

HRTF를 이용한 2채널 스테레오 음원을 수정 및 편집 할 수 있는 입체음향 저작도구 개발

김영식, 김용일, 배명수, 전수민, 이대호

(주)유니텍 지휘지원체계사업단 2팀

e-mail: {kimys, yikim, msbae, smjeon, dhlee}@unitech.co.kr

Development of Stereo Sound Authoring Tool to Modify and Edit 2Channel Stereo Sound Source Using HRTF

Young-Sik Kim, Yong-Il Kim, Myeong-Soo Bae,

Su-Min Jeon, Dae-Ho Lee

Command and support system development Enterprise Organization

UNITECH Co.Ltd.

요 약

컴퓨터를 이용한 가상훈련체계를 구현하는데 있어, 청각적인 요소는 시각적인 요소 다음으로 인간의 중요한 인지 능력을 담당한다. 특히 청각 능력의 향상은 훈련 시의 성과와 밀접한 관련을 가지고 있으며, 훈련 효과 향상에 기여하는 바가 높다. 본 논문에서는 이와 같은 가상훈련체계를 구축하는데 있어 반드시 필요한 음향시스템을 기존에 단순한 재생이 아닌 사용자 혹은 개발자가 필요로 하는 음원을 직접 저작할 수 있는가에 초점을 두었으며, 머리 전달함수(HRTF: Head Related Transfer Function)를 이용한 음원을 수정 및 편집하고 사용할 수 있는 시험체계를 개발하였다. 체계 성능 평가를 위하여 기능 및 청감 테스트를 실시하였다.

ABSTRACT

In implementing a computerized virtual training system, the auditory element is responsible for the human cognitive ability following visual elements. Especially, the improvement of hearing ability is closely related to the performance of the training, and it contributes to improvement of the training effect. In this paper, we propose a sound system that is necessary for constructing such a virtual training system as a test system that can use a sound source using a head related transfer function (HRTF). Functional and auditory tests were performed to evaluate system performance.

키워드 : 가상훈련시스템, 훈련 음향 생성, 시뮬레이터, 훈련체계, 음향시스템

Key word : Virtual training sys, Acoustic sound generation, Simulator, Training sys, Acoustic sound sys

1. 서론

지속적인 기술의 발전에 따라 이내 국내에서도 다양한 시뮬레이터 및 가상훈련장비를 이용한 훈련체계의 이용 빈도가 증가하고 있다. 특히 이러한 가상훈련장비를 사용하는데 있어 가장 많이 활용되고 있는 것이 3D 영상을 이용한 시뮬레이션 기술을 많이 활용하고 있다. 하지만 대다수의 일반인들은 가상훈련장비의 가장 중요한 역할은 눈에 해상하는 3D 영상에 많은 초점을 맞춰서 설명하고 있다. 그리고 3D 영상과 관련된 새로운 기술들과 프로그램들은 지속적으로 개발 보완되고 있으나 사람의 귀에 해당하는 음향관련 사항에 대해서는 간과하고 있는 것이 사실이다.

가상훈련장비를 통하여 훈련을 진행하고 기록된 데이터로

부터 재현 할 때에 사용자에게는 오감 중 시각과 청각이 가장 중요한 역할을 차지한다. 그 중 청각을 통하여 가상으로 재현된 공간에서 실제 존재하는 것처럼 만들어 줄 수 있다.

본 논문의 주요 구성을 다음과 같이 하였다. 2장에서는 개발목적을 살펴본다. 3장에서는 머리전달 함수의 이론을 살펴보고 4장에서는 사용자가 머리전달함수를 이용한 2 Channel 음원을 수정 및 편집할 수 있는 도구 및 UI를 설명하며 5장에서는 결론에 대하여 논의한다.

2. 개발 목적

이번 연구 및 개발은 순수과학적인 음향을 연구하기보다 사용자가 만들고자 하는 가상의 공간에서 하드웨어적인 준비

요소가 없이도 입체 음향을 만들어 사용할 수 있도록 머리전달함수를 이용한 2채널 스테레오 음원의 단점을 수정/편집할 수 있는 음향저작도구를 개발하기로 했다. 가상혼련 장비 등에 사용되는 멀티채널의 경우보다 저비용으로 입체음향을 만들 수 있을 것으로 기대한다. 급변 연구 및 개발의 목표는 위 수정/편집한 음원이 가상혼련장비 등의 일부로 사용할 수 있는가에 집중하였다.

3. 머리전달 함수란

일반적으로 인간의 청각 위치파악은 양 귀의 소리의 차이를 기반으로 소리의 공간적 위치를 감지하는데 이를 바이노럴 현상이라 하며 청각 적 차이, 소리의 압력 레벨, 도착시간, 두 귀에 나타나는 신호의 스펙트럼의 현상에 따라 방향을 결정한다.[5] HRTF는 두 개의 재생 채널을 사용하여 입체 음향의 공간화를 가능하게 한다. 이는 청취자의 귀에 대한 공간상의 모든 정보를 토대로 음향 응답을 나타내기 때문이다. HRTF에는 사람의 몸통, 어깨 및 머리에 의해 생성되는 반사에 의한 정보가 내장되어 있기 때문이다.[1] 하지만 측정은 일반적으로 고가의 장비 뿐 아니라 다른 요소들로 인하여 주로 연구 목적으로 수행되며 광범위하게 사용하는 데에는 한계가 있다.[2] HRTF 정보는 전형적으로 한 쌍의 이어폰 마이크를 통해 외이도의 소리를 포착하여 측정되며 가능한 모든 방향으로 확장하기 위해 소스 또는 청취자를 정밀하게 회전하는 암 또는 턴테이블로 회전시켜 정보를 얻는다.[1] 방향감 제어의 경우 다음과 같다. 방향감이 없는 모노 신호에 대해서 방향감을 부여하는 음상정위 과정은 특정 방향에서 두 외이에 도달하는 물리적 단서로써 측정된 HRTF와의 필터링을 통해 구현된다. 특정 방향의 HRIR과 모노신호에 대한 컨벌루션은 다음과 같다.

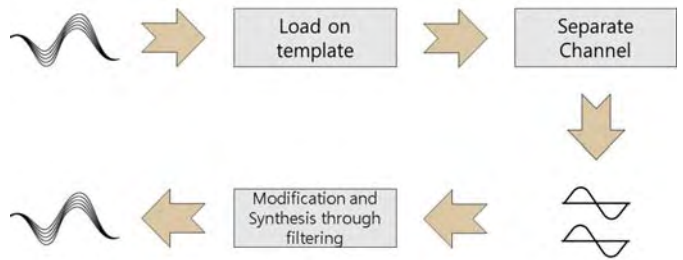
$$y(n) = \sum_{m=0}^{M-1} x(m)h(n-m) \quad (\text{수식 1})$$

여기서 X(n)은 음원샘플, h(n)은 HRIR 필터계수, y(n)은 필터링된 출력신호, M은 HRIR의 탭 개수를 의미한다. 단일 모노 음원에 대해서 특정 방향으로 음상정위 시키기 위해서는 해당 방향의 좌/우 HRTF 필터 계수에 대해서 각각 (수식 1)과 같은 필터링을 수행한다. 그러나 음원이 정면이나 후면 혹은 수직면 상, 즉 두 귀축을 중심으로 원추형 방향에 있는 경우에 대해서 바이노럴 필터링된 결과를 헤드폰으로 들어보면 전/후 방향과 고저에 대한 혼돈 (Cone of Confuse)이 생기게 된다. 이것은 음원이 정면이나 후면 또는 수직면 상에 존재할 대에 방향지각의 주요 소인 ITD나 IID가 좌/우 채널에서 차이가 없기 때문이다. 따라서 HRTF에서 앞/뒤, 상하를 구별할 수 있는 것은 ITD, IID 외에 복합적인 요소들로 이뤄진 스펙트럼 단서에 의한 것인데, 자신의 HRTF가 아닌 더미헤드에서 측정된 HRTF를 이용할 경우에 자기에게 맞는 정확한 스펙트

럼 단서를 제공하지 못하므로 방향지각에 혼돈을 일으킬 수 있다.[3][4]

4. UI구성 및 구현

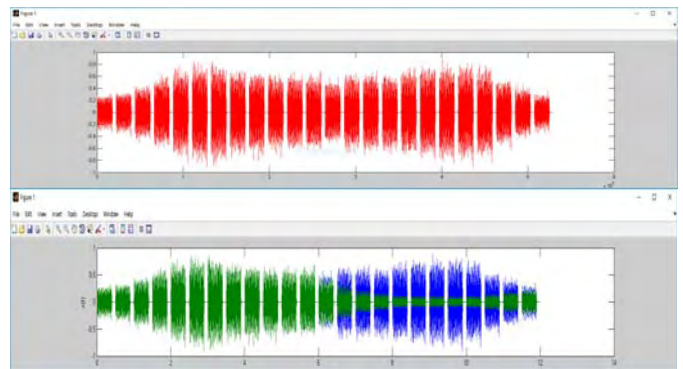
구현한 음원저작도구의 순서도는 아래와 같으며 수정 및 구현, 합성은 편집도구에 해당하며 주 기능은 편집도구에 해당한다.



(그림 1) 저작도구 순서도

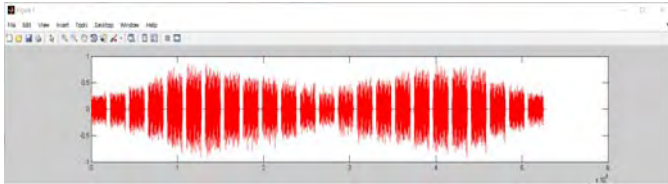
4.1 음원분석

머리전달함수를 이용한 2채널 음원을 분석하기 위해 파형분석을 Matlab을 이용하여 먼저 실험에 착수하였다. 우선 전자음향 파일을 받아 Matlab을 통해 두 신호가 컨벌루션을 통해 입체음향을 구현해 준다는 것을 가지적으로도 확인하였고 다른 부수적인 UI나 프로그램 없이 사람들을 상대로 청취테스트를 했다.



(그림 2) 초기신호(위) 초기신호분석(아래)

위 음향의 결과에 따른 단점은 머리전달함수의 취약점으로 개인 특성마다 맞춘 것이 아니므로 앞/뒤의 구분이 일차적으로 지각할 수 없었다. 이를 개선하기 위하여 머리를 기준으로 앞과 뒤일 경우를 나눠 생각했고 수정이 필요한 대역폭의 진폭을 다운시킴으로서 그 결과는 (그림 3)과 같이 나왔으며 이를 통해 앞과 뒤의 경우를 분간할 수 있는 음원을 저작할 수 있었다.

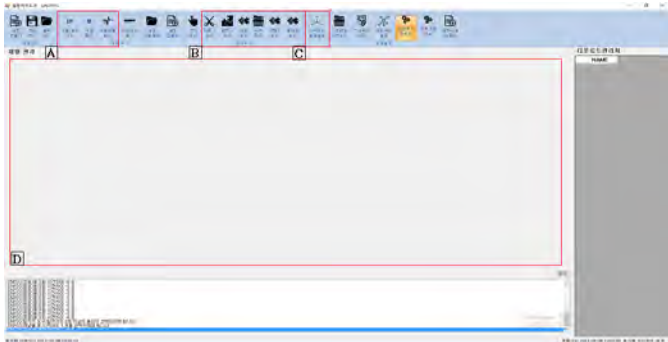


(그림 3) 초기 신호의 신호처리를 통한 개선

4.2 UI구성 및 구현

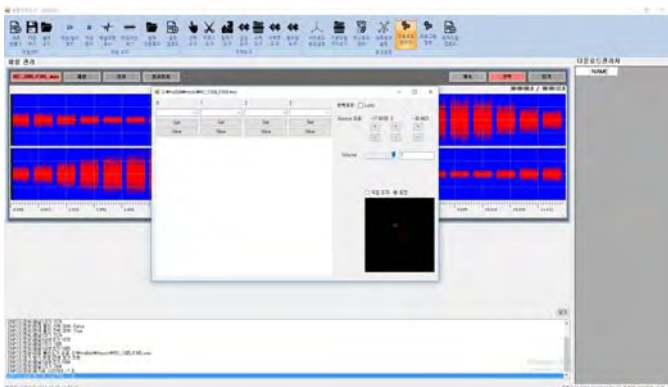
우리는 가장 먼저 머리전달함수의 입체음향의 단점을 하드웨어적인 설치하여 바로 가상장비에 사용하기 위해 음원을 저작할 수 있는 UI를 아래와 같이 구현하였다. 주 메뉴는 아래와 같다.

- A. 음원의 시작과 재생
- B. 음원의 편집
- C. 음원 환경설정 및 가시적 효과
- D. 음원 파일 정보

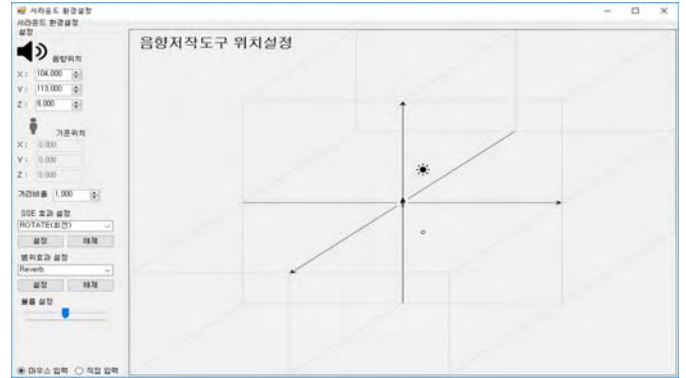


(그림 4) 저작도구 UI구성

채널을 분리하기 위하여 2채널 분리 화면을 구성했고 방향지각의 혼란을 줄이기 위하여 아래와 같이 가시적으로 확인할 수 있는 2차원과 3차원 환경설정을 구현하였다.



(그림 5) 2D 환경



(그림 6) 3D 환경

5. 향후 연구 및 개발과 결론

본 논문에서는 가상훈련장비를 만들 적에 멀티채널에 준하는 2채널 성능을 만들어 내어 가상훈련장비에 적용할 수 있는 음원을 저작할 수 있는가, 사용자가 입체음향을 저작하여 사용하기에 무리가 없는가에 초점을 맞췄고 이에 따른 음향의 변조와 UI로 구성하는 과정을 보여주었다. 우선 멀티채널에 준하는 HRTF를 이용한 2채널 입체음향을 저작하고 테스트를 위해 스피커를 설치 후 5.1채널 입체음향을 들려주어 평가의 기준으로 삼았다. 그 다음으로 수정하지 않은 음원을, 마지막으로 개선한 음원으로 청취테스트를 실시하였다. 수정하지 않은 음원의 분석 결과로 청취테스트를 했을 때 상/하 및 앞/뒤를 지각하는데 어려움이 있다는 결론을 얻을 수 있었고 다채널입체음향에 준하지 못한다는 것을 테스트를 통해 얻을 수 있었다. 처음에는 음원을 그대로 두고 가시적인 것에 비중을 높여 2D화면을 구축하여 다시 청취테스트를 했을 때는 약간의 개선된 결과를 도출 하였지만 지각하는데 혼란은 여전히 남아 있었다. 이를 개선시키기 위해 앞에서 일어나는 일과 뒤에서 일어나는 일을 구분할 수 있도록 상황을 나눴고 진폭의 크기를 변조함으로써 앞과 뒤를 구분할 수 있지만 .wav 파일을 둘로 나뉘어야 한다는 결과를 얻을 수 있었다. 상하의 경우는 3차원의 화면과 음향의 전체적인 크기를 조절하였을 때 구분을 할 수 있었다. 이번 개발의 경우 정적인 상황에서는 HRTF를 이용한 2채널 입체음향의 단점을 저작으로 개선하여 상황에 맞게 대입하여 사용할 수 있다는 결론을 얻었다. 하지만 동적인 상황과 조건이 많은 상황을 표출해야 하는 경우 상황에 따른 저작한 음원의 수가 늘어남에 따라 입체음향의 데이터가 많아진다는 단점과 음원을 바로 사용하지 못하고 사용자가 직접 저작하여 사용해야 하는데 많은 음원을 일일이 다시 수정하기에는 어려움이 있다. 그래서 향후 연구 및 개발에서는 저작된 음원으로 차지되는 비중을 줄이기 위해 샘플링 수를 줄이고 각 음향의 채널마다 잔향의 비중을 줄여 단점을 보완하면 가상훈련장비에 무리 없이 대입하여 사용될 수 있을 것으로 전망하고 있다.

참고문헌

- [1] L. Bonacina, A. Canclini, F. Antonacci, M. Marcon, A. Sarti, S. Tubaro, A LOW-COST SOLUTION TO 3D PINNA MODELING FOR HRTF PREDICTION, Politecnico di Milano Dipartimento di Elettronica, Informazione, e Bioingegneria Via Ponzio 34/5, 20133, Milano, Italy
- [2] Alok Meshram¹, Ravish Mehra¹, and Dinesh Manocha¹ Dept. of Computer Science, University of North Carolina at Chapel Hill, 201 South Columbia Street, NC, 27599-3175 USA, Efficient HRTF Computation using Adaptive Rectangular Decomposition
- [3] M. TOYAMA, M. UCHIYAMA, and H. NOMURA Kogakuin University East-Asia University Hachioji-shi, Tokyo Shimonoseki-shi, Yamaguchi 192-0015 Japan Japan , HEAD RELATED TRANSFER FUNCTION REPRESENTATION OF DIRECTIONAL SOUND FOR SPATIAL ACOUSTIC EVENTS MODELING
- [4] 김시호, 김경훈, 배건성, 최송인, 박만호, HRTF를 이용한 헤드폰 기반의 다채널 입체음향 생성, 논문 2005-42SP-1-10
- [5] HRTF measurement on KEMAR manikin Mengqiu Zhang, Wen Zhang, Rodney A. Kennedy, and Thushara D. Abhayapala Applied Signal Processing Group College of Engineering and Computer Science The Australian National University, Canberra ACT 0200, Australia