
개선된 3차원 오디오 재생을 위한 크로스토크 제거 기법

오승수*, 김기만*

A Cross-talk Cancelling Technique for Improved 3 -Dimensional Audio Reproduction

Seung-Soo Oh, Ki-Man Kim

요 약

일반적으로 라우드 스피커를 이용한 3차원 오디오에서의 크로스토크 제거는 "sweet spot"이라 부르는 청취자의 위치에 따라 성능이 크게 의존하는 것으로 알려져 있다. 따라서 머리의 움직임, 잔향, 머리모양의 차이 같은 혼란을 주는 요소에 강한 새로운 크로스토크 제거기가 제안되었는데 이는 스피커의 대칭 구조와 비대칭 구조를 합성하여 3개의 라우드 스피커가 필요한 구조로 되어있다. 본 논문은 3개의 스피커가 필요한 기존의 크로스토크 제거기 구조를 2개의 스피커만 사용하여 동등한 성능을 구현 할 수 있도록 하였다. 연구된 방법은 청취 실험을 통해 성능을 검증하였다.

ABSTRACT

It has been well known that cross-talk canceller for 3D audio using loudspeakers depends on a listeners position called the sweet-spot. Therefore, new cross-talk canceller was proposed that increases robustness to perturbations such as head movement, reverberations, and different head shapes. It was made up 3 loudspeakers structure to be combined with symmetric and asymmetric speaker geometry. In this paper, we propose new cross-talk canceller using 2 loudspeakers having the same efficiency as existing cross-talk canceller using 3 loudspeakers. The results of the study is verified through the listening tests and also presented a cross-talk cancelling methods for improved 3-D audio production in details.

* 한국해양대학교 전파공학과
접수일자: 2001. 1. 10.

I. 서론

3차원 입체음향 기술은 전문 음향 기술 분야뿐만 아니라 이미 일반 오디오에도 많이 보급되어 있다. 그러나 여전히 보다 완벽하고 충실한 3차원 입체음향 재현에 관한 연구는 꾸준히 전개되고 있다. 이런 분야는 멀티미디어, 가상현실(Virtual Reality)과 같은 엔터테인먼트 분야뿐만 아니라, 시각 장애자를 위한 보조기기(Aid for Blind), 비행 경보 시스템(Pilot Warning System)에 응용 할 수 있다[1][2][3][4].

3차원 오디오의 기본은 청취자의 귀에 실제 음향 환경과 같게 들리도록 음향 신호를 조정하는 것이다. 가장 손쉽게 할 수 있는 방법은 바이노럴 신호(Binaural signal)를 헤드폰을 통해 직접 청취자에게 들려주는 것이지만 청취자가 항상 헤드폰을 착용해야 한다는 제약이 따른다. 라우드 스피커(Loudspeaker)를 통해 바이노럴 신호를 보낸다면 한 쪽 스피커에서 나오는 소리가 반대쪽 귀에 들어가는 크로스토크(Cross-talk)현상이 발생하게 된다. 따라서 이런 크로스토크 현상을 제거하는 것이 라우드 스피커를 이용한 3차원 입체음향 구현에 필수적이다. 그림 1은 기존의 일반적인 크로스토크 제거기를 보여준다. x_L 과 x_R 은 각각 왼쪽과 오른쪽의 바이노럴 신호이고 l_1, l_2 는 라우드 스피커, a_n^L 은 n번째 라우드 스피커에서 청취자의 왼쪽 귀까지의 소리의 전달 함수이다. 크로스토크 신호를 제거할 필터 $h_1 \sim h_4$ 를 구함으로써 청취자는 라우드 스피커를 이용하여 바이노럴 신호를 듣게된다. 따라서 귀에서 들려지는 소리가 y_L, y_R 이라면 크로스토크 제거기는 식 (1)로 표현 할 수 있다.

$$\begin{bmatrix} y_L \\ y_R \end{bmatrix} = \begin{bmatrix} a_1^L & a_2^L \\ a_1^R & a_2^R \end{bmatrix} \begin{bmatrix} h_1 & h_3 \\ h_2 & h_4 \end{bmatrix} \begin{bmatrix} x_L \\ x_R \end{bmatrix}$$

$$y = A H x \dots\dots\dots (1)$$

청취자의 귀에서 정확히 바이노럴 신호가 들리기 위해서는 $AH = I$ (I 는 단위행렬)가 되어야 한다. 그러나 이렇게 구현된 크로스토크 제거기는 청취자의 위치가 조금이라도 변하거나 머리를 움직이게 된다면 전달함수가 변하게 되어 필터의효과가 없어지게 된다. 따라서, 청취자의 움직임에 강한 크로스토크

크 제거기에 대한 연구가 활발하게 되고 있는데 그중 하나가 라우드 스피커 배치의 기하학적 구조를 이용한 청취자 움직임에 강한 크로스토크 제거기이다. 그러나 이 시스템은 일반적인 2채널 오디오 시스템과는 달리 3개의 라우드 스피커를 이용하고 있다 [3].

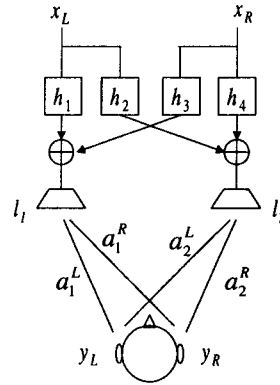


그림 1. 일반적인 크로스토크 제거기.
Fig. 1 The general cross-talk canceller.

하지만 3개의 스피커를 필요로 하는 시스템은 새로운 형태의 하드웨어를 요구하게 되므로 비용이 많이 들게 된다. 따라서 기존의 2개의 스피커를 사용하는 스테레오 시스템에도 적용할 수 있도록 본 논문에서는 3개의 스피커를 필요로 하는 크로스토크 제거기를 2개의 스피커로도 동등한 성능이 나타나도록 구현하였다.

II. 3개의 라우드 스피커를 이용한 크로스토크 제거기

스피커의 대칭, 비대칭 배치와 각 배치형태에 따른 robustness 대역 분석에 따라 두 가지 배치형태를 합성한 새로운 형태의 강한 크로스토크 제거기를 선보였다. 그 특성을 살펴보면 첫째, 주어진 라우드 스피커의 각도에 대해 대칭 배치보다 비대칭 배치의 대역이 더 넓고, 둘째, 대칭 배치는 저주파수 대역의 Robustness를 제공하고, 셋째, 비대칭 배치는 고주파수 대역의 Robustness를 제공한다[3]. 또한 400 Hz이하에서와 5 kHz이상의 신호는 크로스토크 제거

효과가 거의 없으므로 대역저지필터(BSF)를 통과하여 그대로 11과 13스피커로 출력한다. 400 Hz에서 2 kHz의 저주파수대역 신호는 대칭형 배치의 크로스토크 제거기를 이용한다. 대역통과필터1(BPF1)을 통과한 신호는 대칭형 배치의 크로스토크 제거기를 통과하여 11과 13스피커로 출력된다. 2 kHz에서 5 kHz의 고주파수 대역 신호는 비대칭형 배치의 크로스토크 제거기를 이용한다. 대역통과필터2(BPF2)를 지난 신호는 비대칭 배치의 크로스토크 제거기를 지나 왼쪽 신호는 11과 12스피커를 통해, 오른쪽 신호는 12와 13스피커를 통해 출력된다.

III. 제안된 크로스토크 제거기

최근 가정용 극장 시스템과 같은 3개 이상의 멀티스피커를 사용하는 오디오 시스템도 등장했지만, 현재 가장 보편적으로 사용되는 오디오 시스템은 2채널 스피커 시스템을 채용한 경우가 대부분이다. 따라서 앞 절에서 설명한 3개의 스피커를 사용하는 크로스토크제거기를 사용하기 위해서는 새로운 오디오 시스템을 필요로 하는 단점이 있다.

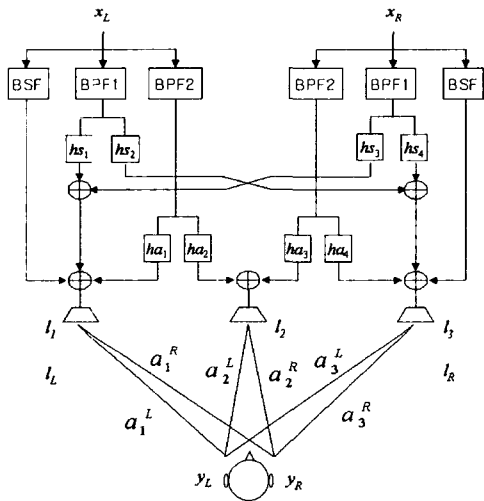


그림 2. 3개의 스피커를 이용한 크로스토크 제거기.

Fig. 2 The robust cross-talk canceller using 3 loudspeakers.

본 논문은 3개의 스피커를 이용한 크로스토크 제

거기를 기존의 오디오 시스템에서도 이용할 수 있도록 2개의 스피커로 구현한다. 따라서, 비대칭 스피커 배치 특성을 이용한 가운데 스피커에서 재생되는 신호를 오른쪽과 왼쪽 스피커로 나오도록 하였다.

n번째 라우드 스피커와 왼쪽 귀 사이의 전달함수를 식 (2)라고 가정한다.

$$a_n^L = e^{-j2\pi\lambda^L d_n^L}, \quad n=1, 2, \dots \quad (2)$$

여기서 λ 는 파장이고 d_n^L 은 n번째 라우드 스피커와 왼쪽 귀 사이의 거리이다. 식 (2)를 가지고 대칭형 스피커배치와 비대칭형 스피커배치 형태에 따른 전달함수를 모델링 할 수 있다.

첫째, 그림 2에서 스피커 배치가 대칭형일 때(라우드 스피커 11, 13)를 생각해보면 이 전달함수는 오른쪽 귀에 대해서도 똑같이 적용할 수 있다. 그림 2에서 대칭형 스피커 배치에 따른 전달함수는

$$a_3^L = a_1^R, \quad a_3^R = a_1^L$$

임을 알 수 있다. 두 귀간 경로 차를 Δ 라 두면, $d_1^R = d_1^L + \Delta$ 이고 이것을 대입하면 다음과 같다.

$$A_s = a_1^L \begin{bmatrix} 1 & e^{-j2\pi\lambda^L \Delta} \\ e^{-j2\pi\lambda^L \Delta} & 1 \end{bmatrix} \quad \dots \quad (3)$$

둘째, 그림 2에서 비대칭 라우드 스피커 배치(라우드 스피커 11, 12)에 대해 생각해보면, 라우드 스피커 12가 청취자의 머리 정 중앙에 있으므로 $a_2^R = a_2^L$ 가 되며 전달함수 행렬은 다음과 같게 된다.

$$A_a = \begin{bmatrix} a_1^L & a_2^L & a_3^L \\ a_1^L e^{-j2\pi\lambda^L \Delta} & a_2^L & a_3^L \end{bmatrix} \quad \dots \quad (4)$$

대칭형, 비대칭형 스피커 배치에 따른 전달함수를 식 (3)와 식 (4)로 나타낼 수 있었다. 전체적으로 그림 2의 3개의 스피커에서 나오는 신호가 청취자의 귀에 도달하는 전달 함수를 A_3 라고 하고 식 (5)와 같이 표현 할 수 있다.

$$A_3 = \begin{bmatrix} a_1^L & a_2^L & a_3^L \\ a_1^R & a_2^R & a_3^R \end{bmatrix} \quad \dots \quad (5)$$

또한 2개의 스피커 1L과 1R를 통해 청취자에게 전

달되는 전달 함수를 A_2 라고 둘 수 있다.

$$A_2 = \begin{bmatrix} a_1^L & a_3^L \\ a_1^R & a_3^R \end{bmatrix} \dots\dots\dots (6)$$

이 전달함수는 2개의 스피커 l_L 과 l_R 에서 재현하고자 하므로 다음과 같은 관계식을 둘 수가 있다.

$$\begin{bmatrix} y_L \\ y_R \end{bmatrix} = \begin{bmatrix} a_1^L & a_2^L & a_3^L \\ a_1^R & a_2^R & a_3^R \end{bmatrix} \begin{bmatrix} l_1 \\ l_2 \\ l_3 \end{bmatrix} \\ = \begin{bmatrix} a_1^L & a_3^L \\ a_1^R & a_3^R \end{bmatrix} \begin{bmatrix} l_L \\ l_R \end{bmatrix} \dots\dots\dots (7)$$

따라서 식 (7)의 관계를 이용하여 식 (8)로 정리할 수 있다.

$$\begin{bmatrix} l_L \\ l_R \end{bmatrix} = \begin{bmatrix} a_1^L & a_3^L \\ a_1^R & a_3^R \end{bmatrix}^{-1} \begin{bmatrix} a_1^L & a_2^L & a_3^L \\ a_1^R & a_2^R & a_3^R \end{bmatrix} \begin{bmatrix} l_1 \\ l_2 \\ l_3 \end{bmatrix} \dots (8)$$

식 (8)을 이용하여 3개의 스피커에서 나오는 신호를 2개의 스피커로 구현해 낼 수 있게 하였다.

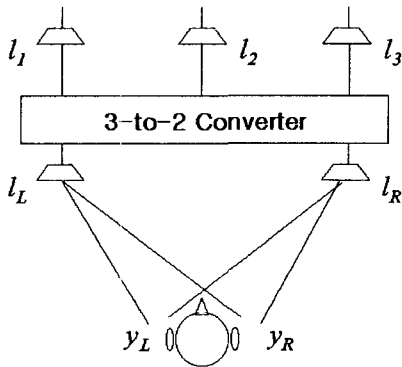


그림 3. 3-to-2 converter.
Fig. 3 3-to-2 converter.

IV. 청취 실험 및 고찰

크로스토크 제거 필터를 구현하기 위해 전달함수는 수직면상 0도, 수평면상으로 0도와 30도 방향에서 측정된 머리전달함수를 사용하였다. 이 머리전달함수는 미국 MIT 대학 Media 연구실에서 측정한

KEMAR HRTF(Head Related Transfer Function)로, 확산음장에서 측정된 128 포인트, 44.1 kHz로 표본화된 것이다[4][5].

기존의 3개의 스피커를 이용한 크로스토크 제거기를 구현하기 위해서, 바이노럴 신호를 120개의 탭을 갖는 FIR filter를 통과시켜 분리한 후, 400 Hz에서 2 kHz대역인 대칭형은 대칭형 크로스토크 제거 알고리즘인 셔플러(shuffler) 행렬을 이용한 방법을 사용하여 역필터를 구현하였으며, 2 kHz에서 5 kHz대역인 비대칭형은 일반 크로스토크 제거 알고리즘을 이용, 왼쪽과 가운데, 오른쪽과 가운데로 나누어서 구현하고, 크로스토크 제거효과가 거의 없는 400 Hz이하 5 kHz는 그대로 두어 전체 크로스토크 제거 회로를 구현하였다.

청취실험은 기존의 크로스토크 제거기와 제안된 방법을 사용하여 실시하였다. 실험을 위해 사용된 입력신호는 바이노럴로 녹음된 음악신호이며 기존의 방법과 제안된 방법을 사용하여 크로스토크를 제거한 후 0도에서 90도의 방향성을 주었다. 실험 방법으로 머리의 움직임을 스피커 정면일 때와 왼쪽으로 10도 돌렸을 때에 기존방법과 제안된 방법으로 크로스토크가 제거된 방향성이 있는 바이노럴 음악신호를 들려주어 소리의 방향을 판정하였다. 실험에 참가한 사람들은 모두 10명이고 이들은 모두 입체음향에 관한 사전 지식이 없었다. 그림 4와 그림 5에 그 결과를 나타내었으며 각 그림에서 표시된 부분의 영역은 인지한 청취자 수에 비례한다.

이 실험을 통해 2개의 스피커를 이용한 기존의 방법과 제안된 방법을 청취하여 비교한 결과, 제안된 방법은 기존의 2개의 스피커를 사용한 방법보다 방향 인지도가 향상되는 것을, 특히 머리의 움직임이 있는 경우 그 효과가 더욱 커지는 것을 입증하였으며, 이는 sweet spot영역이 확대된 것을 의미한다.

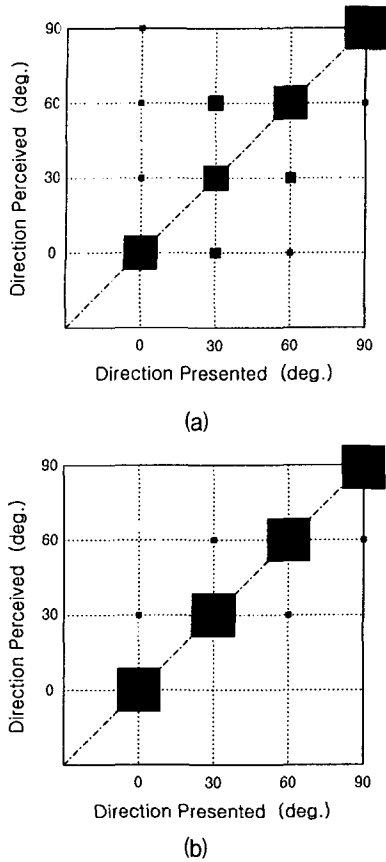


그림 4. 머리를 스피커 정면으로 두었을 때의 청취자의 방향 인지,
 (a) 일반 크로스토크 제거기,
 (b) 제안된 크로스토크 제거기.

Fig. 4 The direction perception when listener's head is located at front of loudspeakers without head turning,
 (a) general cross-talk canceller,
 (b) proposed cross-talk canceller.

V. 결론

기존의 3개의 스피커를 필요로 하는 청취자의 움직임에 강한 크로스토크 제거기는 크로스토크 제거 영역을 크게 확대시켰다. 하지만 일반적인 시스템에서 적용하기 힘든 3개의 스피커를 이용한다는 단점을 가지고 있었다.

본 논문에서는 일반적으로 사용되는 2개의 스피커로 기존의 방법과 같은 성능을 가지도록 크로스토크

크 제거기를 구현하였다. 청취 실험결과, 2개의 스피커만으로도 기존의 방법과 같은 크로스토크 제거영역을 확대할 수 있음을 알 수 있었다.

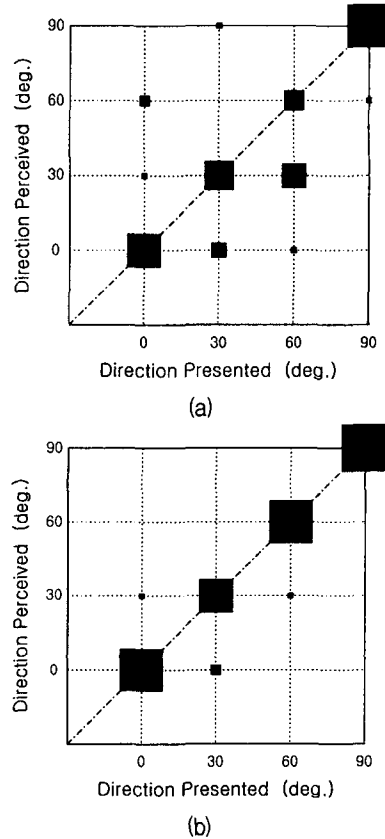


그림 5. 머리를 10도 왼쪽으로 돌렸을 때의 청취자의 방향 인지,
 (a) 일반 크로스토크 제거기,
 (b) 제안된 크로스토크 제거기.

Fig. 5 The direction perception when listener turns head 10 degrees,
 (a) The general cross-talk canceller,
 (b) Proposed cross-talk canceller.

향후 과제로서는 알고리즘을 개선하여 계산량을 줄여 실시간 구현이 가능하도록 해야 할 것이며 하드웨어를 개발하여 실제 오디오 시스템에 적용할 수 있도록 해야 할 것이다.

참고 문헌

[1] 강성훈, "입체음향", 기전연구사, ch. 1-3, 1997.
 [2] D. B. Ward, G. W. Elko, "A Robustness Analysis of 3D Audio Using Loudspeakers", in Proc. IEEE Workshop on Applicat. of signal to Audio and Acoust. (WASPAA-99), New Paltz, NY, USA, Oct. 1999.
 [3] D. B. Ward, G. W. Elko, "A New Robust System for 3D Audio Using Loudspeakers", in IEEE International Conference Acoust. Speech and Signal Processing(2000 ICASSP), 2000.
 [4] W. G. Gardner, "3-D Audio Using Loudspeakers", Degree of Doctor of Philosophy at MIT, Sep. 1997.
 [5] W. G. Gardner, K. D. Martin, "HRTF measurements of a KEMAR" J. Acoust. Soc. Amer., vol.97, no. 6, pp.3907-3908, 1995.
 [6] O. Kirkeby, P. A. Lelson, H. Hamada, "The Stereo Dipole-Virture Source Imaging System Using Two Closely Spaced Loudspeakers", Journal of AES, vol. 46, No. 5, May 1998.
 [7] D. B. Ward, G. W. Elko, " Optimum Loudspeaker Spacing for Robust Crosstalk Cancellation", in IEEE International Conference Acoust. Speech and Signal Processing(1998 ICASSP), 1998.
 [8] B. S. Atal, M. R. Schroeder, " Apparent sound source translator", U.S. Patent 3,236,949, Feb. 1966.
 [9] C. Kyriakakis, P. Tsakalides, T. Holman, "Surrounded by sound", IEEE Signal Processing Mag., vol.16, no. 1, pp. 55-66, Jan. 1999.

오 승 수 (Seung-Soo, Oh)

2000년 한국해양대학교 전파공학과(공학사)
 2000년~현재 : 한국해양대학교 전파공학과 대학원 재학중(석사과정)
 ※주관심분야 : 오디오 신호처리, 오디오 워터마킹, 입체음향, 진기음향

김 기 만 (Ki-Man, Kim)

1988년 연세대학교 전자공학과(공학사)
 1990년 연세대학교 대학원 전자공학과(공학석사)
 1995년 연세대학교 대학원 전자공학과(공학박사)
 1995년~1996년 연세의료원 의용공학교실
 1996년~현재 : 한국해양대학교 전파공학과 조교수
 ※주관심분야 : 오디오 신호처리, 스마트 안테나, 소나 신호처리, 마이크로폰 어레이 시스템, 실시간 DSP 구현 등