

WFIR 구조를 이용한 바이노럴 필터 설계

김동현*, 김기만

한국해양대학교 전파공학과

Binaural Filter Design using Warped FIR Structure

Dong-Hyun, Kim, Ki-Man, Kim

Dept. of Radio Sci.&Eng., Korea Maritime Univ.

silhouet@chollian.net, kimkim@hanara.kmaritime.ac.kr

요약

지금까지 바이노럴 필터 설계 방법들의 대부분은 linear frequency scale 을 이용한 것이지만, 사람의 귀는 non-linear frequency scale 을 가지며 critical band 에 의한 청각정보로 인지한다. 따라서, 이와 같은 특징을 이용하여 좀 더 효율적으로 바이노럴 필터를 설계할 수 있다.

본 논문에서는 frequency warping 을 이용해 non-linear frequency resolution 을 갖는 바이노럴 필터를 제시한다. 또한, 종래의 설계방법에 의한 필터와 warped FIR 구조를 갖는 바이노럴 필터와의 비교정위를 통해 성능의 비교 평가를 수행 한다.

1. 서론

오디오의 역사 중에서 원 음장의 충실 전송(고충실도 및 입체 음향 전송)의 연구는 1920 년대에 시작되었고, 헤드폰에 의한 바이노럴 시스템 또는 3 채널 수음 재생에 의한 입체 음향 실험은 1930 년대로 그 역사를 올라갈 수 있다. 그 목적은 전송계의 충실, 다이내믹 레인지는 물론, 음의 방향 정위와 깊이감의 재현이다.

입체 음상 제어 알고리즘은 구성상 헤드폰 재생을 위한 바이노럴 프로세서와 2 개의 스피커 재생을 위한 트랜스오럴 프로세서 2 가지로 나뉘어진다. 이중 바이노럴 프로세서는 임의의 신호를 임의의 방향에 정위시키는 프로세서로서 마치 입력신호를 바이노럴 녹음으로 얻어진 것과 같은 신호로 변환한다.

auralization 과 바이노럴 기술은 음향 신호처리 분야에 서 매우 중요한 역할을 한다. 사람 또는 더미헤드의 머리 전달함수(HRTF) 측정과 모델은 auralization 과 spatial hearing 연구에 매우 중요한 정보를 제공한다.

가상 현실 분야에서 주목되고 있는 청각 디스플레이는 바이노럴 청취를 이용한 음상의 3 차원 정위가 가능한 시스템이다. 음원으로부터 두 귀의 고막까지의 전달 함수를 미리 몇 군데 위치 관계에 대해서 계측하여 저장 하여 두고, 보간 계산에 의하여 순시치를 구하여, 이것을 기본으로 고막상에서 음압을 실시간으로 계산하면, 실질적으로 존재하지 않는 음원을 근사적으로 생성할 수 있다. 이 시스템을 이용하면 헤드폰을 사용할 경우 3 차원적인 음장을 재현할 수 있다. 대표적인 시스템으로는 convolvotron 이 있다[1].

바이노럴 프로세서에서 해결해야 하는 한가지 문제는 특정 방향에 음상을 정위시키기 위해 프로세서에 많은 계산적 부하가 걸린다는 것이다. 이 문제를 해결하기 위해서는 계산의 효율을 높이는 것과 HRTF 의 적절한 모델을 만드는 것이 필요하다.

본 논문에서는 warped 구조를 갖는 FIR 필터를 이용해 바이노럴 필터를 구성하여 이러한 문제를 해결하려 하며, 기존의 FIR 구조를 가지는 바이노럴 필터가 dynamic interpolation 을 위해 매우 복잡한 계산을 요했던 문제점을 보완하고자 한다.

2. 머리전달함수(HRTF)의 특성

HRTF는 공간에서의 한 점에서부터 청취자의 외이까지의 자유공간 전달 함수로 구성된다. 이것은 HRTF가 음상 정위의 단서(localization cues), 두 귀에 입사하는 음의 시간 차이(interaural time delay), 그리고 레벨차이(interaural level difference)같은 사람의 spatial hearing 특성 해석에 필요한 모든 정보를 포함하고 있음을 의미한다. 따라서, HRTF를 이용해 직접적으로 FIR 필터를 구성하여 바이노럴 프로세서를 구현할 수 있다.

3. Warped Filter 설계

일반적으로 HRTF는 시간 영역에서 windowing 하고 minimum-phase reconstruction에 기초를 둔 FIR 필터로 만들어 졌다.

사람의 청각이 non-linear frequency resolution을 갖는 것은 HRTF의 modeling 또한 같은 방법으로 해석 할 수 있음을 의미한다.

HRTF를 non-linear frequency resolution으로 근사화하는 방법에는 두 가지가 있다. 하나는 높은 주파수보다 낮은 주파수에서 적합하도록 weighting function을 사용하는 것이고, 다른 하나는 non-linear frequency resolution을 갖도록 HRTF 필터를 설계하는 것인데, 이것이 frequency warping이라 불리는 것이다[2].

warping 방법으로 가장 많이 쓰이는 것은 bilinear conformal mapping을 이용하는 것이다. bilinear warping은 unit delay 대신 first order all-pass section을 삽입함으로써 만들 수 있다[2].

$$z^{-1} \Leftarrow D_1(z) = \frac{z^{-1} - \lambda}{1 - \lambda z^{-1}} \quad (1)$$

여기서 λ 는 warping 계수이다.

식(1)에서 critical band 개념을 근거로 한 적절한 λ 값의 선택을 통해 bilinear warping이 심리음향의 Bark scale에 알맞도록 할 수 있다. 표본화 주파수 f_s 에 의존하는 λ 값은 식(2)로부터 구할 수 있다[3].

$$\lambda = 1.0211 \left[\frac{2}{\pi} \tan^{-1} (0.076 f_s) \right]^2 - 0.19877 \quad (2)$$

식(2)를 통해 λ 는 표본화 주파수 48kHz에서 0.7313, 44.1kHz에서 0.7233의 값을 가진다.

식(1)과 (2)로부터 필터의 frequency warping이 이러한 간단한 치환에 의해 구현될 수 있음을 알 수 있다. 그림 1에 각기 다른 λ 를 가진 first-order allpass 구조를 이용한 warping의 효과를 나타내었다[4].

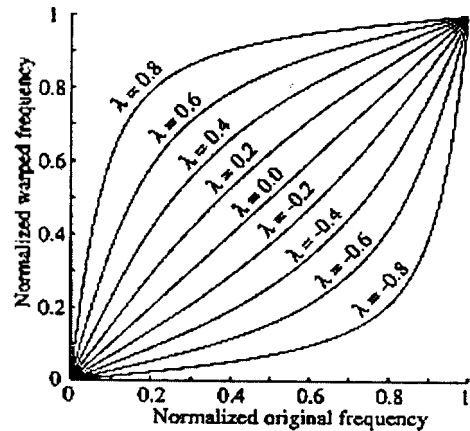


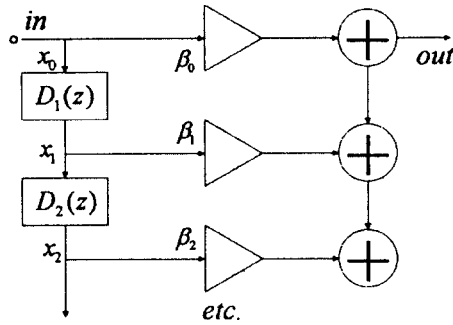
그림 1. Warping parameter λ 의 값에 따른 frequency warping 특성.

그림 2는 frequency warping에 의한 warped 필터의 FIR 구조를 나타낸다[2][4]. 여기서 그림 2 (a)의 D_1 은 unit delay 대신 삽입한 first order all-pass section이며, 그림 2 (b)의 β_i 는 필터 계수이다.

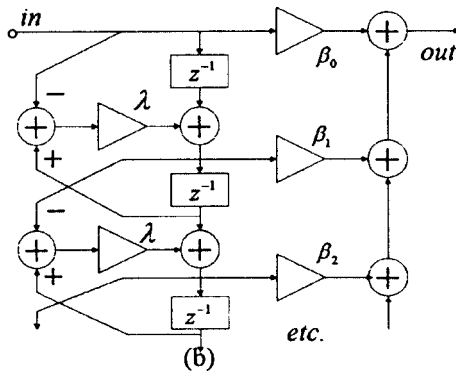
$$B_w(z) = \sum_{i=0}^M \beta_i [D_1(z)]^i \quad (3)$$

여기서 M 은 필터 차수이다[5].

그림과 식들에서 알 수 있듯이 warped FIR 필터는 엄밀히 말해 $z = \lambda$ 에서 M 개의 pole을 갖는 재귀형 구조이다.



(a)



(b)

그림 2. FIR 구조를 가진 warped 필터, (a) allpass delay 를 가진 warped 필터, (b) WFIR 필터의 실질적인 표현

4. Frequency Warping 의 유용성

warped 구조의 필터가 일반적인 구조의 필터와 다른 점은 필터내에 allpass section 이 있다는 것이다. 이로 인해 HRTF 를 직접적으로 FIR 필터의 계수로 사용한 것과는 달리 음상 정위가 약간의 유동성을 갖게 되는데, 이것은 HRTF 의 dynamic interpolation 을 가능하게 한다.

[5]

천천히 이동하는 음원을 재생하게 되면 음상 정위가 HRTF 의 측정 위치에서만 형성되는 부자연스러운 음상을 형성하게 되는데, HRTF 를 직접적으로 FIR 필터의 계수로 사용한 필터에서는 이것을 해결하기 위해서 복잡한 계산을 요했으나, warped 구조의 필터에서는 음상

정위의 유동성으로 인해 비교적 쉽게 이를 구현할 수 있다.

5. 실험

본 실험에서 warped 구조를 갖는 FIR 필터 바이노럴 프로세서와 일반적 구조의 FIR 필터 바이노럴 프로세서 간의 비교 청취를 수행하였다. 사용된 HRTF 는 채널당 128 sample 수를 갖는 MIT Media Lab.의 compact HRTF 이다.

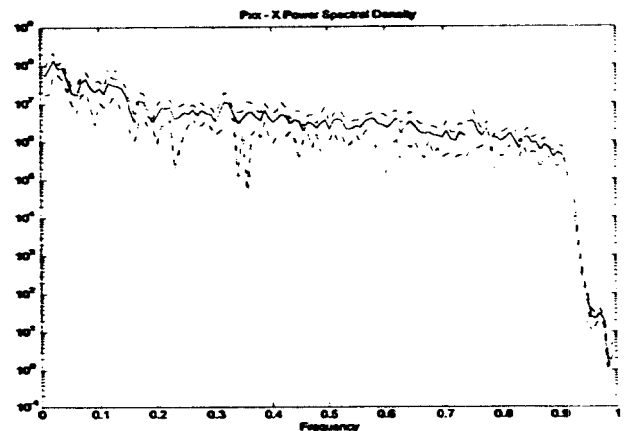
각각의 테스트 음은 0.5 초의 간격을 가지고 세번씩 반복하여 제시된다. 원음으로 사용되는 것은 다음과 같다.

- pink noise (20~20KHz, -12dB)
- 남자와 여자의 목소리 (무향실 녹음, lake dsp demo CD)
- 음악 (Guitar, 무향실 녹음, lake dsp demo CD)

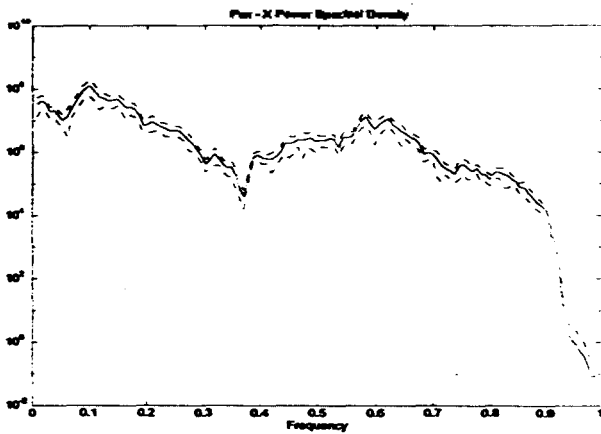
각각의 원음은 wav 파일로 저장되고, 각각의 필터에 대해 신호처리 된 다음 결과를 다시 wav 파일로 저장하였다.

그림 3은 azimuth 0°, elevation 0° 의 HRTF 를 사용한 필터 결과의 스펙트럼이다. 사용된 원음은 pink noise 이며 44.1KHz 로 표본화 되었다. 각각의 그림은 다음과 같다.

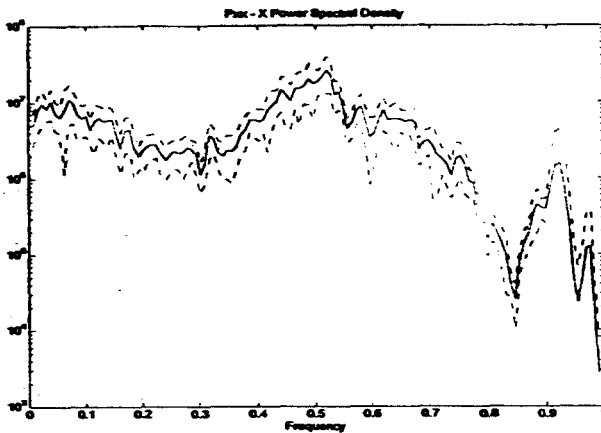
- 그림 3(a) : 실험에 사용된 원음
- 그림 3(b) : 128 tap 을 갖는 일반적인 FIR 필터
- 그림 3(c) : 128 tap 을 갖는 WFIR 필터
- 그림 3(d) : 64 tap 을 갖는 WFIR 필터



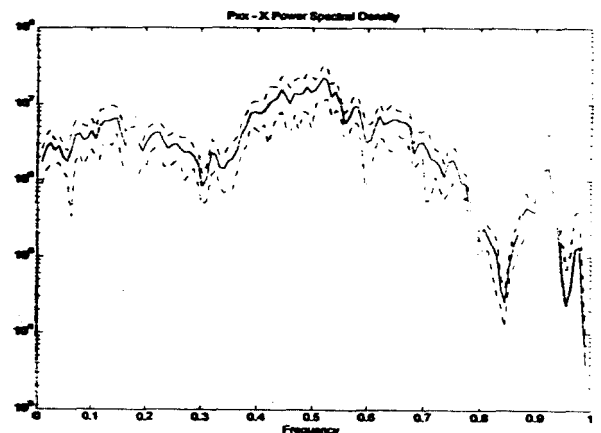
(a)



(b)



(c)



(d)

그림 3. azimuth 0°, elevation 0° 의 HRTF 를 사용한 필터 결과의 스펙트럼

6. 결론

본 논문에서는 frequency warping 을 이용해 non-linear frequency resolution 을 갖는 WFIR 바이노럴 필터를 제시 하였다. 또한, 기존의 FIR 구조의 바이노럴 필터와의 비교 청취를 통해 성능을 비교 평가 하였다. 실험 결과 WFIR 구조의 바이노럴 필터가 기존의 FIR 구조의 필터에 비해 표본화 주파수가 낮을수록, 필터의 차수가 낮을수록 더 나은 성능을 보였다. 앞으로 DSP processor 로 WFIR 필터를 구현하여, dynamic interpolation 에 대해 고찰해 나갈 계획이다.

참고문헌

- [1] 강성훈, 입체음향, 기전연구소, 1997.
- [2] 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 22-25, 1997.
- [3] J. O. Smith and J. S. Abel, "The Bark bilinear Transform", *Proc. IEEE ASSP Workshop*, Mohonk, Newpaltz, 1995.
- [4] H. W. Strube, "Linear prediction on warped frequency scale", *J. Acoust. Soc. Am.*, vol. 68, no. 4, pp. 1071-1076, 1980.
- [5] J. Huopaniemi and M. Karjalainen, "HRTF filter design based on auditory criteria", *Proc. Nordic Acoustical Meeting (NAM 96)*, Helsinki, Finland, pp. 323-330, June 12-14, 1996.