Proceedings of Meetings on Acoustics

Volume 19, 2013 http://acousticalsociety.org/

ICA 2013 Montreal
Montreal, Canada
2 - 7 June 2013

Architectural Acoustics
Session 3aAAa: Virtual Concert Hall Acoustics I

3aAAa6. Sound Cask: Music and voice communications system with three-dimensional sound reproduction based on boundary surface control principle.

Yusuke Ikeda* and Shiro Ise

***Corresponding author's address: Faculty of Engineering, Kyoto University, Kyoto-shi, 615-8530, Kyoto-fu, Japan, ae-yusuke-ikeda@archi.kyoto-u.ac.jp**

To reproduce a highly realistic sound field reproduction, we have developed a three dimensional sound reproduction system based on the boundary surface control principle. We set up an Internet connection among multiple systems, which enables distant speakers to carry out voice telecommunication by simultaneously hearing the same sound field and perceiving the other speakers' simulated positions as if they were in the same location. In this reproduction system, a listener can freely move his head and rotate his body because the system reproduces the sound not only at the point but at the area. In this paper, we introduce a 'Sound Cask', which is a 96-channel sound reproduction system based on the boundary surface control principle. The 'Sound Cask' is used to realize such musical telecommunication as if musical performers were all playing in the same hall. The system provides space large enough to play a small musical instrument such as violin, which allows listening and communication accompanied with natural body movements by the performer. Another specific feature of the system is that loudspeakers are installed aiming every possible direction excluding the floor. The system design is suitable for relocation as it can be assembled or disassembled when needed.

Published by the Acoustical Society of America through the American Institute of Physics

Received 30 Jan 2013; published 2 Jun 2013

INTRODUCTION

The use of voice telecommunications has been steadily increasing since the invention of the telephone. Furthermore, the advent of mobile phones prompted personalization of the telephone, making voice telecommunication an essential part of our lives. However, the practicality of mobile phones, which allow us to telecommunicate anywhere and anytime, does not diminish the importance of actual face-to-face communication. Because the telephone is designed with limitations on the amount of transmittable data and focuses on the verbal feature of voice, telephone communication lacks various aspects of spatial information that is found in real space. Communication involves varied modalities, and focusing on voice only, as in the case of the telephone, does not enable communication with the same perception as when conversation partners are in the same room. The limitations of the telephone can be explained through the properties of a conventional stereophonic audio system which cannot reproduce adequately the spatial information of a sound field.

We proposed and developed a sound field sharing system based on the boundary surface control (BoSC) principle(Ise, 2007). With a multi-channel microphone array and speaker system, BoSC enables us to record and reproduce a sound field in a given area. A conventional conventional 62-channel sound reproduction system reproduces a sound field sufficient to capture a conversation partner's head area and allows one to perceive the speaker's position with high accuracy(Enomoto *et al.*, 2008). We used two sets of such sound reproduction systems and developed a sound field sharing system, which enables one to perceive a conversation as if the speakers are in the same room, by convolving voices in real time with the inverse system and playing background sounds of environments such as a forest or a music hall(Ikeda *et al.*, 2010).

In the conventional 62-channel sound reproduction system, speakers and listeners can make body movements such as rotating their heads and seats. In this paper, we introduce a "Sound Cask" communication system based on the BoSC principle; this system suits the needs of voice and music communication accompanied by body movements.

SOUND FIELD REPRODUCTION BASED ON THE BOUNDARY SURFACE CONTROL PRINCIPLE

In 1993, Ise proposed a 3-D sound reproduction system based on the boundary surface control principle that uses the Kirchhoff-Helmholtz integral equation and the inverse system(Ise, 1993). Figure 1 shows its basic concept.

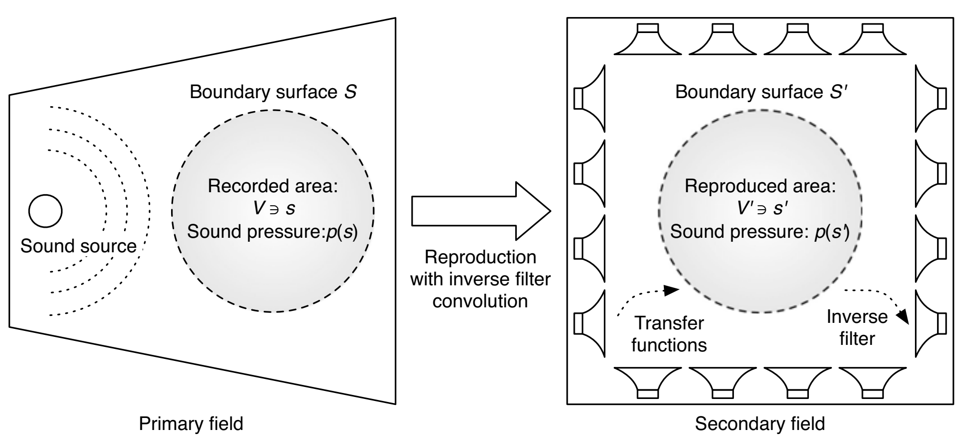

FIGURE 1: Concept of sound field reproduction based on the boundary surface control principle.

We are considering the reproduction of a sound field within a recorded area V in the primary field into a reproduced area V' in the secondary field. Given that V is congruent with V', the following equation holds:

$$|r' - s'| = |r - s| \quad (s \in V, r \in S, s' \in V', r \in S') \tag{1}$$

where S and S' denote a boundary of the recorded area and a boundary of the reproduced area respectively. If we denote the sound pressure in V and V' as $p(s)$ and $p(s')$ respectively, $p(s)$ and $p(s')$ are denoted by the following equations.

$$p(s) = \int\int_S G(r|s)\frac{\partial p(r)}{\partial n} - p(r)\frac{\partial G(r|s)}{\partial n} dS,\ (s \in V) \tag{2}$$

$$p(s') = \int\int_{S'} G(r'|s')\frac{\partial p(r')}{\partial n} - p(r')\frac{\partial G(r'|s')}{\partial n} dS',\ (s' \in V') \tag{3}$$

where n and n' denote normal vectors on S and S' respectively. By applying the Equation 1, we obtain the following relationships of Green's function and its gradient:

$$G(r|s) = G(r'|s') \tag{4}$$

$$\frac{\partial G(r|s)}{\partial n} = \frac{\partial G(r'|s')}{\partial n'} \tag{5}$$

Hence, it follows that if the sound pressure and its gradient on each boundary are equal to each other, then the sound pressures in each area are also equal to each other from Equations 2 and 3. This is expressed as

$$\begin{aligned} &\forall r \in S, \forall r' \in S', \\ &p(r) = p(r'), \frac{\partial p(r)}{\partial n} = \frac{\partial p(r')}{\partial n'} \\ &\Longrightarrow \forall s \in V, \forall s' \in V', \quad p(s) = p(s'). \end{aligned} \tag{6}$$

If we consider this as a boundary value problem, from the uniqueness of the solution it follows that either the sound pressure value or its gradient value are sufficient to determine the value of both(Kleinman and Roach, 1974).

We measure sound pressure by placing microphones onto control points that are a discretized boundary surface in the primary field. Sound pressure in the secondary field has to be checked at control points that are congruent with the ones in the primary field. The loudspeakers, which serve as a secondary sound source, are separated from the control points by being placed onto their outer territory. We measure the transfer function directed from the loudspeakers towards the control points, and we construct the inverse system. The inverse system is convolved with the sound pressure recorded in the primary source, and when played by loudspeakers, the sound pressure reproduced at control points in the secondary field is equivalent to that in the primary field. The use of the inverse system has several advantages including the possibility of making corrections to loudspeaker discrepancies and also making corrections to sound reflection inherent in the secondary field and to acoustic reflection taking place in the secondary field.

SOUND FIELD SHARING SYSTEM

To share auditory space with a distant communication partner as if that partner were in the same room, we developed a sound field sharing system that uses more than two sound field reproduction systems based on the BoSC principle. Figure 2 shows the basic concept of the sound field sharing system.

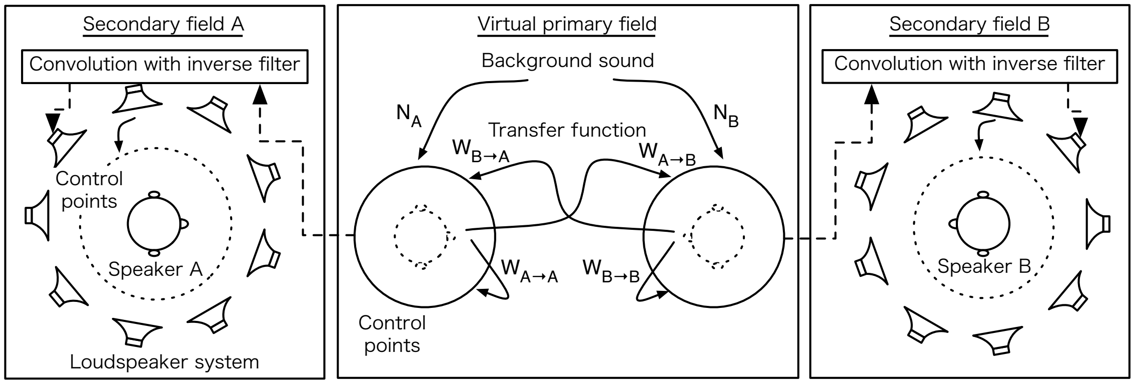

FIGURE 2: Concept of the sound field sharing system.

The sound field sharing system enables speakers who are away form each other (speaker A and B) to communicate as if they were in the same space (virtual primary field). The voice communication and musical performance are recorded in each secondary field and played inside the control points in the virtual primary field. The sound made by each speaker in the secondary field is reflected by the walls or a floor in the virtual primary field and propagated to both the other speaker ($W_{A\to B}$, $W_{B\to A}$) and the original speaker ($W_{A\to A}$, $W_{B\to B}$). For example, if the virtual primary field is in a concert hall, voice or musical performance is affected by reverberation of the concert hall and then propagated to the both. Furthermore, the sounds in the virtual primary field include spatial information such as position of the speakers.

We can also hear background sounds (N_A,N_B) in the virtual primary field . For example, if the virtual primary field is in a forest, we may hear animal calls and rustling leaves as background sounds. The sounds that are recorded by microphones at the control points in the virtual sound field are convolved with the inverse filter and then played by loudspeakers in the secondary fields. Because the sound field inside the control points in the secondary field are the same as those inside the control points in the virtual primary field, speakers can communicate as if they were in the virtual primary field. However, direct sounds that are made inside the control points in the virtual primary field and propagated directly to the control points must be removed from the recorded voice which is propagated back to the original speaker ($W_{A\to A}$, $W_{B\to B}$).

If there is no need to reproduce the background sound (N_A, N_B) in real time, the background sound can be reproduced after recording. For voice communication between systems, the voice is convolved with transfer functions from a loudspeaker to the other microphone array ($W_{A\to B}$, $W_{B\to A}$) and to the speaker's own microphone array ($W_{A\to A}$, $W_{B\to B}$) in the shared sound field. Then, the convolved voices are added and reproduced in the secondary field. However, in the latter transfer function, a direct sound effect must be removed.

In previous studies, we developed a sound field sharing system using the BoSC system with a conventional 62-channel loudspeaker system(Enomoto, 2009; Enomoto *et al.*, 2009). The BoSC system allows us to perceive the direction of the reproduced voice. The system transmits voice direction in a three-person conversation by changing the transfer functions in accordance with the speaker's facing angle(Ikeda *et al.*, 2010). Only 24 loudspeakers were used to reproduce voice between the systems to reduce the number of calculations(Enomoto *et al.*, 2010).

Sound cask

Sound cask feature

The conventional sound field sharing system using the conventional 62-channel loudspeaker system allows speakers to move their body by only rotating their chair. Therefore, to apply the sound field sharing system to communication with much more body movement, such as what might be in a musical performance, we introduced and developed a 96-channel sound field reproduction system based on BoSC principle, called "Sound Cask" from its shape. Figure 3 shows a picture of the sound cask. The sound cask has a regular nonagonal shaped in the

FIGURE 3: Sound cask

horizontal direction. It is also consists of three surfaces in the vertical direction and is designed to avoid parallel sides except for the floor and ceiling planes This characteristic shape has the effect of suppressing acoustic modes inside the sound cask to simplify the design of the inverse filter. With internal dimensions of 1950 mm diameter at the central horizontal plane and 2150 mm height, the sound cask is large enough for a player to play wind and stringed musical instruments inside the system as shown in Fig.3.

A total of 96 full-range loudspeakers (FOSTEX FE103En) are installed omnidirectionally except for the floor plane. Each loudspeaker driver is installed in a closed enclosure (15×15×12 cm). The loudspeakers installed on the wall surface are divided into six layers based on their height, with 9 loudspeakers allotted to each of the top and the bottom layers, and 18 loudspeakers allotted to each of the remaining layers. The average interval between layers is around 350 mm. The average interval between loudspeakers in the horizontal direction is around 540 mm at the top and bottom layers, and around 330 mm at all other layers. Six loudspeakers are installed on the ceiling plane. The loudspeakers in the sound cask are driven by digital amplifiers installed at the bottom of the sound cask.

In a conventional 62-channel BoSC system, loudspeakers are installed around the upper body of a listener. However, in the sound cask, loudspeakers are installed to cover the whole body of a listener. Additionally, the wall parts of the sound cask are modularized and can be dismantled for transporting the system. The system is divisible into nine parts in the horizontal direction, with each part forming a side of a regular nonagon. The walls of the sound cask are divisible into three (top, middle and bottom) parts in the vertical direction. To shorten the reverberation time and thereby simplify the design of the inverse system, we installed sound-absorbing materials on the wall and ceiling, but not on the floor. For sound-absorbing material, we use poriwool (with a thickness of 120 mm and a density of 32 kg/m^3) to achieve both effectiveness and comfort. The noise barrier performance of the system is Dr-20.

FIGURE 4: Microphone array for BoSC in a sound cask to record transfer functions.

Figure 4 shows the BoSC microphone array used for recording sound fields and measuring transfer functions. The BoSC microphone array is designed to have the same configuration as C_{80} fullerene. Eighty omnidirectional microphones (DPA 4060BM) are installed at the node of the fullerene. The diameter of the microphone array is about 46 cm and is large enough to enclose a listener's head. When we record background sound, we set up this BoSC microphone array at the place where we suppose that a speaker or musical performer located. For example, if we suppose two audience members listening to a concert in a sound field sharing system, we need to record the sound by setting the BoSC microphone array at the position of the audience member seats. To create sound field database for the BoSC system, we have recorded numerous sound fields from classical concerts, traditional Japanese Nou performances, pipe organs in a church, natural sounds and etc. For sound field recordings with the BoSC microphone system, we need an 80-channel recording system. A common PC-based multichannel recording system places restrictions on recording place options because it requires an external power supply. To avoid this issue, a portable 80 channel sound recording system has been constructed and used for recording the sound outdoors and at other places(Omoto and Ikeda, 2012).

When we measure the transfer function from each loudspeaker and each microphone of the BoSC microphone array, we set up the BoSC microphone array at the place where we suppose a listener's head is. For example, when we suppose that a listener is seated in a chair, we set up the BoSC microphone array so that the center of the array is level with a height of 120 cm from the floor to measure the transfer functions.

Sound field sharing system using sound casks

The sound field sharing system allows us to freely select sharing sound field from sound field database via the network. Figure 5 shows the flow of sound signals in the sound field sharing system using the sound cask system. The computer obtains the sound field data to be shared

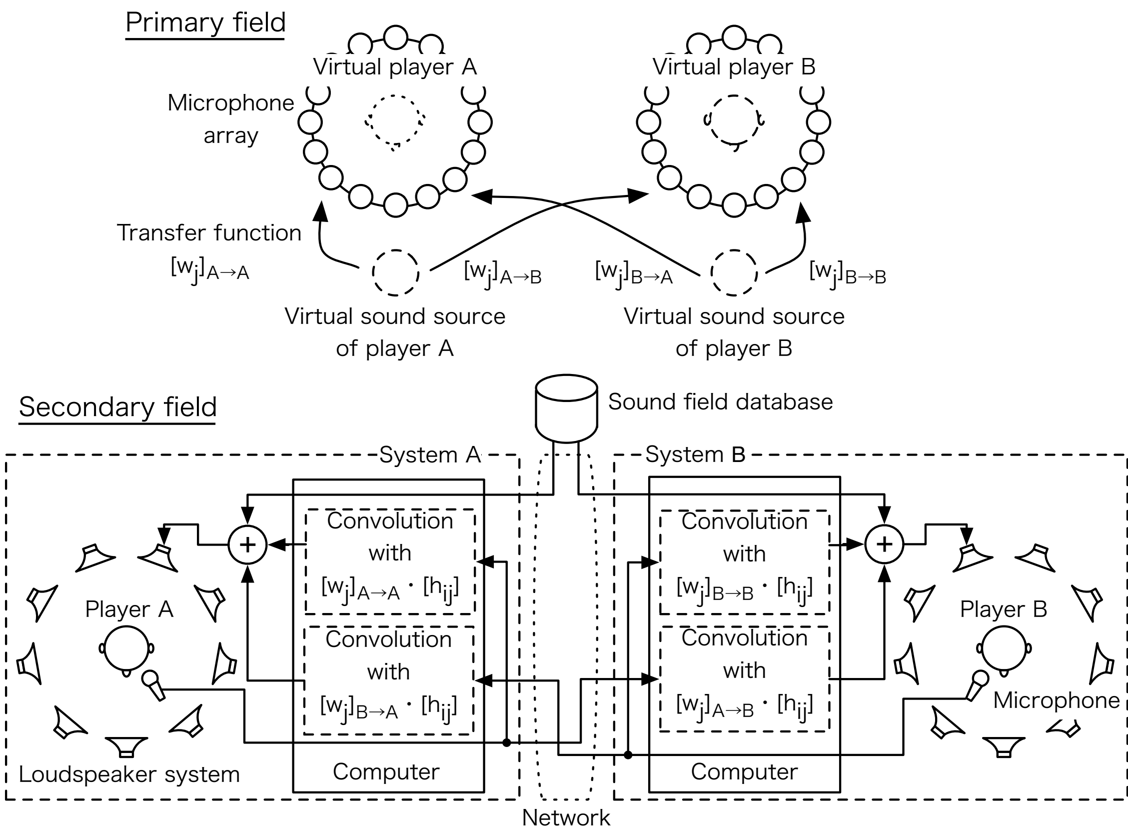

FIGURE 5: Signal flow of the sound field sharing system using sound casks.

with other participants from the recorded sound field database. By preliminarily convolving the sound field data with the inverse system, the system only plays the data to reproduce the sound field.

The sound of voice or a musical performance to be transferred among the systems needs to be convolved with the transfer functions that are measured in the primary sound field. The musical performance recorded inside the system is convolved with the transfer functions from a sound source to control points corresponding to the other listener's virtual position in the primary field ($[w_j]_{A\to B}$, $[w_j]_{B\to A}$). To reproduce the sound reflection of a player's musical performance, we convolved the performance sound with the transfer functions from a sound source to control points corresponding to the player's virtual position in the primary field ($[w_j]_{A\to A}$, $[w_j]_{B\to B}$). Hence, in this case, the direct sound must be removed from the transfer functions. These convolutions involving voice or musical communication must be calculated in real time. Because there are 96 channel loudspeakers as secondary sound sources, we need to calculate a large number of convolutions. To avoid the problem of delay caused by buffering used by the computer to get the sound signals, we can suppress delay time by convolving initial part of the filter using purpose-built hardware. Additionally, the recording of a musical performance inside the system needs an echo canceler.

As a prerequisite for testing the sound field sharing system, we are currently conducting research on the sound reproduction performance of the sound cask from both physical and psychological aspects. Regarding the psychological aspect of the investigation, we are testing sound localization using the sound cask(Kobayashi *et al.*, 2013), and looking into the effect of transmission delay and reverberation time on ensemble performance(Tsubasa *et al.*, 2012). In the future, we will research the effect of the sound field sharing with sound cask on telecommunication by applying the system architecture introduced in this paper.

References

Enomoto, S. (**2009**). "3d sound field recording/reproduction system for telecommunication(in japanese),"architectural acoustics and noise control", Architectural acoustics and noise control **38**, 37–42.

Enomoto, S., Ikeda, Y., Ise, S., and Nakamura, S. (**2008**). "Three-dimentional sound field reproduction and recording system based on boundary surface control principle", in *Proceedings of ICAD*.

Enomoto, S., Ikeda, Y., Ise, S., and Nakamura, S. (**2009**). "3-d sound field reproduction system for the sound field shared communication based on the boundary surface control principle (in japanese)", in *Proceedings of meeting of acoustical society of Japan*, 1411–1414.

Enomoto, S., Ikeda, Y., Ise, S., and Nakamura, S. (**2010**). "Optimization of loudspeaker and microphone configurations for sound reproduction system based on boundary surface control principle", in *Proceedings of International Congress on Acoustics*.

Ikeda, Y., Enomoto, S., Ise, S., and Nakamura, S. (**2010**). "Three-party sound field sharing system based on the boundary surface control principle", in *Proceedings of ICA*.

Ise, S. (**1993**). "A study on the sound field reproduction in a wide area(1) -based on krchhoff-helmholtz integral equation- (in japanese)", in *Proceeings of meeting of acoustical society of japan*, 479–480 (Acoustical Society of America).

Ise, S. (**2007**). "An attempt of sound field sharing system for profound communication. -concept and basic stance of the project- (in japanese)", in *Proceedings of meeting of acoustical society of Japan*, 585–586.

Kleinman, R. and Roach, G. (**1974**). "Boundary integral equations for the three dimensional helmholtz equation", SIAM Review **16**, 214–236.

Kobayashi, M., Ueno, K., Yamashita, M., Ise, S., and Enomoto, S. (**2013**). "Subjective evaluation of a virtual acoustic system: Trials with three-dimensional sound field reproduced by the 'sound cask'", in *Proceedings of International Congress on Acoustics*.

Omoto, A. and Ikeda, I. (**2012**). "Construction of 80-channel mobile sound recording system", in *Proceedings of the AES Japan Section Conference in Sendai*.

Tsubasa, N., Tamaki, W., yusuke, I., yoko, U., and shiro, I. (**2012**). "Study on the effect of sound delay conditions on ensemble performance", in *Proceedings of meeting of acoustical society of japan*.