

공간감 향상을 위한 스테레오 신호의 바이노럴 렌더링

백용현, 박영철

연세대학교 MSP 연구실

(hyen0330@yonsei.ac.kr)

Rendering of Stereo Signal for the Enhancement of Spatial Envelopment

Baek, Yong-Hyun, Park, Young-Choel

Yonsei University, MSP Lab.

요약

본 논문에서는 이어폰 및 헤드폰 환경에서 스테레오 음원의 공간 정보를 보존하고 청취자의 엔벨롭감을 향상시키기 위한 Binaural 렌더링 기법을 제안하였다. 스테레오를 그대로 Binaural 렌더링 하지 않고 업믹스를 통해 스테레오 음원을 5.1채널의 음원을 생성한 뒤에 5.1포맷의 스피커 위치 기반에 의해 렌더링 시킨다. 제안된 알고리즘은 두 가지 단계로 나뉘는데 첫 번째 단계로 스테레오 음원을 주성분 분석법(PCA)을 통하여 primary와 ambient신호로 각각 분리 하고 두 번째 단계로 분리한 primary와 ambient신호를 스테레오의 공간 정보에 기반 하여 머리전달 함수(HRTF)를 이용하여 렌더링 한다. 또한 초기 잔향을 이용한 외재화 기법을 통하여 음원을 머리 밖에 정위시킴으로써 음원의 공간감을 향상시켰다. 제안된 방법은 주관적인 평가를 통하여 기존의 방법보다 우수한 성능을 보임을 확인하였다.

1. 서론

최근 스마트폰이 널리 보급되어 가정이나 개인용 PC에서 뿐만 아니라 모바일 환경에서 영화와 음악을 감상하는 사용자가 급격히 늘어났다. 이러한 모바일 휴대용 기기를 이용할 경우 일반적으로 스피커를 이용하는 것과는 다르게 이어폰(또는 헤드폰)을 이용하게 되어 청취환경이 달라지게 된다. 스피커 환경(스테레오 기준)에서는 자연스러운 청각 환경을 만들어주며 머리 밖에서 음원을 재생할 수 있으나 상호 채널 간섭에 의한 크로스 토크(cross talk)가 발생하고 음원을 청취자의 뒤쪽에 정위시키기 어렵다. 반면에 헤드폰 환경에서는 이러한 채널 간섭에 대한 영향이 없고 재생 공간의 음향 특성을 배제 할 수 있으나 음원이 머리 안쪽에 정위되는 단점이 있다.

본 논문에서는 스테레오 신호의 공간 정보를 보존하고 음상이 머릿속에 내재되는 문제를 해결하고 음원이 청취자를 에워 싸는 엔벨롭감을 향상시키기 위한 binaural 렌더링 알고리즘을 제안한다. 첫 번째로 음원의 정위 측면에서 볼 때, 대부분의 스테레오 콘텐츠들은 레코딩 방법에 따라 다를 수도 있으나 방향성이 있는 음원이 보통 두 스피커 간의 레벨 차에 의해 두 스피커 사이에 정위된다. 그러므로 스테레오 신호는 보통 페닝 기법에 의해 정위 되어 있다고 가정한다. 이것을 그대로 헤드폰으로 음원을 듣게 되는 경우 원래 스테레오 음원이 가지는 공간 정보가 달라지게 된다. 즉, 음상이 정위되는 위치가 달라 지게 된다. 두 번째로 음원이 머릿속에 내재되는 이유는 음원이 발생하는 위치에서부터 청취자의 귀 사이의 전달 경로가 사라지기 때문이며 이것을 머리 밖에 위치시키게 함으로써 공간감을 향상 시킬 수 있다[1]. 세 번째로 스테레오 스피커는 보통 사용자의 정면에 위치하게 되어 음원의 너비가 좁으며 제한적이다. 흔히 5.1 채널 이상의 스피커 포맷에서는 서라운드 스피커를 사용자의 후방에 위치 시킴으로써 음원이 청취자를 에워 싸는 엔벨롭감을 향상 시키며 이는 두 서라운드 채널의 상관

관계가 작을수록 더 효과가 크다[2]. 따라서 binaural 렌더링 시에도 Ambient 같은 신호를 뒤에 위치시켜주어야 한다.

앞에서 언급한 세 가지 목적을 달성하기 위하여 본 논문에서 제안한 알고리즘은 먼저 스테레오 신호를 주성분 분석법(PCA)를 이용하여 페닝 신호와 Ambient 신호로 분리한다. PCA를 통한 공간 정보를 기반으로 5.1채널로 업믹스를 하고 이를 HRTF 필터링을 통하여 5.1 채널 포맷의 위치에 정위 시킨다. 업믹스한 각 채널들의 초기 잔향을 더해줌으로써 음상을 외재화 시킨다. 제안된 알고리즘은 기존의 방법들에 비해 스테레오 신호의 공간 정보 유지, 공간감 및 엔벨롭감에 대한 주관적 평가에서 우수한 성능을 보임을 확인하였다.

본 논문의 구성은 다음과 같다. 2장에서 기존의 알고리즘들에 대해 설명하고 3장에서는 제안한 알고리즘을, 4장에서는 모의 실험 결과를 보이고 5장에서 본 논문의 결론을 맺는다.

2. 기존의 스테레오 신호의 Binaural 렌더링 알고리즘

기존의 음원 분리 기반의 Binaural 렌더링 알고리즘을 살펴보면 팬텀 소스 기반[3]의 기술은 아래와 같이 음원을 분리한다.

$$\begin{aligned} S &= \frac{X_l + X_r}{\sin(\gamma_b) + \cos(\gamma_b)} \\ A &= X_l - \sin(\gamma_b)S \end{aligned} \quad (1)$$

여기서 X_l, X_r 은 입력 스테레오 신호이고 γ_b 는 탄젠트 룰에 의해 구한 페닝 신호의 각도이며 b는 대역 인덱스이다. 스테레오 신호 분석은 단구간 푸리에 변환(STFT)을 통한 주파수 영역에서 이루어지나 수식의 간편화를 위해 주파수 인덱스는 생략하였다. 위의 분리 과정은 두 채널에서 하나의 Ambient만을 분리해 내고 렌더링시 단순히

$A_l = A, A_r = -A$ 로 양쪽 채널에 더하기 때문에 충분한 엔벨롭 감을 청취자에게 주지 못한다. 또한 이 논문에서는 입력 스테레오 신호 분석을 통해 패닝 각도를 찾아내면 그 방향으로 렌더링 한다. 그러나 업믹스에서는 스테레오 스피커 포맷에서 음원의 너비가 좁아지는 단점으로 인해 가운데 채널을 추가 하여 이러한 단점을 극복하기 때문에 최근에 단순히 패닝신호의 각도로 렌더링 하기보다는 업믹스 과정을 통해 전방 3채널을 렌더링 하는 연구가 진행 되었다. 따라서 업믹스 알고리즘 역시 Binaural 렌더링 성능에 영향을 주게 된다.

돌비 업믹스를 이용한 Binaural 렌더링[4]의 경우는 원래 스테레오 음원의 공간 정보는 배제한 채 단순히 두 채널을 서로 더하거나 뺌으로써 센터 채널과 후방 서라운드 채널을 생성하기 때문에 원래 스테레오 음원이 가지는 패닝 각도가 달라지게 된다. 또한 음원이 정확히 정면에 위치한 경우를 제외하고는 Ambient를 정확히 분리 하지 못하고 패닝신호가 Ambient로 빠져나오게 되기 때문에 전방에 위치해야 할 패닝 성분이 청취자의 뒤쪽에 위치하는 문제가 발생한다.

위의 문제점들을 살펴보면 스테레오 입력의 공간정보를 유지하고 정확하게 Ambient를 분리해내어 엔벨롭감을 높여주어야 한다.

3. 제안된 Binaural 렌더링 알고리즘

3-1 스테레오 to 5.1 채널 업믹스

엔벨롭감을 항상 시키기 위해서는 스테레오 신호를 5.1 채널의 다채널 음원으로 업믹스 하고 특히 음원 분리 과정에서 스테레오 각 채널의 Ambient의 상관 관계가 작아야 한다. 본 논문에서는 고유치 비 기반의 패닝신호와 Ambient신호를 분리한다[5].

$$S = \sqrt{\frac{\lambda_1 - \lambda_2}{\lambda_1}} (X_l + X_r) \tag{2}$$

$$A_l = X_l - \left(1 - \sqrt{\frac{\lambda_2}{\lambda_1}}\right) a_l (X_l + X_r)$$

$$A_r = X_r - \left(1 - \sqrt{\frac{\lambda_2}{\lambda_1}}\right) a_r (X_l + X_r)$$

여기서 λ_1, λ_2 는 입력 스테레오 신호로부터 주성분 분석법을 통한 공분산 행렬의 고유치이고 a_l, a_r 는 고유치로부터 구한 패닝 계인이다. 고유치 비를 이용한 음원 분리 방법은 입력 스테레오 신호의 에너지를 보존할 수 있고 스테레오 각 채널에서 상관도가 거의 없는 Ambient를 분리해 낼 수 있다. 또한 각 고유치는 primary 신호와 ambient신호의 에너지와 대응되기 때문에 두 고유치의 비는 primary와 ambient간의 에너지의 비가 된다. 따라서 패닝 신호의 패닝 각도와 무관하게 분리 성능이 강건하게 유지된다.

또한 스테레오 음원이 가지는 공간 정보를 유지하기 위하여 분리해낸 패닝 신호는 패닝 계인으로부터 탄젠트 법칙에 의해 패닝 각도를 구해내고 이것을 기반으로 전방 3개의 채널에 대해 벡터 기반의 패닝 기법(VBAP)을 통해 다시 패닝 한다. 그림 1은 패닝 각도에 따른 3채널의 계인을 그린 그림이다.

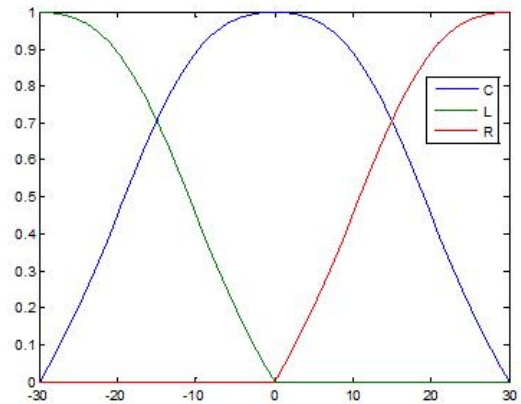


그림 1 Pair-wise VBAP를 이용한 3채널 패닝 계인

3-2 Binaural 렌더링

Binaural 렌더링은 HRTF 필터링을 통해 이루어진다. 3-1에서 업믹스 된 5채널을 각각 ITU 5.1 채널 포맷에 기반 하여 각 스피커 위치에 해당하는 방향으로 필터링 한다. 필터링은 STFT를 통해 주파수 영역에서 수행되며 HRTF 셋은 MIT 데이터베이스 셋을 이용하였다. 그러나 HRTF로 렌더링 할 경우 음색의 변하게 되는 단점이 있고 0dB가 넘는 주파수 응답을 가지는 경우가 있기 때문에 필터링 뒤 레벨도 바뀐다. 5.1채널을 다운 믹스 하는 과정이기 때문에 필터링 하기전의 에너지를 유지해 주어야 한다. 이것을 해결하기 위하여 HRTF를 정규화 하였다. 정규화는 아래 식을 통해 이루어진다.

$$\begin{aligned} H_l &= H_l / \max(|H_l|, |H_r|) \\ H_r &= H_r / \max(|H_l|, |H_r|) \end{aligned} \tag{3}$$

3-3 외재화

Binaural 렌더링 시 앞서 언급하였듯이 음원이 발생하는 지점에서부터 청취자의 고막까지의 전달경로가 사라지기 때문에 음상이 머릿속에 맺히는 음원의 내재화 문제가 발생하게 된다. 음원이 머릿속에서 들리게 되면 현실감과 공감감이 떨어지는 요인이 되기 때문에 내재되는 음원을 머리 밖으로 외재화 시켜야 된다.

외재화에 관한 지금까지 연구에 따르면 외재화를 위해서 초기잔향이 큰 영향을 주며 초기잔향만을 더해 주어도 충분히 외재화 효과를 얻을 수 있다. 또한 초기잔향이 반사되어 들어오는 각도가 음원과 좌우 90°일 때 가장 큰 외재화 효과를 얻을 수 있다. 초기 잔향은 개수는 실험을 통하여 2개만으로도 충분히 외재화 효과를 낼수 있고 그 이상 초기 잔향을 추가한다 하더라도 주관적으로 큰 차이를 느낄 수 없었다.

그림 3은 같은 사실을 바탕으로 한 외재화 알고리즘의 블록도를 나타낸다. 두 개의 초기 잔향만을 생성하여 좌우 각각 음원 각도의 ±90° 방향으로 렌더링 하여 각 채널에 더해준다. 또한 음이 반사되어 들어오는 실제 환경을 모델링하기 위해 흡음 필터를 두어 모델링한 잔향을 사용하였다.

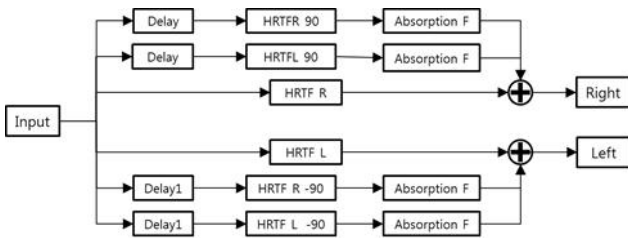


그림 2 외제화 알고리즘 블록도

4. 모의 실험

모의 실험은 돌비 업믹스, 팬텀 소스 알고리즘 그리고 제안된 알고리즘을 음원 정위의 정확도, 엔벨롭감을 주관적 평가를 통해 성능을 측정하였다. 청취 실험에 참여한 청취자의 수는 10명이고 각 청취자에게 먼저 임의의 방향으로 정위시킨 패닝신호를 스테레오 환경에서 들려주고 Binaural 렌더링 된 음원을 들려주는 비교 실험을 통해 진행되었다. 그림 3은 각 알고리즘의 음원 정위 성능을 비교한 그림이다. y축은 점수를 나타내며 -3~3점까지 청취자에게 측정하도록 하였다. 스테레오 패닝 각도를 기준으로 0점에 가까울수록 스테레오와 binaural 렌더링된 음원의 정위가 같음을 의미한다.

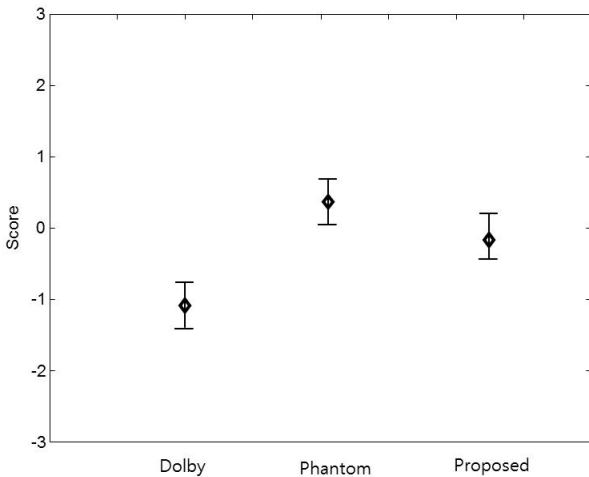


그림 3 음상 정위 주관적 평가 결과

그림 3에서 알 수 있듯이 공간 정보를 이용한 팬텀 소스 및 제안된 알고리즘은 정위에 유사한 성능을 보이나 돌비 기반의 알고리즘의 경우 공간 정보를 보존하지 못하기 때문에 낮은 점수를 기록하였다.

두 번째 실험으로 청취자의 엔벨롭감을 측정하였다. 여기서는 점수가 높을수록 청취자가 느끼는 엔벨롭감이 더 크다는 것을 의미하며 그림 4에서 알 수 있듯이 제안된 알고리즘의 우수한 Ambient 분리 성능으로 인하여 다른 알고리즘 보다 높은 점수를 보임을 알 수 있다.

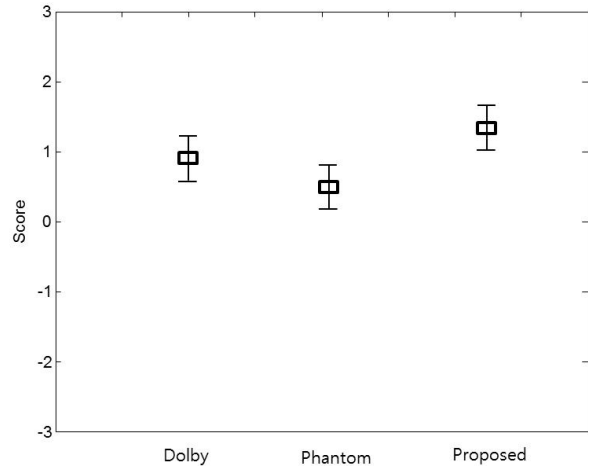


그림 4 엔벨롭감 주관적 평가 결과

6. 결론

본 논문에서는 스테레오 음원의 Binaural 렌더링 알고리즘을 제안하였다. 공간감을 향상시키기 위해 스테레오 음원을 먼저 업믹스 하여 5.1음원을 생성한 뒤 스피커 위치 기반에 따라 렌더링 하였다. 업믹스 시 고유치 비 기반의 음원 분리 알고리즘을 사용하여 두 채널 사이의 상관도가 낮은 Ambient를 분리 해 냄으로써 음원의 엔벨롭감을 높이고 패닝 각도 추정을 통해 입력 신호의 공간 정보를 유지하였다. 또한 Binaural 렌더링 시 문제가 되는 음원의 내재화 문제를 해결함으로써 렌더링 음원의 공간감을 더욱 높였다. 제안된 알고리즘은 기존의 스테레오 binaural 렌더링 알고리즘에 비해 주관적 평가에서 우수한 성능을 보임을 확인하였다.

참고 문헌

- [1]. TS Choi, YC Park, DH Youn, "Efficient Out of Head Localization System for Mobile Applications" AES ,2006
- [2]. G. A. Soulodre, M. C. Lavoie, and S. G. Norcross, "Investigation of Listener Envelopment in Multichannel Surround Systems," AES, 2002
- [3]. J. Breebaart and E. Schuijers. "Phantom materialization: A novel method to enhance stereo audio reproduction on headphones. IEEE Trans. 2008.
- [4].Christos Tsakostas , Andreas Floros, Yiannis Deliyiannis, "Binaural Rendering for Enhanced 3D Audio Perception",In Proceedings of the AudioMostly, 2007
- [5]. 백용현, 이근상, 전세운, 이석필, 박영철, "오디오 업믹스를 위한 효율적인 Primary-Ambient 분리 알고리즘", 방송 공학회 하계 학술 대회, 2012