

가상현실을 위한 실감음향 인터페이스

김용완*, 김현석*, 김현빈*

*한국전자통신연구원 가상현실연구개발센터
e-mail : ywkim@etri.re.kr

Real Sound Interface for Virtual Reality

Yong-Wan Kim*, Hyun-Suk Kim*, Hyun-Bin Kim*
*Virtual Reality Research Center, ETRI

요 약

가상현실은 현실에 구애받지 않고 상상의 세계를 현실과 같이 만들어 내어 인체의 모든 감각기관이 인위적으로 창조된 세계에 몰입됨으로써 현실에 있는 것처럼 느낄 수 있는 사이버스페이스이다. 본 논문에서는 청취자의 머리움직임에 따라 머리 추적기로부터 수신된 정보로서 음향의 방향감 간의 상관모델을 구성하여 시스템내의 S/W 모듈에 전달하는 인터페이스와 이를 토대로 머리 움직임 변화에 부합되는 실감음향을 생성하고, 영상도 함께 동기화하여 몰입감을 증대시키는 실감음향 인터페이스에 대해서 논의한다.

1. 서론

현재 가상현실 시스템의 몰입감을 위해서 HMD(Head Mounted Display), Tracker, Data Glove, 입체음향 시스템, 3D 그래픽 가속보드를 장착한 컴퓨터 등이 유기적인 인터페이스를 유지하며 동작하는 시스템이 개발되고 있다 [4][5].

입 변화에 부합되는 실감음향을 생성하고, 영상도 함께 동기화하여 몰입감을 증대시키는 실감음향 인터페이스에 관하여 논의한다. 이와 같이 가상현실에 입체음향 기술을 적용할 경우 사용자에게 보다 현실감 있는 음향을 제공할 수 있어 현장감과 몰입감을 증대할 수 있다.

그림1은 실감음향 인터페이스 시스템을 보여주고 있으며, 크게 입체음향 생성부분, 영상/음향/머리움직임 동기화 부분, 위치 추적 부분 등으로 구성된다.

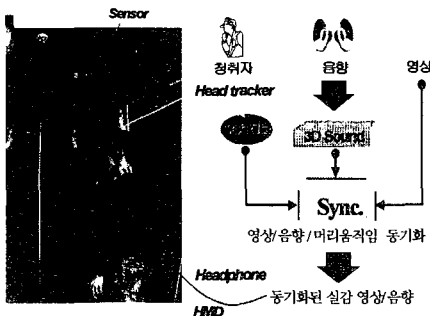


그림 1 실감음향인터페이스

2. 실감음향 인터페이스

현재까지 초음파나 자기장을 이용하여 청취자의 공간상에서의 위치변화를 컴퓨터가 인식케 하여 청취자의 이동에 따라 음원의 변화를 가능하게 하는 실감음향시스템이 연구되고 있으며 Crystal River Engineering사의 Convolutron의 경우 머리 추적장비를 장착하고 90msec의 시간지연을 갖는 실시간에 가까운 장비이나 4개의 DSP 보드를 이용하여 개발된 고가의 장비이며, 그 외에 현재까지 개발된 대표적인 모션 트래킹 장비로는 Polhemus사의 UltraTrak Pro System, Ascension Technology 사의 Motion Star, Logitech 사의 Head Tracker 등이 있다 [3][4][5].

실감음향 인터페이스는 HMD에 부착된 Head Tracker

이에 본 논문에서는 청취자의 머리움직임에 따라 머리 추적기로부터 수신된 정보에 기초하여 음향의 방향감 간의 상관모델을 구성하여 시스템내의 S/W 모듈에 전달하는 인터페이스와 이를 토대로 머리 움직임

를 이용하여 청취자의 머리움직임을 추적하고, 여기서 얻어진 머리위치 정보와 영상 내에서 소리를 내는 이동객체의 위치 정보를 해당 음향에 매핑시킨 음향(실감음향)을 HMD내에 부착된 음향장치로 시간적/공간적으로 동기화시켜 영상과 함께 재생하는 구조로 되어있다.

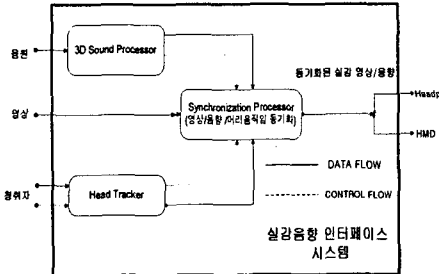


그림 2 실감음향인터페이스의 구성도

그림 2에 보듯이 실감음향 인터페이스 시스템의 각 모듈의 주요 기능은 다음과 같다.

- 1) 그래픽 재생 모듈 : HMD(Head Mount Device)의 영상장치를 이용한 영상 재생 기능
- 2) 오디오 재생 모듈 : HMD(Head Mount Device)의 헤드폰을 이용한 음향 재생 기능
- 3) 머리 움직임 추적 모듈 : 청취자의 머리 움직임을 센싱하여 머리 위치를 추적하는 기능
- 4) 입체음향 렌더링모듈 : 모노/스테레오 음을 입체음향으로 생성하는 기능
- 5) 영상/음향/머리움직임 동기화 모듈 : 영상/머리움직임과 입체음향을 동기화 시켜 영상장치와 헤드폰으로 이미지와 사운드를 보내주는 기능

2.1 머리움직임 추적 모듈

트래킹 시스템은 초음파나 자기장을 이용하여 물체의 공간상의 위치변화를 감지하여 회전값이나 위치값을 획득하는 장치이다. 대표적인 트래킹 시스템으로는 Logitech 사의 Head Tracker, Polhemus 사의 FASTRAK, Ascension Technology 사의 Extended Ranged Bird 등이 있다. 대부분의 시스템은 센서가 2-4 개 정도가 사용되나 더욱 정교한 물체의 움직임 파라미터를 얻어내기 위해서는 그 이상의 센서가 사용되어야 한다. 하지만 그 이상의 데이터를 받아들이기 위해서는 시스템에 부담을 주게 되므로 대부분 2-4 개 정도를 보통 사용하게 된다. 그림 3 에 보는 바와 같이 본 시스템에서는 HMD 에 부착되어 있는 각각 1 개의 InertiaCube 와 SonicDisc 로부터 HMD 의 위치 정보 및 Cue, Raw, Pitch 등의 움직임 파라미터를 얻어내어 입체음향 렌더링 모듈에 전달한다. 입체음향 렌더링 모듈에서는 정의된 객체의 운동으로부터 사운드를 동기화시키고(synchronization), 추출된 움직임 파라미터를 사운드 파라미터로 매핑하고 사운드 렌더링 통하여 입체음향을 생성한다 [1][3].

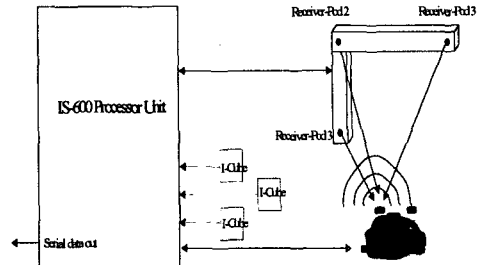


그림 3 머리 움직임 추적 모듈

2.2 입체음향 렌더링 모듈

입체음향 렌더링 모듈은 공간상의 음상정위라던지 동굴이나 콘서트홀같은 공간감에 따라 듣는 느낌이 달라질 수 있도록 하는 음상을 시뮬레이션 하는 것을 말한다. 대표적으로 Crystal River Engineering사의 Beachtron System, Convolver System, Acoustiction Audio 가 있으며 이것들은 HRTF(Head Related Transfer Function)를 좌우측 필터링 처리하여 입체음향을 생성하는 것들이다 [1][2][7][8].

음향관련 재생기술은 멀티채널 마이크로폰을 3차원 공간상에 배치하여 다중음원을 녹음한 후, MPEG-2 오디오나 Dolby AC-3 방식을 이용하여 멀티채널 스피커로 재생하는 방식과 인체의 음향인식 기능을 이용한 입체음향 녹음 및 재생, 더미헤드(인간의 머리과 귀의 구조를 흉내낸 녹음장치)를 통한 음원 녹음, 인간의 머리구조를 기반으로 HRTF(머리전달함수, Head Related Transfer Function)를 측정하여 입체음을 재생하는 방식으로 발전되어 왔다.

입체음향의 구현은 간단하게는 인간의 청각 구조와 유사한 더미헤드를 사용해 좌우 두 귀 채널에 대해 각각 녹음한 음을 헤드폰으로 그대로 재생하면 가능하다. 그러나 일반적으로는 그림 4에서의 흐름과 같이 HRTF를 이용한 음상정위, 실내음장모델을 반영한 음장제어의 신호처리 과정을 거쳐 입체음향을 생성하고 재생함으로써 구현한다.

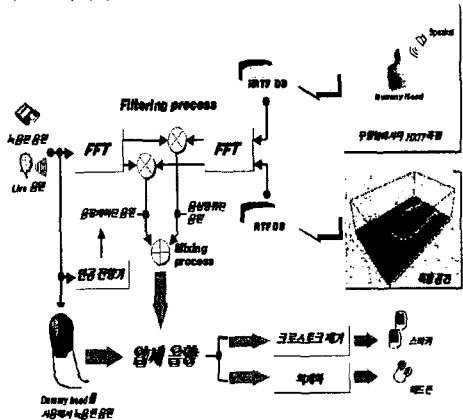


그림 4 입체음향 생성흐름도

2.2.1 입체음상 정위

일반적으로 음상정위의 구현은, 공간 상의 음원이 청취자의 두 귀에 이르기까지의 전달함수를 나타내는 머리전달함수(HRTF)에 의한 바이노럴(Binaural) 필터링 처리와 음원이 청취자로부터 떨어진 거리에 따른 소리세기(Intensity) 변화를 시뮬레이션 하는 이득(Gain) 조절 과정의 구현으로 구성된다. 그림 5는 모노 음원에 대한 2채널 입체음향처리(Binaural Spatialization) 과정을 나타내는 블록도이다.

입력 모노 신호에 대한 바이노럴 HRTF 필터링에 있어서 음상정위 위치를 지정하는 가상음원의 위치 정보가 필요하며 가상음원의 청취자의 수평면(horizontal plane)[9]과 정중면(median plane)[9]에 대한 방위각(azimuth)과 고도각(elevation)을 산정하여 이에 매핑되는 좌우 HRTF 쌍을 인덱싱하게 된다. 본 시스템의 HRTF 필터링에는, 무향실에서 Neumann KU-100 더미헤드를 사용하여 직접 측정된 HRTF 데이터를 사용하였으며 각 방향에 대한 HRTF 필터는 256개의 임펄스 응답 데이터를 FIR 필터 계수로 사용하였다.

또한, 단구간 푸리에변환을 기반으로 한 Overlap-Save 방식 필터링[10]을 구현하였고 FFT 기반 주파수 영역 컨볼루션 알고리즘을 사용하였다. HRTF 필터링을 거친 신호는 가상음원의 청취자로부터의 거리에 대응하는 이득(gain) 조절 과정을 거쳐 음상정위된 입체음향 신호로서 좌우 각 채널에 대해 처리 결과를 출력하게 된다.

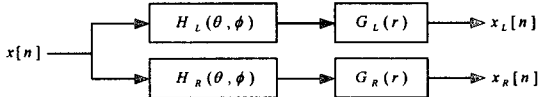


그림 5 2채널 입체음상 정위 블록도

2채널 재생 방식에 있어서 위치음 효과는 비교적 단순한 반면에, 이동음 효과는 구현에 있어서 고려해야 할 점이 많다. 위치음 효과는 단순히 머리 전달 함수와의 컨볼루션을 통해서 구할수 있지만 이동음 효과는 특정 점에서만 측정된 머리전달 함수를 이용하여 이동하는 음을 생성해야 하기 때문이다.

본 논문에서는 입체음향을 2채널로 재생할 때 이동음 효과를 부드럽게 생성되도록 하기 위해 다음의 방법을 사용하였다.

단순히 이동음 효과의 구현은 고정된 HRTF가 아닌 가변적인 HRTF와 가변적인 이득제어를 통해서 한 위치에서 다른 위치로의 이동 음상을 구현할 수 있다. 이를 위해 움직이는 가상음원의 이동 궤적을 정의하고 궤적 상의 각 위치에 대응하여 변하는 HRTF를 연속적으로 매핑하고 필터링하는 것이 필요하다. [10]에 소개된 시스템 구현에서는 가상 음원의 이동 궤적을 선분 세그먼트를 기반으로 구성하고 HRTF를 시간 영역에서 보간하였다. 기존에 사용하던 HRTF는 측정 공간을 매쉬 형태로 분할한 특정 위치에서 측정된 불연속 HRTF이기 때문에 불연속 구간에서의 HRTF는 측정된 주위의 HRTF를 이용하여 보간 기법으로 재생하였던 것이다. 그러나 [10]에서 구현된 이동음 효과는

다음과 같은 문제점들을 지니고 있었다.

- 1) 가변 HRTF 필터링에 있어서 HRTF의 필터링 계수가 바뀌게 되는 경계 부분에서 출력 신호에 클릭음 발생
- 2) 측정되지 않은 위치의 HRTF에 대한 보간 생성 문제: 두 지점간의 HRIR 특성이 급격히 변하는 구간은 보간이 별로 의미를 없을 수 있다.

그래서 HRIR을 보간하는 것보다는 각 지점에서 서로 다른 HRTF를 써서 출력음을 생성하고, 그 출력음을 매끄럽게 보간하는 방법을 사용하기로 한다.

2.2.2 가상음장 제어

어떤 실내공간에서 소리가 방사되면 소리는 모든 방향으로 전파되고 벽, 천정, 바닥, 공기 등에 의해 흡수, 반사되어 우리의 귀에 도달하게 된다. 이때 공간의 형태 및 벽면을 구성하는 재료의 흡음률 등에 따라 전파되는 소리는 공간 환경에 의존하는 왜곡 특성을 보이게 되며 이를 통해 청취자는 거리감이나 공간감을 지각하게 된다. 똑같은 악기를 콘서트 홀이나 리스닝 룸 등 각기 다른 장소에서 연주할 경우, 느껴지는 분위기가 다른 것은 바로 이러한 이유에 기인한 것이다. 일반적으로 음장을 통한 거리감과 공간감의 인지에 필요한 중요 단서중의 하나로 잔향을 들 수 있는데, 이것은 실내에서 발생한 음원의 에너지가 순식간에 없어지지 않고 시간 경과에 따라 점차로 사라지는 물리적인 현상을 일컫는 말로서 그것의 물리량인 잔향 시간은 음장의 분석과 제어에 필요한 지배적인 요소로 알려져 있다. 음파가 실내의 공간 경로를 진행하면서 여러 벽면 등에 부딪칠 때 반사 내지 흡수되는 상황과 소리의 확산 느낌을 제공하는 잔향 효과를 표현하기 위해서 여러 방법이 시도되고 있는데 그림 6은 피드백을 통해 음의 크기와 위상을 변화시키는 빗형 필터(Comb filter)와 잔향의 조밀성을 기하기 위한 전역통과 필터(Allpass filter)를 조합해서 현실감있는 음장을 구현한 시스템을 보여주고 있다. 그림에서 D1~D4, D_a는 각 필터의 시간지연율, G1~G4, G_a는 이득을 그리고 W는 음원과 청취자간의 거리를 모델링하기 위한 가중치값(Weighting factor)을 각각 나타낸다.[6]

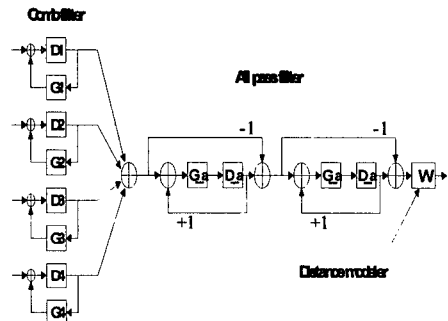


그림 6 가상음장 제어

2.3 동기화된 그래픽/음향 렌더링 시스템

그래픽/음향 렌더링 시스템은 그래픽 영상 내의 움직이는 객체의 위치를 파악하여 그 운동 파라미터를 해석하고 객체의 충돌을 감지함으로써 충돌 위치에 상응하는 음상정위와 그 객체가 속해 있는 공간의 종류에 상응하는 음장감을 생성하는 입체음향을 그래픽과 동기화시켜 렌더링 하는 시스템이다.

이를 위하여 단일 객체의 움직임에 대한 그래픽 영상에 대하여, 그래픽 엔진에서 Event Handler에 전달하는 메시지를 분석하여 관심있는 객체를 결정하고 객체의 위치정보를 그래픽 엔진에 요구한다. 그래픽 엔진에서는 Head Tracker Process Unit과 통신하여 실시간 위치정보 및 Cue, Raw, Pitch등의 움직임 파라미터를 수신한다. Head Tracker Process Unit에 내장된 OS가 적절한 Time Resolution으로 Serial Port에 관련 파라미터를 송신하는 역할을 한다. 그리고 결정된 일련의 위치정보를 이용하여 객체에 대한 운동해석 및 동역학적 해석을 통해 객체의 운동을 정의하고 입체음향 생성에 필요한 움직임 파라미터를 추출한다. 그 후 객체의 실시간 위치 정보를 이용하여 입체음상 정위 기술에 의한 자연스러운 이동음을 렌더링하여 HMD 상에 디스플레이 하면 객체가 실제로 이동하는 듯한 느낌을 가질 수 있게 된다. 또한 Cue, Raw, Pitch등의 정보를 이용하여 회전하는 듯한 느낌도 가질 수 있게 할 수 있다. 그리고 그래픽 지리 정보(동굴, 사막, 바다 등)를 인덱스화하여 HMD 상의 위치정보와 비교하여 공간의 종류를 결정하고 그 공간에 상응하는 음의 흡수, 반사 등의 현상을 고려한 파라미터를 음장감을 생성하는 가상음장제어 모듈에 전달하게 된다. 가상음장제어 모듈에서는 각종지리정보의 음장제어 파라미터를 DB화 시켜놓고 요구된 음장감을 렌더링하게 된다.

그림 7의 그래픽/음향 렌더링 시스템은 파악된 운동 파라미터를 해석하여 충돌 위치와 주변 환경에 상응하는 그래픽을 Head Mouned Display에 디스플레이한 화면이다. 그림에서 보듯이 사막 위를 비행기가 비행하고 있는 환경에서 시연자의 머리움직임에 맞춰 비행기의 위치 및 회전 방향이 바뀌며 또한 비행기의 소음도 입체음향처리되어 위치감을 느낄 수 있는 시스템이다.

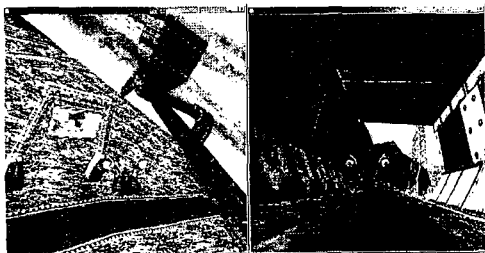


그림 7 그래픽/음향 렌더링 시스템

3. 결론

본 논문에서는 실감음향인터페이스 기술을 정의하고 청취자의 머리움직임 운동을 해석하여 그 음향을 동기화된 그래픽화면과 같이 렌더링하는 시스템에 관하여 논의하였다.

그래픽 및 음향을 제어하여 입체적으로 실감나게 매핑 및 동기화 하는 기술은 시청각분야의 핵심적인 부분으로, 오디오, 비디오, 게임 등 다양한 분야의 산업 제품에 고급화, 고부가가치화를 가져오며 나아가 실감 시청각 정보를 처리할 수 있는 생활 환경 조성에 기반적 역할을 함으로써 삶의 질 향상에 기여할 것이다.

참고문헌

- [1]F. Alton Everest, " The Master handbook of Acoustics" 3rd Edition, Division of McGraw-Hill, Inc.,1994.
- [2]Durand R. Begault," 3D SOUND", Academic Press, Inc., 1994.
- [3]WorldToolKit Reference Manual, Sense8, Corp., 1998.
- [4]<http://www-pablo.cs.uiuc.edu/>
- [5]<http://www.cre.com/>
- [6]Myung Hyun, "The Development of 3D Sound Editor "SoriWave" for Multimedia Contents", AES, 1999.
- [7]강 성훈, 강 경욱, " 입체음향(Spatial Audio)", 기전 연구사, 1997.
- [8]강 성훈, " 방송 음향(Broadcast Sound Engineering)", 기전 연구사, 1997.
- [9] 서울대학교 뉴미디어통신공동연구소, "2채널 스피커에 의한 3차원 음상정위 알고리즘 개발에 관한 연구", 연구보고서, 1997.
- [10] 임정빈, 강성훈, 김천덕, "자연스러운 3차원 음장의 연속 재생을 위한 머리 전달 함수의 보간", 한국음향학회 하계학술발표대회 논문집, pp. 163-166. 1997. 7.