

보편적인 기저함수를 이용한 중앙면상의 머리전달함수 모델링 Modeling of Median-plane Head-related Impulse Responses Using a Set of General Basis Functions

황 성 목† · 박 영 진* · 박 윤 식*

Sungmok Hwang, Youngjin Park and Youn-sik Park

(2008년 2월 15일 접수 ; 2008년 4월 1일 심사완료)

Key Words : Head-related Impulse Response(머리전달함수), Principal Components Analysis(주성분분석법), Virtual Sounds(가상음향)

ABSTRACT

A principal components analysis(PCA) of the median-plane head-related impulse responses (HRIRs) in the CIPIC HRTF database reveals that the individual HRIRs in the median plane can be adequately reconstructed by a linear combination of 12 orthonormal basis functions. These basis functions can be used to model arbitrary median-plane HRIRs, which are not included in the process to obtain the basis functions. Memory size can be reduced up to 5-fold depending on the number of HRIRs to be modeled. To clarify whether these basis functions can be used to model other set of arbitrary median plane HRIRs, a numerical error analysis for modeling and a series of subjective listening tests were carried out using the measured and modeled HRIRs. The results showed that the set of individual HRIRs in the median plane, which were measured in our lab using different measurement conditions, techniques, and source positions, can be modeled with reasonable accuracy. All subjects, involved in the subjective listening test, reported not only the accurate vertical perception but also the front-back discrimination with the modeled HRIRs based on 12 basis functions.

1. 서 론

인간이 소리를 들었을 때 음원의 방향성을 인지하는데 영향을 미치는 주요인자는 양이간 시간차(interaural time difference: ITD), 양이간 레벨차(interaural level difference: ILD), 외이에 의한 스펙트럼 왜곡(spectral distortion)이다^(1~3). 이런 주요인자의 정보는 음원의 음압과 고막에 도달한 음압 사이의 주파수영역 전달함수인 머리전달함수(head-

related transfer function: HRTF)에 모두 담겨 있으므로^(3~9), HRTF는 가상 입체 음향(virtual sounds) 생성을 위해 중요한 역할을 담당한다⁽¹⁰⁾. 그런데 개개인마다 외이와 머리 등의 형상과 크기가 달라서 개개인마다 고유의 HRTF를 지니고 있다. 따라서 보다 현실감 있는 입체 음향 생성을 위해서 청취자 고유의 HRTF를 직접 측정하거나 적절한 방법을 이용하여 얻은 개인화된 HRTF를 사용할 필요가 있는데, 모든 청취자마다 3차원 공간상의 분포된 음원 위치에 해당하는 HRTF database를 측정하여 보유하고 있기 위해서는 큰 저장 공간을 필요로 한다. 저장 공간의 크기를 줄이기 위해 소수의 파라미터만을 사용하여 효과적으로 HRTF를 모델링하는 기법이 필요하다.

† 교신저자: 정회원, KAIST
E-mail: tjdahr78@kaist.ac.kr
Tel: (042)869-3076, Fax: (042)869-8220
* 정회원, KAIST

지난 수 십년동안 HRTF 모델링을 위해서 많은 방법들이 개발되었다^(11~17). 특히 Martens는 처음으로 주성분 분석법(principal components analysis: PCA)을 이용하여 HRTFs를 모델링하였으며⁽¹⁸⁾, Kistler와 Wightman은 10명에 해당하는 5300 HRTFs를 PCA를 통해 얻어진 5개의 기저함수들의 조합으로 적절히 모델링할 수 있음을 밝혔다⁽¹⁹⁾. 즉, 5개의 기저함수와 이에 곱해지는 가중치만을 저장하고 있으면 모델링에 사용된 모든 HRTFs를 생성할 수 있다. Wu 등은 음원의 음압과 고막에 도달한 음압 사이의 시간영역 전달함수(head-related impulse response: HRIR, HRTF과 역 푸리에 변환 관계)를 PCA를 이용하여 모델링하였다⁽²⁰⁾. 그러나 이전의 연구들은 기저함수를 모델링하고자 데이터의 PCA를 통해 구하였는데, 이는 모델링하고자 하는 데이터가 바뀌거나 추가된다면 PCA를 다시 수행해야 하며, 새로운 기저함수들과 이에 곱해지는 가중치 모두가 저장되어야 하므로 계산량과 저장공간이 늘어나게 되는 단점이 있다. 따라서 모델링해야 하는 데이터가 바뀌거나 추가되더라도 PCA를 다시 수행하여 새로운 기저함수를 구할 필요 없이, 특정한 보편적인 기저함수들의 조합으로 모든 사람의 HRTFs 혹은 HRIRs 모델링이 가능하다면 앞서 언급한 계산량과 저장공간이 늘어나는 문제점을 극복할 수 있다.

이에 이 연구에서는 많은 사람의 HRIRs로 구성된 큰 데이터베이스로부터 PCA를 통해 기저함수를 추출하고 이 기저함수들이 PCA과정에 포함되지 않은 임의의 HRIRs를 모델링하는데 사용될 수 있음을 수치적인 오차분석과 주관평가를 통해 밝힌다. 즉, 모든 사람의 HRIRs를 모델링하는데 사용될 수 있는 보편적인 기저함수들을 구하고 이를 검증하는 데 이 연구의 목적이 있다. 이 연구에서는 방위각이 0°인 중앙면(median plane)상의 HRIRs만을 한정해서 다룬다.

2. 보편적인 기저함수

이 장에서는 HRIR을 모델링하는데 사용될 보편적인 기저함수를 구하는 절차를 간략하게 기술한다. 보다 자세한 내용은 본 저자들의 이전연구^(21,22)에 상세히 기술되어 있다.

이 연구에서는 보편적인 기저함수를 CIPIC HRTF database⁽²³⁾를 이용하여 구한다. CIPIC HRTF database는 3차원 공간상의 음원의 위치를 달리하며 측정된 45명의 개인 HRIRs로 구성되어 있다. 이 중에서 이 연구는 중앙면상의 왼쪽 귀에 해당하는 2205 HRIRs(45명 × 49음원위치)만을 이용한다. PCA를 수행하기 전에, 음원에서 방사된 음파가 고막에 가장먼저 도달하는 직접음부터 1.5 ms까지의 응답만을 추출하기 위해 데이터 전처리 과정을 수행한다. 직접음부터 1.5 ms까지의 응답(67 samples, 44.1 kHz sampling rate)은 외이, 머리, 몸통, 어깨 등의 영향을 모두 내포하고 있으며, 이런 전처리 과정을 통해 데이터 사이즈를 줄일 수 있다. 이렇게 전처리된 HRIRs로 구성된 데이터 행렬(Y: 67 × 2205)의 코베리언스 행렬로부터 고유치 해석을 통해 기저함수를 얻게 된다. 즉, k번째 주요 기저함수는 k번째로 큰 고유치에 해당하는 고유벡터가 된다. 이때, 코베리언스 행렬을 구하기 전에 원래 HRIRs의 평균을 구하여 이를 빼주는 작업을 거친다. 모든 HRIRs의 평균은 방향성에 대한 정보는 없고, 실험조건 등에 따라 달라지므로 이를 제거하고 기저함수를 구한 뒤, 추후 HRIR 복원과정에 이 평균을 다시 더 해주면 된다. 이로써 기저함수가 불필요하게 방향성에 무관한 성분에 영향을 받지 않도록 한다. 이렇게 얻어진 기저함수를 특별히 주성분(principal component: PC)라 부르며, 원래의 HRIRs는 이 PC들의 선형결합과 HRIRs 평균의 합으로 복원할 수 있다는 것이 PCA를 이용한 HRIR 모델링의 주요 핵심이다. 많은 PCs를 사용할수록 더욱 정확하게 모델링할 수 있지만, 소수의 파라미터만을 이용하여 모델링하는 것이 목적이므로 모델링에 사용될 PC의 수를 결정해야 한다. PC의 수를 결정하는 기준으로는 percentage variance(% var)와 percentage modeling error(% error)가 사용될 수 있다. % var는 식 (1)과 같이 표현된다.

$$\% \text{ var}(k) = \frac{\sum_{i=1}^k \lambda_i}{\sum_{i=1}^N \lambda_i} \times 100(\%) \quad (\lambda_1 \geq \dots \geq \lambda_k \geq \dots \geq \lambda_N) \quad (1)$$

여기서 $\lambda_i(i=1,2,\dots,N)$ 는 코베리언스 행렬의 i번째 큰 고유치이다. % error는 식 (2)와 같이 표현된다.

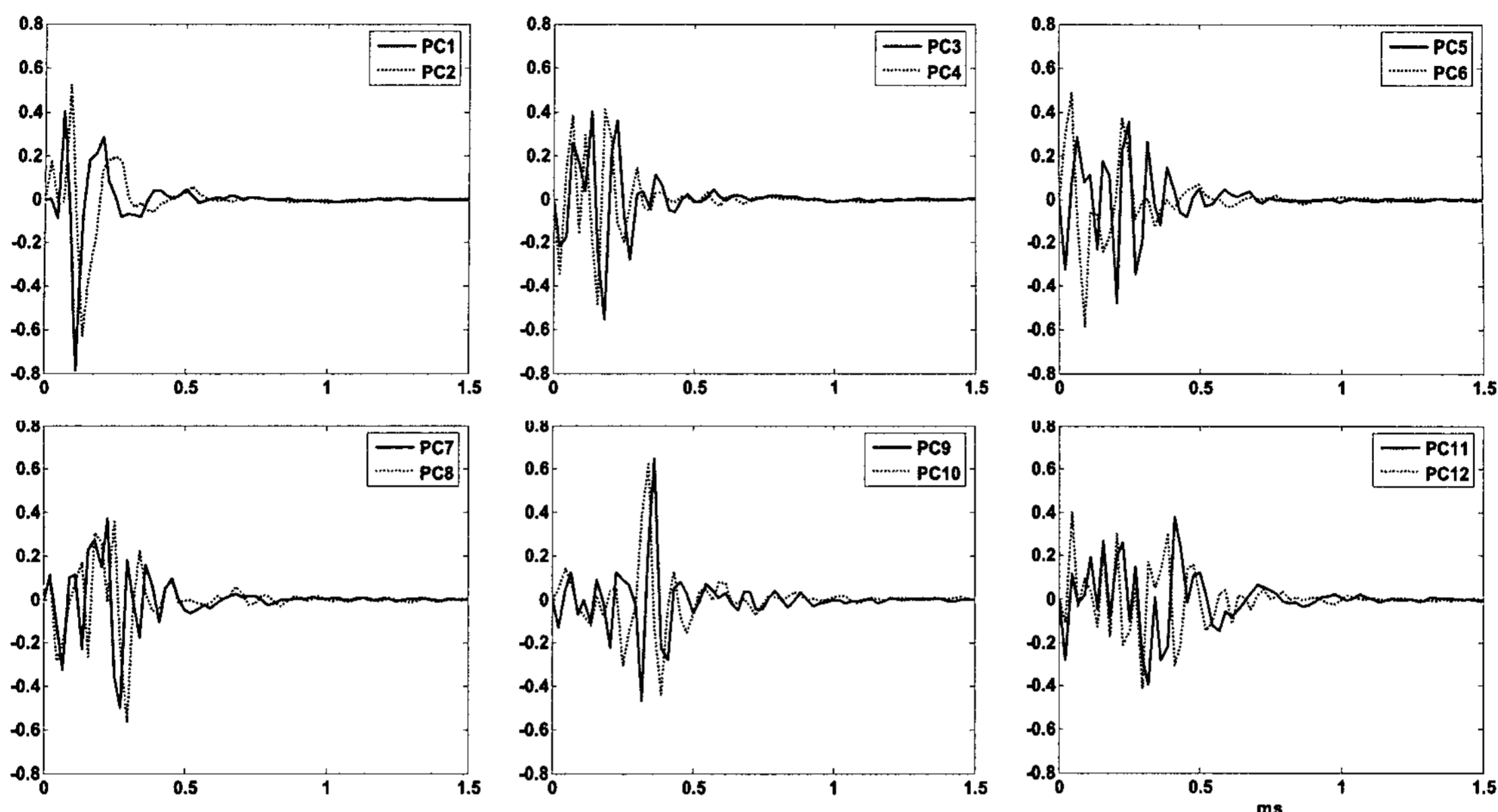


Fig. 1 Twelve PCs obtained from PCA of the left ear 2205 HRIRs in the median plane

$$\% \text{ error } (k) = \frac{\|Y - \hat{Y}_k\|_F^2}{\|Y\|_F^2} \times 100 (\%) \quad (2)$$

여기서 Y 와 \hat{Y}_k 는 각각 원래의 측정된 2205 HRIRs로 구성된 행렬과 k 개 PCs로 모델링된 HRIRs로 구성된 행렬을 나타내며, 하첨자 F 는 행렬의 프로베니우스 놈(Frobenius norm)을 의미한다. Table 1은 PC수의 증가에 따른 $\% \text{ var}$ 와 $\% \text{ error}$ 를 나타낸다. PC의 수가 증가함에 따라 $\% \text{ var}$ 은 지속적으로 증가하고 $\% \text{ error}$ 는 지속적으로 감소함을 확인할 수 있다. 이 연구에서는 90% 이상의 $\% \text{ var}$ 와 5% 이하의 $\% \text{ error}$ 를 기준으로 PC의 수를 결정하였다. 즉, 12개의 PCs를 사용하여 HRIR을 모델링한다. Fig. 1은 사용될 12개의 PCs를 보여준다. PC와 이에 곱해지는 가중치(PCW) 들에 대한 물리적인 의미와 고도각에 대한 의존성 등은 이전연구^(21,22)에 상세히 기술되어 있다.

3. 임의의 HRIRs 모델링

이 장에서는 CIPIC HRTF database의 PCA를 통해 구한 보편적인 기저함수들을 이용하여 PCA과정에 포함되지 않은 임의의 HRIR을 모델링하고 수치적인 오차를 분석한다.

Table 1 Percentage variance and modeling error as a function of number of PCs(k)

k	$\% \text{ var } (k)$	$\% \text{ error } (k)$	k	$\% \text{ var } (k)$	$\% \text{ error } (k)$
1	25.1	36.7	11	88.7	5.5
2	42.4	28.2	12	90.2	4.8
3	52.3	23.4	13	91.4	4.2
4	59.8	19.7	14	92.4	3.7
5	66.5	16.4	15	93.2	3.3
6	72.0	13.8	16	93.9	3.0
7	76.9	11.3	17	94.5	2.7
8	80.5	9.6	18	95.1	2.4
9	83.9	7.9	19	95.6	2.2
10	86.4	6.7	20	96.0	2.0

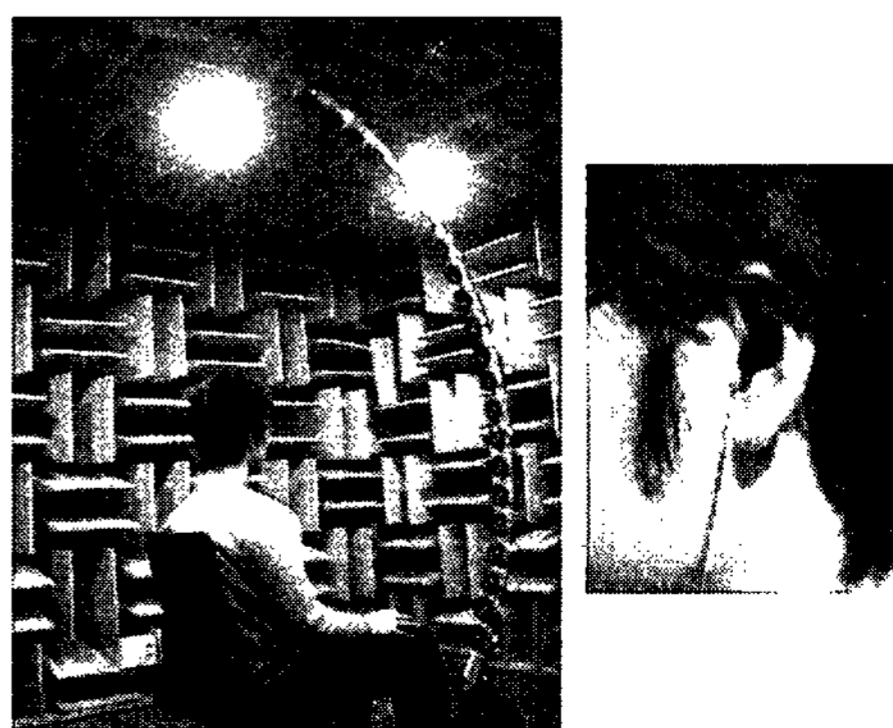


Fig. 2 Experimental setup for the measurement of individual HRIRs in an anechoic chamber

3.1 개인 HRIR 측정

무향실 환경에서 중앙면상에 위치하는 음원의 고도각을 -30° 부터 210° 까지 5° 간격으로 달리하면서 49 음원 위치에 해당하는 HRIRs를 측정하였다. 고도각 0° 는 음원이 피실험자의 정면에 존재하는 경우이고, 90° 와 180° 는 각각 음원이 피실험자의 머리 바로 위, 후면에 존재하는 경우이다. 실험에는 정상적인 청각을 가진 6명의 피실험자(ID: HS, KB, LS, CT, CH, CY)가 참여하였다. HRIR 측정에는 Fig. 2와 같이 피실험자의 머리 중심에서 1.2m 떨어진 지그에 -30° 부터 90° 까지 5° 간격으로 부착된 25개의 스피커(SDT10BNDPA, tweeter) 어레이가 사용되었다. 고도각 90° 이상의 후방음원에 대해서는 피실험자가 앉는 의자를 180° 회전하여 측정하였다. 스피커의 입력은 가청주파수 범위(20 Hz ~ 20 kHz)의 백색잡음을 사용하였고, 바이노널 마이크로폰(B&K Type-4101)을 피실험자의 고막 입구에 위치하여 출력신호를 측정하였다. 입력 및 출력신호는 44.1 kHz로 샘플링하여 1.5초 동안 측정하였다. 이렇게 측정된 신호를 이용하여 HRTF($H_{L,R}$)를 식 (3)과 같이 얻는다.

$$H_{L,R}(f) = \frac{H_{2,L,R}(f)}{H_{1,L,R}(f)} e^{j2\pi fD} \quad (3)$$

여기서 f 는 주파수를 의미하고, $H_{2,L,R}$ 은 스피커 입력신호와 왼쪽(L) 혹은 오른쪽(R) 귀에 위치한 마이크로폰 출력신호 사이의 주파수영역 전달함수이다. 스피커 특성과 마이크로폰 특성으로 인해 실제 스피커의 출력신호 및 마이크로폰 출력신호가 왜곡되므로 이를 보상해주기 위하여 자유음장 조건에서 스피커 입력신호와 마이크로폰 출력신호 사이의 주파수영역 전달함수인 $H_{1,L,R}$ 으로 보상해준다. $e^{j2\pi fD}$ 는 스피커와 머리 중심점 사이의 거리에 해당하는 시간지연을 보상해준다. HRIR은 이렇게 구한 HRTF의 역푸리에 변환으로 구한다.

3.2 수치적인 모델링 오차 분석

CIPIC HRTF database를 이용하여 2장에서 구한 PCs와 중앙면상의 모든 HRIRs의 평균을 이용하여 3.1절에서 측정된 6명의 중앙면 HRIRs를 모델링하였다. Table 2는 각각의 피실험자에 대해서 구한 % error를 나타낸다. % error는 식 (2)와 같이 구하는

데, 이때 사용된 Y 와 \hat{Y}_k 는 각각 피실험자의 측정된 49 HRIRs로 구성된 행렬과 k 개 PCs로 모델링된 HRIRs로 구성된 행렬이다. 12개 PCs를 이용하여 모델링한 경우, 모델링 오차가 12.4%에서 18.7% 사이에 분포한다. 그런데 Table 1에 나타난 것과 같이 12개 PC로 CIPIC HRTF database의 중앙면상의 HRIRs를 5% 이내의 오차로 모델링할 수 있는 것과 비교할 때, Table 2의 모델링 오차가 더 큰 것을 확인할 수 있다. 이는 CIPIC HRTF database의 평균과 이 실험에 참여한 6명의 HRIRs의 평균이 다르기 때문이며 이는 5장에서 자세히 설명한다. Table 2는 모델링에 사용되는 PC 수가 작아질수록 모델링 오차가 급격히 증가함을 보여준다. PC 수에 따른 모델링 정확성은 주파수 영역의 HRTF에서도 확인할 수 있다. 한 예로 Fig. 3는 피실험자 CH의 측정된 HRTFs와 모델링된 HRIRs의 푸리에 변환으로 얻은 HRTFs의 크기(dB 단위)를 비교한 그림이다. 측정된 HRTFs(left-top panel)과 비교할 때, 12개 PC를 사용한 경우(right-top panel) 약 4 kHz 이상에서 나타나는 주파수 특징(spectral peak, notch) 등이 정확하게 모델링됨을 보여준다. 또한, 모델링에 사용되는 PC 수가 감소할수록(bottom panels) 주파수 특징의 정확한 묘사가 어려워짐을 확인할 수 있다.

4. 주관평가

이 장에서는 측정된 HRIRs와 모델링한 HRIRs를 이용한 피실험자의 음상 정위감을 알아보기 위한 주관평가 방법과 결과를 기술한다.

4.1 주관평가 방법

주관평가는 중앙면상에 위치한 음원의 고도각을 -30° 부터 210° 까지 30° 간격으로 9개의 음원 위치에 대해 수행한다. 주관평가에 사용된 장치는 Fig. 4와 같다. 먼저, MATLABTM에서 가청주파수범위의 백색잡음과 측정된 혹은 모델링된 HRIR을 컨볼루션(convolution)하여 각각의 음원 위치에 해당하는 신호를 생성한다. 이 신호는 컴퓨터 사운드카드(Creative SB X-Fi Elite Pro)와 앰프(audio analog verdi settanta)를 거쳐 헤드폰(AKG K1000)에 전달된다. 피실험자는 헤드폰을 통해 소리를 듣고 그 소리가 들리는 방향(고도각)을 Fig. 5와 같은 MATLABTM

GUI 를 이용하여 기록한다. 9개의 음원 위치에 대해 각각 10번씩 반복하여 실험하였고, 총 90번의 신호는 임의의 순서로 배열되었다.

일반적으로 헤드폰과 고막 사이의 주파수영역 전달함수(headphone to ear-canal transfer function : HpTF)가 모든 주파수 대역에서 일정하지 않으므로 헤드폰을 이용한 청취 실험결과에 HpTF이 큰 영향을 미친다⁽²⁴⁾. 또한, HpTF가 개개인마다 다르며 헤드폰을 착용위치가 미세하게 변하더라도 HpTF은 크게 변한다⁽²⁵⁻²⁷⁾. 따라서 정확한 주관평가가 이루어지기 위해서는 HpTF를 정확하게 보상되어야 한다. 이에 이 연구에서는 각각의 피실험자가 헤드폰을 착용할 때마다 HpTF을 측정하여 보상해주었고, 이를 용이하게 하기 위하여 완전 개방형 헤드폰 (AKG K1000)을 사용하였다.

엄밀히 말하면, 중앙면상에서도 왼쪽과 오른쪽귀의 HRIRs는 서로 약간의 차이를 보이지만, 이 연구

에서는 왼쪽과 오른쪽 HRIRs의 대칭성을 가정하여 헤드폰 양쪽 모두 같은 소리를 내도록 하였다. Morimoto의 실험결과에 의하면, 한쪽 귀의 소리가 정확하게 재현된다면 반대쪽 귀를 막은 상태에서도 청취자가 고도각을 정확하게 인지할 수 있다⁽²⁸⁾. 따라서 위의 가정은 고도각 인지 결과에 큰 영향을 주지 않는다.

4.2 주관평가 결과

Fig. 6은 측정된 HRIRs와 모델링된 HRIRs를 이용하여 주관평가를 실시했을 때 각각의 피실험자가 인지한 음원 위치의 분포를 나타낸다. 각 패널의 가로축은 실제 들려준 음원의 위치를 나타내고 세로축은 피실험자가 인지한 음원의 위치를 나타낸다. 따라서 실선으로 나타난 대각선을 따라 응답이 분포하면, 피실험자가 음원의 위치를 정확하게 인지한 것을 의미한다. 원의 크기는 응답의 횟수에 비례한다. 첫 번째 행의 결과를 보면, 측정된 HRIR을 사용한 경우 모든 피실험자가 음원의 위치를 정확하게 인지함을 확인할 수 있다. 이는 이 연구에서 피실험자의 HRIRs가 정확하게 측정되었으며 실시한 주관평가 방법이 타당함을 뒷받침한다. 두 번째 행에 나타난 것처럼 12개 PCs로 HRIRs를 모델링한 경우, 피실험자들의 응답이 측정된 HRIR을 사용한 경우보다

Table 2 Percentage modeling error(% error) for each subject according to the number of PCs(k). The empirical mean of the CIPIC dataset is used

k	HS	KB	LS	CT	CH	CY
4	64.1	50.2	62.2	48.4	38.3	42.4
8	37.4	23.3	36.8	20.9	18.4	25.1
12	15.1	15.0	18.7	14.4	12.4	13.9

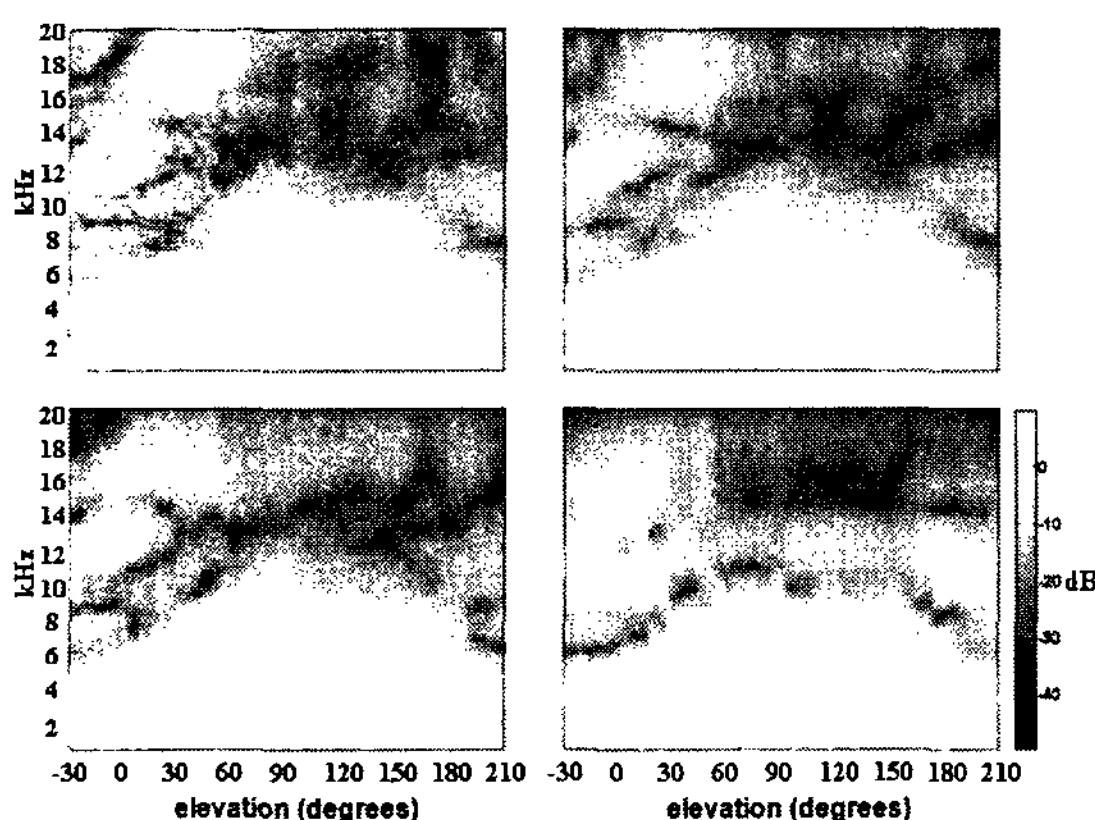


Fig. 3 Log magnitude of measured HRTFs in the median plane(left-top panel) from the left-ear of subject CH and modeled HRTFs obtained when 12 PCs(right-top panel), 8 PCs(left-bottom panel), and 4 PCs (right-bottom panel) are used



Fig. 4 Apparatus for the subjective listening tests

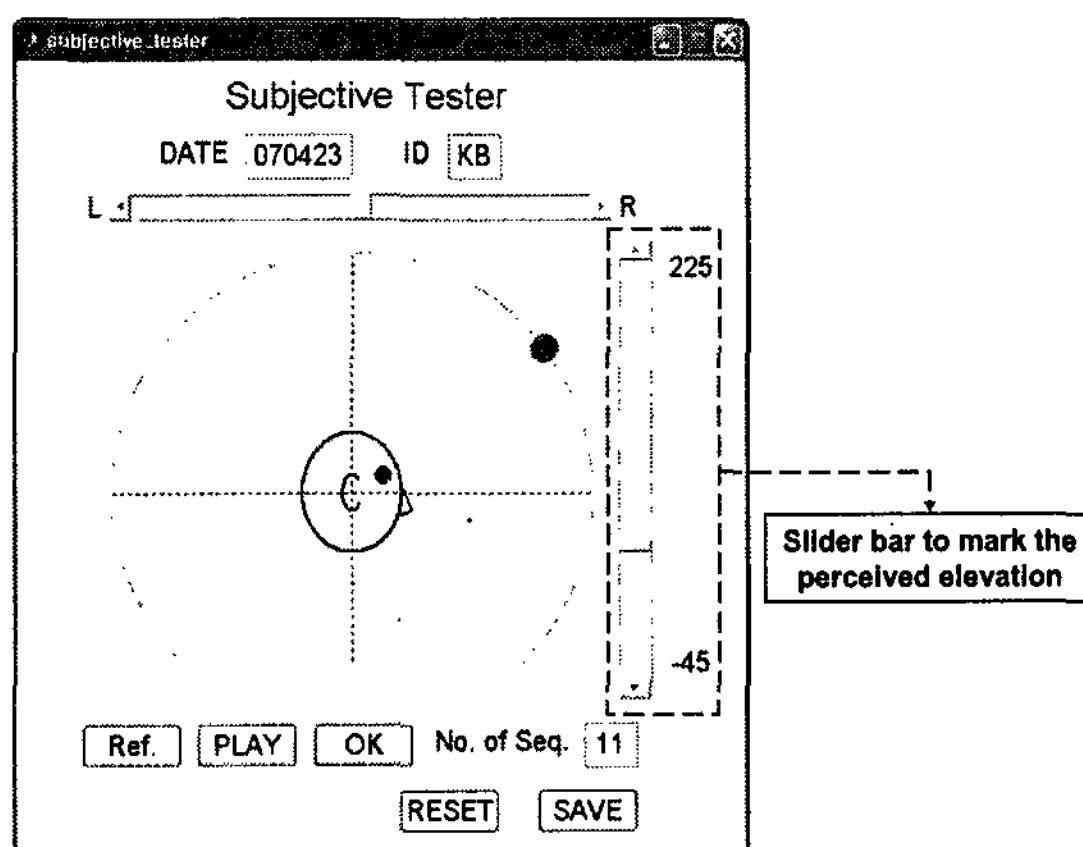


Fig. 5 MATALAB™ GUI for the listening tests

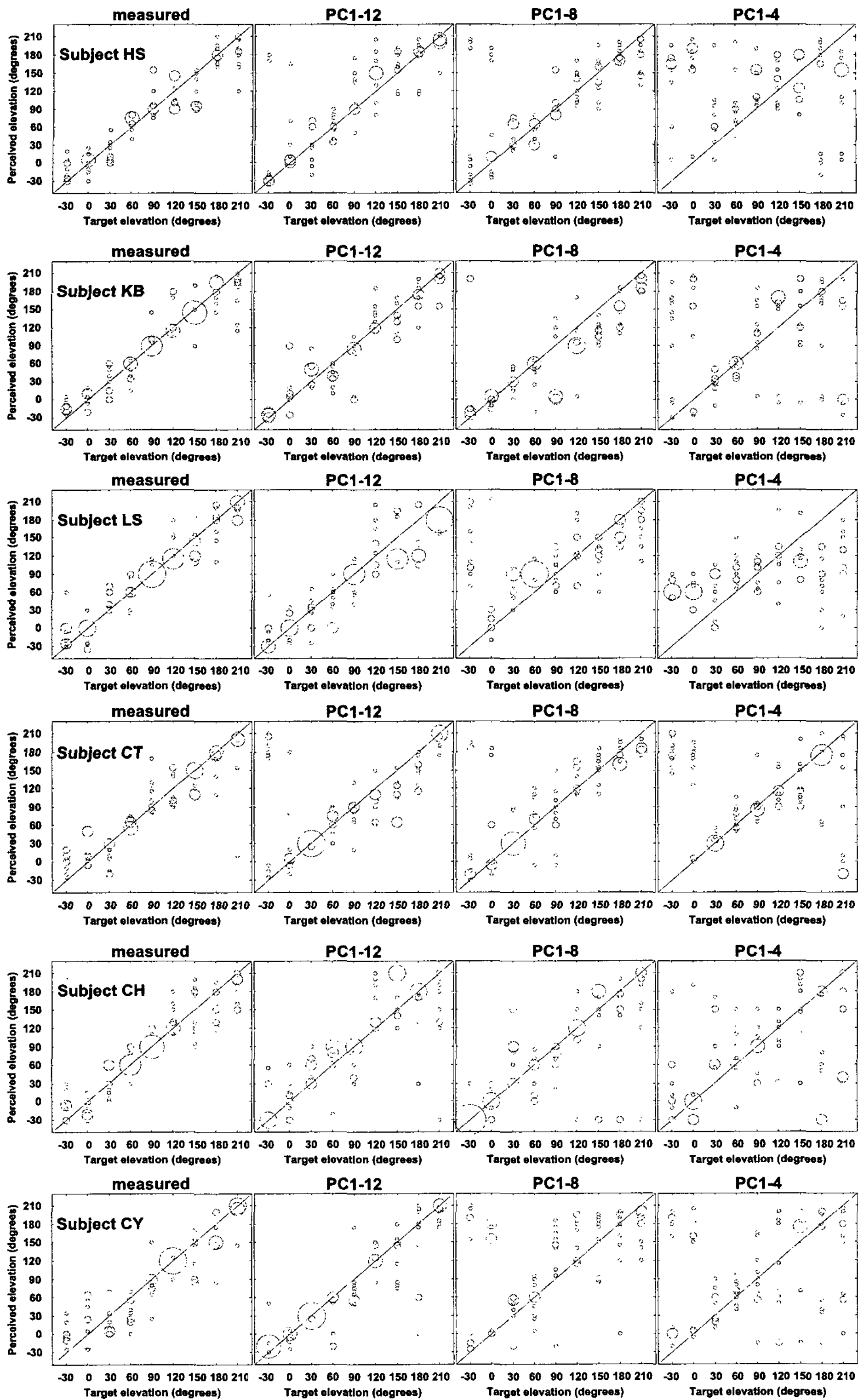


Fig. 6 Distributions of the responses of six subjects with the measured and modeled HRIRs. Each row corresponds to the responses of a single subject. Each column corresponds to the responses of six subjects with the measured or modeled HRIRs as denoted at the top of each column

Table 3 Mean values and standard deviations (in parentheses) of the localization errors for each subject in degrees

	Measured	PCs 1~12	PCs 1~8	PCs 1~4
HS	20.1 (18.6)	25.3 (32.5)	31.0 (38.1)	72.7 (60.6)
KB	16.6 (18.7)	20.0 (22.8)	29.4 (30.8)	60.6 (59.4)
LS	16.5 (17.7)	27.5 (21.9)	42.7 (41.3)	57.4 (38.6)
CT	20.9 (24.1)	31.8 (42.0)	28.9 (36.8)	53.4 (61.5)
CH	22.2 (22.1)	30.7 (33.3)	31.5 (42.6)	48.2 (52.3)
CY	23.3 (21.9)	21.4 (32.2)	49.8 (52.0)	58.4 (57.4)

조금 더 흩어져있고 앞-뒤 혼돈이 약간 발생하지만 대체적으로 모든 피실험자가 음원의 위치를 정확하게 인지함을 보인다. 8개 PCs로 HRIR을 모델링한 경우(세 번째 행)는 이전의 두 경우보다 인지 성능이 떨어짐을 확인할 수 있는데, 특히 앞-뒤 혼돈이 빈번하게 일어난다. 네 번째 행의 결과는 4개 PC로는 피실험자가 음원의 위치를 인지할 수 없고 HRIRs가 부정확하게 모델링됨을 보여준다.

각각의 피실험자에 대해서 음원 위치의 인지 성능을 정량적으로 알아보기 위해 인지오차를 식(4)와 같이 정의한다.

$$e_i = |T_i - P_i| \quad (i = 1, 2, \dots, 90) \quad (4)$$

여기서 T_i 와 P_i 는 각각 i 번째 신호의 실제 음원위치와 피실험자가 인지한 위치를 나타낸다. 각각의 피실험자에 대해 인지오차의 평균과 표준편차를 Table 3에 나타내었다. 측정된 HRIR을 사용한 경우 평균 인지오차가 16.6°에서 23.3°까지 분포하는데 이는 Iida 등⁽²⁹⁾과 Morimoto 등⁽³⁰⁾이 실시한 실험결과와 비슷한 수치이다. 모델링에 사용되는 PC수가 감소할수록 인지오차가 커짐을 확인할 수 있는데, 그 정도는 각각의 피실험자마다 다른 양상을 보인다. 피실험자 HS, KB, CH의 경우 PC 수가 12개에서 8개로 감소할 때 인지오차가 점진적으로 증가하지만 PC 4개를 사용할 때 오차가 급격히 증가함을 보인다. 측정된 HRIR을 사용한 경우와 12개 PCs로 모델링한 경우의 결과를 비교해볼 때, 전반적으로 평균 인지오차와 표준편차가 증가하지만 피실험자 CY는 평균 인지오차는 감소한 반면 표준편차는 증가한 결과를

보였다. 그러나 전반적으로 모든 피실험자가 12개 PCs로 모델링할 때 측정된 HRIR을 사용한 경우와 비교하여 비교적 정확한 인지성능을 보인다.

5. 토 의

이 연구에서 얻은 12개의 보편적인 기저함수를 사용하여 HRIR을 모델링할 경우, 저장 공간의 측면에서 큰 이점이 있다. Fig. 7은 모델링하고자 하는 HRIR의 수에 따른 원래 HRIRs의 메모리 크기와 12개의 PCs로 모델링했을 때 메모리 크기의 비를 보여준다. 모델링하고자 하는 HRIR의 수가 증가할수록 두 메모리 크기의 비가 줄어드는 것을 확인할 수 있으며, 그 비는 약 18%로 수렴한다. 즉, 이 연구에서 얻은 12개의 보편적인 기저함수를 사용하여 모델링하면, 최대 82%의 메모리 절감 효과를 기대할 수 있다. 또한, 모델링해야 하는 HRIRs데이터가 바뀌거나 추가되더라도 PCA를 다시 수행하여 새로운 기저함수를 구할 필요 없이, 이 연구에서 구한 12개의 보편적인 기저함수를 사용하여 모델링 가능하므로 추가적인 계산량이 적다는 장점이 있다.

3.2절에서 12개 PCs로 6명 피실험자의 HRIR을 모델링할 때, 모델링 오차가 CIPIC HRTF database로 12개 PCs를 얻을 때의 5% 오차 범위보다 큰 것을 언급하였다. 이는 Fig. 8에 나타난 것처럼 이 연구에 참여한 6명의 피실험자 HRIRs의 평균값이 CIPIC HRTF database의 그것과 다르기 때문이며, 이는 실험조건, 방법, 실험기구 등의 차이에 기인하는 것이다. 따라서 실제 모델링하고자 하는 데이터의 평균값을 사용하면 수치적인 모델링 오차는 줄어들 수 있다. Table 4는 CIPIC HRTF database에서 얻은 12개 PC와 이 연구에서 측정한 6명 피실험자 HRIRs의 평균을 사용하여 그들의 HRIRs를 모델링한 결과이다. Table 2의 결과(CIPIC dataset의 평균을 사용한 경우)와 비교할 때, 모델링 성능이 향상되는 것을 확인할 수 있다. 12개의 PCs를 사용할 경우 6명에 대한 모델링 오차 평균이 14.9%에서 7.7%로 줄어든다.

그러나 HRIRs의 평균값은 방향성에 대한 정보를 포함하고 있지는 않으므로 청취자의 음원위치 인지 성능에는 큰 영향을 미치지 않는 것으로 판단된다. 4.2절의 주관평가 결과에 나타난 것과 같이

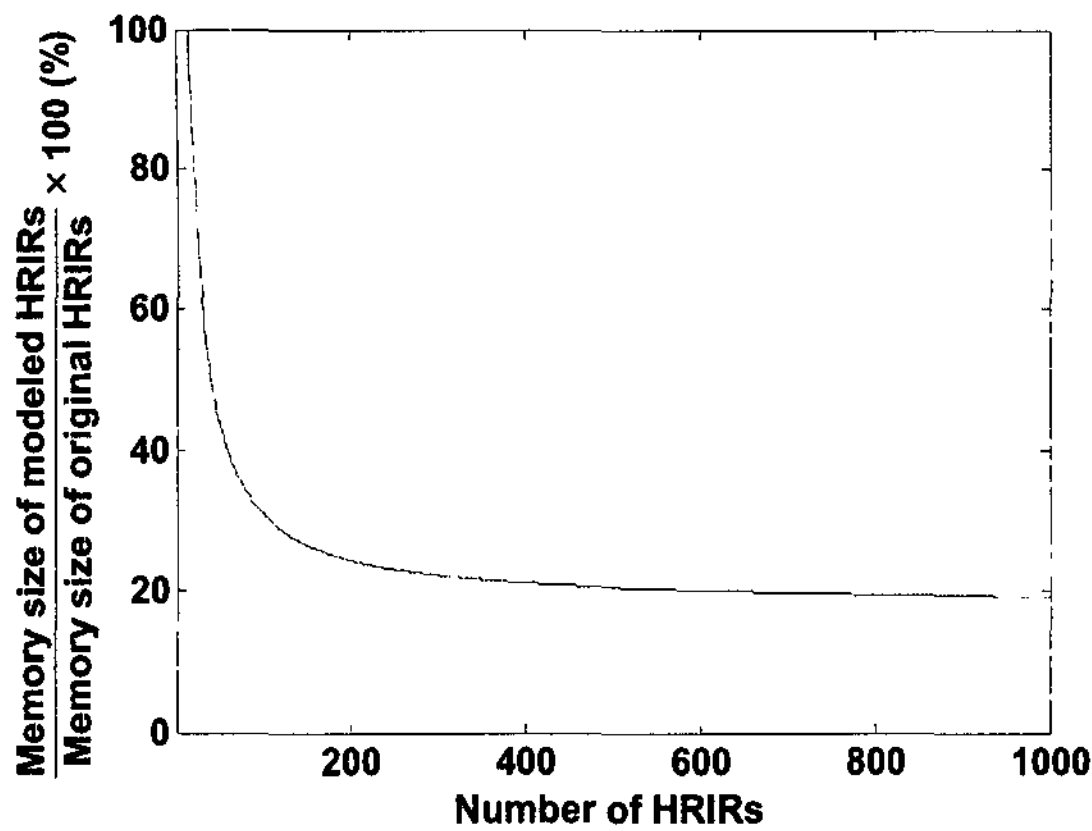


Fig. 7 Ratio of memory size of modeled HRIRs to the one of original HRIRs as a function of the number of HRIRs when the 12 PCs are used for modeling

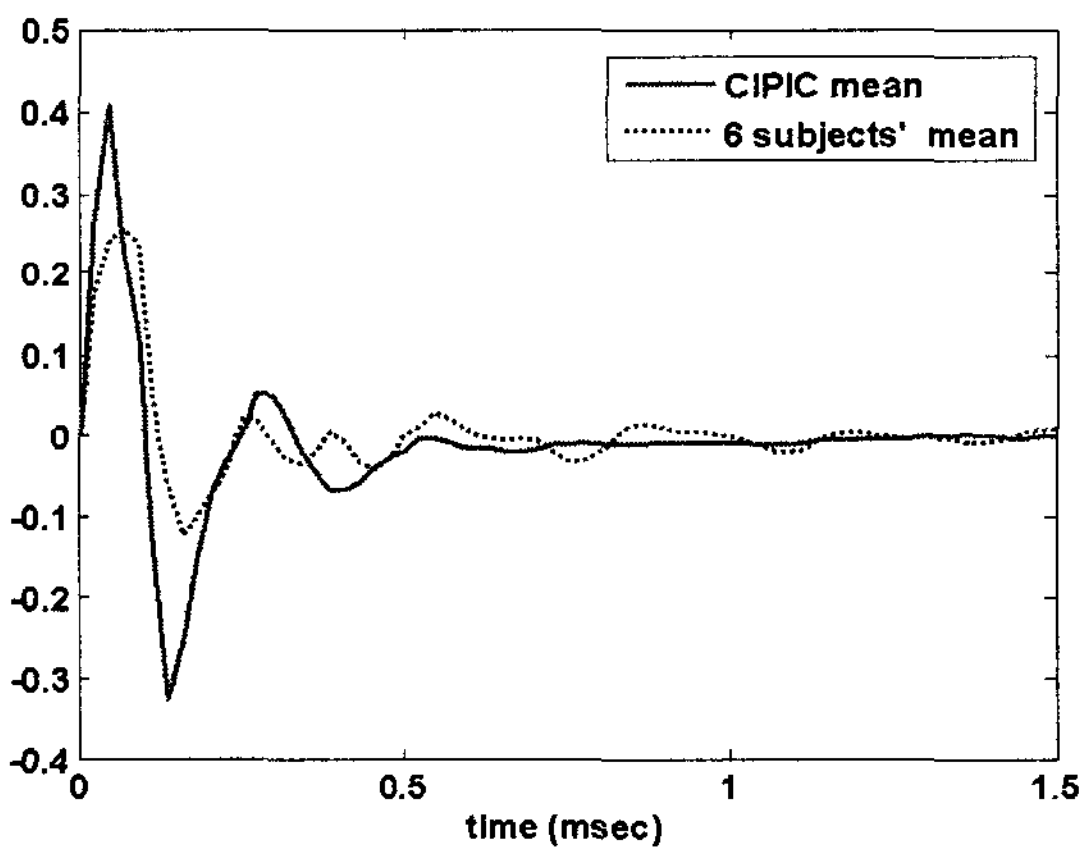


Fig. 8 Empirical means of CIPIC HRTF database and 6 subjects' data

Table 4 Percentage modeling error (% error) for each subject according to the number of PCs(k). The empirical mean of the 6 subjects' data is used

k	HS	KB	LS	CT	CH	CY
4	20.2	20.0	26.4	15.9	15.6	11.8
8	15.2	13.2	18.6	10.7	10.3	9.2
12	8.5	7.7	12.0	6.4	5.9	5.8

CIPIC HRTF database의 평균값을 사용하더라도 12개 PCs를 사용하여 HRIR을 모델링할 때 정확한 인지 성능을 보인 결과는 이를 뒷받침한다.

6. 결 론

이 연구에서는 CIPIC HRTF database의 중앙면 HRIRs를 이용한 PCA로 부터 구한 기저함수가 임의의 개인의 중앙면 HRIR을 모델링하는데 사용될 수 있는 보편적인 기저함수가 됨을 수치적인 오차분석과 주관평가를 통해 밝혔다. 12개 PCs를 사용할 경우, 최소자승오차 관점에서 평균 7.7%의 오차로 모델링할 수 있으며 개인의 측정된 HRIRs를 사용한 경우와 비슷한 음상 정위감을 보인다. PC 수가 감소할수록 모델링 정확도와 음상 정위감은 떨어진다. 이 연구에서 얻은 12개의 보편적인 기저함수를 사용함으로써 최대 82%의 메모리 절감 효과를 기대할 수 있다. 향후 많은 양의 HRIRs 데이터를 다루는 분야에서 이 연구가 모델링 효율 및 저장 공간의 측면에서 이점이 있을 것으로 판단된다.

후 기

이 연구는 한국과학재단의 국가지정연구실사업 (ROA-2005-000-10112-0)과 BK21 사업의 지원으로 수행되었습니다.

참 고 문 헌

- (1) Blauert, J., 1996, Spatial Hearing, MIT Press, Cambridge, MA.
- (2) Brungart, D. S. and Rabinowitz, W. M., 1999, "Auditory Localization of Nearby Sources. Head-related Transfer Functions", J. Acoust. Soc. Am., Vol. 106, pp. 1465~1479.
- (3) Cheng, C. I. and Wakefield, G. H., 2001, "Introduction to Head-related Transfer Functions (HRTFs): Representations of HRTFs in Time, Frequency, and Space", J. Audio Eng. Soc., Vol. 49, pp. 231~248.
- (4) Algazi, V. R., Avendano, C. and Duda, R. O., 2001a, "Elevation Localization and Head-related Transfer Function Analysis at Low Frequencies", J. Acoust. Soc. Am., Vol. 109, pp. 1110~1122.
- (5) Gardner, M. B. and Gardner, R. S., 1973,

“Problem of Localization in the Median Plane: Effect of Pinna Cavity Occlusion”, *J. Acoust. Soc. Am.*, Vol. 53, pp. 400~408.

(6) Hebrank, J. and Wright, D., 1974, “Spectral Cues Used in the Localization of Sound Sources on the Median Plane”, *J. Acoust. Soc. Am.*, Vol. 56, pp. 1829~1834.

(7) Middlebrooks, J. C. and Green, D. M., 1992, “Observations on a Principal Components Analysis of Head-related Transfer Functions”, *J. Acoust. Soc. Am.*, Vol. 92, pp. 597~599.

(8) Raykar, V. C., Duraiswami, R. and Yegumnarayana, B., 2005, “Extracting the Frequencies of the Pinna Spectral Notches in Measured Head Related Impulse Responses”, *J. Acoust. Soc. Am.*, Vol. 118, pp. 364~374.

(9) Shaw, E. A. G. and Teranishi, R., 1968, “Sound Pressure Generated in an External-ear Replica and Real Human Ears by a Nearby Point Source”, *J. Acoust. Soc. Am.*, Vol. 44, pp. 240~249.

(10) Kim, S., Kim, K., Bae, K., Choi, S. and Park, M., 2005, “Headphone-based Multi-channel 3D Sound Generation Using HRTF”, *J. IEK-SP*, Vol. 42, pp. 71~77.

(11) Blommer, M. A. and Wakefield, G. H., 1997, “Pole-zero Approximations for Head-related Transfer Functions Using a Logarithmic Error Criterion”, *IEEE Trans. Speech Audio Process.*, Vol. 5, pp. 278~287.

(12) Brown, C. P. and Duda, R. O., 1998, “A Structural Model for Binaural Sound Synthesis”, *IEEE Trans. Speech Audio Process.*, Vol. 6, pp. 476~488.

(13) Durant, E. A. and Wakefield, G. H., 2002, “Efficient Model Fitting Using a Genetic Algorithm: Pole-zero Approximations of HRTFs”, *IEEE Trans. Speech Audio Process.*, Vol. 10, pp. 18~27.

(14) Haneda, Y., Makino, S., Kaneda, Y. and Kitawaki, N., 1999, “Common-acoustic-pole and Zero Modelling of Head-related Transfer Functions”, *IEEE Trans. Speech Audio Process.*, Vol. 7, pp. 188~196.

(15) Kulkarni, A. and Colburn, H. S., 2004, “Infinite-impulse-response Models of the Head-related Transfer Function”, *J. Acoust. Soc. Am.*, Vol. 115, pp. 1714~1728.

(16) Mackenzie, J., Huopaniemi, J. and Välimäki, V., 1997, “Low-order Modelling of Head-related Transfer Functions Using Balanced Model Truncation”, *IEEE Signal Process. Letters*, Vol. 4, pp. 39~41.

(17) Yim, J., Kim, C. and Kang, S., 1996, “Lower-order ARMA Modeling of Head-related Transfer Functions for Sound-field Synthesis System”, *J. Acoust. Soc. Kor.*, Vol. 15, pp. 37~44.

(18) Martens, W. L., 1987, “Principal Components Analysis and Resynthesis of Spectral Cues to Perceived Direction”, in *Proceedings of International Computer Music Conference*, San Francisco, CA, pp. 274~281.

(19) Kistler, D. J. and Wightman, F. L., 1992, “A Model of Head-related Transfer Functions Based on Principal Components Analysis and Minimum-phase Reconstruction”, *J. Acoust. Soc. Am.*, Vol. 91, pp. 1637~1647.

(20) Wu, Z., Chan, F. H. Y., Lam, F. K. and Chan, J. C. K., 1997, “A Time Domain Binaural Model Based on Spatial Feature Extraction for the Head-related Transfer Function”, *J. Acoust. Soc. Am.*, Vol. 102, pp. 2211~2218.

(21) Hwang, S. and Park, Y., 2007, “Median HRIR Customization Via Principal Components Analysis”, *Transactions of the Korean Society for Noise and Vibration Engineering*, Vol. 17, No. 7, pp. 638~648.

(22) Hwang, S. and Park, Y., 2008, “Interpretations on Principal Components Analysis of Head-related Impulse Responses in the Median Plane”, *J. Acoust. Soc. Am.*, Vol. 123, pp. EL65~71.

(23) Algazi, V. R., Duda, R. O., Thompson, D. M. and Avendano, C., 2001b, “The CIPIC HRTF Database”, in *Proceedings of the 2001 IEEE ASSP Workshop on Applications of Signal Processing to Audio and Acoustics*, New Paltz, NY, pp. 99~102.

(24) Møller, H., 1992, “Fundamentals of Binaural

Technology”, Applied Acoustics, Vol. 36, pp. 171~218.

(25) Møller, H., Hammershøi, D., Jensen, C. B. and Sørensen, M. F., 1995a, “Transfer Characteristics of Headphones Measured on Human Ears”, J. Audio Eng. Soc., Vol. 43, pp. 203~217.

(26) Møller, H., Jensen, C. B., Hammershøi, D. and Sørensen, M. F., 1995b, “Design Criteria for Headphones”, J. Audio Eng. Soc., Vol. 43, pp. 218~232.

(27) Kulkarni, A. and Colburn, H. S., 2000, “Variability in the Characterization of the Headphone Transfer-function”, J. Acoust. Soc. Am., Vol. 107, pp. 1071~1074.

(28) Morimoto, M., 2001, “The Contribution of Two Ears to the Perception of Vertical Angle in Sagittal Planes”, J. Acoust. Soc. Am., Vol. 109, pp. 1596~1603.

(29) Iida, K., Itoh, M., Itagaki, A. and Morimoto, M., 2007, “Median Plane Localization Using a Parametric Model of the Head-related Transfer Function Based on Spectral Cues”, Applied Acoustics, Vol. 68, pp. 835~850.

(30) Morimoto, M., Iida, K. and Itoh, M., 2003, “Upper Hemisphere Sound Localization Using Head-related Transfer Functions in the Median Plane and Interaural Differences,” Acoustical Sci. Technol., Vol. 24, pp. 267~275.