

Development of a Text-Based Communication System with Eye Tracking Technology

Takahiro Hayashi* and Reo Kishi¹

Institute of Science and Technology, Niigata University,
8050 Ikarashi-2-no-cho, Nishi-ku, Niigata-shi, Niigata, 950-2181, Japan
¹Graduate School of Science and Tehnology, Niigata University,
8050 Ikarashi-2-no-cho, Nishi-ku, Niigata-shi, Niigata, 950-2181, Japan

(Received October 26, 2015; accepted February 15, 2016)

Keywords: eye tracking, gaze writing, gaze interface, human sensing, workload

Gaze writing (text entry by eye gaze) is a promising approach to helping people with motor disabilities use a computer and communicate with others. In this paper, we propose a novel gaze-writing system which is equipped with a copy and paste interface. To compensate for the lack of eye tracking accuracy, the proposed interface adopts a two-step selection method to accurately specify a copy text range. By the experimental evaluation of performance and workload, we have confirmed that the two-step selection method can reduce the number of misoperations and the users' workload.

1. Introduction

Eye trackers are devices for electronically locating the point of a person's gaze. Recently, eye trackers have become interactive, which makes it possible to instantly provide feedback to a user about what he or she is looking at. The sophistication of eye tracking technologies has attracted a great deal of interest in the possibility of using a gaze interface to help people with disabilities use a computer and communicate with others.

Owing to the importance of gaze interfaces, various gaze interfaces for text entry have been developed.^(1–4) However, conventional interfaces are just a replacement for a keyboard, which limits the practicality of gaze writing.

Text is an important medium in human-computer communication and human-human communication. When we use text communication tools such as text messaging tools and e-mail applications, we use “copy and paste” in various situations. Although we know the importance of copy and paste for efficient text-based communication, traditional gaze-writing systems are not equipped with a copy and paste interface. It would be therefore, beneficial to develop a gaze interface for text copy and paste to make gaze-writing systems more practical.

Copy and paste is a relatively simple operation for ordinary users who can use a mouse. The operation, however, is difficult for disabled people who cannot use a mouse, especially when gaze is the only way of communicating. In principle, copy and paste could be performed if mouse pointing (for setting a text cursor to the head of a copy range) and drag operations (for selecting text) were replaced by corresponding gaze operations. However, the accuracy of gaze pointing is

*Corresponding author: e-mail: hayashi@ie.niigata-u.ac.jp

still limited because of hardware limitations of eye trackers and physiological characteristics of the eye such as involuntary eye movements. In addition, eye trackers intrinsically cannot distinguish drag operations from just moving operations because they can only track gaze position. To realize copy and paste by eye gaze, it is necessary to develop a software interface that can emulate drag operations and has a mechanism that compensates for the lack of pointing accuracy.

In our previous study,⁽⁵⁾ we developed a gaze-based interface prototype for copy and paste. To compensate for the lack of gaze pointing accuracy, the proposed interface adopts a two-step selection method which is based on the slider interface used in touch-screen smartphones.

In this paper, we mainly focus on the details of the implementation of the proposed interface and on the new findings obtained from the experimental evaluation of the performance and workload of the proposed interface.

The rest of the paper is organized as follows. In § 2, we describe related work, in which we describe conventional studies of gaze interfaces. In § 3, we outline the proposed copy and paste interface. In § 4, we explain the details of the implementation of the proposed interface. In § 5, we present the results of experiments conducted to verify the effectiveness of the proposed interface. Finally, we conclude the paper and describe future work in § 6.

2. Related Work

There are some studies that aimed at overcoming the limitations of gaze pointing accuracy.^(6–8) These studies combine a gaze interface and another interface such as a keyboard, a mouse, or a voice interface. These interfaces, however, cannot be used for motor-disabled people.

Magnification interfaces, which can partially magnify the desktop, are used for selecting a small target.^(9–12) They are practical for selecting small objects that are sparsely placed like icons. They are, however, not suitable for selecting text because they magnify only a particular area of a document and, at the same time, they hide other parts of the document. The contextual information loss disrupts efficient text selection.

The problem of lack of pointing accuracy is also discussed in the field of touch-screen interfaces.^(13,14) Some studies have proposed text copy and paste interfaces for touch-screen smartphones. Figure 1 shows a typical text selection interface used in Android smartphones. This interface adopts a two-step text selection method. In this method, text selection is realized by pressing a single word for a time to initiate the selection and dragging the left and right sliders to adjust the selection. Although the method is effective on a touch-screen interface, it is difficult to simply apply the same mechanism to gaze-based interfaces because it needs drag operations to control the sliders. As

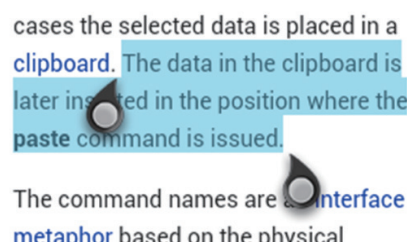


Fig. 1. (Color online) Selection interface used in touch-screen smartphones.

mentioned in § 1, drag operations are not suitable for gaze-based interaction. To solve this problem, our proposed interface adopts a selection mechanism that does not require any drag operations to adjust a text range.

As criteria for evaluating the usefulness of a gaze interface, many studies mainly focus on the performance characteristics of the interface, such as text entry speed. In addition to performance, in this study we evaluate subjective workloads of users.

As workload evaluation methods, the subjective workload assessment technique (SWAT),⁽¹⁵⁾ and NASA task load index (NASA-TLX)⁽¹⁶⁾ have been widely used in human-interface studies.^(17–20) Colle and Reid found that SWAT can more sensitively evaluate changes in difficulty of a task.⁽²¹⁾ These characteristics would be an advantage for using SWAT. However, SWAT requires the user to do a time-consuming pretask called “card sorting”. The purpose of this pretask is to model a user-dependent total workload scale unifying various workload factors. As pointed out in the literature,⁽²²⁾ it takes 30–60 min to finish the pretask. The burden for accomplishing the pretask is a disadvantage of using SWAT.

The NASA-TLX, on the other hand, does not require such pretasks. The NASA-TLX is a questionnaire-based workload evaluation method. It is composed of six questions about six workload factors: mental demand (MD), physical demand (PD), temporal demand (TD), own performance (OP), effort (EF), and frustration level (FL). The user rates each workload factor on a scale from 0 to 100. The total workload score is calculated by a weighted sum of the six scores. Generally, the simple average of the six scores is used as the total workload score.⁽²³⁾ Because it is easy to handle, we use the NASA-TLX in this study for workload evaluation.

3. Proposed Interface

In this section, we outline the proposed gaze interface for copy and paste. Figure 2 shows a screenshot of the interface. With the interface, the user can change the current operation mode (copy/paste mode) [Fig. 2(a)]. In the copy mode, the user can control the text cursor to specify a copy

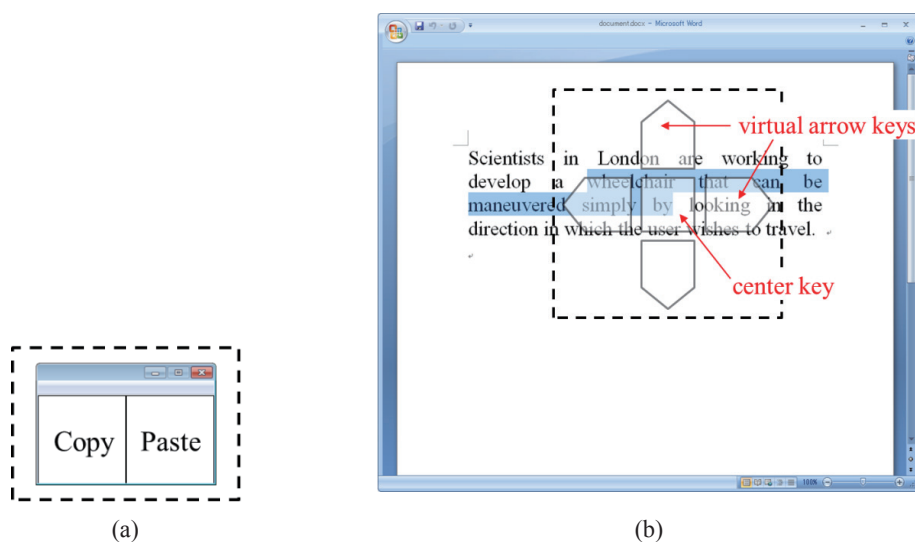


Fig. 2. (Color online) Proposed copy and paste interface.

text range and can copy the specified text to the clipboard. In the paste mode, the user can move the cursor to a desired position and execute “paste” at that position.

The operations of copy and paste with the proposed interface can be divided into the following steps.

Copy operation

- (1) Setting the starting position of a copy range
 - (1.1) Gaze pointing to specify the starting position of a copy range
 - (1.2) Adjusting the starting position
- (2) Marking the copy range
 - (2.1) Gaze pointing to specify the end of the copy range to mark
 - (2.2) Adjusting the marked range
- (3) Copying the marked text to the clipboard

Paste operation

- (4) Setting a pasting position
 - (4.1) Gaze pointing to specify the pasting position
 - (4.2) Adjusting the pasting position
- (5) Pasting the copied text into the pasting position

The proposed interface adopts a two-step selection method which is composed of gaze pointing and adjustment steps.

In the gaze pointing steps [i.e., steps (1.1), (2.1), and (4.1)], the text cursor is set to the position of the character that is nearest to the gazed point. Unlike a mouse that can be released to output a constant coordinate, eye movement is never perfectly stationary and exhibits jitter.^(24,25) In the proposed interface, when gaze remains inside the area of 30×30 pixels area for 1 s, the center of the gaze position is regarded as the coordinate of gaze pointing.

In the adjustment steps [i.e., steps (1.2), (2.2), and (4.2)], the text cursor can be adjusted via the on-screen virtual keys shown in Fig. 2(b). Since these keys are translucently drawn, the user can see the characters under these keys. The virtual arrow keys can be selected by eye fixation like an eye-typing QWERTY keyboard. By gazing at an arrow key, the user can move the cursor in the desired direction.

After adjusting the cursor position, the user can execute copy or paste by eye typing the center key surrounded by the 4 virtual arrow keys [Fig. 2(b)].

4. Implementation

In this section, the details of the implementation of the proposed interface are described. Eye tracking in the proposed interface is provided by Tobii TX 300, an optical eye tracker that senses the infrared light reflected from the eye by a video camera. Tobii TX 300 is a 23 inch LCD monitor of 1920×1080 pixels integrated with eye tracking optics capable of binocular tracking at 0.4° accuracy sampling at 300 Hz. The sampling rate of this device is higher than other devices. Owing to this high sampling rate, the device can more accurately track gaze points than other devices, even if a large head movement occurs.

The specifications of the PC connected to the eye tracker are Intel Core i7-2600 (3.4 GHz) CPU and 4.0 GB memory, and the PC runs the Windows 7 operating system. The gaze coordinates are delivered in real time to the PC via a TCP/IP connection.

The proposed interface is designed as a general input interface, which is suitable for use in any application programs. The user can do copy and paste between different application windows. The copy and paste operations by the proposed interface are realized by emulating mouse and keyboard operations.

Table 1 shows the gaze actions in each step and their corresponding mouse and keyboard event messages sent to the system (operating system). For example, in step (1.1), when the action of gaze pointing is detected, a “mouse click” message is generated. By sending the mouse click message to the system, clicking a left mouse button can be emulated. As a result, the text cursor appears at the “clicked” (gazed) point.

In step (1.2) (the step for adjusting the starting position of the copy range), when the user gazes at a virtual arrow key [Fig. 2(b)], “key press” message is sent to the system with the key code of the arrow key. As a result, the text cursor moves toward the desired direction as if the user had pressed the actual arrow key of a physical keyboard.

In step (2.1) (the step for gaze pointing for specifying the end of a copy range and marking the range), the shift key code is sent to the system in addition to the mouse click message (i.e., the operation of “shift + mouse click” is emulated). As a result, the text range from the starting position specified in steps (1.1) and (1.2) to the current position is marked. Similarly, in step (2.2), the key press message and key code “shift + arrow key” are sent to the system to emulate the adjustment of the marked range.

In step (3) (the step of copy execution), the key press message and key code “ctrl + c” are sent to the system. As a result, the marked text can be copied to the clipboard. The paste operation is realized by the same mechanism, as shown in Table 1.

Table 1
Event messages sent to the system when gaze actions in each step are detected.
(Copy operation)

Step	Gaze action	Event messages system
(1.1)	Gaze pointing for initially setting the starting point of a copy image	“mouse click”
(1.2)	Adjusting the starting position	“key press (LEFT/RIGHT/DOWN)”
(2.1)	Gaze pointing to specify the end of a copy range and marking the range	“key press (SHIFT)” + “mouse click”
(2.2)	Adjusting the marked range	“key press (SHIFT+LEFT/RIGHT/DOWN)”
(3)	Execution of copy	“key press (CTRL+C)”

(Paste operation)

Step	Gaze action	Event messages sent to the system
(4.1)	Gaze pointing for initially setting the pasting position	“mouse click”
(4.2)	Adjusting the pasting position	“key press (LEFT/RIGHT/UP/DOWN)”
(5)	Execution of paste	“key press (CTRL+V)”

5. Experimental Results

5.1 Method

In this section, the experiments we conducted to confirm the effectiveness of our interface are described. Eight able-bodied subjects (P1–P8) participated in the experiments for evaluating the interface. They were novices in gaze writing. Four tasks (Tasks A, B, C, and D) were given to each subject. In Tasks A and B, the subject used the proposed two-step selection interface. In Tasks C and D, the subject used gaze pointing only, i.e., in these tasks, the interface for adjustment [Fig. 2(b)] was not available. The font sizes of the displayed text were set to 20 points for Tasks A and C and 10.5 points for Tasks B and D. Each task is composed of 5 sessions. In each session, four copy targets were given as shown in Fig. 3. In a task, these copied targets were pasted into another window in turn. The time limit for each session was set to 2 min.

As a performance evaluation, we measured the elapsed time from the first gaze action to the completion of a task. In addition, we measured the number of misoperations in a task and the achievement rate of a task. As a workload evaluation, we measured the total workload score with NASA-TLX.

5.2 Results of performance evaluation

Figure 4 shows the average number of misoperations in each task. Here, the number of misoperations is defined as the number of times the copy/paste operations were retried for one copy target. In Fig. 4, the vertical axis represents the average number of misoperations in the five sessions in each task. From the figure, we observe that the numbers of misoperations in Tasks A and B tended to be lower than those of Tasks C and D. This result indicates the usefulness of the adjustment interface. In addition, we observe that there was little difference in the number of misoperations between Tasks A and B. This result indicates that the adjustment interface was effective regardless of font size.

Table 2(a) shows the achievement rates for various tasks. The achievement rate R of a task is defined as $R = n/N$, where n is the number of copy targets the user correctly copied and pasted

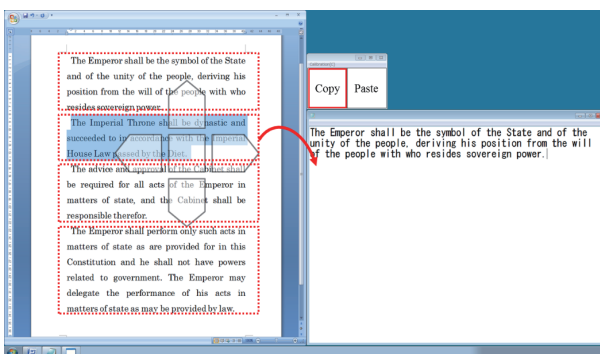


Fig. 3. (Color online) Copy targets in a task.

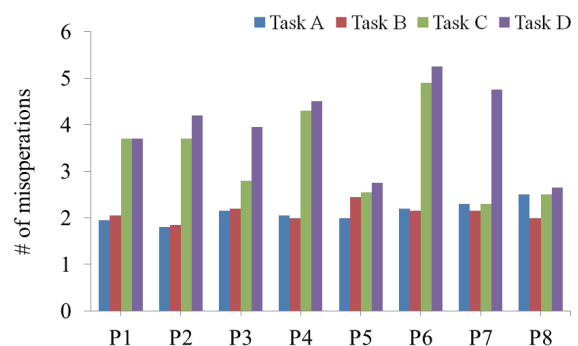


Fig. 4. (Color online) Average number of misoperations in each task.

Table 2

Achievement and improvement rates for various tasks.

(a) Achievement rates.

	P1	P2	P3	P4	P5	P6	P7	P8
R_A	0.95	0.65	0.95	0.95	0.95	1.00	1.00	1.00
R_B	0.80	0.70	1.00	0.85	0.95	0.80	0.95	1.00
R_C	0.95	0.50	0.45	0.50	1.00	0.85	0.95	0.95
R_D	0.65	0.30	0.55	0.15	1.00	0.85	0.75	1.00

(b) Improvement rates.

	P1	P2	P3	P4	P5	P6	P7	P8
R_A/R_C	0.95	0.65	0.95	0.95	0.95	1.00	1.00	1.00
R_B/R_D	0.80	0.70	1.00	0.85	0.95	0.80	0.95	1.00

within the time limit for the task, and N is the total number of copy targets given in the task. Comparing these rates between Tasks A and C, we observe that Task A shows a higher achievement rate than Task C. A similar trend can be observed between Tasks B and D. As a quantitative evaluation, we calculated the improvement rates R_A/R_C and R_B/R_D , where R_A , R_B , R_C , and R_D , respectively, represent the achievement rates for Task A, B, C, and D. From Table 2(b), we find that the improvement rates R_A/R_C and R_B/R_D for subjects P1, P2, P3, P4, P6, P7, and P8 were higher than 1.0. This result indicates that the use of the adjustment interface is indispensable for precise operations for many users and effective regardless of font size.

Although many subjects needed the support of the adjustment interface to complete a task, a few subjects (*e.g.*, P8) were able to complete tasks quickly without the adjustment interface. Figure 5 shows the elapsed times for completing tasks for subject P8. As shown in the figure, the subject completed Tasks C and D more quickly than Tasks A and B, which indicates that subject P8 could accurately set a cursor to a desired position with gaze pointing only. The accuracy of gaze pointing, however, strongly depended on the user. For practical use, the adjustment interface is indispensable. For users who can do accurate gaze pointing, it would be beneficial to provide an option for switching a one-step selection interface (*i.e.*, gaze-pointing-only interface) and a two-step selection interface (*i.e.*, gaze pointing + adjustment interface).

5.3 Results of workload evaluation

Figure 6 shows the total workload scores obtained using the NASA-TLX. The workloads for Tasks A and B for subjects P1, P3, P4, P6, and P8 are lower than those for Tasks C and D. In Tasks A and B, the subjects used the two-step selection interface (gaze pointing + adjustment). In contrast, in Tasks C and D, the subjects used gaze pointing only. These results indicate that the adjustment interface is effective for reducing the user's workload.

Comparing the workloads between Tasks A and B, we can confirm that there was little notable difference between them in many subjects. This result shows that the adjustment interface is effective regardless of font size.

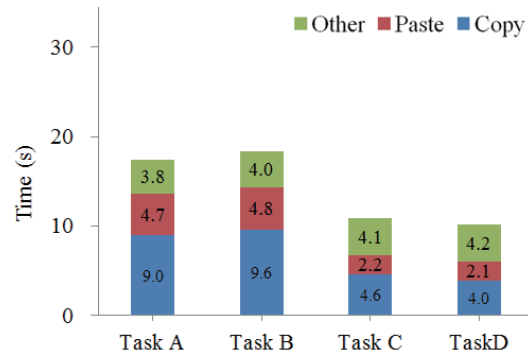


Fig. 5. (Color online) Average time elapsed per operation (P8).

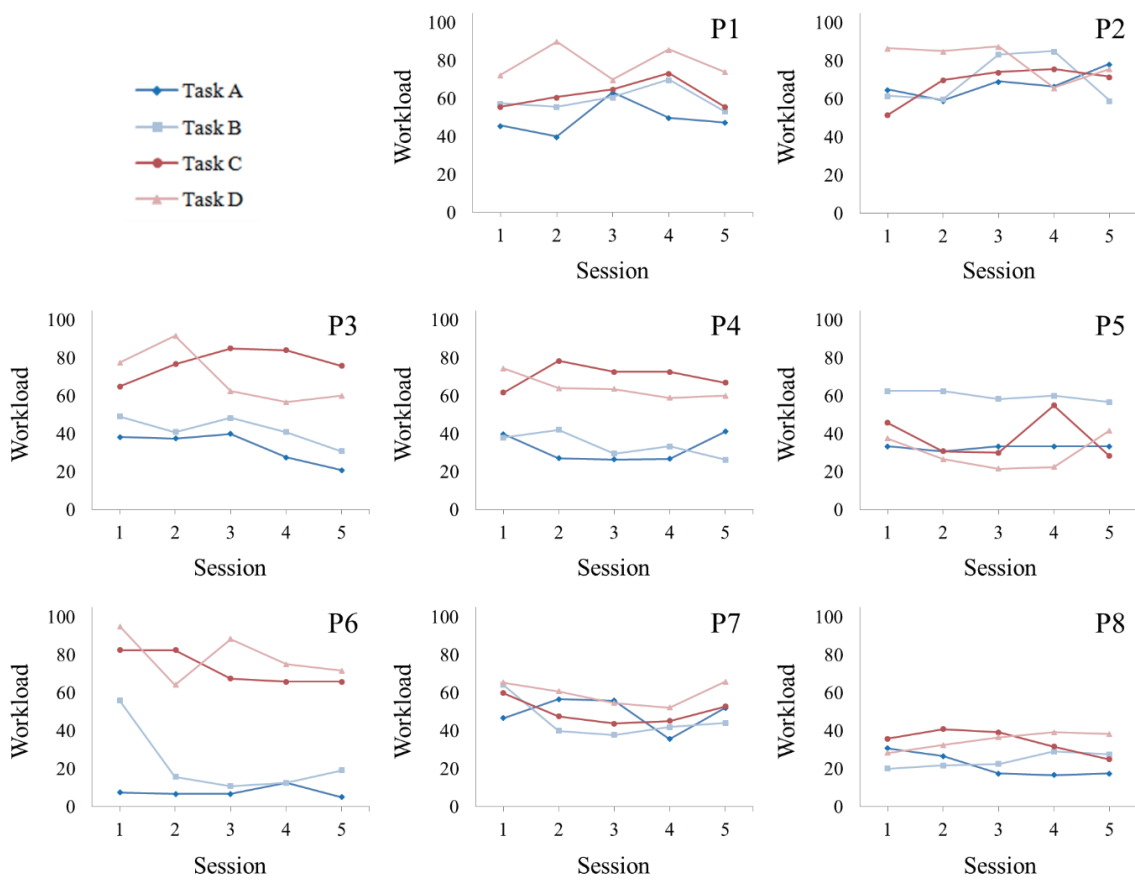


Fig. 6. (Color online) Total workload scores (NASA-TLX).

6. Conclusions

In this paper, we propose a gaze interface for copy and paste to support effective gaze writing. To compensate for the lack of gaze pointing accuracy, we introduced a two-step selection method. From the experimental results of performance and workload evaluations, we confirmed that the two-step selection method is effective for reducing the number of misoperations and the workload.

The two-step selection interface would be effective in various situations other than text selection (*e.g.*, selection of a file from a list). As a future work, we would like to develop gaze interfaces for supporting daily operations required when using a computer.

The proposed interface contributes to the future development of practical gaze-based communication systems that help not only motor-disabled people but also ordinary people to use a computer more conveniently and communicate with others more easily.

References

- 1 P. Majaranta and K. J. Rähkä: Proc. Eye Tracking Res. Appl. (ACM, New York, 2002) pp. 15–22.
- 2 R. J. K. Jacob: Adv. Hum.-Comput. Interact. **4** (1993) 151.
- 3 A. Huckauf and M. H. Urbina: Proc. Eye Tracking Res. Appl. (ACM, New York, 2008) pp. 51–54.
- 4 T. Outi, M. Päivi, I. Poika, and K. Rähkä: Proc. Eye Tracking Res. Appl. (ACM, New York, 2008) pp. 19–26.
- 5 R. Kishi and T. Hayashi: Int. J. Netw. Distrib. Comput. **3** (2015) 125.
- 6 M. Kumar, A. Paepcke, and T. Winograd: EyePoint: Proc. SIGCHI Conf. Hum. Factors Comput. Syst. (ACM, New York, 2007) pp. 421–430.
- 7 M. Yamato, A. Monden, K. Matsumoto, K. Inoue, and K. Torii: Proc. Int. Working Conf. Adv. Visual Interfaces (ACM, New York, 2000) pp. 270–273.
- 8 D. Miniotos, O. Špakov, I. Tugoy, and I. S. MacKenzie: Proc. Eye Tracking Res. Appl. Symp. (ACM, New York, 2006) pp. 67–72.
- 9 M. Ashmore, A. T. Duchowski, and G. Shoemaker: Proc. Graphics Interface (CHCCS, Waterloo, 2005) pp. 203–210.
- 10 M. McGuffin and R. Balakrishnan: Proc. Int. Conf. Hum. Factors Comput. Syst. (ACM, New York, 2002) pp. 57–64.
- 11 D. Miniotos, O. Špakov, and I. S. MacKenzie: Proc. Int. Conf. Hum. Factors Comput. Syst. (ACM, New York, 2004) pp. 1255–1258.
- 12 D. Fono and R. Vertegaal: EyeWindows: Proc. Int. Conf. Hum. Factors Comput. Syst. (ACM, New York, 2005) pp. 151–160.
- 13 C. Chen, S. T. Perraut, S. Zhao, and W. T. Ooi: Proc. Int. Working Conf. Adv. Visual Interfaces (ACM, New York, 2014) pp. 185–192.
- 14 K. Koaral and T. Komuro: Proc. Int. Conf. Hum. Factors Comput. Syst. (ACM, New York, 2013) pp. 1299–1304.
- 15 G. B. Reid and T. E. Nygren: Human Mental Workload (Elsevier Science Publishers, Amsterdam, 1988) p. 185.
- 16 S. G. Hart and L. E. Staveland: Human Mental Workload (Elsevier Science Publishers, Amsterdam, 1988) p. 139.
- 17 T. Hayashi and R. Kishi: Sens. Mater. **26** (2014) 529.
- 18 K. Rähkä and S. Ovaska: Proc. SIGCHI Conf. Hum. Factors Comput. Syst. (ACM, New York, 2012) pp. 3001–3010.
- 19 M. Tungare and M. A. Pérez-Quinones: Extended Abstr. Hum. Factors Comput. Syst. (ACM, New York, 2009) pp. 3431–3436.
- 20 B. Cinaz, B. Arnrich, R. L. Marca, and G. Tröster: Pers. Ubiquitous Comput. **17** (2013) 229.
- 21 H. A. Colle and G. B. Reid: Hum. Factors **40** (1998) 591.
- 22 G. B. Reid, S. S. Potter, and J. R. Bressler: Armstrong Aerospace Medical Research Laboratory Technical Report (AAMRL-TL-89-023, 1989).
- 23 J. C. Byers, A. C. Bittner, and Jr., S. G. Hill: Advances in Industrial Ergonomics and Safety I (Taylor & Francis, 1989) p. 481.
- 24 S. Miyake and M. Kumashiro: Jpn. J. Ergonomics **48** (2012) 295.
- 25 R. J. K. Jacob: Proc. Int. Conf. Hum. Factors Comput. Syst. (ACM, New York, 1990) pp. 11–18.