Proceedings of 20th International Congress on Acoustics, ICA 2010

23-27 August 2010, Sydney, Australia

Theoretical Study of 3D Radiated Sound Field Reproduction System Using Directional Loudspeakers and Boundary Surface Control

Toshiyuki Kimura, Katsuhiro Maki, and Michiaki Katsumoto

National Institute of Information and Communications Technology,
4-2-1 Nukui-kitamachi Koganei, Tokyo, Japan

PACS: 43.60.Ac, 43.60.Sx, 43.60.Tj

ABSTRACT

In order to realize a 3D radiated sound field with which multiple listeners can listen to a sound anywhere around the object without having to wear equipment such as headphones, we proposed a near 3D sound field reproduction system using directional loudspeakers and wave field synthesis. The size of the loudspeaker array, however, is the same as that of the microphone array in the conventional system. Thus, when the size of the loudspeaker array differs, the 3D radiated sound field captured by the microphone array cannot be accurately reproduced. In this paper, the mathematical derivation of the reproduced 3D radiated sound field via inverse filtering based on acoustic transfer functions is described and the 3D radiated sound field reproduction system using directional loudspeakers and boundary surface control is newly proposed.

INTRODUCTION

Several investigations have been carried out on ultra-realistic communication techniques using three-dimensional (3D) audio and video techniques [1]. Applying these techniques enables realization of more realistic communication services than conventional audio and video techniques such as HD video and 5.1ch audio. Realizing these services requires development of audio and video techniques for accurately reproducing a given scene or object.

Ultra-realistic audio is one technique that can help realize ultra-realistic communication. In conventional ultra-realistic audio techniques using multiple loudspeakers, such as wave field synthesis (WFS) [2–4] and 22.2ch audio [5], loudspeakers are placed around the listeners in a large space (e.g., a theater) for realistic representation of a particular scene. In this study, we have developed a 3D audio technique to depict an object itself at a given position, as shown in Figure 1. This technique enables listeners occupying any position around the object to listen to the sound emanating from it.

In order to realize ultra-realistic audio technique that enables multiple listeners to listen to a sound anywhere in its vicinity without having to wear equipment such as headphones, we have proposed a near 3D sound field reproduction system using directional loudspeakers and wave field synthesis [6] and developed the real system by constructing the surrounding microphone array and radiated loudspeaker array [7].

However, while the size of the loudspeaker array is the same as that of the microphone array in the proposed system [6], the size of the radiated loudspeaker array is one-fourth that of the surrounding microphone array in the developed real system [7]. Thus, when the sound recorded by the surrounding microphone array is directly replayed by the radiated loudspeaker array, since the size of the 3D radiated sound field is scaled down to that of the radiated loudspeaker array, the 3D radiated sound field captured by the surrounding microphone array cannot be accurately reproduced in the developed system.

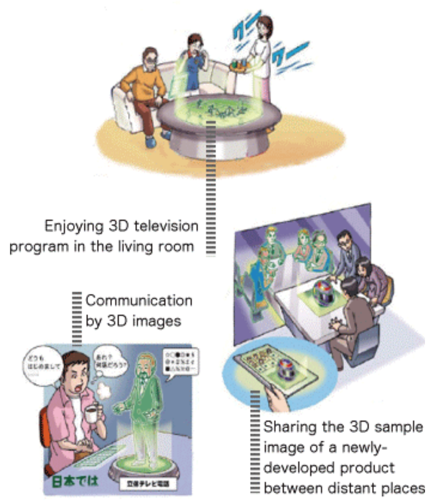

Figure 1. Future image of ultra-realistic communication [1]

On the other hand, the loudspeaker array is not the same size as the microphone array in practical applications of the ultra-realistic communication (e.g., 3D television). Thus, in order to solve the problem of the reproduced 3D radiated sound field being scaled down in the developed system [7], a technique needs to be developed for the 3D radiated sound field to be accurately reproduced even if the loudspeaker array is not the same size as the microphone array.

In this study, to reproduce the 3D radiated sound field when the sizes are not same, a novel 3D radiated sound field reproduction system using directional loudspeakers and boundary surface control [8] is proposed. In the following sections, the theoretical basis of the proposed system is discussed. The 3D radiated sound field can be accurately reproduced via inverse filtering based on acoustic transfer functions, even when the loudspeaker array is not the same size as the microphone array.

THEORETICAL STUDY

Principle of conventional system

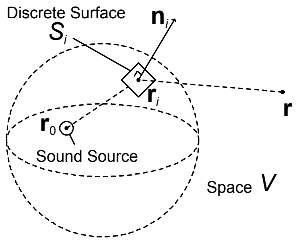

Figure 2. Coordinates in the conventional system

As shown in Figure 2, there are sound sources inside a boundary surface S. If $\mathbf{r}$ denotes the position vector outside a boundary surface S, $\mathbf{r}_i$ is the position vector of S_i (the ith element of a discrete boundary surface), and $\mathbf{n}_i$ is the normal unit vector directed toward the outside of the discrete boundary surface at $\mathbf{r}_i$, $P(\mathbf{r},\omega)$ (the sound pressure at $\mathbf{r}$) can be expressed according to the principle of the conventional system [6]:

$$P(\mathbf{r},\omega) = jk\sum_{i=1}^{M} P(\mathbf{r}_i,\omega)D_s(\mathbf{r}_i\,|\,\mathbf{r}) \\ G(\mathbf{r}_i\,|\,\mathbf{r},\omega)\Delta S_i \quad (\mathbf{r}\in V), \tag{1}$$

where M is the total number of elements in the discrete boundary surface, ΔS_i is the area of S_i, and $D_s(\mathbf{r}_i|\mathbf{r})$ denotes to the directivities of the loudspeakers placed at $\mathbf{r}_i$. $G(\mathbf{r}_i|\mathbf{r},\omega)$ corresponds to the acoustic transfer function from $\mathbf{r}_i$ to $\mathbf{r}$ and is denoted as follows:

$$G(\mathbf{r}_i\,|\,\mathbf{r},\omega) = \frac{\exp(-jk\,|\,\mathbf{r}_i - \mathbf{r}\,|)}{4\pi\,|\,\mathbf{r}_i - \mathbf{r}\,|}, \tag{2}$$

where k ($=\omega/c$) is the wave number, and c is the sound velocity. Eq. (1) shows that the sound pressure in V can be reproduced if the sound pressures $P(\mathbf{r}_i,\omega)$ at the M points are recorded by microphones in the original sound field and directional sound sources with an amplitude $jkP(\mathbf{r}_i,\omega)$ are played at M points in the reproduced sound field.

A diagram of the conventional system is shown in Figure 3. First, the surrounding microphone array, consisting of M omni-directional microphones, is placed around the sound sources in the original sound field and a sound, $P(\mathbf{r}_i,\omega)$, is recorded. The positions of the microphones are on the boundary surface. Second, the radiated loudspeaker array, consisting of M directional loudspeakers, is placed in the reproduced sound field. The positions of the loudspeakers are the same as those of the microphones. The directivity of the loudspeakers is toward the outside of the boundary surface. Finally, the recorded sound, $P(\mathbf{r}_i,\omega)$, is replayed by the radiated loudspeaker array. As a result, because wave fronts are accurately reproduced on the outside of the radiated loudspeaker array, listeners on the outside of the radiated loudspeaker array can feel that the sound sources are being played inside it. In the case of Figure 3, listeners near the piano can feel they are listening to the sound near the piano; the same for the violin.

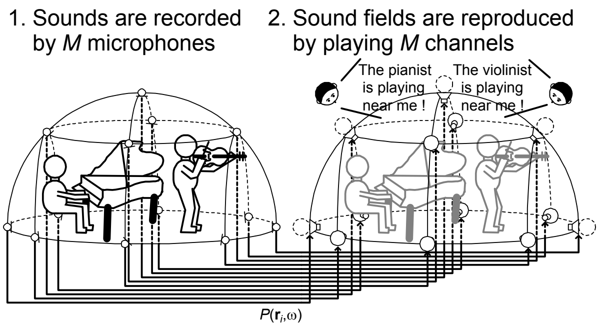

Figure 3. Diagram of the conventional system

However, since the radiated loudspeaker array is the same size as the surrounding microphone array in this system, the size of the reproduced 3D radiated sound field depends on the size of the radiated loudspeaker array. Thus, when the radiated loudspeaker array is not the same size, it cannot accurately reproduce the 3D radiated sound field captured by the surrounding microphone array.

Principle of proposed system

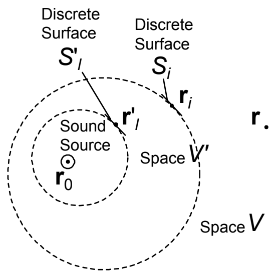

Figure 4. Coordinates in the proposed system

As shown in Figure 4, a boundary surface S' is considered inside the boundary surface S. A space V' is outside a boundary surface S' (i.e., $V\in V'$). Let S'_l (l=1...N) be the lth discrete boundary surface of V'. $\mathbf{r}'_l$ and N are the position vector and total number of discrete surfaces, respectively. To reproduce the 3D sound field in V', the following equation, derived from Eq. (1), is used:

$$P(\mathbf{r},\omega) = jk\sum_{l=1}^{N} P(\mathbf{r}'_l,\omega)D_s(\mathbf{r}'_l\,|\,\mathbf{r}) \\ G(\mathbf{r}'_l\,|\,\mathbf{r},\omega)\Delta S'_l \quad (\mathbf{r}\in V'), \tag{3}$$

where $\Delta S'_l$ is the area of S'_l. Since $\mathbf{r}_i$ is always in V', the following equation can be derived from Eq. (3):

$$P(\mathbf{r}_i,\omega) = jk\sum_{l=1}^{N} P(\mathbf{r}'_l,\omega)D_s(\mathbf{r}'_l\,|\,\mathbf{r}_i) \\ G(\mathbf{r}'_l\,|\,\mathbf{r}_i,\omega)\Delta S'_l \quad (\mathbf{r}_i\in V'). \tag{4}$$

If Eq. (4) is substituted in Eq. (1), the following equation is obtained:

$$P(\mathbf{r},\omega) = jk\sum_{i=1}^{M}\left\{ jk\sum_{l=1}^{N} P(\mathbf{r'}_l,\omega) D_s(\mathbf{r'}_l\,|\,\mathbf{r}_i)G(\mathbf{r'}_l\,|\,\mathbf{r}_i,\omega)\Delta S'_l \right\} D_s(\mathbf{r}_i\,|\,\mathbf{r})G(\mathbf{r}_i\,|\,\mathbf{r},\omega)\Delta S_i$$
$$= jk\sum_{l=1}^{N} P(\mathbf{r'}_l,\omega)\left\{ jk\sum_{i=1}^{M} D_s(\mathbf{r'}_l\,|\,\mathbf{r}_i) G(\mathbf{r'}_l\,|\,\mathbf{r}_i,\omega)D_s(\mathbf{r}_i\,|\,\mathbf{r}) G(\mathbf{r}_i\,|\,\mathbf{r},\omega)\Delta S_i \right\}\Delta S'_l$$
$$(\mathbf{r}\in V, \mathbf{r}_i \in V') \quad , \qquad (5)$$

Thus, by comparing Eq. (5) with Eq. (3), the following equation is derived:

$$D_s(\mathbf{r'}_l\,|\,\mathbf{r})G(\mathbf{r'}_l\,|\,\mathbf{r},\omega) = jk\sum_{i=1}^{M} D_s(\mathbf{r'}_l\,|\,\mathbf{r}_i)G(\mathbf{r'}_l\,|\,\mathbf{r}_i,\omega) D_s(\mathbf{r}_i\,|\,\mathbf{r})G(\mathbf{r}_i\,|\,\mathbf{r},\omega)\Delta S_i$$
$$(\mathbf{r}\in V, \mathbf{r}_i \in V') \quad , \qquad (6)$$

The following scheme is considered: the sound is recorded by M microphones placed at $\mathbf{r}_i$, the recorded signals are processed by M-input N-output filters, and the filtered signals are played by the N directional loudspeakers placed at $\mathbf{r'}_l$. If all the M recorded signals are denoted as $P(\mathbf{r}_i,\omega)$, the N filtered signals $P'(\mathbf{r'}_l,\omega)$ (l=1...N) are expressed as follows:

$$P'(\mathbf{r'}_l,\omega) = \sum_{i=1}^{M} H_{li}(\omega)P(\mathbf{r}_i,\omega) \quad , \qquad (7)$$

where $H_{li}(\omega)$ are the coefficients of the M-input N-output filters. Thus, the reproduced 3D sound field $P'(\mathbf{r}, \omega)$ is written as

$$P'(\mathbf{r},\omega) = \sum_{l=1}^{N} P'(\mathbf{r'}_l,\omega)D_s(\mathbf{r'}_l\,|\,\mathbf{r}) G(\mathbf{r'}_l\,|\,\mathbf{r},\omega)\Delta S'_l$$
$$= \sum_{l=1}^{N}\left\{\sum_{i=1}^{M} H_{li}(\omega)P(\mathbf{r}_i,\omega)\right\} D_s(\mathbf{r'}_l\,|\,\mathbf{r})G(\mathbf{r'}_l\,|\,\mathbf{r},\omega)\Delta S'_l$$
$$= \sum_{i=1}^{M} P(\mathbf{r}_i,\omega)\left\{\sum_{l=1}^{N} H_{li}(\omega) D_s(\mathbf{r'}_l\,|\,\mathbf{r})G(\mathbf{r'}_l\,|\,\mathbf{r},\omega)\Delta S'_l\right\} \quad . \qquad (8)$$

If Eq. (6) is substituted in Eq. (8), the following equation is obtained:

$$P'(\mathbf{r},\omega) = \sum_{i=1}^{M} P(\mathbf{r}_i,\omega)\left[\sum_{l=1}^{N} H_{li}(\omega) \left\{ jk\sum_{n=1}^{M} D_s(\mathbf{r'}_l\,|\,\mathbf{r}_n)G(\mathbf{r'}_l\,|\,\mathbf{r}_n,\omega) D_s(\mathbf{r}_n\,|\,\mathbf{r})G(\mathbf{r}_n\,|\,\mathbf{r},\omega)\Delta S_n \right\}\Delta S'_l\right]$$
$$= jk\sum_{i=1}^{M} P(\mathbf{r}_i,\omega)\left[\sum_{n=1}^{M} D_s(\mathbf{r}_n\,|\,\mathbf{r}) G(\mathbf{r}_n\,|\,\mathbf{r},\omega)\left\{\sum_{l=1}^{N} H_{li}(\omega) D_s(\mathbf{r'}_l\,|\,\mathbf{r}_n)G(\mathbf{r'}_l\,|\,\mathbf{r}_n,\omega)\Delta S'_l\right\}\Delta S_n\right]$$
$$(\mathbf{r}\in V, \mathbf{r}_i \in V') \quad . \qquad (9)$$

If the M-input N-output filters are defined as

$$\sum_{l=1}^{N} H_{li}(\omega)D_s(\mathbf{r'}_l\,|\,\mathbf{r}_n)G(\mathbf{r'}_l\,|\,\mathbf{r}_n,\omega)\Delta S'_l = \begin{cases} 1 & (n=i) \\ 0 & (n\neq i) \end{cases} \quad , \qquad (10)$$

the 3D sound field is reproduced in V; this is indicated by the following equation derived from Eq. (9):

$$P'(\mathbf{r},\omega) = jk\sum_{i=1}^{M} P(\mathbf{r}_i,\omega)D_s(\mathbf{r}_i\,|\,\mathbf{r})G(\mathbf{r}_i\,|\,\mathbf{r},\omega)\Delta S_i$$
$$= P(\mathbf{r},\omega) \quad (\mathbf{r}\in V, \mathbf{r}_i \in V') \quad . \qquad (11)$$

Thus, if the recorded signals are processed by the filters defined as in Eq. (10), the 3D sound field can be reproduced in the listening area (space V) even if the loudspeakers are placed at the position ($\mathbf{r'}_l$), which is not the same as the recorded position ($\mathbf{r}_i$). Since $\mathbf{r}_i$ and $\mathbf{r'}_l$ are points on the boundary surfaces S and S', respectively, it is necessary for the loudspeakers to be placed on the boundary surface (S') which is enveloped by the boundary surface used for recording (S).

Let the matrix form of Eq. (10) be denoted as follows:

$$\mathbf{G}(\omega)\mathbf{H}(\omega) = \mathbf{I} \quad , \qquad (12)$$

$$\mathbf{G}(\omega) = \begin{pmatrix} D_{s11}G_{11} & \cdots & D_{sN1}G_{N1} \\ \vdots & \ddots & \vdots \\ D_{s1M}G_{1M} & \cdots & D_{sNM}G_{NM} \end{pmatrix} \quad , \qquad (13)$$

$$\mathbf{H}(\omega) = \begin{pmatrix} H_{11}(\omega) & \cdots & H_{1M}(\omega) \\ \vdots & \ddots & \vdots \\ H_{N1}(\omega) & \cdots & H_{MN}(\omega) \end{pmatrix} \quad , \qquad (14)$$

$$\mathbf{I} = \begin{pmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{pmatrix} \quad , \qquad (15)$$

where $D_{sli}G_{li}$ denotes $D_s(\mathbf{r'}_l|\mathbf{r}_i)G(\mathbf{r'}_l|\mathbf{r}_i,\omega)\Delta S'_l$. Thus, the filters $\mathbf{H}(\omega)$ are calculated according to the following equation:

$$\mathbf{H}(\omega) = \mathbf{G}^{+}(\omega) \, , \qquad (16)$$

where $\mathbf{G}^{+}(\omega)$ is the Moore-Penrose pseudo inverse matrix of $\mathbf{G}(\omega)$. Since $\mathbf{G}(\omega)$ is the matrix consisting of the acoustic transfer functions from $\mathbf{r'}_l$ to $\mathbf{r}_i$, the calculated filters $\mathbf{H}(\omega)$ are inverse filters of the acoustic transfer functions.

The diagram of the proposed 3D radiated sound field reproduction system based on directional loudspeakers and boundary surface control is shown in Figure 5. First, in the original sound field, M microphones are placed at the points $\mathbf{r}_i$ on the boundary surface S of the control area and the sound $P(\mathbf{r}_i,\omega)$ is recorded. Second, in the reproduced sound field, M microphones are placed at the points $\mathbf{r}_i$, and N directional loudspeakers are placed at the points $\mathbf{r}'_l$ on a boundary surface S' inside the control area. The position of the M microphones is the same as those of the microphones during recording. The directional loudspeakers are then directed toward the outside of the boundary surface S'. Third, the acoustic transfer functions from the N directional loudspeakers to the M microphones $D_s(\mathbf{r}'_l|\mathbf{r}_i)G(\mathbf{r}'_l|\mathbf{r}_i,\omega)\Delta S'_l$ are measured, and the inverse filters $H_{li}(\omega)$ are calculated from the measured acoustic transfer functions. Finally, the recorded signals are filtered by the inverse filters, and the filtered signals $P'(\mathbf{r}'_l,\omega)$ are played by the N directional loudspeakers. As a result, since the 3D radiated sound field is reproduced in the listening area, listeners feel as if they are listening to the sound around sound sources of the original sound field. Since the sound pressure at $\mathbf{r}_i$ is the same as the recorded sound $P(\mathbf{r}_i,\omega)$ in the reproduced sound field, the sound sources can be placed at arbitrary positions inside the control area. Thus, if the sound sources are placed inside the control area and outside the loudspeaker array, listeners can be made to feel as if there are sound images outside the loudspeaker array.

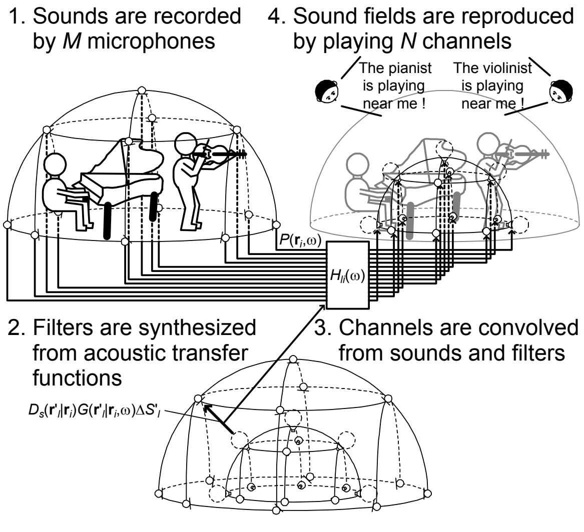

Figure 5. Diagram of the proposed system

When the 3D radiated sound field reproduction system is constructed in a real environment, since the acoustic transfer functions include an initial delay, filters that do not satisfy the causality are obtained if the inverse filters are calculated using Eq. (16). In order to obtain filters satisfying the causality, the inverse filters are calculated using the following equations:

$$\mathbf{H}(\omega) = \mathbf{G}^{+}(\omega)\mathbf{I}(\omega), \tag{17}$$

$$\mathbf{I}(\omega) = \begin{pmatrix} e^{-j\omega T} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & e^{-j\omega T} \end{pmatrix}, \tag{18}$$

where T is the delay time required for the calculation of the inverse filters satisfying the causality. The total delay of the system is T and the reproduced 3D radiated sound field can be expressed as follows:

$$\begin{aligned} P'(\mathbf{r},\omega) &= e^{-j\omega T} jk \sum_{i=1}^{M} P(\mathbf{r}_i,\omega) D_s(\mathbf{r}_i \mid \mathbf{r}) \\ &\qquad G(\mathbf{r}_i \mid \mathbf{r},\omega)\Delta S_i \\ &= e^{-j\omega T} P(\mathbf{r},\omega) \quad (\mathbf{r} \in V, \mathbf{r}_i \in V') \end{aligned} . \tag{19}$$

CONCLUSION

In this study, we proposed a 3D radiated sound field reproduction system using directional loudspeakers and boundary surface control. The system was constructed by using the inverse filters in the conventional system. The proposed system can reproduce a 3D radiated sound field in a listening area even if the loudspeaker array is not the same size as the microphone array.

In a future study, a computer simulation needs to be performed to numerically analyze the reproduced 3D radiated sound field.

REFERENCES

[1] *Website of 3D Spatial Image and Sound Group, Universal Media Research Center, National Institute of Information and Communications Technology,* http://www2.nict.go.jp/x/x171/index_e.html

[2] M. Camras, "Approach to recreating a sound field," *J. Acoust. Soc. Am.* **43**, 1425–1431 (1968)

[3] A. J. Berkhout, D. de Vries, and P. Vogel, "Acoustic control by wave field synthesis" *J. Acoust. Soc. Am.* **93** 2764–2778, (1993)

[4] T. Kimura and K. Kakehi, "Effects of directivity of microphones and loudspeakers in sound field reproduction based on wave field synthesis" in *Proc. 19th ICA*, No. RBA-15-011, pp. 1–6 (2007)

[5] K. Hamasaki, T. Nishiguchi, R. Okumura, Y. Nakayama and A. Ando, "A 22.2 multichannel sound system for ultrahigh-definition TV (UDHTV)" *SMPTE Motion Imaging Journal* **117**(3), 40–49 (2008)

[6] T. Kimura, Y. Yamakata, and M. Katsumoto, "Theoretical study of near 3D sound field reproduction based on wave field synthesis" in *Proc. Acoustics'08*, No. 1763, 4585–4590 (2008)

[7] T. Kimura, Y. Yamakata, M. Katsumoto, T. Okamoto, S. Yairi, Y. Iwaya and Y. Suzuki, "Development of real system in near 3D sound field reproduction system using directional loudspeakers and wave field synthesis" in *Proc. WESPAC*, No. 0164, 1–6 (2009)

[8] S. Ise, "A principle of sound field control based on the Kirchhoff-Helmholtz integral equation and the theory of inverse systems" *ACUSTICA – Acta Acustica* **85**, 78–87 (1999)