

방향 정위된 음원에 시간지연을 이용한 확산감 제어에 관한 연구

방향 정위된 음원에 시간지연을 이용한 확산감 제어에 관한 연구

김 익 형, 정 의 필
울산대학교

전화 : 052-259-1650 / 핸드폰 : 017-705-1986

Sound Diffusion Control for the Localized Sound Image Using Time Delay

Ighyeong Kim and Uipil Chong

School of Computer Engineering and Information Technology,
University of Ulsan

E-mail : upchong@uou.ulsan.ac.kr

Abstract

Many researchers have developed the techniques of an efficient 3-D sound system based on the psycho-acoustics of spatial hearing with multimedia or virtual reality. In this paper, we propose an idea for the improved 3-D sound system using conventional stereo headphones to obtain a better sound diffusion from the mono-sound recorded at an anechoic chamber. We use the HRTF (Head Related Transfer Function) for the sound localization and the wavelet filter bank with time delay for the sound diffusion. We investigate the effects of the 3-D sound depending on the length of time delay at lowest frequency band. Also the correlation coefficient of the signals between the left channel and the right channel is measured to identify the sound diffusion.

I. 서론

최근의 입체음향 시스템은 가전제품이나 멀티미디어 제작도구, 그리고 게임 산업에 이르기까지 다양한 방면으로 발전하고 있으며, 또한 지속적인 연구가 이루어

어지고 있다. 이러한 연구들은 인간이 개선된 양질의 음악을 듣고자 하는 욕구에 의한 것이라 볼 수 있겠다. 컴퓨터 환경체제의 사용자들은 더 이상 전통적인 2 채널 스테레오 방식에 만족하지 않는다. 그리고 4 채널 서라운드 시스템이 인기를 얻고 있긴 하더라도 스피커의 개수와 케이블 양이 증가하기 때문에 사용자들은 추가적인 비용을 감수해야 한다. 이러한 문제를 해결하기 위한 방법으로 본 논문에서는 2 채널 헤드폰을 이용한 개선된 3D Sound 시스템을 제안하려고 한다.

임의의 한 방향에 위치한 음원에서 발생된 소리는 음원의 방향에 따라 일정한 시간차이와 레벨차이를 가지고 청취자의 두 귀에 도달하게 된다. 그리고 소리는 머리, 몸통, 외이 등에 의해서 회절, 반사, 굴절되어 두 귀에 도달한다. 이와 같이 두 귀에 도달하는 소리의 정보를 머리전달함수라고 하며[1], 컨벌루션 처리를 통하여 머리전달함수가 포함하고 있는 방향 정보를 모노음향에 합성할 수 있다. 인간이 소리를 듣는 것만으로 음원의 위치 및 방향 등과 같은 3차원 정보를 인지할 수 있는 이유는 음원으로부터 발생한 소리가 두 귀에 도달하는데 걸리는 시간의 차이(ITD : Interaural Time Difference)와 두 귀에 도달한 소리의 레벨 차이

(IID : Interaural Intensity Difference), 그리고 소리가 공기 중을 진행 할 때의 전파 특성 등에 의한 것이다. 특히 입체 음향을 구현하는데 있어서, ITD는 방위 (Azimuth) 변화에 대한 입체 음상의 단서를 제공하며, IID는 입체 음상의 효과를 더욱 증가시킬 수 있는 방법으로 활용할 수 있다[1]. 본 논문에서는 이들 중에서 IID가 1.5kHz 이상의 주파수 성분에서 반응하는 것을 착안하여 입체음향의 확산감을 제어할 수 있는 단서로 활용하여 웨이브렛 필터 बैं크를 이용해서 음의 확산감 제어를 위한 시스템을 구현한다. 또한 음의 확산 정도를 확인하기 위해서 좌·우 신호의 상관계수 값을 측정한다. 좌·우 신호의 상관계수 값이 0(zero)에 가까울수록 확산감 제어에 좋은 영향을 준다[2].

본문에서는 모노 음을 머리전달함수(HRTF)를 통해서 2 채널 음원으로 생산하는 방법과 논문에서 제안하고자 하는 웨이브렛을 이용한 음의 확산감 제어를 위한 알고리즘을 설명한다. 다음으로 본 연구의 실험 결과를 보여주고, 마지막으로 본 논문의 최종 결과와 향후 연구 방향에 대해서 기술한다

II. 본론

2.1 머리전달함수를 통한 음원 재생

머리전달함수(HRTF)는 자유 음장에서 어느 특정한 위치에 있는 음원으로부터 청취자의 고막까지 음원의 공간 전달을 행하는 전달함수를 의미한다. 본 연구에서는 무향실에서 녹음한 단음과 연주 음을 모노 음원으로 채취해서 머리전달함수(HRTF : Head Related Transfer Function)와 컨볼루션을 통하여 2 채널의 방향 정위(Localization)된 실험 음원을 획득한다.

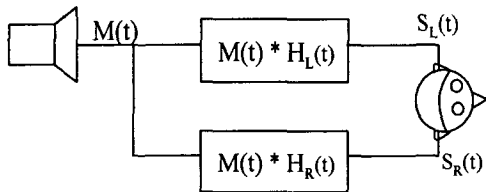


그림 1. 스테레오 음원 획득을 위한 시스템 블록도

그림 1은 실험 음원으로 사용할 스테레오 음을 생산하기 위한 시스템 블록 다이어그램을 보여 준다. 스피커로부터의 모노 음, $M(t)$ 는 오른쪽 및 왼쪽 머리 전달함수($H_R(t)$, $H_L(t)$)와 컨볼루션되어 방향 정위된 원

쪽 음원($S_L(t)$)과 오른쪽 음원($S_R(t)$)을 얻게 된다.

그림 2는 음원으로부터 청취자에게 음이 전달되는 과정을 보여주고 있다. 본 연구는 음상을 정위시키기 위한 방법으로 MIT Media Lab.의 KEMAR 데이터를 이용한 머리전달함수(HRTF)를 사용하고[3], 그 정위된 음상의 확산감을 좋게 하기 위한 방법을 설명한다.

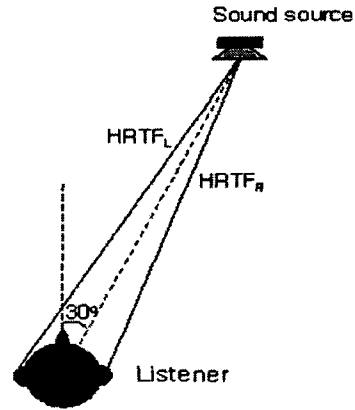


그림 2. 음원에서부터 청취자에게로 음의 전달

2.2 제안한 시스템의 구현

모노 채널의 헤드폰이나 스피커를 통해서 음을 청취할 경우, 음의 파라미터를 조정할 수 있는 방법이 없다. 따라서 머리전달함수와 컨볼루션을 통해 얻어진 2 채널 음을 사용하게 된다. 얻어진 음원에 확산감을 부여하기 위한 방법으로는 웨이브렛 필터 बैं크를 이용한다. 시스템의 구현은 가장 낮은 주파수 대역에 Time Delay를 부여함으로써 음의 확산 정도를 제어한다. 서론에서도 밝힌 바와 같이 머리전달함수에서 음의 정위를 위한 가장 큰 단서는 두 귀 사이의 시간차(ITD)와 레벨차(IID)를 들 수 있다[4].

본 논문에서는 위의 두 가지 단서 가운데에 1.5kHz 이상의 주파수 영역에서 두 귀 사이의 레벨차(IID) 성분에 영향이 주어짐을 착안하여 음의 확산 정도를 조정하려는 것이다[5].

결국 본 연구의 의도는 IID 성분을 조작함으로써 모노 음원으로부터 만든 2 채널의 음을 스테레오 헤드폰을 통해서 청취할 수 있는 더욱 개선된 3D 사운드 시스템을 구현하는 것이다. 무향실에서 녹음된 모노 음(sound source)을 그림 2에서 보는 것처럼 좌·우 머리 전달함수 각각에 적용시켜 원하는 실험 음원을 만든다. 실험에 이용되는 모든 음은 방위각(Azimuth)은 30° , 상승각(Elevation)은 0° , 음원과 청취자와의 거리

방향 정위된 음원에 시간지연을 이용한 확산감 제어에 관한 연구

(Distance)는 1.4m로 고정해서 사용한다. 그림 3은 제안한 시스템의 전체적인 블록 다이어그램을 나타낸다. HRTF(머리전달함수)를 적용한 좌·우 신호들은 웨이브렛 필터 बैं크를 통과하면서 신호의 주파수 대역이 44.1kHz에서 1.5kHz까지 분리가 가능하게 된다[6]. MATLAB software를 이용해서 웨이브렛 필터 बैं크를 구현하고, 음원 실험을 위해 이용한 필터로는 16개의 필터 탭을 가지는 Daubechies 필터를 사용하였다.

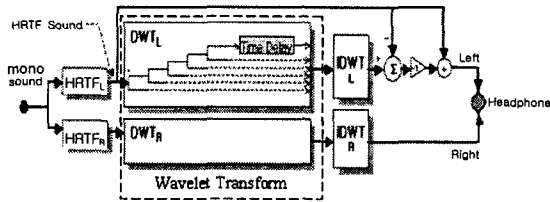


그림 3 제안한 시스템의 전체 블록도

신호의 가장 낮은 주파수 대역인 1.5kHz의 신호를 얻기 위해 5 단계의 웨이브렛 필터 बैं크를 구성한다. 웨이브렛 필터 बैं크의 마지막 단계에서 왼쪽 채널의 가장 낮은 주파수 대역에 약 7.25ms의 시간지연을 시킨다. 웨이브렛 변환을 통해 시간 지연시킨 음을 역 웨이브렛 변환을 통해서 다시 원래의 음으로 만든다. 청취 실험에서 역 웨이브렛 변환을 통해서 만들어진 음과 원래 머리전달함수(HRTF)만을 적용해서 만든 음 사이에 노이즈 신호가 발생하여, 이 노이즈 신호를 감소시키기 위한 방법으로 발생된 노이즈 신호의 위상을 역 위상으로 바꾸어 준다. 이는 본 연구에서 음의 확산 정도를 측정하는 방법으로 좌·우 신호의 상관계수 값을 이용하는데, 이 과정을 통해서 상관계수의 값을 좀 더 0(zero)에 가깝게 만든다[2]. 이렇게 생성된 신호를 HRTF sound 신호에 더해줌으로써 왼쪽 채널의 최종적인 음을 얻어낸다. 최종 음은 오른쪽 채널의 신호와 함께 2 채널 스테레오 헤드폰을 통해서 청취되어진다.

III. 실험 및 청취 평가 결과

3.1 HRTF 음과 최종 음의 상관계수 값 비교

표 1은 시간 지연에 따른 신호의 상관계수 값을 나타낸 것이다. 시간 지연이 길어질수록 음의 확산 정도를 평가하기 위한 상관계수의 값도 함께 증가함을 볼 수 있다. 또한 단순히 시간 지연만을 부과했을 때와 시간 지연 신호에서 발생한 노이즈 신호를 제어한 후에 만들어진 최종 음과의 상관계수 값을 비교했을 때 후자

의 상관계수 값이 훨씬 개선되었음을 확인할 수 있다.

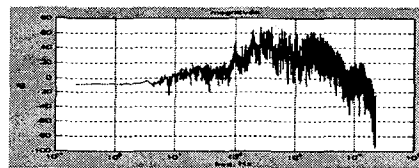
Time Delay	Correlation coefficient	
	Delay time Only	Final sound
7.25ms	0.1077	-0.0487
14.5ms	-0.1184	-0.0616
17.4ms	0.2284	0.1819

표 1. 시간 지연과 상관 계수와의 관계

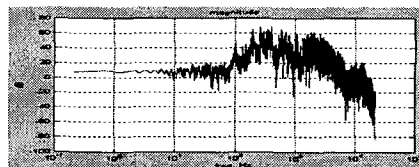
3.2 실험 및 청취 평가

1) 첼로(Cello) 연주 음을 통한 실험 과정

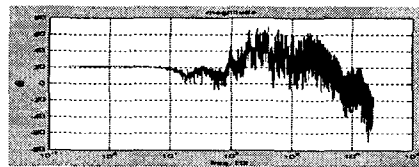
그림 4의 (가)와 (나)의 첼로 연주 음의 파형을 비교할 때 머리전달함수만을 적용하여 만든 음과 시간 지연을 부과했을 때의 파형이 시각적으로도 상당한 차이를 보이고 있음을 확인할 수 있다. (다)의 최종 음의 파형은 이런 신호의 왜곡을 수정한 결과이다.



(가) 머리전달함수만을 적용한 음의 파형



(나) 시간 지연을 포함한 음의 파형



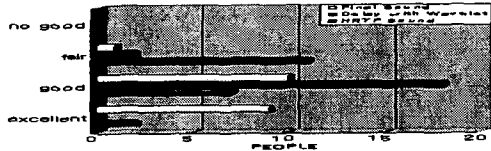
(다) 최종적인 음

그림 4. 첼로 연주 음의 파형 비교

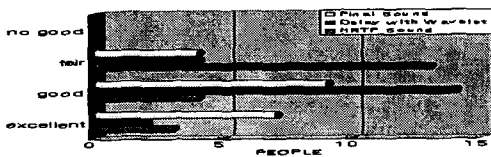
2) 최종 음에 대한 청취 평가 결과

입체음향에 대한 기초적인 지식이 전혀 없는 본 대학교의 재학생들을 대상으로 최종적으로 만든 음에 대한 주관적인 청취 평가를 실시했다. 그림 5는 피실험자들

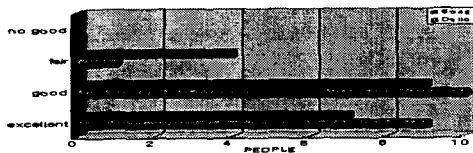
의 음에 대한 확산감 효과의 정도를 그래프로 보여주고 있다. 각각 머리전달함수만을 적용했을 때, 시간 지연의 효과를 포함했을 때, 그리고 노이즈 신호를 감소시키고 만든 최종 음의 분석을 보여주고 있다.



(가) 첼로 연주 음에 대한 청취 평가



(나) 스와니 강(노래 곡)에 대한 청취 평가



(다) 최종 음(첼로, 스와니 강)

그림 4. 연주 음들의 청취 평가 결과

IV. 결론

본 논문에서는 3D 사운드의 확산감 효과를 증가시키기 위하여 웨이브렛을 이용한 음의 확산감 제어 시스템을 제안하였다. 본 연구의 주관적인 평가 결과를 볼 때, 입체 음향을 위한 음의 확산감 제어에 대한 본 논문에서의 알고리즘이 어느 정도 우수한 결과를 얻었다고 본다. 주관적인 청취 평가 결과, 음의 확산감 효과 면에서 본 논문이 제안한 시스템은 고주파수의 음보다는 저주파수의 음에 더 효과적이다. 실제 악기 음의 실험 결과에서도 첼로 연주 음이 트럼펫 음보다도 훨씬 음의 확산 정도가 좋게 나타났다.

음의 확산 정도를 확인하기 위해서 이용한 상관계수 값을 분석해 보면, 원래의 머리전달함수만을 적용한 음보다는 본 논문에서 제안한 시스템의 최종 결과 음이 상관계수 값 면에서도 상당히 개선되었음을 확인할 수 있었다.

본 논문의 향후 진행 과제로는 다양한 음원을 채취하고 테스트하여 공통된 인자를 추출해 내는 것과 사람의 청각 특성을 고려한 청취 공간의 효과를 더욱 높

이기 위하여 컷바퀴 효과(pinna effect)와 연주 홀의 특성을 포함한 연구를 할 것이다. 또한 개발된 음원을 이용한 웹 상에서의 구현(사이버 연주)에 관한 연구도 진행해 나갈 것이다.

본 연구는 2000년도 정보통신 우수시범학교 지원 사업으로 이루어졌습니다.

References

- [1] 이동우, 김영오 외 2, "실시간 입체 음향 제어 시스템의 구현", 한국음향학회지 제 18권 제 3호, pp.79-87, 1999.
- [2] 강성훈, 강경옥 공저, 입체 음향, 기전 연구사, pp.134-141, 1997.
- [3] HRTF Measurements of a KEMAR Dummy-Head Microphone MIT Media Lab Perceptual Computing - Technical Report #280.
- [4] S. Carlile, The Physical and Psychophysical Basis of Sound Localization, in Virtual Auditory Space: Generation and Applications, S. Carlile, Ed. Austin, TX: R. G. Landes, 1996, pp 27-78
- [5] J. Blauert, Spatial Hearing. The psychophysics of human sound localization, revised edition, MIT press, Cambridge MA, 1997
- [6] A. Akansu and R. A. Haddad, Multiresolution Signal Decomposition, Academic Press, pp.304-333,1992.
- [7] Alan V. Oppenheim & Alan S. Willsky, Signal & Systems, pp.128, 1997.
- [8] D. H. Lee, S. J. Kim, U. P. Chong, K. N. Kim and S. D. Lee, 3-D Sound Image Control for Two Channel Headphones, in Proceedings of the ICMT99, pp. 410-415, Pusan, KOREA.
- [9] D. H. Lee, K. N. Kim, S. D. Lee and U. P. Chong, A 3-D Sound Orchestra in the Cyber Space, in Proceedings of KORUS-2000, pp. 12-16, Univ. of Ulsan, KOREA.
- [10] D. Begault, 3-D Sound for Virtual Reality and Multimedia, Academic Press, Boston, MA,1994.
- [11] William G. Gardner, Transaural 3-D audio, MIT Media Lab. Technical Report No.342. 1995.
- [12] 황신, 양진우 외 2, "1/3-옥타브 대역통과필터를 이용한 음상정위기법 성능 향상", 한국음향학회지 제20권 제3호, pp.98-103, 2001.