

# 2개의 Loudspeaker를 이용하는 3D Audio를 위한 Robust Cross-talk 제거기

오승수 · 김기만

한국해양대학교 전자공학과

## A Robust Cross-talk Cancellation System For 3D Audio Using 2 Loudspeakers

Seung-soo Oh · Ki-man Kim

Dept. of Radio Science & Engineering, Korea Maritime University

E-mail : sunnyside@hananet.net

### 요 약

일반적으로 라우드 스피커를 이용한 3D 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 구조 두 가지 스피커 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 시뮬레이션을 통해 성능을 검증 하였다.

### 1. 서 론

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(Multimedia), 가상 현실(Virtual Reality)과 같은 엔터테인먼트 분야 뿐만 아니라, 시각 장애자를 위한 보조기기(Aid for Blind), 비행 경보 시스템(Pilot Warning System)에 응용 할 수 있다[1][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 항상 청취자가 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 필수적인 것이다.

크로스토크 제거기(Cross-talk Canceller)는 1966년 Atal과 Schroeder등에 의해 제시된 이래 지금까지 많은 방법이 연구되어 왔다[4].

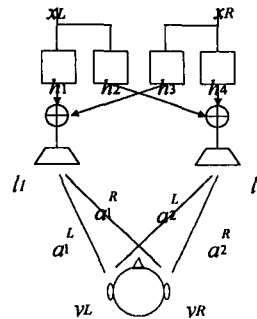


그림 1. 일반적인 크로스토크 제거기 (대칭 라우드 스피커 배치)

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 함수(Transfer Function)이다.

크로스토크 신호를 제거할 필터  $h_1 \sim h_4$ 를 구함으로서 청취자는 라우드 스피커를 이용하여 바이노럴 신호를 듣게된다. 따라서 귀에서 들려지는 소리가  $y_L, y_R$  이라면 크로스토크 제거기는 (1)식으로 표현 할 수 있다.

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

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

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 필터의 효과가 없어지게 된다. 따라서, 청취자의 움직임에 강한 크로스토크 제거기에 대한 연구가 활발하게 되고 있는데 그 중 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.

본 논문에서는 3개의 스피커를 필요로 하는 크로스토크 제거기를 2개의 스피커로도 동등한 성능이 나타나도록 구현하였다.

## 2. 3개의 라우드 스피커를 이용한 Robust 크로스토크 제거기

크로스토크 제거기의 강함을 분석하기 위해서 먼저 적당한 robustness 측정을 정의해야 한다. 가정된 전달함수 모델과 실제 전달함수 사이의 차이는 전달함수 행렬  $A$ 의 불안요소로 고려될 수 있다. 행렬  $A$ 의 불안요소에서 시스템의 robustness는 고유치의 최대값과 최소값의 비인 condition number로 나타난다. 여기서는 robustness 측정 방법으로 전달함수 행렬의 condition number를 사용한다[3].

### 2.1 대칭 라우드 스피커 배치

$n$ 번째 라우드 스피커와 왼쪽 귀 사이의 전달함수를 식(2)라고 가정한다.

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

$\lambda$ 는 파장이고  $d_n^L$ 은  $n$ 번째 라우드 스피커와 왼쪽 귀 사이의 거리이다. 스피커 배치가 대칭형이므로 이 전달함수는 오른쪽 귀에 대해서도 똑같이 적용할 수 있다. 그림1에서와 같은 대칭형 스피커 배치에 대해서 전달함수는  $a_2^L = a_1^R$ ,  $a_2^R = a_1^L$ 임을 알 수 있다. 두 귀간 경로 차를  $\Delta$ 라 두면,  $a_1^R = a_1^L + \Delta$ 이고 이것을 대입하면 다음과 같다.

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

복소수 값을 갖는 행렬에 대해, condition number는 식(4)과 같이 정의된다.

$$\text{cond}\{A\} = \frac{\sigma_{\max}(\sqrt{AA^H})}{\sigma_{\min}(\sqrt{AA^H})} \quad (4)$$

$\sigma_{\max}(\cdot)$ 와  $\sigma_{\min}(\cdot)$ 는 최대, 최소의 singular value이다. 특성방정식의 근을 구함으로써  $A_s, A_s^H$ 의 고유치(eigenvalue)를 구할 수 있다.

$$\sigma_s = 1 \pm \cos(2\pi\lambda^{-1}\Delta) \quad (5)$$

식(4)을 이용하여 condition number를 구하면 다음과 같다.

$$\text{cond}\{A_s\} = \begin{cases} \infty, & \text{if } \cos 2\pi\lambda^{-1}\Delta = \pm 1 \\ 1, & \text{if } \cos 2\pi\lambda^{-1}\Delta = 0 \end{cases} \quad (6)$$

만약 두 귀간의 거리(Interaural distance)  $\Delta$ 가  $\cos(2\pi\lambda^{-1}\Delta) \approx \pm 1$ 라면, 크로스토크 제거기의 성능은 가정한 전달모델이 약간만 변화해도 매우 민감하고,  $\cos(2\pi\lambda^{-1}\Delta) \approx 0$ 이면 변화에 강하게 된다. 따라서 파라미터  $\Delta$ 는 크로스토크 제거기의 robustness를 결정하는 가장 중요한 요소가 된다.

대칭형 스피커 배치에서 두 귀간 경로 차는 스피커 각도에 의해 결정된다. 머리 중앙과 스피커 간의 각도를  $\theta$ 라고 하면, 두 귀간 거리(Interaural distance)  $\Delta$ 는 식(7)과 같다.

$$\Delta = 2r_H \sin \theta \quad (7)$$

$r_H$ 는 머리의 반지름이고 보통 일반 성인의 머리 반지름은 0.0875m이다. 이것을 식(6)에 연관하면 식(8)과 같이 정리할 수 있다.

$$\text{cond}\{A_s\} = \begin{cases} \infty, & \text{if } \sin \theta = i \frac{\lambda}{0.35} \\ 1, & \text{if } \sin \theta = \frac{\lambda}{0.7} + i \frac{\lambda}{0.35} \end{cases} \quad (8)$$

$i \in Z$ 이다.  $\lambda = c/f$ 이므로 다음과 같은 특징이 나타난다. 첫째,  $f$ 가 0가 될수록  $\text{cond}\{A_s\}$ 가 무한대가 되므로 크로스토크 제거는 저주파수 대역에서는 원래 non-robust하다. 둘째, 주어진 라우드 스피커의 각도에 대하여, 시스템이 안정하다면 주파수대역은 제한된다. 셋째, 라우드 스피커의 최적 위치는 주파수에 의존한다는 것이다.

### 2.2 비대칭 라우드 스피커 배치



할 수 있다.

$$A_3 = \begin{bmatrix} a_1^L & a_2^L & a_3^L \\ a_1^R & a_2^R & a_3^R \end{bmatrix} \quad (12)$$

또한 2개의 스피커  $l_L$ 과  $l_R$ 를 통해 청취자에게 전달되는 전달 함수를  $A_2$ 라고 들 수 있다.

$$A_2 = \begin{bmatrix} a_1^L & a_3^L \\ a_1^R & a_3^R \end{bmatrix} \quad (13)$$

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

$$\begin{bmatrix} y_L \\ y_R \end{bmatrix} = \begin{bmatrix} a_1^L & a_2^L & a_3^L \\ a_1^R & a_2^R & a_3^R \end{bmatrix} \begin{bmatrix} l_1 \\ l_2 \\ l_3 \end{bmatrix} \\ = \begin{bmatrix} a_1^L & a_3^L \\ a_1^R & a_3^R \end{bmatrix} \begin{bmatrix} l_L \\ l_R \end{bmatrix} \quad (14)$$

따라서 식(14)의 관계를 이용하여 식(15)로 정리할 수 있다.

$$\begin{bmatrix} l_L \\ l_R \end{bmatrix} = \begin{bmatrix} a_1^L & a_3^L \\ a_1^R & a_3^R \end{bmatrix}^{-1} \begin{bmatrix} a_1^L & a_2^L & a_3^L \\ a_1^R & a_2^R & a_3^R \end{bmatrix} \begin{bmatrix} l_1 \\ l_2 \\ l_3 \end{bmatrix} \quad (15)$$

식(15)을 이용하여 3개의 스피커에서 나오는 신호를 2개의 스피커로 구현해 낼 수 있게 하였다.

#### 4. 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 측정된 KEMAR HRTF(Head Related Transfer Function)로, 확산음장에서 측정된 128포인트, 44.1kHz로 표본화된 것이다[4][6].

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

제안된 방법인 2개의 스피커를 이용한 크로스토크 제거기는 3개의 스피커에서 나오는 신호를 2개의 스피커의 전달함수로 다시 변환하였다.

사용된 바이노럴 입력신호는 핑크잡음이며 0도

에서 90도의 방향성을 주었다. 실험 방법으로 머리의 움직임을 스피커 정면일 때와 왼쪽으로 10도, 오른쪽으로 10도로 돌렸을 때에 방향성이 있는 핑크잡음을 들려주어 소리의 방향을 판정하였다. 9명의 남자와 1명의 여자가 청취실험에 참여하였는데 실험에 참가한 사람들은 모두 입체음향에 관한 사전 지식이 없었다.

이 실험을 통해 3개의 스피커를 이용한 기존의 방법과 제안된 방법의 청취하여 비교한 결과, 제안된 방법으로도 충분히 기존의 방법과 같은 효과를 낼 수 있음을 알 수 있었다.

#### 5. 결 론

기존의 3개의 스피커를 필요로 하는 청취자의 움직임에 강한 크로스토크 제거기는 크로스토크 제거영역을 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 제거기를 구현하였다. 청취 실험결과, 2개의 스피커만으로도 기존의 방법과 같은 크로스토크 제거영역을 확대할 수 있음을 알 수 있었다.

#### 참고문헌

- [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 Applcat. of signal Processing 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] 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
- [6] W. G. Gardner, K. D. Martin, "HRTF measurements of a KEMAR" *J. Acoust. Soc. Amer.*, vol.97, no. 6, pp.3907-3908, 1995.