

입체 음향 재생을 위한 실시간 트랜스오럴 필터 구현

이동우*, 김영오*, 고대식*, 강성훈**

*목원대학교 전자공학과, **대전보건대학 방송제작기술과

Implementation of Real-time Transaural Filter for 3D Sound Reproduction

Dongwoo Lee*, Youngoh Kim*, Deasik Ko*, Seonghoon Kang**

*Mokwon Univ., **Taejon Medical Science College

dwlee@ee.mokwon.ac.kr

요약

본 논문에서는 스테레오 스피커를 통해 입체 음향을 재생할 때 발생하는 크로스 토크(cross talk)를 제거하기 위한 트랜스오럴 필터를 실시간으로 구현하였다. 실시간 처리를 위하여 일반적인 트랜스오럴 필터를 재구성하였으며, 범용 DSP(Digital Signal Processor) 칩인 TMS320C40을 사용하였다. 메인 프로그램은 C-언어를 이용하여 작성하였으며, 빠른 연산을 필요로 하는 컨벌루션 함수는 어셈블리 언어를 이용하여 최적화 하였다. 구현된 트랜스오럴 필터의 평가를 위해서 더미 헤드를 이용하여 녹음된 binaural 신호와 DSP를 이용하여 실시간으로 합성된 binaural 신호를 스테레오 스피커를 통해 재생하여 청취 실험하였다. 청취 실험한 결과 머리 밖에서 음상 정위가 이루어지고 공간감과 거리감은 헤드폰으로 청취할 때 보다 우수하게 나타났다.

1. 서론

최근 들어 입체 음향에 대한 관심이 높아지면서 국내외 대학과 연구소들을 중심으로 입체 음향에 대한 연구가 활발하게 이루어지고 있으며 많은 연구 결과가 발표되고 있다[1,2,3]. 이와 같이 입체 음향에 대한 연구가 활발히 진행되고 있으나 지금까지의 많은 연구가 PC상에서 시뮬레이션만을 통해 진행되거나 오디오의 음질을 낮추어 전체적인 연산량을 줄여서 실시간 처리하였다. 그리고 스테레오 스피커를 통해 입체 음향을 재생할 때 발생하는 크로스 토크(cross talk)를 제거하지 못하여 헤드폰만을 통해서 입체 음향을 재생하였다. 그러나 오디오, 게임기 그리고 가상 현실 등과 같은 분야에서 필

요로 하는 현장감 있는 입체 음향을 제공하기 위해서는 현재 가장 보편화되어 있는 CD 수준의 음질(16bit, 44.1 kHz sampling rate)을 유지하면서 헤드폰과 스테레오 스피커를 통해 실시간으로 입체 음향을 재생할 수 있어야 할 것이다.

본 논문에서는 스테레오 스피커를 통해 입체 음향을 재생할 때 발생하는 크로스 토크를 제거하는 방법으로 일반적인 트랜스오럴(transaural) 필터를 실시간 처리에 적합하도록 재구성하고, 트랜스오럴 필터 연산 중에 연산량이 많은 컨벌루션(convolution) 처리부는 어셈블리 언어를 이용하여 작성하여 범용 DSP(Digital Signal Processor) 칩인 TMS320C40에서 실시간으로 동작하도록 구현하였다.

2.1 트랜스오럴 필터

트랜스오럴 필터의 기본적인 개념은 음원의 방향 정보를 가지고 있는 binaural 신호를 스테레오 스피커를 통해 청취자의 귀까지 전송하는 것이다. 트랜스오럴 필터는 Schroeder와 Atal에 의해 처음 제안되고, Cooper와 Bauck에 의해 개선되었다[4].

그림 1은 청취자가 스테레오 스피커를 통해 소리를 청취할 때의 상황을 나타내고 있다.

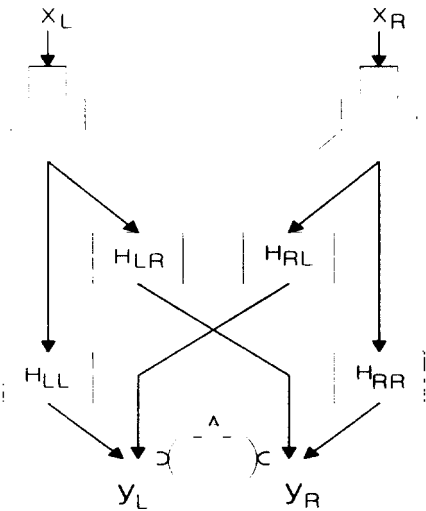


그림 1. 스테레오 스피커로부터 귀까지의 전달함수

x_L 과 x_R 은 스피커로 보내지는 binaural 신호이고, y_L 과 y_R 은 청취자의 양 귀에 도달하는 신호이다. H_{LL} , H_{LR} , H_{RL} , H_{RR} 은 각각 왼쪽 스피커에서 왼쪽 귀까지의 전달함수, 오른쪽 스피커에서 오른쪽 귀까지의 공간 전달함수, 왼쪽 스피커에서 오른쪽 귀까지의 공간 전달함수, 오른쪽 스피커에서 왼쪽 귀까지의 공간 전달함수를 나타낸다. 이와 같은 상태를 벡터 식으로 표현하면 식 (1)과 같다.

$$y = \begin{bmatrix} y_L \\ y_R \end{bmatrix}, H = \begin{bmatrix} H_{LL} & H_{LR} \\ H_{LR} & H_{RR} \end{bmatrix}, x = \begin{bmatrix} x_L \\ x_R \end{bmatrix}, \quad (1)$$

$$y = Hx$$

스피커 X에서 양쪽 귀 Y로의 전달 함수는 H_{ij} 이다. 그러므로, 스피커 X에서 binaural 신호 x 가 방출되면 공간 전달함수 H 를 통과해 양쪽 귀 Y에 신호 y 가 도달하는데, 귀에 도달하는 신호 y 는 스피커 X에서 방출된 신호 x 와 같지 않다. 즉, 오른쪽 귀에는 오른쪽 스피커에서 방출된 신호와 왼쪽 스피커에서 방출된 신호가 함께 도달하며, 왼쪽 귀에 도달하는 신호 역시 마찬가지이다. 그리고 양쪽 귀 Y에 도달한 신호 y 는 전달 함수 H 의 특성을 가지고 있다. 이와 같은 현상을 크로스 토크라고 하며, 크로스 토크를 제거하기 위한 일반적인 트랜스오럴 필터는 그림 2와 같다.

트랜스오럴 필터의 원리는 다음과 같다. x 가 청취자의 귀에 도달하기 원하는 binaural 신호이면, 즉 스피커 X에서 방출된 신호 x 를 원래의 신호 그대로 듣기 위해서는 전달함수행렬 H 는 $\hat{x} = H^{-1}x$ 와 같이 역함수를 취해야 한다. 즉 스피커 X로부터 출력되는 신호 x 와

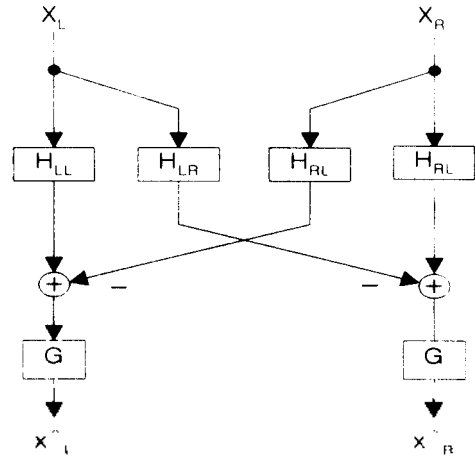


그림 2. 일반적인 트랜스오럴 필터
($G = 1/(H_{LL}H_{RR} - H_{LR}H_{RL})$.)

귀 Y에 도착하는 신호 y 와 같게 만들려면, 전달함수 H 의 역함수를 구해야 한다. 전달함수 H 의 역행렬은 식 (2)와 같다.

$$H^{-1} = \frac{1}{(H_{LL}H_{RR} - H_{LR}H_{RL})} \begin{bmatrix} H_{RR} & -H_{RL} \\ -H_{LR} & H_{LL} \end{bmatrix} \quad (2)$$

이 필터는 채널간 크로스 토크를 제거하는 역할을 하기 때문에 크로스 토크 제거 필터라고도 불린다. 대칭형 스피커로 청취하는 상황에서는 동축상의 항 ($H_i = H_{LL} = H_{RR}$)끼리와 반대 축상의 항 ($H_c = H_{LR} = H_{RL}$)끼리 같다고 볼 수가 있다. 그러므로 식 (2)는 식 (3)과 같이 쓸 수 있다.

$$H^{-1} = \frac{1}{(H_i^2 - H_c^2)} \begin{bmatrix} H_i & -H_c \\ -H_c & H_i \end{bmatrix} \quad (3)$$

2.2 실시간 트랜스오럴 필터의 설계

앞에서 설명한 일반적인 트랜스오럴 필터는 실시간 처리에 부적합하다. 일반적인 트랜스오럴 필터를 시간 영역에서 처리하기 위해서는 최소 8번 이상의 컨벌루션 연산이 필요하고 역 전달함수 H^{-1} 의 성능에 따라 알고리즘 전체의 기능이 좌우되기 때문이다. 그러므로 크로스 토크를 실시간으로 제거하기 위해서는 그림 2와 같은 일반적인 트랜스오럴 필터를 재 구성해야 할 필요성이 있다. 그림 3은 실시간 처리를 위하여 본 논문에서 재구성된 트랜스오럴 필터의 구조이다.

상단부의 H_{LR} , H_{RL} , H_{LR} , H_{RL} 블록은 크로스 토크를 제거하는 부분으로서, X_L 과 X_R 이 입력되면 오른쪽 스피커에서 왼쪽 귀로의 전달함수 H_{RL} 과 왼쪽 스피커에서 오른쪽 귀로의 전달함수 H_{LR} 를 순환하면서 각각 반대쪽의 스피커에서 발생하는 소리를 제거한다. 그리고,

하단부의 H_{LL}^{-1} 와 H_{RR}^{-1} 은 각각 왼쪽 스피커에서 왼쪽 귀로의 역 전달함수와 오른쪽 스피커에서 오른쪽 귀로의 역 전달함수로서 스피커에서 귀까지의 공간의 효과를 제거하기 위한 부분이다.

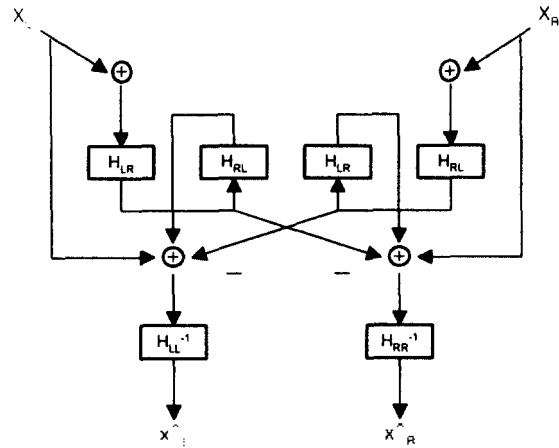


그림 3. 제안된 트랜스오럴 필터

그림 3의 트랜스오럴 필터는 6개의 컨벌루션으로 구현이 가능하며 미리 역 전달함수 H^{-1} 을 계산하여 실시간 처리에 적용이 가능하다.

2.3 실시간 트랜스오럴 필터를 위한 최적화 기법

제안된 트랜스오럴 필터를 실시간으로 구현하기 위하여 부동 소수점 연산용 DSP칩인 TMS320C40을 기반으로 한 디지털 신호 처리 모드(Ariel사의 PC-Hydra)를 사용하였다. 메인 함수는 C-언어를 사용하여 작성하였으며, 연산 양이 많은 컨벌루션 함수는 어셈블리 언어를 사용하여 연산 속도를 증가시켰다. 오디오 인터페이스를 위한 프로그램은 인터럽트 방식에 비해 코드의 양이 적은 플래그(flag) 방식을 이용하였다. 그림 4의 컨벌루션 함수는 C-언어 프로그램에서 실행 할 수 있는 함수로서 어셈블리 언어를 이용하여 최적화 한 것이다. 입력 데이터의 주소(address)와 전달함수의 주소, 그리고 컨벌루션 포인트 수를 입력받아 컨벌루션 처리를 거쳐 부동소수점 연산 결과를 메인(main) 함수로 반환한다.

그림 4의 개별적 처리를 살펴보면 ①은 함수의 이름을 정의하고, ②는 메인 함수로부터 연산에 필요한 인자(argument)를 받을 보조 레지스터(auxiliary register)를 정의한다. ③은 프로그램 실행 코드의 시작을 표시한다. ④는 함수 시작 위치를 표시하며, ⑤는 메인 함수로부터 인자를 받을 때 사용할 보조 레지스터에 이

①	global	_convf
②	FP	.set AR3
③		.text
④	_convf	
⑤	PUSH	FP
⑥	LDI	SP,FP
⑦	LDI	*-FP(2),AR0
⑧	LDI	*-FP(3),AR1
⑨	LDI	*-FP(4),BK
⑩	LDI	BK, RC
⑪	SUBI	2,RC
⑫	LDF	0.0,R2
⑬	MPYF3	*AR0+*(1%,*AR1--(1),R0
⑭	RPTS	RC
⑮	MPYF3	*AR0+*(1%,*AR1--(1),R0
⑯	ADDF3	R0,R2,R2
⑰	ADDF3	R0,R2,R0
⑱	POP	FP
⑲	RETS	
⑳		.end

그림 4. 최적화 설계된 컨벌루션 함수

저장되어 있는 데이터를 대피시킨다. ⑥은 ②에서 정의된 보조 레지스터를 메인 함수에서 전송 받을 인자가 있는 위치를 가리키게 한다. ⑦은 보조 레지스터 AR0에 첫 번째 인자를 저장한다. 트랜스오럴 필터에서는 컨벌루션 처리할 입력 데이터의 위치를 저장한다. ⑧과 ⑨는 각각 전달함수의 위치 또는 전달함수의 역함수(트랜스오럴 필터)의 위치와 컨벌루션 포인트 수를 AR0과 BK에 저장한다. ⑩과 ⑪은 반복 연산을 위해 RC(repeat counter)를 설정한다. ⑫는 R2 레지스터를 초기화한다. ⑬은 첫 번째 입력 데이터와 첫 번째 전달함수 값을 곱해서 R0에 저장하고, ⑭는 RC를 동작시킨다. ⑮는 입력 데이터의 주소와 전달함수의 주소를 증가/감소시키면서 ⑬과 같은 연산을 RC에 저장된 횟수만큼 수행하여 결과를 R0에 저장한다. ⑯의 "||"연산자는 파이프라인 연산자로서 앞줄의 연산과 동시에 현재 연산을 처리하도록 하는 연산자로서, ⑮의 연산이 이루어지는 동시에 결과가 R2에 저장된다. ⑰은 마지막으로 R2에 저장된 최종 결과를 R0에 저장한다[5]. ⑱, ⑲, ⑳은 처리 결과인 R0의 데이터를 부동 소수점 데이터로서 메인 함수에 전송하고 함수를 끝낸다.

3. 실험 및 고찰

제안된 트랜스오럴 필터를 앞에서 언급한 방법을 이

용하여 작성하고 디지털 신호 처리 보드에서 실행시켜 성능을 평가하였다. 그림 5는 구현된 실시간 트랜스오럴 필터의 성능 평가를 위한 실험 환경이다.

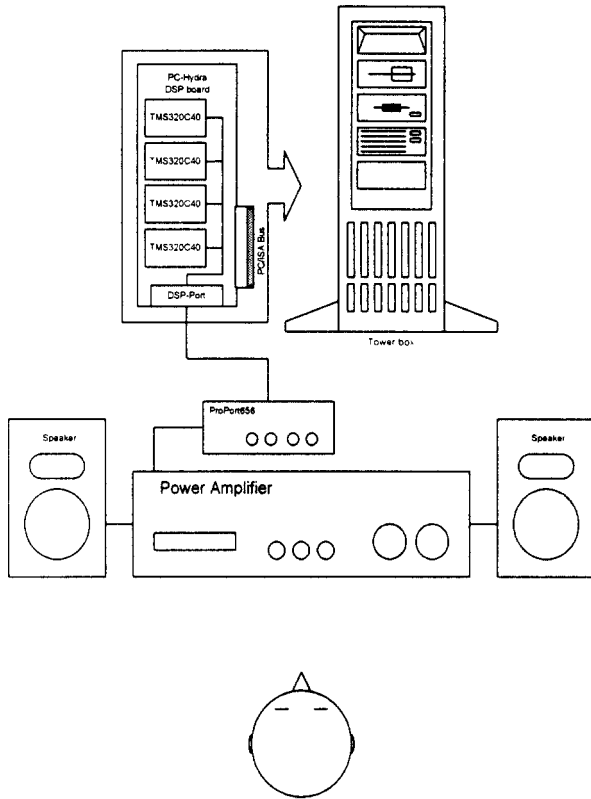


그림 5. 실험 장치도

실험 환경으로, 호스트는 듀얼 펜티엄 PC에 윈도우 NT 4.0을 설치했고, 디지털 신호 처리 보드를 PC/ISA 비스에 장착하였다. 그리고 DSP 컴파일러는 부동 소수점 DSP 컴파일러 V5.0을 사용했으며, 디버깅 소프트웨어는 Code-Composer V3.03을 사용했다. 오디오 입력은 PC의 CD-ROM 드라이브와 사운드 블러스터 32PnP를 통하여 ProPort656으로 입력되며, 트랜스오럴 필터를 통과한 출력은 ProPort656을 통해 파워 앰프(A/V Surround Receiver Pro 700 RDS)를 거쳐 스피커(Boss model 121)로 출력된다. 트랜스오럴 필터의 성능 평가를 위한 입력 binaural 신호는 더미 헤드(Neumann KU-100)를 이용하여 전 방향에서 녹음된 binaural 데이터와 무향실에서 녹음된 음악과 음성을 MIT의 HRTF를 이용하여 실시간으로 변환한 binaural 데이터를 사용했다.

이와 같은 방법으로 청취 실험한 결과 44.1kHz로 입력되는 binaural 신호를 실시간으로 처리하여 결과를 출력했으며, 더미 헤드로 녹음된 binaural 신호와 무향실에서 녹음된 음악을 실시간으로 binaural 처리한 신호

를 입력했을 경우, binaural 신호가 포함하고 있는 방향 정보를 인지할 수 있었다. 즉 전 방향에 대해 음상이 정위되었다. 그러나 정위된 음상은 매우 희미하게 나타났다.

4. 결론

입체 음향을 스테레오 스피커를 통해 청취하면 헤드폰을 통해 청취할 때 보다 현장감이 현저하게 줄어든다. 그것은 두 개의 스피커에서 출력되는 소리가 서로 영향을 주며 두 귀로 청취되기 때문이다. 이러한 현상을 제거하기 위한 방법으로 본 논문에서는 범용 DSP를 기반으로 한 신호 처리 보드를 사용하여 트랜스오럴 필터를 실시간 구현하였다. 실시간 구현된 트랜스오럴 필터를 binaural 신호를 이용하여 청취 실험한 결과 전 방향에 대해 음상 정위가 이루어졌다. 그러나 정위된 음상이 선명하지 못했다. 이와 같이 정위된 음상이 선명하지 못한 것은 역함수의 계산 과정에 오류가 있는 것으로 추정된다.

앞으로 실생활에 적용이 가능한 입체 음향을 구현하기 위해서는 HRTF를 이용한 음상 정위에 대한 연구와 적은 연산으로 크로스 토크를 제거할 수 있는 알고리즘에 대해 연구가 필요하다.

참고문헌

- [1] 김진욱, "실감통신 구현을 위한 3차원 음향 시스템," 석사학위논문 1997
- [2] 안철용, 성평모, "머리전달함수의 Block Convolution을 이용한 3차원 음향합성," 한국음향학회 학술대회 논문집, 1997.7
- [3] 한민수, 강경옥, 강동규, "헤드폰 재생을 위한 3차원 음향 처리 기술," 한국음향학회 학술대회 논문집, 1996.11
- [4] William G. Gardner, "Transaural 3-D audio," MIT Media Laboratory Perceptual Computing Section Technical Report No. 342, July 20, 1995
- [5] Texas Instruments, "TMS320C4x User's Guide," Texas Instruments, 1993