

# 모바일 3D 사운드: 바이노럴 오디오 기술 동향

□ 오현오\*, 이태규\*\*, 전세운\*\*, 윤대희\*\*, 박영철\*\*, 서정일\*\*\*, 이용주\*\*\* / \*일러스표준기술연구소, \*\*연세대학교, \*\*\*한국전자통신연구원

## I. 서론

바이노럴(Binaural)의 문자 그대로의 의미는 “두 개의 귀를 가진(having or relating to two ears)”이다. 특히 오디오 신호처리 영역에서 바이노럴은 사람이 두 개의 귀를 가지고 있기 때문에 발생하는 여러 현상들을 의미하며, 또한 그 현상을 이용하여 두 개의 트랜스듀서를 이용하여 3차원의 입체 음향을 제공하는 방법을 의미한다. 이런 의미에서 모바일 3D 사운드를 위한 기술이라 할 수 있다.

스마트폰의 대중화와 초고속 무선 통신의 영향으로 고품질 멀티미디어를 모바일로 소비하는 사례가 보편화되고 있다. Full HD 혹은 최근의 UHD(Ultra High Definition) TV와 같은 고품질 영상 신호에 대응되는 오디오 콘텐츠는 22.2채널까지 확장 가능한 Immersive 3D 오디오로 대표

된다. 모바일 환경에서 이와 같은 고품질 오디오 콘텐츠를 소비하기 위해서는 스테레오 헤드셋만으로 22.2채널 스피커가 제공하는 3차원 음향을 구현할 수 있는 방법이 요구되며, 바이노럴 렌더링은 이를 실현하기 위한 가장 이상적인 기술로 평가되고 있다.

사람은 두 개의 귀(청각입력센서)를 가지고도 소리에 대한 전방위의 3차원 방향성을 인식할 수 있다. 이는 소리의 방향에 따라 달라지는 음색의 변화를 오래된 학습을 통해 구별하게 된 능력에 근거한 것으로서 방향에 따른 귀에서의 주파수 응답인 머리전달함수(Head Related Transfer Function; HRTF)의 차이를 인식하는 것으로 볼 수 있다. 그런데, 사람마다 서로 다른 머리 크기, 얼굴형 및 귀의 모양 등으로 인해 HRTF는 각 개인에 따라 그 특성이 매우 다르다. 따라서, 하나의 고정된 HRTF로 모든 사람을 만족시키기는 어렵다. 한편, 입체

음향 효과의 극대화를 위해서는 소리의 방향뿐 아니라 원음원이 재생되던 재생공간(Room)의 특성이 반영되어야 하는데, 이를 위한 주파수 응답을 BRTF(Binaural Room Transfer Function)라고 한다. 여러 방향에 대한 BRTF를 충실하게 Rendering하기 위해서는 많은 저장용량과 함께 높은 연산량이 요구된다. 따라서, 이전의 연구들은 BRTF 혹은 HRTF 데이터베이스를 효과적으로 저장하고, 효과적으로 연산처리하는 방법을 많이 다루었다.

스마트폰의 풍부한 저장공간과 높은 연산능력은 복잡한 렌더링 기술을 휴대 기기에서 쉽게 구현 가능하도록 만들고 있으며, 클라우드 서비스 등 다양한 환경의 변화는 바이노럴 렌더링의 대중화에 새로운 전기를 마련하고 있다. 즉, 보다 개인화되고 보다 정교한 BRTF의 사용이 가능해져서 고품질의 바이노럴이 실현될 수 있다. 더불어 스마트폰 사용자들이 고가의 헤드폰 및 이어폰을 사용하는 경우가 증가되면서 이와 같은 고품질 바이노럴에 대한 필요성도 더 높아졌다.

이와 같은 환경변화에 발 맞추어 MPEG(Moving Picture Expert Group)에서는 UHD급 고품질 콘텐츠 재생 환경에서의 새로운 오디오 코덱인 MPEG-H 3D Audio 표준화 과정에서 바이노럴 렌더링을 표준화의 한 부분으로 포함하고 현재 표준 기술 제정을 앞두고 있다. MPEG-H는 차세대 방송용 코덱 기술로 높은 관심을 가진 차세대 코덱 표준이다.

본 논문에서는 바이노럴 오디오 기술에 대한 개요와 함께 여러 이슈들을 설명하고, MPEG-H 표준을 중심으로 한 최근 시장 동향을 살펴본다.

## II. 바이노럴 오디오 기술 개요

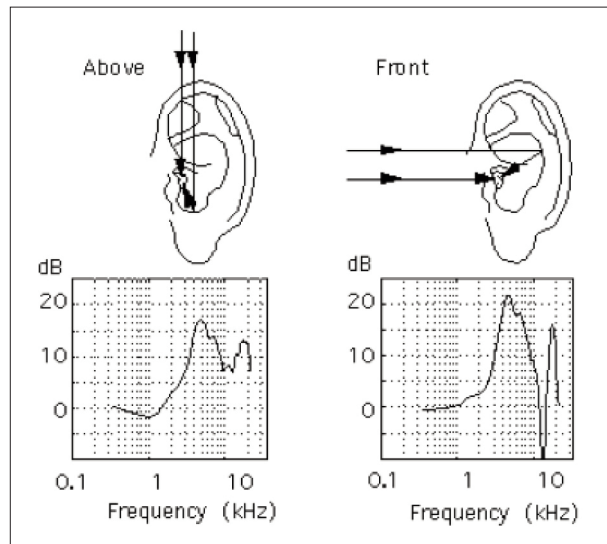
### 1. 바이노럴 3D 사운드

3D 사운드는 3D 영상과 대응되는 기술로 오디오 콘텐츠가 가지고 있는 음장(Sound field)을 재현하는 기술로, 더 구체적으로는 특정 음상(sound image)의 방향감과 거리감, 공간감 등을 사용자로 하여금 그 곳에 있는 것처럼 재현하는 것을 말한다. 따라서 3D 사운드를 통하여 음원이 가지고 있는 현장감을 느낄 수 있으며, 가상 현실에서 필수적인 기술이라고 할 수 있다. 특히 본 논문에서 설명하는 모바일 3D 사운드란 헤드폰의 경우와 같이 두 개의 Transducer(스피커)만을 이용하여 3D 사운드를 제공하는 기술을 말한다. 헤드폰을 이용한 3D 사운드를 적극적으로 이용한 최근의 예로는 2013년 초에 방영된 질레트 광고이다[1]. 해당 광고에서는 3D 사운드가 적용되어 청취자로 하여금 소리를 통하여 실제로 면도를 하는 느낌을 주었다. 특히 유튜브(Youtube)에서 바이노럴 레코딩 또는 다른 오디오 신호처리를 이용한 3D 사운드가 많이 업로드 되고 있다.

이러한 3D 사운드를 지각하기 위한 사람의 인지 방법은 소리에 담겨있는 다양한 단서들에 있다. 이 단서는 크게 두가지로 나눌 수 있는데 첫째는 양쪽 귀에 들어오는 신호의 차이를 인지하는 바이노럴 큐이다. 바이노럴 큐는 주로 양이(兩耳) 신호의 레벨 차이와 시간 차이, correlation 등을 이용하여 공간감을 지각하는데 사용된다. 이를 각각 양이 레벨차(Interaural level difference, ILD), 양이 시간차(Interaural time difference, ITD), 양이 코히어런스(Interaural coherence)라고 정의한다[2]. 만일 오른쪽에서 종소리가 발생한다고 하

면 오른쪽 귀에는 바로 소리가 전달 되는 반면 왼쪽 귀에는 양 쪽 귀 사이사이의 거리에 따른 시간차가 발생한다. 또한 왼쪽 귀에는 머리, 몸통, 외이를 회절하여 들어가야 하기 때문에 감쇄가 발생한다. 이를 Head shadowing이라고 한다. 따라서 사람은 주로 이러한 시간차(ITD)와 레벨차(ILD)를 이용하여 방향을 인지한다. 둘째는 한 쪽 귀로 인지하는 신호 자체의 특성을 이용하여 지각하는 단서이다. 주로 스펙트럼의 특징이 주가 되며 valley나 notch와 같은 형태로 지각하게 된다. 특히 monoral cue의 경우 고도를 인지하는데 사용된다. 예를 들면 정면에서 발생한 소리와 그보다 높은 위치에서 발생한 소리는 외이(귓바퀴)의 공명 주파수가 달라지게 되고 그에 따른 스펙트럼 상의 peak나 notch가 달라지는 특징을 이용하여 고도를 인지하는 것이다(그림 1). 따라서 3D 사운드는 이러한 다양한 단서들이 잘 담기도록 만들어진 소리라고 할 수 있다.

모바일 3D 사운드를 만드는 방법은 크게 바이노럴 레코딩(Binaural recording)과 바이노럴 렌더링(Binaural rendering)이라고 할 수 있다. 바이노럴 레코딩은 해당 공간 자체가 가지고 있는 음장을 취득하기 위해 사용되는 기술이며, 바이노럴 렌더링은 콘텐츠 제작자가 특정 음원을 공간상에 정위시키는 방식을 의미한다. 바이노럴 레코딩은 일반적으로 <그림 2>와 같이 더미 헤드(Dummy head)의 귀에 소형 마이크를 삽입한 후 녹음함으로써 간단하게 이루어 질 수 있다. 더미 헤드를 이용하여 녹음하는 경우 음원과 양이 사이의 전달함수를 통하여 녹음이 된 것이기 때문에 사람의 몸과 더미 헤드가 동일한 구조를 가지고 있다면 동일한 공간감을 지각할 수 있다. 그러나 사람의 몸은 개인에 따라 다른 구조를 가지고 있으며 또한 더미 헤드와 완전히 같을 수 없기 때문에 실제로 바이노럴 레코딩된 음원을 청음할 경우 공간감에서 차이가 존재한다. 따라서 이를 극복하기 위하여 때때로 사람의 귀에



<그림 1> 바이노럴 큐 (음원의 고도를 인지하는 모노럴 큐)

소형 마이크를 장착하여 녹음하기도 한다. 이 경우도 녹음을 한 본인 이외에는 완전한 3D 사운드를 즐길 수 없다. 더불어 바이노럴 레코딩을 방위각, 고도에 따라 정해진 많은 위치에서의 전달함수를 미리 구해두고 (이를 바이노럴 방 충격 응답, Binaural Room Impulse Response, BRIR라고 한다), 특정 음원의 위치를 알고 있을 때, 해당 음원에 BRIR을 컨볼루션하여 3D 사운드를 생성하는 방법을 바이노럴 렌더링이라고 한다. 바이노럴 렌더링 역시 BRIR 측정에 사용된 더미 헤드와 청취자의 동일성의 한계만큼 오차가 발생한다.



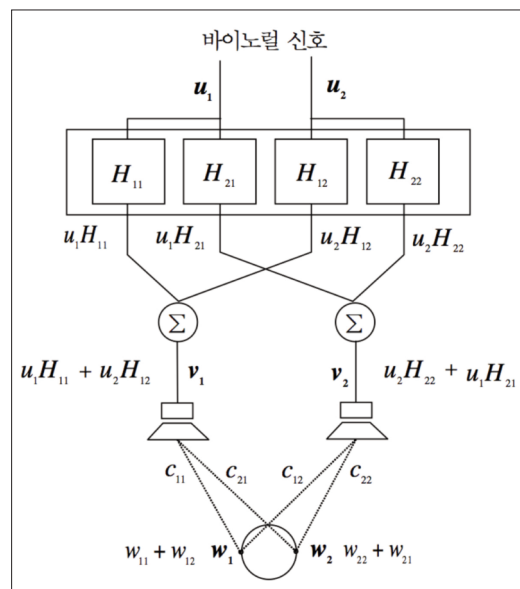
〈그림 2〉 더미 헤드 (Dummy Head)

## 2. Transaural

사운드의 재생 환경 측면에서 오디오 시스템은 binaural system과 transaural system으로 구분할 수 있다[3]. 청자의 귀에서 들리는 음원, 즉, 헤

드폰을 사용하는 경우와 같은 재생 환경을 가정한 ‘binaural’과 달리 스피커를 이용한 음향 시스템을 가정한 재생 환경을 ‘transaural’이라고 부른다. Transaural의 개념은 Schroeder와 Atal에 의해서 초창기 시스템이 제안되어 Cooper와 Bauck에 의해서 보다 개선되었다[4][5]. Transaural 시스템에서는 스피커로부터 좌우 귀에 도달하는 음향 경로(acoustic path)의 영향으로 크로스톡(crosstalk)이 발생하게 되는데, 이를 고려하지 않을 경우, 음원의 방향감과 공간감이 왜곡될 수 있다. 때문에 〈그림 3〉과 같이 크로스톡 음향 경로에 의한 영향을 제거하고, 특정 위치에 있는 청자에게 임의의 음향 효과를 재현하고자 하는 ‘crosstalk cancellation’ 기술이 필요하다[6][7].

〈그림 3〉은 2채널 스피커 환경에서의 crosstalk cancellation 과정을 보여주고 있다. 입력으로 바이노럴 신호  $u_1, u_2$ 가 들어올 때, 이를 바로 스테레오



〈그림 3〉

스피커를 이용하는 transaural system에서 재생하는 경우, 헤드폰을 이용한 경우와 달리, 반대편 스피커의 출력까지 귀에 전달되게 된다. 이 경우, 서로 같거나 혹은 다른 음향 경로를 통과한 두 채널 신호가 더해진 신호가 귀에서 들리게 되는데, 이 신호는 입력 신호  $u_1, u_2$ 와는 다른 특성을 가질 수 있다. 이러한 문제를 해결하기 위해서, 스피커에서 귀까지의 경로  $C(z)$ 를 측정 후, 이를 보상해 주는 역필터(crosstalk canceller)  $H(z)$ 를 설계한다. 이러한 역필터를 통과한 최종 스피커 출력 신호  $v_1, v_2$ 가 스피커를 통해 재생된다. 역필터가 제대로 설계된 경우,  $v_1, v_2$ 는 스피커에서 청취자까지 경로  $C(z)$ 의 특성을 이미 보상한 신호이기 때문에, 이론적으로 귀에서는 입력신호  $u_1, u_2$ 와 동일한  $w_1, w_2$  신호를 들을 수 있다. 즉 원격의 스피커의 재생 신호가 마치 헤드폰을 착용하고 있는 것과 같은 소리를 제공할 수 있다. 그러나 Transaural의 경우도 청자의 위치에 대한 제약과 청자의 머리 움직임에 따른 영향 등 극복해야 할 기술적 과제를 아직 많이 가지고 있으며, 역필터 계산 역시 완전한 해를 갖지 못하는 문제를 가지고 있다. 시중의 많은 상용 Virtual Surround 기술들의 경우 이와 같은 Transaural 기술을 부분적으로 접목하여 구현되는 경우가 많다.

### III. 바이노럴 상용화 이슈 및 최신 기술 동향

이론적으로 완벽한 3D 사운드를 제공할 수 있는 기술임에도 불구하고 바이노럴의 상용화에는 아직까지 여러가지 난제들이 존재한다. 본 장에서는 바이노럴 상용화를 위한 주요 이슈와 함께 여러 접근

방법에 대해 소개한다.

#### 1. Individualization (개인화)

바이노럴 오디오 기술에서 가장 중요한 이슈 가운데 하나는 individual HRTF의 측정이다. 스피커로부터 귀까지의 경로에 대한 측정이 정확히 이루어지지 않을 경우 완전한 바이노럴 렌더링은 불가능하다. 특히, HRTF(Head Related Transfer Function)는 청자의 머리 모양과 크기, 체형, 그리고 귀의 모양 등 사람마다 다른 특성들에 의해 양이 간의 주파수 특성이 달라질 수 있기 때문에, 같은 음원과 음원 환경에 대해서도 사람마다 차이가 날 수 있다. 때문에 individualized HRTF의 정확한 측정은 바이노럴 오디오 기술의 성능을 결정짓는 가장 중요한 요소이다[8]. 시설이 잘 갖춰진 제한된 공간에서 측정하고자 하는 대상 청자의 귀에 마이크를 꼽고 고정된 자세에서 장시간 동안 측정이 이뤄져야 하는 HRTF의 측정 과정은 그 자체가 매우 어려울 뿐 아니라 세상의 모든 사용자에게 대해 이와 같은 DB를 구축한다는 것은 현실적으로 불가능하다. 이에, 일반화된 HRTF를 기본적으로 이용하고 개별적인 HRTF로 근사화하는 ‘individualization’ 기술에 대한 연구도 높은 관심을 보이고 있다.

이를 위해 다수의 측정 결과와 그에 대응하는 신체측 정보 사이의 연관성을 이용하여 역으로 신체측 정보에 대응하는 HRTF의 근사치를 예측하는 방법이 제안되었다[9][10]. 이 방법은 보유하고 있는 데이터 이외의 새로운 청자에 대해서 HRTF를 따로 측정하지 않더라도, 귀나 체형 등의 몇가지 대표값의 측정을 통해서 HRTF를 근사화할 수 있는 장점이 있다. 하지만 이 방법은 신체측 정보의 측정을 위한 부가적인 장비를 필요로 하고, HRTF

Database의 실험군의 특성에 영향을 받기 쉬운 단점을 가지고 있다. 예를 들어 서양인의 데이터를 이용하여 상대적으로 체형이 다른 동양인의 HRTF를 예측하는 경우, 그 오차가 커질 수 있는 것이다. 또한 데이터의 측정과 저장 및 처리에 들어가는 비용이 적지 않다는 문제점이 있다.

또다른 individualization 방법으로는 더미 헤드에서 측정한 HRTF나 평균 두상과 체형을 고려한 HRTF 평균치로부터 실제 사용자의 특정 주파수 특성을 반영하여 바이노럴 오디오 시스템 구현에 적용하기도 한다. 이를 위해서는 먼저 귀의 모양이나 머리 크기, 신체 정보와 HRTF의 주파수 특성과의 관계에 대한 연구가 선행되어야 한다. HRTF에 대한 많은 연구들에서 이에 대한 단서가 밝혀졌고, 많은 양의 HRTF 데이터를 가지고 있지 않더라도 개별화된 HRTF를 근사화할 수 있는 연구가 보고되었다[11][12][13].

## 2. 연산량 이슈

서론에서 언급한 바와 같이 최근의 스마트폰과 클라우드 등 모바일에서의 연산 능력과 저장 공간이 크게 향상되었지만 그럼에도 불구하고 바이노럴 오디오의 품질을 높이기 위한 노력은 여전히 연산량을 극복해야하는 한계를 남긴다. MPEG-H를 예로 할 경우 총 44개의 서로 다른 필터(22입력 채널을 좌/우 출력으로 대응하는 각각의 필터)가 존재하며, 48kHz 표본화 주파수를 사용하는 경우 2초 길이의 BRIR은 각 필터 계수를 96,000개 갖는 셈이다. 1 cycle에 MAC(Multiply & Accumulate)을 연산하는 DSP를 사용한다고 하더라도  $44 \times 96000 \times 48000 = 202,756 \text{MIPS}$  (약 203 GHz)라는 연산량을 요구하는 셈이다. MPEG-H의 바이노럴

기술 표준화가 다른 것보다 연산량의 최적화에 중점을 둔 이유이기도 하다.

한편, 이와 같이 높은 연산량을 우회하는 한 방법으로 프리렌더링이 있다. 즉, 높은 연산량을 요구하는 바이노럴 렌더링을 서버에서 먼저 처리하여 완성된 스테레오 신호를 재생단에 보내주는 방법인데, 이와 같은 방법은 디코더에서의 렌더링 연산량을 0으로 낮춰주는 큰 장점이 있지만, 개인화의 어려움과 On/Off 구현의 불편함 등 또 다른 이슈를 갖게된다. 프리렌더링을 잘 활용한 상용 기술로 DTS사의 Headphone X라는 기술이 있다[14].

## 3. 바이노럴 엠비소닉

Ambisonics의 경우에도 바이노럴 기술에 대한 연구가 이루어지고 있다. 바이노럴 엠비소닉디코딩(binaural ambisonic decoding)이라 부르는 이 기술은 ambisonics 포맷 신호(B-format)에 대해서 헤드폰과 같은 재생 환경을 가정하여 귀에서의 바이노럴 신호를 생성하기 위한 렌더링 기술이다[15]. 때문에 가장 일반적인 바이노럴 엠비소닉 디코딩 기술은 B-format 신호를 ambisonic 디코더를 통해서 스피커 채널 신호로 생성하고, 이를 청자의 청각 특성이 반영된 HRTF를 이용하여 스피커 배치에 따라 다시 렌더링해주는 구조로 이루어진다. 이를 가리켜 가상 스피커(Virtual loudspeaker Binaural or VLB) 기술이라고 부르기도 한다[16]. 이렇게 얻어진 바이노럴 신호는 멀티 채널 스피커 환경의 음향 효과를 재현하게 된다. VLB에 상대되어 HOA 요소신호로부터 바로 바이노럴 신호를 얻는 방법도 존재한다.

바이노럴 엠비소닉 디코딩에는 추가적으로 룸의 특성이 반영되기도 한다. 이는 재생 환경의 공간적

특성을 재현하고자 하는 것으로 BRIR을 이용하여 가상 스피커 채널을 렌더링하는 방법이 대표적이다. 하지만 BRIR은 채널 신호의 방향 이외에도 공간 상의 모든 점에 대해서 정보가 필요한데, 이를 모두 측정하고 저장하는 것은 현실적으로 어렵다. 때문에 무향실에서 측정된 HRTF와 Room IR로부터 얻을 수 있는 초기 반사음(early reflection) 및 잔향음(reverberant sound)을 적용하는 방법이 사용되기도 한다[17]. 이러한 방법은 사용자 편의에 따라 가상 스피커 신호의 합성 이후에 공간적 특성을 임의로 컨트롤할 수 있는 장점을 가지고 있다. 바이노럴 엠비소닉 디코딩은 ambisonic 디코딩된 신호의 특성과 마찬가지로 엠비소닉 신호의 차수(order)에 따라, 그리고 음원 객체의 움직임에 따라 공간적 재현 성능이 달라질 수 있다[18].

#### 4. MPEG-H 3D 오디오에서의 바이노럴

멀티미디어 국제 표준을 주도하고 있는 MPEG에서는 차세대 코딩 표준으로 MPEG-H를 추진하고 있다[19]. 현재 3D Audio 표준화는 MPEG-H의 오디오 시스템을 위한 것으로 두가지 입력 신호 Channel+Object(CO)와 High Order Ambisonic(HOA)에 대하여 독립적으로 표준화를 진행하고 있다. 3D Audio 표준에서는 CO 및 HOA의 Immersive 오디오 콘텐츠를 헤드폰에서 효과적으로 재생하기 위한 기술로써 Binaural rendering 역시 표준의 범위에 포함하고 있다. Binaural rendering은 현재 여러 후보 기술들이 제안된 가운데 경합 중에 있으며, 1차 6개 제안 기술 가운데 실질적으로 4개의 후보 기술이 남아서 최종 표준 결정을 앞두고 있다. 이 가운데 한국전자통신연구원-연세대학교-윌러스표준기술연구소 3개 회사가 공동으로 제안

한 기술의 경우 주파수에 따라 유효한 필터의 길이가 다른 특징을 착안하여 주파수 영역에서의 가변 길이 필터링(Variable Order Filtering in Frequency-domain; VOFF) 기술을 새롭게 제안하였다[20][21]. VOFF을 이용한 바이노럴 렌더링 기술 제안은 6개 제안 기술 가운데 음질과 연산량 모두에서 가장 우수한 성능을 보였으며, 최근 2014년 1월 표준화 회의를 통해서는 CO의 Reference Model 기술을 가지고 있는 독일 Fraunhofer IIS사의 Late Reverberation part 기술과의 결합을 통한 공동 제안이 최종 표준으로 채택되었다[22].

VOFF 기술은 MPEG-H 표준에서뿐 아니라, SBR이 사용되어 QMF domain으로 신호가 출력되는 기존의 HE AAC 기반 코덱의 후처리로써 Binaural Rendering에 적용 가능하여 다양한 확장 application을 고려해볼 수 있는 우수한 실시간 처리 기술이다.

한편, MPEG-H에서 채널 입력의 경우 22,2채널까지 지원하며, 2개의 LFE 채널이 포함될 수 있는데, 바이노럴 렌더링 기술에서는 LFE 채널에 대한 렌더링을 고려하고 있지 않았다. 최근의 실험결과에 따르면, 헤드폰 재생 환경에서 LFE를 포함한 바이노럴 렌더링이 그렇지 않은 경우와 비교할 때 상당한 음질 향상(선호도)를 보이는 것으로 나타났다[23]. 따라서 LFE 채널을 효과적으로 바이노럴 오디오에 포함시키기 위한 기술의 개발이 새롭게 요구된다.

### IV. 결론

홀로그래픽 디스플레이는 여러가지 기술적인 난제에도 불구하고 궁극의 3차원 디스플레이기술로서 관심과 노력이 계속되고 있는 분야이다. 이에 대

응되는 오디오 기술을 꼽는다면 바로 바이노럴이다. 개인화(Individualization)나 고품질을 위한 연산량 이슈 등 여러 기술적 과제에도 불구하고 진화가 계속되고 있으며, 스마트폰, 클라우드 등 모바일

환경에서의 연산 성능 향상 및 고성능 헤드폰 시장의 성장과 함께 고음질 콘텐츠에 대한 니즈의 확대는 바이노럴 렌더링 기술의 발전을 더욱 독려하게 될 것으로 전망된다.

### 참고 문헌

- [1] <https://www.youtube.com/watch?v=0eE4DcHUJ8M>
- [2] D. Begault. 3-D Sound for Virtual Reality and Multimedia . Academic Press, Boston, MA, 1994.
- [3] Jot Jean-Marc, Veronique Larcher, and Olivier Warusfel. "Digital signal processing issues in the context of binaural and transaural stereophony" the 98<sup>th</sup> AES Convention. Audio Engineering Society, 1995.
- [4] Schroeder, Manfred R. "Digital simulation of sound transmission in reverberant spaces." The Journal of the Acoustical Society of America 47 (1970): 424.
- [5] Cooper, Duane H., and Jerald L. Bauck. "Prospects for transaural recording." Journal of the Audio Engineering Society 37.1/2 (1989): 3-19.
- [6] Gardner, William G. 3-D audio using loudspeakers. Springer, 1998.
- [7] 장대영, 서정일, 이태진, 박기윤, 강경옥, "실감방송을 위한 3차원 오디오 기술", 전자통신동향분석 제19권 제4호 2004년 8월.
- [8] Algazi, V. Ralph, et al. "The cipc hrtf database." Applications of Signal Processing to Audio and Acoustics, 2001 IEEE Workshop on the. IEEE, 2001.
- [9] Shimada, Shoji, Nobuo Hayashi, and Shinji Hayashi. "A clustering method for sound localization transfer functions." Journal of the Audio Engineering Society 42.7/8 (1994): 577-584.
- [10] Zotkin, D. Y. N., et al. "HRTF personalization using anthropometric measurements." Applications of Signal Processing to Audio and Acoustics, 2003 IEEE Workshop on.. IEEE, 2003.
- [11] Hu, Hongmei, et al. "Head related transfer function personalization based on multiple regression analysis." Computational Intelligence and Security, 2006 International Conference on. Vol. 2. IEEE, 2006.
- [12] Kahana, Yuvi, et al. "Numerical modelling of the transfer functions of a dummy-head and of the external ear." Audio Engineering Society Conference: 16th International Conference: Spatial Sound Reproduction. Audio Engineering Society, 1999.
- [13] Evans, Michael J., James AS Angus, and Anthony I. Tew. "Spherical harmonic spectra of head-related transfer functions." Audio Engineering Society Convention 103. Audio Engineering Society, 1997.
- [14] <http://www.dts.com/professionals/sound-technologies/headphonex.aspx>
- [15] Collins, Tim. "Binaural Ambisonic Decoding with Enhanced Lateral Localization." Audio Engineering Society Convention 134. Audio Engineering Society, 2013.
- [16] Jot, Jean-Marc, Véronique Larcher, and Jean-Marie Pernaux. "A comparative study of 3-d audio encoding and rendering techniques." Audio Engineering Society Conference: 16th International Conference: Spatial Sound Reproduction. Audio Engineering Society, 1999.
- [17] Noisternig, Markus, et al. "A 3D ambisonic based binaural sound reproduction system." Audio Engineering Society Conference: 24th International Conference: Multichannel Audio, The New Reality. Audio Engineering Society, 2003.
- [18] Gorzel, Marcin, et al. "On the perception of dynamic sound sources in ambisonic binaural renderings." Audio Engineering Society Conference: 41st International Conference: Audio for Games. Audio Engineering Society, 2011.
- [19] ISO/IEC JCT1 SC29 WG11 Output Doc. N13411, "Call for Proposals for 3D Audio," The 103<sup>rd</sup> MPEG Meeting, Geneva, Jan. 2013.



- [20] ISO/IEC JCT1 SC29 WG11 Input Doc. M31271, "Description of Yonsei proposal for MPEG-H 3D Audio Binaural CE," The 106<sup>th</sup> MPEG Meeting, Geneva, Oct. 2013.
- [21] ISO/IEC JCT1 SC29 WG11 Input Doc. M31297, "Description of Yonsei proposal for MPEG-H 3D Audio Binaural CE," The 106<sup>th</sup> MPEG Meeting, Geneva, Oct. 2013.
- [22] ISO/IEC JCT1 SC29 WG11 Input Doc. M32223, "Technical Description of ETRI/Yonsei/WILUS Binaural CE proposal in MPEG-H 3D Audio," The 107<sup>th</sup> MPEG Meeting, San Jose, Jan. 2014.
- [23] ISO/IEC JCT1 SC29 WG11 Input Doc. M32270, "Preliminary CE proposal on LFE binauralization for MPEG-H 3D Audio," The 107<sup>th</sup> MPEG Meeting, San Jose, Jan. 2014.

## 필자 소개



### 오현오

- 1996년 2월 : 연세대학교 전자공학과
- 1998년 2월 : 연세대학교 대학원 전자공학과
- 2002년 8월 : 연세대학교 대학원 전기전자공학과
- 2010년 8월 : LG전자 Digital TV 연구소 책임연구원
- 2011년 3월 : 연세대학교 전기전자공학과 연구교수
- 2014년 현재 : 윌러스표준기술연구소 연구위원
- 주관심분야 : 오디오 신호처리, 오디오 부호화, 객체기반 오디오, 표준화



### 이태규

- 2012년 2월 : 연세대학교 전기전자공학과
- 2014년 현재 : 연세대학교 전기전자공학과



### 전세운

- 2007년 2월 : 연세대학교 전기전자공학과
- 현재 : 연세대학교 전기전자공학과
- 주관심 분야 : 음성 및 오디오 신호처리, 3D 오디오, 멀티채널 오디오 시스템



### 윤대희

- 1977년 2월 : 연세대학교 전자공학과 학사
- 1979년 8월 : Kansas State University 석사
- 1982년 2월 : Kansas State University 박사
- 1985년 9월 ~ 현재 : 연세대학교 전기전자공학과 교수
- 주관심분야 : 디지털 신호처리, 적응 신호처리, 음성 신호처리, 오디오 신호처리

## 필자소개



### 박영철

- 1986년 2월 : 연세대학교 전자공학과 학사
- 1988년 2월 : 연세대학교 전자공학과 석사
- 1993년 2월 : 연세대학교 전자공학과 박사
- 2002년 3월 ~ 현재 : 연세대학교 컴퓨터정보통신공학부 교수
- 주관심분야 : 음성 신호처리, 적응 신호처리, 오디오 신호처리



### 서정일

- 1994년 2월 : 경북대학교 전자공학과 학사
- 1996년 2월 : 경북대학교 전자공학과 석사
- 2005년 8월 : 경북대학교 전자공학과 박사
- 1998년 3월 ~ 2000년 10월 : LG반도체 주임연구원
- 2000년 11월 ~ 현재 : 한국전자통신연구원 실감방송미디어연구부 선임연구원
- 2010년 8월 ~ 2011년 7월 : 영국 사우스햄튼대학 방문연구원
- 주관심분야 : 오디오 부호화, 다채널 음장재현 시스템, 3차원 오디오, 디지털 방송 시스템, 객체기반 오디오 시스템



### 이용주

- 1999년 2월 : 경북대학교 전자공학과 학사
- 2001년 2월 : 경북대학교 전자공학과 석사
- 2001년 2월 ~ 현재 : 한국전자통신연구원 실감방송미디어연구부 선임연구원
- 주관심분야 : 3차원 오디오, 객체기반 오디오 시스템, 오디오 신호처리