Directional Communication using Spatial Sound in Human-Telepresence

Shohei Nagai* **Shunichi Kasahara†** **Jun Rekimoto‡**
*†‡The University of Tokyo †‡Sony Computer Science Laboratory
*shohei.nagai14@gmail.com †kasahara@csl.sony.co.jp ‡rekimoto@acm.org

ABSTRACT
Communication is essential for working effectively with others. We communicate with each other to share their situation and what they are thinking. Especially, using voice is one of the most common ways to communicate. In previous research, we proposed LiveSphere that shares the surrounding environment with a remote person and provides immersive experience to effectively collaborate with each other. This system realizes "human-telepresence" where a person can be in other person and experience the environment. However, the communication in human-telepresence has some problems. In this paper, we propose directional communication with spatial sound to alleviate the problems. We also report on the result of user study.

Author Keywords
Spatial sound; Remote collaboration; Omnidirectional image; First person view streaming; Wearable computer; Human-telepresence

ACM Classification Keywords
H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous

INTRODUCTION
Communication is essential for working effectively with others. We communicate, both verbally and non-verbally, to understand each other's situation and what they are thinking. In remote collaboration, sharing sound and visual information is one of the essential factors for realizing efficient collaboration[3].

We proposed LiveShere[4] to collaborate with a remote person (Figure 1). This system records the environment surrounding a person in the actual situation (referred as to the **Body**) in the form of omnidirectional images and shares them with a remote person (referred as to the **Ghost**). The images are decoupled with ego-motion, so Body and Ghost can observe the shared environment independently. This approaches realizes "human-telepresence" where a person can virtually be in other person and observe the environment as if present.

AH'15, March 09 - 11, 2015, Singapore, Singapore.

http://dx.doi.org/10.1145/2735711.2735818

In human-telepresence, Body does not communicate face to face with Ghost but communicates with Ghost who is virtually present in themselves and independently observe the surrounding environment from the same point of view of Body. In the situation, there are some problems with conventional voice communication.

In this paper, we propose communication using spatial sound in human-telepresence. First, we discuss the existing problems in voice communication in conventional LiveSphere system. Then, we describe the prototype of communication system using spatial sound. We also report on the user study for evaluating the effectiveness of spatial sound and discuss the insights derived from the experiment.

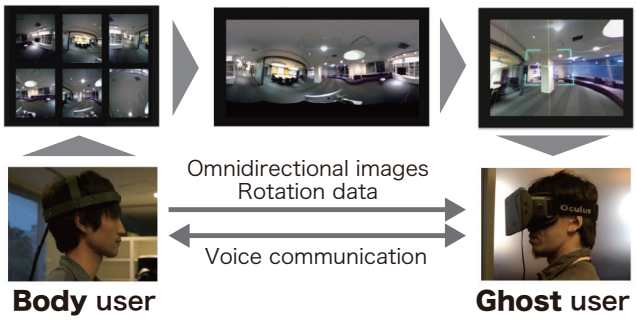

Figure 1. LiveSphere overview: Body user wears the head-worn device and share the environment using omnidirectional images. The capture images are stabilized and transmitted. Ghost user observes the shared environment independently from Body's ego-motion.

COMMUNICATION PROBLEMS
In a previous study, we could found Body and Ghost independently observe and comprehend the environment using LiveSphere. On the other hand, we also observed some problems as shown below.

Difficulty in understanding Ghost's situation
In conventional LiveSphere system, Body communicates with Ghost just with ordinary voice chat, so it is difficult to understand what direction Ghost looks toward and what s/he focuses on. This prevents the understanding of the partner's situation. Actually, in the interview after the task, some participants claim about this issues and even show the fear like "*He (Ghost) is watching everything I'm doing.*"

Difficulty in giving directional guidance
In a collaborative task, they also had difficulties in giving directional guidance. In the task, some participants as Ghost

uses directional expressions like "Turn left." However, such kind of expressions is difficult to pinpoint a specific place and direction. It is observed that Body passed through the direction Ghost expects Body to turn toward. In a local communication, to indicate the specific direction, they can use gesture like pointing to a specific point, or move and speak to from the place on the direction like "Come here." On the other hand, both cannot be used in the conventional system.

COMMUNICATION WITH SPATIAL SOUND

Human has an ability to localize a sound source using their auditory system[1]. Spatial sound uses this human characteristic to allow us to virtually perceive sound from a specific place. Using spatial sound, we can add intuitive directional information to voice communication. This enables Body to understand which direction Ghost is looking toward when talking, which alleviates the above-mentioned problems.

IMPLEMENTATION

We use spatial sound in both Body and Ghost sides. We synthesize spatial sound using openAL. Here, spatial sound is defined with two elements: sound source and its position. We capture Body voice as sound source for Ghost using headset, and vice versa. On Body side, the position of the sound source is defined on the point of a sphere on Ghost's relative head direction from Body's head direction R_{gb}, and vice versa on R_{bg} (Figure 2). These are calculated from Body's head rotation and Ghost's head rotation from the basic direction, both of which are estimated in LiveSphere. At first, R_{bg} is calculated from the head rotation data transmitted from Body and the captured Ghost's head rotation on Ghost side. Then, spatial sound is immediately synthesized by OpenAL using the calculated position of sound and the sound transmitted from Body through the Internet, and provided to Ghost using headset. On Body side, the calculated sound position and the captured sound are sent from Ghost. Then, spatial sound is synthesized and provided to Body.

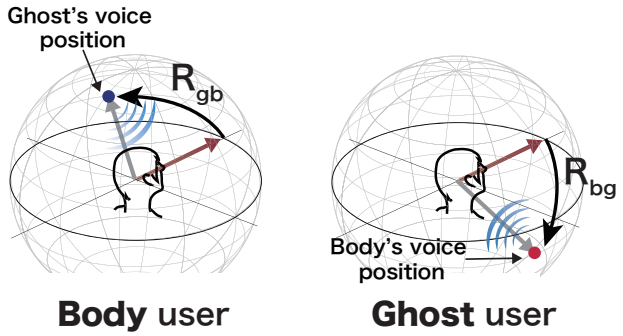

Figure 2. Communication with spatial sound: The position of Ghosts voice on Body side is defined on the point of a sphere on Ghosts relative direction from Bodys head direction and vice versa on Ghost side.

USER STUDY

We made user study for gathering feedback and evaluating how communication using spatial sound alleviating the above-mentioned problems. Six persons participated in the study. In the study, Body participant is in a laboratory and Ghost participant is in a different place. Here, Ghost is told to give guidance for Body to pick up specific objects, whose position was taught to Ghost in advance. All participants play roles of both Body and Ghost. After the task, we interviewed them and recorded their feedback.

RESULT AND DISCUSSION

The result showed Body and Ghost could effectively communicate using spatial sound system. Participants reported that spatial sound was useful to understand the situation of the partners. Some of them comments that they felt "*He was in the same place.*" On the other hand, some participants as Body claimed that it was difficult to estimate which direction the Ghost look toward while s/he said nothing. About this issue, one participant comments "*Noise (included in the transmitted sound data) was a cue to estimate the partner's direction when he said nothing.*" The noise was also spatialized. This implies that, other than human voice, continuous sound directed to the partner's direction is useful to estimate the partner's situation.

One participant reported that *"It would be useful to put a sound marker to indicate a specific direction and object."* Human has an ability to differentiate and provide attention to given sound in the presence of multiple sound sources[2]. This suggests the possibility that Body effectively use such sound marker to understand the partner's indication. Moreover, this can also realize that multiple Ghost users simultaneously give guidance using spatial sound. Then, Body differentiates them and selectively understands what they are saying.

CONCLUSION

In this paper, we propose a novel communication with spatial sound according to the relative head direction in human-telepresence. We developed the prototype of this method and made user study for evaluating this method. The result showed using spatial sound improved the understanding of the partners' situation. Conversely, the study also implied some assignments for the method. We discussed about that and proposed design implications. For future work, we will implement them and make more detailed evaluation.

REFERENCES

1. Blauert, J. *Spatial hearing: the psychophysics of human sound localization.* MIT press, 1997.
2. Broadbent, D. E. *Perception and communication.* Elsevier, 2013.
3. Fussell, S. R., Setlock, L. D., and Kraut, R. E. Effects of head-mounted and scene-oriented video systems on remote collaboration on physical tasks. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems,* CHI '03, ACM (New York, NY, USA, 2003), 513-520.
4. Nagai, S., Kasahara, S., and Rekimoto, J. Livesphere: Sharing the surrounding visual environment for immersive experience in remote collaboration. In *Proceedings of the Ninth International Conference on Tangible, Embedded, and Embodied Interaction,* TEI 15, ACM (New York, NY, USA, 2015), 113-116.