

VR환경에 알맞은 실시간 음장구현에 관한 연구

*채수복, *방승범, *황 신, **고희동, *김순협

*광운대학교 컴퓨터공학과, **한국과학기술연구원 영상미디어연구센터

A Study on Real-Time 3D Sound Rendering for Virtual Reality Environment

*SooBok Chae, *SeungBeum Bhang, *Hwang Shin, **HeeDong Ko, *SoonHyob Kim

*Kwangwoon University, Seoul Korea **Korea Research Institute of Standards and Science

E-mail : luck37@shinbiro.com, kimsh@daisy.gwu.ac.kr

Abstract

본 논문은 VR시스템에 사용되는 실시간 음향제시를 위한 시스템 구현에 관한 것이다. 2개의 Speaker 또는 헤드폰을 사용하여 음상제어, 음장제어의 두 부분으로 구성되어 있다. 음상제어 부분은 각각의 음원의 위치를 정위하고, 음장제어 부분은 레이 트레이싱(Ray Tracing) 기법을 이용하여 음장을 시뮬레이션하고 가상 공간의 음장 파라미터를 추출하여 음원에 적용하면서 실시간으로 음장효과를 렌더링 한다. 이 시스템은 펜티엄-II 333MHz 시스템에서 구현하였다. 최종적으로 청취자는 2개의 스피커 또는 헤드폰을 이용하여 3D음장을 경험하게 된다.

1. 서론

현재의 가상현실 시스템에 관한 연구는 시각적인 면에 많은 발전이 있었다. 이에 비해 여기에 사용되는 음향제시에 관한 연구는 상대적으로 미흡한 실정이다. 현재의 가상현실 시스템의 대부분은 사람의 시각적인 정보, 즉 전방의 일정 영역에 치우쳐있다. 시스템의 완성도 측면에서 본다면 음향 효과적인 면은 시각적인 면에 비해 상대적으로 청취자에게 미치는 영향이 작다고 할 수도 있다. 그러나 보다 효과적인 가상현실 제시를 위해서 청각 시스템을 강화하면 몰입감 및 현실감을 더하고 보이지 않는 곳의 정보를 획득할 수 있다는 큰 장점이 있다.

본 연구에서 사용된 음장 모델링 기법에 관한 연구는 이미 오래 전부터 축소 모형을 이용하거나 컴퓨터 시뮬레이션을 사용한 기하학적 모델링 기법을 사용한 연구가 진행되어져 왔다. 이 시스템에서는 기하학적 모델링 기법 중 레이 트레이싱(Ray tracing) 기법을 사용하여 가상공간의 음장파라미터를 추출한 뒤, Dry Source에 Avatar의 공간위치 정보를 가미한 음장파라미터를 적용하여 실시간 렌더링으로 시스템을 구현하였다.[1]

2. 음장 시뮬레이션

본 논문에서 음장 시뮬레이션은 Auralization Package를 사용하였다. 이 프로그램은 실제환경 또는 원하는 가상공간에 관한 CAD 데이터를 이용, 주어진 실제공간에 대해 Room의 특성을 규정한다. 여기에는 실 공간과 음향적 특성이 같도록 벽체 및 마감재의 흡음률, 각 재료에 따른 흡음계수, 천장, 바닥판, 벽체 등의 반사율, 음원의 위치, 음원의 방향성, 음원의 성질, Receiver의 수와 위치, Receiver의 방향성 등을 고려하여 레이 트레이싱(Ray tracing)기법을 사용하여 공간에 대한 음향특성에 관한 계산을 하게 된다.[5]

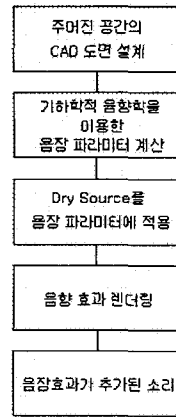


그림 1. 컴퓨터 시뮬레이션을 이용한 가청화 단계의 블록 다이어그램

2.1 레이 트레이싱(Ray tracing)

레이 트레이싱 방법은 음향학적 공간 안에서 음파의 전파를 기하학적으로 표현하기 위해 쓰이는 방법 가운데 하나이다. 어떤 공간의 음향학적 임펄스 응답을 찾기 위해 무한히 작은 크기의 다수의 음선(sound ray) "입자"가 시간 t=0에서 공간 안의 무지향성 음원에서 발사된다. 이러한 입자들은 무지향성으로 균등하게 발사된

다. $t=0$ 이후에 발사는 멈춘다. 이 발사 시간이 기준점(reference point)이다. 각각의 음선 입자는 일정한 에너지를 가지며 각각의 반사 사이에서 직선으로 이동한다.[3]

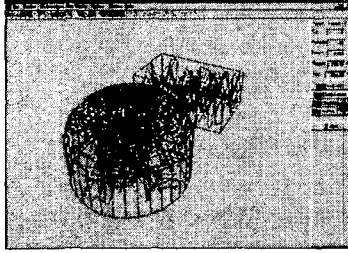


그림 2. 레이 트레이싱 기법을 이용한 시뮬레이션

2.2 시뮬레이션 환경구성

주어진 환경을 시뮬레이션하기 위하여 CAD 데이터를 작성한다. 이 CAD데이터는 실제 환경의 축척을 그대로 사용하여 음장 시뮬레이션 시 실제환경과 거의 유사한 환경을 만들어야 한다.

실시간으로 현실감 있는 음장을 적용시키기 위해서는 적절한 셀의 분할이 중요한 역할을 차지한다. 셀 크기를 정하는 하나의 기준은 초기 반사음 시간이 큰 차이가 나지 않는 가장 큰 셀의 크기를 찾는 것이다. 컴퓨터 시뮬레이션 프로그램을 이용하여 주어진 공간을 모델링한 후 셀 단위로 분할하여 각 셀별로 잔향을 측정 계산하는 방법을 사용하였다. 이러한 방법을 사용할 때 0.5X0.5(m)에서 5X5(m)까지의 셀을 생각 할 수 있다. 만약 0.5X0.5(m)에서 측정된 EDT와 5X5(m)에서 측정된 EDT와의 차이가 거의 없다면 굳이 0.5X0.5(m)셀을 사용할 필요가 없다.

이러한 이유로 육면체의 실내, 회의실 정도 크기의 공간, 대형극장 3가지의 공간을 0.5m 단위로 셀 크기를 조정해가며 비교 실험을 하였다. 실내공간과 회의실 같은 공간은 체적이 작기 때문에 EDT를 측정하는데 있어서나 컴퓨터로 시뮬레이션을 할 때 큰 문제가 없었으나, 대형 극장 같은 경우에는 시뮬레이션을 할 때 셀의 개수가 크게 증가함으로써 계산시간이 너무 오래 걸린다는 점 때문에 1.5X1.5(m)부터 시뮬레이션을 실시하였다. 시뮬레이션 결과 실내공간이나 회의실같이 비교적 체적이 작은 공간에서는 1~1.5m의 크기의 셀이, 대형 극장과 같이 큰공간의 경우에는 2.5~3m 정도의 크기를 가지는 셀이면 작은 셀로 나눈 것과 비슷한 결과를 얻을 수 있었다. 그 이상으로 셀 크기가 커지면 EDT의 분포의 차이가 작은 셀로 나누어 시뮬레이션 했을 경우와 오차가 커지는 것을 발견할 수 있었다.

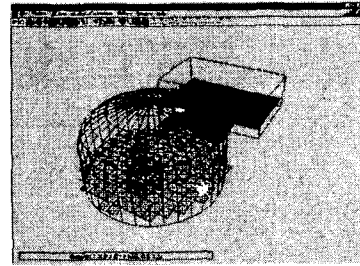


그림 3. 주어진 공간을 셀 단위로 분할한 예

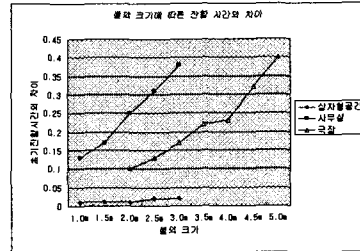


그림 4. 셀 크기에 따른 잔향시간의 차이

2.3 음장 파라미터 추출

실내의 음장을 묘사하기 위해서는 건축음향, 전기음향 등에서 사용되는 실내의 음장을 표현하는 파라미터를 사용한다. 주로 실내의 잔향에 관계된 파라미터를 사용한다. 음장 계산에 의해 추출된 파라미터는 다음과 같다.

2.3.1 EDT(Early Decay Time) 및 잔향시간

EDT(Early Decay Time)는 음원정지 후 10dB 감쇠될 때까지 잔향시간을 나타내는 척도로 임의의 공간의 모양을 묘사하는 단서가 되는 파라미터이다. 또한 RT (Reverberation Time)보다 주관적인 울림과 더 잘 일치한다. EDT는 완전 확산음장에서는 RT와 일치하지만 일반적으로 RT보다 적은 값으로 나타난다.[4][6]

실측정 시에는 실내의 음압 감쇠함수의 앙상블 평균(ensemble average)이 측정점에 있어서 임펄스 응답의 적분 함수로 표현되는 것에 기초한 것이다. 음원 정지 후에 시각 t 에 있어서 잔향 과정의 음압 $h(t)$ 는 대역 잡음의 진폭을 $n(t)$, 측정계를 포함한 임펄스 응답을 $g(t)$ 라고 하면,

$$h(t) = \int_{-\infty}^0 n(\tau)g(t-\tau)d\tau$$

$$\langle h^2(t) \rangle = N \cdot \int_{-\infty}^0 g^2(t-\tau)d\tau = N \cdot \int_0^{\infty} g^2(\tau)d\tau$$

여기에서 $\langle n(\tau)n(\theta) \rangle = N\delta(\theta-\tau)$ 이고, N 은 음원의 파워를 나타낸다. 즉, 임펄스 응답의 제곱에 대한 시간 적분이 잔향 곡선의 각 시점에서 음압 제곱치의 평균

집합이 된다.[8]

2.3.2 직접음과 잔향음과의 비

Intensity는 음원과 청취자와의 거리를 지각하는 가장 기본적인 단서로 룸 모델링 프로그램을 이용하여 직접음과 간접음 경로를 통해 청취자에 도달하는 전체 사운드 Intensity를 고려할 수 있다. 잔향의 유무에 따라, 음압차는 음원과 청취자와의 거리가 같은 경우 3dB 정도의 차이를 보인다. 잔향이 존재하는 경우, R/D (Reverberant-to-DirectSound) ratio의 변화가 Intensity 스케일링 보다 거리에 대한 더 유력한 단서로 알려져 있다.[4][8]

2.3.3 잔향 필터

추출된 파라미터를 실시간으로 가상현실 시뮬레이터에 적용키 위해서는 이것들의 적절한 근사화가 필요하다. PC 환경에서의 실시간 구동을 위하여 인공 잔향기를 사용하여 잔향을 묘사하였다. 인공잔향기는 IIR 필터와 Comb 필터를 사용한 잔향기를 사용한다. 이것에 추출된 파라미터를 적용하여 실시간으로 출력한다. 각 Comb 필터는 임펄스 응답의 초기 잔향을 처리하는 역할을 하며 첫번째 Allpass 필터는 IIR 필터로 구성되어 초기 잔향 뒤의 후기 잔향(Late Reverberation) 또는 확산 잔향(Diffuse Reverberation)을 처리하는 역할을 한다. 마지막에 연결된 Allpass 필터는 이렇게 처리된 소리를 특정 주파수의 변질 없이 출력시키기 위해 사용한다.[6][7]

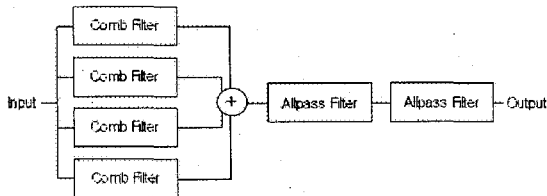


그림 5. Comb 필터로 구현된 잔향기

2.4 Avatar와 음원의 위치 및 방향 제어

본 프로그램에서 Avatar와 음원은 3차원 공간 내에 위치하게 된다. 먼저 Avatar의 경우를 살펴보면, 프로그램 내의 3차원 좌표는 절대 좌표로 설정되어 있고, 좌우는 x축, 상하는 y축, 전후는 z축으로 설정되어 있다. Avatar는 하나의 Object로 좌표계 내에서는 하나의 Point로 인식되어진다. 그리하여 Avatar의 움직임은 좌표계 내에서 좌표점 위치가 변하는 상태로 인식되어진다. 하지만 이 상태에서는 Avatar 자체가 하나의 Point일뿐 무지향성이기 때문에 여기에 3D 벡터를 사용하여 Avatar에 방향성

을 줄 수 있다. 좌우 h, 상하 p, 전후는 r값을 사용하여 좌표계의 어느 방향으로 든 지 방향을 줄 수 있다. 다음으로 음원의 경우를 살펴보면 역시 Avatar와 마찬가지로 x, y, z축의 3D 좌표계로 정위시킨다. 음원의 방향성 역시 h, p, r의 3D 벡터를 사용하여 방향을 3차원으로 조정할 수 있다.

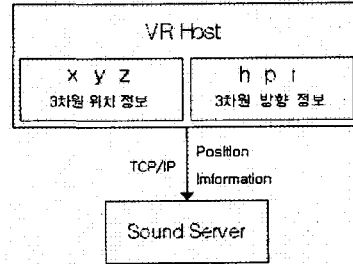


그림 6. 그래픽 서버에서 전달되는 Avatar의 위치정보

2.5 VR 시스템과의 연동

VR시스템과 연동하여 실시간으로 음장을 재현하기 위해 본 프로그램은 TCP/IP 통신 모듈을 사용하여 Avatar의 위치정보를 파악한다.

VR시스템에서는 PC환경의 사운드 서버를 사용하는데, 여기에는 전처리 과정을 통하여 음원의 위치 및 성질이 미리 정의되어 있고 각 셀 마다 음장 파라미터가 테이블로서 작성되어 있다. TCP/IP 통신 모듈을 통해 사용자가 Avatar를 임의의 방향으로 움직일 경우 Avatar가 위치하는 곳에 따라 파라미터 값을 불러들여 위치를 표현하는 음장효과를 실시간으로 렌더링하게 된다.[2]

3. 제안된 시스템

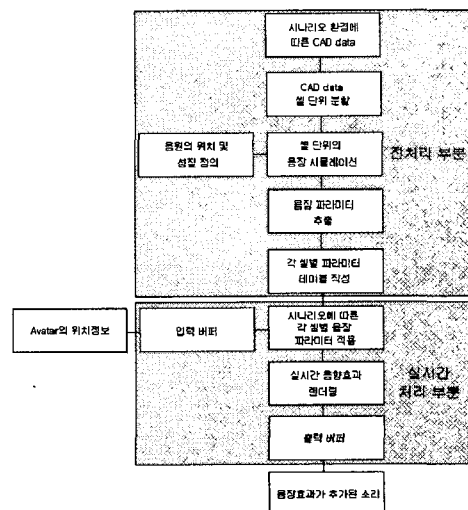


그림 7. 전체 시스템 구성도

제안된 시스템은 MFC를 사용하여 작성하였고, 윈도우 환경에서 실행된다. 이 시스템은 3차원 가상현실 시스템에 음향개선측면에서 실제적인 음장 표현을 위해 사용되는 시스템으로서 현재 개발 중인 VR시스템과 연동하여 실시간으로 정확한 3차원 입체음향을 출력하기 위한 시스템이다. 그래픽을 제공하는 서버와의 통신을 위해 TCP/IP를 사용하여 서버와의 통신을 통해 Avatar의 위치정보를 입력받아서 Avatar의 위치에 따라 파라미터 계수값을 적용시켜 실시간으로 출력한다. 음원은 16비트 44.1kHz로 샘플링 된 모노음원을 사용한다. 20개까지의 음원을 사용할 수 있으며, 각 음원은 전처리 과정을 통하여 음원의 성질을 미리 규정한다.

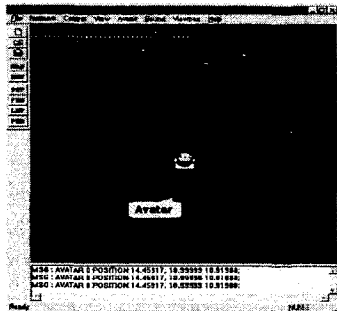


그림 8. Sound Server에서의 프로그램 실행 화면

4. 기존 잔향기와의 음질 평가

잔향기는 잔향시간 또는 1차 초기반사음과 잔향시간을 이용하여 음원에 잔향을 거는 경우가 대부분이다. 가상 현실에서의 음장은 현장감 있고 사실에 가까운 잔향을 재현해야 하므로, 현재의 잔향기를 이용할 경우 문제가 있을 수 있다. 따라서 본 연구에서는 일반적인 상용 잔향기와 본 논문에서 제안된 방법을 이용하여 처리된 잔향을 비교 평가하였다. 실험에 사용한 소리는 악기 소리와 무향실에서 측정된 총소리였으며, 2번씩 청취한 뒤 잔향의 사실성과 듣기 좋은 정도(선호도)의 두 가지를 5단계로 평가하였다. 평가는 높은 숫자일수록 제안된 방법의 사실성이나 선호도가 높은 것으로, 낮을수록 상용 잔향기 쪽이 좋은 것으로 하였다.

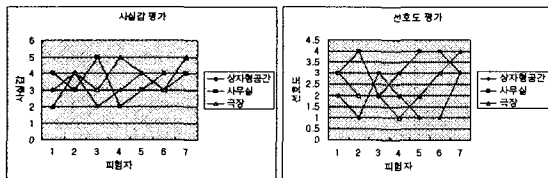


그림 9. 사실감 평가 그림 10. 선호도 평가

결과를 보면 사무실 등 일반적인 환경에서는 제안된 방법이 사실성과 선호도가 높았다. 그러나 극장 등 잔향이 긴 공간에서는 제안된 방법의 사실성이 높았으나, 선

호도는 상용잔향기에 의한 것이 더 높았다. 이것은 상용 잔향기의 경우 울림이 풍부하고, 잔향의 감소 곡선이 완만하여 더 듣기에 좋은 소리를 재현하기 때문인 것으로 평가된다. 반대로 제안된 잔향처리 방법은 초기 반사음 특성과 잔향 시간이 분명하여 사실감은 높지만 잔향의 감소 곡선이 완만하지 않아 선호도가 낮게 나타난다.

5. 결론 및 추후 계획

본 연구는 가상 현실에서의 음장효과를 시뮬레이션하고 실시간으로 재현하기 위한 방법을 제안하였다.

구현된 프로그램은 하나의 독립적인 음향 제작 프로그램이 아닌 VR시스템으로부터 Avatar의 위치를 입력받은 위치의 입체 음장을 실시간으로 재현하는 프로그램이다. 그럼에도 불구하고 기존의 하드웨어를 사용한 잔향기와의 비교 실험에서 거의 동등한 수준의 음장재현 효과를 가지고 있다는 점과 소프트웨어이기 때문에 비교적 비용이 적게 들고, 각각의 셀워치마다 실시간으로 음장을 구현할 수 있다는 것이 큰 장점이라 할 수 있겠다.

추후에는 지금까지 해왔던 연구를 토대로 실 공간과 가상공간에서의 음장을 비교 평가할 예정이다. 또한 지금의 잔향기 필터를 최적화하고, 현재 적용되어 있는 음장 파라미터에 추가로 다른 파라미터를 적용시킴으로써 보다 현실에 가까운 음장을 재현할 수 있는 프로그램을 구현할 것이다.

6. 참고 문헌

- [1] Thomas Funkhouser, Real-Time Acoustic Modeling for Distributed Virtual Environments.
- [2] Ken Pimentel & Kevin Teixeira, Virtual Reality - though the new looking glass, McGraw - 1993.
- [3] Farina - "Pyramid Tracing vs. Ray Tracing for the simulation of sound propagation in large rooms." - In the volume "Computational Acoustics and its Environmental Applications", pp. 109-116, C.A. Brebbia, 1995.
- [4] Crocker, Handbook of Acoustics, 1998.
- [5] Philip Nelson, Measurements of Transient Response of Rooms and Comparison with Geometrical Acoustic Models, 1998.
- [6] Kuttruff, Room Acoustics, 3rd Edition, 1991.
- [7] 이동우, 김영오, 고대식, 강성훈 실시간 입체 음상 제어 시스템의 구현, 한국 음향 학회 학술지, 1999
- [8] 강성훈, 음향 시스템 이론 및 설계, 기전연구사, 1999.