Volume 4 of the Proceedings of HCI International 2003,
pp. 1452-1456, 2003.

Gaze-Based Interaction for Anyone, Anytime

Takehiko Ohno, Naoki Mukawa
NTT Communication Science Laboratories, NTT Corporation
3-1 Morinosato Wakamiya, Atsugi, Kanagawa, 243-0198, JAPAN
takehiko@brl.ntt.co.jp, mukawa@eye.brl.ntt.co.jp

Abstract Gaze-based interaction lets users to operate computers through eye movement, and promises to especially benefit severely disabled people who can barely move their hands or feet. However, existing computer operation by gaze may be neither effective nor comfortable, because of physical restrictions on gaze detection and because eye movement during gaze-based interaction is different from natural eye movement. This article describes the requirements for gaze-based interaction, and presents our approaches to making such interaction useful for anyone, anytime.

1 Introduction

Gaze-based interaction is desired for severely disabled people who can barely move their hands or feet. People who have lost the ability to control their muscles due to amyotrophic lateral sclerosis (ALS), cerebral palsy, or spinal cord injuries, primarily communicate with other people by gaze. However, compared with general communication methods, like conversation and gestures, communication by gaze is difficult because the gaze pattern must be translated into meaningful ideas. Then there is the additional problem of slow transmission speed because translation wastes much time.

Gaze-based interaction has two tasks: computer operation and support of human communication [Fig. 1]. Computer operation includes document browsing (e.g., World Wide Web), creative activities, like drawing pictures and writing, and other activities by using a Graphical User Interface (GUI). Human communication includes the expression of intentions to other persons and a two-way conversation between people. The most essential operation for these tasks is object selection. When a user is able to select objects by gaze, it is possible to execute fundamental operations associated with these tasks.

The study of gaze-based interaction started at the end of the 1980s. However, the proposed interaction techniques were not very useful or helpful to even nondisabled users, let alone disabled ones. Overcoming the limitations that prevents its practical use will require far more research.

In creating gaze-based interaction, the challenge is to develop an environment that maximizes the effect of the task while it minimizes the effort of control for anyone. Moreover, gaze-based interactive environments provide that they offer a more effective and comfortable alternative to other interactive systems for anyone, would also be very useful in improving the interaction for disabled people.

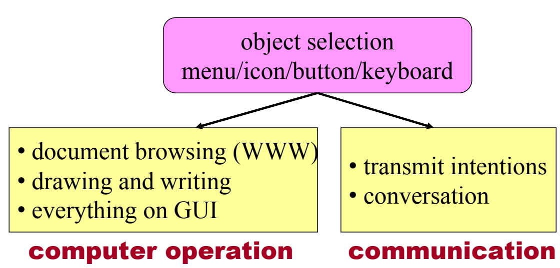
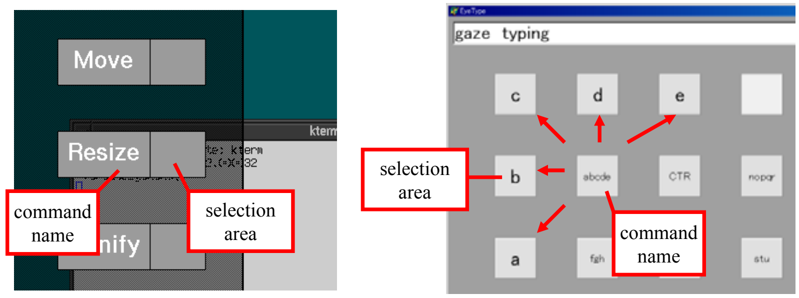


Fig. 1. The category of the gaze-based interaction for disabled users.



(a). Window control (b). Eye keyboard

Fig. 2. Quick Glance Selection Method.

In the rest of this article, we first present the requirements for the selection method and briefly compare existing methods, including ours. Next, we show the requirements for a gaze tracking technique, which is necessary to develop the gaze-based interaction, and our solution. Then, we discuss remaining problems, and finally, mention future work.

2 Gaze Selection Method for Effective Interaction

To maximize the effect of the gaze selection task with minimum effort, we should take into account the features of eye movement. Gaze moves fast, but it is not appropriate to use it for pointing because it is hard to control precisely. It is also difficult to keep the eye on a fixed point for long periods of time. In general, the eyes stay for only 100 to 300 ms on a position before jumping to a different point.

The gaze selection method should (1) minimize fatigue during selection, (2) be intuitive and easy to use, and (3) become easier and faster with repeated use. The first requirement is important when the method is used continuously. In selecting a target by gaze, the eye movement is definitely different from natural eye movement. The eye is not designed to control its direction precisely and fix at a position for a long time. It is also difficult to focus on some point when no indicator is displayed. The second requirement depends on users' skill. If the user is a novice, the system should be intuitive. Appropriate visual feedback also helps the user understand the system's behavior. The third requirement is significant when the method is continually used. With repeated use, users' performance improves because the selection process becomes automated. A bottleneck like dwell time in the selection would prevent this.

To satisfy these requirements, design criteria should be accumulated. Ingenuity of the interface design may significantly improve usability. Here, we clarify the difference between gaze selection and other selection methods. The most salient difference is the way of selection detection. Other selection methods have definite selection events like pressing a button or touching a screen. Gaze selection is different. There is no distinction between just looking at the screen and selecting a target. Therefore, if an object is designed to respond when the user looks at it, it always does so, when for a casual glance. To avoid this problem, several selection methods have been proposed (Majaranta, et al., 2002). Examples are dwell time in selection (Hansen, 1995; Jacob, 1990), and eye blinks and winks for object selection (Shaw, et al., 1990). To distinguish selection from general eye movement, these events should not appear while the user is performing some tasks on the computer. A continuous zooming selection method called *Dasher*, which achieves fast text input, has been proposed (Ward, et al., 2002).

We have proposed a selection method named the *Quick Glance Selection Method* (QGSM). QGSM has separate areas for each menu (Ohno, 1998): a command name area and a selection area [Fig. 2 (a)]. A menu is selected only if the user looks at the selection area, and while the user looks at the command area, an indicator appears at the corresponding selection area. Users who

Tab. 1. The features of four selection methods.

	Dwell time	Blinking and wink	Dasher	QGSM
Minimal fatigue	−	+−	−	+
Easy to use	+	+	−−	+−
Performance improvement	−	−	+	+
Selection speed	−−	−	++	+

do not know menu items look at a command name to find the target menu, and then look at the selection area for selection. If target location is known, the user looks at the selection area directly. We confirmed that this simple method allows very fast selection (Ohno, 1998). QGSM also enables eye-controlled keyboard typing [Fig. 2 (b)]. In this case, five or more characters are assigned to a button (in this figure, they are *a, b, c, d* and *e*). When the user glances at the key, its selection areas appear near it. When the user looks at one of them, the selected character is typed. Otherwise, the selection is canceled. In this design, the eye must be moved twice to select a character, but neither dwell time nor blinks are needed.

The features of those methods are summarized in Table 1. For severely disabled people, fatigue is the most serious problem because they must use the system for a long time and have no alternative means of interaction. Selection speed is also an issue when the user wants to execute most activities by gaze. From this point of view, a zooming approach like Dasher or a consecutive selection approach like QGSM seems to be suitable. However, user evaluations to compare different methods have not been done. More effects are necessary in order to find the best solution and develop more adequate methods.

3 Gaze Tracking System for Gaze-Based Interaction

Gaze tracking systems catch the user's eye to calculate gaze direction and play a key role in gaze-based interaction. However, existing systems are far from comfortable and the systems are inefficient compared with the mouse and keyboard. This is because they were designed to measure users' activity, so that the most important requirement was accuracy. To achieve comfortable gaze-based interaction, a gaze tracking system should (1) have a simple configuration process, including personal calibration, and (2) allow free movement of user's head position. Most existing systems require personal calibration to compensate for gaze direction error. The calibration is done at the beginning of gaze tracking, where the user looks at five to twenty points on the screen. This is a time consuming and burdensome task. A calibration-free method that uses two light sources and the two cameras has been proposed (Shih et al., 2000). However, their algorithm does not completely correct gaze direction error; some residual error remains. To ease the head position restriction, some commercial gaze tracking systems have face tracking mechanism like a pan-tilt camera or movable mirror.

3.1 FreeGaze: An Easy-To-Use Gaze Tracking System

To satisfy those requirements, we developed a gaze tracking system called *FreeGaze* [Fig. 3]. FreeGaze requires the user look at only two points on the screen at the very first time of use (Ohno et al., 2002). Once calibration is finished, the calibration data can be reused; a new calibration is not necessary. To accomplish this advantage, we developed a geometric eyeball model, where major factors of gaze direction error like refraction on the surface of cornea are compensated while the gaze direction is calculated. The head-free type allows free head movement because it detects the user's eye position using a stereo camera mounted on the top of the screen (Ohno et al., 2003). FreeGaze uses the corneal reflection method in which infrared light is used to detect the pupil and the Purkinje image (corneal reflection image).

The system consists of an infrared camera and an infrared LED array. A Windows or Linux PC is connected to the system to calculate the gaze direction and control the camera. The sampling rate is 30 frames per second, and the accuracy is about 0.8 degree in view angle.

(a). Head-fixed type. (b). Head-free type.

Fig. 3. An Easy-to-use gaze tracking system *FreeGaze*.

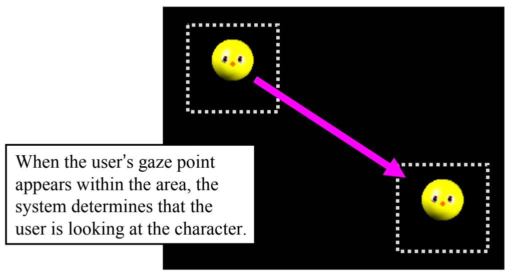


Fig. 4 Simple personal calibration technique.

3.2 Simple Personal Calibration

In general, personal calibration is performed by pressing a key while looking at a marker on the screen. However, for people, can not press keys, so the calibration should be doable without having to do that.

We developed a simple personal calibration technique that requires no keyboard operation. In this procedure, a small agent-like character appears on the upper-left corner of the screen [Fig.4], and jiggles about to attract the user's attention. After the user looks at the character, the character moves to the opposite corner, and as soon as the user looks at it again, the personal calibration is completed. This method is possible because FreeGaze can estimate the user's gaze position approximately even if personal calibration has not yet been performed, and only two points are required for the calibration. FreeGaze and its simple personal calibration technique enable the user to control the computer by gaze with little or no support from other people.

4 Discussion

We described several gaze-based selection methods that will help the disabled communicate with other people and perform some simple selection tasks. So then, is it possible to develop all applications shown in Fig. 1 by these methods? The answer is *"no"*. The most serious problem is the layout limitation. Because gaze is not as precise as a mouse cursor, object size should be larger than the standard icons or menus. Some margin area is necessary between objects to prevent selection error, and how to build complicated applications that are controllable by gaze is still an open problem.

Advancing gaze-based interaction for disabled users in the future requires a method of estimating user's intentions from gaze. When this is done, the user does not have to control her gaze to accomplish the task; rather the task will proceed almost automatically. Several methods of estimating the user's intentions from gaze patterns have already been proposed (Campbell, 2001; Edwards, 1998; Salvucci, 2001). These approaches will improve user performance for complicated tasks with minimum operation burden. Currently, the application areas are limited and sometimes false detection results are output, therefore, further research is required.

5 Conclusion and Future Works

We described the requirements for the selection methods and gaze tracking systems that must be met to realize the gaze-based interaction for disabled people. There are still many problems to be solved. We will continue to develop both gaze tracking systems and gaze-based interaction techniques with the aim of providing universal access to the computer. Worldwide research efforts should be made to overcome the interaction barrier between disabled people and the external environment.

Acknowledgement

We acknowledge Dr. Ishii, Director of NTT Communication Science Laboratories, for his support in this project.

References

Campbell, C.S., & Maglio, P.P. (2001). A Robust Algorithm for Reading Detection, *In Proceedings of the ACM Conference on Perceptive User Interfaces (PUI '01)*.

Edward, G., (1998). A Tool for Creating Eye-Aware Applications that Adapt to Changes in User Behavior , In *Proceedings of the third international ACM conference on Assistive technologies (ASSETS 98)*, 67-74.

Hansen, J.P., Anderson, A.W. & Roed, P. (1995). Eye-Gaze Control of Multimedia Systems, Y.Anzai, K.Ogawa & H.Mori (Ed.), *In Symbosis of Human and Artifact*. *Proceedings of the 6th international conference on human computer interaction*, vol.20A, Elsevier Science Publisher, 37-42.

Jacob, R.J.K. (1990). What You Look At is What You Get: Eye Movement-Based Interaction Techniques. *In Proceedings of CHI'90*, ACM Press, 11-18.

Majaranta, P. & Räihä, K. (2002). Twenty Years of Eye Typing: Systems and Design Issues, *In Proceedings of Eye Tracking Research & Applications Symposium 2002*, 15-22.

Ohno, T. (1998). Features of Eye Gaze Interface for Selection Tasks, Takehiko Ohno, The Third Asia Pacific Computer Human Interaction - APCHI'98, IEEE Computer Society, 176-182.

Ohno, T., Mukawa, N., & Yoshikawa, A. (2002). FreeGaze: A Gaze Tracking System for Everyday Gaze Interaction, In *Proceedings of Eye Tracking Research & Applications Symposium 2002*, 125-132.

Ohno, T., Mukawa, N., & Kawato, S. (2003). Just Blink Your Eyes: A Head-Free Gaze Tracking System, In *CHI 2003 Extended Abstract*, to appear.

Salvucci, D.D., & Anderson, J.R. (2001). Automated Eye-Movement Protocol Analysis, *Human-Computer Interaction*, 16, 39-86.

Shaw, R., Crisman, E., Loomsi, A., & Laszewski, Z. (1990). The Eye Wink Control Interface: Using the Computer to Provide the Severely Disabled with Increased Flexibility and Comfort, *Third Annual IEEE Symposium on Computer-Based Medical Systems*, 105-111.

Shih, S.W., Wu, Y.T., & Liu, J. (2000). A Calibration-Free Gaze Tracking Technique. *In Proceedings of the International Conference on Pattern Recognition (ICPR'00)*.

Ward, D.J. & MacKay, D.J.C. (2002). Fast Hands-free writing by Gaze Direction, *Nature*, Vol.418, 838.