Technical Correspondence

A Multimodal Gaze-Controlled Virtual Keyboard

Hubert Cecotti

***Abstract*—New portable and noninvasive eye-trackers allow the creation of robust virtual keyboards that aim to improve the life of disabled people who are unable to communicate. This paper presents a novel multimodal virtual keyboard and evaluates the performance changes that occur with the use of different modalities. The virtual keyboard is based on a menu selection with eight main commands that allow us to spell 30 different characters and correct errors with a delete button. The system has been evaluated with 18 adult participants in three conditions corresponding to three modalities: direct selection using a mouse, with the eye-tracker to point at the desired command and a switch to select it, and with only the eye-tracker for command selection. The performance of the proposed virtual keyboard was evaluated by the speed and information transfer rate (ITR) at both the command and application levels. The average speed across subjects was 18.43 letters/min with the mouse only, 15.26 letters/min with the eye-tracker and the switch, and 9.30 letters/min with only the eye-tracker. The later provided an ITR of 44.96 and 57.46 bits/min at the letter and command levels, respectively. The results show to what extent a drop of performance can occur when switching between several modalities. While the speed decreases when controlling the virtual keyboard with the eye-tracker only, the system's performance remains functioning for severely disabled people who have their gaze as one of their only means of communication.**

***Index Terms*—Eye-tracker, graphical user interface (GUI) design, human–computer interaction, performance evaluation, virtual keyboard.**

I. Introduction

Thanks to assistive technology, it is possible to propose new adaptive solutions that can improve the independence of people with disabilities. Assistive technology allows disabled people to perform essential daily tasks, which are necessary to live, work, and communicate with family members and friends. A large number of disabilities such as patients with neurolocomotor disabilities or amyotrophic lateral sclerosis are real challenges for care givers and assistive technology [1]. Patients with severe speech and motor impairment, who are not able to speak nor use sign language, require specific human–computer interfaces to communicate with the world [2], [3]. Depending upon the type of disability, communication devices have to be customized in relation to the constraints imposed by the user, from the adaptation of existing devices (e.g., keyboard, joystick), to the creation of advanced technologies (e.g., brain–machine interfaces in the case of locked-in patients [4]). Disabled people who are able to control their gaze can use their eyes as a means for communication (e.g., for controlling a wheelchair [5], [6]). In addition, the information obtained from the eyes reflects the psychological state of the user [7], and therefore, it can be used to detect fatigue.

Manuscript received June 22, 2015; revised October 27, 2015 and February 23, 2016; accepted February 23, 2016. This paper was recommended by Associate Editor P. Remagnino.

The author is with the School of Computing and Intelligent Systems, Ulster University, Londonderry BT48 7JL, U.K. (e-mail: h.cecotti@ulster.ac.uk).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/THMS.2016.2537749

An eye-tracker is a device that measures the person's eye features (e.g., gaze durations, gaze position, pupil size, saccadic velocities, and saccadic amplitudes [8]) and enables the analysis of a person's gaze. Different eye-tracking solutions have been available for decades [9], [10]. Invasive systems use special head wear (e.g., a head-mounted eye-tracking system [11]) or sensors placed on the subject [12], (e.g., electrooculography [13]), whereas noninvasive eye-trackers have no physical contact with the user. Furthermore, a key issue in interfaces with an eye-tracker is the measure of intention, which can be difficult to interpret. This is due to the amount of involuntary eye movements, which lead to the involuntary selection of items (the Midas touch [14]). A solution to this problem is to add some constraints such as the duration of attention on a particular item. In addition, it is possible to point at an item with the eye-tracker and to select the item with other input devices such as a switch. Another issue to manage with the use of the gaze is its unnatural procedure for selection as the gaze is typically used to only point at an item, i.e., an action is not directly implied. Furthermore, the communication of the observed item is achieved through another sense (e.g., motor action with finger-pointing, the voice).

A large number of methods have been discussed in the literature for gaze estimation, eye-tracking [15]–[19], and head movement detection [20]. Head movement detection can be required when the user can move their head during the experiment as opposed to the use of a chinrest, which is built to establish the position of the observer in psychophysical experiments. In fact, a robust eye-tracker should be able to track not only the gaze, but also the user's head position and orientation in order to avoid errors. These methods include adaptive calibration [21] and a limited number of calibration points [22], [23]. Furthermore, the performance is limited by the system itself, i.e., from eye-trackers using a common webcam to high quality eye-trackers with good optics [24]. Different eye-trackers are available in the market, offering various functionalities with a wide price range and precision. As some studies require a high level of precision to measure eye characteristics, such as saccades [25], [26], a precise eye tracking system is expensive. Due to the price, this type of system has not been developed, and authorities may be hesitant to fund expensive communication devices. Currently, a new market has emerged with relatively inexpensive remote camera-based eye-tracker solutions [27], [28]. With this noninvasive solution, the eye-tracker is placed near the computer screen. Hence, the user does not need to wear special glasses or any other device that includes a camera that points directly to one or both eyes. These inexpensive devices lead, therefore, to new interests in human–computer interaction systems using eye-tracking technology.

Despite pioneer works for the use of eye-tracker as a device for computer input [29], [30], the number of applications using eye-tracker for virtual keyboards remains limited. A reliable eye detection and tracking system is a key component in human–computer interaction as it allows us to better understand visual attention and the users' needs. Virtual keyboards represent the state-of-the-art application for testing novel human–computer interaction systems. Novel and adapted virtual keyboards still need to be developed to help a large number of disabled people to communicate and use a personal computer. When using virtual keyboards to assist a disabled person, the particular characteristics of the disabilities, such as the ability to press a button, or to control gaze, should be taken into account. Various systems have been proposed, in particular, brain–computer interface (BCI) using noninvasive electroencephalography (EEG) recordings has been a fundamental active research field for the development of virtual keyboards that allow severely disabled people to communicate [31]. Recent BCI systems include a combination of different modalities, such as eye-tracking [32]. Noninvasive BCIs based on the detection of event-related potentials, such as the P300, have been the pioneer application in this field. Most of the efficient BCIs require the user to control their gaze to select an item. This requirement is related to the detection of even-related potentials where the graphical user interface (GUI) of the P300 speller presents all the letters on the screen, and the user has to select an item by paying attention to the item [33]. BCIs using steady-state visual evoked potential detection are based on the same principle; the user has to look at an item on the screen in order to select it. BCIs based on motor imagery do not require gaze control to spell. However, the GUI for this type of BCI requires gaze control to read instructions presented on the screen, and to follow the feedback of the application. Despite the recent advances in BCI, the accuracy and the information transfer rate (ITR) remain too low to be widely used. In addition, other constraints include the price and the time that is required to prepare the subject to use a system. The BCI user has to wear an EEG cap or helmet in order to use a system. Overall, human–computer interface systems based on EEG signals are considered more difficult and less reliable approaches than eye tracking. Furthermore, the interest of BCI for disabled people is often justified only for a small population, such as locked-in patients, who have BCI as their only means of communication. However, the ability to control the gaze is least affected by disabilities. In fact, severe disabilities, such as spinal cord injuries, do not affect eye movement. Hence, virtual keyboards based on gaze detection can serve a potentially high number of patients and disabled people (e.g., quadriplegia).

A key challenge to provide a reliable portable and affordable solution for a virtual keyboard based on gaze detection is the ability to take into account the constraints of the eye-tracker and human–computer interaction design. For instance, the layout of a regular keyboard can be an obstacle for its use with an eye-tracker because the proximity of the buttons increases the possible confusion of the gaze detection procedure. In this paper, we propose a virtual keyboard with only eight commands that work as a menu to spell 30 different letters. The system includes word completion and an additional command for corrections. The virtual keyboard requires two consecutive steps for enabling a command. First, the user has to point to the item that must be selected. A pointer on the screen can be moved to the chosen location, or a feedback can be provided on the chosen location. Second, the user has to approve the location of the pointer in order to select the corresponding item to enable a command. The main problem is related to the accuracy of the eye-tracker, which may limit the number of commands that can be accessible at any moment as the calibration data should be updated when the user changes his head and body position over time. The present study focuses on human–computer interaction issues with eye tracking and has the following two main contributions: First, it presents a novel robust virtual keyboard using gaze detection that also includes visual and audio feedback, and its evaluation on 18 healthy subjects; second, we evaluate the performance of the virtual keyboard in three conditions to assess the effect of different types of control on the GUI: the computer mouse, the eye-tracker and a switch, and the eye-tracker only. With only the eye-tracker, the user must gaze at the target item for at least a specific period of time (the dwell time), whereas with the eye-tracker and the switch, the user must gaze at the target item and validate the selection with a button. The remainder of the paper is organized as follows: The proposed virtual keyboard is first described in Section II; then, the experimental protocol is detailed in Section III. The results are presented in Section IV, and their implications are finally discussed in Section V.

II. System Overview

The GUI of the virtual keyboard is composed of two main components: the center of the screen where the text is displayed, and the edge of the screen where the possible commands are displayed. A screenshot of the system is depicted in Fig. 1. The virtual keyboard is based on a tree selection with eight commands (c_1 to c_8) (see Fig. 2). A similar principle was applied in a virtual keyboard using the detection of steady-state visual evoked potentials [34]. The tree has two levels. In the first level, five commands (c_1 to c_5) are dedicated to the selection of the letters: "ABCDEF," "GHIJKL," "MNOPQR," "STUVWX," and "YZ_.?!." Each of these items contains six characters. The command $c6$ ("Undo") allows the user to cancel the previous action (e.g., a deleted character, a selected character). The command $c7$ is used for word completion if the current word contains more than four characters; otherwise, the command does not produce an output. The command $c8$ is used to delete a character. By selecting one of the first five commands, the layout is changed: the commands $c1$, $c2$, $c3$, $c4$, $c5$, and $c7$ contain the letters from the command in the upper level of the tree. For instance, if the user selects the first command "ABCDEF," then the commands $c1$, $c2$, $c3$, $c4$, $c5$, and $c7$ become "A," "B," "C," "D," "E," and "F," respectively.

As the user has to look at the items to select them, it may be possible to forget what is currently written in the message in the middle of the screen. In fact, when an experienced user uses a keyboard to write text, the user does not look at the keyboard but only at the text displayed on the screen. The same

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

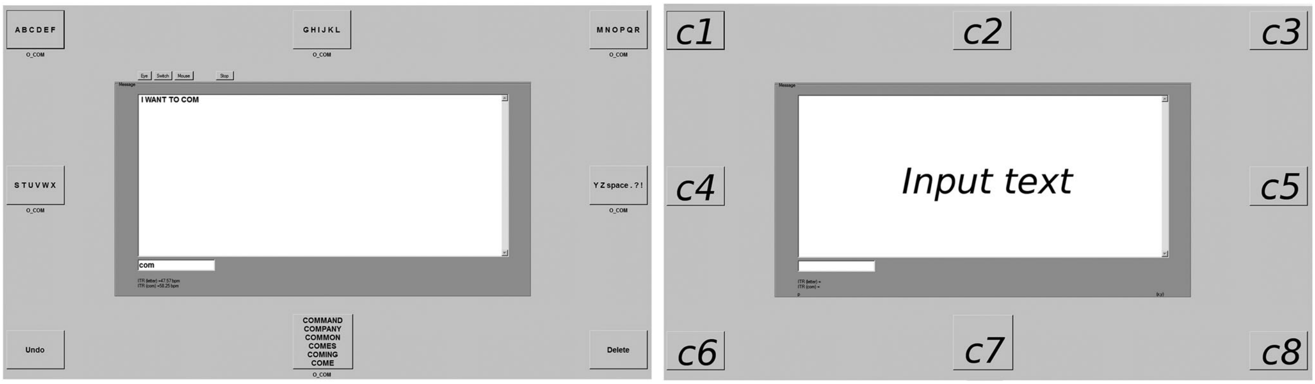

Fig. 1. Representative screenshot of the application (left) with the eight commands (from left to right, top to bottom) and their position (right).

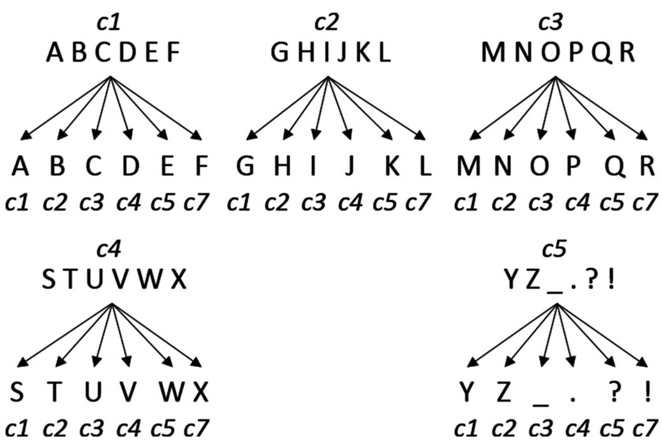

Fig. 2. Tree structure depicting the command tags used for letter selection.

principle applies for the proposed virtual keyboard, but in an opposite way: an experienced user may only pay attention to the items that can be selected, and not the output message that is displayed in the middle of the screen. In order to improve the feedback quality to the user, the last five characters that were written are displayed below each box. This feature is particularly useful during copy spelling when the user has to remember what was already spelled-out without looking at the whole text in the middle of the screen and, hence, changing his gaze location. After the selection of a command, an auditory stimulus is played to highlight the selection of an item. Moreover, visual feedback on the selected item was provided to the user to display the current item selected by the system. Initially, the color of each button representing each command is silver (red $= 192$, green $= 192$, blue $= 192$). When the user pays attention to a button for a duration t, the color changes linearly in relation to the dwell time Δt, and the button becomes greener over time. The color is defined as (red $= v$, green $= 255$, blue $= v$), where $v = 255 \cdot (\Delta t - t)/\Delta t$. During the experiments, the dwell time was set to 2 s. Prior pilot experiments with the author as a subject were carried out with various dwell times (from 0.5 to 2 s). The choice of 2 s was also justified by the novelty of the system to the participants. As it is also important for the subject to have a feedback of the coordinates of the gaze from the eye-tracker, a pointer was continuously displayed on the screen to show the subject the current gaze detection. The goal of this feature was to help the subjects to adapt their position if the detected gaze was not at the expected region on the screen.

The software was written in C# with Microsoft Visual Studio 2012, and the gaze data were acquired with functions from the SDK provided with the eye-tracker [27]. A short calibration was performed prior to each experiment (about 20 s). The virtual speller also includes a word completion system that allows the user to select directly a word after the selection of the first four letters. The prediction is based on the first letters of the current spelled-out word, with a vocabulary obtained from 15 classic free e-books in the Project Gutenberg website [35].

III. Experimental Protocol

A. Subjects

Eighteen healthy volunteer subjects (age $= 28.4 \pm 7.5$, two females, seven with vision correction) participated in the experiments. Subjects did not have prior experience with the application. Subjects with vision correction are reported with the $*$ sign in the subsequent sections. The experimental protocol was reviewed by the Faculty Ethics Filter Committee of Ulster University and was in accordance with the Helsinki Declaration of 1975, as revised in 2000.

B. Materials

The eye-tracker used during the experiments was an Eyetribe [27]. The recorded data were acquired at 30 Hz and contained the coordinates of the gaze and the pupil size for both eyes. Participants were comfortably seated in a chair at about 80 cm from the computer screen (Asus VG248, 24 in, resolution: 1920×1080, 144-Hz refresh rate, 350 cd/m^2). The horizontal and vertical visual angles were approximately 36° and 21°, respectively. Each button on the screen had a size of 4×2.5 cm.

C. Design and Procedure

During the experiment, each subject had to perform two tasks: In the first one, subjects were asked to write a predefined sentence (“THE BLACK CAT EATS A GREEN MOUSE ON THE FLOOR”), which corresponds to 45 characters (90 commands

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

TABLE I
SPELLING PERFORMANCE (MEAN AND STANDARD DEVIATION (SD) ACROSS SUBJECTS) FOR THE MOUSE, THE EYE-TRACKER AND THE SWITCH FOR THE SELECTION, AND THE EYE-TRACKER ONLY

						Time (ms)								
Condition		Speed (letter/min)	ITR (com)	ITR (letter)	R	$c1$	$c2$	$c3$	$c4$	$c5$	$c6$	$c7$	$c8$	all
Mouse	Mean	18.43	106.29	85.39	1.02	1459	1589	1436	2137	1887	3619	2225	4684	1661 ± 971
	SD	3.62	18.92	15.75	0.03	361	283	324	566	440	1515	738	3161	338 ± 231
Eyetracker	Mean	15.26	101.14	71.72	1.17	1775	1680	1479	2236	1843	4019	2312	2674	1810 ± 1031
+ Switch	SD	3.97	23.63	18.70	0.10	497	547	506	519	524	1869	641	1102	478 ± 229
Eyetracker	Mean	9.30	57.46	44.96	1.05	3023	3036	2940	3715	3083	5188	3387	5129	3116 ± 1241
only	SD	1.07	5.25	5.18	0.07	335.25	509	343	784	382	1790	872	1169	291 ± 471

if there are no errors); in the second task, subjects were asked to write two words in free spelling mode. We propose to evaluate the system with three conditions: In the first condition, the user is able to use the mouse to click directly on the items that represent a command on the screen. This condition aims at determining the performance that can be obtained with the GUI, without the eye-tracker. In the second condition, subjects use the eye-tracker to determine the item location to be selected and were required to press a button (space bar) to select the item. In this condition, subjects can press a button as soon as the feedback of the detection is visible, i.e., the color of the observed item becomes green. In the third condition, subjects use the eye-tracker only. The eye-tracker is used to determine both the location of the item and the selection of the item if the users keep their gaze on the selection for a predefined duration, i.e., the dwell time. The word completion function was not used during the experiments in order to provide results that are independent of the input text. Prior to the experiments, subjects were asked to avoid moving their head and body during the tests.

D. Command Selection With the Eye-Tracker

Gaze coordinates in the middle of the screen corresponding to the input text were discarded. To select an item, the user had to pay attention to the corresponding box during Δt ms. During this period, the gaze coordinates must remain in the same region, i.e., the Euclidean distance between the target item (the center of the target button) and the gaze coordinate must always be shorter than the distance between other items and the gaze coordinate. If the coordinate changes from one region to another region before Δt ms, then the timer for the selection is reset.

E. Performance Evaluation

The virtual keyboard performance is measured by the number of letters spelled-out per minute, the ITR at the basic command level (ITR_{com}) and at the letter level ($\text{ITR}_{\text{letter}}$) [36], as well as the mean and standard deviation of the time to produce each command. At the basic command level, the number of possible commands is 8 ($M_{\text{com}} = 8$); it corresponds to the number of items that can be selected by the eye-tracker. At the letter level, the number of commands is 30 ($M_{\text{letter}} = 30$) (the 26 characters [A..Z], space, and {“.,” “?,” “!”}. The ITR is based on the total number of actions (basic commands, letters) and the amount of time that is required to perform these commands. By considering an equiprobability for the different commands and letters, the ITR is defined as follows: $\text{ITR}_{\text{com}} = \log_2(M_{\text{com}}) \cdot N_{\text{com}}/T$, and $\text{ITR}_{\text{letter}} = \log_2(M_{\text{letter}}) \cdot N_{\text{letter}}/T$, where N_{com} is the total number of produced commands to spell N_{letter} characters. In addition, we provide the ratio R between the number of produced commands and the optimal number of commands to spell the input text (the double of the number of characters, including spaces, as each character requires two commands), $R = N_{\text{com}}/2N_{\text{letter}}$, $R \geq 1$. With a dwell time of 2 s, the maximum speed, $\text{ITR}_{\text{letter}}$, and ITR_{com} are 15 letters/min, 73.60 bits/min, and 90 bits/min, respectively.

IV. RESULTS

The virtual keyboard performance is presented for the three conditions in Table I for copy spelling. The condition with the mouse provides the best performance with an average speed of 18.43 ± 3.62 letters/min. This condition is used as a baseline to assess the drop of performance that is achieved by changing the mouse, a common computer device that is familiar to all the subjects, to a different type of modality. With the eye-tracker and the button to select an item on the screen, the average speed is 15.26 ± 3.97 letters/min. With the eye-tracker only, the average speed decreases to 9.30 ± 1.07 letters/min. A Wilcoxon signed-rank test was conducted with a Bonferroni correction applied, showing that the mouse only provides a faster speed than the eye-tracker and the button, which is faster than the eye-tracker only ($p <$ 10e-3). The same pattern of performance is observed for $\text{ITR}_{\text{letter}}$, where the performance drops from 85.39 ± 15.75 with the mouse, to 71.72 ± 18.70 with the eye-tracker and button, to 44.96 ± 5.18 bits/min with the eye-tracker only. However, there is no difference of performance between the mouse and the eye-tracker with the switch for ITR_{com}. Moreover, for R, there is also no difference between the mouse ($R = 1.02 \pm 0.03$) and the eye-tracker ($R = 1.05 \pm 0.07$). These results show that subjects make more errors when they use the eye-tracker with the button compared with the two other conditions ($p <$ 10e-2), but these errors do not impact ITR_{com}. Finally, ITR_{com} is always greater than $\text{ITR}_{\text{letter}}$ due to the commands “Undo” and “Delete” that are present to correct errors, and not to spell letters. The average time to produce a command is 1661 ± 971 ms when subjects use only the mouse and increases to 1810 ± 1031 ms when subjects

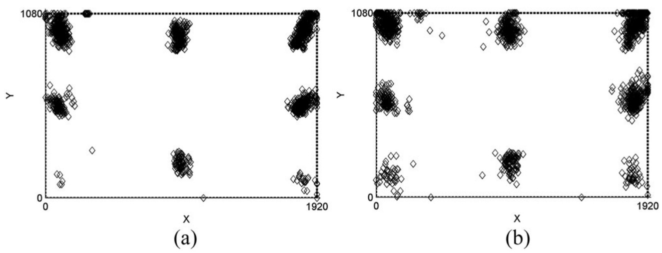

Fig. 3. Coordinates of the decision points obtained from the eye-tracker for all the subjects. (a) Eye-tracker + Switch. (b) Eye-tracker only.

use the eye-tracker with a button. With 3116 ± 1241 ms for the eye-tracker only, it shows that the chosen dwell time of 2 s was on average not too short.

In order to evaluate the gain of speed that may happen when two consecutive commands are on the same location, the different commands were clustered into two groups: those that were identical to the previous command $c_i(t) = c_i(t-1)$ $1 \leq i \leq 8$, and the rest. Pairwise comparisons in the three conditions with a Wilcoxon signed-rank test confirmed a significant difference in the three conditions between the two groups of commands ($p < 10$e-4). The average time to produce a command at the same location as the previous one was 1476 ± 392, 1566 ± 428, and 2731 ± 356 ms for the mouse, the eye-tracker and the switch, and the eye-tracker only, respectively, compared with 1698 ± 333, 1865 ± 505, and 3223 ± 297 ms for other commands. During the free spelling condition, the average number of letters spelled-out was 13.11, with an average speed of 9.48 ± 1.42 letter/min, $\text{ITR}_{\text{com}} = 57.46 \pm 5.25$ bits/min, and $\text{ITR}_{\text{letter}} = 44.96 \pm 5.18$ bits/min. Posthoc analysis did not reveal any difference between free spelling and copy spelling. The corresponding decision points, i.e., the coordinates on the screen that were selected by the subjects, are presented in Fig. 3. They show the relative stability of the gaze detection and a sufficient precision from the eye-tracker with the proposed GUI.

V. Discussion and Conclusion

The performance of a novel virtual keyboard using gaze control with a portable noninvasive eye-tracker has been presented. Three conditions were tested to quantify the change of performance in relation to the modality used. As it was expected, there is significant difference across conditions, with a drop of performance when using the eye-tracker. However, the performance remains high enough (about 10 letters/min) to be used efficiently. In addition, the design of the GUI optimizes the distance between items, hence increasing the robustness of the method to various changes of the head position and the body. Virtual keyboards are difficult to evaluate as the performance depends on the subjects who use the system, their motivation and experience, and the amount of text that is spelled-out during the experiments. In order to avoid the variation of performance related to word completion [37], the system was only tested on the base of two commands per letter. As the present system aims at improving the communication means of disabled people, further experiments will be required to evaluate the performance with patients who may benefit most from the proposed system. When the eye-tracker is only used to point at an item, the space bar can be replaced by any other switch (e.g., eye blinking, a pedal, the detection of voluntary brain responses [38]).

The present multimodal interface does not include head position and orientation detection. This had an effect on gaze detection and the overall performance because subjects did not use a chinrest, and some of them would change their head position throughout the experiment. It is, in fact, natural to slightly orient the head toward the desired item when they are on the left or right side of the screen. An optimal visual angle may exist that takes into account the different items and the potential head movements: a large angle would assure a higher discriminant power between the different locations, but it may increase the amount of head movements. The addition of this information would improve the robustness of the detection as subjects often change their position on the chair despite the request to be in a steady position. For new users, a short dwell time will increase the Midas Touch effect because subjects first need to learn the position of the different items before knowing where to look at to select them. A novice user may be gazing at the item for reasons other than to enter it. Furthermore, a subject mentioned a high blink rate as a side effect of his corrective lenses. In such a case, blinking should not reset the time duration needed to obtain a decision.

In terms of performance, virtual keyboards based on brain response detection that require gaze control, such as the P300 speller and steady-state visual evoked potentials, offer significantly lower performance than the proposed system. The ITR is only about 25 bits/min with P300 spellers [39], and 37.62 bits/min, which is translated in the speller with an average speed of 5.51 letters/min [34], while the proposed system offers an average speed of 9.30 letters/min. The graphical design of an efficient virtual keyboard is a challenging task as it is required to take into account the constraints of the eye-tracker, such as its accuracy and speed. The number of commands can be increased to allow a larger number of letters and the addition of digits; however, a tradeoff must be chosen between the accuracy and the number of items that can be selected at any moment. The current application layout displays the letters from left to right, top to bottom, when a box of letters is split into different boxes. As the selection of a letter requires two commands, and the results have demonstrated a significant decrease of time for two consecutive commands on the same screen location, the layout can be optimized to show at the same location letters that have a high probability to be spelled-out. For instance, the letter "E" has a higher probability than "A" (12.702% for "E" versus 8.167% for "A" [40]); thus, the command "ABCDEF" and "E" could be placed at the same location on the screen. Further analysis should be carried out to determine if the adaptation of the layout has a direct impact on the speed and the gaze path.

In this paper, we have proposed a novel virtual keyboard based on a tree selection that includes word completion, visual, and audio feedback. The system can be used as a benchmark to evaluate different eye-tracking approaches. The system can be used with three modalities: with the mouse, the eye-tracker and a button, and the eye-tracker alone for both pointing at an item and its selection. The system was designed to take into account issues that can occur with a portable eye-tracker, offering a competitive performance with other assistive technology devices. The evaluation on 18 healthy adult participants has demonstrated the robustness of the system in both free spelling and copy spelling. Further works will include an extended menu selection with digits to write numbers and additional items to increase the possibilities of the system. Finally, a new study will be carried out to evaluate the performance with disabled people who will directly benefit from the application.

References

[1] R. Lupu, R. Bozomitu, F. Ungureanu, and V. Cehan, "Eye tracking based communication system for patient with major neuro-locomotor disabilities," in *Proc. 15th Int. Conf. Syst. Theory, Control, Comput.*, Oct. 2011, pp. 1–5.

[2] V. Raudonis, R. Simutis, and G. Narvydas, "Discrete eye tracking for medical applications," in *Proc. 2nd Int. Symp. Appl. Sci. Biomed. Commun. Technol.*, 2009, pp. 1–6.

[3] H. A. Caltenco, L. N. Andreasen Struijk, and B. Breidegard, "Tonguewise: Tongue-computer interface software for people with tetraplegia," in *Proc IEEE Eng. Med. Biol. Soc.*, 2010, pp. 4534–4537.

[4] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, "Brain-computer interfaces for communication and control," *Clin. Neurophysiol.*, vol. 113, pp. 767–791, 2002.

[5] Y. Matsumoto, T. Ino, and T. Ogasawara, "Development of intelligent wheelchair system with face and gazebased interface," in *Proc. 10th Int. Workshop IEEE Robot Human Interactive Commun.*, 2001, pp. 262–267.

[6] D. Purwanto, R. Mardiyanto, and K. Arai, "Electric wheelchair control with gaze direction and eye blinking," *Artif. Life Robot.*, vol. 14, pp. 397–400, 2009.

[7] M. A. Just and P. A. Carpenter, "Eye fixations and cognitive processes," *Cognitive Psychol.*, vol. 8, pp. 441–480, 1976.

[8] R. A. Abrams, D. E. Meyer, and S. Kornblum, "Speed and accuracy of saccadic eye movements: Characteristics of impulse variability in the oculomotor system," *J. Exp. Psychol., Human Perception Performance*, vol. 15, no. 3, pp. 529–543, 1989.

[9] L. R. Young and C. Sheena, "Methods and designs: Survey of eye movement recording methods," *Behavior Res. Methods Inst.*, no. 5, pp. 397–429, 1975.

[10] D. W. Hansen and Q. Ji, "In the eye of the beholder: A survey of models for eyes and gaze," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 3, pp. 478–500, Mar. 2010.

[11] R. G. Lupu, F. Ungureanu, and V. Siriteanu, "Eye tracking mouse for human computer interaction," in *Proc. 4th IEEE Int. Conf. E-Health Bioeng.*, 2013, pp. 1–4.

[12] J. Gips, P. DiMattia, F. X. Curran, and P. Olivieri, "Using EagleEyes—An electrodes based device for controlling the computer with your eyes to help people with special needs," in *Proc. 5th Int. Conf. Comput. Helping People Special Needs*, 1996, pp. 77–84.

[13] H. Manabe, M. Fukumoto, and T. Yagi, "Direct gaze estimation based on nonlinearity of EOG," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 6, pp. 1553–1562, Jun. 2015.

[14] Aristotle, *Aristotle in 23 Volumes*, vol. 21. Cambridge, MA, USA: Harvard Univ. Press, 1944.

[15] E. Guestrin and M. Eizenman, "General theory of remote gaze estimation using the pupil center and corneal reflections," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 6, pp. 1124–1133, Jun. 2006.

[16] W. Hotrakool, P. Siritanawan, and T. Kondo, "A real-time eye-tracking method using time-varying gradient orientation patterns," in *Proc. Int. Conf. Electr. Eng., Electron. Comput. Telecommun. Inf. Technol.*, 2010, pp. 492–496.

[17] C. S. Lin, C. H. Lin, Y. L. Lay, M. S. Yeh, and H. C. Chang, "Eye-controlled virtual keyboard using a new coordinate transformation of long and narrow region," *Optica Applicata*, vol. 38, no. 2, pp. 481–489, 2008.

[18] S. W. Shih and J. Liu, "A novel approach to 3-d gaze tracking using stereo cameras," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 34, no. 1, pp. 234–245, Feb. 2004.

[19] A. Villanueva and R. Cabeza, "Models for gaze tracking systems," *EURASIP J. Image Video Process.*, vol. 2007, p. 23570, 2007.

[20] A. Al-Rahayfeh and M. Faezipour, "Eye tracking and head movement detection: A state-of-the art survey," *IEEE J. Transl. Eng. Health Med.*, vol. 1, pp. 2168–2372, 2013.

[21] Z. Ramdane-Cherif, "An adaptive algorithm for eye-gaze-tracking device calibration," *IEEE Trans. Inst. Meas.*, vol. 57, no. 4, pp. 716–723, Apr. 2008.

[22] A. Villanueva and R. Cabeza, "A novel gaze estimation system with one calibration point," *IEEE Trans. Syst., Man Cybern.*, vol. 38, no. 4, pp. 1123–1138, Aug. 2008.

[23] S. W. Shih, Y.-T. Wu, and J. Liu, "A calibration-free gaze tracking technique," in *Proc. ICPR*, vol. 4, 2000, pp. 201–204.

[24] M. Mehrubeoglu, L. M. Pham, H. T. Le, R. Muddu, and D. Ryu, "Real-time eye tracking using a smart camera," in *Proc. AIPR Workshop*, 2011, pp. 1–7.

[25] R. W. Darwin and E. Darwin, "New experiments on the ocular spectra of light and colours," *Philosophical Trans. Royal Soc. London*, vol. 76, no. 313, pp. 313–348, 1786.

[26] A. E. Bartz, "Eye-movement latency, duration, and response time as a function of angular displacement," *J. Exp. Psychol.*, vol. 64, no. 3, pp. 318–324, 1962.

[27] (2015). The Eye Tribe, Copenhagen, Denmark. [Online]. Available: https://theeyetribe.com/

[28] (2015). Tobii Technology, Danderyd, Sweden. [Online]. Available: http://www.tobii.com/

[29] C. Ware and H. H. Mikaelian, "An evaluation of an eye tracker as a device for computer input," in *Proc. SIGCHI/GI Conf. Human Factors Comput. Syst. Graphics Interface*, 1987, pp. 183–188.

[30] R. J. K. Jacob, "Eye-gaze computer interfaces: What you look at is what you get," *Computer*, vol. 26, no. 7, pp. 65–67, Jul. 1993.

[31] H. Cecotti, "Spelling with non-invasive brain-computer interfaces—Current and future trends," *J. Physiol., Paris*, vol. 105, no. 1ndash;3, pp. 106–114, 2011.

[32] X. Yong, M. Fatourechi, R. K. Ward, and G. E. Birch, "The design of a point-and-click system by integrating a self-paced brain–computer interface with an eye-tracker," *IEEE J. Emerg. Sel. Topics Circuits Syst.*, vol. 1, no. 4, pp. 590–602, Dec. 2011.

[33] L. Farwell and E. Donchin, "Talking off the top of your head: Toward a mental prosthesis utilizing event-related brain potentials," *Electroencephalogr. Clin. Neurophysiol.*, vol. 70, pp. 510–523, 1988.

[34] H. Cecotti, "A self-paced and calibration-less SSVEP based brain-computer interface speller," *IEEE Trans. Neural Syst. Rehab. Eng.*, vol. 18, no. 2, pp. 127–133, Apr. 2010.

[35] (2015). Free ebooks by project Gutenberg. [Online]. Available: https://www.gutenberg.org/

[36] H. Cecotti, I. Volosyak, and A. Graeser, "Evaluation of an SSVEP based brain-computer interface on the command and application levels," in *Proc. 4th Int. IEEE/EMBS Conf. Neural Eng.*, 2009, pp. 474–477.

[37] D. Anson, P. Moist, M. Przywara, H. Wells, H. Saylor, and H. Maxime, "The effects of word completion and word prediction on typing rates using on-screen keyboards," *Assist. Technol.*, vol. 18, no. 2, pp. 146–54, 2006.

[38] Y. Meena, H. Cecotti, K. Wong-Lin, and G. Prasad, "Towards increasing the number of commands in a hybrid brain-computer interface with combination of gaze and motor imagery," in *Proc. 37nd Int. IEEE Conf. Eng. Med. Biol. Soc.*, 2015, pp. 1–4.

[39] G. Townsend, B. K. LaPallo, C. B. Boulay, D. J. Krusienski, G. E. Frye, C. K. Hauser, N. E. Schwartz, T. M. Vaughan, J. R. Wolpaw, and E. W. Sellers, "A novel P300-based brain-computer interface stimulus presentation paradigm: Moving beyond rows and columns," *Clin. Neurophysiol.*, vol. 121, no. 7, pp. 1109–20, 2010.

[40] R. Lewand, *Cryptological Mathematics*. Washington, DC, USA: Math. Assoc. Am., 2000, p. 199.