Rapid Compensation of Temperature Fluctuation Effect for Multichannel Sound Field Reproduction System

Yuki YAI¹, Nonmember, Shigeki MIYABE¹, Hiroshi SARUWATARI¹, Members, Kiyohiro SHIKANO¹, Fellow, and Yosuke TATEKURA¹, Member

SUMMARY In this paper, we propose a computationally efficient method of compensating temperature for the transaural stereo. The conventional method can be used to estimate the change in impulse responses caused by the fluctuation of temperature with high accuracy. However, the large amount of computation required makes real-time implementation difficult. Focusing on the fact that the amount of compensation depends on the length of the impulse response, we reduce the computation required by segmenting the impulse response. We segment the impulse responses in the time domain and estimate the effect of temperature fluctuation for each of the segments. By joining the processed segments, we obtain the compensated impulse response of the whole length. Experimental results show that the proposed method can reduce the computation required by a factor of nine without degradation of the accuracy.

key words: sound field control, temperature fluctuation, computational efficiency

1. Introduction

The transaural stereo [1]–[4] can realize 3-D auditory display with much greater presence than discrete surround systems, which include ordinary stereophonic and wide-spread five-loudspeaker systems for home theatres. This system reproduces the same room transfer functions as those in the primary field, after removing those in the secondary sound field using an inverse filter. This reproduction allows the listener to experience the same sound field as that in the primary sound field. The signal used in this system is recorded using a special technique, called binaural recording [5]. The microphones used for capturing sound in the primary field are set on the ears of a human or a life-size mannequin called a head and torso simulator (HATS). These observed signals are exactly those that a person in the primary sound field would hear. Transaural stereo makes the sound pressure at the listener’s ears the same as the binaural recording signals.

The inverse filter can only compensate the transfer functions that are measured in advance. Since room transfer functions form a time-varying system, the inverse filter should adapt to the variation. Ordinary filter adaptation based on the conventional least-squares-error (LSE) crite-

Manuscript received August 4, 2007.
¹The authors are with the Graduate School of Information Science, Nara Institute of Science and Technology, Ikoma-shi, 630-0192 Japan.
¹¹The author is with Faculty of Engineering, Shizuoka University, Hamamatsu-shi, 432-8561 Japan.
a) E-mail: sawatari@is.naist.jp
DOI: 10.1093/ietfsec/e91-a.6.1329

Copyright © 2008 The Institute of Electronics, Information and Communication Engineers — 30 —
2. Conventional Sound Field Reproduction Using Compensation of Temperature

2.1 Sound Field Reproduction

Here we review the transaural stereo based on the multiple-input/output inverse theorem [11]. Our goal is to reproduce $P$ desired signals $X_p(k)$ at $P$ control points $C_p$ ($1 \leq p \leq P$), where $k$ denotes the index of the frequency bin. Figure 1 shows the configuration of the sound field control. To design an inverse filter that strictly removes the effect of room transfer functions with nonminimum phases, the number $S$ of loudspeakers $L_s$ ($s = 1, \ldots, S$) must be larger than the number $P$ of control points [11]. Then we measure all the room transfer functions $G_{ps}(k)$ between $L_s$ and $C_p$ to form a $P \times S$ matrix $G(k)$ whose entry of the $p$th row and the $s$th column is $G_{ps}(k)$. We design an inverse filter matrix $H(k)$ with a Moore-Penrose generalized inverse matrix of $G(k)$ [12], which satisfies

$$G(k)H(k) = I_P,$$

(1)

where $I_P$ is the $P$-dimensional identity matrix. The loudspeakers play back the output signals of the inverse filter, and the input signals $X(k)$ to the inverse filter $H(k)$ are the desired signals and are denoted by

$$X(k) = [X_1(k), \ldots, X_P(k)]^T,$$

(2)

where $X_p(k)$ is the desired signal at $C_p$. Then the signals $D_p(k)$ ($1 \leq p \leq P$) at the $P$ control points are expressed as

$$D(k) = [D_1(k), \ldots, D_P(k)]^T = G(k)H(k)X(k) = X(k).$$

(3)

This condition shows that the observed signals $D(k)$ are the desired signals $X(k)$, and the desired signals are reproduced. Utilizing this mechanism, we set a HATS at the assumed standing position of the listener, and design the inverse filter of the transfer function measured at the ears of the HATS. Using this inverse filter, binaural recordings can be reproduced at the listener’s ears.

2.2 Linear Warping of Time Axis

A change in temperature causes a change in the transfer function, and then the inverse filter cannot hold the inverse characteristics. This change is due to the change in acoustic velocity, and appears as a rescale (or “warp”) of the impulse responses. Assuming that the temperature changes from $T_0$°C to $T_1$°C, the warping ratio $W$ can be expressed as

$$W = \frac{331.5 + 0.6T_0}{331.5 + 0.6T_1},$$

(4)

where the acoustic velocity at 0°C is 331.5 m/s. Utilizing this model, an estimation method of warped impulse responses has been proposed [9]. Assume that an N-point impulse response $g(n)$ (0 ≤ n < N denotes time sample index) in the time domain is warped into $\tilde{g}(n, W)$ with a warping ratio $W$. At first, we obtain the complex spectrum $G(k)$ (0 ≤ k < N) of $g(n)$ using a fast Fourier transform (FFT):

$$G(k) = \sum_{n=-N/2+1}^{N/2} g(n)e^{-j\frac{2\pi nk}{N}},$$

(5)

where $j = \sqrt{-1}$. An inverse discrete Fourier transform (DFT) with (1/W)-fold linearly scaled phases gives the estimation $\tilde{g}(m, W)$ of the impulse response $\tilde{g}(m, W)$ after the temperature fluctuation as

$$\tilde{g}(m, W) = \frac{1}{M} \sum_{k=-M/2+1}^{M/2} G(k)e^{j\frac{2\pi km}{M}},$$

(6)

where $M$ denotes the maximum integer number which does not exceed $N$ and $WN$. By updating the inverse filter with the impulse responses warped in the above manner, we can construct an adaptive sound field reproduction system with high robustness against temperature fluctuation.

2.3 Adaptive Estimation of Warping Ratio

Although we need precise information on temperature to apply the compensation technique described in Sect. 2.2, thermometers have bias and error that are too large for the estimation of the warping ratio. To address this problem, some researchers have proposed a direct estimation technique of the warping ratio without measuring the temperature itself [8]. To estimate the warping ratio, this method uses a microphone and reproduces signals not only at the listener’s ears but also at the microphone (where silent signal with zero amplitude is reproduced in our experiments).

Figure 2 shows the configuration of the adaptive sound field reproduction. In the following, we assume that the microphone is set at the $P$-th control point $C_p$. Using the estimated transfer function $G_{ps}(k, W)$ from $L_s$ to $C_p$, the estimation $\tilde{D}_p(k)$ of the observed signal can be expressed as
\[
\hat{D}_p(k, W) = \hat{G}_p(k, W)H(k)X(k),
\]
where the S-dimensional row vector \(\hat{G}_p(k, W)\) is denoted as
\[
\hat{G}_p(k, W) = \{\hat{G}_p(1, W), \ldots, \hat{G}_p(S, W)\}.
\]
The squared error of the estimated observation is given by
\[
\varepsilon(W) = \sum_k |\hat{D}_p(k) - \hat{D}_p(k, W)|^2,
\]
where \(\hat{D}_p(k)\) is the observed signal at \(C_p\) after the fluctuation. We estimate \(W\) iteratively by the steepest descent method of \(\varepsilon(W)\) as
\[
W^{(i+1)} = W^{(i)} - \alpha \left. \frac{\partial \varepsilon(W)}{\partial W} \right|_{W=W^{(i)}},
\]
where \(W^{(i)}\) is the estimated warping ratio at the \(i\)th iteration and \(\alpha\) is the step-size parameter. The partial differential in Eq. (10) can be written as
\[
\frac{\partial \varepsilon(W)}{\partial W} = 2 \sum_k \left| \hat{D}_p(k) - \hat{D}_p(k, W) \right|^2
\]
\[
-2 \sum_n \left| \hat{D}_p(n) - \hat{D}_p(n, W) \right|^2
\]
\[
-2 \sum_n \hat{G}_p(n, W) \frac{\partial \hat{G}_p(n, W)}{\partial W} H(n)X(n),
\]
where \(\ast\) denotes the complex conjugate, and \(\hat{G}_p(k, W)/\partial W\) is the Fourier transform of its time-domain expression \(\partial \hat{G}_p(n, W)/\partial W\) given by
\[
\frac{\partial \hat{G}_p(n, W)}{\partial W} = \frac{1}{M} \sum_{k=-N/2+1}^{N/2} G_p(k) e^{-j2\pi kn/W^2}.
\]
where \(M\) is the maximum integer that does not exceed \(N\) and \(WN\).

3. Proposed Method

3.1 Strategy

In conventional DFT operation with \(N\) sample points, computational cost can be reduced from \(O(N^2)\) to \(O(N \log N)\) using an FFT. However, for inverse DFT operations accompanied by phase modifications in Eqs. (6) and (12), an FFT cannot be applied. Since the typical length of the impulse response used for high-quality sound field reproduction is over 5,000 with the sampling frequency of 48,000 Hz, the effect of \(O(N^2)\) is critical. Our proposed strategy for the reduction of computation time is segmentation of the impulse response. By segmenting the \(N\) points into \(L\) parts and warping each of them, the number of computation can be reduced into \(O(LN/2) = O(N^2/L)\). Thus the computational cost reduces according to increase of the number \(L\) of segmentation.

3.2 Segmentation of Warping

The configuration of the proposed method is shown in Fig. 3. First, we segment the impulse response using the hanning window. Next, we warp each of the windowed parts separately. Finally, the warped impulse response with full length is obtained by gathering the warped parts using overlap-and-add method.

Here we discuss segmentation of the impulse response \(g(n)\) with \(N\) points in Eq. (6) into parts with \(T (= N/L)\) points. First, we pad each of both beginning and end of \(g(n)\) with \(T/2\) zeros to prevent loss of the waveform near the ends using the hanning window. The impulse response including the zero padding is denoted by \(g^{\text{log}}(n)\) \((0 \leq n < N + T)\), given by
\[
g^{\text{log}}(n) =\begin{cases} 
0 & (0 \leq n < \frac{T}{2}) \\
g(n - \frac{T}{2}) & (\frac{T}{2} \leq n < N + \frac{T}{2}) \\
0 & (N + \frac{T}{2} \leq n < N + T) 
\end{cases}
\]
Next, a $T$-point hanning window with $T/2$ shifting points is used to segment $\hat{g}_l^{\mathrm{long}}(n)$ into $2L+1$ parts, denoted by $f_l^{\mathrm{win}}(t)$ ($0 \leq l \leq 2L$, $0 \leq t < T$), as

$$f_l^{\mathrm{win}}(t) = \hat{g}_l^{\mathrm{long}}\left(\frac{T}{2}l + t\right)w(t),$$  \hspace{1em} (14)

where $w(t)$ is the hanning window given by

$$w(t) = 0.5 - 0.5 \sin\left(\frac{n\pi T}{T}\right).$$  \hspace{1em} (15)

The overlap of the hanning window with $T/2$ points is used to prevent the circular convolution effect, which occurs when a rectangular window without an overlap is used. Subsequently, we obtain warped parts $f_l^{\mathrm{win}}(t, W)$ ($0 \leq t < U$) using the operation in Eq. (6), as

$$f_l^{\mathrm{win}}(t, W) = \frac{1}{U} \sum_{m=-\frac{U}{2}+1}^{\frac{U}{2}} F_l^{\mathrm{win}}(m)e^{-j\frac{2\pi}{U}m},$$   \hspace{1em} (16)

where $U = T$ if $W \geq 1$, otherwise $U$ is the maximum integer that does not exceed $WT$, and $F_l^{\mathrm{win}}(m)$ denotes the FFT of $f_l^{\mathrm{win}}(t)$. Next, we place $f_l^{\mathrm{win}}(t, W)$ on the correct position; while the starting sample of $f_l^{\mathrm{win}}(t)$ is $IT/2$, $f_l^{\mathrm{win}}(t, W)$ should be placed on the $(W(l - 1) + 1)T/2$-th sample. Although the $Q(l, W)$-point integer sample shift is a simple allocation of the window, the $q(l, W)$-point shift smaller than a sample should be conducted in frequency domain, where $Q(l, W)$ is the maximum integer value smaller than $(W(l - 1) + 1)T/2$ and

$$q(l, W) = \frac{(W(l - 1) + 1)T}{2} - Q(l, W).$$  \hspace{1em} (17)

To obtain the $q(l, W)$-point shift, we obtain the frequency response $F_l^{\mathrm{shift}}(m, W)$ of $f_l^{\mathrm{win}}(t, W)$ using FFT with suitable length of $R$ points which is larger than $WT$. Then the $q(l, W)$-point shifted segment $F_l^{\mathrm{shift}}(m, W)$ in the frequency domain is obtained as

$$F_l^{\mathrm{shift}}(m, W) = F_l^{\mathrm{win}}(m, W)e^{-j2\pi l m/W}. $$  \hspace{1em} (18)

The warped segment positioned in the correct position, denoted by $f_l^{\mathrm{long}}(n, W)$, is obtained by positioning inverse FFT $f_l^{\mathrm{shift}}(t, W)$ of $F_l^{\mathrm{shift}}(m, W)$ on the $Q(l, W)$-th position as

$$f_l^{\mathrm{long}}(n, W) = \begin{cases} f_l^{\mathrm{shift}}(n - \frac{l}{2}, W) & Q(l, W) \leq n < Q(l, W) + R \\ 0 & \text{otherwise}. \end{cases}$$  \hspace{1em} (19)

Finally, the whole warping $\hat{g}_l^{\mathrm{long}}(n, W)$ of $g^{\mathrm{long}}(n)$ is obtained by summing $f_l^{\mathrm{long}}(n, W)$ over $0 \leq l \leq 2L$ as

$$\hat{g}_l^{\mathrm{long}}(n, W) = \sum_{l=0}^{2L} f_l^{\mathrm{long}}(n, W).$$  \hspace{1em} (20)

The warping of $g(n)$ with warping ratio $W$ is obtained by removing the $T/2$ zeros from both the beginning and the end of $\hat{g}_l^{\mathrm{long}}(n, W)$ as

$$\hat{g}(n, W) = \hat{g}_l^{\mathrm{long}}\left(n + \frac{T}{2}, W\right) \text{ for } 0 \leq n < M. \hspace{1em} (21)$$

3.3 Segmentation of Partial Differential

Here we discuss segmentation of the partial differential of $g(n)$ in Eq. (12). First, we obtain $2L+1$ segments $g_{pj}^{\mathrm{win}}(t)$ ($0 \leq l \leq 2L$) of $g_{p}(n)$ in the same manner as Eq. (13) in Sect. 3.2, i.e., using zero padding and hanning windows. Next, using the FFT $G_{pj}^{\mathrm{win}}(m)$ of $G_{pj}^{\mathrm{win}}(t)$, we calculate the segmented partial differential $\partial g_{pj}^{\mathrm{win}}(t, W)/\partial W$ ($0 \leq t < U$) as

$$\frac{\partial g_{pj}^{\mathrm{win}}(t, W)}{\partial W} = \frac{1}{U} \sum_{m=-\frac{U}{2}+1}^{\frac{U}{2}} G_{pj}^{\mathrm{win}}(m, W)e^{-j\frac{2\pi}{U}m},$$  \hspace{1em} (22)

where $\tau_l$ compensates for the effect of segmentation on the differential of the exponential term and

$$\tau_l = \frac{T}{2} W(l - 1).$$  \hspace{1em} (23)

By the overlap and add of $(\partial g_{pj}^{\mathrm{win}}(t, W))/\partial W$ in the same manner as Eqs. (18)-(21), $\partial g_{p}(n, W)/\partial W$ in Eq. (12) is obtained.

3.4 Computational Efficiency

By increasing the segmentation number $L$, we can reduce the number of multiplications. First, we discuss the computational cost of warping using Eq. (6). The conventional method requires $4N^2 + N \log N$ multiplications for an $N$-point FFTs and DFT with phase modification. On the other hand, the proposed method in Sect. 3.2, we calculate the warping, windowing, FFT, shifting and inverse FFT with $N/L$ points for $(2L+1)$ times, where the length of FFT $R = N/L$ in the overlap-add is assumed. Thus the ratio of the number of multiplications between the proposed and conventional methods is

$$\frac{\left[4\left(\frac{N}{L}\right)^2 + 3\frac{N}{L} \log \left(\frac{N}{L}\right) + 3\frac{N}{L}\right] \cdot (2L+1) + \frac{4}{4N^2 + N \log N}. \hspace{1em} (24)$$

For the segmentation of the partial differential using Eq. (12), the conventional method requires $5N^2 + N \log N$ multiplications. The ratio of the number of multiplications between the proposed and the conventional methods is

$$\frac{\left[5\left(\frac{N}{L}\right)^2 + 3\frac{N}{L} \log \left(\frac{N}{L}\right) + 3\frac{N}{L}\right] \cdot (2L+1) + \frac{5}{5N^2 + N \log N}. \hspace{1em} (25)$$
4. Experiments and Discussions

To show the efficacy of the proposed method, we conducted simulations using impulse responses measured in a real room.

4.1 Experimental Conditions

In this experiment, we assume sound field reproduction using 16 loudspeakers and seven control points, where six points are set around the listener's ears and the other is set at a microphone to observe the error. Figure 4 shows the layout of the room where the impulse responses are measured. We set a HATS at the center of the room. We also set a microphone 1.5 m to the left of the HATS, and a loudspeaker used as primary source is placed 1.5 m in front of the HATS. We show the parameters of the measurement in Table 1. To simulate the variation of temperature, we measured impulse responses several times while changing the temperature using air conditioning. In the measurement of the computation time, the computation is performed by MATLAB on a PC equipped with the Intel Xeon 64-bit processor whose clock frequency is 3.0 GHz.

4.2 Evaluation of Proposed Fast Compensation

We evaluated the performance of temperature compensation at the six control points around the HATS. Each of the impulse responses has 16384-point length. We compared the performances of the conventional method and the proposed method for each of the segmentation numbers $L = 2^0, 2^1, \ldots, 2^{12}$.

4.2.1 Accuracy of Compensation

First, we compared the accuracy of the temperature compensation using the proposed method. The evaluation score is the signal-to-noise ratio (SNR) given by

$$\text{SNR [dB]} = 10 \log \frac{\sum_n |y(n)|^2}{\sum_n |y(n) - \hat{y}(n)|^2}$$

(26)

where $y(n)$ and $\hat{y}(n)$ are the desired signal and the signal obtained by reproduction, respectively. A higher SNR indicates higher accuracy. We evaluated the average SNR at the six control points around the ears of the HATS.

Figures 5 and 6 show the performances of the reproduction of the impulse for increase and decrease in temperature. The temperature varies between 25.0°C and 26.0°C. During both the increase and the decrease, the segmentation does not degrade the accuracy of compensation when the segmentation number $L$ is below $2^{10}$. One of the main reasons for the error when using a large segmentation number is the circular convolution effect.

Here we discuss the degradation of accuracy caused by large segmentation number. The warping with Eq. (6) truncates the spectrum around highest frequency of sampling theorem as it uses the spectral values only in the lower frequency than that of $M/2$-th frequency bin. Since the shape of frequency bin in the DFT is not rectangular but sinc-function-like, the error caused by the truncation propagates around the truncated frequency bin. Thus the warping cannot calculate the spectrum of high frequency accurately. The error appears as circular convolution in the time domain and can be reduced to some extent using appropriate window function, as we use hanning window. However, when the

![Fig. 4 Layout of acoustic environment room.](image)

**Table 1** Parameters in the experiments.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sampling frequency</td>
<td>48 kHz</td>
</tr>
<tr>
<td>Reverberation time</td>
<td>160 ms</td>
</tr>
<tr>
<td>Quantization resolution</td>
<td>24 bit</td>
</tr>
</tbody>
</table>

![Fig. 5 Relation between the segmentation number and accuracy of compensation when temperature raised from 25.0°C to 26.0°C.](image)

![Fig. 6 Relation between the segmentation number and accuracy of compensation when temperature fell from 26.0°C to 25.0°C.](image)
large segment number is set, the short width of the window augments the bandwidth of each frequency bin and the effect of the truncation reaches the low frequency. Thus the proposed segmentation cannot calculate the warping accurately with too large segment number. Figure 7 shows an evaluation of accuracy in each frequency bin. Here we use two warping ratios $W_1 = 0.99$ and $W_2 = 1/W_1$. Sinusoids who has spectral value in only a single frequency bin is warped with one ratio, and recovered by warping again with the other. The plot is the SNR of the recover of the original sinusoid. The degradation of accuracy appears in the high-frequency bins.

4.2.2 Computational Efficiency

Figure 8 shows the ratio of the amount of computation of the proposed method to that of the conventional method. The values shown are the computation time normalized by 0.7 s, which is the time required by the conventional method to calculate the warping of a channel of impulse response. In addition, Eq. (24) is plotted as a theoretical curve. In contrast to the decrease in the number of multiplications with the increase in segmentation number $L$, the process becomes complex and the amount of summation and memory allocation increases. Thus in practice, computation time becomes larger than the theoretical value and the reduction of computation time is saturated when $L$ is large. Thus, we should choose the value of $L$ carefully considering the accuracy of the calculation. In this experiment, the saturation appears around $L = 2^9$ whose computation time is about 1/50 that of the conventional method. Fortunately, the experiment in Sect. 4.2.1 shows that the segmentation number does not degrade the accuracy of the compensation.

4.3 Evaluation of Proposed Fast Adaptation

In this section, we evaluate the performance of temperature adaptation using the proposed method discussed in Sects. 3.2 and 3.3.

4.3.1 Computational Efficiency in Partial Differential

Figure 9 shows the relation between the segmentation number and the amount of computation required for the calculation of the partial differential using Eq. (12). The values shown is the computation time normalized by 3.9 s, which is the time required by the conventional method. The theoretical curve is also plotted using Eq. (25). Similarly to the computation time required for the warping discussed in Sect. 4.2.2, the reduction of the computation for the partial differential is saturated by large segmentation number. The saturation appears around $L = 2^9$ and the computation time is only about 1/200 that of the conventional method.

4.3.2 Performance of Adaptation

We compare the accuracy of reproduction using the inverse filter adapted using the proposed method. Table 2 shows the parameters used for the adaptation. The temperature is increased from 23.0°C to 26.0°C. We used music as the sound to be presented to the user while a zero signal is reproduced at the microphone to observe the error. The length of the
music is 30 s.

Figure 10 shows the values of warping ratio $W$ estimated by Eq. (12) and by its segmentation as discussed in Sect. 3.3, at each of the iterations. The upper limit is the best score obtained in a preliminary experiment where we compared many warping ratios that were sampled at a small regular interval. There was little difference among the conventional method and the proposed method with segmentation number below $2^6$ and these values converge near the upper limit.

Figure 11 shows the relation between the segmentation number and the accuracy of reproduction using inverse filter estimated at the 20th iteration. Accuracy is not degraded using a segmentation number $L$ lower than $2^6$, at which the computation time is almost saturated.

Figure 12 shows the total amount of computation required for a single iteration in adaptation. Although all the calculations in adaptation are the same except for the segmentation using Eqs. (6) and (12), the improvement of computational efficiency by the proposed method is apparent. The saturation appears when the segmentation number $L$ is larger than $2^5$, and the time required is about $1/9$ that of the conventional method, whose computation time is 167.7 s. Therefore, it is demonstrated that the proposed method can successfully reduce the amount of computation without the degradation of accuracy.

5. Conclusions

We propose a computational efficient linear warping method of impulse response for adaptation of inverse filter to fluctuation of temperature. The conventional linear warping method of $N$-point impulse response requires large computation of order $O(N^2)$ for calculation of warping and steepest decent of warping ratio. The computational complexity is a bottleneck in the real-time implementation. The proposed method reduces the computational cost into $O(N^2/L)$ by dividing the impulse response of the length $N$ into the segments of the length $N/L$, warping each of the segments, and connecting the warped segments using overlap-and-add method. The computational amount for the calculation of the gradient is also reduced by the segmentation. Experimental results show that the computation time of the warping and the steepest decent are reduced to 1/50 and 1/200 those of the conventional method, respectively. It is also discussed that the segmentation has little loss of accuracy. The reduction of computation successfully removes the bottlenecks, and the total computation time of the adaptation in an iteration is reduced to 1/9 that of the conventional method.

References


Yuki Yai was born in Kagawa, Japan, on April 24, 1982. He received the B.E. degree in electrical and electronic engineering from Shizuoka University, and received the M.E. degree in information and science from Nara Institute of Science and Technology in 2007.

Shigeki Miyabe was born in Nara, Japan, on July 1, 1978. He received the B.E. degree in electrical and electronics engineering from Kobe University in 2003, and received the M.E. and Ph.D. degrees in information and science from Nara Institute of Science and Technology (NAIST) in 2005 and 2007, respectively. He is now a Research Fellow of the Japan Society for the Promotion of Science at Graduate School of Information Science, NAIST. His research interests include sound field control, blind source separation, and array signal processing. He is a member of the IEEE and the Acoustical Society of Japan.

Hiroshi Saruwatari was born in Nagoya, Japan, on July 27, 1967. He received the B.E., M.E. and Ph.D. degrees in electrical engineering from Nagoya University, Nagoya, Japan, in 1991, 1993 and 2000, respectively. He joined Intelligent Systems Laboratory, SECOM CO., LTD., Mitaka, Tokyo, Japan, in 1993, where he engaged in the research and development on the ultrasonic array system for the acoustic imaging. He is currently an associate professor of Graduate School of Information Science, Nara Institute of Science and Technology. His research interests include array signal processing, blind source separation, and sound field reproduction. He received the Paper Awards from IEICE in 2001 and 2006, and from TAF in 2003. He is a member of the IEEE and the Acoustical Society of Japan.

Kiyohiro Shikano received the B.S., M.S., and Ph.D. degrees in electrical engineering from Nagoya University in 1970, 1972, and 1980, respectively. He is currently a professor of Nara Institute of Science and Technology (NAIST), where he is directing speech and acoustics laboratory. From 1972 to 1993, he had been working at NTT Laboratories. During 1986–1990, he was the Head of Speech Processing Department at ATR Interpreting Telephony Research Laboratories. During 1984–1986, he was a visiting scientist in Carnegie Mellon University. He received the Yonezawa Prize from IEICE in 1975, the Signal Processing Society 1990 Senior Award from IEEE in 1991, the Technical Development Award from ASJ in 1994, IPSJ Yamashita SIG Research Award in 2000, and Paper Award from the Virtual Reality Society of Japan in 2001, IEICE paper award in 2005 and 2006, and Inose award in 2005. He is a fellow of the Institute of Electrical and Electronics Engineers (IEEE) and Information Processing Society of Japan, and a member of the Acoustical Society of Japan (ASJ), Japan VR Society, and International Speech Communication Society (ISCA).

Yusuke Tatekura was born in Kyoto, Japan on May 17, 1975. He received the B.E. degree in precision engineering from Osaka University in 1998, and received the M.E. and Ph.D. degrees in information science from Nara Institute of Science and Technology (NAIST) in 2000 and 2002, respectively. He is currently an assistant professor of Shizuoka University. His research interests include sound field control and virtual sound source synthesis. He is a member of the IEEE, the Acoustical Society of Japan, and the VR Society of Japan.