CLOSE-TALKING SPHERICAL MICROPHONE ARRAY USING SOUND PRESSURE INTERPOLATION BASED ON SPHERICAL HARMONIC EXPANSION

Yoichi Haneda[⋆] *Ken'ichi Furuya*[†] *Shoichi Koyama*[‡] *Kenta Niwa*[‡]

[⋆] The University of Electro-Communications, Tokyo, Japan
[†] Oita University, Oita, Japan
[‡] NTT Media Intelligence Laboratory, Tokyo, Japan

ABSTRACT

We propose a novel close-talking spherical microphone array that uses the residual signal between the observed sound pressure and the interpolated sound pressure at the center of the spherical array. The interpolated sound is obtained from the sound pressures observed on the surface of a sphere on the basis of the spherical harmonic expansion, assuming that the sound originates from the outside of the array. If the sound source is close to the spherical array, the array cannot express the spherical wave correctly because the number of microphones is limited. As a result, the residual signal increases. This method is a modified form of the conventional method, which interpolates the sound pressure by using the Kirchhoff integral equation. In contrast with the conventional method, we interpolate the sound at the center of the sphere by using only the average value of the sound pressures on the spherical array surface. The computer simulations were conducted using a 12-element spherical microphone array with radius of 5 cm. These results showed that the performances of both methods were almost equivalent, although the proposed method used half the number of microphones as the conventional method.

***Index Terms*—** close-talking microphone, microphone array, spherical harmonic expansion

1. INTRODUCTION

The microphone array is an important technology used to receive far-field sound in a noisy environment. However, when the noise level is very high, the far-field microphone array may not be able to eliminate the noise. A close-talking microphone is often used for such a case. The close-talking microphone is designed to have a spatial sensitivity that is very high near the microphone; for example, the distance between the microphone and sound source (mouth) is approximately 4 cm, but far-field sound is not acquired. Although a pressure-gradient microphone is usually used as a close-talking microphone, its frequency response varies when the distance between the source and microphone or the direction of the microphone are changed [1].

To overcome this problem, some near-field spherical arrays have been proposed in the spherical harmonic domain [3][4]. These methods use the spatial orthonormal decomposition of the sound with respect to the directional and radial components to emphasize the near-field sound. Radial filtering may be a promising approach to create a true three-dimensional microphone array.

Another close-talking microphone array with orientation invariance has been proposed on the basis of the Kirchhoff integral equation by Date et al. [5]. In this close-talking spherical microphone array, the microphones are arranged on an "open" sphere, and one microphone is placed at the center of the sphere. First, this array interpolates the sound pressure at the center of the sphere by using the surface integral of the sound pressure, the normal derivative of the sound pressure to the boundary surface, and the time derivative of the sound pressure on the boundary spherical surface on the basis of the Kirchhoff integral equation. Then, this array outputs the residual signal between the observed sound pressure at the center of the sphere and the interpolated sound pressure. Theoretically, if the sound source is located outside of the closed region (the spherical array), the output signal is zero. However, when the sound source is very close to the microphone array, the residual signal is increased because the array system may be unable to distinguish whether the position of the sound source is inside or outside the closed region as a result of using a finite number of microphones. The array uses these characteristics to achieve a close-talking microphone.

The Kirchhoff integral equation is also applied to wave field synthesis (WFS) [6]. In this research area, high-order ambisonics (HOA) based on spherical harmonic expansion are also studied [7]. Recently, it was reported that WFS and HOA have a close relationship [8]-[10]. This relationship seems to be applied to the above close-talking spherical microphone array, so we propose a new close-talking spherical microphone array based on spherical harmonic expansion. The proposed spherical array interpolates the sound pressure at the center of the sphere by using the sound pressures observed on the sphere. The array output is the difference between the sound pressure observed at the center and the inter-

978-1-4799-2893-4/14/$31.00 ©2014 IEEE

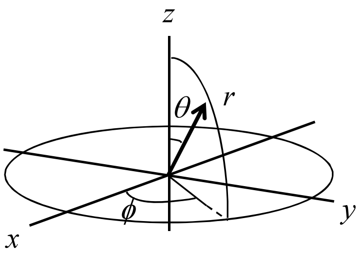

Fig. 1. Definition of spherical coordinates: $x = r\sin\theta\cos\phi, y = r\sin\theta\sin\phi$, and $z = r\cos\theta$.

polated sound pressure in the same way as the conventional method does.

In this paper, section 2 describes a conventional close-talking spherical microphone array based on the Kirchhoff integral equation. In section 3, we propose a novel close-talking spherical microphone array based on the spherical harmonic expansion. Section 4 shows the results of the computer simulation with a spherical array with a radius of 5 cm.

2. CONVENTIONAL CLOSE-TALKING ARRAY BASED ON KIRCHHOFF INTEGRAL

In this paper, we assume a spherical microphone array in which the microphone units are equispaced or equiangular [11][12]. Moreover, the array is open-sphere and acoustically transparent. We use the spherical coordinates shown in Fig. 1. In this figure, r is the radius of the sphere, and θ and ϕ indicate the angles. After that, we use the simple expression of $\Omega = (\theta, \phi)$.

Date et al. has proposed a close-talking spherical microphone array based on the Kirchhoff integral equation in the time domain [5]. The Kirchhoff integral equation shows that the sound pressure at an arbitrary position within a closed region can be calculated by using the sound pressures, normal directional pressures, and time differential of the sound pressures on its surface. Date et al. applied this equation to a spherical microphone array by setting the closed surface and arbitrary position as a spherical surface and the center of the sphere, respectively. Figure 2 shows an illustration of the microphone positions and the boundary surface. This method first interpolates the sound pressure at the center of the sphere by using the surface integral of the sound pressures, normal directional sound pressures, and time differential of the sound pressures on the sphere. The array output signal $y(t,\omega)$ is obtained by subtracting the interpolated signal from the actual observed sound $p_0(t,\omega)$ at the center. When assuming that the sound is a monophonic wave, and we omit the time index t and ω, the output signal y is

$$y = p_0 - \frac{1}{M}\sum_{i=1}^{M}\left[p(r,\Omega_i) + r\frac{\partial p(r,\Omega_i)}{\partial n} + \left(\frac{r}{c}\right)\frac{\partial p(r,\Omega_i)}{\partial t}\right]_{t_0=t-\frac{r}{c}}, \tag{1}$$

Here, r is the radius of the array, $p(r,\Omega_i)$ indicates the sound pressure on the sphere, i is the index of the microphones, M is the number of microphones, and the c is the sound velocity. $[\cdot]_{t_0=t-r/c}$ inserts the time delay of r/c into the calculation results of the second term on the right side of the equation. In this equation, if the sound source is outside of a closed region, the array output is theoretically zero.

In practice, the normal derivative to the boundary surface pressure is obtained from the microphone pair placed normal to the sphere surface with a distance of Δr, as illustrated in Fig. 2. That is, we use

$$\frac{\partial p(r,\Omega_i)}{\partial n} = \frac{p(r_1,\Omega_i) - p(r_2,\Omega_i)}{\Delta r}, \tag{2}$$

where $r_1 = r + \Delta/2$ and $r_2 = r - \Delta r/2$. Simultaneously, the sound pressure $p(r,\Omega_i)$ is obtained as the average of two sound pressures of the microphone pair

$$p(r,\Omega_i) = \frac{p(r_1,\Omega_i) + p(r_2,\Omega_i)}{2}. \tag{3}$$

Moreover, when assuming the incident wave is a plane wave $p(t,\omega) = A(\omega)e^{j(\omega t - kr)}$ ($k = \omega/c$), the time derivative term and the time delay operation of $[\cdot]_{t_0=t-r/c}$ are replaced by $jkrp(r,\Omega_i)$ and e^{-jkr} respectively. As a result, the output signal is

$$y = p_0 - \frac{1}{M}\sum_{i=1}^{M}\left[p(r,\Omega_i) + r\frac{p(r_1,\Omega_i) - p(r_2,\Omega_i)}{\Delta r} + jkrp(r,\Omega_i)\right]e^{-jkr}. \tag{4}$$

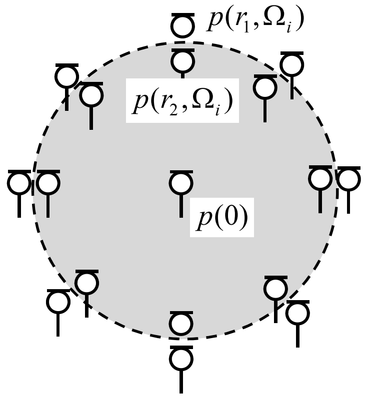

Fig. 2. Illustration of microphone positions and boundary surface in a conventional close-talking spherical array based on Kirchhoff integral equation.

3. PROPOSED METHOD BASED ON SPHERICAL HARMONIC EXPANSION

3.1. Spherical harmonic expansion

The sound pressures $p(r,\Omega)$ observed on the sphere can be transformed to the spherical wave spectrum (spherical harmonic expansion coefficients) $P_{nm}(r)$ by using the spherical harmonic function $Y_n^m(\Omega)$ [2].

$$P_{nm}(r) = \int_\Omega p(r,\Omega) Y_n^m(\Omega)^* d\Omega. \quad (5)$$

Here, $d\Omega = \sin\theta d\theta d\phi$. The sound pressures can be also calculated by using the inverse transform,

$$p(r,\Omega) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} P_{nm}(r) Y_n^m(\Omega). \quad (6)$$

The general solution of the sound wave coming from the outside of the sphere in spherical coordinates is known as [2]

$$p(r,\Omega) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} A_{nm} j_n(kr) Y_n^m(\Omega), \quad (7)$$

where $j_n(\cdot)$ is a spherical Bessel function. Inserting this equation into (5), the spherical wave spectrum at r is

$$P_{nm}(r) = A_{nm} j_n(kr). \quad (8)$$

This equation shows that the spherical wave spectrum $P_{nm}(r)$ of the incident sound from the outside of the sphere can be separated into two parts, the spherical Bessel function $j_n(kr)$ that depends on the radial direction and A_{nm} that is independent of the radial direction. Using this fact yields the important interpolated relationship between the spherical wave spectrum $P_{nm}(r_0)$, which is at the radius of $r_0(< r)$, and $P_{nm}(r)$.

$$P_{nm}(r_0) = \frac{j_n(kr_0)}{j_n(kr)} P_{nm}(r). \quad (9)$$

3.2. Interpolation of sound at the center

From equation (9), the interpolated spherical wave spectrum at the center ($r_0 = 0$) is,

$$P_{nm}(0) = \frac{j_n(0)}{j_n(kr)} P_{nm}(r). \quad (10)$$

Because $j_0(kr) = \frac{\sin kr}{kr}$, $j_0(0)$ is 1. For other order n, for example $n = 1, 2, 3, \cdots$, $j_n(0)$ is 0 [2]. Therefore, we can use the following relations,

$$P_{00}(0) = \frac{1}{j_0(kr)} P_{00}(r), \quad (11)$$

$$P_{nm}(0) = 0 \qquad (n,m) \neq (0,0). \quad (12)$$

Substituting these relationships and $Y_0^0(\Omega) = 1/\sqrt{4\pi}$ into (6), we obtain

$$\hat{p}(0) = \frac{1}{j_0(kr)} \frac{1}{\sqrt{4\pi}} P_{00}(r). \quad (13)$$

By discretizing the spherical surface to M areas, the integral in (5) can be replaced with a summation. Moreover, when considering that the unit area size of the sphere is 4π,

$$P_{00}(r) \sim \frac{1}{\sqrt{4\pi}} \frac{4\pi}{M} \sum_{i=1}^{M} p(r,\Omega_i). \quad (14)$$

If we choose t-design sampling on the sphere (spherical design with equal quadrature coefficients), this equation will have high accuracy [13].

Inserting (14) into (13), the interpolated sound pressure at the center of the sphere can be obtained by

$$\hat{p}(0) = \frac{1}{j_0(kr)} \frac{1}{M} \sum_{i=1}^{M} p(r,\Omega_i). \quad (15)$$

This equation shows that the sound pressure at the center of the sphere can be interpolated only by the average sound pressure on the surface of the sphere.

Finally, the output signal of our proposed close-talking microphone array system, which uses the residual between the observed signal and interpolated signal at the center as the conventional system does is

$$y_{\mathrm{SH}} = p_0 - \frac{1}{j_0(kr)} \frac{1}{M} \sum_{i=1}^{M} p(r,\Omega_i). \quad (16)$$

Note that $j_0(kr)$ becomes 0 for a certain value of kr. For such frequencies, we have to inhibit the array processing. These are known as the "forbidden frequencies" because of the spherical Bessel zeros [2][14]. To overcome this problem, we can choose a spherical array with a small radius to eliminate the Bessel zeroes for speech frequencies. Another solution that uses a dual-sphere array has been proposed [15].

4. COMPUTER SIMULATIONS

We conducted the computer simulations to compare the performance of the conventional and proposed arrays. Figure 3 shows the 12-element spherical microphone array with a radius of 5 cm. In this case, there is no spherical Bessel zero at frequencies below 3.4 kHz. The microphones are located on each surface of a regular dodecahedron. The conventional method had to use the microphone pairs on the surfaces to obtain the normal derivative of the pressures. Therefore, the conventional method requires twice the number of microphones than the proposed method excepting the center microphone. In the conventional method, the distance between the two microphones of the microphone pair was set to 2 mm. The simulations were conducted at 1, 2, and 3 kHz. The sound velocity was set to 340 m/s.

Figure 4 shows contour plots of the spatial responses of the omnidirectional microphone and those of a dipole (pressure-gradient) microphone, which is a conventional close-talking microphone. Figure 5 shows the contour plots of the spatial responses of the arrays. The contours are in 5 dB steps. Because the microphone arrangement is asymmetric for $\phi = 0$ and $\pi/2$, we evaluated the spatial responses

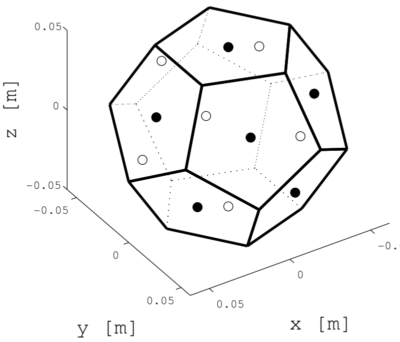

Fig. 3. Twelve-element spherical array

of two surfaces. The right sides of these figures show the spatial responses at $\phi = 0$, and the left sides show the spatial responses at $\phi = \pi/2$. The output level observed at a distance of 5 cm from the microphone array at the north pole position was set to 0 dB. The white area in these figures indicates that the output level was below -40 dB. These results show that the two methods have the same performances, although the proposed method uses approximately half the number of microphones. These results seem to show a similar relationship as that between the WFS and HOA.

The average gain of the spatial response of the omnidirectional microphone was approximately -9.5 dB at 15 cm from the array. On the other hand, the proposed and conventional methods had a spatial average gain at 15 cm of -37 dB, -21 dB, and -11 dB for 1 kHz, 2 kHz, and 3 kHz, respectively. These results showed that the performances of these close-talking microphone arrays were very good, especially below 2 kHz ($kr = 1.84$). The decrease in the performances at high frequencies seems to be due to spatial aliasing.

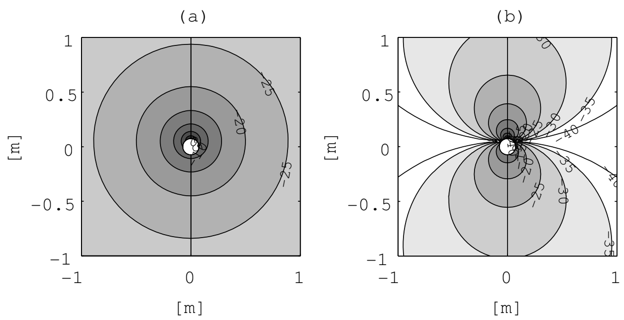

Fig. 4. Simulation results of the spatial responses of (a) omnidirectional and (b) dipole microphones.

5. CONCLUSION

A novel close-talking spherical microphone array using the residual signal between the observed sound pressure and the interpolated sound pressure based on the spherical harmonic expansion was proposed. The basic idea is based on a conventional method, which interpolates the sound pressure at the

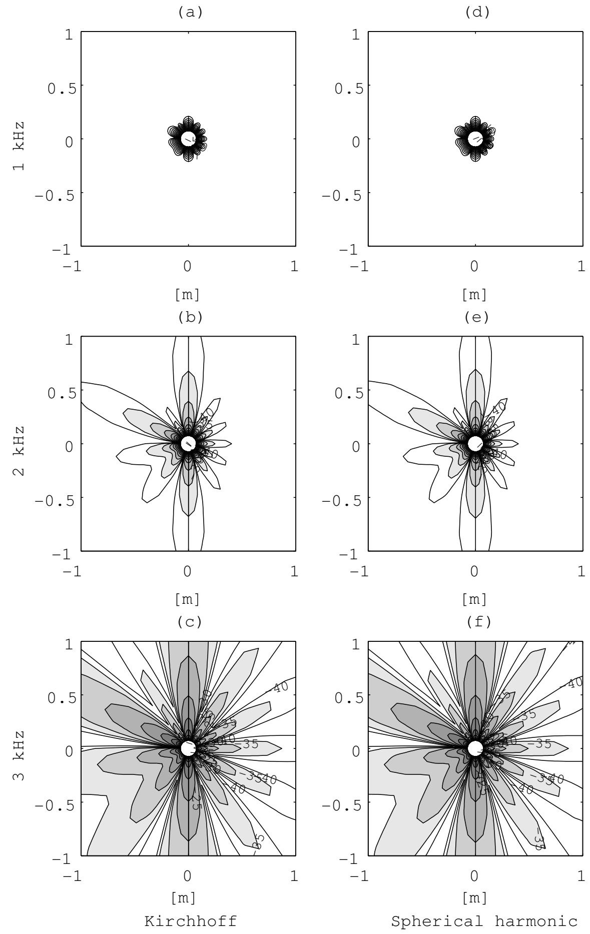

Fig. 5. Simulation results of spatial responses, (a)–(c): conventional, (d)–(f): proposed

center position by using the Kirchhoff integral with the sound pressures, their normal derivatives, and their time derivatives on the boundary surface. In contrast, our proposed spherical array interpolates the sound pressure at the center by using only the sound pressures on the surface and their averaging process. The computer simulations for the 12-element spherical microphone array with a radius of 5 cm showed that both methods had almost the same performance, although the proposed method used approximately half the number of microphones compared to the conventional method. These results seem to be similar to the relationship between WFS and HOA. The computer simulations also showed that the spatial sensitivities of these methods steeply decreased by approximately 11 dB at 2 kHz ($kr = 1.84$) compared with an omnidirectional microphone.

6. ACKNOWLEDGEMENT

This work was supported by JSPS KAKENHI grant numbers 25330187 and 25280060.

7. REFERENCES

[1] Y. Huang, J. Benesty, *Audio Signal Processing*, Kluwer Academic Publishers, Boston, 2004.

[2] E. G. Williams, *Fourier Acoustics*, Academic Press, London, 1999.

[3] J. Meyer, G. W. Elko, "Position independent close-talk microphone," *Signal Process.*, vol. 86, no. 6, 1254–1259, 2005.

[4] E. Fisher, B. Rafaely, "Near-Field Spherical Microphone Array Processing with Radial Filtering," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 19, no. 2, 256–265, 2011.

[5] H. Date, K. Furuya, S. Mikami, "A New Principle of Sound Reception with Spatial Separability," *Acta Acustica United with Acustica*, vol. 72, no. 4, 280–287, 1990.

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

[7] J. Daniel, "Spatial sound encoding including near field effect: Introducing distance coding filters and a viable, new ambisonic format," in *Proc. 23rd Conf. AES*, Copenhagen, Denmark, May 2003.

[8] J. Daniel, R. Nicol, S. Moreau, "Further Investigations of High Order Ambisonics and Wave field Synthesis for Holophonic Sound Imaging," in *Proc. 114th Conv. AES*, Amsterdam, 2003.

[9] S. Spors, J. Ahrens, "A comparison of wave field synthesis and higher-order ambisonics with respect to physical properties and spatial sampling," in *Proc. 125th Conv. AES*, San Francisco, CA, Oct. 2008.

[10] M. Poletti, "Three-Dimensional Surround System Based on Spherical Harmonics," *J. Audio Eng. Soc.*, vol. 53, no. 11, 1004–1025, 2005.

[11] J. Meyer, G. W. Elko, "A highly scalable spherical microphone array based on an orthonormal decomposition of the sound field," in *Proc. IEEE Int. Conf. Acous., Speech, Signal Process. (ICASSP)*, 2002, vol. II, pp. 1781–1784.

[12] T. D. Abhayapala, D. B. Ward, "Theory and design of high order sound field microphones using spherical microphone array," in *Proc. IEEE Int. Conf. Acous., Speech, Signal Process. (ICASSP)*, 2002, vol. II, pp. 1949–1952.

[13] Z. Li, R. Duraiswami, "Flexible and Optimal Design of Spherical Microphone Arrays for Beamforming," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 15, no. 2, 702–714, 2007

[14] B. Rafaely, "Analysis and design of spherical microphone arrays," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 13, no. 1, 135–143, 2005.

[15] I. Balmages, B. Rafaely, "Open-Sphere Designs for Spherical Microphone Arrays," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 13, no. 2, 727–732, 2007.