Swipe&Switch: Text Entry Using Gaze Paths and Context Switching

Andrew Kurauchi
Insper
São Paulo, SP, Brazil

Wenxin Feng
Boston University
Boston, MA, USA

Ajjen Joshi
Affectiva
Boston, MA, USA

Carlos H. Morimoto
University of São Paulo
São Paulo, SP, Brazil

Margrit Betke
Boston University
Boston, MA, USA

ABSTRACT

Swipe-based methods for text entry by gaze allow users to swipe through the letters of a word by gaze, analogous to how they can swipe with a finger on a touchscreen keyboard. Two challenges for these methods are: (1) gaze paths do not possess clear start and end positions, and (2) it is difficult to design text editing features. We introduce *Swipe&Switch,* a text-entry interface that uses swiping and switching to improve gaze-based interaction. The interface contains three context regions, and detects the start/end of a gesture and emits text editing commands (e.g., word insertion, deletion) when a user switches focus between these regions. A user study showed that Swipe&Switch provides a better user experience and higher text entry rate over a baseline, EyeSwipe.

Author Keywords

Eye typing; Text Entry; Gesture-Based Typing; Gaze Swiping; Eye Tracking

CCS Concepts

•**Human-centered computing** → **Text input;** Gestural input;

INTRODUCTION

Text entry by gaze benefits users, such as those with Amyotrophic Lateral Sclerosis (ALS) or immersed in virtual reality (VR), who are not able to manually type on physical or virtual keyboards. Dwell-time-based selection [1, 10, 11, 13, 17], which requires that users hold their gaze on the desired target until it is selected, is one of the most popular designs used to mitigate the so-called Midas touch problem [5].

Gaze gestures are also commonly applied to select keys on a virtual keyboard [2, 3, 4, 15, 16, 18, 19, 20]. Alternatively, Context Switching [12] uses a single saccade between contexts to perform selections. Also relevant to our method are interfaces that are inspired by swipe-based text-entry in mobile devices [6, 7]. Pedrosa et al. [14] introduced Filteryedping, where the user looks at all the letters that compose the word to be typed. EyeSwipe [8] is a text entry method that also uses gaze paths like Filteryedping, but allows dwell-free text entry using dynamic action buttons and reverse-crossing mechanism. In this paper, we introduce Swipe&Switch, a text entry method using gaze *swiping* and *switching*.

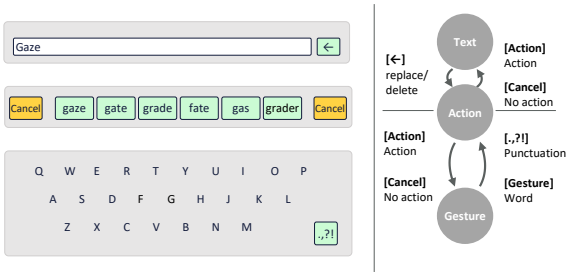

Figure 1. The Swipe&Switch interface is composed of three regions (from top to bottom): text, action, and gesture. The action region changes based on the user's gaze behaviors in the previous region. The previously focused region/buttons are shown in brackets.

SWIPE&SWITCH

Swipe&Switch Interface

Swipe&Switch utilizes switching between different regions to allow explicit starting and ending of a gesture, and text-editing commands in a swipe-based interface. The interface (Figure 1) is composed of three main regions: text, gesture, and action. The design of the three regions can decouple tasks that primarily require gaze *control* (e.g. glancing through the letters in the gesture region) from those that primarily require gaze *perception* (e.g. checking input text in the text region).

The **text region** contains the entered text and a backspace key. The **gesture region** is where the gaze gesture is performed and contains all the characters. The **action region**, placed between the text and gesture regions, is used for command selection and confirmation, and changes dynamically based on where the user previously focused her gaze. With separate regions for different gaze behaviors, the Swipe&Switch design

UIST '20 Adjunct, October 20–23, 2020, Virtual Event, USA

ACM ISBN 978-1-4503-7515-3/20/10.
https://doi.org/10.1145/3379350.3416193

reduces the switching between gaze control and perception modes and enhances the eye-typing experience.

Swipe&Switch provides two ways to select the beginning of the gaze gesture in the gesture region: (1) the first fixation or (2) any other fixation longer than a hidden dwell-time (set to 700 ms during our experiment). The letters within a certain distance to either of these initial fixations are used as first letter candidates. The last letter is determined based on the distance to the last fixation on the gesture region.

The user emits commands to the interface by switching focus from one region to another. When the user moves from the gesture region to the action region the interface shows either candidate words (computed from the gaze path) or punctuation marks (if the user focused on the punctuation key) in the action buttons. When the user moves from the text region to the action region after focusing on the backspace key, the action buttons show candidates to replace the last word and the option to delete it. When the user moves from the action region to any other region, the interface performs the selected action. The action region also has two cancel buttons, one at each end. No action is performed if either of them is selected.

SWIPE&SWITCH EVALUATION

We evaluated Swipe&Switch by comparing it to EyeSwipe[1], in terms of performance and user experience. 12 university students (4 females, 8 males; ages 19 to 28, average 21) were recruited to participate in the experiment. A Tobii EyeX eye tracker was used to collect the gaze information. The keyboard was displayed on a full screen window on a 22-inch LCD monitor (1920 $\times$ 1080 pixels resolution).The length of the square keys (e.g. character keys in EyeSwipe) was approximately 2 degrees (100 pixels) separated by approximately 0.5 degrees (25 pixels). Participants seated at approximately 70 cm from the screen.

Procedure: The participants were encouraged to memorize and type as fast and accurately as possible as many phrases (from MacKenzie and Soukoreff's phrase dataset [9]) as they could. At the end of the experiment they completed a questionnaire about their subjective feedback on the two typing methods and their basic demographic information.

Design: We used a within subjects design with dependent variable text entry rate (wpm), and independent variables session and method. Each participant performed 32 5-minute sessions, totaling 80 minutes with each interface, aside from a short practice section. Participants performed 8 sessions per day, 4 with each method. Half the participants started with EyeSwipe and the other half with Swipe&Switch. In each day the participant started with a method different from the previous day. The first 4 sessions were performed with the first method and the last 4 sessions with the second.

Experimental Results

Text Entry Rate

Participants were able to enter text, on average, faster using Swipe&Switch compared to EyeSwipe. The mean and maximum text entry rates are shown in Figure 2.

[1]Source code available at: https://github.com/toshikurauchi/eyeswipe

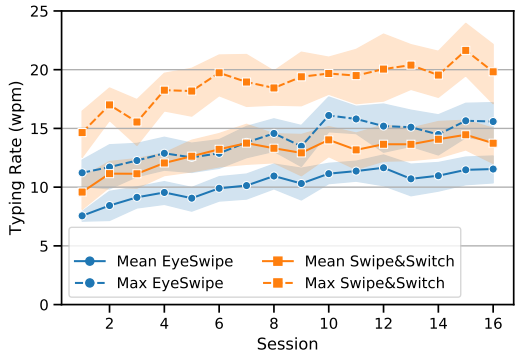

Figure 2. The mean and maximum text entry rates for each session for each interface (with standard deviation bands).

We conducted a within subjects repeated measures ANOVA on the mean (m) and maximum (M) text entry rates with the independent variables method (EyeSwipe and Swipe&Switch) and session (1–16). There were significant main effects of method (m: $F_{1,11} = 153.135$, $p < 0.01$; M: $F_{1,11} = 259.857$, $p < 0.01$) and session (m: $F_{15,165} = 11.071$, $p < 0.01$; M: $F_{15,165} = 7.093$, $p < 0.01$). The interaction between method and session was not significant (m: $F_{15,165} = 0.782$, $p > 0.05$; M: $F_{15,165} = 1.017$, $p > 0.05$).

Subjective Feedback

At the end of the experiment participants answered a questionnaire about their subjective feedback. Participants indicated their perception of performance, learnability, and user experience in a 7-point rating scale (Figure 3). They rated Swipe&Switch higher than EyeSwipe on 5 of the 6 dimensions.

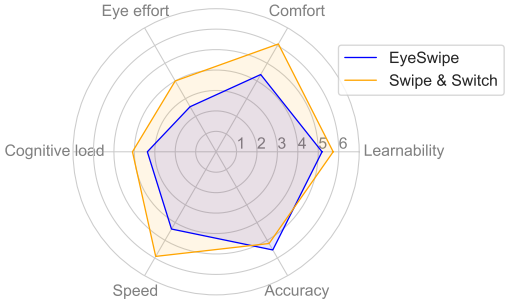

Figure 3. Average responses for all participants to the 7-point rating scale questions. For all six dimensions, the higher the score the better.

DISCUSSION AND CONCLUSIONS

Swipe&Switch uses switching to and from the action region to delimit the swipe gesture. Experimental results suggest that users can enter words approximately 20% faster using Swipe&Switch (13.74 wpm) compared to EyeSwipe (11.54 wpm) on average. Regarding peak velocity, a participant was able to enter a sentence at 33 wpm with Swipe&Switch, compared to 21 wpm with EyeSwipe.Some participants indicated that Swipe&Switch takes practice to get used to. Overall, participants considered Swipe&Switch to be a more "fluid" text entry method.

REFERENCES

[1] Antonio Diaz-Tula and Carlos H. Morimoto. Augkey: Increasing foveal throughput in eye typing with augmented keys. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, pages 3533–3544, New York, NY, USA, 2016. ACM.

[2] Dan Witzner Hansen, Henrik H. T. Skovsgaard, John Paulin Hansen, and Emilie Møllenbach. Noise tolerant selection by gaze-controlled pan and zoom in 3d. In *Proceedings of the 2008 Symposium on Eye Tracking Research & Applications*, ETRA '08, pages 205–212, New York, NY, USA, 2008. ACM.

[3] Anke Huckauf and Mario Urbina. Gazing with peye: New concepts in eye typing. In *Proceedings of the 4th Symposium on Applied Perception in Graphics and Visualization*, APGV '07, pages 141–141, New York, NY, USA, 2007. ACM.

[4] Poika Isokoski. Text input methods for eye trackers using off-screen targets. In *Proceedings of the 2000 Symposium on Eye Tracking Research & Applications*, ETRA '00, pages 15–21, New York, NY, USA, 2000. ACM.

[5] Robert J. K. Jacob. 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, pages 11–18, New York, NY, USA, 1990. ACM.

[6] Per Ola Kristensson and Keith Vertanen. The potential of dwell-free eye-typing for fast assistive gaze communication. In *Proceedings of the Symposium on Eye Tracking Research and Applications*, ETRA '12, pages 241–244, New York, NY, USA, 2012. ACM.

[7] Per-Ola Kristensson and Shumin Zhai. Shark2: A large vocabulary shorthand writing system for pen-based computers. In *Proceedings of the 17th Annual ACM Symposium on User Interface Software and Technology*, UIST '04, pages 43–52, New York, NY, USA, 2004. ACM.

[8] Andrew T. N. Kurauchi, Wenxin Feng, Ajjen Joshi, Carlos H. Morimoto, and Margrit Betke. Eyeswipe: Dwell-free text entry using gaze paths. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, pages 1952–1956, New York, NY, USA, 2016. ACM.

[9] I. Scott MacKenzie and R. William Soukoreff. Phrase sets for evaluating text entry techniques. In *CHI '03 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '03, pages 754–755, New York, NY, USA, 2003. ACM.

[10] Päivi Majaranta, Ulla-Kaija Ahola, and Oleg Špakov. Fast gaze typing with an adjustable dwell time. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '09, pages 357–360, New York, NY, USA, 2009. ACM.

[11] Päivi Majaranta and Kari-Jouko Räihä. Twenty years of eye typing: Systems and design issues. In *Proceedings of the 2002 Symposium on Eye Tracking Research & Applications*, ETRA '02, pages 15–22, New York, NY, USA, 2002. ACM.

[12] Carlos H. Morimoto and Arnon Amir. Context switching for fast key selection in text entry applications. In *Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications*, ETRA '10, pages 271–274, New York, NY, USA, 2010. ACM.

[13] Martez E. Mott, Shane Williams, Jacob O. Wobbrock, and Meredith Ringel Morris. Improving dwell-based gaze typing with dynamic, cascading dwell times. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI '17, pages 2558–2570, New York, NY, USA, 2017. ACM.

[14] Diogo Pedrosa, Maria Da Graça Pimentel, Amy Wright, and Khai N. Truong. Filteryedping: Design challenges and user performance of dwell-free eye typing. *ACM Trans. Access. Comput.*, 6(1):3:1–3:37, March 2015.

[15] Marco Porta and Matteo Turina. Eye-s: A full-screen input modality for pure eye-based communication. In *Proceedings of the 2008 Symposium on Eye Tracking Research & Applications*, ETRA '08, pages 27–34, New York, NY, USA, 2008. ACM.

[16] Sayan Sarcar, Prateek Panwar, and Tuhin Chakraborty. Eyek: An efficient dwell-free eye gaze-based text entry system. In *Proceedings of the 11th Asia Pacific Conference on Computer Human Interaction*, APCHI '13, pages 215–220, New York, NY, USA, 2013. ACM.

[17] Oleg Špakov and Darius Miniotas. On-line adjustment of dwell time for target selection by gaze. In *Proceedings of the Third Nordic Conference on Human-computer Interaction*, NordiCHI '04, pages 203–206, New York, NY, USA, 2004. ACM.

[18] David J. Ward, Alan F. Blackwell, and David J. C. MacKay. Dasher—a data entry interface using continuous gestures and language models. In *Proceedings of the 13th Annual ACM Symposium on User Interface Software and Technology*, UIST '00, pages 129–137, New York, NY, USA, 2000. ACM.

[19] Jacob O. Wobbrock, James Rubinstein, Michael W. Sawyer, and Andrew T. Duchowski. Longitudinal evaluation of discrete consecutive gaze gestures for text entry. In *Proceedings of the 2008 Symposium on Eye Tracking Research & Applications*, ETRA '08, pages 11–18, New York, NY, USA, 2008. ACM.

[20] Wenge Xu, Hai-Ning Liang, Yuxuan Zhao, Tianyu Zhang, Difeng Yu, and Diego Monteiro. Ringtext: Dwell-free and hands-free text entry for mobile head-mounted displays using head motions. *IEEE transactions on visualization and computer graphics*, 25(5):1991–2001, 2019.