

공간음향의 실시간 전달을 위한 시스템 설계

김풍혁, 최덕재

전남대학교 전산학과 인터넷응용기술연구소

dcwork@iat.chonnam.ac.kr, dchoi@chonnam.ac.kr

Design Realtime 3D-Sound Transport System

Pung-Hyeok Kim, Deok-Jai Choi

Dept. of Computer Science, Chonnam National University

요약

초고속 네트워크가 점점 일반화됨에 따라 네트워크를 기반으로 하는 멀티미디어 서비스에 대한 요구도 점점 확대되고 있으며, 동시에 보다 실감 멀티미디어 데이터에 대한 기대 역시 커지고 있다. 입체음향 기술은 음향 부분의 실감성을 향상시키는 방법으로, 네트워크 참여자들의 가상현실 환경 공유 상황에서 체감 품질을 보다 더 향상시키는 역할을 할 수 있다. 본 논문에서는 입체음향 기능들을 네트워크를 통해 전달하고 활용하는데 필요한 요소에 대해 논의하고 그 처리 방법을 제시한다. 아울러 제시한 방법을 적용할 수 있는 활용 분야에 대해서 다룬다.

1. 서론

고성능 네트워크가 일반화되면서 네트워크를 통한 고화질 멀티미디어 데이터 전송은 물론이며 전송하는 멀티미디어 데이터의 실감성을 보다 향상시키고 그 몰입도를 증가시키기 위한 실험들이 활발하다. 네트워크 참여자들의 상호작용을 위한 네트워크 기반의 가상환경 공유는 이러한 데이터를 다루면서 데이터 발생 및 전송 측면에서의 실시간성을 만족시켜야 한다.

본 논문에서는 네트워크로 연결된 참여자들 사이의 가상환경을 공유하는데 있어서, 음향효과의 실감성을 높이기 위해 입체음향 데이터를 네트워크상에서 실시간으로 전달하고 조작하기 위한 방안을 제시한다. 이를 위하여 2절에서는 입체음향을 구현하는 두 갈래의 기술들에 대해 요약 설명한다. 3절에서는 네트워크 기반에서 입체음향을 활용하기 위한 구조와 그 활용분야를 제시하고, 4절에서는 전체적인 결론을 내린다.

2. 배경 기술

입체음향 기능은 크게 멀티채널 방식에 의한 구현과 공간음향 방식에 의한 구현으로 나뉘어진다. 멀티채널 방식은 방송과 같이 주로 저장매체를 기반으로 하는 분야에서, 공간음향 방식은 게임과 같은 실시간 분야에서 널리 활용된다.

멀티채널 오디오는 다수의 트랙에 개별적으로 음원 데이터들을 녹음하고, 재생할 때는 해당 음원 데이터들을 동시에 재생함으로써 입체음향 효과를 얻는 기술이다.

멀티채널 오디오를 저장하는 형식은 구조화 형식과 비구조화 형식으로 나눌 수 있다. 이 가운데 비구조화 형식은 채널 각각을 개별적인 스트림, 혹은 내부적으로 나누어진 샘플들로 구성된 통합 스트림으로 표현하는 방법이다. [1] 그림 1은 비구조화 형식 멀티채널 오디오의

개념을 설명하고 있다.

Digital Theater System (DTS), Audio Coding-3 (AC3), MPEG Advanced Audio Coding (AAC) 등 방송이나 동영상 저장에 관련된 매체에서는 대부분 비구조화 멀티채널 오디오 형식을 표준으로 사용한다. [2] 그림 2는 5.1 채널을 나타내는 멀티채널 오디오의 개념도를 도식한 것이다.

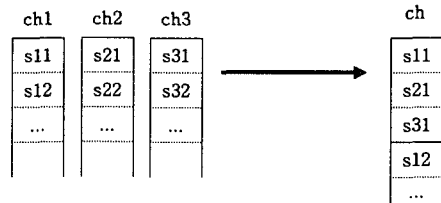


그림 1. 비구조화 멀티채널 오디오 형식.

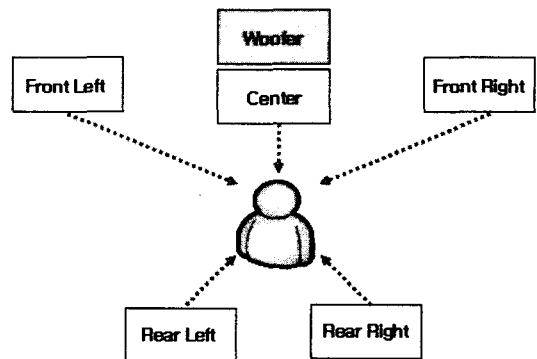


그림 2. 5.1 멀티채널로 표현되는 입체음향.

게임 분야에서는 게임에서 사용될 입체음향들을 미리 만들어 저장할 수 없으므로, 소리 개체 사이의 관계에 따라 사람의 두 귀에 어떤 소리가 들릴 것인지를 실시간으로 렌더링하는 공간음향(3D-Sound) 방식을 사용한다. 이때 두 귀에 들릴 소리가 어떠한 것인지는 머리전달함수(Head Related Transfer Function)에 의해 정해진다. 즉 공간음향 렌더링 엔진의 역할은 음원에 대한 소리를 결정하는 여러 속성들에 따라 소리에 시간차 및 크기변화 등의 효과를 주어 입체감을 느끼도록 하는 것이다.

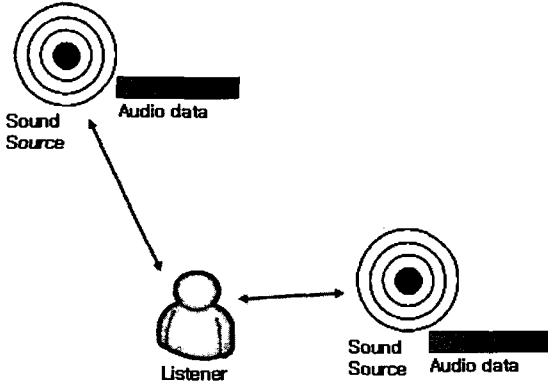


그림 3. 공간음향의 개념도.

그림 3은 공간음향 모델을 도식한 것이다. 음향을 듣기 위해서는 공간상에 소리를 듣는 개체(object)인 청취자(listener)와 소리를 내는 개체인 음원(source)이 정의되어야 한다. 3차원 그래픽에서와 마찬가지로 청취자와 음원 각각은 3차원 공간상에서의 특정 위치가 주어진다. 공간상에서의 위치와 더불어 두 가지 개체들에 여러 가지 부가적 속성들이 정의되어 있으며, 이들 속성들의 관계를 해석함으로써 입체음향을 발생시킨다.

3. 실시간 입체음향 전송 시스템

그림 4는 우리가 제안하고자 하는 실시간 입체음향 전송 시스템의 개념도이다. 음향 데이터에 공간 음향을 구성하기 위한 속성 정보를 적절하게 실어 보냄으로써, 네트워크상에서 입체음향 환경을 실시간으로 공유할 수 있게 된다. 이로써 음향 측면에서 동일한 가상 환경을 다수의 참여자들과 공유하는 한편 전송서버에서는 공유 내용에 필요한 음향 변화를 실시간으로 수행할 수 있게 된다.

입체음향을 구현한 멀티미디어 데이터를 네트워크상에서 활용하는 방안으로써 멀티채널 오디오 형식으로 저장된 동영상 전송하는 방법으로 접근할 수도 있다. 하지만 네트워크상에서 공유하려는 내용이 미리 저장된 것이 아니라 실시간으로 발생하는 것일 경우에는 이 방법은 적절하지 않다. 한편 실시간으로 네트워크 기반으로 입체음향을 활용하는 예는 온라인 게임 분야에서 널리 쓰이는 것이지만, 이들 예는 전적으로 동일한 게임을 하는 환경에서만 유효한 것이다. 즉 동일한 게임에 참여

한다는 것은 참여자들이 필요한 환경을 미리 준비하는 것을 전제로 함을 의미한다. 이는 방송과 같이 방송자와 시청자의 관계처럼 상대방에게 없는 내용을 기반으로 입체음향환경을 공유하려는 경우에는 그대로 적용할 수 없다.

따라서 네트워크 참여자가 시청할 내용에 대한 사전 준비가 없으면서도 동일한 가상 환경을 공유하며, 참여자 각각에게 맞는 차별화된 입체 음향을 제공하기 위해서는 공간음향 기술을 네트워크상에서 이용할 수 있도록 확장하는 방법이 보다 효율적이다.

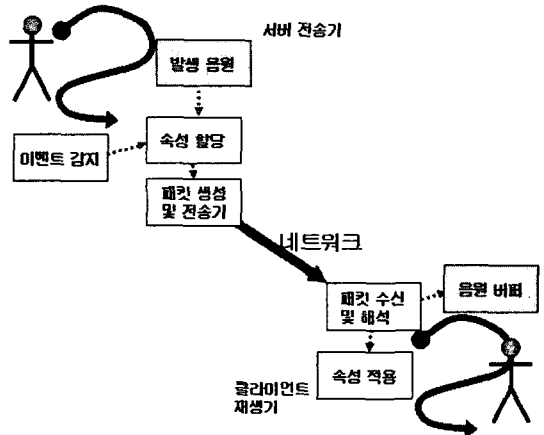


그림 4. 공간음향에 의한 실시간 입체음향 전송.

공간음향 기술은 사용자 측면에서 볼 때 음향데이터 자체를 취급하는 부분과 입체음향 기능을 발휘하는 속성들을 설정하는 부분으로 나뉘어진다. 이를 네트워크 전송 측면에서 생각하면 음향 데이터를 보내는 것과 해당 데이터에 대한 명령을 보내는 것으로 나누어 생각할 수 있다. 이는 RTP (Realtime Transport Protocol) [3] 와 유사하게 보이는 특성이다. 하지만 RTP에서 명령 패킷을 취급하는 RTCP는 RTP로 보내는 데이터 자체에 적용될 내용을 다루는 것이 아니라, 데이터 전송 상태에 대한 통계적인 값을 알리는 부가기능을 수행한다. 공간음향 재생에 필요한 음향 데이터와 그에 대한 속성 조정은 어느 한 쪽만으로 유효하게 사용할 수 있는 정보가 아니라 데이터를 재생하면서 동시에 적용해야 하는 것이다. 따라서 해당 음향 데이터에 대한 속성은 그 데이터를 적재한 패킷과 함께 전송하는 것이 보다 적합하다. 결과적으로 입체음향을 위한 속성들이 음향 데이터에 부가적으로 붙는 헤더처럼 구성된다.

서버 전송기에서 속성 할당기의 역할은 이처럼 해당 음원에 입체음향 효과에 변화를 주는 이벤트가 발생하였을 때, 이벤트 발생기로부터 적절한 데이터를 받아 이에 해당하는 속성 정보를 설정하고 현재 발생한 음향 데이터와 같이 결합시켜 하나의 데이터 패킷을 만드는 역할을 수행하는 것이다. 역으로 수신기 프로그램에서는 받은 패킷으로부터 속성 부분과 음향 데이터 부분을 분리하여 속성 데이터는 환경을 설정하는 명령으로 수행하고

록 만들고, 음향 데이터는 재생용 큐로 분류하여 넣는다. 이 과정은 네트워크 기반 멀티미디어 재생에 관한 여러 가지 사항들과 함께 공간음향 재생 플랫폼에 맞게 동작시킬 수 있어야 한다.

표 1은 DirectX Sound에서 3차원 공간음향을 활용하기 위해 설정해야 하는 주요 속성들이다. [4] 단순히 음원 데이터에 입체적 특성을 표현하기 위한 속성들을 고정헤더 형식으로 전부 포함시키는 방법을 생각할 수도 있으나, 표 1에서 보는 바와 같이 주요 속성만 헤더로 구성해도 적지 않은 크기를 갖게 된다. 특히 샘플링 비율이 낮아지거나 압축 상태의 음향 데이터를 갖는 경우에는 패킷 내에서 이들 속성 헤더가 차지하는 비율이 더욱 커진다. 게다가 변화할 필요가 없는 속성들이 다수 존재할 경우 기억공간을 낭비하는 것은 물론이고 재생하는 쪽에서 속성 적용을 위한 함수 호출을 불필요하게 반복하게 된다.

name	type
position	3d-vector
outside cone angle	double word
outside cone volume	long
inside cone angle	double word
direction	3d-vector
velocity	3d-vector
min distance	3d-vector
max distance	3d-vector

표 1. Source Buffer 에 적용되는 주요 속성.

따라서 제안된 시스템에 원활하게 적용하기 위해서는 음원 데이터에 필요에 따른 가변형 속성 헤더를 갖도록 하고, MPEG 에서의 프레임 구성 형태와 유사한 부가적인 명령 패킷으로 보완하는 구조가 더 효율적이다. 이 구조를 지원하기 위해서 그림 5와 같이 데이터 패킷의 내부에 (1) 현재 패킷이 속성 데이터를 갖는 패킷인가, 음향 데이터만 적재한 패킷인가, (2) 속성 데이터를 가질 경우 속성 데이터 헤더에는 어떠한 값이 포함되어 있는가, (3) 패킷에 적재된 속성 혹은 음향 데이터는 어떤 음원개체(sound source)에 적용되는 것인가 등과 같은 형식들을 식별할 수 있도록 식별번호를 할당하여 이들을 구분한다.

음향 데이터 전용	적용할 음원개체	전송 정보	실제 음향 데이터
속성 패킷 전용	적용할 음원개체	속성 1, 속성 2, ...	전송 정보
속성, 데이터 혼합	적용할 음원개체	속성 1, ...	전송 정보 실제 음향 데이터

그림 5. 전체 데이터 패킷 개념도.

최종 데이터 패킷을 수신하는 수신기는 송신기에서 만든 패킷을 받아 송신기에서 패킷을 만드는 순서의 역방

향으로 적용하여 재생하는 구조를 갖지만 수신기는 패킷의 내용을 해석하는 기능만 수행하므로 송신기에서 이벤트 감지 및 해석에 대응하는 기능은 포함하지 않는다. 한편 수신한 속성, 혹은 음향 패킷의 유효성을 판단하기 위해서 기존 스트리밍 알고리즘 분야에서 적용하는 부가적인 전송 정보가 추가로 필요하다. 이들 전송 정보는 해당 패킷에 대한 음원 개체에 따라 적용 기준이 달라질 수 있으므로 헤더를 구성하는 다른 정보들과 구분하여 적재 및 해석을 수행한다.

4. 결론

본 논문에서는 네트워크상에서 입체음향 환경을 공유하고 제공자가 이를 실시간으로 조작할 수 있는 실시간 공간음향 전송기 및 수신기 구조를 제시하였다. 이것은 멀티채널 오디오를 사용하는 경우보다 실시간 조작성과 다양성을 확보할 수 있다. 이러한 특성은 기존의 주문형 비디오 서비스처럼 미리 준비된 하나의 멀티미디어 데이터를 시청하는 경우보다는 가상현실이나 시청자 참여형 방송, 원격강의와 같이 실시간 조작성 및 수신자들의 참여도가 중시되면서 음향을 통한 채감 만족도를 높이고 싶은 경우에 더욱 강점을 갖는다. 향후 제시한 설계를 더욱 구체화하고 앞서 논의한 활용분야에 실제로 적용, 구현함으로써 제시한 시스템의 효율성을 높이고 실용적인 분야로 발전시킬 예정이다.

참고문헌

- [1] Frode Eika Sandnes, "Efficient Large-scale Multichannel Audio Coding", Euromicro Conference, 2001. Proceedings. 27th 4-6 September, 2001 page:392-399
- [2] 강명수, 김규년, 소리객체에 입체음향 기술을 적용시킨 멀티채널 오디오 형식. 2002년 한국정보처리학회 춘계학술발표논문집 제9권 제1호, 213-216쪽.
- [3] RFC 3550. A Transport Protocol for Real-Time Application.
- [4] MSDN DirectX Sound page. http://msdn.microsoft.com/library/default.asp?url=/library/en-us/directx9_c/dx9_directsound_programming_guide_3d_sounds.asp