A study on 3-D Sound Field Localization System Using Parametric Loudspeaker and Indirect Loudspeakers for Reverberation Reproduction

Tomoyuki WADA[1]; Daisuke IKEFUJI[1]; Masato NAKAYAMA[2]; Takanobu NISHIURA[2]

[1] Graduate School of Information Science and Engineering, Ritsumeikan University, Japan

[2] College of Information Science and Engineering, Ritsumeikan University, Japan

ABSTRACT

3-D sound field reproduction systems can provide listeners with a sense of presence. In this study, we have proposed the 3-D sound field localization system using parametric loudspeakers. This system can easily represent sound images on the wall by using parametric loudspeakers that have a high directivity. In contrast, it is difficult to reproduce reverberant environments such as a concert hall because the parametric loudspeaker can hardly represent a reverberation. In the former system, therefore, we add the reverberation and control the reverberation time by using indirect electrodynamic loudspeakers. However, it can imperfectly reproduce reverberant environments because the reverberant characteristic is not only the reverberation time. In this paper, we newly propose the 3-D sound field localization system using the parametric loudspeaker and indirect electrodynamic loudspeakers for the reverberation reproduction. Specifically, in the reverberation reproduction, we design inverse filters of transfer functions between listening points and indirect electrodynamic loudspeakers. Then, we reproduce the desired reverberant environment by emitting signals convoluted with inverse filters and the impulse response in the desired sound field from indirect electrodynamic loudspeakers. As a result of the objective experiment on the performance of the sound image localization and the reproducibility of the reverberation, we confirmed the effectiveness of the proposed system.

Keywords: 3-D sound field reproduction, Parametric loudspeaker, Indirect electrodynamic loudspeaker, Reverberation

I-INCE Classification of Subjects Number(s): 01.4

1. INTRODUCTION

3-D sound field reproduction systems can provide listeners with a sense of presence. A binaural and a transaural systems (1, 2) have proposed as 3-D sound field reproduction systems. Although the binaural system can reproduce the sound field by using the head-related transfer function (HRTF) of listeners, they feel troublesomeness because this system forces listeners to use a headphone. On the other hand, the transaural system can reproduce the sound field by utilizing inverse filters with multiple remote-loudspeakers. However, it requires a large-scale system with a lot of loudspeakers and high computational costs.

In this study, we have proposed the 3-D sound field localization system using parametric loudspeakers (3). This system represents a high realistic 3-D sound field without the headphone, and it utilizes a small-scale system. It can easily represent sound images on the wall by using parametric loudspeakers that have a high directivity. Thus, listeners can obtain a high realistic sensation by perceiving sound images on the wall. In contrast, it is difficult to reproduce reverberant environments such as a concert hall because the parametric loudspeaker can hardly represent a reverberation. In the former system, therefore, we add the reverberation by using indirect electrodynamic loudspeakers. This system reproduces desired reverberant environments by controlling the reverberation time. However, it can imperfectly reproduce desired reverberant environments because the reverberant characteristic is not only the reverberation time.

In this paper, therefore, we newly propose the 3-D sound field localization system using the parametric loudspeaker and indirect electrodynamic loudspeakers for the reverberation reproduction. To achieve this system, we reproduce an impulse response in the desired sound field (the desired impulse response) at the

[1] (is0095xr,cm000074)@ed.ritsumei.ac.jp

[2] (mnaka@fc,nishiura@is).ritsumei.ac.jp

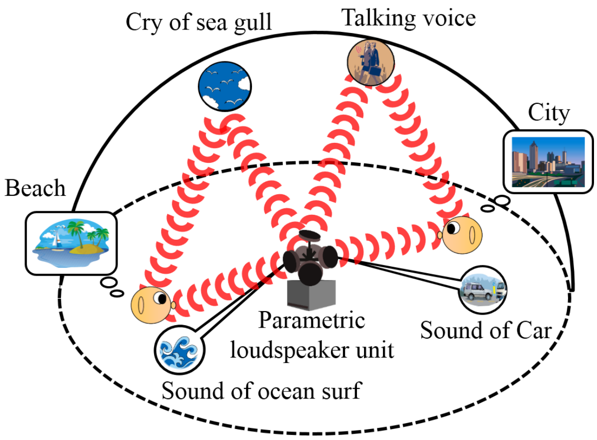

Figure 1 – The overview of the 3-D sound field localization system using parametric loudspeakers

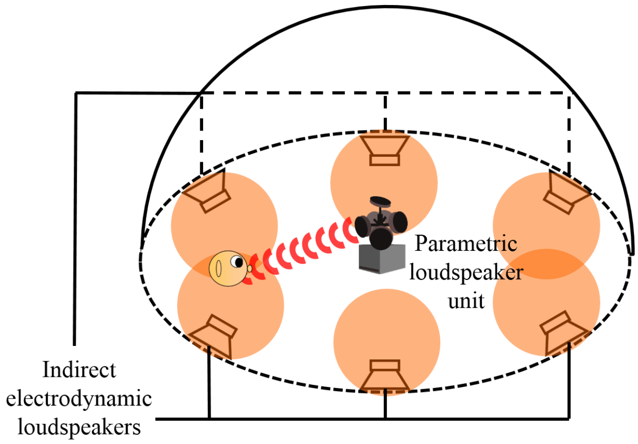

Figure 2 – The overview of the 3-D sound field localization system using the parametric loudspeaker and indirect electrodynamic loudspeakers

listening point. Specifically, the parametric loudspeaker represents the direct sound, and indirect electrodynamic loudspeakers represent the reverberation. In the reverberation reproduction, we design inverse filters of transfer functions between listening points and indirect electrodynamic loudspeakers. Then, indirect electrodynamic loudspeakers emit signals convoluted with inverse filters and the desired impulse response. Finally, we reproduce the desired impulse response at the listening point.

2. FORMER SYSTEM

The 3-D sound field localization system using parametric loudspeakers can easily represent sound images to listeners because the parametric loudspeaker has a high directivity. Figure 1 shows the overview of the 3-D sound field localization system using parametric loudspeakers. This system represents sound images to listeners by that emitted sounds from parametric loudspeakers reflect on walls, floors and ceilings. However, the reflective sound from the parametric loudspeaker hardly diffuses. Thus, the parametric loudspeaker has a difficulty to represent a reverberation. To solve this problem, we previously proposed a method that adds the reverberation by using indirect electrodynamic loudspeakers. The indirect electrodynamic loudspeaker is an electrodynamic loudspeaker for representing an indirect sound such as an indirect illumination that represents indirect light. Figure 2 shows the overview of the achieved system by using this method. In Fig. 2, this system represents the sound image from the parametric loudspeaker and the reverberation from indirect electrodynamic loudspeakers. In the reverberation representation, the former system controls the reverberation time by using an impulse response based on a white noise and an exponential decay curve. Equation (1) defines the observed impulse response $IR_{former}(k)$ at the listening point.

$$IR_{former}(k) = g_0(k) + \sum_{m=1}^{M} n(k) \cdot \mathrm{e}^{-\alpha} * g_m(k), \tag{1}$$

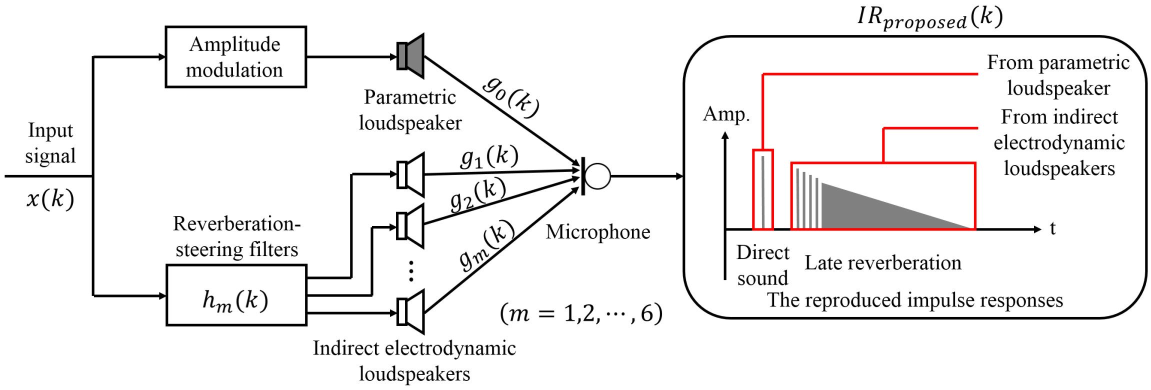

Figure 3 – The principle of the proposed system

where k is time index, M is the number of indirect electrodynamic loudspeakers, $g_0(k)$ is the impulse response between the parametric loudspeaker and the listening point, $g_m(k)$ is the impulse response between the m-th indirect electrodynamic loudspeaker and the listening point, $n(k)$ is the white noise, α is a constant of an exponential slope corresponding to the desired reverberation time, and the symbol * stands for the convolution. The former system controls the reverberation time by determining the value of α. However, it can imperfectly reproduce reverberant environments because the reverberant characteristic is not only the reverberation time.

3. PROPOSED SYSTEM

As mentioned previously, the former system has a difficulty to reproduce reverberant environments with a high accuracy. Therefore, we newly propose the 3-D sound field localization system using the parametric loudspeaker and indirect electrodynamic loudspeakers for the reverberation reproduction. Arrangements of the proposed system is same as that of the former system as shown in Fig. 2, and the proposed system is different from the former system in terms of the representation method of the reverberation. To achieve this system, we reproduce an impulse response in the desired sound field (the desired impulse response) at the listening point. Specifically, the parametric loudspeaker represents the direct sound, and indirect electrodynamic loudspeakers represent the reverberation as shown in Fig. 2. In the reverberation reproduction, we design reverberation-steering filters which consist of inverse filters and the desired impulse response based on the multi-channel LMS (4). The desired reverberation is reproduced at the listening point by controlling each indirect electrodynamic loudspeaker with reverberation-steering filters.

3.1 Principle of the proposed system

In the proposed system, the parametric loudspeaker represents the direct sound, and indirect electrodynamic loudspeakers represent the reverberation in order to reproduce the desired impulse response. Figure 3 shows the principle of the proposed system. In Fig. 3, the parametric loudspeaker emits the input signal modulated by amplitude, and indirect electrodynamic loudspeakers emit the input signal convoluted with reverberation-steering filters. Then, emitted sounds from each loudspeaker should go through each transfer function and should arrive at the listening point. Listeners will observe the signal which is the sum of emitted sounds at the listening point. Equation (2) defines the reproduced impulse response $IR_{proposed}(k)$ at the listening point.

$$IR_{proposed}(k) = g_0(k) + \sum_{m=1}^{M} h_m(k) * g_m(k). \tag{2}$$

In Fig. 3 and Eq. (2), where k is time index, M is the number of indirect electrodynamic loudspeakers, $x(k)$ is the input signal, $h_m(k)$ is the reverberation-steering filter for the m-th indirect electrodynamic loudspeaker, $g_0(k)$ is the impulse response between the parametric loudspeaker and the listening point, $g_m(k)$ is the impulse response between the m-th indirect electrodynamic loudspeaker and the listening point, and the symbol * stands for the convolution. Furthermore, Eq. (3) defines frequency characteristic of reverberation-steering filters.

$$H_m(z) = R(z)/G_m(z), \tag{3}$$

where z is the frequency index, $H_m(z)$ is the signal which is projection of $h_m(k)$ to the axis of the frequency, $G_m(z)$ is the signal which is projection of $g_m(k)$ to the axis of the frequency, and $R(z)$ is frequency characteristic of reverberation in the desired sound field.

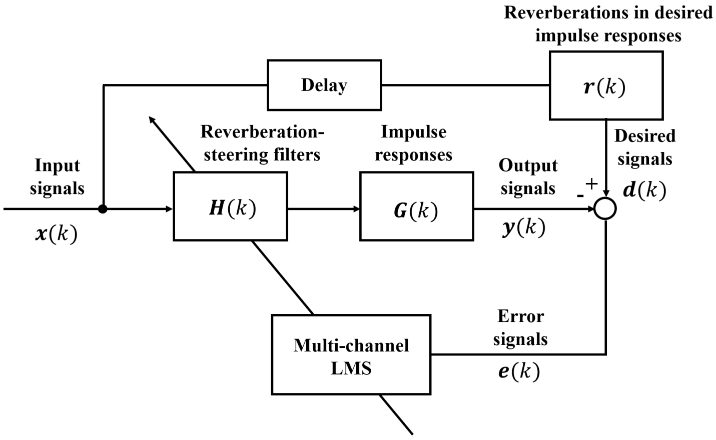

Figure 4 – The design of reverberation-steering filters

3.2 Filter design based on multi-channel LMS

The proposed system employs the multi-channel LMS (4) in order to design filters considered in electro-acoustic characteristics such as the A/D converter and transfer functions between indirect electrodynamic loudspeakers and observed points. The multi-channel LMS is an adaptive algorithm and based on the MINT (Multi-input/output INverse Theorem) (5). It can steer a reverberation at M observed points by using $M+1$ loudspeakers. Figure 4 shows the design of reverberation-steering filters. In Fig. 4, k is time index, $\mathbf{x}(k)$ is the matrix of input signals, $\mathbf{H}(k)$ is the matrix of reverberation-steering filters, $\mathbf{G}(k)$ is the matrix of measured impulse responses, $\mathbf{y}(k)$ is the matrix of output signals, $\mathbf{d}(k)$ is the matrix of desired signals, $\mathbf{r}(k)$ is the matrix of reverberations in desired impulse responses, and $\mathbf{e}(k)$ is the matrix of error signals. In the proposed system, we adaptively design reverberation-steering filters based on the multi-channel LMS. This system represents the desired reverberation at observed points by emitting signals convoluted with reverberation-steering filters from indirect electrodynamic loudspeakers. Equation (4) defines the update of the multi-channel LMS.

$$\mathbf{H}(k+1) = \mathbf{H}(k) + 2\mu\mathbf{G}(k)^T\mathbf{e}(k), \tag{4}$$

where $\mathbf{H}(k)$ is the matrix of reverberation-steering filters, $\mathbf{G}(k)$ is the matrix of measured impulse responses, $\mathbf{e}(k)$ is the matrix of error signals, T represents the transposition, and μ is a constant called the step size in order to control the size of the update of the coefficient. Each matrix is defined as Eqs. (5)～(10).

$$\begin{aligned}
\mathbf{H}(k) &= [\mathbf{h}_1(k)^T, \mathbf{h}_2(k)^T, \ldots, \mathbf{h}_m(k)^T]^T, && (5)\\
\mathbf{h}_m(k) &= [h_{m1}(k), h_{m2}(k), \ldots, h_{mJ}(k)]^T, && (6)\\
\mathbf{G}(k) &= [\mathbf{g}_1(k), \mathbf{g}_2(k), \ldots, \mathbf{g}_l(k)]^T, && (7)\\
\mathbf{g}_l(k) &= [\mathbf{g}_{l1}(k)^T, \mathbf{g}_{l2}(k)^T, \ldots, \mathbf{g}_{lm}(k)^T]^T, && (8)\\
\mathbf{g}_{lm}(k) &= [g_{lm1}(k), g_{lm2}(k), \ldots, g_{lmJ}(k)]^T, && (9)\\
\mathbf{e}(k) &= [e_1(k), e_2(k), \ldots, e_l(k)]^T, && (10)
\end{aligned}$$

where k is time index, l is the index of the observed point, m is the index of the indirect electrodynamic loudspeaker, J represents the total number of filter coefficients.

4. EVALUATION EXPERIMENT

In the proposed system, the parametric loudspeaker represents a sound image, indirect electrodynamic loudspeakers represent a reverberation. Therefore, we carried out the objective evaluation experiment on the performance of the sound image localization and the reproducibility of the reverberation to confirm the effectiveness of the proposed system.

4.1 Experimental conditions

Table 1 shows experimental conditions, and Fig. 5 shows experimental arrangements. This experiment utilized the parametric loudspeaker, six indirect electrodynamic loudspeakers and the dummy head as shown in Fig. 5. The dummy head is the equipment for the binaural recording. The parametric loudspeaker represents the sound image to the dummy head from three directions, θ =-45, 0, -45 [deg.]. Indirect electrodynamic loudspeakers represent the reverberation to the dummy head. Furthermore, the Inter-Aural Cross Coefficient

Table 1 – Experimental conditions

Parametric loudspeaker	MITSUBISHI, MSP-50E
Indirect electrodynamic loudspeaker	BOSE, MODEL-101VM
Dummy head	NEUMANN, KU100
Sampling frequency	192 [kHz]
Quantization	16 [bit]
Ambient noise	$L_A = 39.7$ [dB]
Sound source	White noise (2 [s])
Reverberation time in the experimental room	$\mathrm{T}_{60} = 0.4$ [s]

(IACC) (6) is employed as the evaluation index to evaluate the performance of the sound image localization. The IACC represents the correlation of acoustic signals arriving at left and right ears. It is based on the normalized Inter-Aural Cross correlation Function (IACF) (6). The IACF and the IACC are defined as Eqs. (11) and (12).

$$IACF_{t_1,t_2}(\tau) = \frac{\sum_{t=t_1}^{t_2} p_l(t)\cdot p_r(t+\tau)}{\sqrt{\sum_{t=t_1}^{t_2} p_l^2(t)}\sqrt{\sum_{t=t_1}^{t_2} p_r^2(t)}}, \tag{11}$$

$$IACC_{t_1,t_2} = \max|IACF_{t_1,t_2}(\tau)|, \quad |\tau| \leq 1[ms], \tag{12}$$

where t is time index, t_1 and t_2 are measured time, $p_l(t)$ is the acoustic signal arriving at left ear, $p_r(t)$ is the acoustic signal arriving at right ear, and τ is the time delay.

The reverberation time (7) is employed as the evaluation index to evaluate the reproducibility of the reverberation. This index represents time required for a sound in a room to decay by 60 [dB]. It is derived from the reverberation curve (7) that is defined as Eq. (13).

$$< S^2(t_r) >= \sum_{t=t_r}^{T} p^2(t), \tag{13}$$

where $<>$ is the ensemble average, t_r is the border time between the early reflection sound and the late reverberation, T is the length of the signal, $p(t)$ is the measured impulse response, and $< S^2(t_r) >$ is the reverberation curve.

Moreover, the D (Definition) value (7) is also employed as the evaluation index to evaluate the reproducibility of the reverberation. This index represents the energy ratio of the entire signal and the early reflection sound. It is defined as Eq. (14).

$$D = \frac{\sum_{t=0}^{t_r-1} p^2(t)}{\sum_{t=0}^{T-1} p^2(t)}, \tag{14}$$

where t_r is the border time between the early reflection sound and the late reverberation, T is the length of the signal, and $p(t)$ is the measured impulse response.

4.2 Experimental results

Table 2 shows the result of the objective evaluation experiment for the performance of the sound image localization, and Figs. 6, and 7 show results of the objective evaluation experiment for the reproducibility of the reverberation. Table 2 shows the IACC of each representation direction of the sound image in the former and proposed systems. The range of the IACC is from 0 to 1. The performance of the sound image localization increases depending on the increase of the IACC. In Tbl. 2, the IACC in the proposed system is approximately 0.6 or more in each direction. Therefore, the performance of the sound image localization in the proposed system is higher than that of the former system. Next, Fig. 6 shows the reverberation time in the target environment, the former and the proposed systems. The reproducibility of the reverberation time is higher when the reproduced reverberation time is close to the reverberation time in the target environment.

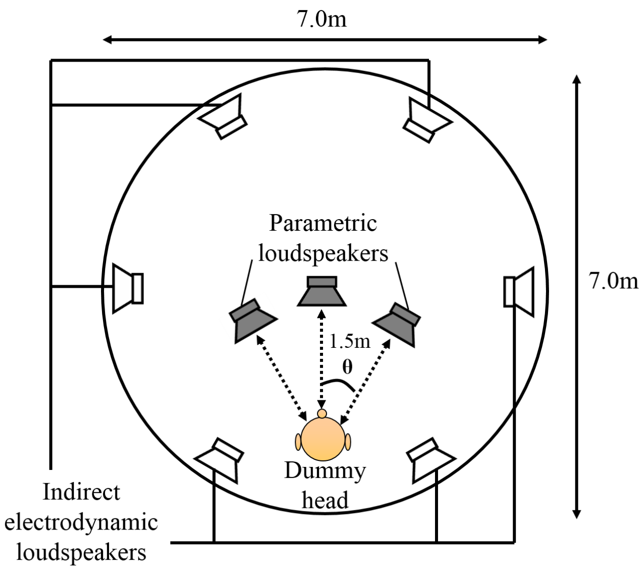

Figure 5 – Experimental arrangements

Table 2 – The IACC of the former and the proposed systems

	The desired target direction of the arrival θ		
	−45 [deg.]	0 [deg.]	45 [deg.]
The former system	0.31	0.36	0.29
The proposed system	0.57	0.94	0.60

In Fig. 6, reproduced reverberation times in the former and proposed systems are close to the reverberation time in the target environment. Therefore, both of the former and the proposed systems can reproduce the reverberation time in the target environment. Figure 7 shows the D value in the target environment, the former and the proposed systems. The blue bar is the D value using 20 [ms] as the border time between the early reflection sound and the late reverberation. Similarly, the red one is the D value using 30 [ms], and the green one is the D value using 40 [ms]. The reproducibility of the D value is higher when the reproduced D value is close to the D value in the target environment. In Fig. 7, the reproduced D value in the proposed system is close to the D value in the target environment than the reproduced D value in the former system. Therefore, the proposed system can accurately reproduce the D value in the target environment than the former system. Accordingly, we confirmed that the proposed system can highly maintain the performance of the sound image localization and accurately reproduce desired reverberant environments.

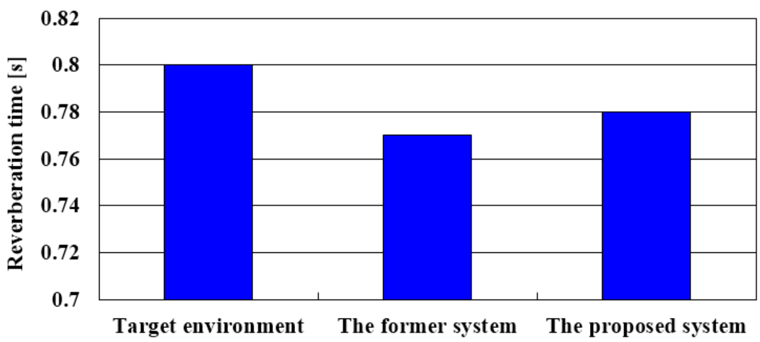

Figure 6 – The reverberation time in the target environment, the former and the proposed systems

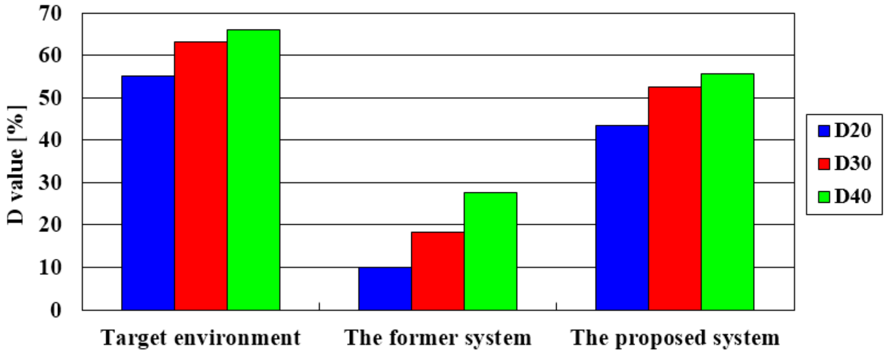

Figure 7 – The D value in the target environment, the former and the proposed systems

5. CONCULUSIONS

In this paper, we proposed the 3-D sound field localization system using the parametric loudspeaker and indirect electrodynamic loudspeakers for the reverberation reproduction by utilizing reverberation-steering filters which consist of inverse filters and the desired impulse response in the reverberation reproduction. Specifically, in the reverberation reproduction, this system independently steers each indirect electrodynamic loudspeaker by designing reverberation-steering filters based on the multi-channel LMS in order to reproduce desired reverberant environments. As a result of the objective evaluation experiment, we confirmed the proposed system can highly maintain the performance of the sound image localization and accurately reproduce desired reverberant environments. We intend to expand a reverberation-reproduced area based on the reverberation time and the D value.

ACKNOWLEDGEMENTS

This work was partly supported by Grants-in-Aid for Scientific Research funded by Japan's Ministry of Education, Culture, Sports, Science, and Technology.

REFERENCES

1. Henrik Moller, “Fundamentals of binaural technology,” Applied Acoustics, Vol. 36, No. 3-4, pp. 171-218, 1992.
2. Jerry Bauck and Duane H. Cooper, “Generalized transaural stereo and applications," Journal of the Audio Engineering Society, Vol. 44, No. 9, pp. 683-705, 1996.
3. Yutaro Sugibayashi, Sota Kurimoto, Daisuke Ikefuji, Masanori Morise and Takanobu Nishiura, “Three-dimensional acoustic sound field reproduction based on hybrid combination of multiple parametric loudspeakers and electrodynamic subwoofer," Applied Acoustics, Vol. 73, No. 12, pp. 1282-1288, 2012.
4. Stephen J. Elliott, Ian M. Stothers, and Philip A. Nelson, “A multiple error LMS algorithm and Its application to the active control of sound and vibration,” IEEE Transactions on Acoustics, Speech and Signal Processing, Vol. 35, No. 10, pp. 1423-1435, 1987.
5. Masato Miyoshi and Yutaka Kaneda, “Inverse filtering of room acoustics,” IEEE Transactions on Acoustics, Speech and Signal Processing, Vol. 36, No. 2, pp. 145-152, 1988.
6. Kenji Fujii, Yoshiharu Soeta and Yochi Ando, “Acoustical properties of aircraft noise measured by temporal and spatial factors,” Journal of Sound and Vibration, Vol. 241, No. 1, pp. 69-78, 2001.
7. Heinrich Kuttruff, “Room Acoustics, Fifth Edition,” Spon Press, 2009.