SICE Annual Conference 2008
August 20-22, 2008, The University Electro-Communications, Japan

Novel Interaction Methods with Mixed Reality Space

Mai Otsuki[1], Asako Kimura[2], Fumihisa Shibata[1], and Hideyuki Tamura[1]

[1] Graduate School of Science and Engineering, Ritsumeikan University, Shiga, Japan
[2] PRESTO, Japan Science and Technology Agency, Saitama, Japan

Abstract: Mixed reality (MR), which merges the real and virtual worlds in real time, is a form of further extension of the conventional virtual reality (VR). There have been many studies in the field of MR; however, very few studies discuss interaction methods. In this study, we propose two novel methods for interaction with an MR space; RealSound Interaction and ToolDevice.

Keywords: Mixed reality, interaction method, input device, sound input, tool device, device set, metaphor of everyday tools.

1. INTRODUCTION

Mixed reality (MR), which merges the real and virtual worlds in real time, is an attractive extension of conventional virtual reality (VR). Interaction with an MR space is a fruitful field, since users can see their hands and interactive devices directly, and hear real sound events in cooperate with virtual multimedia data.

There have been many studies in the field of MR; however, magnetic sensors and/or visual markers are/is used in most of their interaction methods. In addition, to make their interaction intuitive, the producers of each MR attraction should have been designed their interaction method by themselves.

In this study, we propose two novel and intuitive methods for interaction with an MR space; RealSound Interaction and ToolDevice.

RealSound Interaction is an interaction method using the direction and location of “sound events” as input or interaction devices into the MR space. Since the users do not have to use any new and special devices, they can realize an intuitive operation with their familiar sound sources.

ToolDevice is a set of interaction devices using a metaphor of existing tools which are familiar in everyday life. In this method, since the user already knows how to use the original tools and had experience of using them, he/she can manipulate an MR space intuitively with the device.

In this paper, we describe the concepts of these two interaction methods and their prototype systems, and introduce some applications using these methods.

2. REALSOUND INTERACTION

Our approach “RealSound Interaction” is based on the idea that the sound events occurring 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.

We focus on detecting the direction and location of a variety of sound events and using them for 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. In this system, it is possible to use a variety of sound sources, such as handclapping, playing a castanet, or pushing a buzzer.

2.1 Related work

Several techniques using sounds for pointing interaction have been proposed. Migratory cursor [1] operates a cursor by nonverbal vocalizations as well as voice commands. This technique uses voice sounds to control a cursor position, not to indicate a location directly. BLUI [2] and Kirifuki [3] localize a user’s blowing sound on a display and use the information to support GUI operations. The target of these systems is a 2D screen, and the extension to 3D interaction is not easy. In contrast, Fraser et al. [4] used ultrasonic sounds to indicate 3D position of spatial input devices. Ultrasonic sounds allow high-accuracy detection of 3D position. However, since they are inaudible, it is difficult for the user to notice the existence of the sounds.

In RealSound Interaction technique, sound events including handclapping and musical instruments are used as inputs to indicate a direction and location. Since the users do not have to use any new and special devices, they can realize an intuitive operation with their familiar sound sources. In addition, it can leverage audible sounds themselves as the sound feedback so that users could confirm whether the input was executed or not and what kind of input they made. We believe that this allows user-friendly and easy-handled interface. We adopt microphone arrays, more than one microphone apposed, to detect the direction and location of sound events, which are appreciated in the field of acoustics [5][6].

The preceding work “PingPongPlus” by Ishii et al. [7] 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 as the sound events.

On the other hand, our approach is aimed 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.

2.2 Interaction methods using sound event

(1a) *Menu Selection* (using direction of sound source)

Menu items (CG) are displayed around a user, even though he/she is moving. The user could select one item by generating a sound signal toward it.

(1b) *Direction indication*

Computer-generated virtual objects could appear, disappear, move, or start animation in/toward the direction of the sound event generated by a user.

(2a) *Menu Selection* (using location of sound source)

Some graphical menu items are displayed at several fixed points of MR space. A user could select one item by generating a sound signal at the location of the item.

(2b) *2D/3D Pointing*

Computer-generated virtual objects could appear, disappear, move, or start animation at/toward the 2D/3D position of the sound event generated by the user.

2.3 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 (Fig. 1) [8].

2.3.1 Wearable microphone array

A fixed type of linear microphone arrays (Fig. 2 (a)) 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) (Fig.

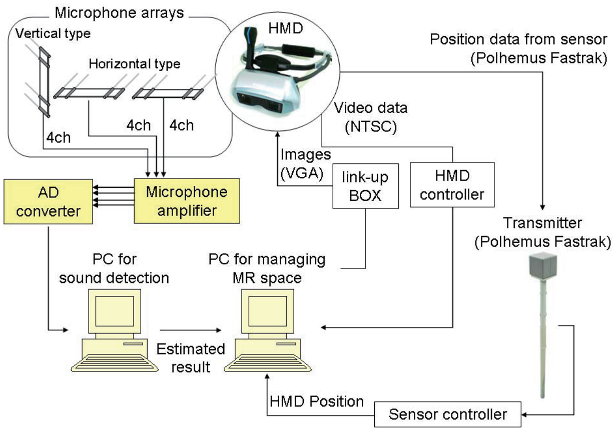


Fig. 1 System configuration

(a) Fixed type (b) Wearable type

Fig. 2 Microphone arrays

2 (b)). Since the user moves freely and faces to his/her target, the array can capture the sound constantly in his/her front direction.

2.3.2 Direction and location estimation of sound events

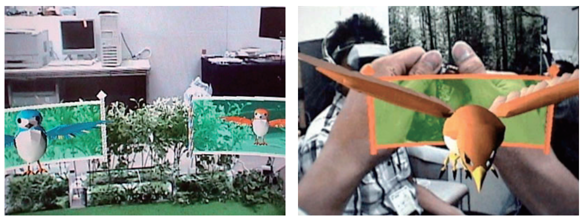
A sound source direction can be estimated by one micro-phone array. CSP (Cross-power Spectrum Phase analysis) method is used for sound source direction estimation algorithm [9]. This method gives a direction of sound source in the horizontal plane.

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

2.4 Application

"Watch the Birdie!" 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 by using



(a) A menu item could be selected by using the sound direction. In this figure, a birdcall is used as the sound device.

(b) When the user generates a sound with the birdcall, the virtual bird selected in Figure 5 (a) flies from the direction to him/her.

(c) The ducklings (CG) gather toward the position of the real mother toy duck where the squawking sound is played.

Fig. 3 Watch the Birdie!

RealSound Interaction. Firstly, along the method (1a), the user selects a virtual bird which he/she wants to call using a real birdcall (Fig. 3 (a)), then, along the method (1b), the birds fly from the direction of another birdcall (Fig. 3 (b)). Secondly, using the method (2b), mother duck (a real toy with a speaker) squawking and then many virtual ducklings run toward their mother (Fig. 3 (c)). The users also can gather them by a handclap.

2.5 User feedback

We have demonstrated the application at the 11th VRSJ Annual Conference (Sept. 2006, in Japan), Interaction 2007 (Mar. 2007, in Japan) and UIST 2007 (Oct. 2007, in USA) to collect users' feedbacks of our preliminary system. Their comments and findings are summarized as follows:

- Operations without any special hardware in hands were fairly favorable for most audiences.
- Interactions using handclap, castanet and buzzer were evaluated that the sounds themselves and the tactile impression for making sounds could be good feedback for the input confirmation.
- Although our interaction method could not allow high-accuracy detection of sound position, most audiences enjoyed a novel modality rather than required the improvement of accuracy.
- Many audiences were interested in using separate sound events, such as different sound sources, for different interactions.
- Without any instruction, most audiences made interaction with the virtual objects in the area within the scope of HMD where high-accuracy detection of microphone array is promised.
- Although the conference rooms were quite noisy, the direction and location estimation was more accurate than we had expected with only some sound baffles.

3. TOOLDEVICE

"ToolDevice" is a set of devices for various operations in an MR space. ToolDevice uses a metaphor of existing tools which are familiar in everyday life. Such tools have good affordance themselves, and at the same time, every user already has the mental model for their operations. We believe that such advantages not only lead users to the correct operation, but provide an intuitive operation.

3.1 Related work

There have been some studies of interaction devices using a metaphor of existing tools.

Ikeda et al. developed devices for transferring invisible multimedia data between computers and real equipments based on existing tools [10]. For example, SyringeDevice enables the users to get music data by pointing its head on a computer display and sucking the music titles up, and play them by pushing out in the cup on a speaker.

Ryokai et al. introduced I/O Brush [11] as a new drawing tool. It looks like a common physical paintbrush,; however, the users cannot only draw on a large display but can use the images captured from anything around them with a built-in camera as its ink.

Kitamura et al. proposed a technique for virtual object manipulation in VR space with chopsticks [12].

In the field of MR, Kato et al. developed table-top augmented reality (AR) systems for virtual interior design and city planning [13][14]. In these systems, the users are able to use a real paddle or a cup to move virtual furniture and buildings.

These studies also realize intuitive manipulations by using a metaphor of existing tools. Expanding on these devices, we aim to develop a set of devices which adapt various operations in an MR space and make the operations simple to use. In our approach, the user can select the most suitable device for the required manipulation from the set.

3.2 Extracted three integral operations

In this study, we firstly listed up the works requiring a wide MR space and that are difficult to operate with a mouse on a 2D display (Fig. 4), such as designing, layout and 3D modeling. Secondly we extracted necessary operations to complete these works "picking and moving," "modeling" and "drawing," and decided to develop ToolDevice for these three concepts as below (Fig. 5).

(1) *Picking and moving*

The users can pick and move 3D virtual objects in a similar way to the real world, for example using a hand or tweezers.

(2) *Modeling*

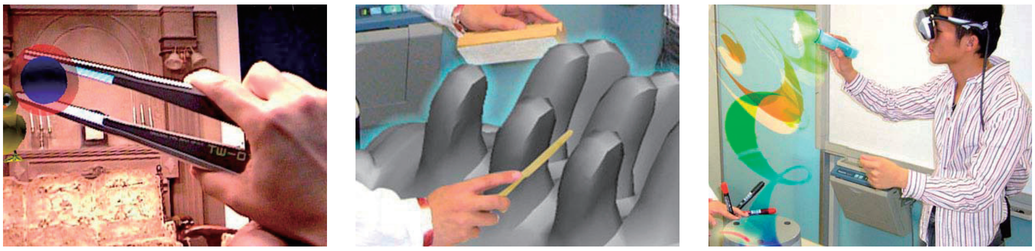
The users can make 3D virtual models in a similar way to that using a spatula or knife in the real world.

(3) *Drawing*

The users can make a drawing on any objects or any

(a) Grab (b) Cut (c) Draw

Fig. 4 Three examples difficult to operate with a mouse.



(a) Picking and moving (b) Modeling (c) Drawing

Fig. 5 Three extracted operations

space as a canvas, whether it is virtual or real, in a similar manner that drawn on a canvas using a brush or pen in the real world.

3.3 System overview

3.3.1 TweezersDevice

The first ToolDevice is for picking and moving operations. We associated these operations with the manipulation of tweezers which is a tool for picking and moving an object in the real world. ToolDevice imitates a shape of an existing tool, and its tactile and audio sensation help the manipulation to become more intuitive. The following electronic components were built into our TweezersDevice (Fig. 6):

- A magnetic sensor (POLHEMUS Fastrak) to detect the position and orientation of TweezersDevice.
- A potentiometer to detect the opening angle of TweezersDevice.
- Color LEDs, a speaker and a vibration motor for the confirmation of a manipulation.
- Force feedback mechanism to present reaction force while picking a virtual object.

As for the force feedback mechanism, we implemented two types of brake; a ratchet brake and a drum brake. The ratchet brake locks the cylinder by a pawl. It can fix the opening width of TweezersDevice depending on a size of the virtual object. The drum brake locks the cylinder by a belt. The additional function with this brake is that it can present softness which depends on a material of virtual object by loosening its belt.

3.3.2 System configuration

We developed the MR system with which the users are able to pick, move, and release virtual objects in MR space using TweezersDevice (Fig. 7).

Based on the outputs from the magnetic sensor and potentiometer, the system discriminates whether the device is able to pick the virtual objects and whether it is picking them.

When the users pick the virtual objects with TweezersDevice, the color LEDs, the speaker, and the vibration motor are switched on. Additionally, the force feedback is provided, whose magnitude is calculated in accordance with the size and material of the virtual object and the opening angle of TweezersDevice.

We use a Panasonic Let'sNote R5 for the control of TweezersDevice and a Canon MR Platform system for managing and displaying the MR space. Users watch the MR space through a HMD (Canon VH-2002). Magnetic sensors 3SPACE FASTRAK (Polhemus)

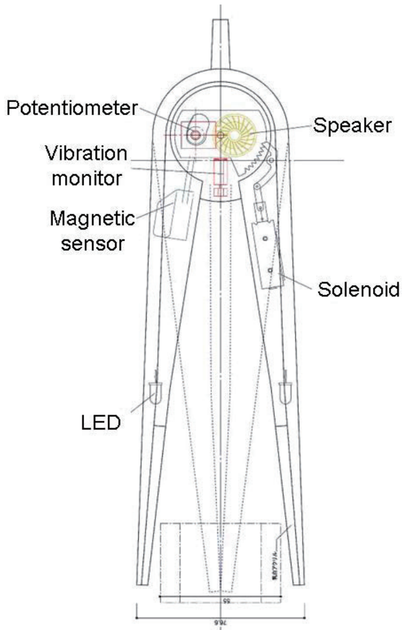


Fig. 6 Internal mechanism of TweezersDevice

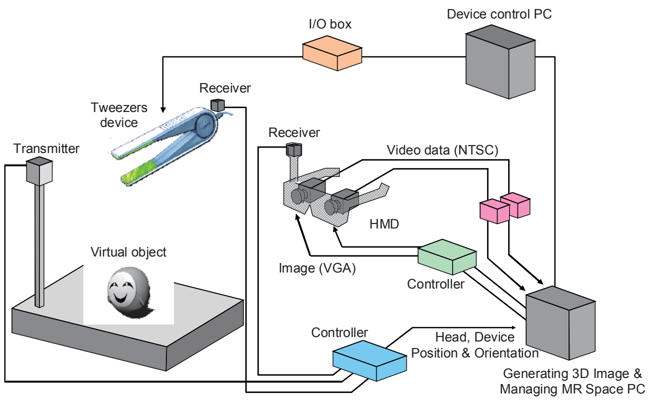


Fig.7 System configuration

detect position and orientation of the HMD and TweezersDevice.

3.4 Application

To evaluate the effectiveness of TweezersDevice, we developed applications in which users can pick, move, and release virtual objects, such as simple spheres or CG characters, staying or moving within an MR space (Fig. 8). In this manipulation, when the user picks one of the objects, the LEDs emit the same color of it, and the sound effect, vibration and reaction force are presented. The intensity of the reaction force is changed depending on the size and material of the object.

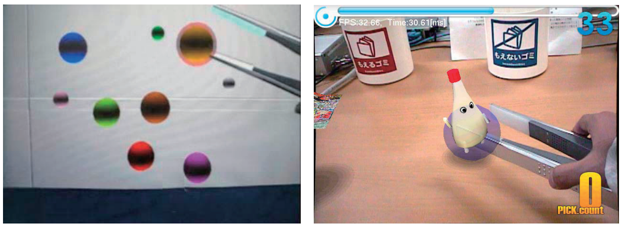


Fig. 8 TweezersDevice application; the users can pick and move the virtual objects in the MR space.

3.5 User feedback

From the experience of the pilot study in our laboratory and demonstrations at several conferences, we got some findings as follows;

- The users could understand how to use TweezersDevice just by observing other's manipulation once, and could pick and move the virtual object easily from the beginning.
- The users could not only move the virtual object in 3D space but also turn it with TweezersDevice easily.
- All of the electric components built into the device for visual, tactile and audio sensation worked well. Especially, the force feedback mechanism using the dram brake was highly evaluated, since it is similar to the feedback of real tweezers and it improves the operational feeling.
- In the device adopting the ratchet brake, the opening width is locked when the users pick the virtual objects, and they can't close the device by more than that. They commented that they felt uncomfortable because this behavior was different from that of real tweezers'.
- Similarly in the case adopting drum brake, the opening width is also locked when the users pick the virtual object. However, if the users close the device more tightly, the belt brake slips and they can get a springy feedback. The users said that this feedback is similar to that of real tweezers', and it felt like picking a real object.
- In this MR system, because computer generated images are superimposed on real scene, TweezersDevice is occluded by virtual objects. Therefore, some users couldn't recognize the anteroposterior relationship between TweezersDevice and the virtual object and became confused.

4. CONCLUSION

In this paper, we proposed our novel interaction methods with MR space; RealSound Interaction and ToolDevice. As for the future of RealSound Interaction, we are planning to include additional sound sources other than harmonic sounds and to realize simultaneous utilization of multi-users. For ToolDevice, we are also planning to evaluate TweezersDevice, and develop second and third generations.

ACKNOWLEDGMENT

This research was supported by the Japan Society for the Promotion of Science through Grants-in-aid for Scientific Research (A), "a mixed reality system that merges real and virtual worlds with three senses" and Precursory Research for Embryonic Science and Technology (PRESTO) Program " Foundation of technology supporting the creation of digital media contents" of Japan Science and Technology Agency (JST).

REFERENCES

[1] Y. Mihara *et al.*: "The migratory cursor: Accurate speech-based cursor movement by moving multiple ghost cursors using non-verbal vocalization," Proc. of ASSETS 05, pp. 76 - 83, 2005.
[2] S. N. Patel *et al.*: "BLUI: Low-cost localized blowable user interfaces," Proc. of UIST 2007, pp. 217 - 220, 2007.
[3] S. Iga *et al.*: "Kirifuki: Inhaling and exhaling interaction with visual objects," Proc. of IWEC2002, pp. 119 - 126, 2002.
[4] M. Fraser *et al.*: "Seconds Matter: Improving distributed coordination by tracking and visualizing display trajectories," Proc. of CHI 2007, pp. 1303 - 1312, 2007.
[5] T. Nagai *et al.*: "Estimation of source location based on 2-D MUSIC and its application to speech recognition in cars," IEEE Proc. of ICASSP 2001, vol. 5, pp. 3041 - 3044, 2001.
[6] K. Nakadai *et al.*: "Robust tracking of multiple sound sources by spatial integration of room and robot microphone arrays," IEEE Proc of ICASSP 2006, vol. 4, pp. 929 - 932, 2006.
[7] H. Ishii *et al.*: "PingPongPlus: Design of an athletic-tangible interface for computer-supported cooperative play," Proc. of ACM CHI 1999, pp. 394 - 401, 1999
[8] 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, pp. 653 - 662, 2007.
[9] M. Omologo *et al.*: "Acoustic event location using a crosspower-spectrum phase based technique," IEEE Proc. of ICASSP 1994, vol. 2, pp. 273 - 276, 1994.
[10]Y. Ikeda *et al.*: "TOOL DEVICE: Handy haptic feedback devices imitating every day tools," Proc. of HCII 2003, pp. 661 - 665, 2003.
[11]K. Ryokai *et al.*: "I/O Brush: Drawing with everyday objects as ink," Proc. of CHI 2004, pp. 303 - 310, 2004.
[12]Y. Kitamura *et al.*: "Virtual Chopsticks: Object manipulation using multiple exact interactions," Proc. of IEEE VR, pp. 198 - 204, 1999.
[13]H. Kato *et al.*: "Virtual object manipulation on a table-top AR environment," Proc. of ISAR 2000, pp. 111 - 119, 2000.
[14]H. Kato *et al.*: "MagicCup: A tangible interface for virtual objects manipulation in table-top augmented reality," Proc. of ART03, pp. 85 - 86, 2003.