A Prototype Gaze-Controlled Speller for Text Entry

Mindaugas Vasiljevas, Justas Šalkevičius, Tadas Gedminas, Robertas Damaševičius
Software Engineering Department
Kaunas University of Technology
Kaunas, Lithuania
robertas.damasevicius@ktu.lt

Abstract—**Eye typing provides a means of communication that is useful for people with disabilities which prevent using their hands for text entry. In this paper, we describe the development of the prototype gaze-controlled speller and discuss its experimental evaluation. Using scrollable virtual keyboard interface, the text input speed of 1.2 wpm was achieved.**

Keywords—gaze tracking; gaze writing; eye typing; hands-free text entry; speller; assistive technology.

I. INTRODUCTION

Communication is central to the human life and experience. With the rise of the electronic means of communication and internet-based social networks as well as wide-spread use of smartphones and tablet PCs, the role of communication the role of texting has increased significantly. According to one report [1], different types of text-based communication (text messaging, e-mail) are the preferred mode of communication for young people (text messaging – 54%, email – 11% vs. for example, cell-phone call – 38%, face-to-face talk – 33%).

To most people, text entry is a simple action. However, over a billion people are estimated to be living with some form of disability or impairment [2]. To those suffering from physical disabilities or age-related impairments, the text entry task may present a significant challenge. For example, in case of such disabilities as amyotrophic lateral sclerosis (ALS) often lead to complete loss of control over voluntary muscles, except the eye muscles. Today's computer systems are not suitable to be used for such people as the input to computers is still fairly limited to mechanical (keyboard, mouse), audio (speech) and tactile (touchpad) inputs. Inability to use a conventional physical input device, such as the mouse or the keyboard, raises the importance of other input modalities such as eyes for connecting persons with severe motor impairments to the digital world. The design of the hardware and software that enables access of handicapped people to ICT services often fails to take into account the user needs [3]. Such limitations raise barriers for people with major or minor disabilities such as elderly people with motor impairments in benefiting from the use of modern ICT applications. Therefore, a large number of individuals are at risk of becoming excluded from the information and knowledge society [4].

To overcome these barriers, new concepts and methods of human-computer interaction (HCI) must be researched and developed in order to efficiently and effectively address the accessibility problems in human interaction with software applications and services while meeting individual requirements of the users in general. Eye typing has been defined as the production of text using the focus of the eye (aka gaze) as a means of input [5]. Eye typing has been known for 30 years now [6], but recently it has received an increased attention from the researchers with the arrival of affordable eye tracking devices on the market.

Systems using gaze-controlled eye typing may be called as *gaze spellers* (using an analogy to brain-controlled BCI spellers [7]). It is a kind of *assistive technology* [8], specifically designed for the purpose of increasing or maintaining the capabilities of people with disabilities, which can be used in *ambient assisted living* (AAL) environments [9] for people with special needs. It has the general aim of bridging the *digital divide* and providing *universal access* [10] to anyone.

The current research is important as the existing eye typing systems still have many limitations (low entry speed, poor usability, etc.) and even small improvements in the design of such systems can lead to significantly improved life quality of impaired people.

The structure of the remaining parts of the paper is as follows. Section 2 discusses the related work. Section 3 describes the developed prototype of gaze-controlled speller. Section 4 describes the experimental results. Finally, Section 5 presents conclusions and discusses future work.

II. RELATED WORK

Several different eye typing systems have been described in research papers. These systems mainly differ in their approach towards presentation and layout of letters in the user interface. A typical example is presenting an on-screen keyboard. The user has to options for action: looking at the desired letter or key for selecting it, and dwelling (i.e., pausing eye movements for a moment) on it for input. Known examples of such systems are GazeTalk [11], ERICA [12], pEYEwrite [13], and Špakov *et al.* [14].

GazeTalk [11] uses a probabilistic character layout strategy to show only 6 most likely next characters on-screen, while next 6 most likely words predicted from the previous words in the sentence are shown. The users have achieved text entry speed from 4 words per minute (wpm) for character-only input to up to 10 wpm using the most likely-words feature.

In ERICA [12], six large on-screen keys were used instead of an entire keyboard due to limited resolution of the eye tracker. A prediction algorithm allowed to decrease eye-typewriting time by 25%.

pEYEwrite [13] groups the letters together in a hierarchical pie structure. To enter a letter, the user first dwells on the pie slice containing the desired group of letters, then dwells on the desired slice in a popup pie menu. Novice entry rates of 7.9 wpm were reported with a dwell time of 400 ms.

Špakov *et al.* [14] use "scrollable keyboards" where one or more rows are hidden to save space combined with keyboard layout optimization according to letter-to- letter probabilities. The users achieved 8.86 wpm speed for the 1-row keyboard, and 12.18 wpm for the 2-row keyboard, respectively.

Other related works include different kinds of text entry systems using virtual keyboard interface. Methods employed in these systems for increasing input systems can be directly transferred to the gaze speller domain, e.g., predictive keyboard layouts in SoftType [15].

AUK [16] uses a 12-key soft keyboard similar to the one used in mobile phones and supports several different entry modes (1-to-10 key, joystick), various layout configurations for different performance levels; integration with additional performance enhancing techniques, such as text prediction and dictionary or prefix-based word disambiguation.

Alternative interfaces for gaze typing include Dasher. Dasher [17] allows users to write by zooming through a world of boxes. Each box represents an individual letter and the size of a box is proportional to the probability of that letter given the preceding letters. The entry rates for Dasher range between 16–26 wpm [17].

Dwell-free eye-typing interface [18] tracks how simply look at or near their desired letters without stopping to dwell on each letter. The users reached a mean entry rate of 46 wpm on a perfect letter recognizer. While dwell-free eye-typing may be more than twice as fast as traditional eye-typing systems, the working prototype of the system still has to implemented that would deal effectively with entry errors.

SMOOVS [19] utilized smooth-pursuit eye movements combined with a two-stage interface that uses a hexagonal layout of letters. The system had achieved the speed of 4.5 wpm, while the users have complained about low comprehensibility of the interface.

Word/phrase prediction or completion is also widely used [20]. As a word is entered, the stem of the current word is expanded to form a list of matching complete words. The list is displayed in a dedicated region of a user interface allowing the user to select the word early. An example is Filteryedping [21] - a key filtering–based approach for supporting dwell-free eye typing that recognizes the intended word by performing a lookup in a word list for possible words that it can form after discarding none or some of the letters that the user has looked at. It sorts the candidate words based on their length and frequency and presents them to the user for confirmation. The method has achieved the rate of 19.8 wpm.

III. DEVELOPMENT OF GAZE SPELLER

A. Usage scenario

Usually eye-tracking interfaces are designed to imitate operation of a standard pointing device such as a mouse. The gaze tracking system, either head mounted or attached in front of the user then tracks the user's gaze and transforms it to the screen coordinates.

During eye typing, the user first locates the letter on a virtual keyboard by moving his/her gaze to it. The gaze tracking device follows the user's point of gaze while software records and analyses the gaze behavior. For input, the user has to fix his/her gaze at the letter for a pre-defined time interval (aka dwell time). When the dwell time has passed, the letter is selected by the system and users can move on to gaze to the next letter. Feedback is typically shown on both on focus and on selection.

B. Advantages and disadvantages

As an input method, gaze has both advantages and disadvantages. It is easy to focus on items by looking at them and target acquisition using gaze is very fast, given the targets are sufficiently large [22]. However, gaze is not as accurate as the mouse partly due to technological reasons as well as some features of the eye [23]. The size of the fovea and the inability of the camera to resolve the fovea position restrict the accuracy of the measured point of gaze [5].

C. Technical limitations

When humans look at things, their fix their gaze on them for 200 to 600 ms [23]. For a computer to distinguish whether the user is looking at an object, a longer interval longer of time is needed. Usually, 1000 ms is long enough to prevent false selections [22]. While requiring the user to fixate for long intervals allows preventing false selections, this may be uncomfortable for most users.

The dwell time also places an upper limit on eye typing speed, e.g., if dwell time is 1,000 ms, the upper limit for typing speed is 12 words per minute (wpm) (considering that 1 word is equal to about 5 characters, for English text).

D. Accessibility/usability requirements and limitations

Accessibility limitations of eye gaze tracking systems have been formulated by Hansen *et al.* [11] as follows:

1. A large portion of the users is not able to get a sufficiently good calibration due to false reflections from glasses, occlusion by eyelids or eyelashes, interference with the ambient light, or low contrast between iris and the pupil.

2. Gaze tracking systems usually require that the user does not move. It is very difficult for most people and impossible for people with involuntary, e.g. spastic, movements.

4. People's eyes tend to dry out due to eye fatigue and long exposure to strong light.

5. Present eye -tracking systems are only for stationary and indoor use.

The requirements for interfaces for impaired users are [24]: 1) Limited access to details: complex and vital details of the system have to be hidden to avoid user overwhelming and trapping. 2) Self-learning: detected common patterns in the behavior of the user should be used to automatically create rules or shortcuts that speed and ease up the use of the system. 3) System interruption: Impaired users have in most cases no idea how the system is working, therefore easy cancellation of system's activities must be ensured.

According to Lopes [25], user interface for persons with disabilities must: support variability allowing to provide the means to adapt to user-specific requirements; support of a wide range of input devices and output modes; provide minimal user interface design; promote interaction and retain user attention on the tasks; and provide strong feedback mechanisms that may provide rewarding schemes for correct results.

E. Architecture

The architecture of the developed prototype gaze speller system is quite simple (see Fig. 1). It consists of the gaze tracking device (Eye Tribe), which is connected to a PC via USB 3.0 connection. On the PC, the core modules are responsible to calibration procedure and gaze feedback.

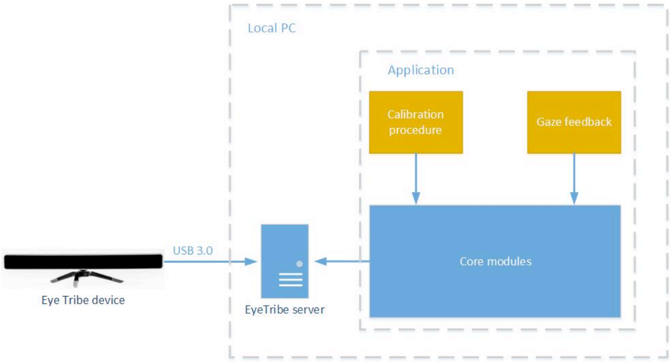
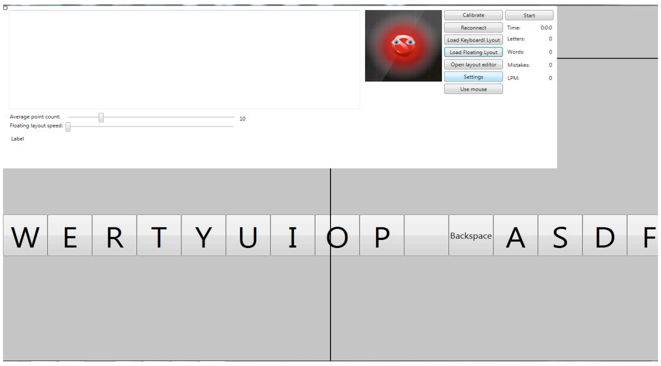


Fig. 1. Architecture of the gaze speller prototype system

F. Interface

The primary driving motive for designing a user interface for a gaze speller is usability as good user experience would also enhance the user acceptance of the system. Our developed interface was inspired by Špakov *et al.* [14] and is based on the concept of "scrollable keyboard" (see Fig. 2).



Time: 0:0:0
Letters: 0
Words: 0
Mistakes: 0
LPM: 0
Average point count: 10
Floating layout speed:
Label

Fig. 2. Interface of the developed gaze speller

Current implementation uses standard QWERTY layout mapped to a single scrollable line of letters. Feedback is ensured by the black line which always stays on the center of the screen while the one-line keyboard moves underneath it depending on the horizontal position of the gaze. Letter selection for input is provided by eye dwelling. Additional menu buttons are provided for calibration, connection to the gaze tracking device, loading of alternative keyboard layouts, and setting program options. Layout editor has been implemented for designing other keyboard layouts.

Finally, the operation of the system can be imitated using a mouse if a gaze tracking device is disconnected.

IV. EXPERIMENTS

A. Apparatus

The eyeTribe eye tracker (tracking range 45cm – 75cm, tracking area 40cm x 30cm at 65cm distance) was connected to a HP Ultrabook notebook running Microsoft 8 OS 64-bit with a Intel Core i5-4202Y 1.60 GHz CPU and 4 GB RAM. The application was displayed on a 14" LCD display with LED backlight and screen resolution of 1920x1080 (see Fig. 3). The eyeTribe eye tracker communicates with notebook via USB 3.0 interface.

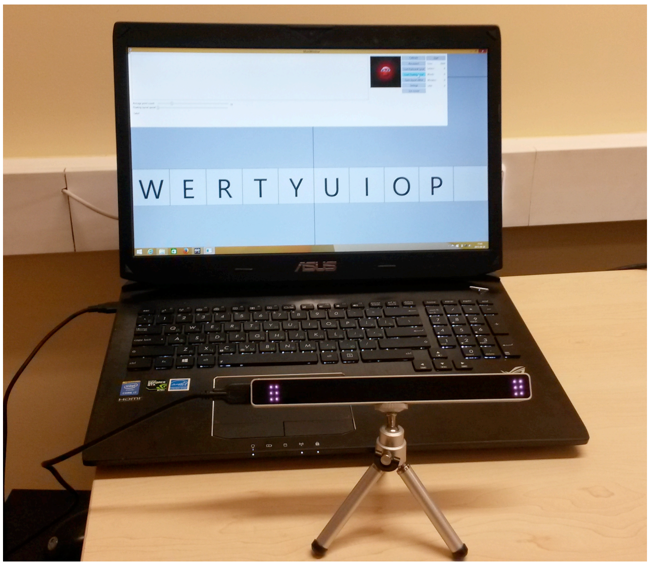


Fig. 3. Deployment of the eye tracking system.

B. Procedure

Prior to collecting data, the experimenter explained the task and demonstrated the software. The experiment was carried out with one disabled person, who could not control his legs and his hand movements are limited. The participant was instructed on the method of text entry, early word selection, error correction, and the audio feedback. He was instructed to enter the given phrases as quickly and accurately as possible and make corrections only if an error is detected in the current or previous word. The participant was allowed to enter a few trial phrases to become familiar with the gaze-controlled selection and correction methods.

For the experiment, we used a fragment of the well-known novel “Alice in Wonderland” by Lewis Carroll (Charles Lutwidge Dodgson):

“*The rabbit-hole went straight on like a tunnel for some way, and then dipped suddenly down, so suddenly that Alice had not a moment to think about stopping herself before she found herself falling down a very deep well.”*

The text consists from 219 characters (including spaces).

A volunteer participant was recruited, who had no prior experience using an eye tracker, to enter the text.

C. Performance metrics

Typing speed is measured in wpm, where a word is any sequence of five characters, including letters, spaces, punctuation, etc. When measuring accuracy, both corrected errors and errors left in the final text are taken into account.

Keystrokes per character (KSPC) measures the average number of keystrokes used to enter each character of text. KSPC is an accuracy measure reflecting the overhead incurred in correcting mistakes.

Error rate is calculated by comparing the text written by the participant with the presented text.

D. Results

The mean for typing speed achieved was 1.2 wpm. This is quite typical for traditional dwell-based eye typing, but is still too slow for fluent text entry. However, the experiment showed that the participant improved with practice over the four blocks of input. The error rate is quite low overall, as the participant generally chose to correct errors during the text entry.

TABLE I. RESULTS OF EXPERIMENT

Speed, wpm	KSPC	Error rate
1.2	1.44	0.01

E. Evaluation

We can compare the input speed of the developed gaze speller with other text typing systems using both traditional and alternative input methods and modalities. Average computer users achieve 33 wpm text entry speed [26] while using standard PC and a keyboard. An average user of the “T9 input method” on a 12-key mobile phone keypad can produce up to 10 wpm [27]. The speed achieved using the Brain Computer Interface (BCI) or Neural Computer Interface (NCI) spellers and electroencephalogram (EEG) / electromyogram (EMG) data as input is in range of 0.2-2.55 wpm, while the eye-blink based EMG speller developed by the authors of this paper achieved 2.4 wpm [28, 29]. Other gaze tracking based text entry spellers report up to 12 wpm speed for dwell-based interfaces [14] and 20 wpm for dwell-free interfaces [21].

The prototype gaze speller described in this speller is still in the early stage of development and its performance is in the lower range of the similar systems. However, there is much space for improvement still left.

V. CONCLUSSION AND FUTURE WORK

This paper has presented a new hands-free text entry system using gaze as the only source of input. Gaze speller is designed to assist the severely motor impaired individuals who are unable to create motion input, but are able to voluntarily control their eyes.

Further research is needed to perform more extensive experiments using a large group of participants (both healthy and impaired), to analyze more efficient letter layouts based on letter frequency and letter/word prediction, to implement adaptive control of dwell time, to evaluate usability of the user interface using common usability evaluation procedures such as NASA-TLX [30], to assess user learnability vs. fatigue with gaze speller in prolonged sessions, and, possibly, integrate several different input modalities (e.g., also using EMG signals) for text entry tasks.

ACKNOWLEDGMENT

The authors would like to acknowledge the contribution of the COST Action IC1303 – Architectures, Algorithms and Platforms for Enhanced Living Environments (AAPELE).

REFERENCES

[1] Interpersonal Communication: A First Look. SAGE. https://us.sagepub.com/sites/default/files/upm-binaries/52575_Gamble_(IC)_Chapter_1.pdf

[2] M. Dawe, “Desperately Seeking Simplicity: How Young Adults with Cognitive Disabilities and Their Families Adopt Assistive Technologies,” CHI '06: Proc. of the SIGCHI conference on Human factors in computing systems, pp. 1143-1152, 2006.

[3] M.C. Domingo, “Review: An overview of the Internet of Things for people with disabilities,” J. Netw. Comput. Appl. 35, 2, pp. 584-596, 2012.

[4] P. Gregor, and A. Dickinson, “Cognitive dificulties and access to information systems: an interac-tion design perspective,” Journal of Universal Access to Information Society, Vol. 05, pp. 393-400, 2006.

[5] P. Majaranta, I. Scott MacKenzie, Anne Aula, and Kari-Jouko Räihä, “Effects of feedback and dwell time on eye typing speed and accuracy”, Universal Access in the Information Society 5(2), pp. 199-208, 2006.

[6] P. Majaranta, and K.-J. Raiha, “Twenty years of eye typing: systems and design issues,” Proc. of the ACM Symposium on Eye Tracking Research and Applications—ETRA 2002, pp. 15–22. ACM, New York, 2002.

[7] I. Scott MacKenzie, and K. Tanaka-Ishii, “Text Entry Systems: Mobility, Accessibility, Universality,” Morgan Kaufmann Publishers Inc., 2007.

[8] A. Gillespie, C. Best, and B. O'Neill, “Cognitive function and Assistive Technology for Cognition: A Systematic Review,” Journal of the International Neuropsychological Society, 18, pp. 1-19, 2012.

[9] A. Dohr, R. Modre-Opsrian, M. Drobics, D. Hayn, G. Schreier, “The Internet of Things for Ambient Assisted Living,” in 7th Int. Conf.on Information Technology: New Generations (ITNG), pp. 804-809, 2010.

[10] C. Stephanidis, G. Salvendy, D. Akoumianakis, A. Arnold, N. Bevan, D. Dardailler, P.L. Emiliani, I. Iakovidis, P. Jenkins, A. Karshmer, P. Korn, A. Marcus, H. Murphy, C. Oppermann, C. Stary, H. Tamura, M. Tscheligi, H. Ueda, G. Weber, and J. Ziegler, “Toward an Information Society for All: HCI challenges and R&D recommendations,” Int. Journal of Human-Computer Interaction 11(1), pp. 1–28, 1998.

[11] J.P. Hansen, D.W. Hansen, and A.S. Johansen, “Bringing gaze based interaction back to basics,” Proc. of HCI International 2001, pp. 325–328. Erlbaum, Mahwah, NJ, 2001.

[12] M. Ashmore, A.T. Duchowski, and G. Shoemaker, “Efficient eye pointing with a fisheye lens,” Proc. of Graphics Interface 2005, pp. 203–210, 2005.

[13] A. Huckauf, and M. H. Urbina,”Gazing with pEYE: new concepts in eye typing,” ACM SIGGRAPH Symposium on Applied Perception in Graphics and Visualization, APGV 2007, p. 141, 2007.

[14] O. Špakov, and P. Majaranta, “Scrollable Keyboards for Casual Eye Typing,” PsychNology Journal 7(2), Special issue on "Gaze Control for Work and Play", pp. 159-173, 2009.

[15] P.E. Jones, “Virtual keyboard with scanning and augmented by prediction,” Proc. of the 2nd European Conference on Disability, Virtual Reality and Associated Technologies, pp. 45–51, 1998.

[16] A. Mourouzis, E. Boutsakis, S. Ntoa, M. Antona, and C. Stephanidis, “An accessible and usable soft keyboard,” Proceedings of HCI International 2007, pp. 961–970. Berlin: Springer, 2007.

[17] O. Tuisku, P. Majaranta, P. Isokoski, and K.-J. Raiha, “Now dasher! dash away! longitudinal study of fast text entry by eye gaze,” Proc. of the 5th ACM Symposium on Eye-Tracking Research & Applications, ACM Press, pp. 19–26, 2008.

[18] P.O. Kristensson and K. Vertanen, “The potential of dwell-free eye-typing for fast assistive gaze communication,” Proc. of the Symposium on Eye Tracking Research and Applications (ETRA '12), ACM, pp. 241-244, 2012.

[19] O. Lutz, A. Venjakob, and S. Ruff, “SMOOVS: Towards calibration-free text entry by gaze using smooth pursuit movements,” Journal of Eye Movement Research 8(1):2, pp. 1-11, 2015.

[20] J. Miro, and P.A. Bernabeu, “Text entry system based on a minimal scan matrix for severely physically handicapped people,” Proc. of the 11th Conference on Computers Helping People with Special Needs—ICCHP 2008, pp. 1216–1219. Springer, Berlin, 2008.

[21] D. Pedrosa, M. da Graça Pimentel, and K.N. Truong, “Filteryedping: A Dwell-Free Eye Typing Technique,” Proc. of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '15). ACM, pp. 303-306, 2015.

[22] C. Ware, and H.H. Mikaelian, “An evaluation of an eye tracker as a device for computer input,” Proceedings of CHI/GI ’87. ACM Press, pp. 183–188, 1987.

[23] R.J.K. Jacob, “Eye tracking in advanced interface design,” in: W. Barfield, and T.A. Furness (eds.), Virtual environments and advanced interface design. Oxford University Press, pp. 258–288, 1995.

[24] A. Marinc, C. Stocklöw, A. Braun, C. Limberger, C. Hofmann, and A. Kuijper, “Interactive personalization of ambient assisted living environments,” Proc. of the 2011 Int. Conf. on Human interface and the Management of Information - Volume Part I (HI'11), LNCS vol. 6771, Springer-Verlag, pp. 567-576, 2011.

[25] J.B. Lopes, “Designing user interfaces for severely handicapped persons,“ Proc. of the 2001 EC/NSF Workshop on Universal accessibility of ubiquitous computing: providing for the elderly (WUAUC'01),” ACM, pp. 100-106, 2001.

[26] C.M. Karat, C. Halverson, D. Horn, and J. Karat, “Patterns of entry and correction in large vocabulary continuous speech recognition systems”, Proc. of the SIGCHI conference on Human Factors in Computing Systems (CHI '99), pp. 568–575, 1999.

[27] A. Cockburn, and A. Siresena, “Evaluating Mobile Text Entry with the Fastap Keypad,” British Computer Society Conference on Human Computer Interaction, pp. 77-80, 2003.

[28] M. Vasiljevas, R. Turčinas, and R. Damaševičius, “EMG Speller with Adaptive Stimulus Rate and Dictionary Support,” Proc. of FeDCSIS'2014: Federated Conference nn Computer Science and Information Systems, pp. 233-240, 2014.

[29] M. Vasiljevas, R. Turcinas, and R. Damasevicius, “Development of EMG-based speller,” Proceedings of INTERRACCION 2014: XV International Conference on Human Computer Interaction.

[30] T. Hayashi and R. Kishi, “Utilization of NASA-TLX for Workload Evaluation of Gaze-Writing Systems,” Proc. of the 2014 IEEE International Symposium on Multimedia (ISM '14), pp. 271-272, 2014.