Sonoliards: Rendering Audible Sound Spots by Reflecting the Ultrasound Beams

Naoya Muramatsu* sh.mn.nat@gmail.com
Kazuki Ohshima* s1613108@u.tsukuba.ac.jp
Ryota Kawamura * s1721663@s.tsukuba.ac.jp
Ooi Chun Wei* made3102@gmail.com
Yuta Sato* s.yutassie@gmail.com
Yoichi Ochiai * wizard@slis.tsukuba.ac.jp
***University of Tsukuba (1-1-1 Tennodai, Tsukuba, Ibaraki, Japan)**

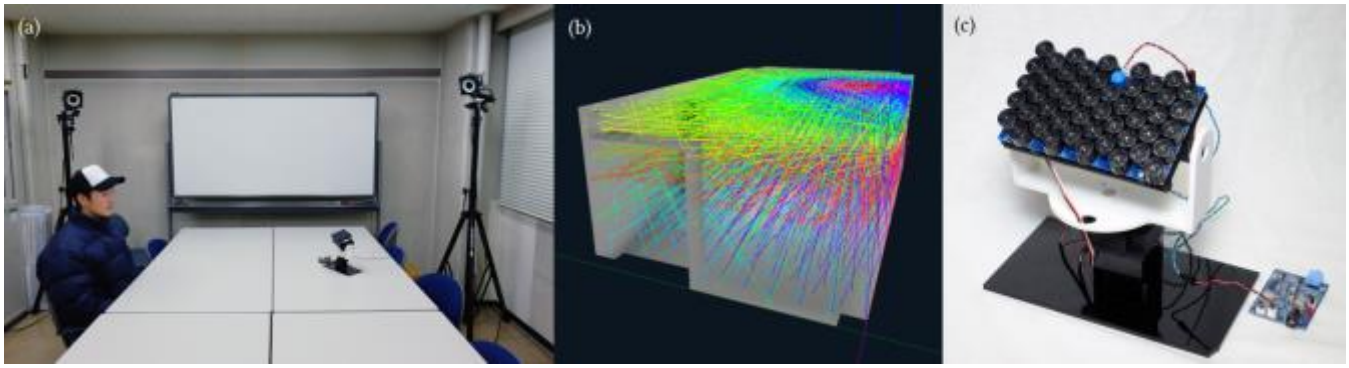

Figure 1: (a) Our system set up. (b) Rendering of simulation result. (c) Parametric speaker with two-axis servo motors.

AUTHOR KEYWORDS
Ultrasound beams; Parametric Speaker; Acoustic Field Generation; Acoustic Simulation; Ray Tracing Method.

ACM Classification Keywords
H.5.1. Information interfaces and presentation (e.g., HCI): Multimedia Information Systems.

INTRODUCTION

We propose a real-time indoor dynamic acoustic field generation system of a spot audio towards a particular person.

Personalized sound propagation is becoming important in the ubiquitous computing context. Many audio interaction systems that are appearing in mobile interactions and personal recognition by audio recognition techniques are nowadays increasing its accuracy. In these fields, continuous sound propagation for individuals is highly desired. However, sound waves have low directivity thus widely received by unintentional hearers. To solve this problem, continuous efforts have been made, such as audible-sound directions loudspeakers and ultrasound-based super directional loudspeakers.

Studies using sound propagation are found in various research fields. Tung et al. estimates the force applied to the smartphone via structure-borne sound propagation using its built-in sensor instead of specialized sensors used in conventional methods [1]. This enables the application of expressive interfaces. Also, Ochiai et al. proposed a method of spatial audio rendering using ultrasound and realized a higher audience selectivity [2]. Conventional 3D sound technology based on head related transfer function (HRTF) could not reproduce realistic sounds in the virtual space, but the problem can be solved by sound rendering. The sound rendering pipeline shown in the study consists of sound synthesis, sound propagation, and sound generation. This makes it possible to reproduce realistic sounds such as sounds occurred when people drop something. Nowadays various types of ultrasound-based super directional loudspeakers are available on the market and is used in many places such as museum, platforms, galleries, and these places need to avoid sound contaminations.

Hence, spot audio techniques have been explored by generating the ultrasound beams toward the target person in a certain area. However, everyone in the area can hear the sound. To solve this problem, our system generates an acoustic field dynamically which can only be heard by the target person using motion capture system and parametric-speaker with two-axis servo motors. Furthermore, with the proposed system, we accomplished real-time tracking and

UIST'17 Adjunct, October 22–25, 2017, Québec City, QC, Canada.

ACM ISBN 978-1-4503-5419-6/17/10.
DOI: http://dx.doi.org/10.1145/3131785.3131807

sound rendering up to 15 fps, allowing practical applications which rely heavily on continuous sound propagation to be deployed.

SYSTEM OVERVIEW

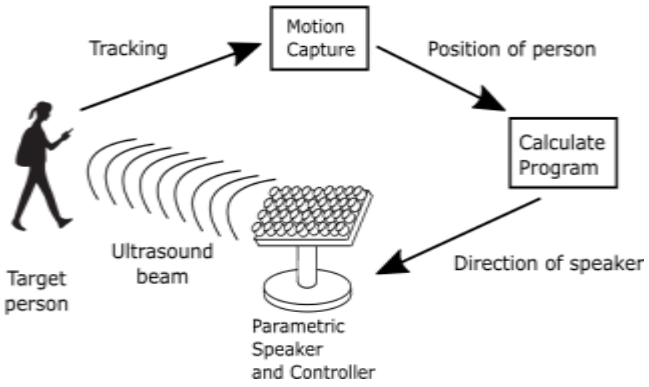

Figure 2. System overview of Sonoliards.

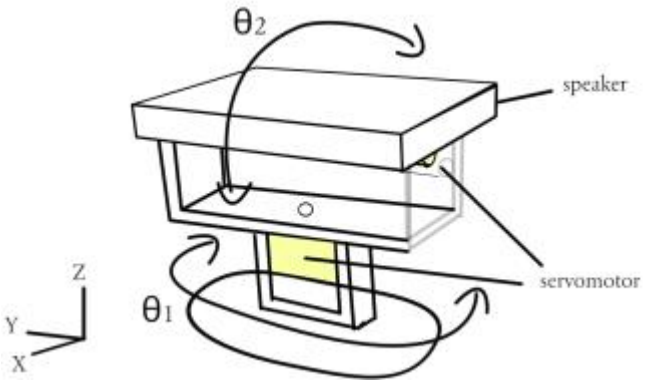

Figure 3. Structure of two axis servo direction controller we used.

The purpose of this system is to create a highly directional sound reproduced with a parametric-speaker and to generate an acoustic field to reach only a specific person in real time by using the reflection indoors. In order to realize this, it is crucial to acquire the position of the person relative to the loudspeaker in the room then calculation is done to determine the direction of the loudspeaker. This system consists of three parts: motion capture, two-axis loudspeakers with servo control, and algorithm for the calculation of the loudspeaker direction (see Figure 1(a) and Figure 2).

Motion capture is used to obtain the position of the head of a person indoors. This system uses six motion capture cameras. In order to acquire the position of the head, the user of this system needs to attach a marker of motion capture on his head via a hat or the like.

The direction of the loudspeaker is controlled by two axis servo motors, and its movable range is the upper half of the sphere (see Figure 1(c) and Figure 3). It moves according to the angle sent from the program.

The program uses the ray tracing method to calculate the direction of the loudspeaker so that the sound can only be heard by the target person. The algorithm is based on the assumption that the sound as a form of light particles in the calculation of ray tracing method, the number of sound particles is determined by assuming at least one sound particles is present in the volume of a person head regardless of his position in the simulated space (Figure 1(b)) after the sound particles are being emitted. In simulation, the propagation of the emitted sound beam to the surroundings from the source is calculated, and the collision of sound beam on the target person head is detected. At this time, we determine the direction of the loudspeaker to the initial direction of the sound beam which has the least reflection number and the shortest traveling distance to the target or the least likely to collide with people other than the target person. Real-time tracking and rendering of acoustic field is realized by continuously repeating the initialization of sound beam, target head position tracking, determination of the direction of sound beam in simulation, and shifts in direction of the loudspeaker accordingly.

When we run the program using an indoor 3D model where the number of wall surfaces is about 20, the operation speed of this system is about 15 fps. In the experiments, we allowed the subjects to freely roam through the room and got the same impression that most of them felt as if the sound followed them in a real time.

APPLICATION

The main feature of this system is that you can use the reflection to avoid obstacles and deliver the sound with high accuracy, so you can play the sounds that only reach specific people.

For example, this can be applied to raise individual attention during the examination. Also, it is possible to generate an acoustic field reaching a position where sound cannot be delivered by a conventional speaker system, such as a specific position on the opposite side of the wall. This eliminates the dependency relation between the space design and the speaker position, and enables interior space design with higher degree of freedom. Such a high degree of freedom in space design is thought to be very useful for the design of amusement facilities and so on.

Also, by combining this system with multiple speakers, it is possible to deliver different sounds to different people at the same time. We expect to realize the system that implements different guidance and warnings for each person.

It is not implemented yet now, but when you deliver the sound to the target person, it can be used for personalized guidance or even in horror amusement facilities where it can designate the direction of where you deliver the sound.

REFERENCES

1. Tung, Y.-C., and Shin, K. G. Expansion of human-phone interface by sensing structure-borne sound propagation. In *Proceedings of the 14th Annual International Conference on Mobile Systems*, Applications, and Services, MobiSys '16, ACM (New York, NY, USA, 2016), 277–289.

2. Ochiai, Y., Hoshi, T., and Suzuki, I. Holographic whisper: Rendering audible sound spots in three-dimensional space by focusing ultrasonic waves. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI '17, ACM (New York, NY, USA, 2017), 4314–4325.