

3차원 오디오 시스템에서 합과 차 여파기를 이용한 새로운 광대역 간섭신호 제거 방법

A New Robust Acoustic Crosstalk Cancellation Method with Sum and Difference Filter in 3D Audio System

김 래 훈*, 임 준 석**, 성 광 모***
(Lae-Hoon Kim*, Jun-Seok Lim**, Koeng-Mo Sung***)

*서울대학교 전기컴퓨터공학부 음향공학연구실, **세종대학교 전자공학과, ***서울대학교 전기컴퓨터공학부
(접수일자: 2001년 4월 9일; 수정일자: 2001년 5월 16일; 채택일자: 2001년 5월 19일)

본 논문은 차일환 명예회장의 정년퇴임을 기념하여 현정되었습니다

현재 라우드스피커를 이용하는 3차원 오디오 시스템에서의 간섭신호를 제거함으로써 최적 청취 영역 (sweet spot)을 넓히는 여러 가지 방법들이 제안되어 있다. 그러나 기존의 방법들은 간섭신호를 제거할 수 있는 주파수가 제한되어 최적 청취 영역 형성이 제한되는 단점을 공통적으로 가지고 있다. 특히, 저주파 대역의 간섭신호를 제거하지 못하는 본질적인 문제점을 가지고 있다. 본 논문에서는 기존의 시스템보다 훨씬 넓은 주파수 대역에서 간섭신호를 효과적으로 제거할 수 있는 새로운 광대역 간섭신호제거 방법을 제안한다. 이 방법은 세 개의 라우드스피커를 이용한 방법에 합과 차 여파기를 적용하였다.

핵심용어: 오디오 신호처리, 신호 재현, 개인성 분석, 다채널 간섭신호 제거기, 3차원 입체음향 재생 시스템
투고분야: 음향 신호처리 분야 (1.3), 전기음향 분야 (3.4)

There are some methods to enhance the 'sweet spot' in loudspeaker-based 3D audio systems. Most of them can be only applied to narrow frequency band inherently. In this paper, we introduce the more robust 3D sound reproduction system, which has far wider robust bandwidth. The system applies a sum and difference filter to the conventional three loudspeaker-based one.

Keywords: *Audio signal processing, Sound reproduction, Robustness analysis, Multichannel crosstalk canceller, 3D sound reproduction system*

ASK subject classification: *Acoustic signal processing (1.3), Electro-acoustics (3.4)*

I. 서 론

최근 가상현실 시스템 (Virtual Reality System), 홈씨어터 (Home Theatre), 멀티미디어 PC (Multimedia PC) 등이 발달하면서 충실히 3차원 입체음향을 재생할 수 있는 방법에 대해 많은 관심이 생겨나고 있다[1]. 충실히

책임저자: 김래훈 (laihong1@acoustics.snu.ac.kr)

132-302 서울시 관악구 신림동 산 56-1

서울대학교 뉴미디어통신공동연구소 음향공학 연구실(132동 302호)
(전화: 880-8427; 팩스: 886-0791)

입체음향 재생을 위해서는 청취자의 귀 근처에서 음향 신호를 정확하게 제어하는 것이 필수적이다. 이를 위한 가장 쉬운 방법은 헤드폰을 통해 바이노럴 신호를 재생하는 방법이다. 그러나 헤드폰을 사용할 경우에는 마치 머리 속에서 소리가 들리는 듯하게 느끼는 머리 내 음상 정위 등을 포함한 여러 가지 문제점들이 있다[2]. 이런 문제들은 라우드스피커 시스템을 사용하여 오른쪽 신호는 오른쪽 귀에만 들리게 하고, 왼쪽 신호는 왼쪽에만 들리게 하도록 하면 극복할 수 있다. 그러나 스피커 시스템만으

로 완벽한 분리가 불가능하여 왼쪽 및 오른쪽 신호가 서로 섞이는 간섭신호(crosstalk)현상이 생긴다. 이러한 간섭신호를 제거할 때 제거 효과는 정해진 청취 위치를 포함하는 최적 청취 영역(sweet spot)이라고 불려지는 제한된 영역에서 나타난다. 그러나 지금까지 제안된 연구 결과에서는 이런 최적 청취 영역도 제한된 주파수 영역에서만 효과적으로 나타난다. 그러므로 충실한 3차원 음향 재생을 위해서는 보다 넓은 주파수 영역에 걸쳐서 최적 청취 영역 형성하는 간섭 제거기가 매우 중요하다. 이를 위해서는 주파수 변화에 좀 더 강인한 특성을 갖는 간섭 신호 제거기가 필요하다. 근래에 라우드스피커와 청취자 사이의 각도 변경을 이용하여 최적 청취 영역을 넓히고자 하는 몇몇 연구들이 있었다[3-5]. 그러나 이들 연구에서도 저주파에서 간섭신호 제거되지 않는 문제점을 가지고 있다.

본 논문에서는 그러한 주파수 대역폭 제한 문제를 완화 할 수 있는 방법을 제안한다. 또한 제안하는 간섭신호 제거방법이 기존의 방법보다 넓은 주파수 대역에서 강인성이 향상되었음을 시뮬레이션 결과 비교를 통해 보인다.

II. 간섭신호 제거 방법에서의 강인성 분석

음향 재생 시스템에서 왼쪽 귀와 오른쪽 귀에 도달하는 신호 \hat{P}_L, \hat{P}_R 은 아래의 수식과 같이 표현된다.

$$\begin{bmatrix} \hat{P}_L \\ \hat{P}_R \end{bmatrix} = AH \begin{bmatrix} P_L \\ P_R \end{bmatrix}, \quad (1)$$

여기서 A 는 음향 전달 함수이고, H 는 간섭기호 제거 행렬이다. 그리고 P_L 과 P_R 은 각각 왼쪽, 오른쪽으로 입력되는 신호이다. 실제 전달 함수는 원래 위치에서의 청취자의 머리 이동이나 청취자 개인마다의 머리전달함수의 차이로 인해 이상적인 경우와 차이를 보인다. 따라서 위의 모델과 실제간의 차이에서 오는 문제들에 가능한 영향을 적게 받는 시스템이 필요하다. 이런 시스템의 강인성은 선형 시스템 아론에서 아래와 같이 정의 되는 행렬 조건 상수에 반영된다고 알려져 있다[6].

$$cond\left(A = \frac{\sigma_{\max}(\sqrt{AA^H})}{\sigma_{\min}(\sqrt{AA^H})}\right), \quad (2)$$

여기서 $\sigma_{\max}(\sqrt{AA^H})$ 와 $\sigma_{\min}(\sqrt{AA^H})$ 은 각각 가장 큰, 가장 작은 특이값(singular values)을 뜻한다. 위 수

식에서 A 가 음향 전달 함수라고 할 때 주파수에 따라서 서로 다른 값의 행렬 조건 상수가 존재한다. 행렬 조건 상수 값이 매우 높을 때는 $H = A^{-1}$ 에 의하여 디자인되는 간섭신호 제거기 구현이 불가능해진다. 반대로 값이 작으면 간섭신호 제거기를 구현할 수 있게 된다[3,4]. 따라서 가능한 넓은 주파수 범위에 걸쳐서 작은 행렬 조건 상수 값이 유지되도록 하여 간섭신호 제거기가 넓은 범위에 걸쳐서 존재하도록 할 필요가 있다.

III. 세 개의 라우드스피커를 이용한 기존의 간섭신호 제거 방법

그림 1의 세 개의 라우드스피커를 이용한 기존의 간섭신호 제거기에서 n 번째 ($n=1, 2, 3$) 라우드스피커와 m 번째 ($m=1, 2$) 청취자의 귀 사이의 전달함수 A_{mn} 은 $e^{-jkr_{mn}}/r_{mn}$ 으로 표현될 수 있다. 여기서 k 는 파수(wave number)이고 r_{mn} 은 n 번째 라우드스피커와 m 번째 청취자의 귀 사이의 거리이다. 이 전달함수 모델은 청취자의 머리의 영향이나 귀에 도달하기까지 생기는 감쇄를 무시한 간략화된 모델이다. 그러나 이 모델은 양 귀 도달 시간차(interaural time delay)를 모델링하고 있고, 대부분의 실제적인 경우의 라우드스피커 위치(스피커와 청취자의 각도 $\theta \leq 30^\circ$)에서는 머리를 공간상에 존재하는 두 점(본 논문의 경우)으로 가정할 때와 구(球)로 가정할 때의 양 귀 도달 시간

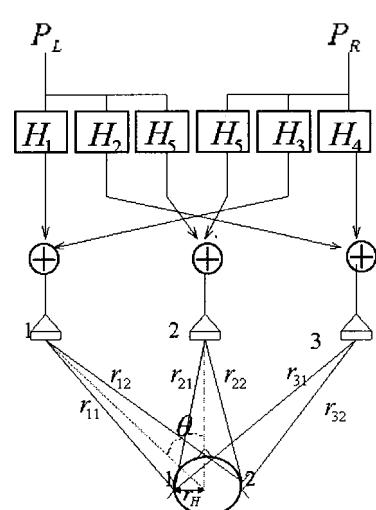


그림 1. 기존의 세 개의 라우드스피커를 이용한 간섭신호 제거기의 구조

Fig. 1. Schematic diagram of the conventional crosstalk canceller for three loudspeaker system.

차가 거의 같다[3]. 또한 양 귀 도달 시간차는 간섭신호 제거기의 강인성을 결정하는 가장 중요한 요소이므로, 이러한 간략화된 모델은 충분히 가치있는 중요한 의미를 갖는다[3].

시스템의 구조를 보면 두 개의 라우드스피커가 청취자 앞에 대칭적으로 놓여있고 세 번째 라우드스피커는 청취자의 바로 앞 중앙에 놓여 있다. 즉 그림 1에서 $r_{11} = r_{32}$, $r_{12} = r_{31}$, $r_{21} = r_{22}$ 이다. 이때 음향 전달 험수 행렬 A는 아래와 같이 표현된다.

$$A = \begin{bmatrix} e^{-jkr_{11}}/r_{11} & e^{-jkr_{21}}/r_{21} & e^{-jkr_{31}}/r_{31} \\ e^{-jkr_{12}}/r_{12} & e^{-jkr_{22}}/r_{22} & e^{-jkr_{32}}/r_{32} \end{bmatrix} = \begin{bmatrix} A_1 & A_2 & A_3 \\ A_3 & A_2 & A_1 \end{bmatrix}. \quad (3)$$

이러한 시스템에서 AA^H 의 고유값들은 다음과 같다.

$$\begin{aligned} \epsilon_{1,2} &= |A_1|^2 + |A_2|^2 + |A_3|^2 \\ &\pm (2|A_1||A_3|\cos(k\Delta) + |A_2|^2), \end{aligned} \quad (4)$$

여기서 $\Delta = r_{12} - r_{11} \approx 2r_H \sin \theta$ 는 양 귀에 도달하는 신호들의 경로차를 의미하며, θ 는 머리 정중앙과 라우드스피커 사이의 각도를 의미하고, r_H 는 인간 머리의 반지름으로 성인 남자의 머리 반지름의 평균값인 0.0875m를 사용했다. $\cos(k\Delta) = \cos(2\pi f/c \times \Delta)$ 이므로 다음과 같은 결과를 얻을 수 있다. f 는 주파수이고 c 는 음속이다.:

- (1) 저주파에 가까워 질수록 $\cos(k\Delta)$ 값이 영(zero) 값에 가까워지고 이에 따라 $\text{cond}(A)$ 이 무한대로 발산해 가므로 저주파에서의 간섭신호 제거는 본질적으로 강인하지 못하다.
- (2) 고주파에서도 $f/c \times \Delta$ 값이 정수배가 되는 곳에서 $\cos(k\Delta)$ 값이 영(zero) 값에 가까워지므로 주어진 라우드스피커의 어느 각도에서도 제한된 주파수 영역에서만 시스템이 안정화될 수 있다.
- 식 (3)에서 A_2 가 빠지면 두개의 스피커를 사용한 시스템의 전달 특성과 동일해 지므로 이 두개의 스피커를 사용한 시스템 역시 전술한 단점을 가지고 있다.

IV. 새롭게 제안된 방법

새롭게 제안된 시스템의 구조는 그림 2를 통해 확인할

수 있다. 이 시스템은 기존 시스템에는 없는 합과 차 여파기를 포함하고 있다. 음향 재현의 방식은 아래와 같은 수식으로 표현될 수 있다.

$$\begin{bmatrix} P_L \\ P_R \end{bmatrix} = \begin{bmatrix} e^{-jkr_{11}}/r_{11} & e^{-jkr_{21}}/r_{21} & e^{-jkr_{31}}/r_{31} \\ e^{-jkr_{12}}/r_{12} & e^{-jkr_{22}}/r_{22} & e^{-jkr_{32}}/r_{32} \end{bmatrix} \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} h_1 & h_3 \\ h_5 & h_6 \\ h_2 & h_4 \end{bmatrix} \begin{bmatrix} P_L \\ P_R \end{bmatrix} \triangleq AHb. \quad (5)$$

음향 전달 특성은 다음과 같이 표현된다.

$$A = \begin{bmatrix} e^{-jkr_{11}}/r_{11} & -e^{-jkr_{31}}/r_{31} & e^{-jkr_{21}}/r_{21} \\ e^{-jkr_{12}}/r_{12} & -e^{-jkr_{32}}/r_{32} & e^{-jkr_{22}}/r_{22} \\ e^{-jkr_{31}}/r_{31} & -e^{-jkr_{11}}/r_{11} \\ e^{-jkr_{32}}/r_{32} & -e^{-jkr_{12}}/r_{12} \end{bmatrix}. \quad (6)$$

AA^H 의 고유값들은 $\epsilon_{1,2} = a \pm b$ 이라 할 수 있고 여기서 a , b 는 아래와 같다.

$$\begin{aligned} a &= 2|e^{-jkr_{11}}/r_{11} - e^{-jkr_{31}}/r_{31}|^2 + |e^{-jkr_{21}}/r_{21}|^2 \\ &= 2\{(1/r_{11})^2 + (1/r_{31})^2\}(1/r_{21})^2 - 4\cos(k\Delta)/(r_{11}r_{31}) \\ b &= (e^{-jkr_{11}}/r_{11} - e^{-jkr_{31}}/r_{31})(e^{jkr_{31}}/r_{31} - e^{-jkr_{11}}/r_{11}) \\ &+ (e^{jkr_{11}}/r_{11} - e^{-jkr_{31}}/r_{31})(e^{-jkr_{31}}/r_{31} - e^{-jkr_{11}}/r_{11}) + (1/r_{21})^2 \\ &= -2\{(1/r_{11})^2 + (1/r_{31})^2\} + (1/r_{21})^2 \\ &+ 4\cos(k\Delta)/(r_{11}r_{31}) \end{aligned} \quad (7)$$

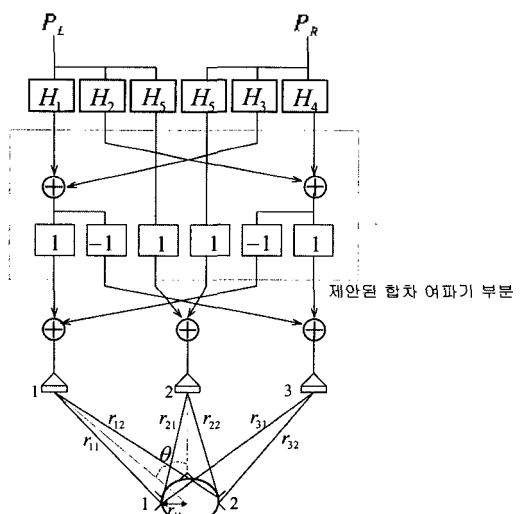


그림 2. 제안된 세 개의 라우드스피커를 이용한 간섭신호 제거기의 구조

Fig. 2. Schematic diagram of the proposed crosstalk canceller for three loudspeaker system.

이런 경우에

$$cond[A] = \begin{cases} \sqrt{\frac{2[(1/r_{11})^2 + (1/r_{31})^2 - 4\cos(k\Delta)/(r_{11}r_{31})]}{(1/r_{21})^2}} \\ \text{if } 2[(1/r_{11})^2 + (1/r_{31})^2 - 4\cos(k\Delta)/(r_{11}r_{31})] \geq (1/r_{21})^2 \\ \sqrt{\frac{(1/r_{21})^2}{2[(1/r_{11})^2 + (1/r_{31})^2 - 4\cos(k\Delta)/(r_{11}r_{31})]}} \\ \text{if } 2[(1/r_{11})^2 + (1/r_{31})^2 - 4\cos(k\Delta)/(r_{11}r_{31})] \geq (1/r_{21})^2 \end{cases} \quad (8)$$

식 (8)에서 $\cos(k\Delta)$ 값이 영(zero) 값에 가까워진다 하더라도 $cond[A]$ 값은 발산하지 않고 특정 값으로 수렴한

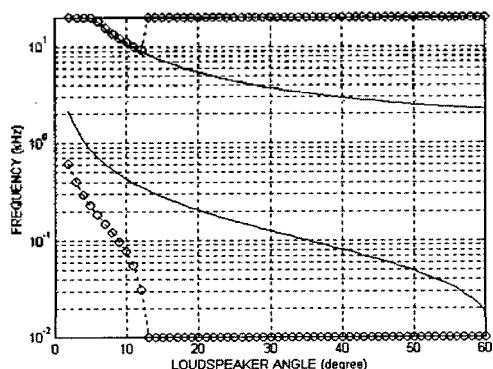


그림 3. $r_{11} = r_{21} = 0.5m$ 일 때의 기존의 시스템(굵은 선)과 제안된 시스템(점선과 원)의 'robust bandwidth' (위선과 아래선 사이에 해당됨)

$$r_{11} = r_{21} = 0.5m$$

Fig. 3. Robust bandwidth (between lower bound and upper bound) for a conventional 3-channel system (solid) and the newly proposed system (dashed&circle). $r_{11} = r_{21} = 0.5m$.

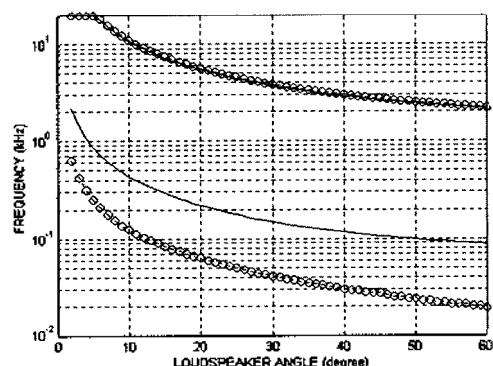


그림 4. $r_{11} = r_{21} = 3m$ 일 때의 기존의 시스템(굵은 선)과 제안된 시스템(점선과 원)의 'robust bandwidth' (위선과 아래선 사이에 해당됨) $r_{11} = r_{21} = 3m$

Fig. 4. Robust bandwidth (between lower bound and upper bound) for a conventional 3-channel system (solid) and the newly proposed system (dashed&circle). $r_{11} = r_{21} = 3m$.

다. 즉 이 시스템은 기존의 방법에서는 본질적으로 불가능했던 저주파에서 까지도 안정화된 간섭신호 제거를 할 수 있음을 알 수 있다. 또한, r_{11} , r_{21} , θ 값을 조정하면 고주파에서도 적정치 이하로 낮출 수 있다. 이러한 결과는 이어지는 시뮬레이션 결과를 통해 확인할 수 있다.

V. 시뮬레이션 결과 및 고찰

시뮬레이션 결과인 그림 3와 그림 4에서는 각각 다른 거리에서의 라우드스피커 세 개를 이용하는 기존의 시스템과 여기서 새로 제안하는 시스템 각각의 장인한 간섭신호 제거기 사용 대역폭을 라우드스피커 각도에 따라 함께 보여주고 있다. 그림 3은 $r_{11} = r_{21}$ 값이 0.5m일 때의 결과이고, 그림 4는 $r_{11} = r_{21}$ 값이 3m일 때의 결과이다. 그리고, 인간은 음높이를 비선형적으로 인식하므로 두 그림에서 모두 로그 스케일을 사용했다. 시스템의 장인함을 비교하기 위해서 여기서는 사용 가능 대역폭(행렬 조건 상수가 어느 값 이하를 갖는 연속적인 주파수 대역을 의미한다. 여기서는 그 값을 참조 논문과 같은 값 10으로 설정했다[5].)의 정도를 비교한다.

그림 3의 제안된 방법에서는 사용 가능 대역폭이 12° 이상 값을 갖는 라우드스피커 각도에서 전 가청 주파수 대역에 이르는 것을 볼 수 있다. 여기서 라우드스피커와 청취자 사이의 거리가 0.5m이므로 이러한 구조는 멀티미디어 PC용 3차원 입체음향 재생시스템에 활용 가능하다. 그림 4에서는 비록 그림 3의 결과보다 고주파 특성은 비교적 좋지 않지만 기존의 시스템보다는 좋은 저주파 특성을 확인할 수 있다. 이러한 성능 저하는 라우드스피커와 청취자 사이의 거리가 3m로 멀어진 것에서 기인한다. 이 구조는 바이노럴 녹음 또는 바이노럴 합성 음원을 사용할 수 있는 가상 현실 시스템이나 흡씨어터에 적합하다.

시뮬레이션 결과는 제안된 방법이 넓은 사용 가능 대역폭을 가지고 있음을 보여준다. 또한 이를 이용한 구조는 멀티미디어 PC로부터 흡씨어터에 이르기까지 그 활용범위가 다양할 수 있을 것이다.

VI. 결론

본 논문에서는 합과 차 여파기를 사용하여 기존보다 더 넓은 영역에 걸쳐서 간섭신호를 제거할 수 있는 방법

을 제안하였다. 본 제안 방법을 사용함으로써 기존의 방법보다 더 넓은 대역에 걸쳐서 3차원 효과를 내는 시스템을 구성할 수 있음을 시뮬레이션을 통하여 보였다. 또 멀티미디어 PC같은 작은 규모뿐 아니라 홈씨어터같은 보다 큰 규모에 대해서도 넓은 주파수 영역에 대해서 간섭 신호를 제거할 수 있음을 함께 보였다.

참 고 문 헌

1. C. Kyriakakis, "Fundamental and Technological Limitations of Immersive Audio System", *Proceedings of IEEE*, 5, pp. 941–951, 1998.
2. 이신렬, 김래훈, 방희석, 성광모, "바이노럴 합성을 위한 3채널 HRTF 측정," *한국음향학회 학술발표대회 논문집 제19권 제 2(s) 호*, 337–340, 2000.
3. D. B. Ward, and G. W. Elko, "Effect of Loudspeaker Position on the Robustness of Acoustic Crosstalk Cancellation," *IEEE Signal Processing Letter*, 6, pp106–108, 1999.
4. J. Yang, and W. Gan, "Speaker placement for robust virtual audio display system," *Electronics Letters*, 36, pp. 683–685, 2000.
5. D. B. Ward, and G. W. Elko, "A new robust system for 3D audio using loudspeakers," *Proc. IEEE ICASSP00*, 2, pp. II781–II784, 2000.
6. G. H. Golub and C. F. Van Loan, *Matrix computation*, The Johns Hopkins Univ. Press, Baltimore, MD, 3rd Edition, 1996.

저자 약력

● 김 래 훈 (Lae-Hoon Kim)

1975년 7월 15일생
1999년 2월: 서울대학교 전기공학부 졸업
2001년 2월: 서울대학교 전기컴퓨터공학부 석사
2001년 3월~현재: 서울대학교 전기컴퓨터공학부 박사과정
※ 주관심분야: 음향신호처리

● 임 준 석 (Jun-Seok Lim)

한국음향학회지 제20권 제1E호 참조

● 성 평 모 (Koeng-Mo Sung)

한국음향학회지 제20권 제1E호 참조