ISRA 2019 AMSTERDAM

PROCEEDINGS of the
International Symposium on Room Acoustics

15 to 17 September 2019 in Amsterdam, Netherlands

NAG NEDERLANDS AKOESTISCH GENOOTSCHAP

Direction-of-arrival distribution analysis of reflected sounds using spherical microphone array

Yuto IZUMI[1]; Makoto OTANI[2]

[1, 2] Kyoto University, Japan

ABSTRACT

Direction-of-arrival (DoA) of reflected sound is an important factor to characterize the spatial impression of room acoustics, including auditory source width and listener envelopment. Impulse response consists of direct sound, early reflections, and late reverberation. For the late reverberation, diffuse sound field is often assumed, although it has been suggested that DoA distribution of late reverberation is directionally biased, namely, anisotropic. The purpose of this study is to analyze the DoA distribution of late reverberation by employing a spherical microphone array and high-order plane wave decomposition, or spherical harmonics expansion, and decay cancellation. The DoA distributions were analyzed at one multipurpose hall using a spherical microphone array. Certain anisotropic DoA distributions were shown at any receiver position, which is more prominent as time goes by. In addition, the results implied that the direction toward which the DoA distribution may be biased depends on the acoustic absorption of the room as well as relative positioning between the receiver and the stage.

Keywords: Direction-of-arrival, Plane wave decomposition, Decay cancellation

1. INTRODUCTION

It is well known that apparent source width (ASW) and listener envelopment (LEV) are two of the main spatial components of spatial impression. Direction-of-arrival (DoA) of reflected sound is one of the important factors to characterize both of them (1). In particular, it is assumed that DoA distribution is significant as a factor of LEV. Bradley and Soulodre have shown that LEV is related to the level of late lateral reflections (2). Moreover, Hanyu et al. have reported that the early reflections arriving from above and frontal directions to a listener contribute to LEV and suggested that DoA distribution is necessary for evaluating LEV (3).

For the late reverberation, diffuse sound field is often assumed. However, Suzuki et al. have reported that DoA distribution of late reverberation is not isotropic using C-C method (4–6). The C-C method enables analysis of sound field at wideband frequencies, whereas it is not able to estimate the DoA of multiple sound waves that arrive simultaneously from different directions because it estimates a sound intensity vector at each time segment. In real sound fields, especially in the late reverberation, many sound waves arrive simultaneously or at very close intervals. For this reason, it can be considered that the C-C method is not suitable to analyze DoA distribution of late reverberation.

A spherical microphone array is one of the powerful devices to enable a spatial analysis of room impulse responses. DoA has been analyzed using a spherical microphone array (7–9). The plane wave decomposition is a method to represent the sound field as a sum of plane waves (10). Using this method, the analysis can be performed in more detail, while requiring a large number of microphones. Few studies using spherical microphone array have focused on the late reverberation as opposed to the early reflections (11–14). Several studies have analyzed the DoA distributions including not only the early reflections but also the late reverberation at multiple concert halls (9,15), however, they focused on a time segment of only 250 ms after the arrival of the direct sound.

This paper introduces a method to analyze DoA distribution of late reverberation based on plane wave decomposition and decay cancellation, and then demonstrates analysis results from measurements performed in one multipurpose hall.

[1] izumi.yuto.25a@st.kyoto-u.ac.jp
[2] otani@archi.kyoto-u.ac.jp

Peer reviewed

2. PLANE WAVE DECOMPOSITION

2.1 Plane wave decomposition

The degree n and order m spherical harmonics are defined by (16)

$$Y = \sqrt{\frac{2n+1}{4\pi}\frac{(n-m)!}{(n+m)!}} P_n^{|m|}(\cos\theta)e^{im\phi} \tag{1}$$

where ϕ, θ are the azimuth and elevation, respectively, $i = \sqrt{-1}$, and $P_n^{|m|}(\cdot)$ represents associated Legendre function. Spherical harmonics are orthogonal and satisfy the following equation,

$$\int_{\Omega} Y_{n'}^{m'}(\phi,\theta) Y_n^m(\phi,\theta)^* \Omega = \delta_{nn'}\delta_{mm'} \tag{2}$$

Consider the sound pressure of a unit-amplitude plane wave which arrives from the direction (ϕ_l, θ_l) with wave number k. The sound pressure measured at (ϕ, θ, r) is given by (16)

$$p_l(\phi,\theta,r) = \sum_{n=0}^{\infty}\sum_{m=-n}^{n} b_n(kr) Y_n^m(\phi,\theta) Y_n^{m*}(\phi_l,\theta_l) \tag{3}$$

where b_n is a radial function. For a rigid spherical array, b_n is given by

$$b_n(kr) = 4\pi i^n \left(j_n(kr) - \frac{j_n'(kr)}{h_n'(kr)} h_n(kr) \right) \tag{4}$$

where j_n, h_n are the spherical Bessel and Hankel functions, respectively, and j_n', h_n' are their derivatives, respectively.

In the following, the formulation of plane wave decomposition is introduced with reference to (10,13). Assuming that the sound field is composed by infinite number of plane waves, the sound pressure measured at (ϕ, θ, r) can be written as

$$p(\phi,\ \theta, r) = \int_{\Omega_l \in S^2} w_l\, p_l(\phi,\theta,r) d\Omega_l \tag{5}$$

where w_l is the plane wave coefficients. Using spherical harmonics decomposition, w_l in the spherical harmonics domain can be written as

$$w_l = \sum_{n=0}^{\infty}\sum_{m=-n}^{n} w_{nm} Y_n^m(\phi_l,\theta_l) \tag{6}$$

Using Eq. (6) and the orthogonality equation, Eq. (2), Eq. (5) is represented as

$$p(\phi,\theta,r) = \sum_{n=0}^{\infty}\sum_{m=-n}^{n} w_{nm} b_n(kr) Y_n^m(\phi,\theta) \tag{7}$$

Furthermore, the spherical harmonics decomposition of $\mathrm{p}(\phi, \theta, r)$ can be written as

$$p(\phi,\theta,r) = \sum_{n=0}^{\infty}\sum_{m=-n}^{n} p_{nm} Y_n^m(\phi_l,\theta_l) \tag{8}$$

where p_{nm} is calculated using Eq. (2) as

$$p_{nm} = \int_{\Omega} p(\phi,\theta,r) Y_n^m(\phi,\theta)^* d\Omega \tag{9}$$

From Eqs. (7) and (8),

$$w_{nm} = p_{nm}/b_n(kr) \tag{10}$$

Substituting Eq. (10) in Eq. (6) yields

$$w_l = \sum_{n=0}^{\infty}\sum_{m=-n}^{n} \frac{p_{nm}}{b_n(kr)} Y_n^m(\phi_l,\theta_l) \tag{11}$$

Here w_l represents the amplitude and phase of the plane wave arriving from the direction (θ_l, ϕ_l). Therefore we can estimate the DoA distribution from this equation.

2.2 Truncation order

In most practical situations, w_l is calculated from the following formula:

$$w_l = \sum_{n=0}^{N} \sum_{m=-n}^{n} \frac{p_{nm}}{b_n(kr)} Y_n^m(\phi_l, \theta_l) \tag{12}$$

where N is the truncation order. In case that the microphones are distributed equally, to ensure a truncation of spherical harmonics decomposition to N, it is necessary to use M microphones, where (17):

$$M \geq (N+1)^2, \tag{13}$$

and the spatial resolution ψ_0 [rad] is given by

$$2\psi_0 \simeq 2\pi/N. \tag{14}$$

The radial function b_n, Eq. (4), are plotted in Fig. 1, demonstrating that the amplitude approaches to zero as the order of $b_n(kr)$ increases when kr is small. To prevent divergence of $1/b_n(kr)$ in Eq. (12), $b_n(kr) \gg 0$ is required. Moreover, in order to avoid the spatial aliasing, the upper frequency is limited by $kr < N$.

For the reasons mentioned above, we analyze the DoA distributions with combinations of frequencies and truncation orders shown in Table 1, and Fig. 2 shows their directivity patterns. It is depicted that the higher the truncation order is, the sharper the directivity is.

3. DECAY CANCELLATION

In this paper, we take an average of amplitudes of the plane waves in each direction over time for a period of interest, such as from 100 ms to 1000 ms after the direct sound arrives, in order to analyze the DoA distribution in the late reverberation. However, the results reflect only the early part in the time window because the amplitude attenuates as time passes. Then we introduce decay cancellation (18).

Consider the decay cancellation of the impulse response $p(t)$. The decay curve $E_s(t)$ of $p(t)$ is given by

$$E_s(t) = \int_t^{\infty} p(\tau)^2 d\tau \tag{15}$$

Then the decay cancelled impulse response $g(t)$ is defined as

$$g(t) = p(t)/\sqrt{E_s(t)} \tag{16}$$

It should be noted that the decay cancellation enables us to extract only DoA information by removing the decay effects, but, at the moment, it is not clear that this approach is suitable for evaluation of subjective spatial impression.

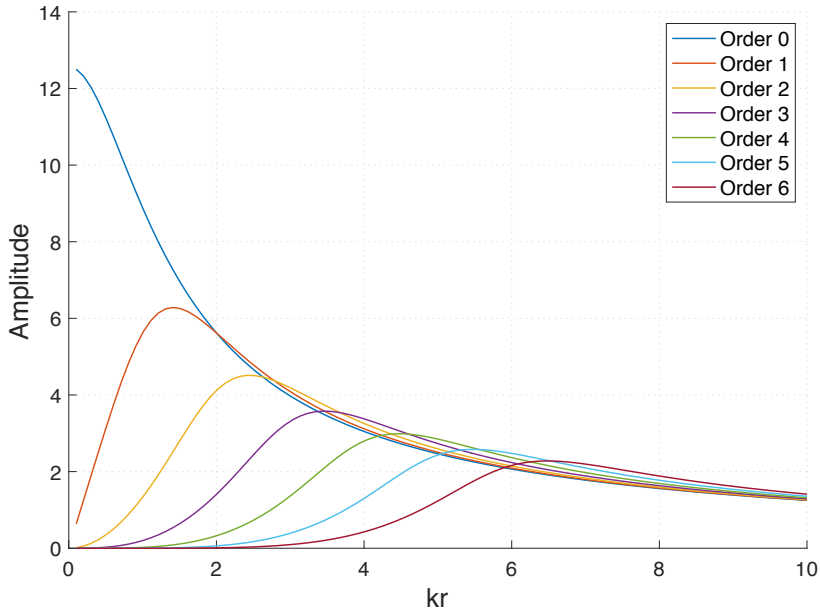

Figure 1 – Amplitude of $b_n(kr)$

Table 1 – Combination of frequency and truncation order

Frequency [Hz]	Order	Frequency [Hz]	Truncation order
1,000	2	4,000	4
2,000	2	5,000	5
3,000	3	6,000	6

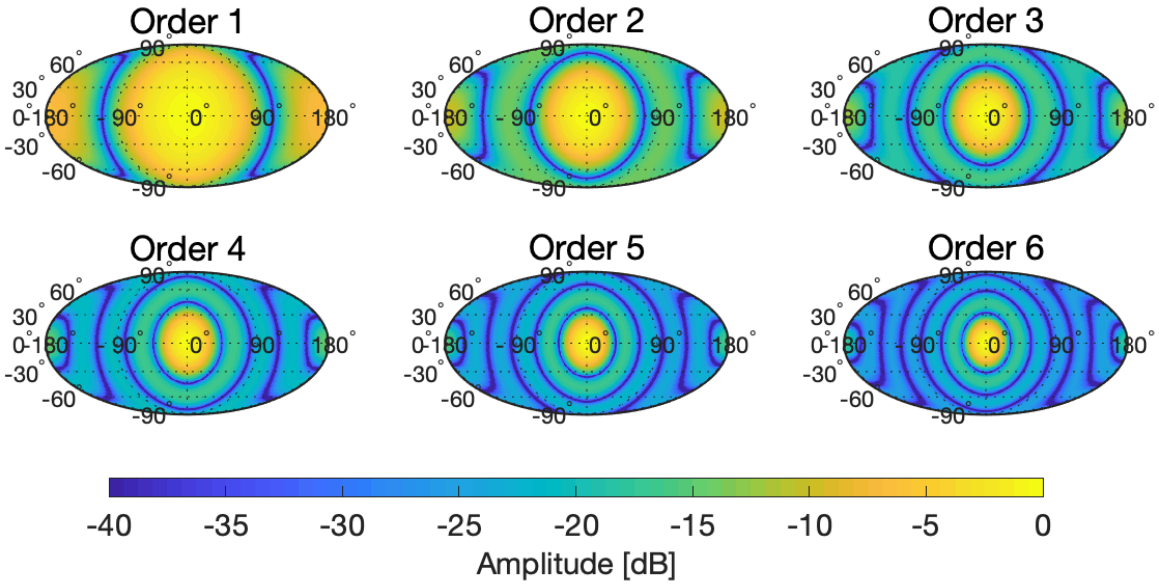

Figure 2 – Directivity pattern. The scale of colormap is normalized to the maximum.

4. MEASUREMENTS

4.1 Measurement set

A spherical rigid microphone array, with 64 omnidirectional microphones arranged equally on a sphere based on Fibonacci-spiral (19) and 50 mm radius, was used in this measurement. Impulse responses were measured in a multipurpose hall, Hardy Hall in Doshisha University, Japan. In the measurement, 9 measurement positions were chosen, which are shown as P1-P9 in Fig. 3. The center of the microphone array was placed 1.2 m above the floor at each position. A loudspeaker (Genelec, 1037C) was placed on a stand to be 45 degrees upwards at the center of the stage, 5.3 m from the rear wall of the stage. The height of the bottom of the loudspeaker from the floor was 0.25 m. A 20-seconds log-TSP (Time Stretched Pulse) is output from the loudspeaker. It should be noted that the sound field is not driven with an omnidirectional source, which might cause a bias of DoA distribution. The sampling frequency is 48,000 Hz. Figure 4 shows the reverberation time (T30), which is calculated by the omnidirectional component of impulse responses which were measured using the spherical microphone array and an impulse response which was measured using an omnidirectional microphone in another day, respectively. The SNRs for RIRs measured at P2 with the spherical microphone array and the omnidirectional microphone are about 55 dB and 80 dB, respectively. The differences between estimated reverberation times between two microphones may be attributable to the difference of SNR between the microphones and/or measurement dates.

4.2 Analysis procedure

In the first step, the impulse responses measured using the spherical microphone array are converted to time-frequency domain by using short time Fourier transform (STFT). Here, 256-samples Hanning window with 50 % overlapping is applied. Second, plane wave decomposition, mentioned in Section 2, is applied to the signals. Third, the decay is canceled by using decay cancellation as described in Section 3. In the decay cancellation, the decay curve in Eq. (12) is calculated by using the omnidirectional component of the signals obtained in the previous step. Finally, the signals are extracted for a period of interest and then are averaged. The results are shown in color maps using Mollweide projection, in which latitude and longitude correspond to elevation and azimuth,

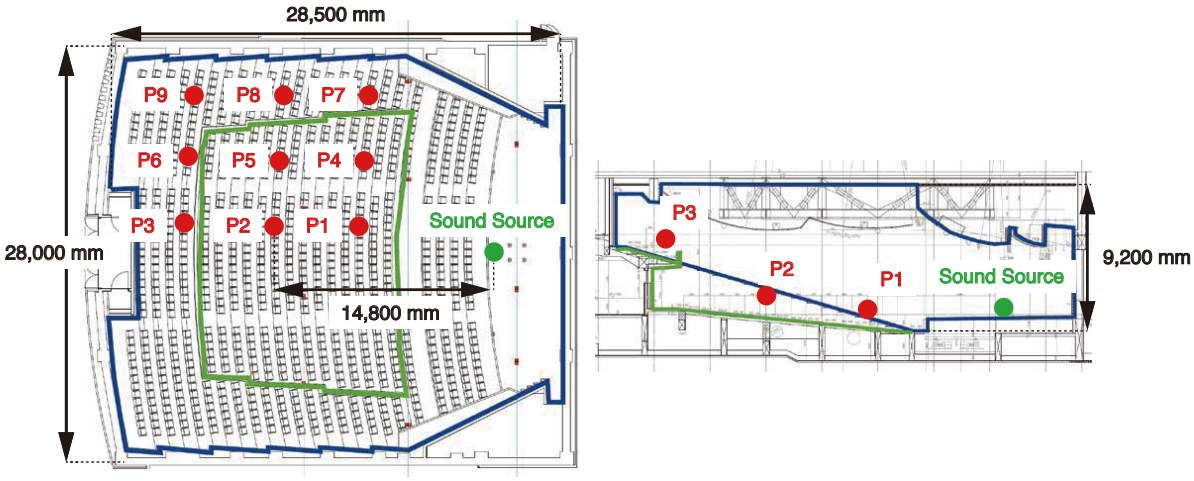

Figure 3 – Hardy Hall, left: plan, right: section

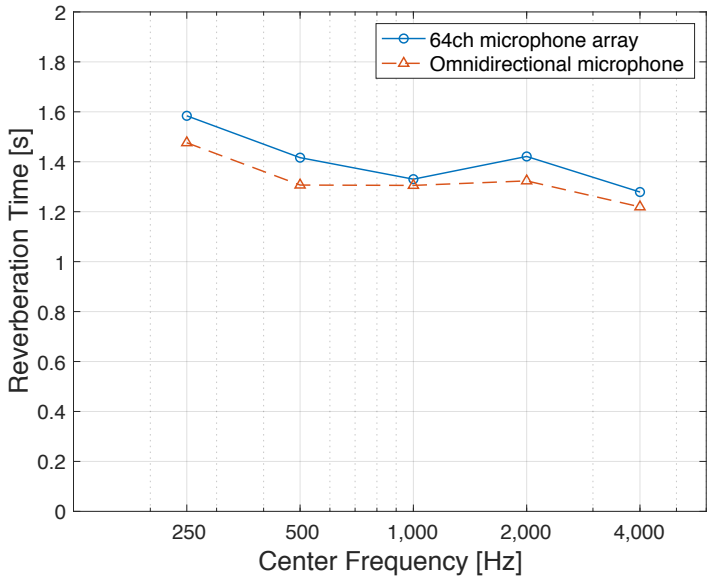

Figure 4 – Reverberation Time

respectively; a direction of 0 degrees azimuth and 0 degrees elevation corresponds to frontal direction; the mark x represents the source direction.

4.3 Results

4.3.1 Comparison among frequencies

Figure 5 depicts the DoA distributions analyzed at each frequency bin from 1,000 Hz to 6,000 Hz with intervals of 1,000 Hz. The impulse responses are measured at P2, and they are analyzed with the time window which includes the signals from 100 ms to 1,000 ms after the arrival of direct sound. The scale of colormap is normalized to the maximum. It can be found that the distributions are biased toward the front with 0 degrees elevation for 3,000 Hz or more than 3,000 Hz. On the other hand, for 1,000 Hz and 2,000 Hz, the distributions are almost uniform, which are attributed to the less sharp directivity pattern at the lower order as indicated in Fig. 2.

4.3.2 Comparison among measurement positions

Figure 6 depicts the DoA distributions measured at P1 ~ P9, respectively. The impulse responses are analyzed at 5,000 Hz with truncation order of 5 and the time window from 100 ms to 1,000 ms.

The comparison among these results illustrates that the directions toward which the DoA distributions are biased are slightly different among the measurement positions. Furthermore, the direction of the sound source is included in the region in which plane waves with larger energy arrive although the measurement positions are different. The elevations of the DoA distributions are almost 0 degrees at 90 and −90 degrees azimuth. At 0 degrees azimuth, on the other hand, the elevations of the DoA distributions are lower than 0 degrees for P3, P6, and P9.

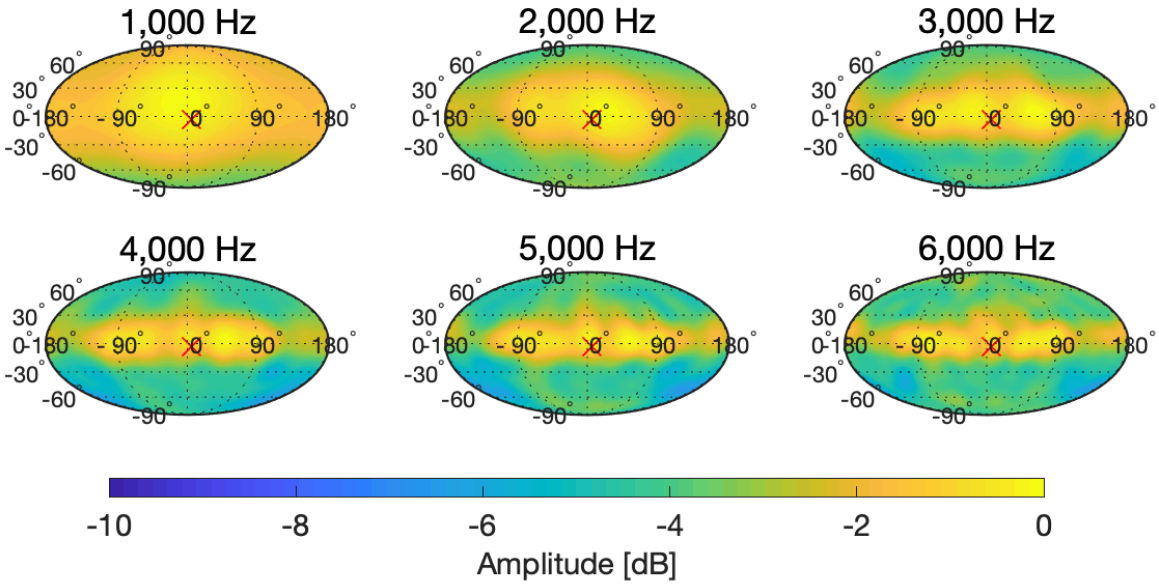

Figure 5 – Comparison among frequencies (P2, 100 ~ 1,000 ms). The scale of colormap is normalized to the maximum. The cross represents the source position.

4.3.3 Comparison among temporal segments

Figure 7 depicts the time transition of the DoA distributions with 80 ms time windows. The impulse responses were measured at P3 and are analyzed at 5,000 Hz with truncation order of 5.

The figure suggests that the anisotropic DoA distribution is more prominent as time passes, and the directions from which plane waves with larger energy arrive are similar at each time segment after a certain period of time.

4.4 Discussions

In Figure 5, the DoA distributions for 1,000 Hz and 2,000 Hz are almost uniform. However,

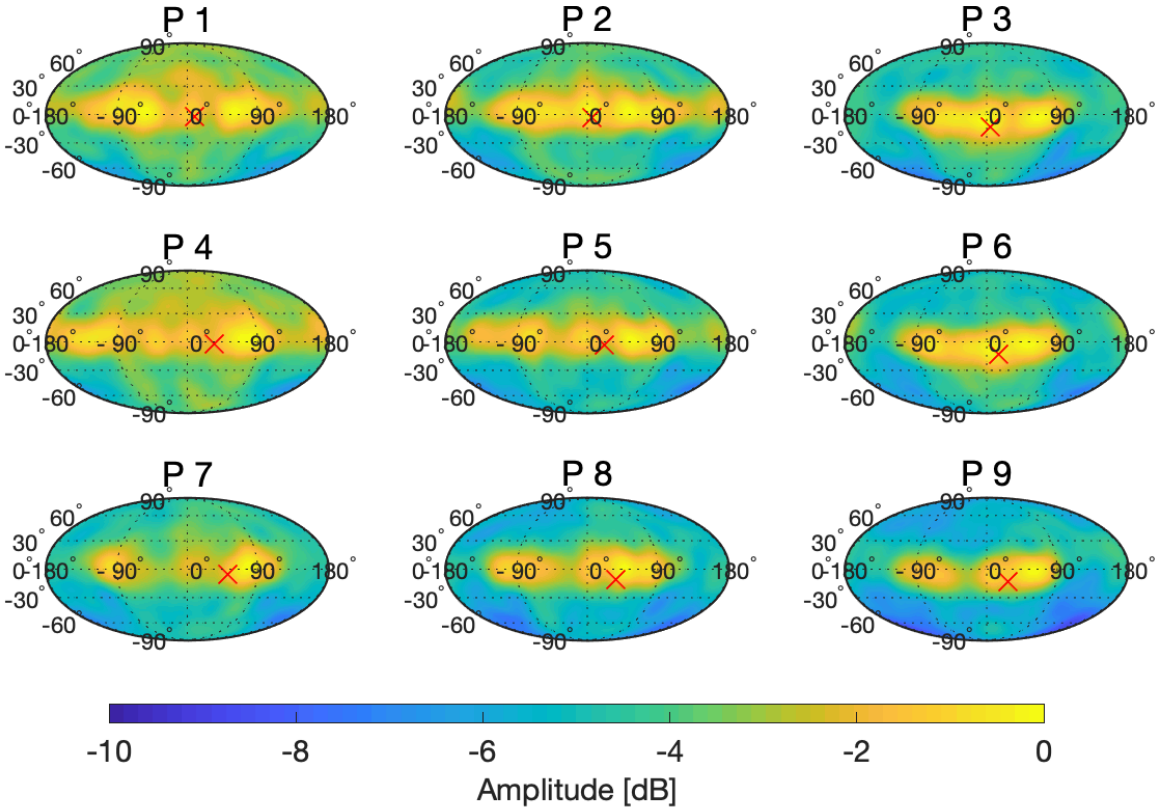

Figure 6 – Comparison among measurement places (5,000 Hz, order 5th, 100 ~ 1,000 ms). The scale of colormap is normalized to the maximum. The cross represents the source position.

according to Eq. (14), the spatial resolution is about 180 degrees with 2nd order, which means that the analysis is not capable to distinguish two incoming plane waves whose relative angle is closer than the resolution angle. Therefore, it is possible that the results with the low order do not present the DoA distributions accurately.

The comparison of the results among measurement positions in Figure 6 implies that the directions toward which the DoA distributions may be biased depending on the direction of the sound source. Furthermore, the DoA distributions are biased toward front hemisphere. There are upholstered seats behind the microphone array at around 0 degrees elevation. Thus, it can be considered that the DoA distributions of late reverberation may be affected by the acoustic absorptions of the room.

In a previous work, the DoA distribution of late reverberation is analyzed with C-C method in a concert hall (5). In that case, it has been reported that the DoA distribution is biased toward front hemisphere and diagonally upward direction. The elevations toward which the DoA distributions are biased are different from the results obtained in this work. However, it is difficult to discuss the causes because the IRs were measured at different rooms and with a different method.

5. CONCLUSIONS

In this paper, a method to analyze DoA distribution of late reverberation has been introduced, which employs a spherical microphone array, plane wave decomposition, and decay cancellation. The analysis from measurements performed in one multipurpose hall has been conducted. The results showed that the DoA distributions of late reverberation are anisotropic at any frequency and any measurement position, which is more prominent as time passes. Moreover, the results implies that the DoA distributions may be affected by the acoustic absorption of the room as well as relative

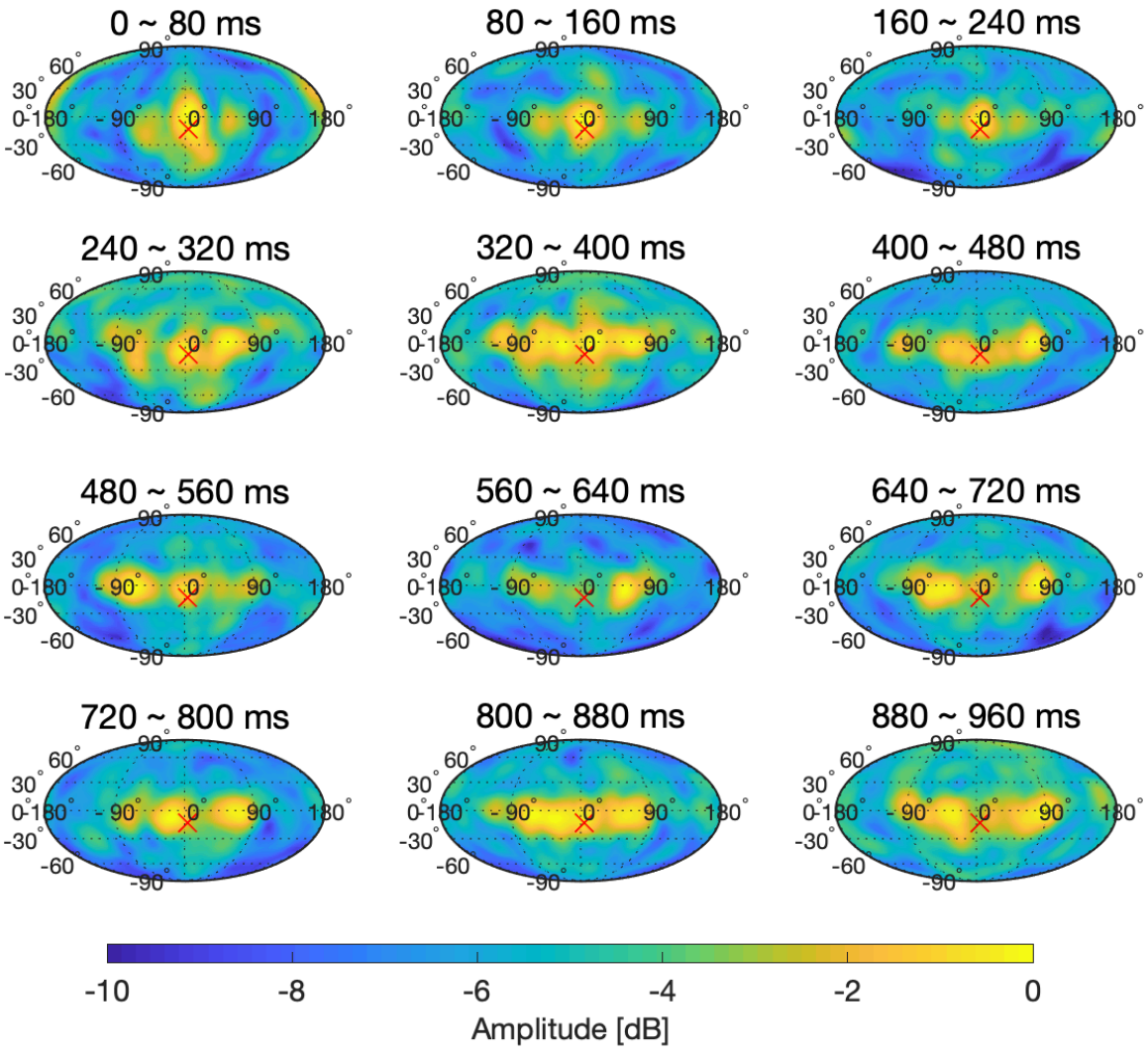

Figure 7 – Comparison among time segments (P3, 5,000 Hz, order 5th). The scale of colormap is normalized to the maximum. The cross represents the source position.

positioning between the receiver and the sound source.

The present results were derived only from one multipurpose hall with a non-omnidirectional source. To find a general tendency of DoA distribution, the analysis in other rooms should be performed with an omnidirectional source. Additionally, future studies include verification of the relationship between DoA distribution of late reverberation and the auditory impression.

ACKNOWLEDGEMENTS

The authors are grateful to Prof. Takao Tsuchiya, Doshisha University, and Mr. Hiraku Okumura, Yamaha Corporation, for collaborating at the measurements in Hardy Hall. This work was partially funded by a JSPS Grant (No.19H04153 and 19H04145).

REFERENCES

1. Morimoto M, Fujimori H, Maekawa J. Discrimination on between auditory source width and envelopment. J. Acoust. Soc. Jpn. 1990;46(6):449–57.
2. Bradley JS, Soulodre GA. The influence of late arriving energy on spatial impression. J. Acoust. Soc. Am. 1995;97(4):2263–71.
3. Hanyu T, Kimura S, Chiba S. A new objective measure for evaluation of listener envelopment focusing on the spatial balance of reflections. J. Arch. Plann. Env. Eng. 1999;64(520):9–16.
4. Hanyu T, Inage D, Sekiguchi K. Measurement of directional information of sound field by 4-channel cardioid microphones. In Proc. Meet. Acoust. Soc. Jpn. 2008;1123–4.
5. Suzuki R, Hoshi K, Hanyu T. Analysis of directional distribution of arriving sound energy using instantaneous intensity and decay-cancelled impulse response. In: Summaries of technical papers of annual meeting. Archi. Inst. Jpn. 2014;251–2
6. Suzuki R, Hoshi K, Hanyu T. Analysis of frequency characteristics of room sound field using the decay-cancelled impulse response. In: Summaries of technical papers of annual meeting. Archi. Inst. Jpn. 2015;
7. Rafaely B. Phase-mode versus delay-and-sum spherical microphone array processing. IEEE Signal Process Lett. 2005;12(10):713–6.
8. Khaykin D, Rafaely B. Acoustic analysis by spherical microphone array processing of room impulse responses. J. Acoust. Soc. Am. 2012;132(1):261–70.
9. Gover BN, Ryan JG, Stinson MR. Measurements of directional properties of reverberant sound fields in rooms using a spherical microphone array. J. Acoust. Soc. Am. 2004;116(4):2138–48.
10. Meyer J, Elko G. A highly scalable spherical microphone array based on an orthonormal decomposition of the soundfield. In: 2002 IEEE International Conference on Acoustics, Speech, and Signal Processing. 2002. p. II-1781-II–1784.
11. de Araujo FH, Castro Pinto FA de N, Boscher Torres JC. Room reflections analysis with the use of spherical beamforming and wavelets. Appl Acoust. 2018;131:192–202.
12. Sun H, Mabande E, Kowalczyk K, Kellermann W. Localization of distinct reflections in rooms using spherical microphone array eigenbeam processing. J. Acoust. Soc. Am. 2012;131(4):2828–40.
13. Park M, Rafaely B. Sound-field analysis by plane-wave decomposition using spherical microphone array. J. Acoust. Soc. Am. 2005;118(5):3094–103.
14. Rafaely B. Plane-wave decomposition of the sound field on a sphere by spherical convolution. J. Acoust. Soc. Am. 2004;116(4):2149–57.
15. Pätynen J, Tervo S, Lokki T. Analysis of concert hall acoustics via visualizations of time-frequency and spatiotemporal responses. J. Acoust. Soc. Am. 2013;133(2):842–57.
16. Williams EG. Fourier Acoustics: Sound Radiation and Nearfield Acoustical Holography. Elsevier; 1999. p.190-193, p.224-230.
17. Rafaely B. Analysis and design of spherical microphone arrays. IEEE Trans. Speech Audio Process. 2005;13(1):135–43.
18. Hanyu T. Analysis Method for Estimating Diffuseness of Sound Fields by Using Decay-Cancelled Impulse Response. Build. Acoust. 2014;21(2):125–33.
19. Kaneko S, Suenaga T, Akiyama H, Sekine S. A microphone array for 3-dimensional sound field recording, based on the Fibonacci-spiral. In Proc. Meet. Soc. Jpn. 2017;715–6.