

디지털 초고음질 음원 서비스에 대한 고찰

민병갑 EBS 라디오 기술감독

바야흐로 UHD 초고화질 TV가 대중에게 팔리고 있는 시대가 왔습니다. 아날로그 TV, 컬러 TV 서비스의 시작은 역사책이나 나올 아득한, 그야말로 역사가 되었고, 이 아날로그 방송 서비스가 종료되고 HDTV로 전면 전환된 것은 이제 만 2년이 지났습니다. 물론 1997년에 그 시절의 정보통신부에서 지상파디지털방송추진협의회 운영이 시작되고, ATSC 방식이 결정되었으니 HDTV 규격 디지털 전환의 시작은 오래되었다고 할 수 있군요.

그런데 이런 변화의 과정에 비해서 UHD 서비스 도입과 서비스 과정은 엄청나게 짧은 시간에 이루어지는 것 같습니다. 필자 개인적으로도 6MHz에서 UHD 화질의 방송 송수신이 가능하다는 것 자체가 마법같이 느껴지는데, 제 주위의 일반 시청자 중 일부는 이렇게 빨리 서비스가 가능해졌다는 얘길 들으면, “이런 게 되면 애초에 좋은 것으로 하지, TV 팔아먹으려고 그러냐?!”하는 말씀도 하곤 합니다.

물론 이런 얘기 들으면, 방송사 직원으로서 “사실 이런 과정이 압축 기술이 빨리 발전했고, TV 제조사도 좋은 기술을 빨리 적용하고 어찌 고저찌고.....” 라는 비 공대생에게 절대 알아 먹히지 않을 말을 하고 있습니다. 실상 저도 잘 알지 못하니 설명을 못 하는 것이 당연하겠지요.

어쨌든 이렇게 영상 분야에서의 디지털 포맷은 엄청난 발전이 끊이지 않는데 비해서, 음향 분야에서는 디지털 하면 예나 지금이나 다들 MP3만 말씀하시고 떠올립니다. 사실 그동안 음향분야에서도 많은 포맷이 개발되었지만, 여러 가지 이유에 따라 아직도 MP3가 가장 많이 쓰이고 있는 것이 현실이죠.

그런데 음향에서 디지털 매체가 대중에게 널리 퍼진 것은 CD가 원조라고 할 수 있고, CD는 디지털 저장매체 이후 음악을 담을 수 있는 가장 매력적인 매체로 오랜 기간 동안 생명력을 가져왔습니다. 하지만 이제는 이 얇고 작은 CD조차도 MP3에 밀려서 생명력을 점차로 잃어가고 있는 것이고요.

한 가지 재미있는 것은 CD를 카세트테이프의 연장선으로 생각하시고, 조그마한 기기에서 제목이 함께 디스플레이 되는 MP3를 디지털 음원의 시작으로 알고 계시는 분들이 은근히 많다는 사실입니다.

이런 얘기를 듣고 나누다 보면 ‘역시 미디어는 디지털이고 아날로그고 간에 포맷이 중요한 것이 아니라 **쉽고 편하고 재밌게** 대중에게 다가설 수 있는 것이 갑이다라는 생각이 들기도 합니다. 뛰어난 기술보다는 안에 담기는 콘텐츠가 더 중요하다는 말입니다 (MP3가 폭발적인 성장을 한데는 불법 공유도 크게 한몫했다고 생각하지만, 지금은 논외로 하겠습니다). 그런데 지금 이야기하고자 하는 것이 콘텐츠는 아닙니다. 음향분야의 UHD라고 할 수 있는 초고음질 포맷인 24bit / 96KHz 이상의 음질과 이 음질을 담아낸 대표적인 코덱인 HFPA

와 FLAC입니다.

사실 음악에 관심이 많은 사람은 한 번이라도 들어봤을 법한 유명한 고음질 포맷이 하나 더 있었습니다.

SACD

이 SACD(Super Audio Compact Disc)는 델타-시그마($\Delta\Sigma$) 변조라는 1비트 DSD(Direct Stream Digital) 방식을 그대로 사용하는 방식입니다. 이 델타 시그마 변조 자체는 SACD를 위해서 개발된 포맷은 아니고, A/D, D/A 컨버팅을 하는데 있어서 기기 특성에 덜 의존되며 선형성을 보존하기 위해서 개발된 이론이며 대부분의 컨버터에 쓰이는 A/D, D/A 변환 방식이라고 합니다. CD에 담기는 PCM의 64배에 이르는 샘플링(2.8224MHz)과 1비트(On/Off) 방식의 양자화로 고음질을 구현하는 방식인데, 아무리 오버샘플링을 한다 하더라도 이 방식이 베이스밴드 비전문가인 저로서는 쉽게 와 닿지는 않네요.

이에 반해 PCM 방식은 데이터를 굉장히 쉽게 표현할 수 있고, 따라서 소프트웨어적으로 처리가 쉽다 보니 컴퓨팅 능력이 발달한 지금 시대에서는 고음질 처리가 상대적으로 쉽게 이뤄질 수 있습니다.

바로 이런 점에서 녹음과 재생에 있어서 전용 기기가 필요한 SACD가 시장의 선택은 받지 못했지만, 핸드폰에서 24bit PCM 음원이 처리되는 지금 시대에 이르러서야 고음질 음원 시장이 조금씩 퍼지고 있는 상황입니다.

그런데 여기까지 상황을 보면 정말 아이러니합니다. 일반인들은 구분도 못 하는 차이를 위해서 상위 규격의 포맷과 기기가 나오고, 그런데 이 기기가 일반인들의 선택을 받지 못해서 살아남지 못하고 있는 현실인 것이죠. 그런데 너무나 당연합니다. SACD의 경우를 놓고 보면, 고음질을 감상하기 위해서 SACD로 제작된 디스크와 SACD 구동 기기(특이한 포맷을 사용하다 보니 당연히 구동기기도 비싸고, 음반 자체도 비쌌다)를 구입 해야만 들을 수 있는 “불편함”이 필수기 때문입니다.



SUPER AUDIO CD

[그림 1] SACD Logo

그나마 불편함이 약간은 해소되고 출시된 것이 바로 HFPA(High Fidelity Pure Audio)입니다.

HFPA

이미 오래된 이야기지만 post-DVD 영상 고화질 매체에서 승리한 진영은 소니의 Blu-ray 디스크 진영입니다. 현재의 블루레이는

Primary Audio	CODEC	LPCM	Dolby Digital	Dolby Digital Plus	Dolby Lossless	DTS digital surround	DTS-HD	DRA	DRA Extension
	Max. bitrate	27.648 [Mbps]	640 [kbps]	4.736 [Mbps]	18.64 [Mbps]	1.524 [Mbps]	24.5 [Mbps]	1.5 [Mbps]	3.0 [Mbps]
	Max.ch	8(48kHz, 96kHz), 6(192kHz)	5.1	7.1	8(48kHz, 96kHz), 6(192kHz)	5.1	8(48kHz, 96kHz), 6(192kHz)	5.1	7.1
	bits/sample	16, 20, 24	16 - 24	16 - 24	16 - 24	16, 20, 24	16 - 24	16	16
	Sampling frequency	48kHz, 96kHz, 192kHz	48kHz	48kHz	48kHz, 96kHz, 192kHz	48kHz	48kHz, 96kHz, 192kHz	48kHz	48kHz, 96kHz

[그림 2] Specification of BD-ROM Primary audio streams

자연스레 24비트 192KHz 2채널의 운용이 가능했던 DVD Audio처럼, Bluray Audio 매체가 등장했고, 이 중, 유니버설뮤직이 주축이 되어 만든 포맷이 바로 HFPA Bluray Disc입니다. 2013년 등장한 이 포맷은 SACD와 비교가 되는데, 음질만 놓고 보면 DVD-A, SACD 등과 같이 모두 24bit / 192KHz급으로 음질적으로 크게 개선된 것은 없어 보입니다. 하지만 오디오 코딩은 기본적으로 Linear PCM이며, 무손실 압축인 Dolby True HD나 DTS-HD Master Audio 코덱을 이용합니다.

이 HFPA가 기대되는 이유를 들여보자면,

1

SACD와 달리 일반 블루레이 플레이어에서 재생이 가능하다 보니, 기기 보급률이 이미 높다는 점.

2

복제 방지 기술을 기반으로 음원 신뢰성이 높다는 점을 대표적으로 꼽을 수 있을 것 같습니다.

첫째는 쉽게 이해할 수 있는 내용이고, 둘째는 출처를 알 수 없는 고음질 포맷과 달리 음반사가 주축이 되어 생산한 진정한 고음질 마스터본이 음원에 담긴다는 뜻입니다. 스튜디오에서부터 녹음되고 마스터링 된 24bit 포맷의 음반을 그대로 들을 수 있다는 뜻입니다.

그런데 개인적으로 HFPA 또한 CD만큼의 호응을 얻을 수 있을까 하는 의문에는 다소 부정적인 견해를 가지고 있습니다. 음질의 문제가 아닌 서비스 접근성의 문제로 말이죠. 물론 앞서 얘기했듯이 블루레이 플레이어가 대중화되었기 때문에, 감상이 SACD 출시 상황에 비해서 쉬운 것은 맞습니다. 그러나 시장이 형성되기 위해선 다수의 대중에 어필할 수 있어야 한다고 생각되는데, 블루레이 플레이어가 있다고 하더라도 HFPA 음반을 구입해서 감상할 청취자가 과연 얼마나 될까 하는 의문이 있습니다.

잘 아시다시피 오디오를 효과적으로 감상하기 위해서 스피커의 역할이 몹시 중요한데, 제가 방문했던 집 중 많은 분들이 5.1ch 스피커 다섯 개를 TV 옆에 나란히 놓고 감상하고 계시 정도로 일반인에게 있어서 제대로 된 음악 감상 환경 구성이 쉬운 편이 아닙니다.

그리고 복제 방지 기술에 관해서도 녹음 방법이 너무나 쉬운 오디오 분야에서 과연 기대하는 만큼의 효용이 있을까 하는 의문이 드는 것도 사실이구요.

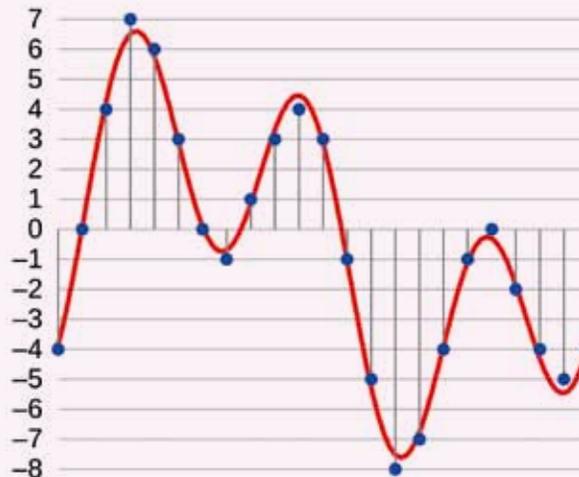


[그림 3] HFPA 공식 Facebook의 음반 소개

여기서 SACD부터 HFPA까지 거의 비슷하게 지원하는 고음질 포맷인 24bit와 192KHz PCM의 의미를 잠깐 살펴보겠습니다.

모두 잘 아시다시피 CD 음원은 16bit/44.1KHz PCM입니다.

이는 아날로그 신호를 X축은 1초당 44,100개로 쪼개고, Y축은 $2^{16}=65,536$ 개로 쪼개서 디지털화했다는 뜻이죠. 간단히 얘기하면, CD 음질 디지털 파일이면 초당, X축 44,100칸, Y축 65,536칸의 격자를 만들고 아래의 [그림 4]와 같은 물결무늬(소리원본)를 그려 넣고 물결무늬가 지나가는 자리 가까이에 있는 위치들을 데이터로 저장하고 있는 것이며, 소리 자체가 담기는 것은 아닙니다. (아래 그림은 4bit 양자화입니다)



[그림 4] 음원의 4bit 디지털화

여기서 44,100은 인간 최대 가청 주파수인 22KHz를 두 배로 만들고(나이퀴스트 이론), 여유분을 조금 더했다는 설을 인터넷에서 금방 찾을 수 있는데, 기술이라는 것이 이렇게 허술하게 만들어지진 않습니다. 답은 44,100Hz가 인간이 들을 수 있는 소리를 다 담을 수 있는 최소 주파수 22KHz의 나이퀴스트 이론 주파수인 44KHz를 초과하면서, 그 당시 Sony와 Philips가 주도하던 비디오 규격과 맞출 수 있는 주파수기 때문입니다. 좀 더 기술적으로 얘기하면, 과거 아날로그 주사방식 NTSC 방식 525라인 60Hz와 PAL 방식 625라인 50Hz에

서 각 “Hz×유효라인”의 최소 공배수의 배수 중 첫 숫자가 44,100이기 때문입니다.

NTSC 유효주사수 : $490/2 = 245 \times 180 = 44,100$

PAL 유효주사수 : $588/2 = 294 \times 150 = 44,100$

(2로 나누는 이유는 비월주사 interlace 방식이기 때문입니다)

따라서 이런 방식으로 샘플링하면, NTSC나 PAL 방식 어떤 것이라도 CD랑 같이 동작하는데 싱크 문제가 발생하지는 않을 것입니다.

그런데 인간이 들을 수 있는 소리는 22,000단계(44KHz)면 충분한데 왜 192KHz까지 나왔을까요?

먼저는 일단 영상이 디지털화되고, 규격이 다양해지면서 44.1KHz를 고집할 필요성이 없어진 데다가 44.1KHz가 과연 충분한 음질을 제공하고 있는가에 대한 의문이 생긴 것입니다.

44.1KHz를 더 작게 잡아. 22.05KHz라도, 음악이든 말이든 다 알아들을 수는 있습니다. 그런데 [그림 4]처럼 디지털화를 하는데 있어서 각 구간의 간격이 멀어지다 보면(낮은 비트수와 낮은 샘플레이트), 그 사이는 컴퓨터가 특수한 알고리즘에 근거해서 값을 예상합니다. 물론 이 값을 산출하는 방식과 알고리즘은 여러 방식이 많이 있습니다. 어쨌든 컴퓨터가 실제로는 담기지 않은 값을 계산해서 넣다 보니 실제 소리와는 어떻게든 달라지고, 이는 음질에 영향을 미치게 됩니다. 좀 오버해서 표현하면 디지털이 아날로그를 절대 따라갈 수 없는 것입니다.

그리고 이렇게 직접적으로 Waveform에 관한 부분 외에도 A/D, D/A 변환이나 양자화 비트수 변환 등에서 쓰이는 Dithering 기술이라는 것이 있는데, 이는 간단히 설명하면 잡음을 넣어서 작은 소리의 해상력을 올리는 (말도 안 되는 것 같은) 기술입니다.

4KHz 근처의 소리를 가장 잘 들을 수 있는 사람 청각 특성이나 마스킹 효과 등으로 삽입된 잡음(디터 신호)이 가려지긴 하지만, 분명한 것은 전체 잡음도는 증가한다는 것입니다. 물론 이 잡음을 잡아낼 수 있는 사람은 드물긴 할 겁니다.

어쨌든, 라디오라는 저음질(?) 매체를 통해서 소리를 전달하는 방송 제작 현업자 입장으로, 스튜디오에서 발생하는 생생한 목소리와 음악을 청취자 분들에게 그대로 전달하고 싶은 욕심이 없다면 거짓이겠지요. 그래서 양자화 오류나 디터링 잡음 등에 가장 원초적으로 대응할 수 있는 방법이 바로 비트 수와 샘플레이트를 최대한으로 끌어올려서 녹음하고 송출하는 것입니다. 그것이 바로 24bit / 192KHz 기술로 각종 음반에서 고음질 서비스 얘기만 나왔다 하면 지원하는 규격입니다.

이 24bit, 192KHz는 1초를 X축을 192,000단계로 쪼개고, Y축은 $2^{24}=16,777,216$ 단계로 쪼개서 저장한다는 것인데, 양자화 단계에 있어서 그만큼 오류도 줄어들고, 디터링 과정에서의 잡음도 최소화됩니다.

그런데 아이러니한 점이 있습니다.

바로 인터넷에 24bit, 192KHz만 검색해보시면, 많이 보실 수 있는 정보로 “과연 24bit/192KHz 필요한가”류의 기사입니다. 사실 인터넷에서 쉬운 경로로 구할 수 있는 24bit 음악 중에서 많은 음원들은 CD에서 추출하거나 LP를 재녹음한 음원입니다. 이렇게 CD에서 추출한 음악을 24bit/192KHz로 만드는 것은 그야말로 호박에 줄을 긋고 수박이라고 말하는 것과 같습니다.

24bit/192KHz라는 그릇에 담았지만 내용은 16bit/44.1KHz 데이터이기 때문입니다.

제대로 된 효과를 보기 위해서는 스튜디오에서 24bit/192KHz로 제작된 음원을 그대로 듣는 것입니다.

이런 부분을 HFPA에서는 복제방지 기술을 탑재하고, 음반사에서 직접 HFPA 블루레이 디스크를 판매하며 음원 신뢰성을 확보하는 방식을 선택하고 있습니다. 또 한 번 말씀드리지만, 개인적으로 이런 방식이 정상적이고 확실한 방법임에는 동의하지만, 서비스 확대 차원에서는 오히려 벽으로 작용하지 않을까? 하는 생각을 가지고 있습니다.

앞으로 시간이 흐르면서 음원들이 기본적으로 24bit 수준으로 녹음/재생 되는 시점이 언젠가는 올 것으로 믿습니다. 과거엔 컴퓨터 성능이 충분하지 못해서 재생이 불가능했지만, (현재 UHD/4K 영상을 범용 기기에서 실시간 디코딩을 할 수 없는 것과 비슷합니다) 현재는 FLAC이라는 공개 무손실 압축 코덱을 통해서 일반 기기에서도 감상이 되고 있습니다. 무려 스마트폰에도 이 기능이 탑재되기 시작했습니다. 그런데 사실 디지털 파일을 감상하는 데 있어서 컨버터의 역할은 몹시 중요합니다. 여기서 컨버터의 성능은 논외로 하고, 24bit/192KHz 실시간 디코딩이 스마트폰에서도 가능해졌다는 것 자체도 큰 의미가 있다고 생각합니다. 여기서 만약 A/D 컨버터, 즉 녹음에 있어서도 24bit 처리가 스마트폰에서 이루어질 만큼 대중화된다면 24비트 포맷이 디지털 음향 표준으로 자리 잡을 것입니다.

앞서 말씀드린 HFPA의 성공에 의문을 갖는 이유도 이렇게 FLAC이라는 공개 포맷을 통해서 24bit / 192KHz가 범용적으로 디코딩이 되는 현실에서 FLAC을 통한 고음질 시장이 클 것이라 생각되지, 과연 일반 소비자들께서 HFPA 매체를 구입할까 하는 생각이 들기 때문입니다. 고화질이든 고음질이든 일반 대중을 대상으로 콘텐츠 서비스를 하는 입장에서는 쉽고 편하고 좋은 방법이 최고입니다. (물론 합법의 테두리 안에서입니다)

EBS 라디오의 현재 송출 표준은 48KHz입니다. 개인적으로도 48KHz면 충분하지라고 하고 다녔는데, 차세대 신규 시스템을 기획하면서 24bit / 192KHz 청취 환경을 EBS 제작기술부(음향팀)에 구축한 적이 있습니다. 담당자로서 오랜 시간 감상할 기회가 자연스럽게 생겼는데, 처음 들 때는 솔직히 쉽게 느낄 수 없었지만 들으면 들을수록 해상도나 공간감의 질감 등 미묘한 부분에서 차이가 점점 느껴졌습니다.

모 음향 현업자의 표현으로는 "더러운 안경 쓰다가 깨끗이 닦아서 다시 쓰는 느낌"이라고 하는데, 사실 (가는 귀가 쉽게 먹는) 음향 종사자는 물론이고 일반 감상 위주의 애호가분들도 사실 48KHz만 듣고 이게 192KHz인지 48KHz인지 구분하긴 쉽지 않을 겁니다. 특히나 팝 계열의 음악은 더욱더 의미가 낮은 것 같고, 국악, 재즈, 클래식 등의 아날로그 음악이나 공개방송에서는 큰 효과를 낼 수 있다고 생각합니다.

현재 EBS FM에서는 많은 방송 콘텐츠를 생산해내고 있는 데 반해서, 스토리지 용량은 한정적이라 EBS FM 표준 제작 규격은 당분간은 48KHz로 유지할 것으로 예상됩니다. 하지만 EBS의 차세대 FM 제작 시스템의 편집 소프트웨어만이라도 32bit-384KHz까지 지원 가능하도록 설계되어 있고, 이를 통해서 초고음질의 오디오 콘텐츠를 생산뿐 아니라 청취자 귀 까지 24비트를 송출할 수 있는 날이 오기를 기대합니다.

“잘 듣는 것”이 “잘 보는 것” 보다는 다소 어려운 점이 많아서 음향 제작 현업이 꽤 소외되는 느낌이 있지만, EBS 음향 엔지니어들 모두는 묵묵히 청취자, 시청자를 위해서 최고의 소리를 추구하고 있습니다. 감사합니다. 🎧

그림출처

[그림 1] http://en.wikipedia.org/wiki/Super_Audio_CD

[그림 2] 2.B Audio Visual Application Format Specifications for BD-ROM Version 2.4 (April 2010)

[그림 3] <https://www.facebook.com/HFPureAudio1>

[그림 4] http://en.wikipedia.org/wiki/Digital_audio