

UHDTV를 위한 차세대 오디오 표준: MPEG-H 3D Audio

□ 이태진, 백승권 / ETRI

I. 서론

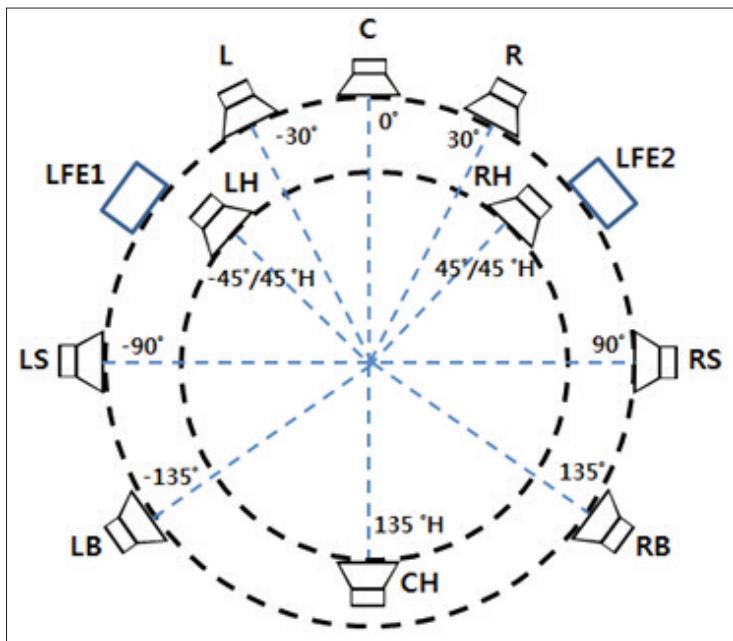
방송 환경이 점차 고화질로 발전해 가면서, HDTV를 넘어서는 UHDTV와 같은 차세대 방송서비스에 대한 관심이 증대되고 있다.

UHDTV를 위한 오디오 서비스는 몰입감(Immersive)과 개인화(Personalization)라는 키워드로 정의할 수 있다. 몰입감은 <그림 1>과 같이 10.2채널[1]과 같은 수직면까지 제공하는 다채널 오디오 기술, 헤드폰 환경에서 3D 입체음향을 재생하는 기술 등을 활용하여 기존 스테레오나 5.1채널 서비스를 넘어서는 현장감과 사실감을 제공해 주는 것을 의미한다. 개인화는 방송 환경에서 개인의 취향에

따른 오디오 제어 및 선택이 가능한 서비스로 스포츠 중계에서 아나운서/해설자의 목소리만을 크게 하여 좀 더 명료한 방송을 청취하거나, 관중과 현장의 소리를 크게하여 좀 더 현장감 있는 방송을 청취 할 수 있다[2].

UHDTV 오디오 서비스를 위한 표준 기술은 MPEG의 MPEG-H 3D Audio와 Dolby의 AC-4 등이 있다. 두 기술 모두 몰입감과 개인화라는 UHDTV 오디오 서비스를 제공할 수 있지만, 본 논문에서는 차세대방송표준포럼에서 선정한 국내 지상파 UHDTV 서비스를 위한 오디오 표준인 MPEG-H 3D Audio에 대해서 소개하고자 한다.

* 본 원고는 2016년도 정부(미래창조과학부)의 재원으로 정보통신기술진흥센터의 지원을 받아 수행된 연구임 (No.B0101-16-0295, 초고품질 콘텐츠 지원 UHD 실감방송/디지털시네마/사이니지 융합서비스 기술 개발)



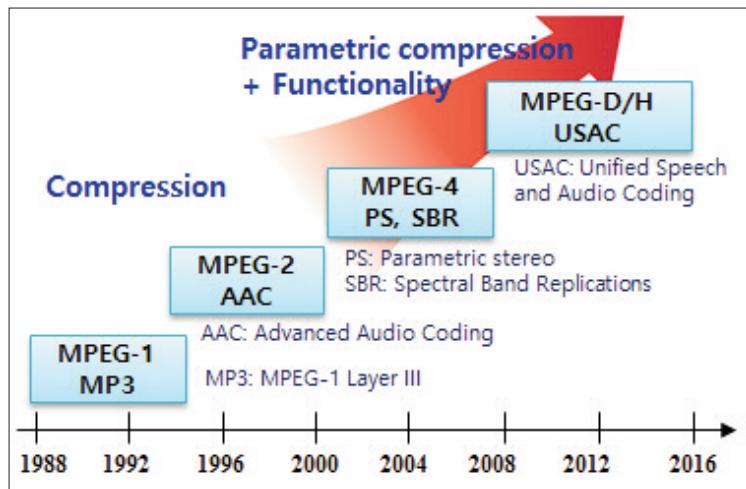
<그림 1> 10.2채널 스피커 구성

II. MPEG 오디오 표준

오디오 압축 기술은 디지털 오디오 신호를 음질 손실을 최소화하면서 데이터량을 줄이는 기술로 1980년대 후반부터 기술개발 및 표준화를 진행해 오고 있다. MPEG에서는 1992년 MPEG-1 Audio 표준(ISO/IEC 11172-3)[3]을 시작으로 2016년 하반기 표준화가 예정되어 있는 MPEG-H 3D Audio (ISO/IEC 23008-3)[4]까지 다양한 압축률과 기능을 제공하는 오디오 압축 기술에 대한 표준화를 진행해 왔다.

<그림 2>는 MPEG의 오디오 압축 기술에 대한 역사를 간략하게 정리한 것이다. 1.41 Mbit/s의 정보량을 가지는 CD를 192kbps로 압축하기 위한 MPEG-1 오디오 표준은 우리에게 MP3라는 표준으로 잘 알려져 있으며, DTV 방송과 모바일 오디오

재생장치 등에 많이 활용되고 있는 MPEG 오디오 표준 중 가장 성공적으로 시장에 서비스되고 있다. AAC(Advanced Audio Coding)이라는 표준명으로 잘 알려져 있는 MPEG-2 오디오 표준(ISO/IEC 13818-7)은 128kbps로 CD를 압축하기 위한 기술로 DAB, DRM 등에서 오디오 표준으로 서비스되고 있다[5]. 고주파 대역이나, 스테레오에 대한 정보를 파라미터로 표현하여 64kbps까지 CD를 압축하기 위한 MPEG-4 오디오 표준(ISO/IEC 14496-3)은 HE-AAC(High-Efficiency AAC) v1/v2로 알려져 있으며 DVB, DAB 등에서 사용되고 있다[6]. 스테레오 기반의 오디오 서비스에서 5.1채널과 같은 다채널 오디오 서비스가 확장되어 가면서, 다채널 오디오에 대한 효율적인 압축 기술에 대한 시장의 요구에 따라 MPEG-D MPEG-Surround (ISO/IEC 23003-1) 표준이 개발되었다[7]. 또한



〈그림 2〉 MPEG 오디오 압축 기술의 역사

다채널뿐 아니라 다각체 오디오 신호에 대한 압축 기술인 MPEG-D SAOC(ISO/IEC 23003-2)가 2002년에 표준화 되었다[8].

인간의 음성 생성 모델을 기반으로 하고 있는 음성 부호화 기술과, 인간의 청각 모델을 기반으로 하고 있는 오디오 부호화 기술은, 음성 통신과 음악 방송 등 각각의 독립적인 서비스 영역에서 독자적으로 기술 발전을 이루어 왔다. 하지만 최근 방송과 통신이 융합하는 방향으로 기술이 발전하면서, 음성 통신과 음악 방송으로 구분되던 서비스 구조가 깨어지고, 더 이상 음성과 오디오 신호를 별도의 콘텐츠로 분리하는 것이 어렵게 되었다. 따라서 하나의 부호화기로 음성과 오디오 신호 모두를 처리할 수 있는 새로운 부호화 기술이 필요하게 되었다[9]. 하지만 기존의 부호화 기술은 음성이나 오디오 신호의 특성을 기반으로 서로 다른 목표에 따라 개발되었기 때문에, 음성과 오디오 신호 모두에 대해 우수한 음질을 제공할 수 없다. 이에 따라 새로운 부호화 기술에 대한 표준이 요구되었고,

MPEG에서는 2007년 CfP를 시작으로 표준화를 진행하여 2012년 MPEG-D USAC(Unified Speech and Audio Coding) 표준을 승인하였다 (ISO/IEC 23003-3)[10].

USAC에 대한 표준화 이후, UHDTV 환경에서 다수의 스피커를 이용하여 공간상의 임의의 위치에 음원을 위치시킬 수 있고, 정확한 음상정위가 가능한 새로운 오디오 부호화 기술에 대한 시장의 요구에 따라 2013년 1월 103차 MPEG 회의에서부터 CfP를 시작으로 MPEG-H 3D Audio에 대한 표준을 진행하였다[11]. Ⅲ장에서는 MPEG-H 3D Audio의 요구사항에서부터, 현재까지의 표준화 현황에 대해서 기술한다.

III. MPEG-H 3D Audio 표준화 현황

MPEG-H 표준은 AV 모두에 대해 몰입감과 현장감이 현저하게 증가한 새로운 표준으로 비디오

<표 1> MPEG-H 3D Audio 서비스 시나리오

서비스 시나리오	특징
홈 씨어터 (Home theater)	<ul style="list-style-type: none"> UHD 환경에서 FOV(Field of View)의 증가에 따른 몰입감과 현장감이 증대되는 상황에서, 오디오는 청취자 주위를 스피커로 둘러싸서, 기존보다 더 많은 채널을 제공하여 비디오와 동일한 수준의 몰입감과 현장감 제공 헤드폰 환경에서도 다채널 스피커와 동일한 수준의 몰입감과 현장감 제공
개인형 TV (Personal TV)	<ul style="list-style-type: none"> 고화질을 제공하는 개인형 3D 디바이스 환경에서 홈 씨어터에 비해 상대적으로 작은 display 주위에(상/하, 좌/우) 스피커를 배치하여 향상된 몰입감과 현장감 제공(헤드폰 환경도 고려)
모바일 TV (스마트폰)	<ul style="list-style-type: none"> 현대인의 삶에서 스마트폰은 매우 중요한 역할을 하며 이러한 스마트폰 환경에서 비디오 청취는 자연스러운 일이기 때문에, MPEG-H 3D Audio도 이를 고려해야 함 MPEG-H 3D Audio는 다채널 환경이 기본이므로, 이를 모바일 환경에서 전송하기 위해 저 비트율, 저 복잡도 오디오 기술이 필요
오디오 전용 서비스	<ul style="list-style-type: none"> 비디오가 없는 오디오 단독 청취환경에 대한 고려도 필요하며, 이 경우 더 높은 수준의 몰입감과 현장감을 제공



<그림 3> MPEG-H 3D Audio 평가를 위한 다채널 오디오 재현 환경

의 경우 HD를 넘어서는 UHD 환경을 고려하고 있고, 오디오의 경우 몰입감이 현저하게 증가한 환경에서 개인화 서비스를 제공한다. 2012년부터 새로운 오디오 표준에 대한 논의를 시작하여, 홈 씨어터, 개인형 TV, 모바일 TV와 오디오 전용 서비스와 같은 4개의 서비스 시나리오를 선정하였다[12]. 4개의 모든 시나리오에 대해 <그림 3>과 같이

22.2채널과 같은 다채널 오디오 신호에 대한 압축과 다양한 재생환경에 적응적으로 재생이 가능한 Flexible rendering/downmix 기능을 고려하고 있다.

4개의 서비스 시나리오를 기반으로 요구사항을 정의하여, 기본적인 요구사항 6개와 부가 요구사항 2개를 선정하였다.

〈표 2〉 MPEG-H 3D Audio 요구사항

기본 요구사항	특징
High quality	• 고음질 application을 위해 비트율을 증가시키면 CD 수준의 음질을 제공해야 함
Localization and Envelopment	• 주어진 청취환경에서 정확한 음상정위, 높은 몰입감과 현장감을 제공해야 함
Rendering on setups with fewer loudspeakers	• 비트스트림은 Reference 렌더링 환경 보다 더 적은 수의 스피커 환경에서 디코딩/렌더링을 제공하고 Reference 환경과 비슷한 음질을 제공해야 함
Flexible Loudspeaker Placement	• Reference 렌더링과 다른 위치에 있는 스피커 배치에서도 높은 몰입감과 현장감을 제공해야 함
Latency	• 생방송을 고려하여 충분히 낮은 latency를 제공해야 함
입력 오디오 신호	• 채널, 객체, HOA 입력 신호를 지원해야 함
부가 요구사항	특징
복잡도	• 디코더는 디코딩/レン더링이 저복잡도로 가능해야 함
Interactivity	• 개인의 interactive application을 위해 객체 오디오 신호를 제어할 수 있어야 함

MPEG-H 3D Audio는 아래와 같이 3개의 입력 오디오 신호 포맷을 지원한다.

- 고차 다채널 오디오(High-order Multi-channel Audio): NHK 22.2채널과 같이 상층 채널(Height Channel)을 포함하여 스피커들이 3차원으로 배치되는 다채널 오디오 재생 시스템을 지원하여야 한다. 고차 다채널 오디오 신호는 전방위 음향장면을 표현할 수 있으며, 적어도 9.1 이상의 채널로 구성되어 있다.
- 객체기반 오디오(Object-based Audio): 다채널 오디오 배경음과 분리되어 있는 오디오 객체 음원들을 전송하고 단말에서 제어할 수 있어야 한다. 오디오 객체는 자유롭게 음향 장면을 구성할 수 있으며, 사용자와 상호작용 서비스를 가능하게 한다.
- 장면기반 오디오 신호 (HOA-based Audio): 음향 장면기반 오디오 신호를 전송할 수 있어야 한다. 음향 장면기반의 오디오 신호는 특정 오디오 획득장비를 통해 획득될 수 있으며, 오디오 채널 신호와 장면을 표현하는 메타데이터

로 구성된다. MPEG에서 장면기반 오디오 신호를 획득하기 위한 수단으로 HOA(High Order Ambisonics) 포맷을 채택하였으며, HOA 포맷은 오디오 채널 신호와 채널 신호를 배치하기 위한 메타데이터 정보가 필요하다.

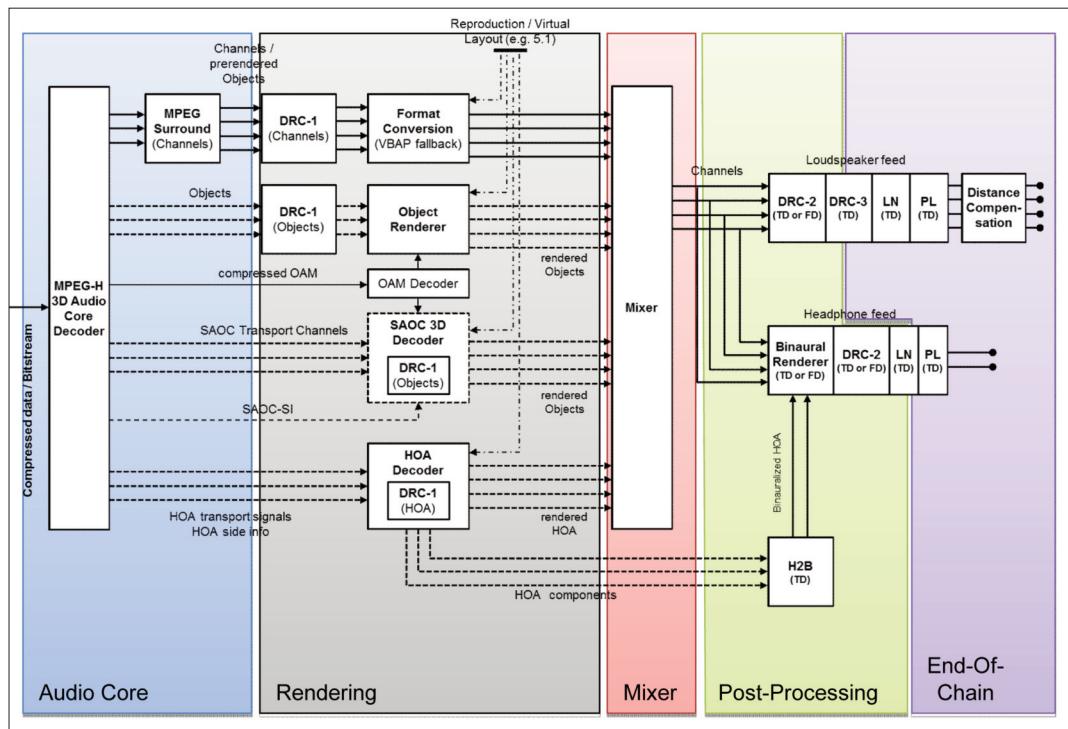
지금까지 설명한 MPEG-H 3D Audio의 요구사항과 서비스 시나리오를 기반으로 먼저 홈 씨어터와 오디오 전용 서비스를 위한 Phase 1에 대한 표준화를 2013년 103차 MPEG 회의에서부터 시작하여 2014년에 완료하였다[13]. MPEG-H 3D Audio Phase 1은 512kbps의 비트율에서 최대 22.2채널의 오디오 신호에 대한 압축 기술과 헤드폰과 다양한 스피커 환경에 대한 렌더링 기술을 포함하고 있다.

MPEG-H 3D Audio Phase 2는 개인형 TV나 스마트폰과 같이 모바일 환경(128kbps 이하)을 고려한 것으로, 2014년 4월 108차 MPEG 회의에서부터 표준화를 진행하였고, Phase 1과 Phase 2를 합한 MPEG-H 3D Audio에 대한 표준을 2016년 하반기에 승인할 예정이다.

IV. MPEG-H 3D Audio 기술

MPEG-H 3D Audio 인코더는 다채널 오디오 신호, 다객체 오디오 신호, HOA 신호 등 기본적으로 고차 오디오 채널 신호를 압축하여 전송하는 기능을 담당한다. 기존의 스테레오 및 5.1 채널보다 다루는 입력 채널수가 고차이므로, MPEG-H 3D Audio 인코더의 압축 성능도 중요하기 때문에, 코어 압축 모듈로 USAC 기술을 채택하였다. USAC은 MPEG-H 3D Audio 이전에 표준화된 최신 오디오 압축 기술로 음성/음악에서 고른 성능을 보이며 압축 효율 측면에서도 매우 뛰어난 성능을 제공한다. MPEG-H 3D Audio는 새로운 압축 모듈을 개발하고 탑재하는 대신 USAC 기술을 활용하되,

고차 채널신호 압축효율을 높이기 위하여 MPEG Surround 기술을 접목하였다. MPEG Surround 기술도 이미 표준화가 완료된 기술로 MPEG-H 3D Audio에 USAC과 더불어 활용된 기술이다. 오디오 객체신호는 MPEG SAOC 기술을 활용할 수 있다. HOA 기반 장면 오디오 정보는 MPEG-H 3D Audio에 새롭게 소개된 포맷으로 HOA를 위한 인코딩/디코딩 모듈이 필요하며, 종전의 MPEG 오디오 표준기술로 HOA기반 오디오 채널 신호를 복원할 수 없었으므로 새로운 HOA 인코딩/디코딩 모듈이 필요하다. HOA는 채널 신호와 메타데이터로 표현할 수 있으므로 채널신호는 USAC으로 인코딩/디코딩하고 추가된 HOA 모듈에서 전송된 메타데이터를 활용하여 장면정보를 복원할 수 있다.



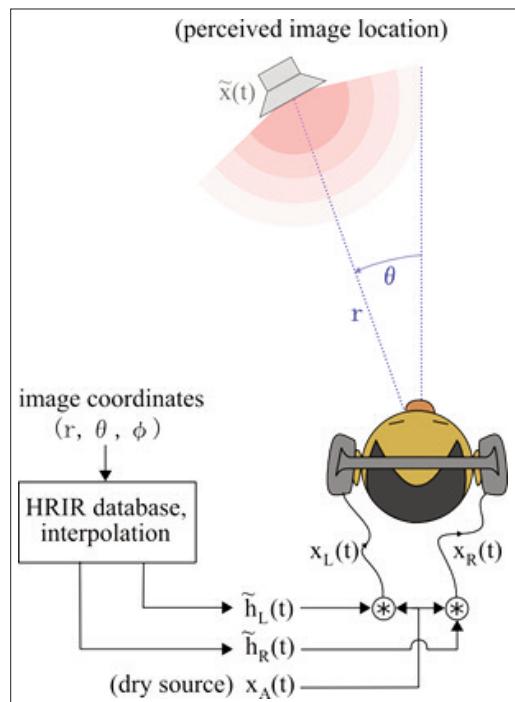
〈그림 4〉 MPEG-H 3D Audio 디코더 구조

MPEG-H 3D Audio의 주요 특징을 정리하면, 다양한 형태의 입력 오디오 포맷을 지원하기 위하여 MPEG-Surround, SAOC, USAC 등 MPEG 오디오 표준기술을 활용하였으며, HOA 메타데이터 처리를 위한 HOA 메타데이터 처리 모듈과, 다양한 형태의 재생환경에 적응하기 위한 렌더링 모듈만이 새롭게 개발되어 포함된 것이다. 즉, 압축하고 전송하는 측면에서의 MPEG-H 3D Audio 기술은 최대한 기존 MPEG 기술을 활용한 것으로 상호간에 유연한 동작을 위하여 하나의 코덱 구조에 적절하게 조합되어 구현되었다. <그림 4>는 MPEG-H 3D Audio 디코더 구조이다.

MPEG-H 3D Audio의 주요 특징은 다양한 형태의 입력 오디오 포맷에 대한 고효율 압축과 다양한 재생환경에 최적의 3D Audio를 재현하기 위한 렌더링 기술이다. 예를 들어 22.2 채널의 콘텐츠가 사용자 단말에 전달되었다 하더라도, 상황에 따라서 사용자는 스테레오 혹은 5.1 채널로 22.2 채널 콘텐츠를 소비하기를 희망한다고 했을 때, MPEG-H 3D Audio는 이를 충족시킬 수 있는 기능을 제공하여야 한다. 이를 위해 필요한 것이 렌더링 기술이다. MPEG-H 3D Audio 렌더링 기술은 콘텐츠 형태와 사용자 단말환경에 무관하게 최상의 3D Audio를 제공하기 위한 기술이다. <그림 4>에서 format converter가 본 기능을 수행한다. format converter는 입력채널 레이아웃과 재생 출력 레이아웃간의 변환장치로 수동적(passive) 믹싱 과정을 통해 출력 채널을 생성하는 대신에, 원 콘텐츠의 효과를 최대한 반영하기 위한 능동적(active) 믹싱과정을 지원한다. 예를 들어, 22.2 채널을 단순하게 다운믹스하여 스테레오 채널 신호를 생성할 수 있겠으나, 채널간의 상관관계, 위치 정보들을 고려하여 다운믹싱을 수행하면 보다 원

콘텐츠 음향장면에 가까운 스테레오 신호를 얻을 수 있다.

사용자가 헤드폰 환경을 선호하게 된다면 MPEG-H 3D Audio는 원 콘텐츠의 효과를 최대한 반영할 수 있도록 헤드폰을 위한 스테레오 신호를 생성할 수 있다. 이를 위해 MPEG-H 3D Audio는 바이노럴 렌더링 모듈을 탑재하고 있다. 바이노럴 렌더링 모듈은 디코딩된 고차 다채널 신호를 스테레오 신호로 변환하는 과정에서 <그림 5>와 같이 공간상의 스피커 위치에서 발생하는 전달함수를 적용하여 스테레오 신호를 생성하는 모듈이다. 스피커 위치에서 발생하는 전달함수를 BRIR(Binaural Room Impulse Response)이라고 하며, 채널 수 만큼의 BRIR이 존재하며 이를 활용하여 바이노럴 스테레오 신호를 생성한다. 원리는 단순하게



디코딩된 채널 신호와 해당하는 BRIR을 콘볼루션(convolution)하는 것이다. 그러나 채널수가 증가함에 따라 콘볼루션 연산량이 급격하게 증가할 수밖에 없으며, MPEG-H 3D Audio에서는 이를 효율적으로 최적화된 연산량으로 처리할 수 있는 툴을 추가하였다.

마지막으로, MPEG-H 3D Audio에서 지원하는 기능으로 DRC(Dynamic Range Control) 모듈이 있다. MPEG-H 3D Audio는 기본적으로 다양한 포맷의 오디오 신호를 처리한다. 따라서 상호간에 발생하는 오디오 신호의 레벨차가 존재할 수 있다. DRC의 첫 번째 요구되는 기능은 상호 포맷간의 레벨차를 맞추어 주어야 한다. 따라서 프로그램별로 다양한 형태의 오디오 포맷이 존재하더라도 최종 출력레벨의 차이를 느낄 수 없도록 DRC가 오디오 포맷별로 출력 레벨을 조절할 수 있어야 한다. 기본적으로 DRC기능은 출력 신호의 레벨을 자동적으로 일정레벨로 맞추어 줄 수 있어야 한다. 예를 들어, 방송콘텐츠 별로 상이한 오디오 레벨로 제작되었다 하더라도, 사용자 단말에서는 사용자가 볼륨조절을 개별적으로 조절하지 않아도 동일한 라우드니스(loudness)를 가질 수 있도록 DRC가 처리할 수 있어야 한다. 이는 최근 들어 많은 이슈가 되어왔던 사항으로, 사용자가 채널별로 혹은 콘텐츠 별로 인지하는 라우드니스 레벨 차이가 상이하여 시청 시, 오디오 볼륨을 상시 조절해야 하는 번거로움을 배제하기 위한 기술이다. MPEG-H 3D Audio는 코덱 단에서 원천적으로 DRC를 통해 라우드니스를 제어하여 출력신호를 제공함으로써, 사용자의 볼륨조절의 번거로움을 최소화 시킬 수 있다[14].

V. MPEG-H 3D Audio 프로파일

MPEG 표준은 RM(Reference Model) 설정 이후에도 성능향상을 위해 다양한 툴들을 추가하기 때문에, 활용분야에 따라 채널수, 비트율, 사용 툴 등에 대해 프로파일로 정의한다. 각 프로파일은 서비스 목적에 따라 선택적으로 복잡도를 제공할 수 있고 기능 및 지원 가능한 채널수를 명시하고 있다. MPEG-H 3D Audio는 3개의 프로파일을 지원하며 각각의 특징은 아래와 같다.

- **Main 프로파일:** Main 프로파일은 Phase1에서 개발된 툴들에 대한 프로파일로 복잡도의 제약 사항을 두지 않고, 고품질 및 다기능을 지원하기 위하여 최대한 필요한 툴들을 모두 포함시키고 있으며, 채널수 및 객체수도 가능한 제약을 두지 않고 있다. Main 프로파일에 포함된 툴들은 앞서 열거한 USAC, SAOC, MPEG Surround, HOA 등으로 기존 표준기술들을 그대로 포함시킨 superset 프로파일이다.
- **Low-Complexity(LC) 프로파일:** MPEG-H 3DA가 고음장감을 제공하고 다기능을 지원함으로써 새로운 오디오 코덱 형상을 제시한 것은 사실이나, Main 프로파일은 실제 서비스 단말에 지원하기 위해서는 높은 복잡도로 인하여 다소 부담이 되는 것은 사실이다. 이를 해결하기 위하여 Phase 2 기술 개발 기간 동안에 저 복잡도를 갖는 툴, 혹은 기존 툴을 저 복잡도로 단순화한 기술들을 표준에 반영하였다. LC 프로파일은 이러한 툴들을 조합하여 저 복잡도를 갖는 프로파일을 제공하는 것으로 저 복잡도를 갖는다 하더라도 음질 성능 및 기본적인 기능은 Main 프로파일과 크게 다르지 않다.

〈표 3〉 MPEG-H 3D Audio LC 프로파일의 레벨

레벨	최대샘플링율	최대 채널수 (비트스트림/디코더)	최대 스피커수 /구성(예)	최대 디코더 객체 수	채널+객체 구성(예)	최대 HOA 차수	최대 HOA 차수 + 객체(예)
1	48000	10/5	2/2.0	5	2채널+3객체	2	2차+3객체
2	48000	18/9	8/7.1	9	6채널+3객체	4	4차+3객체
3	48000	32/16	12/11.1	16	12채널+4객체	6	6차+4객체
4	48000	56/28	24/22.2	28	24채널+4객체	6	6차+4객체
5	96000	56/28	24/22.2	28	24채널+4객체	6	6차+4객체

- **High 프로파일:** High 프로파일은 Main 프로파일과 LC 프로파일의 superset 프로파일이다. LC 프로파일이 저 복잡도를 위한 프로파일이기는 하나, 새롭게 추가된 툴들로 인하여 Main 프로파일에 기술되지 않은 툴들이 발생하였다. 이에 따라 MPEG-H 3D Audio의 모든 툴들이 포함된 High 프로파일을 새롭게 정의하여 MPEG-H 3D Audio Phase 1과 Phase 2의 모든 툴들을 반영하였다.

LC 프로파일에 대해서는 〈표 3〉과 같이 스테레오에서 22.2채널에 대해 레벨을 정의하고 있다.

VI. 국내 지상파 UHDTV 표준화 현황

국내를 포함하여 일본, 미국, 유럽 등에서 초고선명 TV에 대한 기술개발과 표준화가 활발하게 진행 중에 있다. ATSC(Advanced Television Systems Committee)에서는 ATSC 3.0이라는 프로젝트를 2012년도부터 진행해왔으며, 차세대 UHD 방송표준을 제정하는 것을 목적으로 하고 있다. 비디오의 경우, UHD 영상을 위한 압축 방식으로 MPEG의

HEVC(High Efficiency Video Coding) 기술이 채택되는 것에 대하여 이견이 없었으나, 오디오의 경우 이해관계자들의 첨예한 대립으로 인하여 오디오 코덱 표준 선정의 어려움이 있었지만, ATSC 3.0에서는 MPEG-H 3D Audio와 AC-4 두 개의 오디오 표준 코덱을 포함하고 나라별로 오디오 표준을 선정하는 방향으로 논의가 되고 있다.

국내에서도 2015년 지상파 UHDTV 방송을 위한 700MHz 주파수 대역을 각 방송사가 할당 받음으로써, 본격적인 UHDTV 방송을 위한 작업에 착수하였다. ATSC와 마찬가지로 비디오 코덱 선정은 HEVC로 무난하게 결정되었으나, 오디오의 경우 Dolby의 AC-4, MPEG-H 3D Audio, DTS 기술이 경합하여, 차방포럼에서는 UHDTV 오디오 서비스를 실현하기 위한 렌더링 기능을 포함한 MPEG-H 3D Audio 기술을 잠정 표준 RM(Reference Model)으로 선정하고 표준 초안을 작성하였다. 현재 국내 지상파 UHDTV 서비스를 위한 정합테스트를 준비 중에 있으며, 올 9월까지 MPEG-H 3D Audio 기반의 국내 지상파 UHDTV에 대한 정합 테스트를 마무리한 후 2017년 2월부터 본격적인 MPEG-H 3D Audio 기반의 UHDTV 오디오 서비스를 제공할 예정이다.

VII. 맷음말

본 논문에서는 차세대 방송환경에서 5.1채널을 넘어서는 몰입감과 개인형 방송 서비스를 제공하기 위한 UHDTV 오디오 기술인 MPEG-H 3D Audio의 기술의 표준화 현황 및 기술의 개요에 대해 살펴보았다. 국내에서는 최대 16트랙을 지원하는 UHDTV 오디오 서비스를 MPEG-H 3D Audio를 이용하여 서비스할 예정이므로, 청취자는 오디오 채널에 대한 제어 및 선택을 통한 개인화된 방송과 기존 5.1채널을 넘어서는 몰입감과 현장감을 제공하는 새로운 오디오 서비스를 받을 수 있다.

물론 대부분의 방송 프로그램이 스테레오 형태로 제작되고 있기 때문에, 5.1채널 이상의 서비스와 객

체기반의 개인화 서비스가 보급될 때까지는 많은 시간이 필요할 것으로 예상된다. 하지만 먼저 객체기반의 서비스를 통해 청취자가 특정 오디오 신호의 볼륨을 제어한다든지 특정 오디오 채널을 선택하는 것과 같은 개인화된 방송을 먼저 시작하고, 가정 내에서 사운드바와 같이 다채널 오디오를 쉽게 서비스 받을 수 있는 환경이 구비되면 5.1채널을 넘어서는 몰입감/현장감을 제공하는 오디오 서비스로 단계적으로 접근한다면, MPEG-H 3D Audio 표준의 모든 기능을 활용하여 가정에서 10.2채널의 다채널 오디오 서비스와 객체기반 서비스를 제공받는 시점이 우리가 생각하는 것보다 훨씬 빨라질 수도 있다.

참고 문헌

- [1] 이태진 외, "UHDTV를 위한 10.2채널 기반 다채널 오디오 재현 기술", 방송공학회 논문지 제 17권 제 5호, 2012년 9월
- [2] 오현오 외, "ATSC 3.0 오디오 코덱 표준화 동향", 방송과미디어 제 20권 4호, 2015년 10월
- [3] ISO/IEC 11172-3:1993, "Information Technology – Coding of moving pictures and associated audio for digital storage media at up to about 1.5 Mbit/s – Part 3: Audio"
- [4] ISO/IEC 23008-3:2015, "Information Technology – High efficiency coding and media delivery in heterogeneous environments – Part 3: 3D Audio"
- [5] ISO/IEC 13818-7:2006, "Information Technology – Generic coding of moving pictures and associated audio information – Part 7: Advanced Audio Coding(AAC)"
- [6] ISO/IEC 14496-3:2009, "Information Technology – Coding of audio-visual objects – Part 3: Audio"
- [7] ISO/IEC 23003-1:2007, "Information Technology – MPEG Audio Technologies – Part 1: MPEG Surround"
- [8] ISO/IEC 23003-3:2010, "Information Technology – MPEG Audio Technologies – Part 2: Spatial Audio Object Coding(SAOC)"
- [9] 이태진 외, "MPEG 오디오 최신 표준: USAC 기술", 방송공학회 논문지 제 16권 제 5호, 2011년
- [10] ISO/IEC 23003-3:2012, "Information Technology – MPEG Audio Technologies – Part 3: Unified Speech and Audio Coding"
- [11] 서정일 외, "차세대 실감방송 서비스를 위한 MPEG-H 3D AUdio 표준화 동향" 전기의 세계, 제 64권 제 2호, 2015년
- [12] ISO/IEC JTC1/SC29/WG11/N13411, "Call for Proposals for 3D Audio", January 2013, Geneva, CH
- [13] ISO/IEC 23008-3:2015, MPEG-H (High efficiency coding and media deliver in heterogeneous environments), Part 3: 3D Audio, 2015.
- [14] 백승권, "MPEG-H 3D Audio 배경 및 진행과정", 방송과 기술, 5월호, 2016년.

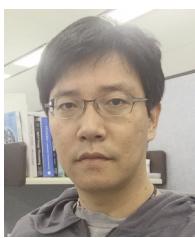
필자 소개

이태진



- 2014년 2월 : 충남대학교 전자전파정보통신공학과 (공학박사)
- 2002년 10월 ~ 2003년 9월 : 일본 Tokyo Denki University, 방문연구원
- 2000년 5월 ~ 현재 : ETRI 책임연구원, 오디오연구실 실장
- 2015년 3월 ~ 현재 : 과학기술연합대학원대학교 이동통신 및 디지털방송공학전공 부교수
- 주관심분야 : 오디오 부호화, 실감음향, 오디오 신호처리

백승권



- 1999년 2월 : 한국항공대학교 전자공학과 (공학사)
- 2001년 2월 : 한국과학기술원 정보통신공학부 (공학석사)
- 2005년 8월 : 한국과학기술원 정보통신공학부 (공학박사)
- 2005년 8월 ~ 현재 : ETRI 선임연구원
- 주관심분야 : 오디오 부호화, 오디오 신호처리