

장면 및 음원 복잡도 축소에 의한 3차원 사운드 재현의 실시간화 기법

성숙정⁰, 이정선, 오수진, 남양희
이화여자대학교 대학원 디지털미디어학부

monica810@ewhain.net⁰, vivarena@hanmail.net, bluerhino@hanmail.net, yanghee@ewha.ac.kr

The Realtime method of 3D Sound Rendering for Virtual Reality : Complexity Reduction of Scene and Sound Sources

SukJeong Seong⁰, JeongSeon Yi, SuJin Oh, YangHee Nam
Division of Digital Media, The Graduate School of Ewha Womans Univ.

요 약

실감 재현이 중요한 가상현실 응용에서는 사용자에게 고급 그래픽 환경을 제시하고 사용자의 인터랙션에 즉각적인 피드백을 제공함으로써 실제감과 몰입감을 증대시키는 연구가 진행되어왔다. 실제감, 공간감 전달을 위해 시각과 청각을 함께 활용하는 것이 효과적이거나, 가상공간의 특징을 반영한 3차원 사운드 재현 연구는 국내외 통틀어 초기단계에 머물러 있다. 실제감과 공간감을 반영한 3차원 사운드의 재현을 위해서는 음원의 전파, 반사, 잔향 등의 계산이 사용자의 인터랙션에 따라 새롭게 계산되어야한다. 그러나 사운드 전파경로와 공간을 이루는 모든 폴리곤들과의 충돌을 검사하며 반사 등을 계산하는 것은 실시간성이 중요한 가상현실응용에서는 무리가 따르므로 실시간성을 보장하기 위한 계산량 축소가 요구된다. 본 논문에서는 다수의 음원이 존재하는 복잡한 가상공간에서의 3차원 사운드를 재현하기 위하여 사운드 전파 계산에 필요한 최소한의 정보를 가지는 오디오 원 그래프로 공간을 재구성하고 다수의 음원을 대상으로 음원 축소 및 군집화를 적용하여 3차원 사운드효과를 실시간으로 재현하는 알고리즘을 제안한다.

1. 서론

실감 재현이 중요한 가상현실 응용에서는 사용자에게 고급 그래픽 환경을 제시하고 사용자의 인터랙션에 즉각적인 피드백을 제공함으로써 실제감과 몰입감을 증가시키는 연구가 진행되어왔다. 실제감, 공간감 전달을 위해 시각과 청각을 함께 활용하는 것은 효과적이거나 가상현실 응용에서는 그래픽 분야의 속도와 환경개선에 그 초점이 맞춰져왔으며, 3차원 사운드 재현의 경우는 막대한 실시간 계산량의 증대, 범용 하드웨어의 비버 수 제한 등의 이유로 지금까지는 많은 연구가 이루어지지 않았다. 실시간 인터랙션에 따라 청취자에게 전달되는 사운드는 음원과 청취자의 위치 관계에 따라 달라지므로 매번 새롭게 계산되어야 한다. 또한 공간감 등의 3차원 사운드 재현에 필요한 전파, 반사, 잔향 등의 요소들 역시 공간과 청취자, 음원의 관계에 따라 계산되어야 한다. 이처럼 3차원 사운드의 실시간 재현만으로도 많은 계산량이 요구 되고 있으나 가상현실 응용에서는 3차원 사운드 재현 외에도 그래픽 렌더링, 사용자 인터랙션 등의 다른 요구들도 동시에 처리해야 함을 고려하면 계산량은 더욱 증가되므로 실시간 계산량을 최소화 하는 것이 필요하다. 본 논문에서는 이와 같은 계산량 축소를 위해 사운드 전파에 중요한 구멍(Hole&Gap)을 보존하면서 사운드 전파에 필요한 최소의 정보를 가지는 오디오 원 그래프를 이용하여 그래픽 정보를 단순하게 표현하고 음원의 대상축소와 군집화를 함께 적용한 알고리즘을 제안한다.

2. 관련연구

게임, 가상현실응용에 적용 가능한 3차원 사운드 재현을 위한 다각적인 연구가 시도되어왔으나 대부분의 연구들은

실시간 계산량의 문제를 해결하기보다 많은 제약조건을 둔 부분적 연구이거나, 특별한 DSP 칩, 고가의 잔향제조기 등을 사용한 연구였다[1][2][3]. 3차원 사운드를 재현하기 위한 실시간 계산량을 줄이는 연구들도 시도되었는데 Funkhouser는 공간을 분할하여 사운드가 공간의 경계면에 도달 할 확률이 높은 반사경로의 그래프를 전처리를 통해 구성함으로써 실시간 계산을 간편하게 하였다. 하지만 이는 움직이는 음원, 인터랙티브하게 변화하는 공간에는 비효율적인 단점을 가지고 있다[4]. 또한 사운드 전파 측면에서 그래픽을 단순화하여 계산량을 줄이는 Chris의 연구에서는 기존 그래픽 단순화 방법에서 구분 하지 못했던 사운드 전파특성에 중요한 구멍(Hole & Gap)을 보존하면서 복잡한 폴리곤 모델을 단순화 하는 방법을 사용하고 있으나 실제적인 사운드 재현에 대한 연구는 반영되지 않았으며 음원의 수가 증가하게 되면 현재 범용되는 오디오 하드웨어에서의 채널 수 제한으로 인해 사운드를 계산하여 재생하지 못한다는 한계를 가지고 있다[5]. 다수의 움직이는 음원을 대상으로 음원의 위치와 크기를 반영하여 음원을 군집화 하는 Nicolas의 연구에서는 청취자의 위치와 방향, 음원의 위치의 방향, 음의 크기를 고려하였으나 7만개 이하의(반사계산시에는 300여개) 폴리곤으로 이루어진 환경을 대상으로 하는 등 공간 복잡도에 제한을 두고 있었다[6].

3. 가상환경에서의 실시간 3차원 사운드 재현

복잡한 가상환경에서 다수의 음원을 대상으로 실시간 3차원 사운드를 재현하기 위하여 본 논문에서는 오디오 계산에 필요한 공간정보의 축약과 대상 음원의 축소 방법을 동시에 사용하는 방법을 제안한다.

본 연구는 대학 IT연구센터 육성 지원사업의 연구결과로 수행되었음

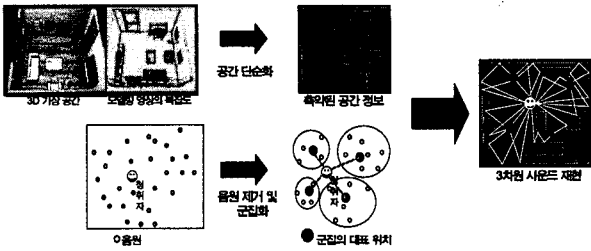


그림 1. 장면 및 음원 복잡도 축소를 통한 실시간 3차원 사운드 재현

3장에서 가상공간을 구성하는 오브젝트의 정보들을 사운드 전파에 필요한 정보들로 최소화하고 오브젝트 재질에 따른 흡음률을 함께 표현할 수 있는 오디오 씬 그래프 구조로 가상공간을 재구성하는 그래픽 단순화와, 음원의 수가 많을 때 음원을 대상으로 청취자의 인지에 영향을 끼치지 않는 유사음원들을 제거하고 음원을 군집화 하는 방법을 설명한다.

3.1 그래픽 단순화

가상공간을 구성하는 수십만 개의 폴리곤을 대상으로 사운드 전파와의 충돌, 반사 등의 계산을 수행한다면 실시간 가상현실용에는 무리가 있으므로 오브젝트들의 특징을 최대한 반영하면서 많은 폴리곤을 단순화하는 방법이 필요하다. 그림 2의 (a)처럼 오브젝트 전체를 포함하는 바운딩 박스의 형태로 표현하는 하는 방법이 가장 기본이 되지만, 사람이 인터랙션 하기에는 작은 공간이라도 사운드는 통과할 수 있기 때문에 사운드 전파에서는 오브젝트내의 빈 공간(Gap & Hole)을 보존하는 것이 매우 중요하다(그림 2의 (b)).



그림 2. 오브젝트의 바운딩 박스의 예[5]

본 논문에서는 사운드 전파 특성을 반영하여 오브젝트 내의 빈 공간(Gap&Hole)을 보존하면서 몇 개의 바운딩 박스의 집합으로 축약하는 Chris의 방법을 이용하여 복잡한 오브젝트를 단순화 한다[5]. 이렇게 단순화된 오브젝트 정보와 사용자가 정의하는 오브젝트의 재질에 따른 흡음률을 가진 공간구조 관계를 표현한 오디오 씬 그래프를 구성한다.

3.1.1 오디오 씬 그래프(Audio Scene Graph)

게임이나 가상현실과 같은 컴퓨터그래픽 분야에서는 그래픽 렌더링 효율을 높이기 위해서 씬 그래프(Scene Graph)라는 구조를 사용한다. 오브젝트의 공간정보들로 표현된 씬 그래프를 이용하여 현재 사용자의 시야에서 볼 수 있는 오브젝트만을 렌더링 하는데 사용한다. 오디오 씬 그래프는 기존의 그래픽 씬 그래프와 흡사하지만 사운드 재현에 필요한 최소의 정보를 표현한다는 측면에서 차이가 있다. 3차원 사운드 재현 계산은 음원의 전달 경로에서 공간을 구성하는 오브젝트와의 충돌, 전달매개체의 흡음률에 따라 달라진다.

오디오 씬 그래프는 그림 3과 같이 전체적인 가상공간을 Part A, B처럼 분할된 공간별로 구성한다. 이 단위공간은 3차원 그래픽분야에서 사용되는 BSP(Binary Space partitioning)를 이용하여 얻는다[7]. 또한 분할된 하나의 공간(Part A)은 공간정보와 공간에 속하는 오브젝트의 정보들(위치, 오브젝트 재질에 따른 흡음률, 오브젝트를 표현하는 바운딩 박스의 집합)을 표현한다(그림3의 A1 obj). 이를 위해서 가상공간을 분

할하여 공간의 위치관계를 구조화할 뿐 아니라 분할된 공간에 포함된 오브젝트 특성을 표현할 수 있는 오디오 씬 그래프 구조를 제안한다.

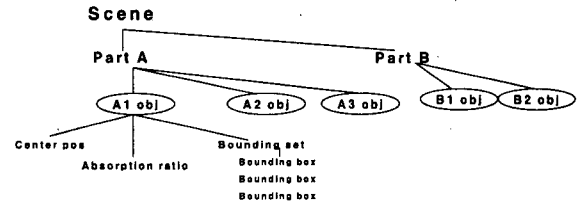


그림 3. 오디오 씬 그래프

진처리로 수행되는. 그래픽 단순화단계의 결과로 표현되는 오디오 씬 그래프 구조는 이후 실시간으로 변화되는 인터랙션에 의해 변화하는 사용자의 특성과 음원간의 관계에 따라 분할된 공간을 대상으로 초기 반사를 계산 할 수 있을 뿐 아니라 바운딩 박스로 표현된 오브젝트 정보와 재질별 흡음률을 이용한 공간을 구성하는 오브젝트와 음원 사이의 반사까지 계산 가능하다. 오디오 씬 그래프는 공간을 단순화 하여 재구성하고 사운드전파에 필요한 최소의 정보만을 표현하여 모든 오브젝트 정보와의 3차원 사운드 계산에 비하여 실시간 계산량을 축소하여 실시간 사운드 렌더링에 효과적으로 사용 가능하다. 본 논문에서는 초기반사를 적용하기 위하여 이미지 스텝 방법을 적용하고[8], 잔향효과는 오디오 API에서 정의한 함수를 이용한다[9].

3.2 음원군집화

3차원 사운드 재현계산(반사, 잔향 등)은 앞서 언급 한 것과 같이 실시간 계산량을 증가시킨다. 그러나 가상공간에 존재하는 모든 사운드를 재생하더라도 사용자는 모든 음원을 지각하지는 못한다. '작은 소리는 큰소리에 묻혀 인지되지 않는다.'는 음향학 마스킹 효과에 기반 하면 '현재 사용자의 위치를 기준으로 일정거리 이상의 먼 곳에 위치하기 때문에 작은 크기로 들리는 음원, 위치가 가까운 음원이라 하더라도 음 세기가 작은 음원들은 사용자에게 인지되지 않는다. 이를 이용하여 재생되어도 사용자가 인지하지 못하는 음원은 재생 대상에서 제외하여 일차적으로 실시간 계산량을 축소한다. 또한 사용자의 현재 위치를 기준으로 하여 음원의 특징(크기, 방향, 위치)의 유사도에 따라 음원을 군집화하고 하나의 군집에 포함된 음원들의 특성을 반영한 군집의 대표 값(대표위치, 방향, 음원)을 구한다. 이를 이후 반사등의 계산에서 하나의 음원처럼 사용하여 실시간 계산량을 줄이는 Nicolas의 알고리즘을 사용한다[6].

본 연구에서는 44100Hz로 샘플링된 음원으로 사운드 데이터베이스를 구축하고, 이들의 실시간 계산을 위한 전처리 단계로 각 음원의 음압 및 음의 저하비율을 이용한 잡음 정도를 계산한다[7].

청취자 인식 측면에서의 음원 대상 축소방법은 가상환경에 포함된 모든 음압의 합과 잡음 정도를 이용하여 청취자가 인식 할 수 있는 음의 크기인 마스킹 임계값을 구한다. 현재 청취자의 위치와 방향을 기준으로 하여 청취자에게 인식되는 음원의 세기 순으로 음원과 마스킹 임계값을 비교하면서 임계값을 만족하지 못하는 음원들을 재생대상에서 제외한다.

음원축소단계를 거친 후 남아있는 음원을 대상으로 청취자와 음원 사이의 거리, 방향, 청취자가 인지하는 음의 크기를 기준으로 K-means 방법, K-center방법 등으로 군집화 한다. 하나의 군집에 포함되는 음원들의 특징(크기, 방향, 위치)을 대표하는 군집의 대표 위치 값을 구하고, 군집에 포함된 음원

을 믹싱하여 마치 다채널을 가진 하나의 음원처럼 재생에 이
용한다.

4. 구현 및 결과

가상환경에서의 실시간 3차원 사운드 테스트 베드는 그림
4와 같이 구성 하였다. 실시간 계산을 효과적으로 처리하기
위하여 그래픽엔진과 사운드 엔진을 분리했으며, 그래픽 엔진
에서는 입력되는 가상공간에 대한 그래픽 단순화를 전처리로
수행한다. 그 결과로 만들어진 오디오 씬 그래프가 사운드 엔
진으로 전달되어 사운드 전파계산에 이용된다. 사운드 엔진에
서는 사운드 데이터베이스를 대상으로 음원의 특징 추출을
위한 전처리 단계를 수행하고, 그래픽 엔진에서 전달받은 오
디오 씬 그래프와 실시간으로 입력되는 사용자와 음원의 정
보를 대상으로 음원의 대상 축소 및 군집화 과정을 실시간으
로 계산하여 3차원 사운드를 재현한다.

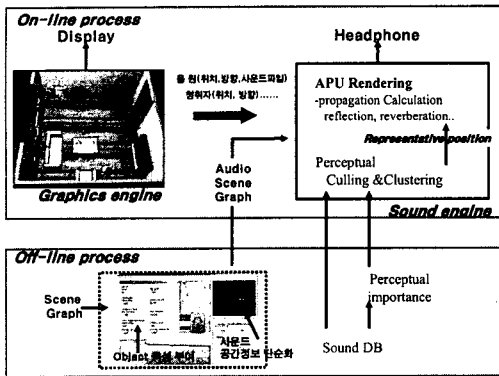


그림4. 시스템 구성도

실험환경은 펜티엄4, 1.7GHz, 512 RAM, ATI Fire GL
8800 . Creative Sound Blaster Live를 사용하였다. 오디오
API는 DirectX 3D Sound를 이용하였다[10].

4.1 그래픽 단순화 결과

본 연구에서는 14만 여개의 폴리곤으로 이루어진 사찰 모
델을 대상으로 그래픽 단순화를 적용하였다. 그 결과 사찰의
지붕, 불상 등의 많은 폴리곤으로 표현된 오브젝트들이 약
530여 개의 바운딩 박스들로 간략화 되었다. 또한 그림 5(c)
에서 열려진 문 등의 빈 공간이 보존되었음을 확인할 수 있
다.

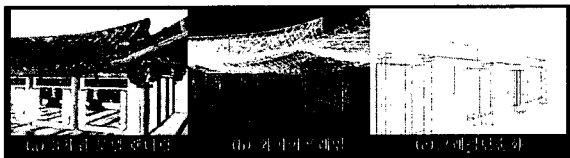


그림 5. 그래픽 단순화 결과

그래픽 단순화의 결과로 재구성된 공간 구조는 오디오 씬
그래프 형식으로 표현되어 이후 음원의 전파계산을 위해 사
용한다.

4.2 음원 단순화 결과

100개 이상의 음원을 대상으로, 실시간으로 입력되는 사용
자의 정보를 반영하여 음원을 재생대상에서 제외하는 단계를
거치면 최소 40% 이상의 음원 축소율을 보인다. 또한 미리
정의한 몇 개의 군집으로 재생대상 음원을 군집화하면 실시

간 계산량은 더욱 감소된다.

표 1에서는 재생대상음원을 100개로 하여 대상 축소를
적용시킨 전후의 초기반사(1차반사) 계산에 소요되는 시간을
비교하였다. 초기반사의 구현은 그래픽 단순화 결과로 표현되
는 오디오 씬 그래프의 바운딩 면과 공간 정보, 음원의 정보
를 이용한 이미지 소스 방법을 적용하였다.

	음원 수/ 군집 수	초기반사 계산시간(ms)
단순화 전	100/4	1068
단순화 후	61/4	661

표 1. 음원단순화 결과

<그래픽 단순화 결과 바운딩 면 538개를 대상으로 적용한 결과임>

5. 결론 및 향후연구

본 논문에서는 다수의 음원이 존재하는 복잡한 가상공간
에서 3차원 사운드를 재현하기 위해서, 공간 단순화를 적용하
고 음원의 군집화를 적용한 축소 방법을 사용하여 실시간 계
산량을 줄이는 방법을 제안 했다. 향후연구로는 다양한 가상
환경에 대한 실험과 오브젝트 재질에 따른 음원의 초기 반사
계산의 확장, 간향효과의 실시간 재현과 입력되는 3D 모델을
대상으로 가상공간을 최적화하는 대표평면을 자동으로 추출
하는 방법을 연구하고, 음원 축소 및 군집화의 적용 전후를 비
교한 사용자 테스트를 통하여 군집화 알고리즘의 신뢰도를
판단하고자 한다.

참고문헌

- [1] 명현, 김기홍의 4명, 입체음향 생성에 있어서 자연스러운 이동음
효과 구현, 정보과학회 논문지: 소프트웨어 및 응용 제 28권
제 10호, 2001.
- [2] Ritta V., User interaction and authoring of 3D sound Scenes in
the carrouso EU project, Audio engineering society, 2003.
- [3] Savioja L., Huopaniemi J., et 2. Real-Time Virtual Audio
Reality, Proceedings of the ICMC, pp. 107-110, 1996.
- [4] Funkhouser T. et 6, A beam tracing method for interactive
architectural acoustics, Acoustical Society of America, 2004.
- [5] Chris J., Nadia T., Significant facet Retrieval for real-time 3D
sound Rendering in complex virtual Environments, VRST,
2003.
- [6] Nicolas T, Emmanuel G, George D, Perceptual audio Rendering
of complex virtual Environments, ACM Siggraph, 2004.
- [7] Ted P., Andreas S., A Review of Algorithms for Perceptual
Coding of Digital Audio Signals, Proc. of International
Conference on Digital Signal Processing(DSP), pp.179-205,
1997.
- [8] Naylor B. F., Construction good partitioning trees, Graphics
Interface 93, 1993.
- [9] Boriah J. Extension of the image model to arbitrary polyhedra,
J. of the Acoustical Society of America 75, 1984.
- [10] Direct Sound 3D, Microsoft Direct X homepage,
<http://www.microsoft.com/windows/directx/default.asp>.