

논문 2010-47SP-5-1

# 헤드폰 청취환경에서의 실감 오디오 재현을 위한 음상 외재화 기법

( A Sound Externalization Method for Realistic Audio Rendering in a  
Headphone Listening Environment )

김 용 국\*, 전 찬 준\*, 김 홍 국\*\*, 이 용 주\*\*\*, 장 대 영\*\*\*, 강 경 옥\*\*\*

( Yong Guk Kim, Chan Jun Chun, Hong Kook Kim, Yong Ju Lee, Dae Young Jang,  
and Kyeongok Kang )

## 요 약

본 논문에서는 헤드폰 재생 환경에서의 머리 밖 음상정위를 위한 음상 외재화(externalization) 기법을 제안한다. 제안된 기법에서는 기존의 머리전달함수(HRTF) 또는 초기 반사음 등을 이용한 외재화 기법들에서 발생하는 정위된 음성의 음색 왜곡을 줄이는 것에 그 초점을 맞춘다. 즉, 제안된 음상 외재화 기법은 모델 기반의 HRTF와 잔향 기법을 결합하고, 전방 음상 외재화의 성능 향상을 위하여 decorrelation 및 spectral notch 필터링 기법 등을 포함한다. 제안된 음상 외재화 기법의 성능을 평가하기 위하여 백색잡음, 효과음, 음성 및 오디오 등 다양한 장르의 음원을 이용하여, 평가자의 주관에 의한 청취평가를 수행하였다. 제안된 음상 외재화 알고리즘은 성능평가 결과에서 기존의 방법에 비해 더 좋은 외재화 거리 성능을 보였다.

## Abstract

In this paper, a sound externalization method is proposed for out-of-the-head localization in a headphone listening environment. In order to reduce timbre distortion by the conventional methods using a measured a head-related transfer function (HRTF) or early reflections, the proposed method integrates a model-based HRTF with reverberation. In addition, for improving frontal externalization performance, techniques such as decorrelation and spectral notch filtering are included. To evaluate the performance of the proposed externalization method, subjective listening tests are conducted by using different types of sound sources such as white noise, sound effects, speech, and music. It is shown from the test results that the proposed externalization method can localize sound sources farther away from out of the head than the conventional method.

**Keywords:** 음상 외재화, 실감 오디오, 3D 오디오, 머리전달함수, 음상 정위

\* 학생회원, \*\* 평생회원-교신저자, 광주과학기술원 정보통신공학부

(School of Information and Communications, Gwangju Institute of Science and Technology)

\*\*\* 정회원, 한국전자통신연구원 실감음향연구팀

(Realistic Acoustics Research Team, Electronics and Telecommunications Research Institute)

※ 본 연구는 방송통신위원회, 지식경제부 및 한국산업평가관리원의 IT 원천기술 개발 사업의 일환으로 수행한 연구로부터 도출되었음[2008-F-011, 차세대 DTV 핵심기술 개발].

접수일자: 2010년6월19일, 수정완료일: 2010년8월4일

## I. 서 론

최근 멀티미디어 기술의 발달, 특히 음향 기술의 급격한 발달과 더불어 고품질 오디오에 대한 요구와 함께 보다 현실감 있는 오디오를 재생하기 위한 실감 오디오 기술 개발이 요구되고 있다. 이러한 요구를 만족시키기 위해 사용자의 가상현실 증강 및 실감나는 오디오 효과를 제공해줄 수 있는 3차원 오디오에 대한 연구가 활발히 진행되고 있다. 3차원 오디오 기술은 개발자가 일반 음향에 특정 가상 공간정보를 담아 사용자가 위치해 있지 않은 가상의 환경에 대한 방향감과 거리감을 느낄 수 있도록 하는 기술이다<sup>[1]</sup>. 이러한 3차원 오디오 기술은 크게 머리전달함수(head-related transfer function; HRTF)나 초기 반사음(early reflection) 등을 이용하여 음원이 실제 위치해 있지 않은 곳에 음상을 정위시키는 음상정위(sound image localization) 기술, 공간감을 주기 위한 입체음장(virtual sound field) 기술, loudspeaker에서 3차원 오디오 재생을 위한 가상 채널 확장 기술, 헤드폰 재생환경에 대해 음상을 머리 밖으로 뻗히게 하는 음성 외재화(externalization) 기술, 그리고 loudspeaker에서의 좌우 채널에 의해 나타나는 상호간섭을 제거(crosstalk cancellation)하는 기술 등으로 나눌 수 있다. 이중에서도 음성 외재화 기술은 PMP, MP3 플레이어, DMB 등 휴대단말 장치들을 이용하여 영상 및 음향 등 멀티미디어를 소비하는 경우가 증가하고 있는 추세에 맞추어 그 중요성이 증대되고 있다. 즉, 휴대단말 장치들은 스피커보다는 헤드폰이나 이어폰을 통해 오디오를 듣는 경우가 일반적이며, 이때 헤드폰이나 이어폰에서 발생된 오디오 신호가 스피커 환경에서와 같이 소리가 전달되는 동안 벽면 및 사물에 의한 반사 및 회절 등에 따른 음향 특성 변화 없이 고막까지

직접 전달됨으로써 소리가 머리 내부에 맺히게 되는 내재화 현상(in-head localization)이 발생하게 된다<sup>[1~2]</sup>. 이는 사용자가 느끼는 현실감 및 공간감을 떨어뜨리게 되는 주요인이 된다. 뿐만 아니라 음상이 머릿속에 멎힘으로써 사용자는 스피커로 오디오를 들을 때보다 피로감을 더 쉽게 느끼게 되는 문제 또한 발생하게 된다<sup>[3]</sup>. 이와 같은 음성 내재화 문제를 해결하기 위하여 음성 외재화 기술이 필요하게 된다.

음성 외재화 기술은 그림 1과 같이 음상이 머리 외부에 멎히도록 하여 휴대단말 장치 사용자들이 이어폰이나 헤드폰을 통해서도 마치 영상이 재생되는 단말장치와 동일한 위치에 놓인 스피커를 통해 오디오를 듣는 것과 같은 효과를 만들어 낼 수 있다. 음성 외재화 기술을 실현하기 위한 다양한 방법이 최근 연구되고 있다<sup>[2~6]</sup>. 하지만 기존의 외재화 알고리즘은 원음의 음색을 왜곡시키거나 사람마다 다른 인체 특성으로 인해 느끼게 되는 공간감의 편차가 심하다는 단점이 있다<sup>[3]</sup>. 그 뿐만 아니라 위에서 언급한 대부분의 외재화 방법들이 정면에 음상을 정위시키고자 하는 경우, 측면에 비해 외재화 효과가 현격히 떨어지는 문제 역시 존재한다.

본 논문에서는 수학적으로 모델링된 HRTF 모델 및 image method 기반 인공잔향을 오디오에 추가하여 음상을 사용자의 머리 밖에 멎히도록 하는 외재화 알고리즘을 제안한다. 또한 전방 음성 외재화 성능을 향상시키기 위하여 decorrelation 및 spectral notch 특성을 추가적으로 적용한 외재화 알고리즘을 제안한다.

본 논문의 구성은 다음과 같다. II장에서는 음성 외재화 알고리즘에 사용된 머리전달 함수 모델 및 인공잔향 생성 방법, 전방 음성 외재화 성능을 향상시키기 위한 decorrelation 및 spectral notch 필터링에 대해 기술한다. III장에서는 이를 전체적으로 통합한 외재화 알고리즘을 제안한다. 그리고 IV장에서는 제안된 알고리즘의 성능을 검증하기 위한 성능평가 방법 및 그 결과에 대해 설명하고, 마지막으로 V장에서 결론을 맺는다.

## II. 헤드폰 환경에서의 실감오디오 재현을 위한 음성 외재화 기법

### 1. 머리전달함수(HRTF) 모델

머리전달함수는 음원에서 청취자의 고막까지 소리가 전파될 때까지의 음의 전달경로에 의한 임펄스 응답으로 정의된다<sup>[7~8]</sup>. 이러한 머리전달함수는 음원에서 고막

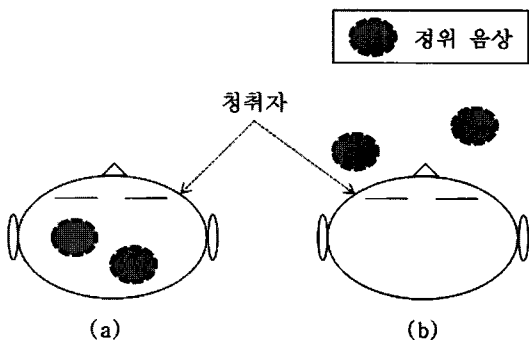


그림 1. (a) 내재화 현상과 (b) 외재화 효과  
Fig. 1. (a) in-head localization and (b) externalization.

까지 경로 상의 머리, 귀, 흉부 등에 의한 모든 반사 및 회절에 대한 효과를 포함하고 있다. 일반적으로 머리 전달함수는 사람의 머리 및 상체와 유사하게 디자인되고, 고막 근처에 마이크가 장착된 더미헤드를 이용하여 머리 전달함수를 측정하는 방법이 일반적이다. 하지만 더미헤드를 이용하여 실제 측정된 머리 전달함수로 입체음향을 생성하거나 외재화 처리를 목적으로 사용하는 경우, 사람과 더미헤드의 불일치로 인한 특정 사용자가 느끼는 입체음향 정위(localization) 성능 및 외재화 효과가 다르게 된다<sup>[9]</sup>. 따라서 정위 성능 및 외재화 성능을 높이기 위해 청취자 개개인에 적합한 머리 전달함수를 각각 측정해야 하지만, 측정에 따른 높은 비용 및 실제 측정의 어려움이 존재한다.

이러한 단점을 극복하기 위한 방법 중의 하나로, 기존 머리 전달함수를 근거로 하여 머리 전달함수를 수학적 모 델링하는 방법이 연구되어 왔다<sup>[10-11]</sup>. 머리 전달함수 모 델을 이용하면 방위각, 고도 및 거리를 변수로 가지기 때문에 각각의 경우에 대하여 머리 전달함수를 측정하지 않아도 될 뿐만 아니라 알고리즘 적용에 있어서도 메모리 사용면에서 효율적이라 할 수 있다. 또한 청취자의 신체조건을 변수로 표현하여, 청취자에 적합한 머리 전달함수를 모델링할 수 있다.

본 논문에서 사용된 머리 전달함수 모델은 Kistler와 Wightman<sup>[10]</sup>에 의해 제안된 구조적 머리 전달함수 모델(structural HRTF model)에 근거한다. 구조적 머리 전달함수 모델의 전체 구조는 그림 2와 같다. 그림 2에서 보는 바와 같이 머리 전달함수 모델은 청취자의 머리 크기 및 음원의 방위각, 고도를 입력받게 된다. 또한 본 논문에서는 음상정위 방향에 따라 각각 다른 구조를 적용한다. 즉, 측면 음상정위의 경우에는 head model 및 pinna model의 순서로 머리 전달함수 모델을 적용하며, 정면 음상정위의 경우에는 head model 및 pinna model의 적

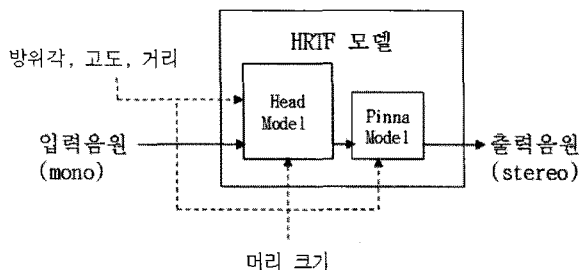


그림 2. 음상 외재화를 위한 머리 전달함수 모델  
Fig. 2. Schematic diagram of a structured HRTF model for sound externalization.

용 순서가 바뀐 구조적 머리 전달함수 모델을 적용한다.

구조적 머리 전달함수 모델을 통한 음원의 처리과정은 다음과 같다. 첫째, 입력된 음원에 대하여 head model을 통해 머리에 의해 발생하는 head shadow 특성을 반영하게 된다. 이는 머리에 입사되는 음원신호에 대하여 인간이 음원의 위치를 인지하는데 사용되는 양 귀 간 상호 시간차이(interaural time difference; ITD)와 상호크기차이(interaural level difference; ILD)를 반영하게 된다. Head shadow 효과는 아래의 식 (1)과 같은 one-pole/one-zero 전달함수로 표현된다.

$$H(z, \theta) = \frac{(2\alpha(\theta) + \frac{2c}{a}T) + (\frac{2c}{a}T - 2\alpha(\theta))z^{-1}}{(2 + \frac{2c}{a}T) + (\frac{2c}{a}T - 2)z^{-1}} \quad (1)$$

여기서  $a, c, T$ 는 머리의 반지름, 소리의 속도, 샘플링 주기를 각각 의미한다. 그리고  $\alpha(\theta)$ 는 음원의 입사 방향각  $\theta$ 에 따라 식 (2)와 같이 주어진다.

$$\begin{aligned} \alpha_L(\theta) &= 1 - \sin(\theta) \\ \alpha_R(\theta) &= 1 + \sin(\theta) \end{aligned} \quad (2)$$

둘째, head model을 통과한 출력 음원은 pinna model을 통해 귓바퀴에 의해 변하게 되는 음향 특성을 반영한다. 귓바퀴에 의한 반사로 인해 크기 및 도달시간이 바뀌게 된 오디오 신호는 식 (3)으로 근사화될 수 있다. 즉,

$$s_{out}(n) = s_{in}(n) + \sum_{k=1}^N \rho_k s_{in}(n - \tau_k(\theta, \phi)) \quad (3)$$

여기서  $s_{out}(n)$ 은 pinna model을 통과한 출력신호를,  $s_{in}(n)$ 은 pinna model에 입력되는 입력신호를 나타내며,  $N$ 은 반사음 차수,  $\theta$ 와  $\phi$ 는 방향각과 고도를 각각 나타낸다.  $\tau_k(\theta, \phi)$ 는 귓바퀴에 의한 도달 지연시간을 의미하며 식 (4)로 근사화된다.

표 1. 식 (3)과 식 (4)에서의 pinna model 계수  
Table 1. Pinna model coefficients defined in Eqs. (3) and (4).

$k$	$\rho_k$	$A_k$	$B_k$	$D_k$
1	0.85	1	2	1
2	-1	5	4	0.5
3	0.15	5	7	0.5
4	-0.35	5	11	0.5
5	0.25	5	13	0.5
6	1	5	15	0.5

$$\tau_k(\theta, \phi) = A_k \cos(\theta/2) \sin(D_k(\pi/2 - \phi)) + B_k \quad (4)$$

식 (3) 및 식 (4)에 사용되는 각 계수, 즉  $A_k, B_k, D_k$  그리고  $\rho_k$ 의 값은 표 1과 같으며 이는 음상정위 성능을 최대화하기 위하여 반복적 청취실험을 통해 도출되었다.

## 2. Image method를 이용한 인공 잔향 생성

일반적으로 특정 위치에서 발생한 소리가 사람의 고막에 도달되는 동안 2.1절에서 언급한 머리전달함수에 해당하는 사람의 인체에 의한 소리의 반사 및 회절이외에도 소리가 공간의 벽이나 천정, 사물 등에 부딪혀 반사 및 회절되어 사람의 고막에 도달하게 된다. 즉, 음원이 도달하는 경로 상에서의 음향 변화로 인해 사람은 소리로부터 공간감 및 거리감을 인지할 수 있다<sup>[7]</sup>. 소리가 전달되는 경로는 음원으로부터 고막으로 직접 입사되는 직접음(direct source)과 벽이나 사물에 부딪혀 입사되는 반사음(reflected source)으로 크게 나눌 수 있다. 이러한 잔향에 해당하는 반사음을 인공적으로 생성하기 위해 다양한 방법들이 사용되고 있다<sup>[12~13]</sup>.

특히, 외재화 알고리즘의 적용에 있어서, 무향실에서 측정된 실측 머리전달함수나 머리전달함수 모델만을 음상정위에 사용할 경우, 생성된 음원에 공간 잔향 성분은 배제되어 있으므로 청취자는 결과음원으로부터 공간감 및 거리감이 현격히 떨어지게 된다. 따라서, 음상을 효과적으로 외재화시키기 위해서는 이러한 잔향 성분을 추가시켜주는 과정이 반드시 필요하다. 따라서 본 논문에서는 머리전달함수 모델을 통해 방향 정위된 음원에 대해 거리감을 부여하여 음상을 외부에 뻗히도록 하기 위하여 image method<sup>[13]</sup>를 이용하여 인공 잔향효과를 생성한다. 즉, 머리전달함수에 의해 방향정보가 부가된 음원에 image method를 이용하여 생성한 룸 임펄스 응답을 처리함으로써 직접음과 함께 인공적으로 생성한 잔향이 더해지게 된다. 이를 통해 생성된 결과 신호로부터 청취자는 거리감 및 공간감을 느낄 수 있게 된다.

## 3. Decorrelation을 이용한 전방 음상 외재화 성능 향상 기법

일반적으로 한 개의 스피커에서 발생한 음원이 공간을 통해 사람의 양 귀로 전달되는 동안 벽 및 사물, 인체 등에 부딪혀서 들어오는 반사음 특성에 의해 양 귀로 입사되는 두 신호의 상호상관도(correlation)가 저하

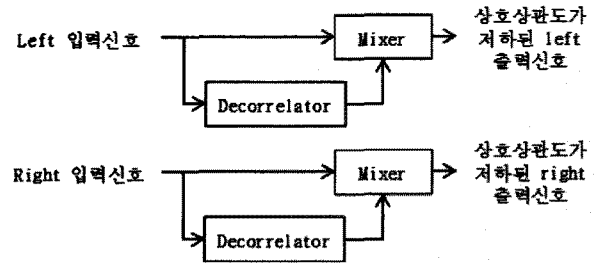


그림 3. 음상 외재화를 위한 decorrelation 및 mixing 기반 상호상관도 제어

Fig. 3. Procedure of reducing the correlation by combining a decorrelator and a mixer.

되는 현상이 발생한다<sup>[8,14]</sup>. 따라서 공간 현상에 따라 특정 위치에 가상으로 음상을 정위시키는 3차원 오디오 기술에서는 음원의 좌/우 채널 신호간의 상호상관도에 따라 음상정위의 성능 차이가 발생한다. 특히, 측면 음상정위의 경우, HRTF 처리에 의한 head shadow effect에 의해 좌/우 출력신호의 상호상관도가 저하되는 효과가 있다. 하지만 정면 음상정위의 경우에는 head shadow effect에 따른 좌/우 채널 신호의 상호상관도 차이가 존재하지 않게 된다. 그러므로 정면이나 후면의 경우에는 상호상관도가 1에 가까운 동일한 신호가 생성된다. 또한 상호상관도를 적절히 낮추어 줌으로써, 중앙에 정확히 맺히는 음상을 좌/우로 확산시키게 되고, 이를 통해 청취자는 음상이 중앙에 정확히 맺히는 음원에 비해 보다 더 외재화 효과를 잘 느끼는 것으로 알려져 있다<sup>[14]</sup>.

이러한 현상을 보완하기 위하여 본 논문에서는 그림 3과 같이 좌/우 신호를 decorrelation시킴으로써 상호상관도의 변화 효과를 얻는다. 즉, 상호상관도에 따른 인지 단서를 음상 외재화에 적용함으로써 전방 외재화 성능 향상을 도모함과 동시에, 특정 음원으로부터 양 귀로 소리가 전달될 때 잔향 등의 청취 공간 특성에 의해서 좌/우 채널 신호간의 상호상관도가 떨어지는 효과를 모델링할 수 있다. 하지만 단순히 전체 대역에 대하여 decorrelation을 적용할 경우, 원음에 비하여 외재화 효과는 높아지지만 음상의 좌/우 확산으로 인한 음상이 틀어지는 현상이 생기는 단점이 존재한다. 이러한 단점을 해결함과 동시에 외재화 성능을 더욱 향상시키기 위하여 대역별 decorrelator를 설계할 필요가 있다. 본 연구에서는 청취실험을 통해 전체 대역 중 저대역에 해당하는 절반은 그대로 bypass시키고, 나머지 고대역에 해당하는 절반에 대하여 decorrelation을 적용시킨 음원의

경우 다른 방법들에 비해 더 높은 외재화 효과를 얻을 수 있을 뿐만 아니라, 음질 및 음상정위 측면에서도 전체 대역에 대해 decorrelation을 적용한 알고리즘에 비해 개선효과를 얻을 수 있음을 관찰하였다.

#### 4. Spectral notch 특성을 이용한 전방 음상 외재화 성능 향상 기법

일반적으로 실측 머리전달함수를 스펙트럼 상에서 분석할 경우 각 방위각 및 고도별로 대역별 스펙트럼 특성이 다르게 나타나는 것을 관찰할 수 있다. 따라서 본 논문에서 제안된 알고리즘에서는 고도가 달라짐에 따라 HRTF 스펙트럼 상에 나타나는 spectral notch 위치를 그림 4와 같이 실제 subject로부터 측정된 CIPIC HRTF DB<sup>[15]</sup>로부터 추출하고 이를 적용함으로써 전방 음상 외재화 성능 개선을 도모한다. Notch 위치에서의 극히 개인적인 특성을 배제하기 위하여 CIPIC HRTF DB에서 제공하는 전체 45명 (좌/우 귀에 해당하는 총 1350 개 데이터)의 subject로부터 측정된 모든 HRTF에 대해 각 고도별 평균 스펙트럼을 도출한다. 이를 통해 개인적인 신체차이에 의한 특성을 배제하고 청취자들의 평균적인 신체 특성에 따라 발생하는 notch 특성을 적용하는 것이 가능하다. 따라서 본 논문에서는 정면에 음상을 정위시키고자 하는 경우 CIPIC HRTF로부터 추출한 정면 방위각에 대한 평균 notch 특성을 적용한다. 본 논문에서 추출된 notch 위치는 9,474 Hz, 14,470 Hz, 그리고 17,743 Hz이다.

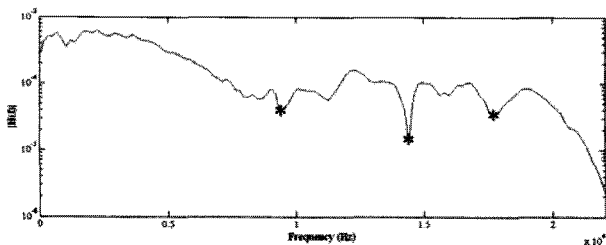


그림 4. HRTF의 평균 스펙트럼 및 notch 위치  
Fig. 4. Average HRTF spectrum obtained from the CIPIC HRTF database, where \* represents a notch frequency.

### III. 제안된 HRTF 모델 및 잔향 기반 음상 외재화 알고리즘

본 논문에서 제안된 전방 음상 외재화 알고리즘의 전체 구조는 그림 5와 같다. 음상 외재화를 적용하기 위

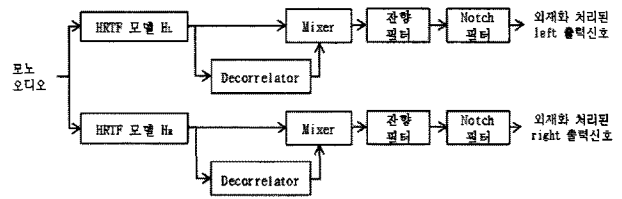


그림 5. 제안된 외재화 알고리즘의 전체 구조  
Fig. 5. Overall structure of the proposed sound externalization algorithm.

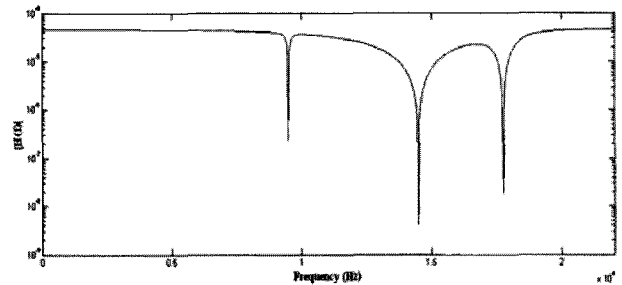


그림 6. Notch 특성을 적용하기 위한 notch filter의 스펙트럼  
Fig. 6. Magnitude response of a notch filter for the proposed sound externalization algorithm.

하여 먼저, HRTF 모델을 통해 원하는 방위각에 음상을 정위 시킨다. 다음으로 2.2절에서 기술한 바와 같이 음상 정위된 음원에 대하여 공간감 및 거리감 부여를 통해 음상을 외재화 시키기 위하여 잔향 필터를 통과시킨다. 본 논문에 사용된 잔향은 20m×20m×3m 크기의 공간에서 음원이 10m, 18m, 1.5m, 왼쪽 마이크 및 오른쪽 마이크가 9.8m, 10m, 1.5m 지점 및 10.2m, 10m, 1.5m에 위치하며, 반사음 차수와 반사 계수는 각각 24와 0.6인 환경을 가정하여 생성하였다.

음상을 정면(0° 방위각)에 정위시키고자 하는 경우 정면 음상정위 성능향상을 위하여 HRTF로 처리한 출력신호에 대해 decorrelation 및 mixing을 수행하고, 소리가 정면 음상정위 시 발생하는 spectral notch 특성을 적용하기 위한 notch 필터를 적용하였다. 즉, 정면 음상정위의 경우, HRTF 모델과의 convolution 과정을 통해 음상정위가 수행된 출력 신호에 대해 decorrelation 필터로 convolution을 수행한다. 이를 통해 얻은 decorrelation된 신호와 decorrelation 필터 입력 신호를 mixing함으로써 좌/우 채널이 상호상관도가 저하된 스테레오 신호를 얻게 되고, 각 신호에 대해 image method를 통해 얻은 공간 임펄스 응답을 부가시켜 준다. 최종적으로, 공간 임펄스 응답이 부가된 출력신호에 대하여 spectral notch를 생성하기 위해 그림 6과 같이

설계된 notch 필터와 convolution을 수행함으로써, 정면 음상 외재화 성능이 개선된 신호를 최종적으로 얻을 수 있었다.

### IV. 성능평가

본 논문에서 제안된 음상 외재화 알고리즘의 성능 검증을 위하여 실제 음상 거리 측정을 통한 외재화 거리에 대한 성능평가를 수행하였다. 외재화 거리에 대한 성능평가 과정은 다음과 같다. 먼저 시험자는 청취자에게 내재화된 효과와 외재화 효과를 인지시키기 위하여 동일한 원음에 대하여 헤드폰 및 1m 거리만큼 떨어진 스피커를 통해 재생되는 소리를 들려준다. 다음으로 각 음원에 대하여 원음 및 외재화 알고리즘을 처리한 음원을 각각 들려준 후 그림 7과 같이 음원이 정위되는 위치를 청취자가 직접 가리키도록 한다. 청취자가 가리킨 지점에 대하여 시험자는 청취자의 머리 중심으로 부터의 거리를 직접 측정하고 이에 대한 통계를 통해 음상 외재화 거리 성능을 평가하는 방법을 수행하였다<sup>[6]</sup>. 이를 위하여 백색잡음, 효과음, 음성, 음악 음원을 가지고 0°, 30°, 60°, 그리고 90° 각각에 대해 외재화 거리에 대한 성능평가를 실시하였다. 본 평가를 위하여 20대의 청각 장애가 없는 16명의 청취자가 실험에 참여하였으며, 헤드폰은 AKG 사의 K240이 실험에 사용되었다.

그림 8은 0°, 30°, 60°, 그리고 90°에 대한 각도별 음상정위 거리의 평균 및 각 실험 콘텐츠별 음상 정위 거리의 평균과 95% 신뢰구간을 나타낸다. 그림에서 y축은 머리의 중심으로부터 손으로 가리킨 지점까지의 거리를 의미한다. 그림에서 보는 바와 같이 제안된 외재화 알고리즘의 경우 정면에서 약 25cm, 측면에서 약 40cm 내외의 외재화 거리 성능을 보임을 알 수 있다.



그림 7. 외재화 거리 성능 측정  
Fig. 7. Snapshot of measuring the perceptual distance.

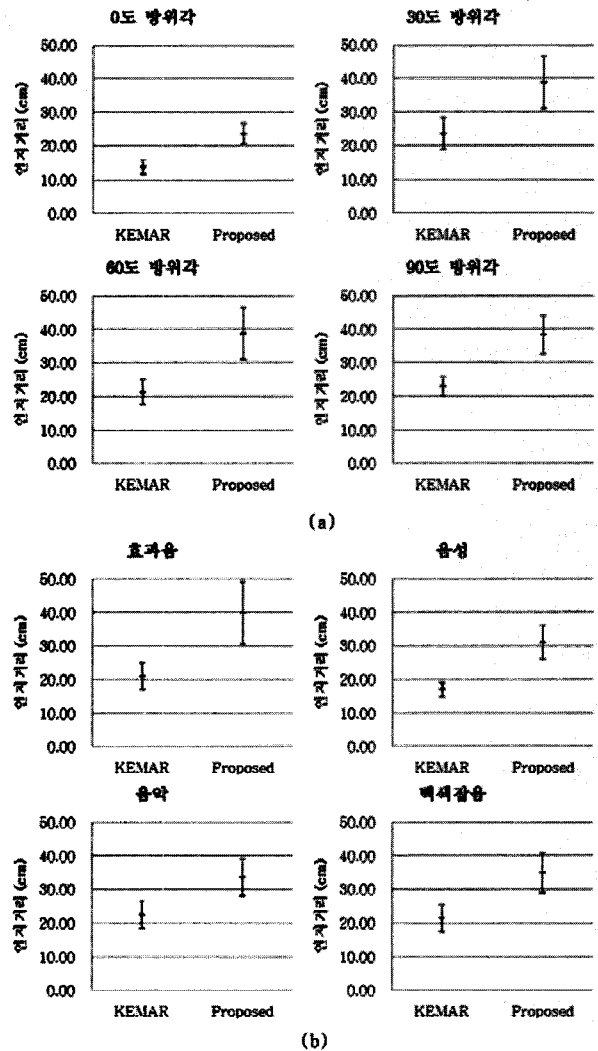


그림 8. 외재화 거리 성능 평가 결과; (a) 음상정위 각도별 외재화 거리 성능 (b) 콘텐츠별 외재화 거리 성능

Fig. 8. Comparison of the perceptual distances between the KEMAR HRTF method and the proposed externalization method for different (a) azimuth and (b) audio sources.

또한 대표적 실측 HRTF인 KEMAR HRTF<sup>[16]</sup>를 이용하여 음상을 정위시킨 음원과 비교할 경우, 제안된 방식이 더 높은 외재화 거리 성능을 보인다는 것을 알 수 있었다. 그리고 콘텐츠별 외재화 거리 성능의 경우, 평균적으로 가장 높은 외재화 거리 성능을 보이는 것을 알 수 있었다.

### V. 결론

본 논문에서는 헤드폰 기반 청취 환경에서의 실감 오디오 재현을 위한 머리전달함수 모델 및 image

method 기반 외재화 알고리즘을 제안하였다. 먼저 수학적으로 모델링된 머리전달함수를 이용하여 음상의 방향을 정위시키고, image method를 이용해 인공잔향을 생성 및 적용시킴으로써 거리감 및 공간감을 부여 하도록 설계하였고, 각 모듈에 대하여 생성 파라미터 최적화를 수행하였다. 추가적으로 외재화 알고리즘의 전방 음상 외재화 성능을 향상시키기 위하여, decorrelation 및 spectral notch 필터링을 적용하였다. 외재화 거리 성능의 주관적 청취평가를 통해, 제안된 방법이 기존의 방식에 비해 높은 외재화 거리 성능을 보인다는 것을 알 수 있었다.

### 참 고 문 헌

- [1] D. R. Begault, *3-D Sound for Virtual Reality and Multimedia*, Academic Press, Cambridge, MA, 1994.
- [2] D. J. M. Robinson and R. G. Greenwood, "A binaural simulation which renders out-of-head-localisation with low-cost digital signal processing of head-related transfer functions and pseudo reverberation," in *Proc. of 104th AES Convention*, Amsterdam, Netherlands, Preprint 4723, May 1998.
- [3] 서정일, 이용주, 장인선, 유재현, 강경욱, "청취환경 차이에 따른 3차원 오디오 기술 개발 동향," *한국방송공학회지*, 제13권, 제1호, pp. 82-96, 2008년 3월.
- [4] P. Rubak, "Headphone signal processing system for out-of-the-head localization," in *Proc. of 90th AES Convention*, Paris, France, Preprint 3063, Feb. 1991.
- [5] T. Choi, Y. C. Park, and D. H. Youn, "Efficient out of head localization system for mobile applications," in *Proc. of 120th AES Convention*, Paris, France, Preprint 6758, May 2006.
- [6] Y. G. Kim, H. K. Kim, Y. J. Lee, D. Y. Jang, I. Jang, and K. Kang, "A sound externalization algorithm based on a structural head-related transfer function model and reverberation," in *Proc. of International Technical Conference on Circuits/Systems, Computers and Communications*, Jeju island, Korea, pp. 1316-1319, July 2009.
- [7] J. Blauert, *Spatial Hearing*, MIT Press, 1997.
- [8] 강성훈, 강경욱, *입체 음향*, 기전연구소, 1997년 7월.
- [9] W. G. Gardner, *3-D Audio Using Loudspeakers*, Kluwers Academic Publishers, 1998.
- [10] D. J. Kistler and F. L. Wightman, "A model of head-related transfer functions based on principal components analysis and minimum-phase reconstruction," *J. Acoust. Soc. Am.*, vol. 91, no. 3, pp. 1637-1647, Mar. 1992.
- [11] C. P. Brown and R. O. Duda, "An efficient HRTF model for 3-D sound," in *Proc. of IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, New Paltz, NY, pp. 298-301, Oct. 1997.
- [12] J. A. Moorer, "About this reverberation business," *Computer Music Journal*, vol. 3, no. 2, pp. 13-28, June 1979.
- [13] J. B. Allen and D. A. Berkley, "Image method for efficiently simulating small-room acoustics," *J. Acoust. Soc. Am.*, vol. 65, no. 4, pp. 943-951, Apr. 1979.
- [14] G. S. Kendall, "The decorrelation of audio signals and its impact on spatial imagery," *Computer Music Journal*, vol. 19, no. 4, pp. 71-87, Winter 1995.
- [15] V. R. Algazi, R. O. Duda, D. M. Thompson, and C. Avendano, "The CIPIC HRTF database," in *Proc. of IEEE Workshop on Applications of Signal Processing to Audio and Electroacoustics*, New Paltz, NY, pp. 99-102, Oct. 2001.
- [16] W. G. Gardner and K. D. Martin, "HRTF measurements of a KEMAR," *J. Acoust. Soc. Am.*, vol. 97, no. 6, pp. 3907-3908, June 1995.

— 저 자 소 개 —



**김 용 국**(학생회원)  
 2006년 전남대학교 전자컴퓨터정보통신공학부 학사 졸업  
 2008년 광주과학기술원 정보통신공학부 석사 졸업  
 2008년~현재 광주과학기술원 정보통신공학부 박사과정.

<주관심분야 : 3D 오디오, 오디오 신호처리>



**전 찬 준**(학생회원)  
 2005년 한국기술교육대학교 전자공학과 학사 졸업  
 2009년~현재 광주과학기술원 정보통신공학부 석사과정.  
 <주관심분야 : 3D 오디오, 오디오 신호처리>



**김 흥 국**(평생회원)-교신저자  
 1988년 서울대학교 제어계측공학과 학사 졸업  
 1990년 한국과학기술원 전기 및 전자공학과 석사 졸업  
 1994년 한국과학기술원 전기 및 전자공학과 박사 졸업

1990년~1998년 삼성종합기술원 전문연구원  
 1998년~1998년 MMC Technology 선임연구원  
 1998년~2003년 AT&T Labs-Research Senior Member Technical Staff  
 2003년~현재 광주과학기술원 정보통신공학부 교수.

<주관심분야 : 음성인식, 다채널 음성 및 오디오 처리, 3D 오디오>



**이 용 주**(정회원)  
 1999년 경북대학교 전자공학과 학사 졸업  
 2001년 경북대학교 전자공학과 석사 졸업  
 2001년~현재 한국전자통신연구원 방통융합미디어연구부 선임연구원.

<주관심 분야: : 3차원 오디오, 오디오 신호처리>



**장 대 영**(정회원)  
 1991년 부경대학교 전자공학과 학사 졸업  
 2000년 배재대학교 컴퓨터공학과 석사 졸업  
 2008년 배재대학교 컴퓨터공학과 박사 졸업

1991년~현재 한국전자통신연구원 방송통신융합연구부문 책임연구원  
 2004년 동경전기대학/DiMagic Co. Ltd. 방문연구원.

<주관심분야 : 객체기반/3차원 음향 신호처리>



**강 경 옥**(정회원)  
 1985년 부산대학교 물리학과 학사 졸업  
 1988년 부산대학교 물리학과 석사 졸업  
 2004년 한국항공대학교 항공전자공학과 박사 졸업

1991년~현재 한국전자통신연구원 책임연구원 및 팀장  
 2006년 영국 Southampton 대학 방문연구원.  
 <주관심 분야 : 음향 신호처리, 3차원 오디오, 디지털 방송 시스템, 객체기반 오디오 시스템>