

가상현실 환경에서의 3 차원 사운드 생성을 위한 거리 변화에 따른 구조적 머리전달함수 모델

이영한, 이길호, 김홍국
광주과학기술원, 정보통신공학과
{cpumaker,ghlee,hongkook}@gist.ac.kr

A Range Dependent Structural HRTF Model for 3D Sound Generation in Virtual Environments

Young Han Lee, Gil Ho Lee, and Hong Kook Kim
Dept. of Information and Communications
Gwangju Institute of Science and Technology

요약

본 논문에서는 가상현실 환경에서 방위각, 고도, 거리 등의 위치 정보를 인지할 수 있는 3 차원 사운드를 생성하기 위한 구조적 머리전달함수(Head-Related Transfer Function, HRTF) 모델을 제안하고 이를 구현한다. 이를 위해 우선 기존에 연구된 방위각과 고도에 관한 head model 과 pinna model 을 근간으로 거리의 변화를 고려하도록 하는 HRTF 모델을 제안한다. 제안된 거리 변화에 따른 HRTF 모델은 거리 차이에 따른 음압의 변화 모델과 근거리에서 사람의 머리에 의해 생기는 shadowing 모델로 구성된다. 본 논문에서는 제안한 모델로 부터 mono 사운드를 방위각, 고도, 거리를 인지할 수 있는 stereo 사운드로 변환하여 생성하는 3 차원 사운드 생성기를 구현하였고, 일반인을 대상으로 거리에 대한 청취 실험을 통하여 제안한 모델의 성능을 측정하였다. 그리고 제안된 모델을 가상현실의 실감모델인 MP3 에 구현하여 그 효과를 입증하였다.

Keyword : HRTF, Range dependence, Structural model, Virtual reality

1. 서론

가상현실 연구는 지난 수십 년 동안 많은 발전을 이루어 왔다. 가상현실 실현을 위한 중요한 구성요소로서의 3 차원 사운드는 영상과 결합할 경우 가상현실의 몰입도를 높여주는 역할을 한다 [1]. 다시 말해서 가상현실 구현의 완성도를 높일 수 있다는 측면에서 가상현실에서의 3 차원 사운드는 매우 중요한 요소로 자리매김을 하고 있다.

그 동안 3 차원 사운드에 관한 연구는 주로 음원(sound source)의 위치와 청취자의 귀 사이의 관계를 찾는 데 집중되어 왔다 [2]. 일반적으로 이러한 음원과 청취자 사이의 관계를

머리전달함수 (Head-Related Transfer Function, HRTF)라 부르고 이것을 통해 3 차원 사운드를 생성한다.

인간은 양 귀의 고막에서 들리는 음원 크기의 차이인 ILD (Interaural Level Difference)와 고막에 도달하는 시간 차이인 ITD (Interaural Time Difference)로 음원의 위치를 인지한다 [2]. 모든 방위각(azimuth)과 고도(elevation), 거리(distance)에 따라 ITD 와 ILD 는 변하게 되고, 이 변화를 이용하여 인간은 후천적 학습으로 음원의 위치를 인지한다. 하지만 ITD 나 ILD 는 음원의 위치 뿐만 아니라 음원의 주파수 성분에 따라 달라지며, 머리의 크기, 형태, 귀의 위치, 귓바퀴의 모양 등 개인의 신체조건에 민감하게 변화한다.

신체조건에 따라 HRTF 가 달라지기 때문에 많은 연구자들은 통계를 통해 인간의 표준 머리모델을 만들고 그것을 이용하여 방위각, 높이, 거리에 따른 HRTF 를 측정해왔다. 하지만 이와 같이 측정된 HRTF 를 이용한 가상현실 구현은 모든 각도 및 모든 거리에 대한 HRTF 가 필요하다는 단점이 있다. 그리고 위에서 언급한 것처럼 개인의 신체조건에 민감하기 때문에 표준 머리모델을 대상으로 측정을 했음에도 불구하고, 음원의 위치 인지 효과는 개인차가 심하다. 따라서 개인에 적합한 HRTF 를 측정해야 한다. 그러나 개인에 맞는 HRTF 를 모두 측정한다는 것은 현실적으로 큰 어려움이 있다. 이러한 단점을 보완하기 위해 제안되어온 방법이 HRTF 모델링이다. 많은 연구자들은 측정된 HRTF 를 근거로 HRTF 를 수학적으로 표현하는 HRTF 모델링 방법을 연구해왔다 [3][4][5]. 이러한 모델링 방법은 방위각, 고도, 거리를 변수로 가지기 때문에 모든 경우에 대하여 HRTF 를 측정하지 않아도 될 뿐만 아니라 3 차원 사운드를 생성하는데 있어서 메모리를 효율적으로 사용할 수 있다. 뿐만 아니라, 인간의 신체조건을 변수로 표현하여, 청취자에 따라 적합한 HRTF 를 만들 수 있다.

HRTF 를 모델링하는 방법 중에서 본 논문에서는 구조적 머리전달 함수 모델 (structural HRTF model)을 기반으로 하여 거리 변화에 따른 모델을 제안한다. 구조적 HRTF 모델은 그림 1 과 같이 머리에 의한 효과, 귀에 의한 효과 등 여러 개의 블록으로 나뉘어지며 이를 통해 3 차원 사운드를 생성한다.[3]. 먼저 입력으로 들어온 모노 입력 신호는 Head Shadow & Delay 블록을 거치면서 방위각에 대한 HRTF 가 적용된다. 그리고 나서, 고도를 느끼게 해주는 컷바퀴 효과(pinna echo) 및 청취환경(room echo), 흉부 효과(shoulder echo)를 거쳐 측정된 HRTF 를 이용한 3 차원 오디오와 유사한 효과를 얻게 된다. 이를 이용하면 방위각, 고도에 대한 HRTF 를 구현할 수 있지만, 거리 변화에 대한 HRTF 변화는 고려하지 않았다..

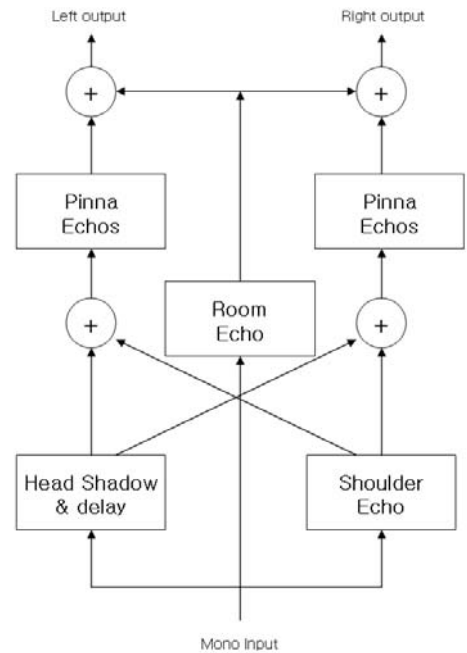


그림 1. 기존 구조적 머리전달 함수 모델

가상현실 실현을 위한 3 차원 음향 생성에 있어서 방위각, 고도에 따른 음의 변화가 필요하지만, 거리에 따른 변화를 구현없이는 가상현실에 적용하였을 때 사용자의 몰입도가 떨어지게 된다. 따라서 본 논문에서는 가상현실 사용자가 가상현실에 몰입하는 것을 돕기 위한 거리 변화에 대한 HRTF 모델을 제안한다

본 논문의 구성은 다음과 같다. 제 2 장에서는 거리에 따른 HRTF 의 특징을 살펴보고, 제 3 장에서는 HRTF 의 특징을 이용하여 거리 변화에 따른 구조적 머리전달함수를 제안한다. 그리고 제 4 장은 제안한 구조적 머리전달함수를 이용하여 생성한 3 차원 사운드를 통해 제안한 모델의 성능을 알아보고, 제 5 장에서 결론은 맺도록 한다.

2. 거리 변화에 따른 HRTF 의 특징

먼저, 인간은 음원의 거리를 인지하기 위해 음의 크기와 잔향(Reverberation)을 이용한다. 하지만 이 중에서 잔향은 청취자의 가상현실 환경, 즉 운동장, 산속, 콘서트 홀 등에 따라 크게 바뀌기 때문에 본 논문에서는 거리 변화에 있어서

음의 크기 만을 고려하기로 한다.

거리 변화를 고려한 HRTF 모델을 만들기 위해 우선 공간을 구분할 필요가 있다. 그 이유는 청취자가 인지하는 거리에 따라 모델링하는 방법을 다르게 해야만 하기 때문이다. 다시 말해서, 가까운 거리와 먼 거리는 다른 방법에 의해 모델링해야 할 필요가 있다 [6]. 일반적으로 머리에서 1 meter 내의 거리를 근거리 지역(proximal region)으로, 그리고 1 meter 밖의 거리를 원거리 지역(distal region)으로 나눈다. 일반적으로 원거리에 대한 모델링 방법은 그림 2(a)에서 보이는 바와 같이 ITD 는 유지한 반면 ILD 만에 변화를 준다. 이를 위해 inverse square law 를 적용하여 거리 변화에 따른 HRTF 를 모델링한다. 하지만 음원이 근거리에 있다고 가정을 한다면 모델링하는 방법은 달라지게 된다. 그림 2(b)에서 보는 것과 같이 근거리일 경우는 ITD 가 변하게 되고, ILD 는 shadowing 되는 영역으로 인하여 inverse square law 를 적용하기 어렵다. 따라서 본 논문에서는 거리 변화를 고려한 HRTF 모델을 공간에 따라 크게 근거리와 원거리로 나누어 제안한다.

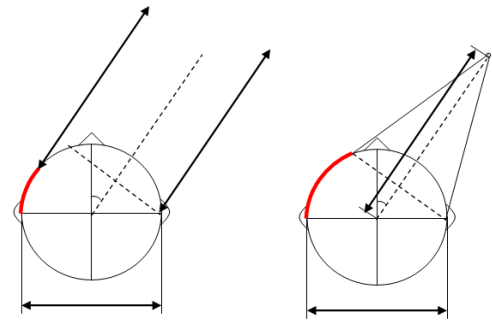
2-1 원거리 HRTF

인간이 음원의 거리를 인지하기 위해 앞에서 언급한 것처럼 ILD 와 ITD 를 이용한다. 하지만 음원의 위치가 머리에서 충분히 멀다고 가정할 경우 같은 방위각과 고도를 가지는 신호는 거리가 다르더라도 동일 한 ITD 를 가지게 된다고 가정할 수 있다. 그리고 구조적 HRTF 모델의 경우 head shadow & delay 를 모델링을 하기 위해 충분히 멀리 떨어져 있는 음원에서 발생하는 신호를 모델링한다고 가정을 하면, 전체적인 음의 크기 조절을 통해 만족할 만한 결과를 얻을 수 있다. 이를 위해 inverse square law 를 사용한다

2-2 근거리 HRTF

근거리에서의 HRTF 특징을 설명하기 위해 음원에 가까이 있는 귀를 near ear, 멀리 있는 귀를 far ear 로 명명한다. 근거리의 경우 그림 2 에서

보는 것처럼 음파의 진행경로가 달라지기 때문에 원거리와 다르게 ITD 에도 역시 변화가 생긴다.



(a) 원거리에서 ITD 차이 (b) 근거리에서 ITD 차이

그림 2. 거리에 따른 ITD 변화

하지만 ITD 의 변화는 인간이 인지하지 못 할만큼 미미하기 때문에 ITD 의 변화가 없다고 가정하고 거리에 대한 모델링을 한다 [6]. ILD 은 기본적으로는 Inverse square law 를 따른다. 하지만 근거리의 경우 원거리와는 다르게 머리에 의해 far ear 에 음영(shadow) 효과가 나타나게 되고 이로 인해 음의 크기가 작아지게 되고, 고주파 성분에서는 감쇄가 일어난다 [7].

3. 거리 변화에 따른 구조적 머리전달함수 모델 제안

본 논문에서 제안한 거리 변화에 따른 구조적 머리전달함수 모델은 다음 그림 3 과 같이 기존 구조적 머리전달함수의 후처리 단으로 연결(cascading)된다.

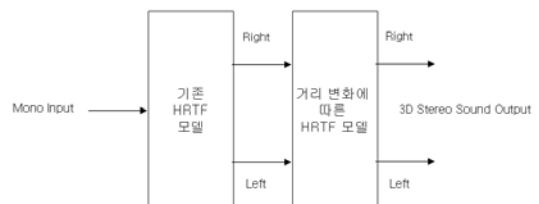


그림 3. 제안한 구조적 머리전달함수 모델

세부적으로 제안한 모델은 크게 3 개의

블록으로 나누어 진다.

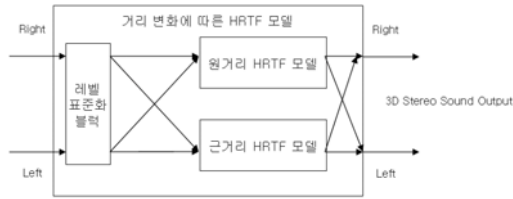


그림 4. 제안한 모델의 세부 블록

그림 4 에서 보는 것과 같이 제안한 모델은 음의 레벨을 표준화하는 블록 (Level Normalization Block)과 제 2 장에서 설명한 바와 같이 원거리 HRTF 모델 블록과 근거리 HRTF 모델 블록으로 나뉜다. 그리고 WAVE 파일 형식의 표현범위의 한계를 고려하여 표현 가능한 거리의 범위는 0.25m 에서 5m 로 제한한다.

3-1 레벨 표준화 블록

거리 변화에 따른 구조적 HRTF 모델의 입력은 16 비트 WAVE 파일이다. WAVE 파일은 -32768 에서 32767 까지 음원 신호의 표현이 가능하기 때문에 제안하는 HRTF 모델에서 음의 크기를 조절할 경우 표현 가능한 범위를 넘는 경우(overflow)가 발생하게 된다. 이러한 overflow 문제를 해결하기 위해 WAVE 파일 샘플 중 최대값이 4096 이 넘을 경우 최대값이 4096 이 되도록 전체적으로 음원 신호의 크기를 조절하여준다.

3-2 원거리 HRTF 모델

거리 변화에 따른 원거리 HRTF 모델은 기존 구조적 HRTF 모델로 생성된 left/right stereo 사운드에 inverse square law 를 적용한다. 이를 위한 사운드 크기를 다음 식을 통해 얻는다.

$$\frac{I_m}{I_{ref}} = \left(\frac{d_{ref}}{d_m} \right)^2 \quad (1)$$

여기서 I_m 는 거리에 따른 음의 크기이고,

d_m 는 음원과의 거리, d_{ref} 는 기준 거리인 1m,

I_{ref} 는 3-1 절의 레벨 표준화 블록에 의해 표준화된 음의 크기이다.

3-3 근거리 HRTF 모델

근거리 HRTF 모델의 경우 머리에 의해 shadowing 효과를 나타내는 블록과 음원의 level 을 조정하기 위한 블록으로 나뉜다. 다음 그림 5 는 근거리 HRTF 모델을 나타낸 것이다.

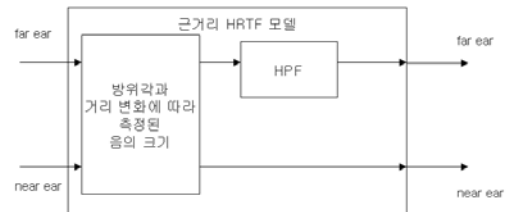


그림 5. 근거리 HRTF 모델

근거리에서는 far ear 의 음의 크기가 가까워질수록 작아지는 것과 고주파가 감쇄하는 특성이 있다 [7]. 근거리에서 far ear 에 들리는 음의 크기를 효율적으로 구하기 위하여 다음과 같은 수식을 이용하여 모델링 한다.

$$\frac{I_{m, far ear}}{I_{ref}} = \left(\frac{d_m}{d_{ref}} \right) \quad (2)$$

여기서, $I_{m, far ear}$ 는 far ear 음의 크기이다

측정된 HRTF 로 근거리에 대한 음의 크기를 조절해준 후 far ear 에 대해 고주파 감쇄를 위해 HPF 를 far ear 에 적용한다. 이를 위해 차단 주파수를 4 kHz 를 가지는 2 차 IIR Butterworth 필터를 사용하였다.

식 (3)은 near ear 를 위한 사운드 크기 조절 방식을 나타낸다. Near ear 의 경우 측정된 HRTF 음의 크기가 거리 비율에 비례하고 각도에 의해 영향을 받음을 알 수 있다.

$$\frac{I_{m,near\ ear}}{I_{ref}} = \left(\frac{d_{ref}}{d_m} \right) \times \frac{1}{2^{\frac{|azimuth|}{90}}} \quad (3)$$

여기서, $I_{m,near\ ear}$ 는 near ear 음의 크기이고 azimuth 는 방위각을 나타내며 정면을 0 도로 하여 near ear 로 향하는 각도로써 범위는 -90 도 ~ +90 도로 한다.

4. 거리 변화에 따른 HRTF 모델 구현 및 성능 분석

본 논문에서 제안한 방식으로 구현된 HRTF 모델을 실감 모델 중 하나인 MP3 플레이어 조작에 관한 가상현실에 다음과 같이 적용하였다. 먼저, MP3 로 압축을 위한 음원 데이터를 수집한 후, 16 비트 44.1 kHz, mono 형식의 WAVE 파일로 저장한다. 그리고 가상현실에서 사용자가 MP3 플레이어를 조작하다가 그로 인한 이벤트가 발생한다면, 가상현실 시스템으로부터 구현한 HRTF 모델은 MP3 플레이어의 현재의 위치, 즉 방위각, 고도, 거리를 입력 받는다. 이 때 위치 정보는 구좌표계(spherical coordinate)로 표현되며, 각 축은 방위각, 고도, 거리에 해당한다. 이렇게 입력된 위치 정보로부터 3 차원 사운드를 생성한다. 이러한 HRTF 모델은 앞에서 언급한 것처럼 모든 위치에 대한 HRTF 를 필요로 하지 않기 때문에, 메모리나 저장공간 측면에서 효율적이다.

가상현실에서 MP3 플레이어 출력음을 생성하는 거리변화에 따른 구조적 HRTF 모델의 성능을 테스트하기 위해 본 논문에서는 제안한 HRTF 모델을 이용하여, 일반인을 대상으로 거리에 대한 청취 실험을 하였다. 실험에서 고도는 고정된 채, 방위각 0 도, 30 도, 60 도, 90 도, 거리 0.25m, 0.50m, 1m, 1.5m, 2m 에 대하여 헤드폰을 이용하여 청취 실험이 이루어 졌다. 실험 방법은 제안된 HRTF 모델로 생성한 3 차원

사운드를 1m 기준음과 비교 청취 후 생성된 3 차원 사운드의 거리를 예측하도록 하였다. 실험에 사용한 음원은 16 bits, 44.1 kHz, mono WAVE 파일 형식으로 모든 주파수 성분을 고루 가지고 있는 총소리를 이용하였다. 일반인 10 명을 대상으로 실험을 진행하였으며 실험 결과는 표 1 과 같다.

<표 1. 3 차원 사운드를 이용한 거리 인지 실험>
(단위 : m)

생성방위각	생성 거리	인지 평균	인지오차
0 도	0.2	0.32	0.11
	0.5	0.70	0.20
	1.5	1.81	0.33
	2.0	2.22	0.28
우현 30 도	0.2	0.35	0.14
	0.5	0.65	0.15
	1.5	1.55	0.33
	2.0	2.02	0.44
우현 60 도	0.2	0.34	0.13
	0.5	0.67	0.17
	1.5	1.32	0.28
	2.0	2.08	0.26
우현 90 도	0.2	0.34	0.13
	0.5	0.65	0.15
	1.5	1.41	0.31
	2.0	1.87	0.25

실험 결과에서 알 수 있듯이 생성한 3 차원 사운드와 사람이 인지한 거리는 근거리에서 최대 0.4m, 평균 0.15m 의 오차를 가지고, 원거리에서는 최대 0.5m, 평균 0.31m 의 오차를 가지는 것으로 나타났다. 이는 실제로 인간이 거리를 인지할 때 1m 에서 5m 사이에서 최대 0.55m 의 오차를 나타내는 것을 감안한다면 만족할 만한 결과라고 할 수 있다 [2]. 오른쪽으로 60 도 일 경우 가장 근사한 결과를 얻을 수 있었고, 방위각에 따른 큰 편차를 보이지 않음을 알 수 있다.

5. 결론

본 논문에서는 실감 모델의 위치 변화에 따른 효과적인 3 차원 사운드를 생성하기 위해 거리 변화에 따른 구조적 머리전달함수를 제안하였다. 또

한 이 머리전달함수로 생성한 3 차원 오디오를 이용하여 제안한 알고리즘의 성능을 측정하고 그 오차를 분석하였다. 이는 기존 구조적 HRTF 모델에 거리 변화에 따른 사운드를 제공하므로 가상현실 사용자의 몰입도를 높이는데 효과적으로 평가되었다.

6. Acknowledgement

본 연구는 광주과학기술원 실감방송 연구센터(RBRC)를 통한 정보통신부 대학 IT 연구센터(ITRC) 사업의 지원, 광주과학기술원 실감콘텐츠 연구센터(ICRC)를 통한 과학기술부 특정연구개발 사업의 지원과 광주과학기술원[GIST] 기관고유사업 지원에 의해 수행되었습니다.

참고문헌

- [1] 강성훈, “오디오와 심리음향,” 국제음향공학워크샵, pp. 1 - 11, 1996 년 7 월.
- [2] J. Blauert, *Spatial Hearing*, Revised Edition, MIT Press, Cambridge, MA, 1997.
- [3] C. P. Brown and R. O. Duda , “An efficient HRTF model for 3-D sound,” in *Proc. WASPAA*, Oct. 1997.
- [4] Kistler, D.J. 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, 1992.
- [5] N.M. Cheung, S. Trautmann, and A. Horner, "Head-Related Transfer Function Modeling in 3-D Sound Systems with Genetic Algorithms," *Journal of the Audio Engineering Society*, Vol 46, No 6, pp. 531-539, 1998
- [6] R. O. Duda and W. L. Martens, “Range dependence of the response of a spherical head model,” *J. Acoust. Soc. Am.*, vol. 104, no. 5, pp. 3048-3058, Nov. 1998.
- [7] D. S. Brungart and W. M. Rabinowitz, “Auditory localization of nearby sources. Head-related transfer functions.” *J Acoust Soc Am*, Vol 106, No 3, pp. 1465-

1479. 1999.

[8] D.R. Begault, *3D Sound for Virtual Reality and Multimedia*, Academic Press, Cambridge, MA, 1994.

[9] W. G. Gardner and K. D. Martin, “HRTF Measurements of a KEMAR,” *J. Acoust. Soc. Amer.*, Vol. 97, No 6, pp. 3907-3908, Jun. 1995.