

Adaptive Compensation of Temperature Fluctuation Effect in Sound Reproduction System

Yosuke TATEKURA, Hiroshi SARUWATARI and Kiyohiro SHIKANO

Graduate School of Information Science, Nara Institute of Science and Technology

8916-5 Takayama-cho, Ikoma, Nara, 630-0101, JAPAN

yosuke-t@is.aist-nara.ac.jp

ABSTRACT

We describe a method of compensating temperature fluctuation effect by linear-time-warping processing in a sound reproduction system. This technique is applied to impulse responses of room transfer functions, to achieve a high-quality sound reproduction system, particularly one that treats high-frequency components. The impulse responses are measured before and after temperature fluctuation, and the former are converted to the latter by the proposed process. We design inverse filters for the system, and evaluate the improvement of the reproduction accuracy and spectrum distortion. Also, we propose an adaptive algorithm for the warping ratio estimation, using the observed signal of reproduced sound which is obtained at only one control point, and we can determine a suitable warping ratio which improves the reproduction accuracy at every control point by about 14 dB.

1 Introduction

To achieve a sound reproduction system with loudspeaker reproduction, it is important to design inverse filters which cancel the effects of room transfer functions (RTFs). Generally speaking, in a design of inverse filters for such sound reproduction systems, we often consider the RTFs as a time-invariant system. However, the time invariance of the RTFs cannot be guaranteed in a real acoustic environment. Thus, fluctuation of the RTFs degrades the reproduction accuracy of the inverse system. Any conventional design methods, which include the off-line method by MINT [1] or the use of the least-norm solution [2], can be used to easily design inverse filters, however, these filters cannot compensate the fluctuation of impulse responses of RTFs. On the other hand, on-line processing in inverse filter design, such as the multiple error LMS algorithm [3], can achieve a system that can adapt to fluctuation of impulse responses of RTFs. A demerit of conventional on-line methods is that microphones must be set at on all control points, which makes the sound reproduction system complicated, and inconvenient for the listeners.

Temperature fluctuation is the most important factor in the modification of the propagation time of sound because the speed of sound is a function of temperature. If the room temperature changes by 1°C, the speed of sound changes by about 0.2%. Accordingly, the propagation time of sounds is modified even if transfer channels of sound do not change; RTFs also change by about 0.2%. In addition, the exact control of high-frequency components is difficult because the

wavelength becomes shorter as the frequency increases. By degrading reproduction accuracy, sound localization accuracy is also degraded. For this reason, we should rescale (or "warp") the time axis of the impulse responses of RTFs to compensate the room temperature fluctuation.

In this paper, we apply a time-warping method [4] to the multichannel sound reproduction system [5] to obtain high-quality sound, particularly of high-frequency components. The room temperature is not always equable, and it is difficult to measure it accurately. Accordingly, it is not guaranteed that the suitable warping ratio can be obtained only from results of using a thermometer. Therefore, we must obtain the suitable warping ratio without precise measurement of the temperature in order to devise a warping method for the sound reproduction system. Hence, we newly propose an adaptive algorithm for the warping ratio estimation, which uses an observed signal measured by one of the control points.

2 Linear Warping Process

2.1 Rescaling of Time Axis

Consider the warping of the N -point-length impulse response of RTF, $g(n)$, into $g'(m)$ with warping ratio C_{th} . First, $g(n)$ is transformed into a frequency domain by DFT; this can be given by

$$G(k) = \sum_{n=0}^{N-1} g(n)e^{-j\frac{2\pi k}{N}n} \quad (0 \leq k \leq N-1), \quad (1)$$

where $G(k)$ is a frequency-domain representation of $g(n)$.

Next, we determine $g'(m)$ at every sampled point m ($0 \leq m \leq M-1$), where M is the maximum integer which does not exceed $C_{th} \cdot N$. If $C_{th} \geq 1$, $g'(m)$ can be given by

$$g'(m) = \frac{1}{N} \left\{ \sum_{k=0}^{\frac{N}{2}-1} G(k)e^{-j\frac{2\pi k}{C_{th}N}m} + \sum_{k=\frac{N}{2}}^{N-1} G(k)e^{j\frac{2\pi(k-N)}{C_{th}N}m} \right\}; \quad (2)$$

otherwise $g'(m)$ can be given by

$$g'(m) = \frac{1}{M} \left\{ \sum_{k=0}^{\frac{M}{2}-1} G(k) e^{-j \frac{2\pi k}{C_{\text{th}}^2 N} m} + \sum_{k=N-\frac{M}{2}}^{N-1} G(k) e^{j \frac{2\pi(k-N)}{C_{\text{th}}^2 N} m} \right\}. \quad (3)$$

The theoretical value of the warping ratio is defined as the ratio of two sound speeds

$$C_{\text{th}} = \frac{331.6 + 0.6T_0}{331.6 + 0.6T_t}, \quad (4)$$

where T_0 is the original room temperature and T_t is the room temperature after fluctuation at time t .

2.2 Adaptive Algorithm for Warping Ratio Estimation

The compensation of impulse responses of RTFs based on the theoretical value of the warping ratio cannot be guaranteed to be suitable because there are bias and errors in the measurement of temperature using thermometers. Each thermometer shows the temperature at only one point in the room, and temperatures in the transfer channel of sound are not always uniform. Therefore, in this section, we propose an automatic method of estimating the warping ratio using the observed signal from a microphone placed at only one of the control points.

In this method, the warping ratio which minimizes the squared error e between observed signal $d(n)$ and estimated signal $\hat{d}(n)$ is sought. The squared error e is defined as

$$e = \sum_n \left\{ d(n) - \hat{d}(n) \right\}^2, \quad (5)$$

where

$$\hat{d}(n) = \text{IFFT} \left\{ G'(k, C_{\text{th}}) H(k) X(k) \right\} \quad (6)$$

and $\text{IFFT} \{ \cdot \}$ denotes the inverse FFT operator. $G'(k, C_{\text{th}})$ is $G(k)$ warped with warping ratio C_{th} , $H(k)$ is inverse filter of $G'(k, C_{\text{th}})$, and $X(k)$ is the input signal of the system.

By using the squared error e and the i -th warping ratio $C_{\text{th}}^{(i)}$, the $(i+1)$ -th warping ratio $C_{\text{th}}^{(i+1)}$ can be updated by the steepest descent method, as shown below;

$$C_{\text{th}}^{(i+1)} = C_{\text{th}}^{(i)} - \alpha \frac{\partial e}{\partial C_{\text{th}}} \Big|_{C_{\text{th}}=C_{\text{th}}^{(i)}}, \quad (7)$$

where α is a step-size parameter. In the second term on the right-hand side of Eq. (7), the partial differentiation can be given by

$$\begin{aligned} \frac{\partial e}{\partial C_{\text{th}}} &= 2 \sum_n \left\{ d(n) - \hat{d}(n) \right\} \cdot \frac{\partial}{\partial C_{\text{th}}} \left\{ d(n) - \hat{d}(n) \right\} \\ &= 2 \sum_n \left\{ d(n) - \hat{d}(n) \right\} \\ &\quad \cdot \frac{\partial}{\partial C_{\text{th}}} \left\{ -\text{IFFT} \left[G'(k, C_{\text{th}}) H(k) X(k) \right] \right\} \\ &= 2 \sum_n \left\{ d(n) - \hat{d}(n) \right\} \\ &\quad \cdot \left\{ -\text{IFFT} \left[\frac{\partial}{\partial C_{\text{th}}} G'(k, C_{\text{th}}) H(k) X(k) \right] \right\}, \quad (8) \end{aligned}$$

where

$$\begin{aligned} \frac{\partial}{\partial C_{\text{th}}} G'(k, C_{\text{th}}) &= \frac{\partial}{\partial C_{\text{th}}} \text{FFT} \left\{ g'(m, C_{\text{th}}) \right\} \\ &= \text{FFT} \left\{ \frac{\partial}{\partial C_{\text{th}}} g'(m, C_{\text{th}}) \right\}, \quad (9) \end{aligned}$$

where $\text{FFT} \{ \cdot \}$ denotes the FFT operator. If $C_{\text{th}}^{(i)}$ is greater than or equal to 1, the partial differentiation $\partial g'(m, C_{\text{th}}) / \partial C_{\text{th}}$ in Eq. (9) can be given by

$$\begin{aligned} \frac{\partial g'(m, C_{\text{th}})}{\partial C_{\text{th}}} &= \frac{1}{N} \left\{ \sum_{k=0}^{\frac{N}{2}-1} G(k) e^{-j \frac{2\pi k}{C_{\text{th}}^2 N} m} \cdot j \frac{2\pi k m}{C_{\text{th}}^2 N} \right. \\ &\quad \left. + \sum_{k=\frac{N}{2}}^{N-1} G(k) e^{j \frac{2\pi(k-N)}{C_{\text{th}}^2 N} m} \cdot j \frac{2\pi(k-N)m}{C_{\text{th}}^2 N} \right\}; \quad (10) \end{aligned}$$

otherwise $\partial g'(m, C_{\text{th}}) / \partial C_{\text{th}}$ can be given by

$$\begin{aligned} \frac{\partial g'(m, C_{\text{th}})}{\partial C_{\text{th}}} &= \frac{1}{M} \left\{ \sum_{k=0}^{\frac{M}{2}-1} G(k) e^{-j \frac{2\pi k}{C_{\text{th}}^2 N} m} \cdot j \frac{2\pi k m}{C_{\text{th}}^2 N} \right. \\ &\quad \left. + \sum_{k=N-\frac{M}{2}}^{N-1} G(k) e^{j \frac{2\pi(k-N)}{C_{\text{th}}^2 N} m} \cdot j \frac{2\pi(k-N)m}{C_{\text{th}}^2 N} \right\}. \quad (11) \end{aligned}$$

In practice, first, the adaptation algorithm is performed in particular control point for error observation. Next, the warping ratio obtained by the algorithm is applied to other control points for which their reproduced signals are not observed by the system. By using above-mentioned procedure, we can compensate the temperature fluctuations in all control points.

3 Performance of Adaptive Algorithm for Warping Ratio Estimation

In this paper, we assume that the number of secondary acoustic sources of the reproduction system is 16 and the number of control points is 6. We evaluate that if the optimized warping ratio matches the room characteristics after temperature fluctuation by the proposed method.

3.1 Measurement of Impulse Response

The impulse responses used in this study are measured in an acoustic experiment room, in which the reverberation time is about 0.15 seconds. The arrangement of the apparatus is shown in Fig. 1. Six positions, at the two ears of the HATS and 0.05 m in front and behind them, are selected as control points at which the microphones are set, as shown in Fig. 2. Time stretched pulse [6] (TSP) was used as the sound source signal for measurement, where TSP signal length is 131072-point, sampling frequency is 48000 Hz, and addition for averaging is 4 times.

3.2 Numerical Condition

Here, we attempt to warp the impulse responses of RTF, TF_t , in those of RTF, TF_0 . The room temperature for TF_t is 27.6°C on average, and that for TF_0 is 26.2°C on average. The theoretical value of the warping ratio, C_{th} , is calculated to be 1.00240 using Eq. (4).

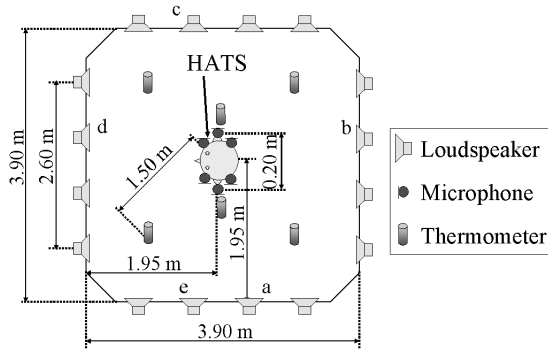


Figure 1: Layout of acoustic experiment room.

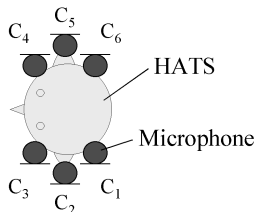


Figure 2: Arrangement of microphones.

The proposed algorithm searches for the warping ratio by using the only observed signal at C_4 . We apply the warping ratio in C_4 to the compensation of temperature fluctuation in other control points (C_1 , C_2 , C_3 , C_5 , and C_6). Hence, we perform the supervised adaptation in only C_4 . In contrast, we perform the unsupervised adaptation in other control points. The initial solution of the warping ratio is set to be 1.00000, and the process is renewed 20 times. The step-size parameter α is 0.00001. The input signal of the system $X(k)$ is an impulse response of the band-pass filter of which the passband range is 150–10000 Hz. Hereafter, we define TF'_t as TF_t warped by the theoretical warping ratio and TF''_t as the TF_t warped by the suitable warping ratio which is obtained by the adaptive method.

Using the impulse responses TF_0 , TF_t , TF'_t and TF''_t , we design inverse filters for the multichannel sound reproduction system. The inverse filters are designed by the least-norm solution in the frequency domain. The impulse responses, in which the signal length is 7200 points, are transformed into the frequency domain by FFT with a length of 32768, and inverse filters are determined. These inverse filters are filtered by a band-pass filter in which the passband range is 150–10000 Hz, and are then transformed into the time domain by a 32768-point inverse FFT. We define the inverse filters of TF_0 , TF_t , TF'_t and TF''_t as INV_0 , INV_t , INV'_t and INV''_t , respectively.

3.3 Experimental Result

Figure 3 shows a behavior of the updated warping ratio in the iterative adaptation, and the warping ratio converges at about 1.00195.

The relationship between the number of iterations and squared error from the desired signal at C_4 is shown in Fig. 4. The squared error is defined by Eq. (5). This result shows that the error can be reduced by about 18 dB. Figure 5

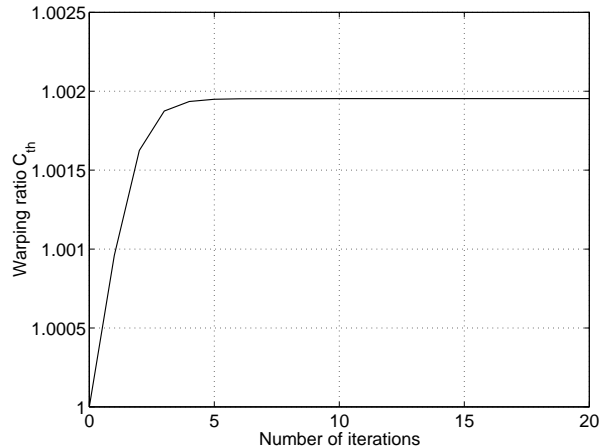


Figure 3: The relationship between the number of iterations and the warping ratio.

Table 1: Reproduction accuracy at every control point

inverse filters	SNR [dB]					
	C_1	C_2	C_3	C_4	C_5	C_6
INV_t	9.9	8.8	8.4	8.8	8.7	8.9
INV'_t	16.2	15.0	15.5	15.3	15.6	15.5
INV''_t	23.4	22.7	22.3	22.7	24.0	24.1

shows the relationship between the number of iterations and squared error at other control points for which their reproduced signals were not observed by the system, except at C_4 . The errors are found to converge almost monotonically at all control points, and reduced by about 18 dB on average.

Table 1 shows the SNRs. The original sources are obtained by convolving TF_0 with INV_0 , which are ideal responses of this reproduction system. Before the warping process, the SNR is 8.9 dB on average. After the warping process, we can improve the SNR by 14.3 dB on average. These results indicate that we can realize a high-quality sound reproduction system by using this adaptation algorithm.

The spectral characteristics of the system impulse responses are evaluated. Figure 6 shows the spectral characteristic of the system impulse response observed at C_2 . These results reveal that more error was amplified with higher frequency before using the proposed method, but the spectral characteristics became almost flat after compensation.

To evaluate the spectral characteristics numerically, we calculate the spectrum distortion (SD); this can be given by

$$SD [dB] = \sqrt{\frac{1}{n} \sum_{k=1}^n \left(20 \log \frac{|D_j(\omega_k)|}{|\hat{D}_j(\omega_k)|} \right)^2}, \quad (12)$$

where $|D_j(\omega)_k|$ is the magnitude of response of the system obtained by convolving TF_0 with INV_0 at C_j , and $|\hat{D}_j(\omega)_k|$ is that of the system obtained by convolving TF_0 with inverse filters, except for INV_0 .

Table 2 shows the SDs. By using the theoretical warping ratio, we can improve the SDs by about 3.4 dB and by about 3.9 dB using the suitable warping ratio. These results indicate that we must obtain suitable and exact warping ratios

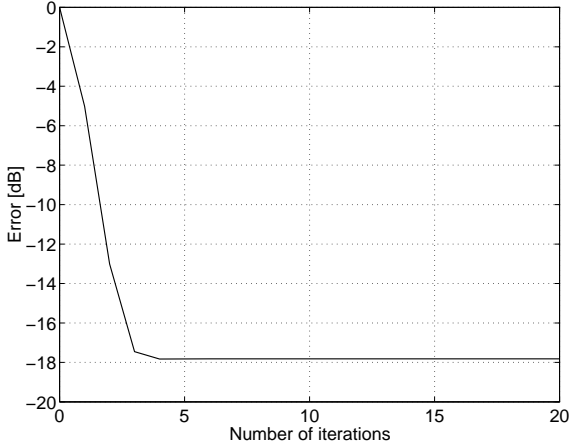


Figure 4: The relationship between the number of iterations and the squared error at C_4 . In C_4 , the reproduced signal was observed by the system for the warping ratio estimation.

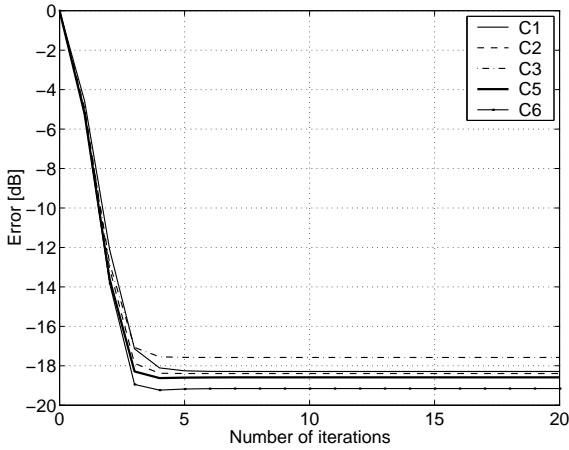


Figure 5: The relationship between the number of iterations and the squared errors at all control points except C_4 . In these control points, their reproduced signals were not observed by the system.

to realize a high-quality sound reproduction system.

4 Conclusion

We described a compensation method for temperature fluctuation by linear time warping to avoid the degradation of reproduction accuracy of a sound reproduction system. In numerical simulation using real environmental data, we could improve the reproduction accuracies using the warping process with a suitable warping ratio, in comparison with those before using the process. Also we proposed an adaptive algorithm for warping ratio estimation. By using the adaptive algorithm, we can improve the reproduction accuracy at every control point by about 14 dB.

5 Acknowledgment

This work was partly supported by CREST (Core Research for Evolutional Science and Technology) in Japan.

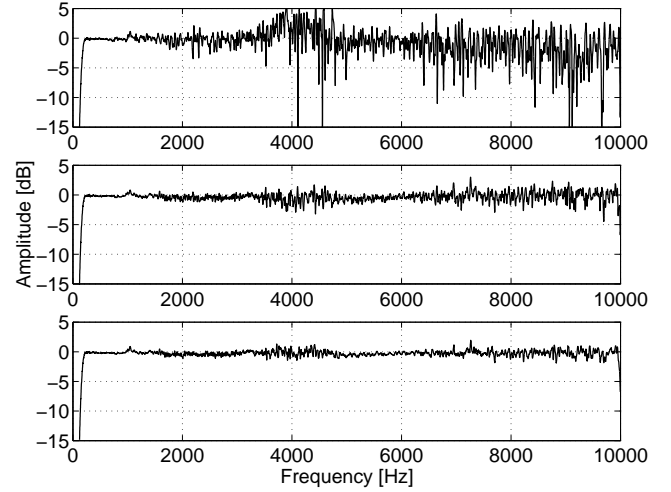


Figure 6: Frequency characteristic of system impulse response at C_4 . The top figure is the waveform with INV_t , the middle figure is that with INV'_t , and the bottom figure is that with INV''_t .

Table 2: Spectrum distortion at every control point

inverse filters	SD [dB]					
	C_1	C_2	C_3	C_4	C_5	C_6
INV_t	4.6	5.7	5.2	5.1	4.8	4.6
INV'_t	1.8	1.8	1.8	1.7	1.5	1.4
INV''_t	1.0	1.1	1.1	1.1	1.0	0.9

References

- [1] M. Miyoshi and Y. Kaneda, "Inverse filtering of room acoustics," *IEEE Trans. ASSP*, vol.36, no.2, pp.145–152, 1988.
- [2] A. Kaminuma, S. Ise and K. Shikano, "A method of designing inverse system for multi-channel sound reproduction system using least-norm-solution," *Proc. of Active99*, vol.2, pp.863–874, 1999.
- [3] S. J. Elliott, I. M. Stothers and P. A. Nelson, "A Multiple Error LMS Algorithm and Its Application to the Active Control of Sound and Vibration," *IEEE Trans. ASSP*, vol.35, no.10, pp.1423–1434, 1987.
- [4] T. Hikichi and F. Itakura, "Time variation of room acoustic transfer functions and its effects on a multi-microphone dereverberation approach," *Workshop on Microphone Arrays: Theory, Design & Application*, CAIP, 1994.
- [5] A. Kaminuma, S. Ise and K. Shikano, "Robust sound-reproduction-system design against the head movement," *Proc. of The seventh Western Pacific Regional Acoustics Conference (WESTPRAC VII)*, vol.1, pp.489–492, 2000.
- [6] Y. Suzuki, F. Asano, H.-Y. Kim, and T. Sone, "An optimum computer-generated pulse signal suitable for the measurement of very long impulse responses," *J. Acoust. Soc. Am.*, vol.97, no.2, pp.1119–1123, 1995.