Development of an Effective Gaze Writing System with an Interface for Text Copy and Paste

Reo Kishi
Graduate School of Science and Technology,
Niigata University, Japan

Takahiro Hayashi
Institute of Science and Technology,
Niigata University, Japan
hayashi@ie.niigata-u.ac.jp

Abstract

While many interfaces for gaze writing (text entry by eye gaze) are available today, these interfaces are just a replacement of a keyboard. For more efficient gaze writing, this paper proposes a novel gaze writing system which has two kinds of gaze writing interfaces: (1) an interface for text entry and (2) an interface for text copy and paste. The interface for text entry was selected from conventional interfaces based on experiments of workload evaluation of beginner users. The interface for text copy and paste, which is newly developed in our study, enables a user to specify a text range and paste the text to a desired position with only gaze interaction. To overcome the lack of accuracy of gaze pointing, the proposed system adopts the model of two step cursor control, which processes initial cursor setting and adjustment in turn. The experimental results have shown that the proposed interface makes gaze writing more efficiently.

Keywords: gaze interface, gaze writing, text copy and paste, subjective workload, eye tracking.

1. Introduction

Recently eye trackers have become interactive, which makes it possible to instantly provide computer-enhanced feedback to someone about what he or she is gazing at. The increased sophistication of eye tracking technologies have generated a great deal of interest in the possibility of gaze interface for ensuring people with disability easily communicate with others.

Gaze writing (text entry by eye gaze) is an promising approach of communication for motor disabled people who cannot either talk or use a keyboard or a mouse. Due to the importance, many gaze interfaces for text entry have been developed [1-4].

Many studies have evaluated gaze-writing systems based on the text entry speed and error rate of mistyping. However, those studies have paid less attention to the workload of a user [5, 6]. Disabled people cannot always handle such an interface properly even if able-bodied persons can do it. Hence, developing a gaze interface considering the mental workloads of users is crucial for practical use.

Besides, the purpose of the conventional interfaces is just a replacement of a keyboard, which limits the usability and flexibility of gaze writing.

From observation of our daily activities with text communication tools such as messaging tools, e-mails applications and notes applications, we have noticed that we perform copy and paste in various situations, for

example, copying a URL, copying the model number of a product to find the specification of the product from the web, copying a part of other person's message for replying e-mail, and so forth. Therefore, it would be beneficial for efficient gaze writing to introduce a gaze interface for text copy and paste.

While copy and paste is a relatively simple operation for ordinary users who can use a mouse, the operation is not so easy for users who cannot use a mouse. In principle, copy and paste with eye gaze could be performed when simply replacing mouse operations such as pointing and dragging with corresponded gaze operations. However, due to the lack of accuracy of gaze pointing, this approach is not practical. For actualizing copy and paste with gaze interaction, it is necessary to design an interface taking into account the physiological characteristics of eye movement.

This paper proposes and evaluates a system for effective gaze writing with support of a text copy and paste interface.

The system is composed of two main interfaces: (1) the interface for text entry and the interface for text copy and paste. The interface for text entry adopted in the system is selected from conventional text entry interfaces based on workload evaluation. The interface for text copy and paste is newly developed. In order to compensate for the lack of accuracy of gaze pointing, the proposed system adopts a model of two step cursor control, which is inspired from the slider control for text selection with finger pointing used in touch screen smartphones.

The rest of the paper is organized as follows. In Section 2, we describe related works, where typical gaze writing systems are reviewed. In Section 3, we experimentally evaluated the workloads of two types of conventional gaze writing systems. In Section 4, we present the proposed gaze writing system. In Section 5, we describe the implementation of our gaze writing system. In Section 6, we show a demonstration of text copy and paste using the proposed system. In Section 7, we show the experiments conducted for confirming the effectiveness of the proposed interface for text copy and paste. Finally, we conclude the paper and describe future works in Section 8.

2. Related Works

Many interfaces for text entry have been developed. These interfaces can be categorized into dwell-time-based interfaces and dwell-time free interfaces [7].

The primary technique used for gaze writing is eye typing. With eye typing, letters are selected from an on-screen keyboard. A virtual keyboard with QWERTY layout is the most common interface [1, 2]. To eye type, the user looks at a letter on an on-screen keyboard. If the user's gaze remains fixed on the same letter for a set time period (the dwell time) the system assumes the user intended to write that letter. The length of dwell time typically ranges from 400 to 1000ms [2]. Shorter the dwell time, the faster text entry. However, too short dwell time will increase the amount of unintended selections. Generally, average text entry speed of eye typing is 6-7 WPM (words per minute).The same mechanism of interaction with dwell time is introduced in our prototype of copy and paste interface to change operation modes.

While Eye typing is a simple approach for gaze writing, the text entry speed is limited by dwell time. To solve the problem, various dwell-time free gaze-writing systems have been developed [3, 4].

Dasher [4] is a well-known dwell time free system. Dasher is also known as a part of the GNOME desktop software in UNIX systems. Dasher was originally developed as a text entry method with a pointing device such as a mouse. As an extension, the mode of gaze inputting was added.

In Dasher interface, characters are vertically displayed on the right side of a window in alphabetical order. When the user gazes at a desired letter, the letter zooms in and moves towards the center of the window. When the letter crosses the center line of the window, the letter is entered (see [4]).

It has been reported that the gaze-writing speed with Dasher by well-trained users reaches about 25 WPM. There is no static element in Dasher, which enables such fast inputting. However, it also requires continuous responses to users. The fully-detailed analysis of the workload of Dasher, especially for beginners, has not been reported.

Conventional text entry interfaces are just a replacement of a keyboard, which limits the usability and flexibility for text communication in various situations as explained in Section 1. For more effective text entry, the proposed system supports text copy and
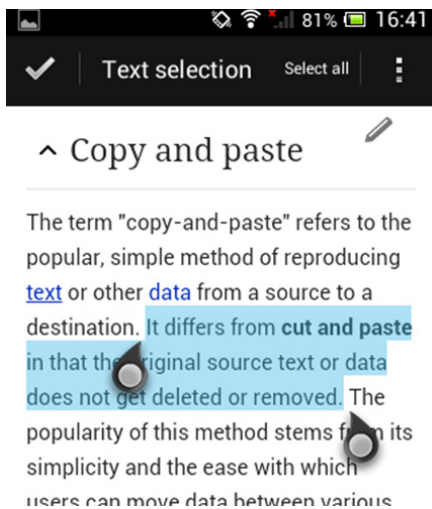
Fig. 1. Conventional copy and paste interface used in touch screen smartphones

paste in addition to key input. To our knowledge, there is no gaze interface dedicatedly designed for text copy and paste. If gaze pointing had a practical accuracy, text copy and paste could be performed easily by replacing mouse operations such as pointing and dragging with corresponded gaze operations. However, the accuracy of gaze pointing is limited because of hardware limitations of eye trackers and characteristics of eye movements like involuntary eye movement, which makes it difficult to copy and paste by eye gaze.

To improve the accuracy of gaze pointing, some studies adopted a multimodal pointing technique combining eye gaze and other input methods such as keyboard typing, mouse button clicking and speech inputs [8-10]. However, these techniques cannot be used for disabled people in the situation that eye gaze is the only interface available for communication.

Interface for locally magnifying the display at the point of the user's gaze have been used for selecting a small target [11-14]. Such magnification interfaces are suitable for selecting sparsely placed objects like icons. However, magnification interfaces are not suitable for text selection because of the disadvantage that information outside the zoomed region is lost. Since it is necessary to select multiple text lines having a wide area for copy and paste operation, the contextual information loss occurs with a high probability, which disrupts efficient text selection.

Gaze pointing and finger pointing have the same problem of the lack of pointing accuracy. Some studies have proposed text copy and paste interfaces on touch screen smartphones [15, 16]. Fig. 1 shows a typical text selection interface used in smartphones. The interface adopts a two-step text selection approach. In this approach, text selection is actualized by long pressing a single word to initiate the selection and dragging the left and right sliders to adjust the selection. Our system applies this two-step text selection approach to text copy and paste based on gaze operations.

3. Workload Evaluation of Gaze Interfaces for Text Entry

The proposed gaze writing system is composed of two gaze interfaces: an interface for text entry and an interface for text copy and paste. In this section, we focus on the interface for text entry adopted in our system. As described in Section 2, many interfaces for text entry have been developed. We chose the best interface from the existing ones based on workload evaluation as follows.

In order to choose the best interface from existing ones, we experimentally compared the workloads of the two typical interfaces, namely, QWERTY on-screen keyboard as a representative of dwell-time-based interface and Dasher as a representative of dwell-time free interface.

To evaluate the workload of the two interfaces, we used NASA-TLX [17]. NASA-TLX, which was developed by NASA Ames Research Center, is a questionnaire-based method for assessing the subjective workload of a user working with human-machine systems. The questionnaire of NASA-TLX is composed of 6 questions regarding workload factors MD (mental demand), PD (physical demand), TD (temporal demand), OP (own performance), EF (effort) and FR (frustration level). Each workload factor is scored by a user on a scale of 0 to 100.

NASA-TLX asks the six workload factors with the six questions. For example, the question about MD is "How much mental and perceptual activity was required?". Since NASA-TLX is designed for using it in various situations, the phrases appearing in the questions have a wide range of meanings. For example, the phrase "mental and perceptual activity" include
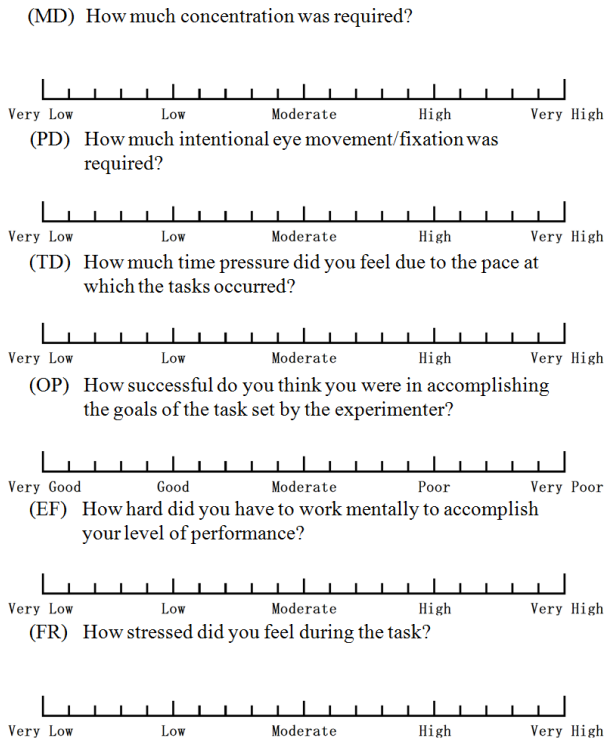
(MD) How much concentration was required?

Very Low Low Moderate High Very High

(PD) How much intentional eye movement/fixation was required?

Very Low Low Moderate High Very High

(TD) How much time pressure did you feel due to the pace at which the tasks occurred?

Very Low Low Moderate High Very High

(OP) How successful do you think you were in accomplishing the goals of the task set by the experimenter?

Very Good Good Moderate Poor Very Poor

(EF) How hard did you have to work mentally to accomplish your level of performance?

Very Low Low Moderate High Very High

(FR) How stressed did you feel during the task?

Very Low Low Moderate High Very High

Fig. 2.Partly Modified NASA-TLX for workload evaluation of gaze writing systems

various concepts of mental and perceptual activities, e.g., "thinking", "memorizing", "concentrating", etc.

For applying NASA-TLX to the workload evaluation in gaze writing, we concretized these questions as shown in Fig. 2. For example, in our question about MD, the expression "mental and perceptual activity" is replaced with more concrete expression "concentration". In addition, in order to give examinees a common criterion for scoring, we attached labels "very low", "low", " moderate", "high" and "very high" to divisions 0, 25, 50, 75 and 100 of the scale, respectively.

In the experiment, 8 able-bodied university students who are novices at gaze-writing systems participated. The first 4 participants, P1, P2, P3 and P4, used a QWERTY on-screen keyboard and the next 4 participants, P5, P6, P7, and P8, used Dasher.

The experiment consists of 8 sessions. In each session, 10 tasks are given to the participants. In each task, one simple English phrase composed of about 300 characters is displayed, where the phrase is randomly selected from a predefined phrase set [18].
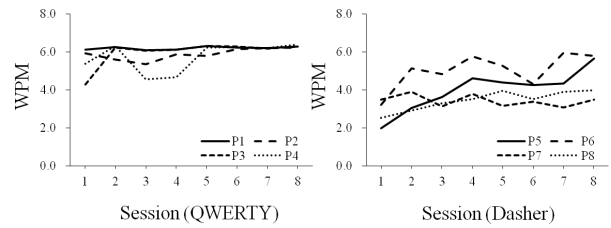

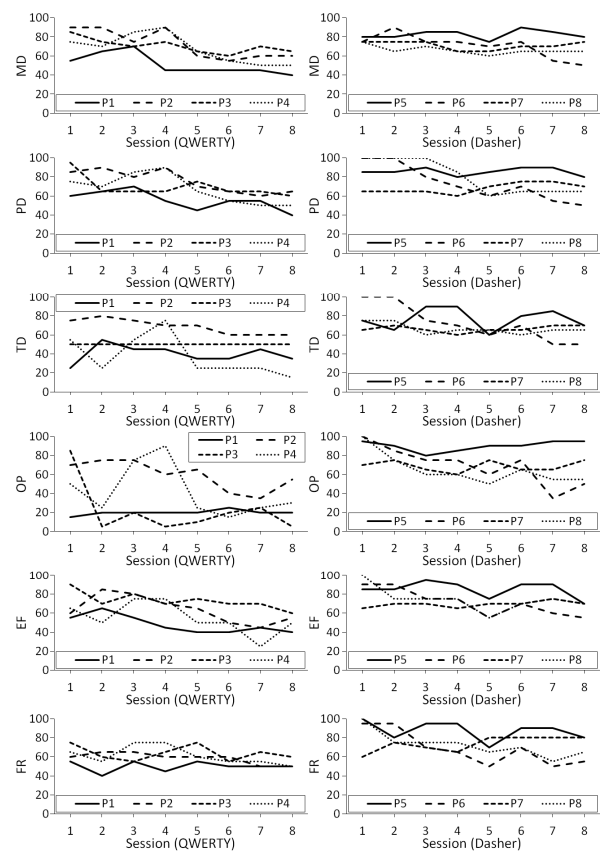Fig. 3. Text entry speeds of QWERTY and Dasher users

Fig. 4. Workloads of QWERTY and Dasher users

The participants input the displayed phrase within 1 minute. If a participant inputs all the phrases correctly within the set time limit, the text entry speed reaches about 6 WPM (words per minute). This paper adopts the rule of 1 word= 5 characters.

Fig. 3 shows the text entry speeds (WPM) of QWERTY users and Dasher users. From the results, we can confirm that totally, the text entry speeds of QWERTY users tend to be higher than Dasher users.

Fig. 4 shows the scores of the 6 workload factors evaluated by the participants. From the results, we can confirm that in any of the 6 workload factors, QWERTY's workloads are lower than Dasher. While there is a report that expert users of Dasher can do texting faster than QWERTY [4], our experimental results indicate that QWERTY is easier to use than Dasher for beginners.

If a user was skillful with it, the workload of Dasher could decrease. However, the high workload is a potential factor that prevents beginners from using Dasher. Similarly, it would be difficult for disabled people to use Dasher.

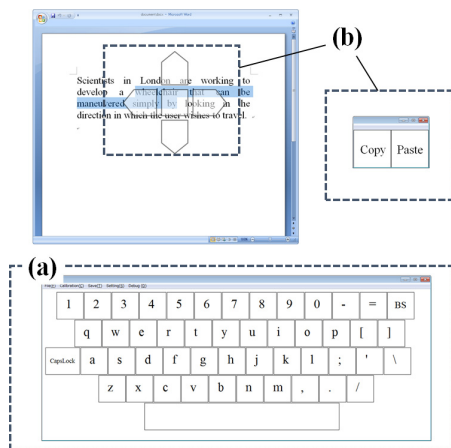## 4. The Proposed System

4.1. *Outline*

Fig. 5. Screenshot of the proposed interface

Fig. 5 shows a screenshot of our gaze writing system. As shown in the figure, the system has an interface for text entry (Fig. 5(a)) and an interface for text copy and paste (Fig. 5(b)). The interface for text copy and paste is newly developed.

4.2. *Interface for text entry*

As shown in the previous section, we experimentally compared the workloads of the two types of interfaces, namely, a QWERTY on-screen keyboard and Dasher. The experimental results indicate that QWERTY on-screen keyboard is the best choice for beginners. From the findings, we adopted a QWERTY on-screen keyboard as the text entry interface of our system.

4.3. *Text copy and paste operations*

In principle, copy and paste could be performed by replacing mouse operations such as pointing and dragging with corresponded gaze operations. However, this approach is not practical because of the lack of accuracy of gaze pointing. Text objects are small and so tightly arranged that the space between objects is extremely narrow. Due to the characteristics of text objects, the operation of text selection is more difficult than selection of other types of objects like icons. For actualizing text selection with gaze pointing, it is necessary to develop a dedicated interface suitable for the style of gaze input. While many interfaces for text entry are available today, to our knowledge, there is no dedicated interface for text copy and paste with gaze interaction. Therefore, we newly developed an interface for text copy and paste.

The operations of copy and paste can be divided into some phases as follows.
(Copy operation)
    (1) Set the text cursor to the start position of the copy range to set mark
    (2) Set the text cursor to the end position of the copy range to select text
    (3) Copy the selected text to the clipboard
(Paste operation)
    (1) Set the text cursor to the paste position
    (2) Paste the copied text to the set position

At the beginning of a copy operation, a user specifies the start and end positions of the copy range to select text. Similarly, in the paste operation, a user specifies the paste position. The common operation needed in these phases is setting a text cursor to a desired position. Therefore, in development of a copy and paste interface, considering how efficiently a user can control a cursor is important.

4.4. *Interface for text copy and paste*

In our system, a user can change the current mode, namely the copy mode or the paste mode, using the interface shown in Fig. 5(b). In the copy mode, a user moves the cursor twice for specifying the start and end positions of the copy range. In the paste mode, the user moves the cursor for specifying the paste position.
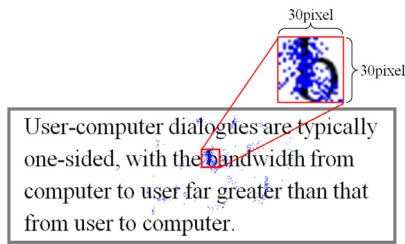
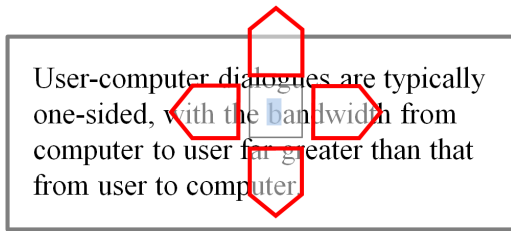Fig. 6. Involuntary eye movement during eye fixation

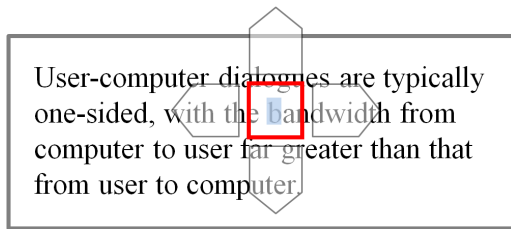Fig. 7. The interface for adjusting the position of the cursor

Fig. 8 The button for specifying the cursor position

Our prototype adopts two step cursor movement model, which is similar to the conventional two step text selection model used in touch screen smartphones. As explained in Section 2, the conventional model requires dragging operations for adjusting the sliders specifying the copy range (Fig. 1). However, drag operations are not suitable for gaze interaction. Therefore, our prototype adopts a model which does not require drag operations for adjusting a cursor.

Our model is composed of two gaze operations: *initial cursor setting* and *adjustment*.

In the step of initial cursor setting, the text cursor is set to the position of the character nearest to the gaze point. However, unlike a mouse that can be released to output a constant coordinate, eye movement is never perfectly stationary and exhibits jitter [19, 20]. Fig. 6 shows fixation jitter occurred by involuntary eye movement. In our prototype, when gaze remains inside the 30×30 pixels area for 1 second, the center of the eye positions is regarded as the position the user wants to point.

To compensate for the inaccurate pointing, as the second step, the cursor position is adjusted with the on-screen cursor keys shown in Fig. 7. The cursor keys are translucently drawn at the current cursor position, by which a user can see the characters under the cursor buttons. The cursor keys can be selected by eye fixation. By selecting a cursor key, a user can move the cursor to the desired direction.

After adjusting the cursor position, the user can perform copy or paste by eye fixation on the center button shown in Fig. 8.

5. Implementation

This section describes the implementation of our prototype. Eye tracking is provided by the Tobii TX300 eye tracker. Tobii TX300 is a 23 inch 1920×1080 pixels LCD monitor integrated with eye tracking optics capable of binocular tracking at $0.4°$ accuracy sampling at 300 Hz. The specification of the PC connected with the eye tracker is Intel Core i7-2600 (3.40Ghz) CPU and 4.0 GB memory and it is running on Windows 7 operating system.

The gaze coordinates are delivered to the PC via a TCP/IP connection. The PC receives real-time gaze data via the TCP/IP connection. The real-time control of eye tracking can be achieved by using Tobii Analytics SDK. Text copy and paste is implemented by generating mouse events and keyboard events via calling Windows API at the timing when individual gaze operations are performed.
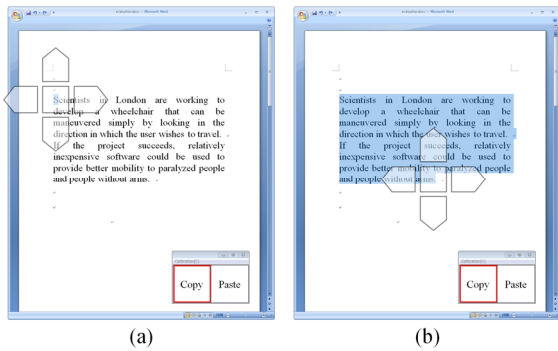
(a)   (b)

Fig. 9. Copy mode. (a) Setting the initial cursor to the start position of a text region to copy. (b) Setting the cursor to the end position of the text region
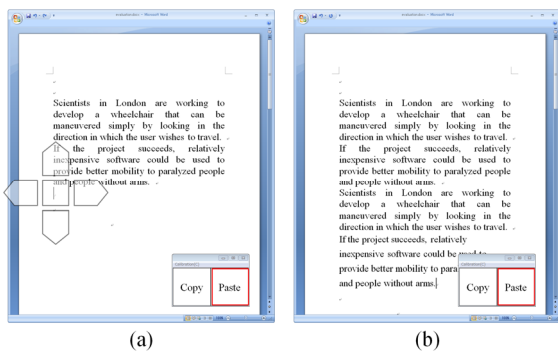

(a)   (b)

Fig. 10. Paste mode. (a) Setting the cursor to the paste point. (b) Result of pasting.

6. Demonstration

Fig. 9 shows the screenshots when actually performing a copy operation with the proposed interface. In the copy mode, a user specifies the start position of the copy range at first as shown in Fig. 9(a). For adjusting cursor position, the interface for virtual cursor key is available. Next, the user specifies the end position of the copy range as shown in Fig. 9(b). By the cursor movement, the user can specify the copy range. Finally, when the user gazes the center button, the text in the copy range is copied to the clipboard.

Fig. 10 shows the screenshots when performing a paste operation with the proposed interface. In the paste mode, the user moves the cursor position to the paste position as shown in Fig. 10(a). After that, the user can perform pasting the copied text to the desired position by gazing the center button as shown in Fig. 10(b).
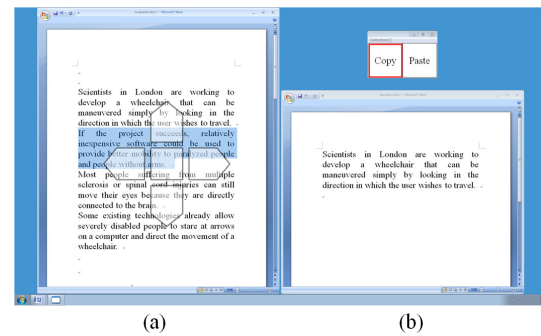

(a)   (b)

Fig. 11. The screenshot of the task windows

Group A	P1	P2	P3	P4	P5
mean elapsed time (sec) (copy operation)	29.00	40.75	24.25	19.00	24.75
mean elapsed time (sec) (paste operation)	4.50	4.75	6.25	4.25	4.75
number of retries	0	0	0	0	1
task completion rate	1.00	1.00	1.00	1.00	1.00

Group B	P6	P7	P8	P9	P10
mean elapsed time (sec) (copy operation)	53.50	62.67	48.50	—	21.50
mean elapsed time (sec) (paste operation)	8.25	3.33	6.50	—	2.75
number of retries	11	14	11	—	2
task completion rate	1.00	0.75	1.00	0.00	1.00

Fig. 12. Experimental results

7. Experimental results

This section describes the experiments conducted for evaluating our system. As explained in Section 4, our system is composed of the interface for text entry and the interface for text copy and paste. Since the evaluation of the interface for text entry has been described in Section 3, in this section we focus on the evaluation of the interface for text copy and paste.

10 able-bodied subjects participated in the experiments for evaluating the interface for text copy and paste. They are university students and novices in gaze interfaces. They are divided into 2 groups, A and B. 5 subjects (P1, P2, P3, P4 and P5) in group A used

the proposed interfaces for text copy and paste which have the adjusting function of cursor position. In group B the other 5 subjects (P6, P7, P8, P9 and P10) control cursor using only eye fixation without the support of adjustment interface.

In a task, each subject is given a text copy and paste task 4 times. In a task, the target text is displayed in the text edit region of a window as shown in Fig. 11(a). The font of the text is 24 pt. Times New Roman. The text is composed of 133 words 4 English sentences. In each task, subjects copied each sentence from the top in turn, and then pasted the copied text to another window as shown in Fig. 11 (b).

In the experiments, we measured the elapsed times from the start to end of copy operations, the elapsed times of paste operations, the total number of retries of copy and paste operations and the task completion rate.

Fig. 12 shows the experimental results. From the comparison of the elapsed times of copy operations between group A and B subjects, we confirm that subjects in group A were able to complete the tasks more quickly than the subjects in group B except for P10. As shown in Fig. 12, subject P9 in group B could not complete the copy and paste task in any sentence due to the difficulty of eye fixation at a desired point. This is because the subject could not point a cursor to a desired position by only eye fixation. The task completion rate of subject P7 was less than 1, which means the subject could not complete some tasks, e.g., deeming the tasks too difficult. Unlike group B, all the subjects in group A could complete all the tasks correctly and quickly.

Besides, from the results that the numbers of retries in group A were greatly less than group B, the model of two step cursor control adopted in our system is effective for efficient gaze writing.

8. Conclusions

This paper proposed a system for effective gaze writing with support of a gaze interface for text copy and paste. In order to deal with the lack of accuracy of gaze pointing, the interface for text copy and paste adopts the model of two step cursor control : initial cursor setting and adjustment. From the experimental results, we have confirmed that the gazed interface for text copy and paste is effective for efficient gaze writing.

Gaze is an indispensable way of communication for motor-disabled people who cannot either talk or use a keyboard or a mouse. However, using gaze as an input device is not so easy due to the lack of accuracy. Therefore, it is crucial to design gaze interfaces considering physiological characteristics of eye movement.

The model of two step cursor control adopted in the proposed system can be applied to other fundamental computer operations such as file selection. Based on the findings of this study, we are planning to develop various gaze interfaces to use a computer more easily and effectively.

Acknowledgements

References

1. P. Majaranta and K. J. Räihä, Twenty Years of Eye Typing: Systems and Design Issues, *Proc. Eye Tracking Research and Applications* (2002), pp. 15-22.
2. R. J. K. Jacob, Eye Movement-Based Human-Computer Interaction Techniques: Toward Non-command Interfaces, *Advances in Human Computer Interaction*, **4** (1993), pp. 151-180.
3. A. Huckauf and M. H. Urbina, Gazing with pEYEs: Towards a Universal Unput for Various Applications, *Proc. Eye Tracking Research and Applications*, (2008) pp. 51--54.
4. T. Outi, M. Paivi, I. Poika and R. Kari-jouko, Now Dasher! Dash Away! Longitudinal Study of Fast Text Entry by Eye Gaze, *Proc. Eye Tracking Research and Applications* (2008), pp. 19-26.
5. K. Räihä, S. Ovaska, An Exploratory Study of Eye Typing Fundamentals: Dwell Time, Text Entry Rate, Errors, and Workload, *Proc. SIGCHI Conf. Human Factors in Computing Systems* (2012), pp. 3001-3010.
6. T. Hayashi, R. Kishi, Workload Evaluation of Gaze-Writing Systems, *Sensors and Materials*, **26**(7) (MYU, Tokyo, 2014), pp 529-538.
7. N. Bee and E. Andre, Writing with Your Eye: A Dwell Time Free Writing System Adapted to the Nature of Human Eye Gaze, *Proc. Workshop Perception and Interactive Technologies for Speech-Based Systems* (2008), pp. 111-122.
8. M. Kumar, A. Paepcke and T. Winograd, EyePoint: Practical Pointing and Selection Using Gaze and Keyboard, *Proc. SIGCHI Conf. Human Factors in Computing Systems* (2007), pp 421-430.
9. M. Yamato, A. Monden, K.i. Matsumoto, K. Inoue and K.

Torii, Button Selection for General GUIs Using Eye and Hand Together, *Proc. Int. Working Conf. Advanced VisualInterfaces* (2000), pp. 270-73.

10. D. Miniotas, O. Spakov, I. Tugoy and I. S. MacKenzie, Speech-Augmented Eye Gaze Interaction with Small Closely Spaced Targets, *Proc. Eye Tracking Research and Applications Symposium* (2006), pp. 67-72.

11. M. Ashmore, A. T. Duchowski and G. Shoemaker, Efficient Eye Pointing with a FishEye Lens, *Proc. Graphics Interface* (2005), pp. 203-210.

12. M. McGuffin, and R. Balakrishnan, Acquisition of Expanding Targets, *Proc. Int. Conf. Human Factors in Computing Systems* (2002), pp. 57-64.

13. D. Miniotas, O. Špakov, I.S. MacKenzie, Eye Gaze Interaction with Expanding Targets, *Proc. Int. Conf. Human Factors in Computing Systems* (2004), pp. 1255-1258.

14. D. Fono and R. Vertegaal, EyeWindows: Evaluation of Eye-Controlled Zooming Windows for Focus Selection, *Proc. Int. Conf. Human Factors in Computing Systems*(2005), pp. 151--160.

15. C. Chen, S. T. Perraut, S. Zhao and W. T. Ooi, BezelCopy: An Efficient Cross-Application Copy-Paste Technique for Touchscreen Smartphones, *Proc. Int. Working Conf. Advanced Visual Interfaces* (2014), pp. 185-192.

16. K. Koaral and T. Komuro, A Zooming Interface for Accurate Text Input on Mobile Devices, *Proc. Int. Conf. Human Factors in Computing Systems* (2013), pp. 1299-1304.

17. S. G. Hart and L. E. Staveland, Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research, *Human Mental Workload*, eds. P.A. Hancock and N. Meshkati (Elsevier Science Publishers, 1988), pp. 139-183.

18. I. S. MacKenzie and R. W. Soukoreff, Phrase Sets for Evaluating Text Entry Techniques, *Extended Abstracts on Human Factors in Computing Systems* (2003), pp. 754-755.

19. R. J. K. Jacob, What You Look at is What You Get: Eye Movement-based Interaction Techniques, *Proc. Int. Conf. Human Factors in Computing Systems* (1990), pp 11-18.

20. G. Buscher, A. Dengel and L. Elst, Eye Movements as Implicit Relevance Feedback, *Proc. Int. Conf. Human Factors in Computing Systems* (2008), pp 2991-2996.