

특정 피실험자의 맞춤형 머리전달함수에 대한 불특정 다수를 상대로 한 주관평가

A subjective listening test of specific subject's customized head-related transfer functions
on various subjects

조현* · 박영진† · 박윤식*

Hyun Jo, Youngjin Park and Youn-sik Park

1. 서 론

머리전달함수(HRTF, Head-Related Transfer Function)란 음원의 음압과 사람의 고막 위치에서의 음압 간의 음향학적 전달함수로, 사람이 음원위치추정에 활용한다고 알려진 주요 인자인 양이간 시간차(ITD, Interaural Time Difference), 양이간 레벨차(ILD, Interaural Level Difference), 스펙트럼 인자 등을 포함한다.^(1,2)

머리전달함수를 활용하여 가상입체음향(VAD, Virtual Auditory Display)을 구현하는 방식은 머리전달함수의 특성이 개인별로 다르므로 인해 머리전달함수의 맞춤화가 다양한 방식으로 연구되어왔다. 그 중, Hwang 이 제안한 머리전달함수의 맞춤화 기법은 사용자에 맞는 스펙트럼 인자를 개개인 별로 제공하기 위하여 청취자의 개인별 머리전달함수를 직접적으로 측정하지 않고도 9 개의 파라미터 튜닝을 통하여 가상입체음향 생성에 적합한 머리전달함수를 찾아냄으로써 머리전달함수 측정의 시간적 금전적 비용을 줄일 수 있었다.⁽³⁾ Hwang 의 선행연구에서 9 명의 맞춤화된 머리전달함수의 주파수 스펙트럼은 피실험자 LY, HS 를 제외하고 90 도 이하의 고도각 위치에 대해서 그 형상이 비슷하다. 개개인 별 맞춤형 머리전달함수는 특히 Iida 가 제안했던 1 피크, 2 노치를 활용한 머리전달함수의 모델링과 유사하다.⁽⁴⁾ 고도각 별로 5kHz 부근의 피크를 공유하면서 0 도 위치의 고도각에서는 2 개의 노치를 활용하고, 고도각 40 도 부근에서는 1 개의 노치를 활용한다면 개개인별 맞춤형 머리전달함수를 잘 모델링할 수 있다. (그림 1. 위) 즉, 정성적인 분석만으로 평가한다면 90 도 이하의 고도각 위치에서 모든 사람들의 맞춤형 머리전달함수를 동일시 볼 수 있다는 의미이다. 이는 역설적으로 Hwang 이 제안한 방식으로 머리전달함수를 맞춤화할 경우, 개인별 노치의 양상이 비슷하기 때문에 대표적인 피실험자의 맞춤형 머리전달함수만 사용하더라도 다른 대부분의 피실험자에게 비슷한 고도감을 구현할 수 있다는 의미이기도 하다. 따라서 본 연구에서는 이러한 가정이 맞는지 주관평가를 통하여 확인하였다. Hwang 이 제안하였던 방법을 바탕으로 피실험자 CH 의 맞춤형 머리전달함수를 다른 피실

험자에게 들려주었을 때 얼마만큼의 훌륭한 고도감 구현이 가능한지 평가하였다. 또한 추가적으로 CIPIC 머리전달함수 데이터 베이스에 측정된 피실험자 별 머리전달함수의 시간영역 평균을 구하여 이를 주관평가 하였다. 시간영역 평균의 경우에는 Iida 가 제안하였던 뚜렷한 노치 및 피크가 존재하고 그 양상이 Hwang 의 방법으로 구한 개인별 맞춤형 머리전달함수와 유사하기 때문에 이를 평가하였다.

2. 주관평가

(1) 주관평가 절차

주관평가는 Hwang 이 수행한 방식과 동일하게 정지한 음원에 대한 주관평가를 중앙면상 고도각 0 도부터 180 도 까지 30 도 각도 차이로 총 7 위치에 대해 평가하였다.⁽³⁾ Hwang 의 연구에서 피실험자였던 피실험자 CY, KB, KD, LS, PJ 가 참가하였다. 주관평가는 2 회 실시하였으며, 1 회 당 총 70 번의 신호(7 각도, 10 번의 반복)에 대해 평가하였다. 1 회의 평가가 완료되는 데는 ‘헤드폰 특성 보상 시간’, ‘주관평가에 대한 설명의 시간’ 등을 포함하여 약 30 분 정도의 시간이 소요되었다.

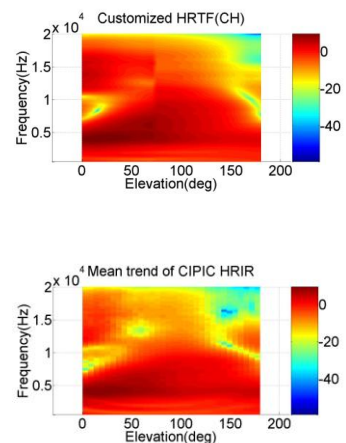


그림 1. Hwang 의 방법으로부터 계산된 피실험자 CH 의 맞춤형 머리전달함수. (위) CIPIC 머리전달함수 데이터 베이스의 시간영역 평균값. (아래)

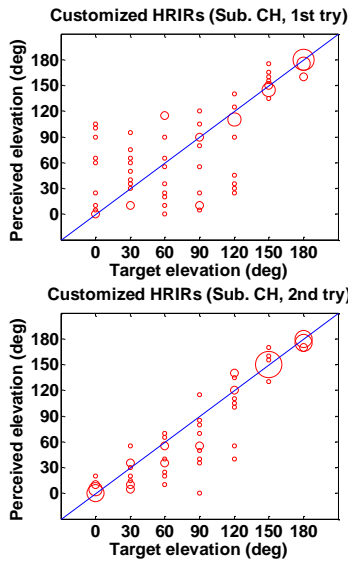


그림 2. 피실험자 KB 에게 피실험자 CH 의 맞춤형 머리전달함수를 들려주었을 때의 고도각 인지 결과.

(2) 주관평가 결과

그림 2는 피실험자 KB 의 주관평가 결과이다. 앞쪽에 위치한 음원에 대해 노치패턴이 피실험자 별로 유사하다 할지라도 정확한 고도각 인지가 되지 않는 것을 알 수 있다. (KB 의 첫 번째 실험) 표 1.은 실제 각도와 인지한 각도 간의 절대 오차를 Hwang 의 실험결과⁽³⁾와 비교해서 정리한 것이다. CH 의 맞춤형 머리전달함수를 피실험자에게 들려주면, Kemar 의 머리전달함수를 들려주었을 때보다는 향상된 성능을 보이지만, 피실험자 자신의 머리전달함수나 피실험자 자신의 맞춤형 머리전달함수를 들었을 때의 성능이 확보되지 않음을 확인할 수 있다.

표 2.는 고도각 오차와 앞뒤혼돈 확률 (front-back confusion rate)에 대한 결과이다. 고도각 오차는 절대 오차에서 앞뒤혼돈으로 인해 발생한 고도각 오차 값을 보상한 값이다. 특히 피실험자 CY, LS, PJ 가 CH 의 맞춤형 머리전달함수를 청취하였을 경우, 앞뒤혼돈 확률이 피실험자가 개인의 맞춤형 머리전달함수를 들었을 경우에 비해서 증가하는 것을 확인할 수 있다. 한편, 시간영역 평균 머리전달함수를 사용한 경우는 대부분의 경우에서 절대 오차, 고 표 1. 실제 각도와 인지한 각도간의 절대 오차. 모든 각도에 대한 평균 값. (괄호 안의 수치는 분산 값)

	Individual HRIRs	Customized HRIRs	Kemar HRIRs	CH's customized HRIR (1)	CH's customized HRIR (2)	Mean HRIR (1)	Mean HRIR (2)
CY	18.7 (17.8)	16.8 (17.1)	57.9 (45.1)	31.0 (23.9)	33.7 (35.4)	54.1 (48.5)	51.0 (54.7)
KB	27.7 (28.1)	13.9 (17.2)	40.9 (37.8)	29.1 (30.7)	16.1 (19.1)	32.6 (28.9)	38.8 (37.9)
KD	29.2 (26.8)	35.1 (31.2)	42.6 (46.2)	30.0 (34.9)	32.9 (32.7)	30.6 (29.7)	38.6 (38.0)
LS	16.9 (14.8)	20.5 (15.0)	30.6 (27.6)	31.2 (23.9)	27.9 (19.9)	30.1 (22.9)	26.3 (23.2)
PJ	25.1 (21.3)	32.9 (27.5)	43.4 (38.2)	42.4 (42.3)	35.6 (40.1)	28.3 (26.3)	47.3 (46.1)

표 2. 모든 각도에 대한 고도각 오차. (괄호 안의 수치는 앞뒤혼돈 확률)

	Individual HRIRs	Customized HRIRs	Kemar HRIRs	CH's customized HRIR (1)	CH's customized HRIR (2)	Mean HRIR (1)	Mean HRIR (2)
CY	18.4 (1.4)	16.2 (1.4)	28.3 (40.0)	25.4 (14.3)	23.6 (18.6)	24.5 (34.3)	17.4 (35.7)
KB	24.9 (4.3)	9.6 (7.1)	26.3 (17.1)	23.4 (12.9)	14.4 (2.9)	24.4 (14.3)	22.6 (25.7)
KD	25.6 (10.0)	26.6 (14.3)	23.6 (21.4)	22.3 (11.4)	27.7 (8.6)	24.6 (12.9)	26.8 (18.6)
LS	16.5 (1.4)	19.8 (2.9)	23.4 (14.3)	27.1 (11.4)	22.9 (11.4)	26.0 (14.3)	22.7 (10.0)
PJ	24.8 (2.9)	30.4 (4.3)	23.6 (31.4)	26.7 (20.0)	19.9 (20.0)	22.9 (10.0)	22.9 (25.7)

도각 오차, 앞뒤혼돈 확률이 Kemar 의 머리전달함수를 사용했을 때와 그 수준이 비슷하다.

3. 결론

Hwang 이 제안한 방식으로 머리전달함수를 맞춤화할 경우, 개인별 노치의 양상이 비슷하나, 피실험자 CH 의 맞춤형 머리전달함수를 대표 값으로 사용하여 주관평가를 실시한 결과, 고도각의 절대오차, 앞뒤혼돈 확률이 본인들의 개인별 머리전달함수 혹은 본인들의 맞춤형 머리전달함수를 사용한 경우에 비해 현저히 증가하였다. 따라서 가상입체음향의 성능을 Hwang 이 제안했던 수준으로 구현하고자 하는 경우에는 개개인 별 파라미터 튜닝이 필요하다.

후 기

본 연구는 두뇌 한국 21프로젝트, 지식경제부 및 정보통신연구진흥원의 IT핵심기술개발사업[2008-F-044-01, 전자파, 음향 및 건물 환경을 개선하는 지능형 건설 IT 융합 신기술 개발] 및 한국과학재단을 통해 교육과학기술부의 국가 지정연구실 사업으로부터 지원받아 수행되었습니다 (R0A-2005-000-10112-0).

참고문헌

- (1) D. R. Begault, 1994, *3-D sound for virtual reality and multimedia*, Academic press.
- (2) J. Hebrand and D. Wright, 1974, "Spectral cues used in the localization of sound sources on the median plane," *J. Acoust. Soc. Amer.*, Vol. 56, pp. 1829-1834.
- (3) S. Hwang, Y. Park, and Y-S. Park, 2008, "Modeling and customization of head-related impulse responses based on general basis functions in time domain," *Acta Acustica united with Acustica*, Vol.94, pp. 965-980.
- (4) K. Iida, M. Itoh, A. Itagaki, and M. Morimoto, 2007, "Median plane localization using a parametric model of the head-related transfer function based on spectral cues," *Applied Acoustics*, Vol. 68, pp. 851-863.