ROBUST CROSSTALK CANCELLATION FOR 3D SOUND USING MULTIPLE LOUDSPEAKERS

Da-Jhuang Su and Shih-Fu Hsieh[1]

Department of Communication Engineering
National Chiao Tung University, Hsinchu, Taiwan 300, Republic of China
Tel: 886-3-5731974, E-mail: djsu.cm93g@nctu.edu.tw, sfhsieh@mail.nctu.edu.tw

ABSTRACT

Crosstalk cancellation is needed for 3D sound playback over two loudspeakers. A critical problem is the degraded spatial hearing due to head movements. Previous research has shown that multiple loudspeakers arrangement is more robust to such perturbation than two loudspeakers setup. In this paper, we will propose novel simplified shuffler structures of crosstalk cancellers for multiple loudspeakers. Their computational costs are lower than conventional setup. Perturbation analysis and computer simulation will demonstrate the advantages of using multiple loudspeakers.

***Index Terms*—**3D sound, crosstalk cancellation

1. INTRODUCTION

A virtual sound source in 3D space can be realized by delivering binaural signals through headphones. However, in many applications, e.g., home entertainment environment, spatialized teleconferencing, many listeners prefer not to wear headphones. If loudspeakers are used to deliver binaural signals, the crosstalk signal arriving at each ear from the other loudspeaker must be canceled[1,7]. Conventional crosstalk canceller is proposed by Atal and Schroeder in 1963. It has demonstrated that the performance of crosstalk canceller suffered from head movement [4]. The approach proposed in [1] is to track the listener and adjust the loudspeaker signals to maintain the binaural transmission so that a more robust crosstalk canceller is possible. Recent works [2, 3] have demonstrated improved performance if a number of loudspeakers are used which exceeds the number of points in the listening space. In such case, the 3D sound reproduction has better immunity to the head movement. In this paper, we will investigate using multiple loudspeakers for 3D sound. A simplified shuffler crosstalk canceller, realized with only two filters, will be proposed under setups of three and four loudspeakers. It requires only two filters to realize. Besides its economical realization, acoustical perturbation analysis also shows it is robust to head movements.

2. SYSTEM MODEL

2.1 Crosstalk Cancellation

Consider the block diagram in Fig. 1, where two binaural signals x_i , *i=1,2,* pass through a 3x2 crosstalk canceller $\mathbf{C}_{3\times2}$; three loudspeaker signals y_j , *j=1,2,3,* are fed into 3 loudspeakers and through a acoustic channel $\mathbf{G}_{2\times3}$ to get two ear signals s_i , *i = 1, 2*. With assumed symmetric setup, we have

$$\mathbf{G}_{2\times3} = \begin{bmatrix} g_1 & g_3 & g_2 \\ g_2 & g_3 & g_1 \end{bmatrix} \tag{1}$$

In order to reproduce binaural signals at both ear, we have

$$\mathbf{G}_{2\times3} \cdot \mathbf{C}_{3\times2} = \mathbf{I}_{2\times2} \tag{2}$$

Based on (2), we can find out the crosstalk canceller $\mathbf{C}_{3\times2}$. Due to symmetric loudspeakers setup, there are three variables to be solved by two equations. Thus, the solutions of $\mathbf{C}_{3\times2}$ is not unique. In the following, we will introduce some possible solutions of crosstalk canceller.

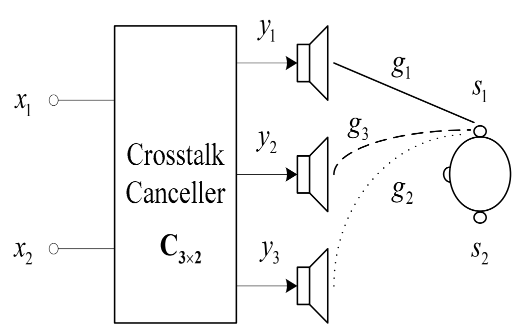

Fig. 1. Three loudspeakers setup

[1] This work was supported by the NSC and MediaTek research center at Chiao Tung University.

2.2 Design of Crosstalk Canceller

A direct forward type crosstalk canceller for three loudspeakers has been suggested in [6] that uses six filters to implement the pseudoinverse of $\mathbf{G}_{2\times3}$, which is undesirable in view of computational cost.

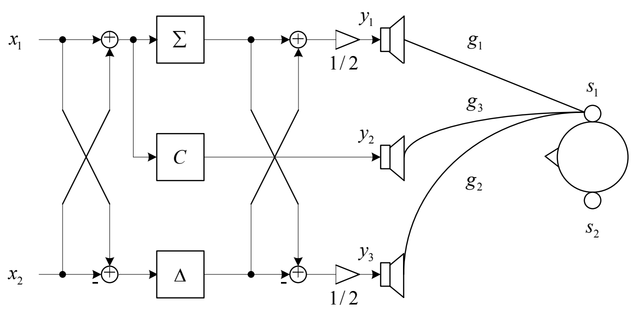

Fig. 2. Shuffler structure of crosstalk canceller

A shuffler form [6] using only 3 filters, as shown in Fig. 2, is then proposed by factorization of the crosstalk canceller matrix $\mathbf{C}_{3\times2}$, where

$$\begin{cases} \Sigma = \dfrac{g_1+g_2}{\left(g_1+g_2\right)^2+2g_3^2} \\ C = \dfrac{g_3}{\left(g_1+g_2\right)^2+2g_3^2} \\ \Delta = \dfrac{1}{g_1-g_2} \end{cases} \tag{3}$$

Now we will derive a novel simplified shuffler form. Due to symmetry in (1), we can reduce (2) into 2 equations:

$$C_1g_1+C_2g_3+C_3g_2=1 \tag{4}$$

$$C_1g_2+C_2g_3+C_3g_1=0 \tag{5}$$

By taking sum and difference of (3) and (4), we have

$$C_1-C_3=\frac{1}{g_1-g_2} \tag{6}$$

$$\left(C_1+C_3\right)\left(g_1+g_2\right)+2C_2g_3=1 \tag{7}$$

To simplify (7), we may choose

$$C_2=\frac{1}{2g_3} \tag{8}$$

In general, $g_1+g_2\neq0$, therefore,

$$C_1+C_3=0 \tag{9}$$

From (6) and (9), two identical filters of opposite sign are

$$C_1=-C_3=\frac{1}{2\left(g_1-g_2\right)} \tag{10}$$

We have arrived at one solution for $\mathbf{C}_{3\times2}$

$$\mathbf{C}_{3\times2}=\frac{1}{2}\begin{bmatrix} \dfrac{1}{g_1-g_2} & \dfrac{-1}{g_1-g_2} \\ \dfrac{1}{g_3} & \dfrac{1}{g_3} \\ \dfrac{-1}{g_1-g_2} & \dfrac{1}{g_1-g_2} \end{bmatrix} \tag{11}$$

We than further factorize $\mathbf{C}_{3\times2}$ into

$$\mathbf{C}_{3\times2}=\frac{1}{2}\begin{bmatrix} 0 & 1 \\ 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} H & 0 \\ 0 & \Delta \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \tag{12}$$

where

$$\begin{cases} H=\dfrac{1}{g_3} \\ \Delta=\dfrac{1}{g_1-g_2} \end{cases} \tag{13}$$

The novel simplified shuffler form is shown in Fig. 3. It can be seen that there are only two filters needed to perform crosstalk cancellation and has less computation than the previous structures. Besides, the filters of simplified shuffler form in (13) are easier to implement than those in (3).

Similarly, we can obtain a simplified shuffler form for four loudspeakers as shown in Fig. 4, where

$$\begin{cases} \Sigma_{23}=\dfrac{1}{g_2+g_3} \\ \Delta_{14}=\dfrac{1}{g_1-g_4} \end{cases} \tag{14}$$

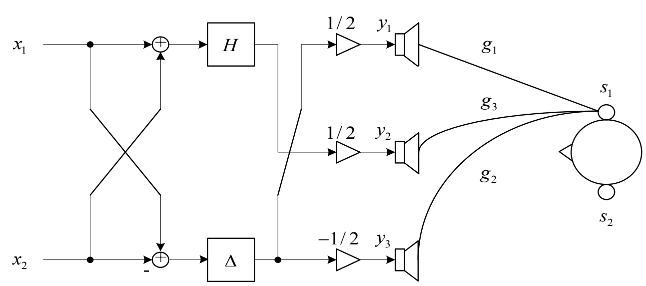

Fig. 3. Simplified shuffler structure of crosstalk canceller for three loudspeakers

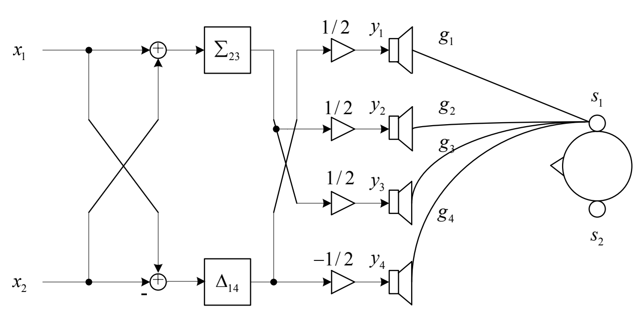

Fig. 4. Simplified shuffler structure of crosstalk canceller for four loudspeakers

3. ACOUSTICAL CHANNEL PERTURBATION

The main disadvantage of crosstalk canceller system is that it is critically dependent on the listener's head being in a fixed design position, the so called "sweet-spot". Many studies have shown that the lateral movement away from the design position of as little as a few centimeters results in loss of the 3D audio effect. While head moves, the acoustical channel **G** will be perturbed to become $\mathbf{G}_+=\mathbf{G}+\delta\mathbf{G}$ so that $\mathbf{G}_+\cdot\mathbf{C}\neq\mathbf{I}_{2\times 2}$. If **C** can be adjusted by $\delta\mathbf{C}$ to satisfy perfect crosstalk cancellation, we have

$$(\mathbf{G}+\delta\mathbf{G})(\mathbf{C}+\delta\mathbf{C})=\mathbf{I}_{2\times 2} \quad (15)$$

From matrix theory in [5], we can get the fractional crosstalk canceller's compensation is bounded by

$$\frac{\|\delta\mathbf{C}\|}{\|\mathbf{C}\|}\leq\|\mathbf{G}^{-1}\|\cdot\|\delta\mathbf{G}\|=cond\{\mathbf{G}\}\cdot\frac{\|\delta\mathbf{G}\|}{\|\mathbf{G}\|} \quad (16)$$

which implies that the larger the condition number $cond\{\mathbf{G}\}$ is, the more sensitive to acoustic channel perturbation the crosstalk canceller **C** is. Conversely, a smaller $cond\{\mathbf{G}\}$ implies the crosstalk canceller **C** has better immunity to channel error, i.e., **C** is more robust. Hence, $cond\{\mathbf{G}\}$ serves as a useful robustness measure for crosstalk canceller.

Fig. 3 shows the condition number of using different numbers of loudspeakers at different frequencies. Here, the robust bandwidth indicates the band so that the condition number is below a specified value, say 3. It can be seen that the robust bandwidth increases (improves) as more loudspeakers are used. Thus, using multiple loudspeakers for 3D sound can result in a robust crosstalk canceller.

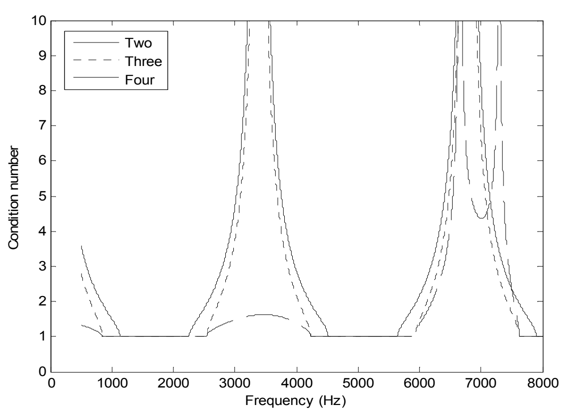

Fig. 5. Robust bandwidths for 2, 3, and 4 loudspeakers

4. SIMULATION RESULTS

4.1 Crosstalk Cancellation Performance

In our simulation, the side loudspeakers are located at $\pm 30°$. For 3 loudspeakers, a center loudspeaker is added. For 4 loudspeakers, an additional pair of loudspeakers are located at $\pm 15°$. We let x_1 be an impulse and x_2 =0. Two performance measures are defined. The signal power ratio after crosstalk cancellation is applied is defined as

$$R_c=10\log\left\|\frac{s_1}{s_2}\right\|^2\Bigg|_{withC's} \quad (4.1)$$

and the total error is defined as

$$\varepsilon=\sqrt{\|\underline{\mathbf{e}}_1\|_2^2+\|\underline{\mathbf{e}}_2\|_2^2} \quad (4.2)$$

where $\underline{\mathbf{e}}_1$ and $\underline{\mathbf{e}}_2$ represent the error between the actual ear signals and the desired ear signals. Fig. 6 and 7 show the performance comparison using different numbers of loudspeakers with different FIR filter orders. We can see that the signal power ratio R_c using three and four loudspeakers simplified shuffler (SS) form is larger than two loudspeakers shuffler form.

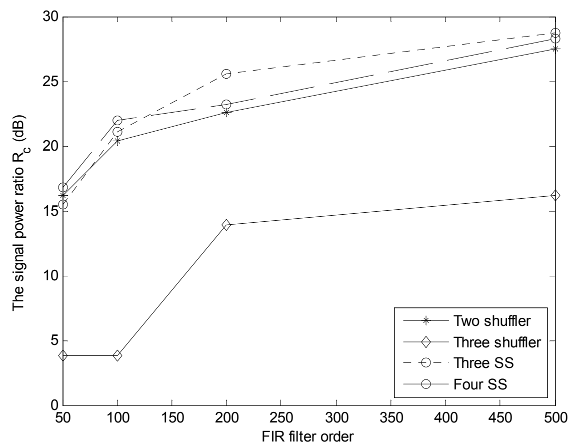

Fig. 6. Comparison of signal power ratios using different numbers of loudspeakers

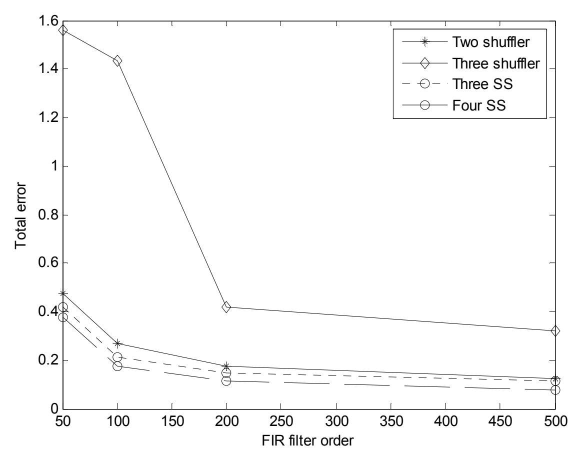

Fig. 7. Comparison of total errors using different numbers of loudspeakers

Similar results for total errors are given in Fig. 7. We can see that simplified shuffler structures using three or four loudspeakers have similar performance and both are much better than that using only two loudspeakers. The performance of three loudspeakers shuffler form is worse than two loudspeakers shuffler form. It is because the filters $\sum$ and C of three loudspeakers shuffler form in (3) require a larger FIR filter to model.

4.2 Perturbation Performance

In this section, we will investigate which structure of crosstalk canceller has better immunity to acoustic channel perturbation. In our simulation, we choose a set of head positions. In Fig. 8 and Fig. 9, we compare the signal power ratio R_c and total error while the head moves around. In these plots, location A represents the initial head position; B, C represents the head movement to right; and D, E to left. It can be seen that both the signal power ratio and total error of different structure crosstalk canceller degrade while head moves to either lateral directions. In addition, we can see that the performance of crosstalk canceller using multiple loudspeakers is better. It means that the use of multiple loudspeakers for 3D sound can result in better immunity of crosstalk canceller to head movements.

5. CONCLUSIONS

We have proposed simplified shuffler forms using three and four loudspeakers for robust crosstalk cancellation. They require only two filters to realize the crosstalk canceller filters, compared to the 6-filters' direct form and 3-filters' shuffler form. In addition, the filters of simplified shuffler form need fewer FIR taps to implement than the shuffler form and its performance is comparable to the direct form implementation. Perturbation analysis has also demonstrated that the crosstalk canceller using multiple loudspeakers is more robust.

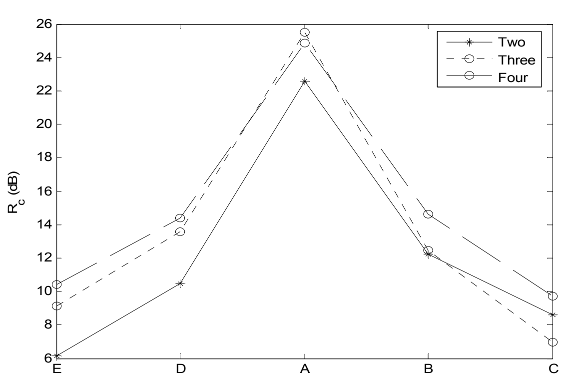

Fig. 8. Comparison of signal power ratios due to head movement

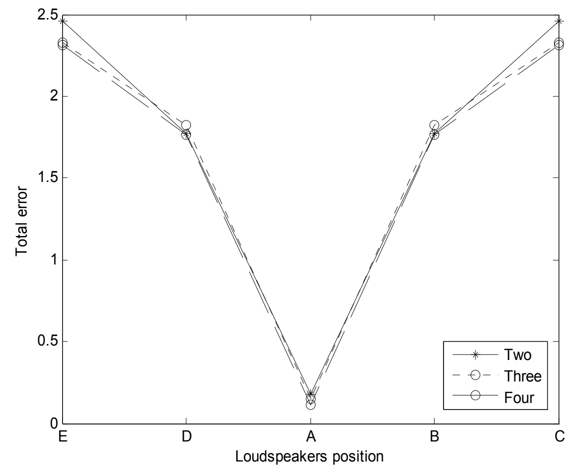

Fig. 9. Comparison of total errors due to head movement

6. REFERENCES

[1] W. G. Gardner, *3D Audio Using Loudspeakers*, Kluwer Academic Pub: Boston, 1998.

[2] J. J. Lopez and A. Gonzalez, "Experimental evaluation of cross-talk cancellation regarding loudspeakers' angle of listening," *IEEE Signal Processing Lett*., vol. 8, no. 1, pp. 13-15, Jan. 2001.

[3] J. Yang, W. S. Gan and S. E. Tang, "Development of virtual sound imaging system using triple elevated speakers," *IEEE Trans. Consumer Electron*., vol. 50, no. 3, pp. 916-922, Aug. 2004.

[4] D. B. Ward and G. W. Elko, "Optimum loudspeaker spacing for robust crosstalk cancellation," *ICASSP'*98, Seattle, vol. 6, pp. 3541-3544, May 1998.

[5] G. H. Golub and C. F. Van Loan, *Matrix Computations, 2nd ed*., Johns Hopkins, 1989.

[6] J. L. Bauck and D. H. Cooper, "Generalized transaural stereo and applications," *J. Audio Eng. Soc.*, vol. 44, no. 9, pp. 683-685, Sept. 1996.

[7] P.Y. Wang and S.F. Hsieh, "Virtual-loudspeakers-based multichannel sound system," *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, Mohonk Mountain House, Oct 19-22, 1997.