A Real-World Pointing Device Based on an Optical Communication System

Yuichi Mitsudo

116-2 Kamedanakano-cho, Hakodate-shi, Hokkaido, Japan
Future University Hakodate
mitsudo@fun.ac.jp

Abstract. In the present paper, a new augmented reality environment that is based on an optical communication system is described. Optical communication devices have been used in several studies on ubiquitous computing. A novel physical structure of an optical communication system that enables the user to select the optical signal by simply pointing its transmitter with his/her finger is developed. In such an environment, the optical transmitter can be treated as a visual tag, referred to as a GhostTag, that includes continuous data, such as audio files. In addition, the PointSpeech application, which provides the user with audio assistant data via GhostTag, is presented herein.

1 Introduction

In the present paper, a new augmented reality environment, which enables the attachment of continuous data to real-world objects as an annotation, is presented. In some studies, RF tags or visual tags are used to attach small amounts of data to real-world objects[1,2,3]. The annotations are stored in the tags as a visual pattern on paper or as digital data in small circuits. The user must be close to the tag and readout the annotations. In the present system, the annotation is placed within the visual angle of the real-world object (Fig.1). The visual angle within which an annotation is placed is referred to as the **GhostTag**, and the user reads out the annotation by a pointing behavior, referred to as **TeleClick**.

This visual-angle-based annotation system is similar to a GUI system, in which the GhostTag corresponds to an icon, and the index finger corresponds to the cursor. TeleClick, in which the user overrup his/her index finger on the visual angle of object(Fig.1, Fig.3) is the method used to select an annotation.

2 GhostTag

GhostTag is essentially the visual angle around a tag. GhostTag can be generated by a simple combination of a transmitter, photo-detector and an image sensor. Figure 2 shows the system used to generate and read GhostTags.

An optical transmitter(tag), which transmits optical signal, is attached to a real-world object. When an image sensor observes the this transmitter, a photodetector, which is directed toward the surface of the image sensor, detects the

R. Shumaker (Ed.): Virtual and Mixed Reality, LNCS 5622, pp. 70–79, 2009.

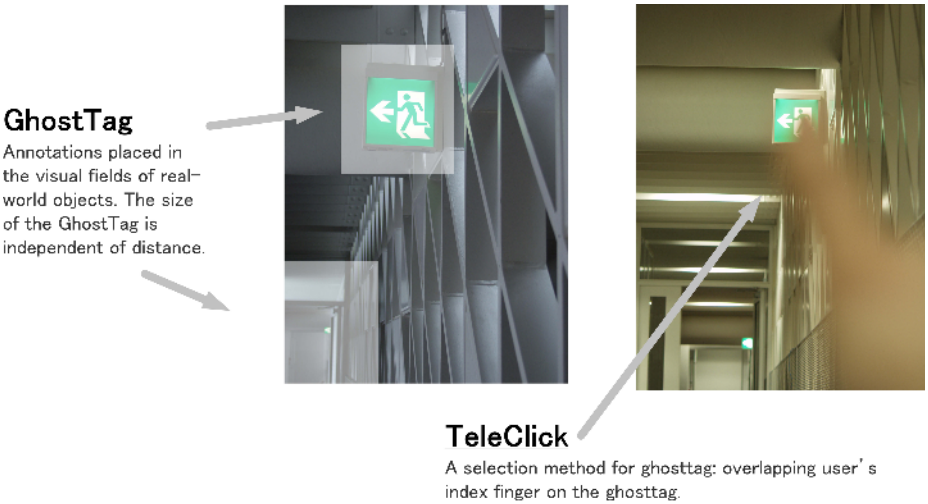

Fig. 1. GhostTag: continuous data placed within in the visual angle, TeleClick: selection method of GhostTag

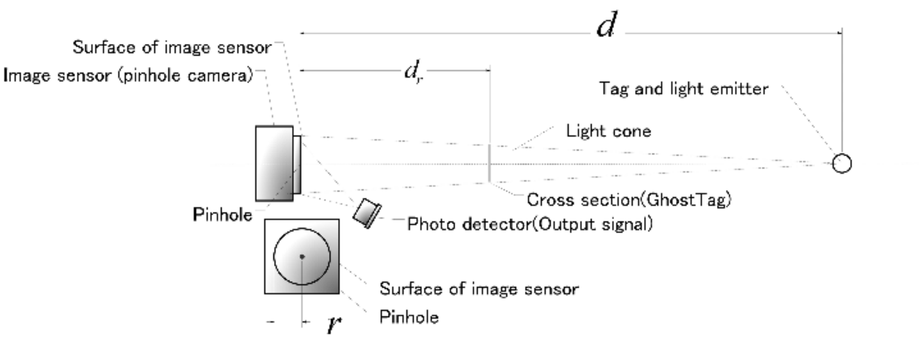

Fig. 2. Schematic diagram of the GhostTag setup

optical signal reflected by the surface of the image sensor. The optical signal is a cone-shaped and is symmetric about an axis formed by a line running from the center of the image sensor to the center of the light emitter. The cross section of this cone is GhostTag. The center of GhostTag is a light emitter on the tag, and the size of GhostTag is determined by the distance between the image sensor and cross section d_r. The visual angle of GhostTag, θ_r, is determined primarily by d_r.

Generally, optical communication systems that use transmitters and receivers have time-domain resolution, rather than spatial resolution. Owing to its simple structure, in which the photodetector is directed toward the image sensor and senses the signal reflected from the surface of the image sensor, the proposed system has both time-domain and spatial resolutions.

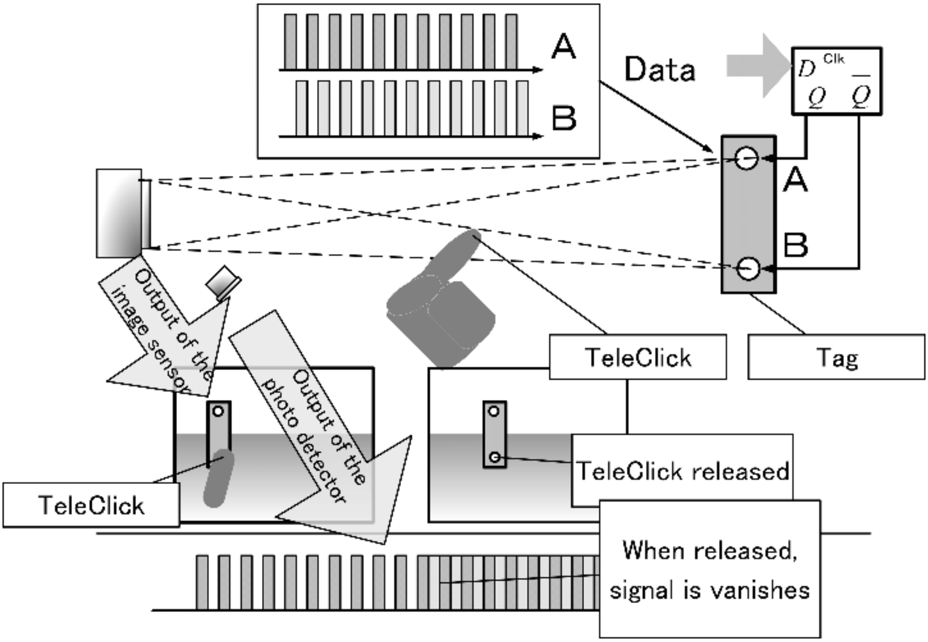

Fig. 3. Diagram of TeleClick operation

3 PointSpeech

PointSpeech is an application system that provides audio assistance to the user. The user obtains sound data from GhostTag by TeleClicking. In the following, the method used to detect a TeleClick by the user and read out the contents of the TeleClicked GhostTag is described. The optical transmitter, attached to real world object as a tag, has two LEDs: one for transmitting data signals (Q) and another for transmitting inverted data signals ($\bar{Q}$) (Fig. 3). When a receiver receives both these signals, it outputs a flat signal (data signal is eliminated). When a user blocks an inverted signal by TeleClicking as described above, the data signal can be observed by the receiver. In PointSpeech system, Pulse Width Modulation(PWM) is used to translate auditory data.

3.1 RealEyeCommunicator/VisionCommunicator

Receiver of PointSpeech system is implemented through RealEyeCommunicator(REC)/VisionCommunicator(VC), a receiver part(image sensor and photodetector) of GhostTag generator. VC includes a cell phone CCD as the image sensor , and the other receiver uses the eye of the user as the image sensor. By implementing through REC/VC system, PointSpeech does not require image processing to select and readout the data from GhostTag, althogh it is a

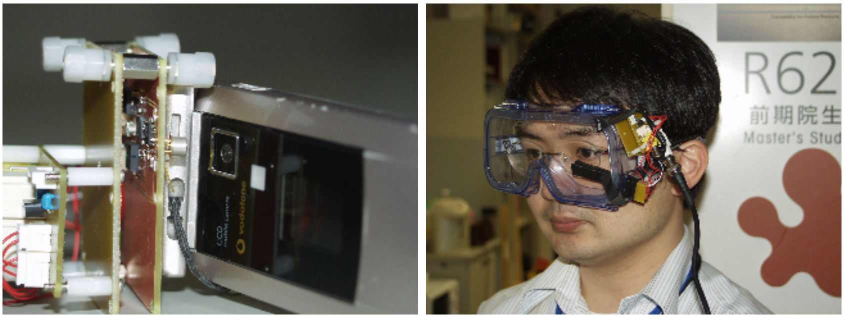

Fig. 4. A view of VisionCommunicator/RealEyeCommunicator

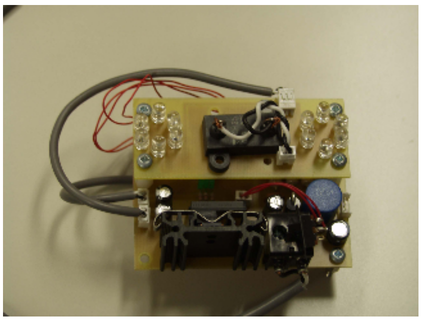

Fig. 5. A transmitter of VC/REC

vision-based pointing system. This enables the use of high-speed signals as an annotation, and helps to employ numerous optical systems, including the eye of the user as an image sensor.

4 System Implementation

PointSpeech was implemented in the proposed optical system. Audio data is broadcast from transmitters, and the user selects the transmitted signals by pointing at the transmitter. Two types of receivers, REC and VC are used.

5 Experiments

In this section, the visual angle of GhostTag and its pointing time are described through experiments. In the following subsections, an optical translator, for which the LED distance is 6.0 cm, is used as a tag. The tag emits a digital optical signal, the content of which is a sound generated by pulse width modulation (PWM).

Pulse width modulation is a popular digital modulation for conveying analog information over communication channels. Pulse width modulation uses a rectangle wave that is pulse width modulated such that the average value is equal to the source signal. By the principle of PWM, the following characteristics can be lead. The user can retrieve same sound weather which LED to be TeleClicked. In the following section, the visual angle of the GhostTag and the time required to TeleClick are obtained for several cases.

5.1 Visual Angle Required for Read out

The visual angle of GhostTag required in order for VisionCommunicator or RealEyeCommunicator to be used as a receiver is shown.

VisionCommunicator. The following experiment shows the visual angle of GhostTag when VisionCommunicator is used. The experiment is performed when $d_r = 0.4$m and a cursor of 16×10^{-3} m in diameter is used, and both the distance d and the position of the cursor are manipulated (Fig.6).

In this experiment, the integrated cell-phone camera (Fig. 4) with a resolution of 160×120 is integrated into the system. The experimental value is shown in Figure 11, and its view is shown in Fig. 7.

Before the cursor touches the optical signal, the voltage output by the receiver is flat. When the signal appears, that cursor position is left (right) edge of the GhostTag. The horizontal length of GhostTag is measured by this method.

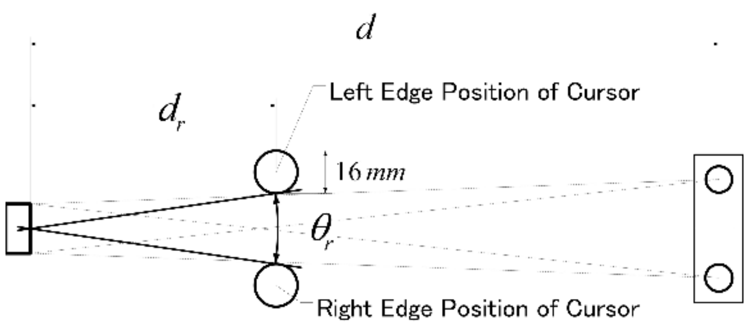

Fig. 6. Experimental setup

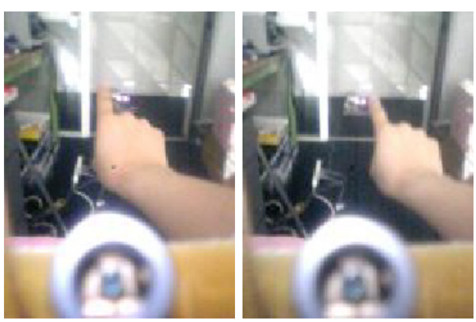

Fig. 7. Left and right edges of GhostTag

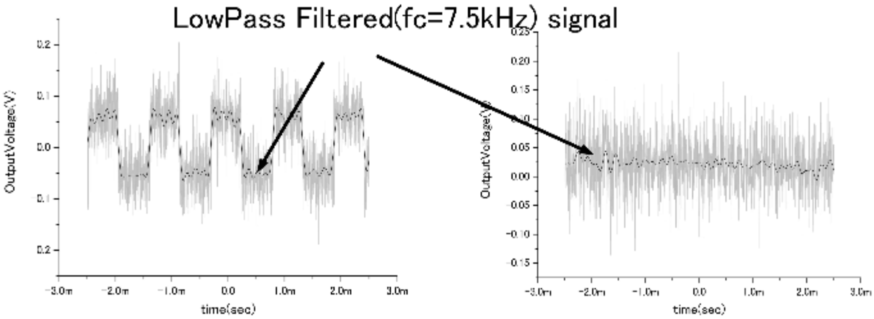

Fig. 8. Examples of output signals: clicked signal (left) and normal signal (right)

Figure 8 shows the normal and "clicked" signals.

Based on these experimental results, VisionCommunicator has sufficient spatial resolution to detect a TeleClick performed by the index finger of the user.

RealEyeCommunicator. Next, an experiment using the RealEyeCommunicator (REC) as a receiver. The REC uses the human eye as its image sensor. A glass-like interface is constructed and a photo-diode is attached near the eye of the user to receive a reflected optical signal from the eye (Fig. 4). Then, a ruler is placed behind the transmitter, and the position of the cursor is measured (Fig. 9). A graph of the position and the received signals when d =1.5m and d_r =0.4m is shown in Fig. 10. Both edges (W8 and E6) have a flat signal, and enough signal ($\geq$ 0.005V) to make sounds appeares from W7 to E4. The code

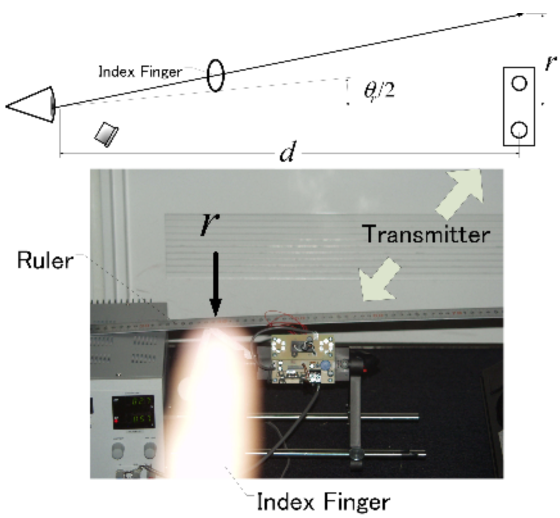

Fig. 9. A view of VisionCommunicator/RealEyeCommunicator

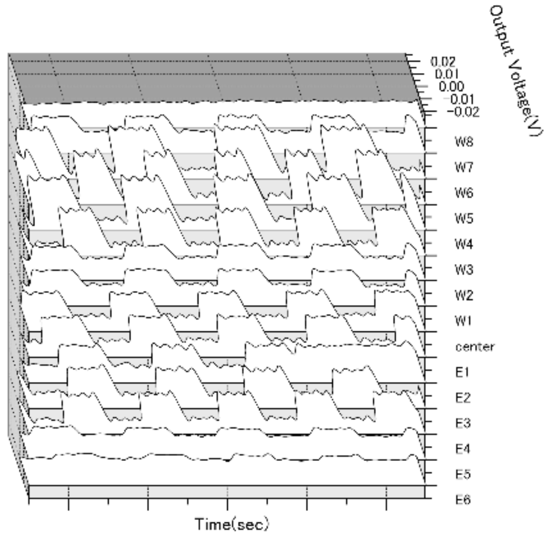

Fig. 10. A view of VisionCommunicator/RealEyeCommunicator

of W7 or E4, means 8.0×10^{-2}m left(or 6.0×10^{-2}m right) of the center. The visual angle of GhostTag(θ_r) can be calculated to

$$\theta_r = \arctan \frac{7}{150} + \arctan \frac{4}{150} - \arctan \frac{0.8}{40} = 3.1° \quad (1)$$

when $d = 150$cm. The last part of this equation, $\arctan \frac{0.8}{40}$ is the visual angle of index finger. The visual angles in case of d=2.0, 2.5m is presented in Fig.11. The visual angles for d =2.0m and 2.5m are shown in the following figure (Fig. 11)

5.2 Time Required for Pointing and TeleClicking

The movement of TeleClicking, overupping his/her fingertip on a target in view, is similar to natural pointing. To evaluate this novel movement, an experiment is conducted to compare the pointing time of TeleClick to the times for the other pointing methods. Five subjects participated in this experiment. A black circle is displayed on the screen, and the subjects are instructed to point to or TeleClicked the targets as quickly as possible(Fig. 12).The distance between the screen and the subject is 2.5m, and the diameter of the target is 0.6×10^{-2}m (visual angle: 2.5°). The experiments were carried out as follows: natural pointing, TeleClicking with natural view, TeleClicking with cell phone camera view, and laser pointing (Fig. 12) for five subjects. As shown in Fig. 12, the pointing times for natural pointing, natural TeleClicking, and laser pointing are similar, but the pointing time for the cell phone system is approximately double compared to the other pointing methods. One reason of delay is the low refresh rate of the cell-phone CCD.

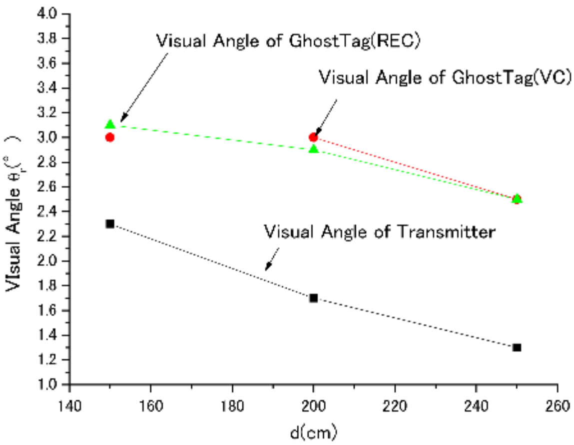

Fig. 11. Visual angles of GhostTag of VC and REC

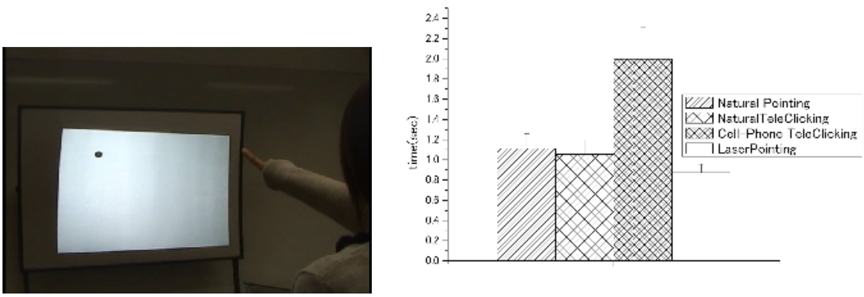

Fig. 12. Experimental view and results

6 Previous Work

Pierce present a interaction techniques called "Stickey Finger", overlapping user's index finger to the object in head-tracked immersive virtual world[4]. In this work, some other interaction techniques, Head Crasher, Framing, Lifting Palm, using 1st person view, are presented. Interactive Sight and TeleClick, which using user's viewfield, is presented in the authors previous work[5,6]. These system includes a receiver of Real Eye Communicator, but transmitter is not developed. A interaction method to optical signal with TeleClick behavior, has been developed in this paper. Ayatsuka presents finger point interaction with visual code, overlapping index finger on the code, in their paper[3].

In some studies, laser pointer, or narrow angle LED were mounted to user's index finger, to detect the target of natural pointing[7,8,9], and evaluating its pointing(motion) time. The pointing time of hand-held laser pointer is reported in several papers[10,11].

7 Conclusions

In the present paper, a real-world streaming data translator, that allows continuous data to be placed in a small visual angle to a real-world object, called PointSpeech, were developed. PointSpeech is based on a cell phone camera (VisionCommunicator) or the human eye (RealEyeCommunicator) with an added optical communication system, which receives optical signals from real-world tags. These optical signals, modulated in both temporal and spatial domains, appear its data part by detecting user's contact. The combination of transmitters and a receiver has sufficient ability to resolute user's contact with his/her index finger, this system can be used as a real-world pointing device, even though it is a data translator. The experimental results indicate that these systems, vision communicator and real eye communicaotr, have very similar characteristics with respect to resolution, data translation ability, and pointing ability. However, the systems differ in their practical applications. The pointing times for these systems are measured and compared to other pointing methods such as natural pointing and laser pointing methods. The results for the eye system are similar to those for the natural pointing and laser pointing methods. However, the time required by the cell phone system is double that of the other methods. In the present paper, a novel optical system that has both the ability to translate audio signals and the ability to distinguish a transmitter which visual angle is 2.8° is presented. These abilities allow the system to be used as a pointing device, and enables to realize PointSpeech system, to attach sound data on the image field of the user.

Acknowledgement

The present study was supported by a Grant-in-Aid for Young Scientists B(20700113).

References

1. Want, R., Fishkin, K.P., Gujar, A., Harrison, B.L.: Bridging physical and virtual worlds with electronic tags. In: CHI 1999: Proceedings of the SIGCHI conference on Human factors in computing systems, pp. 370–377. ACM Press, New York (1999)
2. Rekimoto, J., Ayatsuka, Y.: Cybercode: designing augmented reality environments with visual tags. In: DARE 2000: Proceedings of DARE 2000 on Designing augmented reality environments, pp. 1–10. ACM Press, New York (2000)
3. Ayatsuka, Y., Rekimoto, J.: Active cybercode: a directly controllable 2d code. In: CHI 2006: extended abstracts on Human factors in computing systems, pp. 490–495. ACM Press, New York (2006)
4. Pierce, J.S.: Image plane interaction techniques in 3d immersive environments. In: Proceedings of the 1997 symposium on Interactive 3D graphics (1997)
5. Mitsudo, Y., Mogi, K.: Real eye communicator: An eye-mediated real world pointing device. In: Proceedings of the First International Conference on Pervasive Computing (Short paper), in an informal companion volume of short papers, Zurich, pp. 112–118 (2002)

6. Mitsudo, Y.: Interactive sight: A new interaction method for real world environment. In: HCII 2003, vol. 2, pp. 746–750 (2003)
7. Sibert, J., Gokturk, M.: Natural pointing techniques using a finger-mounted direct pointing device. In: SIGGRAPH 1998: ACM SIGGRAPH 98 Conference abstracts and applications, p. 121. ACM, New York (1998)
8. Gokturk, M., Sibert, J.L.: An analysis of the index finger as a pointing device. In: CHI 1999: extended abstracts on Human factors in computing systems, pp. 286–287. ACM, New York (1999)
9. Tsukada, K.: Ubi-finger: Gesture input device for mobile use. In: Proc. of 5th Asia Pacific Conference on Computer Human Interaction, November 2002, vol. 1, pp. 388–400 (2002)
10. Dan, R., Olsen, J., Nielsen, T.: Laser pointer interaction. In: CHI 2001: Proceedings of the SIGCHI conference on Human factors in computing systems, pp. 17–22. ACM, New York (2001)
11. Myers, B.A., Bhatnagar, R., Nichols, J., Peck, C.H., Kong, D., Miller, R., Long, A.C.: Interacting at a distance: measuring the performance of laser pointers and other devices. In: CHI 2002: Proceedings of the SIGCHI conference on Human factors in computing systems, pp. 33–40. ACM, New York (2002)