

특집논문-06-11-4-07

## 머리전달함수의 심리음향적 특성을 이용한 전/후 음상정위 특성 개선

구교식<sup>a)</sup>, 차형태<sup>a)†</sup>

## Improvement of front/back Sound Localization Characteristics using Psychoacoustics of Head Related Transfer Function

Kyo-Sik Koo<sup>a)</sup>, and Hyung-Tai Cha<sup>a)†</sup>

## 요약

일반적으로 3D 입체음향을 생성하기 위해서는 소리가 음원으로부터 청자의 두 귀에 이르는 정보를 포함하고 있는 머리전달함수를 사용한다. 그러나 HRTF의 비개인화적인 특성 때문에 양쪽 귀까지의 거리가 같아지는 혼돈원추상에서는 앞/뒤 방향지각에 대한 혼돈을 주게 됨으로써 입체감을 저하시키게 된다. 이에 본 논문에서는 인간의 청각특성에 따른 여기에너지를 이용하여 앞/뒤 음상정위의 특성을 개선하는 방법을 제안하고자 한다. 서로 대칭되는 각 HRTF의 여기 에너지를 계산하고 각 bark 대역에 따른 에너지의 비율을 추출하는 방식으로써 각 HRTF의 스펙트럼 특성을 부각시키는 방법을 제안하였으며 청감테스트를 통해서 제안한 방식이 기존 방법보다 방향감을 개선시킴을 확인할 수 있었다.

## Abstract

HRTF DB, including the information of the sounds which is arrived to our ears, is generally used to make a 3D sound. But it can decline some three-dimensional effects by the confusion between front and back directions due to the non-individual HRTF depending on each listener. In this paper, we propose a new method to use psychoacoustic theory that reduces the confusion of sound image localization. And we make use of an excitation energy by the sense of hearing. This method is brought HRTF spectrum characteristics into relief to draw out the energy ratio about the bark band. Informal listening tests show that the proposed method improves the front-back sound localization characteristics much better than the conventional methods.

Keywords : HRTF, 3D Audio, Psychoacoustic

## I. 서론

멀티미디어 기술이 발전함에 따라서 가상현실이라는 단어가 큰 이슈가 되고 있다. 가상현실이란 어떤 특정한 환경

· 상황을 컴퓨터를 이용하여 시뮬레이션 함으로써 그것을 사용하는 사람이 재현된 공간 안에 실제로 존재하는 것처럼 만들어 주는 일을 뜻한다<sup>[10]</sup>. 구체적인 예로서는 비행기 조종 훈련이나 수술 실습부터 우리가 실제로는 겪을 수 없는 원자로 내에서의 작업이나 깊은 심해, 우주 저편에서의 활동 등 다양하며 이는 인간의 오감 중에서 시각, 청각, 촉각에 해당하는 그래픽, 사운드 및 운동감 기술의 실시간적인 상호작용으로 구현되어질 수 있다. 즉 인간의 오감을 만족

a) 숭실대학교 전자공학과, 멀티미디어시스템 연구실  
Soongsil University, School of Electronics Engineering, Multimedia System Lab

† 교신저자 : 차형태(hcha@ssu.ac.kr)

\* 본 연구는 숭실대학교 교내연구비 지원으로 이루어졌음.

시키는 멀티모달기술이라 할 수 있으며 각 기술 중에서 사운드 기술은 가상현실 구현에서 가장 현실감을 느끼게 해주는 부분으로 3차원 입체음향이라는 말로서 대변할 수 있다.

일반적으로 입체음향이란 모노음에 음의 고저, 음색뿐만 아니라 방향감 및 거리감까지 포함시켜 가상의 음원을 3차원 공간상에 임의로 위치시키는 것을 의미한다<sup>[1]</sup>. 이 때 인간이 공간감과 거리감을 인지하는 주된 단서로는 음원으로부터 소리가 두 귀에 도달할 때 생겨나는 시간, 레벨 및 스펙트럼의 차 등이 있다. 더불어 머리, 몸통, 컷바퀴 등의 청자의 신체 특성에 따른 소리의 변화나 공간의 반사면에 따른 음의 회절 및 반사, 왜곡 등도 포함될 수 있다. 이처럼 소리가 각 방향에서 인간의 두 귀에 도달하는 정보를 포함하고 있는 것을 머리전달함수(HRTF: Head Related Transfer Function) 이라고 하며 모노 음원과의 컨벌루션을 통하여 원하는 방향에 음상을 정위시키고 거리감까지 조절할 수 있게 된다<sup>[1-4][6-11]</sup>. 그러나 청자들의 신체적인 특징이나 여러 요인들이 개인마다 특성이 고유하므로 비개인화된 HRTF를 사용하게 된다면 청취자는 상/하, 앞/뒤 방향지각에 혼돈을 줄 수 있다<sup>[7]</sup>. 특히 그림 1과 같이 음원과 청취자의 두 귀 사이의 거리가 같은 혼돈원추상에서는 시간이나 레벨의 차로서는 각 방향의 차이를 구분할 수가 없게 된다. 이를 해결하기 위한 가장 좋은 방법은 각각 개인에 대해 HRTF를 측정하여 이용하는 것이지만 현실적으로 불가능하다.

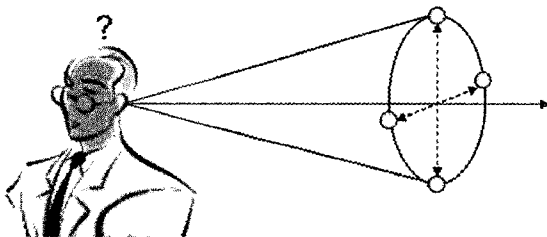


그림 1. 혼돈 원추  
Fig. 1. Cone of confusion

따라서 본 논문에서는 앞/뒤 음상정위의 혼돈을 개선하기 위해 심리음향적인 특성을 이용하는 알고리즘을 제안하고자 한다.

논문의 구성은 다음과 같다. 먼저 II장에서 HRTF를 이용한 입체음향 구현에 대하여 설명하고 III장에서 심리음향에

대해 설명한다. 이어서 IV장과 V장에서 제안한 알고리즘에 대한 설명 및 시뮬레이션에 관한 결과를 제시한다. 마지막으로 VI장에서 결론을 맺는다.

## II. HRTF를 이용한 입체음향 구현

### 1. HRTF (Head Related Transfer Function)

인간은 두 귀를 이용하여 음원의 위치를 지각할 수 있다. 단순히 소리를 듣는 것만으로도 방향을 판별할 수 있는 단서로는 음원으로부터 발생한 소리가 두 귀에 도달하는데 걸리는 시간차(ITD: Interaural Time Difference) 레벨차(ILD: Interaural Level Difference) 및 위상차(IPD: Interaural Phase difference) 등을 들 수 있다. 더불어 인간의 신체적인 특징이나 공기 중에서의 음의 진행, 반사에 의한 특성 변화 등도 방향 지각을 위한 요인에 포함될 수 있다.<sup>[1-4][6-11]</sup>

이런 특성을 종합적으로 포함하고 있는 것을 머리전달함수(HRTF: Head Related Transfer Function) 이라고 하며 시간 영역에서의 임펄스 응답인 HRIR(Head Related Impulse Response)의 형태로 제공된다. HRTF를 이용하여 입체음향을 구현시 ITD는 방위 변화에 대한 입체 음상의 단서를 제공하고 ILD는 입체 음상에 대한 효과를 더욱 증가시킬 수 있는 방법으로 활용할 수 있다.

본 논문에서는 MIT Media Lab.에서 KEMAR 더미 헤드를 사용하여 측정한 HRTF DB를 사용하였으며 그림 2는 고도 0°, 방위각 60°에 해당하는 HRIR DB를 나타낸다.

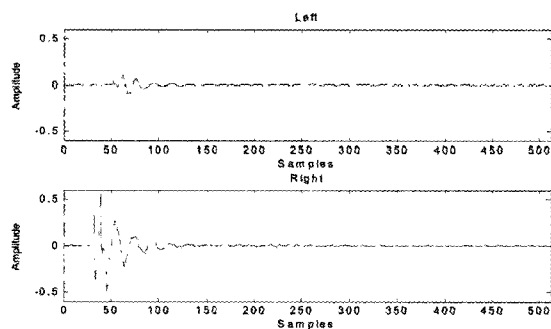


그림 2. 60°에서의 HRIR DB  
Fig. 2. HRIR DB of azimuth 60°

### 2. 입체음향 생성

3차원 입체음향 생성의 목적은 음색의 왜곡이나 음질저하 없이 헤드폰이나 스피커를 기반으로 한 시스템에서도 입체감을 느낄 수 있도록 하는 것이다. 5.1ch나 7.1ch 등의 다채널 재생 시스템에서 입체음향은 서라운드 출력 채널을 증가시킴으로 인해 그 효과를 증가시킬 수 있으나 설치비용 및 복잡성의 증가와 공간적인 제약이 뒤따른다는 단점이 있다. 따라서 최근에는 HRTF를 이용한 방법이 이슈가 되고 있다.

일반적으로 입체음향을 생성하기 위한 기술로는 음상정위 기술 및 음장재현기술 등을 들 수 있다. 음상정위 기술이란 가상의 음원을 공간상의 특정한 위치에 위치시키기 위하여 실측된 HRTF를 이용하여 모노음에 방향 및 거리감을 부여하는 기술이다. 입체음향을 생성하는 방법 중에 가장 대표적인 방법으로 모노음과 HRTF와의 컨벌루션을 수행한다. 식1은 컨벌루션 연산에 관한 식을 보여준다.

$$y[n] = \sum_{i=0}^{i=N-1} x[n-i] \times h[i] \quad (1)$$

여기서  $x[n]$ 은 입력 신호,  $h[n]$ 은 머리전달함수,  $y[n]$ 은 출력신호를 나타낸다.

최종적으로 구현된 입체음향은 스피커와 헤드폰에서 재생되는데 일반 스테레오 음상은 단순히 청취자의 앞면에서 지각되지만 이 기술을 이용하면 그림 3과 같이 원하는 위치에서 음상을 지각하게 할 수 있다<sup>[10]</sup>.

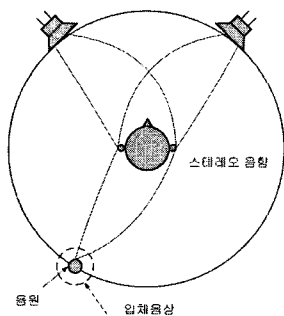


그림 3. 입체음상정위  
Fig. 3. 3D Sound ocalization

음장재현기술이란 잔향을 인위적으로 제어하여 특정 실내에 음원이 존재하는 것과 같은 효과를 구현하는 기술이다<sup>[13]</sup>. 잔향은 음의 재생이 끝난 후에도 반사면으로부터 반사가 계속되어 울리는 음으로 공간의 크기나 구조, 반사면의 재질 등에 의해 다른 특성을 나타내므로 잔향을 이용하여 공간감 및 거리감을 생성할 수 있다.

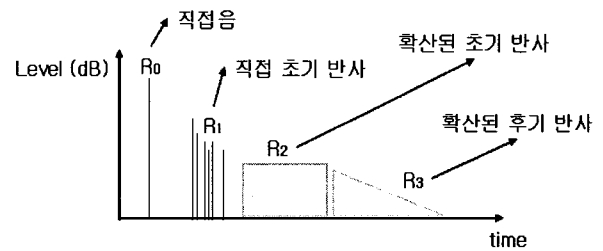


그림 4. 잔향  
Fig. 4. Reverberation

### III. 심리음향 (psychoacoustics)

일반적인 음성 및 오디오 신호와 같은 음향 신호에 있어서 수신기는 사람의 귀가 되고 사운드에 대한 사람의 지각은 귀의 마스킹 현상 등의 영향을 받게 된다. 심리음향 분야는 이러한 청각적인 특성, 특히 내이에서의 시간 또는 주파수 성분들에 대한 해석 능력 등을 이해하는 분야로 임계 대역 주파수 분석 등을 행하게 된다. 내이의 기저막에서 소리의 주파수 분석이 이루어진다. 기저막을 길게 펼쳐보면 하나의 긴 관으로 되어 있는데 관의 각 위치에 따라 공진 주파수가 달라서 각각의 주파수 성분에 따라 기저막의 특정 위치에서 자극에 대한 반응을 보이게 된다. 일반적으로 임의의 두 개의 순음 성분의 주파수 차이를 천천히 변화시킬 때, 그러한 변화의 차를 청자가 지각하게 되는 순간의 주파수 차이 폭을 임계 대역(critical band)라고 하며 이러한 임계 대역은 바크 율(Bark ratio)로 표현한다<sup>[11]</sup>. 심리 음향의 모든 연산은 이러한 바크 스케일(Bark scale)을 기준으로 이루어지게 된다. 표 1에서는 임계 대역과 해당하는 중심 주파수, 그리고 임계 대역의 대역폭을 나타내고 있다.

표 1. 임계 대역  
Table 1. Critical Band

Critical band [Bark]	Center frequency [Hz]	Critical Bandwidth [Hz]
1	50	-100
2	150	100-200
3	250	200-300
4	350	300-400
5	450	400-510
6	570	510-630
7	700	630-770
8	840	770-920
9	1000	920-1080
10	1175	1080-1270
11	1370	1270-1480
12	1600	1480-1720
13	1850	1720-2000
14	2150	2000-2320
15	2500	2320-2700
16	2900	2700-3150
17	3400	3150-3700
18	4000	3700-4400
19	4800	4400-5300
20	5800	5300-6400
21	7000	6400-7700
22	8500	7700-9500
23	10500	9500-12000
24	13500	12000-15500
25	19500	15500-

각 임계 대역에서 주파수 변별력을 고려하여 각 대역에 해당하는 신호들에 대한 임계대역 에너지와 마스킹 특성을 고려한 청각 자극 에너지를 계산할 수 있다<sup>[5]</sup>.  $x(n)$ 이 입력 신호일 때 어떤 임계 대역에 대한 전력 스펙트럼을  $X(\omega, i)$ 라고 정의할 수 있다. 이 때  $i$ 는 시간 프레임에서의 인덱스를 의미하며  $X(\omega, i)$ 로 인한 기저막에서의 임계 대역 에너지는 식 2로 표현 가능하다.

$$X_a(z, i) = a_0(z)X(z, i) = a_0(z) \sum_{\omega_z}^{\omega_{zh}} X(\omega, i) \quad (2)$$

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

여기서  $\omega_{zh}$ 와  $\omega_{zl}$ 은 전체 임계 대역  $Z$ 에서의 임의의 임계 대역  $z$ 에 대한 저주파 경계와 고주파 경계를 나타낸다.

$a_0(z)$ 는 외이(外耳)에서부터 중이(中耳)까지의 다양한 전달 특성 요소(transmission factor)들에 의한 주파수 의존 감쇠 특성을 나타내는 감쇠 성분이다. 단, 이때  $a_0(z)$ 는 외이부터 중이까지의 감쇠 특성을 나타낸 것이므로, 신호를 직접 내이(內耳)로 전달하는 경우 또는 난청 환경과 같이 이러한 특성의 변화가 있을 경우를 제외하고는 그 영향을 무시할 수 있다. 그림 6은 임계 대역 에너지를 나타낸다.

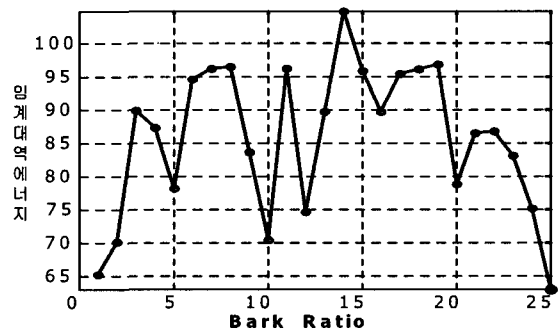


그림 6. 임계 대역 에너지  
Fig. 6. Critical Band Energy

식 (2)의  $X_a(z, i)$ 는 기저막(basilar membrane)에서의 신호 에너지의 확산 현상을 나타내는 지각적인 에너지 확산 함수(spreading function)를 통해 컨벌루션(convolution)을 취함으로써 기저막에서의 신호 에너지에 대한 응답을 나타낼 수 있다. 임계 대역의 신호 에너지에 대한 지각 패턴을 나타내는 자극 에너지(excitation energy)를 다음과 같이 표현할 수 있다<sup>[5]</sup>.

$$X_e(z, i) = \sum_{v=0}^{Z-1} [SF(v, z)X_a(v, i)] \quad (3)$$

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

여기서, 확산 함수  $SF(v, z)$ 는 다음과 같이 표현되어 질 수 있다.

$$SF(v, z) = 15.81 + 7.5(\Delta z + 0.474) - 17.5\sqrt{1 + (\Delta z + 0.474)^2} \quad (4)$$

여기서,  $\Delta z = z - \nu$ 는 바크 인덱스(Bark index)로 표현된 임계 대역 간 차이 값이며 청각 자극에너지는 임계 대역 에너지보다 변화가 급격하지 않다.

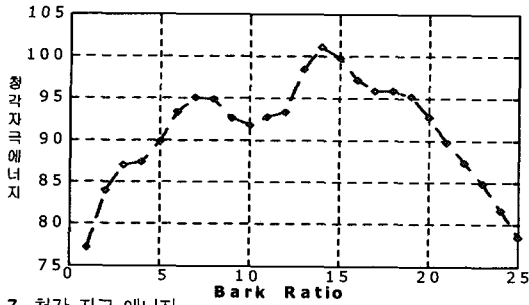


그림 7. 청각 자극 에너지  
Fig. 7. Excitation Energy

#### IV. 앞/뒤 음상정위감 개선

##### 1. 기존 알고리즘

비개인화된 HRTF로 인한 앞/뒤 방향의 혼돈을 개선하기 위한 기존의 알고리즘으로는 각 방향에 대하여 특정 주파수 대역을 감쇄 및 증폭시키는 방법을 들 수 있다<sup>[8]</sup>. 이는 앞/뒤 방향지각에 관련된 특징이 특정 주파수 대역에 있다는 심리음향의 연구결과를 바탕으로 한 것이다.

그림 8과 같이 HRTF를 5개의 대역으로 나누어 각 방향에 따라 저주파 대역은 8dB, 고주파 부분은 12dB를 증폭/감쇄시킨 후 원 HRTF와 컨벌루션을 시행하여 개선된 HRTF를 획득하게 된다.

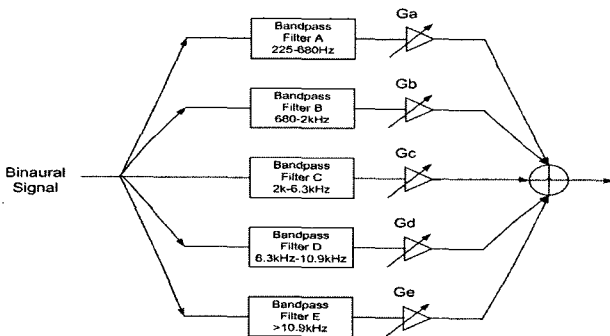


그림 8. 필터 뱅크 방식  
Fig. 8. Filterbank method

두 번째 방법으로는 인간의 신체 특성 중에서 귀의 돌출된 정도의 차이가 앞/뒤 방향 지각에 관계한다는 것으로 그 특성을 이용하는 방법이다<sup>[14]</sup>. 돌출된 정도가 큰 것이 앞/뒤 방향을 잘 지각하는 것으로 판단하여 큰 각도에 대한 주파수 특성을 부각시키게 된다.

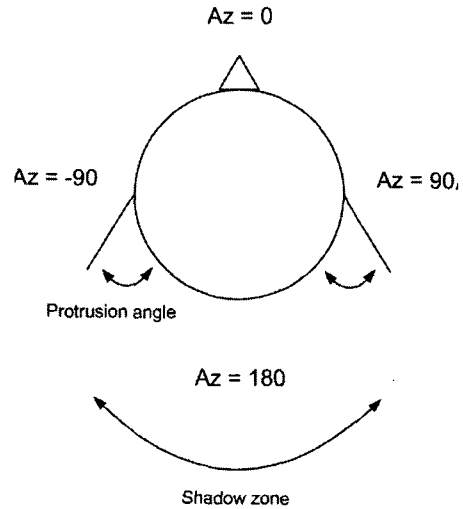


그림 9. 귀의 돌출각도를 이용  
Fig. 9. Using protrusion angle

세 번째 방법으로는 혼돈이 일어나는 두 방향의 주파수 스펙트럼의 차를 이용하는 방법이 있다<sup>[7]</sup>. 먼저 서로 대칭되는 방향의 스펙트럼차를 이용하여 가중치 함수를 생성한 후 원 HRTF에 적용함으로써 HRTF를 수정하게 된다.

$$Diff_p(f) = |HRTF_p(f)| - |HRTF_{p\_symm}(f)| \quad (5)$$

$$HRTF'_p(f) = HRTF_p(f) \times W_p(f)$$

$$W_p(f) = \begin{cases} \log_2(2^{Diff_p(f)} + 1), & Diff_p(f) \geq 0 \\ 2^{Diff_p(f)}, & Diff_p(f) < 0 \end{cases} \quad (6)$$

##### 2. 심리음향을 이용한 HRTF의 개선

소리가 귀에 도달하면 고막에서 역학적인 에너지로 변환되고 중이를 통해 증폭되게 된다. 증폭된 에너지는 기저막에서 전기적인 자극으로 변환되어 청신경계에 전달되어진

다. 이 때 마스킹 효과나 절대가청한계 등의 인간의 청각 특성으로 인해 지각적으로 무관한 정보는 인간의 청각에 영향을 끼치지 못하게 된다. 따라서 본 논문에서는 인간의 청각특성을 기반으로 한 알고리즘을 제안한다.

먼저 음상을 정위시킬 방향,  $a$ 의 머리전달함수,  $HRTF_a$ 의 주파수 스펙트럼 특성을 확인하기 위하여 Fourier Transform을 적용한다. 다음으로 방향지각에 혼돈을 일으키는  $a$ 의 대칭방향,  $b$ 의 머리전달함수,  $HRTF_b$ 의 주파수 스펙트럼 특성을 확인하기 위하여 Fourier Transform을 적용한다.

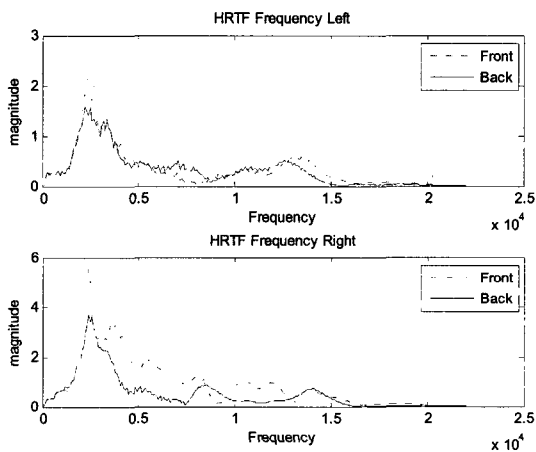


그림 10. 머리전달함수의 스펙트럼 (30° & 150°)  
Fig. 10. Frequency Spectrum of HRTF (30° & 150°)

그림 10은 앞/뒤 방향에 따른 머리전달함수의 주파수 스펙트럼을 보여주고 있다. 많은 연구 결과에서 음상 지각에 대한 단서로는 스펙트럴 큐라 불리는 고막에 도달하는 신호의 스펙트럼 변화를 들고 있다. 음상 방향은 특정 주파수 대역에 의존하게 되는데 이 대역을 결정대역이라 한다. 즉 음원 방향에 관계없이 고막에 도달하는 신호의 스펙트럼에서 강조된 주파수 대역과 일치하는 방향 결정대역의 방향에 음상이 정위되게 된다.

다음으로 각 머리전달함수가 청각에 끼치는 영향을 확인하기 위하여 각 임계 대역,  $z$ 에서의 임계대역 에너지,  $cband\_energy_a(z,i)$ ,  $cband\_energy_b(z,i)$ 를 계산한다. 위에서 설명한 것처럼 머리전달함수는 각 방향에 대해서 방향

결정대역이 존재한다. 따라서 머리전달함수가 고막에 도달하면 우세한 방향 결정대역에 따라 청각 기관의 임계대역에 끼치는 영향이 달라진다.

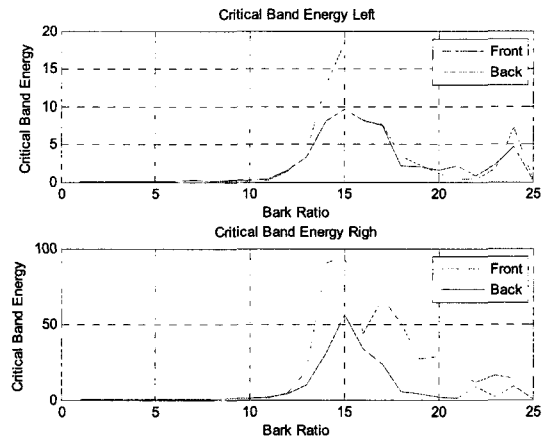


그림 11. 임계대역 에너지 (30° & 150°)  
Fig. 11. Critical Band Energy of HRTF (30° & 150°)

이어 임계대역에  $cband\_energy_a(z,i)$ 를 이용하여 각 임계대역에 대한 청각 자극 에너지,  $excitation\_energy_a(z,i)$ 를 계산한다. 마찬가지로  $b$ 방향에 대한 청각자극 에너지,  $excitation\_energy_b(z,i)$ 를 계산한다.

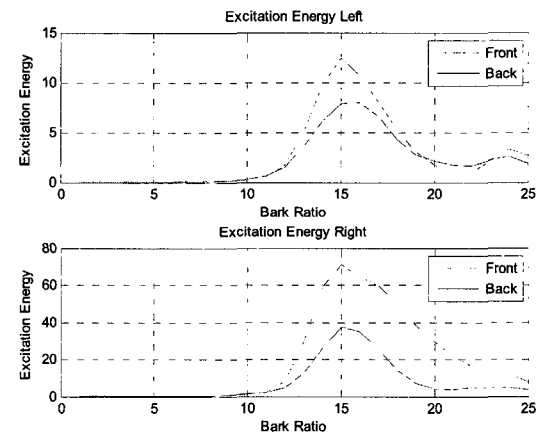


그림 12. 청각자극 에너지 (30° & 150°)  
Fig. 12. Excitation Energy of HRTF (30° & 150°)

그림 12를 보게 되면 머리전달함수는 저주파 부분에서는 각 방향에 대한 주파수 스펙트럼 차이가 크지 않으므로 청

각자극 에너지의 차이도 크지 않지만 고주파 부분에서는 주파수 스펙트럼의 차이가 커지므로 인간의 귀에 끼치는 영향도 달라지게 됨을 알 수 있다.

그리고 음상을 정위시킬 방향, a와 혼돈을 일으키는 대칭 방향, b의 각 청각 자극 에너지의 비율을 계산하여 각 임계 대역에 대한 가중치,  $rate_a(z,i)$ 를 생성하게 된다.

$$rate_a(z,i) = \frac{excitation\_energy_a(z,i)}{excitation\_energy_b(z,i)} \quad (7)$$

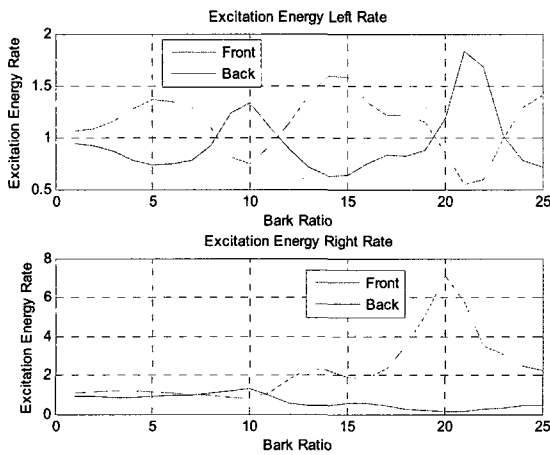


그림 13. 청각자극 에너지 비율 (30° & 150°)  
Fig. 13. Rate of Excitation Energy (30° & 150°)

이렇게 구해진 가중치를 원 HRTF<sub>a</sub> 주파수 스펙트럼에 적용하여 각 방향에 따른 우세한 대역을 증폭시켜 준다. 그러면 스펙트럼 상의 피크나 노치 같은 성분이 강조되어 결과적으로는 혼돈방향의 두 HRTF의 청각자극에너지 차이가 부각되는 효과를 얻게 된다.

$$HRTF'_a(f_z,i) = HRTF_a(f_z,i) \times rate_a(z,i) \quad (8)$$

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

여기서  $f_z$ 는 각 임계 대역에 해당하는 주파수 범위를 의미한다. 기존의 방식에서는 모든 HRTF에 대하여 고정된 대역만을 증폭시키지만 본 알고리즘은 각 방향의 HRTF에 대하여 유동적으로 청각에 영향을 미치는 부분만을 제어함으로써 방향감을 극대화시킬 수가 있게 된다.

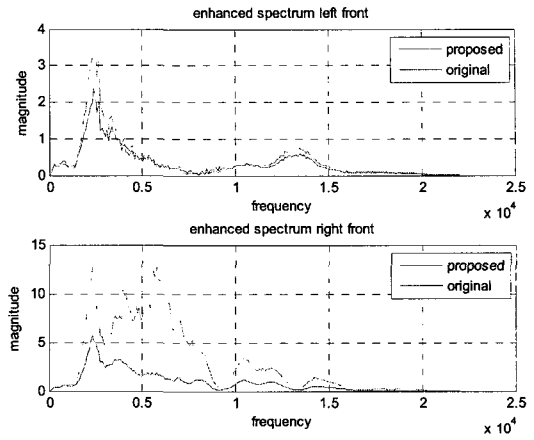


그림 14. 증폭된 머리전달함수 (30° & 150°)  
Fig. 14. Amplified HRTF (30° & 150°)

증폭된 HRTF'<sub>a</sub>는 원 HRTF<sub>a</sub>와의 컨벌루션 연산을 통해 방향감이 개선된 머리전달함수를 생성하게 되며 개선된 머리전달함수는 최종적으로 모노음과의 컨벌루션 연산을 수행하여 방향감이 개선된 3D 입체음향을 생성하게 된다.

$$enHRTF(f,i) = HRTF'_a(f,i) * HRTF_a(f,i) \quad (9)$$

$$3D \text{ 입체음향} = enHRTF * Sound\_Source \quad (10)$$

제안된 알고리즘의 Block Diagram은 그림 15와 같다.

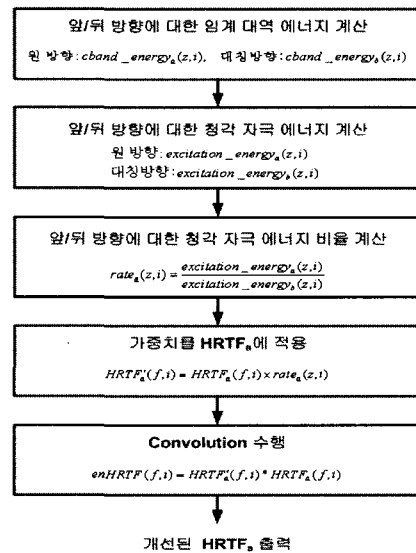


그림 15. 제안된 알고리즘  
Fig. 15. Proposed algorithm

### V. 시뮬레이션 및 청감테스트

본 논문은 HRTF를 이용하여 3D 입체 음향을 생성하는 과정에서 발생하는 앞/뒤 음상정위 특성에 대한 혼돈을 개선하기 위한 알고리즘을 제안한다.

표 2. 사운드 파일  
Table 2. Sound Source

Source 1	영화 효과음(폭탄 폭발음)
Source 2	영화 효과음(비행기 엔진소리)
Source 3	음성

시뮬레이션에서 사용되는 사운드 신호는 영화 DVD에서 추출된 44.1kHz, 16bits 오디오 신호를 사용하였으며 HRTF DB는 MIT Lab.에서 측정한 것을 사용하였다. 사운드 신호재생을 위해 청취자의 귀에 밀착되는 헤드폰(audio-technica ATH-ES7)을 사용하였으며 청취자가 사전에 음원의 종류를 인지할 수 없도록 하기 위하여 청취자와 실험에 사용된 컴퓨터와 거리를 두도록 하였으며 무순서로 음원을 재생하였다.

제안된 알고리즘의 성능을 평가하기 위하여 0°/180°, 30°/150°, 45°/135°, 60°/120°의 네 방향에 대하여 청감테스트를 실시하였다. 비교대상은 원 HRTF, 주파수 대역별로 가중치를 주는 방식 및 제안된 알고리즘이며 테스트 방법은 오디오 전문 분야에 종사하지 않는 남녀 10명을 대상으로 하여 방향감이 가장 떨어지는 신호를 1로 기준을 삼아서 개선 정도를 점수로 나타내도록 하였다.

[표 3]과 [표 4]에서 확인할 수 있듯이 제안된 방법이 기존의 방법에 비해서 효과가 있음을 확인할 수 있다. 이는 제안된 방법을 이용하여 원 HRTF의 스펙트럼에서 각 방향에 따라 인간의 청각에 끼치는 영향이 큰 대역을 조절해 줌으로서 결과적으로 방향 결정대역이 갖는 에너지를 증가시켜주어 그에 따른 청각에 미치는 영향을 조절할 수 있기 때문이다.

더불어 본 알고리즘이 음질에 미치는 영향을 확인하기 위하여 음질테스트를 실시하였다. 비교대상은 원 HRTF, 주파수 대역별로 가중치를 주는 방식 및 제안된 알고리즘이며 테스트 방법은 오디오 전문 분야에 종사하지 않는 남녀 10명을 대상으로 하여 음질이 가장 떨어지는 신호를 1

로 하여 각 방향에 대한 결과를 제시하도록 하였다.

표 3. 전방감 테스트 결과  
Table 3. Forward test result

방향	0°			30°		
source	S1	S2	S3	S1	S2	S3
원HRTF	2.1	1.6	3	3.1	2.5	2.2
기존방식	3.8	3	3.4	4	4	3.1
제안방식	4.5	3.9	3.7	4.1	3.9	3.9
방향	45°			60°		
source	S1	S2	S3	S1	S2	S3
원HRTF	2.7	2.9	3.5	1	2	1
기존방식	3.9	3.5	4	3	2.6	3
제안방식	4	3.8	4.5	3.4	4.9	3.5

표 4. 후방감 테스트 결과  
Table 4. Backward test result

방향	120°			135°		
source	S1	S2	S3	S1	S2	S3
원HRTF	2.5	2.4	2	2.5	3	2.7
기존방식	3.3	3	2.5	3.4	3.8	3.3
제안방식	4	3.7	3.5	4	3.8	3.6
방향	150°			180°		
source	S1	S2	S3	S1	S2	S3
원HRTF	2.3	2	3.1	1.9	2.5	1
기존방식	3.2	3.5	4.2	3.3	2.9	3
제안방식	4.1	4.2	3.9	3.2	3.6	3.5

표 5. 테스트 평균 결과  
Table 5. Test average result

방향	0°/180°			30°/150°		
source	S1	S2	S3	S1	S2	S3
원HRTF	2.0	2.1	2.0	2.7	2.3	2.7
기존방식	3.6	3.0	3.2	3.6	3.8	3.7
제안방식	3.9	3.8	3.6	4.1	4.1	3.9
방향	45°/135°			60°/120°		
source	S1	S2	S3	S1	S2	S3
원HRTF	2.6	3	3.1	1.8	2.2	1.5
기존방식	3.7	3.7	3.7	3.2	2.8	2.8
제안방식	4	3.8	4.1	3.7	4.3	3.5



표 6. 음질테스트 결과  
Table 6. Sound quality test result

방향	0°/180°			30°/150°		
source	S1	S2	S3	S1	S2	S3
원HRTF	3.3	3.7	3.6	3.4	3.2	3.9
기존방식	3.1	4	3.5	3.6	3.3	3.8
제안방식	3.4	4.3	3.8	3.6	3.5	4.2
방향	45°/135°			60°/120°		
source	S1	S2	S3	S1	S2	S3
원HRTF	3.8	3.1	3	3.9	3.5	3.9
기존방식	3.8	3.6	3.3	4.1	4	3.6
제안방식	4.1	3.9	3.5	4	3.9	4

[표 6]은 각 방향에 대한 음질 테스트 결과를 나타내고 있으며 제안된 알고리즘을 통해 생성된 입체음향의 음질이 기존 알고리즘의 음질보다 개선되었음을 알 수 있다.

## VI. 결론 및 향후 연구 방향

본 논문에서는 HRTF를 이용한 3D 입체음향 생성과정에서 심리음향을 이용한 앞/뒤 방향지각의 혼돈을 개선하기 위한 알고리즘을 제안한다. 바이노럴 시스템에서 HRTF를 이용하여 3D 입체음향을 생성할 경우 시간이나 레벨의 차 등 소리가 음원에서 인간의 두 귀에 도달하는 정보를 이용하게 된다. 그러나 음원과 청취자의 두 귀 사이의 거리가 같아지는 혼돈원추 상에서는 시간이나 레벨의 차로서는 각 방향의 차이를 구분할 수 없게 된다.

이에 각 시스템에서의 신호 에너지 표현인 임계 대역 에너지와 각 임계 대역 에너지에 대한 지각적인 응답을 나타내는 자극 에너지를 이용하여 각 방향에 대해 인간의 귀에 영향을 크게 끼치는 부분을 수정하는 알고리즘을 제안하였다. 이러한 청각적 자극 에너지를 이용한 개선 방식은 사운드가 인간의 청각에 미치는 영향을 직접 제어함으로써 주파수 스펙트럼을 이용한 기존의 방식에 비해 정확한 방향감을 생성할 수 있었다. 또한 제안된 알고리즘의 개선 결과를 확인하기 위해서 오디오 비전문가를 대상으로 청감테스

트를 실시하여 기존의 알고리즘에 비해 앞/뒤 방향지각의 혼돈 문제가 개선됨을 확인할 수 있었다.

본 알고리즘은 사람이 단지 2개의 귀로 입사되는 음향으로 모든 음감을 지각한다는 점을 감안하여 음원으로부터 인간의 고막까지를 전달 경로로 모델링하는 실감음향기술에 있어서 현실감을 더해줄 것이다. 그러나 이러한 청각 모델링이 상당히 주관적인 특색을 가지므로 모든 사람들에게 동일한 효과를 느끼게 하기에는 한계가 있으므로 청각 모델링의 오차를 낮추기 위한 여러 가지 연구가 동반되어야 할 것이다.

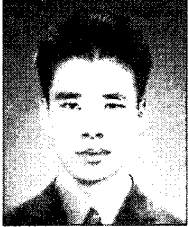
## 참고 문헌

- [1] 윤대희, "오디오 신호처리 기술 동향", 전자공학회지, 제31권 6호, 2004
- [2] 김시호, 김경훈, 배건성, 최승인, 박만호, "HRTF를 이용한 헤드폰 기반의 다채널 입체음향 생성", 전자공학회논문지, 제42권 제1호, 2005
- [3] 김익형, 정의필, "웨이블릿을 이용한 입체음향의 확산감 제어", 신호처리시스템학회 논문지, 제4권 4호, 2003
- [4] Durand R. Begault, "3-D Sound for Virtual Reality and Multimedia", NASA, 2000
- [5] 구교식, 차형태, 유일현, "필터 계수 확장을 이용한 적응 지각 필터의 속도 개선", 신호및시스템학회, 2005
- [6] Navarun Gupta, Armando Barreto and Carlos Ordonez, "Spectral Modification of head-related transfer functions for improved virtual sound spatialization", IEEE, 2002
- [7] 김경훈, 김시호, 배건성, 최승인, 박만호, "헤드폰 기반의 입체음향 생성에서 앞/뒤 음상정위 특성 개선", 한국통신학회 논문지 Vol.29 No.8c, 2004
- [8] 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, 1998
- [9] 명현, 김기홍, 김기호, 김용완, 김현빈, 김풍민, "입체음향 생성에 있어서 자연스러운 이동음 효과의 구현", 한국정보과학회 제28권 제10호, 2001
- [10] 김기홍, 김용완, 명현, 정혁, 김기호, "3D Sound 기술", 정보과학회지 제19권 제5호, 2001
- [11] 김학진, '크로스토크 제거를 통한 입체 음향 구현에 관한 연구', 전자공학회논문지 제 41권 SP편 제 2호, 2004
- [12] 서정국, 차형태, '잡음 에너지 제어를 통한 지각필터 성능 개선', 한국음향학회지 제 24권 제1호, 2005
- [13] 강성훈, 강경욱, 입체음향, 기전연구사, 1997
- [14] Ming Zhang, Kah-Chye Tan, M.H.Er, "A Refined Algorithm of 3D Sound Synthesis", ICSP98, 1998

---

저 자 소 개

---



구 교 식 (Kyo-Sik Koo)

- 2005년 2월 : 송실대학교 정보통신공학부 졸업
- 2005년 3월~현재 : 송실대학교 전자공학과 석사과정
- 주관심분야 : 오디오 신호처리, 심리음향, 디지털 신호처리



차 형 태 (Hyung-Tai Cha)

- 1993년 : The University of Pittsburgh (공학박사)
- 1993년~1996년 : 삼성전자 신호처리 연구소 선임연구원
- 1996년~현재 : 송실대학교 정보통신공학부 부교수
- 2001년~현재 : 신호 및 시스템학회 이사
- 2001년~현재 : 방송공학회지 편집위원
- 주관심분야 : Multimedia Systems and Applications Audio and Video Signal Processing, Communication System, ASIC Implementation of Digital System