

입체음향 생성을 위한 3채널 바이노럴 녹음기법에 관한 연구.

이신렬, 성평모

서울대학교 전기공학부

A study on 3-channel binaural recording technique
for 3D sound generation.

E-Mail : sinlyul@acoustics.snu.ac.kr

Abstract

입체음향 생성을 위한 기존의 방법은 크게 바이노럴 녹음기법과 머리전달함수(HRTF)를 이용한 바이노럴 합성 기법으로 나눌 수 있다. 현재 바이노럴 기법은 기존 스테레오 시스템에 비해 공간감, 몰입감 측면에서는 탁월한 효과가 있지만, 음질의 저하와 정면 음상 정위가 잘 되지 않는다는 치명적인 단점 때문에 프로페셔널 오디오 분야에서는 거의 사용되지 않고 몇몇 PC 게임용으로만 사용되고 있다. 본 논문은 정확한 정면 음상 정위를 위해 "3채널 더미헤드를 이용한 바이노럴 녹음기법"을 제안하고, 기존 스테레오 녹음기법과의 호환성 유지를 위해 녹음 현장에서 직접 사용될 수 있는 3채널 더미헤드를 사용한 "Weighted Diffuse-field equalization 기법"에 대해 제안하며, 3 채널 더미헤드를 이용하여 기존 HRTF 데이터를 대체할 수 있는 정면 음상 정위에 강인한 "3 채널 더미헤드 HRTF 측정 기법"에 대해 제안한다.

1. 서론

바이노럴 녹음이 스테레오 녹음의 한계를 뛰어넘는 차세대 녹음 방식의 표준이 될 것이라는 것에 대해서는 많은 사운드 엔지니어들이 동의하고 있지만 재생 시 나타나는 "정면 음상 정위의 어려움", "스피커와의 호환성", "머리 내 음상 정위" 등의 문제점 때문에 더미헤드가 만들어진지 반세기가 지난 지금에도 바이노럴 녹음은 일부에서만 사용되고 있다. 이는 바이노럴 녹음이 단순히 더미헤드를 사용하는 것으로 끝나는 것이 아니라 녹음 이후 다양한 신호처리가 요구되기 때문이다. 지금까지 입체음향에 대한 많은 연구들은 이러한 바이노럴 녹음 기법보다는 "트랜스오럴 필터"등을 사용한 입체음향 재생기법에 대해 많은 관심을 가져왔다. 이 결과 어떤 음악을 트랜스오럴 필터를 통해 재생시키면 어느 정도 입체음향을 느낄 수 있는데 반해 어떤 음악은 별로 효과가 없어 트랜스오럴 필터가 제대로 작동을 하는지에 대해서도 판단할 수 없게 된다. 이는 기존의 스테레오 녹음 방식인 "원 포인트 녹음 방식"에서 "AB 방식"에 이르는 다양한 녹음 기법에 대한 정확한 이해와 많은 실험을 바탕으로 바이노럴 녹음 방식을 이해해야 할 것이다. 본 논문은 "정면 음상 정위의 어려움", "스피커와의 호환성", "머리 내 음상 정위" 등의 기존의 바이노럴 녹음 시에

나타났던 문제점들을 해결할 수 있는 "기존 스테레오 녹음방식"과 "Ambisonic 녹음 방식"을 결합하여 바이노럴 녹음방식에 접목시킨 "입체음향 생성을 위한 3채널 바이노럴 녹음기법"을 제안한다.

2. 3채널 더미헤드를 이용한 바이노럴 녹음.

바이노럴 녹음은 입체음향을 생성하는 가장 손쉬운 방식이다. 개인의 경우 무 지향성 마이크를 이용하여 자신의 귀 근처에서 자신의 HRTF로 간편하게 녹음할 수 있으며, 상업용인 경우 더미헤드를 이용하여 공연장에서 직접 녹음할 수 있다.

2-1. 기존 바이노럴 녹음시의 문제점

바이노럴 녹음은 이러한 제작상의 간단함에도 불구하고 다양한 문제점을 안고 있다.

① "머리 내 음상 정위"

더미헤드의 주파수응답은 입사 음의 방향에 따라 매우 다르게 나타난다. 헤드폰 재생 시 귀는 이러한 주파수 응답 변이로 방향감, 깊이감, 음원의 높이를 판단한다. 만약 이러한 공간 정보가 더미헤드 마이크와 청취자의 귀 사이의 오차를 포함해서 헤드폰을 통해 재생되면 이러한 오차는 청취자를 혼란시키고 "머리 내 음상 정위"를 유발한다. 이는 자신의 머리를 사용하지 않는 경우에는 필연적으로

나타나는 현상이며 정확한 Equalization을 통해 “머리 내 음상 정위 현상”을 줄일 수 있다.

② “정중면 음상 정위” : Directional band
정중면 내에서의 음의 방향 지각은 음원의 실제 위치보다는 음원의 주파수 성분에 더욱 의존한다. 이는 양 귀 시간차가 없는 정중면 내에서의 정확한 음상 정위가 더미헤드 녹음만으로는 어렵다는 것을 나타낸다. 즉 정확한 Equalization과 부가적인 액센트 마이크를 사용함으로써 안정된 음상을 정위시킬 수 있다.

③ “Hole in the middle”효과.
시간차를 이용한 스테레오 마이킹 기법에서 두 마이크가 서로 멀리 떨어져 있을 때 역 자승 법칙 (Inverse square Law)에 의해 일어난다. 오케스트라 녹음 시 중앙에 배치된 악기는 이 영향 때문에 다소 작게 들리고 바깥 양쪽 악기 쪽으로 양분되어 들리게 되며, “Comb filter 효과” 영향으로 음이 부자연스럽고 위상이 변형되어 들리게 된다. 이렇게 중앙 이미지가 근접한 스피커로 치우치는 현상을 “Hole in the middle”이라고 한다. 이러한 문제점은 무 지향성 마이크를 이용한 “AB 방식” 스테레오 녹음에서도 동일하게 나타나며, 이를 해결하기 위해 1930년대 Bell 연구소에서 액센트 마이크를 중앙 채널에 부가시킴으로써 “Hole in the middle”현상을 없앨 수 있었고 정면 음상을 더욱 뚜렷하게 정위시킬 수 있었다. 이 방식은 현재 오디오/비디오 분야의 멀티 채널 재생 표준으로 자리 잡고 있다. 또한 이 현상은 더미헤드 녹음 시 비대칭을 Equalization 처리함으로써 크게 줄일 수 있다.

④ 정확한 음상 재현의 어려움.
더미헤드를 사용한 바이노럴 녹음은 기존의 스테레오 녹음기법에서 “무 지향성 마이크를 사용한 Near-coincident 녹음기법”과 유사하다. 이는 One-point 녹음인 “Ambisonic”기법에 비해 다음과 같은 문제점을 내포하고 있다.

㉠ “Near-coincident 녹음기법”은 comb filter 효과로 인해 정확한 음상 정위가 어렵고 경우에 따라서는 정면 음이 수척해지거나 속이 빈소리로 종종 들린다.

㉡ 양 헤드폰 사이 혹은 두 스피커 사이의 음상 정위가 어렵다. 이는 현재 대부분의 스테레오 녹음이 one-point 녹음의 일종인 “Panpotted Mono”방식을 사용하고 있으며, 이 방법은 양 헤드폰 사이 혹은 두 스피커 사이의 음상 정위에 가장 효과적이다.

㉢ “Dummy-Recorder-Headphone 시스템”의 시간차가 청취자의 양 귀 시간차와 다를 경우 입체음향 효과는 크게 저하된다. 그리고 헤드폰을 통해 기존 스테레오를 수년간 청취한 것이 우리의 귀를 더미헤드에 의해 제공되는 양 귀 시간차를 받아들이지

못하도록 혼란시킨다. 이는 자신의 머리를 사용하지 않고 더미헤드로 녹음할 경우 더미헤드의 양 귀 사이의 시간차가 자신과 조금이라도 다를 경우 심각한 음질의 저하와 음상 정위 저하를 가져오게 된다. 이는 양 귀 시간차를 없애고 저주파수 L-R 신호 보강을 위한 “Blumlein shuffling filter”를 사용하면 어느 정도 개선된다.

2-2. “3채널 더미헤드를 이용한 바이노럴 녹음 기법”

이러한 문제점을 보완하기 위한 “3채널 더미헤드를 이용한 바이노럴 녹음 기법”은 기존의 더미헤드의 정면에 액센트 마이크를 부가한 새로운 더미헤드를 사용하여 바이노럴 녹음을 한다. 이 방식의 장점은 다음과 같다.

- ㉠ 정확한 정면음상을 필요로 하는 오디오/비디오 분야에서 입체음향 녹음 시 유용하다.
- ㉡ 멀티 채널 확장이 용이하다.
- ㉢ “Sweet spot”이 넓어진다.
- ㉣ 정확한 음상 재현이 가능하다.

3. 3채널 더미헤드를 이용한 “Weighted Diffuse-field Equalization” 기법

Equalization은 측정과 녹음과정에서 변환기의 영향을 제거하기 위해 사용되어진다. 하지만 Equalization은 이러한 목적 이외에도 기존 스테레오와의 호환성 유지와 헤드폰 제작을 위해서 더욱 중요한 의미를 가진다. 현재까지의 대표적인 Equalization 기법인 “Free-field equalization”과 “Diffuse-field equalization”에 대해 알아보고 기존 스테레오 녹음기법과의 호환성 유지할 수 있는 녹음현장에서 직접 사용될 수 있는 3 채널 더미헤드를 사용한 “Weighted Diffuse-field Equalization 기법”에 대해 제안한다.

3-1. “Free-field equalization” 과 “Diffuse-field equalization”

Equalization 선택 기준은 어느 방향이 청취 시 가장 중요한 전송 경로인지에 대한 다양한 가정에서 그 기원을 찾을 수 있다. 기존 상황은 청취실의 특성을 묘사하는데 사용되어진다. Free-field Equalization(이하 FF-EQ)는 무향실에서 청취자의 정면에 위치한 단일 음원을 가정한 것이다. 이는 “Free-field Equalized 헤드폰”의 설계 배경이 된 아이디어이다. 또한 0°와는 다른 방위각을 FF-EQ로 사용될 수 있다. 이 주장은 일반적인 스테레오 장치에서 스피커로부터 청취자에게 도달하는 직접음이 대개 20~50° 사이의 입사각을 가진다는 데에 따른 것이다. 이와는 반대로 적절한 흡수율을 가진 청취

실에서도 직접음은 청취자에게 도달하는 음의 일부 분에 지나지 않으며 청취 지점에서의 음은 대부분 반사음을 포함한다고 말할 수 있다. 이러한 가정은 청취 위치에 대한 스피커의 위치가 청취실의 반경에 비해 다소 멀 때 성립된다. 이러한 기준상황은 Diffuse-field로 주어지게 되고 이를 "Diffuse-field Equalization"(이하 DF-EQ)라 한다. 이러한 설계 기준을 따르는 헤드폰을 "Diffuse-field Equalized 헤드폰"이라고 부른다. FF-EQ에 대한 DF-EQ의 장점은 다음과 같다.

- ① FF-EQ는 특정 방향의 HRTF를 포함하고 있으며 청각은 이를 "Directional Cue"로 해석할 수 있다. 이에 반해 DF-EQ는 어느 특정 방향으로부터의 Cue를 포함하고 있지 않다.
- ② Theile는 인간이 DF-EQ를 FF-EQ보다 선호한다는 청취 평가 결과를 내놓고 헤드폰 제작의 기준을 기존의 FF-EQ에서 DF-EQ로 바꾸었다.
- ③ DF-EQ는 더미헤드 녹음에서 기존 스테레오 마이크 기법과 호환성이 개선된다.
- ④ DF-EQ는 측정 HRTF와 헤드폰 응답으로부터 방향성과 관련이 없는 부분을 제거하는 가장 강한 기법이며 서로 다른 머리와 다른 장비를 사용함으로써 나타나는 편차를 크게 감소시킬 수 있다.
- ⑤ DF-EQ는 3D 사운드 시스템에서 HRTF를 "Minimum-phase 바이노럴 합성 필터"로 구현 가능하게 한다.

3-2. DF-EQ를 구하는 방법

이러한 바이노럴 측정 및 녹음 시 사용되는 DF-EQ 곡선을 구하는 방법은 다음과 같다.

- ① Stationary excitation in diffuse field.
이 방법은 확산 음장을 실제로 만드는 것으로 잔향실 내에서 stationary 여기신호의 응답들을 사용하여 DF-EQ곡선을 구하는 것으로 간단하고 신속하게 측정값을 얻을 수 있지만 이러한 음장은 룸의 용적에 의존되는 따라서 Mode가 중첩되지 않는 저 주파수에서는 정확한 추정 값을 얻기 어렵다. 이는 더 큰 잔향실을 사용하면 해결할 수 있지만 반대로 공기의 흡수가 음장이 균일한 고 주파수 한계 범위를 줄어들게 한다. Kuhn은 모든 주파수 범위에서 "Diffuse-field HRTF" 추정 값을 유도하기 위한 두 개의 잔향실 사용을 제안했다. 또한 Free-field 내에서 음파를 중첩함으로써 확산 음장을 만드는 방법에 대한 연구를 통해 Veit는 "uncorrelated narrow-band noise" 입력 시 확산 음장을 생성하는 무향실 내에서의 8개 스피커 장치를 제안했다. 이러한 접근 방법은 넓은 주파수 범위에 유효하다. 그러나 장치내의 음향 인텐시티 벡터의 측정은 공간적으로 제한된 지역에 대해 음장이 diffuse 하게된다. 이는 피험자의 머리에 대해서 너무 작아 그 안에

위치시키기가 어렵다는 단점이 있다.

② Averaged frequency response measurements in free field.

이 방법은 음장의 물리적인 생성에 의존하지 않고 모든 방향으로 입사하는 자유 음장 내의 평면 파에 의한 변형된 신호의 중첩에 의해 DF-EQ를 유도해 내는데 목적을 두고 있다.

이 방법은 대부분의 경우 무향실인 자유 음장 내에서 측정된 HRTF의 Power Averaging을 통해 DF-EQ를 얻는다. DF-EQ 곡선의 정확한 추정은 입사 음의 연속적인 분포를 요구한다. 그러나 HRTF는 이산적인 입사 음으로만 측정되어질 수밖에 없기 때문에 측정 지점 수는 더미헤드의 DF-EQ 비교 평가를 통해 어려워진다. 이 방법은 아주 많은 수의 측정을 필요로 하며 추정 값의 정확성에 대한 제한 없이는 측정지점의 개수를 줄일 수 없다. 게다가 청취자 혹은 더미헤드 주위로 음원을 옮기는 것이 많은 비용과 시간 그리고 장비를 요구한다.

3-2. "Weighted Diffuse-field Equalization" 기법

DF-EQ는 잔향실, 무향실과 같은 특정 공간 긴 측정 과정이 필요로 하기 때문에 많은 비용과 시간이 필요하다. 이러한 한계를 극복하기 위해 녹음 공간에서 바로 Equalization 필터를 얻을 수 있는 새로운 추정 기법을 제안한다. "Weighted Diffuse-field Equalization"(이하 Weighted DF-EQ)기법의 장점은 다음과 같다.

- ① 더미헤드를 잔향이 있는 음장 내에서 음원으로 부터 상당히 멀리 배치시킨다면 DF-EQ는 유지될 수 있을 것이다. 그러나 이러한 마이크의 위치는 음원으로부터 너무 멀리 떨어져 있어 잔향을 너무 많이 포함하게된다.
- ② 실시간 주파수 분석기를 사용한 음장 분석에서 대부분의 음의 에너지는 더미 헤드의 전방에서 입사하며, 더미헤드 응답에서는 상당한 양의 좌우 비대칭이 존재한다. 이는 기존의 DF-EQ 곡선과는 상당한 편차를 나타내며 기존의 좌우 동일한 곡선을 사용할 경우 재생 시 많은 문제점을 포함하게 된다.
- ③ 다양한 음악 장르와 공연장에 따라 EQ 곡선이 조금씩 다르게 나타나며 이는 기존의 DF-EQ 곡선 보다는 녹음 현장에서 음악에 따른 EQ 곡선을 구하는 것이 가장 효과적이다.
- ④ 공연장에서 더미헤드 측정 결과 "10¹/3 Octave Averaging Free-field" 곡선이 실시간 주파수 분석기 곡선과 가장 가까웠고 그 다음으로 수평면내의 30° 곡선, 그 다음으로 DF-EQ 곡선이 가까웠다. 이 결과는 바이노럴 녹음 또는 합성 시 DF-EQ에 비해 "Weighted DF-EQ"이 보다 나은 음질을 제공할 수 있다는 가능성을 제시한다.

“Weighted DF-EQ기법”은 액센트 마이크를 장착한 3채널 더미헤드를 사용하여 정면 음상 정위에 강인하고 음질의 저하가 없으며 스테레오 호환성도 뛰어난 가장 자연스러운 “Weighted DF-EQ 곡선”을 찾을 수 있다. 또한 잔향실, 무향실과 같은 특정공간이 필요한 한계를 극복할 수 있으며 녹음이 진행되는 바로 그 공간에서 녹음 음원에 따른 “Weighted DF-EQ 곡선”을 구할 수 있다. “Weighted DF-EQ 곡선”은 녹음 공간에서 측정 임펄스 응답의 후부 잔향 주파수 분석을 통해 유도된다. 이를 위해서는 다음과 같은 parameter를 고려해야 한다.

- ① 시작 시각: 직접음과 초기 반사음 도착 이후.
- ② 끝 시각: 고 주파수에서 noise level 이하로 유용한 신호가 떨어지기 이전.
- ③ 간격: sharp한 주파수 resolution을 제공하기에 충분히 긴 시간.

4. 3채널 더미헤드를 이용한 HRTF 측정 기법
 측정 HRTF를 사용하는 현재의 바이노럴 합성 방식은 시스템의 성능 향상에도 불구하고 바이노럴 녹음이 안고 있는 “정면 음상 정위의 어려움”, “스피커와의 호환성 문제”, “머리 내 음상 정위”의 문제점을 그대로 가지고 있다. 이러한 문제점을 극복하기 위해 최근에 등장한 것이 “HRTF refinement 기법”이다. Ming Zhang은 “Cone-of Confusion”내의 음상 정위 개선을 위해 Pinnae 효과를 강조하고 서로 다른 방향으로부터 입사하는 음의 차이를 증대시키기 위해 스펙트럼 Peak부분은 유지시키면서 Notch 부분을 증폭시켜 스펙트럼 차를 증폭시켰다. 이 결과 “Front-back reversal”을 줄일 수 있었고 이동 음의 정위 능력을 향상시킬 수 있었다. 하지만 이 실험은 음상 정위를 위해 음질의 상당한 저하를 감수 해야했으며 트랜스오럴 필터 사용 시 음질의 저하는 더욱 심해진다.

4-1. “3채널 더미헤드 HRTF 측정 기법”

단일 지향성 액센트 마이크를 이용한 3채널 더미헤드는 Notch를 강조하는 방식과는 달리 정면 음을 자연스럽게 강조하여 음질 저하를 막을 수 있고 후면 음은 단일 지향성 마이크를 사용하기 때문에 그대로 유지할 수 있어 스펙트럼 차를 증폭시킬 수 있다. “3채널 더미헤드 HRTF 측정 기법”의 장점은 다음과 같다.

- ① Pinnae 효과를 증대시킴으로써 “Front-back”현상을 줄일 수 있다.
- ② 이동 음의 정위 능력을 향상시킬 수 있다.
- ③ 기존 측정 HRTF를 “3 채널 더미헤드 HRTF”로 단순히 교체함으로 시스템의 교체 없이 기존 바

이노럴 합성기의 문제점을 개선시킬 수 있다.

5. 결론 및 향후 계획

바이노럴 녹음 및 합성 기법은 입체 음반 제작, HMD를 이용한 가상 현실 게임 및 시뮬레이터 제작, 차세대 방송인 입체 방송 등에 필수적인 부분이다. 제안한 “3채널 더미헤드를 이용한 입체음향 제작기법”은 기존 더미헤드 녹음의 문제점을 보완하여 바이노럴 녹음의 현장감, 사실감에 스테레오 녹음의 고음질 화를 더할 수 있었다. 바이노럴 녹음의 대중화를 가로막는 가장 큰 걸림돌은 바이노럴 녹음상의 문제점 이외에도 녹음에서 후처리까지 누구나 손쉽게 사용할 수 있는 통합된 장비가 없다는 것이다. 오디오 신호처리를 전공한 엔지니어 수준만이 이해할 수 있는 바이노럴 녹음 및 후 신호처리를 3채널 더미헤드는 녹음과 녹음 현장에서의 실시간 Equalization, 음장을 인위적으로 제어할 수 있는 BSF filter를 하나로 묶어 누구나 쉽게 사용할 수 있게 했으며 후 처리 과정을 극소화 할 수 있었다. 향후 계획은 다음과 같다.

- ① 청취평가를 통한 액센트 마이크 선택 및 개인 값 조절.
- ② 청취평가를 통한 3채널 더미헤드 사양 선택.
- ③ “3채널 더미헤드 측정 HRTF”를 이용한 바이노럴 합성기 청취 평가.
- ④ One point 녹음기법과 바이노럴 녹음기법을 결합한 다양한 녹음 기법 연구.
- ⑤ 가상현실 게임, 시뮬레이터, 입체 방송 등 다양한 장르에 적합한 입체음향 콘텐츠 제작에서부터 신호처리, 스피커 배치에 이르는 입체음향 통합 solution 개발.

본 논문에서 제안한 “3채널 더미헤드를 이용한 바이노럴 녹음기법”, 3채널 더미헤드를 사용한 “Weighted DF-EQ기법”, “3채널 더미헤드 HRTF 측정 기법” 및 제반 장치는 현재 “A&D engineering”과 함께 특허 출원중이다.

6. 참고 문헌

1. H. Moller, “Design criteria for headphones”
2. G. Theile, “On the standardization of the frequency response of High-Quality studio headphones”
3. J. Jot, “Equalization methods in binaural technology”
4. D. Griesinger, “Equalization and spatial equalization of Dummy-head recording for loudspeaker reproduction”
5. R. Streicher, “The new stereo soundbook”
6. H.W. Gierlich, “Processing Artificial-Head Recording”
7. R. Allen, “Whatever happened to Ambisonic”