Theoretical Study and Numerical Analysis of Near 3D Sound Field Reproduction Based on Wave Field Synthesis

Toshiyuki Kimura[1], Yoko Yamakata[1], and Michiaki Katsumoto[1]

[1]*National Institute of Information and Communications Technology, Koganei, Tokyo, 184-8795, Japan*

Correspondence should be addressed to Toshiyuki Kimura (t-kimura@nict.go.jp)

ABSTRACT

The development of near 3D sound field reproduction techniques is an important part of achieving ultra-realistic communication for applications such as 3D television and 3D teleconferencing. In this paper, the principle of a near 3D sound field reproduction technique using wave field synthesis is defined from the Kirchhoff-Helmholtz integral equation and two methods – dipole control and directional point control – are proposed. The performance of the two methods was evaluated using computer simulation, which showed that the performance of the dipole control method is very well, while the performance of the directional point control method is enough if the directivity of the loudspeakers was unidirectional or shotgun.

1. INTRODUCTION

We have been investigating ultra-realistic communication techniques as shown in Fig. 1 [1]. If video and audio can be more realistically reproduced in a 3D space by applying 3D video and audio techniques, more realistic forms of communication (e.g. 3D television, 3D teleconferencing, etc.) will be possible than those currently provided by conventional video and audio techniques (HD video and 5.1 ch audio). 3D sound field reproduction techniques capable of providing the aural components of ultra-realistic communication include binaural [2, 3], transaural [4, 5, 6], stereo dipole [7], wave field synthesis [8, 9, 10], and boundary surface control techniques [11, 12, 13]. In this paper, we focus on wave field synthesis and propose a 3D sound field reproduction technique based on wave field synthesis.

Wave field synthesis is a 3D sound field reproduction technique for reproducing wave fronts of a control area in a listening area based on Huygens' principle. Microphones placed on the boundary of a control area record the original sound and loudspeakers placed on the boundary of the listening area then play the recorded sound. The position of the loudspeakers is the same as that of the microphones. Multiple listeners can listen to the sound anywhere in the listening area without having to wear a device such as headphones because this technique reproduces the sound field of a 3D space rather than the sound field of binaural positions.

Conventional wave field synthesis systems have been constructed based on the Kirchhoff-Helmholtz integral equation, which mathematically defines Huygens' principle [14]. However, in these systems, the loudspeakers are placed around the listeners and the sound field of the inside of the loudspeaker array is reproduced, as shown in Fig. 2(a). As a result, listeners cannot listen to sounds around the sound sources.

In the proposed system, however, because the loudspeakers are placed around the sound sources and the sound field of the outside of the loudspeaker array is reproduced, as shown in Fig. 2(b), listeners can listen to sounds around the sound sources. Thus, as shown in Figs. 2(a) and 2(b), the direction of the microphones and loudspeakers in the proposed system is the opposite to that in the conventional system. Although the condition for reproducing wave fronts has been theoretically studied in the conventional system [10], it was not theoretically studied in the proposed system. In addition, while we assume that the sound field reproduced in the proposed system is the near sound field, where the distance between the sound sources and listeners is less than one meter, the condition for reproducing wave fronts in the near sound field was not investigated in the proposed system.

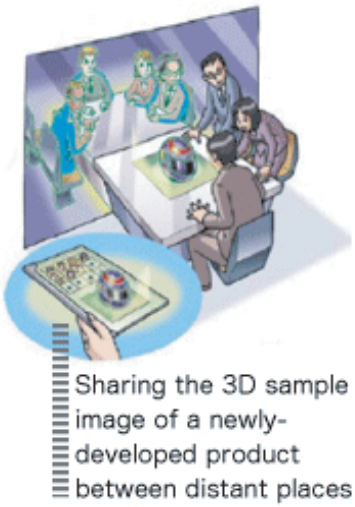

Fig. 1: Future image of ultra-realistic communication using 3D video and 3D audio [1].

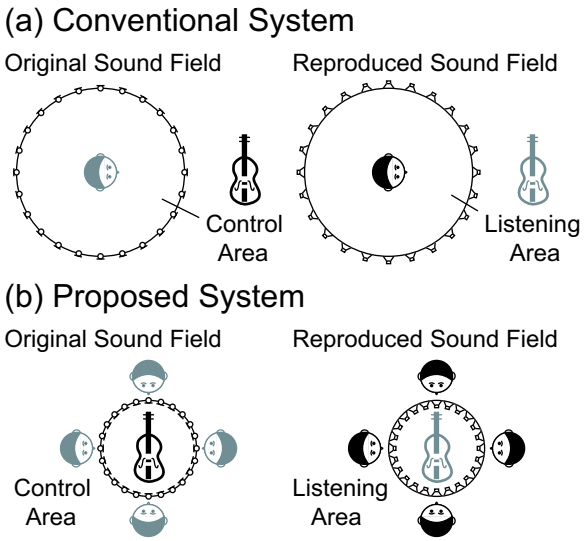

Fig. 2: Original sound field and reproduced sound field in conventional and proposed systems.

In this paper, near 3D sound field reproduction techniques that enable listeners to listen to sounds around the sound sources are proposed based on wave field synthesis. The principle of the near 3D sound field reproduction technique is defined on the basis of the Kirchhoff-Helmholtz integral equation and two methods – dipole control and directional point control – are proposed in Section 2. The performance of the two methods was evaluated using computer simulation in Section 3.

2. PRINCIPLE OF NEAR 3D SOUND FIELD REPRODUCTION TECHNIQUE

2.1. Kirchhoff-Helmholtz Integral Equation

As shown in the left of Fig. 3, if sound sources are surrounded by a continuous boundary surface S, and $\mathbf{r}_S$ and $\mathbf{r}$ are the position vectors on S and on space V (the outside of S), $P(\mathbf{r},\omega)$ (sound pressure at $\mathbf{r}$) is denoted as follows:

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

where k is a wave number, and $\mathbf{n}_S$ is the normal unit vector toward the outside of the continuous boundary surface at $\mathbf{r}_S$. Eq. (1) shows that the sound pressure of the space V is reproduced if the monopole sources (amplitude $\frac{\partial P(\mathbf{r}_S,\omega)}{\partial \mathbf{n}_S}$) and dipole sources (amplitude $-P(\mathbf{r}_S,\omega)$) are played at $\mathbf{r}_S$.

However, to construct the system the boundary surface S must be discretized since the monopole and dipole sources are not placed continuously on S. As shown in the right of Fig. 3, if $\mathbf{r}_i$ is the position vector on 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 in S_i, Eq. (1) can be converted as

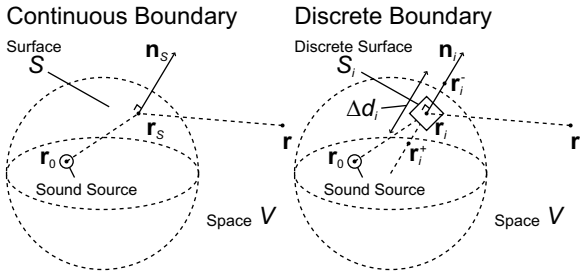

Fig. 3: Coordinates in Kirchhoff-Helmholtz integral equation.

follows:

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

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

2.2. Dipole Control Method

Sound pressure gradients $\frac{\partial P(\mathbf{r}_i,\omega)}{\partial \mathbf{n}_i}$ and sound pressures $P(\mathbf{r}_i,\omega)$ were approximated by the sound pressures at neighbor points as follows:

$$\frac{\partial P(\mathbf{r}_i,\omega)}{\partial \mathbf{n}_i} \approx \frac{P(\mathbf{r}_i^+,\omega) - P(\mathbf{r}_i^-,\omega)}{\Delta d_i}, \quad (3)$$

$$P(\mathbf{r}_i,\omega) \approx \frac{P(\mathbf{r}_i^+,\omega) + P(\mathbf{r}_i^-,\omega)}{2}, \quad (4)$$

where, as shown in the right of Fig. 3, $\mathbf{r}_i^+$ and $\mathbf{r}_i^-$ are the position vector of the neighboring points at the inside and outside of $\mathbf{r}_i$, and $\Delta d_i (= |\mathbf{r}_i^+ - \mathbf{r}_i^-|)$ is the distance between the neighboring points. The dipole and monopole sources were also approximated by two monopole sources at neighboring points as follows:

$$\frac{\partial}{\partial \mathbf{n}_i}\left(\frac{e^{-jk|\mathbf{r}-\mathbf{r}_i|}}{|\mathbf{r}-\mathbf{r}_i|}\right) \approx \frac{1}{\Delta d_i}\left(\frac{e^{-jk|\mathbf{r}-\mathbf{r}_i^+|}}{|\mathbf{r}-\mathbf{r}_i^+|} - \frac{e^{-jk|\mathbf{r}-\mathbf{r}_i^-|}}{|\mathbf{r}-\mathbf{r}_i^-|}\right), \quad (5)$$

$$\frac{e^{-jk|\mathbf{r}-\mathbf{r}_i|}}{|\mathbf{r}-\mathbf{r}_i|} \approx \frac{1}{2}\left(\frac{e^{-jk|\mathbf{r}-\mathbf{r}_i^+|}}{|\mathbf{r}-\mathbf{r}_i^+|} + \frac{e^{-jk|\mathbf{r}-\mathbf{r}_i^-|}}{|\mathbf{r}-\mathbf{r}_i^-|}\right). \quad (6)$$

If Eqs. (3)-(6) are substituted for Eq. (2), the following equation is derived:

$$P(\mathbf{r},\omega) = \frac{1}{4\pi}\sum_{i=1}^{M}\left\{P(\mathbf{r}_i^+,\omega)\frac{e^{-jk|\mathbf{r}-\mathbf{r}_i^-|}}{|\mathbf{r}-\mathbf{r}_i^-|} - P(\mathbf{r}_i^-,\omega)\frac{e^{-jk|\mathbf{r}-\mathbf{r}_i^+|}}{|\mathbf{r}-\mathbf{r}_i^+|}\right\}\frac{\Delta S_i}{\Delta d_i}. \quad (7)$$

Eq. (7) shows that the sound pressure of the space V is reproduced if the monopole sources (amplitude $P(\mathbf{r}_i^+,\omega)$) are played at M points (position $\mathbf{r}_i^-$) and the monopole sources (amplitude $-P(\mathbf{r}_i^-,\omega)$) are played at M points (position $\mathbf{r}_i^+$).

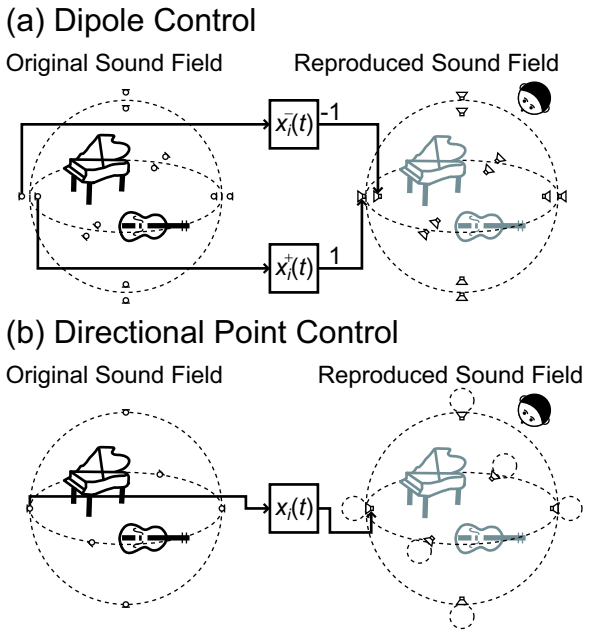

Fig. 4: Near 3D sound field reproduction system based on dipole control and directional point control methods.

A diagram of the near 3D sound field reproduction system based on Eq. (7) (called the "dipole control method") is shown in Fig. 4(a). Firstly, M pairs of omnidirectional microphones were placed on the boundary surface around the sound sources in the original sound field and audio signals ($x_i^+(t)$ and $x_i^-(t)$) were recorded. Secondly, M pairs of omnidirectional loudspeakers were placed at the same position as the microphone pairs in the reproduced sound field and the processed audio signals ($x_i^+(t)$ and $-x_i^-(t)$) were played. As a result, since wave fronts were reproduced on the outside of the boundary surface, listeners on the outside of the boundary surface experienced the feeling that the sound sources were being played on the inside of the boundary surface.

2.3. Directional Point Control Method

The sound pressure gradients $\frac{\partial P(\mathbf{r}_i,\omega)}{\partial \mathbf{n}_i}$ were approximated

as follows:

$$\begin{aligned}\frac{\partial P(\mathbf{r}_i,\omega)}{\partial \mathbf{n}_i} &= \frac{\partial}{\partial \mathbf{n}_i}\left(\frac{Ae^{-jk|\mathbf{r}_i-\mathbf{r}_0|}}{|\mathbf{r}_i-\mathbf{r}_0|}\right)\\ &= -\frac{Ae^{-jk|\mathbf{r}_i-\mathbf{r}_0|}}{|\mathbf{r}_i-\mathbf{r}_0|}\left(\frac{1}{|\mathbf{r}_i-\mathbf{r}_0|}+jk\right)\cos\theta_{i0}\\ &= -P(\mathbf{r}_i,\omega)\left(\frac{1}{|\mathbf{r}_i-\mathbf{r}_0|}+jk\right)\cos\theta_{i0}\\ &\approx -jkP(\mathbf{r}_i,\omega)\cos\theta_{i0}\\ &\qquad\left(\text{if } k \gg \frac{1}{|\mathbf{r}_i-\mathbf{r}_0|}\right), \quad (8)\end{aligned}$$

where $\mathbf{r}_0$ and A are the position vector and amplitude of the sound sources in an original sound field, and $\cos\theta_{i0}(=\frac{\mathbf{n}_i\cdot(\mathbf{r}_i-\mathbf{r}_0)}{|\mathbf{n}_i||\mathbf{r}_i-\mathbf{r}_0|})$ denotes the angle between the vector $\mathbf{n}_i$ and the vector $\mathbf{r}_i-\mathbf{r}_0$. The dipole sources were also approximated as follows:

$$\begin{aligned}\frac{\partial}{\partial \mathbf{n}_i}\left(\frac{e^{-jk|\mathbf{r}-\mathbf{r}_i|}}{|\mathbf{r}-\mathbf{r}_i|}\right) &= -\frac{e^{-jk|\mathbf{r}-\mathbf{r}_i|}}{|\mathbf{r}-\mathbf{r}_i|}\left(\frac{1}{|\mathbf{r}-\mathbf{r}_i|}+jk\right)\cos\theta_i\\ &\approx -jk\frac{e^{-jk|\mathbf{r}-\mathbf{r}_i|}}{|\mathbf{r}-\mathbf{r}_i|}\cos\theta_i\\ &\qquad\left(\text{if } k \gg \frac{1}{|\mathbf{r}-\mathbf{r}_i|}\right), \quad (9)\end{aligned}$$

where $\cos\theta_i(=\frac{\mathbf{n}_i\cdot(\mathbf{r}_i-\mathbf{r})}{|\mathbf{n}_i||\mathbf{r}_i-\mathbf{r}|})$ denotes the angle between the vector $\mathbf{n}_i$ and the vector $\mathbf{r}_i-\mathbf{r}$. If Eqs. (8)-(9) are substituted for Eq. (2), the following equation is derived:

$$\begin{aligned}P(\mathbf{r},\omega) = \frac{jk}{4\pi}\sum_{i=1}^{M} P(\mathbf{r}_i,\omega)\frac{e^{-jk|\mathbf{r}-\mathbf{r}_i|}}{|\mathbf{r}-\mathbf{r}_i|}\\ (\cos\theta_i-\cos\theta_{i0})\Delta S_i \quad (10)\end{aligned}$$

This equation is known as the Fresnel-Kirchhoff diffraction formula [15]. In the proposed system, since the direction of vector $\mathbf{r}_i-\mathbf{r}_0$ is almost the same as that of vector $\mathbf{n}_i$, it can be approximated as $\cos\theta_{i0}\approx 1$. Thus, (10) can be converted as follows:

$$\begin{aligned}P(\mathbf{r},\omega) &\approx \frac{jk}{4\pi}\sum_{i=1}^{M} P(\mathbf{r}_i,\omega)\frac{e^{-jk|\mathbf{r}-\mathbf{r}_i|}}{|\mathbf{r}-\mathbf{r}_i|}(\cos\theta_i-1)\Delta S_i\\ &\approx \frac{jk}{4\pi}\sum_{i=1}^{M} P(\mathbf{r}_i,\omega)D_i\frac{e^{-jk|\mathbf{r}-\mathbf{r}_i|}}{|\mathbf{r}-\mathbf{r}_i|}\Delta S_i, \quad (11)\end{aligned}$$

where $D_i(=\cos\theta_i-1)$ corresponds to the directivity of the monopole sources placed at $\mathbf{r}_i$. Eq. (11) shows that the sound pressure of the space V is reproduced if the directional monopole sources (amplitude $P(\mathbf{r}_i,\omega)$) are played at M points (position $\mathbf{r}_i$).

A diagram of the near 3D sound field reproduction system based on Eq. (11) (called the "directional point control method") is shown in Fig. 4(b). Firstly, M omnidirectional microphones were placed on the boundary surface around the sound sources in the original sound field and audio signals ($x_i(t)$) were recorded. Secondly, M directional loudspeakers were placed at the same position as the microphones in the reproduced sound field and the recorded audio signals ($x_i(t)$) were played. The directivity of the loudspeakers was toward the outside of the boundary surface. As a result, because wave fronts were reproduced on the outside of the boundary surface, listeners on the outside of the boundary surface experienced the feeling that the sound sources were being played on the inside of the boundary surface.

The system based on the directional point control method requires half the number of microphones and loudspeakers compared with the system based on the dipole control method. However, the approximations used in Eqs. (8) and (9) ($k \gg \frac{1}{|\mathbf{r}_i-\mathbf{r}_0|}$ and $k \gg \frac{1}{|\mathbf{r}-\mathbf{r}_i|}$) are valid if the distances ($|\mathbf{r}_i-\mathbf{r}_0|$ and $|\mathbf{r}-\mathbf{r}_i|$) are very long. Thus, since it is hard to say that the approximations are valid when the distances are less than one meter such as the system shown in Fig. 4, the performance of the directional point control method may degrade compared with that of the dipole control method. In Section 3, Computer simulation is performed in order to evaluate whether the performance of the directional point control method is enough to construct the practical system.

3. COMPUTER SIMULATION

3.1. Simulation Environment

As shown in Fig. 5, 162 control points were placed on a sphere with a radius of 0.4 m and 162 synthesis points were placed on a sphere with a radius of 0.8 m. The position of 162 points corresponds to the vertex of a Class I Method 1 icosahedral geodesic dome following 4 frequencies [16].

The sound source signal $s(t)$ was a sine-wave signal with an amplitude of A and frequency of f as follows:

$$s(t) = A\sin 2\pi f t. \quad (12)$$

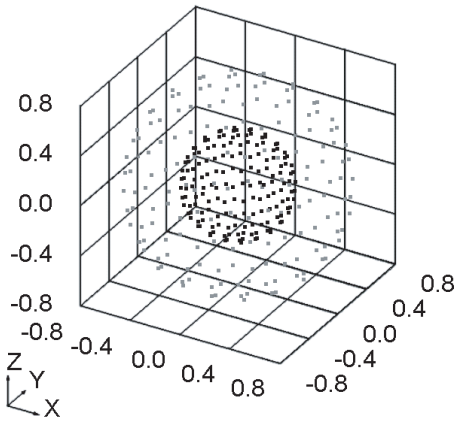

Fig. 5: Position of control points (black) and synthesis points (gray) used in computer simulation.

Let $\mathbf{r}$ be the position vector of a synthesis point. The $p_0(\mathbf{r}, f, t)$ (sound pressure at the synthesis point $\mathbf{r}$ in the original sound field) is denoted as follows:

$$p_0(\mathbf{r}, f, t) = \frac{1}{d_0} s\left(t - \frac{d_0}{c}\right) = \frac{A}{d_0} \sin\left\{2\pi f\left(t - \frac{d_0}{c}\right)\right\}, \quad (13)$$

where $d_0 (= |\mathbf{r} - \mathbf{r}_0|)$ is the distance between the sound source and the synthesis point, $\mathbf{r}_0$ is the position vector of the sound source, and c is the sound velocity.

In the case of the dipole control method, the $x_i^+(t)$ and $x_i^-(t)$ (recorded signals of the ith microphone pair) were denoted as follows:

$$x_i^+(t) = \frac{1}{d_{i0}^+} s\left(t - \frac{d_{i0}^+}{c}\right) = \frac{A}{d_{i0}^+} \sin\left\{2\pi f\left(t - \frac{d_{i0}^+}{c}\right)\right\}, \quad (14)$$

$$x_i^-(t) = \frac{1}{d_{i0}^-} s\left(t - \frac{d_{i0}^-}{c}\right) = \frac{A}{d_{i0}^-} \sin\left\{2\pi f\left(t - \frac{d_{i0}^-}{c}\right)\right\}, \quad (15)$$

where $d_{i0}^+ (= |\mathbf{r}_i^+ - \mathbf{r}_0|)$ and $d_{i0}^- (= |\mathbf{r}_i^- - \mathbf{r}_0|)$ are the distance between the sound source and the ith microphone pair, and $\mathbf{r}_i^+$ and $\mathbf{r}_i^-$ are the position vector of the ith microphone pair defined as follows:

$$\mathbf{r}_i^+ = \mathbf{r}_i - \frac{\Delta d_i}{2} \mathbf{n}_i, \quad (16)$$

$$\mathbf{r}_i^- = \mathbf{r}_i + \frac{\Delta d_i}{2} \mathbf{n}_i, \quad (17)$$

where $\mathbf{r}_i$ is the position vector of the ith control point, $\mathbf{n}_i$ is the normal unit vector toward the outside of the discrete boundary surface at $\mathbf{r}_i$, and $\Delta d_i (= |\mathbf{r}_i^+ - \mathbf{r}_i^-|)$ is the distance between the microphones. The $p(\mathbf{r}, f, t)$ (sound pressure of the synthesis point $\mathbf{r}$ in the reproduced sound field) was calculated from $x_i^+(t)$ and $x_i^-(t)$ as follows:

$$\begin{aligned} p(\mathbf{r}, f, t) &= \sum_{i=1}^{M} \left\{ \frac{1}{d_i^-} x_i^+\left(t - \frac{d_i^-}{c}\right) - \frac{1}{d_i^+} x_i^-\left(t - \frac{d_i^+}{c}\right) \right\} \\ &= \sum_{i=1}^{M} \left[\frac{A}{d_i^- d_{i0}^+} \sin\left\{2\pi f\left(t - \frac{d_i^- + d_{i0}^+}{c}\right)\right\} \right. \\ &\quad \left. - \frac{A}{d_i^+ d_{i0}^-} \sin\left\{2\pi f\left(t - \frac{d_i^+ + d_{i0}^-}{c}\right)\right\} \right], \quad (18) \end{aligned}$$

where M is the total number of loudspeaker pairs, and $d_i^+ (= |\mathbf{r} - \mathbf{r}_i^+|)$ and $d_i^- (= |\mathbf{r} - \mathbf{r}_i^-|)$ is the distance between the ith loudspeaker pair and the synthesis point.

In the case of the directional point control method, the $x_i(t)$ (recorded signal of the ith microphone) was denoted as follows:

$$x_i(t) = \frac{1}{d_{i0}} s\left(t - \frac{d_{i0}}{c}\right) = \frac{A}{d_{i0}} \sin\left\{2\pi f\left(t - \frac{d_{i0}}{c}\right)\right\}, \quad (19)$$

where $d_{i0} (= |\mathbf{r}_i - \mathbf{r}_0|)$ is the distance between the sound source and the ith microphone, and $\mathbf{r}_i$ is the position vector of the ith microphone. The $p(\mathbf{r}, f, t)$ (sound pressure of the synthesis point $\mathbf{r}$ in the reproduced sound field) was calculated from $x_i(t)$ as follows:

$$\begin{aligned} p(\mathbf{r}, f, t) &= \sum_{i=1}^{M} \frac{D_i}{d_i} x_i\left(t - \frac{d_i}{c}\right) \\ &= \sum_{i=1}^{M} \frac{D_i A}{d_i d_{i0}} \sin\left\{2\pi f\left(t - \frac{d_i + d_{i0}}{c}\right)\right\} \quad (20) \end{aligned}$$

where M is the total number of loudspeakers, $d_i (= |\mathbf{r} - \mathbf{r}_i|)$ is the distance between the ith loudspeaker and the synthesis point, and D_i is the directivity of the ith loudspeaker.

Since the arrival direction of sound sources at the synthesis point $\mathbf{r}$ isn't evaluated by the $p_0(\mathbf{r}, f, t)$ and $p(\mathbf{r}, f, t)$ (sound pressures of the synthesis point $\mathbf{r}$ in the original sound field and the reproduced sound field), the sound intensity vectors at the synthesis point $\mathbf{r}$ are also calculated. Since the direction of sound intensity vectors corresponds to the arrival direction of the sound sources [17], the arrival direction of sound sources at the synthesis point $\mathbf{r}$ can be evaluated by the sound intensity vectors.

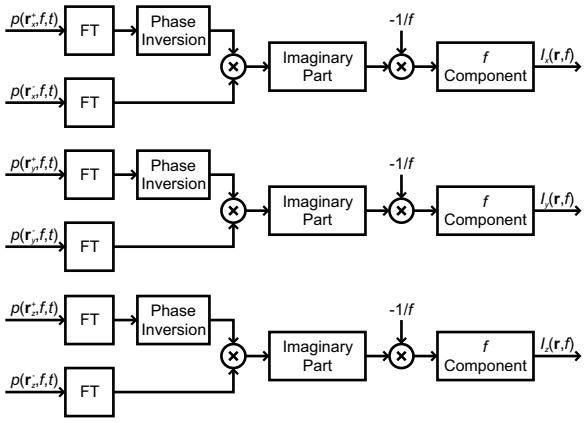

Fig. 6: Block diagram of calculation of sound intensities in computer simulation (reproduced sound field).

The sound intensity vectors were calculated using the cross-spectral method, as shown in Fig. 6 [17]. Note that $I_x(\mathbf{r},f)$, $I_y(\mathbf{r},f)$ and $I_z(\mathbf{r},f)$ in Fig. 6 are the x, y and z components of the sound intensity vectors $\mathbf{I}(\mathbf{r},f)$, and $p(\mathbf{r}_x^+,f,t)$, $p(\mathbf{r}_x^-,f,t)$, $p(\mathbf{r}_y^+,f,t)$, $p(\mathbf{r}_y^-,f,t)$, $p(\mathbf{r}_z^+,f,t)$ and $p(\mathbf{r}_z^-,f,t)$ in Fig. 6 are the sound pressure at six points ($\mathbf{r}_x^+$, $\mathbf{r}_x^-$, $\mathbf{r}_y^+$, $\mathbf{r}_y^-$, $\mathbf{r}_z^+$ and $\mathbf{r}_z^-$). The position vectors of the six points were set as follows:

$$\mathbf{r}_x^+ = \mathbf{r}+(\Delta,0,0)^T, \mathbf{r}_x^- = \mathbf{r}-(\Delta,0,0)^T, \tag{21}$$

$$\mathbf{r}_y^+ = \mathbf{r}+(0,\Delta,0)^T, \mathbf{r}_y^- = \mathbf{r}-(0,\Delta,0)^T, \tag{22}$$

$$\mathbf{r}_z^+ = \mathbf{r}+(0,0,\Delta)^T, \mathbf{r}_z^- = \mathbf{r}-(0,0,\Delta)^T, \tag{23}$$

where Δ is 0.001 m. The calculated sound intensity vector $\mathbf{I}(\mathbf{r},f) = \{I_x(\mathbf{r},f), I_y(\mathbf{r},f), I_z(\mathbf{r},f)\}^T$ was evaluated in section 3.2. The $\mathbf{I}_0(\mathbf{r},f) = \{I_{x0}(\mathbf{r},f), I_{y0}(\mathbf{r},f), I_{z0}(\mathbf{r},f)\}^T$ (sound intensity vector in the original sound field) was also calculated.

Parametric conditions are shown in Table 1. The $\mathbf{r}_i$ and $\mathbf{r}$ (position vector of the control points and synthesis points) were set in a three-dimensional coordinate as follows:

$$\mathbf{r}_i = \begin{pmatrix} r\cos\theta_i\cos\phi_i \\ r\sin\theta_i\cos\phi_i \\ r\sin\phi_i \end{pmatrix} \qquad (i = 1...M), \tag{24}$$

$$\mathbf{r} = \begin{pmatrix} R\cos\theta_j\cos\phi_j \\ R\sin\theta_j\cos\phi_j \\ R\sin\phi_j \end{pmatrix} \qquad (j = 1...N), \tag{25}$$

Amplitude of sound source (A)	1
Frequency of sound source (f)	125, 250, 500, 1000, 2000, 4000, 8000, 16000 Hz
Position vector of sound source ($\mathbf{r}_0$)	$(0,0,0)^T$ $(0.3,0,0)^T$ $(0,0.3,0)^T$ $(0,0,0.3)^T$
Sound velocity (c)	340 m/s
Total number of control points (M)	162
Radius of control points (r)	0.4 m
Total number of synthesis points (N)	162
Radius of synthesis points (R)	0.8 m
Normal unit vector ($\mathbf{n}_i$)	$\mathbf{r}_i/\lvert\mathbf{r}_i\rvert$
Neighbor distance of pairs (Δd_i)	0.002 m
Directivity of loudspeakers (D_i)	Omnidirectional, Bidirectional, Unidirectional, Shotgun

Table 1: Parametric conditions in computer simulation.

where θ_i and ϕ_i are the azimuth and elevation angle of the ith control point, and θ_j and ϕ_j are the azimuth and elevation angle of the jth synthesis point. The value of the azimuth and elevation angles of control points and synthesis points is shown in Table 2. The maximum interval of the control points is about 13 cm, which is less than half of the wavelength of 1000 Hz sound ($= \frac{340\,\mathrm{m}}{1000\,\mathrm{Hz}} = 34$ cm). Thus, the spatial sampling theorem for reproducing a wave front at sound frequencies under 1000 Hz is satisfied.

The D_i (directivity of the ith loudspeaker) is defined as follows:

$$(\text{Omnidirectional})D_i = 1, \tag{26}$$

$$(\text{Bidirectional})D_i = \cos\theta_{is}, \tag{27}$$

$$(\text{Unidierctional})D_i = \frac{1}{2}(1+\cos\theta_{is}), \tag{28}$$

$$(\text{Shotgun})D_i = \begin{cases} \cos\theta_{is} & (|\theta_{is}| \le 90°) \\ 0 & (|\theta_{is}| > 90°) \end{cases}, \tag{29}$$

where $\cos\theta_{is} = \frac{\mathbf{n}_i\cdot(\mathbf{r}-\mathbf{r}_i)}{|\mathbf{n}_i||\mathbf{r}-\mathbf{r}_i|}$. Directional patterns of loud-

i, j	θ_i, θ_j	ϕ_i, ϕ_j	i, j	θ_i, θ_j	ϕ_i, ϕ_j	i, j	θ_i, θ_j	ϕ_i, ϕ_j	i, j	θ_i, θ_j	ϕ_i, ϕ_j	i, j	θ_i, θ_j	ϕ_i, ϕ_j	i, j	θ_i, θ_j	ϕ_i, ϕ_j
1	0.00	90.00	28	-90.00	41.11	55	-44.27	16.05	82	18.00	0.00	109	63.73	-16.05	136	162.00	-41.11
2	-162.00	75.45	29	-18.00	41.11	56	8.27	16.05	83	36.00	0.00	110	116.27	-16.05	137	-175.61	-46.35
3	-90.00	75.45	30	54.00	41.11	57	27.73	16.05	84	54.00	0.00	111	135.73	-16.05	138	-148.39	-46.35
4	-18.00	75.45	31	126.00	41.11	58	80.27	16.05	85	72.00	0.00	112	-126.00	-26.57	139	-103.61	-46.35
5	54.00	75.45	32	-126.00	31.72	59	99.73	16.05	86	90.00	0.00	113	-54.00	-26.57	140	-76.39	-46.35
6	126.00	75.45	33	-54.00	31.72	60	152.27	16.05	87	108.00	0.00	114	18.00	-26.57	141	-31.61	-46.35
7	-126.00	63.43	34	18.00	31.72	61	171.73	16.05	88	126.00	0.00	115	90.00	-26.57	142	-4.39	-46.35
8	-54.00	63.43	35	90.00	31.72	62	-170.77	14.55	89	144.00	0.00	116	162.00	-26.57	143	40.39	-46.35
9	18.00	63.43	36	162.00	31.72	63	-153.23	14.55	90	162.00	0.00	117	-142.04	-30.15	144	67.61	-46.35
10	90.00	63.43	37	-178.04	30.15	64	-98.77	14.55	91	180.00	0.00	118	-109.96	-30.15	145	112.39	-46.35
11	162.00	63.43	38	-145.96	30.15	65	-81.23	14.55	92	-134.77	-14.55	119	-70.04	-30.15	146	139.61	-46.35
12	-162.00	58.28	39	-106.04	30.15	66	-26.77	14.55	93	-117.23	-14.55	120	-37.96	-30.15	147	-126.00	-58.28
13	-90.00	58.28	40	-73.96	30.15	67	-9.23	14.55	94	-62.77	-14.55	121	1.96	-30.15	148	-54.00	-58.28
14	-18.00	58.28	41	-34.04	30.15	68	45.23	14.55	95	-45.23	-14.55	122	34.04	-30.15	149	18.00	-58.28
15	54.00	58.28	42	-1.96	30.15	69	62.77	14.55	96	9.23	-14.55	123	73.96	-30.15	150	90.00	-58.28
16	126.00	58.28	43	37.96	30.15	70	117.23	14.55	97	26.77	-14.55	124	106.04	-30.15	151	162.00	-58.28
17	-139.61	46.35	44	70.04	30.15	71	134.77	14.55	98	81.23	-14.55	125	145.96	-30.15	152	-162.00	-63.43
18	-112.39	46.35	45	109.96	30.15	72	-162.00	0.00	99	98.77	-14.55	126	178.04	-30.15	153	-90.00	-63.43
19	-67.61	46.35	46	142.04	30.15	73	-144.00	0.00	100	153.23	-14.55	127	-162.00	-31.72	154	-18.00	-63.43
20	-40.39	46.35	47	-162.00	26.57	74	-126.00	0.00	101	170.77	-14.55	128	-90.00	-31.72	155	54.00	-63.43
21	4.39	46.35	48	-90.00	26.57	75	-108.00	0.00	102	-171.73	-16.05	129	-18.00	-31.72	156	126.00	-63.43
22	31.61	46.35	49	-18.00	26.57	76	-90.00	0.00	103	-152.27	-16.05	130	54.00	-31.72	157	-126.00	-75.45
23	76.39	46.35	50	54.00	26.57	77	-72.00	0.00	104	-99.73	-16.05	131	126.00	-31.72	158	-54.00	-75.45
24	103.61	46.35	51	126.00	26.57	78	-54.00	0.00	105	-80.27	-16.05	132	-126.00	-41.11	159	18.00	-75.45
25	148.39	46.35	52	-135.73	16.05	79	-36.00	0.00	106	-27.73	-16.05	133	-54.00	-41.11	160	90.00	-75.45
26	175.61	46.35	53	-116.27	16.05	80	-18.00	0.00	107	-8.27	-16.05	134	18.00	-41.11	161	162.00	-75.45
27	-162.00	41.11	54	-63.73	16.05	81	0.00	0.00	108	44.27	-16.05	135	90.00	-41.11	162	0.00	-90.00

Table 2: Azimuth and elevation angles of control points and synthesis points.

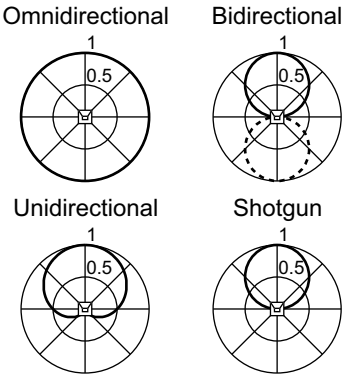

Fig. 7: Directional patterns of loudspeakers.

speakers are shown in Fig. 7. Note that the solid and dashed lines denote the positive and negative values of the directivity.

3.2. Simulation Results

The results of the simulation of the sound pressures and sound intensity vectors are shown in Figs. 8-11 when the frequency of the sound sources was 1000 Hz. The color of the arrows in Figs. 8-11 denotes the $p_0(\mathbf{r}, f)$ and $p(\mathbf{r}, f)$ (root mean square of the sound pressures in the original sound field and reproduced sound field) defined as follows:

$$p_0(\mathbf{r}, f) = \sqrt{\frac{1}{T}\int_0^T \{p_0(\mathbf{r}, f, t)\}^2 dt}, \tag{30}$$

$$p(\mathbf{r}, f) = \sqrt{\frac{1}{T}\int_0^T \{p(\mathbf{r}, f, t)\}^2 dt}, \tag{31}$$

where T(=1 sec) is the time length. It is indicated that the performance of the proposed methods in the sound pressure distribution is good if the color of the arrows in the proposed methods is the same as that of the arrows in the original sound field. Note that before plotting the $p_0(\mathbf{r}, f)$ and $p(\mathbf{r}, f)$ were converted to dB units as follows:

$$p_0(\mathbf{r}, f)[\mathrm{dB}] = 20\log_{10}\frac{p_0(\mathbf{r}, f)}{P_0(f)}, \tag{32}$$

$$p(\mathbf{r}, f)[\mathrm{dB}] = 20\log_{10}\frac{p(\mathbf{r}, f)}{P(f)}, \tag{33}$$

where the $P_0(f)(= \frac{1}{N}\sum_{\mathbf{r}} p_0(\mathbf{r}, f))$ and $P(f)(= \frac{1}{N}\sum_{\mathbf{r}} p(\mathbf{r}, f))$ are the values compensating the difference of the sound pressure distributions between the original sound field and the reproduced sound field.

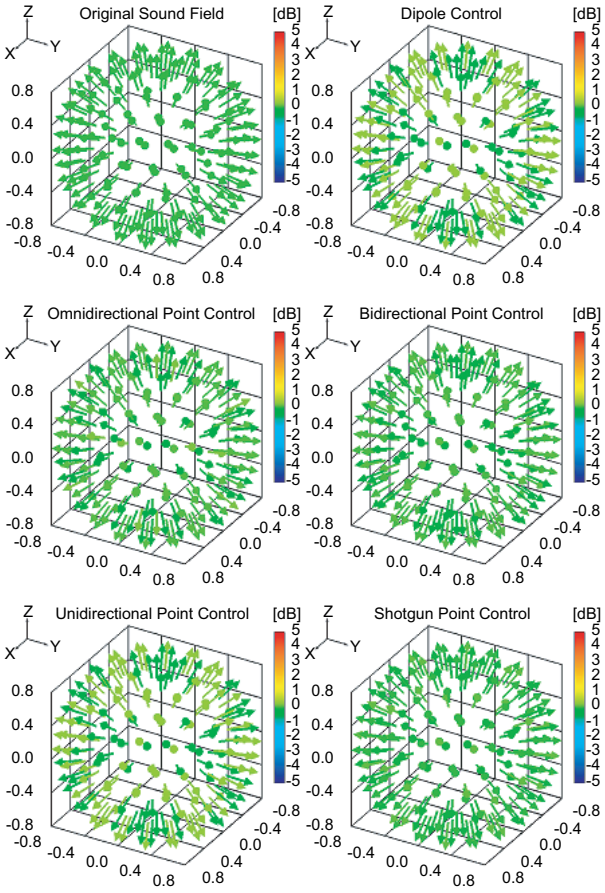

Fig. 8: RMSs of sound pressures and sound intensity (f=1000 Hz, $(0,0,0)^T$).

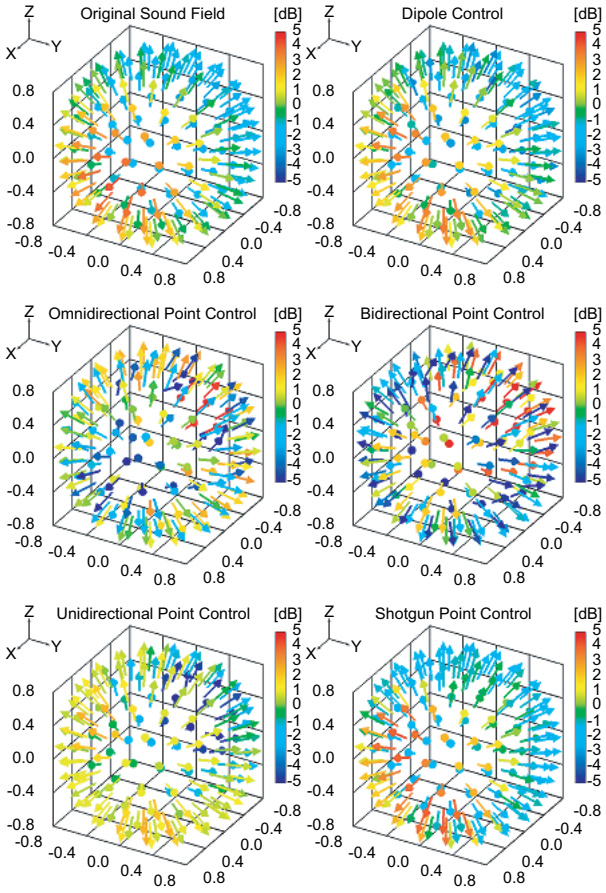

Fig. 9: RMSs of sound pressures and sound intensity vectors (f=1000 Hz, $(0.3,0,0)^T$).

The direction of the arrows in Figs. 8-11 denotes the $\mathbf{I}_0(\mathbf{r},f)$ and $\mathbf{I}(\mathbf{r},f)$ (the sound intensity vectors in the original sound field and reproduced sound field). It is shown that the performance of the proposed methods in the arrival direction distribution of sound sources is good if the direction of the arrows in the proposed methods is the same as that of the arrows in the original sound field. Note that the $\mathbf{I}_0(\mathbf{r},f)$ and $\mathbf{I}(\mathbf{r},f)$ were normalized so that the length of vectors is one before plotting.

In the bidirectional point control method, the directions of the sound intensity vectors were different from those of the sound intensity vectors in the original sound field when the position of the sound source was not central. Thus, it is considered that the performance of the bidirectional point control method was not good.

In the omnidirectional point control method, even if the position of the sound source was not central, the directions of the sound intensity vectors were almost the same as those of the sound intensity vectors in the original sound field, while the RMSs of the sound pressures were different from those of the sound pressures in the original sound field. Thus, it is considered that listeners could localize the direction of sound images in the omnidirectional point control method when they stopped at a particular listening position and listened to the sounds.

In the other methods, even if the position of the sound source was not central, the RMSs of the sound pressures were almost the same as those of the sound pressures in the original sound field. Thus, it is considered that listeners experienced a realistic sensation in the other methods

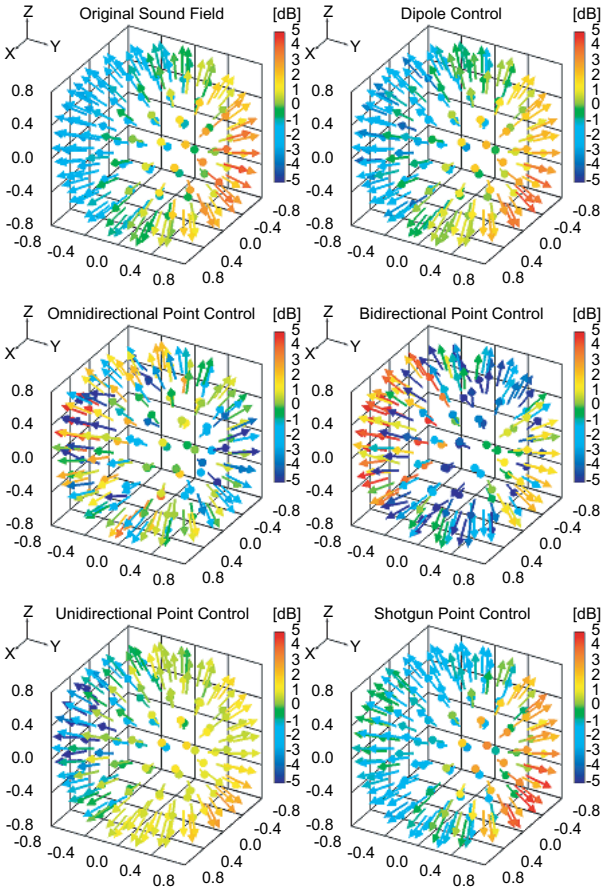

Fig. 10: RMSs of sound pressures and sound intensity vectors (f=1000 Hz, $(0,0.3,0)^T$).

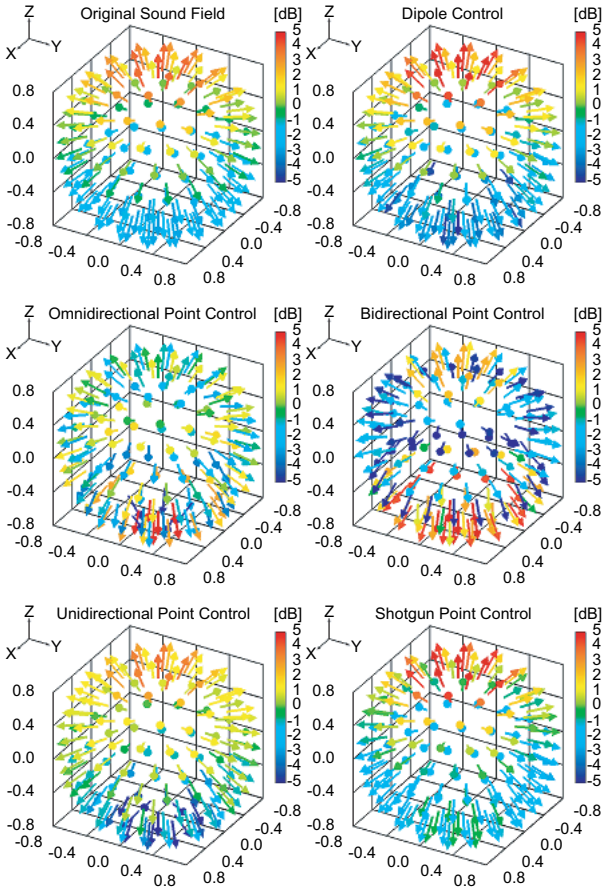

Fig. 11: RMSs of sound pressures and sound intensity vectors (f=1000 Hz, $(0,0,0.3)^T$).

when they moved around the loudspeaker array and listened to the sounds.

To quantitatively evaluate the performance at all the frequencies of the sound sources, two measures were applied. The first was the SNR of the RMSs of sound pressures defined as follows:

$$\mathrm{SNR}(f) = 10\log_{10}\frac{\sum_{\mathbf{r}}\{p_0(\mathbf{r},f)\}^2}{\sum_{\mathbf{r}}\{p(\mathbf{r},f)-p_0(\mathbf{r},f)\}^2}, \qquad (34)$$

where $p_0(\mathbf{r},f)$ and $p(\mathbf{r},f)$ are the RMSs of the sound pressures in the original and reproduced sound fields, respectively. It is indicated that the performance of the proposed methods in the sound pressure distribution is good if the SNR of the RMSs is high. Note that $p_0(\mathbf{r},f)$ and $p(\mathbf{r},f)$ were averaged in the whole $\mathbf{r}$ before calculating the SNRs.

The second measure used was the intensity direction error (IDE) defined as follows:

$$\mathrm{IDE}(f) = \sqrt{\frac{1}{N}\sum_{\mathbf{r}}\left[\cos^{-1}\left\{\frac{\mathbf{I}(\mathbf{r},f)\cdot\mathbf{I}_0(\mathbf{r},f)}{|\mathbf{I}(\mathbf{r},f)||\mathbf{I}_0(\mathbf{r},f)|}\right\}\right]^2}, \qquad (35)$$

where $\mathbf{I}_0(\mathbf{r},f)$ and $\mathbf{I}(\mathbf{r},f)$ are the sound intensity vectors in the original and reproduced sound fields, respectively, and N(= 162) is the total number of synthesis points. It is shown that the performance of the proposed methods in the arrival direction distribution of sound sources is good if the intensity direction error is close to zero.

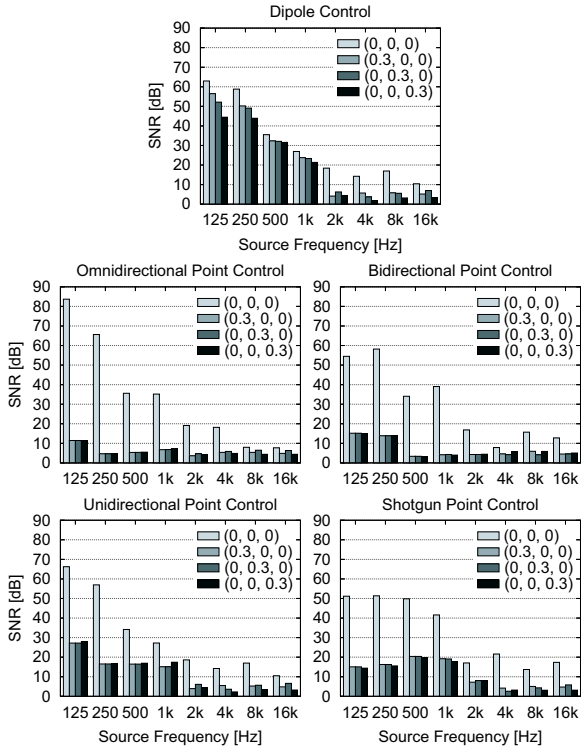

Fig. 12: SNRs in proposed methods.

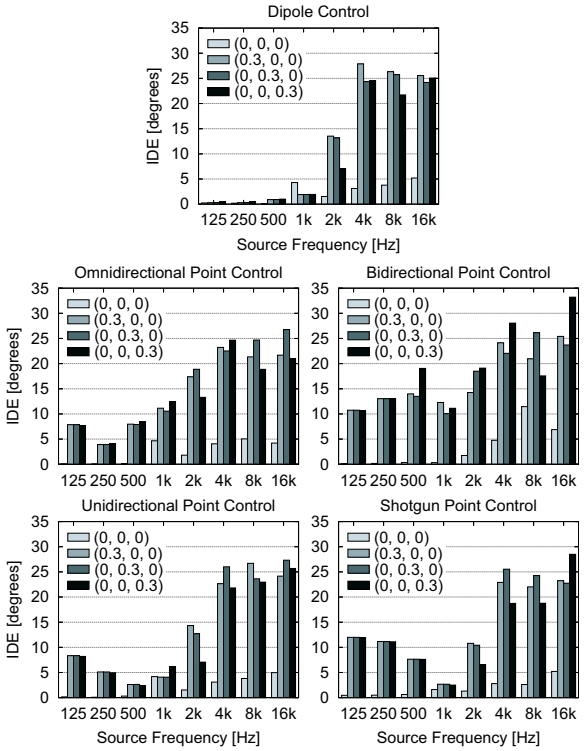

Fig. 13: Intensity direction errors in proposed methods.

The results of the SNRs and intensity direction errors for all the proposed methods are shown in Figures 12-13. In all the proposed methods, when the frequency of the sound sources was more than 2000 Hz, the SNRs were always less than 20 dB and the intensity direction errors were more than 20 degrees. This was due to the fact that the spatial sampling theorem to reproduce a wave front over 2000 Hz sound was not satisfied since the interval of the control points was more than half of the wavelength of over 2000 Hz sound.

In the dipole control method, the SNRs were more than 21.3 dB and the intensity direction errors were less than 4.3 degrees at frequencies of less than 1000 Hz. Thus, if systems were constructed using this method, listeners could localize the direction of sound images and experience a more realistic sensation when they moved around the loudspeaker array since the sound pressures and arrival directions of sound sources were reproduced well.

In the bidirectional point control method, when the position of the sound source was not central, the SNRs were less than 16 dB at frequencies of less than 1000 Hz and the intensity direction errors were more than 13 degrees. Thus, if systems were constructed using this method, the performance is not good since the sound pressures and arrival directions of sound sources were not reproduced.

In the omnidirectional point control method, when the position of the sound source was not central, the SNRs were less than 12 dB at frequencies of less than 1000 Hz, although the intensity direction errors were less than 12.5 degrees. Thus, if systems were constructed using this method, listeners could localize the direction of sound images since the arrival directions of sound sources were sufficiently reproduced. However, they would not experience a realistic sensation when they moved around the loudspeaker array since the sound pressures were not reproduced.

In the shotgun point control method, the SNRs were more than 14.3 dB and the intensity direction errors were less than 12.0 degrees at frequencies of less than 1000 Hz. In the unidirectional point control method, the SNRs

were more than 15.0 dB and the intensity direction errors were less than 8.4 degrees at frequencies of less than 1000 Hz. Thus, if systems were constructed using these methods, listeners could localize the direction of sound images and experience a realistic sensation when they moved around the loudspeaker array since the sound pressures and arrival directions of sound sources were sufficiently reproduced.

4. CONCLUSION

In this paper, near 3D sound field reproduction techniques using wave field synthesis were proposed as methods of achieving ultra-realistic communication in applications such as 3D television and 3D teleconferencing. The principle was derived from the Kirchhoff-Helmholtz integral equation and two methods – dipole control and directional point control methods – were proposed. Computer simulation was used to evaluate the performance of the two methods. The results showed that the dipole control method performed very well, while the directional point control method performed satisfactorily if the directivity of the loudspeakers was unidirectional and shotgun.

Future work will include manufacturing microphone and loudspeaker arrays based on the proposed methods and evaluating the performance of the systems in a real environment using both acoustical measurement and subjective assessment.

5. REFERENCES

[1] *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] J. Blauert, *Spatial Hearing*, pp. 372–392, MIT Press, Cambridge, Mass, revised edition, 1997.

[3] S. Takane, Y. Suzuki, T. Miyajima, and T. Sone, "A new theory for high definition virtual acoustic display named ADVISE," *Acoust. Sci. & Tech.*, vol. 24, no. 5, pp. 276–283, September 2003.

[4] M. R. Schroeder, D. Gottlob, and K. F. Siebrasse, "Comparative study of european concert halls: Correlation of subjective preference with geometric and acoustic parameters," *J. Acoust. Soc. Am.*, vol. 56, no. 4, pp. 1195–1201, October 1974.

[5] M. Miyoshi and N. Koizumi, "Transaural system using multiple loudspeakers," in *Proc. DAGA*, 1991, pp. 781–783.

[6] J. Bauck and D. H. Cooper, "Generalized transaural stereo and applications," *J. Audio Eng. Soc.*, vol. 44, no. 9, pp. 683–705, September 1996.

[7] O. Kirkeby, P. A. Nelson, and H. Hamada, "The 'Stereo Dipole': A virtual source imaging system using two closely spaced loudspeakers," *J. Audio Eng. Soc.*, vol. 46, no. 5, pp. 387–395, May 1998.

[8] M. Camras, "Approach to recreating a sound field," *J. Acoust. Soc. Am.*, vol. 43, no. 6, pp. 1425–1431, June 1968.

[9] A. J. Berkhout, D. de Vries, and P. Vogel, "Acoustic control by wave field synthesis," *J. Acoust. Soc. Am.*, vol. 93, no. 5, pp. 2764–2778, May 1993.

[10] 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," in *Papers of AES 13th Regional Conv.*, Tokyo, Japan, July 2007, number 0037, pp. 1–8.

[11] S. Ise, "A principle of sound field control based on the Kirchhoff-Helmholtz integral equation and the theory of inverse systems," *ACUSTICA - Acta Acustica*, vol. 85, no. 1, pp. 78–87, January/February 1999.

[12] S. Takane, Y. Suzuki, and T. Sone, "A new method for global sound reproduction based on Kirchhoff's integral equation," *ACUSTICA - Acta Acustica*, vol. 85, no. 2, pp. 250–257, March/April 1999.

[13] S. Ise, M. Toyoda, S. Enomoto, and S. Nakamura, "The development of the sound field sharing system based on the boundary surface control principle," in *Proc. Int. Cong. Acoust.*, September 2007, number ELE-04-003, pp. 1–7.

[14] M. Born and E. Wolf, *Principles of Optics*, pp. 418–421, Cambridge University Press, Cambridge, UK, 7th(expanded) edition, 1999.

[15] M. Born and E. Wolf, *Principles of Optics*, pp. 421–425, Cambridge University Press, Cambridge, UK, 7th(expanded) edition, 1999.

[16] H. Kenner, *Geodesic Math and How to Use It*, University of California Press, Berkeley, CA, second paperback edition, 2003.

[17] F. J. Fahy, *Sound Intensity*, Spon Press, UK, 1995.