

청각특성을 이용한 머리전달함수의 저차 모델링에 관한 연구

김 동 현*, 김 기 만**

A Study on Low Order Modeling of Using Auditory Characteristics

Dong-Hyun Kim*, Ki-Man Kim**

Abstract

Head Related Transfer Functions(HRTFs) are the source of information for research of spatial hearing and describe the relations between the perception of a sound and the sound source. However, one of the problems in 3-D sound synthesis is the computational load of accurate HRTF approximation. To overcome this, both computationally efficient and perceptually relevant digital models of HRTFs have to be created. In general, IIR filters have the advantage over FIR filters of representing transfer functions with lower orders, but are more difficult to design. Some current techniques that use IIR filters to model HRTFs are based on Prony's method, Yule-Walker method, BMR(balanced model reduction) method.

This thesis proposes a method for low order IIR model using auditory characteristics from high order FIR impulse response. Listening tests have been performed to determine the subjective preference of different HRTF filter design. The results of listening tests showed that proposed IIR model outperform conventional method.

* 한국해양대학교 전파공학과 석사과정 전파공학 전공

** 한국해양대학교 전파공학과 조교수

1. 서 론

임의의 시간에 어떤 특정한 공간에서 발생한 음상(sound image)을 다른 시간, 공간에서 충실히 재현시키는 음상 제어기술은 최근 멀티미디어(Multimedia), 가상 현실(Virtual Reality), 원격 통신(Telecommunication), 시각 장애자를 위한 보조기기(Aids for Blind), 실감 통신(Actual Feeling Communication) 등 많은 분야에서 주목을 받고 있다[1].

입체 음상 제어 알고리즘은 구성상 헤드폰 재생을 위한 바이노럴 프로세서(binaural processor)와 2 개의 스피커 재생을 위한 트랜스오럴 프로세서(transaural processor) 두가지로 나뉘어진다. 이중 바이노럴 프로세서는 모노(mono) 신호를 임의의 방향에 정위시키는 프로세서로서 마치 입력신호를 바이노럴 녹음으로 얻어진 것과 같은 신호로 변환한다. 트랜스오럴 프로세서는 바이노럴 프로세서에 스피커 재생시 발생하는 누화(crosstalk)를 제거하기 위한 누화제거기(crosstalk canceller)가 추가된 것이다[1][2].

가상 현실 분야에서 주목되고 있는 청각 디스플레이는 바이노럴 청취를 이용한 3 차원 음상 정위 시스템이다. 음원에서 두 귀의 고막까지 전달 함수를 미리 몇 군데 위치에서 측정해 저장하고, 고막상에서 음압을 계산하면, 실질적으로 존재하지 않는 음원을 근사적으로 생성할 수 있으며, 이 시스템을 이용하면 헤드폰을 사용할 경우 3 차원적인 음장을 재현할 수 있다. 이를 시뮬레이션하기 위해서는 여러 가지 파라미터, 즉 공간전달함수, 머리전달함수(Head Related Transfer Function ; HRTF)등이 필요하게 되는데, 이중 HRTF 는 자유공간에서 정위된 음원으로부터 사람의 귀로 전달되는 음향적 과정을 표현하며 사람이 음원의 위치를 판단하는 정보를 포함하고 있다.

HRTF 는 일반적으로 무향실(anechoic chamber)의 임의의 어떤 위치에서 사람의 귀 또는 더미헤드(dummy head)에 설치된 마이크로폰의 임펄스 응답을 측정하여 얻을 수 있으며, 측정된 임펄스 응답을 직접 필터의 계수로 사용하여 FIR(Finite Impulse Response)형태의 HRTF 필터를 설계할 수 있다[2]. 그러나 측정된 임펄스 응답은 시간영역에서 윈도우링(windowing)하고, 최소위상변화를 갖도록 재구성해도 약 44.1 kHz 의 표본화 주파수(sample rate)에서 적어도 100~200 개의 샘플을 가지므로, 실시간 FIR 필터를 구현하기 위해서는 매우 많은 계산이 따른다. 따라서, 원래의 HRTF 가 음의 지각에 미치는 특징을 포함하며 실시간 계산이 가능하도록 계산의 효율을 높인 적절한 HRTF 모델의

개발이 필수적이라 할 수 있다[3][4].

IIR(Infinite Impulse Response)필터는 FIR 필터보다 낮은 차수로 모델링이 가능하므로 HRTF 필터를 실시간 계산이 가능하도록 낮은 차수로 설계 할 때 많이 사용되고 있으며, 이러한 설계 방법으로는 Prony 방법, Yule-Walker 방법, BMR 방법 등이 있다[3]. 일반적으로 BMR 방법을 이용하여 IIR filter 를 설계하는 것이 Prony 방법과 Yule-Walker 방법을 사용할 때 보다 좀더 효율적으로 설계 가능하다.

따라서, 본 논문에서는 BMR(Balanced Model Reduction)방법을 이용하여 높은 차수를 가지는 FIR 형태의 HRTF 필터를 IIR 형태로 설계하고, 입체음향 시스템의 실시간 구현시 필수적인 계산의 효율을 높이는 방법을 제시한다. 또한 IIR 형태로 근사화된 HRTF 필터의 주파수 패턴을 분석하여, 음상정위에 큰 영향을 미치는 주파수 대역을 더 나은 해상도(resolution)를 갖도록 설계하여 저차 모델 설계시 발생하는 오차가 음상정위에 미치는 영향을 최소화하였다.

2. 바이노럴 프로세서와 HRTF 필터 설계

2.1 바이노럴 프로세서

그림 2-1 에서 시간 영역에서의 모노음 x_m 은 왼쪽, 오른쪽 귀의 HRTF(H_l , H_r)와 convolution 을 통해 특정 방향에 정위된 가상의 음 y_l 과 y_r 을 생성한다 [3].

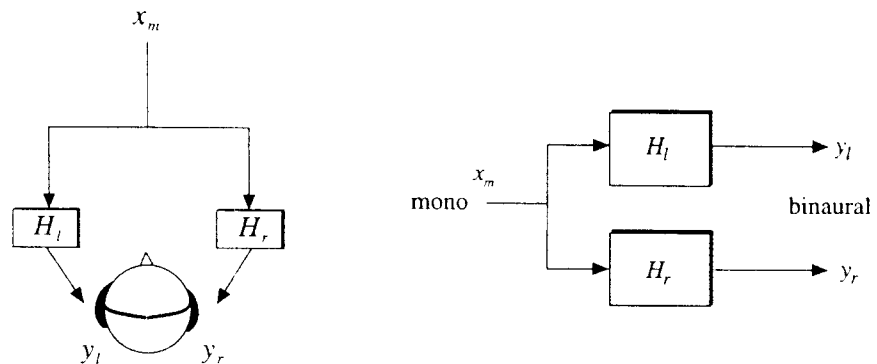


그림 2-1. 바이노럴 프로세싱(monophonic to binaural).

$$y_l = x_m * H_l \quad (2-1)$$

$$y_r = x_m * H_r$$

방향에 따른 HRTF의 데이터 베이스를 사용하고 헤드폰에 대한 보정을 하여 모노 음원을 바이노럴로 바꾸는 바이노럴 프로세서는 그림 2-2 과 같다[6].

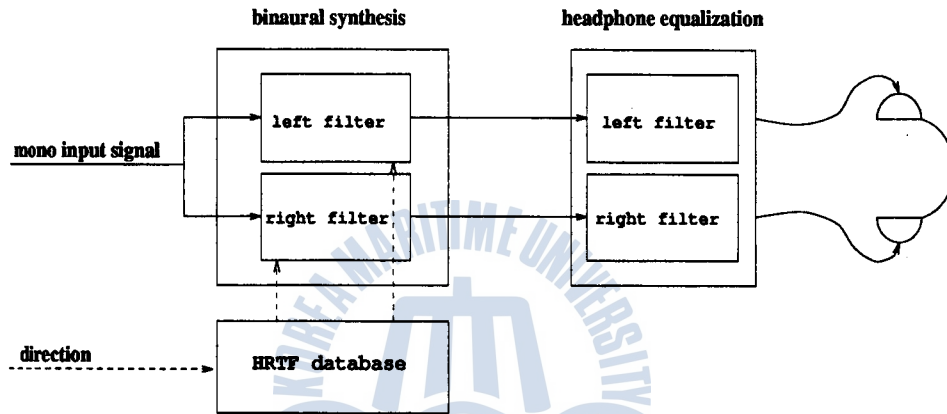


그림 2-2. 일반적인 바이노럴 프로세서 구조.

HRTF는 사람의 머리, 몸통, 귀바퀴의 산란과 회절에 의한 선형 시 불변 시스템의 결과이므로 측정된 임펄스 응답은 FIR 필터로서 직접적으로 구현될 수 있다. 그러나, 데이터 손실 없이 HRTF를 측정하려면 약 5msec가 필요하고 44.1 kHz 표본화 주파수에서 약 220개의 tap 수를 갖는다. 이것을 실시간으로 처리하려면 좌우 두 채널을 convolution 해야 하므로 초당 약 2000만번의 계산을 할 수 있는 프로세서가 필요하다.

이러한 복잡한 계산을 줄이기 위해서는 바이노럴 프로세서의 HRTF 필터를 저차로 모델링하는 방법이 필요하다. 일반적으로 측정된 HRTF를 시간영역에서 windowing하고 최소위상변화를 갖도록 재구성하는 방법을 사용하지만, 이 역시 100 ~ 200개의 샘플 수를 가진다[4][6]. 이보다 더 저차(low order)로 모델링하는 방법으로 두 가지가 제시될 수 있는데, 그 하나는 측정된 HRTF를 resampling하는 방법이다. 만약 측정시 44.1 kHz로 표본화된 HRTF를 그 절반으로 resampling 한다면 22.1 kHz의 표본화 주파수를 가질 것이고 샘플의 수는

50 ~ 100 개로 줄어든 것이다. 하지만 이 방법을 사용하려면 사용되는 음원 역시 22.1 kHz 의 표본화 주파수를 가져야 하므로, 현재 음의 녹음 추세로 볼 때 적당하지 않다. 제시될 수 있는 다른 한가지 방법은 FIR 형태의 HRTF 필터를 IIR 형태로 변형하는 것이다. 일반적으로 FIR 필터를 IIR 필터로 모델링하는 방법에는 Prony 방법, Yule-Walker 방법 등이 있으나, 본 논문에서는 이보다 좀 더 효율적 형태인 Balanced Model Reduction 방법을 사용한다[6].

2.2 Balanced Model Reduction 을 이용한 HRTF 필터 설계

먼저, N 개의 계수값를 가지는 FIR 필터는 다음과 같은 전달 함수 형태로 나타낼 수 있다.

$$\begin{aligned} F(z) &= b_0 + b_1 z^{-1} + \dots + b_n z^{-n} \\ &= b_0 + F_1(z) \end{aligned} \quad (2-2)$$

필터 $F(z)$ 를 state-space 형태로 변환하면 다음과 같다.

$$A = \begin{bmatrix} 0 & 0 & \dots & 0 & 0 \\ 1 & 0 & \dots & 0 & 0 \\ & & \dots & & \\ 0 & 0 & \dots & 1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad (2-3)$$

$$C = [b_1 \quad b_2 \quad \dots \quad b_{n-1} \quad b_n] \quad , \quad D = b_0$$

이제 FIR 형태의 (A, B, C) 시스템에 대해서 다음과 같은 형태의 Hankel matrix 를 구할 수 있다.

$$H = \begin{bmatrix} b_1 & b_2 & \dots & b_n \\ b_2 & b_3 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ b_n & 0 & \dots & 1 \end{bmatrix} \quad (2-4)$$

$$(2-5)$$

식 2-18 의 H matrix 는 symmetric matrix 이므로 다음과 같이 Singular Value

Decomposition(SVD)할 수 있다.

$$H = V \Lambda V^T$$

여기서 Λ 는 singular value 를 대각항으로 갖는 대각행렬(diagonal matrix)이고 V 는 대응하는 고유벡터이며, $V V^T = I$ 이다.

n 차수 FIR 필터 (A, B, C)의 Hankel matrix 가 식 2-5 와 같이 분해된다면, k 차수 reduced balanced 시스템은 다음과 같은 시스템으로 나타낼 수 있다.

$$(A_k, B_k, C_k) \quad (2-6)$$

여기서 각각의 A_k, B_k, C_k 는 다음과 같다.

$$\begin{aligned} A_k &= V_k^T A V_k \\ B_k &= V_k^T B \\ C_k &= C V_k \end{aligned}$$

따라서, 앞서 필터 $F(z)$ 를 state-space 형태로 변환한 식 2-3 을 적용하면 다음과 같이 좀더 간단한 형태로 나타낼 수 있다[7].

$$A_k = V(2:n, 1:k)^T A V(1:n-1, 1:k) \quad (2-7)$$

$$B_k = V(1, 1:k)^T$$

$$C_k = C V(1:n, 1:k)$$

지금까지 FIR 필터에 적합한 형태의 BMR 방법을 정리하면 다음과 같다 [8][10].

1) Hankel matrix H 를 만든다.

- 2) H matrix 를 식 2-5 처럼 SVD 한다.
- 3) Hankel singular value 를 나타내어 reduced order 를 결정한다.
- 4) A_k, B_k, C_k 행렬을 계산한다.
- 5) state-space 로 표현된 시스템을 전달함수형태로 변환한다.

3. 청각특성을 이용한 HRTF 모델링

일반적으로 HRTF 필터를 어느 정도 이상 저차로 설계하게 되면 설계시 발생된 오차로 인해 원래의 HRTF 가 가지는 음의 지각 특징을 잃어 버리게 되며, 이 경우 설계된 필터는 더 이상 HRTF 필터로서의 역할을 할 수 없게 된다. 따라서 본 논문에서는 저차로 설계된 HRTF 필터의 주파수 패턴을 분석, 음상정위에 큰 영향을 미치는 주파수 대역을 좀 더 나은 해상도를 갖도록 설계하여 저차 모델 설계시 발생하는 오차가 음상정위에 미치는 영향을 최소화할 수 있는 HRTF 필터의 설계방법에 대해 제시한다.

azimuth 가 어떤 임의의 각도에서 다른 각도로 변할 때 128 tap FIR HRTF 필터가 가지는 스펙트럼의 오차는 우리가 azimuth 의 변화를 지각하는 단서로 볼 수 있으며, 설계된 필터 역시 그와 비슷한 오차 값을 가질 때 그에 상당하는 azimuth 의 변화를 느낄 수 있다. 따라서 저차 모델링 필터에서 azimuth 가 변할 때 발생하는 주파수 스펙트럼상의 오차를 128 tap FIR HRTF 필터의 경우와 비교하여 오차가 심하게 발생하는 부분의 해상도를 높여줌으로써 음상 정위의 정확성을 향상 시키고자 하였다.

저차로 설계된 필터와 128 tap FIR HRTF 필터를 비교해 볼 때 공통적으로 약 9kHz 이하의 주파수 대역에서 대부분의 오차가 발생하므로, HRTF 필터를 9 kHz 이하와 이상의 대역으로 각각 분해하여 9 kHz 이하의 주파수에서 좀 더 나은 해상도를 갖도록 설계함으로써 설계된 필터의 음상 정위의 정확성을 향상 시킬 수 있다.

9 kHz 이상의 주파수 성분만을 포함하는 왼쪽 귀의 HRTF 를 H_{la} , 9 kHz 이하의 주파수 성분만을 포함하며 H_{la} 보다 나은 해상도를 갖도록 설계된 왼쪽 귀의 HRTF 를 H_{lb} 라 하고, 9 kHz 이상의 주파수 성분만을 포함하는 오른쪽 귀의 HRTF 를 H_{ra} , 9 kHz 이하의 주파수 성분만을 포함하며 H_{ra} 보다 나은 해상

도를 갖도록 설계된 오른쪽 귀의 HRTF 를 H_{rb} 라 할 때, 바이노럴 프로세싱은 다음 식과 같다.

$$y_l = (x_m * H_{la}) + (x_m * H_{lb}) \quad (3-1)$$

$$y_r = (x_m * H_{ra}) + (x_m * H_{rb})$$

여기서, x_m 은 모노 입력음이고, y_l 과 y_r 은 특정방향에 정위된 가상의 음이다.

청각특성을 이용하여 설계된 식 3-1 과 같은 구조의 HRTF 필터는 그림 3-1 과 같다.

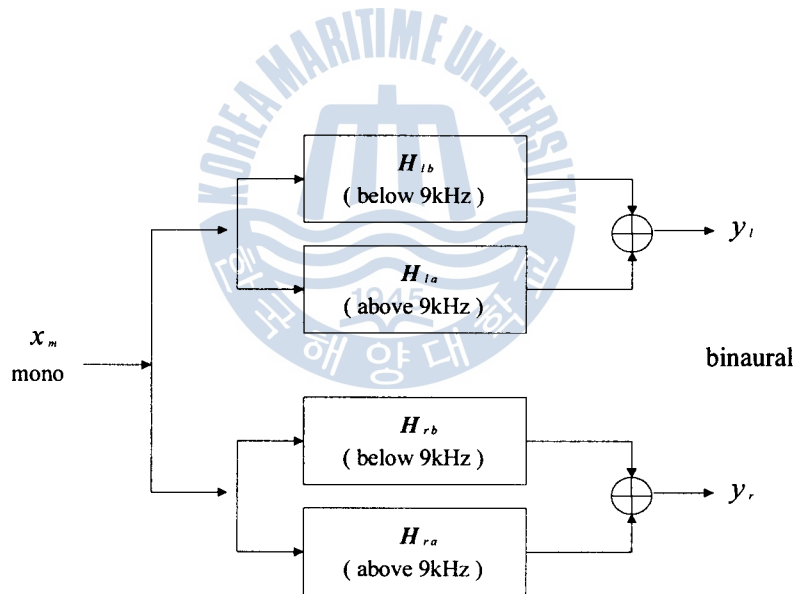


그림 3-1. 청각특성을 이용하여 설계된 HRTF 필터 구조.

4. 결 론

본 논문에서는 BMR 방법을 이용하여 높은 차수를 가지는 FIR 형태의 HRTF

필터를 IIR 형태로 설계하고, 입체음향 시스템의 실시간 구현시 필수적인 계산의 효율을 높이는 방법을 제시하였다. 또한 IIR 형태로 근사화된 HRTF 필터의 주파수 패턴을 분석하고, 음상정위에 큰 영향을 미치는 주파수 대역을 더 나은 해상도를 갖도록 설계하여 저차 모델 설계시 발생하는 오차가 음상정위에 미치는 영향을 최소화하였다.

시뮬레이션과 청취 실험 결과, 제안한 방법으로 설계된 10/10 tap + 13/13 tap IIR HRTF 필터가 약 9 kHz 이하의 대역에서 보다 나은 특성을 갖도록 설계됨으로 인해 우수한 특성을 보였으며, 특히 머리 뒤쪽에 음상정위가 맺히는 경우에 다른 설계 방법보다 우수한 특성을 보였다. 이것은 HRTF의 낮은 주파수 영역에서 앞뒤 관정에 많은 정보를 포함하고 있음을 의미한다.

HRTF 필터 설계시 가장 문제가 되는 것은 역시 필터의 차수 결정이라 할 수 있는데, 필터가 높은 차수를 가지는 경우 성능은 좋아지지만 결과적으로 계산량이 방대해지므로 입체음향 시스템의 실시간 구현이 어렵다[9].

제안한 방법에서 필터의 설계시 발생하는 많은 계산들은 HRTF가 방향에 따라 데이터 베이스화되어 저장되기 때문에, 이것은 일종의 HRTF 필터 진치리 과정으로 볼 수 있으므로 문제가 되지 않는다[11].

제안된 HRTF 설계방법에서 가장 중요한 것은 해상도를 향상시킬 주파수 대역의 설정이다. 모든 방향에서 일정한 주파수 대역의 해상도를 향상시키기 보다는 각각의 방향에서 음상정위에 보다 많은 정보를 포함하고 있는 적절한 주파수 대역을 찾아 이 부분의 해상도를 높이는 방법이 보다 이상적일 것이므로, 사람의 청각특성에 대한 보다 많은 연구가 이루어 진다면 보다 양호한 성능을 갖도록 필터를 설계할 수 있을 것이다.

앞으로 입체 음상 제어 시스템의 완벽한 구현을 위해서는 방향감 이외에 거리감과 공간감을 제어할 수 있도록 하여야 할 것이며, 헤드폰 청취를 위한 바이노럴 프로세서 이외에, 두 개의 스피커 청취를 위한 트랜스오럴 시스템 구현도 이루어져야 할 것이다.

참 고 문 헌

- [1] 강성훈, *입체음향*, 기전연구소, ch. 1-3, 1997.
- [2] D. Begault, *3-D Sound for Virtual Reality and Multimedia*, Academic Press, ch. 2-4, 1994.

- [3] J. Huopaniemi and M. Karjalainen, "Review of digital filter design and implementation methods for 3 - D sound," *102nd Audio Engineering Society (AES) Convention*, preprint no. 4461, Munich, Germany, March 1997.
- [4] D.Kistler and F. 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] Bill Gardner and Keith Martin, "HRTF Measurements of a KEMAR Dummy - Head Microphone," *MIT Media Lab Perceptual Computing - Technical Report #280*, May 1994.
- [6] J-M. Jot, V. Larcher, and O. Warusfel, "Digital signal processing issues in the context of binaural and transaural stereophony," in *Proc. 98th Audio Engineering Society Convention*, Paris, France, Feb. 1995.
- [7] Bartłomiej Beliczynski, "Approximation of FIR by IIR Digital Filters: An Algorithm Based on Balanced Model Reduction," *IEEE Trans. Signal processing*, vol. 40, no. 3, pp. 532 - 541, Mar. 1992.
- [8] Jonathan Mackenzie, Jyri Huopaniemi and Vesa Välimäki, "Low-order modeling of head-related transfer functions using balanced model truncation," *IEEE Signal processing letters*, vol. 4, no. 2, pp. 39-41, Feb. 1997.
- [9] 김동현, 김기만, "WFIR 구조를 이용한 바이노럴 필터 설계", 한국음향학회 학술발표대회, 제 17 권, 1(s)호, pp. 193-196, July 1998.
- [10] 김동현, 김기만, "BMT 를 이용한 HRTF 필터 설계", 한국해양정보통신학회 추계종합학술대회, pp. 424-427, Nov. 1998.
- [11] 김동현, 김기만, "머리전달함수의 저차 모델링에 관한 연구", 한국음향학회 학술발표대회, 제 17 권, 2(s)호, pp. 35-38, Nov. 1998.