USING EYE GAZE AND A DISAMBIGUATION ALGORITHM TO ENTER WORDS

Barbara Millet[1], Mohan Sridharan[2], and Yulin Wang[1]

[1]Department of Industrial Engineering
[2]Department of Computer Science
Texas Tech University
2500 Broadway
Lubbock, TX 79409
Corresponding author's e-mail: barbara.millet@ttu.edu

Use of eye typing technology has been investigated since the early 1980s (Majaranta & Räihä, 2002). Typically, an eye typing system is equipped with an eye tracker that measures the user's eye behaviors (Majaranta & Räihä, 2007). The eye tracker records the user's eye movements by the corneal-reflection method using a standard desktop computer with an infrared camera mounted beneath a display monitor (Poole & Ball, 2005). The computer system then uses real-time eye behaviors to trigger text entry. Eye typing text entry can greatly benefit the motor impaired, as it allows them to bypass use of traditional mice and keyboards which require limb movements and muscle strength for control.

To assess eye typing performance, word per minute (WPM) is a widely used measure of the eye typing speed. It does not consider the number of keystrokes or gesture during text entry, but only the length of the resulting transcribed string and how long it takes to produce it (Majaranta, MacKenzie, Aula, Räihä, 2006). In contrast to calculating entry speed, measuring accuracy is more challenging. A simple measure of error rate is to obtain the number of characters in error as a percentage of the length of the presented string (Mackenzie & Soukoreff, 2002). However, a more thorough analysis involves categorizing the types of errors. Error rates (ER) are categorized as uncorrected ER, corrected ER, and total ER. Uncorrected errors are those that remain in the in the transcription. Corrected errors are any characters that were backspaced during entry (Soukoreff & Mackenzie, 2003). Corrected errors will not appear in the transcribed string and requires the analysis of the input stream. The Total Error Rate combines both the corrected and uncorrected ERs.

As text entry involves learning, it is important to understand the effects of learning on user performance with any new text entry method. Learning, however, has different effects on entry rates and error rates (Isokoski, 2004). Error rates are typically very high initially, but then either stay the same or quickly fall to more tolerable levels. Entry rates, however, tend to improve following power law of learning (Isokoski, 2004).

Additionally, researchers have proposed a variety of measurements specific to eye typing systems. For instance, Read-Text Events (RTE) counts the changes in gaze direction from the virtual keyboard to the typed text field (Majaranta, Aula, & Räihä, 2004). The ideal value of 0 indicates that the user is confident to proceed without verifying the transcribed text. Similarly, Re-Focus Events (RFE) measures the average number of times a user re-focuses on a key in order to select it. The ideal value of 0 indicates that the user focuses on each key only once (Majaranta et al., 2006). Another measure is the Upper Limit Text Entry Speed. This assesses the upper limit placed by the system on eye typing speed (Ashtinai & MacKenzie, 2009, 2010). Current Information Rate, based on information theory, is used to specify the number of bits (per character) required for a given text (Ward, Blackwell & MacKay, 2002). The factor of time, such as the time to perform a gesture (Drewes & Schmidt, 2007), the time interval between the moment a key was selected and the moment the gaze left the key, as well as the dwell time threshold to trigger text entry, are also measured in many eye typing investigations. These measures, however, tend be used as a function of the eye typing technique employed.

According to the user behaviors they are based on, eye typing techniques can be categorized as fixation text entry, eye gestural text entry, blinks text entry (Ashtiani & MacKenzie, 2010), and continuous gaze text entry (De Luca, Weiss, & Drewes, 2007). When using eye fixation (or dwell time) to enter text, the user keeps his/her attention on the target letter for some time in order to trigger the input. Such systems are typically set up with an eye tracker and an on-screen keyboard. The eye tracker follows the user's gaze and the associated software records and analyses the gaze behavior. The system determines which letter the user wants to type (Majaranta et al., 2006) based on the direction of the gaze. However, most of these systems can only achieve between 6 to 15 WPM. These low text entry rates are a consequence of system limitations. One of the key concerns stems from the required dwell time. The dwell time should be kept at 500±100 ms (Špakov & Miniotas, 2004), as too short a dwell time hinders target searching, while too long a dwell time may increase user fatigue (Majaranta et al., 2006; Špakov & Miniotas, 2004). This reliance on dwell time significantly limits achievable typing speeds. Another key concern centers on screen target sizes. Eye trackers are not perfectly designed so they still have accuracy issues with sensor lags. Our eyes' instability and the "Midas Touch Problem" (Jacob, 1991) will increase difficulties for eye trackers to track the exact area where the user is looking (MacKenzie, 2010). Although bigger targets may be helpful, they will occupy space needed for other on-screen applications. Moreover, it is not natural for humans to hold gaze on a target as we always tend to move away our fixations once we find the target (Jocab, 1991).

To develop systems that can surpass dwell time limitations, researchers have proposed techniques using eye gestural input. It follows the concept that eye gestures can be taken as a pointing device and as an alternative for mouse input (Drewes & Schmidt, 2007). It uses an infrared LED of the eye tracker to create a reflection spot on the eyeball and a video camera to detect this spot. Then, the computer system calculates the eye-gaze direction by measuring the reflection spot and the center of the pupil (Drewes & Schmidt, 2007). For example, EyeWrite (Wobbrock, Rubinstein, Sawyer, & Duchowski, 2008), Eye-S (Porta & Turina, 2008), Quickwriting (Perlin, 1998; Bee & Andre, 2008), and pEYEwrite (Urbina & Huckauf, 2007; Urbina & Huckauf, 2010) are typical systems using this technique. For EyeWrite and Eye-S, the sequences of using eye gestures to "write" a letter are defined and stored in the system. The user has to remember these sequences to enter text. In studies, participants have been able to reach approximately 5 WPM using EyeWrite and 7 WPM when using Eye-S. In contrast, Quikwriting and pEYEwrite do not require that the user remember the sequences. Quikwriting is designed with an interface divided into eight equally sized sections around a central resting area. To type a letter, the user looks the target letter and then moves his/her gaze from the center to one of the outer sections before looking back to the center. pEYEwrite is designed with a hierarchical pie menu. The first level of the hierarchical pie menu contains letters located separately in six sub-areas. Once a sub-area is selected, the next level of pie menu will pop up and a single letter from the selected group can be specified. Novice users of Adaptive Quickwriting (an update of Quickwriting) can achieve 5 WPM. For users of pEYEwrite, average entry rates are 7.4 WPM for novices and 13.5 WPM for experts with error rates of 3.29% to 5.41%.

To facilitate communication for the most severely motor-impaired person, who can only perform extremely limited eye movements, Ashtiani and MackKenzie (2009) proposed a new technique that uses blinks to trigger text entry. A prototype and its updated version named BlinkWrite2 (Ashtiani & MackKenzie, 2010) were developed using a "scanning ambiguous keyboard". For BlinkWrite2, 26 English letters are distributed across three on-screen keys in addition to a space key. Participants achieved speeds of 5.3 WPM with an error rate of 10- 20%. Since this technique was designed specifically for the severely motor-impaired person, it does not emphasize fast typing with high accuracy.

Another eye typing technique described in the literature uses continuous eye gaze. Continuous gaze typing systems do not require the user to fixate on a target for a specific amount of time. Dasher, a typical system using such technique, detects the user's gaze continuously while zooming towards the target letter. It can lead to fast text entry but at the cost of learning time. Ward, Blackwell, and Mackay (2002) claim that when using Dasher participants should be able to achieve up to 25 WPM for novice users and 34 WPM (with less than 5% error rate) for expert users. However, this finding does not necessarily indicate fast entry rates for all eye typing systems using this technique. Tuisku, Majaranta, Isokoski, and Räihä (2008), for example, reported an average typing speed of 17.3 WPM using Dasher. Furthermore, StarGazer (a 3D interface using pan and zoom), leads only to 8.2 WPM with an error rate of 1.23% (Hansen, Skovsgaard, Hansen, & Møllenbach, 2008). User satisfaction is another issue for this technique. For example, novice users claimed that the dynamic interface of Dasher changed so fast that it was overwhelming and even frustrating to some extent (Wobbrock et al., 2008).

In our study, we propose a prototype eye typing system using scan paths to enter words. This technique uses the participant's continuous eye trajectories as trigger signals for word entry. In other words, users enter words using eye gaze, scanning from letter to letter. Compared with the discrete single-letter entry, our technique allows for input of more than one character at a time. We hypothesized that participants will perceive our technique as a fast and natural approach for eye typing. Furthermore, due to participant familiarity with the QWERTY keyboard layout and the relatively short duration of the experiment (30-40 minutes), we expected participant preference for the QWERTY layout over the alphabetical. In our study, we investigate the effects keyboard layout (QWERTY vs. Alphabetical) on participants' eye trajectories and their typing performances, as well as their perceived comfort using our new technique.

In our pilot, 10 voluntary participants (5 male, 5 female) entered 9 words with each of the two keyboard layouts. Eye tracking was provided by the desktop ASL Eye Trac6 unit. Text entry rates with each layout were about 9 WPM. Results indicate that the mean WPM was not associated with keyboard layout. A word by keyboard interaction effect, however, was significant ($F(8,64)=2.079$, $p= 0.05$), with faster entry rates on 5 of the 9 words when using the alphabetical keyboard layout. An algorithm was developed to classify the scan paths and was evaluated against the collected data set. Over the set of experimental trials, the classification accuracy was ~91% with the QWERTY layout (min: 85% and max: 100%); and ~83% with the alphabetical layout (min: 75% and max: 85%). These results indicate that the QWERTY layout may lead to better disambiguation between words.

REFERENCES

Ashtiani, B. & MacKenzie, I. S. (2010). BlinkWrite2: An improved text entry method using eye blinks. *Proceedings of the ACM Symposium on Eye Tracking Research and Applications*, Austin, TX, Mar 22-24, 339-345.

Bee, N., & André, E. (2008). Writing with your eye: A dwell time free writing system adapted to the nature of human eye gaze. *Proceedings of the Perception in Multimodal Dialogue Systems*, Kloster Irsee, Germany, June 16 -18, 111-122.

De Luca, A. D., Weiss, R., & Drewes, H. (2007). Evaluation of eye-gaze interaction methods for security enhanced PIN-entry. *Proceedings of the 19th Australasian conference on Computer-Human Interaction: Entertaining User Interfaces*. Adelaide, Australia, November 28-30, 199-202.

Drewes, H., & Schmidt, A. (2007). Interacting with the computer using gaze gestures. *Proceedings of the Human-Computer Interaction – Interact 2007 11th IFIP TC 13 International Conference*, Rio de Janeiro, Brazil, September 10-14, 2007.

Hansen, D. W., Skovsgaard, H. H. T., Hansen, J. P., & Møllenbach, E. (2008). Noise tolerant selection by gaze-controlled pan and zoom in 3D. *Proceedings of the 2008 Symposium on Eye Tracking Research & Applications,* Savannah, Georgia, March 26-28, 205-212.

Isokoski, P. (2004). Manual text input: Experiments, models, and systems. Unpublished doctoral dissertation, University of Tempere, Finland.

Jacob, R. J. K. (1991). The use of eye movements in human-computer interaction techniques: What you look at is what you get. *ACM Transactions on Information Systems (TOIS)*, 9(2), 152-169.

MacKenzie, I. S., & Soukoreff, R. W. (2002). Text entry for mobile computing: Models and methods, theory and practice. *Human-Computer Interaction*, 17, 147-198.

MacKenzie, I. S., & Ashtiani, B. (2009). BlinkWrite: efficient text entry using eye blinks. *Universal Access in the Information Society*, 10(1), 69-80.

MacKenzie, I. S. (2010). An eye on input: research challenges in using the eye for computer input control. *Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications,* Austin, Texas, Mar 22-24, 11-12.

Majaranta, P., Aula A., & Räihä, K. (2004). Effects of feedback on eye typing with a short dwell time. *Proceedings of ETRA 2004*. San Antonio, Texas, March 22-24, 139-146.

Majaranta, P., MacKenzie, I. S., Aula, A., & Räihä, K. (2006). Effects of feedback and dwell time on eye typing speed and accuracy. *Universal Access in the Information Society,* 5(2), 199-208.

Majaranta, P. and Räihä K. (2002). Twenty years of eye typing: systems and design issues. *Proceedings of the 2002 Symposium on Eye Tracking Research & Applications*. New Orleans, Louisiana, March 25-27, 15-22.

Majaranta, P. and Räihä K. (2007). Text entry by gaze: Utilizing eye tracking. In I. S. MacKenzie and K. Tanaka-Ishii (Eds.), *Text Entry Systems Mobility, Accessibility, Universality* (pp.176-187). CA, US: Morgan Kaufmann.

Perlin, K. (1998). Quikwriting: continuous stylus-based text entry. *Proceedings of the 11th annual ACM Symposium on User Interface Software and Technology*, New York, NY, November 1-4. 215- 216.

Poole, A., & Ball, L. J. (2005). Eye tracking in human-computer interaction and usability research. In C. Ghaoui (ed.), *Encyclopedia of Human Computer Interaction*. Idea Group, Pennsylvania, 211-219.

Porta, M., & Turina, M. (2008). Eye-S: a full-screen input modality for pure eye-based communication. *Proceedings of the 2008 Symposium on Eye Tracking Research & Applications,* Savannah, Georgia, March 26-28, 27-34.

Soukoreff, R. W., & MacKenzie, I. S. (2003). Input-based language modeling in the design of high performance input systems. *Proceedings of Graphics Interface 2003* (pp. 89-96). Toronto: Canadian Information Processing Society.

Špakov, O. & Miniotas, D. (2004). On-line adjustment of dwell time for target selection by gaze. *Proceedings of the Third Nordic Conference on Human-Computer Interaction*. Tampere, Finland, October 23-27, 203-206.

Tuisku, O., Majaranta, P., Isokoski, P., & Räihä, K. (2008). Now Dasher! Dash away!: Longitudinal study of fast text entry by eye gaze. *Proceedings of the 2008 Symposium on Eye Tracking Research & Applications,* Savannah, Georgia, March 26-28, 19- 26.

Urbina, M. H. and Huckauf, A. (2007). Dwell-time free eye typing approaches. *Proceedings of Annual Conference on Communication by Gaze Interaction 2007,* Leicester, UK, September 3-4, 65- 70.

Urbina, M. H., and Huckauf, A. (2010). Alternatives to single character entry and dwell time selection on eye typing. *Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications*, Austin, Texas, Mar 22-24, 315-322.

Ward, D. J., Blackwell, A. F., & MacKay, D. J. C. (2002). Dasher: A Gesture-Driven Data Entry Interface for Mobile Computing Human-Computer Interaction. *Human-Computer Interaction*, 17 (2), 199- 228.

Wobbrock, J. O., Rubinstein, J., Sawyer, M. W., & Duchowski, A. T. (2008). Longitudinal evaluation of discrete consecutive gaze gestures for text entry. *Proceedings of the 2008 Symposium on Eye Tracking Research & Applications,* Savannah, Georgia, March 26-28, 11-18.