

3차원 입체음향을 위한 머리전달함수의 IIR 필터 구현

안현진, 김현태, 박장식, 손경식
부산대학교 전자공학과, 동의대학교 멀티미디어공학과, 동의공업대학 영상정보과

IIR Filter Implementation of HRTF to Synthesize 3D Sound

Hyun Jin Ahn, Hyun Tae Kim, Jang Sik Park and Kyung Sik Son
Dept. of Electronics Eng., Pusan National University

lovedresuc@pusan.ac.kr

Dept. of Multimedia Eng., Dongeui University

htaekim@dongeui.ac.kr

Dept. of Visual Technology, Dongeui Institute of Technology

jsipark@dit.ac.kr

요약

입체음향을 구현하는 방법에는 5 혹은 6 개의 스피커를 사용하는 5.1 채널 음향재생 시스템과 2 개의 스피커를 사용하는 방법 있다. 2 채널 입체음향재생 시스템은 라우드 스피커 재생보다 헤드폰 재생에 사용되고 있다. 2 채널로 입체음향을 재생하기 위하여 사람의 머리와 귀바퀴를 FIR 필터로 모델링한 머리전달함수(HRTF, head related transfer function)을 이용한다. FIR 필터로 측정된 HRTF는 실시간 처리에 있어서 계산량이 많은 단점이 있다. 본 논문에서는 2 채널 입체음향시스템을 효과적으로 구현하기 위하여 FIR HRTF를 IIR필터로 구현하는 방법을 제안한다. FIR 필터계수로부터 IIR 필터계수를 구하는 알고리즘은 균형화 모델 감소법(balanced model reduction)을 이용한다. 본 논문에서는 IIR 필터로 차수를 줄인 HRTF로 입체음향을 재생하는 시뮬레이션을 하였다.

1. 서론

단일 채널음향 재생에서 시작한 음성 및 음향 분야는 2채널의 스테레오 음향 재생에서 벗어나 현장감 있는 음향을 제공하기 위하여 3 차원 입체음향 시스템으로 발전하고 있다. 입체음향은 정위감과 거리감 그리고 공간감을 지각할 수 있도록 공간 정보가 부가된 음향이다. 사람은 두 귀로 전후, 좌우 그리고 상하 방향의 소리의 위치를 파악할 수 있다.

최근 널리 활용되고 있는 입체음향 시스템은 5.1 채널 방식이다. 5.1 채널은 전방에 3 개의 스피커, 후방의 좌우에 2 개 그리고 저음을 재생하기 위하여 1 개의 저음재생 스피커(woofer speaker)를 사용한다. 그러나

5.1채널은 설치하는 것이 불편하고 비용이 발생할 뿐만 아니라 헤드폰으로는 입체음향을 재생하는데 한계

가 있다. 2 채널로 입체음향을 재생하는 방법로 사람의 머리와 귀바퀴를 모델링하여 사람이 두 귀로 방향과 거리를 인지하는 원리와 같은 방법으로 입체음향을 재생하는 것이 HRTF를 이용한 입체음향 재생방법이 있다. 2 채널로 정위감과 공간감을 제공하기 위해서는 HRTF(head-related transfer fuction)로 원음의 방향을 제어하고 RTF(room transfer fuction)로 공간감을 제어한다. 2 채널 입체음향 재생은 5.1 채널 방식에 비하여 설치하는 것이 용이하고 헤드폰으로도 입체음향을 재생할 수 있지만 라우드 스피커로 재생할 때는 크로스토크(cross-talk)를 제거하여야 하고 고차의 FIR 필터로 측정된 HRTF 계산처리를 하여야 한다.

본 논문에서는 2 채널 입체음향시스템을 효율적으로

구현하기 위하여 FIR 필터 구현된 HRTF를 IIR 필터로 변환하여 계산량을 줄이고자 한다. 본 논문에서는 FIR 필터로 구현된 HRTF를 IIR필터로 변환하기 위하여 균형화모델 감소법(balanced model reduction)을 이용하였다. 시뮬레이션을 통하여 균형화모델 감소법으로 512 차의 FIR 필터로 된 HRTF를 32 차, 16 차로 줄였을 때 오차가 거의 없음을 확인할 수 있다.

2. 머리전달 함수

입체 음향은 영화에서는 서라운드(Surround)라는 방식으로 제공된다. 그리고 최근 TV 혹은 음향 시스템에 표준으로 제공되고 있는 5.1채널 방식은 시청자의 전방 정면, 오른쪽, 왼쪽으로 3채널, 시청자의 후방 오른쪽, 왼쪽으로 2채널, 그리고 부족한 저음을 보강하는 LFE(Low Frequency Effect)채널로 구성된다. 3차원 입체음향을 재생하는데 있어서 서라운드의 채널을 증가시킴으로써 설치 비용이 증가하고 시스템 요구 사항의 복잡성을 해결하고 헤드폰 혹은 이어폰을 착용한 상태에서도 가상음원이 적용할 수 있도록 HRTF를 적용하여 3차원 입체음향을 재생한다. HRRF는 자유공간에서 정위된 음원으로부터 사람의 귀로 전달되는 음향적 과정을 나타낸다. 음원의 위치를 파악하는 중요한 정보를 포함하는 임펄스응답(impulse response)이다.

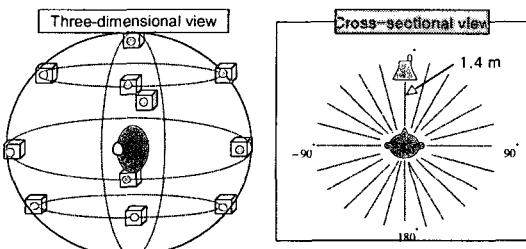


그림 1. HRTF의 측정 방법

측정된 HRTF는 인간의 청각 기관을 모델링한 더미 헤더 마이크를 이용하여 360° 의 방위각(azimuth angle)과 180° 의 고도각(elevation angle) 사이를 특정한 간격을 두고 각각의 위치에 대한 좌우 양쪽의 임펄스 응답을 구한 것이다. HRTF의 모델링에 대한 많은 연구가 진행되었으며 특히 MIT Media Lab에서 KERMA 더미 헤드(dummy head)를 이용하여 그 주위의 710 곳의 HRTF를 측정한 것을 널리 사용하고

있으며 최근 우리나라에서도 우리나라 사람들의 체형에 맞는 HRTF를 측정하기도 하였다.

우리가 관심을 가지는 2 채널 입체음향 재생기술은 크게 두가지 핵심 기술로 나눌 수 있다. 하나는 음이 들려오는 위치를 인위적으로 조정하는 음상정위기술이고 또 하나는 청취하는 공간의 크기 및 환경을 조정하는 음장 제어 기술이다.

음상 정위 기술은 실측된 HRTF를 이용하여 방향성이 없는 모노음에 디지털 신호처리 과정을 거쳐 방향, 거리에 관한 정보를 더해주는 것이며, 음장제어는 반사음과 잔향을 인공적으로 제어하여 특정 실내에 음원이 있는 것과 같은 음향효과를 생성하는 기술이다.

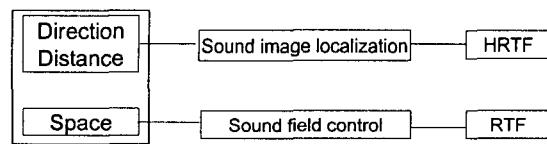


그림 2. 입체음향을 구현하기 위한 기술과 그에 대응한 함수



그림 3. 입체음향을 구현하는 과정

청취자가 음원에 대한 공간적 단서를 지각할 수 있는 것은 머리 전달계의 고유특성에 의하여 두, 귀에 입사한 두 신호간의 차이가 발생하기 때문이다. 두 신호차에 대한 특성은 머리 전달함수에 내포되어 있기 때문에 이를 이용하면 입체화되지 않은 단순음에 공간적 정보가 부과된 바이노럴 타입의 입체음을 생성할 수 있다.

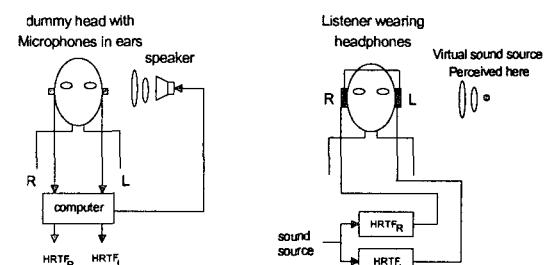


그림 4. 헤드폰을 이용한 입체음향의 재생

3. 머리전달 함수의 IIR 필터 구현

FIR필터로 구현되어 있는 HRTF를 균형화모델 감소법으로 IIR 필터로 구현한다. 균형화모델 감소법은 전달함수(transfer function)를 상태공간형태(state space form)로 변환시키고 Hankel singular value를 구한다. 그리고 Hankel singular value로부터 에너지가 작은 값을 소거하고 전달함수로 변환하여 IIR 필터 계수를 구한다. 표 1은 전달함수를 상태공간형태로 변환하는 과정을 나타낸다.

표 1. 전달함수의 상태공간형태로의 변환

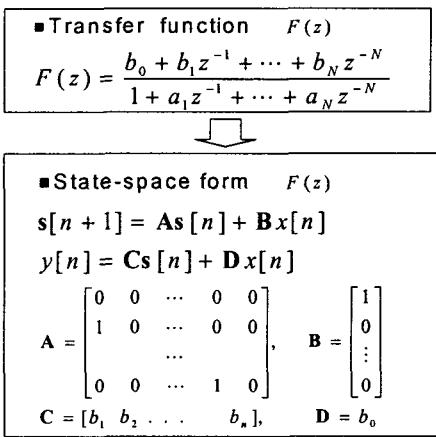
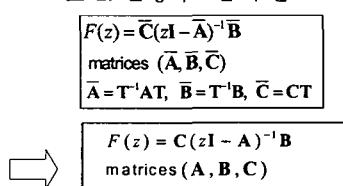


표 1에서 상태공간형태 변환된 상태방정식에서 식 (1)과 같이 Hankel 행렬을 구성한다.

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

그리고 상태방정식과 출력방정식의 행렬과 벡터를 이용하여 표 2와 같이 균형화 모델을 구현한다.

표 2. 균형화모델 구현



$$T = V\Lambda^{-1/2}$$

$$\bar{A} = \Lambda^{1/2}V^TAV\Lambda^{-1/2}$$

$$\bar{B} = \Lambda^{1/2}V^TB$$

$$\bar{C} = [CV\Lambda^{-1/2} \quad CV\Lambda^{-1/2}]$$

아래 식(2)와 같이 Hankel 행렬을 decompose를 한다.

$$H = V\Lambda V^T \quad (2)$$

Hankel singular value를 나열하고 그 값에 따라 필터계수를 결정하고 그에 대한 각각의 행렬 결정한다.

$$\begin{aligned} \Lambda &= \begin{bmatrix} \Lambda_k & 0 \\ 0 & \Lambda_{n-k} \end{bmatrix} \quad V = [V_k \quad V_{n-k}] \\ \bar{A} &= \begin{bmatrix} \Lambda_k^{1/2}V_k^TAV_k\Lambda_k^{-1/2} & \Lambda_k^{1/2}V_k^TAV_{n-k}\Lambda_k^{-1/2} \\ \Lambda_{n-k}^{1/2}V_{n-k}^TAV_k\Lambda_k^{-1/2} & \Lambda_{n-k}^{1/2}V_{n-k}^TAV_{n-k}\Lambda_{n-k}^{-1/2} \end{bmatrix} \\ \bar{B} &= \begin{bmatrix} \Lambda_k^{1/2}V_k^T B \\ \Lambda_{n-k}^{1/2}V_{n-k}^T B \end{bmatrix} \\ \bar{C} &= [CV_k\Lambda_k^{-1/2} \quad CV_{n-k}\Lambda_{n-k}^{-1/2}] \end{aligned}$$

위 상태방정식과 출력방정식에서 $\Lambda_k^{1/2}V_k^TAV_k\Lambda_k^{-1/2}$, $\Lambda_k^{1/2}V_k^TB$, $CV_k\Lambda_k^{-1/2}$ 항은 IIR 필터를 형성하는데 영향을 주고 식(3)과 같이 상태공간형태를 전달함수로 바꿔주는 것이다.

$$G_k = CV_k\Lambda^{-1/2}(zI - \Lambda_k^{1/2}V_k^TAV_k\Lambda_k^{-1/2})^{-1}\Lambda_k^{1/2}V_k^TB \quad (3)$$

4. simulation 결과

그림 5에서 알 수 있듯이 30 차 주위에서 그 값이 현저하게 줄어듬을 알 수 있다. 따라서 32차를 IIR 필터 차수로 선택할 수 있음을 보여준다.

그림 6은 FIR 필터 512 차 HRTF와 IIR 필터의 주파수 특성을 비교한 것이다. IIR 필터 32 차로 변환한 경우에는 저주파영역에서 다소 차이가 있지만 오차가 커지 않다. 16차, 8차, 4차로 차수를 줄일수록 오차가 크게 되었다. 에너지가 작은 Hankel singular value 값을 제거하기 때문이다.

그림 7은 변환한 IIR 필터 계수의 안정성을 분석하기 위하여 pole-zero 다이어그램이다. 극이 모두 단위원 안으로 들어오는 것으로 스템은 안정한 것을 확인할 수 있다.

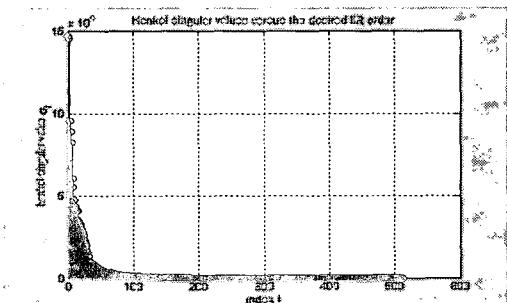


그림 5. Hankel singular value 나열

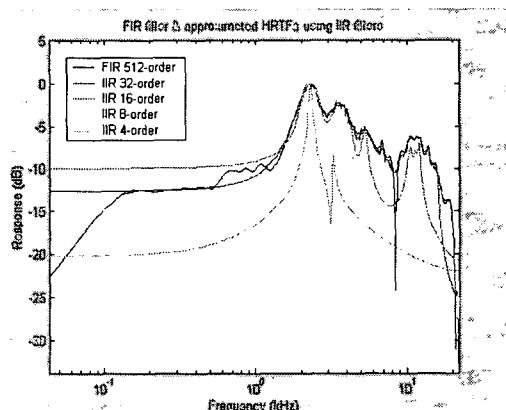


그림 6. FIR필터와 구현한 IIR 필터의 결과 비교

5. 결론

일반적으로 FIR필터는 일반적으로 계산량이 많다는 단점이 있기 때문에 실시간 처리하는데 다소 어려움이 있다. IIR 필터는 필터계수에 따라 불안정해질 수도 있지만 비슷한 성능을 갖는 FIR 필터에 대하여 계산량을 현저히 줄일 수 있다. 본 논문에서 임펄스 응답으로 측정된 HRTF, 즉 FIR 필터로 구현된 HRTF를 IIR 필터로 변환시켜 계산량을 줄임으로써 2채널 3차원 입체음향을 실시간으로 구현하고자 하였다. 본 논문은 균형화모델 감소기법을 이용하여 FIR필터를 IIR필터로 구현 할 수 있었다. 512차 FIR 필터로 된 HRTF를 32차로 줄인 경우가 가장 비슷한 결과를 얻을 수 있었고 차수를 줄여 갈수록 오차가 점점 커지

지만 16 차까지는 정위감을 제공할 수 있었다. 32 차인 경우도 저주파영역에서 주파수 특성이 차이가 있지만 정위감을 제공하는데 있어 저주파 성분은 회절을 하기 때문에 음성 정위에 크게 영향을 주지 않는다.

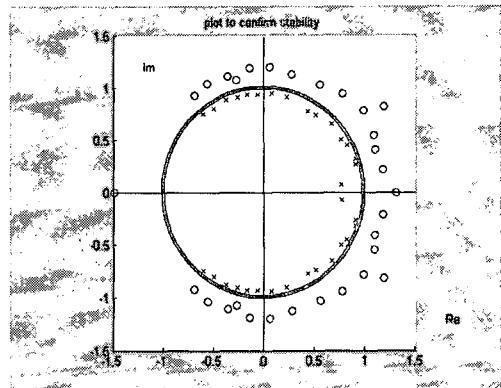


그림 7. stability를 검증하기 위한 그림

[참고문헌]

- [1] C. E. Schmid, "Design of IIR/FIR Filters Using Frequency Domain Bootstrapping Technique and LPC Method," IEEE Trans., ASSP-31, pp.28-34, 1983.
- [2] A. Bester and E. Zeheb, "Reduced Order IIR Approximation to FIR Digital Filters," IEEE Trans, ASSP-39, pp. 2540-2544, 1991
- [3] B. Beliczynski, I. Kale and G. E. Cain, "Approximation of FIR by IIR digital filter: an algorithm based balanced model reduction," IEEE Trans., ASSP-40, (3), pp.532-542, 1992
- [4] B. S. Chen, S. C. Peng and B. W. Chiou, "IIR Filter Design via Optimal Hankel Norm Approximation," IEE proc. G,139, (5), pp. 586-590, 1992.
- [5] R. W. Aldhaheri, "Model Order Reduction Via Real Schur Form Decomposition," Int.J. Control .53, pp.709-716, 1991.
- [6] K. V. Fernando and H. Nicholson, "On the Cross-gramian for Symmetric MIMO Systems." IEEE Trans, CAS-32, pp.487-489, 1985
- [7] B. Gardner and K. Martin, "HRTF Measurements of a KERMA Dummy-head Microphone," MIT Medea Lab. Perceptual Computing-Technical Report #280, pp.1-7 1994