MobiGaze: Development of a Gaze Interface for Handheld Mobile Devices

figure 1. MobiGaze

Takashi Nagamatsu
Kobe University
5-1-1, Fukae-minami, Higashi-nada, Kobe 658-0022, JAPAN
nagamatu@kobe-u.ac.jp

Michiya Yamamoto
Kwansei Gakuin University
2-1 Gakuen Sanda 669-1337 JAPAN
michiya.yamamoto@kwansei.ac.jp

Hiroshi Sato
Kwansei Gakuin University
2-1 Gakuen Sanda 669-1337 JAPAN
atp20548@ksc.kwansei.ac.jp

CHI 2010, April 10–15, 2010, Atlanta, Georgia, USA.
ACM 978-1-60558-930-5/10/04.

Abstract
Handheld mobile devices that have a touch screen are widely used but are awkward to use with one hand. To solve this problem, we propose MobiGaze, which is a user interface that uses one's gaze (gaze interface) to operate a handheld mobile device. By using stereo cameras, the user's line of sight is detected in 3D, enabling the user to interact with a mobile device by means of his/her gaze. We have constructed a prototype system of MobiGaze that consists of two cameras with IR-LED, a Windows-based notebook PC, and iPod touch. Moreover, we have developed several applications for MobiGaze.

Keywords
Gaze, Eye tracking, Stereo camera, Mobile device

ACM Classification Keywords
H.5.2 [Information Interfaces and Presentation]: User Interfaces–Input devices and strategies; I.4.9 [Image Processing and Computer Vision]: Applications

General Terms
Design, Measurement

figure 2. Difficulty of usage of mobile touch screen device by one hand.

Introduction

Today's handheld mobile devices that have a touch screen are used widely, and users can interact with these devices intuitively.

When using a mobile device, the user holds the device in one hand and touches the screen using the other hand. However, when the user is holding something or has one hand occupied with something else, the device is difficult to control with the single hand that is available. In this common circumstance, most users touch the screen using their thumb, as shown in figure 2. It is difficult, however, to touch the entire area of the display with one's thumb; when one touches the top area of the screen, the device becomes unstable and there is a possibility that the user will drop the device. Moreover, as the displays on devices become increasingly larger, one-hand-control becomes even more difficult. In addition to this, the large area of the display is blocked by the thumb, which creates further problems.

To solve these problems, we propose a gaze tracking technique. Gaze tracking can be a means of pointing on a mobile device that does not require the use of both hands and involves no blocking of the screen. If a mobile device has a gaze tracking function, only one hand is needed. Pointing with one's gaze is also quicker than a track ball or pointing stick, and the technique is easily learned.

Eye-tracking devices can be divided into two categories: head-mounted and remote. Because a head-mounted gaze tracker is not suitable for everyday life, we decided to construct a remote type gaze tracker for a mobile device.

A number of problems arise with the introduction of gaze tracking technology to a mobile device. Because a mobile device is small, the measurements made by a gaze tracking system require a high degree of accuracy in order to indicate objects on a small screen. Moreover, the user moves the hand which holds the mobile device. Since the camera is attached to the mobile device, this hand movement causes quick changes in the direction the camera is pointing. A successful gaze tracking method must allow free hand movement while still providing the requisite accuracy.

Several studies have sought to detect the point of gaze (POG) of mobile devices. Lukander constructed a system which makes possible the measurement of the POG on the screen surface of a handheld mobile device. His system used a commercial head-mounted eye tracker and a magnetic positional tracker [3]. Drewes et al. investigated how gaze interaction can be used to control applications on handheld devices. They used a commercial tabletop eye-tracker in combination with a mobile phone attached to a screen [1]. Thus far, there has been no gaze tracking system that can be handheld and that has no need for headgear.

In this paper, we describe a gaze estimation method for MobiGaze and the hardware implementation that we have constructed. We also introduce some applications for this new interaction device.

Estimation of Point of Gaze

A method that successfully realizes gaze tracking on a mobile device must allow free movement of the hand. For this reason, we adopted a stereo gaze tracking method[5]. This method can measure the position of

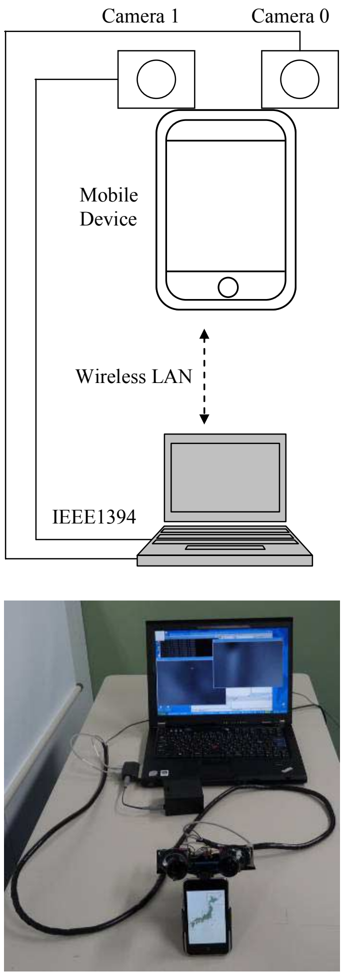

figure 4. System configuration and prototype of the system

the center of the cornea and the vector along the line of sight relative to the two cameras that are employed.

Each of the two cameras used in this method has a light source attached to it, and each light source is positioned near the nodal point of the camera.

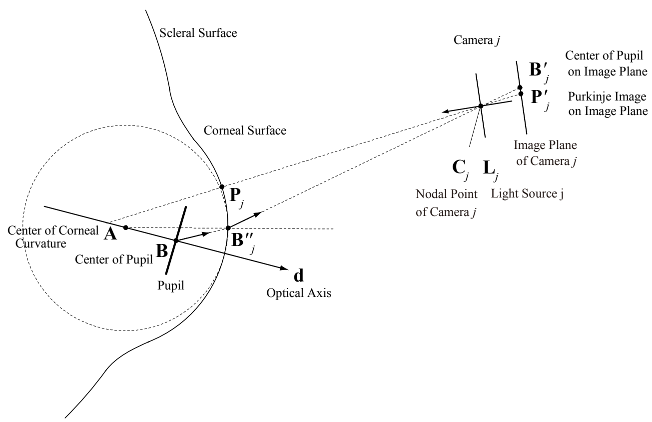

figure 3. Cross section of eyeball showing center of corneal curvature, along with center of pupil, position of light source, and nodal point of camera.

Figure 3 shows a cross section of the eyeball. The geometric center line of the eye is called the optical axis. $\mathbf{A}$ is the center of the corneal curvature. $\mathbf{L}_j$ and $\mathbf{C}_j$ denote the position of the light source j and the nodal point of camera j, respectively; $\mathbf{C}_j$ is assumed to the same as $\mathbf{L}_j$. The value of $\mathbf{C}_j$ $(=\mathbf{L}_j)$ is determined by calibrating the camera beforehand. $\mathbf{A}$, $\mathbf{B}$, $\mathbf{B}'_j$, $\mathbf{B}''_j$, $\mathbf{C}_j$, $\mathbf{L}_j$, $\mathbf{P}_j$, $\mathbf{P}'_j$, and the optical axis of the eye are coplanar. The normal vector of the plane is $(\mathbf{C}_j-\mathbf{B}'_j)\times(\mathbf{P}'_j-\mathbf{C}_j)$, and the plane is expressed as

$$\{(\mathbf{C}_j-\mathbf{B}'_j)\times(\mathbf{P}'_j-\mathbf{C}_j)\}\cdot(\mathbf{X}-\mathbf{C}_j)=0$$

where $\mathbf{X}\ (=(\mathrm{x},\mathrm{y},\mathrm{z})^{\mathrm{T}})$ is a point on the plane. We obtain two planes when we use two cameras ($j=0, 1$). The optical axis of the eye can be determined from the intersection of the two planes, which must not be coplanar.

The line of sight connecting the fovea and the POG is called the visual axis. The user calibration that determines the offset between the optical and visual axes is performed by using the one-point calibration method described in [4].

Hardware Implementation

Because current mobile devices do not have two cameras or the power to process the eye image, we constructed a prototype system using a Windows-based notebook PC for image processing, etc.

Figure 4 shows the system configuration at this time. The system consists of two synchronized monochrome IEEE-1394 digital cameras (Firefly MV, Point Grey Research Inc.), a mobile device (iPod touch 3.1, Apple Inc.), and a Windows-based notebook PC (Windows XP, Intel Core 2 Duo T7500 2.2GHz). The two cameras and the mobile device are connected to the notebook PC via IEEE 1394 and a wireless LAN, respectively.

Each camera is equipped with a 1/3'' CMOS image sensor whose resolution is 752 × 480 pixels. A 25-mm lens and an IR filter are attached to each camera. One infrared LED is attached to each camera near the camera's nodal point. These cameras are positioned on the top of the mobile device. The software on the PC was developed using OpenCV 1.0. The refresh rate was approximately 10 fps.

Benefits of Gaze-and-Touch Interface on a Mobile Device

As shown in figure 5, when a user holds a mobile device, he/she can at most touch the small green area near the thumb. For this reason, we propose a new interaction technique which employs a gaze-and-touch interface.

A gaze-and-touch interface easily avoids the famous Midas-touch problem [2]. The user moves the cursor by eye and selects an object by simply touching anywhere on the screen. The user can also perform a swiping action, in addition to the usual touch action.

For information process on a mobile device, a gaze-and-touch interface can assemble the user's interests as recorded by both POG as well the screen that has been touched. A mobile device with this new interface can thus also serve as a good personalized information device.

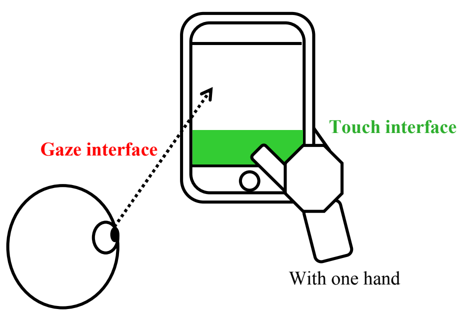

figure 5. Gaze-and-touch interface allows new interaction.

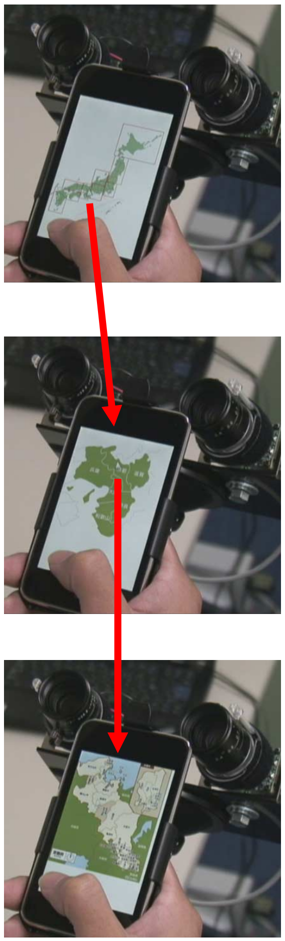

figure 7. Map browser: The region pointed by gaze is zoomed up by touching anywhere on the screen.

Application Example: Map browser

We developed a map browser as an example application for MobiGaze. Figure 6 shows the top page of a map browser on which a map indicates the islands of Japan.

When a user gazes at a region (Figure 7 (top)) and touch the screen, a detailed map is displayed (Figure 7 (middle)). Ultimately, the user can obtain a close-up map (Figure 7 (bottom)).

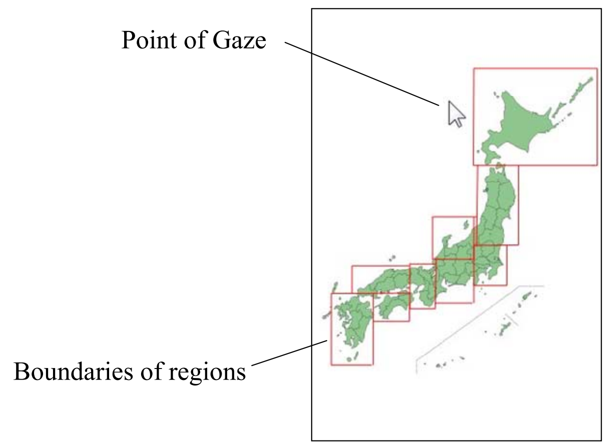

figure 6. Top page of map browser.

Extended prototype: Dual-screen interface

Some mobile devices, such as Nintendo DS, have a dual display. When using such a dual display device, it is no longer possible to operate the mobile device with one hand. With a gaze tracking technique, however, the user can easily interact with a dual screen. We have constructed a prototype system for a dual screen device, which is shown in figure 8.

figure 8. Dual screen type: It is no longer possible to operate such a mobile device with one hand.

A map browser also works on a dual screen device. Here, when a user swipes the lower screen, the maps in the upper screen and lower screen move in synchronization. In this way, even if the display of a mobile device becomes much larger and the area that the thumb can touch remains small, much information can be acquired or transmitted by selecting several buttons or with several swiping gestures.

We have also developed a new browsing application, as shown in figure 9. It provides buttons to select categories on the upper screen, and a contents area on lower screen. First, a user gazes at a category and then taps anywhere on the lower screen. One of the categories is selected and its content is displayed. The user can proceed to read by employing a swiping gesture. In this way, the quick and easy interaction of the gaze interface and the familiar touch interface are effectively combined.

Banner ads are also found on the upper screen. By analyzing the user's gaze, the desired banner ads can be displayed based on the history acquired not by intentionally selecting them, but by simply looking at them.

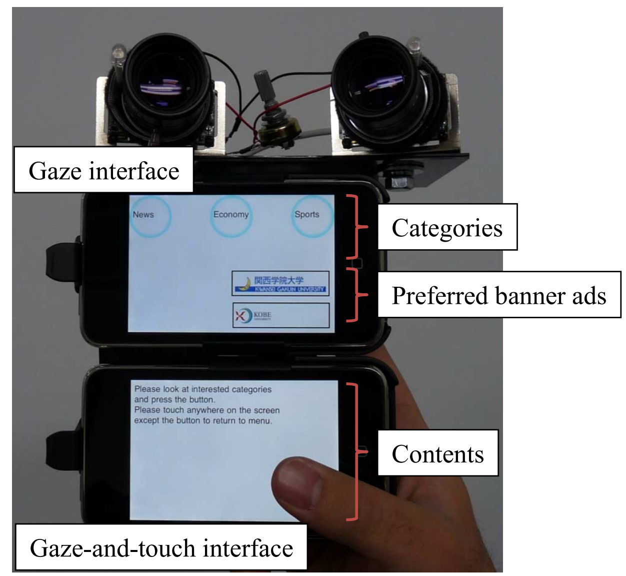

figure 9. New interface design for a dual screen type gaze tracking mobile device

Discussion

Since mobile devices are small, the usability of MobiGaze depends on its accuracy of measurement. Generally, the accuracy of a gaze tracker is 0.5 -1.0°, because "a user generally need not position his eye more accurately than the width of the fovea (about one degree) to see an object sharply [2]." This represents the limitation of the accuracy of gaze measurement. However, if the user can move the mobile device closer to his/her face, the accuracy of measurement in the view angle is same, but the accuracy of measurement in the screen coordinate system in pixels can become more accurate. MobiGaze can therefore realize a high and effective resolution.

Since it is expected that in the near future, mobile devices will have sufficient power and will be equipped with stereo cameras, gaze tracking will be able to work in the mobile device itself.

The current implementation does not allow everyone to use MobiGaze with ease, and some people cannot use it. That said, because mobile devices are for personal use, individuals can choose whether or not to adopt this technology. As well as an accelerometer or a GPS that are equipped with mobile devices of recent date, the MobiGaze will expand the potentiality of mobile devices.

In another application, if an E-book reader is equipped with gaze tracking technology, the history of his/her reading can be recorded and the data can be used to anticipate important passages for the reader, abstract selected information, etc.

Future work in the development of MobiGaze includes improving the performance of its gaze measurement.

Conclusion

In this paper, we have described MobiGaze, which is a gaze interface for handheld mobile devices. With the use of stereo cameras, a user's line of sight is detected in 3D, which enables the user to interact with a mobile device by means of his/her gaze. We have constructed a prototype system of MobiGaze that consists of two cameras with IR-LED, a Windows-based notebook PC, and iPod touch. We have also developed several applications for MobiGaze which take advantage of its gaze-and-touch interface.

References

[1] Drewes, H., Luca, A.D. and Schmidt, A. Eye-gaze interaction for mobile phones. in *Proceedings of the 4th international conference on mobile technology, applications, and systems* (2007), 364-371.

[2] Jacob, R.J.K. The use of eye movements in human-computer interaction techniques: what you look at is what you get *ACM Transactions on Information Systems*, *9* (2) (1991), 152-169.

[3] LUKANDER, K. A system for tracking gaze on handheld devices *Behavior Research Methods*, *38* (4) (2006), 660-666.

[4] Nagamatsu, T., Kamahara, J. and Tanaka, N. 3D Gaze Tracking with Easy Calibration Using stereo Cameras for Robot and Human Communication. in *Proceedings of the 17th International Symposium on Robot and Human Interactive Communication (IEEE RO-MAN) 2008* (2008), 59-64.

[5] Nagamatsu, T., Iwamoto, Y., Kamahara, J., Tanaka, N. and Yamamoto, M. Gaze Estimation Method based on an Aspherical Model of the Cornea: Surface of Revolution about the Optical Axis of the Eye. in *Proceedings of the 2010 Symposium on Eye Tracking Research & Applications* (2010).