

Workload Evaluation of Gaze-Writing Systems

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

¹Department of Information Engineering, Faculty of Engineering, Niigata University,
8050 Ikarashi-2-no-cho, Nishi-ku, Niigata-shi, Niigata 950-2181, Japan

(Received January 14, 2014; accepted April 23, 2014)

Key words: gaze writing, workload assessment, NASA-TLX, eye typing, eye gesturing, continuous writing

In this paper, we present the workload evaluation of three types of eye-based text entry methods: (1) eye typing, (2) eye gesturing, and (3) continuous writing. As metrics for workload evaluation, we used the NASA task load index (NASA-TLX), which was developed by NASA for assessing the workload of users working with human-machine systems. Experimental results have shown that with eye typing, the user can enter text fast with a low workload, and that eye gesturing and continuous writing need time to bring users to a certain level, and a higher workload is needed than eye typing for maintaining a high text entry speed.

1. Introduction

Gaze writing (text entry by eye gaze) is an indispensable way of communication for motor-disabled people who cannot either talk or use a keyboard or a mouse. Owing to its importance, gaze communication has been actively researched for over 30 years.

Since gaze writing could provide motor-disabled people a means of communication, this technology should be considered as an assistive technology, which facilitates life activities. Currently, people have recognized that the evaluation of workload is a key point in the research and development of human-machine interfaces, in search of higher levels of comfort, satisfaction, efficiency, and safety in working with a human-machine system. However, research studies aimed at evaluating workload in gaze writing are still scarce, whereas there are many research studies focused on the evaluation of text entry speed and error rate of mistyping. Owing to the absence of evaluation from users' point of view, addressing the problem of workload is necessary for the development of more practical gaze writing systems.

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

Workload is a subjective degree of how tiring it is to write by eye gaze. This makes workload assessment difficult. This would be the main reason why the workload has not been explored in detail in previous studies.

The goal of our study is to systematically evaluate the workloads of different types of gaze-writing methods. As the first step to the goal, we applied the NASA task load index (NASA-TLX),⁽¹⁾ which is a questionnaire-based workload assessment method, to the workload assessment of gaze-writing methods. In this paper, we show the experimental results of comparing three types of gaze-writing methods: (1) eye typing, (2) eye gesturing, and (3) continuous writing.

The rest of the paper is organized as follows. In § 2, we discuss related works, where we will explain the three gaze-writing methods: eye typing, eye gesturing, and continuous writing. In § 3, we describe the NASA-TLX, where we explain how the NASA-TLX is used to evaluate the workload of a user. In § 4, we show the obtained experimental results. In § 5, we provide conclusions.

2. Gaze-Writing Methods

2.1 Classifications

Many gaze-writing methods have been developed.^(2–13) Bee and Andre⁽²⁾ categorized these methods into three types: (1) eye typing,^(3–5) (2) eye gesturing,^(6–10) and (3) continuous writing.^(11–13) Following the classifications, we chose three gaze-writing systems for workload evaluation. In the next three subsections, we describe the characteristics of these methods.

2.2 Eye typing

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. 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 (dwell time), the system assumes the user intended to write that letter. The dwell time typically ranges from 400 to 1000 ms.⁽¹⁴⁾ The shorter the dwell time, the faster the text entry. However, a very short dwell time will increase the amount of unintended selections. It is reported that the text entry rate reaches 20 words per minute (WPM) by a well-trained user under a well-adjusted dwell time. Generally, the average text entry speed is 6–7 WPM.

2.3 Eye gesturing

Eye gesturing was inspired by pen gesturing like graffiti⁽¹⁵⁾ used in palm devices. Eye gesturing uses eyes as an input device instead of a pen device. It associates an eye-movement pattern with a particular action. Slow text entry is the main problem of eye gesturing. The reason for the slow text entry is that multiple eye movements are needed to perform a desired action. Every saccade (fast movement of eye to point out a particular region of a screen), which lasts between 30–120 ms, is followed by fixation, which lasts 200 ms on average before a new saccade can be started. Generally, it takes

at least 1 s to complete the gesture of letter selection. In addition, for effective input, the user has to memorize eye-movement patterns. The workload of memorizing eye-movement patterns is a potential problem of eye gesturing.

pEYEWrite^(7,10) is an effective text entry method based on eye gesturing. pEYEWrite adopts two hierarchical pie menus⁽¹⁶⁾ (also known as radial menus). A pie menu is made of several “pie slices” as shown in Fig. 1. On the first level, the pie menu consists of six slices that contain groups of five letters [Fig. 1(a)]. On the second level, each pie slice corresponds to a letter [Fig. 1(b)]. To enter a letter, the user moves his/her gaze toward the outer border of a slice that contains the desired letter. The pie menu on the second level opens immediately. The target letter is selected by glancing again through its respective selection border. Figure 2 shows an example of eye gesture with pEYEMenu. This figure illustrates the eye-movement patterns to input the word “hello”. pEYEWrite does not need a dwell time for text input because the selection is immediately performed when the gaze crosses the selection border. From this characteristic, the performance characteristics of pEYEMenu are superior to those of other eye-gesturing-based text entry systems. Urbina and Huckauf⁽¹⁰⁾ reported that the mean text entry speed of pEYEWrite was 7.85 WPM, whereas an expert achieved 12.33 WPM.

2.4 Continuous writing

With continuous writing, moving letters is followed by the user’s gaze with pursuit movement. This movement is the trigger of the letter selection. Dasher⁽¹²⁾ is a typical text entry system based on continuous writing. It is also known as part of the

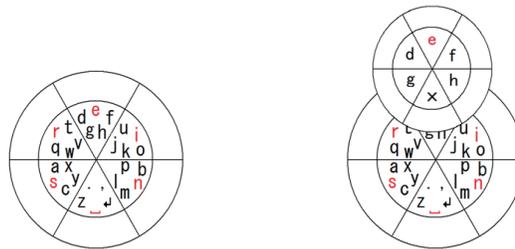


Fig. 1. (Color online) Pie menu interface of pEYEWrite.

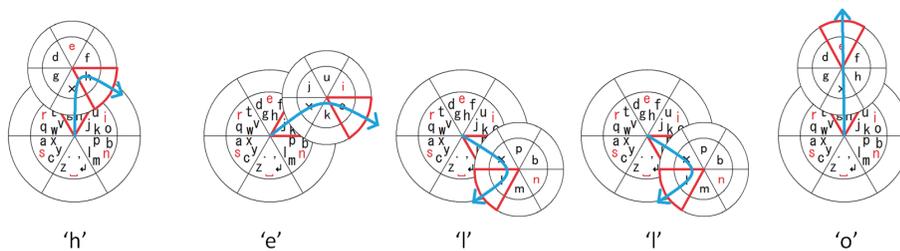


Fig. 2. (Color online) Eye gesturing for inputting word “hello”.

GNOME desktop software in UNIX systems. Figure 3 shows the interface of Dasher. In the interface, characters are vertically displayed on the right side of the window in alphabetical order [Fig. 3(a)]. When the user gazes at a desired letter, the letter zooms in and moves towards the center of the window [Fig. 3(b)]. When the letter crosses the center line of the window, the letter is entered [Fig. 3(c)]. Dasher was originally developed as a text entry method with a pointing device such as a mouse, a touch screen or a joystick. It has been extended for gaze inputting. Generally, text entry with continuous writing is faster than traditional eye-typing methods.⁽¹²⁾

3. Workload Evaluation with NASA-TLX

In this section, we describe the NASA-TLX,⁽¹⁾ which is used in this paper to evaluate the workloads of the above-mentioned three gaze-writing methods. The NASA-TLX was developed by the NASA Ames Research Center for assessing the subjective workload of a user working with human-machine systems. In the NASA-TLX, the total workload is composed of six subscales: mental demand (MD), physical demand (PD), temporal demand (TD), own performance (OP), effort (EF), and frustration level (FL). The meaning of each subscale factor is explained in Table 1. These six subscales are rated by the user within a 100-point range. The descriptions for each measurement in Fig. 4 can help participants answer accurately.

The total workload is defined by the weighted average of the six subscale scores. To determine the weight of each subscale, the participant chooses the subscale that is more relevant to the workload. In total, 15 pairwise comparisons are needed to determine the weights. The number of times each subscale is chosen is set as the weight of the subscale. This is multiplied by the subscale score for each dimension and then divided by 15 to obtain a total workload score from 0 to 100.

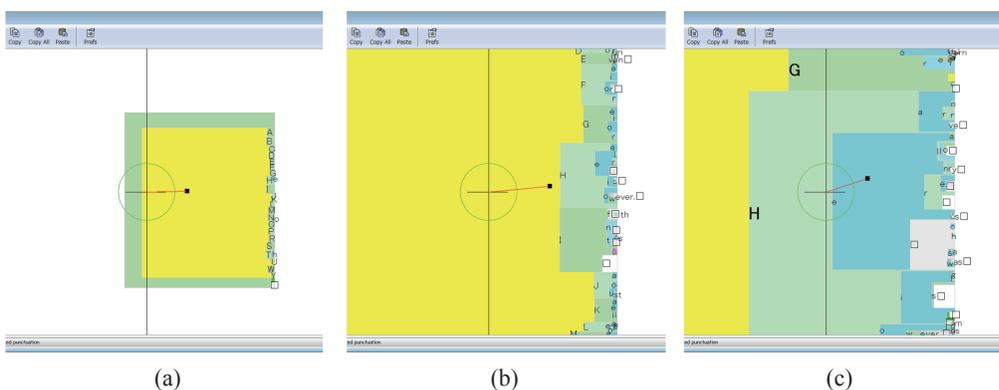


Fig. 3. (Color online) Interface of Dasher.

Table 1
Six subscale factors of NASA-TLX.

Factors	Descriptions
Mental demand	How much mental and perceptual activity was required (<i>e.g.</i> , thinking, deciding, calculating, remembering, looking, and searching)? Was the task easy or demanding, simple or complex, or exacting or forgiving?
Physical demand	How much physical activity was required (<i>e.g.</i> , pushing, pulling, turning, controlling, and activation)? Was the task easy or demanding, slow or brisk, slack or strenuous, or restful or laborious?
Temporal demand	How much time pressure did you feel due to the rate or pace at which the tasks or task elements occurred? Was the pace slow and leisurely or rapid and frantic?
Performance	How successful do you think you were in accomplishing the goals of the task set by the experimenter (or yourself)? How satisfied were you with your performance in accomplishing these goals?
Effort	How hard did you have to work (mentally and physically) to accomplish your level of performance?
Frustration level	How insecure, discouraged, irritated, stressed and annoyed versus secure, gratified, content, relaxed, and complacent did you feel during the task?

4. Experimental Results

4.1 Method

We experimentally compared the performance characteristics and workloads of the three gaze-writing methods, *i.e.*, eye typing, eye gesturing, and continuous writing. As representative systems based on the three text entry methods, we used a QWERTY onscreen keyboard for eye type, pEYWrite, and Dasher, respectively. We implemented these systems with a Tobii TX300 eye-tracking device.

Eighteen able-bodied university students participated in the experiment. Six participants (P1, P2, P3, P4, P5, and P6) used the eye-typing method (QWERTY onscreen keyboard), the next six participants (P7, P8, P9, P10, P11, and P12) used the eye-gesturing method (pEYWrite), and the last six participants (P13, P14, P15, P16, P17, and P18) used the continuous-writing method (Dasher). All participants are novices in gaze writing.

The experiment consists of ten sessions. Each session consists of two phases: 10 min test phase and 10 min break. The calibration of the eye tracker is performed before the first session and in each break time. The first six participants (P1, P2, P3, P4, P5, and P6), who use the eye-based QWERTY onscreen keyboard, adjust their dwell time before the first session starts. In each session, the participants enter English phrases, which are randomly selected from a predefined phrase set.⁽¹⁷⁾ The phrase set contains 500 English phrases. Each phrase is composed of easy words such as “video camera with a zoom lens”. During the 10 min test phase, the participants enter as many English phrases as possible. If the participant notices mistyping, the error must be corrected by deleting and entering the correct letter.

Based on data of the 18 participants, we measured the text entry speed, error rate of mistyping, and workload. The text entry speed is measured in words per minute (WPM). Here, one word means five characters. The phrase “words per minute” (17 characters) is regarded as 3.4 words. The error rate is measured in keystrokes per character (KSPC).⁽¹⁸⁾ If the text is entered without errors (mistyping), KSPC will be 1.00. A KSPC greater than 1 suggests that the user entered an incorrect character and it was deleted. The workload is measured with the NASA-TLX. Each participant answers the questionnaire of the NASA-TLX after each session.

4.2 Results

Figure 4 shows the text entry speeds of the 18 participants. Participants who used the eye-typing method were able to enter text the fastest. The text entry speed achieved was 5–9 WPM even in the 1st session. However, little improvement of the text entry speed was observed during the ten sessions. The text entry speeds of the other two methods (eye gesturing and continuous writing) are relatively low and range from 2 to 4 WPM in the 1st session. However, the speeds gradually increase when sessions are repeated. Until the 10th session, the speeds reach 4–8 WPM, which are comparable to that of eye typing.

Figure 5 shows the error rate of mistyping. The continuous-writing method has a high KSPC. However, with the continuous writing method, the user can easily delete it by only moving the eye toward the right side of the screen. That is, the action of deleting a character can be finished immediately. This is the reason why the text entry speed of the continuous-writing method is as high as the eye-typing method in the end.

Figure 6 shows the total workloads of the three methods. The total workload of the eye-typing method gradually decreases during the ten sessions. In contrast, little improvement of the total workloads of the eye-gesturing and continuous-writing methods was observed. These results indicate that with the eye-typing method, the user can enter text fast with a low workload, and that the eye-gesturing and continuous-writing methods need much training time to bring users to a certain level and need a high workload to maintain a fast text entry.

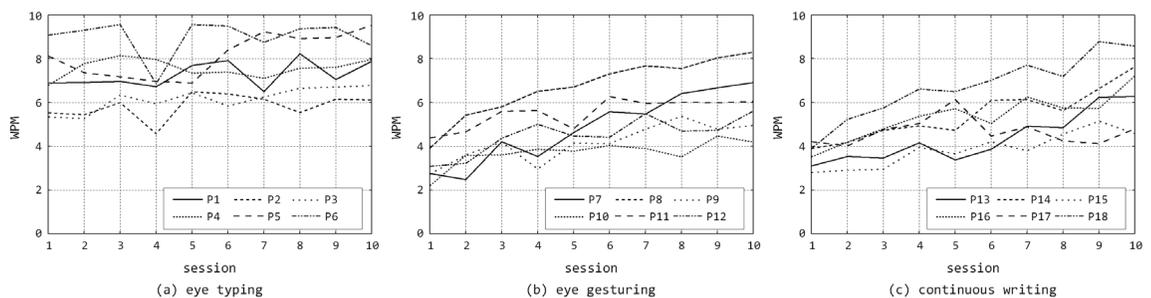


Fig. 4. Text entry speeds.

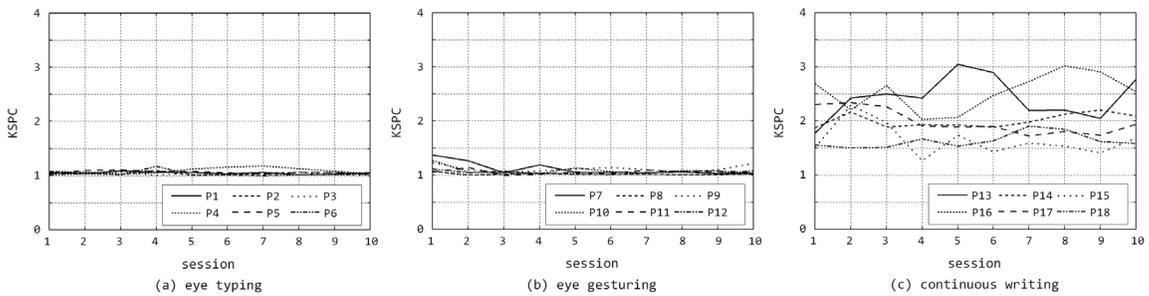


Fig. 5. Error rates of mistyping (KSPC).

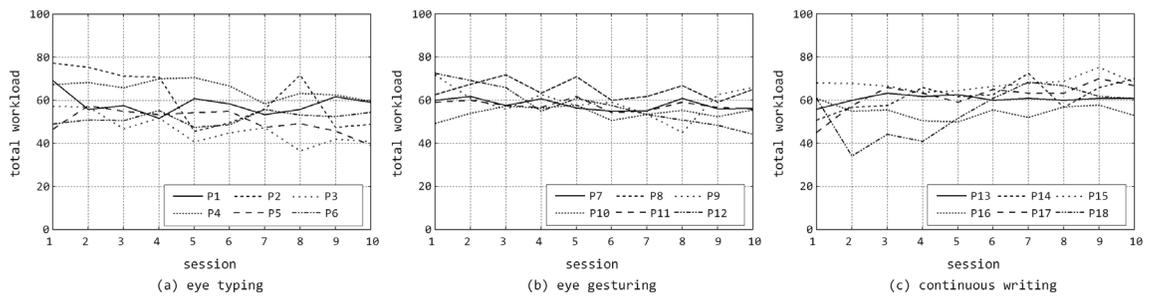


Fig. 6. NASA-TLX total workloads.

Among the six subscale factors in the NASA-TLX, two subscale factors, namely, physical demand (PD) and effort (EF), were rated high by all the participants. The other two subscale factors, namely, own performance (OP) and frustration level (FL), were given relatively low weights as a result of the pairwise comparison between subscale factors. As shown in Fig. 7, mental demand was highly rated by the participants who used the continuous-writing method. As shown in Fig. 8, temporal demand (TD) was highly rated by the participants who used the eye-gesturing method and the participants who used the continuous-writing method.

We interviewed the participants about the main reasons for the experienced workload. Participants P1–P6, who used the eye-typing method, stressed that the intentional eye fixation for letter selection demanded concentration and caused eye fatigue. Participants P7–P12, who used the eye-gesturing method, were sometimes confused which pie-menu slice contains the letter they want to enter. This indicates that memorizing an eye-movement pattern is a workload factor for beginner users of the eye-gesturing method. Participants P13–P18, who used the continuous-writing method, experienced a high temporal demand of catching a letter from the stream of letters on the screen. This phenomenon would lead to high mental and temporal demands in the continuous writing method.

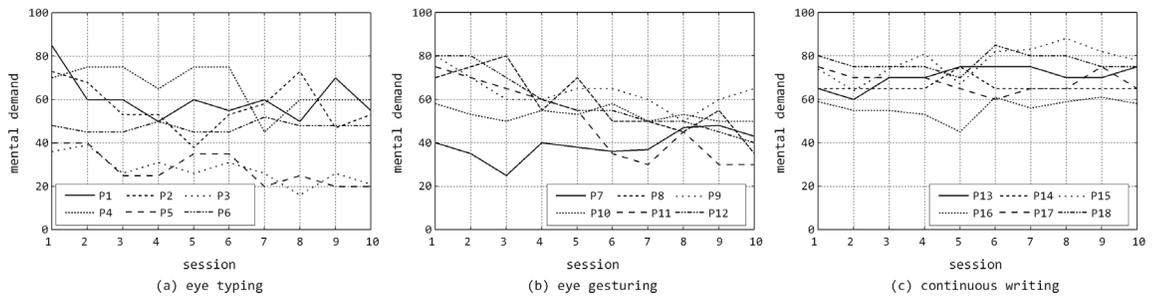


Fig. 7. Mental demand.

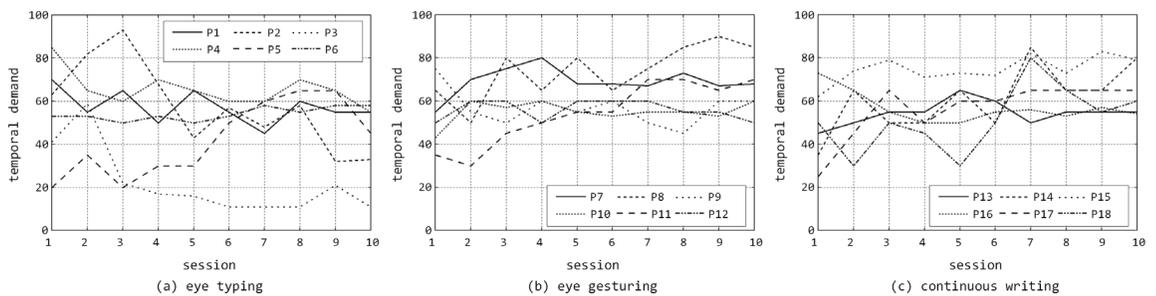


Fig. 8. Temporal demand.

The experimental results can be briefly summarized as follows:

- (1) The text entry speed of eye typing cannot be improved markedly even when eye typing is performed repeatedly. However, the workload gradually decreases in many cases.
- (2) The text entry speeds of eye gesturing and continuous writing increase when the user adopts the systems. However, maintaining a high text entry speed requires a high workload.

These results clarify that from the viewpoint of workload, the use of gaze-writing systems that enable users to quickly input texts is not always the best choice. The results bring a new perspective in the research and development of gaze-writing systems.

5. Conclusions

In this paper, we focused on the workload evaluation of three gaze-writing methods: eye typing, eye gesturing and continuous writing. We experimentally compared the workloads of the gaze-writing methods with those obtained using the NASA-TLX. Experimental results show that with the eye-typing method, beginner users can enter text fast with a low workload, and that the eye-gesturing and continuous writing methods need time to bring users to a certain level and need a higher workload to maintain a high text entry speed.

The NASA-TLX would be a promising approach to the workload assessment of gaze-writing methods. The findings of this study are still limited. A longitudinal study is needed to further analyze the workloads of gaze-writing methods.

Acknowledgements

This work was supported by research grants from NS Promotion Foundation for Science of Perception and Uchida Energy Science Promotion Foundation.

References

- 1 S. G. Hart and L. E. Staveland: Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research, eds. P. A. Hancock and N. Meshkati: Human Mental Workload (Elsevier Science Publishers B. V., Amsterdam, 1988), pp. 139–183.
- 2 N. Bee and E. Andre: Proc. Workshop Perception and Interactive Technologies for Speech-Based Systems, 2008, pp. 111–122.
- 3 D. Miniotas, O. Spakov and G. Evreinov: Symbol Creator: Proc. Int. Conf. Human-Computer Interaction, 2003, pp. 137–143.
- 4 P. Majaranta and K. J. Riih a: Proc. Symp. Eye Tracking Research and Applications, 2008, pp. 15–22.
- 5 J. P. Hansen, and K. T rning, A. S. Johansen, S. Anders, K. Itoh and H Aoki: Proc. Symp Eye Tracking Research and Applications, 2004, pp. 131–138.
- 6 M. Porta and M. Turina: Proc. Symp. Eye Tracking Research and Applications, 2008, pp. 27–34.
- 7 M. H. Urbina and A. Huckauf: Proc. Conf. Communication by Gaze Interaction, 2007, pp. 65–70.
- 8 J. O. Wobbrock, J. Rubinstein, M. W. Sawyer and A. T. Duchowski: Proc. Symp. Eye Tracking Research and Applications, 2008, pp. 11–18.
- 9 A. Huckauf and M. H. Urbina: Proc. Symp. Eye Tracking Research and Applications, 2008, pp. 51–54.
- 10 M. H. Urbina and A. Huckauf: Proc. Symp. Eye Tracking Research and Applications, 2010, pp. 315–322.
- 11 D. J. Ward and D. J. C. Mackay: Nature **418** (2002) 838.
- 12 T. Outi, M. P ivi, I. Poika and R. Kari-Jouko: Proc. Symp. Eye Tracking Research and Applications, 2008, pp. 19–26.
- 13 D. W. Hansen, H. H. T. Skovsgaard, J. P. Hansen and E. M lllenbach: Eye Tracking Research and Applications, 2008, pp. 205–212.
- 14 R. J. K. Jacob: Interaction Techniques: Toward Non-command Interfaces, Advances in Human Computer Interaction, Vol. 4, eds. H. R. Hartson and D. Hix (Alex Publishing Co., Norwood, 1993), pp. 151–190.
- 15 A. Butter and D. Pogue: Piloting Palm: The Inside Story of Palm, Handspring, and the Birth of the Billion-Dollar Handheld Industry (Wiley, New York, 2002), pp. 62–66.
- 16 J. Callahan, D. Hopkins, M. Weiser and B. Shneiderman: Proc. Conf. Human Factors in Computer Systems, 1988, pp. 95–100.
- 17 I. S. MacKenzie and R. W. Soukoreff: CHI '03 Extended Abstracts on Human Factors in Computing Systems, 2003, pp. 754–755.
- 18 R. W. Soukoreff and I. S. MacKenzie: Proc. Conf. Human Factors in Computer Systems, 2003, pp. 113–120.

About the Authors

Takahiro Hayashi received his BE and ME degrees from Kanazawa University, Japan, in 1998 and 2000, respectively, and a PhD degree in engineering from the same university in 2003. Currently, he is an associate professor in the Department of Information Engineering, Niigata University, Japan. His research interests include intelligent systems, interactive systems, and multimedia databases. He is a member of IEEE, IEICE (Japan), and IPSJ (Japan).

Reo Kishi received his BE degree from Niigata University, Japan, in 2014. Currently, he is a master course student in the Graduate School of Science and Technology, Niigata University. His research interests include eye-tracking technologies and interactive systems.