# SOUND FIELD CREATION BASED ON SIMULTANEOUS EQUATIONS METHOD

*Kensaku Fujii*[1], *Tadashi Ujino*[1], *and Mitsuji Muneyasu*[2]

[1]Division of Computer Engineering, University of Hyogo
2167 Shosha, 671-2280, Himeji, Japan
phone: +(81)-79-267-4992, fax: +(81)-79-267-4992, email: fujiken@eng.u-hyogo.ac.jp
[2]Faculty of Engineering Science, Kansai University
3-3-35 Yamate-cho, 564-8680, Suita, Japan
phone: +(81)-6-6368-0788, fax: +(81)-6-6388-8843, email: muneyasu@ipcku.kansai-u.ac.jp

## ABSTRACT

*In this paper, we propose a sound field creation system based on the simultaneous equations method used for active noise control systems. Conventional studies on such sound field creation mainly focus on the transaural system reproducing recorded signals at listening points. We differently purpose creating the same sound field as a target that. In the proposed system, inverse filters forming the sound field are derived by using the simultaneous equation method, which is characterized by auxiliary filters estimating the error between the created and the target sounds. The auxiliary filters can be also used for detecting the lack of taps of the inverse filters and the lack of delays given to the target sound field. The latter lack not only increases the error but also generates echoes preceding the created sound, which annoys listeners. The performance of the proposed system and the lack detection are verified by computer simulations.*

## 1. INTRODUCTION

Transaural system [1] is well known as a means precisely reproducing a binaural recorded sound, which can be formed by canceling cross-talks between two loudspeakers and two microphones placed at listening points. The cross-talks can be canceled by providing the recorded sound signals processed with inverse filters to the loudspeakers. Moreover, the inverse filters can be derived by utilizing an adaptive filter technique [2]. The study on the transaural system accordingly results in the issue of the derivation of the inverse filters. Various methods are hence proposed [3, 4, 5, 6].

On the other hand, the acoustic paths from the loudspeakers to the microphones are assumed to be continuously changing in practical systems. Several methods adapting to the changes, such as by compensating temperature fluctuation [7] and by applying the perturbation method or the equation error method to the derivation of the inverse filters [8, 9], are hence proposed. Each of the methods, however, has a drawback, respectively. The first method cannot deal with the change of the reverberation of the acoustic paths, the second generates an extra noise, and the third involves a bias error originating in its configuration.

In this paper, we propose applying the simultaneous equations method [10] to the derivation of the inverse filters. The configuration of active noise control systems is almost the same as that of the transaural system. Delays, inverse filters and acoustic paths in the transaural system are corresponding to primary paths, noise control filters and secondary paths in the active noise control system, respectively.

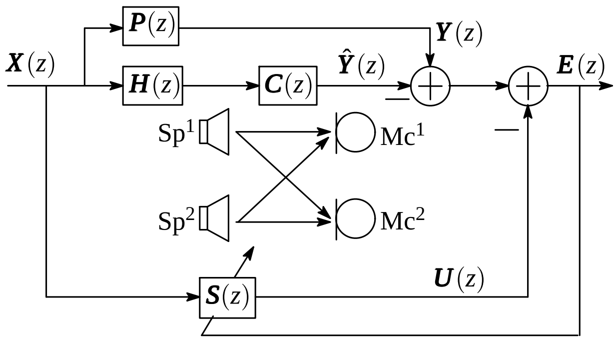



Figure 1: configuration of proposed system.

This means that the inverse filters can be derived with the same procedure as that in the active noise control system.

The simultaneous equations method is characterized by auxiliary filters used for estimating the performance of the active noise control system. When the simultaneous equations method is applied to the transaural system, the delays can be replaced with desired impulse responses. This means that the proposed method can create arbitrary sound fields. In addition, the coefficients of the auxiliary filters inform the lacks of the delays and the taps of the inverse filters. By using computer simulations, we verify that both the lacks can be detected by the coefficients of the auxiliary filters and that the proposed method can provide arbitrary sound fields.

## 2. PRINCIPLE

Figure 1 shows the configuration of the two channel sound field creation system proposed in this paper. The proposed system consists of three filter sets, $\boldsymbol{P}(z)$, $\boldsymbol{H}(z)$, $\boldsymbol{S}(z)$, and an acoustic path set, $\boldsymbol{C}(z)$. The recorded signal pair,

$$\boldsymbol{X}(z) = \begin{bmatrix} X^1(z) & X^2(z) \end{bmatrix}^T, \tag{1}$$

is provided to the inverse filter set,

$$\boldsymbol{H}(z) = \begin{bmatrix} H^{11}(z) & H^{12}(z) \\ H^{21}(z) & H^{22}(z) \end{bmatrix}, \tag{2}$$

the target sound field filter set,

$$\boldsymbol{P}(z) = \begin{bmatrix} P^{11}(z) & P^{12}(z) \\ P^{21}(z) & P^{22}(z) \end{bmatrix}, \tag{3}$$

and the auxiliary filter set,

$$\boldsymbol{S}(z) = \begin{bmatrix} S^{11}(z) & S^{12}(z) \\ S^{21}(z) & S^{22}(z) \end{bmatrix}. \tag{4}$$

The inverse filter set thereby creates a sound pair,

$$\hat{\boldsymbol{Y}}(z) = \boldsymbol{C}^T(z)\boldsymbol{H}^T(z)\boldsymbol{X}(z), \tag{5}$$

through the acoustic path set,

$$\boldsymbol{C}(z) = \begin{bmatrix} C^{11}(z) & C^{12}(z) \\ C^{21}(z) & C^{22}(z) \end{bmatrix}, \tag{6}$$

from two loudspeakers $\mathrm{Sp}^1$ and $\mathrm{Sp}^2$ to two microphones $\mathrm{Mc}^1$ and $\mathrm{Mc}^2$ placed at listening points. At the same time, the target sound field filter set yields a target sound pair,

$$\boldsymbol{Y}(z) = \boldsymbol{P}^T(z)\boldsymbol{X}(z). \tag{7}$$

The auxiliary filter set is used for deriving the optimum inverse filter set, $\boldsymbol{H}_{opt}(z)$, satisfying the relation,

$$\hat{\boldsymbol{Y}}(z) = \boldsymbol{Y}(z), \tag{8}$$

which is equal to

$$\boldsymbol{P}(z) = \boldsymbol{H}_{opt}(z)\boldsymbol{C}(z). \tag{9}$$

This relation means that the optimum inverse filter set is derived from

$$\boldsymbol{H}_{opt}(z) = \boldsymbol{P}(z)\boldsymbol{C}^{-1}(z). \tag{10}$$

To derive the optimum inverse filter set, the simultaneous equations method [10] first identifies the estimation error between the created sound pair, $\hat{\boldsymbol{Y}}(z)$, and the target sound pair, $\boldsymbol{Y}(z)$, by using the auxiliary filter set, $\boldsymbol{S}(z)$. When a tentative inverse filter set, $\hat{\boldsymbol{H}}(z)$, is given to the proposed system, the auxiliary filter set yields the following relation,

$$\boldsymbol{S}(z) = \boldsymbol{P}(z) - \hat{\boldsymbol{H}}(z)\boldsymbol{C}(z). \tag{11}$$

after the identification error,

$$\boldsymbol{E}(z) = \left\{\boldsymbol{Y}(z) - \hat{\boldsymbol{Y}}(z)\right\} - \boldsymbol{U}(z), \tag{12}$$

has been minimized, where

$$\boldsymbol{U}(z) = \begin{bmatrix} U^1(z) & U^2(z) \end{bmatrix}^T \tag{13}$$

is the output pair provided by the auxiliary filter set.

In active noise control systems, $\boldsymbol{P}(z)$ and $\boldsymbol{C}(z)$ are unknown. The simultaneous equations method accordingly gives two different inverse filter sets to the active noise control system at a proper interval, and then derives $\boldsymbol{H}_{opt}(z)$ by using the estimated two auxiliary filter sets. However, $\boldsymbol{P}(z)$ is known in the proposed system, and only $\boldsymbol{C}(z)$ is unknown. In this case, the only unknown, $\boldsymbol{C}(z)$, can be directly derived from (11) as follows:

$$\boldsymbol{C}(z) = \hat{\boldsymbol{H}}^{-1}(z)\left\{\boldsymbol{P}(z) - \boldsymbol{S}(z)\right\}, \tag{14}$$

which can be rewritten to

$$\boldsymbol{C}^{-1}(z) = (\boldsymbol{P}(z) - \boldsymbol{S}(z))^{-1}\hat{\boldsymbol{H}}(z). \tag{15}$$

Substituting $\boldsymbol{C}^{-1}(z)$ into (10) finally yields

$$\boldsymbol{H}_{opt}(z) = \boldsymbol{P}(z)\boldsymbol{C}^{-1}(z) = \boldsymbol{P}(z)(\boldsymbol{P}(z) - \boldsymbol{S}(z))^{-1}\hat{\boldsymbol{H}}(z). \tag{16}$$

Incidentally, giving the simple delay set,

$$\boldsymbol{P}(z) = \begin{bmatrix} z^{-D} & z^{-D} \\ z^{-D} & z^{-D} \end{bmatrix} \tag{17}$$

to the proposed system forms a transaural system.

## 3. TRANSFORM TO FILTER COEFFICIENT

$\boldsymbol{H}_{opt}(z)$ shown by (16) gives a transfer function set expressed with $z$-transform. For applying the above principle to practical systems, transforming the transfer function set to a coefficient set is required. In this paper, we apply a frequency domain technique used in [10] to the transformation. In the frequency domain, (16) is rewritten to

$$\boldsymbol{H}_{opt} = \boldsymbol{P}(\boldsymbol{P} - \boldsymbol{S})^{-1}\hat{\boldsymbol{H}}, \tag{18}$$

where the elements of

$$\boldsymbol{P} = \begin{bmatrix} \boldsymbol{P}^{11} & \boldsymbol{P}^{12} \\ \boldsymbol{P}^{21} & \boldsymbol{P}^{22} \end{bmatrix}, \tag{19}$$

$$\boldsymbol{S} = \begin{bmatrix} \boldsymbol{S}^{11} & \boldsymbol{S}^{12} \\ \boldsymbol{S}^{21} & \boldsymbol{S}^{22} \end{bmatrix} \tag{20}$$

and

$$\hat{\boldsymbol{H}} = \begin{bmatrix} \hat{\boldsymbol{H}}^{11} & \hat{\boldsymbol{H}}^{12} \\ \hat{\boldsymbol{H}}^{21} & \hat{\boldsymbol{H}}^{22} \end{bmatrix} \tag{21}$$

are discrete Fourier transforms (DFTs) of the coefficient vectors given to the target sound field, the auxiliary and the inverse filters, respectively.

Here, it should be noted that DFT is applied to the modified coefficient vectors,

$$\tilde{\boldsymbol{p}}^{mn} = \begin{bmatrix} p^{mn}(0) & \cdots & p^{mn}(I-1) & \underbrace{0 \ \cdots \ 0}_{I} \end{bmatrix}^T, \tag{22}$$

$$\tilde{\boldsymbol{s}}^{mn} = \begin{bmatrix} s^{mn}(0) & \cdots & s^{mn}(I-1) & \underbrace{0 \ \cdots \ 0}_{I} \end{bmatrix}^T \tag{23}$$

and

$$\tilde{\boldsymbol{h}}^{mn} = \begin{bmatrix} \hat{h}^{mn}(0) & \cdots & \hat{h}^{mn}(J-1) & \underbrace{0 \ \cdots \ 0}_{2I-J} \end{bmatrix}^T, \tag{24}$$

where $m = 1,2$ and $n = 1,2$ denote the column and the row numbers of the filter sets. Circular convolutions provided by DFT are thereby converted to linear convolutions. Practically, the coefficient vectors,

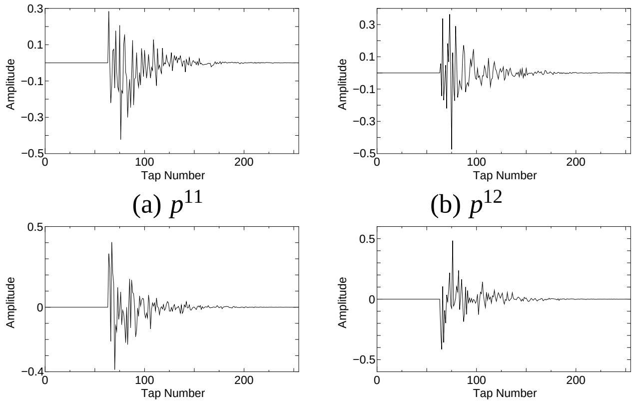
$$\boldsymbol{p}^{mn} = \begin{bmatrix} p^{mn}(0) & \cdots & p^{mn}(I-1) \end{bmatrix}^T, \tag{25}$$

$$\boldsymbol{s}^{mn} = \begin{bmatrix} s^{mn}(0) & \cdots & s^{mn}(I-1) \end{bmatrix}^T \tag{26}$$

and

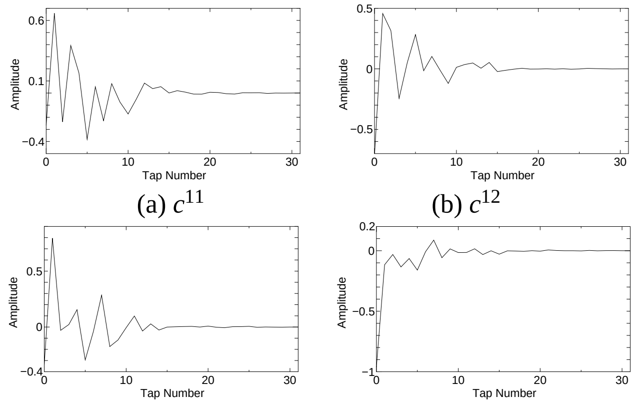
$$\hat{\boldsymbol{h}}^{mn} = \begin{bmatrix} \hat{h}^{mn}(0) & \cdots & \hat{h}^{mn}(J-1) \end{bmatrix}^T \tag{27}$$

are given to the filters, where $I$ and $J$ are the number of taps of the filters. The coefficients of the optimum inverse filters can be derived by extracting the former $J$ components of each element obtained by applying inverse discrete Fourier transform (IDFT) to $\boldsymbol{H}_{opt}$




(a) $p^{11}$ (b) $p^{12}$ (c) $p^{21}$ (d) $p^{22}$

Figure 2: Coefficients of target sound field filters




(a) $c^{11}$ (b) $c^{12}$ (c) $c^{21}$ (d) $c^{22}$

Figure 3: Impulse responses of acoustic paths

## 4. UPDATING PROCEDURE

In practical use, the acoustic paths are assumed to be continuously changing. The proposed system is accordingly designed so as to estimate $\boldsymbol{H}_{opt}(z)$ repeatedly. The current coefficient vectors of the optimum inverse filters can be repeatedly estimated by executing the following procedure:

(1) Give a coefficient vector set, for example,

$$\begin{cases} \hat{\boldsymbol{h}}^{11} = [\ a \quad 0 \quad \cdots \quad 0\ ]^T \\ \hat{\boldsymbol{h}}^{12} = [\ 0 \quad 0 \quad \cdots \quad 0\ ]^T \\ \hat{\boldsymbol{h}}^{21} = [\ 0 \quad 0 \quad \cdots \quad 0\ ]^T \\ \hat{\boldsymbol{h}}^{22} = [\ b \quad 0 \quad \cdots \quad 0\ ]^T \end{cases} \tag{28}$$

to the inverse filter set, where $a$ and $b$ are arbitrary constant.

(2) Calculate the DFT of $\tilde{\boldsymbol{h}}^{mn}$.

(3) Estimate $\boldsymbol{s}^{mn}$.

(4) Calculate the DFT of $\tilde{\boldsymbol{s}}^{mn}$.

(5) Calculate $\boldsymbol{H}_{opt}$ by using (18).

(6) Extract the former $J$ elements of IDFT of $\boldsymbol{H}_{opt}$

(7) Give the extracted $\hat{\boldsymbol{h}}^{mn}$ to the inverse filter set.

(8) Go back to (2).

By repeating the estimation process from (2) to (7), the proposed system can automatically create the sound pair approximated to the target sound pair, even if the acoustic paths change. On the other hand, during the acoustic paths are invariant, $\boldsymbol{S}(z) = 0$ is detected by the auxiliary filter set. The inverse filter set is thereby held optimal.

## 5. VERIFICATION BY COMPUTER SIMULATION

This paper next verifies the performance of the proposed system by using computer simulations. Figures 2 and 3 show the coefficients of the target sound field filters and the impulse response samples of the acoustic paths, where they are given by exponential decay normal random numbers. Naturally, the sound creation path set,

$$\hat{\boldsymbol{P}} = \hat{\boldsymbol{H}}\boldsymbol{C}, \tag{29}$$

generates some group delays. In this simulation, the delays are estimated to be 64 sample times.

In the proposed system, various adaptive algorithms can be used for identifying the error between the target and the created sounds. This paper uses a well-known algorithm,

$$\boldsymbol{s}_{j+1}^{mn} = \boldsymbol{s}_{j+1}^{mn} + \mu \frac{e_j^n \boldsymbol{x}_j^m}{\sum_{i=1}^{2} \boldsymbol{x}_j^{m^T} \boldsymbol{x}_j^m}, \tag{30}$$

for the identification, where $\boldsymbol{s}_j^{mn}$ is an auxiliary filter coefficient vector given at $j$ sample time, $\boldsymbol{x}_j^m$ is a recorded signal vector, $\mu$ is a constant called step size, and

$$e_j^n = \left\{ y_j^n - \hat{y}_j^n \right\} - u_j^n \tag{31}$$

is the difference between the output of the auxiliary filter, $u_j^n$, and the estimation error, $\left\{ y_j^n - \hat{y}_j^n \right\}$. Here, $\hat{y}_j^n$ is the created sound signal detected by microphone Mc$^n$, $y_j^n$ is the desired signal yielded by the target sound field filters. By using the above algorithm, the coefficients of the auxiliary filters are updated $2^{18} = 262,144$ times, and then the coefficients of the optimum inverse filters are calculated by applying IDFT to (18).

For simplifying the verification, we use

$$x_j^1 = r_j^0 + r_j^1 \tag{32}$$

and

$$x_j^2 = r_j^0 + r_j^2 \tag{33}$$

as the recorded signals, where $r_j^0$, $r_j^1$ and $r_j^2$ are white noises. In practical systems, the recorded signals are strongly correlated between each other. The degree of the cross-correlation can be estimated with the power ratio of $r_j^0$ to $r_j^m$. In this simulation, we assume the power ratio to be 20 dB. In addition, the other conditions are

(1) Number of taps of inverse filters; $J = 224$,

(2) Number of samples of acoustic paths; $K = 32$,

(3) Number of taps of auxiliary filters; $I = 256$,

(4) Number of samples of desired paths; $I = 256$,

(5) Duration of DFT; $2I = 512$,

(6) Power ratio of recorded signal to environmental noise; –40 dB,

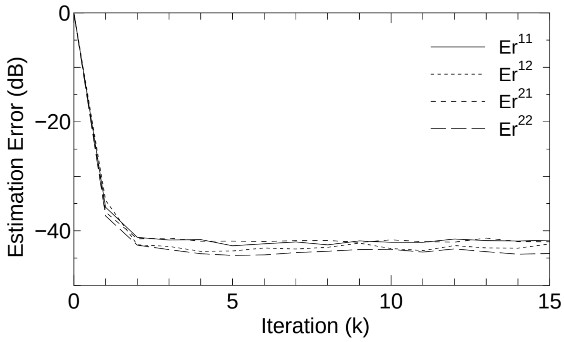



Figure 4: Transition of estimation errors where number of taps is 224 and sample time of inserted delays is 64.

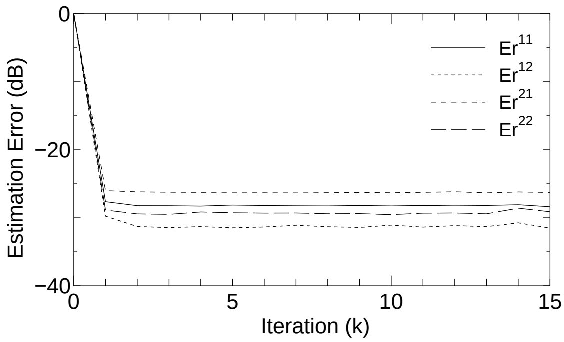



Figure 5: Transition of estimation errors where number of taps is 224 and sample time of inserted delays is 32.

(7) Step size; $\mu = 1.0$.

Figure 4 shows the transition of the estimation errors calculated using

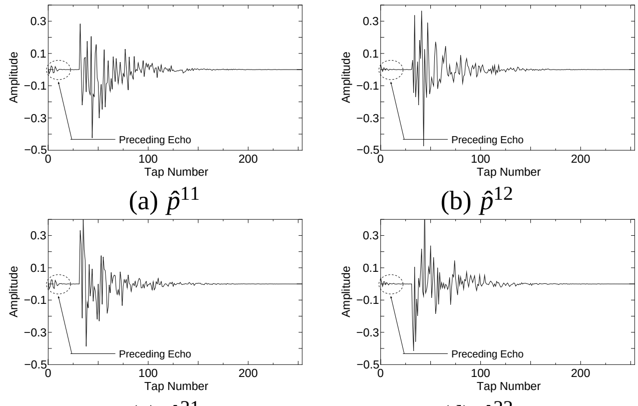
$$E_k^{mn} = 10\log_{10} \frac{\sum_{m=1}^{2}\sum_{n=1}^{2}\sum_{i=0}^{I-1}\{p^{mn}(i) - \hat{p}_k^{mn}(i)\}^2}{\sum_{m=1}^{2}\sum_{n=1}^{2}\sum_{i=0}^{I-1}\{p^{mn}(i)\}^2}, \qquad (34)$$

where, $\hat{p}_k^{mn}(i)$ is the $i$th impulse response sample of $mn$ element of the sound creation path set obtained by repeating the estimation process $k$ times. The estimation error, $E_k^{mn}$, presents the difference between the target and the created sounds. This result shows that the proposed system can create the sound approximated to the target sound.

## 6. DELAY AND NUMBER OF TAPS

The sound creation paths naturally involve some group delays, which requires inserting sufficient delays in the target sound path filters. The lack of the inserted delays degrades the performance of the sound creation system. Fig. 5 shows the transitions of the estimation errors calculated inserting the delay of 32 sample times in the target sound field filters, where the other simulation conditions are the same as those in Fig. 4. In comparison with the results shown in Fig. 4, the estimation errors increase by more than 10 dB.

In addition, Fig. 6 shows the impulse responses of the sound creation paths calculated by reducing the delays to 32 sample times. In Fig. 6, we can see some echoes preceding the substance of the created sound, which annoys listeners.




(a) $\hat{p}^{11}$ (b) $\hat{p}^{12}$ (c) $\hat{p}^{21}$ (d) $\hat{p}^{22}$

Figure 6: Impulse responses of sound creation paths where number of taps is 224 and sample time of inserted delays is 32.

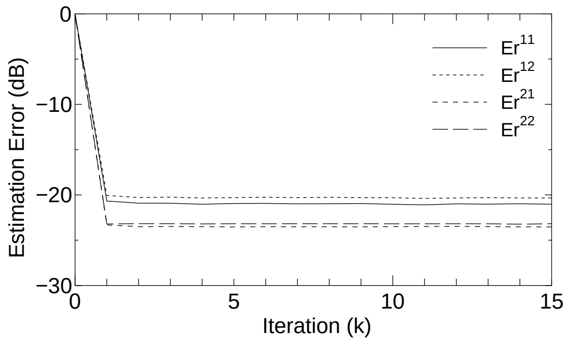
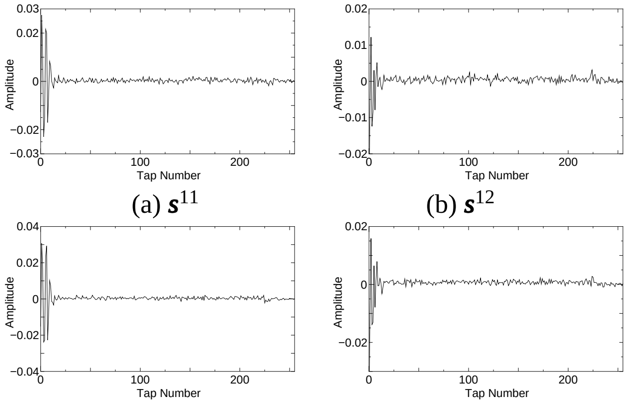



Figure 7: Transition of estimation errors where number of taps is 128 and sample time of inserted delays is 64.

The lack of taps given to the inverse filters also increases the difference between the target and the created sounds. Figure 7 shows the the transitions of the estimation errors calculated reducing the number of the taps of the inverse filters from 224 to 128. In this example, the decreases of the estimation errors stop at about –20 dB, which means the degradation of 20 dB in comparison with the case of 224 taps.

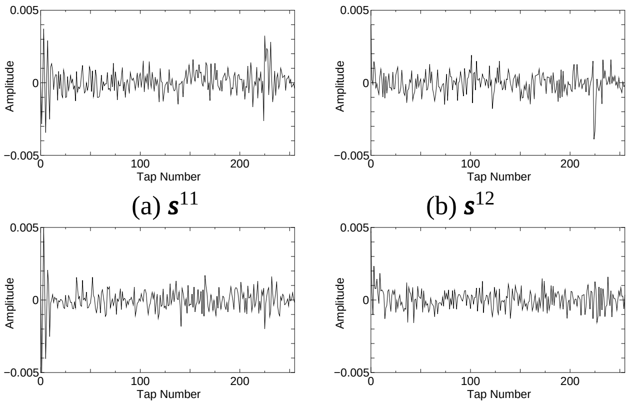
Both the lacks thus increases the estimation errors. For reducing the estimation errors, sound field creation systems are required detecting which of the lacks causes the degradation and then increasing properly the delays or the taps. In the proposed system, the coefficients of the auxiliary filters inform which of them is insufficient. Figures 8 and 9 show the coefficients of the auxiliary filters obtained when the inserted delays are 32 and 64 sample times, respectively. Clearly, the former part of the coefficients shown in Fig. 8 is about ten times of those in the case of 64 sample time delays, which indicates the generation of the preceding echoes.

The lack of the taps given to the inverse filters can be also detected by the coefficients of the auxiliary filters. Figure 10 shows the coefficients of the auxiliary filters calculated reducing the number of the taps from 224 to 128. Apparently, the latter part of the coefficients of the auxiliary filters is more than those in case of 224 taps, which shows the lack



(a) $\boldsymbol{s}^{11}$ (b) $\boldsymbol{s}^{12}$ (c) $\boldsymbol{s}^{21}$ (d) $\boldsymbol{s}^{22}$

Figure 8: Coefficients of auxiliary filters where number of taps is 224 and sample time of inserted delays is 32.



(a) $\boldsymbol{s}^{11}$ (b) $\boldsymbol{s}^{12}$ (c) $\boldsymbol{s}^{21}$ (d) $\boldsymbol{s}^{22}$

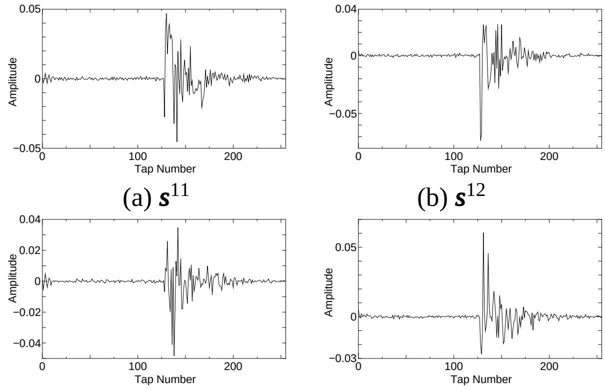
Figure 9: Coefficients of auxiliary filters where number of taps is 224 and sample time of inserted delays is 64.

of the taps. The coefficients of the auxiliary filters thus inform the difference in the impulse responses of the sound creation paths and the target sound fields. In the proposed system, the amount of the inserted delays and the number of the taps can be adjusted by monitoring the coefficients of the auxiliary filters.

## 7. CONCLUSION

In this paper, we have proposed the sound field creation system based on the simultaneous equations method and verified the performance of the proposed system by using computer simulations. In addition, we have shown that the lack of the delays to be inserted in the target sound field filters and the lack of the taps given to the inverse filters can be detected by monitoring the coefficients of the auxiliary filters. This means that the proposed method can adjust the amount of the inserted delays and the number of the taps.

In the future works, we will apply the proposed principle to experimental systems, and verify that the proposed system practically works.



(a) $\boldsymbol{s}^{11}$ (b) $\boldsymbol{s}^{12}$ (c) $\boldsymbol{s}^{21}$ (d) $\boldsymbol{s}^{22}$

Figure 10: Coefficients of auxiliary filters where number of taps is 128 and sample time of inserted delays is 64.

## REFERENCES


[1] J. Bauck and D. H. Cooper,“Generalized transaural stereo and applications,” *J. Audio Eng. Soc.*, vol. 44, pp. 683-705, Sept. 1996.

[2] P. A. Nelson, H. Hamada and S. J. Elliott,“Inverse filters for Multi-channel sound reproduction,” *IEICE Trans. Fundamentals*, vol. E75-A, pp. 1468-1473, Nov. 1992.

[3] T. Tokuno, O. Kirkeby, P. A. Nelson and H. Hamada,“Inverse filter of sound reproduction system using regularization,” *IEICE Trans. Fundamentals*, vol. E80-A, pp. 809-820, May 1997.

[4] H. Nakajima, M. Miyoshi and M. Tohyama,“Sound field control by indefinite MINT filters,” *IEICE Trans. Fundamentals*, vol. E80-A, pp. 821-824, May 1997.

[5] A. Kaminuma, S. Ise, and K. Shikano, “A method of an inverse filter design for multi-channel sound reproduction system using minimum-norm-solution in frequency domain,” *J. Acoust. Soc. Jpn*, vol. 57, pp. 175-183, March 2001.

[6] Y. Furumi, H. Hokari, and S. Shimada, “A study on loudspeaker arrangement in multi-channel transaural system for sound image reproduction,” *IEICE Trans. Fundamentals*, vol. J83-A, pp. 920-923, July 2000.

[7] Y. Takekura, H. Saruwatari, and K. Shikano,“Sound reproduction system including adaptive compensation of temperature fluctuation effect for broad-band sound control,” *IEICE Trans. Fundamentals*, vol. E85-A, pp. 1851-1860, Aug. 2002.

[8] T. Tazaki, K. Tukamoto, Y. Kajikawa, and Y. Nomura, “A study on sound reproduction system using the perturbation method,” *IEICE Tech. Rept.*, EA2005-8, pp. 13-18, May 2005.

[9] T. Horiuchi, H. Hokari, S. Shimada, and T. Inada, “Adaptive estimation of transfer function for sound localization using stereo earphone-microphone combination,” *IEICE Trans. Fundamentals*, vol. E85-A, pp. 1841-1850, Aug. 2002.

[10] K. Fujii, M. Muneyasu, and Y. Fujita, “Frequency domain simultaneous equations method for active noise control systems,” in *Proc. EUSIPCO 2006*, Florence, Italy, September 4-8. 2006.