

개선된 머리전달함수를 이용한 3차원 입체음향 성능 개선 연구

A Study on Enhancement of 3D Sound Using Improved HRTFs

구 교 식*, 차 형 태*
(Kyo-Sik Koo*, Hyung-Tai Cha*)

*숭실대학교 전자공학과

(접수일자: 2009년 6월 15일; 수정일자: 2009년 7월 21일; 채택일자: 2009년 8월 3일)

인간의 청각기관은 소리의 방향과 거리인지에 있어서 여러 가지의 정보들을 복합적으로 이용한다. 이러한 양 귀에 들어오는 소리세기 차이, 위상의 차이, 그리고 주파수 스펙트럼의 차이 등의 정보들을 종합적으로 포함하고 있는 것이 머리전달함수(HRTF: Head Related Transfer Function)이다. 2채널 음향재생 시스템에서 이 머리전달함수를 이용하여 다채널 입체음향을 재생하는 방법이 많이 쓰이고 있다. 그러나 머리전달함수의 비개인화적인 특성 때문에 양쪽 귀까지의 거리가 같아지는 혼돈원주 상에서는 앞/뒤 방향 지각에 대한 혼돈을 주게 됨으로써 입체감을 저하시키게 된다. 본 논문에서는 입체음향을 생성하기 위해 머리전달함수를 사용하는 과정에서 발생하는 혼돈원주 문제점을 해결하기 위한 알고리즘을 제안한다. 각 머리전달함수의 주파수 차 및 인간의 청각 특성을 이용하여 각 방향에 해당하는 머리전달함수의 주파수 스펙트럼 특성을 조절하였다. 본 알고리즘의 성능을 평가하기 위하여 음상정위 테스트 및 청감 테스트를 실시하였으며 실험 결과 헤드폰을 기반으로 한 2채널 입체음향 시스템에서 향상된 입체음향을 재생하였다. 더불어 구현된 입체음향의 음질의 열화도 적음을 확인할 수 있었다.

핵심용어: 머리전달함수, 3D 입체음향, 음상정위, 심리음향, 가상현실

투고분야: 음향 신호처리 분야 (1, 2)

To perceive the direction and the distance of a sound, we always use a couple of information, Head Related Transfer Function (HRTF) contains the information that sound arrives from a sound source to the ears of the listener, like differences of level, phase and frequency spectrum, For a reproduction system using 2 channels, we apply HRTF to many algorithms which make 3d sound, But it causes a problem to localize a sound source around a certain places which is called the cone-of-confusion, In this paper, we proposed the new algorithm to reduce the confusion of sound image localization, The difference of frequency spectrum and psychoacoustics theory are used to boost the spectral cue among each directions, To confirm the performance of the algorithm, informal listening tests are carried out, As a result, we can make the improved 3d sound in 2 channel system based on a headphone, Also sound quality of improved 3d sound is much better than conventional methods.

Keywords: Head related transfer function, 3D Sound, Sound localization, Psychoacoustics, Virtual reality

ASK subject classification: Acoustic Signal Processing (1,2)

I. 서론

나날이 발전하는 기술로 인하여 디지털 시대가 도래한 현재, 멀티미디어 기술은 신호처리 기술의 발달과 저렴한 DSP의 등장으로 인하여 우리 사회의 커다란 이슈로 자리매김 하였다.

이중에서도 가상현실 분야는 현재 가장 활발히 연구가 진행되고 있는 분야 중 하나라 할 수 있다. 가상현실이란 인간의 지각계에 투영된 영상이나 전기 음향 등의 인공 매체를 통해서 음향 환경 정보를 전달하여 피험자에 대한 물체나 현상의 존재를 가상적으로 지각시키는 기술이다 [1]. 여러 형태의 가상현실 시스템 중에 Goggle 형의 입체 디스플레이와 손이나 손가락의 움직임을 검출하는 글로브를 장착하여 피험자 자신의 움직임을 컴퓨터에 전달하여 그것에 대응되는 3차원 공간을 생성하고 시각적

책임저자: 차 형 태 (hcha@ssu.ac.kr)
156-743 서울시 동작구 상도5동 511 숭실대학교 전자공학과
(전화: 02-820-0711; 팩스: 02-821-7653)

인 가상 세계를 생성하여, 이 가상 세계를 체험하는 시스템이 점점 주목받기 시작하고 있다. 가상현실을 구현하는 오감에 대한 기술 중 가장 현실감을 느끼게 해주는 기술로는 사운드 기술을 들 수 있다 [12]. 가상현실에서 음향은 피험자에게 3차원 영상에 대응되는 공간적인 음장을 지각시키기 위한 3차원 가상 음향 환경을 제공하며 이는 입체음향이란 말로서 대변할 수 있다. 일반적으로 입체음향이란 모노음의 음의 고저, 음색, 음원의 방향이나 거리감을 조절하여 청자로 하여금 현장감이나 입체감을 가지게 하는 음향이다 [2]. 이것에 의하여 음원의 위치, 주위의 환경, 자신의 위치, 물체의 움직임, 음상의 형태와 표정 등을 얻을 수 있고, 각태일 파티 효과도 얻어진다. 2채널 시스템에서 입체음향을 구현하는 가장 대표적인 방법은 머리전달함수를 이용하는 방법이다. 머리전달함수는 자유 공간 상에서 음원과 귀 간의 경로를 모델링하는 함수로서 음상의 공간화와 관련된 중요한 음향 파라미터이다 [3]. 이는 헤드폰과 같은 양이 재생 환경에서 음원의 방향성을 포함하는 공간감을 표현하는데 널리 사용되고 있다. 일반적으로 머리전달함수는 더미 헤드 또는 청취자에게 양이 마이크로폰을 착용시켜 직접 측정한다. 그러나 모든 사람마다 머리전달함수를 측정하고 모델링 하는 것은 많은 시간과 노력을 요구하는 일이기 때문에 많은 시스템에서는 더미헤드와 같은 표준화 된 모델에 대하여 측정한 자료를 사용한다. 그 결과로 입체음향의 성능이 떨어지는 현상이 발생한다 [2]. 특히 혼돈원추 상에서는 앞/뒤, 상/하 방향을 제대로 인지할 수 없게 된다. 따라서 본 논문에서는 이러한 단점들을 보완하여 2채널 시스템에서 보다 성능이 좋은 입체음향을 구현하기 위한 알고리즘을 제안하고자 한다. 본 논문의 2절에서는 머리전달함수를 이용한 입체음향 생성에 관해 기술하고 제 3절에서는 제안된 알고리즘에 대해 설명한다. 이어 제 4절에서 테스트를 통한 실험 결과를 제시하며 마지막으로 5절에서 결론을 맺는다.

II. 머리전달함수를 이용한 2채널 입체음향 생성

헤드폰을 이용하여 가상음향을 재생하는 경우 사람이 음을 인지하는 요소가 청각기관으로 입사하는 음파에만 의존한다고 가정할 수 있다. 그렇다면 청각기관에서의 음을 직접 녹음하여 그 음압을 원 음장에서의 청취자 청각기관에 녹음된 음압의 진폭과 위상을 같게 재생시켜

준다면 실제의 원음과 동일하게 느껴지게 될 것이다. 이것을 바이노럴 (Binaural) 재생이라고 하며 더미헤드 또는 실제 사람의 귀에서 녹음하고 헤드폰으로 재생하는 방법을 말한다. 주로 헤드폰이나 이어폰을 이용하여 재생하게 되는데 이를 위해서는 원하는 위치에 대한 귀의 응답, 또는 기준 지점에서부터 귀까지의 정보수집이 필수적이다. 인간의 청각 시스템이 음원의 수평면에 대한 음의 방향을 지각할 수 있는 주요인은 Interaural Time Difference (ITD)와 Interaural Intensity Difference (IID)이며 수직면에 대한 음의 고저는 청취자의 정면이나 뒤에서 재생되는 음에 대해 방향을 인지할 수 있는 주 파라메타는 스펙트럼 요소들이라고 알려져 왔다 [2][11]. 더불어 입사파가 내이에 도달하기 까지 몸통, 머리, 외이와의 상호 작용과 직접음의 반사와 회절에 의한 음의 스펙트럼 변화 등 다양한 요인들이 작용하게 된다. 따라서 이와 같은 모든 사항을 고려하여 방향감을 재생하는 것은 현실적으로 불가능하다. 그러나 인간의 청각 시스템은 각 귀에서의 지향성 의존 필터링에 의해 방향을 인지할 수 있으며 이 필터링에 대한 주파수 응답을 머리전달함수라고 한다. 즉 어떠한 특정 점에서의 음향 신호가 귀까지 도달하는 전달 경로의 특성을 머리전달함수라고 말하며 즉 사람이 음원의 위치를 파악하는 중요한 정보를 포함하고 있다. 이런 특성으로 인하여 머리전달함수는 가상음향에서 빼놓을 수 없는 역할을 하며 이를 취득하게 되면 어떠한 음원이라도 머리전달함수와 컨벌루션 연산을 통하여 청취자에게 가상음원의 효과를 줄 수 있다 [4].

$$\begin{aligned} Z_L &= S * HRTF_L(\theta, \phi) \\ Z_R &= S * HRTF_R(\theta, \phi) \end{aligned} \quad (1)$$

여기서 θ 과 ϕ 는 특정한 방위각과 고도각을 의미하며, S와 Z는 모노음원, 생성된 입체음향 신호를 나타내고 있다. 또한 그림 1은 머리전달함수를 이용하여 거리감, 공간감, 방향감을 제어하는 3D 입체음향 시스템 구성도를 보여주고 있다.

이 시스템을 통해 구현된 입체음향을 청취할 때 사람은 3차원 공간상에 존재하는 음상을 지각하게 된다. 그 지각된 음상의 위치 (방향과 거리)를 판단하는 것을 음상 정위 (Localization)라고 한다. 앞에서 언급한 것처럼 일반적으로 음상 정위는 두 귀간의 시간차 (ITD)와 두 귀간의 레벨 (IID)차 등에 의하여 이루어진다 [1][5]. 그러나 그림 2에 나타난 것처럼 혼돈원추 같은 특수한 경우 정확한 방향을 지각할 수 없다.

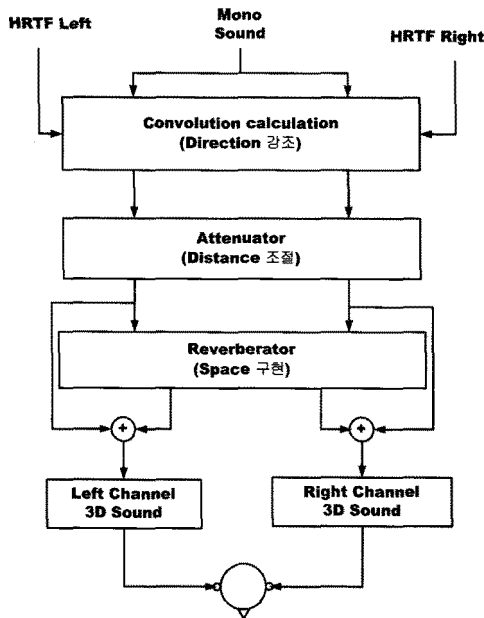


그림 1. 3차원 입체음향 시스템 구성도
Fig. 1. Block diagram of proposed algorithm.

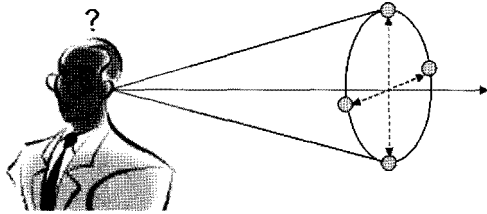


그림 2. 혼돈 원추
Fig. 2. Cone-of-confusion.

혼돈원추란 IID와 ITD를 이용해 상/하나 앞/뒤 방향을 정확하게 판별할 수 없는 영역을 나타낸 것으로서 비 개인화된 HRTF DB로 인해 발생하게 된다. 이를 해결하기 위한 가장 좋은 방법은 청취자 개개인마다 머리전달함수를 측정하여 사용하는 것이지만 이는 매우 어려운 일므로 표준화된 머리전달함수를 개선하여 음상정위의 오차를 최소화시키려는 연구가 활발히 진행되고 있다 [12].

III. 제안한 알고리즘

음상정위에 관련된 많은 연구결과에서 살펴보면 두 귀 사이의 레벨 차이는 저주파 대역보다 1.5 kHz 이상에서 변화가 뚜렷하게 나타나므로 고주파 대역의 신호가 음상정위의 실마리가 된다. 두 귀 사이의 시간차는 머리전달함수의 위상 특성을 나타내는 단서로서 1.5 kHz 이하의 낮은 주파수대역에서 음상정위의 주요 단서가 된다 [5].

더불어 정중면 전체에 걸쳐 올바른 음상 정위를 얻기

위해서는 음원 신호로 5~10 kHz의 주파수 성분이 포함되어 있어야 하며 전 방향의 음상정위에 대해서 콧바퀴가 크게 관여함을 알 수 있다. 이를 통해 음상 방향은 음원 방향에 관계없이 각 주파수 대역에 의존하고 있으며 이와 같이 음상의 방향을 정하는 대역을 방향 결정 대역이라고 한다. 그리고 각 음원 방향에 관계없이 고막에 도달하는 신호의 스펙트럼에서 강조된 주파수 대역과 일치하는 방향 결정 대역의 방향에 음상이 생기게 된다 [1].

이런 특성을 기반으로 하여 바이노럴 시스템에서 방향 지각의 혼돈을 개선하기 위한 기존의 연구는 각 방향에 대한 머리전달함수의 주파수 특성인 스펙트럴 큐를 강조하는 연구가 주를 이루고 있다.

Tan과 gan은 머리전달함수를 5개의 주파수 대역으로 나누어 병렬로 구성된 필터 बैं크를 통과시킴으로서 방향 감을 강조하고자 하였으며 [7] Gupta는 실험을 통해 앞/뒤 방향 지각에 관련한 단서가 인간의 신체 특성 중에서 귀의 돌출된 정도와 관계가 있다고 밝혔다 [8]. 이밖에 Zhang과 Tan 등은 머리전달함수의 주파수 스펙트럼 크기에 따라 생성된 가중치 함수를 적용하고자 하였다 [9]. 최근에는 음상을 정위하고자 하는 방향과 혼돈 방향 머리 전달함수의 주파수 차를 이용하는 방법이 Kim 등에 의해 제안되었다 [6]. 그러나 기존의 연구들은 머리전달함수가 각 방향에 따라 다른 특성을 가지고 있음에도 불구하고 고정 대역만을 개선하거나 다양한 인간의 신체특성을 모두 모델링하기에는 무리가 있다. 또한 단순히 주파수 대역만을 조절함으로써 해당 대역에 주파수 에너지가 과다하게 존재할 경우 정확한 음상의 정위가 불가능하며 음질까지 손상되어 청자에게 불쾌감을 주는 경우가 발생할 수도 있다.

따라서 음상을 정위하고자 하는 방향에 따라 다양하게 스펙트럼 큐를 제어할 수만 있다면 혼돈원추의 문제점을 극복할 수 있다.

이에 본 연구에서는 머리전달함수가 인간의 귀에 미치는 청각 자극 에너지를 계산한 후 이를 바탕으로 하여 머리전달함수를 개선하는 알고리즘을 제안하고자 한다.

제안된 알고리즘의 Block Diagram은 그림 3과 같다.

먼저 음상을 정위시키길 원하는 a 방향의 머리전달함수, HRTF_a 및 이와 혼돈 방향인 b 방향의 머리전달함수, HRTF_b에 푸리에 변환을 적용하여 주파수 영역으로 변환한다.

다음으로 HRTF_a(f)와 HRTF_b(f)의 스펙트럼 차를 고려한다. 이는 심리음향의 특성을 적용하기 위한 전처리 작업으로서 혼돈을 일으키는 방향의 HRTF 쌍의 유사 성

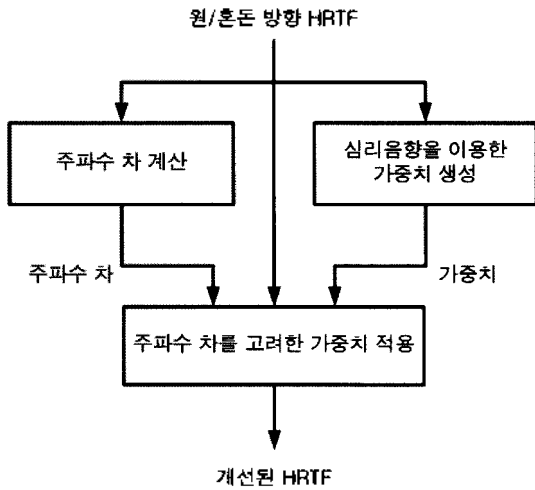


그림 3. 제안된 알고리즘의 Block diagram
Fig. 3. Block diagram of proposed algorithm.

분을 제거하고 해당 방향의 스펙트럼 특성을 확인하기 위함이다. 먼저 식 (2)와 같이 정위시킬 방향 a의 HRTF_a(f)와 앞/뒤 쌍을 이루는 대칭 방향의 HRTF인 HRTF_b(f)와의 크기 차 diff_a(f)를 계산한다.

$$diff_a(f) = |HRTF_a(f)| - |HRTF_b(f)| \quad (2)$$

그림 4는 고도 0°에서 방위각 30°와 150°에 대한 앞 방향의 HRTF 스펙트럼 차를 나타내는 것이다. 1.5 kHz 이하에서는 방위각이 변해도 스펙트럼 차이는 크지 않지만 1.5 kHz 이상의 대역에서는 피크나 노치로 인하여 큰 변화를 나타냄을 확인할 수 있다. 또한 머리전달함수는 음상을 정위시키려는 방향에 따라 우세한 주파수 대역을 가지고 있기 때문에 그 우세한 주파수 대역이 앞/뒤의 방향 지각을 결정하는 스펙트럼 단서가 된다. 따라서 그 특성을 강조할 경우 방향감이 강조될 것이다.

다음으로 a 방향 머리전달함수의 임계 대역에서 주파수 변별력을 고려하여 각 임계 대역에 해당하는 신호들에 대한 임계 대역 에너지, cband_ae(z)를 계산한다. 입력 신호 x(n)에 해당하는 전력 스펙트럼을 x(ω, i)라고 정의할 때, 식 (3)을 통해 기저막에서의 임계 대역 에너지를 계산할 수 있다 [12].

$$X_a(z, i) = a_0(z)X(z, i) = a_0(z) \sum_{\omega_{a,i}}^{\omega_{a,i+1}} (\omega, i) \quad (3)$$

$$0 \leq z \leq Z-1$$

여기서 i는 시간 프레임에서의 인덱스를 의미하며 ω_{a,i}와 ω_{a,i+1}은 전체 임계 대역 Z에서의 임의의 임계대역 z에 대한

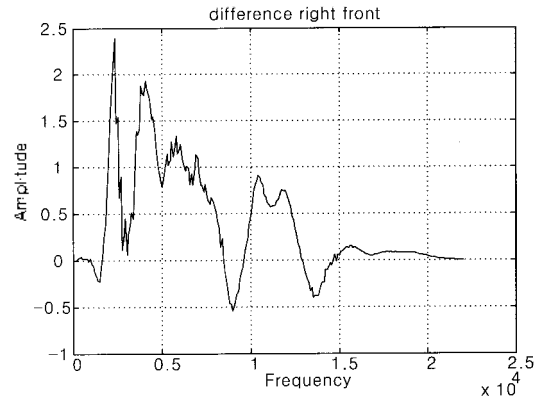


그림 4. 주파수 스펙트럼 차 (고도각 0°, 방위각 30° & 150°)
Fig. 4. Difference of frequency spectrum (Ele. 0°, Azi. 30° & 150°).

저주파 경계와 고주파 경계를 나타낸다. a₀(z)는 외이에서부터 중이까지의 다양한 전달 특성 요소 (transmission factor)들에 의한 주파수 의존 감쇠 특성을 나타내는 감쇠 성분이다.

이어 식 (3)에서 계산된 임계 대역 에너지는 기저막 (basilar membrane)에서의 신호 에너지의 확산 현상을 나타내는 지각적인 에너지 확산 함수 (spreading function)를 통해 권별루션 연산을 수행한다. 이를 통해 기저막에서의 신호 에너지에 대한 지각 배전을 나타내는 자극 에너지 (excitation energy), exci_e(z)를 식 (4)와 같이 표현할 수 있다.

$$X_c(z, i) = \sum_{v=0}^{Z-1} [SF(v, z) X_a(v, i)] \quad (4)$$

$$0 \leq z \leq Z-1$$

확산 함수 SF(v, z)는 다음과 같이 표현되어질 수 있다. 여기서 Δz = z-v는 바크 인덱스로 표현된 임계 대역 간 차이이다.

$$SF(v, z) = 15.81 + 7.5(\Delta z + 0.474) - 17.5 \sqrt{1 + (\Delta z + 0.474)^2} \quad (5)$$

마찬가지로 혼돈을 일으키는 b 방향의 임계 대역 에너지, cband_be(z) 및 청각 자극 에너지, exci_e(z)를 계산한다.

계산된 a와 b 방향의 청각 자극 에너지는 원 머리전달 함수에 적용될 가중치를 계산하기 위하여 청각 자극 에너지 비율, rate_a(z)를 생성한다.

$$rate_a(z) = (exci_{e_a}(z)/exci_{e_b}(z))^{1/2} \quad (6)$$

그림 5는 고도각 0°에서 방위각 30°와 150°에 해당하는 청각 자극 에너지 비율을 보여준다. 추출된 비율은 가중치로서 원 머리전달함수에 적용된다. 그러나 청각 자극 에너지는 전/후 방향 마스크 효과를 고려한다. 따라서 특정 주파수 대역의 차이가 크지 않더라도 옆의 대역에서 강한 에너지가 존재할 경우 주파수 마스크에 의한 영향을 받을 수 있다. 이는 강한 에너지와의 거리가 가까울수록 그 영향은 더 커지게 된다. 따라서 단순히 청각 자극 에너지만 고려할 경우 가중치를 제대로 적용할 수 없게 된다. 이에 우세한 대역만을 강조하기 위하여 앞에서 계산한 주파수 차를 고려하는 부분을 추가시켜 준다. 일정 임계값 이상 주파수 차가 존재하는 부분에는 심리음향을 이용해 생성한 가중치를 적용시키지만 주파수 차가 크지 않은 부분은 원 신호 그대로 두게 된다. 더불어 청각 에너지 비율이 1보다 작은 경우는 각 방향에 대한 실험 결과 혼돈 방향의 주파수가 큰 대역과 유사하게 생성되는 경우가 대부분이기 때문에 그대로 가중치를 적용시켜 준다. 이는 방향감을 강조시키기 위한 부분이기도 하지만 최소한의 수정을 통해 음질의 변화를 최소화하려는 효과도 포함하고 있다.

그림 6은 임계 대역에서 청각 에너지를 이용하여 생성된 가중치와 주파수 차를 나타내고 있다. 각 임계 대역에 해당하는 주파수 스펙트럼 차를 분석하여 보면 23번째 임계대역 (9.5 kHz~12 kHz)은 앞방향이 우세한 대역이다. 그러나 8 kHz~10 kHz 근처에 존재하는 큰 에너지로 인하여 영향을 받음으로서 청각 에너지 비에서는 뒷방향이 우세한 것으로 판단되었다. 더불어 24번째 임계 대역 (12 kHz~15 kHz)에서는 앞 방향이 우세한 대역 및 뒷방향이 우세한 대역이 혼재되어 있지만 청각 자극 에너지 비율은 뒷방향이 우세한 것으로 확인할 수 있다. 따라서 단순히 가중치만을 적용하게 된다면 혼돈 방향의 특성까지 같이 증폭시켜 주게 되어 정위감이 떨어지는 문제가 발생할 수도 있다. 그러나 본 논문에서는 주파수 차를 같이 고려해 줌으로서 음상을 정위하고자 하는 방향의 특성만을 강조해 주게 됨으로 그 성능을 향상시킬 수 있다.

$$w_a(f_c, i) = \begin{cases} HRTF_a(f_c, i) \times w_a(f_c, i) \\ 1, (diff_a(f_c, i) < 0.2) \text{ and } (rate_a(z) > 1) \\ rate_a(f_c, i), \text{ else} \end{cases} \quad (7)$$

여기서, f_c 는 각 임계 대역에 해당하는 주파수 범위를 의미한다.

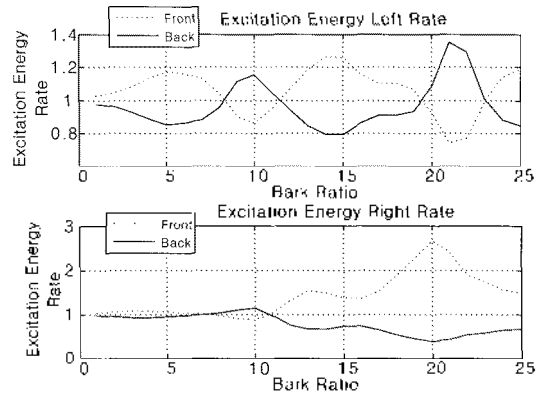


그림 5. 청각 자극 에너지 비율 (고도각 0°, 방위각 30° & 150°)
Fig. 5. Excitation energy rate (Ele. 0°, Azi. 30° & 150°).

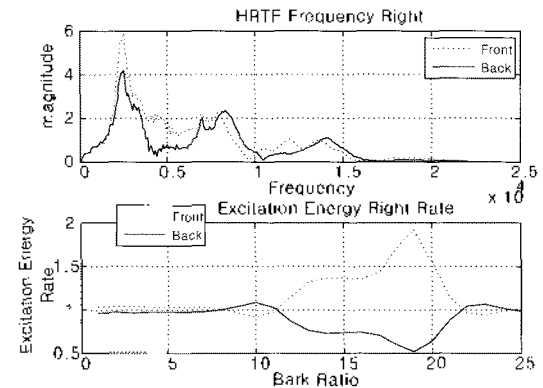


그림 6. 주파수 스펙트럼과 청각 자극 에너지 비율 (고도각 0°, 방위각 30° & 150°)
Fig. 6. frequency difference & Excitation energy rate (Ele. 0°, Azi. 30° & 150°).

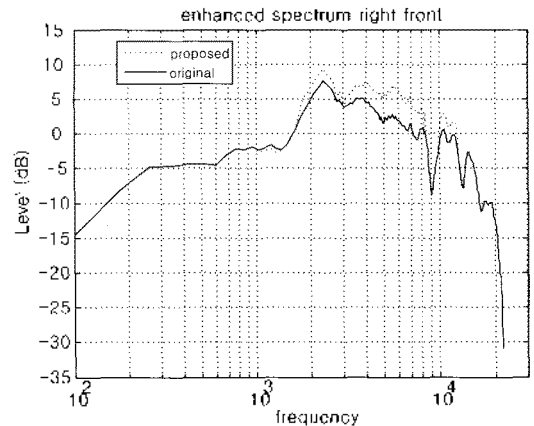


그림 7. 원 HRTF & 수정된 HRTF (고도각 0°, 방위각 30°)
Fig. 7. original HRTF & enhanced HRTF (Azi. 0°, Ele. 30°)

그림 7은 최종적으로 개선된 머리전달함수의 주파수 스펙트럼을 나타낸다. 방향감이 개선된 머리전달함수는 모노음에 컨벌루션 연산을 통해 적용되어 음상을 원하는 위치에 정위시키게 된다. 그 결과로 인간은 소리의 위치를 정확하게 지각할 수 있으므로 게임이나 시뮬레이터와

같은 가상현실 공간 안에서 보다 더 현실감을 느낄 수 있을 것이다.

IV. 실험 및 결과고찰

본 논문은 머리전달함수를 이용하여 보다 현실적인 입체음향을 생성하는 과정에서 발생하는 혼돈원추의 문제를 해결하기 위한 알고리즘을 제안한다. 제안된 알고리즘의 성능을 확인하기 위하여 헤드폰 환경에서 음상정위 테스트 및 청감 테스트를 수행하였다. 테스트는 일반 생활 잡음이 존재하는 사무실 환경에서 수행되었으며 사용된 사운드 신호들은 오디오 CD 및 영화 DVD에서 추출된 44.1 kHz, 16 bits/sample 사운드 신호를 이용하였다. 사용된 머리전달함수로는 웹에 공개되어 있는 MIT의 Media Lab.에서 KEMAR dummy head를 사용하여 측정된 것을 사용하였다. 테스트 장비로는 청취자의 귀에 밀착되는 헤드폰 (audio-technica ATH-ES7), Pentium computer (Intel core2duo E8500)를 사용하였다. 그리고 재생시스템과 피 실험자 간의 거리를 두고 랜덤한 순서로 입체음향을 들려주어 청취자가 음원의 종류를 사전에 알 수 없도록 하였다.

비교대상은 원 머리전달함수, 필터뱅크 방식 및 제안한 방식으로서 방위각 0°/180°, 30°/150°, 45°/135°, 60°/120°의 네 방향에 대해서 테스트를 실시하였다. 입력 사운드 신호로는 표 2와 같이 Classic, Vocal, Helicopter sound의 3가지를 사용하였다.

그림 8은 시뮬레이션에 사용된 음원 중 Classic sound

를 보여주고 있다. Classic sound는 각 약기의 특성에 따라서 저주파에서 고주파 대역까지 전 대역에 걸쳐서 에너지가 존재하고 있다. 따라서 각 방향에 따른 음상정위의 주파수 단서들을 고루 포함하고 있다. 이외에 Vocal sound는 인간의 음성 신호를 3차원 공간상에 정위시키는 상황에서의 성능을 확인하기 위해 사용하였다. 마지막으로 저주파 대역에의 에너지가 큰 신호에서의 성능 테스트를 위해 Helicopter sound를 사용하였다.

기존의 머리전달함수를 사용하여 생성된 입체음향은 그림 9로 확인할 수 있다. 서문에서 제기한 것처럼 기존의 머리전달함수를 사용한 방법은 머리전달함수의 비개인화 문제 때문에 혼돈원추의 문제점을 가지고 있다. 따라서 이를 해결하기 위해 많은 연구가 진행되고 있다.

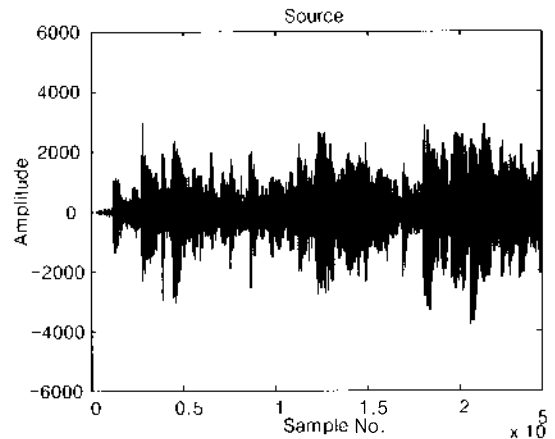


그림 8. 사용된 사운드 음원 - 클래식
Fig. 8. Used sound source for the simulation - Classic.

표 1. 사용된 머리전달함수
Table 1. HRTF DB for simulation.

샘플 수	512 (full) / 128 (compact)		
샘플링 주파수	44.1 kHz		
샘플당 비트수	16 bits/sample		
수평각	0° ~ 360°		
고도각	-40° ~ 90°		
측정 대상	KEMAR dummy head		

표 2. 사용된 음원
Table 2. Sound sources for simulation.

Sample No.	1	2	3
이름	Classic	Vocal	Helicopter
샘플링 주파수	44.1 kHz		
샘플당 비트수	16 bits/sample		

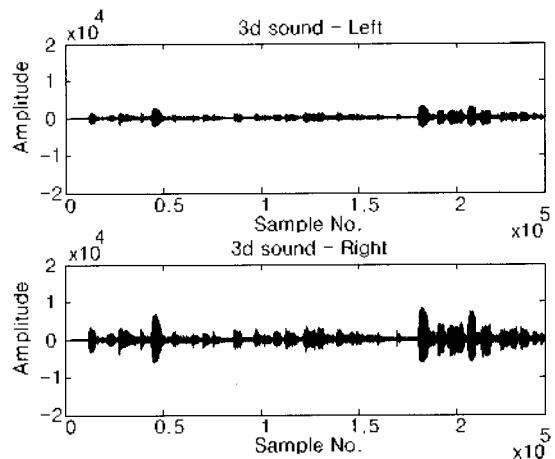


그림 9. 원 HRTF를 이용한 입체음향 생성 (고도각 0°, 방위각 30°)
Fig. 9. Creating 3D sound using original HRTF (Azi. 0°, Ele. 30°).

4.1. 개선된 입체음향의 생성

본 단원은 각 방법을 통해 개선된 머리전달함수를 사용해 새롭게 생성된 입체음향 사운드를 보여주고 있다. 첫 번째로 혼돈원추 상에서의 음상정위감을 개선하기 위하여 사용된 필터뱅크 알고리즘을 고려한다. 필터뱅크 알고리즘은 앞/뒤 방향지각에 관련된 특징이 특정 주파수 대역에 있다는 심리음향의 연구결과를 바탕으로 한 것이다 [7]. 그림 10과 같이 머리전달함수를 5개의 대역으로 나누어 각 방향에 따라 저주파 대역은 8 dB, 고주파 대역은 12 dB를 증폭/감쇄시킨 후 원 머리전달함수와 컨벌루션을 시행하여 개선된 머리전달함수를 획득하게 된다.

그림 11은 필터뱅크 알고리즘을 적용하여 개선된 입체음향이다.

생성된 사운드를 살펴보면 오른쪽 채널이 심하게 증폭됨을 알 수 있다. 이는 5개로 나뉜 머리전달함수의 주파수 대역에서 증폭되는 대역에 사운드 음원의 주파수 에너지가 강하게 존재할 경우 과도하게 강조되어 ITD를 손상 시킴으로서 정확한 방향감을 상실하게 되고 음질 또한 손상되게 된다.

그림 12는 주파수 차를 이용하여 생성된 입체음향이다 [6]. 이 방법은 각 방향에 따라 주파수가 우세한 대역을 강조함으로써 방향감을 정확하게 해준다는 장점이 있다. 그러나 인간의 청각특성을 적용하여 볼 때 강조된 주파수 에너지로 인한 마스킹 현상이 발생한다. 이에 옆의 주파수 대역까지 영향을 끼치게 되므로 ITD가 크게 달라지게 된다. 더불어 옆 주파수 대역이 혼돈 방향의 방향 결정 내역일 경우 인간의 청각에 미치는 영향이 달라지므로 방향감이 떨어지게 된다.

그러나 제안된 알고리즘을 이용하여 생성된 입체음향은 머리전달함수의 각 주파수 에너지가 인간의 청각에 미치는 영향을 고려한다. 그 결과로 각 임계 대역마다 효과적으로 가중치를 생성 및 적용할 수 있으므로 보다 현실적인 느낌을 줄 수 있다. 더불어 주파수 차를 이용하여 가능한 음질의 왜곡을 줄일 수 있다는 장점도 가지고 있다.

4.2. 음상정위 테스트

본 논문에서 제안된 알고리즘의 성능을 평가하기 위한 첫 번째 테스트로 청취자가 개선된 음상의 방향을 지각하는지에 대한 평가를 수행하였다. 앞에서 말한 3가지 방법에 따라 생성된 앞/뒤 방향의 입체음향 3가지를 들려준 후 앞/뒤 방향에 대해 인지하는 정확도를 측정하였으며 20대의 일반 남녀 10명을 대상으로 하였다.

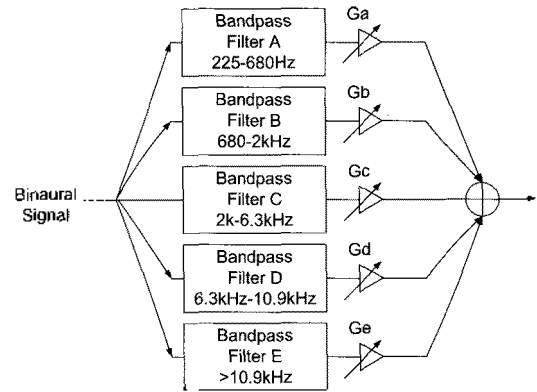


그림 10. 필터뱅크 알고리즘
Fig. 10. Filterbank algorithm.

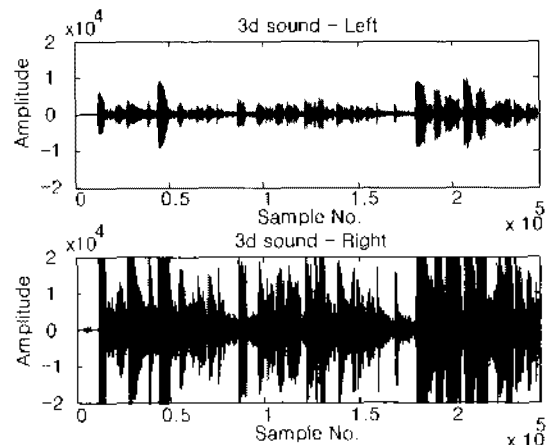


그림 11. 필터뱅크 방법을 이용한 입체음향 생성 (고도각 0°, 방위각 30°)
Fig. 11. Creating 3D sound using filterbank method (Azi. 0°, Ele. 30°).

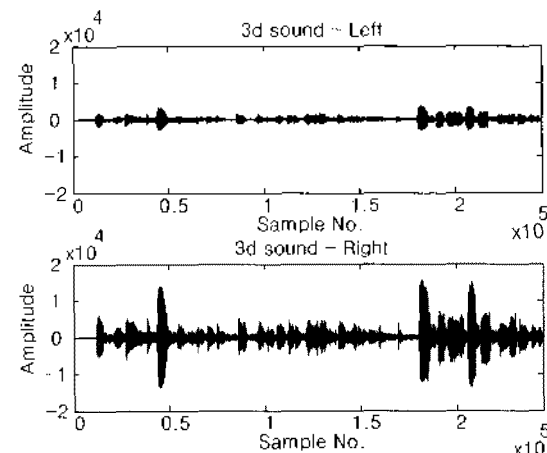


그림 12. 주파수 차를 이용한 입체음향 생성 (고도각 0°, 방위각 30°)
Fig. 12. Creating 3D sound using differences of frequency (Azi. 0°, Ele. 30°).

그림 14는 3가지 방법을 이용하여 생성된 입체음향에 대한 음상정위 테스트 결과를 나타내고 있다. 표 2에서

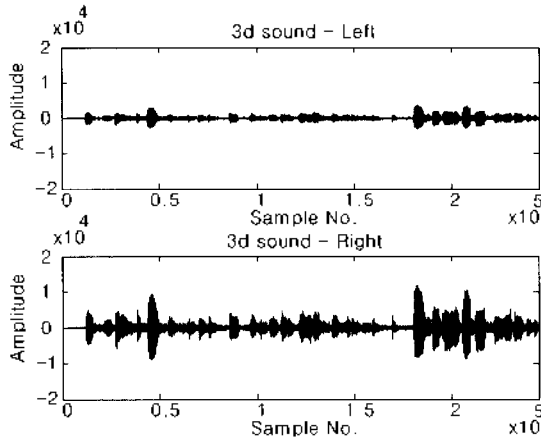


그림 13. 제안된 알고리즘을 이용한 입체음향 생성 (고도각 0°, 방위각 30°)
 Fig. 13. Creating 3D sound using proposed algorithm (Azi. 0°, Ele. 30°).

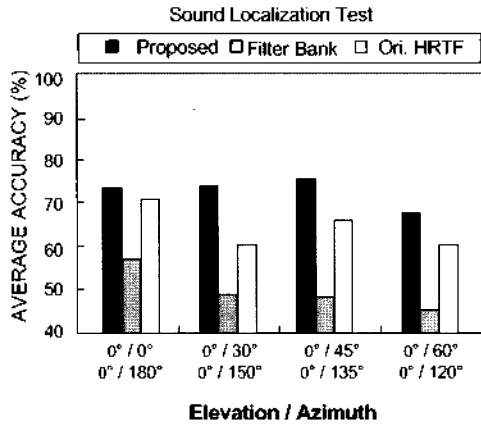


그림 14. 음상정위 테스트 결과
 Fig. 14. Sound localization test result.

제시한 음원에 대한 테스트를 수행한 후 각 혼돈방향에 따른 평균값을 제시하였다. 필터뱅크 방식을 이용하여 생성된 입체 음향은 음상을 정위시키려는 각도가 커질수록 좋지 않은 결과를 보여주고 있다. 테스트 결과에서 확인할 수 있듯이 제안된 방법이 기존의 방법에 비해서 성능이 개선되었음을 알 수 있다. 이는 제안된 알고리즘을 이용하여 원 머리전달함수에서 인간의 청각 기관에 끼치는 영향이 큰 대역을 조절함으로써 결과적으로 방향감을 강조할 수 있기 때문이다.

4.3. 청감 테스트

성능을 평가하기 위한 두 번째 테스트로는 제안된 알고리즘을 통해 생성된 입체음향의 음질 변화 정도를 판단하기 위해 청감 테스트를 실시하였다. 평가 기준으로는 CCIR Recommendation에서 음질 평가를 위해 채택한 방법인 Degradation Category Rating (DCR) Method를 사

표 3. 청감 테스트 점수

Table 3. Degradation Category Scale.

5	Degradation is inaudible.
4	Degradation is audible but not annoying.
3	Degradation is slightly annoying.
2	Degradation is annoying.
1	Degradation is very annoying.

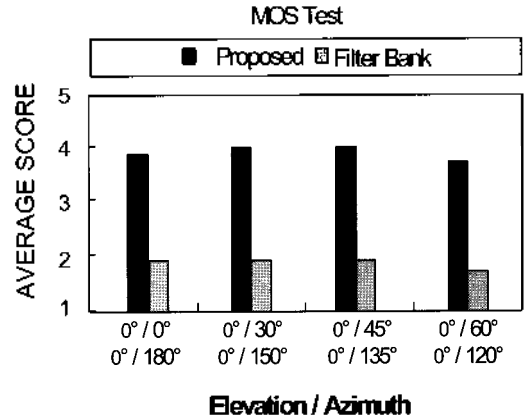


그림 15. 청감 테스트 결과
 Fig. 15. MOS test result.

용하였다. DCR Method는 표 3과 같은 Degradation Category Scale를 통하여 수정된 사운드 신호의 음질이 원 사운드 신호의 음질과 관련하여 얼마나 충실한지를 비교한다 [10].

그림 15는 테스트가 수행된 4가지 방향에 대한 청감 테스트 결과를 나타내고 있다. 음상정위 테스트와 마찬가지로 각 혼돈방향에 대한 평균값을 제시하였다.

테스트 결과를 보면 기존의 필터뱅크 방식을 통해 생성된 입체음향은 음질의 열화가 심함을 알 수 있다. 이는 방위각이 정면에서 좌측 또는 우측으로 감에 따라 머리전달함수의 에너지가 증가함을 무시하고 고정된 값으로 에너지를 조절한다. 따라서 과도하게 증폭/감쇄 되는 현상이 발생하게 된다. 그 결과로 음질이 심하게 손상되어 청자에게 피로감 및 불쾌감을 주게 된다. 특히 저자의 기존 논문과의 결과를 비교해 보면 본 논문에서 수행된 필터뱅크 방식의 결과가 좋지 않게 나오고 있다 [2]. 이는 서로 다른 음원을 시뮬레이션에 적용했기 때문이다. 즉 필터뱅크 방식은 음원에 따라서 그 성능이 크게 차이가 나게 되어 음원에 종속적임을 확인할 수 있다.

그러나 본 논문에서 제안하는 방법은 머리전달함수가 인간의 청각에 미치는 영향을 이용함으로써 음상을 정위하려는 방향이 변함에 따라 능동적으로 가중치를 생성할 수 있게 된다. 더불어 혼돈 방향 간의 주파수 차를 고려함

으로서 효율적인 정위감 구현이 가능할 뿐 아니라 음질의 변화도 최소화 할 수 있다. 그 결과로 청자에게 원하는 방향에 음상이 존재함을 지각시킬 수 있게 된다.

V. 결론

본 논문에서는 헤드폰 재생 시스템에서 비 개인화 된 머리전달함수를 이용하여 입체음향을 생성할 때 발생하는 음상정위의 혼돈 문제를 개선하기 위한 알고리즘을 제안한다. 머리전달함수가 인간의 청각에 미치는 영향을 고려하여 가중치를 생성하고 이를 원 머리전달함수에 적용한다. 이때 보다 정밀한 음상정위를 위하여 혼돈방향 두 머리전달함수의 주파수 차를 고려하게 된다. 제안된 방식을 통해 생성된 입체음향은 음상정위 및 청감 테스트에서 개선된 결과를 나타냄을 확인할 수 있었다. 또한 기존의 개선 방법과 비교할 때 음질의 손상이 적음을 확인할 수 있었다. 그러나 본 알고리즘은 헤드폰 환경이라는 상황에서만 한정되어 있고 청취자에 따라 주관적인 결과 들 나타내고 있다. 따라서 모든 시스템에서 성능을 나타 내기에는 한계가 있다. 그러므로 향후 진행될 연구 과제는 스피커를 이용하는 재생 시스템에서 보다 현실감 있는 청각모델링 구현에 대한 연구가 진행될 예정이다.

감사의 글

본 연구는 숭실대학교 교내 연구비 지원으로 이루어졌음.

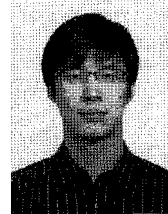
참고 문헌

1. 강성훈, 강경욱, *입체음향*, 기진연구사, 1997.
2. 구교식, 차형태, "스테레오 시스템을 위한 머리전달함수의 개선," *한국음향학회지* 27권, 4호, 207-214쪽, 2008.
3. 이기승, 이석필, "원형 머리 모델을 이용한 머리 전달 함수의 보 간," *한국음향학회지* 27권, 7호, 333-341쪽, 2008.
4. 박장식, 김현태, "합성 스테레오 방식 3차원의 입체음향의 실시간 구현을 위한 머리전달 함수의 IIR 필터 설계," *한국콘텐츠학회지*, 5권, 6호, 74-86쪽, 2005.
5. Durand R, Begault, *3-D Sound for Virtual Reality and Multimedia*, NASA, 2000.
6. 김경훈, 김사호, 배건성, 최승인, 박만호, "헤드폰 기반의 입체음향 생성에서 앞/뒤 음상정위 특성 개선," *한국통신학회지* 29권 8호, 1142-1148쪽, 2004.
7. Chong-Jin Tan, Woon-Seng Gan, "User-defined spectral manipulation of HRTF for improved localisation in 3D sound

- systems," *Electronics letters*, Vol. 34 No. 25, pp. 2387-2389, 1998.
8. Navarun Gupta, Armando Barreto and Carlos Ordonez, "Spectral Modification of head-related transfer functions for improved virtual sound spatialization," *ICASSP '02*, Vol. 2, pp. 1953-1956, 2002.
9. Ming Zhang, Kah-Chye Tan and M.H. Er, "A Refined Algorithm of 3D Sound Synthesis," *ICSP98*, pp.1408-1411, Oct. 1998.
10. ITU-T P.800, *Methods for subjective determination of transmission quality*, Aug, 1996.
11. 김학윤, 김희석, "3차원 음상제어 알고리즘," *산업과학연구*, 15권, 2호, 1-6쪽, 1997.
12. 구교식, 차형태, "머리전달함수의 심리음향적 특성을 이용한 전/후 음상정위 특성 개선," *한국방송공학지*, 11권, 4호, 448-457쪽, 2006.

저자 약력

• 구 교 식 (Kyo-Sik Koo)



2005년 2월: 숭실대학교 정보통신전자공학부 (공학사)
 2007년 2월: 숭실대학교 전자공학학 (공학석사)
 2007년 2월~현재: 숭실대학교 전자공학과 박사 과정 재학
 ※주관심분야: 오디오 및 음성 신호처리, 통신 신호 처리

• 차 형 태 (Hyung-Tai Cha)



1993년: The University of Pittsburgh (공학박사)
 1993년~1996년: 삼성전자 신호처리 연구소 선임 연구원
 1996년~현재: 숭실대학교 정보통신전자공학부 부 교수
 2001년~현재: 한국 신호 및 시스템학회 이사
 2009~ 현재: 한국 방송공학회 이사
 ※주관심분야: Multimedia Systems and Applications Audio and Video Signal Processing, Communication System, ASIC Implementation of Digital System