

[멀티미디어응용] MPEG-H 3D Audio Phase 1 & Phase 2 표준화 현황

멀티미디어 국제 표준을 주도하고 있는 MPEG에서는 최대 22.2채널까지, 입력 신호로는 채널(Channel)신호, 객체(Object)신호, 고차 앰비소닉스(High-Order Ambisonics) 신호까지 지원 가능한 차세대 고품질 오디오 코덱으로 MPEG-H 3D 오디오를 표준화하고 있다. MPEG-H 3D 오디오는 다시 UHD 방송 표준에 대응하기 위한 주요 기능들에 대해 확정하는 Phase 1과 모바일까지 영역을 확장하는 Phase 2로 나뉘어 서로 다른 일정으로 표준화를 진행하고 있다. Phase 1은 지난 2015년 5월 ISO/IEC 23008-3으로 국제 표준(IS)이 정식 승인되어 이후 북미 UHD 방송표준인 ATSC 3.0의 Audio Codec 표준 후보로 제출되어 현재 Dolby의 AC-4와 최종 단계 경합을 벌이고 있다. ATSC를 겨냥하여 MPEG에서는 Fraunhofer-IIS, Technicolor, Qualcomm 3사가 중심이 되어 MAA(MPEG-H Audio Alliance) 라는 연합을 형성하고 이 연합을 중심으로 방송 표준화 논의를 계속하고 있다.

MPEG-H 3D Audio는 오디오 신호를 표현하는 형식에 따라 채널, 객체, 고차 앰비소닉스(HOA) 3가지를 각각 또는 통합적으로 전송할 수 있는 종합적인 오디오 표준인 것이 특징이다. 채널, 객체에 HOA까지 통합은 MPEG-H 3D Audio 표준이 차세대 Immersive Audio를 위한 종합적이고 총체적인 솔루션이 되고자 하는 취지도 있겠으나, 보다 현실적으로는 ATSC 3.0 표준을 비롯한 방송 표준에서의 선정 및 상용화의 관점에서 보다 많은 이해관계자가 MPEG 진영에 동참할 수 있도록 하여 Dolby, DTS 등의 시장 경쟁 우위를 확보하기 위한 의도도 있는 것으로 해석된다. HOA는 현재 Technicolor, Orange 이외에도 Qualcomm 에서 적극적인 상용화 노력을 기울이고 있다.

한편, 방송 표준에 있어서 중요하게 요구되는 기술 가운데 한가지가 DRC(Dynamic Range Control)이다. 이는, 입력 오디오 신호마다, 방송 채널마다, 콘텐츠마다, 그리고 사용자 환경에 따라 제각각인 소리에 대한 재생 크기, 동적 영역을 통제할 필요에 의해 제기되었다. MPEG에서는 지난 2012년 Apple의 첫번째 발의에 따라 MPEG-D 의 부속 표준으로 DRC의

표준화가 진행되었으며, 마찬가지로 최근 국제 표준이 완료되었다. 더불어 DRC 기술은 MPEG-H 3D 오디오 표준의 일부로도 포함되었다. 전술한 것과 같이 DRC 기술은 신호의 Dynamic range를 일정하게 조절하는 기술로써 방송 시청자의 불편을 해소할 수 있는 점에서 방송 응용에서 높은 관심을 가지고 있는 기술이다. 아울러, DRC가 MPEG-H에 포함됨에 따라, MPEG 진영은 Qualcomm에 이어 Apple까지 우군을 더 확보할 수 있는 계기가 되었다.

폴란드, 바르샤바에서 진행된 이번 제112차 MPEG 미팅에서는 IS 발행이 완료된 MPEG-H 3D Audio Phase 1에 대해서는 Reference Software 업데이트, Conformance Test 의 Initiation 등의 표준 후속 관리가 진행되었으며, 한편, 이번 회의를 통해 가장 큰 이슈는 Phase 2가 Phase 1과 통합되어 보다 의미 있는 표준으로 대두되는 분위기가 만들어졌다는 점과, 일명 System Data Carriage 기술(삼성전자 제안)이 이해관계에 따라 대척점에 있는 기관 및 단체들간의 논쟁으로 이슈가 계속되었다는 점을 들 수 있다.

첫 번째 사안의 경우, 지금까지 MPEG Audio 그룹의 분위기는 방송 표준 대응 MPEG-H 3D Audio 는 Phase 1으로 사실상 일단락하고, Phase 2는 “bitstream compatible”을 전제로 개선할 수 있는 기술들을 부분적으로 검토하는 것으로 정리되어 있었다. 그런데, ATSC 등 유관 표준 단체의 일정 등을 고려할 때, 그리고, Phase 1 표준이 아직 시장에 적용된 사례(legacy decoder)가 없다는 점을 전제할 때, Phase 2에 혁신적인 기술들을 추가하고, 이를 Phase 1과 통합한 형태로 MPEG-H 3D Audio라고 불러도 될 것 같다는 의견이 대두되었다(오디오 서브그룹 좌장이 이와 같이 리드하였고, 특별히 반대하는 기관은 없었다.). 그리고, 미리 알고 있기라도 한 듯, MPEG-H 3D Audio의 main proponent인 Fraunhofer IIS 및 컨소시엄 기관들이 Phase 2에 새로운 기술들을 CE로서 다수 기고하였고, 대부분 채택되었다. 이와 같은 기술들의 채택을 통해, 자연스럽게 Phase 2와 Phase 1은 bitstream 호환성에 대한 이슈가 제거되었다. 따라서, 의미 있는 기술은 다음 113차 미팅부터 Phase 2 표준이 완료되기 전까지 새로운 기술로써 기고 가능하다. 채택될 경우, 방송표준에 사용되는 기술로 길이 열릴 수도 있다.

두 번째 사안의 경우, 방송 시스템 및 콘텐츠 전체의 사업모델 관점에서 쟁점이 되었던 기술이다. 현재, 우리나라, 북미를 비롯하여 대부분 나라에서 지상파 방송의 TV 직접 수신 경우보다 IPTV 등 케이블과 인터넷망 등을 통한 전송이 보편화되는 추세이다. 이렇게 형성된 에코시스템 내에서 TV 방송국의 주요 수익모델 중 한가지인 광고에 대한 수익권을 지역 케이블 등 중계사업자가 가져가게 되었고, 이에 대해 TV 방송국 입장에서 수익권을 제고할 수 있는 기술 중 한가지로서 본, System Data Carriage가 이슈가 되었다. 제안된 기술을 이용할 경우, STB(Set-Top-Box)에서 TV로 전송된 오디오 비트열 내에, 해당 오디오뿐 아니라 영상 등 콘텐츠 전체에 대한 관련 부가 정보를 수신할 수 있는 URL 데이터 등을 포함할 수 있다(STB와 TV는 HDMI로 연결되는데, HDMI를 통해 bitstream 자체가 전송되는 것은 오디오뿐이다.). 따라서, 케이블랩스 등 STB 기반의 네트워크 사업자들은 본 기술을 부정적인 입장일 수밖에 없으며, 본 기술을 통해 타 경쟁 오디오 코덱과 차별화하려는 MPEG-H 3D Audio 표준 옹호 진영과 이를 반대하는 진영 간에 liaison 등을 이용한 큰 충돌이 진행되고 있다. 이번 112차 회의에서는 이해관계자들의 충돌을 조정하는 과정에서 본 기술이 담긴 표준의 표준화 일정이 한 단계 delay되어 issue가 되었으며, 따라서 후속 표준화 회의를 통해 논의가 더 진행될 것으로 보인다.

한편, MPEG-H 3D Audio 표준은 본격적으로 Rendering 기술이 표준화의 scope 안으로 들어갔다는 점에서도 시사하는 바가 크다. 기존의 코덱 표준의 scope은 통상적으로 decoding된 raw wave file에 해당하는 오디오 신호를 생성하는 과정까지이다. Interoperability를 제공하는데 필요한 최소한의 범위를 고려한 것이다. 그런데, 기술이 고도화되고, 시스템이 복잡해지면서, 앞서 표준의 scope이라고 생각된 범위까지 만으로 기술을 정의할 경우, 그 이전 혹은 이후의 작업을 각 구현자들이 독자적으로 구현하기에는 복잡성이 높아 만족할 만한 품질을 제공하기 어려운 한계에 이르는 문제가 발생하고 있다. 해당 표준을 만든 해당 기술의 전문가가 더 완성된 형태의 end-to-end solution을 제공하는 것이 구현자의 재량에 맡기는 것보다 해당 기술의 상용화를 더 쉽게 만들 수 있다는 여론이 형성되었다. 또 한가지 관점은 User Experience에 대한 표준화의 관점으로 생각할 수도 있다. 결국 MPEG-H 3D Audio 표준이 전달하는 콘텐츠를 소비하는 최종점은 청취자의 “귀”일 수 있고, 따라서 “귀”에서의 품질이 중요한데, 기존 표준은

이를 위해 필요한 Rendering이나 DRC와 같은 기술은 정의하지 않고, decoding까지만 정의해왔던 것이다. 그 결과 품질이 통제되지 않은 제품이 양산되거나 시장 수용성이 떨어지는 형태의 결론에 이르렀다. 이에 대한 변화로써, MPEG-H 3D Audio는 Format Converter라 불리는 target speaker layout에 대한 적응을 위한 기술, Headphone Virtualization을 위한 Binaural Rendering 기술, 종합적 DRC 기술까지 모두 표준에서 정의하고 있다. 이와 같은 패러다임 변화가 어떤 시장 수용성을 가져올 지 또한 향후 지켜볼 대목인 것 같다.

오현오 (윌러스표준기술연구소 연구위원, henney.oh@wilusgroup.com)