Acoust. Sci. & Tech.

PAPER

Three-dimensional Radiated Sound Field Display System Using Directional Loudspeakers and Wave Field Synthesis

Toshiyuki Kimura[1*], Yoko Yamakata[2], Michiaki Katsumoto[1], Takuma Okamoto[3], Satoshi Yairi[4], Yukio Iwaya[3] and Yôiti Suzuki[3]

[1]*National Institute of Information and Communications Technology, 4–2–1, Nukui-kitamachi, Koganei, Tokyo, 184–8795 Japan*
[2]*Kyoto University, Yoshida-honmachi, Sakyo-ku, Kyoto, 606–8501 Japan*
[3]*Tohoku University, 2–1–1, Katahira, Aoba-ku, Sendai, 980–8577 Japan*
[4]*Sendai National College of Technology, 48, Nodayama, Medeshima-shiote, Natori, Miyagi, 981–1239 Japan*

Abstract: Three-dimensional (3D) radiated sound field display systems are important toward realizing ultra-realistic communications systems such as 3D television. In this paper, a 3D radiated sound field display system using directional loudspeakers and wave field synthesis is proposed. The proposed system is based on the Fresnel-Kirchhoff diffraction formula, which is the approximate form of the Kirchhoff-Helmholtz integral equation. The proposed system was developed by constructing the surrounding microphone array and radiated loudspeaker array. To evaluate the performance of the developed system, the positions of sound images were estimated. The results indicated that the 3D sound field is accurately controlled outside of the radiated loudspeaker array by sharpening the radiation directivity of the loudspeaker units and that the sound images are not correctly estimated in the developed system because the radiation directivity of loudspeaker units is not sufficiently sharp.

Keywords: Radiated sound field, Sound field display, Directional loudspeaker, Wave field synthesis, Sound image position estimation

PACS number: 43.58.Fm, 43.58.Ta, 43.60.Dh, 43.60.Sx

1. Introduction

In recent years, ultra-realistic communications techniques have been investigated [1]. If, by applying these techniques, realistically reproduced three-dimensional (3D) video and audio can appear in a 3D space, as shown in Fig. 1, and several people can view an object anywhere in its vicinity without having to wear equipment such as glasses, this will enable more realistic forms of communication (e.g., 3D television) than those currently offered by conventional video and audio techniques (HD video and 5.1-channel audio).

To realize 3D audio systems that are compatible with 3D video systems for 3D television, 3D sound field reproduction systems must be developed that enable multiple listeners to listen to an object's sound anywhere in its vicinity without the need for headphones. Wave field synthesis [2–4] is a 3D sound field reproduction technique for reproducing wave fronts from a control area in a different area (the listening area), according to the Kirchhoff-Helmholtz integral equation [5]. Multiple listeners can listen to the replayed sounds anywhere in the listening area without wearing devices such as headphones because the technique reproduces the sound field of a 3D space rather than that of binaural positions.

A 3D sound field reproduction system using wave field synthesis for 3D television consists of three systems, as shown in Fig. 2. In the 3D radiated sound field capture system, the multichannel sounds are recorded by microphones placed around sound sources. The recorded multichannel sounds consist of direct sounds from sound sources. In the 3D radiated sound field analysis system, the sound source signals and object information (such as the position and orientation of sound sources) are extracted from the recorded multichannel sounds. In the 3D radiated sound field display system, the multichannel audio signals are synthesized from extracted sound source signals and the object information and played by the loudspeakers. In this paper, we investigate

Fig. 1 Future image of ultra-realistic communications such as 3D television using 3D video and 3D audio [1].

* t-kimura@nict.go.jp

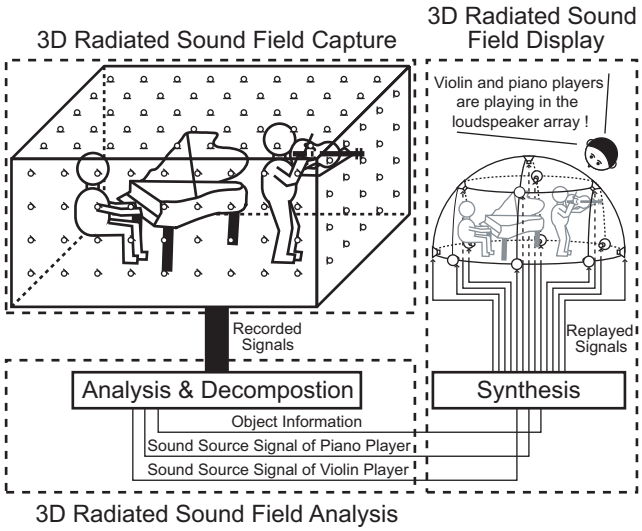

Fig. 2 Component systems in the 3D sound field reproduction system using wave field synthesis for 3D television.

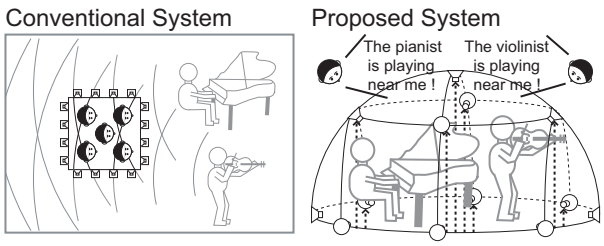

Fig. 3 Reproduced sound fields in conventional and proposed 3D sound field display systems.

a 3D radiated sound field display system using wave field synthesis.

In conventional 3D sound field display systems using wave field synthesis, the loudspeakers are placed around the listeners and the sound scene is displayed inside the loudspeaker array, as shown on the left side of Fig. 3. However, in the proposed systems, the loudspeakers are placed around the sound sources and the radiated sound field of the sound sources is displayed outside the radiated loudspeaker array, as shown on the right side of Fig. 3. Though the conditions for displaying accurate wave fronts in the conventional systems have been theoretically studied [4], similar study has not been performed for the proposed systems.

A spherical loudspeaker array radiating sounds to the outside has been already proposed [6, 7]. However, in this array, one sound source is placed at the center of a sphere in order to reproduce the radiation directivity of sound sources, and only sound pressures of the boundary on the outside of a spherical loudspeaker array are controlled. On the other hand, in our proposed system, multiple sound sources can be placed in the radiated loudspeaker array, and sound pressures of the space on the outside of the radiated loudspeaker array

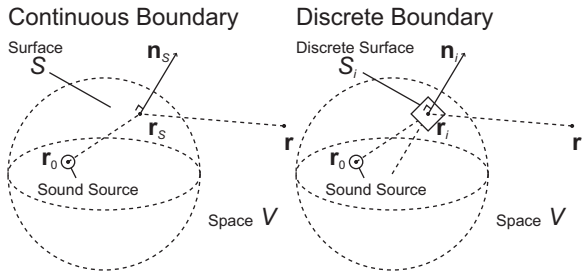

Fig. 4 Coordinates in Kirchhoff-Helmholtz integral equation.

can be controlled.

This paper proposes 3D radiated sound field display systems based on wave field synthesis to enable multiple listeners to listen to sounds in the vicinity of sound sources. In Section 2, the theoretical basis of the proposed system is discussed. The 3D sound field on the outside of the radiated loudspeaker array can be accurately controlled if directional loudspeakers are applied based on the Kirchhoff-Helmholtz integral equation. The diagram of the proposed 3D radiated sound field display system using directional loudspeakers and wave field synthesis is also described. Section 3 describes the specifications of the developed prototype system. In Section 4, in order to evaluate the performance of the developed system, the positions of sound images are estimated.

2. Algorithm of Proposed System

2.1. Kirchhoff-Helmholtz Integral Equation

As shown on the left in Fig. 4, if the sound sources are surrounded by a continuous boundary surface S, and $\mathbf{r}_S$ and $\mathbf{r}$ are the position vectors of S and of space V (the outside of S), respectively, $P(\mathbf{r}, \omega)$, which is the sound pressure at $\mathbf{r}$, can be denoted as follows:

$$P(\mathbf{r},\omega) = \frac{1}{4\pi}\oint_S \left\{ \frac{\partial P(\mathbf{r}_S,\omega)}{\partial \mathbf{n}_S} \frac{e^{-jk|\mathbf{r}-\mathbf{r}_S|}}{|\mathbf{r}-\mathbf{r}_S|} - P(\mathbf{r}_S,\omega)\frac{\partial}{\partial \mathbf{n}_S}\left(\frac{e^{-jk|\mathbf{r}-\mathbf{r}_S|}}{|\mathbf{r}-\mathbf{r}_S|}\right)\right\} dS \quad (\mathbf{r}\in V), \quad (1)$$

where $k(=\omega/c)$ is a wave number, c is a sound velocity, and $\mathbf{n}_S$ is the normal unit vector directed toward the outside of the continuous boundary surface at $\mathbf{r}_S$. If the Sommerfeld radiation condition [5] is satisfied, Eq. (1) shows that the sound pressure in V can be controlled if monopole sound sources with an amplitude $\frac{\partial P(\mathbf{r}_S,\omega)}{\partial \mathbf{n}_S}$ and dipole sound sources with an amplitude $-P(\mathbf{r}_S, \omega)$ are played at $\mathbf{r}_S$.

However, in order to develop a sound field display system, S must be discretized since the monopole and dipole sources are not placed continuously along S. As shown on the right in Fig. 4, if $\mathbf{r}_i$ is the position vector of S_i (ith element of a discrete boundary surface), and $P(\mathbf{r}_i, \omega)$ (sound pressure at $\mathbf{r}_i$) and $\frac{\partial P(\mathbf{r}_i,\omega)}{\partial \mathbf{n}_i}$ (sound pressure gradient at $\mathbf{r}_i$) are constant for S_i, Eq. (1) can be rewritten as follows:

$$P(\mathbf{r},\omega)=\frac{1}{4\pi}\sum_{i=1}^{M}\left\{\frac{\partial P(\mathbf{r}_i,\omega)}{\partial \mathbf{n}_i}\frac{e^{-jk|\mathbf{r}-\mathbf{r}_i|}}{|\mathbf{r}-\mathbf{r}_i|}-P(\mathbf{r}_i,\omega)\frac{\partial}{\partial \mathbf{n}_i}\left(\frac{e^{-jk|\mathbf{r}-\mathbf{r}_i|}}{|\mathbf{r}-\mathbf{r}_i|}\right)\right\}\Delta S_i\ (\mathbf{r}\in V), \quad (2)$$

where M is the total number of elements in the discrete boundary surface, ΔS_i is the area of S_i, and $\mathbf{n}_i$ is the normal unit vector directed toward the outside of the discrete boundary surface at $\mathbf{r}_i$. Eq. (2) shows that the sound pressure in V can be controlled if monopole sound sources with an amplitude $\frac{\partial P(\mathbf{r}_i,\omega)}{\partial \mathbf{n}_i}$ and dipole sound sources with an amplitude $-P(\mathbf{r}_i,\omega)$ are played at M points in the position $\mathbf{r}_i$.

However, from a practical standpoint, it is very difficult to realize sound field display systems according to Eq. (2) because the monopole sound source and dipole sound source cannot be placed at the same position $\mathbf{r}_i$. In Section 2.2., approximations are introduced to Eq. (2) in order to realize practical sound field display systems.

2.2. Fresnel-Kirchhoff Diffraction Formula

The sound pressure at $\mathbf{r}_i$ can be denoted as follows:

$$P(\mathbf{r}_i,\omega)=\frac{Ae^{-jk|\mathbf{r}_i-\mathbf{r}_0|}}{|\mathbf{r}_i-\mathbf{r}_0|},$$

where $\mathbf{r}_0$ and A are the position vector and the amplitude, respectively, of the point sound sources in the original sound field. The sound pressure gradient $\frac{\partial P(\mathbf{r}_i,\omega)}{\partial \mathbf{n}_i}$ was approximated as follows:

$$\begin{aligned}\frac{\partial P(\mathbf{r}_i,\omega)}{\partial \mathbf{n}_i}&=\frac{\partial}{\partial \mathbf{n}_i}\left(\frac{Ae^{-jk|\mathbf{r}_i-\mathbf{r}_0|}}{|\mathbf{r}_i-\mathbf{r}_0|}\right)\\&=-\frac{Ae^{-jk|\mathbf{r}_i-\mathbf{r}_0|}}{|\mathbf{r}_i-\mathbf{r}_0|}\left(\frac{1}{|\mathbf{r}_i-\mathbf{r}_0|}+jk\right)\cos\theta_{i0}\\&=-P(\mathbf{r}_i,\omega)\left(\frac{1}{|\mathbf{r}_i-\mathbf{r}_0|}+jk\right)\cos\theta_{i0}\\&\approx -jkP(\mathbf{r}_i,\omega)\cos\theta_{i0}\left(\text{if } k\gg\frac{1}{|\mathbf{r}_i-\mathbf{r}_0|}\right),\end{aligned} \quad (3)$$

where $\theta_{i0}(=\cos^{-1}\frac{\mathbf{n}_i\cdot(\mathbf{r}_i-\mathbf{r}_0)}{|\mathbf{n}_i||\mathbf{r}_i-\mathbf{r}_0|})$ denotes the angle between the vector $\mathbf{n}_i$ and the vector $\mathbf{r}_i-\mathbf{r}_0$. The dipole sound sources were also approximated as follows:

$$\begin{aligned}\frac{\partial}{\partial \mathbf{n}_i}\left(\frac{e^{-jk|\mathbf{r}-\mathbf{r}_i|}}{|\mathbf{r}-\mathbf{r}_i|}\right)&=-\frac{e^{-jk|\mathbf{r}-\mathbf{r}_i|}}{|\mathbf{r}-\mathbf{r}_i|}\left(\frac{1}{|\mathbf{r}-\mathbf{r}_i|}+jk\right)\cos\theta_i\\&\approx -jk\frac{e^{-jk|\mathbf{r}-\mathbf{r}_i|}}{|\mathbf{r}-\mathbf{r}_i|}\cos\theta_i\left(\text{if } k\gg\frac{1}{|\mathbf{r}-\mathbf{r}_i|}\right),\end{aligned} \quad (4)$$

where $\theta_i(=\cos^{-1}\frac{\mathbf{n}_i\cdot(\mathbf{r}_i-\mathbf{r})}{|\mathbf{n}_i||\mathbf{r}_i-\mathbf{r}|})$ denotes the angle between the vector $\mathbf{n}_i$ and vector $\mathbf{r}_i-\mathbf{r}$. If Eqs. (3)-(4) are substituted in Eq. (2), we derive the following equation in $\mathbf{r}\in V$:

$$P(\mathbf{r},\omega)=\frac{jk}{4\pi}\sum_{i=1}^{M}P(\mathbf{r}_i,\omega)\frac{e^{-jk|\mathbf{r}-\mathbf{r}_i|}}{|\mathbf{r}-\mathbf{r}_i|}(\cos\theta_i-\cos\theta_{i0})\Delta S_i\ (\mathbf{r}\in V). \quad (5)$$

This equation is known as the Fresnel-Kirchhoff diffraction formula [8]. In the proposed system, since the direction of vector $\mathbf{r}_i-\mathbf{r}_0$ is almost the same as that of vector $\mathbf{n}_i$ when sound sources are placed near the center of surrounded area by boundary surface, it can be approximated as $\cos\theta_{i0}\approx 1$. Thus, Eq. (5) can also be written as follows:

$$\begin{aligned}P(\mathbf{r},\omega)&\approx\frac{jk}{4\pi}\sum_{i=1}^{M}P(\mathbf{r}_i,\omega)\frac{e^{-jk|\mathbf{r}-\mathbf{r}_i|}}{|\mathbf{r}-\mathbf{r}_i|}(\cos\theta_i-1)\Delta S_i\\&=\frac{jk}{4\pi}\sum_{i=1}^{M}P(\mathbf{r}_i,\omega)D(\theta_i)\frac{e^{-jk|\mathbf{r}-\mathbf{r}_i|}}{|\mathbf{r}-\mathbf{r}_i|}\Delta S_i,\end{aligned} \quad (6)$$

where $D(\theta_i)(=\cos\theta_i-1)$ corresponds to the directivity of the point sound sources placed at $\mathbf{r}_i$. Eq. (6) shows that the sound pressure in V can be controlled if directional point sound sources with an amplitude $P(\mathbf{r}_i,\omega)$ are played at M points in the position $\mathbf{r}_i$.

Unlike the case of Eq. (2) described in Section 2.1., the sound source placed in the position $\mathbf{r}_i$ is only the directional point sound source. Thus, if Eq. (6) is applied, the practical sound field display systems can be constructed. Although $2M$ microphones and loudspeakers are needed in the case of Eq. (2), the necessary number of microphones and loudspeakers is M in the system based on Eq. (6).

2.3. Diagram of Proposed System

A diagram of the proposed 3D radiated sound field display system is shown in Fig. 5. First, the radiated loudspeaker array consisting of M directional loudspeaker units is placed in the listening room. The directivity of the loudspeaker units is toward the outside of the boundary surface. Second, multichannel audio signals, $P(\mathbf{r}_i,\omega)$, are synthesized by convolving the sound source signals to the acoustic transfer functions from sound sources to M omnidirectional microphones. The positions of the M omnidirectional microphones are the same as those of the loudspeaker units. Third, synthesized multichannel audio signals are played from the radiated loudspeaker array. Since the 3D sound field on the outside of the radiated loudspeaker array is accurately controlled according to Eq. (6), listeners on the outside of array can resultantly feel that the sound sources are being played on the inside of the array. In the case of Fig. 5, listeners close to the piano can feel that they are listening to the sound near the piano, and the same for violin.

3. Development of Proposed System

3.1. Surrounding Microphone Array

For the surrounding microphone array, we used a surrounding microphone array room at the Research Institute of Electrical Communication, Tohoku University [9]. The reverberation time of the room was approximately 150 ms, and 157 omnidirectional microphones (Brüel & Kjær: Type

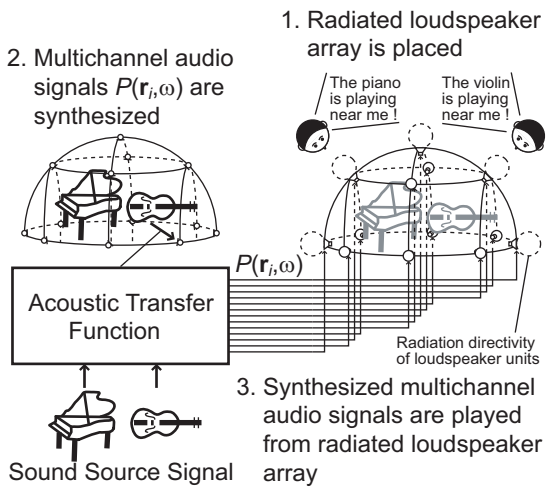

Fig. 5 Diagram of proposed 3D radiated sound field display systems using directional loudspeakers and wave field synthesis.

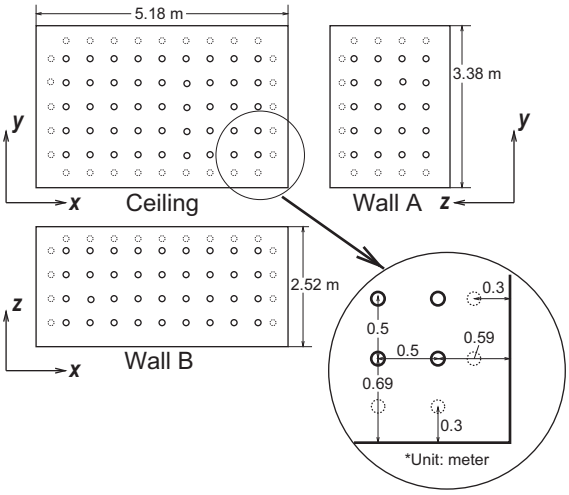

Fig. 7 Arrangement of microphones in the surrounding microphone array [9].

Fig. 6 Image of surrounding microphone array placed in the room [9].

Fig. 8 Image of radiated loudspeaker array.

4951) were attached to the five planes of the room, as shown in Fig. 6. These microphones mainly record the direct sound from sound sources placed in the room because sound absorption panels are also attached to all of the room's planes.

The arrangement of microphones is shown in Fig. 7. To the two narrow sidewall planes (Wall A), 20 (= 5×4) microphones are attached. To the two wide sidewall planes (Wall B), 36 (= 9×4) microphones are attached. And to the ceiling plane, 45 (= 9×5) microphones are attached. The interval between the microphones is 0.5 m. The microphones are connected to 10 microphone preamplifiers for 16 channels (Brüel & Kjær: Type 2694). We have recorded the 157-channel sound of the string quartet by using this array in the 3D radiated sound field capture system [10].

3.2. Radiated Loudspeaker Array

For the radiated loudspeaker array, we newly manufactured that shown in Fig. 8. Its size is one-fourth that of the surrounding microphone array. It has a rectangular enclosure, 157 loudspeaker units, and a stand. The size of the rectangular enclosure is 1.145 m (width) × 0.695 m (depth) × 0.555 m (height), made of plywood and aluminum panels. The 157 loudspeaker units, with a size of 1 inch (AURASOUND: NSW1-205-8A suitable), are directly attached to the five planes of the rectangular enclosure. The directivity of each loudspeaker unit is toward the outside of the rectangular enclosure because the sound radiated from each unit to the inside of the enclosure does not leak to its outside.

The arrangement of the loudspeaker units is shown in Fig. 9. To the two narrow sidewall planes (Wall A), 20 (= 5×4) loudspeaker units are attached. To the two wide sidewall planes (Wall B), 36 (= 9×4) loudspeaker units are attached. And to the ceiling plane, 45 (= 9×5) loudspeaker units are attached. The interval between the loudspeaker units is 0.125

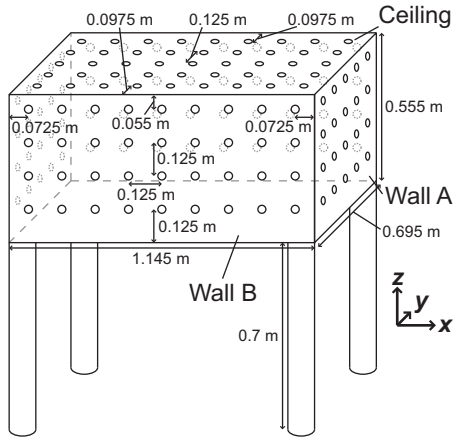

Fig. 9 Arrangement of loudspeaker units in the radiated loudspeaker array.

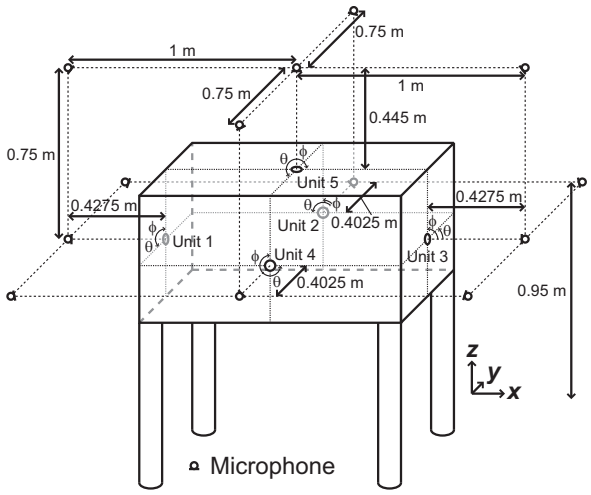

Fig. 10 Positions of the typical loudspeaker units and microphones used in the radiation directivity measurement.

m, which is less than half the wavelength of 1000-Hz sound waves (= $\frac{340\,\mathrm{m}}{1000\,\mathrm{Hz}}$=34 cm). Thus, the spatial sampling theorem that pertains to the reproduction of wave fronts of sound waves with frequencies below 1000 Hz is satisfied. These loudspeaker units are connected to a preamplifier with 157 channels (custom-made). The radiated loudspeaker array is elevated 0.7 m by the stand.

3.3. Radiation Directivity of Loudspeaker Units

To simulate the radiation directivity of loudspeaker units of the developed radiated loudspeaker array in the performance evaluation of Section 4, the radiation directivity of loudspeaker units in the developed radiated loudspeaker array was measured.

First, impulse responses from loudspeaker units to microphones were measured. Microphones (DPA: 4060) were placed at the 13 positions around five typical loudspeaker units selected from the 157 units as shown in Fig. 10. The reverberation time, temperature, and background noise level of the measured room were respectively about 180 ms, 25.5°C, and 24.8 dB(A). Time-Stretched Pulse (TSP) signals [11] (sampling frequency: 48 kHz, length: 65536 samples) were emitted from the playing equipment (Digidesign: ProTools HD) and five typical loudspeaker units, and captured by 13 microphones and the recording equipment (Digidesign: ProTools HD) installed in a PC (Apple: Mac Pro). The sound pressure level was set to 89.7 dB(A) at a frontal direction of 0.5 m from the loudspeaker units. Measured impulse responses (sampling frequency: 48 kHz) were obtained by processing the recorded audio signals on a computer. The number of synchronous repetitions was 16. In order to reduce the effect of reflections of the measured room, the length of measured impulse responses was truncated at an initial 256 samples.

Secondly, audio signals were synthesized by convolving measured impulse responses to a sound source signal on a computer. The sound source signal was three octave-band noises, of which the central frequency is 250, 500, and 1000 Hz. The length, sampling frequency and quantization bit of the signal were respectively 1 s, 48 kHz, and 16 bits.

Finally, $D_i(\theta, \phi)$ (measured radiation directivity of the ith loudspeaker unit) is calculated from audio signals as follows:

$$D_i(\theta,\phi) = 20\log_{10}\frac{d_{i,0,0}}{d_{i,\theta,\phi}}\sqrt{\frac{\int_0^T \{s_{i,\theta,\phi}(t)\}^2 dt}{\int_0^T \{s_{i,0,0}(t)\}^2 dt}}, \qquad (7)$$

where T(=1 s) is the time length, and $s_{i,\theta,\phi}(t)$ and $d_{i,\theta,\phi}$ denote the synthesized audio signal and the distance between the ith loudspeaker unit and microphones when the azimuth and elevation angles of the radiation direction in the ith loudspeaker unit is θ and ϕ. The θ and ϕ are defined as shown in Fig. 10.

The results of the measured radiation directivity in five typical loudspeaker units are shown in Fig. 11. In all units, the radiation directivity in the backward direction decays more than 20 dB compared with that in the frontal direction.

4. Performance Evaluation

Although it is most common to evaluate the performance of the developed system by the time-variant of sound pressures at several thousands of observation points, it spends a lot of times for measuring the time-variant of sound pressures at several thousands of observation points. In this paper, in order to reduce the measuring time, sound pressures and sound intensity vectors are measured at seventeen observation points. Since the direction of sound intensity vectors corresponds to the arrival direction of sound sources when the number of sound sources is one [12], the performance of the developed system can be evaluated by estimating the po-

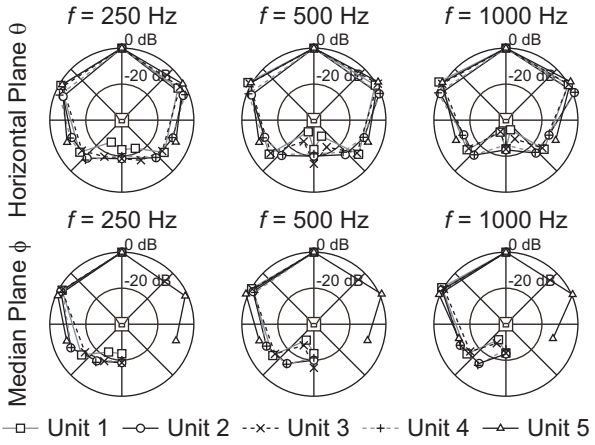

Fig. 11 Results of measured radiation directivity in five typical loudspeaker units of the radiated loudspeaker array.

sitions of sound images based on sound pressures and sound intensity vectors at seventeen observation points.

4.1. Procedure

As shown on the upper part of Fig. 12, 30 sound sources were placed in the surrounding microphone array in the original sound field. The original sound field was a free field with no reflection sounds. On the other hand, as shown in the lower part of Fig. 12, the radiated loudspeaker array was constructed in the reproduced sound field. The reproduced sound field was a free field in order to simplify the evaluation. Seventeen observation points were placed outside the radiated loudspeaker array.

The sound source signal $s(t)$ was an octave-band noise with central frequency f_{cent}. The $x_i(t)$ (microphone signal of the ith microphone) was denoted as follows from the sound source signal $s(t)$ and the impulse response between the sound source and the microphone $g_{i0}(t)$:

$$x_i(t) = g_{i0}(t) * s(t) = \frac{1}{|\mathbf{r}_i - \mathbf{r}_0|} s\left(t - \frac{|\mathbf{r}_i - \mathbf{r}_0|}{c}\right), \tag{8}$$

where $*$ denotes the convolution operation, $\mathbf{r}_i$ and $\mathbf{r}_0$ are the position vectors of the ith microphone and the sound source, respectively, and c is the sound velocity. The $p(\mathbf{R}_l, f_{\mathrm{cent}}, t)$ (instantaneous sound pressure of the lth observation point $\mathbf{R}_l$) was calculated from $x_i(t)$ as follows:

$$\begin{aligned} p(\mathbf{R}_l, f_{\mathrm{cent}}, t) &= \sum_{i=1}^{M} \frac{D_{si}}{|\mathbf{R}_l - \mathbf{r}_i|} x_i\left(t - \frac{|\mathbf{R}_l - \mathbf{r}_i|}{c}\right) \\ &= \sum_{i=1}^{M} \frac{D_{si}}{|\mathbf{R}_l - \mathbf{r}_i||\mathbf{r}_i - \mathbf{r}_0|} s\left(t - \frac{|\mathbf{R}_l - \mathbf{r}_i| + |\mathbf{r}_i - \mathbf{r}_0|}{c}\right), \end{aligned} \tag{9}$$

where M is the total number of loudspeaker units, and D_{si} is the radiation directivity of the ith loudspeaker unit.

The sound intensity vectors were calculated using the cross-spectral method, as shown in Fig. 13. Note that

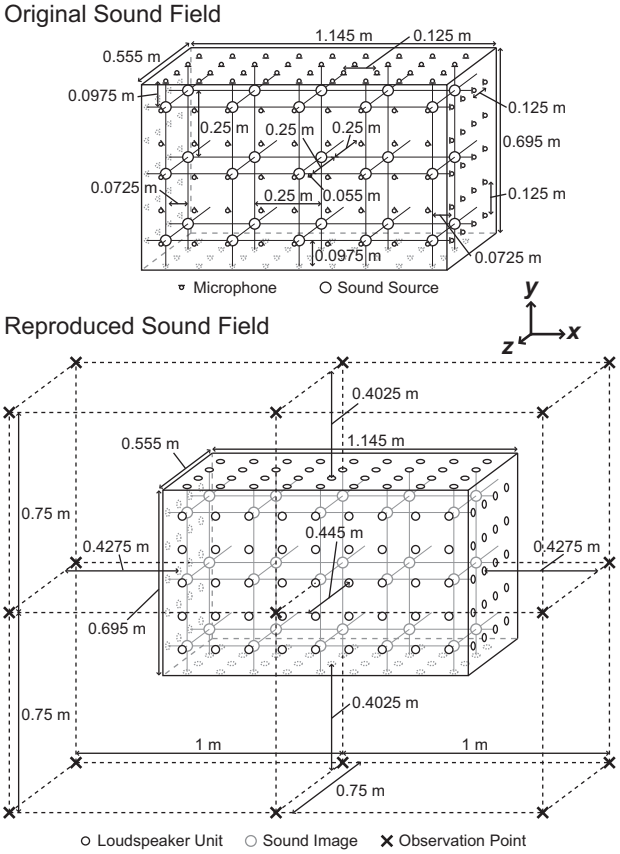

Fig. 12 Positions of the sound sources, sound images, microphones, loudspeaker units, and observation points used in the performance evaluation.

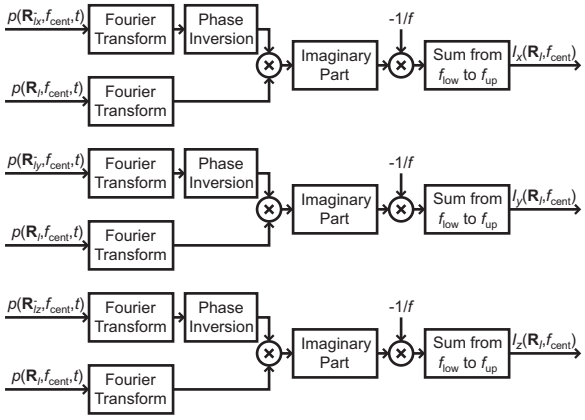

Fig. 13 Block diagram showing the calculation of sound intensities in the computer simulation.

$I_x(\mathbf{R}_l, f_{\mathrm{cent}})$, $I_y(\mathbf{R}_l, f_{\mathrm{cent}})$ and $I_z(\mathbf{R}_l, f_{\mathrm{cent}})$ in Fig. 13 are the x, y and z components of the sound intensity vectors $\mathbf{I}(\mathbf{R}_l, f_{\mathrm{cent}})$, and $p(\mathbf{R}_{lx}^-, f_{\mathrm{cent}}, t)$, $p(\mathbf{R}_{ly}^-, f_{\mathrm{cent}}, t)$, and $p(\mathbf{R}_{lz}^-, f_{\mathrm{cent}}, t)$ in Fig. 13 are the instantaneous sound pressure at three points ($\mathbf{R}_{lx}^-$, $\mathbf{R}_{ly}^-$ and $\mathbf{R}_{lz}^-$). The position vectors of the three points were set as follows:

$$\mathbf{R}_{lx}^- = \mathbf{R}_l - (\Delta, 0, 0)^T, \tag{10}$$

$$\mathbf{R}_{ly}^- = \mathbf{R}_l - (0, \Delta, 0)^T, \tag{11}$$

$$\mathbf{R}_{lz}^{-} = \mathbf{R}_l - (0, 0, \Delta)^T, \tag{12}$$

where Δ is 0.01 m.

The sound image positions $\mathbf{r}_E = (r_{Ex}, r_{Ey}, r_{Ez})^T$ were estimated from the calculated instantaneous sound pressures $p(\mathbf{R}_l, f_{\text{cent}}, t)$ and sound intensity vectors $\mathbf{I}(\mathbf{R}_l, f_{\text{cent}})$ according to following equation:

$$\mathbf{r}_E = \frac{1}{FN} \sum_{f_{\text{cent}}} \sum_{l=1}^{N} \left\{ \mathbf{R}_l - \alpha \frac{\mathbf{i}(\mathbf{R}_l, f_{\text{cent}})}{p(\mathbf{R}_l, f_{\text{cent}})} \right\}, \tag{13}$$

where $\mathbf{i}(\mathbf{R}_l, f_{\text{cent}})(= \frac{\mathbf{I}(\mathbf{R}_l, f_{\text{cent}})}{|\mathbf{I}(\mathbf{R}_l, f_{\text{cent}})|})$ denotes the direction of sound intensity vectors and $p(\mathbf{R}_l, f_{\text{cent}})$ is the RMS of the sound pressures calculated from the instantaneous sound pressure $p(\mathbf{R}_l, f_{\text{cent}}, t)$ in $\mathbf{R}_l$ according to following equation:

$$p(\mathbf{R}_l, f_{\text{cent}}) = \sqrt{\frac{1}{T} \int_0^T \{p(\mathbf{R}_l, f_{\text{cent}}, t)\}^2 dt}, \tag{14}$$

where $T(=1$ s) is the time length. F and N denote the number of octave-band noises and observation points used in the estimation of the sound image positions, respectively. α (coefficient for minimizing the error of the estimated positions of sound images) is derived as following equation (see Appendix):

$$\alpha = \frac{\displaystyle\sum_{f_{\text{cent}}} \sum_{l=1}^{N} \left\{ \frac{\mathbf{i}_{\text{avg}}}{p_{\text{avg}}} - \frac{\mathbf{i}(\mathbf{R}_l, f_{\text{cent}})}{p(\mathbf{R}_l, f_{\text{cent}})} \right\}^T (\mathbf{R}_{\text{avg}} - \mathbf{R}_l)}{\displaystyle\sum_{f_{\text{cent}}} \sum_{l=1}^{N} \left| \frac{\mathbf{i}_{\text{avg}}}{p_{\text{avg}}} - \frac{\mathbf{i}(\mathbf{R}_l, f_{\text{cent}})}{p(\mathbf{R}_l, f_{\text{cent}})} \right|^2}, \tag{15}$$

where $\frac{\mathbf{i}_{\text{avg}}}{p_{\text{avg}}}$ and $\mathbf{R}_{\text{avg}}$ are defined as follows:

$$\frac{\mathbf{i}_{\text{avg}}}{p_{\text{avg}}} = \frac{1}{FN} \sum_{f_{\text{cent}}} \sum_{l=1}^{N} \frac{\mathbf{i}(\mathbf{R}_l, f_{\text{cent}})}{p(\mathbf{R}_l, f_{\text{cent}})}, \tag{16}$$

$$\mathbf{R}_{\text{avg}} = \frac{1}{FN} \sum_{f_{\text{cent}}} \sum_{l=1}^{N} \mathbf{R}_l. \tag{17}$$

4.2. Condition

Parametric conditions are shown in Table 1. The $\mathbf{r}_0 = (r_{0x}, r_{0y}, r_{0z})^T$ and $\mathbf{R}_l = (R_{lx}, R_{ly}, R_{lz})^T$ (position vector of the sound sources and the observation points) were set in a three-dimensional coordinate as shown in Tables 2–3. The $\mathbf{r}_i$ (position vector of microphones and loudspeaker units) was set in a three-dimensional coordinate as follows:

$$\mathbf{r}_i(i = 1\sim40) = \begin{pmatrix} 0.5725\times(-1)^{i-1} \\ 0.125\times Q(\text{fix}(\frac{i-1}{2}), 4) - 0.25 \\ 0.125\times R(\text{fix}(\frac{i-1}{2}), 4) - 0.125 \end{pmatrix}, \tag{18}$$

$$\mathbf{r}_i(i = 41\sim112) = \begin{pmatrix} 0.125\times Q(\text{fix}(\frac{i-41}{2}), 4) - 0.5 \\ 0.3475\times(-1)^{i-41} \\ 0.125\times R(\text{fix}(\frac{i-41}{2}), 4) - 0.125 \end{pmatrix}, \tag{19}$$

$$\mathbf{r}_i(i = 113\sim157) = \begin{pmatrix} 0.125\times Q(i - 113, 5) - 0.5 \\ 0.125\times R(i - 113, 5) - 0.25 \\ 0.305 \end{pmatrix}, \tag{20}$$

Table 1 Parametric conditions in the performance evaluation.

Number of observation points (N)	17
Number of octave-band noises (F)	3
Central frequency of octave-band noises (f_{cent})	250, 500, 1000 Hz
Lower frequency of octave-band noises (f_{low})	$f_{\text{cent}} \div \sqrt{2}$
Upper frequency of octave-band noises (f_{up})	$f_{\text{cent}} \times \sqrt{2}$
Sound velocity (c)	340 m/s
Number of microphones and loudspeaker units (M)	157
Radiation directivity of loudspeaker units (D_{si})	Omnidirectional, Decay 20 dB, Unidirectional, Shotgun, Real

Table 2 Position coordinates of sound sources in the performance evaluation.

Index	r_{0x}	r_{0y}	r_{0z}	Index	r_{0x}	r_{0y}	r_{0z}
1	0.5	0.25	0.25	16	0.5	0.25	0
2	0.5	0	0.25	17	0.5	0	0
3	0.5	-0.25	0.25	18	0.5	-0.25	0
4	0.25	0.25	0.25	19	0.25	0.25	0
5	0.25	0	0.25	20	0.25	0	0
6	0.25	-0.25	0.25	21	0.25	-0.25	0
7	0	0.25	0.25	22	0	0.25	0
8	0	0	0.25	23	0	0	0
9	0	-0.25	0.25	24	0	-0.25	0
10	-0.25	0.25	0.25	25	-0.25	0.25	0
11	-0.25	0	0.25	26	-0.25	0	0
12	-0.25	-0.25	0.25	27	-0.25	-0.25	0
13	-0.5	0.25	0.25	28	-0.5	0.25	0
14	-0.5	0	0.25	29	-0.5	0	0
15	-0.5	-0.25	0.25	30	-0.5	-0.25	0

where $Q(u, v)$ and $R(u, v)$ denote the quotient and remainder when u is divided by v.

Five conditions were set in the radiation directivity of loudspeaker units D_{si}. In the four conditions ("Omnidirectional," "Decay 20 dB," "Unidirectional" and "Shotgun"), sound pressures are calculated according to Eq. (9) by using simulated radiation directivity patterns shown in Fig. 14. Under the "Decay 20dB" condition, the radiation directivity of the loudspeaker units is simulated on the basis of the results of acoustical measurements performed in Section 3.3..

Table 3 Position coordinates of observation points in the performance evaluation.

l	R_{lx}	R_{ly}	R_{lz}	l	R_{lx}	R_{ly}	R_{lz}	l	R_{lx}	R_{ly}	R_{lz}
1	1	0.75	0.75	7	-1	0.75	0.75	13	0	0.75	0
2	1	0	0.75	8	-1	0	0.75	14	0	-0.75	0
3	1	-0.75	0.75	9	-1	-0.75	0.75	15	-1	0.75	0
4	0	0.75	0.75	10	1	0.75	0	16	-1	0	0
5	0	0	0.75	11	1	0	0	17	-1	-0.75	0
6	0	-0.75	0.75	12	1	-0.75	0				

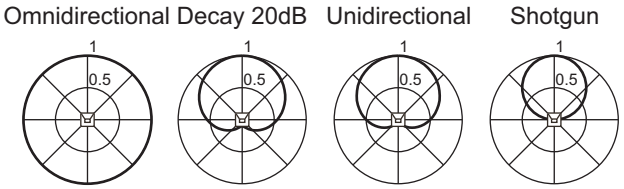

Fig. 14 Simulated radiation directivity patterns of the loudspeaker units used in the performance evaluation.

Definitional equations are denoted as follows:

$$\text{(Omnidirectional)} \quad D_{si} = 1, \tag{21}$$

$$\text{(Decay 20dB)} \quad D_{si} = 0.55 + 0.45\cos\theta_{si}, \tag{22}$$

$$\text{(Unidirectional)} \quad D_{si} = \frac{1+\cos\theta_{si}}{2}, \tag{23}$$

$$\text{(Shotgun)} \quad D_{si} = \begin{cases} \cos\theta_{si} & (|\theta_{si}| \le 90^\circ) \\ 0 & (|\theta_{si}| > 90^\circ) \end{cases}, \tag{24}$$

where $\cos\theta_{si} = \frac{\mathbf{n}_{si}\cdot(\mathbf{R}_j - \mathbf{r}_i)}{|\mathbf{n}_{si}||\mathbf{R}_j - \mathbf{r}_i|}$ and $\mathbf{n}_{si}$ (the directional vector of the ith loudspeaker unit) is defined as follows:

$$\mathbf{n}_{si} = \begin{cases} ((-1)^{i-1}, 0, 0)^T & (i = 1\sim40) \\ (0, (-1)^{i-41}, 0)^T & (i = 41\sim112) \\ (0, 0, 1)^T & (i = 113\sim157) \end{cases}. \tag{25}$$

In the "Decay 20 dB" and "Unidirectional" conditions, the loudspeaker units do not radiate a sound to the behind direction. In the "Shotgun" condition, the loudspeaker units do not radiate a sound to the lateral and behind direction. Thus, it is indicated that the radiation directivity of the loudspeaker units is sharper in order of "Shotgun," "Unidirectional," "Decay 20 dB" and "Omnidirectional".

On the other hand, in the "Real" condition, sound pressures are calculated by using measured impulse responses as follows:

$$p(\mathbf{R}_l, f_{\text{cent}}, t) = \sum_{i=1}^{M} g_{li}(t) * x_i(t), \tag{26}$$

where $g_{li}(t)$ denotes the measured impulse response between loudspeaker units and observation points. Measurement conditions are the same as those performed in Section 3.3..

In order to evaluate the performance of the sound image position estimation method, the sound source positions were directly estimated in the original sound field. The instantaneous sound pressure $p_0(\mathbf{R}_j, f_{\text{cent}}, t)$ at the jth observation point $\mathbf{R}_j$ in the original sound field is denoted as follows:

$$p_0(\mathbf{R}_l, f_{\text{cent}}, t) = \frac{1}{|\mathbf{R}_l - \mathbf{r}_0|} s\left(t - \frac{|\mathbf{R}_l - \mathbf{r}_0|}{c}\right). \tag{27}$$

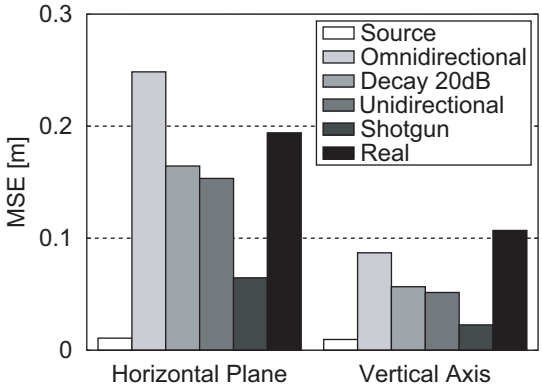

Fig. 15 Results of mean square errors in the performance evaluation.

4.3. Results

In order to compare the errors of estimated positions with the difference limen of the perceived direction in past studies, two types of mean square errors (MSEs) between input sound image positions $\mathbf{r}_0$ and estimated sound image positions $\mathbf{r}_E$ (MSE in the horizontal plane and the the vertical axis) are calculated according to following equations:

(Horizontal plane)

$$\text{MSE} = \sqrt{\frac{1}{30}\sum_{\mathbf{r}_0}\{(r_{Ex} - r_{0x})^2 + (r_{Ey} - r_{0y})^2\}}, \tag{28}$$

(Vertical axis)

$$\text{MSE} = \sqrt{\frac{1}{30}\sum_{\mathbf{r}_0}(r_{Ez} - r_{0z})^2}. \tag{29}$$

MSE in the horizontal plane corresponds to the difference limen of the left, right, up and down directions when listeners listen to a sound from the upper side of the radiated loudspeaker array. MSE in the the vertical axis corresponds to the difference limen of the up and down directions when listeners listen to a sound from the lateral side of the radiated loudspeaker array.

The results of MSEs in the horizontal plane and the vertical axis are shown in Fig. 15. The "Source" condition denotes that the sound source positions were directly estimated in the original sound field. In both types of horizontal plane and vertical axis, the MSEs in the "Source" condition are sufficiently small (0.011 m and 0.010 m). Thus, the sound image position estimation method used in this paper can accurately estimate the position of sound images.

In the five conditions ("Omnidirectional," "Decay 20 dB," "Unidirectional," "Shotgun" and "Real"), the value of MSEs in the vertical axis is smaller than that of MSEs in the horizontal plane. Therefore, it is indicated that the errors of sound image position estimation in the developed system mainly arise in the horizontal plane.

In the four conditions ("Omnidirectional," "Decay 20 dB," "Unidirectional" and "Shotgun"), the MSEs are small if the radiation directivity of loudspeaker units is sharper. Particularly, the values of MSEs are sufficiently small in the "Shotgun" condition (0.065 m and 0.023 m). This indicates that the 3D sound field is accurately controlled outside of the radiated loudspeaker array by sharpening the radiation directivity of the loudspeaker units.

Based on the theoretical consideration described in Section 2, the best radiation directivity of loudspeaker units is "Unidirectional" condition. However, because the loudspeaker units do not radiate a sound to the lateral direction in the "Shotgun" condition while the loudspeaker units radiate a sound to the lateral direction in the "Unidirectional" condition, the MSEs of the "Shotgun" condition is smaller than those of the "Unidirectional" condition. This fact is the effectiveness of the performance evaluation, which is not derived from the theoretical consideration.

The MSE of the "Real" condition is 0.107 m in the vertical axis. This value is more than that of the "Omnidirectional" condition (0.087 m). This value corresponds to the angular error of 12.1° (= $\tan^{-1}(\frac{0.107}{0.5})$). Note that 0.5 m denotes the smallest distance between sound sources and observation points. This angular error value is less than that of the difference limen of the general audience in the front direction in the ventriloquism effect (20° [13]). Thus, it is considered that the general audience can not discriminate the angular difference of the sound images in the vertical direction if the audio-visual system using the proposed system is constructed.

On the other hand, the MSE of the "Real" condition is 0.194 m in the horizontal plane. This value is close to that of the "Decay 20 dB" condition (0.164 m). This is because the radiation directivity of loudspeaker units in the developed system is near "Decay 20 dB" condition described in Section 3.3.. Thus, it is expected that the MSE become smaller if the radiation directivity of loudspeaker units is sharper in the developed system.

5. Conclusion

In this paper, 3D radiated sound field display systems using directional loudspeakers and wave field synthesis were proposed as methods of achieving ultra-realistic communications in applications such as 3D television. The principle was derived from the Fresnel-Kirchhoff diffraction formula, which is the approximate form of the Kirchhoff-Helmholtz integral equation. The proposed system was developed by constructing the surrounding microphone array and radiated loudspeaker array. In order to evaluate the performance of the developed system, the positions of sound images were estimated when sound sources were octave-band noises. The results indicated that the 3D sound field is accurately controlled outside of the radiated loudspeaker array by sharpening the radiation directivity of the loudspeaker units. However, the sound images were not correctly estimated in the developed system because the radiation directivity of loudspeaker units was not sufficiently sharp.

A subject for future work is development of a system with better performance in the developed system by sharpening the radiation directivity of the loudspeaker units. On the other hand, a cubic 3D visual display system, in which several people can view an object anywhere in its vicinity without having to wear equipment such as glasses, has also been developed [14]. Therefore development, by combining the proposed system and the system described above, is needed for a 3D visual and audio system in which several people can view an object and listen to a sound anywhere in its vicinity without having to wear equipment such as glasses and headphones.

REFERENCES

[1] K. Enami, "Research on ultra-realistic communications," *ECTI Trans. Elect. Eng., Electro., Comm.*, 6(1), 22–25 (2008).

[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 on accuracy of synthesized wave fronts in sound field reproduction based on wave field synthesis," *Papers of AES 13th Regional Conv.*, 0037, pp. 1–8 (2007).

[5] B. B. Baker and E. T. Copson, *The Mathematical Theory of Huygens' Principle* (Oxford University Press, London, 1950), pp. 23–26.

[6] R. Avizienis, A. Freed, P. Kassakian and D. Wesel, "A compact 120 independent element spherical loudspeaker array with programable radiation patterns," *Papers of AES Conv.*, 6783, pp. 1–7 (2006).

[7] A. Schmeder, "An exploration of design parameters for human-interactive systems with compact spherical loudspeaker arrays," *Proc. Ambisonics Symp.*, 10, pp. 1–11 (2009).

[8] B. B. Baker and E. T. Copson, *The Mathematical Theory of Huygens' Principle* (Oxford University Press, London, 1950), pp. 72–74.

[9] T. Okamoto, R. Nishimura and Y. Iwaya, "Estimation of sound source positions using a surrounding microphone array," *Acoust. Sci. & Tech.*, **28**, 181–189 (2007).

[10] 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," *Proc. WESPAC*, 0164, pp. 1–6 (2009).

[11] Y. Suzuki, F. Asano, H. Y. Kim and T. Sone, "An optimum computer-generated pulse signal suitable for the measurement of very long impulse responses," *J. Acoust. Soc. Am.*, **97**, 1119–1123 (1995).

[12] F. J. Fahy, *Sound Intensity* (Spon Press, UK, 1995).

[13] S. Komiyama, "Subjective Evaluation of Angular Displacement between Picture and Sound Directions for HDTV Sound Systems," *J. Audio Eng. Soc.*, **37**, 210–214 (1989).

[14] R. Lopez-Gulliver, S. Yoshida, M. Makino, S. Yano and H. Ando, "gCubik+i Virtual 3D Aquarium: Natural Interface between a Graspable 3D Display and a Tabletop Display," *Trans. Virtual Reality Soc. Jpn.*, **15**, 147–156 (2010).

Appendix Derivation of α in Eq. (13)

The estimated position of sound images in each observation point is defined as follows:

$$\mathbf{r}(\mathbf{R}_l, f_{\text{cent}}) = \mathbf{R}_l - \alpha \frac{\mathbf{i}(\mathbf{R}_l, f_{\text{cent}})}{p(\mathbf{R}_l, f_{\text{cent}})}. \tag{A·1}$$

Thus, the squared error of the estimated positions of sound images σ^2 is calculated as follows:

$$\begin{aligned}
\sigma^2 &= \sum_{f_{\text{cent}}} \sum_{l=1}^{N} |\mathbf{r}(\mathbf{R}_l, f_{\text{cent}}) - \mathbf{r}_E|^2 \\
&= \sum_{f_{\text{cent}}} \sum_{l=1}^{N} \left| \mathbf{R}_l - \alpha \frac{\mathbf{i}(\mathbf{R}_l, f_{\text{cent}})}{p(\mathbf{R}_l, f_{\text{cent}})} \right. \\
&\qquad \left. - \frac{1}{FN} \sum_{f'_{\text{cent}}} \sum_{l'=1}^{N} \left\{ \mathbf{R}_{l'} - \alpha \frac{\mathbf{i}(\mathbf{R}_{l'}, f'_{\text{cent}})}{p(\mathbf{R}_{l'}, f'_{\text{cent}})} \right\} \right|^2 \\
&= \sum_{f_{\text{cent}}} \sum_{l=1}^{N} \left| \left\{ \frac{\mathbf{i}_{\text{avg}}}{p_{\text{avg}}} - \frac{\mathbf{i}(\mathbf{R}_l, f_{\text{cent}})}{p(\mathbf{R}_l, f_{\text{cent}})} \right\} \alpha - (\mathbf{R}_{\text{avg}} - \mathbf{R}_l) \right|^2 \\
&= \sum_{f_{\text{cent}}} \sum_{l=1}^{N} \left[\left| \frac{\mathbf{i}_{\text{avg}}}{p_{\text{avg}}} - \frac{\mathbf{i}(\mathbf{R}_l, f_{\text{cent}})}{p(\mathbf{R}_l, f_{\text{cent}})} \right|^2 \alpha^2 \right. \\
&\qquad - 2 \left\{ \frac{\mathbf{i}_{\text{avg}}}{p_{\text{avg}}} - \frac{\mathbf{i}(\mathbf{R}_l, f_{\text{cent}})}{p(\mathbf{R}_l, f_{\text{cent}})} \right\}^T (\mathbf{R}_{\text{avg}} - \mathbf{R}_l) \alpha \\
&\qquad \left. + |\mathbf{R}_{\text{avg}} - \mathbf{R}_l|^2 \right],
\end{aligned} \tag{A·2}$$

where

$$\frac{\mathbf{i}_{\text{avg}}}{p_{\text{avg}}} = \frac{1}{FN} \sum_{f'_{\text{cent}}} \sum_{l'=1}^{N} \frac{\mathbf{i}(\mathbf{R}_{l'}, f'_{\text{cent}})}{p(\mathbf{R}_{l'}, f'_{\text{cent}})}, \tag{A·3}$$

$$\mathbf{R}_{\text{avg}} = \frac{1}{FN} \sum_{f'_{\text{cent}}} \sum_{l'=1}^{N} \mathbf{R}_{l'}. \tag{A·4}$$

The differential of the squared error is calculated as follows:

$$\begin{aligned}
\frac{d\sigma^2}{d\alpha} = \sum_{f_{\text{cent}}} \sum_{l=1}^{N} \left[2 \left| \frac{\mathbf{i}_{\text{avg}}}{p_{\text{avg}}} - \frac{\mathbf{i}(\mathbf{R}_l, f_{\text{cent}})}{p(\mathbf{R}_l, f_{\text{cent}})} \right|^2 \alpha \right. \\
\left. - 2 \left\{ \frac{\mathbf{i}_{\text{avg}}}{p_{\text{avg}}} - \frac{\mathbf{i}(\mathbf{R}_l, f_{\text{cent}})}{p(\mathbf{R}_l, f_{\text{cent}})} \right\}^T (\mathbf{R}_{\text{avg}} - \mathbf{R}_l) \right].
\end{aligned} \tag{A·5}$$

Because the squared error of the estimated positions of sound images is minimized when the differential is zero (i.e., $\frac{d\sigma^2}{d\alpha} = 0$), α is derived from Eq. (A·5) as follows:

$$\alpha = \frac{\displaystyle\sum_{f_{\text{cent}}} \sum_{l=1}^{N} \left\{ \frac{\mathbf{i}_{\text{avg}}}{p_{\text{avg}}} - \frac{\mathbf{i}(\mathbf{R}_l, f_{\text{cent}})}{p(\mathbf{R}_l, f_{\text{cent}})} \right\}^T (\mathbf{R}_{\text{avg}} - \mathbf{R}_l)}{\displaystyle\sum_{f_{\text{cent}}} \sum_{l=1}^{N} \left| \frac{\mathbf{i}_{\text{avg}}}{p_{\text{avg}}} - \frac{\mathbf{i}(\mathbf{R}_l, f_{\text{cent}})}{p(\mathbf{R}_l, f_{\text{cent}})} \right|^2}. \tag{A·6}$$

Cover Letter for Acoustical Science and Technology

(1) **Title of paper**
Three-dimensional Radiated Sound Field Display System Using Directional Loudspeakers and Wave Field Synthesis

(2) **Full name(s) of author(s)**
Toshiyuki Kimura[1*], Yoko Yamakata[2], Michiaki Katsumoto[1], Takuma Okamoto[3], Satoshi Yairi[4], Yukio Iwaya[3] and Yôiti Suzuki[3]

(3) **Affiliation(s)**
[1]National Institute of Information and Communications Technology, 4–2–1, Nukui-kitamachi, Koganei, Tokyo, 184–8795 Japan
[2]Kyoto University, Yoshida-honmachi, Sakyo-ku, Kyoto, 606–8501 Japan
[3]Tohoku University, 2–1–1, Katahira, Aoba-ku, Sendai, 980–8577 Japan
[4]Sendai National College of Technology, 48, Nodayama, Medeshima-shiote, Natori, Miyagi, 981–1239 Japan

(4) **Approximately five keywords**
Radiated sound field, Sound field display, Directional loudspeaker, Wave field synthesis, Sound image position estimation

(5) **PACS number**
43.58.Fm, 43.58.Ta, 43.60.Dh, 43.60.Sx

(6) **Short running title**
3D Radiated Sound Field Display System

(7) **Category of article:**
PAPER

(8) **Mailing address**
〒 619–0288 京都府相楽郡精華町光台 2–2–2
独立行政法人 情報通信研究機構
ユニバーサルコミュニケーション研究所 多感覚・評価研究室
木村敏幸

(9) **Classification**
Electroacoustics

(10) **Number of pages**
TEXT: 10 FIGURES: A·0 TABLES: A·0