김재현
돌비 코리아 대표

AC-4, 새로운 방송환경을 위한 차세대 오디오 포맷

차세대 방송환경 속으로

디지털 엔터테인먼트는 과거와는 비교도 안 될 빠른 속도로 진화를 거듭하고 있다. 소비자들은 기존에 존재하지 않았던 새로운 경험을 갈구하고 있으며, 이에 대한 열망은 점점 더 커지고 있다. 이러한 추이는 오디오 시장에서도 나타나고 있으며, 그 결과 새롭고 혁신적인 오디오 전송 기술의 니즈로 이어지고 있다.

비디오 코딩 기술의 진화는 오디오 코딩 기술을 재정의하는 중요한 기회로 작용하고 있다. 2013년 1월 표준으로 제정된 HEVC(High Efficiency Video Coding) 압축 표준은 엔터테인먼트 경험에 새로운 지표를 제시하면서 오디오 산업에도 새로운 기술의 필요성을 시사했다. HEVC의 압축률은 H.264 AVC(Advanced Video Coding) 보다 약 2배 정도 뛰어나다. 또, 압축률의 개선과는 별도로 HEVC는 UHD(Ultra High Definition) TV의 요구사항인 고해상도를 부호화할 수 있는 효율적인 전송도 지원한다. 계산량의 경우, AVC 대비 2배 정도 복잡하지만 현재 수준의 CPU 프로세싱 능력으로도 충분히 계산이 가능하다. 기존의 HDTV 수신을 지원했던 TV, 셋톱박스(set-top-box, STB) 등의 다양한 재생기기들은 UHD급의 영상을 복호화하기 위해 하드웨어의 변경이 불가피했던 만큼 이는 새로운 오디오 전송 규격의 제안과 사용의 아주 좋은 기회가 될 수 있다.

묵시적으로나마 소비자들은 이 같은 변화에 걸맞은 새로운 UHD급의 오디오 경험도 원하고 있기 때문에, HD에서 UHD의 진화는 단지 비디오 해상도의 진화로 끝나서는 안 된다. 차세대 오디오 시스템에 대한 돌비 래버러토리스(Dolby Laboratories, 이하 돌비)의 비전은 이런 시장의 기대에 충분히 대응하고 있다. HEVC 서비스를 위한 하드웨어 혹은 IC(Integrated Circuit, 집적 회로)의 변경은 오디오 측면에서도, 차세대 오디오 전송 기술로 교체할 수 있는 아주 좋은 기회이다.

AC-4란?

AC-4(Audio Coding - 4)는 돌비에서 개발한 차세대 오디오 포맷으로 작년 4월 유럽 전기통신 표준협회(European Telecommunications Standards Institute, ETSI)를 시작으로 전 세계 표준화 기구에 표준화 활동을 하는 새로운 오디오 전송 기술이다. AC-4는 브로드밴드, 방송, 인터넷 등의 다양한 전송환경이 요구하는 압축효율을 제공하며 새로운 오디오 경험(Audio Experiences)을 제공할 수 있는 유연성을 갖고 있다.

오디오 압축 효율

AC-4는 각각 다른 전송 및 저장환경에서 요구하고 있는 다양한 시나리오에 맞춰 압축효율을 지원하도록 설계됐다.

특히 돌비는 오디오 산업에서 널리 쓰이는 특정 채널과 비트레이트(bitrate)를 중심으로 새로운 부호화 도구(Coding Tools)를 개발했다. 이를 통해 스테레오 음원을 32kbps로, 5.1 채널 음원은 96kbps로 부호화할 수 있다.

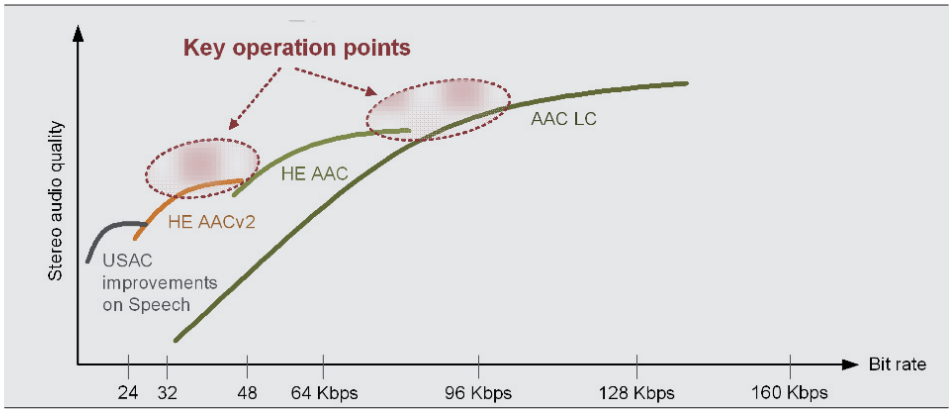


그림 1. Key Stereo and 5.1 Operating Points

다음 표는 기존의 돌비 디지털 플러스(Dolby Digital Plus, E-AC3)와 AC-4의 대표적인 비트레이트에 대한 음질평가 결과를 정리한 것이다. 전체적으로 비트레이트를 기존 대비 50%만 소비해도 비슷한 결과를 얻을 수 있었다.

(단위: kbps)	돌비 디지털 플러스		돌비 AC-4		향상율
	좋음 (Good)	매우 좋음 (Excellent)	좋음 (Good)	매우 좋음 (Excellent)	
모노 (대화소리)	48	64	24	32	50%
모노 (일반)	64	96	40	56	~40%
스테레오	96	160	56	96	~40%
5.1 채널	192	256	96	160	40-50%
7.1 채널	384	640	160	256	~60%
몰입형 오디오	384	896	256	384	30-60%
전체					~50%

AC-4는 또한 극저 비트레이트로 음성신호를 부호화할 수 있어, 음성신호 위주의 서비스인 뉴스나, 장면 묘사 트랙(scene descriptive track), 제작자 코멘트(directors' comment) 등의 서비스에도 효율적이다.

객체화 오디오 전송(Objective Audio for immersive, personal)

지금의 채널 기반 오디오 전송과는 달리 의미 있는 독립 음원 하나하나를 객체화해 오디오를 가공하는 객체기반 오디오(Object based audio)는 기존의 채널 기반 오디오보다 몰입감 있는 오디오 경험을 전달하며, 그 응용방법이 다양하다. 기존의 채널 기반 오디오(channel based audio)는 제작 당시 사용됐던 스피커의 배치를 재생단(Playback side)에서 그대로 사용할 때 원래의 의도가 잘 전달되지만, 사용자의 청취환경은 이와 다른 경우가 많다. 또한, 채널 기반 오디오는 채널의 개수가 콘텐츠 제작자의 의도를 제한하는 요소로 작용해온 반면에 객체 기반의 오디오는 제작자가 원하는 오디오 경험을 객체화된 오디오의 부가 데이터(Object Audio metadata)로 저장 및 재생한다. 해당 정보는 스피커의 개수와 위치가 상이한 재생환경에서도 제작자의 의도가 그대로 반영된 오디오 신호를 재생하도록 지원한다. 이런 방법은 제작자에게 기존 시스템보다 표현의 자유도를 높여줄 수 있다는 점에서 주목할 만하다. 또한, 객체화 오디오 기술은 오디오 신호 하나하나를 독립적으로 가공 및 전송할 수 있어 개인화 오디오(personalized audio) 같은 응용서비스에도 사용될 수 있다.

김재현
돌비 코리아 대표

손쉬운 비디오-오디오 편집

기존 오디오 인코딩의 처리단위인 하나의 오디오 프레임(audio frame)은 코덱의 설계에 따라 그 고유한 길이를 갖고 있었다. 이는 비디오 신호와 결합돼 하나의 동영상 스트림을 구성하는데 오디오의 한 프레임의 길이와 비디오의 프레임률(frame rate)이 맞지 않아 접합(splicing)이나 삽입(insertion)과 같은 추가적인 처리가 필요했다. 하지만 AC-4 오디오 코딩 시스템은 인코더(encoder)에서 입력되는 비디오 신호의 프레임률 값을 참고해 해당 값을 기준으로 오디오 프레임 길이를 결정하므로, 데이터를 보다 손쉽고 정확하게 가공할 수 있게 됐다. 이런 기능은 오디오 코덱 중 AC-4만이 가진 특징이다.

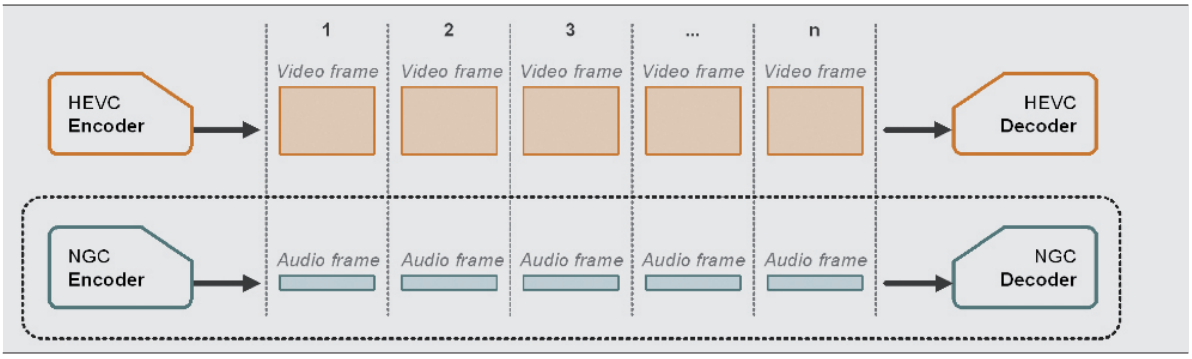


그림 2. Video Framing

다이내믹 레인지(Dynamic range) 제어와 다이얼로그 개선(Dialogue enhancement) 기술

오디오 신호 전송 시 다이내믹 레인지(Dynamic range)를 제어하는 것은 무척 중요하다. 이는 오디오가 재생되는 재생 시스템이나 사용자가 청취하는 환경이 다양하기 때문이다. 더구나 요즘은 기존 거실에 설치된 TV 외에도 방송이나 기타 콘텐츠를 소비하는 방법과 환경이 더욱 다양해지고 있기 때문에 양질의 오디오 데이터를 전송하는 데 있어 주변 상황에 최적화된 다이내믹 레인지를 제어하는 것이 매우 중요하다.

돌비 디지털(Dolby Digital, AC-3), 돌비 디지털 플러스(Dolby Digital Plus, E-AC-3) 등 기존 돌비의 오디오 신호 전송 규격과 같이 AC-4에서도 제작자의 의도가 잘 반영되도록, 다양한 재생장치 및 청취환경에 대응할 수 있는 다이내믹 레인지 제어를 제공하고 있다. 또한, 관련 데이터가 오디오 신호의 유통과정에서 훼손되는 것을 방지하기 위해 증명된 메타데이터(authenticated metadata)라는 특수한 방법을 활용하고 있다.

AC-4의 또 하나의 특징은 다이얼로그 개선(Dialogue enhancement) 기능이다. 청취자가 드라마나 영화를 감상할 때 사람의 대화 오디오 신호만을 독립적으로 조절할 수 있는 인터페이스를 기반으로 개개인의 사용자에 최적화된 콘텐츠 제공이 가능하다.

오브젝트 오디오를 이용한 방송

몰입형 오디오(Immersive audio)

채널 기반 오디오 대비 객체화 기반 오디오가 가질 수 있는 장점은 의미있는 오디오 단위의 독립적인 제어가 가능하다는 것이다. 이런 특징은 보다 몰입적인(immersive) 오디오 제작에 좋은 도구가 될 수 있다. 콘텐츠 제작자 입장에서 본다면 원하는 소리 하나하나를 객체화해 구분할 수 있고 이를 제작자의 의도대로 공간과 시간에 직관적으로 배열할 수 있다. 소리 표현에 대한 자유도가 채널 기반 오디오 대비 매우 증가한 것이다.

재생환경에서도 다양한 스피커 구성에 맞게 의도된 소리가 재생되도록 동작하는 오브젝트 오디오 렌더러(object audio renderer)가

존재하므로 청취환경의 유연성 측면에서 기존의 채널 기반의 오디오보다 유연해졌다고 말할 수 있다.

돌비는 이미 영화산업에서 돌비 애트모스(Dolby Atmos)라는 몰입형 오디오 기술을 소개했으며, 현재 많은 돌비 애트모스 영화들이 제작되어 개봉 했거나 개봉을 앞두고 있다. AC-4는 이런 콘텐츠들을 가정 환경에 전달하는 중요한 수단이 되므로 몰입형 오디오 방송에서 중요한 역할을 할 것이다.

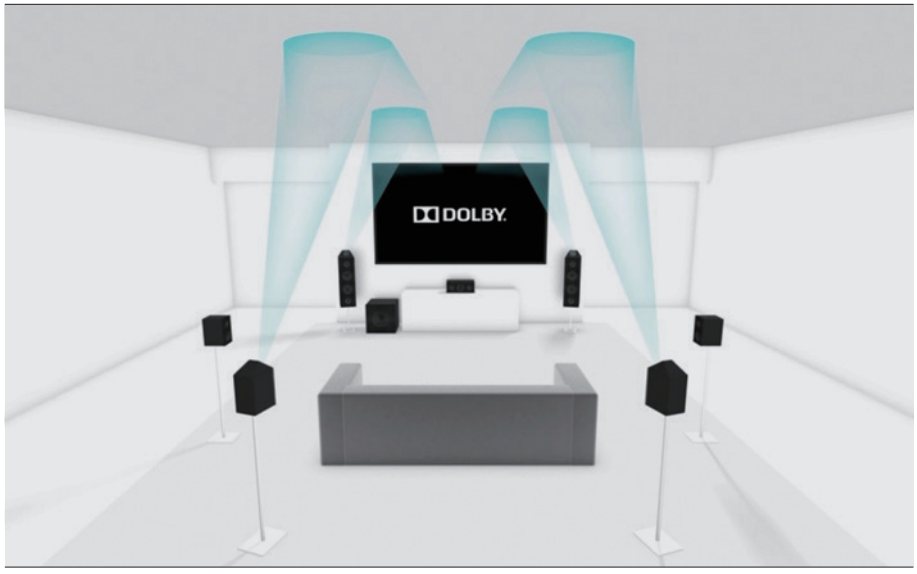
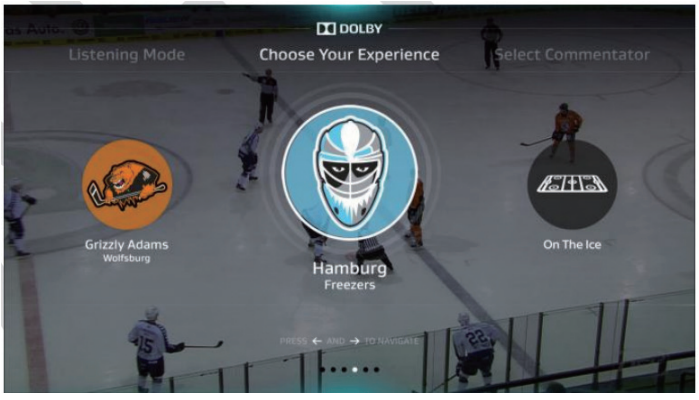


그림 3. 몰입형 오디오 전달을 위한 7.1.4 ch 스피커 배치 예

개인화 오디오(Personalized audio)



또 다른 방송의 예는 개인화 오디오다. 객체단위의 오디오를 제어는 다음과 같은 스포츠 경기 등에 적용될 수 있다. 스포츠 경기중계 중 전송되는 오디오는 주로 중계석의 중계 내용과 해설 내용, 경기장 현장의 소리가 될 수 있다. 여기에 운동선수나 혹은 의미 있는 장소에 마이크를 배치하여 보다 다양한 종류의 오디오를 획득하고 이 모든 소리를 객체단위로 전송하면 청취자는 방송 시나리오 별로 듣고 싶은 오디오만 청취할 수 있다. 예를 들면 A 팀을 응원하는 중계만 들을 수 있다든지, 혹은 중계 및 해설 내용은 제거하고 마치 경기장에서 관람하는 것처럼 현장의 오디오만 들을 수 있는 새로운 형태의 개인화 오디오 서비스가 가능해진다.

결론, 준비된 차세대 오디오 포맷

2012년 돌비 애트모스(Dolby Atmos) 기술을 선보인 돌비는 객체화 오디오 기술을 통해 몰입형 경험의 가능성을 제시했으며, 해당 기술을 가정환경으로 확장하는데 노력을 기울이고 있다. 이 과정에서 습득한 다양한 경험을 바탕으로 돌비는 AC-4라는 새로운 오디오 방송규격을 개발했다. AC-4는 단순히 압축률이 좋은 오디오 코덱이 아니라, 시장의 새로운 니즈를 반영한 전혀 새로운 경험을 가능케 하는 차세대 오디오 전송 규격이다.