Input Method Using Divergence Eye Movement

Shinya Kudo
The University of Electro-communications
1-5-1 Choufugaoka
Choufu, Tokyo, Japan
kudo@kaji-lab.jp

Hiroyuki Okabe
The University of Electro-communications
1-5-1 Choufugaoka
Choufu, Tokyo, Japan
h.okabe@kaji-lab.jp

Taku Hachisu
The University of Electro-communications
1-5-1 Choufugaoka
Choufu, Tokyo, Japan
JSPS Research Fellow
hachisu@kaji-lab.jp

Michi Sato
The University of Electro-communications
1-5-1 Choufugaoka
Choufu, Tokyo, Japan
JSPS Research Fellow
michi@kaji-lab.jp

Shogo Fukushima
The University of Electro-communications
1-5-1 Choufugaoka
Choufu, Tokyo, Japan
JSPS Research Fellow
shogo@kaji-lab.jp

Hiroyuki Kajimoto
The University of Electro-communications
1-5-1 Choufugaoka
Choufu, Tokyo, Japan
Japan Science and Technology Agency
kajimoto@kaji-lab.jp

CHI 2013 Extended Abstracts, April 27–May 2, 2013, Paris, France.
ACM 978-1-4503-1952-2/13/04.

Abstract
A gaze input interface offers hands-free operation by using the view-point position as the cursor coordinates on the display. However, the selection operation of a button is indistinguishable from viewing; this is known as the Midas touch problem. We propose a new input method that measures divergence eye movement, thereby enabling users to “press” a button by moving their viewpoint forward. Comparison of our method and the conventional blinking input method confirms that input speed and accuracy are similar.

Keywords
Divergence eye movement; gaze input interface; depth input

ACM Classification Keywords
H.5.2. [Information interfaces and presentation]: Usability Testing and Evaluation

General Terms
Human Factors

Introduction

Gaze input has been proposed as an intuitive input method for a personal computer (PC) [1]. By moving their eyes, users can move the cursor and select buttons. However, there is a well-known Midas touch problem [2], meaning that this method misinterprets straightforward gazing as a selection operation.

Several solutions have been proposed for dealing with this problem. Hansen et al. [3], Murata et al. [4], and Dario [5] reduced incorrect selection by using gaze time, while Shaw et al. achieved this by using blinking [6]. However, these methods have a trade-off between selection precision and selection speed. Zhai et al. [7] and Yamato et al. [8] avoided incorrect selection by using the user's gaze as cursor movement only and using a mouse for selection; however, use of a mouse is not applicable in some applications that originally required gaze input.

Our goal is to realize intuitive selection by using an action other than gazing or blinking. In this paper, we propose a new input method that measures divergence eye movement, thereby enables users to "press" a button by moving their viewpoint forward.

Method

Divergence eye movement is a type of eye movement that occurs when a person moves his/her viewpoint forward (Figure 1). Pfeiffer et al. tried to estimate 3-dimensional (3D) positions of viewpoints by detecting convergence [9]. Sato et al. used the same information to investigate cross-sections of 3D models [10]. We apply this technique as the selection operation by detecting the forward movement of the viewpoint. We call this method the "depth input method".

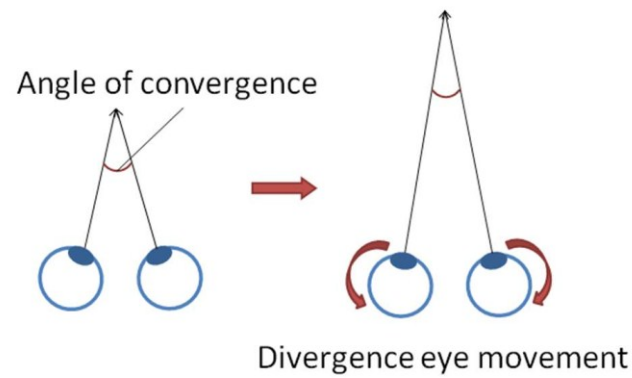

Figure 1. Angle of convergence and divergence eye movement

First, we obtain the viewpoints of the two eyes on the display plane using an eye-tracking device. When divergence eye movement occurs, the viewpoint of the right eye moves to the right and that of the left eye moves to the left. By setting the distance between the two viewpoints as *dx*, the interpupillary distance as *dp*, and the distance between the eyes and display as *d*, the depth of viewpoint is calculated as follows (Figure 2).

$$Depth = \frac{d}{dp - dx} dx \quad (1)$$

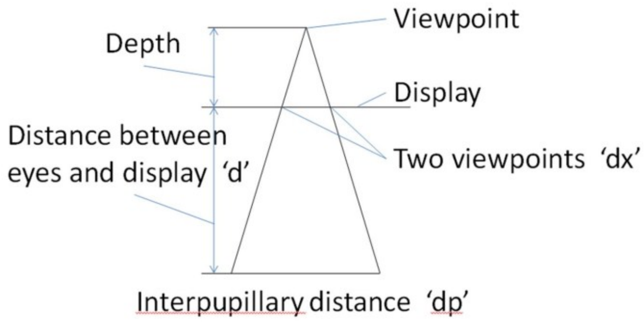

Figure 2. Estimation of the depth viewpoint position

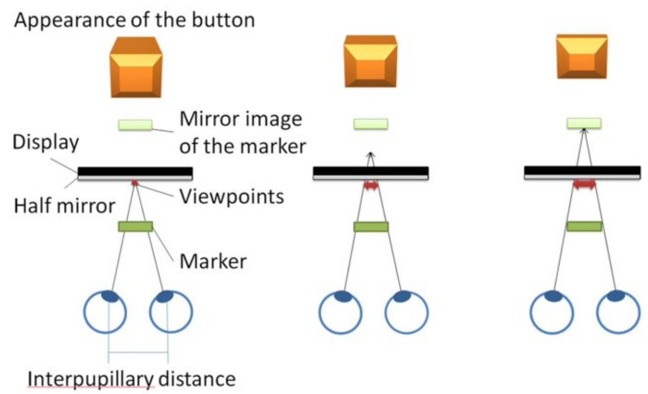

Figure 3. Presentation technique

A 3D-shaped button is presented on the display, moving vertically to the display according to the change in estimated depth of viewpoint (Figure 3).

A preliminary experiment revealed that intentionally staring at empty space is difficult without training. Therefore, we provided a visual marker on the rear of the display using a half mirror on the display and a marker in the front. Adjusting the angle of convergence becomes easier by watching a mirror image of the marker. This setup can be simplified by using a 3D monitor, but this is beyond the scope of this paper.

Experiment

We conducted experiments to verify the operability of the proposed method, compared with the conventional blinking input method.

Experimental setup

The experimental setup is shown in Figure 5. A chin support was set up on the desk 700mm in front of the display to fix the eye position. A half mirror was fixed to the display. A visual marker was placed 300mm in front of the display so that the mirror image of the marker from the half mirror was 300 mm behind the display. The distance between the two viewpoints (*dx*) was 20 mm when the participants looked at the image of the marker.

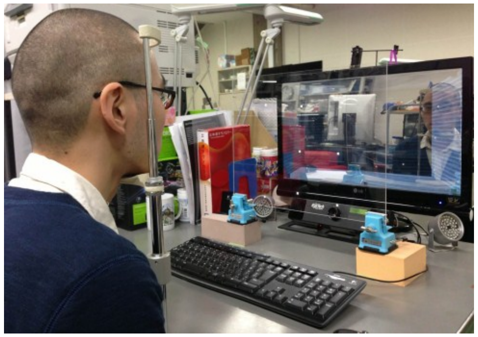

Figure 4. Overview of the experiment

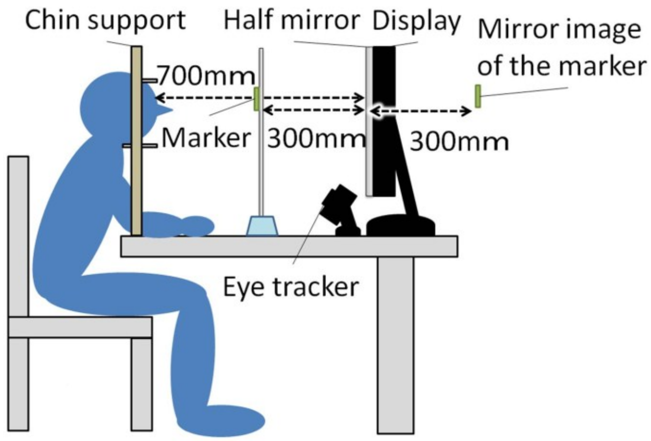

Figure 5. Experiment setup

Interpupillary distance (*dp*) was set to 65mm. We used a desktop type eye tracking device (TM3 EyeTech Digital Systems), which uses the pupil-corneal reflection method. Viewpoint resolution was 1.0 deg. We conducted a calibration of the viewpoints on the display plane for each participant before the experiment. The display resolution was 1920 × 1080 and viewing angle 35.7 deg ×23.3 deg. We used OpenGL for drawing. An image of the actual experiment is shown in Figure 4.

Experimental condition

We compared the depth input method and the conventional blinking input method. In the former method, a button is pressed as the participant moves his/her viewpoint forward to the image of the marker. When the viewpoint reaches the image, it is regarded as a selection operation. In the blinking type method, the button is pressed when the participant's eyes remain closed for 0.5 s. (This duration is typically used to discriminate between natural blinking and selection [6]).

Nine buttons numbered 1 to 9, 40 mm square (3.27 deg in terms of viewing angle) were arranged in a 3 × 3 format on the display (Figure 6). Five randomly chosen numbers were displayed above the central button. Participants were asked to sequentially press the buttons corresponding to the five displayed numbers ignoring incorrect input. They were also asked to complete the task as quickly and accurately as possible. Input of five numbers was regarded as one trial, and each participant conducted ten trials for each method.

Figure 6. Display during experiment

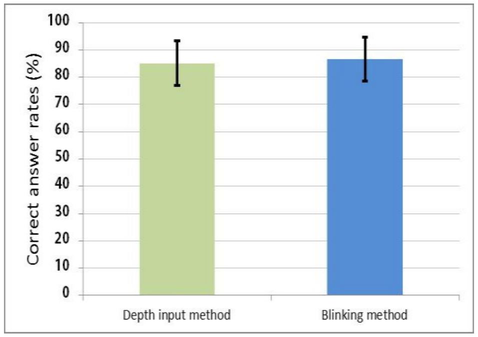

Figure 7. Correct answer rates (%).

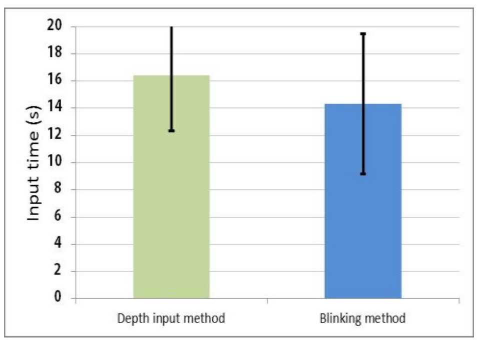

Figure 8. Input time (s).

Procedure
First, the participants practiced the depth input method for about 1-2 min to familiarize themselves with the forward movement of their viewpoints. After the experiment started, participants were instructed to enter the five numbers, while the input speed and accuracy were recorded. Participants were given a 5-second break before starting the next trial. After completion of the ten trials, participants were given a 5-min break. Thereafter, they carried out the experiment using the blinking input method in the same way as the depth input method. The participants comprised four males (one male with glasses, two males wearing soft contact lenses and the other is naked eye) and 4 females (one female with glasses, two females wearing soft contact lenses and the other is naked eye), respectively, aged between 21 to 28 years.

Results
The correct answer rates and input times for all participants are shown in Figure 7 and Figure **8**.

Correct answer rates for both methods were close to 90%, no significant difference was observed (*Chi-square test: X-squared = 0.26, df = 1, p-value = 0.61 > 0.05*). Input times for the blinking input method were mostly 2.1 s faster than those of the depth input method. However, there was no significant difference between the input times of the two methods (t-test: *t = 0.9093, df = 13.354, p-value = 0.3793 > 0.05*).

Discussion
The experiment revealed that the depth input method produces similar performance to the conventional blinking input method. As the blinking input method distinguishes selection and natural blinking by the duration of blinking, it prevents users from closing their eyes for a while, which is a natural behavior when the eyes become fatigued. Conversely, the depth input method can be distinguished from other eye movements, but still, retains some intuitiveness.

During the experiment, we observed a relatively large deviation among the participants. This was probably due to the difference in accuracy of the calibration as a result of wearing glasses or contact lenses. The depth input method requires more accurate calibration than the blinking input method because it uses the distance between the two viewpoints. Therefore, any difference in the calibration accuracy significantly affects the result. Additionally, the calibration was done on a 2D display plane only, which is not sufficient for 3D viewpoint movement. Pfeiffer et al. carried out a 3D calibration on a 3D display [9], which should be done in our system.

After the experiment, there was a comment that eye fatigue occurred during the depth input method rather than the blinking input method. This is due to the unnaturally large eye movements in the depth direction. There was also a comment that the button appeared in double vision during input, which was not comfortable. These problems should be solved by shortening the depth input length through more precise calibration, and by using a 3D display.

Conclusion
In this paper, we proposed a new input method that measures divergence eye movements, enabling user to "press" a button by moving his/her viewpoint forward. We compared this method with the conventional

blinking input method and confirmed that input speed and accuracy are comparable. Thus, the depth input method achieves similar performance to the conventional blinking input method. In the future, we will continue to improve the depth input method by using a 3D display, thereby making it easier to understand.

References

[1] Sibert, L.E. and Jacob, R.J.K. Evaluation of eye gaze interaction. In *Proc. CHI 2000*, ACM Press (2000), 281-288.

[2] Jacob, R.J.K. What you look at is what you get: eye movement-based interaction techniques. In *Proc. CHI 1990*, ACM Press (1990).

[3] Hansen, J.P., Anderson, A.W., and Roed, P. Eye-gaze control of multimedia systems. In *Symbiosis of human and artifact, Vol. 20A*. Anzai, Y., Ogawa, K., and Mori, H. (Eds.), Elsevier Science (1995), 37-42.

[4] Murata, A. and Miyake, T. Effectiveness of eye-gaze input system: identification of conditions that assures high pointing accuracy and movement directional effect. In *Fourth International Workshop on Computational Intelligence & Applications*, IEEE (2008), 127-132.

[5] Dario D. Inferring intent in eye-based interfaces: tracing eye movements with process model. In *Proc. CHI 1999*, ACM Press (1999), 15-20.

[6] Shaw, R., Crisman, E., Loomis, A., and Laszewski, Z. The eye wink control interface: using the computer to provide the severely disabled with increased flexibility and comfort. In *Third Annual IEEE Symposium on Computer-Based Medical Systems*, IEEE (1990), 105-111.

[7] Zhai, S., Morimoto, C., and Ihde, S. Manual and gaze input cascaded (MAGIC) pointing. In *Proc. CHI 1999*, ACM Press (1999), 246-253.

[8] Yamato, M., Monden, A., Matsumoto, K., Inoue, K., and Torii, K. Quick button selection with eye Gazing for general GUI environment. In *Proc*. *International Conference on Software: Theory and Practice* (2000).

[9] Pfeiffer, T., Latoschik, Marc, E., and Wachsmuth, I. Evaluation of binocular eye trackers and algorithms for 3D gaze interaction in virtual reality environments. *Journal of Virtual Reality and Broadcasting*, 5, 16 (2008).

[10] Sato, M. and Kajimoto, H. Dynamic stereograms based on eye convergence for displaying multilayered images. *In SIGGRAPH ASIA*, ACM Press (2012).