

[멀티미디어응용] MPEG의 새로운 오디오 표준화 과제는 “제대로” 된 3차원 오디오

멀티미디어 국제 표준의 대표명사라 할 수 있는 MPEG에서 다음 세대 멀티미디어 콘텐츠를 위한 표준으로 MPEG-H를 준비하고 있다. MPEG-H는 “고효율 (압축)부호화와 동적인 미디어 전송(High efficiency coding and dynamic media delivery)”이라는 주제를 가진 프로젝트로서 기존에 큰 성공을 거둔 표준 기술인 MPEG-2와 MPEG-4의 뒤를 잇는 새로운 코덱 표준 패밀리가 될 것으로 기대된다. 여기서 멀티미디어 코덱이라 함은 오디오 코덱과 비디오 코덱 그리고 이를 잘 포장하여 전송하는데 사용되는 시스템을 통틀어 일컫는다 하겠다. 이 가운데, 비디오는 HEVC(High Efficiency Video Coding)라는 이름으로 이미 표준화가 거의 완료 단계에 이르렀고, 시스템 역시 MMT(MPEG Media Transport)라는 이름으로 표준화가 한참 진행 중에 있다. 이에 비해 오디오는 현재 표준화를 시작하는 단계인 CfP(Call for Proposal; 표준 기술 제안 요청서)도 발행하기 직전인 초기에 있다.

MPEG-H에서 오디오가 다른 부문에 비해 뒤늦게 시작하게 된 이유는, 직전의 주요 오디오 표준 기술인 USAC(Unified Speech Audio Coding)의 완성 시점을 고려했기 때문이라 해석할 수도 있겠으나, 비디오 경우와 다르게 오디오 코덱의 경우 이미 압축 기술 자체는 수렴 단계에 이르러, 새로운 코덱이 나온다고 해도 종전의 2배가 좋아지는 식의 혁신은 기대하기 힘들기 때문이기도 하다. 바꿔 말하면, 이미 완성한 기존의 MPEG 계열의 오디오 코덱들(AAC(Advanced Audio Coding), MPEG Surround, SAOC(Spatial Audio Object Coding))과 그 밖에 시장에서 널리 사용되는 상용 오디오 코덱들(Dolby사의 AC3, DTS사의 DTS 등)이 차세대 멀티미디어 시장에서 사용하기에 큰 불편함이 없기 때문이기도 할 것이다.

이와 같은 고민과 MPEG-H 프로젝트가 추구하는 주제를 고려하여, MPEG 오디오 서브그룹에서는 지난 2012년 2월 99차 회의서부터 이번 101차 회의(2012년 7월)까지 3회에 걸쳐 MPEG 내외부의 전세계 전문가들을 초빙한 오디오 워크숍을 개최하였다. 이 워크숍의 주제는 본 MPEG-H 오디오 코덱의 주제이기도 한 3차원(3D) 오디오이다.

3D 오디오라는 단어는 이미 20년은 족히 시장에서 사용되었을 만큼 식상하다. 소비자에게 사용자에게 새로운 느낌을 제공하기 쉽지 않을 것 같은 주제어이다. 그럼에도 불구하고, MPEG은 ‘3D 오디오’라는 단어를 빼어 들었다. 다만, 기존의 3D 오디오와 구별되는 몇가지 차별점을 포함한 단어라 할 수 있다.

우선, 3D(Dimension, 차원)는 너비, 깊이, 높이 이렇게 3개의 차원을 포함함을 말한다. 기존 오디오 시장에서 3D와 함께 널리 사용되던 용어인 서라운드(Surround)는 너비와 깊이 2개의 차원을 가진 2D 오디오라 할 수 있다. 그러니, 과거에 3D 오디오의 보편적 의미는 아직

2D였다고 할 수 있다. 그 이전의 스테레오(2개의 스피커를 통해 제공하는 사운드)가 제공하던 음장에서 깊이의 한계를 보강하여 더 나은 2D 오디오를 제공했다고 할 수 있겠다.

이에 대비하여 진정한 3D는 서라운드 높이에 대한 차원을 더함으로써, 진정한 임장감(Immersion, 라이브 공연 등의 현장에 있는 듯한 실제감)을 제공하는 경우를 말한다. 스피커를 사람 귀 높이에만 평면적으로 배치하는 것이 아니라 그 위쪽에 하나의 스피커 계층을 두고, 또한 그 아래 하나의 스피커 계층을 두어서 높이에 따른 소리의 정위감을 제공한다. 이와 같은 스피커 배치의 한 예로 22.2채널 포맷이 있다. 이는 NHK에서 최초로 제안하였으며, 이번 MPEG-H 오디오 표준화에서 다루고자 하는 기본 포맷이기도 하다.

높이 채널에 대한 중요성을 높이 여기는 이유로 UHDTV 사용 환경을 생각해볼 수 있다. UHDTV, 즉, 초고해상도 TV가 제공하는 해상도를 기존의 HDTV와 단순 비교하면 최대 16배 큰 화면을 제공할 수 있다. 즉, 거실의 한 벽면 전체가 하나의 TV용 디스플레이가 될 수 있는 셈이다. 이에 대응하는 시청자의 시청거리를 고려하면, 화면 상의 객체가 좌우뿐 아니라 높낮이에서도 차별적인 소리를 제공할 수 있어야 할 것이다. 또한, 가까운 시청거리는 좌우각이 100도에 이르는 넓은 시청각을 요구하는데, 이에 대응하기 위해서는 기존보다 많은 수의 스피커를 통해 화면 구석 구석에 소리를 배치시킬 수 있어야 하는 문제도 있다.

이렇게 등장한 22.2채널 포맷은 다시 새로운 기술적 질문을 제기하였다. 기존 서라운드 오디오의 대표적인 포맷인 5.1채널의 경우도 아직까지 덕네 보급 및 사용률이 높지 않은 상황에서, 과연 각 가정에 22.2채널에 대응하는 스피커를 구비하게 될 것인가 하는 것이다. 또한, 22.2채널 스피커를 모두 구비한다고 하여도, 각 스피커의 정확한 위치와 거리를 확보하여 청취공간을 구성하는 것은 0.1%의 극상층을 제외한다면, 사실상 불가능하다. 따라서, 비록 22.2채널로 생산된 콘텐츠라고 하더라도 이보다 적은 수의 스피커, 다른 위치에 배치된 스피커, 심지어는 헤드폰으로 청취하는 사용자에게도 가급적 원 제작자의 의도에 충실한 사운드 장면을 제공할 수 있어야 한다. MPEG-H는 이 문제를 표준화의 핵심 화두로 던지고 있다. 즉, 기존의 오디오 코덱과 같이 단순히 압축을 얼마나 더 잘하느냐가 아니라, 실제 재생 환경에서 쓸모가 있는 오디오 코덱을 만들고자 한다.

이와 같은 고민 속에 자연스럽게 등장한 것이 객체 오디오 신호이다. 여기서 객체란 사운드 장면을 구성하는 각 음원들을 의미한다. 즉, 기타, 드럼, 보컬이 3인조 밴드라면, 기타, 드럼, 보컬이 각각 하나의 객체라고 볼 수 있다. 기존의 오디오 신호를 다루는 패러다임은 이와 같은 객체들이 한데 어우러진 사운드 장면을 재생단의 스피커에 1:1로 매칭되는 채널이라는 단위로 묶어서 처리하는 것이었다. 이와 같은 채널 대신, 객체 단위로 신호를 저장하고, 전송하게 된다면, 그래서 사용자단에서 이를 다시 재구성(렌더링이라고 부름) 할 수 있다면, 앞서 제기한 문제, 즉,

사용자의 스피커 구성 환경에 독립적으로 제작자가 원하는 사운드 장면을 제공해야 한다는 필요를 비교적 손쉽게 해결할 수 있다. 즉, 3인조 밴드의 각 객체 신호를 독립적으로 보내주었다면, 사용자가 22개 스피커를 가지든, 헤드폰으로 청취하든 그 환경에 맞게 렌더링 할 수 있다.

객체 오디오 신호에 대한 코덱은 사실 앞서 2006년에 표준화를 시작하여 2009년에 완성한 SAOC(Spatial Audio Object Coding)라는 MPEG 표준이 있다. 따라서, 지금까지 MPEG의 표준화 관행(?)으로 추정하건데 이번 3D 오디오 표준화 시에 기존의 SAOC 코덱의 틀들을 최대한 재활용(?)하고자 할 것이다. 그렇지만, SAOC가 압축 효율을 극대화하기 위해 패러메트릭(Parametric) 부호화 기법을 사용하고 있어서, 초고음질을 지향하는 3D 오디오에 맞지 않는 부분이 있을 것이다. 그리고, 이런 부분들은 새로운 표준화의 아이템이 될 것으로 예상된다.

3D 오디오 표준화의 논의 가운데, 채널과 객체 이외에 새로운 입력 신호로 거론된 것이 HOA(High Order Ambisonics) 신호이다. 마이크를 통해 오디오 신호를 취하고, 이를 다시 재생하는 일련의 과정에 있어서, 특수하게 제작한 마이크와 이를 표현하는 특수한 저장 방법을 이용하는 HOA는 기존의 채널 혹은 객체의 패러다임과는 또 다른 형태로 사운드 장면을 표현하는 포맷이다. 이 기술이 비록 학계에서는 많이 연구되고 있으나, 이 분야에 선구적으로 투자를 하고 있는 프랑스의 Orange Lab(프랑스 텔레콤), 독일의 Technicolor(구 Thomson Deutsche Lab) 등을 제외하고는 대응할 만한 저변 기술을 가지고 있는 곳이 없는 편이다. 그런 이유로 본 표준화의 대상으로 포함하고자 하는 논의 과정에서 조차 많은 갈등을 빚었다. 특히, MPEG 오디오 표준화 회의의 종주기업이라고 할 수 있는 독일의 프라운호퍼 연구소(Fraunhofer IIS, 이제는 International Audio Lab.이라는 회사를 별도로 세웠다)에서도 강한 거부감을 가지고 있는 포맷인지라, 비록, 현재 표준화의 대상으로 포함은 되어 있으나, 표준화가 완성된 시점에서는 어떤 모습으로 위치하게 될 지는 관심 있게 지켜봐야 할 것 같다. 물론, HOA를 연구하고 있는 국내외 연구기관이 있다면, 이번 표준화가 그 기술을 널리 사용하도록 저변화 할 수 있는 좋은 기회라고 생각할 수도 있다.

향후 추진일정

현재까지 결정된 3D 오디오 표준화의 일정은 다음과 같다. 우선 다음 미팅인 102차 회의(2012년 10월)에서 CfP를 발행할 예정이다. 그리고, 참여기관이 각자의 기술을 제출하고 이에 대한 음질평가를 통해 RM(레퍼런스 모델)을 결정하는 것은 104차 회의(2013년 4월)로 예정되어 있다. 그리고, 105차 회의(2013년 7월)에 가면, 본 고에서 얘기하는 3D 오디오의 실체가 무엇인지 그 내부를 들여다볼 수 있는 RM 소스 코드와 WD(Working Draft) 문서를 열람하게 될 것이다. 그리고 1~2년정도의 협업(Collaboration) 기간을 거치면서 표준 기술을 완성하면, 3D Audio에 대한 국제 표준 문서가 발행될 것으로 예상된다.

오현오 (굿데이랩 대표, ho@gooddaylab.com)