

고현장감 오디오 서비스를 위한 하이브리드 오디오 기술동향

Hybrid Audio Technology Trends for Immersive Sound Service

장대영 (D. Jang) 오디오연구실 책임연구원
이용주 (Y.J. Lee) 오디오연구실 책임연구원
유재현 (J. Yoo) 오디오연구실 선임연구원
이태진 (T. Lee) 오디오연구실 책임연구원

* 본 연구는 미래창조과학부 및 정보통신기술진흥센터의 정보통신·방송 연구개발 사업의 일환으로 수행하였음[R0126-16-1034, 채널/객체 융합형 하이브리드 오디오 콘텐츠 제작 및 재생기술 개발].

현재, 극장 사운드를 중심으로 객체기반 오디오 기술을 포함하는 하이브리드 포맷의 사운드 서비스가 점차 확산되고 있으며, 미국, 유럽, 한국 등에서는 차세대 방송용 오디오에 객체기반 오디오 기술의 도입을 적극적으로 고려하고 있다. 객체기반 오디오 기술은 콘텐츠의 제작단계에서 재생환경을 고려할 필요가 없고, 현장의 음향을 신호와 3차원 공간정보로 구분하여 음향공간의 정보를 그대로 표현함으로써, 재생단말에서 3차원 공간정보를 활용하여 재생할 수 있게 한다. 이러한 객체기반 실감음향 기술개발을 위해서는 편리한 제작 및 3차원 공간정보 표현 기술이 필요하며, 청취환경에서는 객체기반 실감음향 콘텐츠를 제작자의 의도대로 렌더링할 수 있는 재생 및 제어 기술이 필요하다. 이에 객체기반 실감음향 기술이 포함되는 하이브리드 오디오 기술의 현황에 대하여 살펴보고자 한다.

방송·전파·위성 & 스마트 미디어
기술 특집

- I. 머리말
- II. 하이브리드 오디오란?
- III. 하이브리드 오디오 편집
기술동향
- IV. 하이브리드 오디오 재현
기술동향
- V. 하이브리드 오디오 표준화
현황
- VI. 맺음말

I. 머리말

실감음향 기술은 오랜 역사를 통하여 조금씩 발전했으나, 최근까지는 채널을 늘리는 방법 외에는 실감음향 성능을 획기적으로 개선할 방법이 마땅하지 않았다. 5.1채널, 7.1채널 오디오 이후에 9.1, 10.2, 15.1, 22.2, 31.1채널 등 다양한 채널방식들이 제안되어 보다 몰입감 있는 실감음향을 재생하기 위해 노력하였지만, 문제는 콘텐츠를 제작하는 일이 복잡해짐에 따라 콘텐츠 확보 및 서비스에 한계가 드러나면서 어느 방식 하나 주도권을 잡지 못하는 실감음향의 춘추전국 시대를 맞이하게 되었다[1].

이러한 국면에 하나의 커다란 파문을 일으키는 실감음향 방식이 객체기반 오디오를 포함하는 하이브리드 포맷의 오디오(이하 하이브리드 오디오) 방식이다. 헐리웃을 중심으로 미국 기업인 Dolby의 Atmos, DTS의 DTS:X에 이어 IOSONO를 인수한 유럽계 Barco의 AuroMax까지 가세하면서 또 다른 오디오 방식 전쟁을 예고하고 있다.

이에 본고에서는 하이브리드 오디오 서비스의 가장 중요한 축인 콘텐츠 제작을 위한 편집도구와 서비스를 위한 재현기술을 중심으로 현재까지의 동향을 살펴보고자 한다.

II. 하이브리드 오디오란?

하이브리드 오디오는 채널기반 오디오와 객체기반 오디오를 동시에 포함하는 오디오로서, 국제적으로 통용되는 표현은 하이브리드 포맷 오디오라는 표현으로 많이 사용되며, 의미가 좀 다르지만 몰입형 사운드(Immersive Sound)로 불리는 경우도 있다.

오디오 콘텐츠의 교환을 위한 포맷형식은 전통적으로 채널기반의 오디오 포맷이 주로 사용되고 있으며, 아직까지도 대세를 이루고 있다. 채널기반의 오디오 포맷은

스피커의 개수와 스피커의 배치방법이 가장 중요하며, 스피커의 개수 및 배치방법에 따른 콘텐츠 제작이 이루어지고, 동일한 스피커의 개수 및 배치방법에 따른 청취 환경에서 청취되기를 기대하게 된다.

하지만, 일반적인 가정의 청취환경에서는 스피커의 개수나 배치를 정해진 규정에 맞추어 설치하는 것이 어렵다. 그러다 보면 제작자의 의도대로 오디오 콘텐츠의 3차원 이미지를 재현하는 것이 어려워지며, 결국 3차원 오디오의 청취를 포기하게 되는 사례가 많다. 이러한 문제를 해결하기 위해 등장한 것이 객체기반의 오디오 포맷이라 할 수 있으며, 현장의 음원신호와 음원의 위치정보를 함께 패키징함으로써, 청취환경의 스피커 배치에 따라 객체음원들을 제작자의 의도대로 재현할 수 있게 된다.

하이브리드 포맷의 오디오란 채널기반의 오디오 포맷과 객체기반의 오디오 포맷이 혼합되어 있다는 의미이며, 국내에서는 이를 하이브리드 오디오라고 부르고 있다. 객체기반의 음원은 명확한 위치가 감지되어야 하는 중요한 객체음원을 선별하여 객체기반 오디오 포맷으로 제작하며, 그 밖의 위치정보가 중요하지 않은 사운드들을 채널기반의 오디오 포맷으로 제작하여 함께 패키징하는 것이다. 단말에서는 하이브리드 오디오 콘텐츠를 해석하여 객체기반 음원신호와 공간정보를 포함하는 메타데이터를 추출하여 렌더링하게 되고, 이를 채널기반 오디오 신호와 믹싱하여 재생한다.

객체기반 오디오 기술은 20세기 말 MPEG-4로부터 본격적으로 출발되었다. Moving Picture Experts Group(MPEG) 오디오 기술을 주도하고 있는 독일 FhG에서 Euro Project Creating, Assessing and Rendering in Real-time Of high-quality aUdio-viSual environments(CAROUSO)[2]의 객체기반 3차원 음향 결과물을 활용하여 창업한 IOSONO는 객체기반 오디오 방식과 Wave Field Synthesis(WFS) 기술을 적용하여 수

평면을 커버하는 스피커어레이를 통하여 실감음향을 재생하는 기술을 상용화하였다. IOSONO 시스템은 초기에 극장 및 테마파크에 적용되어 호평을 받았지만, 상대적으로 막대한 설치비용을 부담스럽게 생각하는 극장주들에게 효과적으로 어필하지 못하였고, 더는 확산시키지 못하였다.

한국전자통신연구원에서는 2002년부터 객체기반 오디오 기술을 연구하고 있으며[3], 2007년 (주)오디즌을 통하여 객체기반 오디오 기술을 음악에 적용한 대화형 음악 서비스, Music2.0 기술을 상용화하였고, 이를 기반으로 MPEG-A Interactive Music Application Format(IMAF) 표준화를 주도하는 한편, MPEG-H 3D Audio 표준화에도 객체+채널 오디오 기술을 제안하며 국제표준화에 적극적으로 참여하고 있다. Music2.0 서비스는 일부 매니아층을 확보하는 성과를 거두기도 했지만, 저작권 협상이 쉽지 않아 콘텐츠를 확보하는데 어려움을 겪으며, 서비스가 활성화되지는 못하였다.

영국 BBC에서는 워블던 테니스 실황을 해설 음성과 배경음이 구분된 객체기반 오디오 방식의 방송을 인터넷을 통해 중계하면서, NetMix라는 툴을 이용하여 사용자가 음성과 배경음을 적절히 조정할 수 있도록 하는 객체기반 방송시험을 수행하였다. 본 시험을 통하여 해설음성과 배경음의 음량조정에 대한 개인의 선호도를 조사한 결과, 각자의 임의대로 조정한 소리를 더 선호하는 경향을 발표한 바 있다.

한편, SRS를 인수한 DTS는 객체기반 오디오 포맷인 Multi-Dimensional Audio(MDA)를 확산시키기 위한 연합체를 구성하여 2012년 1월 Consumer Electronics Show(CES)에서 공개하였으며, 이후 곧바로 Dolby는 2012년 4월 CinemaCon에서 유사한 포맷인 Atmos를 전용 콘텐츠 제작 도구, 극장용 오디오 프로세서와 함께 발표하였고, Atmos 콘텐츠를 제작하여 공급하면서 한 발 앞선 상용화에 성공하였다.

Atmos는 특히 Objects로 표현된 객체기반 오디오 방

식만을 사용하지 않고, 배경음에 대하여 Beds로 표현된 채널기반 오디오 방식을 함께 사용함으로써, 객체기반 오디오 방식의 콘텐츠 제작 및 렌더링의 어려움을 다소 해결할 수 있는 절충안을 명시적으로 제시하였는데[4], 이전의 기술들도 완전한 객체기반 오디오 기술의 구현 상 어려움을 채널과 객체를 함께 사용하는 방식을 통해 해결하고 있는 것으로 밝혀져 있다.

Dolby Atmos는 (그림 1)과 같이 2015년 말까지 1,200개의 극장, 300개의 영화 타이틀, 100개의 후반제작 스튜디오를 확보하였으며, 계속 확대하며 주도권을 잡아가고 있다.

2015년 4월 Convention of Cinema Industry(CinemaCon) 전시회에서는 채널기반 오디오 포맷인 Barco의 Auro-3D 사운드에 IOSONO의 3차원 음향 렌더링 기술을 적용한 Barco의 객체기반 오디오 방식인 AuroMax가 공개되어 (그림 2)와 같이 인프라를 확보하고 있으며, DTS도 2015년도에 MDA 기술을 보완하여 DTS:X로 상



(그림 1) Dolby Atmos의 전 세계 인프라확보 규모[5]



(그림 2) Auro-3D의 전 세계 인프라확보 규모

〈출처〉: <http://www.barco.com>, 2015, 9.

용화하여 본격적으로 서비스를 시작함으로써, 객체기반 오디오 방식의 주도권을 차지하기 위한 일대 격전이 예고되고 있다[5].

이러한 가운데, 오디오 콘텐츠 포맷의 주도권을 확보하기 위해서는 콘텐츠 및 상영관의 확보가 무엇보다도 중요하다는 것을 알 수 있으며, 보다 편리한 콘텐츠 제작기술 및 우수한 렌더링 기술을 통하여 새로운 오디오 방식의 시대에 새로운 기술 경쟁력을 확보할 수 있을 것으로 전망된다. 이에 Ⅲ장에서 V장을 통하여 하이브리드 오디오 콘텐츠 편집기술 및 콘텐츠 재생을 위한 재현 기술과 하이브리드 오디오 표준화의 최근 동향을 살펴 보고자 한다.

Ⅲ. 하이브리드 오디오 편집 기술동향

기존의 대표적인 오디오 편집툴로는 Pro tools, Nuendo, Pyramix 등이 있는데, 이들을 통해 편집한 오디오의 출력은 채널 오디오 신호만을 포함하고 있으므로, 채널기반의 편집기능을 수행한다고 볼 수 있다.

앞서 기술한 것과 같이 객체 오디오는 객체 오디오 신호와 객체 오디오의 렌더링 정보를 포함하는 메타데이

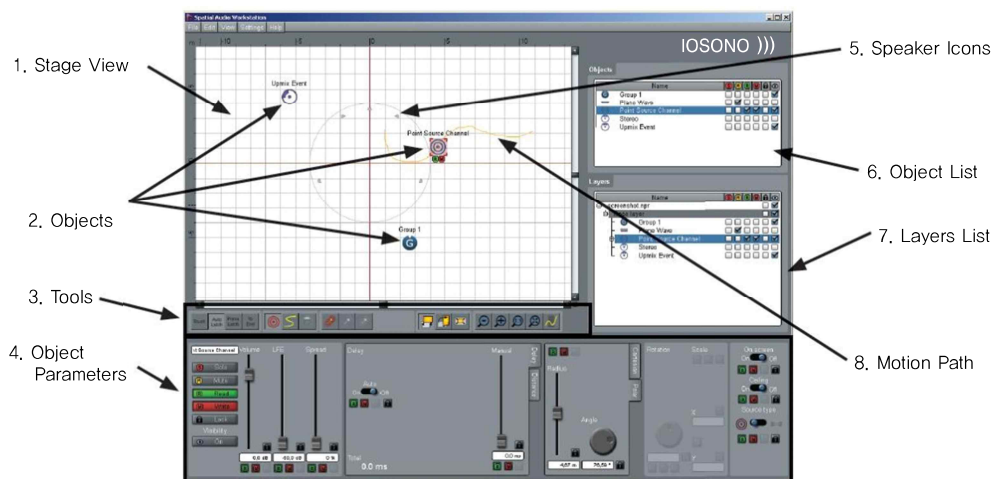
터로 구성되는데, 기존의 저작도구에서는 이러한 객체 오디오 신호와 메타 데이터를 함께 출력하는 기능을 지원하지 않는다.

하지만, 기존의 오디오 편집도구의 저작방식을 보면 객체기반 오디오 처리방식이 이전부터 활용이 되고 있음을 알 수 있다. 채널기반의 편집툴에서도 객체 오디오 신호의 렌더링 정보와 유사한 오토메이션(automation) 정보를 편집툴 내에서 저장하고 편집할 수 있으며, 이에 따라 렌더링된 오디오 신호를 재생할 수 있다. 그러나, 이러한 정보는 편집툴 내부에만 저장되며, 이를 파일형태로 저장하는 방법은 제공하고 있지 않으므로, 객체기반 오디오 저작에 활용되지는 못하고 있다.

객체 오디오를 지원하는 하이브리드 오디오 서비스를 제공하기 위하여 위와 같은 편집툴에 플러그인(Plug-in) 형태로 메타데이터를 생성하여 활용할 수 있도록 하는 툴들이 개발되고 있다.

객체 오디오의 편집에 대한 개념을 포함하는 편집툴로는 IOSONO 콘텐츠의 제작을 위하여 개발된 Spatial AudioWorkstation을 들 수 있다.

(그림 3)은 Spatial Audio Workstation의 Graphic



(그림 3) IOSONO의 Spatial Audio Workstation

<출처>: <http://www.iosono-sound.com>

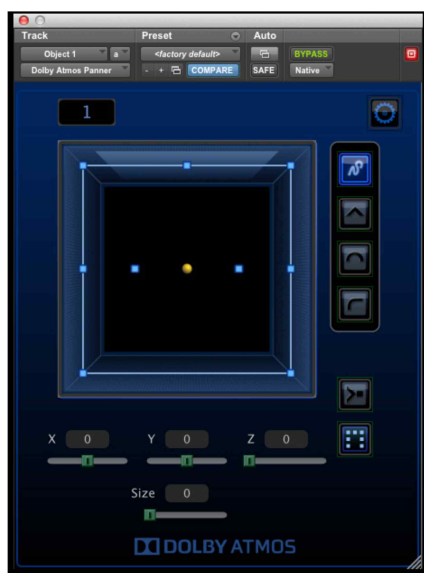
User Interface(GUI)를 나타낸 것이다. 이 편집툴은 하나의 화면에서 여러 개의 객체를 제어할 수 있고, 이러한 제어정보를 저장할 수 있으며, 생성된 제어정보는 다양한 출력 채널환경에서 활용될 수 있도록 개발되었다[6].

하지만, 제어정보가 별도의 파일로 생성하는 것이 아니라 프로젝트 내에 저장되며, 출력될 오디오 신호를 렌더링 할 때에 사용된다는 측면에서는 기존의 편집도구와 유사하다고 할 수 있다. 이러한 이유로 독립적인 객체 오디오 편집툴이라고 설명하는 것은 어렵다고 할 수 있다.

Dolby에서는 Atmos 콘텐츠의 편집을 위한 편집 시스템을 개발하였다. Atmos 편집 시스템에는 2가지 주요한 툴이 있는데, ‘Dolby Atmos Monitor application’과 ‘Dolby Atmos Panner Plug-in’이다.

이와 함께 ‘Dolby Rendering and Mastering Unit(RMU)’이 있는데, RMU에서 ‘Monitor application’이 실행되며, 이 외의 다양한 편집 관련 설정 및 기능을 수행한다.

‘Dolby Atmos Panner Plug-in’은 프로툴즈에서 동



(그림 4) Dolby Atmos Panner Plug-in

〈출처〉: www.dolby.com



(그림 5) DTS MDA Creator

〈출처〉: www.listen.dts.com

작하는 플러그인으로서, RMU와 객체 오디오의 렌더링 정보를 포함하는 메타데이터를 주고받는 기능을 수행한다. (그림 4)는 Dolby의 ‘Atmos Panner Plug-in’의 GUI를 나타낸 것이다.

Dolby RMU, Dolby Panner plug-in과 상용 장치인 mixing console을 활용하여 Atmos 콘텐츠의 제작이 가능하다.

DTS에서도 (그림 5)와 같이 객체 오디오를 지원하는 편집도구인 ‘MDA Creator’를 개발하여 콘텐츠 편집에 활용하고 있다. MDA Creator는 Pro tools의 플러그인 형태로 제공되는데, 다양한 재생 채널환경을 지원하며, 객체 오디오를 편집하고 저장할 수 있다. 객체 오디오를 위한 메타데이터는 DTS 자체 포맷인 MDA 포맷에 따라 저장된다.

Barco에서도 채널 오디오와 객체 오디오가 포함된 Auro-3D 콘텐츠의 제작을 위한 저작도구를 개발하였다. ‘Auro-3D® Authoring Tools(AAT)’과 ‘AuroMatic® Pro 2D/3D’가 이에 해당하는데, Pro tools에서 플러그인 형태로 동작한다[그림 6] 참조].



(그림 6) Auro3D Panner GUI

<출처>: www.auro-3d.com

국내에서도 객체 오디오를 편집할 수 있는 편집도구의 개발이 진행되고 있는데, Dolby와 DTS사가 개발한 방식과 같은 플러그인 형태로 개발되고 있다.

국내에서 개발 중인 객체 오디오 편집도구의 경우 다양한 재생 채널환경을 지원하며, 객체 오디오 신호를 위한 메타 데이터는 자체 포맷에 따라 오디오 신호와는 독립된 파일로 저장된다. (그림 7)은 국내에서 개발 중인 객체 오디오 편집도구를 나타낸 것이다[7].

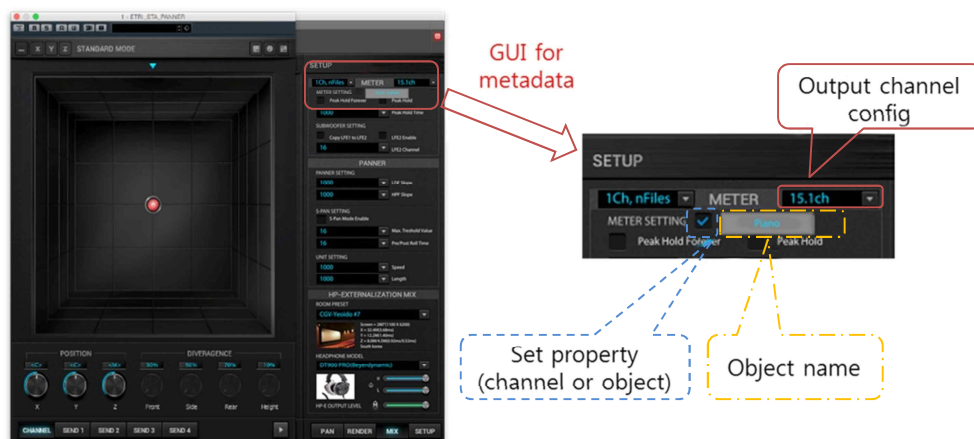
지금까지는 극장 오디오 관련 대규모 업체에서 하이브리드 오디오 관련 편집도구가 개발되고 있으나, 최근에는 극장뿐만 아니라 방송에서도 객체 오디오가 포함

되는 하이브리드 오디오가 표준으로 채택되고 있어, 객체 오디오를 지원하는 편집도구는 지속해서 개발될 것으로 예상된다.

IV. 하이브리드 오디오 재현 기술동향

하이브리드 오디오는 여러 기관들에서 관련 연구가 이루어지고 있는데, 재현기술에 관한 동향을 살펴보면 다음과 같다.

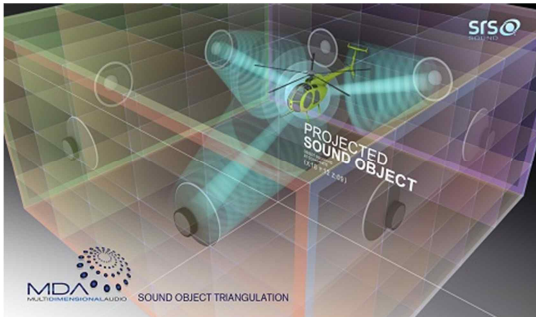
Dolby는 CinemaCon 2012 행사에서 기존 5.1/7.1채널 시네마 사운드의 시장 지배력을 D-Cinema 시장에서 유지하기 위해 ‘객체+채널’의 새로운 극장용 오디오 시스템으로서 객체 사운드에 대해서는 최대 64채널을 지원하는 Atmos[8]를 소개하였다. (그림 8)과 같이 채널기반 신호인 Beds를 기반으로 객체신호를 더해 콘텐츠가 제작되는데, 다양한 극장 재생환경에 맞추어 최적의 사운드를 재생할 수 있고 특별한 음향효과를 줄 수 있다는 점이 큰 특징으로 헐리웃 영화를 중심으로 많은 영화를 Atmos 사운드 포맷으로 제작하였다. 국내에서도 영화나 방송 등의 환경에서 채널신호와 객체신호를 함께 사용하여, 실감을 제공하기 위한 하이브리드 오디오 기술에 대한 관심을 증폭시켰고 해당 시스템이 설치된 상영관이 점차 늘어나고 있다.



(그림 7) 국내 개발 중인 Panner Plug-in[7]



(그림 8) Dolby Atmos 사운드의 개념[8]



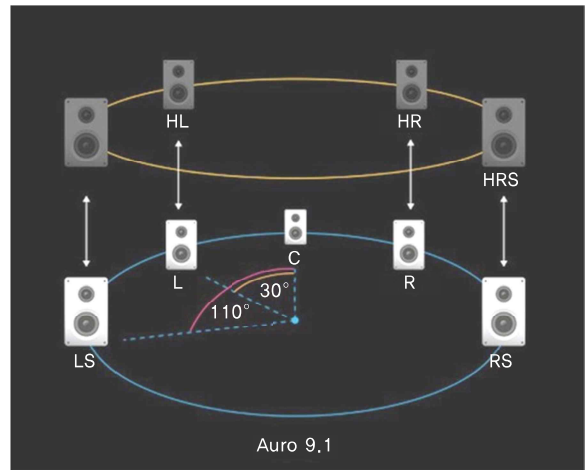
(그림 9) MDA 사운드 합성 개념도

〈출처〉: technewspedia.com/3d-audio-alliance-3daa-and-multi-dimensional-audio-md-audio-accelerated-gpu/

DTS 또한 Atmos와 유사하지만 오픈 포맷으로 MDA (Multi Dimensional Audio)[9]를 개발하여, 하이브리드 오디오의 극장, 블루레이 등에서의 재현 기술에 대해 Dolby와 대응하고 있다. 시네마 서버 제작사, 콘텐츠 제공자 등을 중심으로 한 얼라이언스를 구성한 MDA는 채널기반의 Beds를 재생하기 위한 채널 포맷을 특정하지 않는 기술 개념(그림 9) 참조를 제공함으로써, 기존의 극장에서도 활용할 수 있다고 홍보하고 있다.

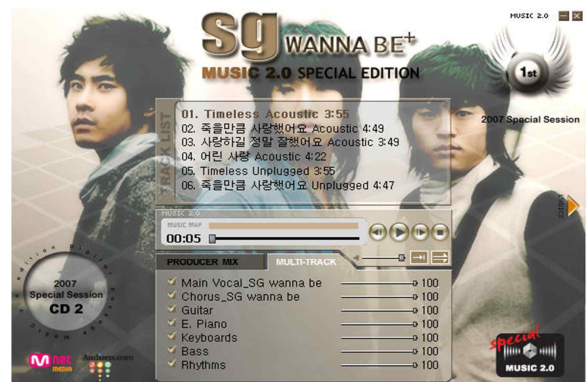
Barco는 유럽에서 International Telecommunication Union(ITU) 5.1채널에 4개의 높이(height) 채널을 추가한 포맷인 Auro 9.1(그림 10) 참조를 기본으로 Auro 3D 사운드 포맷을 통해 홈씨어터(9.1, 10.1채널) 및 시네마(11.1, 13.1채널) 시장 진출을 도모하고 있는 상황이다. Auro3D는 MDA와 그 행보를 같이 하고 있다가 2015년에는 독립을 선언하면서 AuroMax라는 하이브리드 오디오 방식을 공개하였다. IOSONO를 인수하면서 음장합성 기술의 응용을 통한 3차원 오디오 렌더링 기술을 접목하여 객체기반 음원을 충실히 재현하고 있다.

한편 ETRI에서는 음원 객체 오디오 신호를 활용하여



(그림 10) Auro9.1 포맷의 채널 배치

〈출처〉: www.barco.com



(그림 11) Music2.0 디지털 음반[7]

이용자가 음원 객체의 위치와 방향, 볼륨 등을 자유로이 배치할 수 있는 실감 오디오 기술을 개발한 경험을 가지고 있으며, 이를 기반으로 하이브리드 오디오 재현 기술의 시초라 할 수 있는 MUSIC2.0 음반을 출시하여 국내 음반시장에 큰 반향을 일으켰다. 이 음반은 (그림 11)과 같은 형태로 사용자에게 제공되는 서비스로 PC상에서 전용 플레이어를 통해 음악을 들으면서 해당 곡 내에 포함되어 있는 악기나 가수의 목소리들을 조정하면서 청취가 가능한 음반이다.

독일 프라운호퍼 연구소는 오래 전부터 WFS 기술을 연구하여 객체기반 오디오 신호를 사용하는 개념을 일찍이 소개한 바 있으며 이를 IOSONO로 상품화하여 독

일과 미국을 비롯한 여러 국가의 영화 상영관이나 콘서트홀, 클럽 등에 해당 시스템을 설치하여 서비스한 바 있다.

V. 하이브리드 오디오 표준화 현황

하이브리드 오디오와 관련된 표준화가 주로 논의되는 곳은 ITU-R Study Group(SG) 6이다. 이 SG6의 BS파트는 'Broadcasting Service'를 메인 표준화 이슈로 하고 있으며 세부적으로 오디오 메타데이터에 관한 표준화가 논의되는 곳은 WP 6B와 6C이다. WP6B에서는 'Broadcast service assembly and access'에 대해서, 그리고 WP6C에서는 'Programme production and quality assessment'에 대한 표준화를 다루고 있다. 다시 말해 현재 ITU-R에서 오디오 표준화가 이루어지는 부분은 방송 애플리케이션에 집중하고 있다고 말할 수 있다. 특히 두 Working Party 모두 'Audio Metadata' 및 'Multichannel Sound Technology'를 연구하는 라포쳐 그룹이 형성되어 있고 여기에서 해당 논의가 활발하게 이루어지고 있다. 그리고 ITU-R에서 제정되는 표준은 Standard가 아닌 Recommend라는 표현으로 배포되고 있다는 점이 다른 표준화 단체와는 차이가 있는 부분이다.

2014년 Recommendation ITU-R BS.2051[10]을 통하여 제정된 Advanced Sound System은 객체기반 포맷 신호에 대해서도 다루고 있지만 오디오 메타데이터에 대해서는 요구사항만을 언급하고 있어 이를 충족시키기 위한 다양한 논의가 이루어지고 있는데, 그중에서 대표적인 것이 Audio Definition Model(ADM) 모델이다. 이는 European Broadcasting Union(EBU)에서 먼저 2014년 표준화가 이루어진 포맷으로(EBU TECH 3364[11]) 2015년 6월 ITU-R 권고로도 제정이 되었다[12]. 이 메타데이터 모델은 오디오 파일을 포맷파트와 콘텐츠 파트로 구분하여 해당 정보를 안정적으로 서술

하는 것이 목적이다. 포맷파트에서는 채널기반 포맷 신호인지 객체기반 포맷신호인지에 대한 구분, 객체기반 신호라면 공간상의 렌더링 정보, 재생 타이밍 정보 등이 기록되며, 콘텐츠 파트에서는 이렇게 모인 신호들이 영화 콘텐츠인지, 음악 콘텐츠인지 등에 대한 정보를 제공한다. 특히, ADM element로서 다양한 정보를 제공할 수 있는데, 가령, Type Definition이라는 속성에는 하위 항목으로서 DirectSpeakers Matrix, Objects, High Order Ambisonic(HOA), Binaural을 제공하고 있다. 각각, 채널기반 포맷신호를 사용하여 오디오 신호와 스피커 재생 신호를 1:1 매핑시키는 경우, 다운 믹스 혹은 업믹스를 위하여 채널신호가 특정 계수 값을 가지고 더해지거나 빼지는 경우, 객체기반 포맷신호를 사용하는 경우, 'Higher Order Ambisonic' 신호를 재생하는 경우, 헤드폰 입체 음향 신호를 재생하는 경우에 해당한다. 또, audioBlockFormat 속성에는 rtime과 duration 항목을 제공하여 특정 객체의 재생시작 시각과 지속시간 정보를 제공한다. 그리고 position element를 통해 직교 좌표계나 극좌표계로 객체의 공간상 위치정보를 제공할 수 있다. 다시 말해, 사용자 혹은 재생 단말은 객체기반 포맷신호를 전송 받거나 혹은 미디어로부터 읽어들이면, ADM에 정의되어 있는 정보를 이용해서 주어진 재생 시스템 하에서 어떻게 해당 객체를 렌더링할 것인지 결정하여 재생할 수 있게 된다.

이뿐만 아니라 Broadcast Wave Format(BWF)[13]에 기초하여 다채널, 대용량의 'Advanced Sound System'까지 포함할 수 있도록 새로운 chunk를 포함하여 만든 웨이브 파일 포맷을 변경하고자 하는 BW64(Broadcast Wave 64)[14], 실시간 방송 등에서의 활용을 위한 오디오 신호 패키징 정보이자 비트스트림 포맷을 다루는 MDA가 주로 논의되고 있다.

하이브리드 오디오 관련 기타 표준으로는 MPEG에서 MPEG-H 3D Audio라는 차세대 오디오를 위한 부호화

및 렌더링 기술에 대해 표준화하였으며, 이를 기반으로 북미 차세대 방송 표준인 Advanced Television Systems Committee(ATSC) 3.0 표준에 Dolby의 Atmos를 기반으로 한 AC-4 기술의 대체표준으로 반영시킨 바 있다. AC-4 기술은 기존 방송용 AC3 기술의 구조를 따르면서 확장하여 하이브리드 오디오 서비스를 제공할 수 있도록 하고 있다.

한편, 극장 및 블루레이용 하이브리드 오디오 방식이 Dolby의 Atmos, DTS의 DTS:X, Barco의 AuroMax로 분화되어 있어, 극장 시스템 설치 및 콘텐츠 제작에 있어 번거로움이 많다. 이러한 상황을 조금이나마 완화시키기 위해 Society of Motion Picture & Television Engineers(SMPTE)에서는 극장 및 블루레이용 하이브리드 오디오 시스템의 통합을 위해 'Interoperable Immersive Audio'라는 표준을 제정하고, 콘텐츠 제작 절차 및 극장 시스템의 구성에 대한 논의가 계속 진행되고 있다[5].

VI. 맺음말

지금까지 하이브리드 오디오 서비스를 위한 하이브리드 오디오 콘텐츠 편집기술과 하이브리드 오디오 콘텐츠 재현술의 최근 동향과 서비스 현황에 대하여 살펴보았다.

영화산업에서는 차세대 오디오 기술로서 Dolby의 Atmos, DTS의 DTS:X, Barco의 AuroMax가 치열한 주도권 경쟁을 벌이고 있으며, 향후 영화용 오디오 포맷으로 자리잡을 것이 확실시 되고 있다. 방송 산업에서도 미국의 ATSC 3.0 표준화에 객체기반 실감음향 기술이 포함되어 있으며, 유럽의 DVB에서도 차세대 Ultra-High Definition TV(UHDTV)를 위한 오디오 기술에 객체기반 실감음향 기술을 도입하여야 한다는 공감대가 형성되고 있으며, 이를 반영하여 MPEG-H 3D Audio 표준은 객체를 포함하는 실감음향의 압축 부호화뿐만

아니라 이례적으로 오디오 채널 포맷변환 및 바이노럴 렌더링 기술을 표준화 아이템에 포함시켜 대부분의 표준화가 완료되었다[15].

하이브리드 오디오 기술은 기존의 실감음향효과를 개선할 수 있는 가능성이 크며, 그동안 채널기반의 오디오 콘텐츠 제작기술과 재생기술이 서로 종속적 관계를 통해 함께 발전해 왔지만, 산업과 시장의 융통성 측면에서 많은 제한사항을 가질 수밖에 없었던 오디오 시장이 객체기반 실감음향 기술을 통해 제작기술과 재생기술이 종속성을 탈피함으로써 각각 독립적으로 발전할 수 있는 계기가 마련되었다는 측면에서 향후 오디오 시장의 획기적인 변화가 오리라는 추측도 해 볼 수 있다. 하이브리드 오디오 기술은 향후 음향뿐만 아니라 영상 및 오감 기술의 객체기반 서비스의 시발점이 될 것이라 조심스럽게 예측해 본다.

용어해설

ITU 가장 잘 알려진 통신관련 국제 표준화 기구의 하나로 주로 통신, 방송 관련 표준화를 제정하여 권고하고 있음.

ATSC 미국, 캐나다를 중심으로 하는 북미 방송서비스를 위한 표준을 제정하는 기관으로 현재 차세대 방송서비스를 위한 ATSC3.0 표준을 제정하고 있음.

MPEG 1980년도부터 매체의 디지털화에 따른 A/V 부호화 기술의 표준화를 위해 결성된 단체로서 MP3, A/V, HEVC 등 표준을 제정한 바 있음.

SMPTE 영화, TV방송시스템 및 관련 기기의 규격을 제정하는 기구로서 SMPTE 표준 시험 차트나 방송용 1인치(2.54cm) VTR 규격 등이 유명함.

약어 정리

ADM	Audio Definition Model
ATSC	Advanced Television Systems Committee
BWF	Broadcast Wave Format
CAROUSO	Creating, Assessing and Rendering in Real-time Of high-quality aUdio-viSual envirOnments
CES	Consumer Electronics Show
CinemaCon	Convention of Cinema Industry
EBU	European Broadcasting Union
GUI	Graphic User Interface

HOA	High Order Ambisonic
IMAF	Interactive Music Application Format
ITU	International Telecommunication Union
MDA	Multi-Dimensional Audio
MPEG	Moving Picture Experts Group
RMU	Rendering and Mastering Unit
SG	Study Group
SMPTE	Society of Motion Picture & Television Engineers
UHDTV	Ultra-High Definition TV
WFS	Wave Field Synthesis
WP	Working Party

참고문헌

- [1] 장대영 외, “초고해상도(UHD) 사운드 기술의 현재와 미래,” 방송공학회지, 제17권 제4호, 2012. 10, pp. 47-59.
- [2] European project CARROUSO Deliverables, <http://www2.idmt.fraunhofer.de/projects/carrouso/index.html>
- [3] D.Y. Jang et al., “Object-based 3D Audio Scene Representation,” 115th AES Convention Paper, no. 5956, Oct. 2003.
- [4] Dolby Atmos White Paper, “Next-Generation Audio for Cinema,” Dolby Laboratories Inc., 2012.
- [5] Toward an Open-standard Surround-sound Format, <https://www.editorsguild.com/fromtheguild.cfm?FromTheGuildid=454>
- [6] IOSONO, “Spatial Audio Workstation 2 – Operation Manual,” IOSONO GmbH, Germany
- [7] 장대영, 이태진, “객체기반 실감음향 기술개발,” 한국방송공학회 하계학술대회 논문집 T3.2-4, 2015. 7.
- [8] Dolby Atmos Audio Technology, <http://www.dolby.com/us/en/brands/dolby-Atmos.html>
- [9] DTS: X Technology with Multi Dimensional Audio, <http://listen.dts.com/pages/dts-x>
- [10] Recommendation ITU-R BS.2051, Advanced sound system for programme production, <http://www.itu.int/rec/R-REC-BS.2051/en>
- [11] EBU TECH 3364 Audio Definition Model – Metadata Specification, Jan. 2014, <https://tech.ebu.ch/docs/tech/tech3364.pdf>
- [12] Recommendation ITU-R BS.2076, Audio Definition Model, <http://www.itu.int/rec/R-REC-BS.2076/en>
- [13] Recommendation ITU-R BS.2051, File format for the exchange of audio programme materials with metadata on information technology media, <http://www.itu.int/rec/R-REC-BS.1352/en>
- [14] SMPTE, “Immersive Sound for Cinema,” SMPTE Standards Update Webcast, 2014.
- [15] J. Herre et al., “MPEG-H 3D Audio – The New Standard for Coding of Immersive Spatial Audio,” *IEEE J. Selected Topics in Signal Processing*, vol. 9, no. 5, Aug. 2015, pp. 770-779.