

이머시브 오디오의 다운믹스 청취 환경에 따른 공간감 변화분석(7.1.4채널 스튜디오와 5.1채널 아파트 거실의 정량적, 정성적 교차 분석을 중심으로)

윤오성
한국영상대학교 음향제작학과
e-mail:solid-state@pro.ac.kr

Analyzing Spatial Perception in Downmixed Immersive Audio(A Mixed-Methods Comparison of a 7.1.4 Studio and a 5.1 Home Environment)

O-Sung Yun
Dept. of Music & Sound Technology, Korea University of Media Arts

요약

최근 3차원 몰입형 오디오(Immersive Audio) 콘텐츠가 대중화되고 있으나, 이를 소비하는 주요 공간인 공동주택(아파트) 환경은 전문 스튜디오와 물리적, 음향학적으로 큰 괴리를 지닌다. 본 연구는 층간소음 문제로 인해 청취 음압이 제한되고 다채널 스피커 배치가 어려운 아파트 거실 환경에서, 낮은 음압과 5.1 채널 다운믹스(Downmix) 재생이 청취자의 몰입감에 미치는 영향을 교차 분석하였다.

연구 결과, 청취자가 아파트 환경에서 수용 가능한 평균 체감 음압은 62.3 dBA로 음악 스튜디오 표준(82dB SPL)에 비해 현저히 낮았으나, 동시 측정된 물리적 저주파 에너지는 평균 73.9 dBC로 나타나 약 11.6 dB의 압력 차이가 발생함을 확인하였다. Object 의존도와 주파수 마스킹이 높은 Pop, Rock 장르는 이머시브 효과의 왜곡이 컸으나, 클래식 및 어쿠스틱 타악기 장르는 낮은 음압의 5.1 다운믹스 환경에서도 높은 명료도와 공간감을 유지하였다. 본 연구는 주거 환경의 물리적 한계 속에서 몰입형 오디오의 실효성을 극대화하기 위한 믹싱 전략 및 재생 환경의 대안을 제시한다는 점에서 의의를 가진다.

이 논문은 2026년도 한국영상대학교 지원을 받아 수행하였음.

의 제약을 받게 된다.

1. 서론

그리고 대부분의 가정에서는 사운드바를 활용하여 5.1 채널 이하로 다운믹스하여 재생하는 것이 현실이다. 이는 창작자가 의도한 3차원 공간 정보의 손실과 음원 소스 간의 주파수 마스킹을 필연적으로 발생시킨다.

1.1 연구의 배경

최근 영상 매체 및 음원 스트리밍 플랫폼을 중심으로 Dolby Atmos를 비롯한 객체 기반 몰입형 오디오(Object-based Immersive Audio)의 보급이 급증하고 있다. 창작자들은 7.1.4 채널 이상의 다채널 스피커와 82dB SPL 수준의 레퍼런스 음압이 확보된 전문 스튜디오 환경에서 입체적인 정위감과 포위감(Envelopment)을 섬세하게 디자인한다.

그러나 실제 소비자가 콘텐츠를 경험하는 홈 엔터테인먼트 환경, 특히 국내 주거 형태의 대다수를 차지하는 아파트 거실은 제작 환경과 많은 차이가 있다. 층간 소음 분쟁에 대한 우려로 인해 청취자는 물리적 진동을 유발하는 음압을 임의로 대폭 낮추어 청취할 수밖에 없다. 이러한 저음압 청취 환경은 인간의 청각 특성인 등청감 곡선에 의해 저역대 및 고역대 주파수 인지력의 심각한 저하를 유발한다. 나아가 공간 및 비용의 제약으로 천장 스피커 설치가 어렵고, 거실의 특성상 사이드 스피커를 거치하는데 공간

1.2 문제 제기

문제는 7.1.4 채널의 3차원 정보를 5.1 채널의 2차원 평면으로 강제 병합(Fold-down)하는 과정에서 발생하는 공간 정보의 왜곡과 인지적 혼란이다. 제작자가 의도한 수직적 객체(Object) 이동과 상부 공간감은 다운믹스 알고리즘에 의해 기존의 수평 채널(Ear-level)로 다시 할당되는데, 이때 물리적인 하이트(Height) 스피커를 대체하기 위해 수평 채널 간의 음압과 위상 차이를 조작하여 가상의 음상인 팬텀 이미지(Phantom Image)를 생성하게 된다.

이는 곧 청취자가 허공에 소리가 존재한다고 느끼도록 뇌의 청각 피질을 속이는 일종의 착청(Auditory Illusion) 현상에 기인한다 [1]. 즉, 다운믹스 환경에서의 입체감은 물리적 실체가 아니라 왜

곡해서 듣는 뇌의 복잡한 연산 결과물인 셈이다.

그러나 아파트 거실과 같은 비대칭적 청취 공간에서는 이러한 인지적 환상이 쉽게 무너진다. 불규칙한 룸 모드(Room Mode)와 초기 반사음이 뇌가 팬텀 이미지를 정확하게 결상(Imaging)하는 과정을 교란하기 때문이다. 결과적으로 객체의 정위감이 모호해질 뿐만 아니라, 수평 채널로 과도한 정보량이 밀집되어 심각한 주파수 마스킹(Masking) 현상을 유발한다. 이처럼 5.1 다운믹스가 청취자의 인지적 피로도와 공간감 왜곡에 미치는 복합적인 부작용이 많음에도 불구하고, 실제 주거 환경을 반영한 실증적 분석은 미비한 실정이다.

1.3 연구의 목적

본 연구는 공동 주택이라는 특수한 공간적, 사회적 제약 속에서 몰입형 오디오가 청취자에게 전달되는 과정의 한계를 정량적, 정성적으로 분석하는 것을 목적으로 한다. 이를 위한 구체적인 연구 목표는 다음과 같다.

1. 주거 환경의 현실적 음압 한계 규명: 청취자가 이웃에게 피해를 주지 않는 선에서 자율적으로 설정한 하우스 레퍼런스 레벨(dBA)과 층간 소음의 원인이 되는 저주파 물리 에너지(dBC) 간의 격차를 측정하여 아파트 환경의 구조적 한계를 수치화한다.
2. 채널 축소에 따른 이머시브 체감도 분석: 전문 스튜디오 환경의 7.1.4 채널 재생과 아파트 환경을 가정한 5.1 채널 다운믹스 재생 시, 청취자가 체감하는 객체 정위감, 포위감, 밸런스의 차이를 분석한다. 이는 향후 가정용 재생 환경을 고려한 다운믹스 호환성 확보 전략을 수립하는 데 목적이 있다.

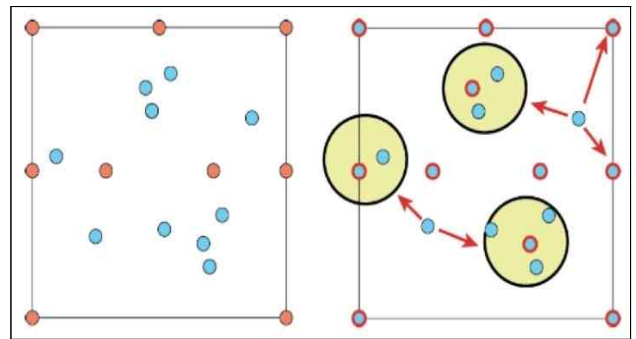
1.4 이론적 배경

전통적인 채널 기반 오디오와 달리, Dolby Atmos로 대표되는 객체 기반 오디오는 개별 사운드 소스를 3차원 공간 내의 좌푯값(Metadata)을 가진 객체로 취급하여 재생 환경에 맞게 실시간 렌더링한다. 그러나 7.1.4 채널로 디자인된 콘텐츠가 물리적 천장(Height) 스피커가 부재한 5.1 채널 환경으로 다운믹스될 경우, Height 채널 정보가 수평 채널로 병합된다. 이는 창작자가 의도한 수직적 이동 궤적의 상실과 함께, 한정된 스피커로 여러 객체의 소리가 물리면서 발생하는 주파수 마스킹 및 공간적 분리도 훼손을 일으킨다. [그림1]처럼 Spatial Coding에서는 Bed도 Object로 변환되어 클러스터에 포함된다. 이러한 이유로 서라운드 또는 Height 채널에 배치된 소리는 스테레오에서 최대 4.5dB 까지 감소할 수 있다[2].

인간의 청각은 주파수와 음압(SPL)에 따라 감도가 비선형적으로 나타나며, 이는 플레처-먼슨의 등청감 곡선으로 설명된다. 이 원리에 따르면, 음압이 낮아질수록 저역대와 고역대에 대한 인간의 청각 민감도는 급격히 하락한다. 82dB SPL의 스튜디오 레퍼런

스 환경에서 완벽한 밸런스로 믹싱 된 곡이라 할지라도, 공동 주택의 층간 소음 방지를 위해 볼륨을 낮춘 저음압 환경(약 65dBA 내외)에서는 베이스와 서브우퍼 대역의 에너지가 실제 출력보다 훨씬 작게 체감된다. 이 과정에서 발생하는 물리적 압력(dBC)과 체감 볼륨(dBA)의 괴리는 몰입형 오디오의 역동성을 저해하는 주요 요인이 된다.

또한, 전통적인 음향기기 평가 연구는 스피커 채널, 공간의 체적, 재생 음압 등 단일 변인만을 통제하고 비교하는 실험실 환경(Laboratory setting)을 지향해 왔다. 그러나 이러한 완벽히 통제된 무향실 및 스튜디오 연구 결과는, 여러 제약 조건이 동시에 발적으로 겹쳐진 일반 가정환경에서 발생하는 복합적인 음질 열화 현상을 설명하지 못한다. 따라서 실제 소비자가 몰입형 오디오를 경험하는 생태학적 타당성을 확보하기 위해서는 개별 변인의 통제를 넘어 현실의 제약 조건들을 하나의 복합 환경 패키지(Complex Environment Package)로 구성하여 교차 분석하는 접근이 요구된다.



[그림1] Spatial Coding 없음(좌), 있음(우). 출처: Dolby

2. 본론

2.1 연구 방법 및 실험 설계

본 연구는 기기의 물리적 성능 평가가 아닌 주거 공간이라는 복합적 제약 환경에서의 이머시브 오디오 소비 경험 규명에 목적이 있다. 따라서 전통적인 단일 변인 통제 방식을 탈피하여, 스피커 채널과 음압이라는 변인을 묶은 환경 패키지(Environment Package)를 기준으로 실험을 설계하였다.

2.1.1 주요 분석 변인

가) 독립 변인(환경 패키지) - 비교군: 7.1.4 채널 시스템을 기반으로 한 이상적인 다채널 재생 환경. Dolby사의 인증을 받은 한국영상대학교 Dolby Atmos Studio.

실험군: 5.1 채널 다운믹스 시스템과 청취자가 층간 소음을 고려해 스스로 설정한 저음압(하우스 레퍼런스 레벨)이 결합된 아파

트 거실 환경.

나) 종속 변인(주관적 청취 경험) - 특정 객체의 정위감(Localization), 공간의 포위감(Envelopment), 악기 간의 음색 및 명료도(Timbre & Clarity), 사운드 스테이지의 전후 거리감(Depth)을 5점 척도로 측정하였다.

다) 조절 변인(음원 특성) - 환경적 제약이 콘텐츠에 미치는 영향을 확인하기 위해 믹싱 밀도(Density)가 높고 객체 이동이 잦은 팝/록 장르와, 대체적으로 Bed 의존도가 높은 어쿠스틱/클래식 장르를 교차 배치하였다.

2.1.2 실험 환경 및 시스템 캘리브레이션

본 실험은 실제 홈 청취 환경을 충실히 모사하기 위해 연구자의 아파트 거실에서 진행되었다(112m²). 생태학적 타당성을 확보하면서도 데이터의 객관성을 유지하기 위해 C-weighting을 기준으로 각 스피커의 주파수 응답 및 출력 레벨이 동등하게 재생되도록 핑크 노이즈를 사용하여 시스템 캘리브레이션을 엄격히 수행하였다.

2.1.3 청취자 주도적 음압 측정(하우스 레퍼런스 레벨)

피험자(N=7)를 대상으로 본격적인 청취 평가 전, 아파트 거실에서 중간 소음을 유발하지 않는 최대 물입 볼륨을 스스로 설정하도록 지시하였다. 피험자별 자율 볼륨 설정에 따른 음압 편차를 정확히 비교하기 위해, 곡의 다이내믹에 따른 오차를 통제하고 전 피험자에게 같은 타임스탬프(Whiplash 특정 구간 60초)를 자극원(Reference Stimulus)으로 제공하였다. 이때 GVDA GD151B 음압계를 사용하여 인간의 청감 특성을 반영하는 A-weighting(dBA)과 물리적 저주파 에너지를 포함하는 C-weighting(dBC, Slow Response)을 동시 측정하여 그 격차를 기록하였다.

연구 참여자는 청각 훈련을 2년 이상 이수한 음향전공 대학생 6명(남 3, 여 3)과 성인 여성 1명으로 구성하였다. 이들은 미세한 주파수 변화와 위상 왜곡을 인지할 수 있는 훈련된 청취자(Golden Ears) 집단이다.

평가는 1명씩 진행하였으며, 바이어스(Bias) 방지를 위해 스테레오와 바이노럴 버전을 통한 사전 청취 과정을 거친 후, 각 환경에서 오픈 테스트 형태로 진행되었다.

[표1] 피험자 명단

	이름	성별	나이
1	박00	여	21
2	이00	남	25
3	남00	남	25
4	차00	여	21
5	장00	여	23
6	정00	남	21
7	박00	여	38

2.1.4 주관적 청취 평가 및 설문 절차

음악 장르와 믹싱 특성이 다른 5곡(aespa 'Whiplash', Roxy Music 'Avalon', Stewart Copeland 'White Throated Sparrow', Beethoven 'Piano Concerto No.2', RATM 'Bulls On Parade')을 실험 음원으로 선정하였다. 피험자는 본인이 직접 설정한 저음압 환경에서 각 곡을 7.1.4 채널과 5.1 채널 다운믹스로 번갈아 청취하였다. 이후 5점 척도로 구성된 설문지를 통해 상단 객체의 이동 궤적, 포위감, 명료도 등에 대한 블라인드 평가를 독립적으로 수행하였다.

2.2 연구 결과 및 논의

본 실험에서는 피험자(N=7)가 아파트 거실이라는 주거 환경을 가정하고 직접 설정한 적정 볼륨에서, 층간 소음의 주요 원인이 되는 저주파 물리 에너지(dBC)와 청취자가 인지하는 체감 볼륨(dBA) 간의 격차를 측정하였다. 측정 결과는 [표 2]와 같다.

[표2] 피험자의 dBA와 dBC 비교

피험자	dBA	dBC	격차 (dBC - dBA)
1	62.9	75.2	+ 12.3
2	65.4	79.7	+ 14.3
3	61.4	74.5	+ 13.1
4	60.9	69.8	+ 8.9
5	62.4	71.6	+ 9.2
6	62.4	74.2	+ 11.8
7	60.8	72.2	+ 11.4
평균	62.3	73.9	+ 11.6

측정 결과, 피험자들이 거실 환경에서 시끄럽지 않다고 느끼며 수용한 볼륨의 평균은 62.3 dBA로 나타났다. 이는 스튜디오 레퍼런스 레벨(음악 82dB SPL)에 비해 약 19.7dB 이상 현저히 낮아진 수치로, 일반적인 공동 주택 거주자가 이머시브 오디오를 소비하는 현실적인 한계를 잘 보여준다.

가장 주목할 만한 점은 dBA와 dBC 간의 격차다. 청취자가 수용한 체감 볼륨(평균 62.3 dBA) 이면에는 평균 73.9 dBC라는 상당히 높은 물리적 저주파 에너지가 존재하고 있었으며, 두 수치 간의 평균 격차는 11.6dB에 달했다. 특히 피험자 2번의 경우, 65.4 dBA의 볼륨 설정에서 79.7 dBC가 측정되어 무려 14.3 dB의 압력 차이가 발생했다.

이는 청취자가 거실에서 듣기에 무리가 없다고 느끼는 볼륨조차도 실제로는 서브우퍼를 포함한 재생 시스템이 방음벽이나 바닥 구조물을 진동시킬 수 있는 거대한 물리적 에너지를 방출하고 있음을 의미한다. 다운믹스 환경에서 낮은 볼륨으로 청취할 때 발생하는 이 같은 음압 격차는 이후 진행된 5.1 시스템의 LFE 만족도 및 포위감 설문 결과와도 유의미한 상관관계를 가진다.

두 개의 설문지 응답 데이터(각각 7명의 피험자 응답)를 추출하여 5.1 채널 환경과 7.1.4 채널 환경 간의 항목별 평균 점수 비교

통계를 산출했다.
5점 척도로 진행된 설문 통계 분석 결과는 [표3] 와 같다.

[표3] 피험자의 dBA와 dBC 비교

문항 (평가항목)	5.1 채널 (평균)	7.1.4 채널 (평균)	격차 (7.1.4 - 5.1)
Whiplash: 신스 사운드의 상단 공간 이동 궤적이 뚜렷한가?	2.43	4.00	+ 1.57
Avalon: 색소폰의 정위감이 잘 느껴지는가?	3.00	4.57	+ 1.57
Avalon: 리버브/이펙트가 공간 전체를 감싸는 포위감이 충분한가?	3.29	4.57	+ 1.28
Bulls On Parade: Rock Sound의 밀도가 더 넓게 확장되어 들리는가?	2.71	3.86	+ 1.15
Piano Concerto No.2: 피아노와 오케스트라 사이의 전후 거리감이 명확한가?	3.57	4.43	+ 0.86
Whiplash: 강한 킥/베이스가 보컬을 방해하지 않고 분리되는가?	3.29	4.14	+ 0.85
Piano Concerto No.2: 콘서트홀의 반사음과 잔향이 사실적으로 구현되는가?	3.43	4.29	+ 0.86
White Throated Sparrow: 다양한 타악기들이 3차원 공간 내에서 겹치지 않는가?	3.86	4.57	+ 0.71
White Throated Sparrow: 타악기 특유의 타격음이 공간의 울림과 조화로움인가?	3.57	4.29	+ 0.71
Bulls On Parade: 기타 스크래칭 사운드가 객체로서 역동적으로 움직이는가?	2.71	3.43	+ 0.71

가장 큰 점수 차이(+1.57점)를 보인 항목은 Whiplash의 상단 신스 사운드 궤적과 Avalon의 색소폰 정위감이다.

5.1 채널 시스템은 수평적인 공간 배치에는 어느 정도 대응하지만, 물리적인 Height 스피커의 부재로 인해 다운믹스 시 상단 객체의 뚜렷한 이동 궤적을 표현하는 데 매우 제한된다는 것이다. (5.1채널에서는 상단 궤적 점수가 2.43점으로 나타남). 결과적으로 Height 채널의 부재로 인한 객체의 이동 및 정위감이 많이 상실되게 나타났다.

반면에, White Throated의 타악기 겹침 방지(+0.71)나 조화로움(+0.71), Beethoven의 거리감(+0.86)과 잔향감(+0.86)은 5곡 중 가장 점수 격차가 적었다. White Throated의 5.1 점수는 3.86점, Beethoven은 3.57점으로, 상단 스피커가 없고 볼륨이 낮은 조건에서도 보통(3점)을 훌쩍 뛰어넘는 상당히 긍정적인 평가를 받았다. 이는 [3]The Effect of Sound Level on Perception of Reproduced Soundscapes에서도 언급된 현장

음량보다 9.5dB 낮은 수준에서 재현된 사운드 스케이프는 실제 현장 경험과 유사한 주관적 인상을 제공했다는 내용과 유사한 결과 값이다.

또한, 이러한 현상은 곡의 믹싱 특성에 기인한 것으로 해석할 수 있다. 타악기나 클래식 음원은 다이내믹 레인지가 넓고 악기 간의 주파수 마스킹이 적어서 낮은 음압과 제한된 채널 수 환경에서도 각 악기의 트랜지언트와 정위감이 비교적 온전히 보존된다. 그리고 클래식 음원의 이머시브 효과는 뚜렷한 객체의 궤적보다는 홀의 잔향과 수평적 베드(Bed) 확장에 의존하므로 5.1 시스템의 수평 서라운드만으로도 청취자에게 충분한 포위감을 제공할 수 있었던 것으로 생각한다. 결론적으로, 주거 환경의 한계(저음압, 다운믹스) 속에서 이머시브 오디오의 실효성을 확보하기 위해서는 곡의 밀도를 낮추고 마스킹을 최소화하는 믹싱 전략이 유리함을 시사한다.

실험 후 피험자와의 인터뷰에서는 대다수가 음압에 대한 문제를 가장 많이 언급하였다.

3. 결론

본 연구는 3차원 몰입형 오디오(Immersive Audio)가 전문 스튜디오를 벗어나 공동 주택(아파트)이라는 일상적 주거 공간에서 소비될 때 발생하는 물리적, 음향학적 한계를 규명하고, 재생 시스템(7.1.4 채널 vs 5.1 채널) 및 음원 특성에 따른 청취자의 체감도 변화를 교차 분석하였다. 본 연구를 통해 도출된 주요 결론은 다음과 같다.

첫째, 주거 환경의 층간 소음 딜레마와 음압의 한계 수치화. 청취자들이 아파트 거실 환경에서 자율적으로 설정한 하우스 레퍼런스 레벨의 평균 체감 음압은 62.3 dBA로 측정되었다. 그러나 이면의 저주파 물리 에너지는 평균 73.9 dBC에 달해, 약 11.6dB의 상당한 압력 차이가 존재함이 증명되었다. 이는 청취자가 체감상 듣기 편안하다고 느끼는 수준의 볼륨에서도, 재생 시스템의 서브우퍼와 저역대 에너지는 주거 구조물에 층간 소음을 유발할 수 있는 막대한 물리적 에너지를 방출하고 있음을 시사한다.

둘째, Height 채널과 사이드 채널 부재에 따른 공간감 축소. 물리적 상단 스피커와 사이드 채널이 배제된 5.1 채널 다운믹스 환경에서는 스튜디오 표준인 7.1.4 채널 대비 객체의 수직 이동 궤적과 공간 전체를 감싸는 포위감(Envelopment) 평가 점수가 최대 1.57점(5점 만점 기준) 하락하였다. 이는 5.1 채널 기반의 다운믹스나 보급형 홈시어터 시스템이 수평적 방향성을 제공할 수는 있으나, 창작자가 의도한 3차원 객체의 역동성과 입체감을

온전히 재현하기에는 명확한 한계가 있음을 확인시켜 준다.

셋째, 음원 장르 및 믹스 밀도(Density)에 따른 다운믹스 호환성 저음압 및 채널 축소 환경이 몰입형 오디오 체감에 미치는 악영향은 곡의 믹싱 특성에 따라 큰 편차를 보였다. 팝, 록과 같이 주파수 밀도가 높고 상단 객체 의존도가 높은 곡은 낮은 볼륨에서 심각한 마스킹 효과가 발생하여 입체감이 크게 훼손되었다. 반면, 어쿠스틱 타악기나 클래식 음원은 악기 간 여백이 충분하고 트랜지언트가 명확하여, 5.1 채널의 낮은 음압 환경에서도 각 악기의 정위감과 포위감을 비교적 우수하게 유지하였다.

결론적으로, 아파트 거실 환경에서 몰입형 오디오의 효과를 누리기 위해서는 단순히 채널 수를 늘리는 물리적 접근을 넘어, 주거 환경의 저음압 특성과 다운믹스 알고리즘을 고려한 소프트웨어적 접근 및 믹싱 전략이 필수적이다.

4. 제언 및 향후 연구 방향

본 연구의 결과를 바탕으로 홈 엔터테인먼트 오디오 산업의 발전과 후속 연구를 위해 다음과 같이 제언한다.

사용자가 볼륨을 낮출 때 단순히 전체 게인을 줄이는 것을 넘어 등청감 곡선을 반영하여 공간감의 뼈대가 되는 중고역대 객체의 명료도는 유지하되, 구조물 진동을 유발하는 초저역대(Sub-bass) 에너지만 선택적으로 감쇠시키는 진보된 야간 모드(Night Mode) 알고리즘의 개발이 필요하다.

향후 연구 방향으로 본 연구는 실제 물리적 스피커가 배치된 5.1 채널 환경을 변인으로 설정하였으나, 최근 일반 가정에서는 사운드바의 상향 방사형(Up-firing) 스피커를 통한 가상 돌비 애트모스 재생이나 헤드폰을 이용한 바이노럴 렌더링 방식이 널리 사용되고 있다. 따라서 후속 연구에서는 본 실험의 기준 음압(약 60~65dBA)을 동일하게 적용하여, 가상 서라운드 시스템과 물리적 5.1 채널 간의 몰입감 차이를 비교 분석한다면 몰입형 오디오 소비 환경에 대한 더욱 종합적인 통찰을 얻을 수 있을 것이다.

참고문헌

- [1] Diana Deutsch, "Musical Illusions and Phantom Words", EIDOS, 박정미, 박종화 옮김, 2023년.
- [2] https://docs.hybrik.com/tutorials/dolby_atmos/dolby.
- [3] Anugrah Sabdono Sudarsono, Yiu W. Lam, William J. Davies "The effect of sound level on perception of reproduced soundscapes", Applied Acoustics 저널, Vol. 110, 9월, 2016년.