THE INTEGRATION OF ELECTROMYOGRAM AND EYE GAZE TRACKING INPUTS FOR HANDS-FREE CURSOR CONTROL

Craig A. Chin and Armando Barreto
Department of Electrical and Computer Engineering, Florida International University
Miami, FL 33174

ABSTRACT
Electromyogram (EMG)-based and eye gaze tracking (EGT)-based hands-free cursor control input systems have been developed in the past as independent forms of cursor control. Each form of control possesses its own advantages and disadvantages in terms of usability. This paper presents a novel form of hands-free cursor control that integrates these two inputs in order to provide the user with the ability to manipulate the cursor more efficiently than the individual input systems operating in isolation. An experiment was conducted to compare the performance of this new EMG/EGT input system to EGT and mouse input systems in point-and-click trials. The results showed that while the EMG/EGT system was slower than the EGT system and the mouse, it produced a significantly smaller error rate than EGT input alone and therefore, EMG/EGT input could be considered to be a more usable form of hands-free cursor control when compared to EGT input.

Keywords: Electromyogram, eye gaze tracking, cursor control

INTRODUCTION
Electromyography is the study of muscle function through the monitoring of the electrical signals emitted by the muscle [1]. 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 spatio-temporal summation of these electrical signals results in what is called an electromyogram (EMG) signal. Therefore, the EMG signal provides an effective means of monitoring muscle activity.

EMG signals from muscles in the body have been used in the past for cursor control. This approach has been used in [2, 3], where cranial muscles were targeted. The use of EMG signals from cranial muscles is an approach that would be suitable for individuals who are unable to use their hands because of a motor disability or because their hands would otherwise be assigned to higher priority tasks.

The advantages of EMG-based cursor control are that it provides the user with the ability to perform small, discrete cursor movements, and a robust, stable “clicking” procedure. However, it has been shown that this approach performs slowly compared to a mouse-operated system in point-and-click tests [2, 3], and could potentially become tiresome if the user is required to make large excursions across the screen with this input approach.

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 rapid, ballistic motion that moves the eye from one area of focus of the visual scene to another. Vision is suppressed during a saccade. 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 that 1° in size [4].

Eye gaze tracking (EGT) techniques seek to determine the user’s visual line of gaze by taking video images of the eye in order to establish a relationship between the geometric orientation of visual features of the eye and the line of gaze. The most popular EGT technique at present uses the relative position of
the bright eye (pupil) center and the center of the glint (corneal reflection) to determine the line of gaze [4, 5]. 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.

EGT techniques have been shown to perform faster than a mouse in point-and-click tests [5]. However, this approach has some disadvantages. One such disadvantage is the so-called “Midas Touch” problem [4], which originates when eye gaze dwell time is used to issue a left-click operation. During a human-computer interaction session, situations may arise where a user may only desire to stare at an icon to examine it, rather than to select it. If this user is utilizing an EGT input system that issues left-clicks when the point of gaze dwells in a small area for specific period of time, unintended selections may result. Another disadvantage is the limited accuracy of the approach. This limitation results from the fact that 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 still allows variations of about 1° of visual arc in the direction of gaze [4, 5]. 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 input system, this would severely deteriorate the computer cursor’s stability. There is also the issue of POG offsets that 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 eye gaze tracker decays (becomes less accurate) as the head moves away from its original position [6]. Therefore, the only ways to restore the accuracy of the EGT system are to either place 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 EGT and EMG 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, it was decided to pursue the creation of a bimodal cursor control system that will selectively utilize both types of input from the user to provide a more efficient manipulation of the screen cursor, under a wider range of circumstances. Ideally, the hybrid EMG / EGT system will use incremental (stepping) positional commands derived from the EMG subsystem to effect small cursor displacements within a restricted neighborhood of the current cursor location. Similarly, only the EMG subsystem will be used to determine when the user commands a click operation. In this way the cursor stability and clicking reliability observed in the evaluation of EMG inputs 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 EGT subsystem will be employed. This will reduce the time and effort required to perform these types of cursor manipulations.

**MATERIALS AND METHODS**

**EMG Subsystem Implementation**

The purpose of the EMG subsystem was to detect EMG signals generated during muscle contractions associated with prescribed facial movements, and to translate these signals into cursor actions. Figure 1 displays the placement of the four Ag/AgCl electrodes used to capture the EMG signals from the head of the user. 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. Also, an electrode was placed over the right mastoid as a reference.

![Electrode placement for the EMG cursor control system](image)

Figure 1 Electrode placement for the EMG cursor control system

There was a predefined mapping between cursor actions, facial movements, and muscle contractions for the EMG subsystem (Table 1).

<table>
<thead>
<tr>
<th>Cursor Action</th>
<th>Facial Movement</th>
<th>Muscle Contraction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Left</td>
<td>Left Jaw Clench</td>
<td>Left Temporalis</td>
</tr>
<tr>
<td>Right</td>
<td>Right Jaw Clench</td>
<td>Right Temporalis</td>
</tr>
<tr>
<td>Up</td>
<td>Eyebrows Up</td>
<td>Right Frontalis</td>
</tr>
<tr>
<td>Down</td>
<td>Eyebrows Down</td>
<td>Procerus</td>
</tr>
<tr>
<td>Left-Click</td>
<td>Left &amp; Right Jaw Clench</td>
<td>Left &amp; Right Temporalis</td>
</tr>
</tbody>
</table>

Table 1 Relations between cursor actions, facial movements and muscle contractions

An EMG classification algorithm was created for the purpose of determining if a facial muscle contraction had occurred and if so, which specific muscle was the source of this contraction. Given the one-to-one mapping between muscle contraction and cursor action shown in Table 1, the output of an effective muscle contraction classification algorithm can be used to translate facial movements into cursor actions.

The EMG classification algorithm used the power spectral density (PSD) of the input EMG signals as the basis for classification. It had been observed previously that the four muscles being monitored possessed distinct EMG spectral characteristics, and that this spectral information would be useful for performing classifications [2, 3]. Empirical observations suggested that Mean Power Frequency (MPF) values would be a suitable way of representing the spectral data for this input configuration. The MPF is derived from PSD values as a weighted average frequency in which each frequency component, \( f \), is weighted by its power, \( P \). The equation for the calculation of the MPF is given by:

\[
MPF = \left( \frac{f_0 \times P_0 + \ldots + f_k \times P_k + \ldots + f_{N-1} \times P_{N-1}}{P_0 + \ldots + P_k + \ldots + P_{N-1}} \right)
\]

where \( k = 0, 1, 2, \ldots, N - 1 \) (in our implementation \( N = 256 \)).

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

The EMG classification algorithm derived three features from each PSD estimate calculated for each EMG input to assist in determining which muscle was the source of a contraction. These features were: the maximum PSD magnitude, the sum of all the PSD magnitudes for a given estimate, and the MPF value for the estimate. The four-channel analog-to-digital conversion and all the EMG analysis calculations were implemented, in real-time, in an Innovative Integration ADC-64 board.

**EGT Subsystem Implementation**

The eye tracking system used for the EGT subsystem was an R6-HS Remote Optics system manufactured by Applied Science Laboratories. The EGT subsystem functioned by allowing a beam from near infrared LED’s located on a pan/tilt optics module to illuminate the eye of the user. The eye image that was produced by this illumination was focused and sensed by a video camera also present on the pan/tilt unit. Video image data was fed into the eye tracker control unit which performed image feature recognition and POG estimation. The POG estimates and pupil diameter values were then sent out to the display PC (the computer that interacted with the user) and the interface PC (the computer that interacted with the experimenter). These values were sent to the serial ports of the respective computers. The cursor control application running on the display PC received these values via hardware interrupts that occurred at a rate of 120 Hz.

A fixation identification algorithm was created to obtain a more reliable estimate of the user’s point of attention from the raw POG coordinates produced by the eye tracking system. This algorithm utilized temporal and spatial criteria to determine whether or not a fixation had occurred. More specifically, the algorithm extracted a 100 ms moving window (temporal threshold) of consecutive POG data points (POG<sub>x</sub>, POG<sub>y</sub>), and calculated the standard deviation of the x- and y-coordinates of these points. If both standard deviations were less than the coordinate thresholds associated with 0.5° of visual angle (spatial threshold), then it was determined that the onset of a fixation had occurred, and the point used to represent the fixation were the coordinates of the centroid of the POG samples received during the 100 ms window analyzed, (F<sub>x</sub>, F<sub>y</sub>). If it was determined that a fixation had not occurred, then the window was advanced by one data point and fixation identification was performed again.

**Information Fusion and Cursor Update Algorithm**

A new information fusion and cursor update algorithm determined the effective cursor position (C<sub>x</sub>, C<sub>y</sub>) as a merging of the incremental EMG commands (∆x, ∆y) and the absolute coordinates of a qualified EGT fixation (F’<sub>x</sub>, F’<sub>y</sub>):

\[
\begin{align*}
C_x[n] &= \begin{cases} 
C_x[n-1] + \Delta x[n] & \text{If EMG update} \\
F'_x[n] & \text{If EGT update}
\end{cases} \\
C_y[n] &= \begin{cases} 
C_y[n-1] + \Delta y[n] & \text{If EMG update} \\
F'_y[n] & \text{If EGT update}
\end{cases}
\end{align*}
\]  

(eq. 2)

(eq. 3)

where C<sub>x</sub> and C<sub>y</sub> represent the x- and y-coordinates of a cursor position and n represents 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<sub>x</sub>[n], C<sub>y</sub>[n]) may 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 \(\Delta x\) or \(\Delta y\). The direction of the increment was determined by the output value of the EMG classification algorithm.

An EGT subsystem update involved replacing the previous cursor position with the absolute coordinates of a qualified fixation \((F'_x, F'_y)\). A qualified fixation was determined by taking every new fixation centroid \((F_x, F_y)\) identified by the fixation identification algorithm, and testing it to determine if it signified a new point of user attention, or if it simply was the continuation of previous fixation. This was done by measuring the distance between the current qualified fixation position \((F'_x, F'_y)\) and the \((F_x, F_y)\) under test. This distance was compared to the Euclidean distance defined by the standard deviations in x and in y of the POG points that resulted in the new fixation \((F_x, F_y)\). 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 new qualified fixation point \((F'_x, F'_y)\).

**Design of Point-and-Click Experiment**

An experiment was designed to test whether the EMG/EGT input would produce lower error rates and comparable task times to those recorded for an EGT input system in point-and-click trials. Also, this experiment would use the error rate and task time measures to compare the performance of EMG/EGT input to that of a mouse in completing these trials.

The experiment was created in Visual Basic and each trial was displayed on a 19” monitor. The participant was seated in front of the monitor, such that the eye to screen distance was approximately 29”. Each trial layout contained a square icon labeled “HOME” and a circular icon labeled “TARGET”. There were three target diameters (48 pixels, 66 pixels, 96 pixels), three pointing distances (286 pixels, 578 pixels, 778 pixels), and four angles of approach (NE, SE, SW, NW) chosen for this experiment. These factors were crossed to produce 36 (3 target diameters x 3 distances x 4 angles) unique trial layouts. The distance between the two icons was spaced in such a way that the center of the screen would always bisect the pointing distance.

There were three cursor control techniques used in the experiment: EMG/EGT, EGT, and mouse. 30 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 layouts was repeated twice resulting in 72 trials per participant. The layouts were presented in a random order. Also, there was a practice session prior to each experiment to allow the user to gain some familiarity with the cursor mechanism assigned to them.

**RESULTS**

The trial time and error rate results collected during the experiment were arranged separately and analyzed as mixed design ANOVAs. Both the trial time and error rate data were found to be substantially non-normal in their distributions. This prompted the use of logarithmic transformations of both the trial time \([\log_{10}(X)]\) and error rate \([\log_{10}(X + 1)]\) data sets in advance of the statistical analysis. The test of between-subjects effects for trial time revealed a significant effect for cursor control input \((p < 0.0005)\), and the contrasts for these between-subjects effects also revealed that the EMG/EGT input is significantly slower than both the mouse and EGT inputs (See Figure 2). The between-subjects tests for error rate indicated a significant effect for cursor control input \((p < 0.0005)\) and the contrasts for these between-subjects effects revealed that the EMG/EGT input can be considered
to have a comparable error rate with the mouse input. Also, the EMG/EGT input was found to have a significantly smaller error rate than the EGT input (See Figure 2).

![Figure 2 – Bar charts showing mean trial time (left plot) and error rate (right plot) performance for cursor control inputs used in the experiment, after logarithmic transformations](image)

**CONCLUSIONS**

The results showed that the addition of the EMG-based interaction to the EGT-based interaction has resulted in a reduction of the user’s speed in performing cursor control tasks (4.685 s mean trial time for EMG/EGT compared to 3.070 s for mean trial time for EGT). However, the EMG/EGT input produced a much lower error rate compared to the EGT input (0.135 errors/trial for EMG/EGT and 3.976 errors/trial for EGT). In spite of its slowness, the assertion that the EMG/EGT input is a more usable form of cursor control compared to EGT is justified by the fact that EMG/EGT provides a more reliable left-click operation and the fact that there is no limitation on spatial accuracy (the minimum size icon it can select).

**ACKNOWLEDGMENTS**

This work was sponsored by NSF grants IIS-0308155, CNS-0520811, HRD-0317692 and CNS-0426125.

**REFERENCES**