

## 지능형 실감 오디오 재생 기술

정종진·임태범·이석필 (전자부품연구원)

### I. 서론

급속한 정보 통신 기술의 발달로 인해 멀티미디어 재생 개발 기술들은 단순히 수동적으로 보고 듣는 재생 기술에서 벗어나 청취자 감성, 취향 등에 따라 보다 실감 있고 사용자가 능동적으로 재생할 수 있는 기술로 진화 하고 있다. 또한 멀티미디어 기기시장은 기존의 단순 복합 디지털 기기들이 아닌 사용자 감성 및 취향 제어가 가능한 인간 친화적 지능형 멀티미디어 기기가 주류를 이룰 것이며, 그 중에서도 특히 인간의 청각을 통한 사용자 맞춤형 감성 서비스를 지원하는 실감형 멀티미디어 관련 산업은 고부가가치 프리미엄 시장을 형성 할 것이라 기대된다. 이미 IT 기능이 기존의 '정보의 소통'에서 '감성의 소통'으로 진화 중에 있으며, 또한 미래시대에는 느낌까지 디지털 신호로 전달 가능한 기술이 발달 될 것이다. 이에 맞추어 사람의 감성, 주변 분위기, 섬세한 공간 정보를 전달하는 사실적인 오디오 개발 및 인프라가 구축되어 모든 멀티미디어 제품에 적용된다면, 사용자는 보다 현장감 있게 멀티미디어를 즐길 수 있을 것이다.

오디오는 비디오에 비해 인간의 감성, 주변 시

청 환경에 따라 더 많은 영향을 준다. 이러한 관점에서 오디오 관련 기술은 비디오 분야 기술에 비해 인간의 내적 감성에 호소하는 심리적 측면이 강조되는 연구 분야로서, 이와 관련한 핵심 기술 및 소자 개발에 있어서 고도의 전문성이 필요하며 이를 연구하는 기관의 수에 있어서도 희소성을 지니므로 향후 고부가가치를 창출할 수 있는 매우 중요한 요소 기술이다.

현재 전세계 오디오 관련 제품 시장에서 가장 큰 흐름은 홈시어터, A/V 리시버, PC 및 TV에 네트워크 기술을 접목시켜 고음질의 음악을 손쉽게 즐길 수 있게 발전하고 있다. 또한 발달된 디지털 신호처리 기술로 각종 필터의 구현, 음장 제어, 3차원 오디오 효과 등의 구현 가능해져서 네트워크 기술을 접목시킨 고음질 오디오 서비스 이외에 시장에서의 수요를 견인할 만한 새로운 오디오 기술 및 장치 개발이 필요한 시점이다. 이에 전 세계적으로 향후 오디오 기술의 핵심 요소가 될 실감 오디오 연출 기술에 대한 많은 연구가 진행 중에 있으나 이와 관련한 제품 개발은 걸음마 상태이며, 이에 대한 국내의 기술 개발을 아직은 미비한 상태이다.

이에 본 기고에서는 고품질 및 실감형 오디오

를 추구하는 “지능형 실감 오디오 재생 기술”에 대한 소개와, 이 기술의 핵심 기술 요소인 사용자 능동형 오디오 재생을 위한 멀티 트랙 음원 재생 기술, 사용자 감성 및 현장감 재생을 위한 정보 표현 기술 및 오디오 재생 기술, 사용자의 청각 특성과 환경정보 기반 실감 오디오 전송 기술에 관해 살펴보기로 한다.

## II. 지능형 실감 오디오 재생 기술

### 1. 개요

“지능형 실감 오디오 재생 기술”은 다양한 멀티채널 음향 오디오 구조에 기반 한 차세대 객체 기반 오디오 콘텐츠를 사용자 측면에서 사용자 환경, 특성, 취향 등에 따른 실감 감성 정보를 통해 청취자가 실제 환경과 같은 현장감을 느낄 수

있도록 제공하는 재생 플랫폼 기술로서 실감 있는 오디오(reality), 현장에서 듣는 듯한 오디오(presence)를 사용자가 존재하는 물리적인 공간의 음향학적인 특성은 물론 사용자의 취향 및 신체적, 심리적 환경에 최적화하여 제공한다.

<그림 1>은 다양한 환경에서 취득된 고음질 콘텐츠를 사용자 특성, 취향, 환경 등을 기반으로 한 실감 오디오 처리를 통해, 청취자에게 실제 환경에서 느끼는 감동을 제공할 수 있는 미래 지향형 실감 오디오 재생 기술의 개념을 설명하고 있다. 이 기술이 제공하는 주요 기능은 다음과 같다.

- 음원 취득과정에서 사용자 능동형 오디오 서비스 가 능토록 독립 음원의 분리 및 환경 파라미터 추출을 함께 수행함으로써 광대역 실감 음원 획득을 가능케 하며 공간/음향 부호화를 통해 현장감 있는 능동형 고품질 실



<그림 1> 지능형 실감 오디오 재생 플랫폼 기술 개념도

감 오디오 재현 가능.

- 청취자의 취향, 감성 정보 처리 및 사용자/오디오 정보 통합 표시 UI (User Interface) 를 통해 청취자 맞춤형 고품질 실감 오디오 서비스를 제공가능.
- 청취자의 위치 및 방향, 스피커의 위치 및 수 등 청취자와 연관된 다양한 환경정보를 획득하고, 이를 기반으로 다채널 오디오 스트림의 채널분배(Channel Distribution) 및 음원정위를 통해 청취자에게 최적의 오디오 서비스를 제공가능.
- 오디오 재생용 스피커 간에 다채널 실감 오디오를 고품질로 보장하기 위해 오디오 전송에 특화된 오디오 전용 가변 다채널 인터페이스를 통해 청취자 환경 기반 무선 다채널 오디오 스트림 채널분배를 지원함으로써 청취자에게 기기의 확장성 및 이동성, 관리의 용이성, 사용 공간의 쾌적성 등을 제공 가능.

이러한 기능을 제공 할 수 있는 지능형 실감 오디오 재생 관련 기술은 향후 <표 1>에 명시된 제품들에 적용 가능하리라 판단되며, 적용 및 판매 시점 이후로 국내 및 세계적으로 디지털 오디오 관련 시장의 가파른 성장을 가져오리라 예상된다.

위와 같은 오디오 서비스를 제공하는 미래 지향형 실감 오디오 재생 핵심 기술요소는 1) 사용자 능동형 음원 서비스를 위한 공간음향기술을 이용한 실감 오디오 획득/저장 및 부호화 기술, 2) 청취자의 취향이나 감성정보를 바탕으로 한 맞춤형 실감 오디오 재생 기술, 3) 청취자의 청각특성과 환경정보 기반 실감 오디오 전송 및 표현 기술 등으로 구성되며 각각의 기술들에 대해 자세히 살펴보기로 하자.

<표 1> 관련 기술의 활용방안 및 적용가능 제품

적용 산업	기존 제품	적용기술
가전	- Home Theater System - Speaker - DTV / 3DTV - High-end Audio	- 상황인지 기능 - Active Speaker 기술 - 3D 음장 제어기술 - 청각특성 음원 보정 기술
게임	- 게임기	- 3D 음장 제어기술 - 청각특성 음원 보정 기술 - 환경요소 반영기술
통신	- 휴대용 단말기	- 3D 음장 제어기술 - 청각특성 음원 보정 기술 - 환경요소 반영기술
콘텐츠 산업	- 오디오 콘텐츠 - 멀티미디어 콘텐츠	- 3D 오디오 모델링 기술 - 음원 센싱 나노소자 기술 - 객체기반 다채널 오디오 부호화 기술 - 3D 공간/음향 파라미터 부호화 기술

## 2. 사용자 능동형 오디오 서비스

지금까지의 오디오 서비스는 음원 개발자 중심의 오디오 서비스, 즉 보컬 및 모든 악기가 믹스된 단일음원이기 때문에 사용자는 단순히 오디오 음원 개발자나 음반 제작사가 발매한 단일 음원을 일방적으로 수동적 청취할 수밖에 없다. 하지만 사용자 능동형 오디오 서비스에서는 사용자가 능동적으로 자신이 원하는 음악적 취향에 따라 능동적으로 각각의 객체 기반의 독립 음원을 선택, 감성에 따른 음원 효과 추가, 최적의 음원 청취 위치(Sweet Spot) 변경, 음원 및 스피커 재생 공간 및 위치 변경 재생 등을 할 수가 있다.

현재 이와 같은 사용자 능동형 오디오 서비스는 ETRI (한국 전자 통신 연구원)와 (주)오디즌이 세계 최초로 사용자가 각각의 음원을 제어해 가수와 악기의 소리를 원하는 대로 개별 선택하여 감상할 수 있는 새로운 오디오 기술을 개발하여 이를 실제 음반에 적용한 뮤직 2.0기술을 상



용화하여 실제 음반을 제작 발매하여 서비스 하고 있다<sup>[1]</sup>. 뮤직 2.0은 차세대 신개념 오디오 엔진으로 가수의 목소리와 악기를 개별 선택하여 감상할 수 있는 것이 특징으로 보컬, 피아노, 기타, 드럼 등 자신이 좋아하는 음원 소리를 선택 및 청취가 가능하다. (참고: 웹 2.0이 UCC와 블로그 같이 누구나 손쉽게 데이터를 생산하고 인터넷에서 공유할 수 있도록 한 사용자 중심의 인터넷 세상을 일컫듯이 뮤직 2.0도 능동적이고 이용자 참여가 가능한데서 붙여진 이름이다.) 이 뮤직 2.0을 이용하면 반주음을 모두 없애고 보컬만을 무반주로 감상해 볼 수도 있으며, 마이크와 연결하여 전문 악기 연주자들이 연주한 음악에 본인 목소리를 같이 녹음시켜 별도의 외부음원으로도 재생산이 가능하다.

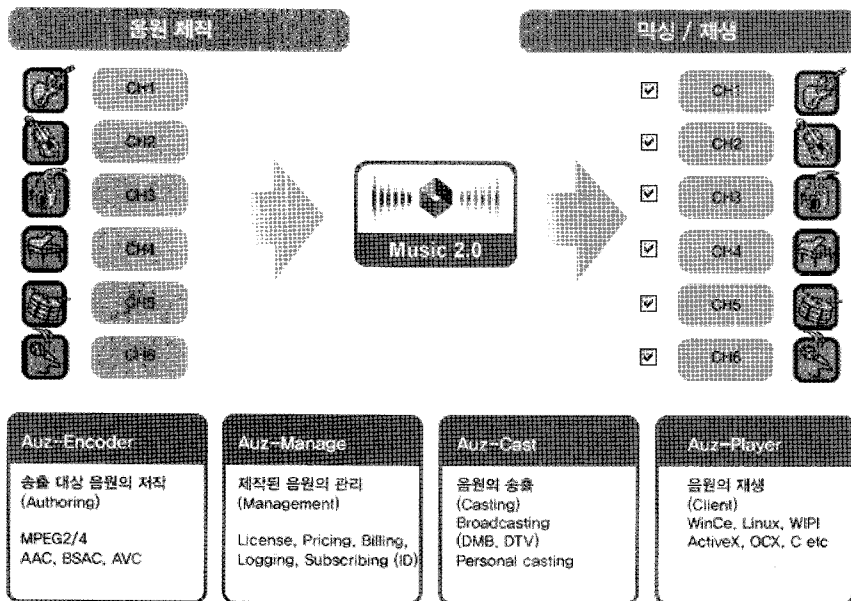
현재 ETRI와 오디즌은 뮤직 2.0 기술과 관련하여 차세대 음악 포맷인 IMAF (Interactive Music Audio Format) 으로 MPEG에서 표준화



〈그림 3〉 뮤직 2.0 플레이어 재생 화면

진행 중에 있다<sup>[2]</sup>.

뮤직 2.0과 유사하게 프랑스의 iKlax 에서도 멀티트랙 음원 재생과 관련한 Audio format 개발 및 MPEG에서의 표준화 진행, 멀티트랙 음원 제작물 제작, 멀티트랙 음원 재생이 가능한 휴대

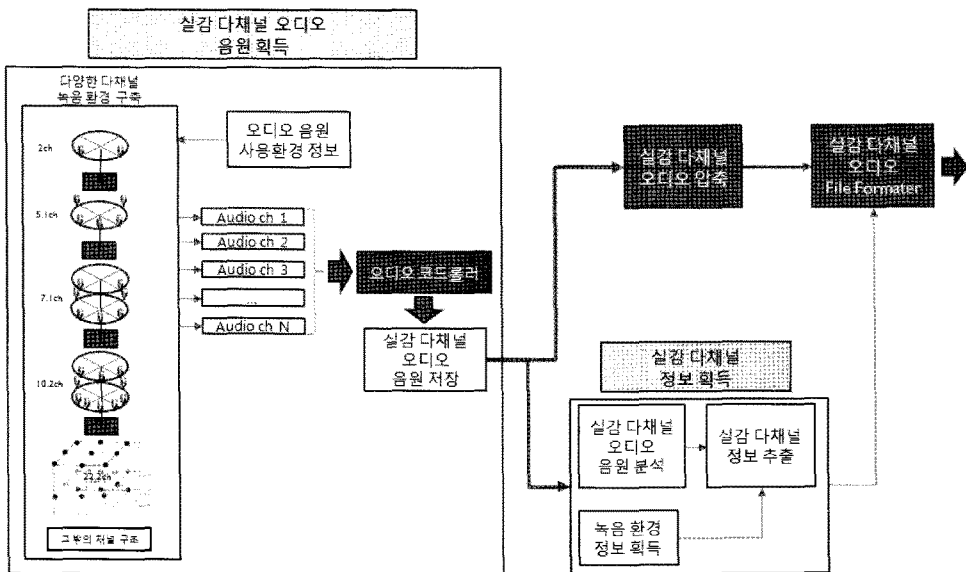


〈그림 2〉 뮤직 2.0 개념도

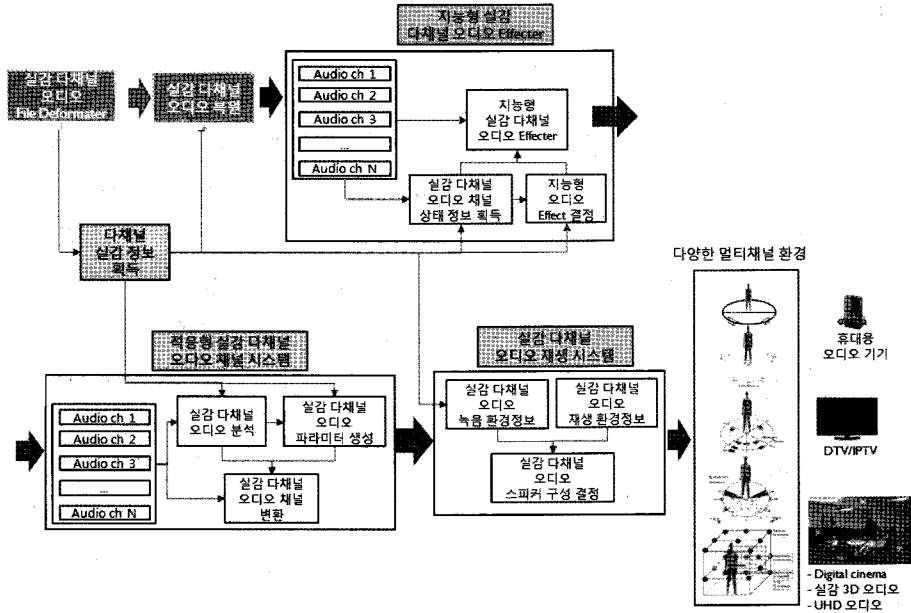
용 기기 개발을 하고 있다<sup>[3]</sup>.

뮤직 2.0과 iKlax 미디어가 단순히 독립음원을 재생기에서 제어하여 몇몇의 독립 음원을 조합하여 능동적으로 사용자가 음악을 선택 재행하는 서비스라면 지능형 실감 오디오 재생 기술의 사용자 능동형 오디오 서비스는 뮤직 2.0 및 iKlax에서의 독립음원 제어를 통한 능동형 재생 서비스 뿐만 아니라, 스테레오 / 5.1채널 오디오를 포함한 다양한 실감 다채널 오디오를 취득하여 다채널 오디오가 갖는 2차원 / 3차원 공간 파라미터, 실감 다채널 오디오 파라미터 등을 분석하여 지능적으로 다채널 오디오를 컨트롤하고, 다채널 오디오를 사용자의 청취 환경 따라 조절함으로써 어떠한 사용자 청취 환경에서도 최상의 실감 다채널을 제공한다. 실감 다채널 오디오 파라미터 추출은 실감 다채널 오디오 녹음 환경 정보, 기존 평면 서라운드를 위한 2차원 공간 파라미터 channel level difference (CLD), inter-channel correlation (ICC), channel pre-

diction coefficient (CPC), 그리고 3차원 서라운드를 위한 3차원 공간파라미터 inter-channel amplitude difference (IAD) 등을 추출하여 지능형 실감 다채널 오디오 효과와 적응형 실감 다채널 오디오 채널 변환에서 실감 다채널 오디오의 공간감/사실감 극대화를 지원한다. 지능형 실감 다채널 오디오 효과 기술은 다채널 오디오 오디오의 각 채널 오디오가 갖는 sound delay, channel gain, pitch position, energy difference 등의 오디오 특성을 분석하여 전체 오디오가 하모니를 이루도록 컨트롤하며, 오디오 채널 믹싱을 수행할 경우 각 오디오 채널의 특성을 고려한 오디오 생성을 통해 잡음 발생 및 오디오 품질 열화를 막는 기능을 수행하며, 적응형 실감 다채널 오디오 채널 변환 기술은 다양한 사용자의 청취 환경에 능동적으로 대처할 수 있도록 사용자의 청취 환경에 맞추어 오디오 오디오의 채널을 확장하거나 축소하는 기술로 실감 파라미터가 적용되어 사용자에게 실감 다채널 오디오



<그림 4> 실감 오디오 재생을 위한 다채널 오디오 획득 및 정보 해석 분석도



〈그림 5〉 실감 다채널 오디오를 위한 지능형 오디오 시스템 및 재생부 블럭도

의 생동감을 확대한다.

### 3. 사용자 능동형 실감 오디오 지원 오디오 포맷

앞서 사용자 능동형 서비스 지원을 위하여 ETRI와 (주)오디즌이 공동 연구 개발한 뮤직 2.0, MPEG에서 표준화 준비중인 IMAF와 iKlax에서 준비 중인 iKlax Audio 포맷에 대해 언급 보았다. 이와 유사하게 사용자 정보 기반 오디오 서비스를 제공하기 위해 미디어 파일 포맷에 대한 연구가 활발히 진행 중인데 이와 관련한 기술들의 동향을 잠시 살펴본다.

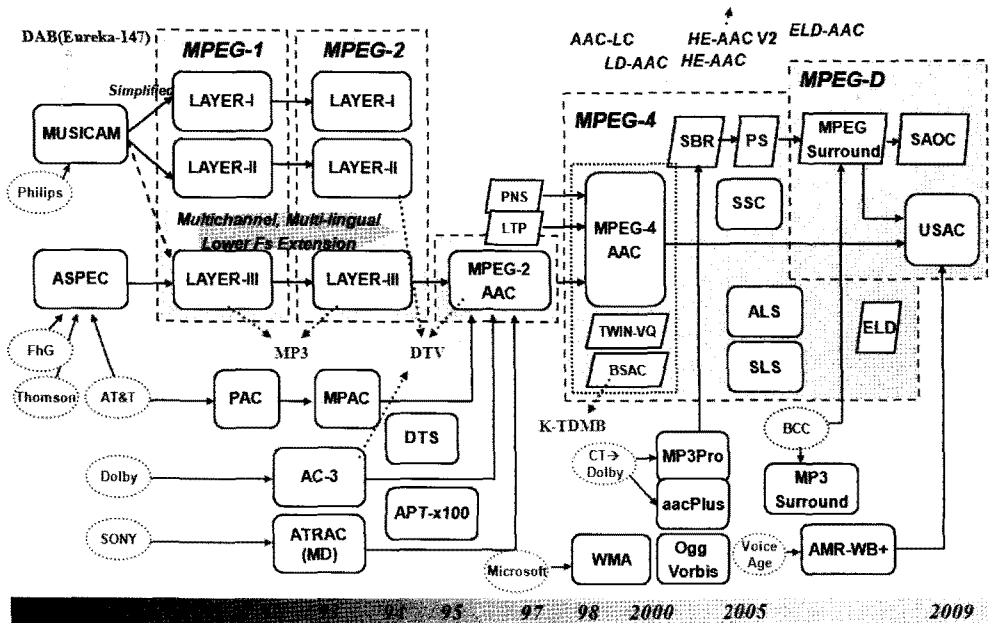
#### 가. 오디오 파일 포맷 기술 동향

MPEG에서는 미디어 파일 포맷의 표준으로써 MPEG-4 파일 포맷<sup>[4]</sup>과 MPEG-21 파일 포맷<sup>[5]</sup>을 각각 제정하고 있다. MPEG-4 파일 포맷

은 오디오나 비디오 등의 미디어 데이터를 효과적으로 저장하고 인터넷에서 스트리밍을 제공하기 위한 파일 포맷으로써 1998년부터 표준화 작업이 시작되었다. MPEG-4 파일 포맷은 세부적으로 오디오 데이터, 비디오 데이터, 메타 데이터, 장면묘사 데이터 등의 부분으로 구성된다. 여기서 오디오나 비디오 데이터는 MP3(MPEG-1 audio layer 3)나 H.264의 비트스트림과 같은 정보를 각각 의미한다. 메타데이터는 미디어의 압축 포맷, 길이, 채널 수, 재생 정보, 편집 정보 등을 제공하여 오디오나 영상을 목적에 맞게 재생하거나 편집하는 등의 정보로 사용된다. 그리고 장면묘사 데이터는 2차원이나 3차원의 오디오나 영상 콘텐츠를 객체별로 묘사하기 위한 정보를 의미한다. MPEG-4 파일 포맷은 비디오, 오디오 등의 다양한 객체를 조절할 수 있도록 객체를 조합해 장면을 묘사해줄 수 있는 영역과 양방향 서비스를 제공할 수 있는 영역을 포함하고

있다. MPEG-4 파일 포맷은 애플 컴퓨터의 QuickTime 파일 포맷을 기반으로 개발되었다. QuickTime 파일 포맷의 구조는 다양한 동영상이나 음성의 비트스트림을 유연하게 다중화할 수 있어 범용의 파일 포맷으로서 ISO (international organization for standardization) 베이스 미디어 파일 포맷<sup>[5]</sup>에 채용되어 MPEG-4 파일 포맷 등 기타 포맷의 근간으로 사용되고 있다. 1998년 ISO는 ISO 베이스 미디어 파일 포맷을 기본으로 MPEG-4 시스템에 필요한 부분을 확장한 형태의 MPEG-4 파일 포맷을 정의하기로 결정하고 2000년에 MPEG-4 파일 포맷을 국제표준으로 최종 승인하였다. MPEG-4 파일 포맷은 표준화 승인 이후부터 동영상을 장시간 기록하는 용도로 디지털 카메라에 사용되었다. 근래에는 제3세대 휴대 전화의 동영상 포맷으로서 채용되고 PDA를 포함한 모바일 동영상 재생기의 주요 포맷이 되어 왔다. 하지만 MPEG-4 파일 포

맷은 MPEG-2가 DVD나 디지털 방송 등 여러 분야에 걸쳐서 활용되고 있는 것과는 달리 MPEG-4 표준의 내용이 매우 복잡하고 너무 높게 책정된 로열티 등의 이유로 근래까지는 다양한 분야에서 폭넓게 활용되고 있지 않았다. 하지만 최근에는 MPEG-4/H.264 AVC (Advanced Video Coding)가 일본 지상파 디지털 방송의 휴대 단말기를 위한 방송으로의 채용되고, Blu-ray Disc나 HD DVD의 비디오 코덱으로서 승인이 되고 있는 등 응용분야가 늘어나고 있는 추세이다. 또한 모바일 응용 분야 등에서 기존의 MPEG-4 표준의 일부 기술을 바탕으로 특정한 분야의 응용에 최적화된 기술을 추가하여 새로운 프로파일(profile)을 제정하려는 연구가 진행 중이다. 이에 따라 MPEG-4 파일 포맷 역시 활용 분야에 맞게 조정되어 사용해야 하기 때문에 포맷에 대한 많은 연구가 진행되고 있다



<그림 6> 고품질 오디오 포맷의 진화도

#### 나. Scene Description 기술 동향

정보통신 기술의 발전에 따라 사용자는 이제 단순 오디오/비디오 위주의 멀티미디어 서비스에서 여러 가지의 미디어가 혼합하여 제공되는 리치미디어(rich media) 서비스를 향유하고 있다.<sup>[6]</sup> 이러한 오디오, 비디오, 그래픽 및 텍스트들로 구성된 멀티미디어 데이터를 통해 보다 다양하고 풍부한 대화형 서비스를 제공하고자 하는 리치미디어 서비스를 위하여 비디오, 오디오, 텍스트, 애니메이션 등을 모두 포함하는 리치 장면과 함께 사용자와의 상호작용(interactivity), 콘텐츠의 다운로드와 동시에 재생이 가능한 스트리밍(streaming) 서비스, 리치 장면의 동적갱신(dynamic update), 다양한 미디어 소스를 지원하기 위한 확장성(extensibility), 이동 단말에서 리치미디어 서비스를 지원하기 위한 규격의 경량성 등이 고려되어야 한다. 이러한 리치미디어 서비스의 구성 요소를 고려한 장면묘사 연구 및 개발 역시 활발하게 진행되고 있다.

장면묘사 서비스를 제공하는 주요 기술로써 Flash<sup>[6]</sup>는 de-facto 표준으로 여겨질 만큼 리치미디어 활용 시장에서 그 활용도는 매우 높다. Macromedia에서 개발한 Flash는 비디오, 오디오, 텍스트, 애니메이션 등 리치 장면을 표현할 수 있을 뿐만 아니라 기본적인 사용자와의 상호작용이 가능하다는 장점이 있지만, 당초 PC 환경을 기반으로 개발되었기에 모바일 환경에는 적합하지 않아 다양한 모바일 단말에 대응하기 위하여 Flash Lite라는 모바일 버전이 또한 새로이 소개되었다<sup>[7]</sup>. 또한 국제표준규격으로 Flash와 같은 리치미디어 서비스를 제공하기 위한 기술로써 국제표준 기구인 W3C (world wide web consortium)의 SMIL (synchronized multi-

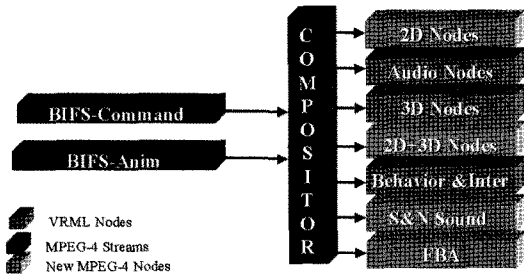
media integration language)<sup>[8]</sup>과 SVG (scalable vector graphics)<sup>[9]</sup> 기술이 있다. SMIL과 SVG는 모바일 환경에 적합한 프로파일을 따로 정의하여 3GPP (the third generation partnership project)와 OMA (open mobile alliance)에서 일부 사용되고 있으며, 두 기술은 download-and-play 또는 progressive download-and-rendering 서비스를 고려하여 설계되었기 때문에 이와 같은 서비스에 제공하는데 적합하다. 또 다른 국제규격인 MPEG-4 BIFs는 리치미디어 콘텐츠를 제작하기 위한 composition coding 기술로서 2D와 3D를 모두 사용한 멀티미디어 제작, 해당 장면에 대한 순차적 업그레이드, 리치미디어의 스트리밍 지원 및 서로 다른 구성 미디어간의 동기화를 지원하고 있을 뿐만 아니라 사용자와의 상호작용 서비스도 지원한다.

상기 Flash, SVG 및 BIFs 등은 PC 환경에서 개발되었기 때문에 모바일 기기에 적용하기에는 과다한 연산량을 갖는 등의 문제점이 존재한다. 이러한 기술적인 제약 사항들을 해결하고 de-facto 표준보다는 국제표준과 같은 open standard의 제정을 통해 보다 폭넓은 시장 확보를 위해서 MPEG에서는 모바일 환경에 적합한 리치미디어 콘텐츠 규격인 LAsER 표준을 제정하였다. LAsER는 모바일 환경을 위하여 상호작용 및 스트리밍, dynamic update등을 고려한 경량의 장면묘사 기술이라 할 수 있다. 이러한 상호작용성 및 스트리밍, dynamic update를 모두 고려한 장면묘사 기술인 BIFs와 LAsER등이 있다.

#### 다. MPEG-4 BIFs, LAsER

BIFs는 binary format for scenes의 약어로서, 직역하면 장면이진부호화를 의미한다<sup>[1]</sup>. 최





〈그림 7〉 MPEG-4 scene description

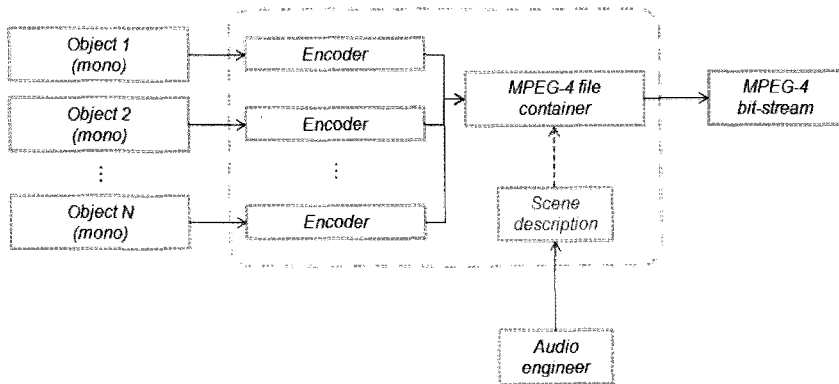
근 방송 기술의 디지털화로 인해 기존의 단순 오디오 및 비디오 중심의 방송 서비스를 탈피하여 음성, 영상, 이미지, 텍스트 등의 다양한 멀티미디어 서비스 제공과 이들 간의 대화형 기능을 제공하는 리치미디어 서비스가 각광을 받고 있다. 이러한 필요성에 따라 MPEG-4는 장면의 양방향 실시간 비디오 데이터 서비스 또는 연동형 데이터방송 구현 기술로써 BIFs를 고안하게 되었다. 즉, BIFs는 2차원 내지 3차원의 음성/영상 콘텐츠의 합성, 저장, 재생을 위한 바이너리 형식을 규정한 것으로써 BIFs를 통해 프로그램과 콘텐츠 데이터베이스가 원활하게 연동될 수 있게 된다. 예를 들어 BIFs는 한 장면에서 어떤 자막을 삽입할 지, 그림을 어떤 형태로 포함할 지, 몇 초 간격으로 얼마동안 그림 등이 나올지를 기술

한다. 또한 특정 장면에 대하여 상호작용을 위한 이벤트의 정의 및 이벤트 처리를 통해 사용자가 BIFs를 통하여 렌더링되는 객체와 상호작용을 할 수 있다.

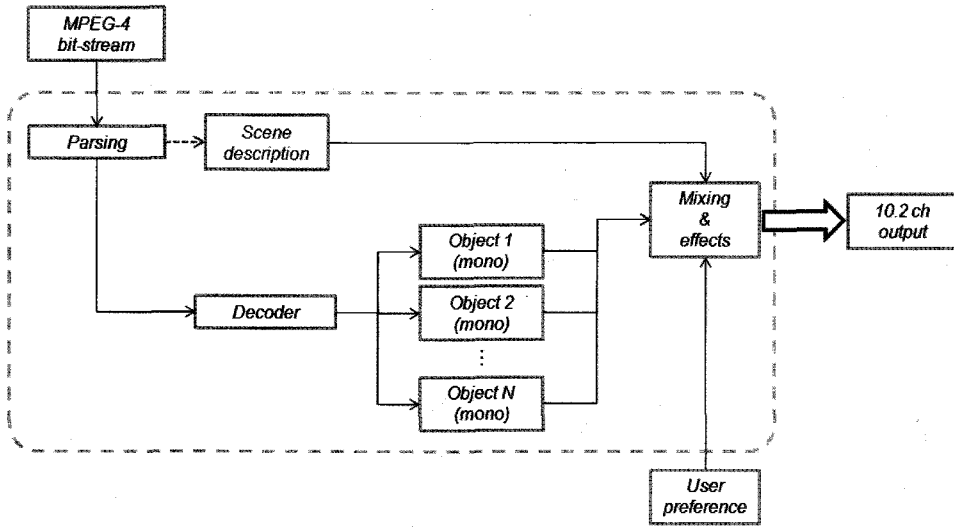
라. 사용자 능동형 실감 오디오 지원 오디오 정의

미래 지향형 실감 오디오 재생 플랫폼에서도 다양한 어플리케이션에 적용이 가능하고, 사용자 능동형 고품질 실감형 오디오 서비스에 적합한 MPEG-4 ALS 기반 실감 오디오 무손실 오디오 포맷 및 복호화기를 지원한다. 이러한 특성의 이 오디오 포맷은 고품질 오디오 오디오를 지원함은 물론 청취자의 청각적 특성, 청취 환경(사용자 스피커 관련 정보 및 위치 정보등), 청취자 감성정보, 객체 오디오 구성 정보, 공간 오디오 정보들을 XML 형식으로 표현하여 오디오 데이터에 포함시키고, 이를 재생단계에서 해석하여 음원에 음향효과를 주어 청취자로 하여금 현장감 있는 오디오를 들을 수 있도록 한다.

〈그림 8, 9〉는 사용자 능동형 실감 오디오 지원을 위한 오디오 포맷의 부/복호화기를 보여주고 있다. 부호화된 음원 개체, 사용자 능동지원



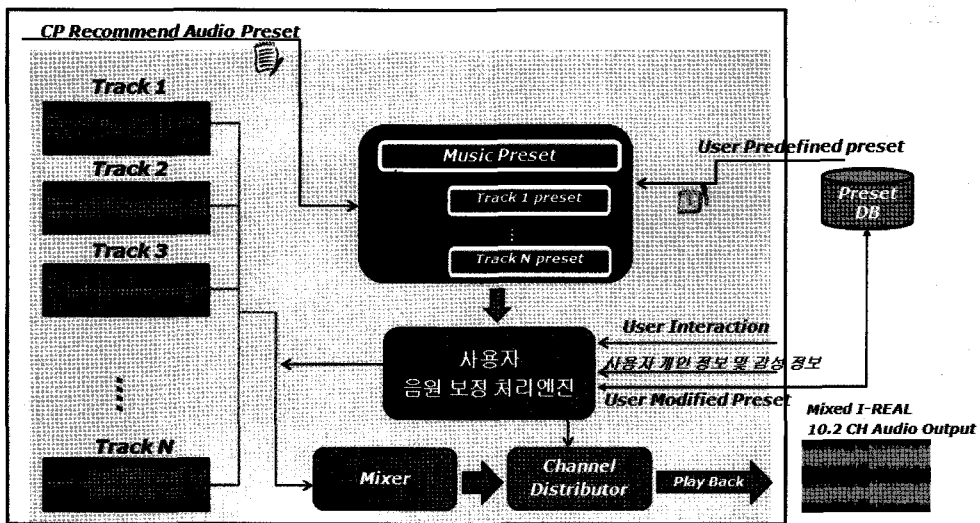
〈그림 8〉 사용자 능동형 실감 오디오 지원 오디오 포맷 부호화기 구조



〈그림 9〉 사용자 능동형 실감 오디오 지원 오디오 포맷 복호화기 구조

을 위한 메타데이터 정보, 장면묘사 정보등이 MPEG-4 파일 포맷에 저장되어 하나의 비트스트림을 형성하게 되고, 이를 복호화기가 해석하여 사용자 정보 및 장면 묘사 정보를 이용해 사용자는 출력되는 음원을 자신이 원하는 대로 조절하여 바뀌가며 여러 실감 음향 효과를 주게 된다.

이와 같은 오디오 포맷을 이용한 사용자 능동형 실감 오디오 재생 서비스는 객체 단위의 독립 음원의 제어 및 재생, 수신단에서 청취자가 장면으로부터 느끼는 방향감 및 공간감을 통해 현실감을 증대 시키는 입체 음상 정위, 음원에 잔향과 같은 특정 공간 특성을 반영한 가상공간효과를 처리해줌으로써 청취자는 음원으로부터 거리감



〈그림 10〉 사용자 능동형 실감 오디오 재생 개념도

및 공간감을 인지하게 할수 있는 가상 공간 효과 재현등의 서비스가 가능해진다.

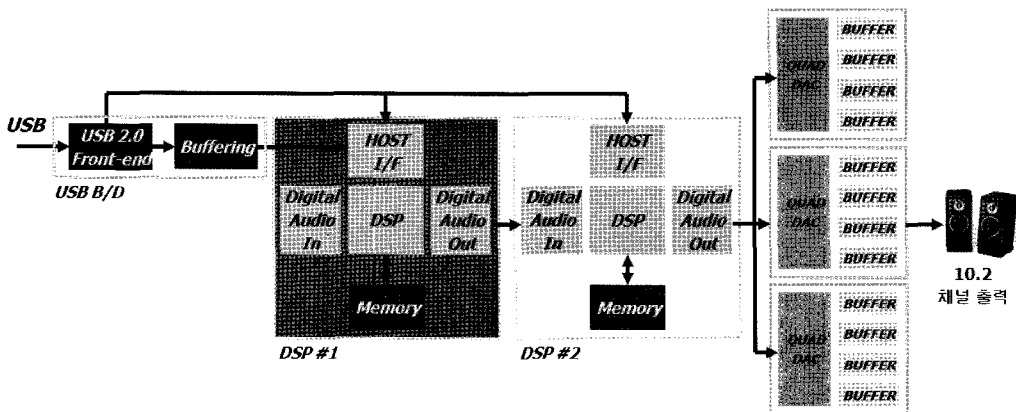
#### 4. 지능형 실감 오디오 재생 플랫폼

지능형 실감 오디오 재생 플랫폼은 사용자가 음악을 능동적이고 실감있게 청취 할 수 있도록 실질적으로 객체 단위의 무손실 독립 음원 처리, 사용자 감성 및 취향, 음원 및 청취 환경의 공간 정보 처리 및 3D 입체 음향 정보의 데이터베이스를 해석하여 그 결과를 바탕으로 실감 음원 재현, 통합 정보 표시 장치를 통한 사용자 제어 명령과 함께 이들을 처리하여 다채널 고품질 무선 전송 모듈로 오디오 출력하는 플랫폼이다. 이와 같은 플랫폼 기능을 구현하는데 있어 다양한 음원 처리를 위해 <그림 11>과 같이 2중 DSP 처리를 하였고, 전처리 DSP에서는 입력된 사용자 능동형 실감 무손실 오디오를 복호화, 제어기로부터 해석된 음원 처리 신호를 이용해 사용자 음원 보정을 해주는 DSP Filter Stage를 구성하였고, 후처리 DSP에서는 다양한 음원 Effect (3D 입체 음향 처리, 입체 음상 정위 처리, 가상 공간 효과 처리 등) 를 처리한다. 또한 사용자의

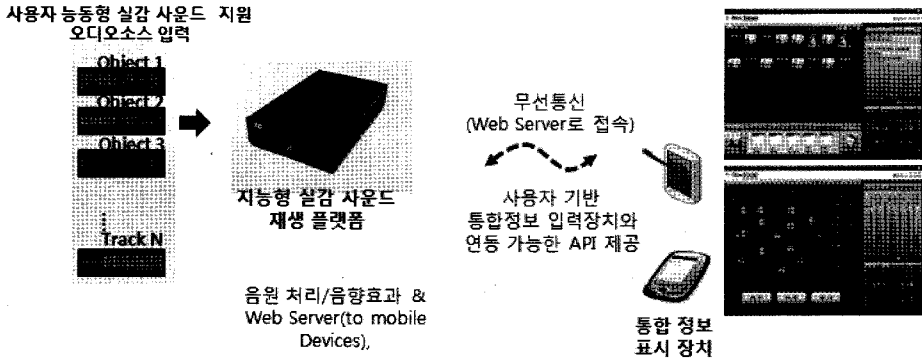
실감 오디오 제어의 편의성을 제공하는 통합 정보 표시 장치와의 연동하기 위해 웹서버를 내장하여 다양한 정보와 응용이 가능한 인터페이스를 제공한다.

#### 가. 통합 정보 표시 장치 활용

단순히 오디오를 선택, 재생, 멈춤 등의 기능만을 수행하는 일반 오디오 기기와는 다르게 지능형 실감 오디오 재생 플랫폼은 사용자의 감성 / 취향 선택, 공간 정보 확인 및 선택, 3D 입체 음향 효과등 음악 이외의 다양한 실감 오디오 정보 등을 제공 할 수 있고 사용자는 이들을 확인 및 선택하여 자신의 취향에 맞게 실감 오디오를 감상 할 수 있다. 하지만 기존의 단순한 기능의 리모콘으로는 지능형 실감 오디오 재생 플랫폼이 제공하는 다양한 기능 수행하기가 불가능하다. 따라서 통합 정보 표시 장치를 활용하여 사용자는 플랫폼이 제공하는 실감 오디오에 대한 정보, 사용자의 취향이나 감성정보를 바탕으로 한 사용자 보정 음원 정보, 이를 통한 실감 오디오 제어가 가능하게 되어 보다 편리하게 실감 오디오를 인지 하고 편리하게 조정, 감상이 가능한 사



<그림 11> 지능형 실감 오디오 재생 플랫폼 Hardware 구조도



〈그림 12〉 통합 정보 표시 장치 활용도

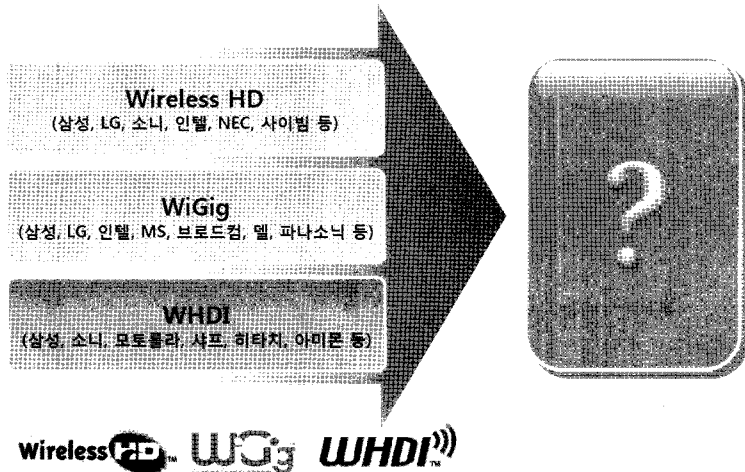
용자 인터페이스를 (UI)를 제공하는 장치이다.

최근 대부분의 개인들이 휴대폰, iPOD 터치(폰), PDA 등등의 고성능의 휴대용 멀티미디어를 1대 이상은 보유하고 있는 상황에서 이러한 통합 정보 표시 기능을 기존의 개인이 보유한 기기에서 별도의 경제적 부담 없이 손쉽게 활용 가능하다. 이러한 통합 정보 표시 장치의 활용은 비단 지능형 실감 오디오 재생 서비스뿐만 아니라 다양한 정보와 기능 및 사용자 어플리케이션을 제공해 주는 멀티미디어 기기 및 가전제품 등에 활용하여 사용자 편의 증대를 도모 할 수 있다.

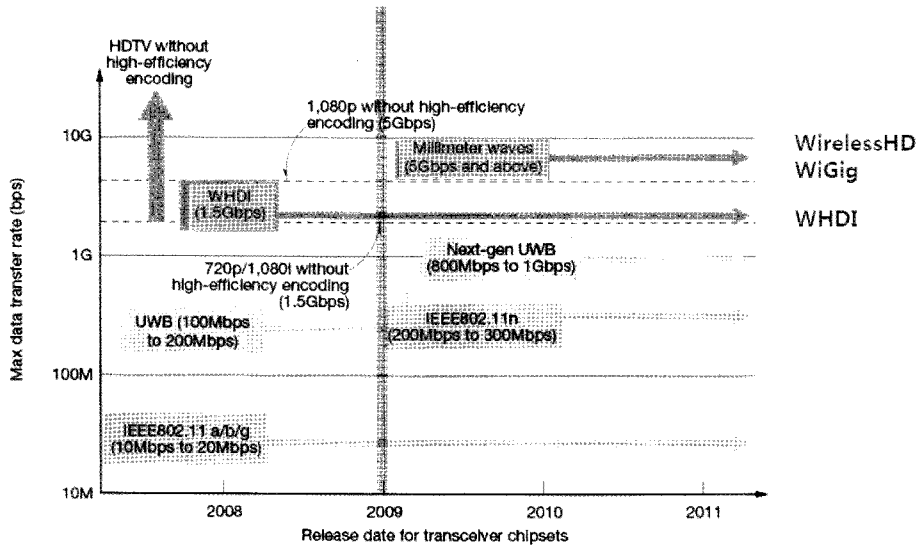
### 5. 다채널 고품질 실감 오디오 무선 전송 기술

#### 가. 개요

전자 기기의 디지털화 및 무선화 추세속 에서도 가정용 오디오 전송 및 재생 분야는 최근 들어 일부 무선화를 진행 중이지만 거의 유일하게 유선 아날로그 방식을 유지하고 있다. 가정에서 주로 사용하는 5.1 채널 홈시어터만 보더라도 사방에 설치된 6개의 스피커에 유선을 연결해야 하



〈그림 13〉 무선 A/V 전송 기술별 참여 업체 현황



〈그림 14〉 무선 A/V 전송 규격 표준화 진행상황

며, DVD, CD 플레이어 등등 다른 오디오 가전기기들의 유선까지 더해져 많이 혼잡해진다. 이에 전 세계적으로 많은 주요 가전업체를 위주로 고품질의 입체 오디오를 보장하며 재생 플랫폼과 스피커 간에 전송을 위해, 최근에는 고품질 디지털 데이터 처리와 스피커 운용상의 편리성을 고려하여 이를 무선 인터페이스로 구현하려는 연구 개발에 박차를 가하고 있고, 각종 A/V 디지털 기기들을 무선으로 연결하여 쓸 수 있는 시대가 열리면서 자사의 관련 기술을 국제 표준으로 완성하기 위해 치열하게 경쟁하고 있다. 관련 자세한 설명 및 동향은 이번 특집호의 “무선 오디오 전송 기술 동향”에 설명되어져 있다.

#### 나. 다채널 고품질 실감 오디오 무선 전송 기술

지능형 실감 오디오 재생 기술에서의 다채널 실감 오디오는 10.2 채널 192kHz/24비트로 무선 전송 되어 실감음이 재생된다. 하지만 다채널

오디오 인터페이스 분야에서는 가전기기 관련 대기업에서 기존의 무선 네트워크 표준을 기반으로 오디오 전송에 적용하거나 일부 소형 업체에서 오디오 전송용의 단순 무선 연결 기술을 이용하고 있어 고품질의 실감 오디오를 보장하지 못 한다. 따라서 이를 보장하기 위해서는 특화된 사운드 전용 가변 다채널 인터페이스 및 여러개의 스피커를 동시에 재생할 수 있는 동기 기술을 적용하여 통해 청취자 환경 기반 다채널 사운드 스트림 채널분배를 지원함과 동시에 추후 청취자에게 기기의 확장성 및 이동성, 관리의 용이성, 사용 공간의 쾌적성 등을 제공할 수 있다.

### III. 결론

본 기고에서 소개된 지능형 실감 오디오 재생 기술을 통해 사용되는 다양한 환경에서 취득된 고음질 오디오 콘텐츠를 사용자 특성, 감성 / 취향 / 재생환경 등을 기반으로 한 실감 오디오 처

리를 통해 실제 환경에서 느끼는 현장 오디오를 감상 할 수 있다. 객체 기반의 독립 음원과 이와 관련한 사용자 감성/취향 정보, 공간 및 입체 음향 정보 등이 믹스된 음원을 사용자는 통합 정보 표시 장치를 통해 실감 오디오에 대한 많은 정보를 확인 / 제어 할 수 있고 이를 바탕으로 실감 오디오에 대해 다양한 응용이 가능하다.

이러한 실감 오디오 서비스는 지금의 정형화된 음악서비스에서의 단순히 음악의 볼륨만을 조절할 수 있는 음악 서비스에서 벗어나 청취자 또는 사용자가 직접 음악 콘텐츠에 실려진 다양한 오디오 음원 소스의 특징을 선택적으로 조절하여 자신만의 새로운 오디오 콘텐츠를 생성 할 수 있는 새로운 개념의 음악 서비스의 한 예가 될 수 있다. 이에 대한 연구가 계속 발전되면 오디오 음원 생산자, 재생기 생산자, 사용자에게 모두 수익과 흥미를 가져다 줄 뿐만 아니라 더 나아가 음반 시장 및 오디오 관련 시장의 큰 변화를 줄 것 이라 예상된다.

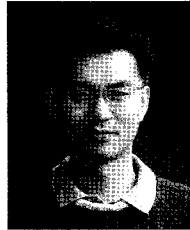
#### 참고문헌

- [1] 홍진우, 강경욱, ETRI, 음향 선별 UCC 오디오 기술 개발, 보도자료, 2007.10
- [2] Inseon Jang, Jeongil Seo, Kyeongok Kang, Hui Yong Kim (ETRI), Kevin Seung Chul Ham (Audizen Inc), MPEG2008 / M15626, A proposal for technical specification of Interactive Music AF.
- [3] www.iklax.com
- [4] ISO/IEC 14496-14:2003, Information technology - Coding of audio-visual objects - Part 14: MP4 file format, Nov., 2003.
- [5] ISO/IEC 21000-9, Information technology -Multimedia framework (MPEG-21), Part 9: File Format, July, 2005.
- [6] Adobe Flash, <http://www.adobe.com/products/flash>.
- [7] Adobe Systems, Flash Player for Mobile Devices Delivers High-Impact Video and Dynamic Web Content white paper, Oct., 2007.
- [8] W3C, Synchronised Multimedia Integration Language(SMIL 2.0) - [Second Edition], <http://www.w3c.org/TR/2005/REC-SMIL2-20050107>, Jan., 2005.
- [9] W3C, Scalable Vector Graphics(SVG) 1.1 Specification [Recommendation], <http://www.w3c.org/TR/2003/REC-SVG11-20030114>, Jan., 2003.
- [10] S. G. Chen, R. Jiang, "A new fast filtering algorithm based on algebraic composition," in proc. of SiPS 99 1999 IEEE Workshop, pp.742-750, 1999.
- [11] 조충상, 김제우, 신화선, 최병화, "MPEG-4 ALS 복호화기를 위한 고성능 필터링 방법," 2008 전자공학회 추계학술대회논문집, 2008년 11월.
- [12] C.S. Cho, J.W. Kim and B.H. Choi, "A Low Complexity MPEG-4 ALS Coding for High Quality Object Audio System," IEEE Trans. CE, Dec., 2008 Submitted.
- [13] K. Park, J. Seo, J. Wee, W. Jeon, J. Paik, "Wireless audio transmitting

apparatus, speaker and system and controlling method”, EP08173036.8(유럽 출원 특허)

- [14] D.H. Kim, K.H. Kim, S. Liu, and J. H. Kim, “A TMO-Based Approach to Tolerance of Transmission Jitters in Tele-Audio Services”, Computer System Science & Engineering, 17(6), Nov., 2002, pp.325-333.

### 저자소개



임 태 범

1995년 2월 서강대학교 물리학과  
1997년 2월 서강대학교 전자계산학과 대학원  
1997년 1월~2002년 7월 대우전자 영상기술 연구소  
2009년 12월~현재 전자부품연구원

주관심 분야 : Digital 방송, Digital 데이터 방송, 맞춤형 방송, IPTV, IP Streaming, Audio Codec

### 저자소개



정 종 진

1997년 2월 성균관대학교 공과대학 학사  
2002년 8월 성균관대학원 공과대학 석사  
2002년 6월~현재 전자부품연구원

주관심 분야 : Digital TV Solution, Audio Codec, DRM



이 석 필

1990년 2월 연세대학교 전기공학과  
1992년 8월 연세대학교 전기공학과 학사  
1997년 8월 연세대학교 전기공학과 공학박사  
1997년 7월~2002년 2월 대우전자 영상연구소 팀장  
2009년 12월~현재 전자부품연구원 디지털미디어연구 센터 센터장

주관심 분야 : 디지털방송통신융합시스템, A/V 신호처리