

대화형 방송을 위한 객체기반 3차원 오디오 방송 시스템

유재현* · 이용주* · 이태진* · 장인선*
서정일* · 백승권* · 강경옥* · 장대영*

1. 서론

객체기반 3차원 오디오 서비스란 독립적으로 제어가 가능한 '객체별 오디오 신호'와, 전체 객체 오디오 신호를 하나의 장면(scene)으로 표현한 '3차원 장면정보'를 이용하여, 청취자가 각각의 오디오 객체 및 3차원 장면정보를 적절히 제어하면서 오디오를 청취할 수 있는 서비스이다. 객체기반 3차원 오디오 기술은 음원을 청취자의 전/후/좌/우 및 상/하 모든 곳에 위치시킬 수 있는 3차원 오디오 기술과, 혼합된 오디오 신호에서 특정 오디오 신호만을 객체화하여 제어할 수 있는 객체기반의 오디오 기술을 융합하여, 객체 오디오 신호를 3차원 오디오 공간상에서 제어할 수 있는 현실감 있는 멀티미디어 서비스 기술이다[1,2,3].

기존의 방송에서는 아래의 그림 1과 같이, 믹싱된 하나의 오디오 신호만이 전송되지만, 객체기반의 3차원 오디오 서비스에서는 그림 2와 같이 오디오 신호를 음원에 따라 여러 개의 독립된 객체로 전송하여, 객체별로 제어할 수 있다. 따라서 객체기반 3차원 오디오 서비스는 기존의 수동적인 오디오 서비스와 달리 자신만의 새로운 오디오 장면을 구성할 수 있는 인터랙티브 멀티미디어

서비스라 할 수 있다.

한편, 객체기반 3차원 오디오 서비스에서는 각 오디오 객체의 특성을 사용자가 일일이 제어할 수 있는 제어의 자유도는 매우 높은 반면, 이를 위한 사용자 인터페이스가 복잡하다는 단점이 있다. 즉, 오디오의 음압, 3차원 공간상의 위치 등과 같은 오디오 객체의 특성을 사용자가 제어하기 위해서는 각각을 설정할 수 있는 GUI가 제공되어야 한다. 키보드나 마우스와 같이 자유도가 높은 인터페이스를 제공하는 PC 환경에서는 이러한 GUI가 큰 문제가 되지 않지만, 리모컨과 같이 자유도가 높지 않은 인터페이스를 사용하여야 하는 DTV(Digital TV), DAB(Digital Audio Broadcasting), DMB, DVD player, mp3 player 등과 같은 재생 환경에서는 사용자가 객체기반 3차원 오디오 서비스를 통해 각 오디오 객체의 특성을 제어하는 것이 어려울 수 있다.

본 논문에서는 이와 같은 문제점을 해결하여, 리모컨 등과 같은 비교적 간단한 UI를 가지는 실시간 방송환경에서도 편리하게 객체기반 3차원 오디오 서비스를 이용할 수 있도록 하는 방법을 제안한다. 이를 위해 본 연구에서는 미리 잘 구성된 여러 개의 preset 오디오 장면을 저작하고, 이를 객체기반 오디오와 함께 전송하였으며, 사용자는 오디오 객체의 특성을 일일이 제어하지 않고, 미리 저작된 여러 개의 preset 오디오 장면들 중

* 한국전자통신연구원 전파방송연구단 방송미디어연구그룹
※ 본 연구는 정통부 SmarTV 3D AV 기반기술개발 연구 지원으로 수행되었음

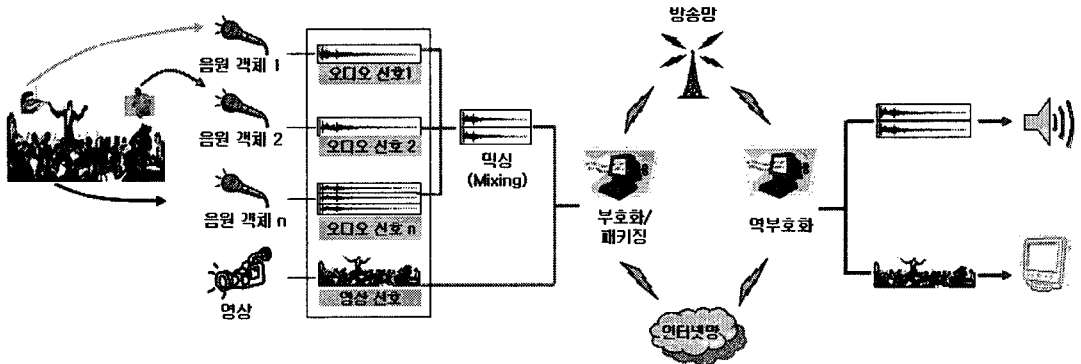


그림 1. 일반적인 오디오 서비스 개념도

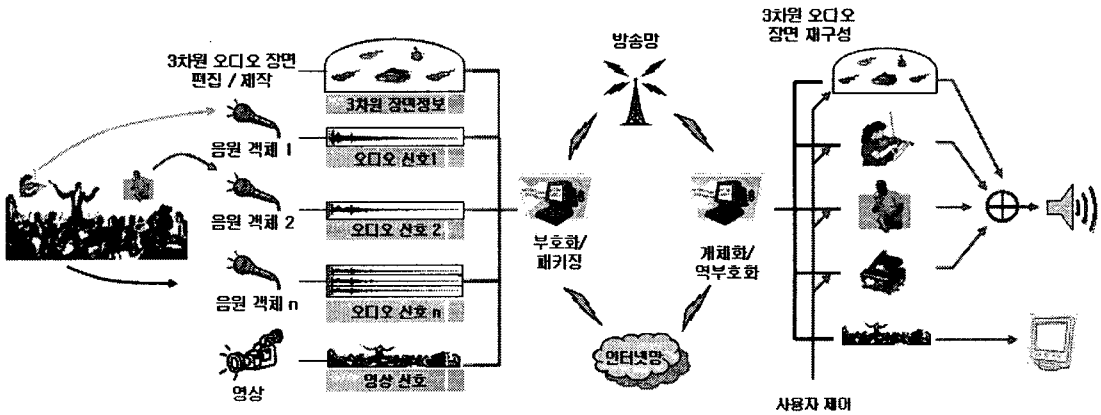


그림 2. 객체기반 오디오 서비스 개념도

하나를 선택하여 재생할 수 있도록 하였다. 이러한 방법은 사용자에게 있어서, 각 객체를 별도로 조작하는 것보다 편리하게 서로 다른 여러 개의 오디오 장면을 청취할 수 있다는 장점을 가진다. 또한, preset 오디오 장면을 제공하면서도, 각 오디오 객체를 제어하는 기능을 단말에서 제공함으로써, 재생 환경에 따라 오디오 객체 제어의 자유도는 그대로 유지할 수도 있다. 즉, 간단한 GUI를 가지는 재생 환경에서는 미리 저장된 preset 오디오 장면들 중 하나를 선택하여 시청하도록 하고, 자유도가 높은 GUI를 가지는 환경에서는 각 오디오 객체의 특성을 하나하나 제어할 수 있게 된다. 본 논문에서는 이와 같은 preset 오디오 장면을

이용한 객체기반 3차원 오디오 서비스가 간단한 GUI를 가지는 실시간 재생 환경에서도 잘 적용된다는 것을 확인하기 위하여, 지상파 DMB 환경에서 객체기반 3차원 오디오 서비스를 제공하는 실험을 수행하고, 이를 통해 효율적으로 객체기반 3차원 오디오 서비스를 제공할 수 있음을 확인하였다. 본 논문의 구성은 다음과 같다. 객체기반 3차원 오디오 서비스의 기술적인 배경이 되는 MPEG-4 AudioBIFS(Binary Format for Scenes) 규격과 입체 음향 기술에 대해 2장에서 기술하고, 3장에서 지상파 DMB 환경에서 객체기반의 3차원 오디오 서비스하기 위한 객체기반 3차원 오디오 방송 시스템의 각 모듈별 기능에 대해 설명한다. 그

리고 4장에서, 제안한 방법에 따른 객체기반 3차원 오디오 서비스에 대한 실험에 대해 기술하고, 5장에서 결론을 맺는다.

2. 객체기반 3차원 오디오 서비스를 위한 기반 기술

객체기반 오디오 방송 서비스를 위해서는 각각의 오디오 객체를 다중화 하는 기술과 이를 3차원 공간상에 배치하는 기술이 필요하다. 본 객체기반 3차원 오디오 방송 시스템에서는 MPEG-4 시스템 규격에 따라 객체별로 생성된 오디오 신호의 특성과 편집자가 의도하는 다양한 오디오 장면을 기술하고, 다중화 하였다.

한편, 객체기반 3차원 오디오 서비스를 위해서는 MPEG-4 시스템을 이용하여 기술된 3차원 오디오 장면을 재현하기 위한 입체 음향 기술이 필요하다. 입체 음향 기술은 오디오 음원이 녹음된 그 공간에 있는 것과 같은 현실감을 주는 것을 목적으로 하는 것으로서, 음원이 3차원 공간상의 특정 위치에 있는 것과 같은 공간감을 주는 음상 정위 기술과, 현실과 같은 잔향을 부여하는 잔향 효과가 대표적인 기술이라 할 수 있다.

2.1 MPEG-4 Audio BIFS

MPEG-4 시스템[4]은 부호화하고자 하는 대상을 객체 단위로 분할하고, 객체의 속성에 부호화기를 선정하여 압축률을 높이고, 사용자의 입력을 참조하여 객체별로 서로 다른 효과를 줄 수 있는데, 이 중 오디오 객체의 특성 및 3차원 장면을 위해 사용되는 것이 Audio BIFS이다. 여기에서 객체라는 용어는 장면 상에서 각 요소(ES, Elementary Stream)로 참여하는 AV 데이터를 가리킨다. 특히 이 때 객체를 물리적 시공간에 배치하기 위해서는 객체 자체에 대한 정보 이외에

별도로 장면정보가 필요하다. 장면정보는 노드(node)로 이루어진 그래프로 표현되는 계층구조(scene graph)를 가지며, 사용자 단말에서는 이와 같은 객체 정보와 장면 정보를 이용하여 3차원 오디오 장면을 구성하게 된다. 각 노드에는 유형에 따라 미리 정의된 필드가 하위 구조로 속하게 되며 이러한 특정 유형의 필드를 이용한 객체기술자(OD, object descriptor)를 참조함으로써 해당 객체와 장면 그래프를 연결할 수 있게 된다. 또한 IOD(Initial OD)가 장면 서술에 관련된 OD로서 함께 사용된다.

오디오 장면을 구성하기 위해서 음원의 음압과 위치를 나타내는 Sound2D 노드, 음원의 지향성을 표현하는 DirectiveSound 노드 등 많은 노드들을 사용하는데, 본 논문에서 제안하는 단말에서는 WideSound 노드 인터페이스에 따른 위치감 및 AcousticScene 노드 인터페이스에 따른 잔향모델을 통해 사용자가 음원의 위치 및 잔향을 조절할 수 있다[4,5].

2.2 음상 정위 입체 음향 기술

객체기반 3차원 오디오 방송 시스템에서는 오디오 객체의 현실적이고 정확한 음상 정위를 위해 HRTF(Head Related Transfer Function)를 사용한다. HRTF는 음원에서부터 청취자의 귀까지 소리가 전달되는 동안 발생하는 상반신, 머리, 귓바퀴 등으로 인한 반사, 회절 등의 효과를 고려하여 측정된 전달함수이다. 그림 3과 같이 단말에서 사용자가 입력한 음상 정위 정보에 따른 HRTF 데이터를 입력 모노 신호와 콘벌루션한 후 역복합 법칙을 이용해 음원과 청취자 사이의 거리(r)에 따라 음압을 감쇠시키고, 이렇게 생성한 음원을 라우드스피커를 통해 재생하기 위하여 크로스토크(cross talk)를 제거하는 역필터링을 수행하였다.

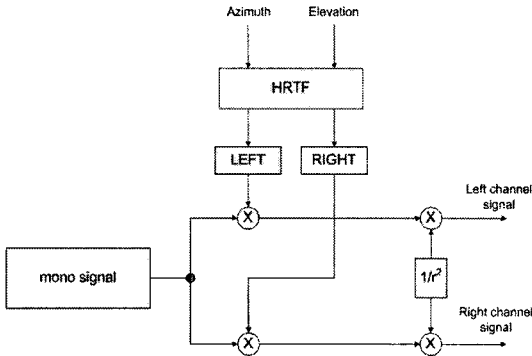


그림 3. 음성 정위 시스템 블록 다이어그램

2.3 잔향 효과

현실감 있는 잔향 효과를 위하여 MPEG-4 AudioBIFS 중 AcousticScene 노드 인터페이스의 잔향모델을 선형시스템으로 구현하였다. 후부 잔향은 all-pass 필터의 일종인 comb 필터를 변형한 IIR 필터[6]를 이용하여 구현하였다. 아래 그림 4에서 g_E 와 g_L 은 각각 초기 반사이득 및 후기 잔향 이득을 나타낸다. g_E 의 값과 함께 그 값을 곱하기 위해 취하는 필터 탭의 위치를 변경함으로써 초기 반사 특성을 제어할 수 있다[1].

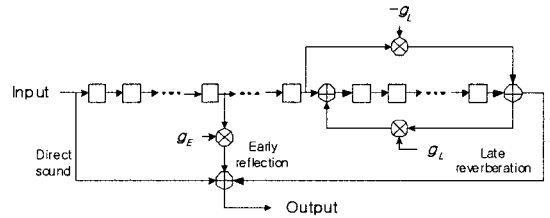


그림 4. 인공잔향기의 구조

3. 객체기반 3차원 오디오 방송시스템

객체기반 3차원 오디오 방송시스템은 그림 5와 같이 전처리모듈, 편집 및 부호화 모듈, 송신모듈로 구성된 서버 시스템과 수신모듈 및 재생모듈로 구성된 수신 단말로 구성된다[7].

다양한 방식으로 획득한 오디오 스트림은 전처리 모듈을 통해 객체기반 시스템에 활용할 수 있는 객체 오디오 신호로 만들어진다. 편집 및 부호화 모듈은 여러 개의 오디오로 구성된 3차원 오디오 장면을 생성하고, 각 오디오 및 비디오 객체를 부호화/다중화하여 MPEG-4 파일을 생성하는 기능을 수행한다. 송신모듈은 MPEG-4 파일을 MPEG-2 TS로 패킷화하고, 이를 지상파 DMB 망을 통해 전송하는 기능을 수행한다. 수신 단말은 지상파 DMB 신호를

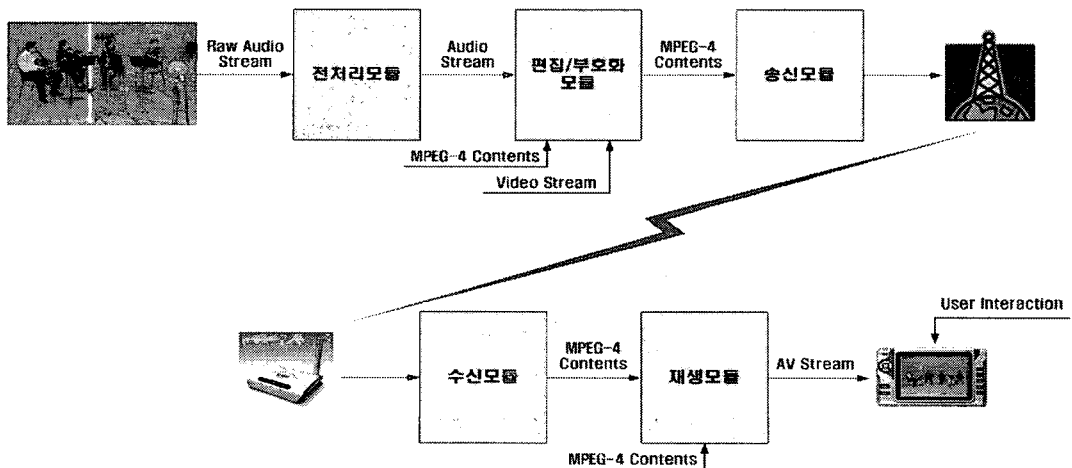


그림 5. 객체기반 3차원 오디오 방송 시스템

수신하고, 이를 재생하는 기능을 수행한다. 이때 수신 단말은 각각의 오디오 객체를 제어할 수 있는 GUI를 제공하며, 여러 개의 preset 오디오 장면들 중 하나를 선택하여 재생하는 기능을 포함한다.

3.1 객체기반 오디오 전처리 모듈

객체기반 3차원 오디오 전처리 모듈에서는 다양한 방식으로 음원을 획득하여 객체별 오디오 신호로 가공하는 과정을 거치게 된다.

3.1.1 객체기반 오디오 획득

객체기반 3차원 오디오를 서비스하기 위해서는 객체별로 전송하고자 하는 음원을 독립적으로 확보해야 한다. 독립적으로 음원을 얻어내는 방법에는 여러 가지가 있을 수 있으나, 최초 오디오 신호의 녹음 시 각 오디오 음원을 개별적으로 녹음하는 방법과 여러 개의 음원이 섞여있는 오디오 신호를 획득하고 이로부터 서비스하고자 하는 음원을 분리하여 획득하는 방법이 일반적이라 할 수 있다.

3.1.2 객체기반 오디오 전처리

다양한 장치를 통해 입력받은 음원으로부터 객체 오디오 신호를 생성한다. 완벽한 객체 오디오 신호의 획득이 어려운 경우, 다수의 마이크로폰을 이용하여 녹음한 혼합음원으로부터 블라인드 음

원분리(Convolutional Blind Separation; CBS) 기술을 이용하여 보다 깨끗한 객체별 오디오 신호를 얻어낸다. 블라인드 음원분리는 원 음원과 녹음환경에 대한 정보 즉, 채널 전달 특성과 같은 사전 정보 없이 마이크로폰의 출력인 혼합신호만을 이용하여 상호 독립적인 원 음원을 추정하는 기술이다. 본 시스템에서는 음원분리의 복잡도를 고려하여 주파수 영역에서 신호의 이차 통계량 및 비안정성에 기반한 블라인드 음원분리 기술을 적용하며 이의 절차는 그림 6과 같다[8,9,10].

본 연구팀에서는 객체기반 3차원 오디오 시스템 적용을 위한 음원분리 기술을 계속적으로 개발 중이며, 현재 개발된 상대 최적화 기반 CBS 알고리즘은 객체 수가 4개인 경우에도 분리도(Signal to Interference; SIR)이 10dB 이상 개선되는 것을 확인하였다[11].

3.2 편집 및 부호화 모듈

객체기반 3차원 오디오 방송시스템의 편집 및 부호화 모듈은 3차원 오디오 장면을 구성한 후, 지상파 DMB 채널을 통해 전송하기 위한 A/V 신호로 부호화하고, MPEG-4 시스템을 이용하여 다중화 하는 기능을 수행한다. 여기에 사용되는 도구가 그림 7에 제시한 객체기반 3차원 오디오 편집도구이다.

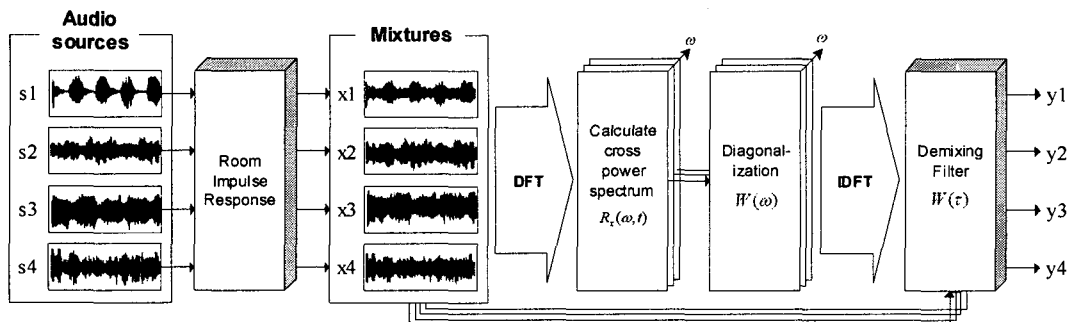


그림 6. 블라인드 음원분리 알고리즘의 절차

기존의 객체기반 3차원 오디오 편집도구에서는 하나의 오디오 장면만을 제작하였다. 편집도구를 이용하여 만들어진 하나의 오디오 장면은 AudioBIFS의 형태로 오디오 데이터와 함께 다중화 하여 전송, 사용자 단말에서 하나의 오디오 장면만을 구성할 수가 있었고 그 외 별도의 장면을 만들어내기 위해서는 반드시 사용자가 단말 S/W에서 직접 모든 오디오 객체를 조정해야만 했다.

본 논문에서는 여러 가지의 오디오 장면을 만들어 preset 형태로 저장하는 방식을 이용하여 하나의 장면밖에 만들 수 없었던 기존의 문제점을 해결하였다. 즉, 각 오디오 객체의 수, 음압, 위치 등의 속성을 서로 다르게 편집한 다양한 종류의 오디오 장면을 구성하여 순서대로 preset0, preset1, preset2, ... 와 같은 식으로 각각 저장한다. 오디오 장면의 구성은 콘텐츠의 내용에 따라 달라질 수 있는데, 오디오 구성 및 비디오 내용 등을 참조하여 편집자가 콘텐츠에 가장 적합하다고 생각하는 오디오 장면을 구성할 수 있다. 가령, 그림 8과 같이 바이올린 2개와 첼로 1개, 비올라 1개로 이루어진 합주 콘텐츠를 제작한다고 하면 각 악기 소리를 가운데로 모아놓은 오디오 장면, 각 악기가 3차원 공간상에 개별적으로 배치되어 입체

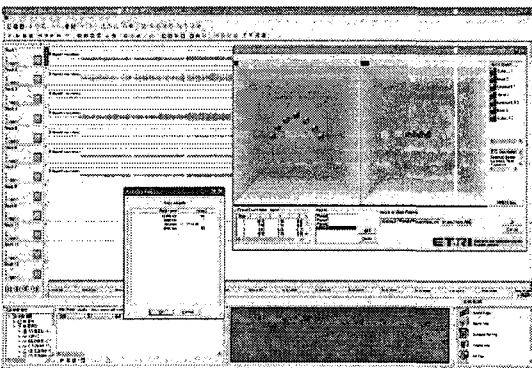


그림 7. 객체기반 3차원 오디오 편집도구

음향을 나타내는 오디오 장면, 악기 소리 중 하나 혹은 그 이상의 신호를 제거한 오디오 장면 등 다양한 3차원 오디오 장면을 구성할 수 있게 된다. preset으로 만들 수 있는 오디오 장면의 개수에는 제한은 없지만 너무 많은 preset을 제작하는 것은 오히려 사용자에게 혼란을 줄 수 있으므로 콘텐츠에 적합한 3~4가지가 적당하다. 사용자는 단말을 통해 이와 같은 다양한 preset 중에서 본인이 선호하는 preset을 하나 선택하여 간편하게 오디오 장면을 구성, 감상할 수 있게 된다. 이러한 preset 오디오 장면 정보는 전체 오디오 장면의 개수, 오디오 장면에 포함된 오디오 객체의 수, 개별 오디오 장면에 속한 각 오디오 객체의 음상 정위, 음압 등의 속성에 대한 정보를 담고 있다.

본 논문에서는, 상기와 같이 preset 정보를 이용하여 오디오 장면 정보를 변경할 수 있는 기능에 대한 실제 방송 환경에서의 적합성 여부를 검증하기 위하여 지상파 DMB 환경을 이용한다. 따라서 오디오 및 비디오 데이터는 지상파 DMB 규격에

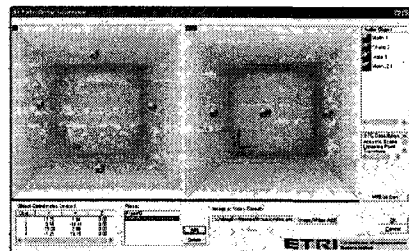
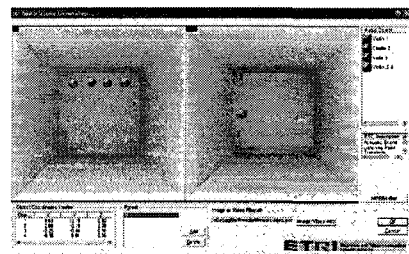


그림 8. 4개의 오디오 객체를 이용한 preset 오디오 장면 설정 화면

맞도록 각각 BSAC(Bit Sliced Arithmetic Coding) [12]과 AVC(Advanced Video Coding) [13]로 부호화한 후, 객체의 속성에 대한 정보를 담은 OD와 BIFS, 그리고 부가 데이터 형태인 preset 정보와 함께 MPEG-4 파일 형태로 다중화 된다.

3.3 송신 모듈

지상파 DMB 송신모듈은 객체기반 오디오 DMB 미디어 처리기를 통해 MPEG-4 파일로 저장된 객체기반 오디오 콘텐츠를 MPEG-2 TS (Transport Stream) 으로 패키징하는 기능과 이를 지상파 DMB망을 통해 전송하는 기능을 수행한다.

객체기반 오디오 DMB 미디어 처리기는 그림 9와 같은 구조를 가지며, 기존의 DMB 미디어 처리기와는 달리, 하나 이상의 오디오 객체를 포함하는 MPEG-4 콘텐츠를 MPEG-2 TS로 패키징하는 기능을 포함한다. 이때, 서로 다른 오디오 객체는 서로 다른 PID(Packet Identifier)를 가지도록 하였다.

이렇게 생성된 MPEG-2 TS는 상용 지상파 DMB 송출장비를 통하여, 지상파 DMB 망을 통해 송출된다.

3.4 수신 단말

수신 단말은 지상파 DMB 신호를 받아 이를 역다중화하고 복호화 하여 오디오 및 비디오를 재생한다. 수신 단말은 일반적으로 하나의 시스템으로 구성되나, 기능적으로 RF망을 통해 전송되는 신호를 수신하여 MPEG-2 TS를 추출하고, 이를 역패킷화하여 MPEG-4 콘텐츠를 복원하는 수신 모듈과 MPEG-4 콘텐츠를 재생하는 재생모듈로 나눌 수 있다.

3.4.1 수신 모듈

지상파 DMB 수신 모듈은 그림 10과 같이 지상파 DMB 신호를 입력 받아 이를 복조하여 MPEG-2 TS를 추출하고, 이를 역패킷화하는 기능을 수행한다.

3.4.2 재생 모듈

재생 모듈은 수신 모듈로부터 받은 MPEG-4 파일 스트림을 역다중화하여 BSAC, AVC 와 OD, BIFS 스트림 등의 MPEG-4 스트림을 복호화하며 BIFS와 OD 및 preset을 통한 3차원 오디오 장면 구성 정보를 이용하여 오디오 장면을 구성, 오디오와 비디오를 재생한다[1].

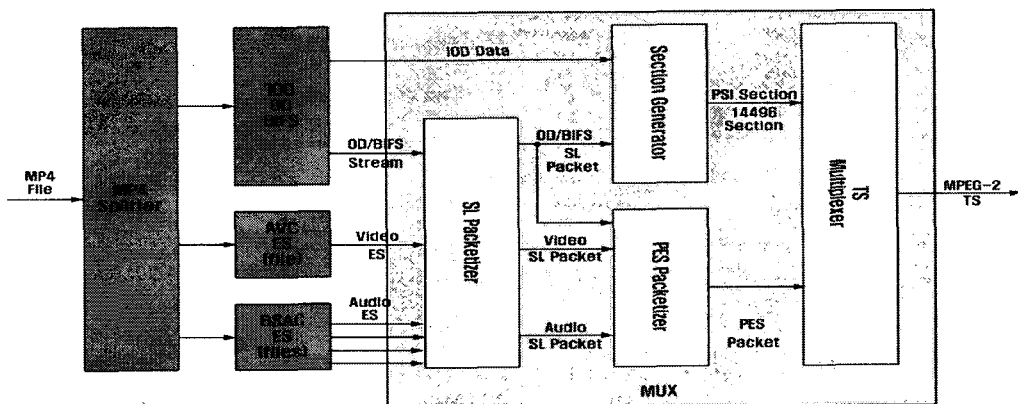


그림 9. 객체기반 3차원 오디오를 위한 객체기반 오디오 DMB 미디어 처리기 구조도

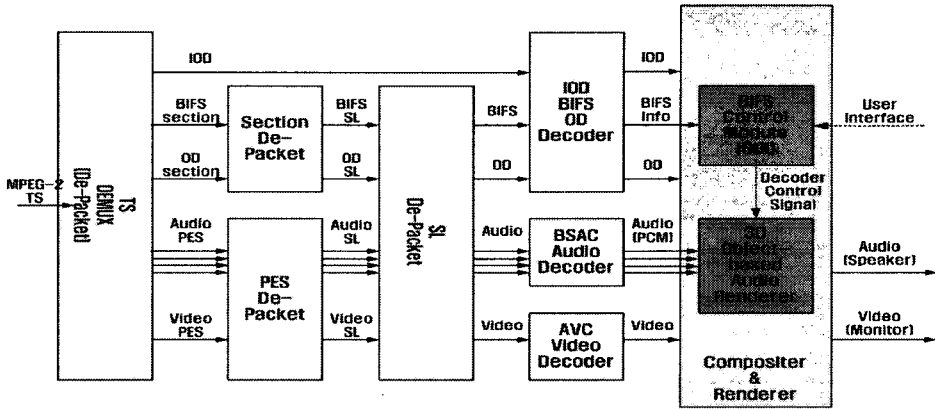


그림 10. 객체기반 3차원 오디오 지상파 DMB 수신 단말 구조도

3.4.3 단말 S/W

사용자는 전술한 수신 모듈과 재생 모듈을 가지고 있는 그림 11과 같은 수신 단말 S/W를 이용하여 오디오와 비디오를 듣고 볼 수 있다.

단말 S/W에서 제공하는 그림 12와 같은 GUI를 이용하면 각 객체의 음상 정위 정보, 음압의 정도, 잔향 효과 등의 특성을 제어할 수 있어 사용자 본인이 원하는 대로 오디오 객체를 제어할 수 있다.

이 인터페이스를 호출하여 각 오디오 객체에 대해서 음상 정위 기능을 활성화시키고 텍스트 입력이나 슬라이드 바를 이용하여 3차원 공간상의 특정 위치로 설정해주며 원하는 음압으로 조정, 잔향의 정도를 조정해주면 하나의 3차원 오디오 장면이 설정된다. 또 다른 오디오 장면을 설정하고자 하는 경우에는 이와 같은 작업을 반복해야 한다. 이렇게 여러 개의 오디오 객체를 하나하나 제어한다는 것은 리모콘 등과 같이 비교적 간단한 UI를 가지는 실시간 방송 환경에서는 조작 자체의 번거로움과 조작에 걸리는 시간 등의 측면에서 사용자에게 매우 불편한 일이다.

따라서 편집도구에서 마련한 몇 가지의 preset 오디오 장면정보를 활용하면 보다 쉽게 오디오 장면을 제어할 수 있게 된다. 단말 S/W의 Prs 버튼을 누르면 그림 13과 같은 리스트가 나타나면서 전송되어온 preset의 리스트가 보여지고 이 중에서 하나를 선택하면, 각각의 preset 오디오 장면 정보에는 각 오디오 객체의 음상 정위, 음압 등의 속성에 대한 정보를 담고 있으므로 그 정보에 맞는 오디오 장면이 구성되어 재생된다. 별도의 인터페이스를 통하여 가능했던 제어가 한 번의 조작만으로 가능해져 매우 편리하게 장면 구성을 변화시킬 수 있게 된다.

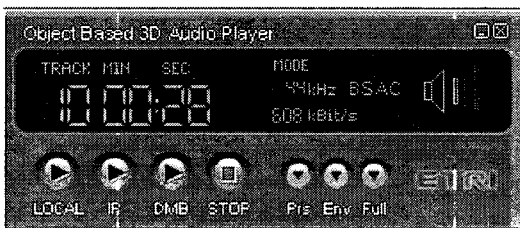


그림 11. 객체기반 3차원 오디오 수신 단말 S/W

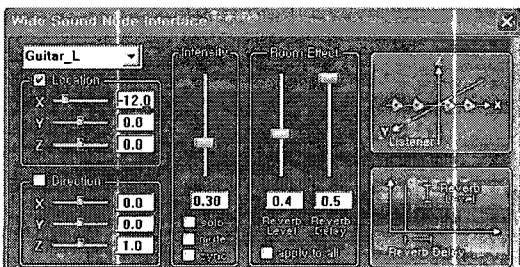


그림 12. 오디오 객체별 속성 제어 인터페이스

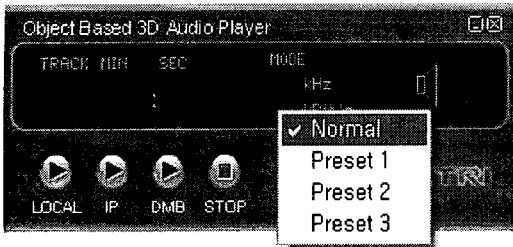


그림 13. preset 장면 구성 정보의 선택

4. 실험 및 고찰

본 객체기반 3차원 오디오 방송 시스템을 테스트하기 위하여 편집도구를 통해 4가지 preset을 가지는 뮤직비디오 콘텐츠를 제작하고 지상파 DMB 송신 환경을 구축, PC에서 수신 및 단말 S/W를 구동하였다. 뮤직비디오에는 최대 64개까지의 악기 소리가 사용되는 경우를 포함하여 상당히 많은 수의 오디오 객체가 사용되는데 본 테스트에서는 드럼과 베이스, 두 개의 기타와 키보드로 이루어진 6개의 악기와 가수의 목소리까지 전체 7개의 오디오 객체가 사용되는 뮤직비디오 콘텐츠를 제작하였다.

본 실험에서는 뮤직비디오 콘텐츠에 적용하기 위해서 그림 14와 같은 4가지의 preset 장면 구성 정보를 제작하였다. 각각은 시계방향으로 normal과 spatialize, vocal, karaoke로 구분되며 normal preset의 경우는 모든 오디오 객체가 공간정보를 가지고 있지 않은 원래의 모노채널 특성대로 청취자 정면 정 가운데에 모든 음상이 정위된다. spatialize preset은 입체 음향 기술을 이용하여 각각의 객체를 미리 지정된 공간상의 위치로 정위시킨다. 각 객체의 공간상의 위치가 모두 다르기 때문에 입체 음향을 청취할 수 있으며 모든 음상이 개별적으로 인지된다. karaoke preset은 재생되는 오디오 객체 중 가수 목소리에 해당하는 오디오

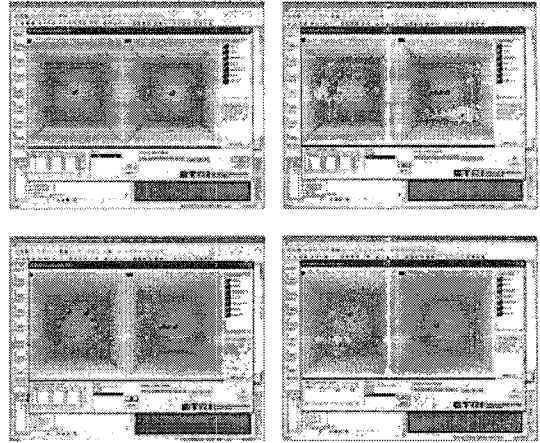


그림 14. 뮤직 비디오 콘텐츠의 preset: Normal, Spatialize, Vocal, Karaoke(시계방향으로)

오 객체만을 제거하여 악기 소리만이 재생되도록 만들어준다. 이 경우 특정 노래의 경음악만을 재생해주기 때문에 마치 노래방에 온 것처럼 사용하는 음악에 맞추어 노래를 부를 수도 있다. vocal preset은 karaoke preset과는 정반대로 가수 목소리를 제외한 모든 악기 소리를 제거하여 사용자가 가수의 노래에 맞추어 악기를 연주할 수도 있다.

기존 방식대로라면 객체별 속성 제어 인터페이스를 사용해야만 이와 같은 오디오 장면 제어를 할 수가 있었다. 가령, 전술한 spatialize preset과 같은 오디오 장면을 구성하고자 한다면 객체별 속성 제어 인터페이스를 호출하여 각 오디오 객체의 음상 정위 효과 기능을 활성화시키고 3차원 공간상의 특정 위치를 설정해주어야 한다. 또 vocal preset과 같은 오디오 장면을 구성하고자 하면 vocal 오디오 객체를 제외한 6개 오디오 객체 하나 하나의 intensity를 0으로 설정해주어야 한다.

이렇듯 7가지의 오디오 객체를 하나하나 제어한다는 것은 시간도 걸리고 그 조작 또한 리모콘 등과 같이 비교적 간단한 UI를 가지는 실시간 방

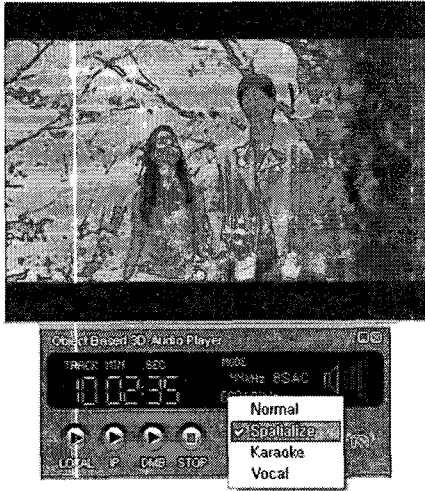


그림 15. 뮤직 비디오 콘텐츠를 이용한 실험

송 환경에서는 조작의 불편함과 조작에 걸리는 시간 등의 측면에서 사용자에게 매우 번거로운 일이다. 따라서 이 뮤직 비디오 콘텐츠 실험의 경우처럼, 그림 15와 같이 preset 오디오 장면 정보를 활용하여 사용자는 별도의 GUI를 활용하지 않고도 단말 S/W 상의 버튼 하나로 4가지의 오디오 장면을 쉽고 편리하게 제어할 수 있다.

5. 결론

객체기반 3차원 오디오 방송시스템은 단순히 편집자가 만들어 놓은 하나의 오디오 장면만을 청취할 수 있는 기존의 방송 시스템을 탈피하여 사용자가 객체의 속성을 제어하여 다양한 오디오 장면을 만들어낼 수 있는 방송서비스이다. 기존에는 오디오 장면 정보를 변경하기 위하여 별도의 GUI로 객체 제어 유저 인터페이스를 활용할 수밖에 없었는데, 자유도가 높지 않은 인터페이스를 제공하는 DTV, DMB, mp3player 등과 같은 환경에서는 그 조작이 용이하지 않기 때문에, 본 논문에서는, 보다 쉬운 제어 방법으로 preset을 사용하는 방안을 제시하였다. 편집자가 콘텐츠에 적합한

몇 가지의 오디오 장면을 preset 정보로 제작하여 전송하면 사용자는 이 중에서 본인의 취향에 맞는 것을 선택하여 감상하면 된다. 객체별 속성 제어 인터페이스도 제공하기 때문에 보다 세부적인 제어를 원하는 사용자는 이를 통해 본인만의 독창적인 오디오 장면을 만들어낼 수도 있다. 따라서 preset 기능을 활용하게 되면, 뮤직 비디오 콘텐츠를 이용해 실시간 지상파 DMB 환경에서 테스트를 한 바와 같이, 사용자에게 보다 쉽고 효율적인 객체 제어 방법을 제공할 수 있게 된다.

참고 문헌

- [1] 박기윤, 이태진, 강경옥, “대화형 3차원 오디오 방송단말 구현,” 한국방송공학회 2004년 학술대회
- [2] Taejin Lee et al., “An Object-based 3D Audio Broadcasting System for Interactive Services,” AES 118th Convention, 2005.
- [3] 유재현, 박기윤, 이태진, 강경옥, “대화형 방송을 위한 객체기반 3차원 오디오 재생 단말의 구현,” 한국음향학회 2005년 춘계학술대회
- [4] ISO/IEC 14496-1, “Information Technology - Coding of Audio-Visual Objects, Part 1: Systems,” 2002.
- [5] ISO/IEC 14472-1, “Virtual Reality Modeling Language,” 1997.
- [6] M. R. Schroeder, “Natural Sounding Artificial Reverberation,” Journal of the Audio Engineering Society, vol. 10, no. pp. - 233, 1962.
- [7] 이태진, 유재현, 장인선, 이용주, 서정일, 장대영, 강경옥, 홍진우, “지상파 DMB 기반의 객체기반 3차원 오디오 방송 시스템,” 한국음향학회 2005년 추계학술대회
- [8] S. Choi, H. Hong, H. Glotin and F. Berthommier, “Multichannel signal separation for cocktail party speech recognition: A dynamic recurrent network,” Neurocomputing, vol. 49, no. 1-4, pp. 299-314, Dec. 2002
- [9] I. Jang, S. Kim, S. Choi and K. Kang,

"F-SEONS: A second-order frequency-domain algorithm for noisy convolutive source separation," in Proc. IEEE Int'l Symposium on Circuits and Systems, Kobe, Japan, May 23-26, 2005.

- [10] L. C. Parra and C. Spence, "Convolutive blind source separation of non-stationary sources," IEEE Trans. Speech and Audio Processing, pp. 320-327, May 2000.
- [11] I. Jang, K. Nam, S. Beack, K. Kang and S. Choi, "Blind source separation using relative optimization," in Proc. WESPAC IX, Korea, 2006.
- [12] ISO/IEC 14496-3, Information Technology - Coding of Audio-Visual Objects, Part 3: Audio, Dec. 1999.
- [13] ITU-T Rec. H.264 | ISO/IEC 14496-10 AVC: Advanced Video Coding for Generic Audio-visual Services, Mat. 2003.



이 용 주

- 1999년 경북대학교 전자공학과(공학사)
- 2001년 경북대학교 대학원 전자공학과(공학석사)
- 2001년~현재 한국전자통신연구원 전파방송연구단 방송미디어연구그룹 선임연구원
- 관심분야 : 대화형 방송 시스템, 음향신호처리, 객체기반 3차원 오디오 신호 처리 등



유 재 현

- 2003년 홍익대학교 전자전기공학부(공학사)
- 2005년 서울대학교 대학원 전기컴퓨터공학부(공학석사)
- 2005년~현재 한국전자통신연구원 전파방송연구단 방송미디어연구그룹 연구원
- 관심분야 : 오디오 신호 처리, 가상 음상 정위 기술, 객체기반 3차원 오디오 신호 처리 등



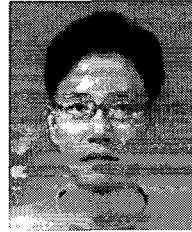
이 태 진

- 1996년 전북대학교 전자공학과(공학사)
- 1998년 전북대학교 대학원 전자공학과(공학석사)
- 1998년~2000년 (주)모벤스 영상음성기술연구팀 연구원
- 2002년~2003년 일본 Tokyo Denki Univ. 객원연구원
- 2000년~현재 한국전자통신연구원 전파방송연구단 방송미디어연구그룹 선임연구원
- 관심분야 : 오디오 신호처리, 객체기반 3차원 오디오 신호 처리 등



장 인 선

- 2001년 충북대학교 전기전자공학부(공학사)
- 2004년 포항공과대학교 대학원 컴퓨터공학과(공학석사)
- 2004년~현재 한국전자통신연구원 전파방송연구단 방송미디어연구그룹 연구원
- 관심분야 : 오디오 신호처리, 음원분리, 객체기반 3차원 오디오 신호처리, MPEG, DMB 등



백 승 권

- 1999년 한국항공대학교 항공전자공학과(공학사)
- 2001년 한국정보통신대학교 대학원 전자공학과(공학석사)
- 2005년 한국정보통신대학교 대학원 전자공학과(공학박사)
- 2005년~현재 한국전자통신연구원 전파방송연구단 방송미디어연구그룹 선임연구원
- 관심분야 : 오디오 신호처리, 음원분리, 객체기반 3차원 오디오 신호처리, MPEG, DMB 등



서 정 일

- 1994년 경북대학교 전자공학과(공학사)
- 1996년 경북대학교 대학원 전자공학과(공학석사)
- 2005년 경북대학교 대학원 전자공학과(공학박사)
- 1998년~1999년 LG 반도체 근무
- 1999년~2000년 현대 전자 근무
- 2000년~현재 한국전자통신연구원 전파방송연구단 방송미디어연구그룹 선임연구원
- 관심분야 : 음향신호처리, 음성신호처리, 객체기반 3차원 오디오 신호처리, MPEG, DMB 등



강 경 옥

- 1985년 부산대학교 물리학과(이학사)
- 1988년 부산대학교 대학원 물리학과(이학석사)
- 2004년 한국항공대학교 대학원 항공전자공학과(공학박사)
- 1991년~현재 한국전자통신연구원 전파방송연구단 방송미디어연구그룹 책임연구원
- 관심분야 : 오디오 신호처리, MPEG-7과 TV-Anytime 기반 맞춤형 방송 등



장 대 영

- 1991년 부경대학교 전자공학과(공학사)
 - 2000년 배재대학교 대학원 컴퓨터공학과(공학석사)
 - 1991년~현재 한국전자통신연구원 전파방송연구단
방송미디어연구그룹 책임연구원
 - 관심분야 : 오디오 신호처리, 음원분리, 객체기반 3차원
오디오 신호처리, MPEG, DMB 등
-
-