to determine whether the POG points remained within a 0.5° dispersion threshold. In a preliminary evaluation, his eye-tracking technique was used to perform object selection interactions with a dwell time of 150 to 250 ms and was found to be quite effective in performing these tasks [7].

More recently, Sibert, along with Jacob and Templeman, provided a more formal evaluation of Jacob's EGT system [10–11]. The evaluation consisted of two experiments that required participants to select circular targets with the EGT system and with the mouse. The EGT system used a dwell time of 150 ms as a selection criterion. The observation of mean selection time in both experiments showed that the EGT system was faster than the mouse and that the difference was statistically significant.

The primary advantage of EGT systems, as shown by these researchers, is that they perform faster than a mouse in point-and-click tests. However, this approach has some disadvantages. One is the so-called "Midas Touch" problem [7]. This problem originates when eye-gaze dwell time is used to issue the left-click operation. During a human-computer interaction session, situations may arise

where a user may only desire to stare at an object to examine it, rather than to select it. If a user is using an EGT system with dwell time-enabled left-clicks, unintended selections will result when the user examines this object for a period that exceeds the dwell time threshold. Another disadvantage is the limited accuracy of the approach. This limitation occurs because the eye only needs to focus incoming light anywhere in the fovea to achieve the higher level of visual acuity available in that region of the retina. However, this requirement still allows variations of about 1° of visual arc in the direction of gaze [7]. The lax nature of this physical constraint limits the accuracy with which the line of gaze can be estimated. Furthermore, if the small jittery motions exhibited by the eye during a fixation were directly translated into cursor movements by an EGT-based system, this would severely deteriorate the computer cursor's stability. Another issue is that POG offsets may occur after the original calibration of the EGT system. These offsets are caused by minor movements of the head from its original calibration position. Morimoto and Mimica have shown experimentally that the calibration mapping of a remote EGT system decays (becomes less accurate) as the head moves away from its original position [12]. Therefore, the only way to restore the accuracy of the EGT system is either to shift the head of the user back to its original position or to recalibrate the system at the present position of the user's head.

The complementary strengths of EMG and EGT input modalities make them well-suited for integration into a more robust cursor control system that will provide computer access to individuals who are unable to use their hands. Therefore, this project pursued the creation of a bimodal cursor control system that will selectively use both types of input from the user to more efficiently manipulate the screen cursor under a wider range of circumstances. We envisioned such a system to require the user to coordinate his or her eye and facial movements in accordance with the following protocol, if the user desired to perform a point-and-click operation:

1. Regional placement of the cursor: The user will change his or her line of gaze so the icon to be selected will be focused on the fovea. The EGT subsystem will capture and process this action to update the cursor position to one that resides in the general area of the icon. If the updated cursor position coincides with the inner area of the icon, then the user may move directly to step 3. If this is not the case, then the user should proceed to step 2.

2. Refinement of cursor position: With the updated cursor position falling outside of the boundaries of the desired icon, the user must now refine the cursor position by using facial movements to move the cursor up, down, left, or right, while still maintaining his or her gaze on the icon. The EMG subsystem will translate the facial movements into the corresponding cursor movements, while having the user maintain his or her eye gaze on the icon will ensure that the EGT subsystem will not update the cursor position to one that resides outside of the boundaries of the icon. Even though the requirement to maintain eye gaze on the icon may seem like a constraint to the user, one should bear in mind that this would be a natural consequence of user attention on the icon.
3. Icon selection: Once the cursor is located within the boundaries of the icon, the user may select it via a specific facial movement. The EMG subsystem will detect this specific facial movement and translate it into the desired left-click operation.

With this protocol, the cursor stability and clicking reliability observed in the evaluation of the EMG subsystem will be inherited by the hybrid system. On the other hand, when the user needs to perform a long cursor displacement on the screen, the output of the EGT subsystem will be used to define the new cursor position. This alteration of control modalities will allow the user to take advantage of the speed achieved by using EGT-based systems.

One should note that other work has integrated non-standard forms of computer input to create enhanced user interaction. One approach is to use a gaze and speech multimedia interface. The protocol for such an input would involve using the EGT input to locate the object to be manipulated and then using speech commands to initiate an icon manipulation procedure (e.g., click, move, or drag-drop). Optimal synchronization of the gaze and speech input streams have been investigated by Kaur et al. [13]. In addition, Zhang et al. have examined the effectiveness of one-, two-, and three-word phrases along with the optimal radius to use for the EGT operative region [14]. These investigations showed that gaze-speech interfaces can overcome the susceptibility of gaze-based interaction to unintended selections, as well as improve the accuracy and speed of speech-recognition systems, by allowing for simpler vocabularies. However, these interfaces did not improve the limited spatial accuracy inherent in EGT systems.

Trejo et al. have performed some important work in integrating EMG and EEG signal inputs for human-computer interaction [15]. They have successfully completed EMG-based interfaces to control a virtual flight stick and to monitor user typing. They have also created an EEG-based interface that performs 1-D control of the mouse. While an EMG/EEG input may be alternatively useful in the long run, the slow speed of operation of current EEG-based forms of cursor control render the EMG/EEG input less usable than an EMG/EGT approach at present [16].

Recently, Surakka et al. have explored the combined use of EGT and EMG for cursor control [17]. However, their system uses voluntary gaze direction (EGT) to perform object pointing and direct thresholding of a single (bipolar) EMG channel to command a click. The EMG signal is obtained from electrodes on the forehead that detect when the subject contracts the corrugator supercilii muscle by frowning. Since their system uses a single EMG signal exclusively for clicking, it clearly does not attempt to alleviate the lack of pointing accuracy in the EGT system. Their analysis revealed that the mouse was faster than the new input in performing object pointing and selecting over short distances. However, the regression slopes derived from Fitts' law analysis suggest that the input may be faster than the mouse over long distances, that is, beyond 800 pixels.

METHODS

An analysis of the operational requirements of this integrated EMG/EGT cursor control system suggested that complete functionality and effective operation of the system necessitated the performance of three key tasks in a continuous fashion:

1. Reliable EMG input assessment: Muscular contractions must be correctly identified.
2. Reliable EGT fixation estimation: EGT fixations must be properly localized when they occur.
3. Reliable estimation: The user's intent must be reliably estimated for cursor manipulation and the resulting effective cursor position must be updated in the GUI.

These tasks and their interrelation are described in **Figure 1**. The remainder of this section describes the implementation of the EMG/EGT system according to this task categorization and also details the experimental design and analysis procedures used to evaluate the performance of the EMG/EGT system.

EMG Subsystem Implementation

Placement of Electrodes

Figure 2 displays the placement of four silver/silver chloride electrodes that are applied to the head of the user to capture the EMG signals. The figure indicates that electrodes were placed over the right frontalis muscle, the left temporalis muscle, the right temporalis muscle, and the procerus muscle, respectively. An electrode was placed over the right mastoid as a reference. Note that the frontalis and temporalis electrodes, as well as all the connecting wires, can be conveniently hidden and secured under a sports headband.

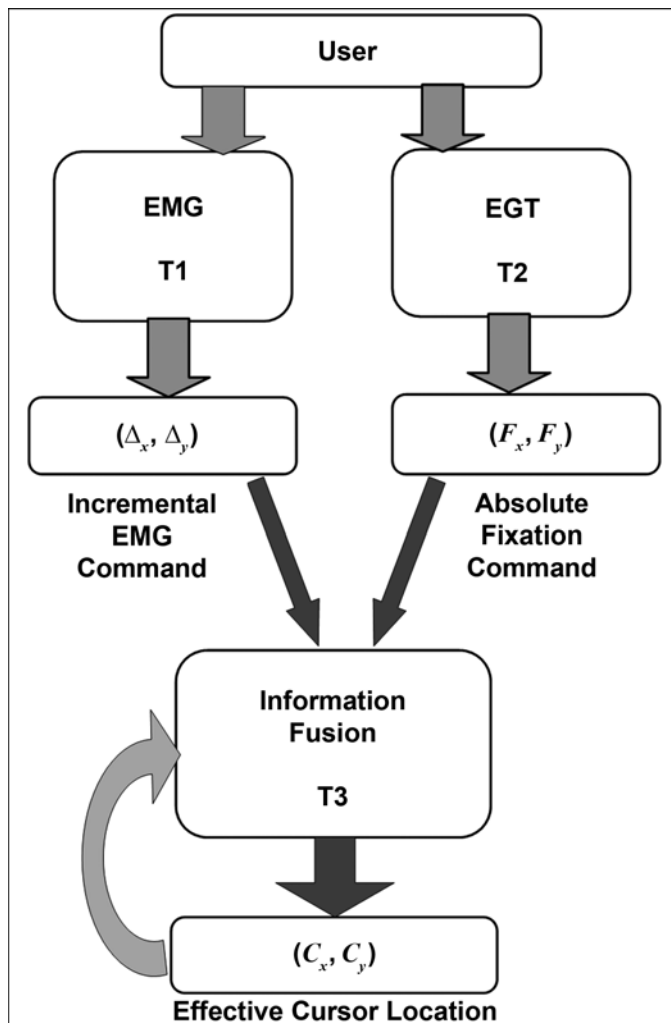


Figure 1.

Conceptual depiction of functionality of integrated EMG/EGT cursor control system, on basis of three key tasks (T1, T2, and T3). EGT = eye-gaze tracking, EMG = electromyogram.

Hardware Components of EMG Subsystem

The hardware components of the EMG subsystem are presented in **Figure 3**. The set of four EMG signals were inputted into Grass® P5 Series AC preamplifiers (Grass Technologies Product Group, Astro-Med Inc; West Warwick, Rhode Island). The ADC64™ DSP/AD board (a digital signal processing and analog-to-digital conversion board) (Innovative Integration; Simi Valley, California) performed analog-to-digital conversion on each signal at a sampling rate of 1.2 kHz and then applied the EMG classification algorithm to these digitized signals in real time. The board was connected to the computer's processor through the peripheral component interface bus. The output of the classification algorithm was sent to the host application via hardware interrupts. These interrupts occurred once every 213 ms (256 samples/1,200 Hz).

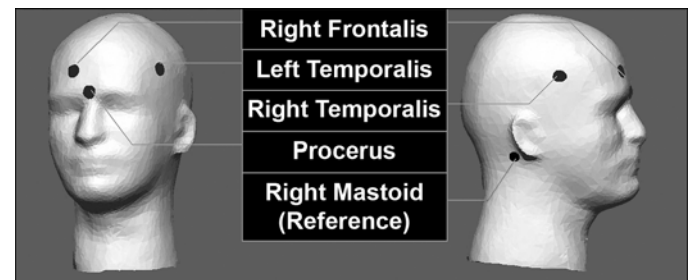


Figure 2.

Electrode muscle placement for electromyogram cursor control system.

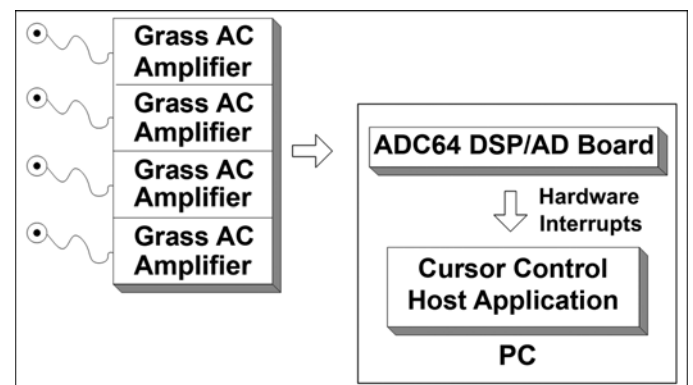


Figure 3.

Block diagram of hardware components of electromyogram subsystem, Grass P5 Series AC amplifiers (Grass Technologies Product Group, Astro-Med Inc; West Warwick, Rhode Island), and ADC64™ DSP/AD (digital signal processing and analog-to-digital) board (Innovative Integration; Simi Valley, California). PC = personal computer.

EMG Classification Algorithm for Muscle Contraction Identification

The desired relations between muscle contractions, facial movements, and cursor actions for the EMG subsystem are given in **Table 1**. The EMG classification algorithm determines whether a facial muscle contraction had occurred and, if so, which specific muscle had contracted. Given the correspondence between each muscle contraction (facial movement) and a cursor action shown in **Table 1**, the output of an effective muscle contraction classification algorithm can be used to provide real-time cursor control.

In spite of the intended one-to-one correspondence between EMG electrodes and muscles monitored, because of the volume conduction in the head, contraction of one muscle may cause significant EMG signals to appear in more than one electrode. Therefore, we used the spectral information in the various EMG channels to resolve this ambiguity.

Research had previously observed that the four muscles being monitored possessed distinct EMG frequency characteristics and that this frequency information would be useful for performing classifications [5]. Empirical observations suggested that mean power frequency (MPF) values would be suitable to represent the frequency data for this input configuration, and MPF values were used as feature inputs to the classification algorithm. The MPF is derived from PSD values, where a PSD plot describes the power distribution of a signal over a given frequency range. More specifically, the MPF is a weighted average frequency in which each frequency component f is weighted by its PSD value P . The equation for the calculation of the MPF is given by

$$\text{MPF} = \left(\frac{f_0 \times P_0 + \dots + f_k \times P_k + \dots + f_{N-1} \times P_{N-1}}{P_0 + \dots + P_k + \dots + P_{N-1}} \right), \quad (1)$$

where $k = 0, 1, 2, \dots, N - 1$. (A typical upper limit used is $N = 256$.)

EMG recordings, taken from a test group of five individuals, revealed that each muscle type had a characteristic range of MPF values: The frontalis muscle had the majority of its frequency content below 200 Hz, with an MPF in the range 40 to 165 Hz. The temporalis muscles had a significant portion of their frequency content above 200 Hz, with an MPF in the range 120 to 295 Hz. The procerus muscle had an intermediate frequency content when compared with the frontalis and temporalis muscles, with an MPF in the range 60 to 195 Hz.

Table 1.

Relationship between muscle contractions (facial movements) and resultant cursor actions.

Muscle Contraction	Facial Movement	Resultant Cursor Action
Left Temporalis	Left jaw clench	Left increment
Right Temporalis	Right jaw clench	Right increment
Right Frontalis	Eyebrows up	Up increment
Procerus	Eyebrows down	Down increment
Left and Right Temporalis	Left and right jaw clench	Left-click

The EMG classification algorithm derived three features from each PSD estimate, calculated for each EMG input, that helped determine which muscle(s) had contracted. These features were the maximum PSD magnitude, the sum of all PSD magnitudes for a given estimate, and the MPF value for the estimate. The EMG classification algorithm performed two types of decision processes: (1) for the detection of single muscle contractions and (2) for the detection of the simultaneous contraction of two muscles.

The cursor actions left, right, up, and down are produced by the predominant contraction of a single muscle (temporalis, frontalis, or procerus). For the algorithm to correctly identify this kind of contraction, a criterion placed on each feature calculated from the PSD estimate, for the electrode (muscle) in question, must be satisfied. These criteria are as follows:

1. The maximum PSD magnitude must exceed the threshold set for that electrode.
2. The sum of the PSD amplitudes for the given electrode must exceed the PSD sums of the other electrodes.
3. The MPF must fall into a range consistent with the muscle associated with the electrode.

The left-click cursor action required the simultaneous contraction of the left and right temporalis muscles. The criteria that must be satisfied for the correct classification of this simultaneous contraction are as follows:

1. The maximum PSD magnitude thresholds must be exceeded for both temporalis electrodes.
2. The PSD sums for both temporalis electrodes must be greater than the other two PSD sums.
3. The PSD sums for both temporalis electrodes must indicate a fairly balanced bilateral contraction, that is, each PSD sum must be greater than 20 percent of the total of both PSD sums.
4. The MPFs from both temporalis PSDs must fall into a range consistent with the temporalis muscle.

Barreto et al. found empirically that neck movements (flexion, extension, and rotation) would often cause unintended cursor actions (primarily left-click and up actions) to be issued if the classification algorithm were based solely on the PSD sum feature [5]. Therefore, an advantage of including MPF features in the analysis was that they made the output of the classification algorithm significantly less responsive to signals due to spurious neck movements.

In addition to 2-D directional control, the EMG classification algorithm also provided control of the speed of the cursor in the four directions specified in **Table 1**. The size of the increment that the cursor moved in either the horizontal or vertical direction could be increased if a contraction was maintained continuously for specific time periods. This relationship between contraction time (seconds) and increment size (pixels) for cursor speed control can be illustrated as—

- 0.213 to 0.640 s $\pm$ 1 pixel.
- 0.853 to 1.280 s $\pm$ 5 pixels.
- 1.493 to 3.413 s $\pm$ 10 pixels.
- >3.413 s $\pm$ 20 pixels.

EGT Subsystem Implementation

Hardware Components of EGT Subsystem

The eye-tracking system used for our EGT subsystem was an R6-HS Remote Optics system (Applied Science Laboratories; Bedford, Massachusetts). In this system, a beam from near-infrared light-emitting diodes, located on a pan/tilt optics module, illuminated the eye of the user. The eye image that this illumination produced was focused and sensed by a video camera also present on the pan/tilt unit. Video image data were fed into an eye tracker control unit that performed feature recognition and POG estimation. The POG estimates were transmitted, in real time, to the display computer (the computer that interacted with the user). The cursor control application running on the display computer received these estimates via hardware interrupts that occurred at a rate of 120 Hz.

Fixation Identification Algorithm

The algorithm determined whether a fixation had occurred by evaluating the input data on the basis of the criteria set by us. More specifically, the algorithm extracted a 100 ms moving window (temporal threshold) of consecutive POG data points (POG_x , POG_y) and calculated the standard deviation (SD) of the x - and y -coordinates

of these points. If both SD values were less than the coordinate thresholds associated with 0.5° of visual angle (spatial threshold), then the onset of a fixation had occurred and its horizontal coordinates (F_x) and vertical coordinates (F_y) would be determined by the centroid of the POG samples received during the 100 ms window analyzed. If a fixation was determined not to have occurred, then the window was advanced by one data point and fixation identification was attempted again. One should note that this algorithm was designed to accommodate blinks (loss of data) of up to 200 ms in duration.

Information Fusion and Cursor Update Algorithm

The information fusion and cursor update algorithm determined the effective cursor position as a merging of the incremental EMG commands (Δ_x, Δ_y) and the absolute coordinates of a qualified EGT fixation (F'_x, F'_y) by

$$C_x[n] = \begin{cases} C_x[n-1] + \Delta_x[n] & , \text{ if EMG update} \\ F'_x[n] & , \text{ if EGT update} \end{cases} \quad (2)$$

$$C_y[n] = \begin{cases} C_y[n-1] + \Delta_y[n] & , \text{ if EMG update} \\ F'_y[n] & , \text{ if EGT update} \end{cases} \quad (3)$$

where C_x and C_y = the x - and y -coordinates of a cursor position and n = a discrete index used to describe the progression of cursor updates through time. The merging of the outputs of the two subsystems implied that the current cursor position ($C_x[n]$, $C_y[n]$) could be updated by either the EMG or EGT subsystem at any time.

An EMG subsystem update involved changing the previous cursor position ($C_x[n-1]$, $C_y[n-1]$) by an increment of Δ_x or Δ_y . The direction of the increment, if any, was determined by the output value of the EMG subsystem.

The EGT subsystem determined a qualified fixation by taking every new fixation centroid (F_x, F_y) identified by the fixation identification algorithm and determining whether it signified a new point of user attention or if it simply was the continuation of previous fixation. To do this, the EGT subsystem measured the distance between the current qualified fixation position (F'_x, F'_y) and the new fixation centroid (F_x, F_y) under test. The EGT subsystem compared this distance with the Euclidean distance defined by the SD values in x and in y of the POG points that resulted in (F_x, F_y). If the distance from (F'_x, F'_y) to (F_x, F_y) was greater than this threshold, then (F_x, F_y) was acknowledged as representing the new point of user attention, and it became the updated qualified fixation point (F'_x, F'_y).

Design of Experiments

We designed two experiments to test the effectiveness and efficiency of EMG/EGT system performance and to compare the performance of the system with other forms of computer input. Nondisabled adult volunteers were used as subjects for these tests (30 volunteers for experiment 1 and 15 for experiment 2). They operated the EMG and EMG/EGT interfaces without using their hands. A system layout of the various components of the hybrid EMG/EGT system is shown in **Figure 4**. The Florida International University Institutional Review Board approved the use of human subjects for this study.

Experiment 1

Experiment 1 was designed to test whether the EMG/EGT-based input would produce lower error rates and comparable task times with those recorded for EGT-based input in point-and-click trials. Also, this experiment would use the error rate and task time measures to compare the performance of an EMG/EGT-based input with that of a mouse (used normally by a nondisabled subject) in completing these trials.

Using Microsoft's Visual Basic (Redmond, Washington), we created a purpose-specific program for this experiment. Each trial was displayed on a 19 in. (48 cm)

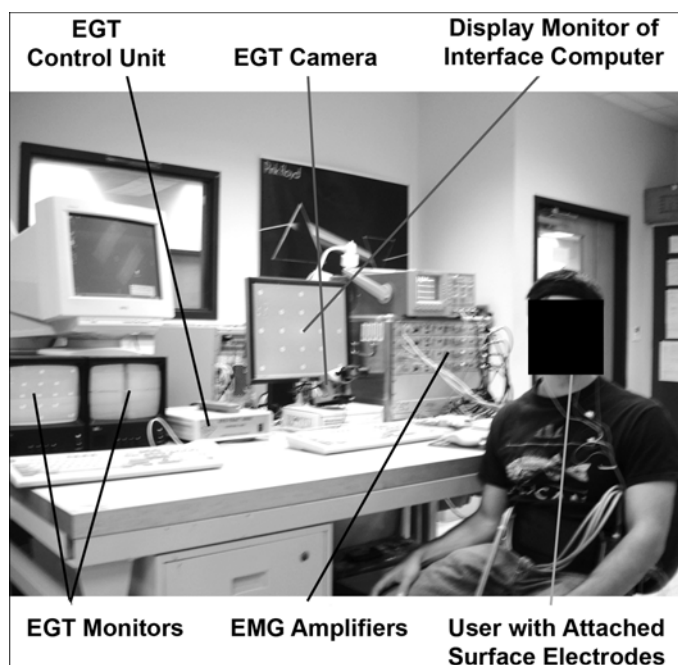


Figure 4. EMG/EGT system components, including experimental instruments. EGT = eye-gaze tracking, EMG = electromyogram.

monitor at a resolution of 1280×1024 pixels. The participant was seated in front of the monitor, such that the eye-to-screen distance was approximately 75 cm. The layout of an example trial is shown in **Figure 5**. Each layout contained a square icon labeled "HOME" and a circular icon labeled "TARGET." Three target diameters (48 pixels, 66 pixels, 96 pixels), three pointing distances (286 pixels, 578 pixels, 778 pixels), and four directions of approach (northeast, southeast, southwest, northwest) were chosen for this experiment. These factors were crossed to produce 36 (3 target diameters $\times$ 3 distances $\times$ 4 directions) unique trial conditions. The placement of the two icons was arranged so that the center of the screen would always bisect the distance between them.

Three cursor control techniques were used in the experiment: EMG/EGT, EGT, and mouse. An eye-gaze dwell time threshold of 350 ms was used to issue left-clicks for the EGT technique. Thirty participants were grouped according to the cursor control technique they would use to perform the experiment, that is, 10 participants for each cursor control technique. For a given trial, a subject was instructed to click the home icon, move the cursor to the target icon, and then click the target icon. The movement time and any selection errors (clicking outside the target icon) were recorded for each trial. Each of the 36 unique trial conditions was repeated twice resulting in 72 trials per participant. The layouts were presented in a random order. Prior to the performance of the 72 timed trials, subjects

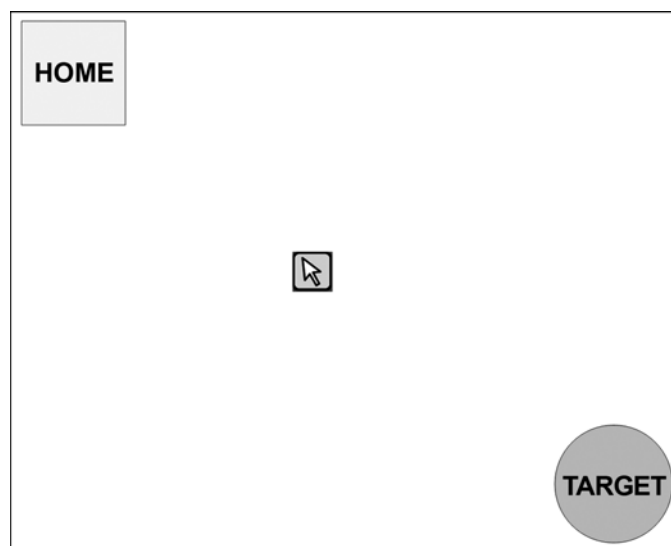


Figure 5. Example point-and-click trial layout on computer monitor screen for experiment 1.

were allowed to practice with their corresponding cursor control technique until they reported that they felt comfortable with the operation of their assigned input mechanism.

Experiment 2

Experiment 2 was designed for testing whether the EMG/EGT-based input could produce a lower error rate than the EGT-based input in point-and-click trials when the source of error was exclusively due to unintended gaze-based selections. To test this premise, we used only large icons in this experiment. We did this to minimize EGT selection errors derived from EGT limitations in accuracy and to assess mainly errors associated with the use of gaze-based selection as a clicking mechanism. The gaze-based dwell time threshold for the EGT system was set to 350 ms. An example of the trial layout used in experiment 2 is shown in **Figure 6**.

Each trial displayed a green circle labeled “START” separated by a center-to-center horizontal distance of 578 pixels (13.0°) from a red target circle. The diameter of each circle was 96 pixels (2.2°). At this size, EGT-based selection errors caused by accuracy limitations were not expected to be predominant. The red target circle was labeled “Y” or “N.” For a given trial, the “START” circle was presented on either the left or right side of the screen, with the target circle located on the opposite side. Both circles were equidistant from the center of the screen.

The trial objective was to have the user select the “START” circle and then move the cursor toward the target circle. The user must then select the target only if a “Y” label was displayed within it but not if an “N” label was displayed. If no target selection was made within 7 s for either kind of target, then the trial would time out. This trial design required that a user examine the target before selecting it. Under these circumstances, unintended selections could possibly occur if a gaze-based selection for an EGT-based input was used.



Figure 6.

Example point-and-click trial layout on computer monitor screen for experiment 2.

For a given experiment, the participant was required to use two cursor control techniques (EGT and EMG/EGT) in a repeated measures design. The cursor control techniques were presented to the participants in a random order. Each cursor control technique had two sessions of data collection. Also, the participants were given a practice session before using each technique to develop their skill in using the technique and also given 5-minute breaks between sessions to minimize the effects of fatigue.

In a session, each of the four unique trial layouts was repeated eight times for a total of 32 trials. This process resulted in a total of 128 trials ($32 \text{ trials} \times 2 \text{ techniques} \times 2 \text{ sessions}$) per participant. Fifteen individuals participated in the experiment.

Data Analysis Methods

For experiment 1, we statistically analyzed the two dependent variables of trial time and error rate separately using mixed design analyses of variance (ANOVAs). These analyses were done to investigate the effects of the various factors on each variable. These analyses were accompanied by orthogonal contrasts of the cursor control techniques for both error rate and trial time.

The parametric tests that were performed in the analysis of the data in experiment 1 are considered valid only if the data satisfy certain assumptions. Two such assumptions are normality and homoscedasticity. Normality refers to how closely a distribution of observed results approximates a normal distribution. A normal distribution is one that is symmetric, unimodal (has a single peak), and bell-shaped. It can be defined by two parameters: its mean (μ) and its SD (σ). Homoscedasticity or homogeneity of variance refers to a case where two or more populations (groups of values) have equal variances.

For experiment 2, we found that the data could not be made to satisfy the assumptions of normality and homoscedasticity by performing data transformations and outlier removal. Therefore, we performed nonparametric tests on the data, since such tests do not require that parametric assumptions be satisfied before analysis. One of the forms of analysis used called the Friedman test is a nonparametric test that is functionally similar to a repeated measures ANOVA. This test involves ranking each block of the experimental results and then using these ranks to determine the average rank of each treatment level. The test statistic that is based on these averages is used to determine whether a statistical difference exists between the treatment levels [18]. The Wilcoxon signed rank test (the other nonparametric test used) is a nonparametric analog of

the paired t -test; i.e., it is used to compare the results of two related samples by taking differences between corresponding results from the two related samples. The absolute values of these differences are ranked, and the ranks corresponding to the positive differences are summed. The last stage of the Wilcoxon test analyzes this sum to determine whether a statistical difference exists between the two samples [18].

The Friedman test was used to analyze the differences between treatments across the 15 subjects that participated in experiment 2. In addition to the Friedman test, a number of Wilcoxon signed rank tests were performed to allow for pairwise comparisons of the different treatment conditions.

STATISTICAL RESULTS

Experiment 1

The mixed design ANOVAs used to analyze the time and error rate results are based on the parametric assumption of normality. Both the trial time and error rate data were found to be substantially nonnormal in their distributions. This finding was circumvented by applying logarithmic transformations to both the trial time [$\log_{10}(x)$] and error rate [$\log_{10}(x + 1)$] data sets.

The tests of between-subjects effects for trial time revealed a significant effect for cursor control technique ($p < 0.001$), and the contrasts for these effects revealed that the EMG/EGT technique was significantly slower than both the mouse ($p < 0.001$) and EGT ($p < 0.001$) techniques. A bar chart representing the transformed mean trial time data is given in **Figure 7**. Also, for results to be shown in a “real-world” context, the marginal means of the cursor control techniques for the untransformed trial time data are given in **Table 2**.

The tests of between-subjects effects for error rate also displayed a significant effect for cursor control technique ($p < 0.001$), and the contrasts for these effects revealed that the EMG/EGT technique had a significantly smaller error rate than the EGT technique ($p < 0.001$).

The contrasts also showed that the error rate produced by the EMG/EGT technique was comparable with that of the mouse ($p = 0.206$). A bar chart of the transformed mean error rate data is given in **Figure 8**, and the marginal means of the cursor control techniques for the untransformed error rate data are given in **Table 3**.

Experiment 2

We examined the data collected for experiment 2 to determine how many selections of “N” label targets occurred for each session. We interpreted these selections as selection errors and divided the total of these errors by the total of “N” label targets presented for each session (16). This produced a selection error proportion for each session, and each subject performed two sessions. Therefore, we recorded four such treatment values (2 cursor control techniques $\times$ 2 sessions) for each subject participating in the experiment. The results of this preprocessing procedure were then analyzed with the Friedman test.

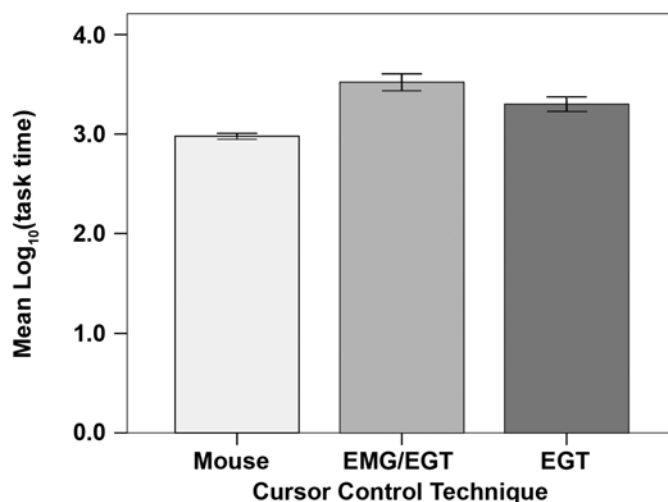


Figure 7.

Mean $\log_{10}(\text{task time})$ values for cursor control techniques (error bars = 95% confidence interval) for experiment 1. EGT = eye-gaze tracking, EMG = electromyogram.

Table 2.

Marginal means of cursor control technique variable for untransformed time data for experiment 1.

Cursor Control Technique	Mean $\pm$ Standard Error (ms)	95% Confidence Level	
		Lower Bound	Upper Bound
Mouse	983.92 $\pm$ 379.92	204.39	1,763.46
EMG/EGT	4,683.97 $\pm$ 379.92	3,905.42	5,464.50
EGT	3,069.81 $\pm$ 379.92	2,290.27	3,849.34

EGT = eye-gaze tracking, EMG = electromyogram.

The Friedman test revealed that the difference between the ranks of each treatment condition was significant ($p < 0.001$). The Wilcoxon signed rank test in turn revealed that these differences were due to effects of the cursor control techniques, because significant differences were only found between treatments that involved different techniques (Table 4). Additionally, Figure 9 shows that the mean error rate was lower for the EMG/EGT technique compared with the EGT technique.

DISCUSSION

The statistical results of experiment 1 have formalized some interesting observations regarding the EMG/EGT

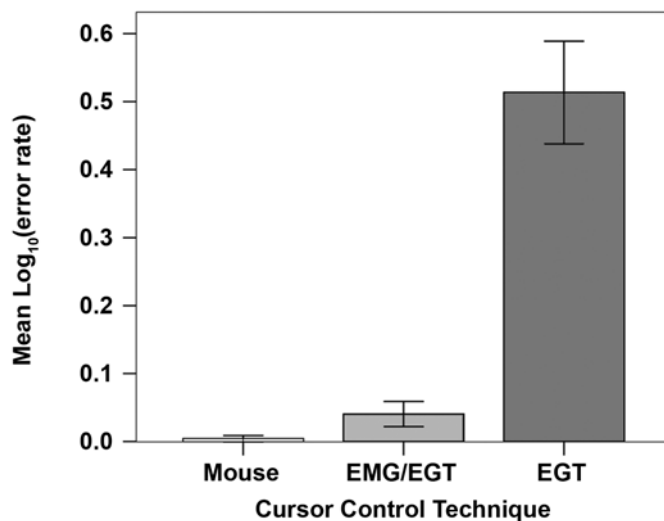


Figure 8.

Mean $\log_{10}(\text{error} + 1)$ values for cursor control techniques (error bars = 95% confidence interval) for experiment 1. EGT = eye-gaze tracking, EMG = electromyogram.

Table 3.

Marginal means of cursor control technique variable for untransformed error data (in errors/trial) for experiment 1.

Cursor Control Technique	Mean $\pm$ Standard Error	95% Confidence Interval	
		Lower Bound	Upper Bound
Mouse	0.01 $\pm$ 0.25	-0.49	0.52
EMG/EGT	0.14 $\pm$ 0.25	-0.37	0.64
EGT	3.98 $\pm$ 0.25	3.47	4.48

EGT = eye-gaze tracking, EMG = electromyogram.

system. First, the addition of the EMG-based interaction to the EGT-based interaction seemingly reduced the user's speed in performing cursor control tasks (4.7 s mean trial time for EMG/EGT compared with 3.1 s for mean trial time for EGT). One may understand this slowing effect more by considering the empirically observed distinct mechanisms used by EGT and EMG/EGT users when performing these trials. Note that the target diameters for this experiment were set at three different values, out of which only the largest target (96 pixel diameter) was found to be reliably selectable by the EGT input. This lack of reliability for selecting the other target sizes (48 pixels and 66 pixels) was due to the inherent low level of accuracy of EGT-based inputs, coupled with the occurrence of POG offsets due to minor head movements. These inaccuracies in the EGT system would force the user to shift his or her gaze around the intended target to eventually select it. This compensatory activity performed by the EGT users increased trial time and decreased target diameter. When EMG/EGT users were confronted with these smaller target sizes, they used EMG-based control to make up for the lack of accuracy exhibited by the EGT subsystem instead of the compensatory eye movements that EGT users made. Unfortunately, because EMG/EGT users were required to coordinate eye movements with facial movements, a task time penalty was incurred, in addition to the task time associated with eye-based control when operating in isolation.

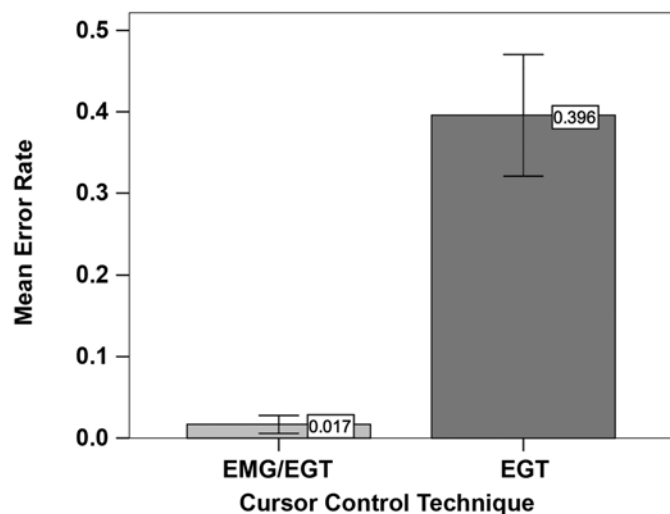
A review of only the trial time results might lead one to conclude that no benefit to integrating EMG and EGT modalities exists. However, the benefits of this integration are strongly validated by the error rate results. The mean error rates for the three cursor control techniques were 0.01 errors/trial for the mouse, 0.14 errors/trial for EMG/EGT, and 3.98 errors/trial for EGT (Table 3). The contrasts of these mean values showed that the EMG/EGT system produced significantly fewer errors than the EGT input ($p < 0.001$) for the target sizes used in this experiment. In fact, the EMG/EGT input produced an error rate similar to that produced by the mouse input ($p = 0.206$ for contrast). Again, the large difference in error rate values between the EMG/EGT and EGT inputs may be attributed to the different approaches employed by the users of the respective systems when selecting the smaller target sizes. The unnatural shifting of eye gaze in the region of these smaller targets, which was used by EGT users to compensate for its inaccuracy, often resulted in unintended left-clicks being issued in the region surrounding the target. These left-clicks were recorded as errors by the Visual Basic program

Table 4.

Wilcoxon signed rank test results for EMG/EGT and EGT cursor control techniques of experiment 1.

Statistical Descriptor	(EMG/EGT + Sess 2) – (EMG/EGT + Sess 1)	(EGT + Sess 1) – (EMG/EGT + Sess 1)	(EGT + Sess 2) – (EMG/EGT + Sess 1)	(EGT + Sess 1) – (EMG/EGT + Sess 2)	(EGT + Sess 2) – (EMG/EGT + Sess 2)	(EGT + Sess 2) – (EGT + Sess 1)
z Score	-0.632	-3.413	-3.417	-3.419	-3.306	-1.364
Asymp. Sig. (2-tailed)	0.527	0.001	0.001	0.001	0.001	0.173

Asymp. Sig. = asymptotic significance, EGT = eye-gaze tracking, EMG = electromyogram, sess = session.

**Figure 9.**

Mean error rates (errors/total session trials) for EMG/EGT and EGT cursor control techniques of experiment 2 (error bars = 95% confidence interval). EGT = eye-gaze tracking, EMG = electromyogram.

used to present the trials to the user. The reason for these unintended left-clicks can be traced to target selection being based on dwell time for the EGT modality. When the user activated the EMG-based input to compensate for the lack of accuracy of EGT-based input, the user's control of the interaction with the computer was enhanced in two critical ways: (1) the deliberate execution of small cursor movements became possible and (2) unintended selections were significantly reduced. These two advantages provided by EMG/EGT control resulted in a more reliable icon selection mechanism, which is especially suited for high-resolution environments.

We conducted experiment 2 primarily to determine, through statistical analysis, whether the EMG/EGT system was less susceptible to selection errors than an eye-tracking system that used gaze dwell time as the basis for its selection operation. The result of this experiment would seem intuitive because the selection operation for the EMG/EGT technique was performed by the EMG-

monitored action of clenching both sides of the jaw simultaneously and did not depend on gaze time. The statistical results supported this assumption with the EGT technique producing a mean gaze-based selection error rate of 0.396 and the EMG/EGT technique having an error rate of 0.017 (**Figure 9**).

A secondary reason for conducting experiment 2 was to see how prone to errors a gaze dwell time selection system would be for tasks that could elicit such errors. This type of experiment had not been conducted previously by the proponents of gaze dwell time-based EGT selection techniques [8,10–11]. In their experiments, the targets presented to the user were always required to be selected; i.e., no decision was necessary. As discussed previously, the gaze time was set to 350 ms for the EGT system used in this experiment. This setting was empirically found to be the best compromise between the speed of selection and the ability to avoid unintended selections, while remaining within range of gaze dwell times reported by Sibert et al. [10–11] and Ware and Mikaelian [8], i.e., 150 to 400 ms. The gaze-based selection error rate produced by this technique was approximately 40 percent, which implies that EGT techniques that use dwell time to directly issue left-clicks would not be recommended for environments where unintended selections based on gaze are possible.

CONCLUSIONS

A hybrid EMG/EGT system was created that has the following key performance features:

1. It does not require the use of hands to perform computer cursor operations.
2. It gives users the ability to modify cursor position pixel by pixel; i.e., the system does not have a spatial accuracy limitation.
3. It provides a reliable left-click operation.

Together, these features attract the EMG/EGT system to a user who requires a hands-free form of cursor control that can execute point-and-click operations in a high-resolution window, icon, menu, and pointing-device environment or in an Internet browser application.

Feature 1 makes the EMG/EGT system a viable option for providing individuals with motor disabilities access to computers through GUIs. Features 2 and 3 are the advantages of using this hybrid system instead of an EGT system that uses gaze dwell time to execute selections.

While the synthesis of a hybrid control system comprising EGT and EMG subsystems did not fully inherit the speed of the EGT system in commanding cursor movements (experiment 1), the increased reliability achieved in the much smaller number of unintended selections (experiment 2) is likely to have an important affect on the quality and comfort of the interaction of users with actual GUIs and Web applications. This finding is particularly relevant since many of these applications may include graphic elements (buttons, active area of hyperlinks, etc.), which cannot be resized in the end-user computer. The additional accuracy found for the EMG/EGT hybrid system under these circumstances is likely to alleviate the potential for user frustration that performance limitations in standard EGT systems may cause.

Although the performance and usability of the EMG/EGT device have been significantly enhanced, work still remains to be done in making this system accessible to computer users outside a laboratory environment. One area in which system usability may be improved is in minimizing/removing the calibration procedure for the EMG subsystem that is a precursor to each user session. One way to do this procedure is to employ a standard supervised classification algorithm to obtain a generalized solution from a feature set derived from a large enough (e.g., 100 or more) representative user population. This automated calibration method would eliminate the need for manual EMG calibration before system use.

Another improvement to the system would be developing a commercial EMG integration kit, which can be used with currently available commercial EGT systems. This kit would consist of (1) a data acquisition board; (2) specialized biosignal amplifiers with settings specifically chosen for this application; and (3) a software package that contains the EMG, EGT, and information fusion modules described previously. All that would be required from the commercial EGT is a stream of POG data.

Finally and probably most importantly, we plan to continue user testing with individuals with motor disabilities. We believe that these tests will provide the greatest insights as how best to enhance system form and function.

ACKNOWLEDGMENTS

This material was based on work supported by National Science Foundation grants IIS-0308155, CNS-0520811, HRD-0317692, and CNS-0426125.

The authors have declared that no competing interests exist.

REFERENCES

1. The National Spinal Cord Injury Association [homepage on the Internet]. Rockville (MD): The Association; c2003–08 [updated 2005 Sep 9; cited 2007 Jul 24]. Fact sheets; [about 2 screens]. Available from: <http://spinalcord.org/news.php?dep=17&page=94spinstat.php>
2. International Campaign for Cures for Spinal Cord Injury Paralysis [homepage on the Internet]. Charlottesville (VA): The Association; c1998–2008 [updated 2004 Jul; cited 2008 Jan 30]. Global summary of spinal cord injury, incidence and economic impact; [1 screen]. Available from: <http://www.campaignforcure.org/globalsum.htm>
3. Basmajian JV, DeLuca CJ. Muscles alive: Their functions revealed by electromyography. 5th ed. Baltimore (MD): Williams & Wilkins; 1985.
4. Chang GC, Kang WJ, Luh JJ, Cheng CK, Lai JS, Chen JJ, Kuo TS. Real-time implementation of electromyogram pattern recognition as a control command of man-machine interface. *Med Eng Phys*. 1996;18(7):529–37. [PMID: 8892237]
5. Barreto AB, Scargle SD, Adjouadi M. A practical EMG-based human-computer interface for users with motor disabilities. *J Rehabil Res Dev*. 2000;37(1):53–63. [PMID: 10847572]
6. Kim JS, Jeong H, Son W. A new means of HCI: EMG-MOUSE. In: Proceedings of the 2004 IEEE International Conference on Systems, Man, and Cybernetics, Vol. 1; 2004 Oct 10–13; The Hague, the Netherlands. Piscataway (NJ): IEEE; 2004. p. 100–104.
7. Jacob RJ. The use of eye movements in human-computer interaction techniques: What you look at is what you get. In: Maybury MT, Wahlster W, editors. Readings in intelligent user interfaces. San Francisco (CA): Morgan Kaufmann Publishers Inc; 1998. p. 65–82.
8. Ware C, Mikaelian HH. An evaluation of an eye tracker as a device for computer input. In: Proceedings of the SIGCHI/GI Conference on Human Factors in Computing Systems

- and Graphics Interface; 1987 Apr 5–9; Toronto, Canada. New York (NY): Association for Computing Machinery; 1986. p. 183–88.
9. Hutchinson TE, White KP Jr, Martin WN, Reichert KC, Frey LA. Human-computer interaction using eye-gaze input. *IEEE Trans Syst Man Cybern.* 1989;19(6):1527–34.
 10. Sibert LE, Jacob RJ. Evaluation of eye gaze interaction. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*; 2000 Apr 1–6; The Hague, the Netherlands. New York (NY): Association for Computing Machinery; 2000. p. 281–88.
 11. Sibert LE, Jacob RJ, Templeman JN. Evaluation and analysis of eye gaze interaction. *NRL Report*. Washington (DC): Naval Research Laboratory; 2001.
 12. Morimoto CH, Mimica MR. Eye gaze tracking techniques for interactive applications. *Comput Vis Image Underst.* 2005;98(1):4–24.
 13. Kaur M, Tremaine M, Huang N, Wilder J, Gacovski Z, Flippo F, Mantravadi CS. Where is “it”? Event synchronization in gaze-speech input systems. In: *Proceedings of the 5th International Conference on Multimodal Interfaces*; 2003 Nov 5–7; Vancouver, Canada. New York (NY): Association for Computing Machinery; 2003. p. 151–58.
 14. Zhang Q, Imamiya A, Mao X, Go K. A gaze and speech multimodal interface. In: *Proceedings of the 24th International Conference on Distributed Computing Systems Workshops*; 2004 Mar 23–24; Hachioji, Japan. Piscataway (NJ): IEEE (Computer Society); 2004. p. 208–13.
 15. Trejo LJ, Wheeler KR, Jorgensen CC, Rosipal R, Clanton ST, Matthews B, Hibbs AD, Matthews R, Krupka M. Multimodal neuroelectric interface development. *IEEE Trans Neural Syst Rehabil Eng.* 2003;11(2):199–204. [\[PMID: 12899274\]](#)
 16. Wolpaw JR, Birbaumer N, McFarland DJ, Pfurtscheller G, Vaughan TM. Brain-computer interfaces for communication and control. *Clin Neurophysiol.* 2002;113(6):767–91. [\[PMID: 12048038\]](#)
 17. Surakka V, Illi M, Isokoski P. Gazing and frowning as a new human-computer interaction technique. *ACM Trans Appl Percept.* 2004;1(1):40–56.
 18. Petrucci JD, Nandram B, Chen M. *Applied statistics for engineers and scientists*. Upper Saddle River (NJ): Prentice Hall; 1999.

Submitted for publication March 28, 2007. Accepted in revised form July 16, 2007.