PROCEEDINGS of the
23rd International Congress on Acoustics

9 to 13 September 2019 in Aachen, Germany

Reproduction of multichannel audio by frame loudspeaker array

Yuta KAKUZAKI[1]; Akio ANDO[2]

[1] University of Toyama, Japan

[2] University of Toyama, Japan

ABSTRACT

Multichannel audio reproduces a sound field by loudspeakers surrounding a listener, which necessarily occupies the living environment in the home. The purpose of this study is to reproduce the sound by loudspeakers placed around the image display, which is referred to as "frame loudspeaker array". In this paper, we focused on the reproduction of the frontal channels of 22.2 multichannel by the array. Each channel signal of the multichannel audio was distributed into loudspeakers with the weighting coefficients that minimized the square error between the original sound field by the channel and the reproduced sound field by the array. The sound fields were represented by the tenth order spherical functions. The signal was passed through the auditory band pass filter and changed its amplitude and delay according to the weighting coefficient calculated for the center frequency of the filter. The band pass signals were then summed up to generate the full band input signal for loudspeakers. The informal listening test showed that this method could preserve the spatial feature of the original multichannel audio.

Keywords: Sound field synthesis, Loudspeaker array, Multichannel audio

1. INTRODUCTION

Multichannel audio has increased the number of loudspeakers to provide a higher sense of reality. It usually requires loudspeakers surrounding a listener. For example, 22.2 multichannel uses three layers of loudspeakers to reproduce a three-dimensional sound field [1]. However, the number of loudspeakers must be limited in order to avoid occupying the living environment in the home. This problem has become serious because the regular broadcasting of 22.2 multichannel has been started with 8K video from December 2018 in Japan. One of the methods to solve this problem may be a use of loudspeaker array set up in the frontal direction.

Over the past few decades, a considerable number of studies have been made on synthesizing a sound field by the loudspeaker array. A wave field synthesis (WFS) is known as a method to recreate a sound filed [2-4]. Particularly, the WFS can recreate a 3D sound field using a line loudspeaker array with a 2.5D Rayleigh integral [4], which models 3D sound propagation on the basis of the stationary phase approximation. However, it may not reproduce the vertical sound stage of 22.2 multichannel. On the other hand, a plane loudspeaker array is not available by the existence of the image display.

In this study, we used a loudspeaker array set on the frame of the image display, which is referred to as "frame loudspeaker array" [5]. We focused on the reproduction of the frontal eleven channels of 22.2 multichannel by the array in this paper. A sound field created by each frontal loudspeaker of 22.2 multichannel (hereinafter referred to as "target field") was synthesized by the frame loudspeaker array using the spherical control area centering on the listener. In the synthesis, the square error of sound pressures of both sound field, target field and synthesized field, was minimized over the control area using the least squares method. The sound fields were represented by the tenth order spherical functions. The resultant weighting coefficient for each loudspeaker of the array was divided into the amplitude and delay information, where the delay was obtained after eliminating the discontinuities of the arctangent function. Each channel signal of 22.2 multichannel was passed through the auditory band pass filter and changed its amplitude and delay according to the weighting coefficient calculated for the center frequency of the filter. The band pass signals were then summed up to generate the full band input signal for loudspeakers.

[1] m1971017@ems.u-toyama.ac.jp
[2] andio@eng.u-toyama.ac.jp

2. SOUND FIELD CONTROL

2.1 Frame Loudspeaker Array

The frame loudspeaker array used in this study was designed so that an image display whose size was 1.78m (W)×1.00m (H) could be installed at the center. The number of loudspeakers was twenty considering the size of loudspeaker. The numbers were seven in horizontal and five in vertical arranged at equal intervals. Figure 1 shows the experimental frame loudspeaker array. In this paper, it is assumed that loudspeakers are arranged as shown in Fig. 2.

Figure 1 – Frame loudspeaker array

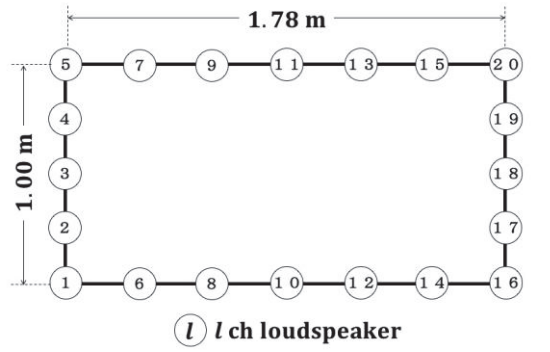

Figure 2 – Frame loudspeaker array for simulation

2.2 Experimental setup

The listening position was set to d [m] in front of the array. The control area is a spherical area centered on the listening point and its radius was set to R [m]. In this study, assuming that the volume of the spherical control area is $1\,\mathrm{m}^3$, the radius R was set to 0.62 m. The relationship between the array and the control area is shown in Fig. 3.

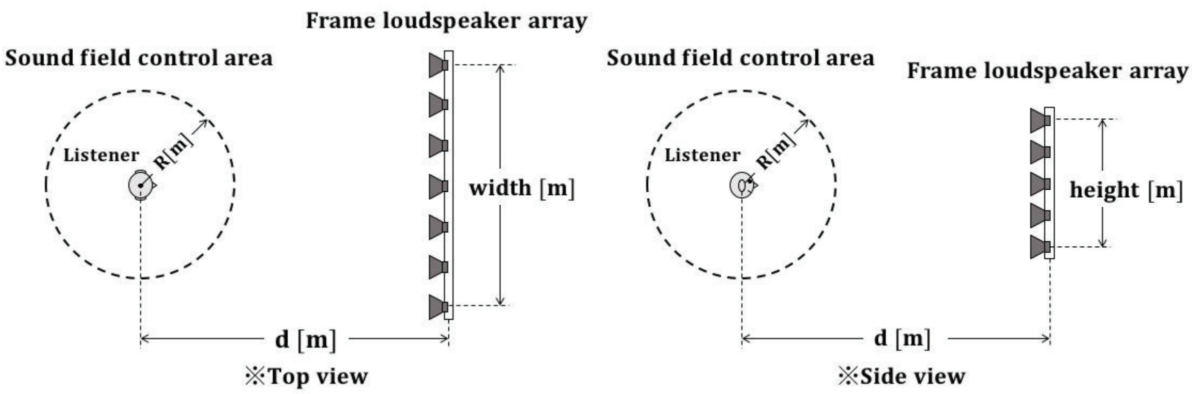

Figure 3 – Experimental setup

2.3 Sound Field Control Function

Each sound source is assumed to be a monopole sound source and is represented by spherical functions. The sound pressure of the target sound field $p_s(\mathbf{r})$, which is a sound field by one of the channels of the multichannel audio, and the sound pressure of the synthesized sound field $p_r(\mathbf{r})$ are shown as

$$p_s(\mathbf{r}) = Q\frac{e^{-ik|\mathbf{r}-\mathbf{r}_s|}}{4\pi|\mathbf{r}-\mathbf{r}_s|} \cong ikQ\sum_{n=0}^{D} j_n(kr)h_n^{(2)}(kr_s)\sum_{m=-n}^{n} Y_n^m(\theta_s,\varphi_s)\overline{Y_n^m(\theta,\varphi)}\,, \tag{1}$$

$$p_r(\mathbb{r}) = \sum_{l=1}^{L} w_l \frac{e^{-ik|\mathbb{r}-\mathbb{r}_l|}}{4\pi|\mathbb{r}-\mathbb{r}_l|} \cong ik \sum_{l=1}^{L} w_l \sum_{n=0}^{D} j_n(kr) h_n^{(2)}(kr_l) \sum_{m=-n}^{n} Y_n^m(\theta_l, \varphi_l) \overline{Y_n^m(\theta, \varphi)}\,, \tag{2}$$

respectively, where $w_l (l = 1, \cdots, L)$ are weighting coefficients that should be solved so as to minimized the error function [6]

$$E(\mathbb{w}) = \iiint_V \{p_r(\mathbb{r}) - p_s(\mathbb{r})\}^2 dV\,. \tag{3}$$

Note that $\mathbb{r}_s = (r_s, \theta_s, \varphi_s)$ is a position of a target sound source, $\mathbb{r}_l = (r_l, \theta_l, \varphi_l)$ is a position of the lth-channel loudspeaker of the array, $\mathbb{r} = (\mathrm{r}, \theta, \varphi)$ is a position of the observation, k is wave number, D is the maximum order of the spherical functions, L is the number of sound sources in the array, Q is the amplitude of the target sound source, j_n is a spherical Bessel function of the first kind, $h_n^{(2)}$ is a spherical Hankel function of the second kind, and Y_n^m is a spherical harmonics function.

Using the least squares method, the weighting coefficients can be obtained by a linear equation of the following form:

$$\mathbb{U}\mathbb{w} = \mathbb{s}\,, \tag{4}$$

where $\mathbb{w}$ is a vector of the weighting coefficient:

$$\mathbb{w} = \begin{bmatrix} w_1 \\ \vdots \\ w_L \end{bmatrix} = \begin{bmatrix} \alpha_1 + i\beta_1 \\ \vdots \\ \alpha_L + i\beta_L \end{bmatrix}, \tag{5}$$

and $\mathbb{U}$ is a matrix whose elements are spatial parameters of the synthesized field. Note that the vector $\mathbb{w}$ is frequency dependent, meaning that it should be calculated for each frequency. The synthesized sound field is controlled by the weighting coefficient calculated by

$$\mathbb{w} = \mathbb{U}^{-1}\mathbb{s}\,. \tag{6}$$

Figure 4 shows the expected sound field controlled by this method.

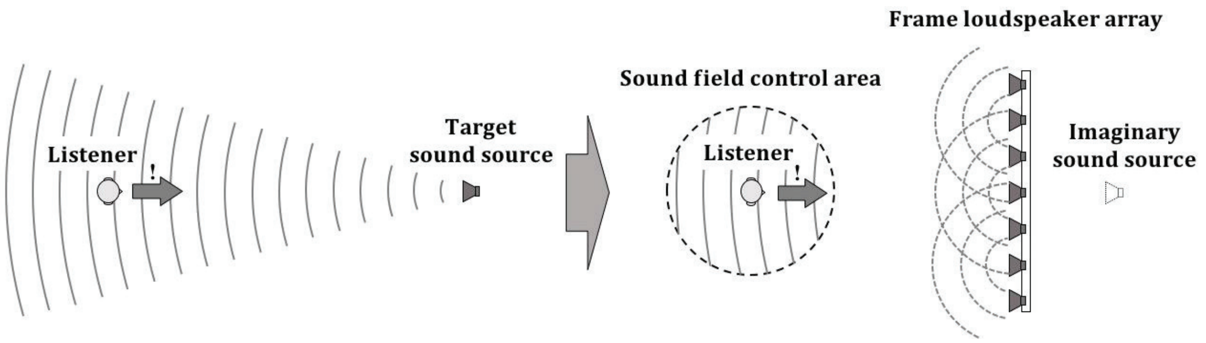

Figure 4 – Sound field control by frame loudspeaker array

Since the weighting coefficients calculated by Eq. (6) depends on the wave number, it changes with the frequency of the target sound source. An example of the frequency characteristic of the coefficient is shown in Fig. 5. It was found to be unstable behavior (1 to 6) in the low frequency band and stable behavior (rotating 7 to 8) in the high frequency band. This is caused by the fact that the condition number of the matrix in Eq. (6) increases in the low band because the long wavelength brings similar lows and columns in the matrix $\mathbb{U}$. Section 3 describes how to improve it. The weighting coefficient vector $\mathbb{w}$ controls the synthesized sound field. Therefore, it will be referred to as the sound field control function hereinafter.

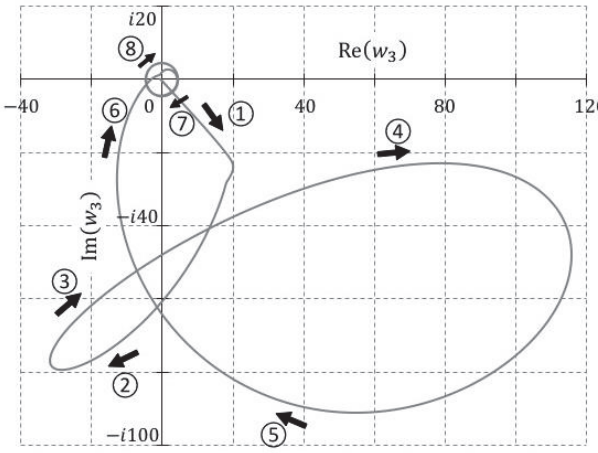

Figure 5 – Behavior of sound field control function

2.3.1 Amplitude Estimation and Control

Sound field control function determines the amplitude of each sound source. Denoting the real and imaginary parts of the function of lth-channel loudspeaker by α_l and β_l, respectively, the amplitude is obtained by

$$A_l = \sqrt{{\alpha_l}^2 + {\beta_l}^2}\ . \tag{7}$$

Amplitude control is performed by applying this equation to the corresponding sound source.

2.3.2 Delay Estimation and Control

Sound field control function determines the delay of each sound source. The phase angle for lth-channel may be given using an inverse tangent function as

$$\varphi_l = tan^{-1}\frac{\beta_l}{\alpha_l}\ . \tag{8}$$

This may cause phase inversion due to the nature of the arctangent function. To avoid the phase inversion, we decided the phase using the equation

$$\varphi_l = \begin{cases} tan^{-1}\frac{\beta_l}{\alpha_l} & (\alpha_l \geq 0 \cap \beta_l \geq 0、\alpha_l \geq 0 \cap \beta_l \leq 0) \\ tan^{-1}\frac{\beta_l}{\alpha_l} - \pi & (\alpha_l < 0 \cap \beta_l < 0) \\ tan^{-1}\frac{\beta_l}{\alpha_l} + \pi & (\alpha_l < 0 \cap \beta_l > 0)\ . \end{cases} \tag{9}$$

Furthermore, the relationship between the signal phase and the target source frequency should be considered. The range of possible values of the phase angle defined by Eq. (9) is $-\pi < \varphi_l < \pi$. However, according to the sound field control function shown in Fig. 5, the locus rotates around the origin, meaning that its angle should be beyond the interval $[-\pi, \pi]$. Therefore, we propose a method to extend the possible range of the phase angle using the phase shift of $2n\pi$ $(n = \cdots, -2, -1, 0, +1, +2, \cdots)$ in order to add the rotation to the phase angle. The resultant phase angle is converted into a delay time by the equation

$$\tau_l = -\frac{\varphi_l}{2\pi f}\ . \tag{10}$$

3. INPROVEMENT OF CONDITION NUMBER OF MATRIX

Sound field control function is sometimes unstable at low frequency as mentioned in section 2. The reason is that the condition number of the matrix $\mathbb{U}$ (shown in the Eq. (6)) becomes large in the low frequency band. In this section, we explain the reason and how to fix it.

Since the wave length of the sound wave is much longer than the size of the control area in the low frequency, the phase of the wave was approximately equal over the control area. As the result, each element of the matrix $\mathbb{U}$ approaches the similar value, and the condition number of the matrix becomes large. To avoid such instability of the matrix inversion, we propose a method that improves the condition number by properly determining the radius of the control area.

We changed the radius of the control area from 0.62m to 2.0m and observed the condition number of the matrix, and the amplitude and the delay. Figure 6 shows a reproduction set up used for the simulation. In the simulation, the distance from the center of the loudspeaker array to the center of the control area was set to 2.0m, and the imaginary source was set to 0.5m behind the array. The center of the control area, the center of the loudspeaker array, and the imaginary sound source are set on a straight line.

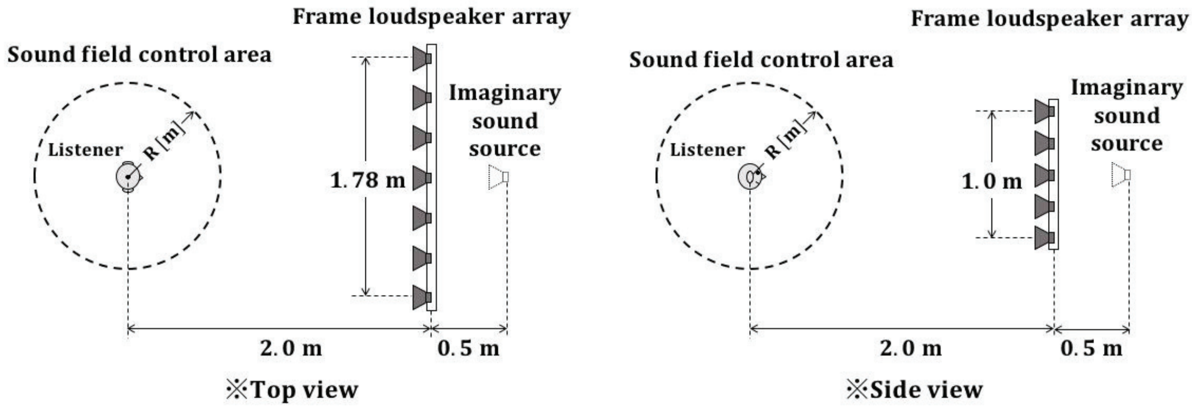

Figure 6 – Listening environment for the simulation

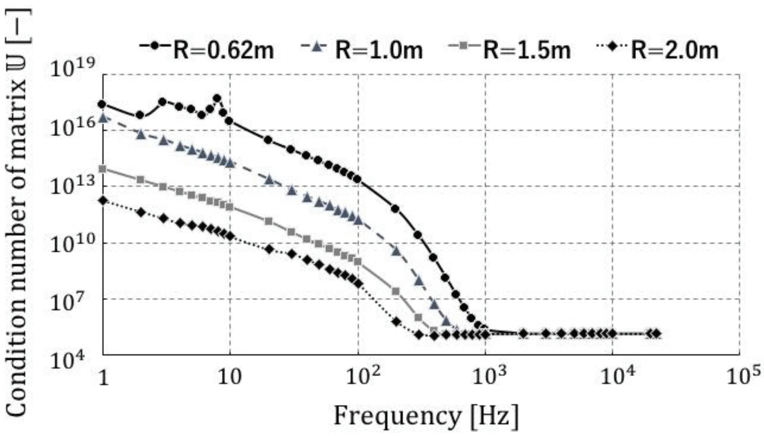

Figure 7 – Change of condition number of matrix $\mathbb{U}$ with respect to frequency when radius of control area R is changed

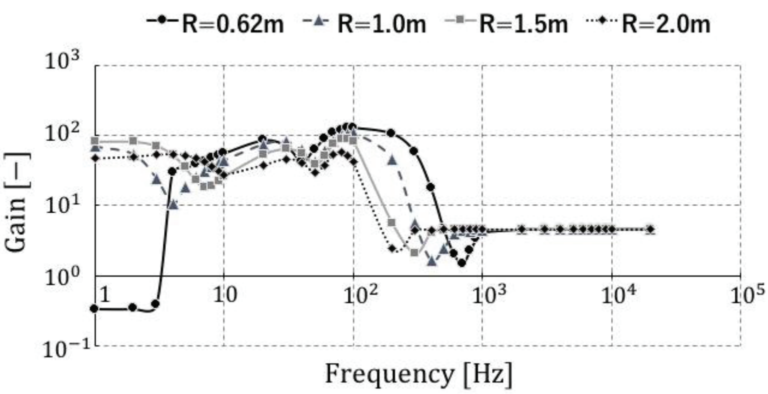

Figure 8 – Change of gain with respect to frequency when radius of control area R is changed

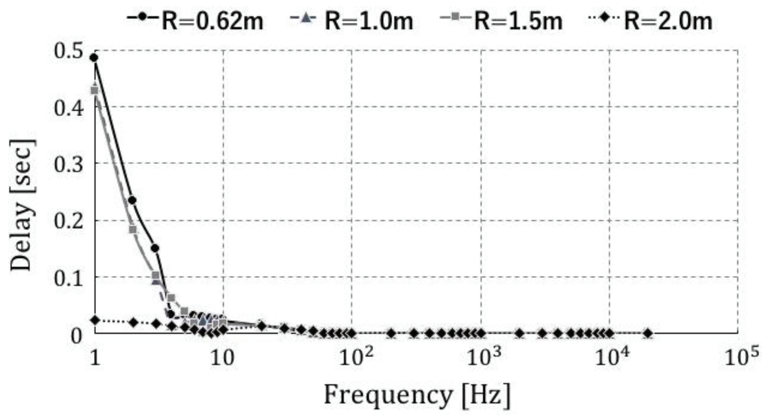

Figure 9 – Change of delay with respect to frequency when radius of control area R is changed

Figures 7, 8 and 9 show the condition number of the matrix $\mathbb{U}$, the amplitude of the weighting coefficient of the third loudspeaker A_3, the delay of the loudspeaker τ_3. The reason for choosing the third loudspeaker is because it has the largest amplitude in the simulation conditions. According to Fig. 7, as the radius of the control area increases, the condition number of the matrix $\mathbb{U}$ decreases particularly in low frequency band. Therefore, the calculation of the sound field control function becomes stable if the radius of the control area is increased. As shown in Fig. 8, when the radius of the control area increases, the frequency at which the gain becomes constant is lower. This frequency is almost the same as the frequency at which the condition number becomes constant as shown in Fig. 7. According to Fig. 9, when the radius of the control area is increased, the delay time becomes constant not only in the high band but also in the low band. From the above results, it was found that the sound field control function can be calculated more accurately by increasing the radius of the control area. However, it is necessary to set the radius of the control area so that the loudspeaker array and the video screen do not enter the area in consideration of the singularity of the sound pressure of the synthesized sound field.

4. MULTICHANNEL AUDIO REPRODUCTION

In this section, a method of reproducing the multichannel audio by creating imaginary sound sources at the channel positions is described. Signal processing for creating imaginary sound source is shown in Fig. 10.

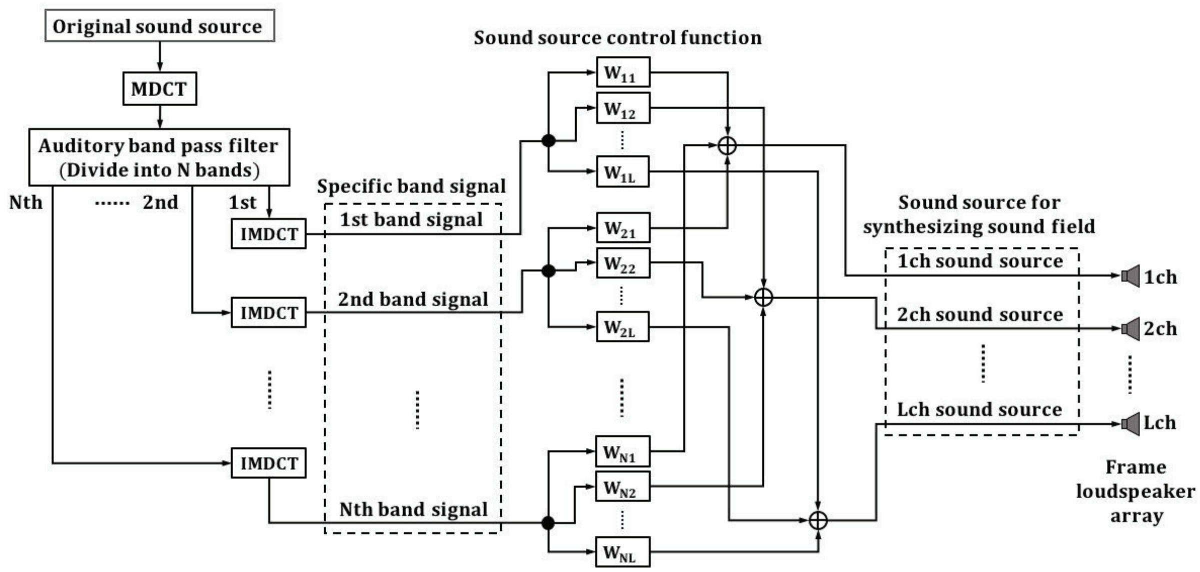

Figure 10 – Signal processing for creating imaginary sound source

An imaginary sound source can be created at any position by applying a sound field control function to the source. The function changes according to the frequency of the target sound source because it depends on the wave number. Since an audio signal has various frequency components, it is difficult to apply one function to all components. Therefore, we propose a method of dividing an audio signal into an auditory band based on MPEG-2 AAC [7], and applying a sound filed control function corresponding to the center frequency of each band to each band signal. In this study, the 69 auditory bands defined by MPEG-2 AAC is reconstructed into 23 bands such that the three consecutive subbands were merged into one.

4.1 Signal Processing for Creating Imaginary Sound Source

This section describes the signal processing procedure shown in Fig. 10. First, the audio signal is divided into band signals using the MDCT (Modified Discrete Cosine Transform), the auditory band pass filter and IMDCT (Inverse MDCT). Second, the imaginary sound source positions are set according to the loudspeaker positions of the multichannel audio. The sound field control function is calculated based on the source position. The amplitude and the delay are estimated for the center frequency of each subband. Third, each band signal is distributed into all loudspeakers of the array through applying the corresponding amplitude and delay. Finally, all band signals for each loudspeaker are summed up to generate the input signal for the loudspeaker. A desirable sound field may be created within the control area by reproducing the audio signal with the loudspeaker array.

4.2 Multichannel Audio by Imaginary Sound Source

In this section, we propose a method to reproduce the multichannel audio by multiple imaginary sound sources. The imaginary sound source is generated at the channel position of the multichannel audio in the manner of section 4.1. A multichannel sound field is synthesized by simultaneously reproducing all source signals. Figure 11 shows this situation for two-channel case.

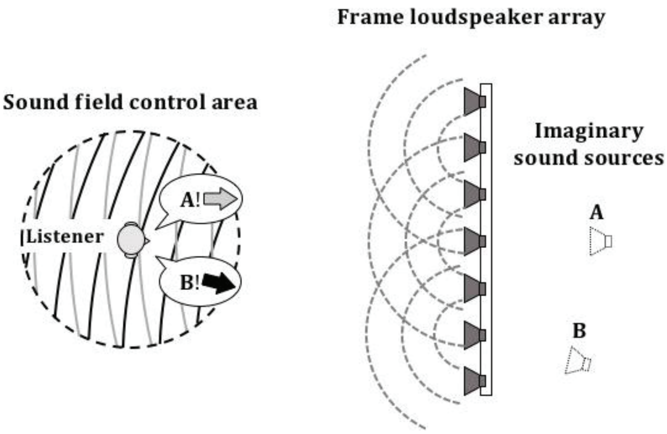

Figure 11 – Multichannel synthetic sound field

5. DISCUSSION

In this paper, the control method for a frame loudspeaker array was proposed. However, the comparison between the plane loudspeaker array, which has loudspeakers also inside the frame, and the frame loudspeaker array has not been studied yet. Such study may reveal the possibility and limitation of the frame loudspeaker array.

The method of expanding the control area was introduced to reduce the condition number of the matrix $\mathbb{U}$. Expanding the control area may degrade the sound field in the original area. Although this possible degradation should be studied by both the subjective and objective evaluations, it still remains unsolved.

The informal listening test showed that this method could preserve the spatial feature of the original multichannel audio. However, the formal listening test should be conducted. Such investigation will be carried out in the future.

6. CONCLUSION

This paper describes a method of reproducing a target sound field by controlling a frame loudspeaker array using a sound field control function optimized by the least squares method. The function varies depending on the target sound source frequency and has inaccuracies in the low band. The latter is due to the fact that the condition number of the matrix $\mathbb{U}$ used in the calculation becomes large. To reduce the condition number, we proposed a method to increase the radius of the spherical control area according to the source frequency. This method suppressed the condition number and improved the accuracy of the source field control function. In addition, decreasing the condition number brought the phenomenon that the amplitude and delay of the sound field control function converged to the fixed value. We also proposed a method to reproduce the multichannel audio signal with the frame loudspeaker array. The reproduction of channels other than the frontal channels of 22.2 multichannel audio will be studied in the next step.

ACKNOWLEDGEMENTS

This work was supported by Hoso Bunka Foundation 2017 Grant.

REFERENCES

1. Hamasaki K, Nishiguchi T, Okumura R, Nakayama Y and Ando A. A 22.2 Multichannel Sound System for Ultrahigh-Definition TV, SMPTE Motion Imaging J., pp. 40-49, 2008.
2. Berkhout A J. A holographic approach to acoustic control, J. Audio Eng. Soc., vol. 36, no. 12, pp. 977-995, 1988.
3. Berkhout A J, de Vries D, Vogel P. Acoustic control by wave field synthesis, J. Acoust. Soc. Am., vol. 93, no. 5, pp. 2764-2778, 1993.
4. de Vries D. Wave Field Synthesis, AES monograph. Audio Engineering Soc. 2009.
5. Ando A, Tokioka A. Control of loudspeaker array by minimizing fluctuation of frequency response and synthesized wave front, Proc. ICA2013, 055001, 2013.
6. Yamabayashi D, Yamoto T and Ando A. Control of frame loudspeaker array using spherical harmonic expansion, Proc. ASJ 2016 Autumn meeting, pp. 409-412, 2016.
7. ISO/IEC JTC1/SC29/WG11 standard. Text of ISO/IEC 13818-7:2005 (MPEG-2 AAC 4th edition), ISO/IEC JTC1/SC29/WG11, N7126, 2005.