

소리객체에 입체음향 기술을 적용시킨 멀티 채널 오디오 형식

강명수, 김규년
울산대학교 컴퓨터 정보통신공학부
e-mail : enonst@yahoo.com

Multichannel surround audio format using 3D sound technology with sound objects

Myung Soo Kang, Kyoo Nyun Kim
School of Information and Communication Technology, University of Ulsan

요 약

멀티 채널 오디오 시스템 및 입체음향 기술이 대중화 되면서 오디오 데이터의 저장 및 재생 방법에 있어서 실감 현장 음향을 고려하는 것이 일반화 되었다. 이 논문에서는 구조화 오디오 데이터 형식 [1]을 이용해 불필요한 데이터 공간을 제거하고, 거기에 입체 음향 기술을 적용시켜 효과적인 오디오 데이터 처리 방법을 제시하고자 한다. 우리는 구조화 오디오 데이터 형식을 개념적으로 명백히 하기 위해 재사용 가능한 "소리 객체"를 사용하였다. 소리 객체는 오디오 데이터를 처리하기 위한 녹음된 소리 샘플의 기본 단위로서 소리객체에 3D 위치 속성이나, 필드 속성을 지정해 소리를 변화시킬 수 있다.

1. 서론

멀티채널 오디오

멀티채널 오디오는 여러 개의 트랙에 따로 녹음된 오디오 데이터를 동시에 여러 개의 채널로 재생하는 기술이다. 이때 각각의 트랙은 녹음된 원음을 스테레오 패닝, 음압 레벨 그리고 룸 시뮬레이션과 같은 필터 효과에 의해 조절한 후 저장한다.¹

멀티채널 오디오 포맷은 저장 형식에 따라 구조화 포맷과 비구조화 포맷으로 구분할 수 있다. 비구조화 포맷은 각각의 채널들을 멀티플렉싱 된 오디오 스트림으로 저장하는 방식이다[1]. Digital Theater System(DTS), Audio Coding-3(AC-3), MPEG Advanced Audio Coding(AAC)등 대부분의 표준화된 오디오 포맷은 비구조화 포맷으로 저장된다. 비구조화 오디오 데

이터 포맷은 일정한 비트 율로 데이터를 저장해야 하기 때문에 소리 데이터가 전혀 없는 구간도 똑 같은 데이터 공간을 요구한다는 면에서 비 효율적이다. 이와는 달리 구조화 데이터 포맷은 소리 데이터가 있는 구간의 오디오 데이터를 저장하고 재생시 필요한 부가적인 매핑 정보를 저장하는 방법이다.

입체음향

임의의 공간에 위치한 음원에서 발생한 소리는 음원의 방향에 따라 일정한 시간차이와 레벨차이를 가지고 청취자의 두 귀에 도달된다. 그리고 소리는 사람의 머리, 몸통, 외의 등에 의해서 회절, 반사, 굴절되어 두 귀에 도달한다. 이와 같이 두 귀에 도달하는 소리의 정보를 머리전달함수(Head Related Transfer Function)라고 한다. 일반적인 입체음향의 정의는 머리 전달 함수를 이용해 원 음장을 풍부하게 재현하고 음의 고저 및 음색, 그리고 방향 감이나 거리감까지도

*본 연구는 학국과학재단 목적기초연구(R01-2000-000283-0)지원으로 수행되었음.

재생하여 음장 감을 가지게 하는 음향으로 정의 된다 [3].

현장음향

멀티채널 오디오 기술이나 입체음향 기술이 목표로 하는 것은 녹음된 소리를 재생하였을 때의 현장감을 높이는 것이다. 여기에는 3 차원 상에서의 음원의 위치와 움직임 그리고 청취환경에 의한 음장효과 등이 포함되어야 한다. 하지만 DTS 나 AC-3 같은 전통적인 멀티채널 오디오 기술에는 이러한 요소들이 충분히 고려되어있지 않다[4]. 이러한 이유로 어떤 면에서는 2 채널 스피커 시스템에 입체음향 기술을 적용시키는 것이 5.1 채널 서라운드 시스템보다 더 효과 적인 현장 음향을 제공한다[4]. 또 멀티 채널 시스템은 다양한 포맷들 간의 상호 호환성(up and down compatibility)의 문제가 있지만 HRTF 를 이용한 입체음향기술은 2 개 이상의 스피커 환경에 적용이 가능하며 호환성의 문제가 발생하지 않는다. 이런 한 이유로 모노 사운드로 녹음된 원음에 입체음향 기술을 적용 시켜 입체음향 효과를 적용시키는 것이 보다 효율적이라 할 수 있다.

2. 소리 객체에 근거한 오디오 시스템

그림 1 은 우리가 제안하고자 하는 오디오 시스템의 대략적인 구성도 이다.

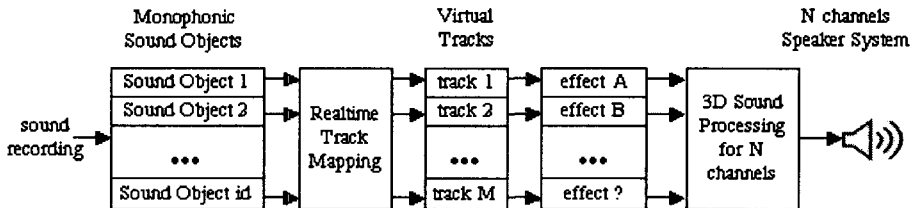


그림 1 시스템 구성도

소리객체(Sound Object)

3D 음향 기술은 음상 및 음장 효과가 부여되지 않은 원음에 입체 음향의 효과를 입히는 것이다. 따라서 특정 음향 시스템을 위해 가공된 형태의 음원 보다는 어떠한 효과도 첨가되지 않은 음원이 더 적합하다. 예를 들어 스테레오나 5.1 채널에 사용되는 음원 소스에는 부족하기는 하지만 입체음향에 영향을 미치는 여러 가지 효과가 이미 들어있다고 할 수 있다. 이러한 이유로 우리는 정확한 입체음향 효과를 주기 위해, 음상 및 음장 효과가 첨가되지 않은 음원을 “소리 객체”로 정의하였다. 소리 객체는 재 사용 가능하다. 여

기서 재사용이라 함은 하나의 소리 객체가 필요에 따라 임의의 시간에 반복해서 재생 가능함을 의미한다. 각각의 소리 객체는 입체음향 처리를 하기 전에 여러 가지 효과를 줄 수 있다.

실시간 트랙 매핑(Realtime Track Mapping)

일반적으로 비구조화 오디오 데이터는 연속적인 스트림 형식을 띄지만 구조화 오디오 데이터는 효율적인 메모리 및 대역폭 활용을 위해 데이터가 없는 구간에 대해서는 데이터를 저장하지 않는다. 따라서 각각의 소리 객체가 재생되는 시간 정보를 저장해 두고 이를 이용해 소리 객체를 실시간으로 가상 트랙에 매핑시켜 주어야 한다. 소리 객체를 트랙으로 매핑 시 모든 트랙에 소리 객체가 할당되는 것은 아니다. 예를 들어 무음 구간에서는 모든 가상 트랙에 대해 소리객체가 할당되지 않는다. 또 오디오 시스템의 데이터 처리능력이나 네트워크의 대역폭에 의해 가상 트랙의 최대 개수는 조정된다.

효과(Effect)

재생되는 각각의 트랙에는 다양한 필터를 이용한 음향 효과를 줄 수 있다. 트랙에 가해지는 음향 효과를 정의하는 필터들은 소리객체와 함께 저장되며 필터들을 서로 구분하는 ID 를 가진다. 소리객체를 트랙에 매핑 시 필터 정보도 같이 전달해 주어 각각의 트

랙에 대한 음향 효과 처리를 할 수 있다.

입체음향 처리(3D Sound Processing)

입체음향 처리는 HRTF 를 이용해 각각의 트랙에 대해 3 차원 공간상에서 음상(거리, 수평방위각, 주직상승각)을 정위함을 의미하며 스피커 시스템에 적당한 입체음향을 생성한다. 또 이 과정에서 필요에 따라 LFE 채널을 분리 후 LEF 채널로 보내 주어야 하는데 그것은 대부분의 멀티 채널 스피커 시스템에서는 저음을 효과적으로 처리하기위한 서브우퍼를 따로 두고 있기 때문이다. 2 채널 이상의 스피커 시스템에 대해 입체음향 구현이 가능하며, 스피커 시스템의 종류에

따라 입체음향 처리 방법이 다르므로 재생시 스피커 환경에 적합한 입체음향 처리가 필요하다. 각각의 트랙에 대해 입체음향 처리가 끝난 데이터는 믹싱 후 스피커로 보내준다.

3. 오디오 데이터

원하는 오디오 시스템을 설계하는 데에는 소리 객체 데이터, 소리 객체 및 필터에 대한 매핑 정보, 효과 및 필터 데이터, 소리객체가 트랙에 매핑 된 후 소리

있다.

그림 1 에서 보는 것처럼 스피커 시스템의 종류에 따라 입체 음향 처리 방법을 서로 다르게 구현해야 한다. 하지만 이러한 방법은 DTS, AC-3 등 기존의 멀티 채널 오디오 코딩에서처럼 오디오 데이터를 특정 스피커 환경에 적합하게 코딩 한 후 다른 스피커 환경을 위한 디코딩을 하는 것보다 음질 및 효율성 면에서 뛰어나다고 할 수 있다.

네트워크 상의 스트리밍 오디오 서비스에서는 네트워크

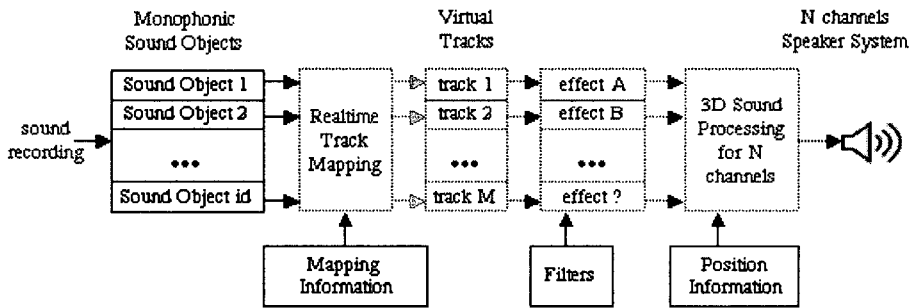


그림 2 오디오 데이터

객체가 가지는 위치정보의 데이터가 필요하다. 그림 2 는 이러한 데이터의 역할을 도식적으로 나타내고 있다. 그림 2 에서 매핑 정보는 소리 객체를 연속된 오디오 스트림으로 만들어 주는 정보와 각각의 트랙에 대해 특정 시간에 적용시키는 필터에 대한 정보를 포함한다. 즉 특정 시간에 재생해야 할 소리객체의 ID 와 특정 시간에 특정 트랙의 소리 객체에 적용해야 할 필터의 ID 를 말한다. 위치정보는 각각의 트랙에 포함된 음원의 3 차원 상에서의 좌표를 의미한다. 소리객체는 데이터 공간을 줄이기 위해 압축된 형태로 저장될 수 있지만 특정 효과를 주기위해 가공된 형태는 아니다. 그렇게 함으로써 모든 형태의 스피커 환경에서 공통으로 사용할 수 있게 된다.

4. 구현 및 활용

우리가 제안하는 멀티채널 오디오 시스템은 실제 구현 방법에 따라 대화식 멀티미디어 콘텐츠나 게임, 스트리밍 오디오 서비스 등에 활용할 수 있다. 또 현재 대부분의 오디오 콘텐츠는 서라운드 녹음이 되어 있지 않은데 이를 활용하는 방법으로 사용 가능하다. 즉 기존의 오디오 콘텐츠를 소리객체 자체로 인식 한다면 매핑 정보와 필터 그리고 위치 정보를 추가 함으로서 기존의 오디오 콘텐츠를 변경하지 않고도 서라운드 스피커 시스템에 적합한 형태로 재생 할 수

의 대역폭을 고려해야 하기 때문에 가상 트랙의 수를 대역폭 내에서 제한 해야 하는 문제점이 발생 할 수 있다. 그 경우에는 대역폭 내에서 Reed Solomon[1] 코딩을 이용해 일정한 비트 율로 데이터를 전송하는 방법과 무음 구간을 가지는 트랙에 대역폭을 초과하는 미래의 데이터를 보내는 방법을 이용할 수 있다.

5. 결론

우리는 오디오 시스템의 현장감을 높이기 위해 입체음향 기술을 적용 시키는 방법을 제시 하였다. 구조화 오디오 형식을 사용하여 데이터 공간을 효율적으로 사용하도록 하였으며, 소리 객체를 특정 스피커 환경을 위해 가공하지 않은 형태로 사용 함으로서 스피커 환경과 오디오 데이터간의 호환성 문제를 극복 하도록 하였다. 또 멀티 채널 형식으로 저장되지 않은 기존의 오디오 데이터를 멀티채널 스피커 환경에 맞게 변환하여 재생 하는 방법을 제시하였다. 오디오 데이터의 네트워크 스트리밍 서비스에서의 문제점의 해결 방안에 대한 추가적인 논의가 필요하고 여기서 제시된 방법의 효율에 대한 추가적인 검증작업이 이루어 져야 한다.

6. 참고문헌

1. F. E. Sandes, "Efficient Large-Scale Multichannel Audio

Coding,” Proceedings of the 27th EUROMICRO Conference: Workshop on Multimedia and Telecommunications, Warsaw, Poland, IEEE Computer Society Press, September, 2001, pp 392-399.

2. AESTD1001.1.01-10, “Multichannel surround sound systems and operations,” AES Technical council document, AES Technical Committee on Multichannel and Binaural Audio Technology, 2001.

3. *Chong, U.P.; Kim, I.H.; Kim, K.N.*, ”Improved 3d sound using wavelets,” Science and Technology, 2001. KORUS '01. Proceedings. The Fifth Russian-Korean International Symposium on , Volume: 1 , Jun 26 - Jul 3, 2001, Page(s): 54 -56

4. *Kraemer, A.*, “Two speakers are better than 5.1 [surround sound].” IEEE Spectrum , Volume: 38 Issue: 5 , May 2001, Page(s): 70 -74