RealSound Interaction: A Novel Interaction Method Using Sound Events in Real World

Mai OTSUKI[1], *Asako KIMURA*[2, 3], *Takanobu NISHIURA*[1], *Fumihisa SHIBATA*[1], *Hideyuki TAMURA*[1]

[1]Graduate School of Science and Engineering,
[2]Research Organization of Science & Engineering,
Ritsumeikan University
1-1-1 Noji-Higashi, Kusatsu, 525-8577,
Shiga, Japan

[3]PRESTO,
Japan Science and Technology Agency
4-1-8, Honcho, Kawaguchi,
332-0012, Saitama, Japan

ABSTRACT

We present a unique and novel method using the direction and location of sound events as input or interaction devices into the mixed-reality space. To estimate the direction and location of sound events, we propose and use a wearable microphone array. This paper describes the outline of our system, the interaction methods using the direction and location of sound events, and finally introduces two applications using this new technique.

ACM Classification: H5.2 [Information interfaces and presentation]: User Interfaces. - Input devices and strategies.

General terms: Design, Human Factors

Keywords: Mixed Reality, Sound Input, Microphone Array, Sound Source Localization, and Interactive Device

INTRODUCTION

In this paper, we present a unique and novel method of interacting with a mixed-reality (MR) space that merges real and virtual worlds in real time. Our approach "RealSound Interaction" is based on the idea that the sound events occur in the real world can work as input or interaction devices into/with an MR space. Here, not only the existence of sound but also the direction and location of the sound events are used for human-computer interaction.

The preceding work "PingPongPlus" by Ishii *et al.* [1] detects a real sound on a ping-pong table and utilizes its location as the source of subsequent operations. This system dedicated to the table-tennis-like electronic entertainment can detect only hitting sounds on as the sound events.

On the other hand, our approach is aiming at detecting the direction and location of a variety of sound events and using them for more general interactive operations such as menu selection or pointing. Consequently, our interaction method could be used for various applications.

In addition, the users can realize an intuitive operation with their sound sources such as handclapping.

SYSTEM OVERVIEW

We developed the system that estimates the direction and location of sound events in the real environment and reflecting them into the MR space (Figure 1) [2].

UIST '07, October 7-10, 2007, Newport, Rhode Island, USA.
ACM 978-1-59593-679-2/07/0010

Wearable microphone array

A fixed type of linear microphone arrays has been investigated in the field of acoustics. One of its drawbacks is that it can work well only in a limited range of the front direction because of the low angular resolution in the crosswise direction. In this research, we newly use a microphone array in a wearable fashion by attaching it onto a head mounted display (HMD) (Figure 2). Since the user moves freely and faces to his/her target, the array can capture the sound constantly in his/her front direction.

Direction and location estimation of sound events

A sound source direction can be estimated by one microphone array. CSP (Cross-power Spectrum Phase analysis) method is used for sound source direction estimation algorithm [3]. This method gives a direction of sound source in the horizontal plane.

Two or more microphone arrays can localize a sound source. In our approach, we used one fixed-type micro-

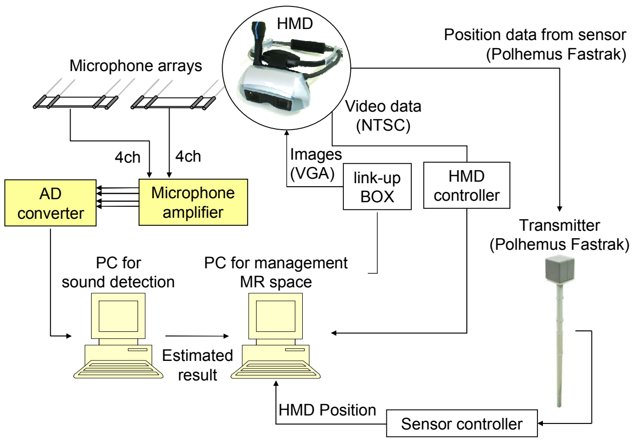

Figure 1: System configuration

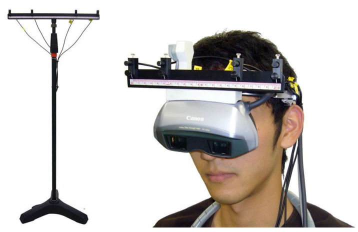

Figure 2: Microphone arrays
(Left: fixed type, Right: wearable type)

Figure 3: Users can select a menu item using a sound direction. In this figure, a birdcall is used as a sound device.

Figure 4: Computer-generated ducklings gather (move) toward the position of real mother duck where squawking sound is played.

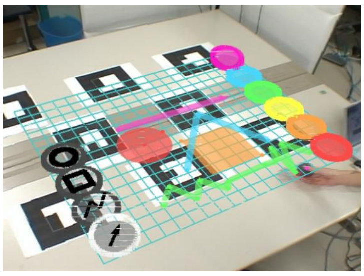

Figure 5: AcousticInk

phone array and one wearable type microphone array. The system that has two microphone arrays arranged in a horizontal direction can estimate the location only on the horizontal plane. It becomes possible to estimate the location of sound events in 3-D space by adding a vertical microphone array.

In order to reduce the estimation caused by users' voice or ambient noise, we gave a weight in appropriate frequency range, and accurately estimated the input sound direction by "weighted CSP analysis" method [3].

In this system, it is possible to use a variety of sound sources, such as handclapping, playing a castanet, or pushing a buzzer. Among them, handclapping seems to be a favorable operation since it is simple and intuitive.

INTERACTION METHODS USING SOUND EVENT

Interaction methods using direction of sound event

(1a) *Menu selection*: Computer-generated graphical menu items are displayed around a user, even though he/she is moving. The user can select one item by generating a sound signal toward it (Figure 3).

(1b) *Direction indication*: Computer-generated virtual objects can appear, disappear, move, or start animation in/toward the direction of the sound generated by a user.

Interaction methods using location of sound event

(2a) *Menu selection:* Some graphical menu items are displayed at several fixed points of MR space. A user can select one item by generating a sound signal at the location of the item.

(2b) *Pointing operation*: Computer-generated virtual objects appear, disappear, move, or start animation at/toward the position of the sound generated by a user (Figure 4).

APPLICATIONS OF REALSOUND INTERACTION

Watch the Birdie!

This is a kind of MR-based entertainment in which a user can enjoy some interactions with virtual birds such as selecting and calling the bird or gathering the ducklings.

Firstly, along the method (1a), the user selects a virtual bird which he/she wants to call using a real birdcall (a device for generating a bird-like imitated sound), then, along the method (1b), the birds fly from the direction of another birdcall. Secondly, using the method (2b), mother duck (a real toy with a speaker) squawking and then many virtual ducklings run toward their mother (Figure 4). The users also can gather them by a handclap.

AcousticInk

Based on RealSound Interaction, we developed a new drawing tool in an MR space, namely "*AcousticInk*" (Figure 5). Along the method (2a), a user can select the shape and color by generating a sound on the virtual palette, and then, along the method (2b), he/she can indicate coordinates by generating a sound in arbitrary location and can draw virtual graphical shapes such as lines or circles.

CONCLUSION

We have developed novel interfaces between the real environment and the MR space through the sound events, and these are unique and useful interactive methods. To realize these methods, we implemented the microphone arrays which are common in the field of acoustics. Accordingly, we can use them not only for ON/OFF by sound events but also for the determination of direction and location of sound events as inputs into the MR space. We have also proposed a new usage of microphone arrays by implementing both traditional fixed type and a new wearable type microphone array attached onto a HMD. These interaction methods are not necessarily used only in MR, but are expected to be widely used in a general system.

ACKNOWLEDGEMENTS

The authors wish to express their sincere thanks to Ms. Akiko Iesaki for her great help in preparing this paper.

REFERENCES

1. H. Ishii, *et al.*: "PingPongPlus: Design of an athletic-tangible interface for computer-supported cooperative play," *Proc. of CHI 99*, pp. 394 – 401, 1999.
2. M. Otsuki, *et al.*: "RealSound Interaction: A novel interaction method with mixed reality space by localizing sound events in real world," *Proc. of HCII 2007*.
3. Y. Denda, *et al.*: "A study of weighted CSP analysis with average speech spectrum for noise robust talker localization," *Proc. of EUROSPEECH 2005*, pp.165 – 168, 2005.