

# 원형 머리 모델을 이용한 머리 전달 함수의 보간

## HRTF Interpolation Using a Spherical Head Model

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

\*건국대학교 정보통신대학 전자공학부, \*\*전자부품연구원 방송통신융합 연구센터  
(접수일자: 2008년 8월 7일; 수정일자: 2008년 9월 25일; 채택일자: 2008년 10월 3일)

본 논문에서는 머리 전달 함수에 대한 새로운 보간 기법을 제안하였다. 제안된 기법은 각 방위각에 대한 머리 전달 함수의 충격파 응답이 인접 방위각에 대한 시간 지연된 충격파 응답의 선형 보간으로 주어진다고 가정하였다. 각 방위각에 대한 충격파 응답의 시간 지연은 방위각, 머리의 물리적 형태, 음원과 머리의 거리 정보를 이용하여 추정될 수 있는 귀와 음원간의 전파시간과 최소 지층 오차를 갖도록 하는 교정값의 합으로 주어진다. 또한 제안된 모델에서는 보간 시 방위각의 간격을 고정 간격이 아닌 가변 간격으로 하였으며 본래의 충격파 응답과 보간된 충격파 응답이 본래의 충격파 응답과 비교하여 청취 상으로 큰 차이가 느껴지지 않고, 보간에 필요한 충격파 응답의 개수가 최소화되는 조건을 만족하도록 결정하였다. 제안된 보간 모델의 유용성을 검증하기 위하여 네이 헤드 및 3명의 사람으로부터 측정된 머리 전달 함수에 대해 제안된 보간 모델을 적용하였다. 머리 전달 함수는 0도의 고도각을 갖는 수평면을 5도 간격의 방위각으로 분할한 총 72개가 사용되었으며, 실험 결과 전체 머리 전달 함수 중 단지 30~40% 만을 사용하고 나머지는 보간에 의해 얻어진 머리 전달 함수를 사용하더라도 청취상의 음원의 위치가 변동되지 않음을 알 수 있었다.

**핵심용어:** 머리전달함수, 보간, 원형머리모델

**투고분야:** 전가음향 분야 (3, 1)

In this paper, a new interpolation model for the head related transfer function (HRTF) was proposed. In the method herein, we assume that the impulse response of the HRTF for each azimuth angle is given by linear interpolation of the time-delayed neighboring impulse responses of HRTFs. The time delay of the HRTF for each azimuth angle is given by sum of the sound wave propagation time from the ears to the sound source, which can be estimated by using azimuth angle, the physical shape of the underlying head and the distance between the head and sound source, and the refinement time yielding the minimum mean square error. Moreover, in the proposed model, the interpolation intervals were not fixed but varied, which were determined by minimizing the total number of HRTFs while the synthesized signals have no perceptual difference from the original signals in terms of sound location. To validate the usefulness of the proposed interpolation model, the proposed model was applied to the several HRTFs that were obtained from one dummy-head and three human heads. We used the HRTFs that have 5 degree azimuth angle resolution at 0 degree elevation (horizontal plane). The experimental results showed that using only 30~50% of the original HRTFs were sufficient for producing the signals that have no audible differences from the original ones in terms of sound location.

**Keywords:** Head related transfer function, Interpolation, Spherical head model

**ASK subject classification:** Electro-Acoustics (3, 1)

### I. 서론

머리전달함수 (Head Related Transfer Function, HRTF)는 자유 공간 상에서 음원 (sound source)과 귀 간의 경로

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)

를 모델링하는 함수로서, 음상의 공간화 (sound localization)과 관련된 중요한 음향 파라미터이다 [1-5]. HRTF는 음원과 귀 사이의 공간 경보에 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. HRTF는 헤드폰과 같은 양어 재생 환경에서

음원의 방향성을 포함하는 공간감을 표현하는데 널리 사용되고 있다 [6].

HRTF는 더미헤드 또는 청취자에게 양이 마이크로폰을 착용시켜 직접 측정하는데, 보통의 경우, 시간 영역의 유한길이디지털 필터 (Finite Impulse Response Digital Filter, FIR filter) 형태로 표현하며 따라서 HRTF는 시간 영역의 충격파 응답 (Head-Related Impulse Response: HRIR)으로 나타낸다. 특정 방위각과 고도각에 대한 1개의 HRTF는 샘플링 주파수가 44.1 kHz인 경우 보통 512개의 샘플로 표현된다 [2]. 현재 반도체 기술의 발전에 따라 대용량 메모리의 사용이 일반화되었지만, 가능한 모든 고도각과 방위각에 대한 음상을 위치시키기 위해서는 매우 많은 HRTF 데이터를 저장해야 한다. 예로서 HRTF를 5도 단위의 방위각과 고도각으로 나타내는 경우 약 5 Mbytes의 메모리를 필요로 한다. 이와 같은 메모리 양은 휴대폰이나 PDA와 같은 휴대용 장치에 있어서 적지 않은 부담으로 작용할 수 있다.

이러한 문제에 따라 HRTF를 보다 적은 정보량으로 표현하는 방법들이 연구되고 있다. 이러한 방법들은 크게 주어진 HRTF를 정보 감축 기법을 통해 압축을 하는 방법 [2, 4, 6, 9, 17]과 보간 및 구획 분할 기법을 통해 몇 개의 대표 HRTF로 모든 방향의 HRTF를 나타낸 방법 [5, 18, 19]으로 나눌 수 있다. 첫 번째 방법의 예로서, Kistler 등은 HRIR에 대해 주요성분분석 (principle component analysis, PCA)을 통해 에너지가 크게 분포하는 몇 개의 기저 벡터 (basis vector)로 HRTF를 나타내었다 [2]. 이와 유사한 국내 연구로서 황 등은 주요성분분석에 바탕을 둔 보간 기법을 비교 연구하였다 [17]. Kistler와 Kulkarni의 연구에서는 공간감 형성에 중요한 역할을 담당하는 HRTF의 위상 정보를 최소위상시스템 (minimum phase system)과 양이시간차 (Interaural Time Difference, ITD)를 이용하여 나타내었다 [1, 4]. 공간 음상의 구현을 위한 연구의 일부로 Kim 등은 HRTF의 위상을 인접한 방위각의 HRTF의 위상값과 두 방위각간의 시간 지연만으로 나타내었다 [9]. 보간에 의한 압축 기법으로, Minnaar 등의 연구에서는 2도 단위의 방위각으로 측정된 HRTF에 대해 주요 성분 에 대한 보간 기법을 제안하였으며, 8도의 간격이상으로 보간을 수행하는 경우 청취 상으로 유의한 차이가 난다고 보고하였다 [5].

이들 연구는 대부분 HRTF의 근사화 오차를 최소화하는 관점에서 보간을 수행하였으며, HRTF의 표현에 필요한 전체 비트수, 즉 정보량 관점에서의 논의는 다소 부족한 편이다. 본 논문에서는 HRTF의 보간 오차와 함께 필

요한 전체 비트수를 함께 고려한 새로운 개념의 HRTF 보간 기법을 제안하였다. 제안된 기법은 각도에 따른 HRTF의 변화율이 가변적일 것이라는 가정에 따라 보간 간격을 가변시키는 불균일 샘플링 (non-uniform sampling) 기법을 사용하였다. 보간 간격의 결정은 최대 보간 오차가 정해진 문턱치를 초과하지 않으면서 샘플된 HRTF의 개수가 최소화되는 관점에서 이루어지며, 이는 동적 프로그래밍 (dynamic programming) 기법을 통해 구현하였다. 최대 보간 오차의 문턱치는 보간된 HRTF가 원래의 HRTF와 동일한 방향감을 갖도록 설정하였다. 또한 본 논문에서는 HRIR의 새로운 보간 모델로서, 각 HRIR을 방위각에 따라 각기 다른 시간 지연을 갖는 충격파 응답으로 모델링하고, 최소 자승 오차 관점에서 보간된 HRIR을 추정하는 기법을 제안하였다. 각 방위각에 따른 시간 지연은 Woodworth와 Schlosberg가 제안한 원형 머리 모델 (spherical head model) [11]을 적용하여, 양쪽 귀에 대한 음파 도달 시간으로 나타내었다. 1개의 모형 머리 (dummy head)와 3인의 실제 머리에서 측정된 HRTF에 대해 제안된 모델을 적용하여 최대 허용 보간 오차에 따른 정보량을 살펴보고, 샘플된 방위각을 각 각의 HRTF에 대해 살펴보았다.

## II. HRIR의 보간

방위각  $\theta_k$ 에 해당하는 오른쪽 귀 또는 왼쪽 귀에 해당하는 HRIR을  $h_{\theta_k}(t)$ 이라 하였을 때, 이를 시간 영역의 선형 보간을 이용하여 나타내면 다음과 같다.

$$\tilde{h}_{\theta_k}(t) = \alpha h_{\theta_1}(t) + (1 - \alpha) h_{\theta_2}(t) \quad (1)$$

여기서  $\theta_1 < \theta_k \leq \theta_2$ , 즉 방위각  $\theta_k$ 에 해당하는 HRIR은 인접된 방위각  $\theta_1$ 과  $\theta_2$ 의 보간에 의해 이루어짐을 알 수 있다.  $\alpha$ 는 보간 상수로서 다음과 같이 나타낼 수 있다.

$$\alpha = \frac{\theta_2 - \theta_k}{\theta_2 - \theta_1} \quad (2)$$

즉, 방위각간의 상대적인 차이의 비율로 주어짐을 알 수 있다. 식 (1)과 같은 단순화된 보간 방법은 매우 큰 보간 오차를 나타내는데, 그 이유는 각 방위각에 대한 HRIR이 서로 개시 시간 (onset time)을 갖기 때문이다



여기서  $h_{\theta_k}(n)$ 은 이산 HRIR (discrete HRIR)의  $n$ -번째 샘플을 나타내며  $W_3$ 는 교정 구간을 나타낸다.  $W_3$ 를 큰 범위로 설정하면 모델 오차가 크게 발생하는 경우에도 교정이 가능하나 최소 자승 오차 합의 반복적인 계산량이 증가되며, 작은 범위가면 계산량은 감소하나 교정할 수 있는 오차의 범위가 감소한다는 단점이 있다. 본 논문에서는 이와, 같은 계산량과 교정의 정밀도를 고려하여 교정 구간을 1-4, 31 사이로 설정하였다. 최종적인 보간된 HRIR은 다음과 같이 나타낼 수 있다.

$$\begin{aligned} \hat{h}_{\theta_k}^*(n) &= \hat{h}_{\theta_k}(n, \Delta_{1,k}^*, \Delta_{2,1}^*) = \alpha h_{\theta_k}(t - \Delta_{1,k}^*) \\ &+ (1 - \alpha) h_{\theta_k}(t - \Delta_{1,k}^* - \Delta_{2,1}^*) \end{aligned} \quad (9)$$

### III. 불균일 간격의 HRTF 보간

본 논문에서 제안된 방법에서는 모든 방위각과 고도각에 대한 HRTF를 저장하지 않고, 일부 방위각과 고도각에 대한 HRTF만을 저장하고, 저장되지 않은 고도각과 방위각에 대한 HRTF는 앞 절에서 제시한 보간식을 이용하여 HRTF를 생성하게 된다. 이 때 보간에 필요한 정보량은 식 (9)의  $\Delta_{1,k}^*$ 와  $\Delta_{2,1}^*$ 로서, 본 논문에서와 같은 교정 구간을 사용한 경우 채널당  $3 \times 2 = 6$  비트가 필요하다. 이는 본래 HRTF가 갖는 정보량인 8192 비트 (512샘플을 샘플당 16비트로 표현하는 경우)에 비해 매우 적은 양임을 알 수 있다. 따라서 정보 감축이 많이 이루어지기 위해서는 되도록 많은 HRTF가 보간에 의해 생성되어야 한다.

보간의 가장 간단한 방법은 보간 간격을 등 간격으로 설정하는 것이다 [5]. 예로서 2:1 보간의 경우, 모든 홀수 번째의 방위각에 해당하는 HRTF를 짝수번째 방위각에 해당하는 HRTF로부터 보간에 의해 얻는다. 그림 2의 왼쪽은 등 간격 보간의 예를 나타낸 것으로, 그림 2에서 회색 점은 검은 점에 해당하는 데이터로 보간이 이루어지는 지점을 나타낸다. 이와 같은 등 간격 보간은 보간이 규칙적인 간격으로 이루어지므로 보간 비율만 알고 있으면 보간 위치에 대한 추가 정보가 필요 없다는 장점이 있다. HRTF의 경우 방위각의 범위에 따라 그 변화 정도가 완만하게 또는 빠르게 나타날 수 있으며, 이에 따라 완만하게 변화하는 영역에서는 보간 간격을 넓혀 정보량을 줄일 수 있으며 급격하게 변화하는 영역에서는 보간 간격을 좁혀 빠르게 변화하는 HRTF의 특성을 적은 오차로 표현할 필요성이 있다. 이러한 특성을 반영한 보간 방법으로 그림 2의 오른쪽에 불균일 간격의 보간 방법이 제시되어 있다. 그림 2에서 보면 보간이 이루어지는 위치와 및 HRTF를 샘플하는 위치가 불규칙하게 분포함을 알 수 있다.

본 논문에서는 HRTF의 정보 감축 기법으로 그림 2의 오른쪽에 제시한 불균일 보간 방법을 적용하였다. 보간에 필요한 HRTF의 방위각 샘플링은 필요한 정보량 (비트 수)을 최소화 하면서 보간 오차가 최소화 되도록 하였다. 그러나 이 두 개의 값이 동시에 최소화될 수 없으므로 아래와 같은 최대값-최소화 기준 (min-max criterion) [13]을 적용하였다.

$$\text{Minimize } R(H), \text{ subject to } D_{\max}(H) \leq D_{\text{thres}} \quad (10)$$

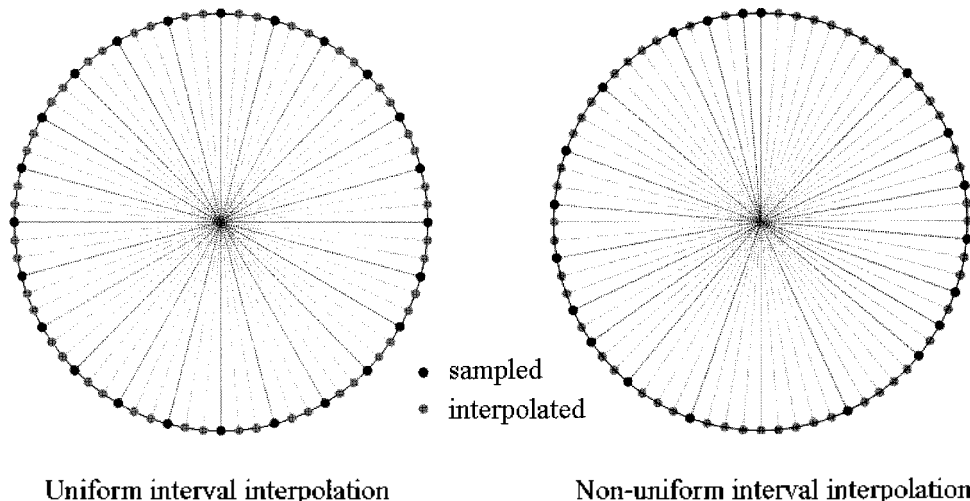


그림 2. 등 간격 보간 (왼쪽)과 불균일 간격 보간 (오른쪽)  
 Fig. 2. Uniform interval interpolation (left), non-uniform interval interpolation (right).

여기서  $H = \{h_{\theta_1}, \dots, h_{\theta_m}\}$ 는 샘플된 HRIR의 집합으로서  $R(H)$ 는 보간에 필요한 전체 비트수를 나타내며  $D_{\max}(H)$ 는 집합  $H$ 에 포함된 HRIR을 이용하여 보간을 수행하였을 때, 최대 보간 오차를 나타내며,  $D_{\max}$ 는 보간 오차의 최대 허용치를 나타낸다.  $R(H)$ 은 집합  $H$ 에 포함된 모든 HRIR을 표현하는데 필요한 비트수와 보간에 필요한 비트수(채널당 6비트)의 합으로 주어지며 HRIR의 표현에 필요한 비트수(8192)가 보간에 필요한 비트수 보다 월등히 크기 때문에  $R(H)$ 를 최소화하는 것은 샘플된 HRIR의 개수를 최소화시키는 것으로 볼 수 있다.

식 (10)의 조건을 만족하는 HRIR의 집합  $H^* = \{h_{\theta_1}, \dots, h_{\theta_m}\}$ 을 구하기 위해 본 논문에서는 그림 3과 같이 모든 방위각에 대한 HRIR을 방위각의 순서대로 1차원으로 나열한 directed graph 형태로 표현하였다. 그림 3에서 그래프의 각 정점은 각 방위각에 대한 HRIR을 나타내며 정점과 정점을 잇는 edge는 정점에 해당하는 HRIR들을 이용하여 보간된 HRIR을 나타낸다. 각 edge는 아래의 가중치  $\omega(h_{\theta_{m-1}}, h_{\theta_m})$ 를 갖는다.

$$\omega(h_{\theta_{m-1}}, h_{\theta_m}) = \begin{cases} \infty & , \text{if } d_{\max}(h_{\theta_{m-1}}, h_{\theta_m}) \geq D_{thr} \\ r(h_{\theta_{m-1}}, h_{\theta_m}) & , \text{otherwise} \end{cases} \quad (11)$$

여기서  $d_{\max}(h_{\theta_{m-1}}, h_{\theta_m})$ 는 방위각의 구간  $[\theta_{m-1}, \theta_m]$  내에 포함된 HRIR을  $h_{\theta_{m-1}}$ 과  $h_{\theta_m}$ 를 이용하여 보간 하였을 때, 최대 보간 오차를 나타내며  $r(h_{\theta_{m-1}}, h_{\theta_m})$ 는  $h_{\theta_{m-1}}$ 과  $h_{\theta_m}$ 의 표현 및 보간에 필요한 전체 비트수를 나타낸다.

Directed graph 상에 존재하는 모든 HRIR의 조합 중에서 조건 (10)을 만족하는 최적의 집합  $H^*$ 을 찾는 것은 보간 오차의 최대값이 문턱치보다 낮으면서, 최소의 정점으로 구성된 경로 (path)를 찾는 문제로 해석할 수 있다. 이와 같은 문제의 해결을 위해 동적 프로그래밍 (dynamic programming) 기법이 적용되었다 [14]. 이 방법은 먼저 각 directed graph 상의 각 정점에 대해 지역 최적 경로 (local optimum path)를 구하고, 역전파 (back-tracking)에 의해 전체적인 경로를 구하는 것이다. 최적 경로를 찾는 전체 과정은 다음과 같다.

$$\begin{aligned} w(m) &= \underset{1 \leq k \leq m}{\operatorname{argmin}} \{ \beta_{k,m} [R(h_{\theta_k}) + r(h_{\theta_k}, h_{\theta_m})] \} \\ R(h_{\theta_m}) &= R(h_{\theta_{m-1}}) + r(h_{\theta_{m-1}}, h_{\theta_m}) \\ D_{\max}(h_{\theta_m}) &= \max \{ D_{\max}(h_{\theta_{m-1}}), d_{\max}(h_{\theta_{m-1}}, h_{\theta_m}) \} \end{aligned} \quad (12)$$

여기서 방위각의 해상도가 5도 인 경우  $1 < m < 72$ 이다.  $R(h_{\theta_k})$ 과  $D_{\max}(h_{\theta_k})$ 는 각각 방위각  $\theta_k$ 에 해당하는 정점까지의 누적 비트수 및 최대 보간 오차를 나타낸다.  $\beta_{k,m}$ 은 보간 오차의 최대치가 문턱치를 초과하는 경우를 판별하기 위한 변수로, 아래와 같이 주어진다.

$$\beta_{k,m} = \begin{cases} \infty & , \text{if } \max \{ D_{\max}(h_{\theta_k}), d_{\max}(h_{\theta_k}, h_{\theta_m}) \} > D_{thr} \\ 1 & , \text{otherwise} \end{cases} \quad (13)$$

$w(m)$ 은 back-tracking 포인터로 방위각  $\theta_m$ 에서 (10)의 조건을 만족하는 경로의 시작 방위각을 나타낸다. 따라서 역순으로 배열된 최적의 방위각은 아래와 같다.

$$h_{\theta_1}, h_{\theta_{w(m)}}, h_{\theta_{w(m)-1}}, \dots$$

그림 3의 예제에서 각 정점에 연결된 경로 중 붉은 선으로 표시된 경로가 지역 최적 경로라면, 최적의 HRIR 집합은  $H^* = \{h_{\theta_1}, h_{\theta_2}, h_{\theta_3}, h_{\theta_m}\}$ 임을 알 수 있다.

보간된 HRIR과 실제 HRIR 간의 보간 오차는 청취 상으로 위치 감 차이를 잘 반영하는 변수를 사용해야 한다. 이러한 조건을 만족하는 변수로 Nicol 등 [8]에 의해 연구된 HRIR 오차 변수와 위상 왜곡 [12] 등이 고려될 수 있는데, 본 논문에서는 실험적으로 청취 상의 오차와 비교적 높은 상관관계를 나타내었던 정규화 된 시간 영역의 로그 오차 함수를 사용하였다. 보간된 HRIR을  $\hat{h}(n)$ 이라 하고 실제 HRIR을  $h(n)$ 이라 했을 때, 정규화 보간 오차는 다음과 같다.

$$D(\hat{h}, h) = 10 \times \log_{10} \left( \frac{\sum_{n=1}^N \{h(n) - \hat{h}(n)\}^2}{\sum_{n=1}^N h^2(n)} \right) \quad (14)$$

### IV. 실험 및 결과

제안된 보간 기법을 실제 HRIR에 적용하여 보간 오차의 문턱치에 따른 필요 정보량, 보간 HRIR의 사용에 따른 공간감의 저하 영향 등을 분석하였다. 실험에는 총 4 종류의 HRIR가 사용되었는데, MIT의 KEMAR [16], 그리고 3명의 사람으로부터 실제 HRIR를 측정하였다 (NJ1, NJ2, NJ3). HRIR의 측정에는 가로, 세로, 높이 각각 3x4x1.8 m 인 완전 무향실에서 type 4101 양이머이크로폰 (B&K,

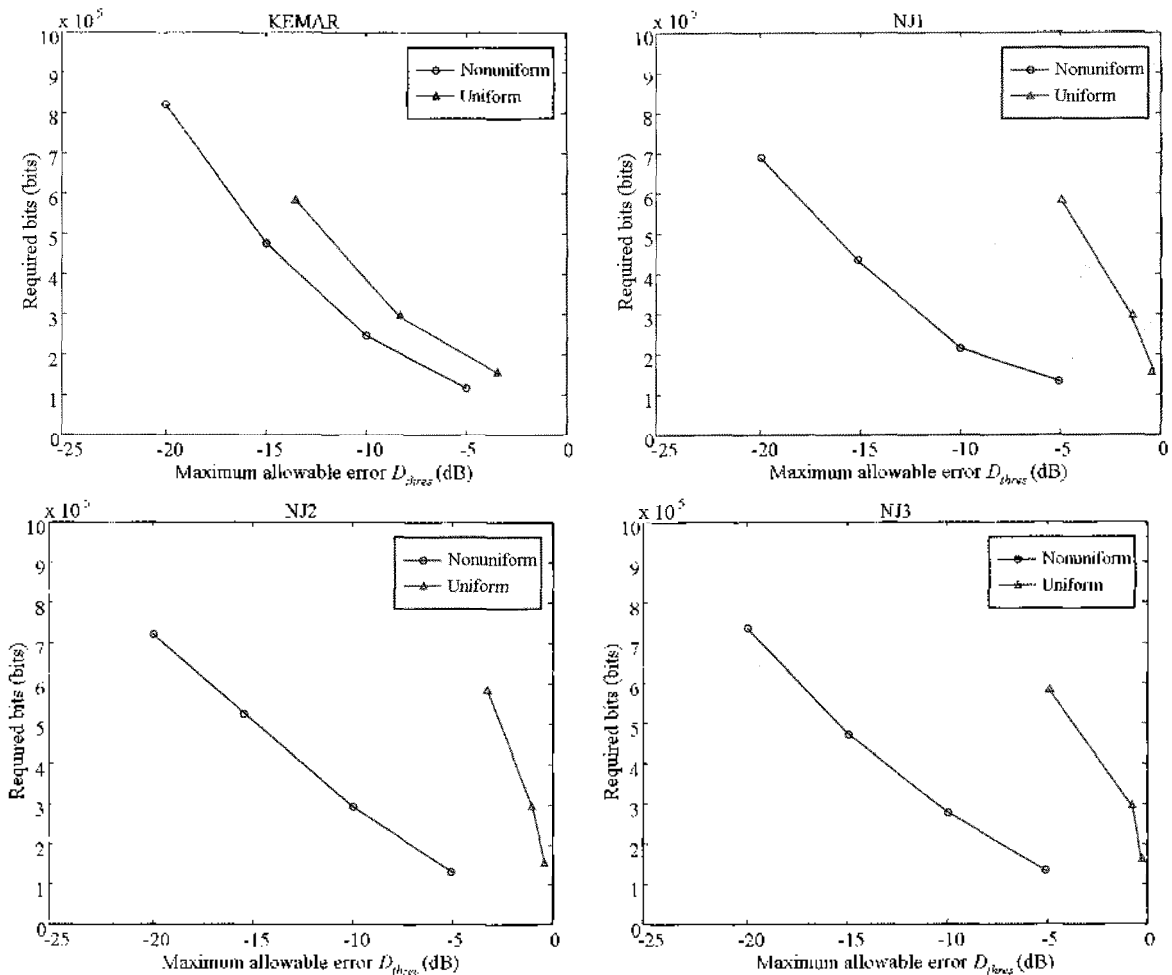


그림 4. 4개 HRTF 데이터에 대한 최대허용오차 대 정보량 (비트수) 특성 곡선  
 Fig. 4. Maximim allowable error-bit rate characteristic curves for 4 HRTFs.

Demark)과 8020A 모니터 스피커 (Genelec, Norway)를 사용하였다. A/D D/A 변환에는 Fireface 400 디지털 오디오 인터페이스 (RME, Germany)를 이용하였다. HRTF의 충격파 응답의 측정에는 최대 길이 열 (Maximum Length Sequence, MLS)기법 [15] 이 사용되었으며 열의 길이는 4096샘플로 설정하여, 이 중 신호를 포함하는 영역 512

샘플만을 충격파 응답으로 사용하였다. 샘플링 주파수는 KEMAR와 동일하게 44.1 KHz 로 설정하였다. HRTF의 측정은 0도의 고도각에 대해서만 이루어졌으며, KEMAR의 경우에도 0도의 고도각에 대한 데이터만을 사용하였다. 방위각은 5도 간격으로 측정되었으며, 따라서  $360/5 = 72$ 개의 HRTF 데이터가 사용되었다.

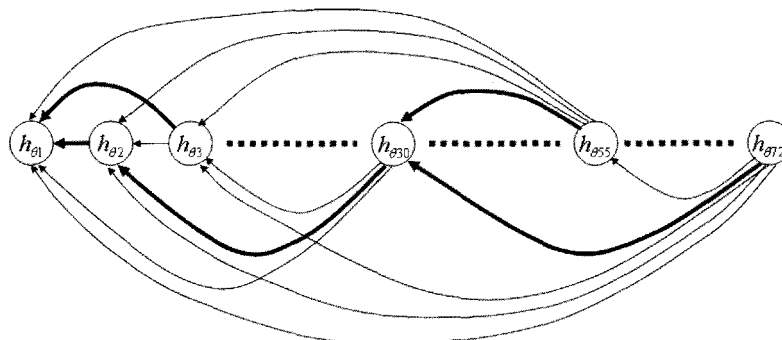


그림 3. HRIR 의 열을 다이렉티드 그래프로 표현한 예  
 Fig. 3. An example of the HRIR sequence using a directed graph.

#### 4.1. 최대 허용 오차에 따른 HRTF의 정보량

4개의 HRTF 데이터 (KEMAR, NJ1, NJ2, NJ3)에 대해 최대 허용 오차  $D_{thres}$ 를  $-20, -15, -10, -5$  (dB)로 변경시키면서 제안된 보간 기법을 적용하였을 때 HRTF의 표현에 필요한 정보량 (비트수)을 살펴보았다. 그림 4는 불균일 보간 시 최대 허용 오차에 따른 HRTF의 비트수와 균일 보간 [5] 시의 최대 오차 대 비트 수를 나타낸 것이다.

그림 4에서 보면 허용 오차가 큰 수록 작은 정보량을 나타내는 특성이 4개의 HRTF에 공통적으로 나타나고 있다. KEMAR 데이터의 경우 다른 데이터에 비해 약간의 높은 정보량을 보이고 있다. 일반적인 비트율-왜곡의 특성은 비선형인 log 함수의 특성을 따르는데, 실험 결과로 나타난 최대 왜곡-정보량의 특성은 선형적인 특성에 가깝게 나타나고 있다. 이는 본 논문에 사용된 왜곡 척도인 식 (14)이 데시벨 단위로 표현되기 때문에, 로그값을 취한 자승 오차와 비트율 간에는 선형적인 특성이 나타나는 것으로 보인다.

그림에서 보면 균일 보간 기법과 비교하여 불균일 보간 기법은 R-D 곡선 상에서 아래쪽에 위치하고 있으며, 이는 불균일 보간 기법이 동일한 비트 수에서 더 작은 최대 오차를 갖음을 의미한다. 이러한 특성은 사람의 머리에서 측정된 HRTF 데이터에 대해 더욱 두드러지게 나타났다. 결론적으로, 불균일 보간 기법은 보다 적은 정보량으로 균일 기법보다 더 작은 왜곡을 갖는 HRTF를 생성할 수 있음을 알 수 있다.

#### 4.2. 동일한 공간감을 얻기 위한 HRTF의 최소 정보량

일반적으로 보간에 사용되는 샘플 HRTF의 개수가 감소하면, 보간 간격이 넓어지며, 이에 따른 보간 오차가 증가하기 때문에 청취 상으로 방향감이 왜곡될 수 있다. 4.1 절에서 제시한 실험 결과에 따르면 최대 허용 오차를 증가시키면 정보량이 감소함을 알 수 있었는데, 최대 허용 오차의 증가에 따라 보간된 HRTF는 본래의 HRTF와 비교하여 그릇된 방향감을 가져올 수 있다.

본 논문에서는 보간 간격을 5도에서 45도 사이로 다양하게 변화시키며 보간된 HRTF를 얻고 이를 유색잡음 (colored noise,  $-3$  dB/Oct)에 통과시켜 보간 왜곡에 따른 청취 상의 방위각 변동을 살펴보았다. 사용된 유색잡음은 1초의 길이를 갖으며 전, 후 100 msec 기간 동안 코사인 램프 함수 (cosine lamp function)을 곱하여 신호의 시작과 종료가 완만하게 변화하도록 하였다. 4개의 HRTF 데이터에 대해 각각 576개의 보간된 HRTF를 사용

하였으며 이들 HRTF를 통과한 유색잡음을 청취자에게 들려주어 방위각의 변동 정도를 살펴보았다. 청취자는 청각에 특별한 이상이 없는 10명으로서 연령대는 20~40대였다. 청취 테스트는 왜곡이 없는 원래의 HRTF로 필터링된 유색잡음을 먼저 들려주고, 다음으로 동일한 방위각을 갖는 보간된 HRTF로 필터링된 신호를 들려주어 두 신호간의 음원 방향이 동일한 지 여부를 테스트자에게 기입하도록 하였다. 청취 테스트는 배경 잡음이 없는 조용한 환경에서, DT990PRO 개방형 헤드폰 (Beyer-dynamic, Germany)을 사용하였다. 실험에 사용된 보간 HRTF는 방위각 당 약 8개이며 보간 왜곡은  $-20\sim-1.6$  사이의 범위를 나타내었다.

청취 테스트 결과 100%의 청취자가 공간감의 차이를 느끼지 않는 최대 오차는 KEMAR 데이터의 경우  $-15.036$ , 3개의 사람 HRTF 데이터에 대해서는 각각  $-15.189$ ,  $-14.356$ ,  $-15.560$ 이었다. 이 값을 제안된 기법의 보간 오차 문턱치  $D_{thres}$ 로 설정하는 경우, 소요되는 HRTF의 정보량은 KEMAR의 경우 475724 bits, 3개 사람 HRTF 데이터에 대해서는 각각 426608 bits, 459352 bits, 508468 bits였다. 이를 HRTF의 원래 정보량 72 (방위각의 개수) $\times$ 512 (1개 HRIR 이 길이) $\times$ 2 (채널 수) $\times$ 16 (HRIR 샘플당 비트 수)=1179648 bits와 비교하면 정보 감축율은 0.36~0.43 정도임을 알 수 있다. 즉, 5도 간격의 방위각으로 측정된 HRTF 데이터에 대해 약 36%~43% 정보만으로 청취 상으로 방향감이 손실되지 않는 HRTF 데이터를 얻을 수 있음을 알 수 있었다.

그림 5에 방향감의 차이가 없도록 허용 최대 오차를 설정한 경우, 샘플된 방위각을 나타내었다. 그림 5에서 검은색 지점은 샘플된 방위각의 위치를, 회색 지점은 보간이 수행되는 방위각의 위치를 나타낸다. 예로서, KEMAR HRTF의 경우 0도에 해당하는 HRIR과 25도에 해당하는 HRIR을 이용하여 사이 각인 5, 10, 15, 20도의 HRIR이 보간에 의해 얻어짐을 나타낸다. 그림 5에서 샘플된 방위각은 영역에 따라 그 간격이 불규칙하게 나타나며 KEMAR 및 3개 사람 (NJ1, NJ2, NJ3)에 대해 각기 다른 패턴임을 알 수 있다. 샘플된 방위각이 이와 같이 측정 대상마다 각기 다르게 나타나는 것은 측정시 사용된 도부소의 모양, 그리고 측정에 참여한 사람의 두상, 귓바퀴의 모양에 따라 음원-귀 간의 경로가 각도에 따라 각기 다르게 나타나는 것에 그 원인이 있는 듯하다. 샘플 간격이 일정하지 않으므로, 본 기법을 실제로 적용하기 위해서는 각 방위각 별로 샘플된 방위각인지, 보간에 의해 계산될 방위각인지에 대한 정보를 추가로 저장할 필요가 있다.

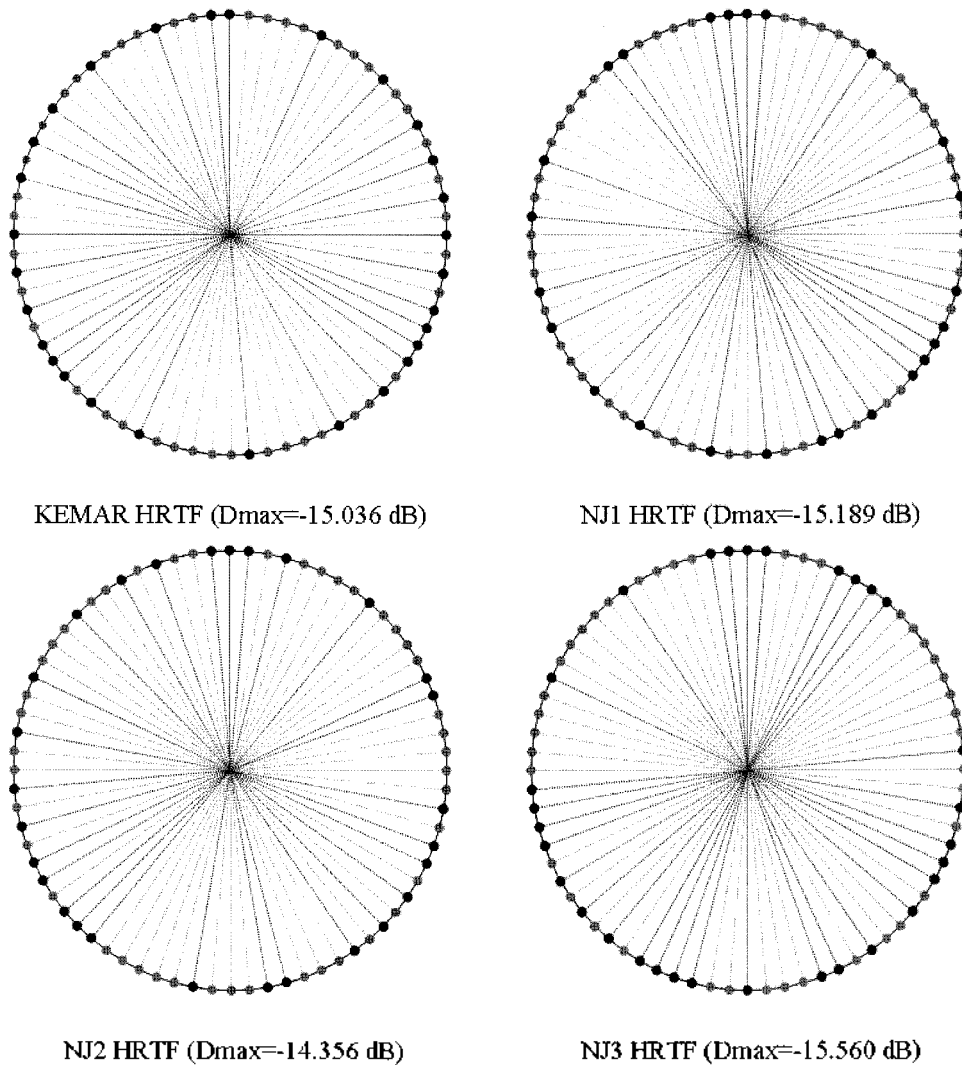


그림 5. 4개 HRTF 데이터에 대한 청취 방향 차이가 없는 HRTF의 최소 샘플링 패턴  
 Fig. 5. The minimum sampling HRTF patterns without audible directional differences for 4 HRTFs.

### V. 결론

본 논문에서는 청취 방향감 생성에 중요한 역할을 담당하는 머리전달 함수의 정보량 감축 기법에 대해 살펴보았다. 머리전달 함수의 충격파 응답을 시간 영역에서 보간하기 위하여 원형 머리 모델에 바탕을 둔 시간 지연 모델이 사용되었으며, 최대 왜곡 최소화 기법에 바탕을 둔 샘플링 기법을 제안하였다. 제안된 기법은 보간으로 인한 오차의 최대값을 문턱치 값 이하로 하면서, 샘플된 HRTF의 개수를 최소화하도록 샘플링이 이루어진다.

1개 모형 머리와 3명의 실제 사람으로부터 얻은 HRTF에 대해 제안된 기법을 적용한 결과, 원래 HRTF의 30~40% 정보량만으로도 방향감이 손실되지 않음을 알 수 있었으며, HRTF의 종류에 따라 샘플 패턴은 각기 다름을 알 수 있었다.

후 후 연구과제보는 청취상의 방향감 차이를 보다 잘 반영할 수 있는 왜곡 모델을 제안하는 것이며, 제안된 기법을 다양한 고도각에 대해 적용하여 결과를 살펴보는 것을 들 수 있겠다. 또한 보다 많은 정보 감축을 위해서는 제안된 보간 기법을 2차원으로 확장하여 방위각 및 고도각의 동시 샘플링에 적용하는 방법을 생각해볼 수 있다.

### 참고 문헌

1. P. F. Holfmann and H. Moller, "Audibility of time differences in adjacent head-related transfer functions", Audio Engineering Society Preprint, 121st Convention, preprint no.6914, 2006.
2. D. 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., 91(3), 1637-1647, 1992.
3. J. Breebaart and A. Kohlrausch, "The perceptual (ir) relevance of HRTF magnitude and phase spectra", Audio Engineering



Society Preprint, 110th Convention, preprint no.5406, 2001.

4. A. Kulkarni, S. K. Isabelle and H. S. Colburn, "On the minimum-phase approximation of head-related transfer functions", Proceedings of IEEE ASSP workshop on applications of signal processing to audio and acoustics, 84-87, 1995.
5. P. Minnaar and J. Plogsties and F. Christensen, "Directional resolution of head-related transfer functions required in binaural synthesis", J. Audio Eng. Soc., 53(10), 919-929, 2005.
6. R. Tamura, Y. Hiraiwa, H. Hasegawa and M. Kasuga, "Research on the creation of a sound field using a 2-loudspeaker mobile phone", Proceedings of IEEE region 10 conference TENCON 2004, 120-123, 2004.
7. S. Hwang and Y. Park, "Time delay estimation from HRTFs and HRIRs", Proceedings of the 8th international conference on motion and vibration control (MOVIC 2006), 2006.
8. R. Nicol, V. Lemaire, A. Bondu and S. Bussan, "Looking for a relevant similarity criterion for HRTF clustering: a comparative study", Audio Engineering Society Preprint, 120th Convention, preprint no. 6653, 2006.
9. S. Kim, S. Jang, D. Kong and S. Bang, "Adaptive virtual surround sound rendering method for an arbitrary listening position", Proceedings of AES 30th international conference, 2007.
10. R. O. Duda and W. L. Martens, "Range dependence of the response of a spherical head model", J. Acoust. Soc. Am., 104(5), 3048-3058, 1998.
11. R. S. Woodworth and G. Schlosberg, *Experimental Psychology*, (Holt, Rinehard and Winston, New York, 349-361).
12. A. Kulkarni, S. K. Isabelle and H. S. Colburn, "Sensitivity of human subjects to head-related transfer-function phase spectra", J. Acoust. Soc. Am., 105(5), 2821-2840, 1999.
13. K.-S. Lee, "Temporal decomposition based on a ratedistortion criterion", IEEE Signal Processing Letters, 11(1), 33-35, 2004.
14. 이기승, "비트율-왜곡 기반 음성신호 시간축 분할", 한국음향학회지, 21(3), 315-322, 2002.
15. D. D. Rife and J. Vanderkooy, "Transfer-Function Measurements using Maximum-Length Sequences", J. Audio Eng. Soc., 37(6), 419-444, 1989.
16. MIT media lab, KEMAR HRTF data, <http://sound.media.mit.edu/pub/Data/KEMAR>
17. 황성목, 박영진, 박윤식, "Comparison of head-related transfer function models based on principle components analysis", 한국소음진동공학학회논문집, 18(6), 642-653, 2008.
18. S. Shimada, N. Hayashi and S. Hayashi, "A clustering method for sound localization transfer functions", J. Audio Eng. Soc., 42(7/8), 577-584, 1994.
19. F. Keyrouz and K. Diepold, "A new HRTF interpolation approach for fast synthesis of dynamic environmental interaction", J. Audio Eng. Soc., 56(1/2), 28-35, 2008.

저자 약력

•이 기 승 (Ki-Seung Lee)



1991년 2월: 연세대학교 전자공학과 공학사  
 1993년 2월: 연세대학교 대학원 전자공학과 (공학 석사)  
 1997년 2월: 연세대학교 대학원 전자공학과 (공학 박사)  
 1997년 3월~1997년 9월: 연세대학교 신호처리 연구센터 선임 연구원  
 1997년 10월~2000년 9월: AT&T Shannon Lab 연구원  
 2000년 11월~2001년 8월: 삼성종합기술원 HCI Lab 전문연구원  
 2001년 9월~2007년 8월: 건국대학교 정보통신대학 전자공학부 조교수  
 2007년 9월~현재: 건국대학교 정보통신대학 전자공학부 부교수  
 \*주관심 분야: 생체 신호 처리, 오디오 신호 처리 등.

•이 석 필 (Seok-Pil Lee)



1990년 2월: 연세대학교 전기공학과 (공학사)  
 1992년 8월: 연세대학교 대학원 전기공학과 (공학 석사)  
 1997년 8월: 연세대학교 대학원 전기공학과 (공학 박사)  
 1997년 7월~2002년 2월: 대우전자(주) 선임연구원  
 2002년 3월~현재: 전자부품연구원 방송통신융합 연구센터장  
 \*주관심 분야: 비디오 신호 처리, 오디오 신호 처리