

입체음향 기반 기술 및 동향

김 현 빈, 명 현, 김 기 홍

한국전자통신연구원 컴퓨터·소프트웨어기술연구소 가상현실연구센터 청각정보연구팀

3D SOUND TECHNOLOGY AND ITS TRENDS

Kim Hyun Bin, Myung Hyun, and Kim Ki Hong

Auditory Information Team, Virtual Reality Research Center, CSTL, ETRI

요 약

최근 PC의 멀티미디어 성능이 향상되고 디지털 신호처리 기술이 발달함에 따라 복잡한 수식과 방대한 정보처리가 요구되는 입체음향(3D Sound) 처리를 PC상에서도 구현하는 것이 가능하게 되면서, 입체음향에 대한 사용자의 관심이 고조되고 있다. 입체음향을 이용한 청각효과는 3D영상의 시각효과에 못지 않은 강한 인상을 줄 수 있는 수단으로써 인정받고 있어 게임, 가상현실, 멀티미디어 콘텐츠 등에서 사용자의 몰입감과 체험 효과를 최대한으로 증대시킬 수 있는 새로운 수단으로 활용되고 있고, 이와 관련하여 입체음향에 대한 연구와 개발이 국내외에서 활발히 진행 중에 있다. 본 논문에서는 2채널 방식의 입체음향 기반기술과 최근의 입체음향 기술개발 동향을 살펴본다.

1. 서론

입체음향은 음원이 발생한 공간에 직접 위치하지 않은 청취자가 재생된 음향을 들었을 때에 음향으로부터 방향감, 거리감 및 공간감 등과 같은 공간적 단서를 지각할 수 있는 음향을 말하며[1], 단순한 모노(Mono)음이나 스테레오(Stereo)음의 신호에 공간적 지각 단서를 부가하여 입체음향신호로 변환하는 기술을 입체음향 생성기술이라 한다. 입체음향 생성 방식은 여러 대의 스피커를 이용하는 서라운드(Surround) 타입의 멀티채널 방식과 2대의 스피커나 헤드폰을 이용하는 바이노럴(Binaural) 타입의 2채널 방식으로 구분할 수 있다.

입체음향관련 기술은 주로 미국, 일본 및 유럽과 같은 선진국을 중심으로 발달된 반면, 국내 기술은 이들 선진국에 비해 기술 수준이 아직 요원한 상태이다. 또한 국내의 영상분야의 3차원 그래픽 기술에 비해 입체

음향 기술은 상대적으로 낙후된 감이 없지 않다. 그러나, 최근에 한국전자통신연구소와 일부 대학에서 입체음향 관련 기술을 연구 중에 있으며, 입체음향을 지원하는 사운드카드나 기존 스테레오방식에서 입체음향을 지원하는 신제품들이 개발되고 있다.

본 논문의 2절에서는 2채널 방식으로 입체음향 생성을 위한 핵심 기술에 대해서 서술하고, 3절에서는 입체음향과 관련한 최근 국내외 기술개발 동향을 살펴본다. 마지막으로 4절에서는 한국전자통신연구원에서 개발 중인 입체음향관련 연구 결과물에 대해서 간단히 소개한다.

2. 2채널 방식의 입체음향 기반기술

음원(Sound Source)이 발생한 공간 내에 있는 청취자의 양쪽 귀에 마이크로폰을 각각 설치하여 녹음한 신호를 바이노럴(Binaural) 신호라 하며, 이 신호를 헤드폰으로 재생할 경우, 현장에서 직접 듣는 것과 같은 음상(Sound Image)을 지각할 수 있다. 음원과 음상은 공간적 특성이 반드시 일치하지 않으며, 음원과 음상이 일치할수록 좋은 음질의 입체음향이 구현되었다고 할 수 있다. 바이노럴 신호에는 음원의 위치, 방향뿐만 아니라 음원을 둘러싸고 있는 공간, 즉 음장(Sound Field)과 관련한 공간적 단서들이 포함되어 있다. 음장이란 음향학적 용어로 가청주파수의 음파가 존재하는 공간을 말한다. 2채널에 의한 방식은 인간이 두 개의 귀로 음향을 지각하는 특성을 이용하여, 음상정위와 음장제어에 의해 생성된 입체음향을 2채널에 의해서 재생하는 방식을 말한다.

바이노럴 타입의 2채널 입체음향을 생성하는 방식으로는 녹음과 필터링에 의한 방법이 있다. 녹음에 의한 방법은 청취자의 양쪽 귀

에 장착한 두 개의 마이크로폰을 이용하여 현장음을 녹음하고 이를 재생하는 방식이다. 모든 사람이 자신의 바이노럴 신호를 녹음할 수 없기 때문에, 현재는 주로 표준치의 머리 모형을 가진 더미헤드(Dummy Head)에 장착한 마이크로폰을 통하여 바이노럴 신호를 녹음하고 이를 일반 청취자에게 들려주는 방식을 이용한다.[2]

필터링에 의한 방법은 단순음을 변형하여 입체음을 생성하는 방식으로, 주로 머리전달함수(Head Related Transfer Function; HRTF)가 필터로서 이용된다. 머리전달함수는 무향실 내에서, 더미헤드를 중심으로 구의 형태로 여러 각도에 배치한 스피커로부터 백색잡음과 같은 임펄스 신호를 방사시켜, 더미헤드의 양쪽 귀에 장착한 마이크로폰으로 측정된 임펄스 응답을 푸리에(Fourier) 변환한 것을 말한다. 이 머리전달함수와 단순음을 컨볼루션(Convolution)하면 원하는 공간상의 위치에 음상을 정위시킬 수 있다.

가. 머리전달함수

청취자가 음원에 대한 공간적 단서를 지각할 수 있는 것은 머리전달계의 고유 특성에 의해서 두 귀에 입사한 두 신호 간에 차이가 발생하기 때문이다. 이 두 신호차에 대한 특성은 머리전달함수에 내포되어 있어, 이를 이용하면 입체화되지 않은 단순음에 공간적 정보가 부가된 바이노럴 타입의 입체음을 생성할 수 있다. 머리전달함수는 입사하는 각도에 따라 달라지기 때문에, 여러 위치에서 임펄스에 대한 머리전달함수를 측정하고, 이를 DB로 구축하는 것이 필요하다. 이 DB로부터 원하는 위치에 해당하는 머리전달함수를 선택하여 단순음과 컨볼루션을 수행하면, 해당 위치에 음상을 정위시킬 수 있다.

MIT Media Lab.에서는 KEMAR 더미헤드를 사용하여 710지점에서 측정한 머리전달함수 DB를 인터넷[3]에서 무료로 공개하고 있다. 한국전자통신연구원에서도 Neumann 더미헤드를 사용하여 측정한 머리전달함수의 DB를 현재 구축하고 입체음향관련 과제에서 활용 중에 있다.

나. 입체음상 정위

청취자가 지각한 음상에 대한 위치를 파악하는 것을 음상정위라 하고, 음상을 공간 상의 특정 장소에 위치시키는 기술을 음상정위

기술이라 한다. 이 기술을 이용하면 고정된 특정 위치에서 소리가 지각되는 “위치음” 효과와 소리가 한 위치에서 다른 위치로 움직이는 “이동음”효과를 생성할 수 있다. 위치음 생성은 해당 위치의 머리전달함수를 단순음과 컨볼루션 연산을 행함으로써 얻을 수 있고, 이동음 생성은 소리가 이동하는 궤적 상에 해당하는 연속적인 머리전달함수들을 단순음과 컨볼루션을 수행함으로써 얻을 수 있다. 이때 측정되지 않은 불연속 공간의 머리전달함수는 인접한 머리전달함수들간의 보간을 통해서 구할 수 있다.

이동음의 경우, 머리전달함수를 이용하는 방식 외에 도플러(Doppler)효과를 이용하는 방식이 있는데, 이는 음원이 가까워지면 주파수가 높은 쪽으로 이동하고, 멀어지면 주파수가 낮은 쪽으로 이동하는 현상을 이용하는 것으로, 소리가 멀어지고 가까워지는 이동음 효과를 낼 수 있다.

다. 가상음장 제어

음원을 둘러싸고 있는 실내 공간의 특성에 따라 동일한 음원이라 할지라도 청취자에게 다른 음향효과를 줄 수 있다. 이는 실내 공간의 크기, 구조, 벽 또는 천정의 재질 등에 의해서 음원에 대한 직접음, 초기반사음, 잔향패턴 및 잔향시간 등이 달라지기 때문이다. 잔향은 음원으로부터 방사가 그친 후에도 천정이나 벽으로부터의 반사가 계속되어 울리는 음으로, 공간감 생성에 주요한 요인임과 동시에 거리감 생성에도 중요한 역할을 하는 것으로 알려져 있다. 반사음과 잔향을 인공적으로 제어하여 특정 실내에 음원이 있는 것과 같은 음향효과를 생성하는 기술을 음장제어기술이라 한다.

음장제어에 가장 많이 사용되는 것 중의 하나로 Schroeder 잔향기[4]가 있는데, 이 잔향기는 병렬로 연결된 다수의 Comb필터와 두 개의 직렬로 연결된 전역통과필터로 구성되어 있다. 여기서 Comb필터는 특정 주파수가 진동하는 효과를 내고, 전역통과필터는 잔향 밀도를 증가시키는 역할을 한다. 이외에 실내의 잔향 특성을 음향학적으로 모사하는 음선추적(Ray Tracing) 방식과 이미지 모델(Image Model)방식이 있다. 음선추적 방식은 음원에서 나오는 음은 모든 방향으로 방사되는 성질을 고려하여 음의 에너지 분포를 구하여 가상 음장을 생성하는 방법이고,

이미지 모델 방식은 빛이 거울에서 반사하는 것과 같이 음파도 실내의 벽면에 부딪혀 한번 반사한다는 전제하에 반사 경로를 구하여 가상 음장을 생성하는 방법이다. 또한 공간 전달함수를 이용하는 방법이 있는데, 이는 특정 공간의 공간적 단서가 내포된 공간전달함수를 측정하여, 이를 단순음과 컨볼루션을 행하여 특정 실내의 음장 특성을 부가하는 방법이다.

라. 입체음향 재생

바이노럴 타입의 입체음향을 재생하는 방식은 2개의 스피커에 의한 재생방식과 헤드폰 재생방식이 주류를 이루는데, 이 두 방식은 서로 장단점이 있다. 먼저 헤드폰 재생방식의 경우, 충실한 입체음향 효과를 구현할 수 있으나, 음상이 청취자의 머리 안에서 지각되는 현상이 발생한다. 실세계의 소리들은 일반적으로 머리의 바깥쪽에서 지각되므로, 머리 바깥 쪽으로 음상을 끌어내기 위한 과정이 필요한데, 이를 외재화(Externalization)라 한다. 청취자에 따라 외재화에 대한 지각 정도가 다르고 주관적이기 때문에, 아직 정립된 방법에 의한 기술은 보고되지 않고 있으나, 정밀도가 높은 머리전달함수 또는 자신의 머리전달함수를 이용할 경우 외재화 효과를 어느 정도 얻는 것으로 보고되고 있다.

스피커 재생방식의 경우, 헤드폰과 같은 내재화 현상은 발생하지 않으나, 좌측 스피커에서 재생된 신호가 청취자의 좌측 귀에만 도달하는 것이 아니고 우측 귀에도 입사되며, 우측 스피커의 신호도 같은 현상이 발생하여, 이 신호들 간에 서로 상호간섭(Crosstalk)을 일으킨다. 이와 같은 현상은 바이노럴 신호에 의한 입체음향 효과를 저하시키는 요인으로 작용한다. 스피커로 방사하기 직전에 Crosstalk 제거를 위한 역필터를 거칠 경우, 상호간섭 없이 입체음향을 스피커로 재생할 수 있다. 이 방식을 트랜스오럴(Transaural)시스템이라 하는데, 충실한 입체음향 효과를 얻기 위해서는 양 스피커와 청취자의 머리 위치를 파라미터로 해서 역필터를 적용해야 한다.

3. 입체음향 개발동향

가. 국외 현황

현재 상용화 수준으로 개발된 해외의 대표

적인 알고리즘 및 제품들로는 다음과 같은 것들이 있다. 싱가포르의 Creative사에서 개발한 멀티미디어용 사운드카드인 AWE64IE의 경우, 스테레오 강화기술을 적용하여 음의 발산영역을 확대하여 공간감을 향상시켰으며, 음원의 위치를 3차원 좌표에 설정하여 해당 위치음을 생성하는 3D Positioning 기능을 갖고 있다. 캐나다의 QSound 알고리즘은 원칙적으로 일상적인 스테레오 재생시스템 장치 이외에 하드웨어를 새로이 추가하지 않고, 2채널의 입체음향 신호를 생성하는 녹음 기술로서 머리전달함수 대신 인간의 청각 신경계를 포함한 전달함수를 모델링하여 입체음향을 녹음하는 것으로 알려져 있다. 미국의 SRS(Sound Retrieval System) Lab.에서 개발한 TruSurround는 5채널의 신호를 2채널로 매핑시켜 음상정위를 실현시킨 기술로서, 기본적으로 머리전달함수를 이용하여 공간 정보를 부가하는 기술을 채용하고 있지만, Crosstalk를 제거하는 과정을 채용하지 않음으로서 계산량을 줄인 것이 특징이다. 미국의 Spatializer Audio Lab.에서는 SRS처럼 라이선스로 기술을 제공하는데, 그 중에 enCOMPASS는 머리전달함수를 이용한 음상정위 기술로, 마이크로소프트사의 DirectSound를 이용한 게임 개발에 이용될 수 있으며, N-2-2는 2채널만으로 멀티채널 효과를 얻을 수 있는 기술로, 다른 유사기술에 비해 청취자의 청취영역이 넓은 것으로 보고되고 있다. 그 외에도 일본의 RSS(Roland Sound Space) 시스템, 호주 LakeDSP사의 Huron DSP 보드, S/W만으로 실시간에 가까운 컨볼루션 연산을 가능하게 해주는 이탈리아 Parma대학의 Aurora라 등을 들 수 있다.

일본의 Hamada교수는 두 개의 스피커가 인접할 경우 스테레오 효과가 떨어지는 것을 개선하기 위하여 스테레오 다이폴 시스템(Stereo Dipole System)을 제안하였다. 미국의 Crystal River Engineering사는 청취자가 머리를 움직일 경우 음파의 입사 각도가 달라져서 음상의 방향감이 달라지는 것에 착안해서, 헤드 트래커로 청취자의 머리 움직임을 추적하여, 머리 움직임에 따른 방향감을 실시간으로 생성하는 시스템인 Convolvotron을 개발하였으나, 헤드폰 전용으로 장비가 고가인 단점이 있다. 미국의 George Washington 대학에서는 사운드 렌

더링(Sound Rendering)이란 기술을 소개하였는데, 이는 특정한 음색을 표현 또는 합성하는 Timber Tree와 Sound Modelling을 이용하여 실제 음향을 생성하는 기술로, 현재 이론적인 연구가 진행중이다. 이외에도 최근 선진국에서는 콘서트홀이나 실내를 건축하기 전에 음장효과를 미리 시뮬레이션하기 위한 가청화(Auralization)기술과 음원이나 소음으로부터 방사되는 음파의 형태를 그래픽으로 처리하여 시각적으로 보여주는 소리 가시화(Sound Visualization) 기술 등이 활발히 연구되고 있다.

나. 국내 현황

입체음향과 관련하여 상용화된 국내 제품으로는, 멀티미디어에 내장되는 사운드카드에 입체음향 지원 기능을 추가한 제품들과 기존의 음향 장비에 간단한 연결만으로 입체음향을 재현하는 장치인 DTEC(주)의 3S(Spectral Surround System)와 라스텔의 RSF(Royal Sound Field)가 있고, 이외에 JG코포레이션의 체감스피커 등이 있다. 경북대 최평교수가 개발하여 DTEC(주)에 기술을 이전한 3S는 주파수 대역을 고·중·저 대역으로 3등분하여 대칭적으로 처리하고, 스테레오 양쪽 신호에 위상적 특성을 가미하여 공간감을 확대하는 기능과 소리의 원근감을 보정하는 기능 등이 있다. 라스텔의 RSF는 DSP를 사용하지 않고 각종 음향 관련 시스템에서 아날로그 다채널 음장을 구현한 기술로, 저렴한 가격으로 DSP칩에 버금가는 입체음향을 구현할 수 있다. JG코포레이션의 체감스피커는 소파, 의자 또는 자켓 등에 얇은 두께의 스피커를 내장하여 소리의 진동을 몸으로 느낄 수 있게 한 것이 특징인데, 입체음향이라고는 할 수 없지만 게임이나 가상 현실 등에서 이용할 경우 실감나는 체감음향을 경험할 수 있으리라 기대된다.

4. 입체음향생성 저작도구 및 실감음향 인터페이스

한국전자통신연구원 청각정보연구팀은 “멀티미디어 콘텐츠용 입체음향 처리 S/W개발” 과제(정통부 출연)를 통하여 입체 음향 처리에 필요한 핵심 요소기술을 확보하고, 청각장비 사용자로 하여금 고급 장비를 사용하지 않고 일반 PC 상에서 순수 소프트웨어만으로

입체음향을 생성할 수 있게 하는 저작도구 “소리결(SoriWave)”을 개발하였다.

본 연구팀에서 개발한 입체음향생성 저작도구(소리결)는 바이노럴 타입의 2채널 입체음향을 생성하는 방식으로 개발되었고 크게 음향편집, 입체음상정위, 가상음장제어 및 크로스톡제거 기능을 지원하는 4가지의 모듈로 구성되어 있으며, 기존의 해외 음향 편집기(CoolEdit Pro, SoundForge 등)에서 지원하는 입체음향 생성 기능에 비해 성능과 사용자의 편의성이 증대되었다.

그림 1은 소리결의 4가지 기본 모듈을 나타낸다.

메인 윈도우인 음향편집 모듈은 File, Edit, View, Remocon, Option, Window, Help 기능을 갖는 7개의 메뉴로 구성되어 있다. 이외에 각종 특수효과(Chorus/Flange, Panning, Pitch Shifting, Vibrato, Fade In/Out, Doppler)를 생성하는 기능을 갖는다.

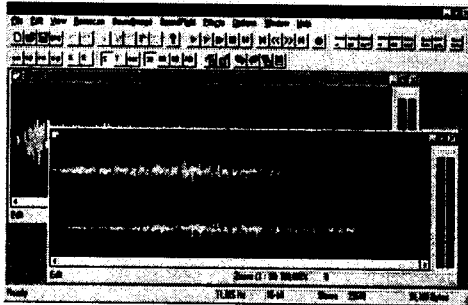
입체음상정위 모듈은 미리 정의되어 있는 위치음/이동음을 생성하는 기능과 사용자가 임의로 정의하는 위치음/이동음을 생성하는 기능을 갖는다. 그림에서는 입체음상정위 모듈의 시스템 정의 위치음/이동음 생성을 위한 사용자 인터페이스 환경을 나타낸다.

가상음장제어 모듈은 잔향, 흡음계수 및 공간의 크기를 시뮬레이션하여 사용자가 원하는 가상공간을 생성하는 기능과 실제 측정된 실내응답함수를 이용하여 생성하는 기능을 갖는다.

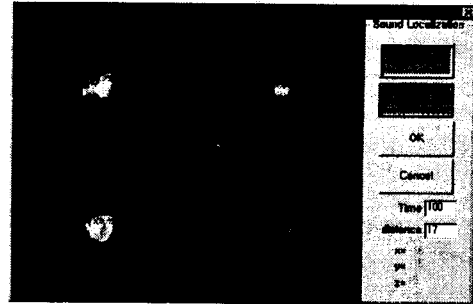
크로스톡제거 모듈에서는 두개의 스피커 위치와 청취자의 위치를 사용자가 임의로 선택하고, 이들 위치에 따라 크로스톡을 제거하는 기능을 갖는다.

저작도구 이외에도 본 연구팀에서는 현재 영상과 음향의 공간적 동기화뿐만 아니라

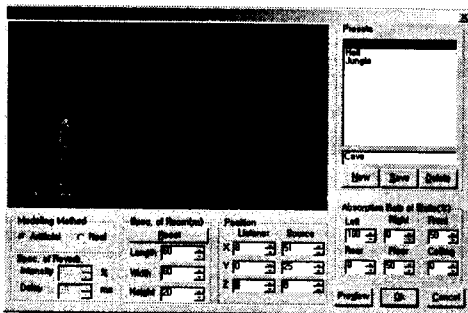
라 청취자의 머리 움직임까지도 함께 고려한 실감음향 인터페이스를 현재 개발 중에 있다. 본 연구팀에서 개발 중인 실감음향 인터페이스는 HMD에 부착된 Head Tracker를 이용하여 청취자의 머리움직임을 추적하고, 여기서 얻어진 머리위치 정보와 영상내에서 소리를 내는 이동객체의 위치 정보를 해당 음향에 매핑시킨 음향(실감음향)을 HMD내에 부착된 음향장치로 시간적/공간적으로 동기화시켜 영상과 함께 재생하는 구조로 되어 있다. 이에 대한 개발이 진행 중으로 가까운 시일에 관련학회에서 발표할 예정이다.



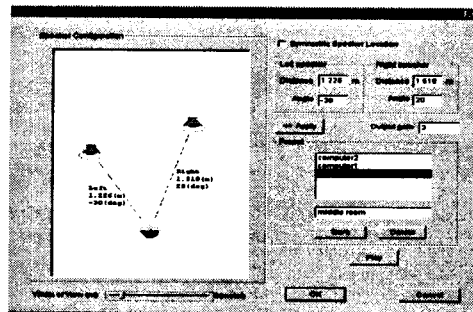
(a) 편집 모듈



(b) 음성 정위 모듈



(c) 음성 제어 모듈



(d) 크로스톡 제거 모듈

그림 1. ETRI에서 개발한 입체음향 저작도구 “소리결”의 GUI 환경

5. 맺음말

앞에서 소개한 입체음향 기술의 활용 분야는 실로 다양하다. 영화산업에서 특수 입체음향 효과는 관객의 영화에 대한 흥미도를 배가시킬 수 있으며, 게임, 가상현실 및 멀티미디어 콘텐츠 분야에서의 입체음향 효과는 사용자의 몰입감을 증대시킬 수 있다. 또한 노래 반주기, 사운드카드 및 일반 오디오 등에서의 입체음향 지원 기능은 제품의 고부가가치를 창출할 수 있으며, 카 오디오 및 가정용 극장시스템(Home Theater)에서 입체음향의 재생은 적은 공간에서도 청취자에게 만족할 만한 음향효과를 충실히 제공할 수 있다. 이 뿐만 아니라 원격회의, 초고속전산망에서의 실감음향 통신, 방송, 교육 및 국방 등 음향이 필요한 모든 분야에서, 입체음향은 1960년대 이후 일반인에게 보급된 스테레오와 서라운드 음향효과에 버금가는 반향을 일으킬 수 있을 것이다.

참고문헌

- [1] 이희중, 서상원, 이명진, 김용완, “입체음향 기술”, *소프트웨어 기술동향*, vol.2, no.2, pp.54-80, 1998.
- [2] Begault, D. R., *3-D Sound for Virtual Reality and Multimedia*, New York, Academic Press Inc., 1994.
- [3] <http://sound.media.mit.edu>
- [4] Schroeder, M. R., “*Natural Sounding Artificial Reverberation*,” *J. Audio Eng. Soc.*, vol.10, pp.219-223, 1962.