

이진 탐색 트리를 이용한 개인화된 머리 전달 함수의 탐색

Selection of Personalized Head Related Transfer Function Using a Binary Search Tree

이 기 승*, 이 석 필**
(Ki-Seung Lee*, Seok-Pil Lee**)

*건국대학교 정보통신대학 전자공학부, **전자부품연구원 방송통신융합 연구센터
(접수일자: 2009년 3월 17일; 수정일자: 2009년 5월 6일; 채택일자: 2009년 6월 11일)

가상 음원의 공간화에 중요한 역할을 담당하는 머리전달 함수는 개인 별로 각기 다른 특성을 갖는다. 개인화된 머리 전달 함수를 얻기 위해서는 많은 시간과 특정한 장치를 필요로하며, 이에 따라 타인으로부터 얻어진 머리 전달 함수를 청취자의 특성에 맞도록 개인화 시키는 방법이 주로 사용된다. 본 논문에서는 머리 전달 함수의 개인화 방법으로 여러 개의 미리 측정된 머리 전달 함수 중에서 사용자에게 가장 적합한 머리 전달 함수를 찾는 방법을 제안하였다. 제안된 방법은 청취자에게 적합한 머리 전달 함수를 빠르고 정확하게 찾기 위해 이진 탐색 트리를 이용한 분할 기법을 사용하였다. 분할 시 왜곡 정도는 사람이 귀에 방향감의 차이를 가장 잘 인지하는 왜곡 척도를 실험적으로 결정하였다. 캘리포니아 데이비스 대학에서 제공하는 CIPIC 머리 전달 함수 데이터 베이스를 이용하여 본 기법의 유용성을 평가하였다. 10명의 실험자가 참여한 청취 테스트에서 제안된 기법을 통해 탐색된 머리 전달 함수는 청취자 본래의 머리 전달 함수와 비교하여 가상 음원의 방향감에 있어서 매우 근접한 결과를 나타내었으며, 기존의 해부학적인 파라미터를 이용한 머리 전달 함수의 탐색 기법과 비교하여 우수한 성능을 나타내었다.

핵심용어: 머리 전달 함수, 개인화, 이진 탐색 트리

투고분야: 전기음향 분야 (3.1)

The head-related transfer function (HRTF), which has an important role in virtual sound localization has different characteristics across the subjects. Measuring HRTF is very time-consuming and requires a set of specific apparatus. Accordingly, HRTF customization is often employed. In this paper, we propose a method to search an adequate HRTF from a set of the HRTFs. To achieve rapid and reliable customization of HRTF, all HRTFs in the database are partitioned, where a binary search tree was employed. The distortion measurement adopted in HRTF partitioning was determined in a heuristic way, which predicts the differences in perceived sound location well. The UC-Davis CIPIC HRTF database set was used to evaluate the effectiveness of the proposed method. In the listening test, where 10 subjects were participated, the stimuli filtered by the HRTF obtained by the proposed method were closer to those by the personalized HRTF in terms of sound localization. Moreover, performance of the proposed method was shown to be superior to the previous customization method, where the HRTF is selected by using anthropometric data.

Keywords: Head-related transfer function, Customization, Binary search tree

ASK subject classification: Electro-Acoustics (3.1)

1. 서론

머리전달함수 (Head Related Transfer Function, HRTF) 는 자유 공간 상에서 음원 (sound source)과 귀 간의 경로를 모델링하는 함수로서, 음상의 공간화 (sound localization)

과 관련된 중요한 음향 파라미터이다 [1]. HRTF는 음원과 귀 사이의 공간 경로에 포함되는 물리적인 구조물들 즉, 두상, 귓바퀴의 모양에 따라 그 특성이 결정되며 머리의 중심부와 음원간의 방위각 (azimuth) 및 고도각 (elevation)에 따라 각기 다른 주파수 응답을 갖는 것으로 알려져 있다 [2]. HRTF는 헤드폰과 같은 양이 재생 환경에서 음원의 방향성을 포함하는 공간감을 표현하는데 널리 사용되고 있다.

책임저자: 이 기 승 (kseung@konkuk.ac.kr)
143-701 서울특별시 광진구 화양동 1번지
건국대학교 정보통신대학 전자공학과 1417호
(전화: 02-450-3489; 팩스: 02-3437-5235)

일반적으로 청취자 본래의 HRTF와는 다른 HRTF를 통해 얻어진 소리는 그 방향감이 변화되는데, 이는 HRTF가 청취자 개개인의 머리 크기, 모양, 귓바퀴의 모양에 따라 각기 다르게 나타나기 때문이다. 따라서 음원의 정확한 공간화를 위해서는 개인별로 측정된 HRTF를 사용해야 한다. HRTF의 측정은 무향실 (anechoic chamber), 양이 마이크로폰, 스피커 연동 도구 등과 같은 특별한 환경과 도구를 필요로 하며, 따라서 일반인이 측정하는 것은 거의 불가능하다. 이와 같은 이유로, 개인화된 HRTF를 사용하는 대신, 이미 측정된 HRTF를 사용자의 특성에 맞게 적절히 변경시키거나 여러 HRTF 중에 사용자에게 가장 적합한 HRTF를 선택하는 방법이 주로 이용되고 있다.

HRTF의 특성을 변경시키는 방법으로 Middlebrooks 은 주파수 영역에서 HRTF를 청취자의 귀 모양에 따라 압축/신장 기법을 제안하였으며 [3], Lee 등은 전·후방에 해당하는 HRTF에 각기 다른 가중치를 도입하여 비개인화된 HRTF의 사용에 따른 전·후방 혼란 (front-back confusion) 문제를 해결하고자 하였다 [4]. 이와 유사한 방법으로 Zhang 등은 가중치를 지수함수화 된 HRTF를 사용하였다 [5].

HRTF가 귀와 머리 모양에 따라 결정된다는 사실을 바탕으로, 귀와 머리의 해부학적 파라미터를 이용하여 HRTF를 개인화하는 방법이 제안되었다. Hu 등은 다중 선형 회귀 방법을 이용하여 해부학적 파라미터로부터 HRTF 변수를 추정하는 방법을 제안하였다 [6]. 복수개의 HRTF에서 청취자에게 가장 적합한 HRTF를 선택하는 방법으로 Zotkin은 해부학적인 파라미터와 이에 대응하는 HRTF가 함께 제공되는 HRTF 데이터 베이스를 이용하여 해부학적 파라미터의 유클리디언 거리가 가장 적게 나타나는 HRTF를 선택하는 방법을 제안하였다 [7]. Hwang 등은 HRTF를 주요성분 (principle component)으로 분석하고 각 주요 성분에 대한 크기를 사용자 스스로 조절하여 적합한 HRTF를 얻도록 하였다 [8].

사용자에게 가장 적합한 HRTF를 선택하는 방법으로 본 논문에서는 CIPIC HRTF 데이터 베이스 [9]에 포함된 각각의 HRTF를 통해 합성된 소리를 직접 듣고 사용자가 선택하는 방법을 제안하였다. 이때 CIPIC 데이터베이스에 포함된 HRTF의 개수는 비교적 많기 때문에 모든 HRTF로 합성된 소리를 일일이 듣고 선택하는 것은 청취자에게 어려운 일이다. 따라서 본 논문에서는 데이터베이스에 포함된 각 HRTF를 이진 트리 (binary tree) [10] 형태로 배치하여, 청취자가 비교적 적은 횟수의 비교만으로 적합한 HRTF를 선택할 수 있도록 하였다. 이진 트리 생성

시 노드 분기는 두 HRTF에 대해 청취 음원의 방향감 차이를 잘 예측할 수 있는 HRTF 왜곡 척도를 사용하였다.

본 논문의 구성은 다음과 같다. 서론에 이어 2장에서는 트리 분할 기법에 대해서 설명하고, 3장에서는 트리 분할에 사용된 왜곡 척도의 선택 방법을 제안한다. 4장에서는 실험 과정 및 실험 결과를 제시하고 5장에서 결론 및 후속 연구를 제시하며 본 논문을 끝맺는다.

II. 이진 트리 기반의 HRTF 선택 방법

특정 개인에 대한 HRTF를 직접 측정하지 않고 적합한 HRTF를 구하는 방법은 그림 1에 나온 제시된 바와 같이 미리 다양한 사람들로 부터 측정된 HRTF 중 가장 적합한 HRTF를 선택하는 것이다. HRTF는 방향에 의존적인 함수이므로, 원하는 방향마다 적합한 HRTF를 개별적으로 선택한다.

이와 같은 HRTF 선택 방법은 비교적 많은 사람들로 부터 측정된 HRTF 중에서 선택하는 경우 보다 적합한 HRTF를 선택할 수 있다. 그러나 다수개의 HRTF로부터 선택하는 방법은 청취자로 하여금 다수의 청취 비교를 요하며 많은 시간과 노력이 필요하다. 예로서 총 45개의 사람으로부터 얻어진 HRTF를 포함하고 있는 UC-Davis 의 CIPIC HRTF 데이터 베이스의 경우 청취 횟수는 45번이 되며 이는 단일 고도각, 방위각에 대한 청취 비교 횟수로서 모든 방위각과 고도각을 고려하면 매우 많은 청취 비교가 수행되어야 한다.

이진 트리 기반의 HRTF 선택은 그림 2에 나타난 바와 같이 반복적인 2진 탐색에 의해 최종적인 HRTF를 선택하게 된다 [10]. 그림 2에서 좌, 우로 분기 (split) 가 일어나는 각 지점을 노드 (node) 라 하며 최상위 노드를 제외하고 각 노드마다 1개의 HRTF가 할당된다. HRTF의

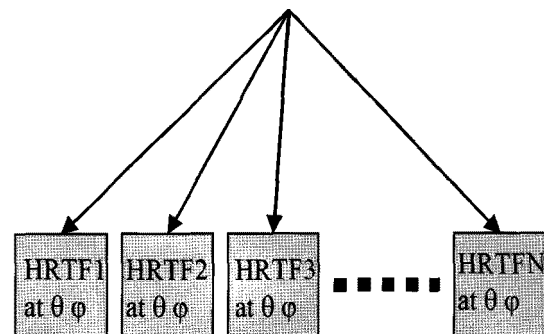


그림 1. 전역 탐색 기반의 HRTF 선택
Fig. 1. Full search HRTF selection.

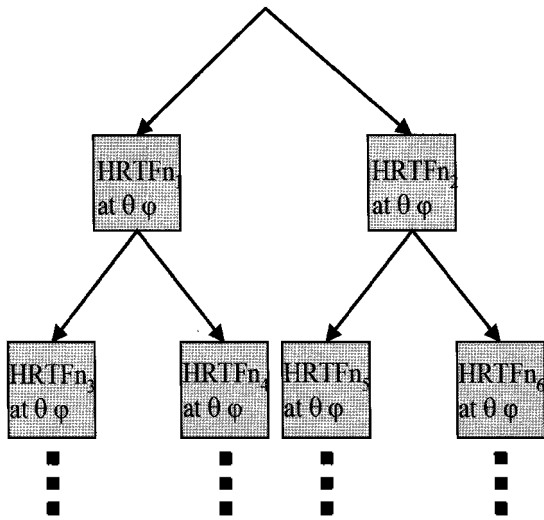


그림 2. 이진 트리 탐색 기반의 HRTF 선택
Fig. 2. HRTF selection based on binary search tree.

선택은 좌, 우 노드에 할당된 HRTF 중 청취 상 지정된 위치에서 발생하는 소리로 좀더 가깝게 인지되는 HRTF를 선택하고 이를 최하위 노드까지 반복적으로 수행함으로써 이루어진다.

이와 같은 이진 트리 기반의 선택 방법은 그림 1의 전역 탐색과 비교하여 적은 탐색 횟수를 요구한다. 예로서 N개의 HRTF 중에 선택하는 경우 전역 탐색은 N번의 청취가 필요한 반면, 이진 탐색 트리에서는 균형 트리 (balanced tree)의 경우 $\log_2 N$ 번 청취가 필요하다.

HRTF 선택을 위한 트리를 생성하는 문제는 주어진 전체 HRTF를 어떻게 반복적으로 분할하고, 각 노드에는 어떠한 HRTF를 할당할 것인가 하는 문제로 요약할 수 있다. 본 논문에서는 상위 노드에서는 청취 상으로 비교적 큰 차이를 갖는 HRTF 간의 선택이 이루어지도록 하고 하위 노드로 갈수록 적은 차이를 갖는 HRTF 간의 선택이 이루어지도록 하였다. 이 방법은 좌, 우로 분기되는 HRTF 간의 차이가 상위 노드에서는 비교적 크고, 하위노드로 갈수록 작아지도록 HRTF를 할당함으로써 구현될 수 있다.

그림 3에 본 논문에서 제안한 노드 분기 방법을 제시하였다. 그림에서 보면 모든 노드는 좌, 우 2개의 하위 노드로 분기되는데 여기서 분기되기 전의 노드를 모 노드 (mother node)로, 분기된 후의 2개 노드를 딸 노드 (daughter node)라 부른다. 제안한 노드 분기 방법은 모 노드에 포함되는 모든 HRTF 들 중에서 청취 상 가장 큰 방향 차이를 나타내는 두 개의 HRTF 조합을 찾는 것이다. 이 두 개의 HRTF를 하위 딸 노드에 대한 대표 HRTF로 간주한다. i -번째 모 노드에 포함되는 HRTF의 집합을 S_i 라 할 때, 좌, 우 딸 노드에 할당되는 HRTF $H_i^{d(i),L}$

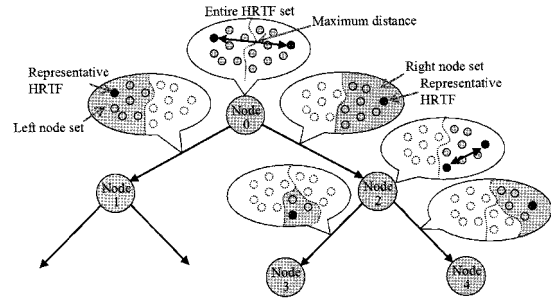


그림 3. 노드 분할을 통한 트리의 생성 과정.
Fig. 3. A procedure of constructing a tree thru node spliting.

과 $H_i^{d(i),R}$ 는 다음과 같이 나타낼 수 있다.

$$H_i^{d(i),L}, H_i^{d(i),R} = \operatorname{argmax}_{H_n, H_m \in S_i} D(H_n, H_m) \quad (1)$$

여기서 $D(H_n, H_m)$ 은 HRTF H_n 과 H_m 간의 청취 상 거리를 나타낸다. 청취 상 거리를 표현하는 왜곡 척도는 다음 장에서 설명될 것이다.

좌, 우 딸 노드에 대한 대표 HRTF가 할당되었으면, 모 노드에 포함된 모든 HRTF를 각 딸 노드에 대한 집합으로 분할 (partition) 한다. 분할은 그림 3에서 제시된 바와 같이 좌, 우 딸 노드에 할당된 HRTF과의 거리를 비교함으로써 이루어지는데, 좌측 딸 노드에 해당하는 HRTF 집합 $S_i^{d(i),L}$ 은 다음과 같이 나타낼 수 있다.

$$S_i^{d(i),L} = \{H_k | H_k \in S_i, D(H_k, H_i^{d(i),L}) < D(H_k, H_i^{d(i),R})\} \quad (2)$$

HRTF 집합 $S_i^{d(i),L}$ 는 다시 모 노드로 간주되어 식 (1)과 (2)에 제시된 방법에 따라 좌, 우 딸 노드로 분할된다. 노드 분할은 노드에 대한 HRTF 집합에 1개의 HRTF만이 포함될 때까지 반복적으로 수행된다. 모든 노드에 대해 1개의 HRTF가 포함되면 최종적인 트리가 생성된 것으로 간주한다. 이와 같은 방법에 의한 트리 생성 방법은 그림 3에 제시된 바와 같이 딸 노드에 포함된 HRTF간의 최대 거리가 모 노드에 포함된 HRTF간의 최대 거리와 항상 작거나 같아지므로 하위 노드로 갈수록 좌, 우 노드의 HRTF 간 거리가 짧아지는 특징을 갖는다. 이는 사용자가 HRTF의 개인화를 위해 다수개의 HRTF 중 자신에게 가장 적합한 HRTF를 선택할 때 초기에는 비교적 큰 차이가 느껴지는 2개의 HRTF 중에 하나의 HRTF를 선택하며, 선택이 진행됨에 따라 유사한 특성을 갖는 HRTF 중에서 사용자에게 더욱 적합한 HRTF를 선택할 수 있음을 의미한다.

III. HRTF 분할을 위한 최적 거리 척도의 선택

앞 장에서 노드의 분기 및 HRTF의 분할을 위해서 두 HRTF간의 거리 척도가 필요함을 알 수 있었다. 두 개의 HRTF에 대한 거리 척도는 주로 HRTF의 근사화와 관련된 연구, 즉 보간(interpolation)과 구획 분할(clustering)과 같은 목적으로 연구되었으며 [11], 본 논문에서와 유사한 목적으로 Shimada 등은 HRTF 개인화를 위한 구획 분할 기법에 루트-캡스트럼을 왜곡 척도로 사용하였다 [12]. 저자 등은 기존 연구에서 HRTF 보간을 위해 수용자 특성 곡선(Receiver Operator Characteristic; ROC) 기반의 왜곡 척도 선택 방법을 제안하였으며, HRIR(Head-Related Impulse Response)를 이용하여 측정되는 왜곡 척도가 비교적 청취상의 차이를 잘 반영함을 알 수 있었다 [13].

본래의 HRTF와 근사화된 HRTF간의 비교가 필요했던 기존의 연구와는 달리, 본 논문에서는 각기 다른 사람으로부터 측정된 HRTF 간의 비교가 이루어지므로 왜곡값은 비교적 크게 나타날 것으로 예상된다. 이에 따라 HRTF를 특징지우는 여러 파라미터들, 즉 크기/위상 스펙트럼, 좌, 우 귀에 도달되는 신호 간의 시간 지연(interaural time delay; ITD), 크기 차(interaural intensity difference; IID), 위상 차(interaural phase delay; IPD)로부터 계산되는 다양한 왜곡 척도를 고려하였다. 본 논문에서는 아래 식으로 주어지는 신호 대 왜곡 비(Signal to Distortion Ratio; SDR)를 두 HRTF 간 거리 척도로 사용하였다.

$$D(X_1, X_2) = \frac{\|X_1 - X_2\|^2}{\|X_1\|^2} \quad (3)$$

여기서 $\|\cdot\|^2$ 는 유클리디언 거리를 나타내며 X_1 과 X_2 는 HRTF H_1 과 H_2 로부터 얻어지는 특징 변수를 나타낸다. 데시벨 단위로 나타낸 왜곡 척도는 다음과 같다.

$$D_{dB}(X_1, X_2) = 10 \log_{10} D(X_1, X_2) \quad (4)$$

본 논문에 사용된 왜곡 측정을 위한 특징 변수로는 HRIR, HRTF의 크기 스펙트럼, HRFT의 펄스스펙트럼 계수, ITD, IID, IPD의 6종류가 사용되었으며 HRIR과 HRTF 크기 스펙트럼에 대해서는 데시벨 단위의 왜곡 척도가, 나머지 특징 변수에 대해서는 선형 왜곡 척도가 사용되었다. 왜곡 척도가 청취상 차이를 얼마나 잘 반영하는가를 나타내기 위해 본 논문에서는 아래와 같은 상호

상관계수(cross correlation coefficient)를 사용하였다.

$$C = \frac{\sum_{m=1}^M \{D(X_{1,m}, X_{2,m}) - \bar{D}\} \{P_D(X_{1,m}, X_{2,m}) - \bar{P}_D\}}{\sqrt{\sum_{m=1}^M \{D(X_{1,m}, X_{2,m}) - \bar{D}\}^2} \sqrt{\sum_{m=1}^M \{P_D(X_{1,m}, X_{2,m}) - \bar{P}_D\}^2}} \quad (5)$$

여기서 $X_{1,m}$ 과 $X_{2,m}$ 은 왜곡 척도의 결정에 사용된 N 개의 HRTF 중 임의로 선택된 두 개의 HRTF 특징 변수를 나타낸다. M 은 전체 조합의 개수로서 ${}_N C_2 = N(N-1)/2$ 가 된다. $P_D(H_1, H_2)$ 는 두 개 HRTF H_1 과 H_2 에 의해 합성된 소리가 청취 상으로 방향감의 차이가 느껴질 확률로서, 실험적으로 구하였다.

본 논문에서는 $P_D(H_1, H_2)$ 를 구하기 위해 정상의 청력을 가진 10명의 피시험자에게 CIPIC HRTF set에 포함된 45 종류의 HRTF 중 임의로 2개를 선택하여 이를 유색 잡음에 통과시킨 신호를 들려주어 인지되는 방향감의 차이를 조사하였다. 방위각과 고도각은 피시험자가 방향의 차이를 비교적 용이하게 구분할 수 있는 고도각=0도, 방위각=0도, +/-40도, +/-80도로서, 총 5개 방향에 대한 HRTF를 사용하였다. 각 피시험자는 5일에 걸쳐 1일 50개의 샘플에 대해 테스트를 시행하였다. 따라서 전체 테스트 샘플은 2500개이다. HRTF 조합 개수는 45개의 HRTF로부터 얻어질 수 있는 모든 조합의 수 보다는 작은 100개를 사용하였으며, 각 조합에 대해 25개의 청취 테스트가 시행되었다. 청취 테스트는 비교적 조용한 환경에서 개방형 헤드폰(Beyerdynamic DT990 PRO, Germany)을 사용하였으며 측정된 헤드폰의 고조파 왜곡은 0.3% (1 kHz)이었다. 청취 시 사용된 헤드폰의 영향을 제거하기 위해 헤드폰의 주파수 응답을 무향실에서 측정하고, 이를 역필터링한 신호를 청취 테스트에 사용하였다. 테스트 실의 크기는 6.7x3.8x2.5 (WDH)로서 평균 잔향시간은 220 msec, 초기 감쇄 시간은 200 msec였다.

실험 결과, 방위각에 따라 약간의 차이가 관찰되었으나 청취상의 차이가 발생할 확률과 가장 높은 상관 계수를 갖는 특징 변수는 IID(상관 계수의 평균 값 0.79)였으며, 특히 0도의 방위각에 대해서 높은 상관 계수(0.83)를 나타내었다. 또한 IID 특징 변수에 대한 상관 계수는 다른 특징 변수에 비해 유의하게 낮은 표준 편차값(0.0344)을 나타내었다. 이러한 실험 결과는 IID로부터 얻어진 왜곡 값이 청취상 방향감의 차이를 비교적 일관되게 잘 예측함을 의미한다. 따라서 본 논문에서는 IID를 이용하여 계산된 왜곡 척도를 거리 측정 함수로 사용하였다.

IV. 실험 및 결과

4.1. 실험 과정

HRTF 선택을 위한 데이터베이스는 UC-Davis에서 제공하는 CIPIC HRTF 데이터베이스 [9]를 사용하였다. 이 데이터베이스에는 총 45명의 사람으로부터 측정된 HRTF를 포함하고 있으며 각 사람에 대한 머리 크기, 귓바퀴의 반경 등 해부학적인 데이터를 함께 포함하고 있다. 각 사람별 모 총 50 종류의 고도각, 총 25 종류의 방위각에서 측정된 HRTF를 포함하고 있으며, 본 논문에서는 수평면에 해당하는 25개의 방위각 (고도각=0도) 중 방위각 0도, +/-10도, +/-20도, +/-30도, +/-40도, +/-55도, +/-80도의 총 13개 방향에 대한 HRTF를 실험 대상으로 삼았다. 제안된 HRTF 선택 방법의 유용성을 평가하기 위해 본 논문에서는 아래의 4가지 방법에 의해 선택된 HRTF를 사용하여 소리를 합성한 후 청취 테스트를 수행하여 방향감에 대한 인지도를 비교 하였다.

- 1) 실제로 측정된 개인 HRTF
- 2) 해부학적인 데이터로부터 선택된 HRTF
- 3) 더미 헤드로부터 측정된 HRTF (MIT KEMAR HRTF) [15]
- 4) 제안된 기법을 통해 선택된 HRTF

1)번에 해당하는 HRTF는 청취 테스트에 참여한 5명의 피시험자로부터 완전 무향 실 (4x3x1.5, WDH)에서 최대 길이 열 (maximum length sequence, MLS) 방법 [14]을 이용하여 측정하였다. 양이 마이크로폰 type 4101 (B&K, Denmark)를 피시험자의 귀에 장착시키고 conditioning amplifier 2693-A (B&K, Denmark)로 증폭한 후 Fire face 400 digital audio interface (RME, Germany)를 이용하여 PC로 data를 전송하였다. 음원 발생을 위한 스피커 시스템으로 8020A (Genelec, Finland) 스피커 시스템이 사용되었으며 음원과 머리 중심부까지의 거리는 약 1.63 m였다. HRTF 측정을 위한 소프트웨어는 Purebits사의 SamplePRO가 사용되었으며 사용된 마이크로폰, 스피커의 특성으로 인한 영향을 감소시키기 위해 취득된 샘플에 대해 마이크로폰, 스피커의 충격파 응답을 역-컨볼루션 하였다.

2)번에서는 CIPIC 데이터베이스에서 제공하는 해부학적 데이터 중에서 헤드/토루소 관련된 데이터 중 측정이 가능했던 12개의 데이터와 귓바퀴와 관련된 데이터 중 7개의 데이터를 사용하여 HRTF를 선택하였다. 해부학적인 데이터의 유클리디언 거리가 가장 짧은 HRTF를 개인

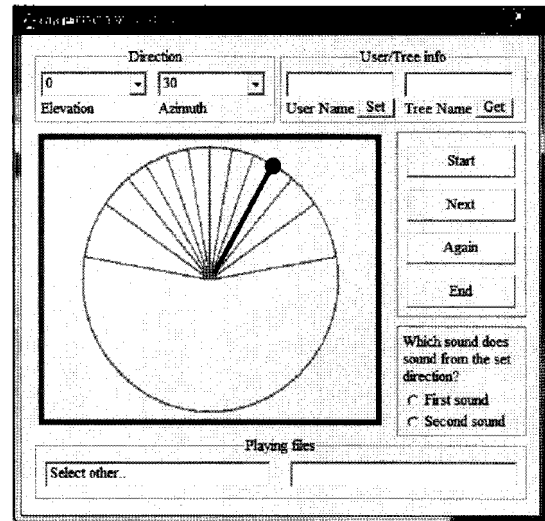


그림 4. HRTF 개인화를 위한 HRTF 선택 프로그램의 실행 화면
Fig. 4. A screen capture describing the HRTF selection program for HRTF customization.

화된 HRTF로 선택하였다. 해부학적 데이터는 디지털 카메라 (EOS 450D, Cannon, F2.8-4 macro lense, Sigma)로 영상을 취득하고 자를 사용하여 측정하였다.

3)번의 HRTF는 사용자와 청취상, 해부학적 구조상 무관한 HRTF를 선택하는 방법이라 할 수 있다. 그림 4는 제안된 방법에 의해 HRTF를 선택하는 프로그램의 실행 화면을 캡처한 것으로, 고도각, 방위각, 개시, 다시 듣기, 평가 버튼 등의 사용자 인터페이스를 포함하고 있으며 현재 청취 중인 소리의 방향은 그림에서 보는 바와 같이 붉은 선과 원으로 표시하였다. 프로그램은 Microsoft Visual C++ 6.0으로 작성하였다. HRTF의 선택에 사용된 음원은 유색 잡음 신호 (-3 dB/oct)가 사용되었으며 청취자는 두 개의 연속된 소리를 듣고 어느 소리가 지정된 방향에 가까운지를 선택하도록 하였다.

청취 테스트에는 헤드폰을 사용하였으며 3장의 최적 거리 함수 선택 시 사용된 실험 조건과 동일한 조건에서 수행되었다.

4.2. 청취 테스트 결과

5명의 피시험자에 대해 13개 방향에 청취 평가가 이루어졌다. 각 피시험자는 4개의 방법으로 선택된 각 HRTF로 합성된 소리를 10번 반복 청취하였다. 따라서 1명의 피시험자가 청취한 전체 소리는 130개이다. 청취 테스트 시의 환경은 3장에서 제시한 HRTF 거리 측정 함수의 선택 시 이용된 환경과 동일하였다. 청취 테스트 시에 사용된 소리는 트리 생성 시 사용된 유색 잡음 신호가 사용되었다 ($f_s = 44.1 \text{ kHz}$).

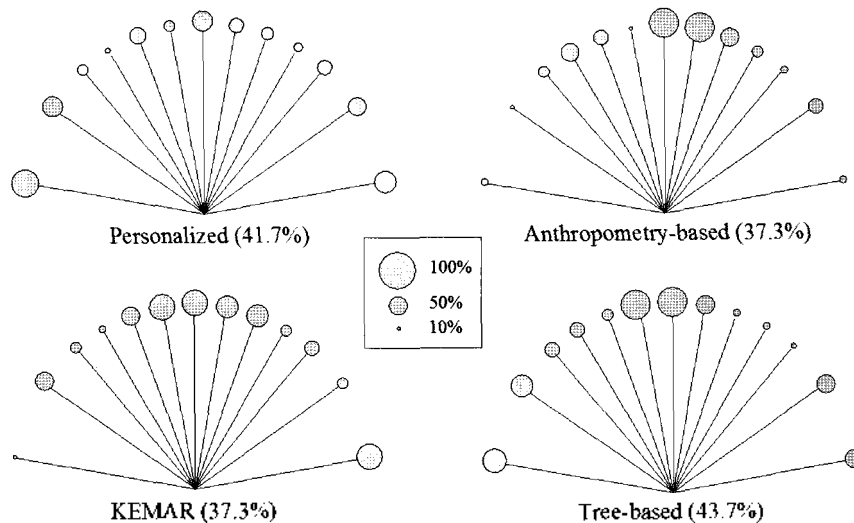


그림 5. 각 HRTF 선택 방법에 대한 방향 별 인식율
 Fig. 5. Correct recognition ratios for each HRTF selection method, for each direction.

표 1. HRTF 개인화 방법 별 실제 방향과 인식 방향과의 평균 절대 오차
 Table 1. Mean absolute error between the real and the recognized directions, for each HRTF personalization method.

HRTF 개인화 방법	평균 절대 오차 (도)
개인별 측정 HRTF 사용	9.31
해부학적 데이터를 이용하여 HRTF 선택	9.87
KEMAR HRTF 이용	11.27
이진 탐색 트리 기반 HRTF 선택	9.89

그림 5에서 4가지 HRTF 개인화 방법에 대한 각 방향 별 인식율을 그림으로 나타내었다. 그림에서 큰 원은 높은 인식율을, 작은 원은 낮은 인식율을 나타내며 괄호안 수치는 전체 인식율을 나타낸다. 이진 탐색 트리를 사용하는 제안된 방법이 가장 높은 전체 인식율을 나타내었으며 특히 좌, 우 양 끝 지점 (방위각 ± 80 도)과 정면 (방위각 0 도)에서 비교적 일관되게 높은 인식율을 나타내었다. KEMAR HRTF와 해부학적인 데이터로 선택된 HRTF는 다른 두 방법에 비해 낮은 인식율을 나타내었으며 특히 좌, 우 끝 지점에서 매우 낮은 인식율을 나타내었다. 개인별로 측정된 HRTF를 사용하는 경우, 가장 높은 인식율을 나타낼 것으로 예상되었으나 제안 기법에 비해 낮은 인식율을 나타내었으며 정면에서의 인식율은 다른 3가지 방법에 비해 다소 낮게 나타났다. 제안된 기법은 20 도 ~ 40 도 영역에서 다소 낮은 인식율을 나타낸 반면 정면과 인접한 방향 (방위각 -10 도 ~ 10 도)에서는 높은 인식율을 나타내었다.

본 논문의 실험 결과가 기존의 HRTF 개인화 기법에서 보고한 인식율과 비교하여 다소 낮게 나온 것은 청취 테스트 시 음원의 방향이 기존 연구에 비해 조밀하게 배열되어 있어 방향의 미세한 차이를 청취자가 구분하기 힘든

것에 원인이 있는 듯하다. 실제로 그림 5의 결과를 살펴 보면 좌, 우 양 끝 및 정면에서는 대체적으로 높은 인식율을 보이는 반면 정면과 측면의 경계에 해당하는 방향에서는 낮은 인식율을 보이고 있다. 이는 개인별로 측정된 HRTF를 사용하는 경우에도 낮은 인식율을 나타낸 이유 중의 하나로서, 음원의 방향을 좀더 넓게 분포 시키는 경우 (예: 0 도, ± 20 도, ± 40 도, ± 80 도) 모든 방법에 있어서 $60 \sim 70\%$ 의 높은 인식율을 나타내었으며 개인화된 HRTF를 사용하는 경우 가장 높은 인식율 (79%)을 나타내었다.

전체 방향에 대한 평균 인식율은 방향이 단순히 맞고 틀림만을 반영한 척도이다. 그러나 오류 중에서도 10 도로 인식되어야 할 소리가 20 도로 인식되는 것과 50 도로 인식되는 것은 서로 다르게 간주되어야 한다. 따라서 본 논문에서는 실제 소리의 방향과 인지된 방향간의 평균 절대 오차 (Mean Absolute Error: MAE)를 구하여 각 방법 간에 비교하였다. 표 1은 이에 대한 결과로서, 개인별 측정된 HRTF가 가장 낮은 값을 갖는 것으로 나타났다. 이는 절대 오차 관점에서는 역시 개인별로 측정된 HRTF가 가장 적합함을 의미한다. KEMAR HRTF를 사용하는 경우에는 가장 큰 값을 나타내었으며, 이는 가장 낮은 인

식율을 나타내었던 이진 결과와 일치하는 것이라 볼 수 있다. 제안 기법과 해부학적 데이터를 이용해 선택하는 방법은 거의 비슷한 평균 절대 오차를 나타내었다.

전체적으로 이진 탐색 트리를 이용하여 적합한 HRTF를 선택하는 방법은 비교적 짧은 시간 (약 13개 방향에 대해 총 5분 정도)에 HRTF 선택을 완료 할 수 있으며 개인별로 직접 측정된 HRTF를 사용하는 경우와 거의 근접한 성능을 나타내었다.

IV. 결론

본 논문에서는 HRTF의 개인화 방법으로 여러 명의 사람으로부터 측정된 HRTF 중에서 선택하는 방법을 제안하였다. 제안된 방법은 HRTF를 통해 합성된 소리를 직접 청취하고 지정된 방향에 가장 가깝게 들리는 소리를 선택하도록 하였다. 이 때 HRTF의 개수가 많으면 청취자가 선택하는 시간이 오래 소요되므로 이진 탐색 트리를 이용하여 빠른 시간에 적합한 HRTF를 선택할 수 있도록 하였다.

이진 탐색 트리의 구성에 핵심적인 역할을 하는 노드 분기 방법으로 본 논문에서는 최대 거리를 갖는 두 개의 HRTF를 좌, 우 말 노드에 할당하고 최소 오차법을 적용하여 모 노드에 포함된 HRTF를 좌, 우 말 노드로 분할하였다. 이와 같은 방법은 탐색 초기에는 청취 상 차이가 크게 구분되는 HRTF를 선택하고 탐색이 진행 되어 감에 따라 좀더 세밀하게 HRTF를 선택할 수 있게된다.

노드 분기를 위한 HRTF의 거리 측정 방법으로, 본 논문에서는 주관적인 청취 차이를 잘 예측할 수 있는 HRTF 파라미터로 왜곡 척도를 구성하였다. 실험 결과, 상호 상관 계수가 가장 높게 나타나는 IID 왜곡 척도를 거리 측정 함수로 사용하였다.

해부학적 데이터를 사용하여 개인화된 HRTF를 선택하는 기존의 방법, 개인별로 측정된 HRTF를 사용하는 방법, 그리고 사용자와는 무관한 HRTF를 선택하는 3가지 방법과의 비교에서 제안된 기법은 방향 인식율에서 가장 우수한 성능을 나타내었으며, 평균 절대 오차 면에서도 개인별로 측정된 HRTF를 사용하는 경우와 근접한 성능을 나타내었다.

감사의 글

본 연구 과정에 포함된 청취 테스트에 참여한 건국대학

교 사범대학 음악교육과 학생 및 강사 선생님들에게 감사의 마음을 전해드립니다.

참고 문헌

1. J. Blauert, *Spatial Hearing: The Psychophysics of Human Sound Localization*, MIT Press, Cambridge, MA, 1983.
2. D. R. Begault, *3D sound for virtual reality and multimedia*, Academic Press, Cambridge, MA, ch. 4, pp. 132-136, 1994.
3. J. C. Middlebrooks, "Individual differences in external-ear transfer functions reduced by scaling in frequency," *J. Acoust. Soc. Am.*, vol. 106, no. 3, pp. 1480-1492, 1999.
4. S.-L. Lee, L.-H. Kim and S.-M. Sung, "Head related transfer function refinement using directional weighting function," *AES 115th Convention*, paper no. 5918, Oct, 2003.
5. M. Zhang, K.-C. Tan and M. H. ER, "Three-dimensional sound synthesis based on head-related transfer functions," *J. Audio Eng. Soc.*, vol. 46, no. 10, pp. 836-844, 1998.
6. H. Hu, L. Chen and Z. Wu, "The estimation of personalized HRTFs in individual VAS," in *Proc. IEEE 4th International Conference on Natural Communication*, pp. 203-207, Oct, 2008.
7. D. N. Zolkın, R. Duraiswami and L. S. Davis, "Rendering localized spatial audio in a virtual auditory space," *IEEE Trans. Multimedia*, vol. 6, no. 4, pp. 553-564, 2004.
8. S. Hwang and Y. Park, "HRIR customization in the median plane via principal components analysis," *AES 3rd Convention*, paper no. 9, June, 2007.
9. *CIPIC HRTF Database Files, Release 1.0*, <http://interlace.cipic.ucdavis.edu>, 2001.
10. Brieman, Friedman, Olsen and Stone, *Classification and Regression Trees*, Wadsworth, Belmont CA, 1984.
11. R. Nicol, V. Lemaire, A. Bondu and S. Busson, "Looking for a relevant similarity criterion for HRTF clustering: a comparative study," *AES 120th Convention*, paper no. 6653, May, 2006.
12. S. Shimada, N. Hayashi and S. Hayashi, "A clustering method for sound localization transfer functions," *J. Audio Eng. Soc.*, Vol. 42, no. 7/8, pp. 577-584, 1994.
13. 이기승, 이석필 "머리전달 함수의 보간에 적합한 왜곡 척도," *한국음향학회지*, 28권, 2호, 85-95쪽, 2009.
14. D. D. Rife and J. Vanderkooy, "Transfer-Function Measurements using Maximum-Length Sequences," *J. Audio Eng. Soc.*, vol. 37, no. 6, pp. 419-444, 1989.
15. *KEMAR HRTF data*, <http://sound.media.mit.edu/pub/Data/KEMAR>, 1994

저자 약력

•이 기 승 (Ki-Seung Lee)
한국음향학회지 제28권 제2호 참조

•이 석 필 (Seok-Pil Lee)
한국음향학회지 제28권 제2호 참조