Sound field reproduction system using narrow directivity microphones and boundary surface control principle

Hiroshi Kashiwazaki¹,* and Akira Omoto²,†

¹Graduate School of Design, Kyushu University, 4–9–1 Shiobaru, Minami-ku, Fukuoka, 815–8540 Japan
²Faculty of Design, Kyushu University, 4–9–1 Shiobaru, Minami-ku, Fukuoka, 815–8540 Japan

(Received 5 July 2017, Accepted for publication 18 October 2017)

Abstract: Boundary surface control (BoSC) is a useful method of reproducing the sound field physically. However, it is challenging in the case of real-time operation. This is mainly due to the calculation cost of a lot of inverse filter convolutions to obtain reproduction signals. This paper proposes a method for reducing of number of inverse filters and implements it in 24-channel narrow directivity (shotgun) microphone array and 24-channel circularly arranged loudspeaker array. Moreover, it provides an experimental evaluation of the reproduction accuracy according to measurement of reproduced wavefront. The accuracy of the reproduced wavefront by the filters, whose number was reduced to less than 1/5 by proposed method, was comparable with the case of full number of filters. Finally, a system aiming at sound field reproduction in a wide frequency range was constructed by a hybrid method of reproducing with an inverse filter in the low range and directly outputting from the speaker in the direction corresponding to the microphone in the high frequency range. We confirmed that real-time processing is possible for this hybrid method by using a convolution plug-in of digital audio workstation software.

Keywords: Sound field reproduction, Boundary surface control principle, Inverse filter matrix, Narrow directivity microphone, Digital audio workstation

PACS number: 43.38.MD, 43.60.Pt [doi:10.1250/ast.39.295]

1. INTRODUCTION

There are several techniques for physical sound field reproduction which reproduce physical quantities such as wavefront, phase relation and the distribution of sound pressures. Typical examples are a first order (B-format) Ambisonics [1–4] and higher order Ambisonics (HOA) [5–11], a wave field synthesis (WFS) [12–16], and a boundary surface control (BoSC) [17–21].

Although there is an attempt to perform real-time playback with B-format Ambisonics [22], the sound field reproduction system with multiple-channel microphone and loudspeakers usually need relatively large scale calculations to obtain reproducing signals.

Also, there is an inevitable bounds for accurate reproduction due to an usage of discrete and limited number of sensors [23]. Typical limitation, called spatial aliasing [24,25], limit the higher limiting frequency, for example. On the other hand, natural characteristics of devices, such as narrow directivity at higher frequencies, might have potential for complemental work.

This paper proposes two techniques useful for realization of simple and accurate reproduction system. (The term “accurate” means that the degree of physical reproduction of original field is high.) The boundary surface control principle [17] is adopted as a basic technique for physical reproduction. Also, the microphone array which consists of narrow directivity is introduced for this purpose.

The first technique is a reduction of number of inverse filters necessary for the implementation of BoSC principle. The second is a hybrid usage of BoSC and a simple reproduction, which emit the recorded sound from roughly corresponding directions. These are the noble usage of narrow directivity especially at high frequency. As described below in detail, the proposed system consists of quite rough approximation of physical reproduction method; BoSC, and the simple psychological reproduction method based on traditional amplitude panning by introducing narrow directional microphones.

*e-mail: hrs@penr.in
†e-mail: omoto@design.kyushu-u.ac.jp
The outline of the system and the reproduction method, BoSC and simple reproduction, used in this study is introduced in the next section. The adopted method for the calculation of inverse filter matrix, and the reduction procedure is shown in the subsequent section. The practical procedure of the hybrid reproduction is also shown. Throughout the examination, the reproduction accuracy is evaluated by similarity index of waveforms measured in the primary field and reproduced field by originally designed microphone array.

2. THE PROPOSED SYSTEM

2.1. Sound Acquiring and Reproducing System

The used system consists of 24-channel narrow directivity (shotgun) microphone array and 24-channel circularly arranged loudspeaker array. These are shown in Figs. 1(a) and 1(b), respectively. The microphone array, named “Hedgehog” due to its appearance has evenly distributed 24 microphones at every 45° in elevation and azimuth angles; i.e., eight microphones at every 45° times three layers.

Fig. 1 System overview; (a) the hedgehog microphone array which consists of 24 narrow directional microphones, and (b) loudspeaker array which consists of 24 loudspeakers (only 21 speakers are shown in the photo). The reproduction system is located in hemi-anechoic chamber.

Although it might be difficult to determine the accurate acoustic center of the shotgun microphone, the distance between adjacent microphones are roughly 140 mm. The frequency characteristics and the directivity are shown in Fig. 2. Since the constructed microphone holder for the array was rather large structure, the directivity characteristics were measured in an anechoic chamber for both cases; measurement of single (one) microphone in which the middle point of the microphone was assumed as center of rotation, and the practical condition in which the center of the whole array was assumed as center of rotation. The sound source was located at around 4 m from the microphone.

Results are shown as polar plots in Fig. 2 in which the left and right parts were correspond to the results of single microphone and with the microphone holder, respectively. No serious differences were observed except the conditions of incidence from the back of the microphone.

Fig. 2 Frequency characteristics and polar plots of directivity of used microphone (AKG C568B). In the polar plot, the results of octave bands from 63 Hz to 8 kHz are shown, in the cases of measurement of one microphone only with simple microphone holder (left) and attached on the holder used in this research (right).
In the design process, the case of 36 microphones at every 30° was firstly planned. However, after measurement of directional characteristics of each microphone (Fig. 2), superpositions of sensitiveness between microphone was concerned. Also, 18 microphone at every 60° could be the candidate. In this case, thinking about the correspondent arrangement of loudspeakers, the left and right sound source would be located at 120° aperture and this might be too wide to create clear phantom images. Considering such conditions, 45° was chosen mostly intuitively.

The reproducing loudspeaker array also has three layers and each layer has eight speakers at 45° step. The assumed radius of the arrangement is 2 m. The elevation angle from the listening position is ±20°. The powered monitor speaker, Genelec 8030A is used for the front three positions (left, center, and right) in middle layer and the remaining 21 speakers are the Genelec 8020C. To correspond conventional stereo reproduction, slightly bigger loudspeakers are introduced at the front position. The whole characteristics of the two different models are quite similar and the substantial differences of the speakers are only the lower limiting frequency; 55 Hz for 8030A and 65 Hz for 8020C, respectively.

2.2. Reproduction Strategy

Two reproduction methods are introduced and their fusion is attempted to achieve accurate reproduction at wider frequency range.

2.2.1. Direct method

The first method is called “Direct,” in which the recorded sound by narrow directional microphone is reproduced from one loudspeaker located almost corresponding direction, i.e., the same azimuth angle and same layer (upper, middle or bottom). Basically, no signal processing is used in this method. The amplitude weighting due to narrow directivity is expected to work as natural amplitude panning. This method can be interpreted as an extension of 6-channel system proposed by Yokoyama [26]. Also, it can be categorized as simplified version of a series of research [27–32], in which the directions of reflection in enclosed space were measured by several methods such as spatial decomposition method, and reproduced by plural loudspeakers mainly using vector based amplitude panning.

As shown in Fig. 2, the directivity is narrower at higher frequency. Therefore, this Direct method is more effective at higher frequency.

2.2.2. Boundary surface control method

The second method is a boundary surface control principle (BoSC, hereafter) proposed by Ise [17]. The conceptual diagram is shown in Fig. 3. This method is based on the Helmholtz-Kirchhoff integral equation and measures distributions of sound pressure and their gradient on arbitrary enclosed surface $S$ in the primary field. Then, the accurate reproduction is achieved by reproducing the distributions of pressure and gradients by the combinations of secondary sources and the inverse filter matrix $H(\omega)$. The acoustic characteristics of secondary field $G(\omega)$ is measured in advance and the inverse filter matrix is calculated to cancel them. Practically, the sound pressures at the surface $S$ in primary field is measured by microphone array which constructs the enclosed space $V$. As already discussed and suggested [33,34], there is no need to measure the gradient of pressures due to the uniqueness of the solution in most cases.

The impulse responses between secondary loudspeakers and the microphone array used for recording are measured in the secondary sound field, in advance. Usually, the inverse filter matrix is calculated in the frequency domain. The recorded signals $X(\omega)$ after passing through this inverse filter matrix $H(\omega)$ are emitted from the secondary loudspeakers. The inner field of enclosed sound source is reproduced at the inner surface $S$ in secondary sound field.
surface, $V$, is then accurately reproduced as $V'$, in the secondary field.

Kimura [35,36] also suggested analytically the advantages of using uni-directional or narrow-directional microphones; these characteristics are effective to simulate the approximated version of principal Helmholtz-Kirchhoff integral equation. As a results, more accurate reproduction were achieved.

2.2.3. Calculation of inverse filter matrix

The procedure for obtaining the inverse filter matrix is introduced here in frequency domain. The $L \times 1$ vector $Y(\omega)$, a reproduced signal in the secondary field, can be expressed as

$$Y(\omega) = G(\omega)H(\omega)X(\omega),$$  

(1)

where $L$ is the number of microphones. The number of loudspeakers is assumed to be $M$ and the $L \times M$ matrix $G(\omega)$ is a transfer matrix, in which the $(\ell, m)$-th elements corresponds to the transfer function between $m$-th loudspeaker and $\ell$-th microphone. The $L \times 1$ vector $X(\omega)$ is a recorded signal at primary field.

The ideal condition in which $Y(\omega) = X(\omega)$, i.e., the perfect reproduction, is realized in the special case of $G(\omega)H(\omega) = I$. However, This is not necessarily satisfied due to the numbers of sensors and loudspeakers, conditions of transfer matrix and so on. Therefore, by introducing Tikhonov regularization method [37–39], the cost function of

$$J(\omega) = \|I - G(\omega)H(\omega)\|^2 + \varepsilon \|H(\omega)\|^2$$  

(2)

is usually assumed and minimization is attempted to obtain an approximated solutions of $H(\omega)$. In the equation above, the $\varepsilon$ is the regularization parameter to cope with the ill-condition of transfer matrix. The optimal solution to minimize the above cost function can be obtained as

$$H(\omega) = [G^H(\omega)G(\omega) + \varepsilon I]^{-1}G^H(\omega).$$  

(3)

There are several method for determining the regularization parameter $\varepsilon$, such as least square, minimization of norm, singular value decomposition, so-called L-curve method, and using genetic algorithm. Details of these standard methods can be found, for example, in [40]. Several trials showed the tendency of larger value at lower frequency and smaller value at around 1 kHz. In addition to such rough estimation, trial and error, and also artificial manipulation such as round off, the heuristically determined regularization parameter $\varepsilon$ shown in Table 1 were used in the following calculations.

2.2.4. Reduction of inverse filters

To calculate the reproducing signals, convolution of the inverse filter is necessary. The number of inverse filter is basically the product of number of loudspeakers and number of microphones. In the proposed system, $24 \times 576 = 576$ filters. In the design process of inverse filter by using Discrete Fourier Transform (DFT), the length of filter is usually chosen to be eight to sixteen times the nominal impulse response, to avoid folding distortion in the inverse DFT. In our case, the impulse response in the secondary field converged at around 10 ms (480 taps). Therefore, the length of 4,096 (more than eight times of 480) was chosen. Since no further truncation was carried out for obtained filters in this case, the convolution of 576 filters with 4,096 taps are necessary. It might be somewhat impractical for on-line or real-time calculation.

The reduction of number of filters is therefore attempted. Omitting frequency index $\omega$, the elements of inverse filter relating to $\ell$-th microphone can be expressed as

$$i_\ell = Gh_\ell$$  

(4)

where $i_\ell$ is a $L \times L$ unit-matrix, i.e., the $L \times 1$ vector of only $\ell$-th element is unity and the others are zero, and $h_\ell$ is $\ell$-th column of inverse filter matrix $H$. In this equation, we assume that only several elements affect dominantly the process of obtaining inverse filter due to the narrow directivity of microphones.

The image of this reduction is shown in Fig. 4. As shown in the figure, except the dominant elements (three elements in the figure) in $h_\ell$, the remains are set to zero. Consequently, the structure is reduced to be

| Freq. [Hz] | 20–125 | 250 | 500 | 1k | 2k | 4k | 8k | –20k |
|-----------|--------|-----|-----|----|----|----|----|------|
| $\varepsilon$ | 0.25 | 0.07 | 0.05 | 0.068 | 0.035 | 0.0125 | 0.025 | 0.05 |

Table 1 The regularization parameter $\varepsilon$ used in the calculation of inverse filter matrix. The frequency means the center frequency of octave bands except the lower (20Hz) and upper (20kHz) limiting frequency.
where the subscript ‘s’ means ‘support’ and it indicate the set of indexes of non-zero elements ($\{i | x_i \neq 0\}$).

In the case shown in the figure, the scale of the equation to be solved is reduced to $\frac{L}{C_2} \frac{1}{C_1} = \frac{L}{C_2} \frac{3}{C_1}$.

The vector of inverse filter applied for the signal recorded at $\ell$-th microphone is then obtained as

$$h_{\ell,s} = [G_H^s G_s + \varepsilon I]^{-1} G_H^s i_{\ell,s}.$$  \hspace{1cm} (6)

This procedure can be interpreted as an approximated method of sparse modeling [41] by using narrow directivity of microphones.

Practically, the maximum number of dominant loudspeakers was assumed to five and the combination of one microphone and one loudspeaker at corresponding direction plus surrounding speakers with shortest distances are assumed to be dominant and used in the calculation. In the upper and lower layers, four loudspeakers are assumed to have an dominant effect, and in the middle layer, five loudspeakers are used in the calculation as shown in Fig. 5. Therefore, $4 \times 8 \times 2 + 5 \times 8 = 104$ filters are necessary in this case. The selection of maximum number being five was arbitrarily prescribed. The other selection possibilities using quantitative evaluation is one of the important topics of our future research.

For such reduced case, it might be possible to realize the on-line calculation in digital audio workstation (DAW) software and convolution reverberator Plug-in. This reduced condition is called “inv104” and the full size inverse filter is called “inv576,” hereafter. The regularization parameters shown in Table 1 were also used in calculation of inv104. Possibilities of choosing another regularization parameter were not examined in this case for simplicity.

### 2.2.5 Hybrid reproduction of Direct and BoSC

As described earlier, the Direct method, which uses directivity of microphones, works well at higher frequency. On the contrary, the BoSC works at lower frequency due to the limited number of microphones. In our system, the adjacent microphone distance is roughly 140 mm, and therefore, the higher limiting frequency for spatial aliasing is around 1.2 kHz. The hybrid reproduction method that both works complementarily has therefore been attempted. The block diagram is shown in Fig. 6.

In the practical calculation shown below, the crossover frequency was set to 3.4 kHz; higher than the limiting frequency of aliasing since the benefit of narrow directional characteristics is prominently at more than, say, 4 kHz. Detailed explanation of this crossover frequency is shown below.

2.2.6 Summary of proposed system

The proposed system in this paper can then be placed as follows;

- simplified version of reproduction system of BoSC compared with conventional ones [19–21]
- practical implementation of Kimura’s idea by using narrow-directional microphone [35,36].

Also, by using the characteristics of narrow directivity of microphones, the following points are extended;

- reduction of number of inverse filters, necessary for BoSC, are attempted,
- hybrid constitution: BoSC in low frequency range, and natural narrow directivity in high frequency range, is attempted.

Practical performance of the system is introduced in the following section.

### 3. PERFORMANCE EVALUATION OF REPRODUCTION SYSTEM

As preliminary examination of the performance of the system, subjective evaluation [42] and physical measurement using sound intensities [43,44] had been carried out. From both results, basic performance of the system was assured. Especially in the subjective evaluation [42], the advantages of hybrid structure was confirmed.

In this paper, for further examination, performance of reproduction is evaluated by the similarity of wavefront in primary and secondary sound fields. For this purpose, 96 channel microphone array arranged into a square grid form was designed and implemented.
3.1. Grid Microphone Array

The array of small sized 96 MEMS (Micro Electro Mechanical System) microphone (Knowles SPU0414HR5H) arranged into 15 mm square grid, was made and named as Grid Microphone Array, GMA. The appearance of GMA is shown in Fig. 7. By introducing appropriate A/D converter (three sets of DirectOut Technologies Andiamo 2, in our case), the sound pressures at $0.12 \text{ m} \times 0.18 \text{ m}$ can be recorded simultaneously.

3.2. Evaluation Procedure

The wavefront both in the primary field and in the reproduced field was measured by the GMA and compared. The block diagram of the measurement is shown in Fig. 8. The primary field is in an anechoic chamber, and the sound source was located at 2 m from the receiving point, the center of Hedgehog microphone array. The square area of $720 \text{ mm} \times 720 \text{ mm}$, at the same hight of the middle layer of Hedgehog microphone array, was assumed as an evaluation area. The propagation of wavefront was observed by the measurement of impulse responses inside the area by using GMA. The GMA covered $120 \text{ mm} \times 180 \text{ mm}$ area at one measurement. The 24 (= $6 \times 4$) times measurements were therefore carried out successively to cover the whole area.

At the same time, the Hedgehog microphone was located and the swept-sine signal was recorded. These recorded signals were then reproduced at the secondary field after processing of above mentioned procedures, Direct, BoSC of full inverse filter inv$576$ and reduced number of inverse filters inv$104$, and their Hybrid.

In the expression of the wavefront, the interpolation method proposed by Yatabe [45,46] was introduced. In this method, the pressures and their normal derivative at arbitrary boundary were estimated by solving the inverse problem using measured sound pressures. These values are then used to estimate the sound pressure at arbitrary inner points. The concept of this interpolation is shown in Fig. 9.

3.3. Wavefront Visualization

Figure 10 shows the results of visualization of measured wavefront. The top row (a) is a results in primary field, and successive rows are for reproduced results of Direct (b), BoSC with full size inverse filter inv$576$ (c), reduced filter inv$104$ (d), and Hybrid (e). In this case, the Hybrid consists of inv$104$ and Direct with crossover frequency of 3.4 kHz. The columns correspond to the frequency of 500 Hz, 1 kHz, 2 kHz, 4 kHz and 8 kHz.

The circular wavefront can be observed in the primary field (a) very clearly. In Direct (b), rough duplication can be observed at narrow region at around the center of evaluation area in 500 Hz and 1 kHz. At higher frequency such as 4 kHz and 8 kHz, there are small regular patterns in the wavefront shape, however, loose circular wavefront shape centered at the position of sound source was reproduced.

As for the BoSC, roughly duplicated shape of wavefront could be observed up to 1 kHz. At higher frequency such as 4 kHz and 8 kHz, there are small regular patterns in the wavefront shape, however, loose circular wavefront shape centered at the position of sound source was reproduced.
2 kHz and 4 kHz. It is also interesting that the shapes of wavefront for inv104 are similar to the Direct case in 4 kHz.

The wavefront of Hybrid are the mixture of inv104 for low frequency and Direct at high frequency.

3.4. Quantitative Evaluation

Visualization of wavefront would be useful for intuitive comparison or grasping the tendencies of each reproduction method. The quantitative evaluation method by using cosine similarity [47–49] is then introduced for further examination. This similarity $\cos \theta_k$ is defined as

$$\cos \theta_k = \frac{|\sum_{x,y} p_k(x,y) q_k(x,y)|}{\sqrt{\sum_{x,y} |p_k(x,y)|^2} \sqrt{\sum_{x,y} |q_k(x,y)|^2}},$$

where $p_k(x,y)$ and $q_k(x,y)$ are the amplitude at $k$-th frequency bin in DFT of impulse responses measured at coordinate $(x,y)$ in primary field and reproduced field, respectively. The superscript ‘$*$’ means complex conjugate. All the measured points $(2,304 = 96 \times 24)$ are used as coordinate $(x,y)$ and the interpolated data are not included in the calculation.

This $\cos \theta_k$ is derived by inner products of complex amplitude $p_k(x,y)$ and $q_k(x,y)$, and normalized by sum of their absolute values. The absolute values in numerator means that the phase information is ignored in this evaluation.

Figure 11 shows the results. In (a), the comparison of inverse filters inv576 and inv104 are shown. Apparently, the full size inverse filter inv576 showed higher value until around 1 kHz; near to the upper limiting frequency for spatial aliasing. In the case of inv104, the value gradually reduced until around 2 kHz, and showed peak at 3 kHz. Both values are then went down to around 0.4 to 0.5 from 4 kHz.

In Fig. 11(b), the cases of Direct and inv104 and Hybrid are shown. The Direct method showed high value at low frequency, but the value decreased from around 500 Hz. After the lowest value at around 2 kHz, the value gradually went up and shaped some peaks at 0.6 to 0.8. These values were clearly higher than the cases of BoSC and therefore, the crossover frequency was set to 3.4 kHz in this research. Consequently, the cosine similarity of Hybrid method reasonably high value in all frequency range.

Although further subjective evaluation is necessary to conclude, the quality of reproducing sound in Hybrid are quite good by introspection report of several subjects, especially in relatively ‘sharp’ sound source location. Even though there are slight differences in the detailed conditions, the preliminary examination of performance of Hybrid method by using inter-aural time and level differences are shown in the reference [42]. In this reference, the best results were obtained for Hybrid method for sound localization test.

3.5. Implementation in DAW Software

In the case of reduced size of inverse filter, inv104 and Hybrid, 104 FIR filters are necessary to reproduce. This scale is acceptable for recent DAW software and Plug-in of convolution reverberator.

Recorded signal at one microphone was send to the four or five ‘Bus’ inside the software; The microphones in upper and lower layer, four Bus are necessary and five Bus for middle layer as shown in Fig. 5. After passing through the corresponding inverse filter by ‘inserted Plug-in,’ the signals are again mixed and emitted from each loudspeaker. Typical screenshot is shown in Fig. 12 with DAW (Nuendo in this case) and plug-in (REVerence) in our playback environment. This playback is usually performed with notebook computer, having Intel Core i7, 2.7 GHz processor and 16 GB memory, for example, without serious load on CPU.

4. CONCLUDING REMARKS

Sound field reproduction system based on boundary
surface control principle and using narrow directional microphones is proposed. The 24-channel Hedgehog microphone array was used as sound acquiring device, and the 24-channel loudspeakers, located in circular and three layers, are used as emitting device.

The boundary surface control principle was used as basic reproduction method. In such ‘wave based’ physical reproduction method, the distance between finite numbers of microphones usually limit the upper limiting frequency of reproduction. The narrow directivity of the microphone was then used to compensate the inevitable degradation of reproduction accuracy in high frequency.

Also, this characteristic was used to limit the number of dominant loudspeakers which affect the reproduction signal at certain direction. This benefit was utilized as reducing the number of inverse filters. In this case, the reduction from full size of 576 filters to 104 was attempted.

The performance evaluations of reproduction was carried out qualitatively and quantitatively, by the visualization of propagating wavefront and calculating the cosine similarity, respectively. The results indicated the superiority of BoSC in low frequency. Also, the Direct reproduction, i.e., the narrow directivity, showed some superiority in high frequency. Consequently, the Hybrid

Fig. 10 Measured wavefront in primary field (a), reproduced field by Direct (b), BoSC with full size inverse filter inv576 (c), the reduced filter inv104 (d), and Hybrid (e).
method of them showed reasonable performance in whole frequency range.

ACKNOWLEDGEMENT

This research is supported by JST/CREST, JSPS KAKENHI JP 25282003, JP 26560012, and JP 17H00811.

REFERENCES

[1] D. H. Cooper and T. Shiga, “Discrete-matrix multichannel stereo,” *J. Audio Eng. Soc.*, 20, 346–360 (1972).
[2] M. A. Gerzon, “Periphony: With-height sound reproduction,” *J. Audio Eng. Soc.*, 21, 2–10 (1973).
[3] M. A. Gerzon, “Surround-sound psychoacoustics,” *Wireless World*, 80, 483–486 (1974).
[4] M. A. Gerzon, “Design of ambisonic decoders for multi-speaker surround sound,” *Proc. 58th Audio Eng. Soc. Convention*, 24 pages (1977).
[5] J. Daniel, “Spatial sound encoding including near field effect: Introducing distance coding filters and a viable, new ambisonic format,” *Proc. 23rd Audio Eng. Soc. Conf.*, Paper Number 13, 15 pages (2003).
[6] M. A. Poletti, “Three-dimensional surround sound systems based on spherical harmonics,” *J. Audio Eng. Soc.*, 53, 1004–1025 (2005).
[7] T. Okamoto, D. Cabrera, M. Noisternig, B. Katz, Y. Iwaya and Y. Suzuki, “Improving sound field reproduction in a small room based on higher-order ambisonics with a 157-loudspeaker array,” *Proc. 2nd Int. Symp. Ambisonics and Spherical Acoustics*, 4 pages (2010).
[8] M. A. Poletti, T. Beiletem and T. Abbayapala, “Higher order loudspeakers for improved surround sound reproduction in rooms,” *Proc. 133rd Audio Eng. Soc. Convention*, 13 pages (2012).
[9] J. Trevino, T. Okamoto, Y. Iwaya and Y. Suzuki, “High order ambisonic decoding method for irregular loudspeaker arrays,” *Proc. ICA 2010*, 8 pages (2010).
[10] J. Trevino, T. Okamoto and Y. Iwaya, “Evaluation of a new ambisonic decoder for irregular loudspeaker arrays using interaural cues,” *Proc. Ambisonics Symp.*, 8 pages (2011).
[11] J. Trevino, S. Koyama, S. Sakamoto and Y. Suzuki, “Mixed-order ambisonsics encoding of cylindrical microphone array signals,” *Acoust. Sci. & Tech.*, 35, 174–177 (2014).
[12] A. J. Berkhout, D. de Vries and P. Vogel, “Acoustic control by wave field synthesis,” *J. Acoust. Soc. Am.*, 93, 2764–2778 (1993).
[13] G. Theile and H. Wittek, “Wave field synthesis: A promising spatial audio rendering concept,” *Acoust. Sci. & Tech.*, 25, 393–399 (2004).
[14] P. Gauthier and A. Berry, “Adaptive wave field synthesis with independent radiation mode control for active sound field reproduction: Theory,” *J. Acoust. Soc. Am.*, 119, 2721–2737 (2006).
[15] P. Gauthier and A. Berry, “Adaptive wave field synthesis for sound field reproduction: Theory, experiments, and future perspectives,” *J. Audio Eng. Soc.*, 55, 1107–1124 (2007).
[16] P. Gauthier and A. Berry, “Adaptive wave field synthesis for active sound field reproduction: Experimental results,” *J. Acoust. Soc. Am.*, 123, 1991–2002 (2008).
[17] S. Ise, “A principle of sound field control based on the Kirchhoff-Helmholtz integral equation and the theory of inverse systems,” *Acta Acust. united Ac.*, 85, 78–87 (1999).
[18] S. Takane, Y. Suzuki and T. Sone, “A new method for global sound field reproduction based on Kirchhoff’s integral equation,” *Acta Acust. united Ac.*, 85, 250–257 (1999).
[19] S. Enomoto, Y. Ikeda, S. Ise and S. Nakamura, “Three-dimensional sound field reproduction and recording system based on the boundary surface control principle,” *Proc. 14th Int. Conf. Auditory Display*, 8 pages (2008).
[20] M. Kobayashi, K. Ueno, M. Yamashita, S. Ise and S. Enomoto, “Subjective evaluation of a virtual acoustic system: Trials with three-dimensional sound field reproduced by the Sound Cask,” *Proc. ICA 2013*, pp. 015077–015077 (2013).
[21] A. Omoto, S. Ise, Y. Ikeda, K. Ueno, S. Enomoto and M. Kobayashi, “Sound field reproduction and sharing system based on the boundary surface control principle,” *Acoust. Sci. & Tech.*, 36, 1–11 (2015).
[22] B. Watson and D. F. Clark, “Real-time ambisonic simulation
of auditorium acoustics within a digital audio workstation environment,” *Proc. 2nd Int. Symp. Ambisonics and Spherical Acoust.*, 2010, 4 pages (2010).

[23] Y. Koyano, K. Yatabe and Y. Oikawa, “Infinite-dimensional SVD and boundary surface control,” *Acoust. Sci. & Tech.*, **129**, 116–125 (2018).

[24] J. Dmochowski, J. Benesty and S. Affes, “On spatial aliasing in microphone arrays,” *IEEE Trans. Signal Process.*, **57**, 1383–1395 (2009).

[25] M. Otani, H. Watabe, T. Tsuchiya and Y. Iwaya, “Effects of spatial aliasing in sound field reproduction: Reproducibility of binaural signals,” *Acoust. Sci. & Tech.*, **38**, 147–153 (2017).

[26] S. Yokoyama, K. Ueno, S. Sakamoto and H. Tachibana, “6-channel recording/reproduction system for 3-dimensional auralization of sound fields,” *Acoust. Sci. & Tech.*, **23**, 97–103 (2002).

[27] J. Merimaa and V. Pulkki, “Spatial impulse response rendering I: Analysis and synthesis,” *J. Audio Eng. Soc.*, **53**, 1115–1127 (2005).

[28] V. Pulkki and J. Merimaa, “Spatial impulse response rendering II: Reproduction of diffuse sound and listening tests,” *J. Audio Eng. Soc.*, **54**, 3–20 (2006).

[29] V. Pulkki, “Spatial sound reproduction with directional audio coding,” *J. Audio Eng. Soc.*, **55**, 503–516 (2007).

[30] J. Vilkamo, J. Vilkamo, T. Lokki and V. Pulkki, “Directional audio coding: Virtual microphone-based synthesis and subjective evaluation,” *J. Audio Eng. Soc.*, **57**, 709–724 (2009).

[31] S. Tervo, J. Patynen and T. Lokki, “Acoustic reflection localization from room impulse responses,” *Acta Acust. united Ac.*, **98**, 418–440 (2012).

[32] S. Tervo, J. Patynen, A. Kuusinen and T. Lokki, “Spatial decomposition method for room impulse responses,” *J. Audio Eng. Soc.*, **61**, 17–28 (2013).

[33] R. Kleinman and G. Roach, “Boundary integral equations for the three dimensional Helmholtz equation,” *SIAM Review*, **16**, 214–236 (1974).

[34] K. Furuya and Y. Ichinose, “Sound field control in a closed space using boundary-pressure method,” *Tech. Rep. IEICE*, EA90-15, pp. 25–32 (1990).

[35] T. Kimura and K. Kakehi, “Effects of microphone and loudspeaker directivity on accuracy of synthesized wave fronts in sound field reproduction with wave field synthesis,” *Trans. Virtual Reality Soc. Jpn.*, **12**, 191–198 (2007).

[36] T. Kimura, “Theoretical study and numerical analysis of 3D sound field reproduction system using directional microphones and boundary surface control,” *Trans. Virtual Reality Soc. Jpn.*, **5**, 231–241 (2010).

[37] A. N. Tikhonov, “Solution of incorrectly formulated problems and the regularization method,” *Soviet Math. Dokl.*, **4**, 1035–1038 (1963).

[38] D. Calvetti, S. Morigi, L. Reichel and F. Sgallari, “Tikhonov regularization and the L-curve for large discrete ill-posed problems,” *J. Comput. Appl. Math.*, **123**, 423–446 (2000).

[39] S. J. Elliott, *Signal Processing for Active Control* (Academic Press, London, 2001).

[40] W. Menke, translated by T. Yanagidani and K. Tsukada, *Geophysical Data Analysis: Discrete Inverse Theory* (Kokon-Shoin, Tokyo, 1997).

[41] M. Elad, *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing* (Springer, New York, 2010).

[42] Y. Saito and A. Omoto, “Performance improvement of 3D sound field reproduction system,” *Trans. Tech. Comm. Arch. Acoust. Acoust. Soc. Jpn.*, AA2016-09 (2016).

[43] M. Nakahara, A. Omoto and Y. Nagatomo, “A simple evaluating method of a reproduced sound field by a measurement of sound intensities using Virtual Source Visualizer,” *AES 143th Convention e-Brief*, #179 (2017).

[44] T. Mikami, M. Nakahara and A. Omoto, “Physical evaluations of reproduced sound fields by measurements of sound intensities using Virtual Source Visualizer,” *AES 143th Convention e-Brief*, #181 (2017).

[45] K. Yatabe and Y. Oikawa, “PDE-based interpolation method for optically visualized sound field,” *Proc. Int. Conf. Speech Signal Process. (ICASSP)*, pp. 4771–4775 (2014).

[46] K. Yatabe and Y. Oikawa, “Optically visualized sound field reconstruction using Kirchhoff-Helmholtz equation,” *Acoust. Sci. & Tech.*, **36**, 351–354 (2015).

[47] A. Singhal, “Modern information retrieval: A brief overview,” *IEEE Data Eng. Bull.*, **24**, 35–43 (2001).

[48] N. Dehak, R. Dehak, P. Kenny and N. Brummer, “Support vector machines versus fast scoring in the low-dimensional total variability space for speaker verification,” *Proc. InterSpeech 2009*, pp. 1559–1562 (2009).

[49] N. Obin and A. Roebel, “Similarity search of acted voices for automatic voice casting,” *IEEE/ACM Trans. Audio*, **24**, 1642–1651 (2016).

**Hiroshi Kashiwazaki** was born in 1994 and graduated from the Department of Acoustic Design, Kyushu University in 2017. He is currently a student of doctoral course, in Communication Design Science Course, Graduate School of Design, Kyushu University.

**Akira Omoto** graduated from the Department of Acoustic Design, Kyushu Institute of Design, in 1987 and received his Ph.D. degree from the University of Tokyo in 1995. From 1987 to 1991, he worked as a research and development engineer at Nittobo Acoustic Engineering Co., Ltd., Tokyo. In 1991, he was appointed as a Research Assistant at Kyushu Institute of Design and was made an Associate Professor in 1997. After integration of universities in 2003, he was made an Associate Professor of Kyushu University and was made full professor in 2014. His research interest includes measurement, evaluation, and control of sound field.