Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Journal of Sound and Vibration 278 (2004) 1196–1204
JOURNAL OF SOUND AND VIBRATION
www.elsevier.com/locate/jsvi

Letters to the Editor

A direct design method of inverse filters for multichannel 3D sound rendering

Sunmin Kim, Youngjin Park*

Department of Mechanical Engineering, Center for Noise and Vibration Control (NoViC), Korea Advanced Institute of Science and Technology (KAIST), Science Town, Taejon 305-701, South Korea

Received 7 March 2003; accepted 12 December 2003

1. Introduction

Multichannel sound reproduction has received a great deal of attention as a dual concept of active noise control since early 1990s. While multichannel 3-D sound system requires complex hardware, it has better performance than the binaural sound system [1]. An inverse filter matrix of acoustic transfer function matrix from multiple loudspeakers to control points surrounding the ears of a listener is necessary to reproduce virtual sound field around the listener.

There are two approaches designing the inverse filters of 3-D sound system. One is the design method in the frequency domain, with which it is easy to obtain the optimal solution and to analyze the solution theoretically, even though the stability of the 3-D sound rendering system may not be guaranteed [2]. The other is the time domain approach, which guarantees the stable and causal solution by approximating the inverse filter in the time domain, while it has complex structure and heavy computational burden [2]. The adaptive technique commonly used in the time domain approach requires the tuning parameters such as a modelling delay and the order of the inverse filter known a priori. The mean squared error of the 3-D sound system is dependent on both the choice of the modelling delay and the number of coefficients in the inverse filters [3]. Lots of time and efforts are required to select the proper tuning parameters. Several methods with low computational requirement and fast convergence speed have been proposed [4,5]. Mourjopoulos [6] proposed the direct inversion method for equalization of room acoustics or loudspeaker. The method is for single-input–single-output system and the inverse filter length becomes very large when the system has many unstable zeros.

In this paper, a direct design method of the inverse filter is proposed for the multichannel 3-D sound rendering. The inverse filter of the system with many unstable zeros has the same length as that of the system with one unstable zero, which is directly designed for multiple-input–multiple-

*Corresponding author. Tel.: +82-42-869-3036; fax: +82-42-869-8220.
E-mail addresses: sunmin21.kim@samsung.com (S. Kim), yjpark@mail.kaist.ac.kr (Y. Park).

0022-460X/$ - see front matter
doi:10.1016/j.jsv.2003.12.014

output system. The tuning parameters can be systematically determined and the inverse filter matrix can be directly obtained with the desired accuracy and low computational burden by using the proposed method.

2. Multichannel sound reproduction

2.1. Inverse problem of multichannel sound system

Multichannel sound reproduction system consisting of multiple loudspeakers generates the desired sounds at multiple control points by determining outputs of loudspeakers based on the optimally designed control filter as shown in Fig. 1.

Fig. 2 shows the block diagram of the multichannel sound reproduction system. In Figs. 2 and 3, x is a sound of virtual source and $\mathbf{d}$ is a desired sound vector, which we want to reproduce at the control points. An error vector $\mathbf{e}$ is defined as the difference between the desired sound vector $\mathbf{d}$ and the reproduced sound vector $\mathbf{y}$ at the control points. $\mathbf{u}$ is an output vector of loudspeakers.

$\mathbf{P}$ represents the desired acoustic transfer function matrix from the virtual sound source to the control points and $\mathbf{C}$ represents the acoustic transfer function matrix from loudspeakers to the control points. The control filter matrix $\mathbf{G}$ is designed to minimize the squared errors. If $\mathbf{G}$ can be expressed as a product of matrices $\mathbf{W}$ and the $\mathbf{P}$, a problem to design the matrix $\mathbf{W}$ becomes an inverse problem of $\mathbf{C}$ exactly when $\mathbf{A}$ is an identity matrix as shown in Fig. 3. Actually, $\mathbf{A}$ is the diagonal matrix consisting of pure delay terms which is necessary to guarantee the causality of the optimal solution because $\mathbf{C}$ has a physical time delay. Therefore, $\mathbf{W}$ becomes the optimal solution when the $\mathbf{W}$ is the delayed inverse filter matrix of $\mathbf{C}$.

2.2. Optimal solution in frequency domain

The optimal solution of the multichannel system derived in the frequency domain [5] is introduced. The cost function is defined as the squared residual errors

$$J(\omega) = \mathbf{e}^{\mathrm{H}}(\omega)\mathbf{e}(\omega). \tag{1}$$

In overdetermined case [2], where the number of loudspeakers is smaller than that of control points, the optimal control input $\mathbf{u}_{opt}(\omega)$ is

$$\mathbf{u}_{opt}(\omega) = (\mathbf{C}^{\mathrm{H}}(\omega)\mathbf{C}(\omega))^{-1}\,\mathbf{C}^{\mathrm{H}}(\omega)\mathbf{A}(\omega)\,\mathbf{d}(\omega). \tag{2}$$

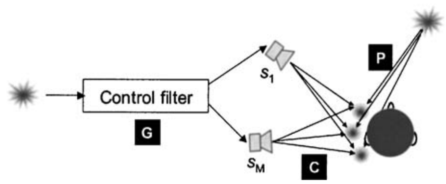

Fig. 1. Multichannel sound reproduction system.

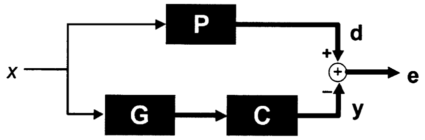

Fig. 2. Block diagram of multichannel sound reproduction system.

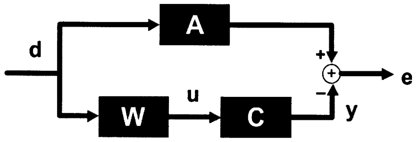

Fig. 3. Inverse problem structure.

In underdetermined and fully determined cases, the optimal solutions can be derived in the similar manners [8]. However, it may not guarantee the stability and causality of 3-D sound system [2].

2.3. Approximate solution in time domain

The causal and stable inverse filter can be obtained from the time domain solution. It is first assumed that the matrix **W** consists of FIR filters of order L. Define a composite vector **w** that is made up of each element of **W**. The length of **w** is $L \times M \times K$ where M and K are the numbers of the control points and the loudspeakers, respectively. To obtain the optimal solution of **w**, matrix inversion should be conducted. One way to determine the coefficients of the matrix is obviously by direct inversion of the matrix. However, this matrix is clearly of higher order, being of dimension $L \times M \times K$. Another approach is to use the modified LMS algorithm, developed for active noise control applications by Elliot et al. [7]

Because the proper order of the FIR filter and the proper time delay of the diagonal matrix **A** should be determined to apply the adaptive algorithm successfully, much time and effort are exhausted especially for large-scale MIMO systems. Here, an effective design method of the inverse filter is proposed which does not need to tune the parameters such as the order of the FIR filter and the modelling delay of the matrix **A**.

3. Direct design of inverse filter

3.1. Inversion of MIMO system

In overdetermined case, the least-squares solution of the inverse filter in the z-domain is as follows:

$$\mathbf{W}(z)_{opt} = (\mathbf{C}^{\mathrm{H}}(z)\mathbf{C}(z))^{-1}\mathbf{C}^{\mathrm{H}}(z)\mathbf{A}(z) = \frac{adj[\mathbf{C}^{H}(z)\mathbf{C}(z)]}{|\mathbf{C}^{H}(z)\mathbf{C}(z)|}\mathbf{C}^{H}(z)\mathbf{A}(z), \tag{3}$$

where z is the z transform variable. The stability of the filter is determined by the denominator polynomial in Eq. (3). Zeros of this polynomial (the poles of the system) always occur in groups of four which are at conjugate reciprocal positions, that is, the denominator polynomial has zeros which are arranged in pairs, with each zero inside the unit circle in the complex z-plane being associated with a zero outside the unit circle. Therefore, any system designed in the frequency domain will not yield a stable system in the time domain [2]. In underdetermined case, the inverse system has poles in groups of four and is always unstable. In fully determined case, the stability of the inverse system depends upon the geometrical arrangement of the loudspeakers and the control points. In general, the determinant in Eq. (3) is a non-minimum phase system and the method designing the stable inversion of the non-minimum phase system is proposed in the next section.

3.2. Separation of non-minimum phase system

The denominator polynomial with time delay in Eq. (3) is defined as $\mathbf{H}(z)$ and is separated into three parts as follows:

$$\mathbf{H}(z) \equiv |\mathbf{C}(z)^{\mathrm{H}}\mathbf{C}(z)| = h_0 + h_1 z^{-1} + \cdots + h_{KL} z^{-KL} = z^{-\Delta}\mathbf{H}_{min}(z)\mathbf{H}_{max}(z), \tag{4}$$

$$\mathbf{H}_{max}(z) = (1 - b_1 z^{-1})\cdots(1 - b_R z^{-1}), \quad |b_i| > 1, \;\; i = 1, \ldots, R, \tag{5}$$

$$\begin{aligned} &\mathbf{H}_{min}(z) = h_\Delta(1 - b_{R+1} z^{-1})\cdots(1 - b_{KL-\Delta} z^{-1}), \\ &|b_j| < 1, \;\; j = (R+1), \ldots, (KL - \Delta), \end{aligned} \tag{6}$$

where K is the number of the loudspeakers and L is the length of FIR filters composing $\mathbf{C}(z)$. Eqs. (5) and (6) represent the maximum phase and the minimum phase subsystems respectively. R is the order of the maximum phase subsystem and Δ is the amount of the time delay.

3.3. Direct stable inversion of maximum phase system

Note that an inverse of the maximum phase system makes the system unstable. The proposed method is focused on how to derive the stable inverse filter of the maximum phase system. An inverse of the maximum phase system with the form of FIR filter becomes an unstable IIR filter. The unstable IIR filter can be approximated to a stable and anti-causal FIR filter. Briefly speaking, the causality of the system is intentionally abandoned in order to guarantee the stability. In the appendix, the stable inversion of the maximum phase system of order 1 is presented in order to help understand the basic idea of the proposed approach. Now, consider the maximum phase system of order R

$$\mathbf{H}_{max}(z) = \prod_{i=1}^{R} (1 - b_i z^{-1}), \quad |b_i| > 1, \tag{7}$$

where $\prod$ is a multiplication operator and the inverse of $\mathbf{H}_{max}(z)$ is

$$\mathbf{H}_{max}(z)^{-1} = \prod_{i=1}^{R} \frac{1}{1 - b_i z^{-1}}, \quad |b_i| > 1. \tag{8}$$

The unstable IIR filter may be approximated to a stable FIR filter of order $R\delta$ according to the appendix. In order to reduce the length of the inverse filter, partial fraction expansion and residue theorem are used as follows:

$$\mathbf{H}_{max}(z)^{-1} = \sum_{i=1}^{R} \frac{c_i}{1 - b_i z^{-1}}, \tag{9}$$

$$c_i = (1 - b_i z^{-1})\mathbf{H}_{max}(z)^{-1}\big|_{z=b_i}, \quad i = 1, 2, \ldots, R. \tag{10}$$

If $\mathbf{H}_{max}(z)^{-1}$ is defined as $x(z)$ over $y(z)$ and the result derived in the appendix is used, the following difference equation can be derived:

$$x(n) = -\sum_{i=1}^{R} \sum_{k=1}^{\delta} \frac{c_i}{b_i^k} y(n+k) + \sum_{i=1}^{R} \frac{1}{b_i^\delta} x(n+\delta). \tag{11}$$

If δ is sufficiently large and the second term of right-hand side of Eq. (11) can be ignored, $x(n)$ is expressed only with $y(n+k)$ as follows:

$$x(n) \cong -\sum_{i=1}^{R} \sum_{k=1}^{\delta} \frac{c_i}{b_i^k} y(n+k). \tag{12}$$

Hence, the inverse of the maximum phase system can be approximated to the stable and anti-causal FIR filter as

$$\mathbf{H}_{max}(z)^{-1} \cong -z^{\delta} \sum_{i=1}^{R} \sum_{k=1}^{\delta} \frac{c_i}{b_i^k} z^{-(\delta-k)}, \tag{13}$$

where δ can be selected using the following equation according to the desired accuracy (*ERROR BOUND*) of the inversion:

$$\left|\frac{1}{b_i}\right|^{\delta} < ERROR\ BOUND, \quad i = 1, \ldots, R. \tag{14}$$

As shown in Eq. (13), the inverse filter has the length of order δ which is the same as the inverse filter length of the maximum phase subsystem with one unstable zero as in the appendix and requires prediction of δ step. Large δ makes the inversion accurate, while the length of the inverse filter increases. Because the future data cannot be generally obtained, the delayed desired sound is reproduced. The delay results from the physical delay which acoustic transfer functions naturally have. Note the modelling delay of the diagonal matrix **A** in Fig. 3. The δ of the proposed approach is related to the modelling delay and the length of FIR filter of the adaptive algorithm.

3.4. Direct design method of inverse filters

The conventional adaptive algorithm may not converge to the optimal solution if the tuning parameters such as the modelling delay, the length of FIR filter and a step size of LMS algorithm are not suitably determined. Therefore, much of time and efforts are wasted to determine the proper parameters by trial and error approach.

Note that the proposed method can adjust the approximation accuracy of the inverse filter matrix by deciding *ERROR BOUND* of Eq. (14) and the inverse filter matrix of **C** can be directly obtained as follows:

$$\mathbf{W}(z) \cong -z^{\Delta+\delta} \frac{\sum_{i=1}^{R} \sum_{k=1}^{\delta} \frac{c_i}{b_i^k} z^{-(\delta-k)}}{\mathbf{H}_{min}(z)} adj[\mathbf{C}^{\mathrm{H}}(z)\mathbf{C}(z)]\mathbf{C}^{\mathrm{H}}(z)\mathbf{A}(z), \tag{15}$$

$$\mathbf{A}(z) = z^{-(\Delta+\delta)}\mathbf{I}, \tag{16}$$

where the length of the inverse filter and the time delay are automatically determined.

The conventional adaptive algorithm requires the parameters which should be properly tuned by trial and error approach, while the proposed method requires only the parameter δ which can be systematically obtained by the considering the accuracy and the computing complexity (filter length). Note that the proposed approach does not require adaptation to find the solution and therefore consumes fraction of computing time of the conventional approach.

4. Computer simulation

Both the proposed and conventional adaptive methods for binaural sound system are simulated to verify the feasibility of the proposed method. Fig. 4 shows the binaural sound system with two loudspeakers symmetrically located at 60° and 1.4 m apart where the inverse filter is called a crosstalk canceller. MIT KEMAR HRTF data [9] were used with a sampling frequency of 22.05 kHz and the determinant of the transfer function matrix had 126 zeros consisting of 116 zeros inside the unit circle and 10 zeros outside the unit circle in the complex z-plane as shown in Fig. 5. It was assumed that a virtual sound source is a white noise located at 30° and 2 m apart.

Fig. 6 shows the desired sound signal and the error signal at left ear of the proposed method. The delayed inverse filter matrix calculated from Eq. (15) had a numerator matrix of order 623, a denominator of order 112 and a time delay of 560 steps (*ERROR BOUND* $= 0.0005$, $\delta = 560$). A sum of mean squared errors normalized with the desired sounds at two ears was −13.1 dB. Fig. 7 shows the result with filter coefficients fixed after convergence of the adaptive method with an FIR filter of order 500 and a modelling delay of order 400. The various step sizes of multiple-error LMS algorithm were used for effective convergence. Here, the result after convergence was

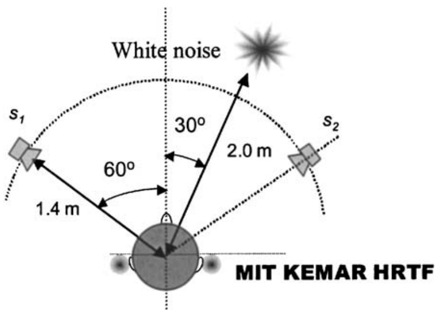

Fig. 4. Binaural sound system.

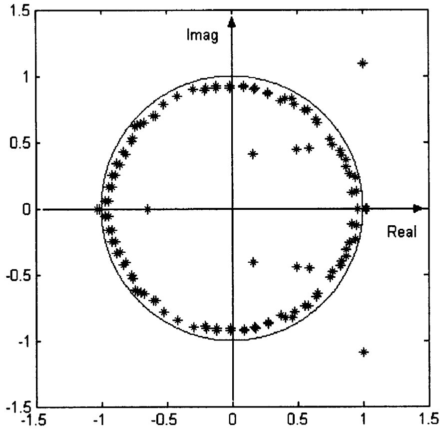

Fig. 5. Poles of the inverse filter: ∗, pole.

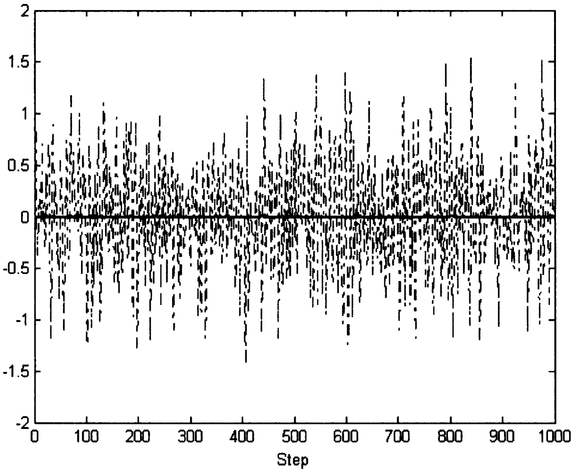

Fig. 6. The proposed method (error bound = 0.0005): -·, desired signal; —, error signal.

presented because the adaptive method rarely converged. The parameters could be properly tuned after lots of time and efforts were consumed. A sum of mean squared errors normalized with the desired sounds at two ears was −5.8 dB.

The inverse filter matrix could be easily and directly obtained by using the proposed method whereas the tuning parameters of the adaptive method could be hardly adjusted by trial and error approach. The proposed method is very useful to obtain the corresponding inverse filter matrices according to various loudspeaker arrangements because less time and efforts are required.

5. Conclusion

The direct designing method of the inverse filter was proposed for use in the multichannel 3-D sound system. The stable and delayed inverse filter matrix could be obtained with the desired

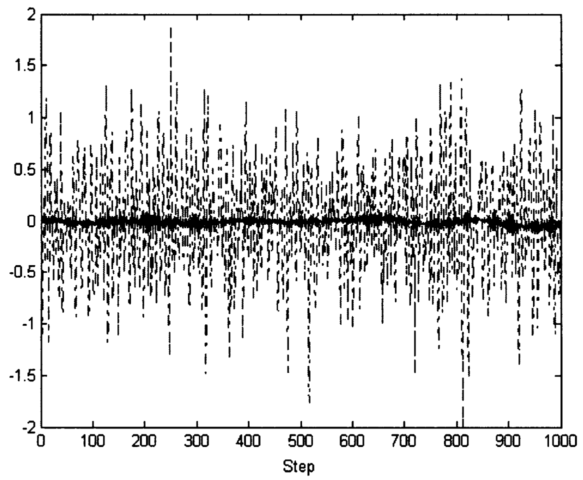

Fig. 7. The result with filter coefficients fixed after convergence of the adaptive method ($L = 500$, $\Delta = 400$): –·, desired signal; —, error signal.

accuracy without heavy computation whereas the conventional adaptive method consumed much of time and effort in adjusting the tuning parameters of LMS algorithm. The computer simulations were carried out to verify the feasibility of the proposed method.

Acknowledgements

The work was partially supported by Virtual Reality Research Center of Korea Science and Engineering Foundation and the Brain Korea 21 project.

Appendix

The inverse filter of the maximum phase system of order 1 can be obtained as follows:

$$\mathbf{H}_{max}(z) = 1 - bz^{-1}, \quad |b| > 1, \tag{A.1}$$

$$\mathbf{H}_{max}(z)^{-1} = \frac{1}{1 - bz^{-1}} \equiv \frac{x(n)}{y(n)}, \tag{A.2}$$

$$x(n) - bx(n-1) = y(n), \tag{A.3}$$

$$x(n) = \frac{1}{b}x(n+1) - \frac{1}{b}y(n+1), \tag{A.4}$$

$$x(n) = -\sum_{k=1}^{\delta} \frac{1}{b^k} y(n+k) + \frac{1}{b^\delta} x(n+\delta) \cong -\sum_{k=1}^{\delta} \frac{1}{b^k} y(n+k), \tag{A.5}$$

$$\mathbf{H}_{max}(z)^{-1} \cong -z^{\delta} \sum_{k=1}^{\delta} \frac{1}{b^k} z^{-(\delta-k)}, \tag{A.6}$$

$$\left|\frac{1}{b}\right|^{\delta} < ERROR\ BOUND. \tag{A.7}$$

References

[1] M. Kleiner, et al., Auralization— an overview, *Journal of Audio Engineering Society* 41 (11) (1993) 861–875.
[2] P.A. Nelson, Active control of acoustic fields and the reproduction of sound, *Journal of Sound and Vibration* 177 (4) (1994) 447–477.
[3] P.A. Nelson, F. Bustamante, H. Hamada, Inverse filter design and equalization zones in multichannel sound reproduction, *IEEE Transactions on Speech and Audio Processing* 3 (3) (1995) 185–192.
[4] A. Gonzalez, J.J. Lopez, Fast transversal filters for deconvolution in multichannel sound reproduction, *IEEE Transactions on Speech and Audio Processing* 9 (4) (2001) 429–440.
[5] M. Bouchard, Y. Feng, Inverse structure for active noise control and combined active noise control/sound reproduction systems, *IEEE Transactions on Speech and Audio Processing* 9 (2) (2001) 141–151.
[6] J.N. Mourjopoulos, Digital equalization of room acoustics, *Journal of Audio Engineering Society* 42 (11) (1994) 884–890.
[7] S.J. Elliott, I.M. Stothers, P.A. Nelson, A multiple error LMS algorithm and its application to active control of sound and vibration, *IEEE Transactions on Acoustics, Speech and Signal Processing* ASSP-35 (1987) 1423–1434.
[8] P.A. Nelson, S.J. Elliott, *Active Control of Sound*, Academic Press, New York, 1992.
[9] B. Gardener, K. Martin, HRTF measurements of KEMAR dummy-head microphone, MIT Media Lab Perceptual Computing, Technical Report No. 280, 1994.