3D Audio-Visual Display Using Mobile Devices

Kotaro Takahashi *
Graduate School of Engineering Kanazawa Institute of Technology
Tomohito Yamamoto †
College of Information Science and Human Communication Kanazawa Institute of Technology

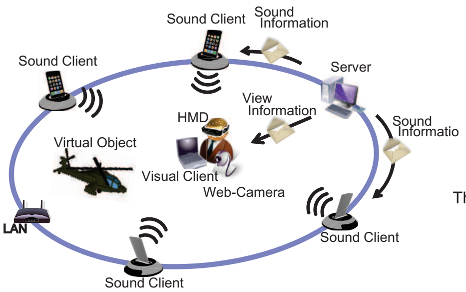

Figure 1: *System Overview*

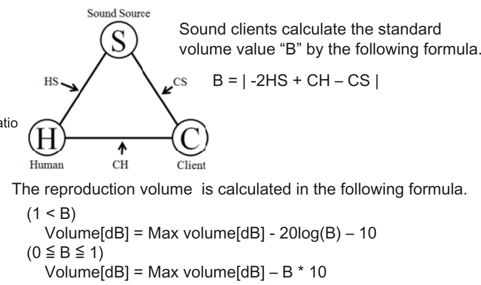

Figure 2: *Method of Sound Reproduction*

Figure 3: *Manipulation of Virtual Object*

1 Introduction

For providing high presence, many kinds of display system have been developed [Hughes et al. 2005]. Typical examples are 3-dimensional display and multi-channel surround speaker system. Moreover, 3D movie such as "Avatar" or 3D TV have been brought to the market. However, these high realistic displays for visual, sound, or both, were usually composed of fixed and expensive equipment.

On the other hand, with the development of mobile and miniaturization technology, potable devises such as a netbook or a smart phone have been widespread. These devises enabled a user to access information anytime and anyplace. If it is possible to build the high presence system by such portable devises, user can experience the high presence every time and everywhere.

In this research, using multiple mobile devises connected to a local area network, we developed the 3D audio-visual display system which could adjust the scale and constitution to user's environment.

2 Our Approach

Our system is composed of multiple terminals such as a PC or an iPhone as shown in Figure 1. Each terminal is connected to LAN and these are divided into one server, one visual client and multiple sound clients. In this system, at first, the server simulates a virtual space which is designed freely by a user. After simulation, the server transmits the information of the virtual space to terminals by UDP multicast. Then, terminals present the high presence to a user using received information.

To reproduce spacial sound, sound clients calculate distance decrement of sound using three positions (a listener's position, a sound source's position and a sound client's position in Figure 2). Using these positions, the system calculates the volume in real time and reproduces a spatial sound by the volume difference of each client [Takahashi et al. 2009].

To represent spacial view, visual client uses a head mount display and a web-camera. For presenting spacial view, the visual client detects user's posture by head tracking sensor, and reflects it to viewing angle in the virtual space. Then, the spacial mixture view which is composed of the virtual space and a real view from web-camera is displayed to a user through the HMD. Using marker tracking, a user can also operate the virtual object as shown in Figure 3. The visual client recognizes a marker in web-camera's view. Then the visual client analyzes a position of the marker and reflects it to the object's position.

Moreover, our system can represent the virtual space to multi-user simultaneously using the global sever which is placed on the internet. To present the virtual space to multi-user, the global server simulates the virtual space, and sends it to servers using TCP. Servers adjust received information to own client's constitution and transmit it to own clients.

This system is basically implemented as a software level , and allows any device to become the terminal if the device can be connected to network. Therefore, using devices such as desktop, laptop, or smart phone which have been widespread, our system can be built relatively inexpensive.

3 Future works

Present system uses earth magnetism sensor for mixture of virtual and real world. In future work, for improving consistency between virtual and real world, the method called monoSLAM will be implemented. Also, present aural display represents surround using only the volume difference of each speaker. For higher expression, additional sound elements such as reverberant and reflected sound should be considered.

References

HUGHES, C. E., STAPLETON, C. B., HUGHES, D. E., AND SMITH, E. M. 2005. Mixed reality in education, entertainment, and training. *IEEE Computer Graphics and Applications 25*, 6, 24 – 30.

TAKAHASHI, K., IKEDA, S., AND YAMAMOTO, T. 2009. Light aural display using network connected multiple computers. In *HCI International 2009 - Posters*, Springer, 401–405.

*e-mail: k-takahashi@eagle-net.ne.jp
†e-mail:tyama@neptune.kanazawa-it.ac.jp

Copyright is held by the author / owner(s).
SIGGRAPH 2010, Los Angeles, California, July 25 – 29, 2010.
ISBN 978-1-4503-0210-4/10/0007