14th European Signal Processing Conference (EUSIPCO 2006), Florence, Italy, September 4-8, 2006, copyright by EURASIP

SOUND REPRODUCTION SYSTEM WITH SIMULTANEOUS PERTURBATION METHOD.

Kazuya Tsukamoto, Yoshinobu Kajikawa and Yasuo Nomura

Department of Electronics, Faculty of Engineering, Kansai University
3-3-35, Yamate-cho, Suita-shi, Osaka 564-8680, Japan
phone: +81-6-6368-0828, fax: +81-6-6330-3770, email: kazu0715@joho.densi.kansai-u.ac.jp
web: http://joho.densi.kansai-u.ac.jp/index-j.html

ABSTRACT

In this paper, we propose a novel sound field reproduction system using the simultaneous perturbation (SP) method and its fast convergence version. In the conventional sound reproduction systems, the preprocessing filters are generally determined and fixed based on transfer functions from loudspeakers to control points in advance. However, movements of control points result in severe localization errors. Therefore, we propose a sound field reproduction system using the SP method which updates the filter coefficients only using error signal. The SP method suffers from the disadvantage of slow convergence, although this method can track the movements of any controlling points. Hence, we also propose an improving method of the convergence speed, which compensates only the delay by using the delay control filters. Simulation results demonstrate that the proposed methods can track the movements of control points and have reasonable convergence speed.

1. INTRODUCTION

Recently, sound field reproduction systems have been actively studied. Sound field control techniques with loudspeakers can be classified into two methods which control large areas and narrow points, respectively.

The method of controlling large areas uses the sound field control theory based on "the Kirchhoff-Helmholtz Integral Equation". A desired sound field can be reproduced in the controlled area by matching the sound pressure on the closed surface which encloses a certain area and the particle velocity of the normal direction to those of a target sound field by using "the Kirchhoff-Helmholtz Integral Equation". However, a great number of loudspeakers are needed according to the sampling theorem of the space if the sound field is controlled over the audio bandwidth of 20kHz[1]. Such an approach is impractical.

On the other hand, the latter method reproduces any desired sound at any desired point (e.g. listener's ears). To do this, the method of controlling points needs to remove the influence of transfer functions between loudspeakers and control points and crosstalk paths, i.e. the so-called crosstalk canceller. The crosstalk canceller can be realized using MINT(Multiple input-output INverse Theorem [2]), LNS(Least Norm Solution [3]), and so on. In this method, the number of loudspeakers can be generally reduced compared with the method of controlling areas. Hence, we focus on the method of controlling points because of the simple implementation.

In this method, the preprocessing filters are generally determined and fixed from transfer functions from loudspeakers to control points, which are measured in advance. Hence, the movements of control points and the changes of environments (e.g. temperature) result in severe localization errors, especially over higher frequency band.

To solve this problem, the preprocessing filters have to track the variation of transfer functions and consequently use an adaptive algorithm. The conventional adaptive algorithms must know the information of transfer functions from loudspeakers to control points. However, it would be difficult to obtain the information of transfer functions in the sound field reproduction systems because the huge computational complexity is required. Therefore, we propose a sound field reproduction system using the simultaneous perturbation (SP) method which has been applied to active noise control systems [4]. The SP method is an adaptive algorithm which updates the filter coefficients using only error signal. Hence, even if the information of transfer functions from loudspeakers to control points is unknown, the preprocessing filters can be updated to track the movements of control points. However, the conventional SP method suffers from the disadvantage of slow convergence. Therefore, we also propose a fast convergence technique which compensates only the delay using the delay control filters.

2. SOUND REPRODUCTION SYSTEM WITH PERTURBATION METHOD AND THE IMPROVING METHOD OF CONVERGENCE SPEED

2.1 Sound Reproduction System with Simultaneous Perturbation Method

Figure 1 shows a conceptual diagram of the sound field reproduction system with the SP method. Even if a listener moves, a desired sound field can be always reproduced by updating the preprocessing filters with the SP method as shown in Fig. 1.

Figure 2 shows a transoral system which consists 4 secondary sound sources and 2 control points. Least Norm Solution (LNS [3]) is used as initial values of the preprocessing filters and their filter coefficients are updated by using the SP method during system operation. The variable perturbation frequency domain time difference simultaneous perturbation(VPFDTDSP) method [4] whose convergence speed is the fastest in the SP methods is used as an updating algorithm. The updating algorithm of VPFDTDSP method is defined as follows:

$$\mathbf{h}_{ij,n+1} = \mathbf{h}_{ij,n} - \mu \Delta \mathbf{h}_{ij,n} \quad (1)$$

$$\Delta \mathbf{h}_{ij,n} = \text{first } N \text{ elements of IFFT } [\mathbf{U}_{ij,n}] \quad (2)$$

$$\mathbf{U}_{ij,n} = \mathrm{diag}[\mathbf{S}_{ij,n}] \sum_{j=1}^{2} \{\mathrm{diag}[\mathbf{E}^*_{j,n}]\mathbf{E}_{j,n} - \mathrm{diag}[\mathbf{E}^*_{j,n-1}]\mathbf{E}_{j,n-1}\}/c_n \quad (3)$$

$$\mathbf{E}_{j,n} = \mathrm{FFT}[0 \cdots 0 \; e_{j,nN+1} \cdots e_{j,(n+1)N}]^T \quad (4)$$

$$\mathbf{S}_{ij,n} = [S_{ij,n}(1) \cdots S_{ij,n}(2N)]^T \quad (5)$$

where $\mathbf{h}_{ij,n}$ is the coefficient vector of the preprocessing filter, n the block time, μ the step-size parameter, $\mathbf{S}_{ij,n}$ a complex vector whose elements are -1 or 1 in both real and imaginary parts, c_n the magnitude of the perturbation, i the secondary sound source number, and j the control point number. $\mathbf{S}_{ij,n}$ is generated so that it has the conjugate symmetry and has following characteristics:

$$\begin{aligned}
&E[\mathrm{Re}\{S_{ij,n}(k)\}^2] = 0, E[\mathrm{Im}\{S_{ij,n}(k)\}^2] = 0 \\
&\mathrm{Re}\{S_{ij,n}(k)\}^2 = 1, \mathrm{Im}\{S_{ij,n}(k)\}^2 = 1 \\
&E[\mathrm{Re}\{S_{ij,n}(k)\}\mathrm{Re}\{S_{ij,m}(k)\}] = 0, (n \neq m) \\
&E[\mathrm{Im}\{S_{ij,n}(k)\}\mathrm{Im}\{S_{ij,m}(k)\}] = 0, (n \neq m) \\
&E[\mathrm{Re}\{S_{ij,n}(k)\}\mathrm{Re}\{S_{ij,n}(l)\}] = 0, (k \neq l) \\
&E[\mathrm{Im}\{S_{ij,n}(k)\}\mathrm{Im}\{S_{ij,n}(l)\}] = 0, (k \neq l) \\
&E[\mathrm{Re}\{S_{ij,n}(k)\}\mathrm{Im}\{S_{ij,n}(l)\}] = 0 \quad (6)
\end{aligned}$$

where $E[\cdot]$ is expectation operation. Moreover, the perturbation added to the preprocessing filter is $c_n \mathbf{s}_{ij,n}$, and c_n and $\mathbf{s}_{ij,n}$ are respectively defined as follows:

$$c_n = \sqrt{\frac{\alpha \sum\limits_{j=1}^{2} \sum\limits_{k=(n-1)N+1}^{nN} e^2_{j,k}}{G^2 \sum\limits_{j=1}^{2} \sum\limits_{k=(n-1)N+1}^{nN} x^2_{j,k}}} \quad (7)$$

$$\mathbf{s}_{ij,n} = \text{first } N \text{ elements of IFFT}[\mathbf{S}_{ij,n}] \quad (8)$$

where G is the mean gain of all transfer functions and α a coefficient that defines a power ratio of the perturbation to the error signal.

2.2 Improving Convergence Speed

In the multi-channel sound reproduction system as shown in Fig. 2, it turns out that the filter coefficients converge to LNS, that is, the filter coefficients after a movement of the listener converge to LNS for the transfer functions after the movement. Figure 3 shows frequency spectra of h_{ij} before and after the movement that the listener moves 10 cm back. It can be seen from Fig. 3 that the frequency spectrum after the movement is similar to that before the movement, except for the delay. Hence, if the difference of the delay can be estimated and applied to the adaptive algorithm, the preprocessing filters could converge faster. Therefore, we also propose an improving method using the delay control filters which compensate only the delay. The procedure of the improving method is as follows:

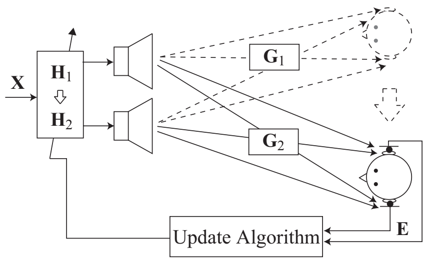


Figure 1: Proposed transoral system.

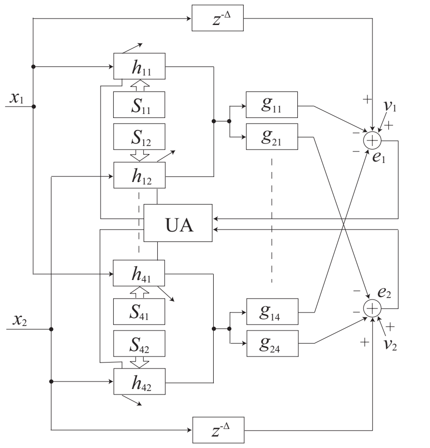


Figure 2: Transoral system using the simultaneous perturbation method.

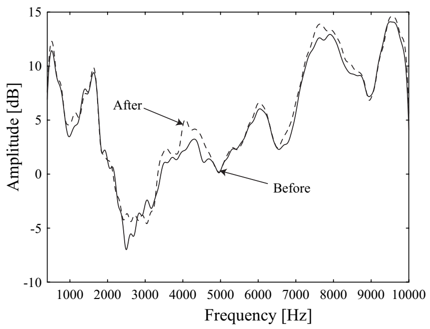


Figure 3: Frequency spectra of h_{ij} before and after variation.

1. Calculate LNS from transfer functions before the system operation.
2. Arrange the delay control filter dc in front of the preprocessing filter, as shown in Figure 4(A). At this time, the delay control filter coefficients are set to 1.0 at the tap of $k = w_1$ and 0.0 at the other taps.
3. Update the preprocessing filter coefficients until the listener moves.
4. If the movement of the listener is detected (Reproduction Accuracy RA$\leq$0), the delay control filter coefficients are updated for several seconds, as shown in Figure 4(B).
5. After updating, calculate w_2 which is the delay after the movement from the peak position of the delay control filter coefficients, except for $k = w_1$.

6. The delay control filter coefficients are set to 1.0 at the tap of $k = w_2$ and 0.0 at the other taps, as shown in Fig. 4(C).
7. Return to Procedure 3.

As mentioned above, the convergence speed is improved by compensating only the delay after the movement.

Moreover, RA in Procedure 4 is defined as follows:

$$\text{RA [dB]} = 10\log_{10}\Big(\sum_{j=1}^{2} d_{j,n}^2 \Big/ \sum_{j=1}^{2} (d_{j,n} - y_{j,n})^2\Big). \quad (9)$$

where $d_{j,n}$ is the desired signal and $y_{j,n}$ the reproduced signal.

Figure 5 shows a transoral system using the delay control filters.

3. COMPUTER SIMULATION

In this section, the validity of the proposed method is verified through computer simulations. The simulation conditions are shown in Table 1. The transfer functions measured in loudspeaker arrangement as shown in Figure 6 are used for the computer simulations. In this figure, the position (b) is 30 cm backward from the position (a), and the position (c) is rotated rightward 90 degrees at the position (a). Figure 7 shows the impulse responses and frequency spectra at each position. In the simulations, Case1 is the case where the listener moves from (a) to (b), and Case2 is the case where the listener's head rotates from (a) to (c). White noise and music which are limited within 400 - 10000 Hz are used for input signals, and the listener moves in 10 seconds.

Moreover, the detection time *Dt* of Procedure 4 of **2.2** is set to 2 seconds in the case of white noise, and 5 seconds in the case of music. Figure 8 shows the convergence properties with and without the delay control filters. In this figure, Ideal RA means the average RA attained by LNS after the movement. Table 2 shows ideal delay values and the corresponding estimated values by the delay control filters.

It can be seen from Case1 of Fig. 8 that the convergence speed can be improved by using the delay control filters. Moreover, the proposed system with the delay control filters is over Ideal RA greatly. It is also found from Fig. 7(A) that the proposed system without the delay control filters loses the information in the first half of LNS. However, the proposed system with the delay control filters does not change the delay of the preprocessing filters. Therefore, the proposed system with the delay control filters can have higher RA than the proposed system without that.

On the other hand, it can be seen from Case2 of Fig. 8 that the convergence speed cannot be improved by using the delay control filters. Moreover, it can be seen from Case2 of Table 2 that the proposed system with the delay control filters cannot detect the delay appropriately. It is also found from Fig. 7(B) that the frequency spectrum of the position (a) is similar to that of the position (b), but is not similar to that of the position (c). Therefore, the proposed system with the delay control filters cannot compensate the delay appropriately. However, the proposed sound field reproduction system using the SP method has the great advantage that it can track the movements of target points.

4. CONCLUSIONS

In this paper, we have proposed a novel sound field reproduction system using the SP method. The proposed system can track any movements of controlling points and consequently has wide controlled areas compared with the conventional one. We also have proposed an improving method of the con-

Table 1: Simulation conditions in sound reproduction system.

Sampling frequency	44100 Hz
Frequency Range	400-10000 Hz
Tap length of dc_{ij}, h_{ij}, g_{ji} (i=1,2,3,4, j=1,2)	256
Stepsize Parameter μ(White noise)	0.0001
Stepsize Parameter μ (Music)	0.00005
Perturbation Magnitude α	0.01
Delay u, w_1	128

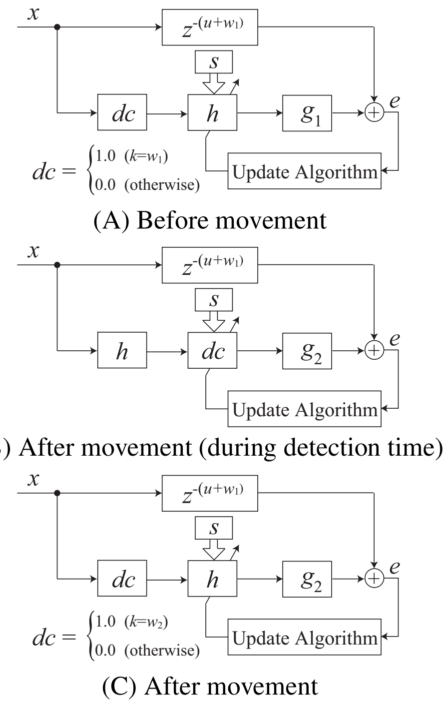


Figure 4: Block diagrams of the proposed method using the delay control filter.

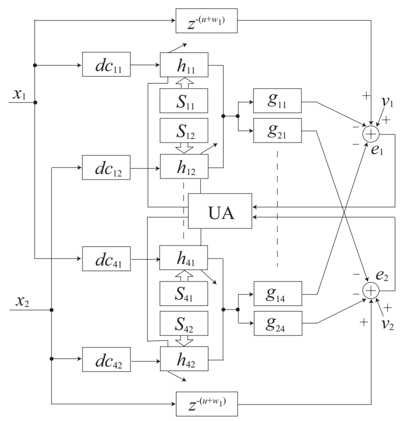


Figure 5: Transoral system using the delay control filters.

Table 2: Ditection result of delay.

Preprocess-ing filter	Case1			Case2		
	Ideal	White	Music	Ideal	White	Music
h_{11}	95	120	95	129	126	82
h_{12}	97	97	97	133	130	201
h_{21}	89	88	253	116	241	221
h_{22}	90	90	90	140	136	110
h_{31}	91	91	91	115	127	15
h_{32}	89	89	253	147	231	0
h_{41}	97	97	98	115	150	129
h_{42}	95	98	157	140	246	189

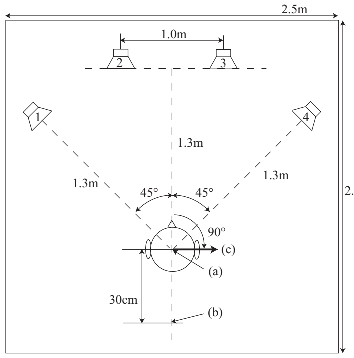
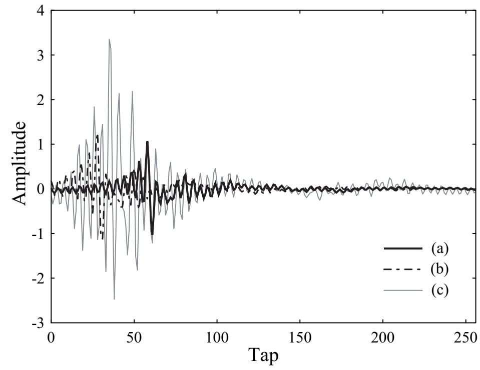
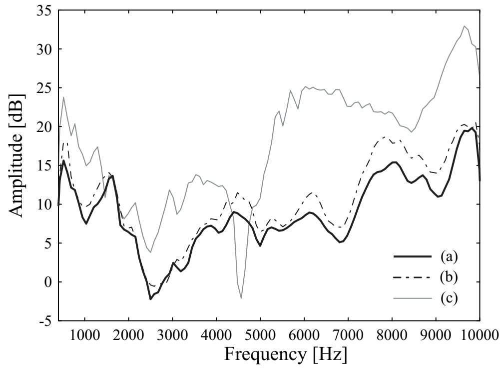


Figure 6: Loudspeaker arrangement.

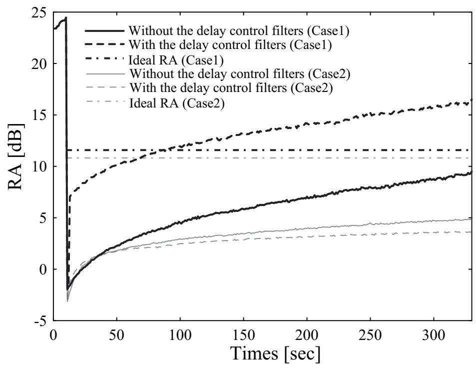


(A) Impulse responses

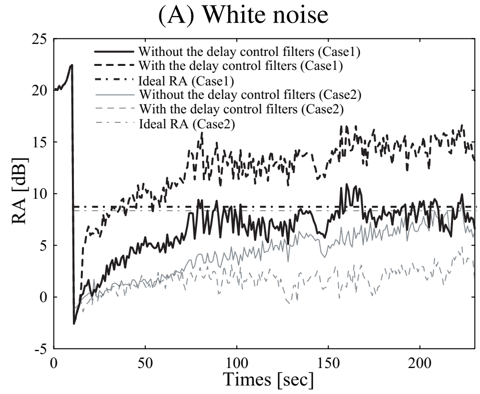


(B) Frequency spectra

Figure 7: Preprocessing filter h_{41} before and after variation.



(A) White noise



(B) Music

Figure 8: Convergence properties of RA.

vergence speed using the delay control filters. The proposed system with the delay control filters can converge faster than that without the delay detection. However, when the frequency spectrum of LNS before the movement is not similar to that after the movement, the proposed system cannot compensate the delay appropriately. We will explore loudspeaker arrangements where the proposed system can operate effectively in the future.

Acknowledgement
This research was financially supported by MEXT KAKENHI(17686018).

REFERENCES

[1] Y. A. Huang and J. Benesty, *Audio Signal Processing*. London: Kluwer Academic Publishers, 2004.

[2] M. Miyoshi and Y. Kaneda, "Inverse Filtering of Room Acoustics," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, vol. 36, no. 2, pp. 142–152, Feb. 1988.

[3] Y. Tatekura, H. Saruwatari, K. Shikano, "An Iterative Inverse Filter Design Method for the Multichannel Sound Field Reproduction System.", *IEICE Trans. on Fundamentals*, vol. E84-A, no.4, pp. 991–998, Apr. 2001.

[4] Y. Kajikawa and Y. Nomura, "Active Noise Control without a Secondary Path Model by Using a Frequency-Domain Simultaneous Perturbation Method with Variable Perturbation," in *Proc. of the 2003 IEEE International Conference on Acoustic, Speech, and Signal Processing*, Hong Kong, China, pp. V-580–V-583, Apr. 2003.