Numerical Analysis of Radiated Sound Field of Spherical Loudspeaker with Variable Radiation Characteristics

Estimation of Size and Shape of Musician's Sound Image by Sound Source Search

Toshiyuki Kimura and Michiaki Katsumoto
Universal Media Research Center
NICT
Nukui-kitamachi 4-2-1, Koganei, Tokyo, Japan
{t-kimura, katumoto}@nict.go.jp

Yoko Yamakata
Academic Center for Computing and Media Studies
Kyoto University
Yoshida-Honmachi, Sakyo-ku, Kyoto, Japan
yamakata@media.kyoto-u.ac.jp

Abstract— **We have developed a three-dimensional (3D) audio system depicting the presence of an object, i.e., a spherical loudspeaker with variable radiation characteristics. In this study, we perform computer simulations for numerical analysis of the 3D sound field generated by the spherical loudspeaker. We also evaluate the size and shape of a musician's sound image depicted by the developed spherical loudspeaker by estimating the position of multiple point sound sources with the help of the sound source search technique. Our results indicate that the general size of the sound image is scaled down to the size of the spherical loudspeaker and that the shape of the sound image is accurately maintained if the radiation directivity of the loudspeaker units is sharply set toward the outside of the loudspeaker.**

Keywords-3D audio; radiated sound field; radiation directivity; computer simulation; sound source search

I. INTRODUCTION

Several investigations have been carried out on ultra-realistic communication techniques using three-dimensional (3D) audio and video techniques [1]. If these techniques are applied, it is possible to realize more realistic communication services than conventional audio and video techniques such as HD video and 5.1ch audio. To realize the abovementioned services, it is necessary to develop audio and video techniques for accurately reproducing a given scene or object.

The ultra-realistic audio technique is the one of techniques 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 a realistic representation of a particular scene. In this study, we have developed a 3D audio technique to depict the presence of an object at a given position. This technique enables listeners occupying any position around an object to listen to the sound emanating from the object.

We developed a 3D audio system that can depict the presence of the object, i.e., a spherical loudspeaker with variable radiation characteristics [6]. A performance of the musician was recorded by 26 microphones placed around him. A spherical loudspeaker consisting of 26 loudspeaker units directly played back the recorded 26-channel audio signals. As a result, several listeners felt as though they could listen to the music, irrespective of the position they occupied around the musician. In this study, for numerical evaluation of the radiated sound field generated by the spherical loudspeaker, the size and shape of the musician's sound image were evaluated. For this purpose, the position of multiple point sound sources was estimated with the help of the sound source search technique. Section 2 presents the simulation environment and the results of the computer simulation performed in this study.

II. COMPUTER SIMULATION

A. Simulation Environment

As shown in Figure 1, M (= 26) microphones were placed at $\mathbf{r}_i$ on the boundary of a sphere of radius r in the original sound field. M (= 26) loudspeaker units and N (= 162) observation points were positioned at point $\mathbf{r}'_i$ on the boundary of a sphere of radius r' and at $\mathbf{R}_j$ on the boundary of a sphere of radius R, respectively, in the reproduced sound field.

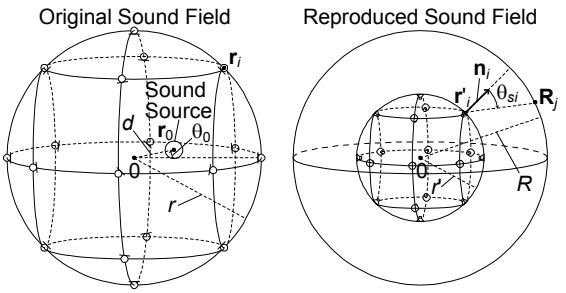

Figure 1. Positions of the microphones, loudspeaker units, and observation points used in computer simulation.

The sound source signal $s(t)$ (= $A\sin 2\pi ft$) is a sine wave of amplitude A and frequency f. $x_i(t)$ (signal recorded by the ith microphone) can be denoted as follows:

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

where $d_{i0}(= |\mathbf{r}_i - \mathbf{r}_0|)$ is the distance between the sound source and the ith microphone. $\mathbf{r}_i$ and $\mathbf{r}_0$ are the position vectors of the ith microphone and the sound source, respectively, and c is the sound velocity. $p(\mathbf{R}_j,f,t)$ (sound pressure at the jth observation point $\mathbf{R}_j$ in the reproduced sound field) is calculated from $x_i(t)$ as follows:

$$\begin{aligned} p(\mathbf{R}_j, f, t) &= \sum_{i=1}^{M} \frac{D_{si}}{d_{ji}} x_i\left(t - \frac{d_{ji}}{c}\right) \\ &= \sum_{i=1}^{M} \frac{D_{si} A}{d_{ji} d_{i0}} \sin\left\{2\pi f\left(t - \frac{d_{ji} + d_{i0}}{c}\right)\right\} \end{aligned}, \quad (2)$$

where M is the total number of loudspeaker units, $d_i(= |\mathbf{R}_j - \mathbf{r'}_i|)$ is the distance between the ith loudspeaker unit and the observation point, $\mathbf{r'}_i$ is the position vector of the ith loudspeaker unit, and D_{si}, is the radiation directivity of the ith loudspeaker unit.

In addition, in this computer simulation, the sound intensity vectors for the observation point $\mathbf{R}_j$ are calculated by using the cross-spectral method, as shown in Figure 2 [7]. $I_x(\mathbf{R}_j,f)$, $I_y(\mathbf{R}_j,f)$, and $I_z(\mathbf{R}_j,f)$ in Figure 2 are the x, y, and z components of the sound intensity vector $\mathbf{I}(\mathbf{R}_j,f) = \{I_x(\mathbf{R}_j,f), I_y(\mathbf{R}_j,f), I_z(\mathbf{R}_j,f)\}^T$, respectively. Further, $p(\mathbf{R}_{jx}^+,f,t)$, $p(\mathbf{R}_{jx}^-,f,t)$, $p(\mathbf{R}_{jy}^+,f,t)$, $p(\mathbf{R}_{jy}^-,f,t)$, $p(\mathbf{R}_{jz}^+,f,t)$, and $p(\mathbf{R}_{jz}^-,f,t)$ in this figure are the sound pressures at points $\mathbf{R}_{jx}^+$, $\mathbf{R}_{jx}^-$, $\mathbf{R}_{jy}^+$, $\mathbf{R}_{jy}^-$, $\mathbf{R}_{jz}^+$ and $\mathbf{R}_{jz}^-$, respectively. The position vectors of these six points are set as follows:

$$\begin{aligned} &\mathbf{R}_{jx}^+ = \mathbf{R}_j + (\Delta,0,0)^T, \mathbf{R}_{jy}^+ = \mathbf{R}_j + (0,\Delta,0)^T, \mathbf{R}_{jz}^+ = \mathbf{R}_j + (0,0,\Delta)^T \\ &\mathbf{R}_{jx}^- = \mathbf{R}_j - (\Delta,0,0)^T, \mathbf{R}_{jy}^- = \mathbf{R}_j - (0,\Delta,0)^T, \mathbf{R}_{jz}^- = \mathbf{R}_j - (0,0,\Delta)^T \end{aligned}, \quad (3)$$

where Δ is 0.001 m.

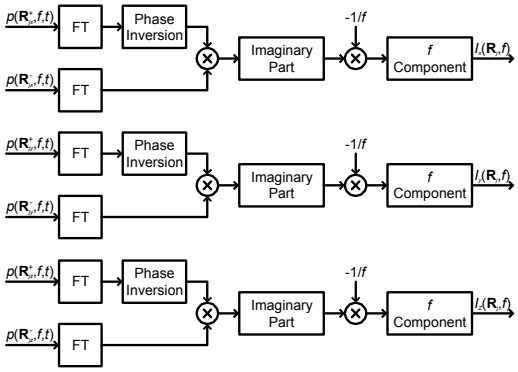

Figure 2. Block diagram showing the calculation of sound intensity in computer simulation.

The parametric conditions employed in the computer simulation are listed in TABLE I. $\mathbf{r}_0$, $\mathbf{r}_i$, $\mathbf{r'}_i$, and $\mathbf{R}_j$ (position vectors of the sound source, microphones, loudspeaker units, and observation points, respectively) were set in a three-dimensional coordinate, as follows:

$$\begin{aligned} \mathbf{r}_0 &= \begin{pmatrix} d\cos\theta_0 & d\sin\theta_0 & 0 \end{pmatrix}^T \\ \mathbf{r}_i &= \begin{pmatrix} r\cos\theta_i\cos\varphi_i & r\sin\theta_i\cos\varphi_i & r\sin\varphi_i \end{pmatrix}^T (i = 1...M) \\ \mathbf{r'}_i &= \begin{pmatrix} r'\cos\theta_i\cos\varphi_i & r'\sin\theta_i\cos\varphi_i & r'\sin\varphi_i \end{pmatrix}^T (i = 1...M) \\ \mathbf{R}_j &= \begin{pmatrix} R\cos\theta_j\cos\varphi_j & R\sin\theta_j\cos\varphi_j & R\sin\varphi_j \end{pmatrix}^T (j = 1...N) \end{aligned}, \quad (4)$$

where θ_i and φ_i (the azimuth and elevation angle of the ith microphone and loudspeaker unit) are set in the manner shown in TABLE II. Note that the origin is the center of the sphere. θ_j and φ_j are the azimuth and elevation angle of the jth observation point. The position of the 162 observation points corresponds to the vertex of a Class I Method 1 icosahedral geodesic dome following four frequencies [8].

TABLE I. PARAMETRIC CONDITIONS IN COMPUTER SIMULATION

Sound source amplitude (A)	1
Sound source frequency (f)	125, 250, 500, 1000, 2000, 4000, 8000, 16000 Hz
Distance of sound sources (d)	0, 0.2, 0.4, 0.6 m
Azimuth angle of sound source (θ_0)	0, 45º
Sound velocity (c)	340 m/s
Number of microphones and loudspeaker units (M)	26
Radius of microphone array (r)	0.8 m
Radius of loudspeaker array (r')	0.085 m
Normal unit vector of loudspeaker units ($\mathbf{n}_i$)	$\mathbf{r'}_i/\|\mathbf{r'}_i\|$
Radiation directivity of loudspeaker units (D_{si})	Omnidirectional, Decay 6 dB, Decay 12 dB, Real
Number of observation points (N)	162
Radius of observation points (R)	1 m

TABLE II. AZIMUTH AND ELEVATION ANGLES OF MICROPHONES AND LOUDSPEAKER UNITS

i	θ_i[º]	φ_i[º]	i	θ_i[º]	φ_i[º]	i	θ_i[º]	φ_i[º]
1	0	0	9	0	45	18	0	–45
2	45	0	10	45	30	19	45	–30
3	90	0	11	90	45	20	90	–45
4	135	0	12	135	30	21	135	–30
5	180	0	13	180	45	22	180	–45
6	–135	0	14	–135	30	23	–135	–30
7	–90	0	15	–90	45	24	–90	–45
8	–45	0	16	–45	30	25	–45	–30
			17	---	90	26	---	–90

Four types of D_{si} (radiation directivity of the ith loudspeaker unit) can be defined as follows:

$$\begin{aligned} &\text{(Omnidirectional) } D_{si} = 1 \text{ (in all } f) \\ &\text{(Decay 6 dB) } D_{si} = \frac{3 + \cos\theta_{si}}{4} \text{ (in all } f) \\ &\text{(Decay 12 dB) } D_{si} = \frac{5 + 3\cos\theta_{si}}{8} \text{ (in all } f) \\ &\text{(Real) } D_{si} = \begin{cases} 1 & (f = 125,250,500) \\ (3 + \cos\theta_{si})/4 & (f = 1\text{k},2\text{k}) \\ (5 + 3\cos\theta_{si})/8 & (f = 4\text{k},8\text{k},16\text{k}) \end{cases} \end{aligned}, \quad (5)$$

where $\cos\theta_{si} = \{\mathbf{n}_i \cdot (\mathbf{r} - \mathbf{r'}_i)\}/\{|\mathbf{n}_i||\mathbf{r} - \mathbf{r'}_i|\}$. Under the "Real" condition, the radiation directivity of the loudspeaker units is

simulated on the basis of the results of acoustical measurements performed using the developed spherical loudspeaker [6].

With the help of the sound source search technique, the sound image position $\mathbf{r}_E$ is estimated from the calculated sound pressures $p(\mathbf{R}_j,f,t)$ and sound intensity vectors $\mathbf{I}(\mathbf{R}_j,f)$, as shown below:

$$\mathbf{r}_E = \frac{1}{FN}\sum_f \sum_{j=1}^{N} \left\{ \mathbf{r}_j + \frac{\mathbf{I}(\mathbf{R}_j,f)}{p(\mathbf{R}_j,f)} \right\}, \tag{6}$$

where N (=162) and F (=8) are the total number of observation points and frequency conditions used in the computer simulation, respectively. Note that $|\mathbf{I}(\mathbf{R}_j,f)| = 1$. $p(\mathbf{R}_j,f)$ is the mean-square sound pressure calculated from $p(\mathbf{R}_j,f,t)$ by using the following equation:

$$p(\mathbf{R}_j,f) = \sqrt{\frac{1}{T}\int_0^T \{p(\mathbf{R}_j,f,t)\}^2 dt}, \tag{7}$$

where T is the period. $\mathbf{r}_E$ is estimated after $p(\mathbf{R}_j,f)$ is normalized so that the variance in the estimated sound image position is minimized. On the other hand, the position of the sound source placed at $\mathbf{r}_0$ is directly estimated in order to evaluate the performance of the sound source search technique used in this computer simulation.

B. Simulation Results

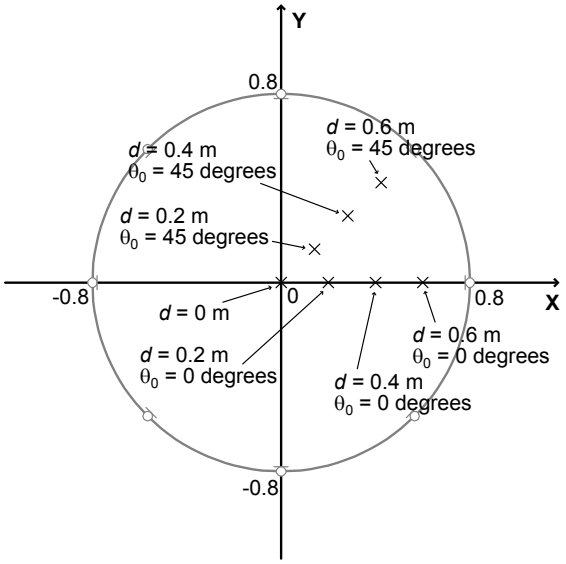

Figure 3. Position coordinates of estimated sound sources in the XY plane.

The estimated position of the sound sources placed at $\mathbf{r}_0$ is shown in Figure 3. Note that only the XY plane is plotted in following figures because the value of the z component of the estimated position is almost zero. The × marks in Figure 3 denote the estimated positions of the sound sources. The gray circle and microphones in this figure denote the microphone array, whose radius is 0.8 m. From the results shown in Figure 3, it is apparent that the estimated position of the sound sources is the same as the inputted position. This indicates that the position of the sound source can be accurately estimated by the sound source search technique used in this computer simulation.

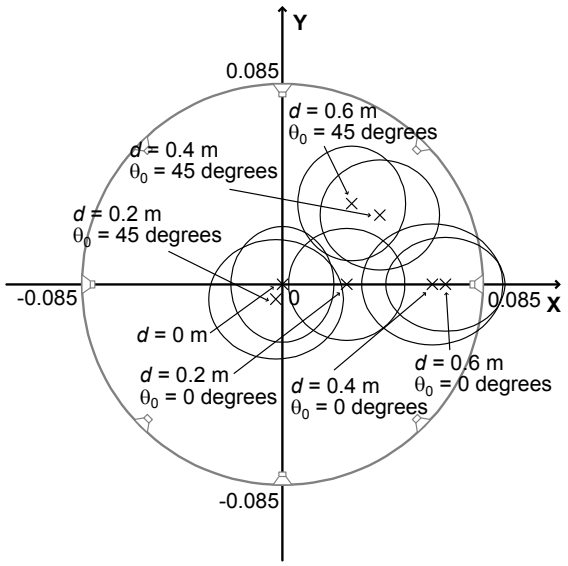

Figure 4. Position coordinates and 95% confidential intervals of estimated sound images in the XY plane. (D_{si}: Real)

Figure 4 shows the estimated position of the sound images under the condition that D_{si} is "Real." The × marks in Figure 4 denote the estimated positions of the sound images. The black ellipses around the marks in Figure 4 denote the 95% confidential interval of the estimated position. The horizontal and vertical lengths of the ellipses correspond to the x and y components of the 95% confidential intervals, respectively. The gray circle and loudspeaker units in Figure 4 denote the spherical loudspeaker, whose radius is 0.085 m.

When the sound source is placed at the center of the microphone array (d = 0 m), the sound image is estimated at the center of the spherical loudspeaker ($\mathbf{r}_E = (0,0,0)^T$). Thus, since the center of the microphone array is the fixed position, the size and shape of the musician's sound image can be evaluated from the center of the spherical loudspeaker.

In Figure 4, the distance between the estimated position and the center of the spherical loudspeaker is always less than that between the sound source position and the center of the microphone array. Thus, if the musician's sound source consists of multiple point sources, the general size of the musician's sound image is considered to be scaled down to the size of the spherical loudspeaker. The confidential intervals of all the estimated sound images are one-fourth the size of the spherical loudspeaker. Thus, it is indicated that the sizes of individual point sound images are scaled up significantly.

The effect of the radiation directivity of loudspeaker units is determined from the estimated results obtained for different radiation directivities (Figures 5–7). When the radiation directivity is "Omnidirectional," although the inputted and estimated azimuth angles of the sound images are equal, the estimated distance between the sound images and the center of the spherical loudspeaker biases to the zero meters. On the other hand, under the conditions "Decay 6 dB" or "Decay 12

dB," the estimated distance between the sound images and the center of the spherical loudspeaker is similar to the inputted distance. These results indicate that the shape of the musician's sound image can be accurately maintained if the radiation directivity of the loudspeaker units is sharply set toward the outside of the loudspeaker.

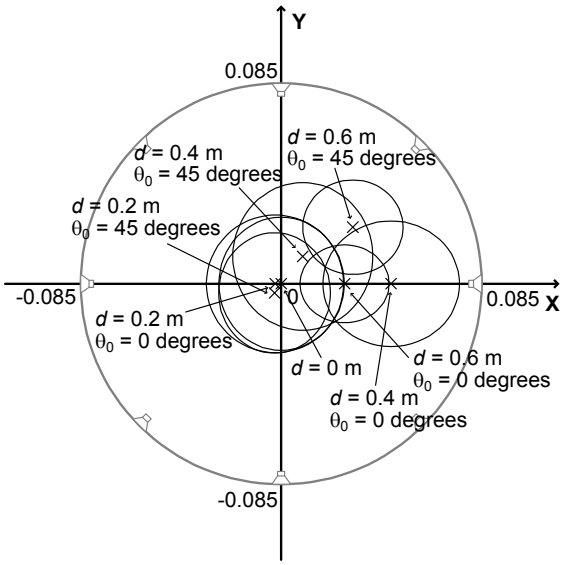

Figure 5. Position coordinates and 95% confidential intervals of estimated sound images in the XY plane. (D_{si}: Omnidirectional)

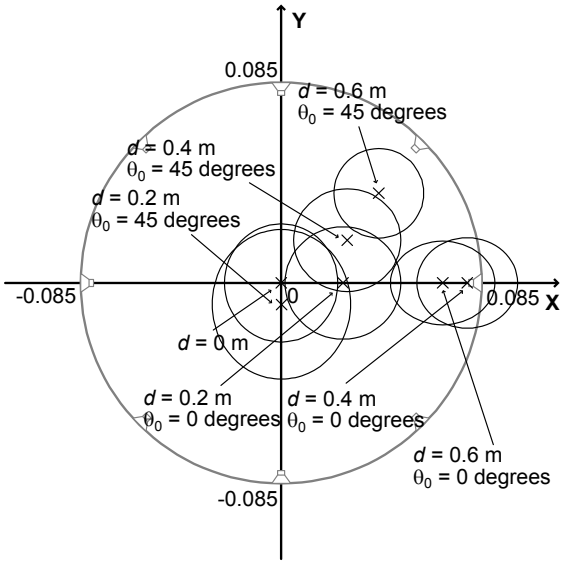

Figure 6. Position coordinates and 95% confidential intervals of estimated sound images in the XY plane. (D_{si}: Decay 6 dB)

III. CONCLUSION

In this study, we performed computer simulations and numerically estimated the radiated sound field reproduced by a spherical loudspeaker. We also estimated the position of multiple sound images by the sound source search technique, and evaluated the size and shape of the musician's sound image depicted by the spherical loudspeaker. Our results indicated that the general size of the musician's sound image was scaled down to the size of the loudspeaker and that the size of individual point sound sources was scaled up significantly. Furthermore, the shape of the musician's sound image was accurately maintained if the radiation directivity of the loudspeaker units was sharply set toward the outside of the loudspeaker.

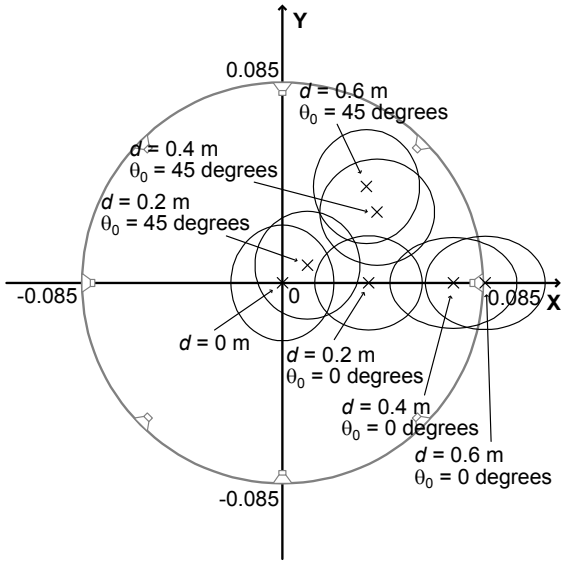

Figure 7. Position coordinates and 95% confidential intervals of estimated sound images in the XY plane. (D_{si}: Decay 12 dB)

In our future study, we plan to perform computer simulations by considering the radiation directivities of multiple point sound sources and numerically evaluate the effect of the developed spherical loudspeaker on the radiation characteristics of the musician's sound image.

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., vol. 43, pp. 1425–1431, June 1968.

[3] A. J. Berkhout, D. de Vries, and P. Vogel, "Acoustic control by wave field synthesis," J. Acoust. Soc. Am., vol. 93, pp. 2764–2778, May 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, 2007, No. RBA-15-011, pp. 1–6.

[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, pp. 40-49, April 2008.

[6] M. Katsumoto, Y. Yamakata and T. Kimura: "Realization of multidimensional speaker with frequency dependence of directivity", Proc. AES Japan Sect. Conf., 2008, No. PS08, pp. 1–6.

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

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