

가상 현실에서의 음장 재현 시스템 구현에 관한 연구

임준희*, 김순협*, 고희동**

*광운대학교 컴퓨터공학과

**한국과학기술원 영상미디어 그룹

A Study on Implementation of the Sound Field Reproduction System for Virtual Reality

Joonhee Yim*, Soon-hoyb Kim, Heedong Ko**

*Dept. of Computer Engineering, Kwangwoon Univ.

*Image and Media Group, KIST

요약

완벽한 가상 현실을 재현하려면 사람의 오감을 실제 세계 수준으로 정교하게 재현해야 한다. 현재 시각 효과는 상당히 발전해 있으나, 청각적, 음향적 부분은 아직도 발전 정도가 상대적으로 낮다. 가상 현실을 위한 음향은 크게 HRTF등을 이용한 음상 정위와, 음장 재현으로 나눌 수 있다. 본 논문에서는 가상 현실 체험자의 이동에 따라 효과적으로 음장을 재현하는 방법에 대해 논하고자 한다. 본 논문에서 제안하는 시스템은 크게 전처리 부분과 실시간 처리 부분으로 구성되어 있으며 전처리 과정에서 음장을 계산하여 실시간 처리 부분에서 체험자의 위치를 추적하여 음장을 재현한다.

1. 서론

완벽한 가상 현실을 재현하려면 사람의 오감을 현실과 구분할 수 없을 만큼 정교하게 재현해야 한다. 사람의 오감에 관한 연구 가운데 가장 앞서 있는 것은 시각에 관한 연구이다. 그리고 청각적 효과인 음향 효과에 관한 연구도 상당히 진전되어 있다. 그러나 그 발전 정도는 시각 효과에 비해 아직 많이 부족한 것이 사실이다. 가상 현실의 음향 효과는 크게 입체 음향 재현과 음장 재현, 두 가지로 구분할 수 있다. 이 둘 가운데 입체 음향에 관한 연구는 현재 상당한 수준에 올라가 있다. 입체 음향을 재현할 수 있는 HRTF를 이용한 방법이 사실상의 표준으로 되고 있다. 그리고 완벽하지는 않아도 청취자와 음원의 위치에 따른 입체 음향 효과를 실시간으로 재현할 수 있는 단계에 와 있다. 그러나 음장 재현은 아직 어떤 묵시적 표준도 존재하지 않으며 실시간 처리를 위한 방법도 아직은 한계가 많다. 현재 임펄스 응답을 이용한 음장 재현은 고정된 음원과 한 명 또는 다수의 청취자 사이의 음장을 재현하는 단계이다[1]. 본 논문에서는 음원이 정지된 상태에서 청취자 또는 가상 현실 체험자가 이동할 때의 음장 변화를 재현하기 위한 방법을

간단히 제시하고 이 방법을 이용하여 구현한 시스템에 관해 논하고자 한다.

2. 기하학적 음향학

현재 많은 음장 시뮬레이션 도구가 기하학적 음향학을 사용하고 있다. 기하학적 음향학은 소리의 전달을 순수하게 에너지의 전달로 생각하는 방법이다[2]. 제안된 시스템에서는 음장 시뮬레이션을 위하여 이미지 모델링과 레이 트레이싱을 사용하였다. 이미지 모델링은 계산이 정확하지만 계산 시간이 오래 걸린다. 레이 트레이싱은 정확도는 조금 떨어지지만 계산 방법이 간단하기 때문에 구현이 쉽다고 계산 시간이 적게 걸리는 장점이 있다 [1].

2.1 이미지 모델링(Image Modeling)

무지향성 음원의 위치 벡터, 반사 표면의 방향 그리고 표면과 좌표계의 원점 사이의 거리를 알면 무지향성 음원의 위치 벡터의 반사 표면에 대한 거울상(mirror image)을 만들어 가상 음원의 위치 벡터를 구할 수 있

다(그림 1)[2].

$$d = b + p \cdot n$$

- 1) b : 원점으로부터 반사 표면까지의 거리
- n : 반사 표면에서 공간 안으로의 법선 벡터
- p : 음원의 위치 벡터

가상 음원의 위치 벡터 r 은 다음과 같다[2].

$$r = p - 2dn \quad 2)$$

2.2 레이 트레이싱(Ray Tracing)

음원에서 발산하는 레이의 단위 위치 벡터와 검출기에 서 추출되는 입자의 에너지는 다음과 같다(그림 2)[2].

$$\begin{aligned} \alpha_{i,j} &= \sin \theta_i \cdot \cos \phi_{i,j} \\ \beta_{i,j} &= \sin \theta_i \cdot \sin \phi_{i,j} \\ \gamma_{i,j} &= \cos \theta_i \end{aligned} \quad 3)$$

$\alpha_{i,j}$: x축 위치 벡터
 $\beta_{i,j}$: y축 위치 벡터
 $\gamma_{i,j}$: z축 위치 벡터

$$e: e_s \frac{(l_2 - l_1)}{R} \exp\left(-m \frac{l_1 + l_2}{2}\right) \quad 4)$$

e_s : 레이 입자가 p_s 에 있을 때의 에너지
 l_1, l_2 : 레이 입자가 검출기와 교차하는 두 점

3. 음장 재현을 위한 시스템 구성

음장 재현을 위해 제안된 시스템은 크게 offline과 online 단계로 구성되어 있다(그림 3). offline 단계에서는 시나리오를 결정하고, 그 시나리오에 따라 공간을 구성하고, 그 공간의 음장을 시뮬레이션하여 임펄스 응답을 구하는 작업을 한다.

3.1 offline 단계

Online 단계에서는 구한 임펄스 응답을 이용하여 입력된 소리에 음장 효과를 추가하여 출력한다. 임펄스 응답을 시뮬레이션하기 위해 레이 트레이싱을 사용하였다. 주어진 공간에서 청취자가 이동할 경우 음원에서 청취자 사이의 음장 또한 변화하게 된다. 이러한 음장의 변화를 재현하기 위해 주어진 공간을 일정한 크기의 격자로 나누어 격자의 중심점에서의 음장을 시뮬레이션 하

여 사용하였다. 이 때 음장의 자연스러운 변화를 위해 적절한 셀의 크기를 시뮬레이션을 통해 결정하였다(그림 4). 이렇게 구한 음장의 임펄스 응답을 저장하여 임펄스 응답 테이블을 구성하고, 이 테이블을 이용하여 음장을 재현한다. 기하학적 음향학은 위상 특성을 고려하지 않아 모두 같은 위상을 가진 임펄스 응답이 나타나게 된다[3]. 컨볼루션을 하여 직접 들어보면 이러한 임펄스 응답을 사용하면 소리가 날카롭게 들렸다. 이러한 단점을 보완하기 위해 임펄스 응답의 각 임펄스에 +1, 0, -1의 세 숫자를 무작위적으로 곱해 위상 특성을 넣어 주었을 때 소리가 더 부드러워 지는 것을 느낄 수 있었다.

3.2 online 단계

online 부분에서 입력 버퍼와 출력 버퍼는 음향 효과를 일으킬 수 있는 어떤 사건(청취자의 이동, 특정 사건에 의한 소리 발생 등)에 대한 응답으로서의 음장 효과 재현이 사람이 느낄 수 있는 한계인 80 - 120ms 이내[4]가 되도록 조절하였다. 컨볼루션 시간을 제외한 입출력에 걸리는 시간은 약 100ms정도이다. 컨볼루션은 입력된 소리와 임펄스 응답을 컨볼루션한다. 이 때 컨볼루션 시간을 줄이기 위해 1-2초 정도 되는 임펄스 응답을 더 작은 단위로 나누어 컨볼루션하였다.

4. 실험 및 결과

실험은 주어진 공간에 따라 음장을 시뮬레이션 하기 위해 사용할 레이 수를 결정하고 주어진 공간을 적절한 크기의 셀로 분할하기 위해 수행하였다. 음장 시뮬레이션에 사용된 소프트웨어는 ODEON 3.1 Combined Edition이다.

4.1 레이 수 설정

주어진 공간의 음장을 시뮬레이션 할 때 사용해야 할 최소 레이 수의 숫자는 다음과 같다[2].

$$N_{\min} = \frac{4(ct_{\max})^2}{r_k^2} \quad 5)$$

N_{\min} : 최소 레이 입자 수
 c : 음속
 t_{\max} : 임펄스 응답의 길이
 r_k^2 : 검출기의 반지름

일반적인 사무실 환경에서 잔향 시간 0.4초[6], 음속 340m/s, 검출기 반지름을 사람 머리의 반지름인 0.17m로 잡을 경우 N_{\min} 은 2.56E+006이다. 그러나 위의 방정식의 결과 값을 음장 시뮬레이션 소프트웨어에 입력해

본 결과 시뮬레이션이 불가능하였다. 그러나 극장이나 복잡한 공간에 100,000개의 레이를 사용하는 경우[5]도 있는 것을 생각하면 위의 수치를 그대로 사용하는 것은 문제가 있다고 할 수 있다. 또한 주어진 공간이 복잡한 경우라도 계산된 N_{min} 값의 1/100 정도인 30,000개의 레이와 1/30정도인 300,000개의 레이를 사용한 시뮬레이션 결과 값이 큰 차이가 없음을 알 수 있었다(그림 4). 시뮬레이션에 사용한 조건은 다음과 같다.

표 1. 시뮬레이션 환경

*1000개 단위로 증가시키며 시뮬레이션
**10000개 단위로 증가시키며 시뮬레이션

레이 수	300 - 30,000* 30,000 - 300,000**
transition 차수	10
임펄스 응답 길이	0.4 - 2초
최대 반사 횟수	2000회
음속	340m/s

실험 결과 방 등의 간단한 공간은 300개, 기둥이 있는 공간 등 조금 더 복잡한 공간은 3,000개, 극장 등 음향학적으로 복잡한 공간은 30,000개의 레이를 사용하면 그 이상의 레이를 사용한 결과와 큰 차이가 없음을 알 수 있었다.

표 2. 사무실 환경의 시뮬레이션 결과

환경		사무실 환경
임펄스 응답 길이		0.4 sec
EDT	300	0.08 sec
	3,000	0.08 sec
	30,000	0.08 sec
잔향시간 (RT_{60})	300	0.5 sec
	3,000	0.5 sec
	30,000	0.5 sec

표 3. 극장의 시뮬레이션 결과

환경		극장 환경
임펄스 응답 길이		2 sec
EDT	300	0.35 sec
	3,000	0.45 sec
	30,000	0.50 sec
	300,000	0.50 sec
잔향시간 (RT_{60})	300	1.75 sec
	3,000	1.75 sec
	30,000	1.75 sec
	300,000	1.75 sec

4.2 셀 크기를 구하기 위한 시뮬레이션

셀 크기는 작게 잡으면 잡을수록 셀 사이의 음장 차이가 작아지는 것이 사실이다. 그러나 셀 크기를 너무 작게 잡을 경우 계산 부하가 높아져 실시간 음장 재현을 하기에 어려운 점이 있다. 따라서 음장 변화를 부드럽게 하면서 계산 부하가 커지지 않는 적절한 타협점을 찾아야 한다. 이 시뮬레이션에서는 계산량을 줄이기 위해 비슷한 결과가 나올 경우 가능한 셀 크기를 크게 잡는 것을 목표로 하였다. 이를 시뮬레이션 하기 위해 주어진 공간을 격자 형태의 셀로 분할하여 각 셀의 EDT를 구하였다(그림 5).

5. 결론 및 향후 계획

본 논문에서는 고정된 음원과 이동하는 청취자 사이의 음장 효과를 재현할 수 있는 시스템을 제안하고 이를 구현하였다. 또한 제안된 시스템에 적절한 음장 시뮬레이션 환경을 구하였다.

그러나 이 시스템으로는 이동하는 음원과 이동하는 청취자 사이에 형성되는 음장을 구하여 실시간으로 재현하는 것이 힘들다는 단점이 있다. 또한 임펄스 응답으로 EDT까지만 구하기 때문에 출력되는 소리의 음색이 gate 이펙터 처리된 것과 비슷하게 바뀌는 단점이 있다. 따라서 앞으로 이를 보완하기 위해 후기 잔향(late reverberation)을 추가하여 음색 변화를 최소화 하는 방법을 구하고 있다. 또한 이동하는 음원과 이동하는 청취자 사이의 음장을 계산하여 재현하는 방법도 연구되어야 할 것이다.

6. 참고 문헌

- [1] Thomas Funkhouser, A Beam Tracing Approach to Acoustic Modeling for Interactive Virtual Environments, SIGGRAPH 98.
- [2] Measurements of Transient Response of Rooms and Comparison with Geometrical Acoustic models, 1998
- [3] Claus Lynge, ODEON Room Acoustics Program Version 3.1, 1998
- [4] Miner & Caudell, Computational Requirements and Synchronization Issues for Virtual Acoustic Display, Presence, Vol. 7, 1998
- [5] Farina, Pyramid Tracing vs. Ray Tracing for the simulation of sound propagation in large rooms.
- [6] JBL Technical Reference.

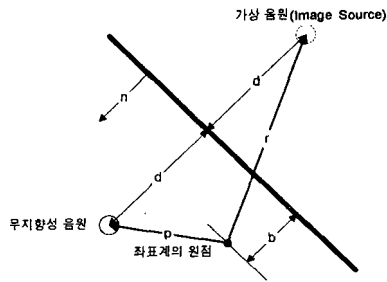


그림 1. 이미지 모델링의 원 음원과 가상 음원

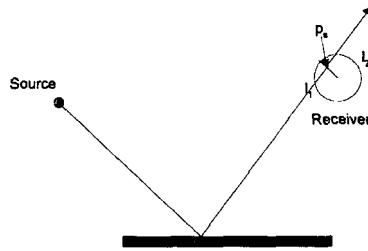


그림 2. 레이 트레이싱 방법

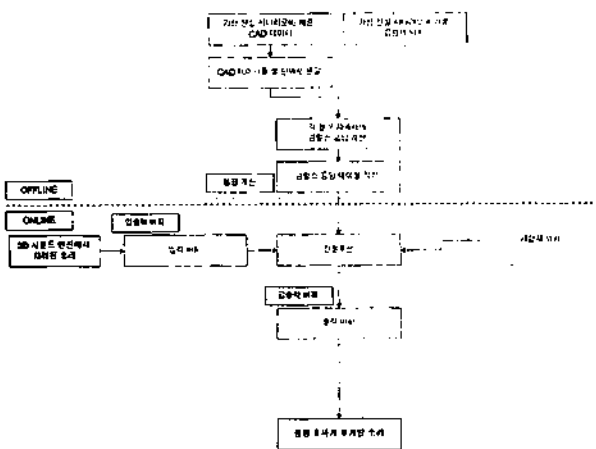


그림 3. 전체 시스템 구성도

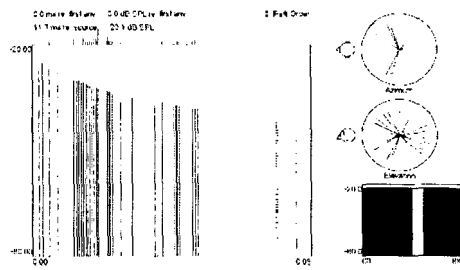


그림 4.a. 300개의 레이를 사용한 에코그램 (사무실)

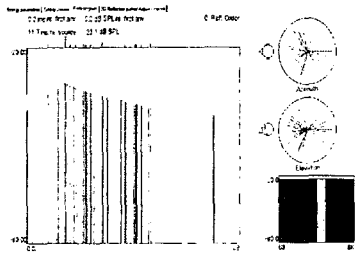


그림 4.b. 30,000개의 레이를 사용한 에코그램(사무실)

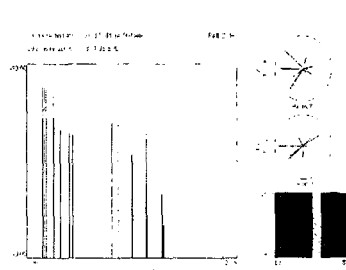


그림 4.c. 300개의 레이를 사용한 에코그램(극장)

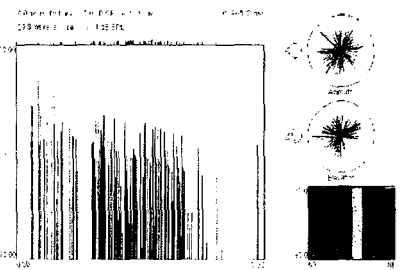


그림 4.d. 30,000개의 레이를 사용한 에코그램(극장)

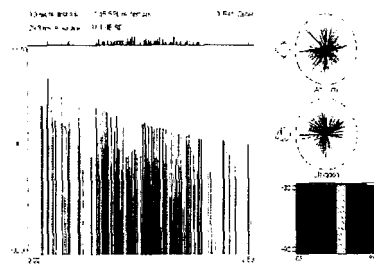


그림 4.e. 300,000개의 레이를 사용한 에코그램(극장)

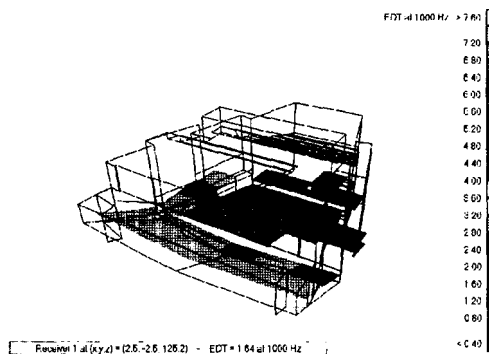


그림 5. 5x5 셀에서의 EDT