Robust Eye Blink–Based Selection Technique for Gaze-Based Interaction

Yasha Iravantchi
Harvard University
33 Oxford St.
Cambridge, MA
yasha@seas.harvard.edu

Krzysztof Z. Gajos
Harvard University
33 Oxford St.
Cambridge, MA
kgajos@seas.harvard.edu

ABSTRACT
The "Midas Touch" problem, or the inability to robustly discriminate between deliberate and accidental selection commands, is a major factor limiting the usefulness of eye gaze–based interaction.In this paper, we present an electromyography (EMG)-based system capable of differentiating between involuntary and deliberate blinks. We then present a study that explores the performance of EMG-based blink as a method of selection compared to prior eye-gaze based methods such as dwell and held-blink. The results of our study show that the EMG-based approach results in the fastest performance out of the three conditions without adversely affecting accuracy. Participants also reported that the EMG-based system was better at detecting when they wanted to execute a click compared to the other two systems. They also found the EMG-based system less exhausting to use.

INTRODUCTION
Eye gaze tracking is an effective hands-free alternative to mouse for controlling pointer position in 2D. However, even though modern eye gaze trackers allow users to accurately move the pointer about the screen, they lack a reliable mechanism for "clicking". A common solution is to use long gaze fixations (or "dwells") as a signal that a click should be executed. Although popular, this approach suffers from what is commonly called the "Midas Touch Problem": a long fixation could be a result of the intention to click or a user's interest in a particular screen element. The "Midas Touch" problem has plagued dwell-based selection systems because the eyes cannot be turned "off" [2], resulting in excessive clicking and high false positive rates. The false positives can be reduced by increasing the required dwell time duration, but that, in turn, makes the interaction slower.

Recent studies have presented a class of solutions to this problem: hybrid methods add a system independent of eye gaze, such as a brain-computer interface (BCI), that allows the user to perform a selection by detecting specific brain activity associated with selection [10, 9, 8]. However, these existing systems often made sacrifices in speed, comfort, or cognitive load in order to improve accuracy. For instance, although the BCI augmented system's accuracy matched long dwell methods, the task completion time exceeded that of the long dwell because the BCI system needs to detect and confirm "select" brain activity once "search" activity commences [10]. An ideal system should be both fast and accurate.

Our approach is to use eye blink as a signal to execute a click. Eye blinks are quick to perform, but there is a challenge of discriminating between deliberate blinks performed specifically to communicate the intention to click and the involuntary blinks that most people naturally perform several times a minute. Our preliminary study showed that involuntary and deliberate blinks have similar durations, so a vision-based solution was not feasible. Instead, we extract EMG signals derived from the muscles surrounding the eyes. We have found that deliberate and involuntary blinks produce substantially different electrical activity in the muscles surrounded the eyes and we were able to leverage these differences to robustly discriminate between deliberate and involuntary blinks.

Using an affordable EEG headset to collect the EMG signal, we have built a real time system for performing selection with eye blinks. We have conducted a study comparing our approach to one where eye blinks were detected using the the eye tracker's camera (where participants had to hold their blinks a little longer than what is natural) and a dwell-based system. The results show that that the EMG-based system resulted in significantly faster performance than either of the other systems. Participants also felt that it behaved more predictably and that using it was less exhausting than the other systems. There were no statistically significant differences in error rates across the three approaches.

PRIOR WORK
Dwell is a popular method for selection using eye gaze interfaces [5]. Once the eye selects a target, it presents a fixation, a period in which the eye stabilizes for a duration of 200–600 ms [5]. Dwell-based systems issue selection commands based on presence of such fixations. However, these systems are susceptible to issuing false commands if the dwell threshold is set too low and hinder performance if set too high. But even with high threshold, they suffer from the "Midas Touch" problem: any time the user fixates on a screen element—even if is just to examine it—a selection is triggered.

One approach to combat the Midas Touch problem has been to include a "clutch", such as a keypress, to engage and disengage commands [7, 6]. Although these systems improve

Every submission will be assigned their own unique DOI string to be included here.

the performance of gaze-based input, they require custom UIs and are no longer "hands-free".

Another approach is to combine eye gaze with a secondary sensing method for selection [10, 9, 8]. Brain-computer interfaces can be used to detect the desire for selection by categorizing specific brain activity into two categories: "search" and "select". If "select" brain activity appears, the hybrid system issues a selection command [10]. Facial muscle–based systems utilize specific muscle movement in the face, such as frowning or the raising of eyebrows to issue selection commands [9]. Lastly, attention based systems look for EEG signal patterns to determine if the user is in an attentive state in order to allow the issue of clicks from an eye gaze interface [8]. However, BCI and attention-based systems have challenges performing selection tasks quickly. Face muscle–based systems, that require users to make face movements, were simply inaccurate: their lowest mean error rate was 15% and 10% for frowning and raising eyebrows, respectively [9].

THE EMG-BASED BLINK DETECTION SYSTEM

We set out to produce a system that can reliably discern between involuntary and deliberate blinks as a means for selection.

A vision based approach (one where the eye tracker camera was used to monitor the duration of the eye blinks) was not effective: there was a substantial overlap in the distribution of blink durations between deliberate and involuntary blinks. In order to accurately discern between the two, blinks would have to be held to about twice the natural duration [1].

Instead, we turned to electromyography (EMG): might there be systematic differences in the electrical activity produced by muscles in the vicinity of a blinking eye depending on whether the blink is deliberate or involuntary? To answer this question, we began by recording the the electric potential over the left eyebrow as a small number of participants were asked to rest (and, naturally, blinked involuntarily) for a period of 25 seconds then executed 5 deliberate blinks. We then passed the collected data through a 38Hz low-pass filter and we computed features such as blink width, peak amplitude, and trough amplitude (see Figure 1).

Our analysis of the data revealed that negative trough amplitude alone was sufficient to robustly differentiate between deliberate and involuntary blinks. Peak amplitude offers a more complete picture of the blink, but it is not entirely robust to blink types that exhibit a significant maximum following the minimum of the depolarization.

In our measurements we determined there was too much variation between users to have a fixed threshold that could differentiate between the two types of blinks for all users. Thus, the system had to be calibrated for each participant. Typical thresholds for blink strength were -800mV but some participants required thresholds as low as -250mV and some as high as -1200mV.

Implementation

We chose the MindWave Mobile dry electrode EEG set as it was affordable (USD 99), had a 512 Hz sampling rate, and had bluetooth capabilities. The MindWave Mobile interfaces with MatLab via OpenVibe, an open-source BCI package, which pre-processes and filters the data before streaming the data via TCP. We perform the feature detection in MatLab, which then sends mouse events to the OS through the Windows API as needed.

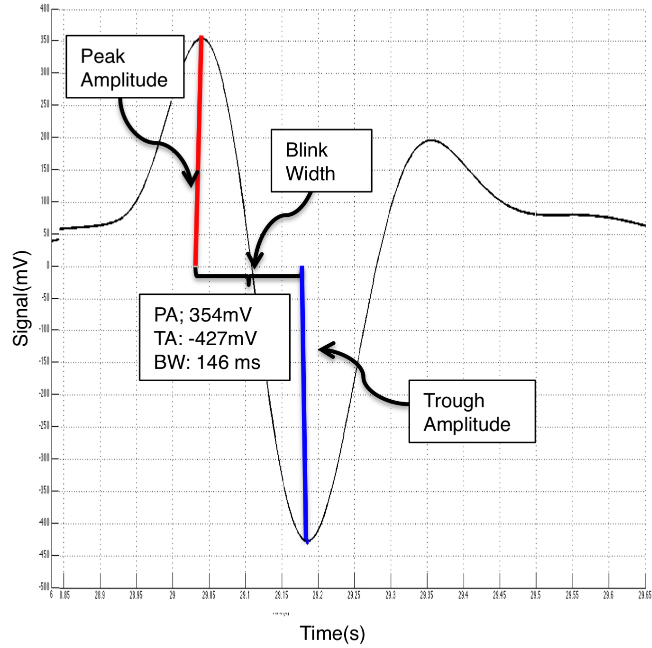

Figure 1. This blink is representative of a deliberate blink. Blue lines represent trough amplitude, red lines represent peak amplitude. PA, TA, and BW stand for top amplitude, bottom amplitude, and blink width, respectively.

EXPERIMENT

We conducted an experiment to compare our approach to dwell and vision-based blink detection.

Participants

Participants were 21 undergraduates (all female) aged 17–22. All participants had to be able to achieve normal vision without glasses (contact lenses were permitted). This was determined by a quick vision test using the GazePoint EyeTracker calibrate function. Participants who could not reliably focus on targets or had vision drift were excluded. Participants were compensated for their participation in this study.

Setup

The study was conducted in a dark room in order to minimize the amount of ambient infrared light that could interfere with the eye tracker. Participants sat approximately 3 feet away from the GazePoint eye tracker and rested their chin on a box to minimize head movement. A 21.5" Dell UltraSharp display with a resolution of 1920x1080 was used. A 15" MacBook Pro running Windows 8.1 was used to conduct the trials. A NeuroSky MindWave Mobile was used to collect the EMG data. OpenVibe v0.18.x and MatLab 2014b versions were used for signal collection and processing. We used a set of pointing tasks based on the ISO 9241-9 standard [4] where circular targets were arranged in a circular configuration. One target at a time was highlighted. As soon as the participant selected that target, another target would be highlighted. The tasks were arranged in 8 blocks of 5 trials each for a total of 40 trials. Target sizes and distances between targets varied between blocks. All participants were presented with the same combination of target sizes and distances and in the same order.

In the dwell condition, in order to click, participants had to fix their gaze on a point for at least 1 second in order to trigger a click. The selection of a 1 second duration was chosen as a result of a pilot study in which participants complained that a 500 ms dwell time was too short.

In the eye tracker blink condition, the participant was informed that, in order to blink, they were required to close their eyes for at least 400 ms in order to trigger a click. In a pilot study, we measured that the average involuntary blink was approximately 200 ms. For the purposes of a held blink, we decided to double the amount of time to 400 ms in order to ensure that involuntary blinks were not captured as click commands.

For the EMG blink condition, the participant was asked to wear a MindWave Mobile EEG headset. The participant was then asked to blink naturally for 15 seconds and then blink with intent five times. The participant was informed that their blinks with intent to click should still feel natural in order to minimize exhaustion.

Procedures

After taking a short vision test in order to evaluate vision competency, the participants were presented with an informed consent form. They were then briefed on the of the study procedures. A short demographics survey was taken prior to the beginning of the experiment.

Next, participants were asked to sit approximately 3 feet in front of a GazePoint eye tracker and a 21.5" display. A box was provided for the participant to rest their chin on to reduce movement during the study.

Prior to each experimental condition, participants were briefed on the selection mechanism. Next, if it was the EMG Blink condition, they were fitted with the MindWave device and the calibration procedure was performed to determine the threshold for detecting deliberate blinks. Then, the eye tracker calibration procedure was performed. Participants were presented with a practice block in each condition so that they could familiarize themselves with the selection mechanism and gaze-based pointer control.

After each condition, each participant was asked to complete an exit survey, in which they assessed how reliably the system detected their intention to perform a selection, their level of exhaustion, overall control, reliability, comfort, and ease of use (see Table 1 for details). The ordering of conditions was counterbalanced across participants. After the participant had completed all three conditions, they were asked to rank the three selection mechanisms based on their overall preference.

Analysis and Design

We performed a within-subjects experiment with Click mechanism (Dwell, Eye Tracker Blink, EMG Blink) as the only factor.

We collected two performance measures:

- Trial completion time (measured from the onset of the target to the successful click; including trials with errors)
- Errors (measured as the number of clicks outside of the valid target)

We also collected 8 subjective measures: an overall preference ranking of the three Click mechanisms and seven additional subjective measures (summarized in Table 1), each recorded on a 7-point Likert scale.

We observed that participants who used the Eye Tracker Blink condition prior to EMG Blink, had a hard time executing their blinks quickly in the EMG Blink condition — their exposure to the Eye Tracker Blink condition affected their subsequent behavior. Similarly, participants who interacted with EMG Blink prior to Eye Tracker Blink did not hold their blinks long enough in the Eye Tracker Blink, despite receiving instructions and having an opportunity to practice, as they were subconsciously expecting a click differentiation based on blink strength, not blink duration. Because of this interference, we used a between-subjects design in the analysis of the performance data: for both trial completion time and the error measures, we only analyzed the data from the first experimental condition experienced by each participant.

We used analysis of variance to analyze the timing data. We log-transformed the timing data prior to the analysis to account for the skewed distributions found in such data. We used a generalized linear model with Poisson distribution to analyze Errors.

We used all data from all three conditions to analyze participants' subjective responses. We used Friedman test to analyze the Likert-scale and the ranking data. In the analysis of the subjective responses, we used the Holm's sequentially-rejective Bonferroni procedure [3] to guard against Type I errors due to multiple hypotheses being tested.

Results

Performance

We observed a significant main effect of condition on task completion times ($F_{2,669} = 38.10, p < .0001$). Post hoc Tuckey HSD analysis revealed that all pairs of conditions were significantly different from each other with the EMG Blink being the fastest (M=1760ms), Dwell the second fastest (M=2000ms) and Eye Tracker Blink the slowest (M=2609ms). These results are summarized in Figure 2.

We observed the largest number of clicks outside the target in the Dwell condition (M=0.28 per trial), and fewer in either EMG Blink (M=0.22) and Eye Tracker Blink (M=0.19) conditions, but this effect was not significant ($\chi^2_{(2,N=672)} = 3.91, n.s.$). These results are summarized in Figure 2.

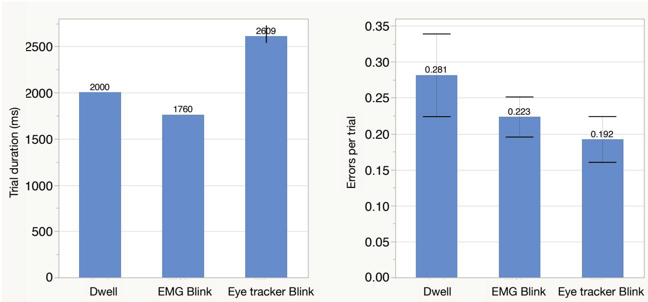

Figure 2. Summary of the performance results. EMG blink resulted in significantly faster overall performance than either Dwell or Eye tracker Blink. There were no statistically significant differences in error rates across the three conditions. Error bars show standard error.

Question	Dwell	EMG	ET	adjusted p
I felt that clicks happened when I did not want them to happen. (7=agree,1=disagree)	4.29	**2.20**	2.32	0.008
How exhausting did you find this method to be? (7=very exhausting, 1=not at all exhausting)	2.71	**2.45**	3.77	0.042
I felt that I had full control over the clicks (7=full control,1=no control)	4.38	**5.90**	5.41	0.100
How reliable did you find this method to be? (7-very reliable, 1=not at all reliable)	4.57	**5.60**	5.00	0.444
I was physically comfortable using this system. (7=very comfortable, 1=not at all comfortable)	4.67	**5.65**	4.86	0.675
It was easy to click using this system. (7=very easy, 1=not easy)	**5.86**	5.85	5.32	1.000
They system did not click at times when I wanted it to. (7=agree, 1=disagree)	3.24	**2.40**	3.55	0.662
How would you rank the systems? (1=best, 3=worst)	2.38	**1.43**	2.05	0.060

Table 1. Summary of the subjective results. Bold font indicates the most positively rated condition for each question. Adjusted p captures p-values adjusted using the Holm's Bonferroni procedure to account for multiple hypotheses being tested.

Subjective Results

On nearly all subjective measures, EMG Blink was rated more positively than either of the two other conditions (these results are summarized in Table 1). The only exception was the Ease of use where EMG Blink was rated minimally below Dwell (M=5.85 for EMG Blink compared to M=5.86 for Dwell). However, only two of these results were statistically significant: participants' perception of the number of false clicks (fewest in EMG Blink, most in Dwell) and of how exhausting it was to complete the task with each interface (least exhausting with EMG Blink, most exhausting with Eye Tracker Blink). The overall preference ranking was marginally significant (EMG Blink ranked highest overall, Dwell lowest).

DISCUSSION AND CONCLUSION

The results of the study show that using an EMG-based mechanism to detect when the user used a deliberate blink to click improved overall performance compared to conditions where 1 second dwell was used to click or where a deliberate blink was detected with a camera (in that condition, the blink had to be held longer than what is natural so that the system could robustly differentiate it from involuntary blinks). Additionally, participants rated the EMG Blink–based more favorably than the alternatives on almost all subjective measures.

One limitation of our approach is our dependence on hardware. The EMG Blink method requires a separate headset. However, the headset that we used is affordable (USD99) and uses a single dry electrode, thus having minimal impact on user comfort.

We envision that the method could, in the future, be integrated with mobile devices such as Google Glass to provide additional robust interaction modality.

REFERENCES

1. Chau, M., and Betke, M. Real time eye tracking and blink detection with usb cameras. Tech. rep., Boston University Computer Science Department, 2005.

2. Glenstrup, A. J., and Engell-Nielsen, T. Eye controlled media: Present and future state. diploma thesis, University of Copenhagen, 1995.

3. Holm, S. A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics 6*, 65-70 (1979), 1979.

4. International Organization for Standardization. 9241-9 Ergonomic requirements for office work with visual display terminals (VDTs)-Part 9: Requirements for non-keyboard input devices, 2000.

5. Jacob, R. J. K. What you look at is what you get: Eye movement-based interaction techniques. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '90, ACM (New York, NY, USA, 1990), 11–18.

6. Kumar, M., Klingner, J., Puranik, R., Winograd, T., and Paepcke, A. Improving the accuracy of gaze input for interaction. In *ETRA '08: Proceedings of the 2008 symposium on Eye tracking research & applications*, ACM (New York, NY, USA, 2008), 65–68.

7. Kumar, M., Paepcke, A., and Winograd, T. Eyepoint: practical pointing and selection using gaze and keyboard. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, ACM (2007), 421–430.

8. Putze, F., Hild, J., Kärgel, R., Herff, C., Redmann, A., Beyerer, J., and Schultz, T. Locating User Attention using Eye Tracking and EEG for Spatio-Temporal Event Selection. In *IUI '13: Proceedings of the international conference on Intelligent user interfaces*, ACM Press (Santa Monica, CA, USA, Mar. 2013).

9. Tuisku, O., Surakka, V., Vanhala, T., Rantanen, V., and Lekkala, J. Wireless face interface: Using voluntary gaze direction and facial muscle activations for human–computer interaction. *Interacting with Computers 24*, 1 (2012), 1–9.

10. Zandera, T. O., Gaertner, M., Kothe, C., and Vilimek, R. Combining eye gaze input with a brain–computer interface for touchless human–computer interaction. *International Journal of Human-Computer Interaction 27*, 1 (Dec 2010), 38–51.