

글로벌 UHD 실감형 오디오 제작 전문가 양성과정 - 2

MPEG-H 3D Audio 제작의 실제 (Post Production과 Live Production)

독일 에를랑겐에 위치한 Fraunhofer IIS(프라운호퍼 IIS)에서 6일 동안 MPEG-H 3D Audio 제작 전문가과정 교육을 수료하였다. Fraunhofer IIS(프라운호퍼 IIS)는 유럽 최고의 응용 연구기관으로 독일 전역에 69개의 연구소 및 연구조직을 두고 있다. MPEG-H 3D Audio에 대한 원천기술을 보유하고 있으며 라이선싱 및 인증 프로그램을 관리하고 있다. 국내 지상파 UHD 방송을 위한 오디오 코덱으로 Fraunhofer IIS의 MPEG-H 3D Audio 기술을 채택하였으며 중국, 북미, 유럽에서도 해당 기술을 표준으로 채택한 바 있다.

MPEG-H 3D Audio 제작의 실제 (Post Production과 Live Production)

MPEG-H 3D Audio의 두 가지 특성은 Immersive Audio(몰입 오디오)와 Object Audio(객체형 오디오)라고 하겠다. Immersive Audio를 구현하기 위하여 Upper level 스피커를 채용하였고 Object Audio는 Switch Group 및 Preset 설정을 통하여 Interactivity(쌍방향성)를 추구한다. Workflow 관점에서 보면 Post Production과 Live Production으로 나눌 수 있는데 각각 어떻게 구현하는지를 설명하도록 하겠다.

Post Production

MPEG-H Audio를 제작하는 Post Production 방법에는 MPEG-H Authoring Tool(MHAT)을 사용하는 방법과 DAW에 MPEG-H Authoring Plug-In을 설치하여 제작하는 방법이 있는데 두 가지 방식의 차이점은 첫 번째 방식은 이미 완성된 Audio Source로 작업을 하는 것이고 두 번째의 방식은 실시간으로 모니터링을 하면서 DAW에서 작업을 수행한다는 것이다. MHAT에서는 Component, Switch Group, Preset이라는 계층적 의미의 용어들이 사용되는데 그 의미는 다음과 같다.

Component	Audio track으로 구성되며 Audio scene의 최소 단위
Group	bitstream 내의 signal elements (from single signal to multiple signal)
Switch Group	상호 배타적인 component들의 그룹 (예: 여러 언어의 commentary 시그널)
Preset	모든 group들의 부분 집합들의 조합

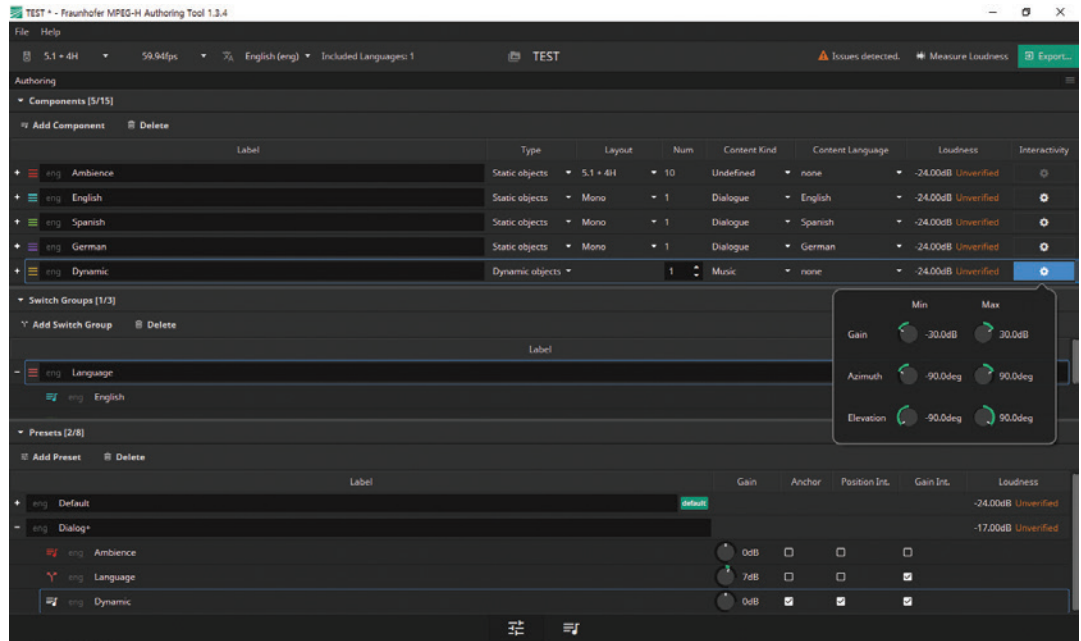


그림 1. MPEG-H Authoring tool (Authoring view)

Components

MHAT(MPEG-H Authoring Tool)에서는 최대 15개의 Component를 지정할 수 있으며 모든 Component에는 실제 Audio file을 지정해 주어야 한다. Add Component 버튼을 클릭하면 새로운 Component를 추가할 수 있다. 이때 팝업되는 Import Sound Track 창에서 추가되는 Component의 특성을 규정하는데 Static Object인지 Dynamic Object 인지, Component를 구성하는 Channel 수와 스피커 layout을 결정해주어야 한다. 또한 MHAT의 경우는 이미 완성된 Audio Source로 작업을 수행하는 stand-alone tool이므로 실제의 Audio file들을 attach 해주어야 하는데 위의 캡처 화면에서 보면 Ambience Component는 5.1+4H이므로 10개의 Audio Input file을 attach 해주어야 하며 English Component는 mono이므로 1개의 Audio Input file을 attach 해주면 된다. Object에는 Static과 Dynamic이 있는데 Component 창의 Interactivity란의 톱니바퀴 모양의 설정 버튼을 눌러보면 그 차이를 알 수 있다. 여기서 각각의 component들의 쌍방향성의 여부 및 정도를 지정하는데 Static의 경우는 Gain 값만 지정이 가능하고 Dynamic의 경우는 Gain, Azimuth, Elevation의 값들을 지정할 수 있다.

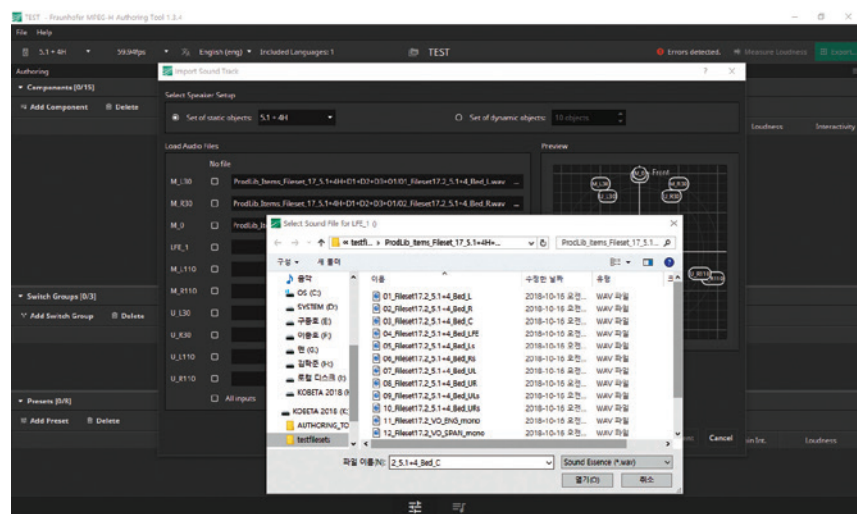


그림 2. Import Sound Track 창

Switch Group

MHAT 창에서 Add Switch Group 버튼을 눌러서 Switch Group을 추가한다. 최대 3개의 Switch Group을 만들 수 있다. [그림 1]에서 보면 Language란 이름의 Switch Group을 생성하였는데 이 이름 위에 English, Spanish, German의 dialog components를 drag & drop 하였다. 이로써 시청자가 리모컨을 통하여 세 개의 언어 중 원하는 것을 선택하여 들을 수 있게 된다. 한 개의 Switch Group에 편입되기 위한 전제조건으로는 동일한 Interactivity 값을 가져야 한다는 것이다.

Preset

Add Preset 버튼을 눌러서 새로운 Preset을 추가한다. 최대 8개의 Preset을 생성할 수 있다. 대표적인 예를 들자면 스포츠 중계를 할 경우 Default와 Dialog+ 라는 이름의 두 개의 Preset을 만든다. Default라는 이름의 Preset에는 Component 창의 Ambience와 Switch Group 창의 Language를 drag & drop 한다. Dialog+라는 이름의 Preset에도 동일하게 해준다. 그다음엔 Preset 창에서 Language Switch Group란의 Gain knob를 +7dB 정도 부스트 해준다. 이후에 시청자가 리모컨을 이용하여 Dialog+ 라는 Preset을 선택할 경우 Default Preset 대비 Dialog Component가 7dB 부스트된 오디오를 들을 수 있게 된다.

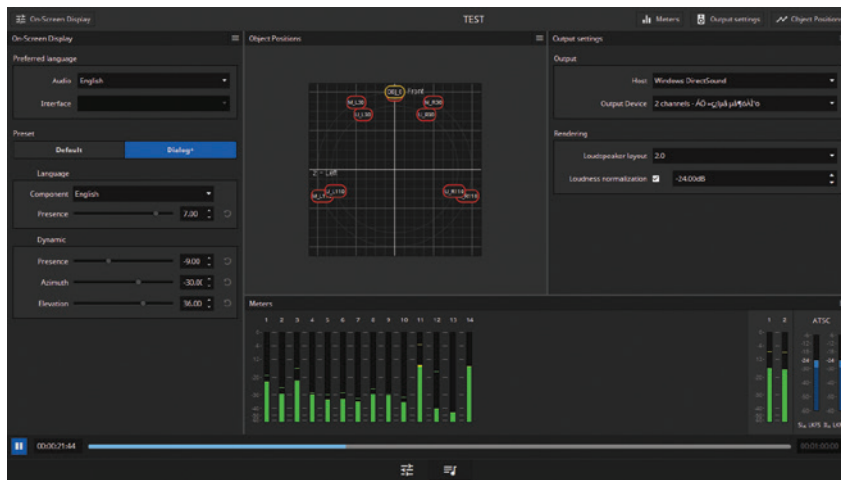


그림 3. MPEG-H Authoring tool (Rendering view)

Live Production

MPEG-H Audio를 Live 환경에서 제작하기 위하여 Authoring과 Monitoring을 수행하는 장비를 AMAU System(MPEG-H Audio Monitoring & Authoring Unit)이라고 한다. 현재 출시된 제품으로는 Jünger 사의 MMA(MPEG-H Multichannel Monitor & Authoring Unit)와 Linear Acoustics 사의 AMS(MPEG-H Authoring & Monitoring System)가 있다. AMAU 장비의 일반적인 Block diagram은 다음과 같다.

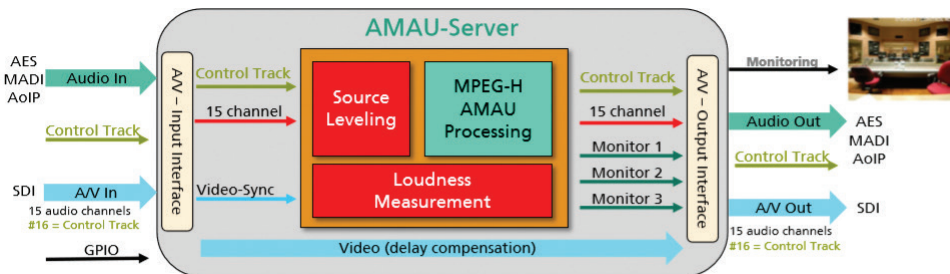


그림 4. General Blockdiagram of AMAU

AMAU 장비는 다음의 세 가지 역할을 수행한다.

MPEG-H Authoring	MPEG-H Monitoring	MPEG-H Loudness
3D Object Arrangement MPEG-H Audio Scene Composition Define Preset Mixes, Languages User Interactivity(Gain, Position) Generation of Metadata	Monitor Mode(Rendering, Bypass) Different Downmixes Preset+Language Selection User Interactivity Control Loudness Normalization	Loudness management Real-Time Loudness Measurement Definition of component Loudness Calculation of Preset Loudness Loudness Adaptation(3rd. party)

AMAU System - MMA(Jünger)



그림 5. AMAU SYSTEM - MMA(Jünger)

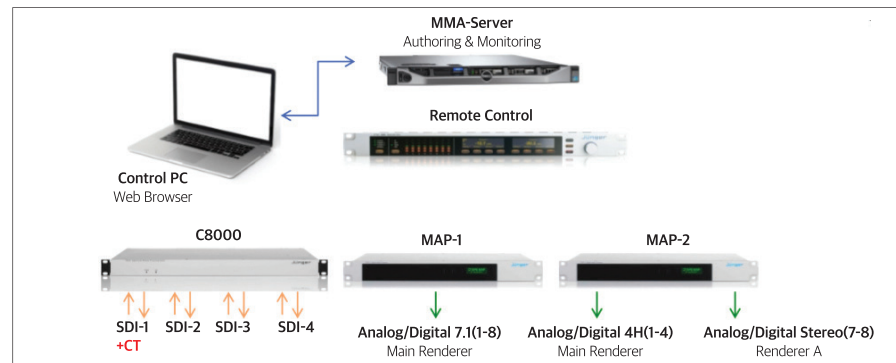


그림 6. Typical MMA Setup

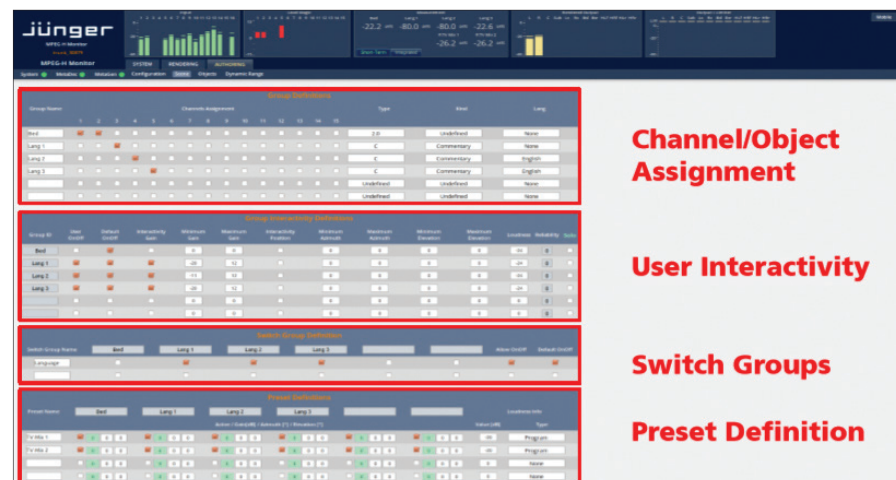


그림 7. MMA-Authoring

AMAU System - AMS (Linear Acoustics)



그림 8. AMAU System - AMS(Linear Acoustics)

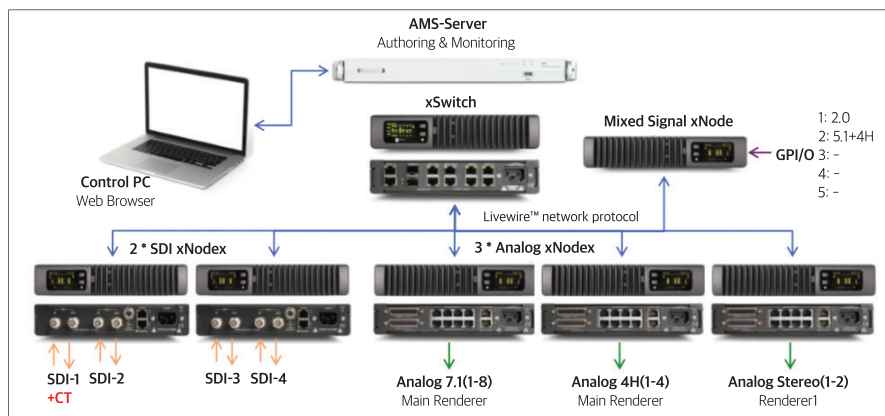


그림 9. Typical AMS Setup

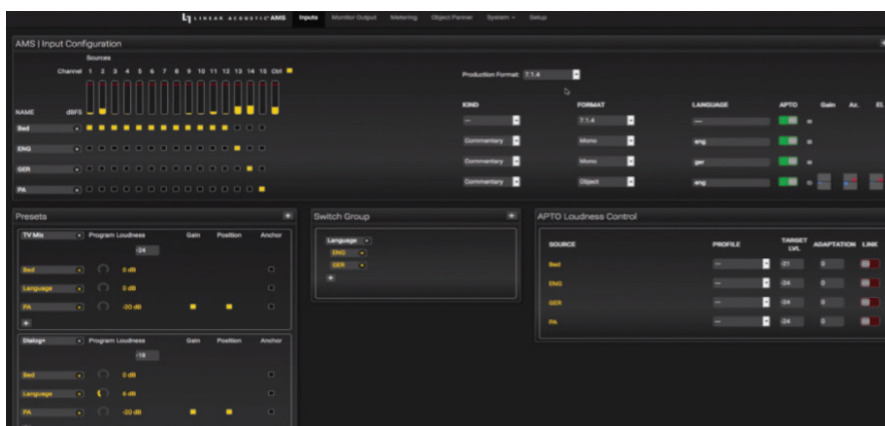


그림 10. AMS Authoring

AMAU 장비인 MMA, AMS에서 Authoring 작업을 수행할 때 [그림 2] MHAT의 Import Sound Track 창에서 음원소스들을 하나하나 Attach 하여주는 대신 All Inputs을 체크하여 Export 하면 Metadata가 생성되는데 이 XML 파일을 Load 하여 사용하면 작업능률을 올릴 수 있다. MMA와 AMS의 Authoring 페이지 사용법은 MHAT과 매우 유사하다.

MPEG-H Audio 제작 Block Diagram

MPEG-H A/V Production - Post Production

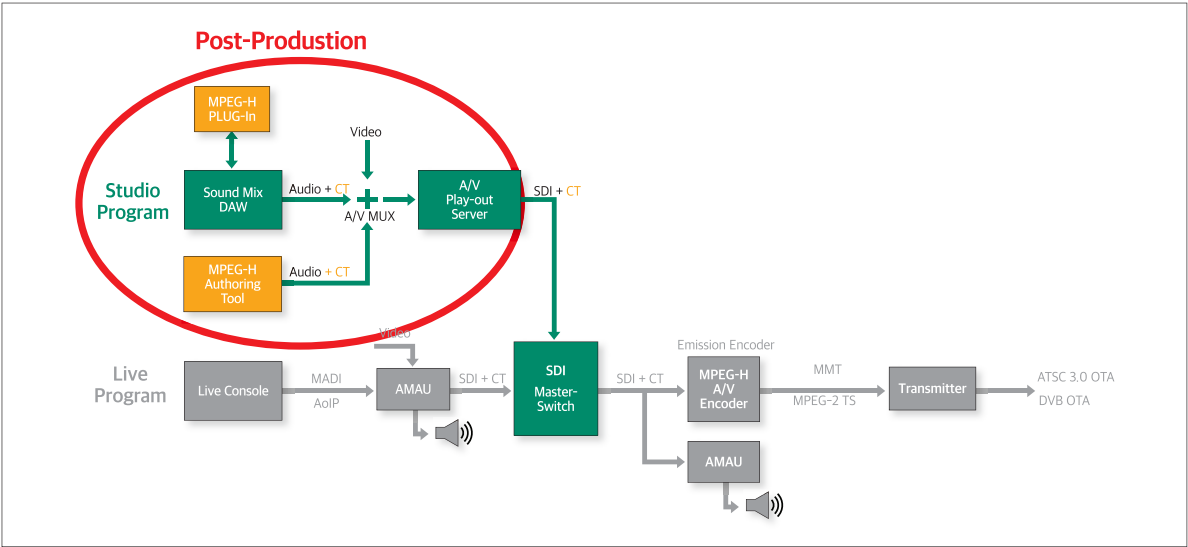


그림 11. Block Diagram of Post Production

역할	Tool
MPEG-H Scene Composition MPEG-H Audio Rendering & Monitoring Generation of MPEG-H Metadata - Control Track	MPEG-H Authoring Tool(MHAT) MPEG-H PLUG-In(SAD-Plug-In of NAT, Fraunhofer IIS) Pyramix MPEG-H Export

MPEG-H A/V Production - Live Production

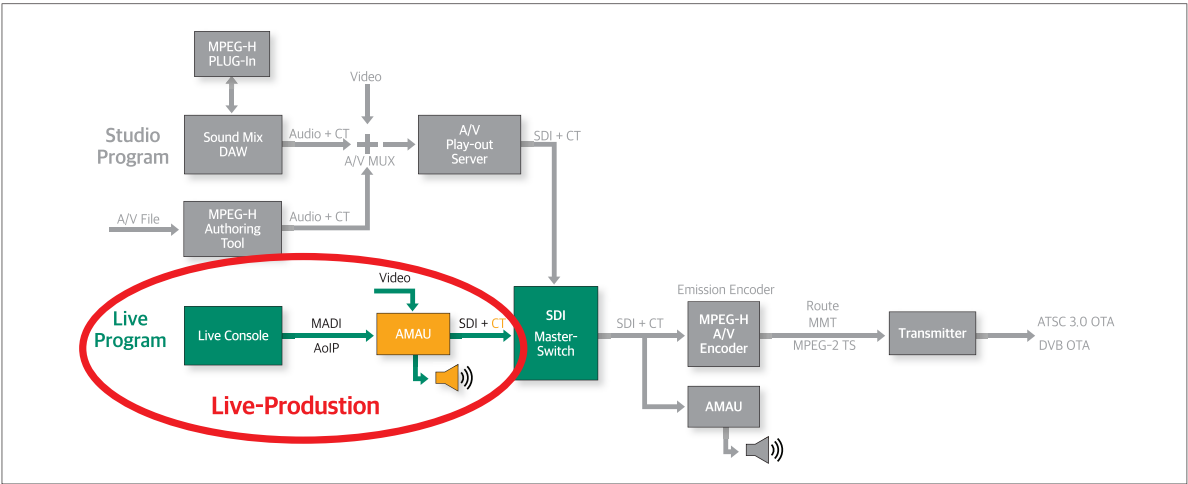


그림 12. Block Diagram of Live Production

역할	Tool
MPEG-H Scene Composition MPEG-H Audio Rendering & Monitoring Generation of MPEG-H Metadata - Control Track	Audio Monitoring & Authoring Unit(AMAU) Audio Concole with MPEG-H Integration(Lawo Immersive Mixing Engine - LIME)

MPEG-H A/V Production - Control & Monitoring

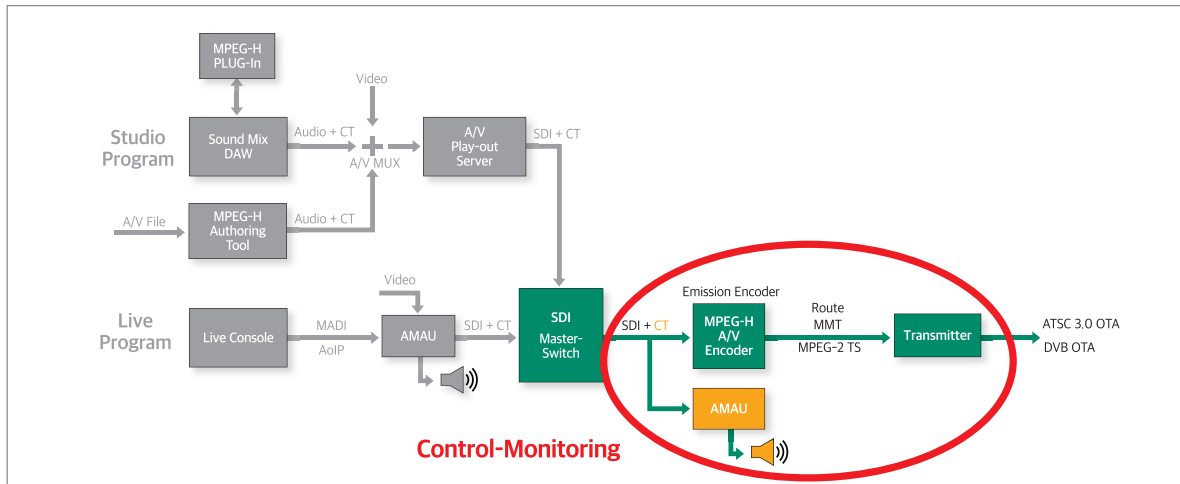


그림 13. Block Diagram of Control & Monitoring

역할	Tool
MPEG-H Audio Rendering & Monitoring	Audio Monitoring & Authoring Unit(AMAU)

MPEG-H 3D Audio를 적용한 제작 실례 (Eurovision Song Contest 2018)

유로비전 송 콘테스트는 유럽 전역의 지원자가 참가하는 콘테스트다. 유럽에서 최대 규모이며 제작비가 가장 많이 드는 방송프로그램이라고 한다. 시청자가 1억 8000만 명에 달한다고 하는데 Audio 제작에 있어서 Immersive & Interactive MPEG-H Live Production을 시도하였다. 실제로 어떠한 Workflow로 제작이 되었는지 소개하고자 한다.

[그림 14]의 Workflow Diagram을 보면 Fraunhofer의 5.0+4H Production Container에 현장의 OB-Van으로부터 Music, Vocal, Hosts-Premix 등의 source가 각각 2ch stereo 신호로 들어오고, Audience 마이크 Signals, 3D Audio를 구현하기 위하여 천장높이에 설치한 4CH의 'Hamasaki-Square' 마이크 Signal 그리고 인터콤을 통하여 여러 언어의 Commentary Signal이 입력된다. AMAU 장비

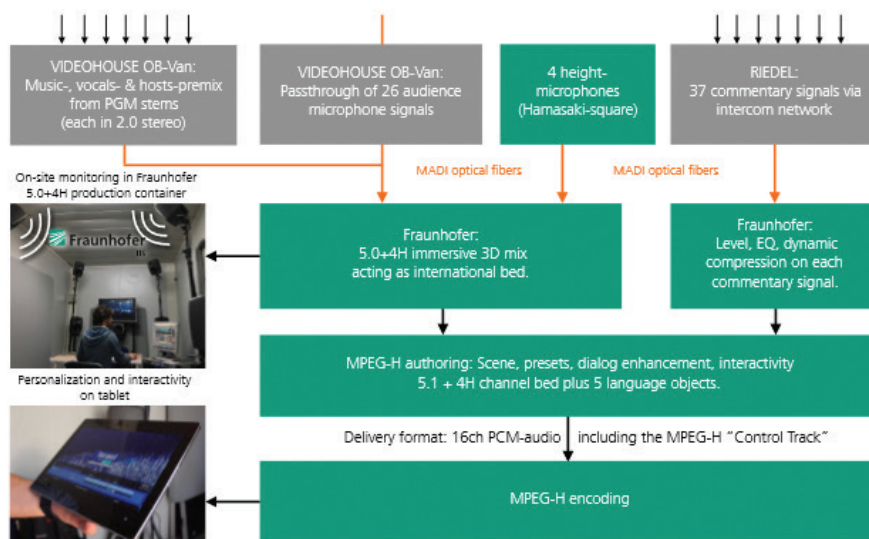


그림 14. Workflow of MPEG-H Live Production(Eurovision Song Contest 2018)

를 사용하여 입력된 이들 Channel로부터 Scene, Presets, Dialog Enhancement, Interactivity를 생성하는 MPEG-H Authoring을 수행한다. 이후 SDI+CT의 신호가 MPEG-H A/V Encoder로 입력되어 Encoding이 수행된다.

맺음말

UHD 방송의 오디오 코덱으로 Fraunhofer의 MPEG-H 3D Audio가 채택된 지금 사람들은 여러 가지 궁금증을 갖게 된다. 기존의 5.1 채널과는 어떻게 얼마나 다른가? 일반 가정에서도 손쉽게 3D Audio를 즐길 수 있을까? 대부분이 Stereo 방송으로도 만족하는데 왜 굳이 3D Audio를 채택하였나? 필자도 주위의 동료들과 대화를 나누어보면 Object Audio를 활용한 AD(Audio Description), 다중 Commentary, Dialog Enhancement에 관하여는 많은 공감감이 있었으나 3D Audio에 관하여는 회의적인 시각들이 많았다. 교육 동안 3D Audio를 체험해보니 분명 기존의 5.1ch과는 확연한 차이를 느낄 수 있었다. 5.1+4H 규격에서는 Middle Layer 스피커 위에 4개의 Upper Layer 스피커를 설치함으로써 머리 위쪽으로부터 소리가 들려오는 3D Audio를 구현한다. 3D Audio가 매우 만족스럽더라도



Fraunhofer 모차르트 룸 (강사: Yannik)

방송타이틀의 부재와 무려 10개의 스피커를 설치해야 한다는 것은 부담감으로 남는다. HD 방송의 5.1ch 시기를 거쳐 오며 방송 Title의 부재 및 가정에서의 스피커 설치에 대한 불편함은 5.1ch 활성화의 큰 걸림돌이 되어왔기 때문이다. 그러나 고무적인 사실은 Soundbar의 설치만으로도 3D Audio를 감상할 수 있고 스마트폰으로 매체를 즐기는 Binaural 환경에서도 3D Audio를 구현할 수 있다는 것이다. 영화, 스포츠, 쇼 장르에 있어서 고품질의 3D Audio를 구현한 방송 타이틀이 많아지고 적정한 가격대의 고품질 Soundbar가 출시된다면 시청자들로부터 환영받는 Audio 기술표준으로 자리를 잡을 수 있을 것이라 생각한다.

독일 에를랑겐에서 꼭 찬 6일간의 일정이었습니다. 수업을 마치고 주위에 식당이 없어 에를랑겐, 뉘른베르크 시내에 나가 저녁 식사 때 한두 잔 마신 독일 생맥주는 참 맛있었습니다. 인솔하여주신 김동한 감독님께 고맙고, 최선을 다하여준 Fraunhofer의 강사진도 너무 고맙습니다. Audio에 대한 식지 않은 열정으로 함께 토론하며 열공한 선후배님들께도 감사의 말씀을 전합니다. 🍷



Fraunhofer 연구소 앞에서