

(19)대한민국특허청(KR)  
(12) 공개특허공보(A)(51) 。 Int. Cl.<sup>7</sup>  
H04R 5/00  
H04R 5/02  
H04R 3/00(11) 공개번호 10-2005-0056241  
(43) 공개일자 2005년06월14일(21) 출원번호 10-2005-7006432  
(22) 출원일자 2005년04월14일  
번역문 제출일자 2005년04월14일  
(86) 국제출원번호 PCT/US2003/030392  
국제출원출원일자 2003년09월26일(87) 국제공개번호 WO 2004/039123  
국제공개일자 2004년05월06일(30) 우선권주장 10/414,261 2003년04월15일 미국(US)  
60/419,734 2002년10월18일 미국(US)(71) 출원인 더 리전트 오브 더 유니버시티 오브 캘리포니아  
미합중국 캘리포니아 94607-5200 오클랜드 프랭클린 스트리트 1111(72) 발명자 알가지 브이 랄프  
미국 캘리포니아 95616, 다비스, 1606 윌로우 레인  
듀다 리차드 오  
미국 캘리포니아 94025, 멘로 파크, 590 바인 스트리트  
탐슨 데니스  
미국 캘리포니아 95617, 다비스, 피 오 박스 74346

(74) 대리인 서원호

심사청구 : 없음

## (54) 동적인 바이노럴 음향 캡처 및 재생 장치

## 명세서

## 기술분야

본 발명은 공간 음향 캡처 및 재생, 특히, 삼차원 공간 음향의 동적인 특성을 캡처하고 재생하기 위한 방법 및 시스템에 관한 것이다.

## 배경기술

공간 음향 캡처 및 재생에 관한 많은 대안의 접근들이 있으며, 전형적으로 사용되는 특별한 접근은 음원이 자연적인 것인지 또는 컴퓨터에 의해 만들어진 것인지에 의존한다. 자연의 소리를 녹음하고 재생하기 위한 공간 음향 기술의 우수한 개관은 *F.Rumsey, Spatial Audio* (Focal Press, Oxford, 2001)에서 찾아볼 수 있으며, 가상 음원의 발생 및 실시간 "연주"를 위한 컴퓨터 활용 방법의 필적할 만한 개관은 D.B.Begault, *3-D Sound for Virtual Reality and Multimedia*(AP Professional, Boston, 1994)에서 찾을 수 있다. 다음은 비교적 잘 알려진 접근의 개관이다.

서라운드 음향(예를 들어, 스테레오, 4채널 방식, 돌비® 등)은 공간 음향을 녹음하고 재생하는데 단연 가장 인기있는 접근이다. 이러한 접근은 개념적으로 간단하다; 즉, 소리가 나길 원하는 곳에 확성기를 놓아두면 음향은 그 위치로부터 나올 것이다. 그러나, 실제로 그렇게 간단하지는 않다. 확성기 사이에 있는 위치, 특히 사이드를 따라서 음향이 나오는 것으로 보이게 하는 것은 어렵다. 같은 음향이 하나 이상의 스피커에서 나온다면 선행음 효과는 소리가 가장 가까운 스피커로부터 나오는 것처럼 보이게 하는 결과를 발생시키고, 스피커에 가까이 앉은 사람들에게 특히 불리하다. 주요 결과는 청취자로 하여금 매우 좁은 "스위트 스팟" 근처에 머물도록 제한하는 것이다. 또한, 복합의 고음질 스피커에 대한 요구는 불편하고 비싸며, 집에서의 사용을 위하여, 많은 사람들이 두개 이상의 스피커의 사용은 수용할 수 없다는 것을 발견한다.

이러한 제한을 줄이기 위한 서라운드 음향을 얻는 대안의 방법들이 있다. 예를 들어, 홈 시어터 시스템은 전형적으로 두 개의 확성기 사이의 공간을 넘어서 음향 스테이지를 확장시키기 위한 음향심리학적 효과를 포함하는 2 채널 믹스를 제공한다. 스피커 신호를 헤드폰 신호로 전환하는 것에 의하여 다수의 확성기에 대한 요구를 회피하는 것 역시 가능하며, 이는 소위 돌비® 헤드폰에서 사용되어지는 기술이다. 그러나 이러한 대안책을 역시 그 자체의 한계를 가지고 있다.

서라운드 음향 시스템은 멀리서 들려오는 음향을 재생하는 데에는 유용하나, 일반적으로, 귀에 속삭이는 것처럼 매우 가까운 음향의 효과를 만들 수는 없다. 결국, 효과적인 서라운드 음향 레코딩을 만드는 것이 전문 음향 기술자의 일이며; 그러한 접근은 원격 회의 또는 아마추어에게는 부적합하다.

다른 접근은 엠비소닉™이다. 널리 사용되고 있지는 않지만, 서라운드 음향에 대한 엠비소닉 접근은 레코딩을 만드는데 있어서 많은 문제점들을 해결한다. (M.A.Gerzon, "Ambisonics in multichannel broadcasting and video," Preprint 2034, 74th Convention of the Audio Engineering Society (New York, Oct.8-12,1983) 및; 이어서 출간된 J.Aud.Eng.Soc., Vol.33, No.11, pp.859-871 (Oct., 1985)) 저장된 구면 조화 함수에 의하여 흔히 있는 음장에 근접하는 방법으로서 추상적으로 서술되어 있다. (J.S.Bamford and J.Vanderkooy, "Ambisonic sound for us," Preprint 4138, 99th Convention of the Audio Engineering Society (New York, Oct.6-9, 1995)). 엠비소닉 레코딩은 삼각각 방향에서 국부적인 압력과 압력 차이를 느끼기 위해서 사운드필드™ 마이크로폰이라 불리는 특수하고, 집약된 마이크로폰 배열을 사용한다. 기본적인 엠비소닉 접근은 복잡함에 있어서의 상응하는 증가와 함께 더 나은 각도 분해능을 조건으로 3 방향 이상으로부터 레코딩을 허용하도록 확장되어진다.

다른 서라운드 음향 방법처럼, 엠비소닉은 확성기의 배열을 위해서 매트릭싱 방법을 사용하며, 이리하여 멀티 스피커 시스템의 모든 다른 장단점들을 가진다. 게다가, 모든 스피커들은 국부적인 압력 구성 성분을 재생하는데 사용된다. 결과로써, 청취자가 스위트 스팟에 위치해 있을 때, 그 구성 성분은 마치 청취자의 머리 안에 있고, 머리 운동이 즐거운 음향 인공물을 소개하는 것처럼 들리는 경향이 있다. (W.G.Gardner, *3-D Audio Using Loudspeakers* (Kluwer Academic Publishers, Boston, 1998), p.18).

웨이브 필드 합성은 매우 실용적이지는 않으나, 또 다른 접근이 된다. 이론적으로, 충분한 마이크로폰과 충분한 확성기로, 레코딩이 만들어지는 공간의 내부를 통하여 존재하는 음압 필드를 재생하기 위하여 서라운드 표면에 마이크로폰에 의하여 캡처되는 음향을 사용하는 것이 가능하다. (M.M.Boone, "Acoustic rendering with wave field synthesis," *Proc.ACM SIGGRAPH and Eurographics Campfire:Acoustic Rendering for Virtual Environment*, Snowbird, UT, May 26-29, 2001)). 이론적인 요구가 어려움에도 불구하고(즉, 수많은 확성기), 100개 이상의 확성기의 배열을 사용하는 시스템이 구조되었고, 효율적이라 말해진다. 그러나 이러한 접근은 분명히 비용면에서 효율적이지 않다.

바이노럴 캡처 역시 다른 접근이 된다. 삼차원적인 음향을 캡처하기 위해서 수백의 채널을 가지는 것이 필수적이지 않다는 것은 잘 알려져 있다; 실제로, 두개의 채널이면 충분하다. 2 채널 바이노럴 또는 3-D 이미지의 입체적인 재생의 어코스티크 아날로그인 "더미-헤드" 레코딩이 공간 음향을 캡처하는데 오래 사용되어 왔다. (J.Sunier, "Binaural overview: Ears where the mikes are. Part I," *Audio*, Vol.73, No.11, pp.75-84 (Nov.1989); J.Sunier, "Binaural overview: Ears where the mikes are. Part II," *Audio*, Vol.73, No.12, pp.49-57 (Dec.1989); K.Genuit, H.W.Gierlich, and U.Kunzli, "Improved possibilities of binaural recording and playback techniques," Preprint 3332, 92nd Convention Audio Engineering Society (Vienna, Mar.1992)). 기본이 되는 아이디어는 간단하다. 음향의 공간적인 특성을 감지하기 위해 인간 뇌에 의하여 사용되는 정보의 근본적인 근원은 왼쪽과 오른쪽 귀의 고막에 도달하는 압력파로부터 나온다. 이러한 압력파가 재생될 수 있다면, 청취자는 본래의 소리가 만들어질 때 마치 존재하는 것처럼 정확하게 그 소리를 들을 것이다.

귀 고막에 도달하는 압력파는 (a) 음향 근원, (b) 청취 환경, 및 (c) 청취자 자신의 몸에 의한 내부 파장의 반사, 회절, 산란을 포함하는 여러 요소들에 의하여 영향을 받는다. 청취자와 정확하게 같은 크기, 모양, 청각적 특성을 갖는 모델 인형이 사람의 귀 고막이 있는 귀 속에 위치한 마이크로폰을 갖추고 있다면, 귀 고막에 도달하는 신호들은 전달되거나 녹음될 수 있다. 그 신호가 헤드폰(헤드폰 드라이버로부터 귀 고막으로의 전달 기능을 바로잡기 위한 적절한 보정을 한)을 통하여 들려질 때, 음향 파형이 재생되고, 그 청취자는 그가 마치 실제로 모델 인형의 방향과 위치에 있는 것처럼 정확한 공간 특성을 가진 음향을 듣는다. 근본적인 문제는 귀 속 공명을 바로잡는 것이다. 헤드폰 드라이버는 귀 속 바깥에 있기 때문에, 귀 속 공명은 두번 나타난다; 한 번은 레코딩에서, 한 번은 재생에서 나타난다. 이것은 소위 말하는 "막힌 관" 레코딩의 사용을 권장하며, 여기에서는 귀 속이 차단되고 마이크로폰이 차단된 입구와 같은 높이이다. (H.Moller, "Fundamentals of binaural technology," *Applied Acoustics*, Vol.36, No.5, pp.171-218 (1992)). 바이노럴 캡처와 함께, 특히, 전화 통신의 이용에서, 실내 반향음이 자연적으로 들린다. 말하는 사람이 마이크로폰에 가깝지 않다면, 주위 환경이 과도하게 공허하고 반향적으로 들리는 스피커폰에서도 공통된 일이다. 바이노럴 픽업으로 들을 때, 이러한 산란한 반향음에 대한 의식이 사라지고 주변은 자연적이고 깨끗하게 들린다.

여전히, 입체 음향 캡처 및 재생과 관련하여 문제점들이 있다. 가장 두드러진 문제는 실제로 항상 중요하지는 않다. 그들은 (a) 모델 인형과 어떤 특정한 청취자의 크기, 모양 및 청각적 특성 사이의 피할 수 없는 불일치, (b) 압력 감지 요소로서의 귀 고막과 마이크로폰 사이의 차이점, (c) 음향 근원의 감지 위치에 시각적 또는 촉각적 신호와 같은 비청각적 요소의 영향을 포함한다. 예를 들어, KEMAR™ 모델 인형에서, 고막 방해물의 영향을 모의 실험하기 위하여 "Zwislocki coupler"를 사용하는데 상당한 노력이 든다. (M.D.Burkhard and R.M.Sachs, "Anthropometric manikin for auditory research," *J.Acoust.Soc.Am.*, Vol.58, pp.214-222 (1975). KEMAR는 노레스 일렉트로닉스, 1151 메이플우드 드라이브, 이타스카, 일리노이스, 60143에 의하여 생산된다.) 그러나, 그것은 마이크로폰이 변환기처럼 귀 고막에 동등하지는 않은 것으로 평가된다.

더욱 중요한 한계는 청취자의 머리의 움직임으로부터 발생하는 동적인 신호의 부족이다. 음향 근원이 모델 인형의 왼쪽에 있다고 가정한다. 청취자는 역시 청취자의 왼쪽으로부터 나오는 것처럼 음향을 들을 것이다. 그러나, 음향이 활발한 동안에 청취자가 근원을 마주보도록 방향을 돌린다고 가정한다. 레코딩은 청취자의 움직임을 감지하지 않으므로, 그 음향은 계속해서 청취자의 왼쪽에서 나오는 것으로 보일 것이다. 청취자의 관점에서, 그것은 마치 음향 근원이 왼쪽에 머무는 공간 내에서 움직이는 것과 같다. 청취자가 움직일 때, 많은 소리의 근원들이 활성화되어 있다면, 그 일은 전체의 청각적인

세계가 청취자와 함께 정확하게 동시적으로 움직이는 것이 된다. 다시 말해서, 레코딩이 만들어 지는 환경에서의 실제로 존재하고 있는 "가상의 현존"을 느끼기 위해서, 정지한 음향 근원은 청취자가 움직일 때 정지한 채로 남아있어야 한다. 다시 말해서, 가상의 음향 근원의 공간적인 위치는 안정적이고, 청취자의 움직임에 독립적이어야 한다.

청취자의 움직임의 결과가 바이노럴 레코딩의 다른 결점에 대하여 책임이 있다고 믿는 이유가 있다. 바이노럴 레코딩을 들을 때, 왼쪽 또는 오른쪽의 소리는 원래 멀리서 온 것처럼 보이나, 바로 앞의 소리는 항상 너무 가까이 있는 것처럼 보인다. 이것이 보편적인 경험이다. 실제로, 몇몇 청취자들은 음향 근원이 그들의 머리 내에 있거나 심지어 등에 있는 것처럼 경험한다. 몇몇 이유들이 이러한 "전면의 구체화"의 손실에 대하여 알려져 왔다. 한가지 논의는 우리가 우리의 정면에서 음향 근원을 보기를 기대하며, 확실한 시각적 신호가 없을 때, 우리의 뒤에 그 근원의 위치를 예상하려는 경향이 있다고 한다. 실제로 현실 상황에서, 음향의 근원이 우리의 앞에 있는지 뒤에 있는지를 말하는 것이 종종 어려우며, 확신이 없을 때 주위를 보기 위하여 몸을 돌리는 이유이다. 그러나 앞/뒤 애매함을 해결하기 위하여 완전히 주위를 돌아볼 필요는 없다. 음향 근원이 수직의 중앙 평면 어느 곳에 위치하고 있다고 가정해 보자. 우리의 신체가 기본적으로 이러한 평면에 대하여 대칭이기 때문에, 두 귀에 도달하는 소리는 본질적으로 같을 것이다. 그러나 우리가 머리를 왼쪽으로 약간 돌린다고 가정해 보자. 근원이 실제로 앞에 있다면, 그 음향은 왼쪽 귀에 도달하기 전에 오른쪽 귀에 도달할 것이고, 반면에 근원이 뒤에 있다면, 그 반대 상황이 일어날 것이다. 두 귀간 도달 시간 차이에서의 이러한 변화는 종종 앞/뒤 애매함을 해결하는데 충분하다.

그러나, 표준의 바이노럴 레코딩과 함께 무슨 일이 일어나는지에 대해서 주목하자. 그 근원이 정면에 있을 때, 우리는 양 왼쪽과 오른쪽 귀에서 같은 신호를 받는다. 레코딩이 청취자의 움직임을 인지하지 않기 때문에, 두 신호는 우리가 우리의 머리를 움직일 때 계속해서 같게 된다. 그러면, 당신이 당신 자신에게 두 귀에 대한 음향이 머리의 움직임에 관계없이 동일하게 남아 있다면 음향의 근원이 어디에 있는 것이 가능한지에 대하여 질문을 하여 본다면, 그 대답은 "머리 내부"이다. 동적인 신호는 매우 강력하다. 표준의 바이노럴 레코딩은 그러한 동적인 신호를 설명하지 못하며, 이는 "정면 붕괴"의 주요 이유이다.

이러한 문제를 해결하는 한 방법은 청취자의 머리가 방향을 바꿀 때, 더미 헤드 방향을 바꾸도록 서보 기구를 사용하는 것이다. 실제로, 그러한 시스템은 Horbach 등(U. Horbach, A. Karamustafaoglu, R. Pellegrini, P. Mackensen and G. Theile, "Design and applications of a data-basedauralization system for surround sound," Preprint 4976, 106th Convention of the Audio Engineering Society (Munich, Germany, May 8-11, 1999))에 의하여 실시된다. 그들은 그들의 시스템이 극히 자연의 음향을 만들어내며, 앞/뒤 혼란을 사실상 제거하였다고 보고하였다. 그들의 시스템이 매우 효율적임에도 불구하고, 한 번에 오직 한 사람의 청취자에 의하여 사용될 수 있고, 레코딩에는 전혀 사용될 수 없다는 것이 분명한 한계이다.

컴퓨터가 만들어낸 음향을 나타내는데 있어서 다음을 장점을 취하기 위해서 헤드 트래킹을 사용하는 많은 가상의 청각 공간 시스템(VAS systems)이 있다:

(i) 청취자의 머리 움직임에 독립하는, 가상의 청각 근원에 대한 안정된 위치, ;

(ii) 양호한 정면 구체화; 및 (iii) 앞/뒤 혼란이 적거나 혹은 없는 것이다. 그러나 VAS 시스템은 (i) 각 음향 근원에 대한 격리된 신호; (ii) 각 음향 근원의 위치에 대한 지식; (iii) 근원이 있는 만큼의 많은 채널들; (iv) 각 근원이 독립적으로 공간에 존재하기 위한 헤드 관련 전달 기능(HRTFs); 및 (v) 실내 에코 및 반향음의 효과에 근접하기 위한 부가적인 신호 조작이 요구된다.

스테레오 또는 서라운드 음향 레코딩과 같이 확성기를 통하여 들려지기 위한 레코딩에는 VAS 기술을 적용하는 것이 가능하다. 이러한 경우에는, 음향 근원(확성기)이 고정되고, 그들의 수와 위치가 알려진다. 레코딩은 분리된 채널을 제공하고, 음향 근원은 모의 실내에 위치한 모의 확성기이다. VAS 시스템은 이러한 음향 신호들을 마치 그들이 컴퓨터가 만들어낸 신호를 나타내는 것처럼 되게 한다. 실제로 이러한 방법으로 서라운드 음향 레코딩에 헤드 트래킹을 사용하는 상업적인 제품들(Sony MDR-DS8000 헤드폰)이 있다. 그러나 그러한 제품들이 할 수 있는 최선의 것은 헤드폰을 통하여 확성기를 듣는 경험을 재창조하는 것이다. 그들은 라이브 레코딩에 쉽게 적용할 수 없고, 총체적으로 원격회의에 부적절하다. 그들은 서라운드 음향 및 앰비소닉 시스템의 많은 문제점들을 그대로 가지고 있으며, 복합의 확성기에 대한 요구를 가진다.

두개 이상의 마이크로폰을 사용하여 라이브 공간 음향을 레코딩하고 재생하는 많은 방법들 역시 있다. 그러나, 우리는 헤드폰 재생을 위하여 디자인되고 청취자의 동적인 움직임에 대응하는 라이브 사운드를 캡처하는 오직 하나의 시스템에 대해서 알고 있다. 우리가 맥그래스 시스템이라 언급하는 그 시스템은 미국 특허 제6,021,206호 및 미국 특허 제6,259,795호에 기술되어 있다. 이 특허 사이의 근본적인 차이점은 전자는 단일의 청취자에 관한 것인 반면에, 후자는 복수의 청취자에 관한 것이라는 점이다. 이들 두 특허는 사운드필드 마이크로폰으로 만들어진 레코딩의 바이노럴 공간화에 관한 것이다(F. Rumsey, *Spatial Audio* (Focal Press, Oxford, 2001), pp. 204-205).

맥그래스 시스템은 (i) 음향이 녹음될 때, 청취자의 헤드의 방향이 알려지지 않고; (ii) 청취자 헤드의 위치가 헤드 트래커와 견주어 지며; (iii) 신호 프로세싱 과정이 멀티채널 레코딩을 바이노럴 레코딩으로 전환하는데 사용되고; 및 (iv) 그 주요 목적이 청취자가 그의 머리를 움직일 때, 위치가 변하지 않는 가상의 근원을 만드는 것을 특징으로 한다. 맥그래스 시스템에서 사용되는 것처럼 앰비소닉 레코딩은 청취자가 없을 때 청취자의 위치에서 발생하는 음장을 캡처하고; 청취자가 있을 때 청취자의 위치에서 음장을 캡처하지 않는 것을 시도한다는 것을 주목하라. 앰비소닉 레코딩은 두 귀간 도달 시간 차이, 두 귀간 도달 수치 차이 및 구형 헤드에 대하여 헤드 관련 전달 기능(HRTFs)에 의하여 도입되는 스펙트럼 변화에 대하여 직접적으로 캡처하지 않는다. 그러므로, 맥그래스 시스템은 복합적인 방향으로부터 들어오는 파장을 재구성하기 위한 녹음된 신호를 사용하고, 분리된 각각의 들어오는 파장을 공간화하기 위하여 HRTFs를 사용해야만 한다. 맥그래스 시스템은 개별화된 HRTF를 사용할 수 있음에도 불구하고, 그 시스템은 복잡하고 그 재구성은 앰비소닉과 관련된 모든 제한을 여전히 받는다.

## 발명의 상세한 설명

본 발명은 앞서 말한 많은 한계들을 극복하고 고정된 바이노럴 레코딩의 (a)머리 회전에 대한 가상의 청각 근원의 위치의 민감도; (b)중앙의 평면 구체화의 약함; 및 (c)심각한 앞/뒤 혼란의 존재와 같은 3가지 가장 심각한 문제점들을 해결한다. 게다가 본 발명은 한 명의 청취자 또는 동시에 듣고 있는 여러 명의 청취자에게 적용할 수 있고, 먼 거리의 청취 및 레코딩에 적용할 수 있다. 결국, 본 발명은 다음의 면에서 공간 음향을 레코딩하기 위한 "보편적인 포맷"을 제공한다. 공간 음향 기술(예를 들어, 스테레오, 4채널 방식, 돌비 6.1, 엠비소닉, 과장 합성 등)에 의하여 만들어진 음향은 본 발명의 형태로 전환될 수 있으며, 본래의 기술이 제공할 수 있는 동일한 공간 효과를 재생하기 위하여 계속적으로 재생될 수 있다. 그러므로 현존하는 레코딩의 실질적인 유산은 질에 있어서의 손실이 거의 없이 보존될 수 있다.

일반적인 용어로, 본 발명은 공간 음향의 동적인 3차원적인 특징을 캡처한다. 여기에서 "모션 트랙 바이노럴" 및 축약형인 "MTB"로 언급되는, 본 발명은 원격의 청취(예를 들어, 전화 통화) 또는 레코딩 및 재생에 사용될 수 있다. 사실상, MTB는 하나 또는 그 이상의 청취자로 하여금 음향이 발생하는(멀리서 듣는 것)공간 뿐만 아니라 음향이 발생하였던(레코딩)공간에서 그들의 귀를 두는 것을 허락한다. 게다가, 본 발명은 각 청취자로 하여금 청취 동안에 독립적으로 그들의 머리를 돌리는 것을 허락하며, 그 결과로서, 다른 청취자들은 다른 방향에서 그들의 머리를 방향지을 수 있다. 그렇게 하여, 본 발명은 헤드 모션의 지각적으로 중요한 효과를 정확하고 효율적으로 나타낸다. MTB는 음향이 발생하는(또는 발생한)공간에서 청취자의 귀를 효율적으로 위치시키고, 청취자의 헤드 모션과 동시에 가상의 귀를 움직이는 것에 의하여 높은 등급의 현실감을 성취한다.

이를 달성하기 위해서, 본 발명은 그 크기가 인간의 머리의 크기에 근접하는 표면에 위치한 복수의 마이크로폰을 사용한다. 단순함을 위하여, 마이크로폰이 탑재되어 있는 표면이 구형이라고 가정할 수 있다. 그러나 본 발명은 그렇게 제한되지 않으며, 다양한 다른 방법으로 실시될 수 있다. 마이크로폰은 균일하게 또는 비균일하게 표면을 덮을 수 있다. 게다가, 마이크로폰의 수는 조금 요구된다.

마이크로폰 배열은 청취자가 있고 싶어하는 청취 공간내 위치에서 전형적으로 놓여진다. 예를 들어, 원격회의를 위해서, 회의 테이블 중앙에 위치될 수 있다. 오케스트라 레코딩을 위해서, 콘서트 홀 내에 가장 좋은 자리에 위치될 수 있다. 홈 시어터를 위해서, 최첨단 영화에서의 최고의 자리에 위치될 수 있다. 마이크로폰에 의하여 캡처된 음향은 레코딩과 원격 청취에 대해서 다르게 다루어진다. 원격 청취 적용에서, 마이크로폰 신호는 청취자에게 직접적으로 보여지나, 레코딩 적용에서, 신호는 멀티 트랙 레코딩에 저장된다.

각 청취자는 그의 헤드 방향을 동적으로 측정하기 위하여 헤드 트래커를 갖추고 있다. 청취자의 머리와 동등체의 근원은 마이크로폰 배열의 동등체의 근원과 일치하는 것으로 항상 생각된다. 그러므로 청취자가 어떻게 이동할지라도, 음향 재생 시스템은 항상 청취자의 귀가 마이크로폰에 관계하여 위치되는 것을 안다. 본 발명의 한 태양에서, 그 시스템은 청취자의 귀에 가장 가까운 두 개의 마이크로폰을 가지고 있으며, 마이크로폰으로부터 청취자의 머리에 있는 한 쌍의 헤드폰으로 적당하게 확장된 신호를 발송한다. 음향 캡처와 함께, 재생 기구를 실시하는 많은 가능한 방법들이 있다. 특히, 오직 헤드폰 청취만이 기술되어 있음에도 불구하고, 헤드폰 대신 확성기를 사용하기 위하여 "혼선 해제(Crosstalk-Cancellation)" 기술을 사용하는 것 역시 가능하다는 것을 주목하여야 한다(G. Gardner, 3-D Audio Using Loudspeakers (Kluwer Academic Publishers, Boston, 1998), incorporated herein by reference).

더 바람직한 태양으로는, 더욱 정교한, 음향심리학적으로 기초한 신호 프로세싱 절차는 마이크로폰 신호의 계속적인 내삽을 허용하며, 그로 인하여, 작은 수의 마이크로폰으로, 청취자가 그의 머리를 움직일 때 발생하는 어떠한 "딸깍거리는 소리" 또는 다른 인공적인 소리를 제거한다.

본 발명의 한 양상에 따르면, 헤드 트래커는 머리를 회전하는 청취자에 대한 보상으로 신호 프로세싱을 변경하는데 사용된다. 간단히, 청취자가 수평면에 대하여 각  $\theta$ 를 통하여 그의 머리 방향을 돌린다고 가정하고, 청취자의 두 귀 중 구체적인 하나에 보내지는 신호를 생각해보자. 한 태양에서, 신호 프로세싱 유닛은 항상 청취자의 귀의 위치에서 가장 가까운 마이크로폰을 사용하기 위하여 마이크로폰 사이를 바꾸는 각  $\theta$ 를 사용한다. 다른 태양에서, 신호 프로세싱 유닛은 가장 가까운 마이크로폰과 다음으로 가까운 마이크로폰으로부터 신호를 내삽하거나 "뺀"하기 위하여 각  $\theta$ 를 사용한다. 또한 다른 태양에서는 신호 프로세싱 유닛은 가장 가까운 마이크로폰과 다음으로 가까운 마이크로폰으로부터 신호를 결합하기 위한 각  $\theta$ 를 교환하는 직선 필터링 절차를 사용한다. 이러한 세번째 태양에서, 아래에 기술되는 보완적인 신호는 현실적인 마이크로폰 또는 실질적인 마이크로폰의 출력을 결합시키는 가상의 마이크로폰으로부터 얻어진다. 한 태양에서, 보완적인 신호는 마이크로폰 배열에 있는 것과는 성질이 다르지만, 같은 음장 내에 위치하는 부가적인 마이크로폰으로부터 얻어진다. 다른 태양에서, 보완적인 신호는 배열 마이크로폰의 특정한 하나로부터 얻어진다. 다른 태양에서, 보완적인 신호는 배열된 마이크로폰 사이의 동적인 변화에 의하여 얻어진다. 다른 태양에서, 보완적인 신호는 동적으로 변화된 배열 마이크로폰의 출력의 스펙트럼 내삽에 의하여 얻어진다. 또 다른 태양에서는, 단일의 보완하는 신호에 대한 위에서 서술된 방법 중의 하나를 사용하여 왼쪽 귀에 하나 그리고 오른쪽 귀에 하나의 두 개의 보완하는 신호가 얻어진다.

본 발명의 한 태양에 따라서, 음향 재생 기구는 음성 출력 장치에 관한 출력기와 청취자의 머리 움직임을 나타내는 신호를 제공하기 위하여 배열된 헤드 트래킹 장치에 관한 입력기를 가지는 신호 프로세싱 유닛을 포함한다. 그 신호 프로세싱 유닛은 청취자의 머리가 음장 내 그리고 마이크로폰의 위치에 위치되어 있다면, 청취자의 귀의 가능한 위치를 나타내는 지점에서 음장을 샘플하기 위하여 위치된 다수의 마이크로폰의 출력을 나타내는 신호를 받기 위하여 배열된다. 그 신호 프로세싱 유닛은 나아가 마이크로폰 출력 신호를 선택하고 헤드 트래킹 장치에 의하여 지시된 것처럼 청취자의 머리의 움직임에 대응하여 하나 또는 그 이상의 선택된 신호를 음성 출력 장치로 나타내기 위하여 배열된다. 음성 출력 장치와 헤드 트래킹 장치는 신호 프로세싱에 선택적으로 직접 연결될 수 있거나 무선으로 될 수 있다.

본 발명의 다른 양상에 따르면, 헤드 트래킹 장치에 의하여 지시된 것처럼 청취자의 머리의 회전에 대응하여, 신호 프로세싱 유닛은 청취자의 머리가 음장 내에 위치되어 있다면 음장 내 청취자의 귀의 위치와 관련하여 복수의 마이크로폰 중에서 가장 가까운 마이크로폰과 다음으로 가까운 마이크로폰으로부터의 출력을 나타내는 신호를 결합하고, 결합된 출력을 음성 출력 장치에 나타내기 위하여 배열된다.

본 발명의 다른 양상에 따르면, 신호 프로세싱 유닛은 각각의 마이크로폰 출력 신호와 연관된 저역 통과 필터 및 청취자의 왼쪽 귀에 대하여 결합된 출력 신호 및 청취자의 오른쪽 귀에 대하여 결합된 출력 신호를 만드는 저역 통과 필터의 출력을 결합하기 위한 서머와 같은 수단을 포함하며, 상기 각 결합된 출력 신호는 청취자의 머리가 음장 내에 위치되어 있다면 음장 내에 청취자의 귀의 위치에 관련하여 가장 가까운 마이크로폰과 다음으로 가까운 마이크로폰으로부터의 출력을 나타내는 출력을 나타내는 포함한다.

본 발명의 다른 양상에 따르면, 신호 프로세싱 유닛은 음장 내에 위치한 실제 또는 가상의 보완적인 마이크로폰으로부터의 출력을 제공하기 위하여 배열된 고역 통과 필터 및 청취자의 오른쪽 귀에 대한 결합된 출력 신호 및 청취자의 왼쪽 귀에 대한 결합된 출력 신호를 가진 고역 통과 필터로부터 출력 신호를 결합하기 위한 서머와 같은 수단을 포함한다. 한 태양에서는 동일한 고주파 신호가 양쪽 귀에 대하여 사용된다. 다른 태양에서는 오른쪽 귀 고역 통과 필터는 음장 내에 위치한 오른쪽 귀 실제 또는 가상의 보완적인 마이크로폰으로부터의 출력을 제공하기 위하여 배열되고, 왼쪽 귀 고역 통과 필터는 음장 내에 위치한 왼쪽 귀 실제 또는 가상의 보완적인 마이크로폰으로부터의 출력을 제공하기 위하여 배열된다. 이 후자의 태양에서 오른쪽 귀 고역 통과 필터로부터 출력 신호는 청취자의 오른쪽 귀에 대한 결합된 출력 신호와 결합되고, 왼쪽 귀 고역 통과 필터로부터 출력 신호는 청취자의 왼쪽 귀에 대하여 결합된 출력 신호와 결합된다.

본 발명의 다른 양상에 따르면, 동적인 바이노럴 음향 캡처 및 재생 장치는 청취자의 머리가 음장 내에 위치되어 있다면 청취자의 귀의 가능한 위치를 나타내는 지점에서 음장을 샘플링하기 위하여 위치된 복수의 마이크로폰을 포함한다. 신호 프로세싱 유닛은 커뮤니케이션 링크 또는 마이크로폰 신호가 녹음된 해독 및/또는 재생 수단을 거쳐 전달되는 신호를 통하여 마이크로폰으로부터 직접 마이크로폰 신호를 받을 수 있다.

본 발명의 목적은 현재 기술을 훨씬 초과하는 현실감 있는 음향 재생을 제공하는 것으로, 다시 말해서, "당신이 거기에 있다."라고 하는 현실감을 제공하는 것이다. 본 발명의 다른 목적은 음향 캡처, 저장 또는 전달 및 재생을 위하여 비교적 적당한 부가의 복잡성을 달성하려는 것이다.

본 발명의 더 나아가는 목적 및 양상은 명세서의 다음 부분에서 개시될 것이고, 상세한 설명은 제한없이 발명의 더 바람직한 태양을 충분히 발표할 목적이다.

도면을 참조하건대, 본 발명은 도 1 내지 도 11에 일반적으로 나타난 장치 및 방법으로 구체화된다. 여기에서 기술된 것뿐만 아니라, 본 발명의 더욱 바람직한 태양은 (1)음향 캡처를 위하여 두개 이상의 마이크로폰을 사용(뒤에 언급하는 것과 같이 오직 두개의 마이크로폰으로 몇몇 유용한 효과가 성취될 수 있음에도 불구하고)하고; (2)청취자의 머리의 방향을 재기 위한 헤드 트래킹 장치를 사용하며; (3)마이크로폰의 출력을 선택적으로 결합시키기 위한 음향심리학적으로 기초된 신호 프로세싱 기술을 사용하는 것으로부터 보여질 것이다.

먼저 도 1 및 도 2를 참조하여, 본 발명에 따르는 바이노럴 동적 음향 캡처 및 재생 시스템 10의 한 태양이 보여진다. 그 보여지는 태양에서, 그 시스템은 복수의 마이크로폰 14을 가지는 원형의 마이크로폰 배열 12, 신호 프로세싱 유닛 16, 헤드 트래커 18 및 왼쪽 20 및 오른쪽 22 헤드폰과 같은 음성 출력 장치를 포함한다. 이들 그림에서 보여지는 마이크로폰 배열은 파노라마식 배열이라 불린다. 뒤에 검토될 것처럼, 전방향성의, 파노라마식의, 그리고 집중식의 적용이라 불리우는 세 종류의 적용이 있다. 오직 예시의 형태로, 본 발명은 다음의 설명에서 파노라마식 적용에 대해서 설명되어진다.

나타난 태양에서, 마이크로폰 배열 12은 청취자의 머리 24의 반경  $b$ 과 거의 동일한 반경  $a$ 을 갖는 원형 주위에 동일한 간격의 8개의 마이크로폰 14 (0 내지 7번으로 번호매겨짐)을 포함한다. 본 발명의 목적이 청취자에게 마이크로폰 배열의 위치에서 그가 실제로 존재하는(존재하였던)느낌을 주는 것이라는 점에서 높이 평가할 만하다. 그렇게하기 위해서, 마이크로폰이 위치되어 있는 원형 둘레는 청취자의 머리의 크기와 대략적으로 같다.

8개의 마이크로폰이 나타낸 태양에서 사용된다. 이에 관하여, 본 발명은 많은 수의 마이크로폰뿐만 아니라 두개의 마이크로폰으로도 그 기능을 할 수 있다. 그러나 오직 두개의 마이크로폰의 사용은 두 귀간 축에 가까운 음향 근원에 대한 최상의 효과를 만드는 8개의 마이크로폰 만큼 실감나는 경험을 만들 수는 없다. 그리고 더 많은 마이크로폰이 사용될 수 있으나 8개의 채널을 가진 레코딩 장치가 쉽게 이용 가능하기 때문에 8이 편리한 숫자이다.

이러한 8개의 마이크로폰에 의하여 만들어진 신호는 왼쪽 20과 오른쪽 22 헤드폰에 보내지는 2개의 신호를 만들기 위하여 신호 프로세싱 유닛 16에서 결합된다. 예를 들어, 도 1에서 보여지는 방향에서 청취자의 머리로, 마이크로폰 #6으로부터의 신호는 왼쪽 귀로 보내질 것이고, 마이크로폰 #2로부터의 신호는 오른쪽 귀로 보내질 것이다. 이것은 표준의 바이노럴 레코딩이 행하는 것과 본질적으로 동일하다.

청취자가 그의 머리를 각  $\theta$ 만큼 회전한 도 2에 그려진 상황을 생각해 보자. 이 각은 헤드 트래커 18에 의하여 감지되고 신호 프로세싱을 변형하는데 사용되어진다. 헤드 트래커는 상업적으로 이용가능하고 헤드 트래커에 대한 자세한 사항은 서술되지 않을 것이다. 헤드 트래커가 회전 운동을 나타내는 출력 신호를 만든다는 것을 아는 것으로 충분하다. 각  $\theta$ 가 정확히  $45^\circ$ 의 배수라면 신호 프로세싱 유닛 16는 단지 청취자의 귀에 맞는 한 쌍의 마이크로폰을 선택할 것이다. 예를 들어 각  $\theta$ 가 정확히  $90^\circ$ 라면 신호 프로세싱 유닛 16는 마이크로폰 #0으로부터의 신호를 왼쪽 귀에, 마이크로폰 #4로부터의 신호를 오른쪽 귀에 보낼 것이다. 다시 말해서, 신호 프로세싱 유닛 16는 도 1에 나타난 "머리 직선"에 대응하는 마이크로폰을 통하여  $90^\circ$ 반시계방향 회전에 대응하는 위치를 갖는 마이크로폰 쌍을 선택할 것이다. 그러나 일반적으로  $\theta$ 는 정확히  $45^\circ$ 의 배수가 아니며, 신호 프로세싱 유닛 16는 아래 기술된 것처럼 헤드폰에 대한 신호를 제공하기 위해서 마이크로폰 출력을 결합해야 한다.

헤드 트래커는 기준 방향에 비교하여 청취자의 머리 방향에 있어서의 변화를 나타내는 신호를 제공한다는 점에서 높이 평가할 만하다. 방향은 대개 3개의 오일러 각(피치, 롤 및 요)에 의하여 나타내어지나, 다른 각 좌표 역시 사용될 수 있다. 측정은 바람직하게는 초당 100번과 같이 높은 샘플링 속도로 만들어지나, 다른 속도가 사용될 수도 있다. 기술여지지 않고, 회전하지 않으며, 직선 앞으로 정의하는 기준 방향은 전형적으로 프로세스의 초기에 준비되어질 것이나, 원할 때면 언제나 청취자에 의하여 변경될 수 있다. 도 1을 참조하여, 청취자의 왼쪽 귀가 마이크로폰 #6의 위치에 있고, 청취자의 오른쪽



쪽 귀가 마이크로폰 #2의 위치에 있다고 하자. 그 후 청취자가 회전하지 않고 걷는다면, 청취자의 위치(및 청취자의 귀의 xyz위치)는 음향 재생에 영향이 없을 것이다. 반면에, 청취자가 그의 머리를 회전한다면, 본래는 항상 청취자의 머리의 중심에 있고 그 방향이 결코 변하지 않는 대응 시스템에서의 초기 위치에 대하여 그의 귀의 위치를 바꾸는 것에 의하여, 신호 프로세싱 유니트 16는 도 2에서 나타낸 것처럼 방향에서의 변화를 보상할 것이다.

일반적으로, 청취자가 이동할 때, 움직임의 병진하고 회전하는 구성성분이 있다. MTB 시스템은 병진의 구성 성분을 무시하는 점에서 높이 평가할 만하다. 청취자의 머리의 중심은 MTB 마이크로폰의 중심과 일치하는 것으로 항상 평가된다. 그러므로 청취자가 어떻게 움직이더라도, 헤드 트래커 18에 의하여 제공되는 신호들은 신호 프로세싱 유니트 16로 하여금 마이크로폰에 대하여 청취자의 귀의 "위치"가 어디인지를 항상 알도록 허락한다. 용어 "위치"는 공간 내에서 한 점의 절대적인 위치를 의미하는 것으로 종종 이해되는 반면에(예를 들어, 몇몇 정의된 기준 축에서 xyz 좌표), 본 발명의 MTB 시스템은 청취자의 귀의 절대적인 위치를 알 필요가 없고 오직 그들의 상대적인 위치만 필요하다는 것을 기억하는 것이 중요하다.

신호 프로세싱 유니트 16가 머리 회전을 나타내기 위한 마이크로폰 신호를 어떻게 결합하는지를 기술하기에 앞서, 도 1 및 도 2가 신호 프로세싱 유니트 16으로 직접 들어가는 마이크로폰 출력을 나타낸다는 것이 중요하다. 그러나, 이 직접적인 연결은 오직 설명의 목적으로 보여지며, 실제의 사용되는 실제의 구성을 반영할 필요는 없다. 예를 들어, 도 3은 원격지 간회의의 구성을 나타낸다. 나타낸 태양에서, 마이크로폰 출력은 신호를 커뮤니케이션 링크 30를 통하여 멀리 위치된 디멀티플렉서/리시버로 전달하는 멀티플렉서/트랜스미터 유니트 26에 공급된다. 그 커뮤니케이션 링크는 무선 링크, 광학상의 링크, 전화 링크 등이 될 수 있다. 그 결과는 마이크로폰으로부터 포착된 음향을 마치 청취자가 마이크로폰 위치에 실제로 있는 것처럼 경험한다는 것이다. 반면에, 도 4는 레코딩 구성을 나타낸다. 나타낸 태양에서, 마이크로폰 출력은 디스크, 테이프, 메모리 카드, CD롬등과 같은 저장 매체 34에 레코딩을 저장하는 레코딩 유니트 32로 공급된다. 뒤의 재생을 위하여, 저장 수단은 신호 프로세싱 유니트 16에 공급되는 컴퓨터/재생 유니트 36에 의하여 접근되어진다.

그러므로, 보여질 수 있는 것처럼, 신호 프로세싱 유니트 16는 음성 입력을 요구하고, 입력은 책, 무선 입력, 광학상 입력, 배선에 의한 연결 등 어떠한 전통적인 형태일 수 있다. 음성 출력뿐만 아니라 헤드 트래커 18에 대한 입력도 같다고 여겨진다. 그러므로 신호 프로세싱 유니트 16와 다른 장치들 사이의 연결 및 여기에서 사용된 것과 같은 용어 "입력" 및 "출력"은 어떤 특정한 형태로 한정되지 않는다는 점이 높이 평가할 만하다.

도 5 내지 도 7을 참고하여, 우리는 이제 본 발명에 따르는 마이크로폰 신호를 결합하는 다른 절차를 기술한다. 간단함을 위하여, 그 기술은 오직 한 귀에 대하여 행해지며, 같은 절차가 필요한 변경을 가하여 다른 귀에 적용된다고 이해하면 된다. 이러한 절차의 각각은 다른 상황에서 유용하며, 각각은 차례로 검토된다.

하나의 절차 100은 도 5에서 보여지며, 여기에서 절차 1로 언급된다. 이 절차에서, 신호 프로세싱 유니트 16는 청취자의 귀의 위치에 가장 가까운 마이크로폰을 항상 사용하면서, 마이크로폰을 바꾸기 위하여 각  $\theta$ 를 사용한다. 이것은 가장 간단한 절차이다. 그러나, 작은 머리 움직임에는 둔감하며, 각각은 성취도를 낮추거나 많은 수의 마이크로폰을 요구하고, 그로 인해 복잡성을 증가시킨다. 게다가, 스위칭은 딸각거리는 소리가 들리는 것을 방지하고자 정교한 필터링과 결합되어야 할 것이다. 머리 방향이 스위칭 경계선을 가로질러 앞으로 움직일때 발생할 수 있는 딸각거리는 소리는 표준의 이력 현상 스위칭 기술을 사용하는 것에 의하여 제거될 수 있다.

다른 절차 120은 도 6에서 보여지며, 여기에서는 절차 2로 언급된다. 이 절차에서, 신호 프로세싱 유니트 16는 가장 가까운 마이크로폰 및 다음으로 가까운 마이크로폰으로부터의 신호 사이에 내삽하거나 "팬"하기 위한 각  $\theta$ 를 사용할 것이다. 마이크로폰 사이를 팬하는 절차 2는 작은 머리 움직임에 민감하고, 몇몇 적용에 있어서 적합하다. 그것은 두 확성기 사이에 가상 근원을 만들기 위한 넓게 팬된 스테레오 레코딩에서 개발된 것과 동일한 원칙에 필수적으로 기초한다(B. J. Bauer, "Phasor analysis of some stereophonic phenomena," J. Acoust. Soc. Am., Vol. 33, No. 11, pp. 1536-1539 (Nov., 1961)). 이러한 원칙을 수학적으로 표현하기 위하여  $x(t)$ 를 시간  $t$ 에서 가장 가까운 마이크로폰에 의하여 포착된 신호라고 하고,  $x(t-T)$ 를 다음으로 가까운 마이크로폰에 의하여 포착된 신호라고 하며,  $T$ 는 하나의 마이크로폰으로부터 다른 마이크로폰으로 음파가 전달되는데 걸리는 시간이다. 간단함을 위하여, 우리는 탐재 표면 주위에서 입사파의 회절로 인한 파형에 있어서의 어떠한 변화를 무시한다. 이러한 변화는 마이크로폰이 알맞게 다른 것의 근처에 있다면 비교적 작을 것이다.

$x(t)$ 가 주파수  $f_{\max}$  이상으로 주파수를 포함하지 않는다면, 시간 지연  $T$ 가  $1/4f_{\max}$  보다 작다면, 그리고 계수  $w$ 가 0에서 1사이라면, 그것은  $(1-w)x(t) + wx(t-T) \approx x(t-wT)$ 로 나타낼 수 있다. 그러므로 귀로의 광선과 가장 가까운 마이크로폰으로의 광선 사이의 각에 따라 패닝 계수  $w$ 를 바꾸는 것에 의하여, 신호의 시간 지연이 두 마이크로폰으로부터 신호의 시간 지연 사이에 대응적으로 있는 신호를 얻을 수 있다.

절차 2에서 에러의 2가지 근원이 있다. 첫번째는  $T > 1/4f_{\max}$  일 때, 근사치가 벗어나는 것이다. 두번째는 두 마이크로폰의 출력이 직선상으로 결합되거나 "혼합될" 때면 언제나 발생하는 분광 착색이다.

신호에 발생하는 한계는 배열된 마이크로폰의 수  $N$ 으로 표현될 수 있다. 원형의 반경을  $a$ , 음향의 속도를  $c$  그리고 인접하는 두 마이크로폰 사이의 거리를  $d$ 라고 한다. 그러면  $d = 2a \sin(\pi/N) \approx 2\pi a/N$  이고,  $T$ 의 최대값은  $d/c$ 이기 때문에, 신호가  $f_{\max} \approx Nc/(8\pi a)$  이상의 중요한 분광학적 내용을 포함한다면 근사치는 벗어난다. ( $T=d/c$ 는 음향 근원이 두개의 마이크로폰을 결합하는 선에 위치되어있는 최악의 경우에 부합한다는 가정을 기억하라. 음향 근원으로서의 방향이 마이크로폰 사이의 선에 직각이라면, 파면은 마이크로폰에 동시에 도달하고, 에러가 발생하지 않는다. 그러나 최악의 경우는 혼한 일로 예를 들어, 근원이 정면으로 앞에 있고, 청취자가 그의 머리를 가장 가까운 마이크로폰 사이에서 귀가 중간 지점으로 위치되도록 회전할때 발생한다. 우리는  $T=d/c < 1/4f_{\max}$  인 환경이  $d$ 가  $1/4$  파장보다 짧은 환경과 동등하다는 것에 주목한다. 샘플링 이론은 우리가 마이크로폰으로 할 것은 공간에서 음향의 파형을 샘플링하는 것이고 근사치에서 벗어난 공간의 샘플링 간격이 너무 클 때 위신호의 결과로서 해석될 수 있다.)

수치  $a=0.0875\text{m}$ ,  $c=343\text{m/s}$ , 및  $N=8$ 을 사용하여, 우리는  $f_{\max} \approx 1.25\text{ kHz}$ 라는 것을 얻는다. 다시 말해서, 8개의 마이크로폰 배열과 함께,  $1.25\text{ kHz}$ 를 넘는 큰 분광학적인 내용물이 있다면 믹싱은 적절하게 연기된 신호를 만드는데 실패할 것이다. 이러한 제한은 마이크로폰 사이의 거리를 줄이는 것에 의하여 일어날 수 있다. 두 마이크로폰의 출력이 선 모양으로 결합될 때, 도달 시간에서의 차이는 관찰할 수 있는 스펙트럼으로 콤팩트 패킷을 소개한다. 콤팩트 패킷의 가장 낮은 주파수 단계는  $f_0=c/2d$ 에서 발생한다. 다시  $d=2\pi a/N$ 이라 가정하고, 우리는  $f_0 \approx Nc/(4\pi a) \approx 2f_{\max}$ 을 얻는다. 우리는  $f_0$ 가 고주파수의 적어도 한 옥타브 위에 있는 것을 원하기 때문에 우리는 에러의 원인이 본질적으로 같은 환경을 이끌고 즉,  $f_{\max} \approx Nc/(8\pi a)$ 위에 중요한 분광학적인 내용물에 대한 요구가 없다. 표 1은 이러한 주파수가  $a=0.0875\text{m}$  및  $c=343\text{m/s}$ 일 때  $N$ 과 어떻게 차이가 나는지 보여준다.

신호가  $f_{\max}$  이상의 중요한 분광학적인 에너지를 가지지 않는다면, 절차 2는 훌륭한 결과를 만든다. 신호가  $f_{\max}$  이상의 중요한 분광학적 에너지를 가진다면, 그리고  $f_{\max}$ 가 충분히 높다면( $800\text{Hz}$ 이상), 절차 2는 여전히 수용가능할지도 모른다. 그 이유는 두 귀간 도달 시간 차이에 대한 인간의 민감도는 고주파수에서 감소하기 때문이다. 이것은 근사치에 있어서의 벗어남은 적절하게 멈춘다는 것을 의미한다. 분광 착색이 상당한 정도라는 것은 진실이다. 그러나 "고주파" 재생이 요구되지 않는 감독 또는 원격회의와 같은 적용에서, 절차 2의 단순성은 그것을 더 바람직한 선택으로 만들 수 있다.

세번째로, 전체적으로 더 바람직한 절차 140은 도 7에서 나타나고, 여기에서는 절차 3으로 언급한다. 이 절차에서, 신호 프로세싱 유닛 16는 가장 가까운 마이크로폰과 다음으로 가까운 마이크로폰으로부터의 신호를 결합하기 위하여 각  $\theta$ 로 변화하는 선모양의 필터링 절차를 사용한다.

절차 3은 음향심리학적으로 동기가 된 선모양의 필터링을 사용하는 신호를 결합한다. 공간 샘플링에 의하여 유발된 문제를 해결하는 적어도 두가지 방법이 있다. 하나는 공간 샘플링 속도를 증가시키는 것이다; 다시 말해서, 마이크로폰의 수를 증가시키는 것이다. 다른 것은 마이크로폰 신호를 결합하기 전에 반위신호 필터를 적용하고, 어떻게 해서든지 고주파수를 저장하는 것이다. 뒤의 접근은 절차 3의 더 바람직한 태양이다.

절차 3은 인간이 고주파수 두 귀간 도달 시간 차이에 민감하지 못하다는 사실을 이용한다. 사인 곡선에서, 두 귀간 상태 민감도는  $800\text{ Hz}$ 이상의 주파수에서 빠르게 떨어진다. 그리고  $1.6\text{ kHz}$ 이상에서는 무시해도 좋다(J.Blauer, Spatial Hearing (Revised Edition), p.149 (MIT Press, Cambridge, MA, 1996) 레퍼런스에 의하여 여기에서 추가됨). 도 8 및 도 9뿐만 아니라 도 7을 언급하면서, 다음은  $N$  마이크로폰 배열, 이 태양에서는  $N=8$ ,에 대한 절차 3과 관련된 프로세싱 과정의 예이다.

1. 블록 142에서,  $x_k(t)$ 는  $k=1, \dots, N$ 인 마이크로폰 배열에서  $k^{\text{th}}$  마이크로폰의 출력이다.

2. 블록 144에서, 대략  $1.0$  내지  $1.5\text{ kHz}$ 사이의 범위에서의 차단주파수  $f_c$  이상에서 날카로운 롤오프를 가지는 저역 통과 필터와의 배열에서  $N$  마이크로폰(예를 들어, 이 태양에서는 8개의 마이크로폰) 각각의 출력을 필터한다.  $y_k(t)$ 는  $k=1, \dots, N$  이고,  $k^{\text{th}}$ 저역 통과 필터의 출력이다.

3. 블록 146에서, 절차 2에서처럼 저역 통과 필터  $z_{LP}(t)$ 를 만들기 위하여 이러한 필터들의 출력을 결합시킨다. 예를 들어, 오른쪽 귀 신호를 고려한다. 오른쪽 귀 28에 대한 광선 30과 가장 가까운 마이크로폰  $14_{\text{closest}}$ 에 대한 광선 32사이의 각을  $\alpha$ 라고 하고, 두 인접 마이크로폰에 대한 광선 사이의 각을  $\alpha_0$ 라고 한다; 예를 들어, 마이크로폰  $14_{\text{closest}}$  및 마이크로폰  $14_{\text{next-closest}}$  가장 가까운 마이크로폰  $14_{\text{closest}}$ 에 대한 저역 통과 필터 200의 출력을  $y_{\text{closest}}(t)$ 라고 하고, 다음으로 가까운 마이크로폰  $14_{\text{next-closest}}$ 에 대한 저역 통과 필터 202의 출력을  $y_{\text{next}}(t)$ 라 한다. 그리고 오른쪽 귀에 대한 저역 통과 출력은  $Z_{LP}(t)=(1-\alpha/\alpha_0)y_{\text{closest}}(t)+(\alpha/\alpha_0)y_{\text{next}}(t)$ 로 주어진다. 왼쪽 귀에 대한 저역 통과 출력은 비슷하게 만들어지며, 왼쪽 귀 신호에 대한 신호 프로세싱 요소는 위에서 기술된 것의 두배이기 때문에 그들은 명쾌함을 위하여 도 9로부터 생략된다.

4. 블록 148에서, 우리는 보완적인 마이크로폰 300을 소개한다. 보완적인 마이크로폰의 출력  $x_c(t)$ 은 보완적인 고역 통과 필터 204로 필터된다. 이러한 고역 통과 필터의 출력을  $Z_{HP}(t)$ 라고 한다. 보완적인 마이크로폰은 분리된 마이크로폰, 배열내에 있는 마이크로폰의 하나, 또는 배열 내에 마이크로폰의 출력을 결합시키는 것에 의하여 만들어지는 "가상의" 마이크로폰일 수 있다. 게다가, 다른 보완적인 마이크로폰은 왼쪽 귀와 오른쪽 귀에 대하여 사용될 수 있다. 보완적인 마이크로폰(들)의 다양한 대체 태양들과 그들의 장점 및 단점이 아래에서 검토된다.

5. 다음으로, 블록 150에서, 고역 통과 필터된 보완적인 신호의 출력은 저역 통과 내삽 신호에 첨가되고, 그 결과적인 신호  $z(t)=z_{LP}(t)+z_{HP}(t)$ 는 헤드폰으로 보내진다. 다시한번, 오른쪽 귀와 왼쪽 귀에 대한 신호는 분리되어 프로세스되어야 하는 것이 관찰된다. 일반적으로  $z_{LP}(t)$ 는 왼쪽과 오른쪽 귀에 대하여 다르다. 아래 대안들 A, B 및 C에 대하여, 신호  $z_{HP}(t)$ 는 두 귀에 대하여 동일하나, 대안 D에 있어서는 다르다.

위에 기술된 신호 프로세싱이 신호 프로세싱 유닛 16에 의하여 수행되어지고, 전통적인 저역 통과 필터, 고역 통과 필터, 애더 및 다른 신호 프로세싱 요소가 사용될 것이라는 점이 높히 평가할 만하다. 부가적으로, 신호 프로세싱 유닛 16는 컴퓨터와 신호 프로세싱을 수행하는 관련 프로그래밍을 포함한다.

절차 3이 훌륭한 결과를 만든다는 것을 주목할 만하다. 절차 1 및 절차 2보다 실행하기에 더 복잡함에도 불구하고, 이 절차가 전체 스펙트럼의 영역을 포함하는 신호를 만들기 때문에 높은 충실도의 재생으로 더 바람직한 태양이다.  $f_c$  이상에서

스펙트럼 구성 성분에 대한 두 귀간 도달 시간 차이(ITD)가 제어되지 않음에도 불구하고, 인간의 귀는 그 주파수 이상의 상태에서는 둔감하다. 반면에 오른쪽/왼쪽 방향에서 음향에 대한 정확한 일시적인 위치 신호에 이르게 되어,  $f_c$  이하의 ITD는 정확할 것이다.

$f_c$  이상에서, 두 귀간 도달 수준 차이(ILD)는 가장 중요한 위치 신호를 제공한다. 고주파ILD는 보완적인 마이크로폰 신호가 어떻게 얻어지느냐에 정확하게 의존한다. 이는 지금 검토될 마이크로폰의 실질적인 설치 및 배치 후에 검토될 것이다.

앞서 언급한 바와 같이, 마이크로폰 배열 내에 마이크로폰은 다른 방법으로 실질적으로 설치될 수 있다. 예를 들어, 딱딱한 철사 또는 로드로 지지하는 것에 의하여 공간에서 효과적으로 매달리게 할 수 있고, 단단한 구형의 표면에 설치될 수 있으며, 또는 단단한 타원체, 절두 실린더, 8각형의 박스와 같은 수직의 축에 대한 회전 표면에 설치될 수 있다.

위에서 기술된 태양이 마이크로폰의 배열을 사용하나, 균일하게 마이크로폰을 간격지을 필요는 없다.

본 발명에 따르면, 우리는 역시 전방향성의, 파노라마식의, 그리고 집중식의 적용이라 불리우는 세 단계의 적용을 구별한다. 그러므로 기술된 태양은 파노라마식 적용의 배경에 있었다.

전방향성의 적용에서, 청취자는 바람직한 방향을 가지지 않고, 마이크로폰은 전체 공간에 걸쳐 균일하게 간격지어져 있다(도시되지 않음). 위에서 기술된 것처럼 파노라마식 적용에서, 청취자의 머리의 수직축은 대개 수직으로 존재하지만, 청취자는 어떠한 방향에 직면하기 위하여 회전하기를 원할 것이다. 여기에서 마이크로폰은 위에서 나타난 것처럼 수평면의 원형 주위에 바람직하게는 균등하게 간격지어진다. 집중식의 적용에서(콘서트, 극장, 영화, 텔레비전 또는 컴퓨터 모니터 보기에 의하여 예시됨), 사용자는 강하게 선호되는 방향을 가진다. 여기에서 마이크로폰은 필요한 마이크로폰의 수를 감소시키거나 더 높은 차단주파수의 사용을 허락하기 위해서 도 10에서 나타난 것처럼 예상되는 귀의 위치 주위에 더 밀도있게 간격지워질 수 있다.

이러한 적용은 대안들 각각과 마이크로폰 배치 및 설치 표면은 다른 내부 마이크로폰 시간 지연 및 다른 분광 착색을 만들 것이다. 특히 자유-공간 매달림은 더 넓은 반경의 요구를 이끌면서, 표면-설치 형태보다 더 짧은 시간 지연을 이끈다. 표면 설치 대안에 있어서, 마이크로폰 픽업은 더이상 전방향성이 될 수 없다. 대신에 그것은 표면에 음향을 흠뻑하는 특성을 물려줄 것이다. 예를 들어, 구형의 표면 또는 절두 실린더 표면에서, 마이크로폰의 동측에 있는 근원에 대하여 고주파 반응은 저주파 반응보다 대략 6-dB 더 클 것이고, 반대측에 있는 근원에 대하여 고주파 반응은 설치 표면의 음향 가리움에 의하여 더욱 약해질 것이다. 설치 표면의 효과가 정확한 두 귀간 도달 시간 차이뿐만 아니라 정확한 두 귀간 도달 수치 차이를 캡처하기 위하여 이용될 수 있다는 것에 주목하자.

다른 설치 배치는 헤드 트래커에 대한 다른 요구를 이끈다는 것을 주목할 만하다. 예를 들어, 방위와 높이 모두 전방향성의 적용을 위하여 끌려가야 한다. 파노라마식 적용에 있어서, 관심의 대상인 음향 근원은 수평면 또는 그에 가까이 위치되어 있다. 이러한 경우에, 마이크로폰을 설치하기 위하여 어떠한 표면을 사용할지라도, 수평면의 원형 주위에 그것들을 위치시키는 것이 더 바람직할 것이다. 이것은 오직 방위각을 측정하는 더욱 간단한 헤드 트래커의 사용을 가능하게 한다.

지금까지 우리는 마이크로폰 배열이 정지한 상태에 있다고 암시적으로 가정하였다. 그러나 MTB 배열이 자전거, 모바일 로봇 또는 사람이나 동물에 설치될 수 없는 이유가 없다. 예를 들어, 헤드 밴드 또는 마이크로폰을 품은 칼라를 착용하고 있는 사람으로부터의 신호는 다른 사람에게 전달될 수 있고, 그는 이동하는 사람이 듣는 것을 들을 수 있다. 모바일 적용에서, MTB 배열에서 위치 트래커를 통합시키는 것이 유리할 수 있다. 배열이 옮겨지는 것뿐만 아니라 회전된다면 MTB 배열의 회전은 회전적으로 안정된 음향 이미지를 유지하기 위하여 청취자의 머리의 어떠한 회전과 결합될 수 있다.

우리는 설치 표면의 크기는 청취자의 머리의 그것과 가까워야 한다고 말하였다. 그러나, MTB의 수중 적용 역시 가능하다. 물에서 음향의 속도가 대략 공기에서 음향의 속도의 4.2배이기 때문에 설치 표면의 크기는 그에 따라서 정해진다. 그것은 중반부에 소개된 두 귀간 도달 시간 차이 및 두 귀간 도달 수치 차이에서의 변화에 대하여 정확할 것이다. 수면 아래 먼 거리의 청취에 있어서, 청취자는 땅 위에, 배 위에, 그리고 물 속에 있을 수 있다. 특히, 다이버는 그의 다이빙 헬멧에 MTB 배열을 가질 수 있다. 다이버는 물속에서 경험하는 자연스럽지 않은 조그만 두 귀간 도달 시간 차이 및 두 귀간 도달 수치 차이때문에 음향 근원의 위치화가 매우 어렵다는 것이 잘 알려져 있다. MTB 배열이 설치된 헬멧은 이러한 문제를 해결해 준다. 다이버가 유일한 청취자라면, 그리고 헬멧이 다이버의 머리와 함께 회전한다면, 두개의 마이크로폰을 사용하는 것으로 충분하며, 헤드 트래킹이 필요하지 않게 된다. 그러나 다이버가 듣는 것을 다른 사람도 듣기를 원한다면, 다수의 마이크로폰 MTB 배열이 필요하다. 마침내 다른 모바일 적용에서와 같이, 회전적으로 안정화된 음향 이미지를 유지하기 위하여 MTB 배열에 부착된 트래커를 사용하는 것이 바람직하다.

구형이 특히 전방향성의 적용에 대하여 이상적인 설치 표면으로 보임에도 불구하고, 다른 표면들이 실제로 더 바람직할 수 있다. 구형의 대칭성은 "브라이트 스팟"의 발생을 초래하고, 이는 음향 근원의 정반대의 측면 위에 부자연스럽게 강한 반응이다. 타원체 또는 절두 실린더는 약한 밝은 지점을 가진다. 실용적인 구조물 및 조립체 배치는 절두 실린더, 및 심지어는 직사각형, 육각형 또는 팔각형 박스가 선호된다. 그러나 간단히 설명하기 위하여, 아래 부분에 대해서는 마이크로폰 배열은 단단한 구형 위에 설치된 것으로 가정한다.

위에서 살펴본 바와 같이, 표면에 설치된 마이크로폰은 표면에 음향을 흠뻑하는 특성을 가진다. 그에 대한 반응으로 생겨나는 비등방성은 그것이 적절한 두 귀간 도달 수치 차이를 이끌기 때문에 실제로 마이크로폰 배열에서 더 바람직하다. 그러나 비등방성은 우리가 마이크로폰으로부터 음향 근원에 대한 방향에 독립적인 정보를 원한다면, 고주파 정보를 나르는 보완적인 마이크로폰에 대하여 문제를 낳는다. 이는 우리에게 절차 3에서 사용되는 보완적인 마이크로폰을 실시하는 대안적인 방법을 고려하게 한다.

보완적인 마이크로폰의 목적은 N 배열 마이크로폰 신호의 저역 통과 필터에 의하여 제거되는 고주파 정보를 회복시키는 것이다. 블록 152에서 나타난 것처럼 도 7B에서는, 이들 보완적인 마이크로폰 신호를 얻는 최소한 다섯가지 방법이 있으며, 그 자체의 장단점이 있다.



대안 A : 단독의 보완적인 마이크로폰 사용하기. 여기에서 단독의 마이크로폰은 고주파 신호를 포획하기 위하여 사용된다. 예를 들어, 구형의 꼭대기에 설치된 전방향성의 마이크로폰이 될 수 있다. 포획이 구형 아래에 있는 음향 근원에 대하여 구형에 의하여 가려지게 됨에도 불구하고, 수평면 내 음향 근원에 대한 균등한 적용을 제공한다.

#### 장점

(1) 개념적으로 단순하다.

(2) 대역폭 효율

보완적인 마이크로폰이 전체적인 오디오 대역폭(CD 질에 대한 22.05 kHz)을 요구함에도 불구하고, 배열 마이크로폰 N 각각은 오직  $f_c$ 의 대역폭을 요구한다. 예를 들어,  $N=8$ 이고  $f_c=1.5$  kHz라면 8개의 배열 마이크로폰은 오직 12 kHz의 대역폭을 함께 요구한다. 그러므로 전체적인 시스템은 정상적인 2 채널 스테레오 CD이상의 대역폭을 요구하지 않는다.

#### 단점

(1) 다른 채널의 요구

이는  $N=8$ 인 배열 마이크로폰의 약점으로, 8 트랙 레코더 및 8 채널 A/D 컨버터가 흔한 상업적인 생산물이지만, 여기에서는 9개의 채널이 필요하기 때문이다.

(2) 비등방성

물질의 보완적인 마이크로폰은 공간의 절반에 대하여 구형의 그림자에 두지 않을 수 있는 장소가 없다.

(3) 부정확한 ILD

같은 고주파 신호가 왼쪽과 오른쪽 귀에 대하여 사용될 때, 고주파 두 귀간 도달 수치 차이(ILD)가 없을 것이다. 이는 높지 않은 주파수 에너지를 가진 음향 근원에 대하여 문제를 일으키지 않을 것이나, 낮지 않은 주파수 에너지는 청취자의 머리의 중심에 위치되려는 경향이 있을 것이다. 게다가, 넓은 밴드 근원에 대하여 충돌하는 신호가 있을 것이다. 이는 전형적으로 위치화의 흐릿함을 증가시키고, "갈라짐 이미지"의 형성에 이른다; 다시 말해서, 두 가지 근원이 있다는 인식, 있어야 할 곳에 저주파 근원 및 머리의 중심부에 고주파 근원이 있다는 인식을 말한다.

대안 B : 배열 마이크로폰 중의 하나의 사용. 보완적인 마이크로폰처럼 배열 마이크로폰 중의 하나를 마음대로 선택하라.

#### 장점

(1) 개념적으로 간단하다.

(2) 대역폭 효율(대안 A와 같다)

(3) 부가적인 채널에 대한 필요가 없다.

#### 단점

(1) 수평면에서 근원에 대한 이방성.

어떤 마이크로폰이 보완적인 마이크로폰으로 선택되더라도, 반대쪽에 있는 근원에 대하여 구형의 음향 가리움에 있을 것이다. 이것은 집중식의 적용에 있어서 수용가능하거나 심지어 바람직함에도 불구하고 전방향성 또는 파노라마식의 적용에서는 수용할 수 없다.

(2) 부정확한 ILD. (대안 A와 동일하다.)

대안 C : 동적으로 변경된 배열 마이크로폰을 사용하기. 청취자의 코에 가장 가까운 마이크로폰을 선택하기 위하여 헤드 트래커 출력을 사용하라.

#### 장점

(1) 부가적인 채널에 대한 필요가 없다.

(2) 이방성 반응이 앞/뒤 구별에서 몇몇 부가적인 개발을 얻는데 사용될 수 있다. 뒤에서의 근원에 대한 머리 가리움은 "킷바퀴 가리움"을 피하기 위하여 몇 단계 치환할 것이다.

## 단점

(1) 더이상의 대역폭 효율이 없다. 어떠한 채널이 보완적인 채널에 사용될 것인지를 알 방법이 없기 때문에, N 채널의 모든 전체 오디오 대역폭에서 전달되고 녹음되어야 할 것이다. 그러나 대역폭 효율은 감독과 같은 단일의 사용자 적응을 위하여 유지될 수 있다. 이는 그 청취자를 위하여 필요한 하나의 전체 대역폭 채널은 마이크로폰에서 마이크로폰으로 동적으로 변경될 수 있기 때문이다.

(2) 대안 D에서 검토되는 것처럼, 스위칭 트랜지언트를 제거하기 위한 부가적인 신호 프로세싱을 요구한다.

(3) 부정확한 ILD(대안 A와 같다.)

대안 D : 두개의 동적으로 변경되는 배열 마이크로폰으로부터 가상의 보완적인 마이크로폰을 만든다. 이 선택은 오른쪽 귀와 왼쪽 귀에 대하여 다른 보완적인 신호를 사용한다. 어떠한 주어진 귀에 대하여, 보완적인 신호는 그 귀에 가장 가까운 두개의 마이크로폰으로부터 얻어낸다. 이는 저주파 신호가 얻어지는 방법과 매우 비슷하다. 그러나 두 마이크로폰(수용할 수 없는 콤팩트 분광 착색을 도입하는) 사이의 패닝 대신에, 우리는 항상 더 가까운 마이크로폰을 선택하면서 그들을 바꾼다. 이러한 방법에서 구현은 자동적으로 정확한 두 귀간 도달 수치 차이를 제공한다.

## 장점

(1) 부가적인 채널에 대한 필요가 없다.

(2) 정확한 ILD

## 단점

(1) 더이상의 대역폭 효율이 없다.( 대안 C와 같다.)

(2) 스위칭 트랜지언트를 감소시키기 위하여 부가의 신호 프로세싱이 요구된다.

(3) 스펙트럼에서의 변화가 들릴 수 있다. 신호가 갑작스럽게 막 변한다면, 청취자는 신호의 불연속에 의하여 만들어진 달각거리는 소리를 자주 듣게 된다. 이것은 특히 머리 위치가 본질적으로 스위칭 경계에 있고 작은 떨림이 머리로 하여금 스위칭 경계를 가로질러 앞뒤로 움직이는 것을 유발하는 것처럼 신호가 재빨리 앞뒤로 변한다면, 불쾌해질 것이다. 스위칭 트랜지언트의 결과로서 일어나는 재빠른 과정은 매우 불쾌한 "재잘거리는" 소리를 만들 것이다. 이러한 재잘거리는 소리가 나는 문제는 이력 현상을 도입하는 표준의 기술에 의하여 쉽게 해결된다; 즉 스위칭 경계는 교차되고, 스위칭 회로는 뒤로 스위칭 되기 전에 원래의 지역으로 다시 최소한의 각 운동을 요구한다. 한 마이크로폰으로부터 다른 마이크로폰으로 스위칭할 때 발생하는 피할 수 없는 불연속은 간단한 크로스 페이딩 기술에 의하여 감소될 수 있다. 즉시로 스위칭하는 것 대신에, 신호는 두번째 신호의 페이드 인 버전으로 첫번째 신호의 페이드 아웃 버전을 첨가하는 것에 의하여 얻어질 수 있다. 그 결과는 첫번째 신호가 페이드 아웃되고 두번째 신호가 페이드 인되는 시간 간격  $T_{fade}$  의 길이에 의존한다. 가상 실험은  $T_{fade}=10ms$  일 때 스위칭 트랜지언트가 매우 희미하고,  $T_{fade}=20ms$  일때 들리지 않는다는 것을 보여주었다. 이러한 숫자들은 헤드 트래커에 대한 데이터 비율과 매우 호환성이 좋으며, 전형적으로 대략 샘플들 사이에 10ms 내지 20ms에 있다. 그러나 특히 근원이 MTB 배열에 가까이 있을 때 가상의 보완적인 마이크로폰이 변화하는 것처럼 스펙트럼에서 변화를 들을 가능성이 여전히 있다.

대안 E : 두 배열 마이크로폰의 스펙트럼들 사이에 내삽하고 일시적인 신호를 재합성하는 것에 의하여 가상의 보완적인 마이크로폰을 만든다. 대안 D와 같이 이 선택은 오른쪽 귀와 왼쪽 귀에 대하여 다른 보완적인 신호를 사용하며, 어떠한 주어진 귀에 있어서, 보완적인 신호는 그 귀에 가장 가까운 두개의 마이크로폰으로부터 얻어진다. 대안 E는 귀에 가장 가까운 두개의 마이크로폰 사이를 변경하는 것보다 오히려 적절하게 내삽하는 것에 의하여 대안 D의 인지할 수 있는 스펙트럼 변화를 제거한다. 문제는 상 제거 효과 없이 마이크로폰 신호의 고주파 부분을 부드럽게 결합시키는 것이다. 고주파에서 상태에 대한 귀의 무감각을 발견하는 기초적인 해결책은 세 단계를 포함한다: (a) 각 마이크로폰으로부터의 신호에 대한 단시간 스펙트럼의 평가, (b) 스펙트럼 사이의 내삽, (c) 스펙트럼으로부터 일시적인 파형의 재합성 등이다. 스펙트럼 분석, 변형 및 재합성에 의한 신호 프로세싱의 대상은 신호 프로세싱계에서 잘 알려져 있다. 이 고전적인 방법은 (a) 푸리에 변환 분석 및 재합성, 및 (b) 필터 뱅크 분석 및 재합성을 포함한다.

## 장점

(1) 부가적인 채널에 대한 필요가 없다.

(2) 정확한 ILD

(3) 스위칭 트랜지언트 또는 스펙트럼 인공물이 없다.

## 단점

(1) 더이상의 대역폭 효율이 없다.(대안 C와 같다.)

## (2) 넓은 계산 요구.

이러한 다섯 대안의 태양들 중 바람직한 것에 대한 적절한 상황은 다음과 같이 요약된다: 대안 A는 대역폭 효과가 우세한 관계일 때 더 바람직할 수 있다; 대안 B는 집중식의 적용에 있어서 좋은 절충점이다; 대안 C는 대역폭에 대한 비용이 수용 가능하다면 먼거리의 청취(원격 회의)에서 매력적이다; 대안 D는 적은 계산 비용에서 대안 E에 가까운 수행을 한다; 그리고 대안 E는 최대의 현실감이 우세한 관계에서 바람직하다.

표 2는 대안 A 및 대안 D에 대한 절차 3뿐만 아니라 절차 1 및 절차 2의 장점 및 단점을 요약하였다.

MTB는 음장 내에 구형과 같은 표면을 삽입하고 청취자의 귀가 위치해 있을 장소 근처에 압력을 느끼는 것에 의하여 청취자의 귀에서 존재할 음장을 캡처하기 위한 시도를 한다는 점을 주목하자. 불충분한 근사치를 만드는데 두가지 방법이 있다.

### 1. 머리 크기의 미스매치.

구형이 청취자의 머리보다 더 작다면, 만들어진 두 귀간 도달 시간 차이는 청취자가 일반적으로 경험하는 것보다 더 작을 것이다. 역으로 구형이 청취자의 머리보다 더 크다면, 만들어진 두 귀간 도달 시간 차이는 정상보다 더 클 것이다. 고정된 위치 에러를 만드는 것에 더하여, 이것은 청취자가 그의 머리를 회전할 때 음향 근원의 위치의 불안정으로 이끈다. 구형이 청취자의 머리보다 더 작다면, 그 근원은 청취자와 약간 회전하는 것으로 나타날 것이고 반면에, 구형이 더 크다면 그 근원은 청취자의 움직임에 반대로 회전하는 것으로 나타날 것이다.

### 2. 컷바퀴 신호의 부재

외이 또는 컷바퀴는 귀고막에 도달하는 음향의 스펙트럼을 변화시키고, 이러한 변화는 방위 및 높이에 따라 다르다. 이러한 스펙트럼 변화는 근원의 높이를 판단하는데 특히 중요한 컷바퀴 신호를 만든다. 그들의 정확한 특성은 복잡하고 사람마다 매우 다르다. 그러나 기본적인 특성은 그 중앙 주파수가 높이와 함께 대칭적으로 변화하는 스펙트럼 노치이다. 스펙트럼 변화는 근원이 머리위를 통과할 때 최소가 된다. MTB 표면은 어떠한 컷바퀴를 포함하지 않기 때문에, 대응하는 스펙트럼 변화가 없다. 변화는 높은 고도에는 대응하지 않기 때문에 대부분의 청취자들은 근원이 그들의 실제 높이에 관계없이 약간 높이 있다고 인지한다.

일반적으로 이러한 두가지 문제를 완전하게 바로잡기 위해서 알려진 절차는 없다. 그러나 특별하지만 중요한 상황에서 유용한 방법들이 있다.

머리 크기의 미스매치는 청취자가 대개 한 방향에서 더 많이 또는 더 적게 보는 집중식의 적용에서 쉽게 바로잡을 수 있다. 구형의 반경을  $a$ 라 하고, 청취자의 머리의 반경을  $b$ 라 하고, 머리 회전 각을  $\theta$ 라 하자. 그러면, 정면에 위치한 근원의 외관 위치는 마이크로폰 데이터를 프로세싱할 때  $\theta$ 의 위치에서  $(b/a)\theta$ 를 사용하는 것에 의하여 안정화될 수 있다. 이러한 간단한 보정이 머리 회전의 작은 각에 대하여 잘 작용한다. 게다가 이 기술을 사용하기 위하여 청취자의 머리 반경을 측정하는 것이 필수적이지 않다. 오직  $\theta$ 의 자리에  $a\theta$ 를 사용할 것을 필요로 하고, 이미지가 안정화될 때까지 청취자로 하여금 계수  $a$ 를 조정하는 것을 허락한다.

관심의 대상인 음향 근원이 수평면에서 더 많이 또는 적게 있다면 컷바퀴 신호의 부재를 보정하는 것도 가능하다. 이러한 경우에, 컷바퀴 전달 기능에 접근하는 필터는 신호 경로에서 각 귀로 도입되고, 사용자는 음향 이미지가 수평면에 나타날 때까지 필터 파라미터를 조정하는 것이 허락된다.

앞서 말한 기술로부터, 본 발명에 대한 일반적인 개념은 (a) 모든 가능한 머리 방향에 대하여 귀의 위치 근처 지점에서 음장을 샘플링하기 위하여 복수의 마이크로폰을 사용하고, (b) 청취자의 귀로부터 각 마이크로폰까지의 거리를 결정하기 위하여 헤드 트래커를 사용하며, (c) 마이크로폰 출력을 저역 통과 필터하고, (d) 청취자의 귀의 위치에서 마이크로폰에 의하여 포착된 신호의 저주파 부분을 평가하기 위하여 저역 통과 필터된 출력을 선 모양으로 내삽(동등하게, 무게, 결합, "팬")하며, (e) 고주파 내용물을 재삽입한다. 이와 같은 일반적인 개념은 다양한 대체적인 방법으로 실시되고 확장될 수 있다. 다음은 그 대안들이다:

1. 매우 작은 수 또는 매우 큰 수의 마이크로폰을 사용한다. 작은 수의 마이크로폰은 저역 통과 필터의 차단 주파수가 적절하게 조절된다면 사용될 수 있다. 심지어 오직 두개의 마이크로폰으로 근원이 마이크로폰의 중앙 평면에 너무 가깝지 않은 동안 동적인 변경의 이점을 얻는 것이 가능하다. 대체적으로 많은 수의 마이크로폰이 경제적으로 사용될 수 있다면, 저역 통과 필터링 및 고주파 회복 단계가 제거될 수 있다. 충분한 마이크로폰으로 내삽 절차가 간단한 스위칭에 의하여 대체될 수 있다.

2. 구형의 전체 표면에 걸쳐 마이크로폰을 붙이고, 청취자의 방위뿐만 아니라 높이를 느끼기 위하여 헤드 트래커를 사용하는 것에 의하여 도 8에서 보여지는 구성을 일반화한다. 가장 가깝고 다음으로 가까운 마이크로폰은 더이상 수평면에 있을 필요가 없으며, 임의의 머리 회전이 허용될 수 있다.

3. 머리 아래에 인공의 토르소를 도입한다. 토르소에 의한 음향의 흠어짐은 높이 및 외형화를 돕는 부가적인 위치 신호를 제공한다. 토르소를 포함하는 것은 마이크로폰 배열을 더 크고 불편하게 만듦에도 불구하고, 그것은 특히 요구되는 적용에 있어서 정당화될 수 있다.

4. 원치 않는 음향 픽업을 버리거나 감소시키기 위하여 마이크로폰 배열에 의한 각각의 마이크로폰을 제자리에 놓는다. 이것은 원치 않는 음향이 오히려 높거나 오히려 낮은 위치에 있을 때와 MTB 표면이 절두 실린더일때 특히 좋다. 이 경우에, 각각의 마이크로폰은 마이크로폰의 수직의 컬럼에 의하여 제자리에 놓여질 수 있고, 그 출력은 수평면 밖에서 민감도를 감소시키기 위하여 결합될 수 있다.

5. 음향 방위 측정기로서 MTB를 사용한다. 두 개의 중심이 같은 MTB 배열, 예를 들어, 도 11에서 보여지는 바와 같이, 머리 크기 구형 402에 설치된 더 작은 배열의 마이크로폰 400 및 구형으로부터 확장된 단단한 로드 406에 설치된 더 큰 배열의 마이크로폰 404를 사용한다. 더 작은 MTB 배열은 일상적으로 사용되며, 청취자는 근원에 직면하기 위하여 회전한다. 그리고 청취자는 더 넓은 MTB 배열로 바꾼다. 청취자가 근원을 직접적으로 지시한다면, 그 근원의 이미지는 중심으로 나타날 것이다. 작은 머리 움직임은 이미지의 증폭된 움직임을 초래할 것이며, 이는 근원을 위치화시키는 것을 더 쉽게 만든다.

공간 음향을 레코딩하는데 많은 대안적인 방법들이 있으며, 특히 서라운드 음향 시스템이 인기가 있다는 점이 중요하다. 헤드폰을 통하여 존재하는 공간 음향 레코딩을 재생하기 위하여 본 발명을 사용할 수 있는 것은 바람직하다.

위에서 언급한 것과 같이, 첨단 기술 서라운드 음향 시스템의 "스위트 스팟"에 마이크로폰 배열을 위치시키는 것에 의하여 존재하는 레코딩을 재녹음하는 직접적인 접근이 있다. 이것은 최상의 청취 경험을 가진 청취자에게 제공할 수 있는 장점이 있다. 반면에, 과거의 상업적인 경험은 하나 이상의 포맷으로 같은 내용을 공중에게 보여주는 것은 바람직하지 못하다는 것을 보여준다.

대안의 접근은 모의 실내에서 모의의 마이크로폰 배열을 자극하기 위하여 모의의 확성기를 사용하는 것에 의하여 재녹음의 과정을 모의 실험하는 것이다. 가장 단순한 상황에서, 구형의 머리 모델(V. R. Algazi, R. O. Duda and D. M. Thompson, "The use of head-and-torso models for improved spatial sound synthesis," Preprint 5712.113th Convention of the Audio Engineering Society (Los Angeles, CA, Oct. 5-8, 2002, 참고 문헌)은 마이크로폰 배열 내에 있는 특정 마이크로폰이 가상의 확성기 각각으로부터 포착하는 신호를 추정하는데 사용될 수 있다. 더 큰 현실감을 위하여, 실내 모델이 실내 반향음의 효과를 모의 실험하기 위하여 사용될 수 있다(D. B. Begault, 3-D Sound for Virtual Reality and Multimedia (AP Professional, Boston, 1994), 참고 문헌). 이 신호 프로세싱 절차는 본래의 레코딩 포맷에 있는 신호를 우리의 MTB(모션 트랙드 바이노럴)포맷에 있는 신호로 변환하는 특별한 실시간 하드웨어를 쉽게 실시될 수 있다. 포맷 컨버터를 통하여 전통적인 재생 유니트로부터 신호를 보내는 것에 의하여 하나 또는 많은 청취자들이 헤드폰을 통하여 CD 또는 DVD를 들을 수 있으며, 머리 움직임에 대한 반응의 장점을 즐긴다.

MTB의 같은 장점은 가상의 음향 공간의 창조와 데이터의 공간 음향 디스플레이로 컴퓨터가 만들어낸 소리를 완벽하게 현실감있게 할 수 있다는 것이다. 요구되는 것은 모의 실험의 MTB 마이크로폰 배열에 의하여 캡처될 음향을 측정하는 것이다. 측정된 마이크로폰 신호는 실질적 마이크로폰으로부터 신호의 위치에 사용될 수 있으며 그 결과로서 하나 또는 많은 청취자들이 헤드폰을 통하여 가상의 소리를 들을 수 있고 머리 움직임에 대한 반응의 이점을 즐긴다. 생생한 현실의 마이크로폰, 녹음된 현실의 마이크로폰 및 모의 실험된 마이크로폰의 사용을 커버하기 위해서, 청구항에서, 우리는 현실의 마이크로폰에 의하여 포착된 신호, 현실의 마이크로폰으로부터 녹음된 신호 및 마이크로폰 출력을 "나타내는" 신호로써 모의 실험 마이크로폰으로 측정된 신호를 언급한다.

그러므로 보여지는 바와 같이, 본 발명의 더 바람직한 태양은 음향 캡처를 위하여 두개 이상의 마이크로폰을 사용하고; 청취자의 머리의 방향을 측정하기 위해서 헤드 트래킹 장치를 사용하며; 마이크로폰의 출력을 결합시키기 위하여 음향심리학적으로 기초된 신호 프로세싱 기술을 사용한다. 본 발명은 어떠한 자연스럽게 발생하는 음향실내 반향음을 포함하여)을 녹음하고, 가상의 소리의 근원에 대한 안정된 위치를 청취자에게 제공하는 소량의 고정된 수의 채널을 사용하여 정적인 바이노럴 레코딩의 주요 한계의 해결 즉, 청취자의 머리 움직임에 대한 독립; 양호한 정면 구체화; 및 적거나 거의 없는 앞/뒤 혼란의 능력을 가지고 있다. 본 발명의 더 나아가 라이브 음향의 레코딩을 다룬다. 라이브 음향으로 모든 음향 근원에 대한 독립의 신호를 얻거나 감각적으로 중요한 에코 및 반향음을 언급하지 않는 것은 어렵거나 불가능하며; 근원의 위치는 대개 알려지지 않는다. 게다가 본 발명으로 작고 고정된 수의 채널들이 있고; 대략의 HRTFs가 마이크로폰 배열에 의하여 자동적으로 만들어지며; 복잡한 실제 실내 에코 및 반향음이 자동적으로 캡처된다.

위에서의 기술이 많은 자세한 내용을 포함하고 있음에도 불구하고, 본 발명의 바람직한 태양의 일부에 대한 기술을 제공한 것만으로 본 발명의 범위를 제한해서는 안된다. 그러므로 본 발명의 범위는 기술계에서 분명한 다른 태양을 모두 포함하고, 본 발명의 범위는 첨부된 청구항 이외 어떠한 것에 의하여 제한되지 않으며, 특이한 구성 요소는 명백하게 지적되지 않는 한 "오직 하나"의 의미로 주장될 수는 없으며 오히려 "하나 또는 그 이상"의 의미를 가진다. 일반 기술계에서 알려진 위에서 기술된 바람직한 태양의 구성 요소들과 구조적, 화학적, 기능적으로 동일체들은 여기에서 참고 문헌에 의하여 기술되어 있으며, 본 청구항에 의하여 포함되는 것으로 주장된다. 게다가, 본 발명에 의하여 해결되는 것으로 보여지는 각각 및 모든 문제들을 나타내는 장치 또는 방법은 필수적이지 않으며 본 청구항에 의하여 포함된다. 게다가, 본 게시물에서의 어떠한 구성요소, 구성성분 및 방법 단계도 청구항에 명확하게 인용되어 있는지에 관계없이 공중에게 바쳐진다. 여기에서의 청구항요소는 문구 "means for"를 사용하여 명확하게 인용되지 않는다면, 어떠한 청구항 요소도 35 U.S.C. 112, 여섯번째 문단의 조항 하에 만들어지지 않는다.

표 1.

표 1

$N$	$f_{\max}$ (Hz)	$N$	$f_{\max}$ (Hz)
2	312		
3	468	10	1,560
4	624	11	1,716
5	780	12	1,872
6	936	13	2,028
7	1,092	14	2,184
8	1,248	15	2,340
9	1,404	16	2,496

표 2.

표 2

장 점	절 차 :	1	2	3A	3D
생생하게 현실적인 음향 재생		x	x	x	x
넓게 적용가능 * ( 재생 및 전화통신 )		x	x	x	x
주요 바이노럴 신호 (ITD,ILD)의 캡처		x	x		x
모든 거리에서 정확한 음향의 재생		x	x	x	x
앞/뒤 혼란의 제거		x	x	x	x
청취자 머리 움직임에 대한 정확한 반응			x	x	x
안정화된 음향 이미지의 생산			x	x	x
동시적인 청취자들의 유지		x	x	x	x
다른 캡처 기술에 대하여 보편적인 포맷의 제공		x	x	x	x
대역폭의 효율적인 사용		x	x	x	
레코딩을 하는데 특별한 기술이 요구되지 않음		x	x	x	x
집약되고 잠재적으로 값싼 재생 시스템		x	x	x	x

\* 게임, 라디오, 텔레비전, 활동 사진, 홈 시어터, 음악 레코딩,  
원격 회의, 감독, 가상의 현실성, 음성 시스템 평가,  
정신음향학적인 보고 등.

단 점

헤드 트래커의 요구	x	x	x	x
최상의 결과를 위한 헤드폰의 요구	x	x	x	x
근원의 높이를 정확하게 재생할 수 없는 가능성	x	x	x	x
완전한 대역폭을 위하여 많은 마이크로폰의 요구	x	x		
머리가 움직일 때 딸깍 소리가 날 수 있음	x			
상당한 분광착색이 날 수 있음		x		
광대역 근원으로 "갈라짐 이미지"를 만들수 있음			x	

## 도면의 간단한 설명

본 발명은 실례의 목적을 가진 다음 도면들을 참고로 더 자세히 이해될 것이다.

도 1은 본 발명에 따르는 동적인 바이노럴 음향 캡처 및 재생 시스템의 한 태양을 나타내는 개략도이다.

도 2는 헤드 트래킹을 도해하는 도 1에서 보여지는 시스템의 개략도이다.

도 3은 원격회의를 위하여 배열된 도 2에서 보여지는 시스템의 한 태양의 개략도이다.



도 4는 레코딩 및 재생을 위하여 배열된 도 2에서 보여지는 시스템의 한 태양의 개략도이다.

도 5는 본 발명에 따르는 헤드 트래킹의 방법의 첫번째 태양을 보여주는 도식이다.

도 6은 본 발명에 따르는 헤드 트래킹의 방법의 두번째 태양을 보여주는 도식이다.

도 7은 본 발명에 따르는 헤드 트래킹의 방법의 세번째 태양을 보여주는 도식이다.

도 8은 도 7에 나타난 방법에 따르는 헤드 트래킹을 보여주는 개략도이다.

도 9는 도 7 및 도 8에 나타난 헤드 트래킹의 방법과 관련이 있는 신호 프로세싱의 한 태양을 보여주는 도식이다.

도 10은 본 발명에 따르는 집중식 마이크로폰 구성을 나타내는 개략도이다.

도 11은 본 발명에 따르는 방향 탐지 마이크로폰 구성의 개략도이다.

## (57) 청구의 범위

### 청구항 1.

신호 프로세싱 유닛을 포함하되,

상기 신호 프로세싱 유닛은 음성 출력 장치에 연결되기 위한 출력을 가지고;

상기 신호 프로세싱 유닛은 헤드 트래킹 장치에 연결되기 위한 입력을 가지며;

상기 신호 프로세싱 유닛은 청취자의 머리가 마이크로폰의 위치에서 음장내에 위치된다면 청취자의 귀의 가능한 위치를 나타내는 지점에서 음장을 샘플하기 위하여 위치되는 다수의 마이크로폰의 출력을 나타내는 신호를 받기 위하여 배치되고;

상기 신호 프로세싱 유닛은 상기 마이크로폰 출력 신호를 처리하고 상기 헤드 트래킹 장치에 의하여 지시된 것과 같이 상기 청취자의 머리의 방향에 대응하여 상기 음성 출력 장치로 바이노럴 출력을 나타내기 위하여 배치된 것을 특징으로 하는 음향 재생 장치.

### 청구항 2.

제 1항에 있어서,

상기 신호 프로세싱 유닛은,

상기 청취자의 머리가 상기 마이크로폰의 위치에서 상기 음장 내에 위치되어 있다면, 상기 음장 내 청취자의 귀의 위치와 관련하여 가장 가까운 마이크로폰 및 다음으로 가까운 마이크로폰으로부터 출력을 나타내는 신호들을 결합하기 위하여 배치된 것을 특징으로 하는 장치.

### 청구항 3.

제 1항에 있어서,

상기 신호 프로세싱 유닛은,

상기 마이크로폰 출력 신호의 각각과 관련된 저역 통과 필터; 및 청취자의 귀에 대하여 결합된 출력 신호를 만들기 위하여 상기 저역 통과 필터의 출력을 결합시키는 수단을 포함하되,

상기 결합된 출력 신호는 상기 청취자의 머리가 상기 마이크로폰의 위치에서 상기 음장 내에 위치되어 있다면, 음장 내 청취자의 귀의 위치와 관련하여 가장 가까운 마이크로폰 및 다음으로 가까운 마이크로폰으로부터의 출력을 나타내는 신호의 조합을 포함하는 것을 특징으로 하는 장치.

### 청구항 4.

제 3항에 있어서,  
 상기 신호 프로세싱 유닛은,  
 상기 음장 내에 위치되는 보완적인 마이크로폰으로부터 출력 신호를 제공하기 위하여 배치된 고역 통과 필터; 및  
 청취자의 귀에 대하여 상기 결합된 출력 신호를 가지는 상기 고역 통과 필터로부터 상기 출력 신호를 결합시키는 수단을 더 포함하는 것을 특징으로 하는 장치.

## 청구항 5.

신호 프로세싱 유닛을 포함하되,  
 상기 신호 프로세싱 유닛은 음성 출력 장치에 연결되기 위한 출력을 가지고;  
 상기 신호 프로세싱 유닛은 헤드 트래킹 장치에 연결되기 위한 입력을 가지며;  
 상기 신호 프로세싱 유닛은 청취자의 머리가 마이크로폰의 위치에서 음장 내에 위치되어 있다면 청취자의 왼쪽과 오른쪽 귀의 가능한 위치를 나타내는 지점에 있어서 음장을 샘플하기 위하여 위치된 다수의 마이크로폰의 출력을 나타내는 신호를 받기 위하여 배치되고;  
 상기 신호 프로세싱 유닛은 마이크로폰 출력 신호를 결합시키고 헤드 트래킹 장치에 의하여 지시된 것과 같이 상기 청취자의 머리의 방향에 대응하여 상기 음성 출력 장치로 바이노럴 출력을 나타내기 위하여 배치된 것을 특징으로 하는 음향 재생 장치.

## 청구항 6.

제 5항에 있어서,  
 상기 신호 프로세싱 유닛은 상기 청취자의 머리가 상기 마이크로폰의 위치에서 음장 내에 위치된다면 상기 음장 내에 청취자의 왼쪽 귀의 위치와 관련하여 가장 가까운 마이크로폰과 다음으로 가까운 마이크로폰으로부터 출력을 나타내는 신호를 결합시키기 위하여 배치되고,  
 상기 신호 프로세싱 유닛은 상기 청취자의 머리가 상기 마이크로폰의 위치에서 상기 음장 내에 위치된다면 상기 음장 내에 청취자의 오른쪽 귀의 위치와 관련하여 가장 가까운 마이크로폰과 다음으로 가까운 마이크로폰으로부터 출력을 나타내는 신호를 결합시키기 위하여 배치된 것을 특징으로 하는 장치.

## 청구항 7.

제 5항에 있어서,  
 상기 신호 프로세싱 유닛은,  
 상기 마이크로폰 출력 신호의 각각과 관련한 저역 통과 필터;  
 청취자의 왼쪽 귀에 대한 결합된 출력 신호를 만들기 위한 상기 저역 통과 필터의 출력을 결합시키기 위한 수단(상기 결합된 출력 신호는 상기 청취자의 머리가 상기 마이크로폰의 위치에서 상기 음장 내에 위치되어 있다면 상기 음장 내에 청취자의 왼쪽 귀의 위치와 관련하여 가장 가까운 마이크로폰 및 다음으로 가까운 마이크로폰으로부터의 출력을 나타내는 신호들의 조합을 포함한다); 및  
 청취자의 오른쪽 귀에 대한 결합된 출력 신호를 만들기 위한 상기 저역 통과 필터의 출력을 결합시키는 수단(상기 결합된 출력 신호는 상기 청취자의 머리가 상기 마이크로폰의 위치에서 상기 음장 내에 위치된다면 상기 음장에서 청취자의 오른쪽 귀의 위치에 관하여 가장 가까운 마이크로폰 및 다음으로 가까운 마이크로폰으로부터 출력을 나타내는 신호의 조합을 포함한다)을 포함하는 것을 특징으로 하는 장치.

## 청구항 8.

제 7항에 있어서,

상기 신호 프로세싱 유닛은,

상기 음장 내에 위치한 왼쪽 귀 보완적인 마이크로폰으로부터 출력을 제공하기 위하여 배치된 왼쪽 귀 고역 통과 필터;

상기 음장 내에 위치한 오른쪽 귀 보완적인 마이크로폰으로부터 출력을 제공하기 위하여 배치된 오른쪽 귀 고역 통과 필터;

상기 왼쪽 귀 고역 통과 필터로부터의 출력과 청취자의 왼쪽 귀에 대하여 상기 결합된 출력을 결합시키는 수단; 및

상기 오른쪽 귀 고역 통과 필터로부터의 출력과 청취자의 오른쪽 귀에 대하여 상기 결합된 출력을 결합시키는 수단을 포함하는 것을 특징으로 하는 장치.

## 청구항 9.

신호 프로세싱 유닛을 포함하되,

상기 신호 프로세싱 유닛은 음성 출력 장치에 연결되기 위한 출력을 가지고;

상기 신호 프로세싱 유닛은 헤드 트래킹 장치에 연결되기 위한 입력을 가지며;

상기 신호 프로세싱 유닛은 상기 청취자의 머리가 상기 마이크로폰의 위치에서 상기 음장 내에 위치되어 있다면 청취자의 귀의 가능한 위치를 나타내는 지점에서 음장을 샘플하기 위하여 위치한 다수의 마이크로폰의 출력을 나타내는 신호를 받고, 상기 마이크로폰 출력 신호를 처리하며, 상기 헤드 트래킹 장치에 의하여 지시된 것과 같이 상기 청취자의 머리의 방향에 대하여 상기 음성 출력 장치로 바이노럴 출력을 나나내는 수단을 포함하는 것을 특징으로 하는 장치.

## 청구항 10.

제 9항에 있어서,

상기 신호 프로세싱 유닛은,

상기 청취자의 머리가 상기 마이크로폰의 위치에서 상기 음장 내에 위치되어 있다면 상기 음장 내에서 청취자의 귀의 위치에 관련하여 가장 가까운 마이크로폰과 다음으로 가까운 마이크로폰으로부터의 출력을 나타내는 신호를 결합시키기 위한 수단을 더 포함하는 것을 특징으로 하는 장치.

## 청구항 11.

제 9항에 있어서,

상기 신호 프로세싱 유닛은,

상기 마이크로폰 출력 신호의 각각과 연관된 저역 통과 필터; 및 청취자의 귀에 대하여 결합된 출력 신호를 만들기 위하여 상기 저역 통과 필터의 출력을 결합하는 수단을 더 포함하며,

상기 결합된 출력 신호는 청취자의 머리가 상기 마이크로폰의 위치에서 음장 내에 위치되어 있다면 상기 음장 내에서 청취자의 귀의 위치에 관련하여 가장 가까운 마이크로폰과 다음으로 가까운 마이크로폰으로부터의 출력을 나타내는 신호들의 조합을 포함하는 것을 특징으로 하는 장치.

## 청구항 12.

제 11항에 있어서,

상기 신호 프로세싱 유닛은,

상기 음장 내에 위치한 보완적인 마이크로폰으로부터 출력 신호를 제공하기 위하여 배치된 고역 통과 필터; 및 상기 고역 통과 필터로부터의 출력 신호와 청취자의 귀에 대하여 상기 결합된 출력 신호를 결합시키는 수단을 더 포함하는 것을 특징으로 하는 장치.

### 청구항 13.

상기 청취자의 머리가 상기 마이크로폰의 위치에서 상기 음장 내에 위치되어 있다면 청취자의 귀의 가능한 위치를 나타내는 지점에서 음장을 샘플하기 위하여 위치된 다수의 마이크로폰; 및 신호 프로세싱 유니트를 포함하되,

상기 신호 프로세싱 유니트는 음성 출력 장치와 연결되기 위한 출력을 가지고;

상기 신호 프로세싱 유니트는 헤드 트래킹 장치와 연결되기 위한 입력을 가지며;

상기 신호 프로세싱 유니트는 상기 마이크로폰 출력 신호를 처리하고, 상기 헤드 트래킹 장치에 의하여 지시된 것처럼 상기 청취자의 머리의 방향에 대응하여 상기 바이노럴 출력을 상기 음성 출력 장치로 나타내는 것을 특징으로 하는 동적인 바이노럴 음향 캡처 및 재생을 위한 장치.

### 청구항 14.

제 13항에 있어서,

상기 마이크로폰은 상기 청취자의 머리의 반경에 근접하는 반경을 가지는 표면 주위에 원형의 배열로 위치된 것을 특징으로 하는 장치.

### 청구항 15.

제 13항에 있어서,

상기 신호 프로세싱 유니트는,

상기 청취자의 머리가 상기 마이크로폰의 위치에서 상기 음장 내에 위치되어 있다면 상기 음장 내에서 청취자의 귀의 위치에 관련하여 가장 가까운 마이크로폰과 다음으로 가까운 마이크로폰으로부터의 출력을 나타내는 신호들을 결합시키기 위하여 배치된 것을 특징으로 하는 장치.

### 청구항 16.

제 13항에 있어서,

상기 신호 프로세싱 유니트는,

상기 마이크로폰 출력 신호의 각각과 연관된 저역 통과 필터; 및 청취자의 귀에 대하여 결합된 출력 신호를 만들기 위하여 상기 저역 통과 필터의 출력을 결합시키기 위한 수단을 포함하고,

상기 결합된 출력 신호는 상기 청취자의 머리가 상기 마이크로폰의 위치에서 상기 음장 내에 위치되어 있다면 상기 음장 내에 청취자의 귀의 위치와 관련하여 가장 가까운 마이크로폰 및 다음으로 가까운 마이크로폰으로부터의 출력을 나타내는 신호들의 결합을 포함하는 것을 특징으로 하는 장치.

### 청구항 17.

제 16항에 있어서,

상기 음장 내에 위치된 보완적인 마이크로폰;

상기 보완적인 마이크로폰으로부터의 출력 신호를 제공하기 위하여 배치된 고역 통과 필터; 및

상기 고역 통과 필터로부터의 상기 출력 신호와 청취자위 귀에 대하여 상기 결합된 출력 신호를 결합시키기 위한 수단을 더 포함하고,

상기 저역 통과 필터에 의하여 제거된 고주파 내용물이 다시 끼워넣어지는 것을 특징으로 하는 장치.

## 청구항 18.

제 17항에 있어서,

상기 보완적인 마이크로폰은,

상기 다수의 마이크로폰내에서 상기 마이크로폰으로부터 분리된 마이크로폰, 상기 다수의 마이크로폰에서 상기 마이크로폰 중의 하나, 상기 다수의 마이크로폰에서 다수의 동적으로 변화된 마이크로폰으로부터의 신호로부터 만들어진 가상의 마이크로폰, 및 상기 다수의 마이크로폰에서 두개의 마이크로폰으로부터 신호의 스펙트럼 내삽에 의하여 만들어진 가상의 마이크로폰으로 구성된 그룹으로부터 선택된 현실 또는 가상의 마이크로폰을 포함하는 것을 특징으로 하는 장치.

## 청구항 19.

상기 청취자의 머리가 상기 마이크로폰의 위치에서 상기 음장 내에 위치되어 있다면 청취자의 왼쪽 및 오른쪽 귀의 가능한 위치를 나타내는 지점에서 음장을 샘플음장을 샘플 위치된 다수의 마이크로폰; 및 신호 프로세싱 유닛을 포함하되,

상기 신호 프로세싱 유닛은 음성 출력 장치에 연결되기 위한 출력을 가지고;

상기 신호 프로세싱 유닛은 헤드 트래킹 장치에 연결되기 위한 입력을 가지며;

상기 신호 프로세싱 유닛은 상기 마이크로폰으로부터 출력을 결합시키고, 상기 헤드 트래킹 장치에 의하여 지시된 것처럼 상기 청취자의 머리의 방향에 대하여 바이노럴 출력을 상기 음성 출력 장치로 나타내기 위하여 배치된 것을 특징으로 하는 동적인 바이노럴 음향 캡처 및 재생 장치.

## 청구항 20.

제 19항에 있어서,

상기 마이크로폰은 상기 청취자의 머리의 반경에 근접하는 반경을 가진 표면 주위에 원형 배열로 위치된 것을 특징으로 하는 장치.

## 청구항 21.

제 19항에 있어서,

상기 신호 프로세싱 유닛은 상기 청취자의 머리가 상기 마이크로폰의 위치에서 상기 음장 내에 위치되어 있다면 상기 음장 내에서 청취자의 왼쪽 귀의 위치에 관하여 가장 가까운 마이크로폰 및 다음으로 가까운 마이크로폰으로부터의 출력을 나타내는 신호를 결합시키기 위하여 배치되고,

상기 신호 프로세싱 유닛은 상기 청취자의 머리가 상기 마이크로폰의 위치에서 상기 음장 내에 위치되어 있다면 상기 음장 내에서 청취자의 오른쪽 귀의 위치에 관하여 가장 가까운 마이크로폰 및 다음으로 가까운 마이크로폰으로부터의 출력을 나타내는 신호를 결합하기 위하여 배치된 것을 특징으로 하는 장치.

## 청구항 22.

제 19항에 있어서,

상기 신호 프로세싱 유닛은 상기 마이크로폰 출력 신호의 각각과 연관된 저역 통과 필터;

청취자의 왼쪽 귀에 대하여 결합된 출력 신호를 만들기 위하여 상기 저역 통과 필터의 출력을 결합시키기 위한 수단(상기 결합된 출력 신호는 상기 청취자의 머리가 상기 마이크로폰의 위치에서 상기 음장 내에 위치되어 있다면 상기 음장 내에서 청취자의 왼쪽 귀의 위치에 관하여 가장 가까운 마이크로폰 및 다음으로 가까운 마이크로폰으로부터의 출력을 나타내는 신호의 조합을 포함한다); 및



청취자의 오른쪽 귀에 대하여 결합된 출력 신호를 만들기 위하여 상기 저역 통과 필터의 출력을 결합시키기 위한 수단 (상기 결합된 출력 신호는 상기 청취자의 귀가 상기 마이크로폰의 위치에서 상기 음장 내에 위치되어 있다면 상기 음장 내에서 청취자의 오른쪽 귀의 위치에 관하여 가장 가까운 마이크로폰 및 다음으로 가까운 마이크로폰으로부터의 출력을 나타내는 신호의 조합을 포함한다)을 포함하는 것을 특징으로 하는 장치.

### 청구항 23.

제 22항에 있어서,

상기 신호 프로세싱 유닛은,

상기 음장 내에 위치한 왼쪽 귀 보완적인 마이크로폰으로부터 출력을 제공하기 위하여 배치된 왼쪽 귀 고역 통과 필터;

상기 음장 내에 위치한 오른쪽 귀 보완적인 마이크로폰으로부터 출력을 제공하기 위하여 배치된 오른쪽 귀 고역 통과 필터;

상기 왼쪽 귀 고역 통과 필터로부터의 출력과 청취자의 왼쪽 귀에 대하여 상기 결합된 출력을 결합시키기 위한 수단; 및

상기 오른쪽 귀 고역 통과 필터로부터의 출력과 청취자의 오른쪽 귀에 대하여 상기 결합된 출력을 결합시키기 위한 수단을 포함하며,

상기 저역 통과 필터에 의하여 제거된 고주파 내용물이 다시 끼워넣어 지는 것을 특징으로 하는 장치.

### 청구항 24.

제 23항에 있어서,

상기 보완적인 마이크로폰은,

상기 다수의 마이크로폰 내에서 상기 마이크로폰들로부터 분리된 하나의 마이크로폰, 상기 다수의 마이크로폰내에서 상기 마이크로폰 중의 하나, 상기 다수의 마이크로폰에서 다수의 동적으로 변화된 마이크로폰으로부터의 신호로부터 만들어진 가상의 마이크로폰, 및 상기 다수의 마이크로폰 내에서 두개의 마이크로폰으로부터의 신호의 스펙트럼 내삽에 의하여 만들어진 가상의 마이크로폰으로 구성된 그룹으로부터 선택된 현실의 또는 가상의 마이크로폰을 포함하는 것을 특징으로 하는 장치.

### 청구항 25.

상기 청취자의 머리가 상기 마이크로폰의 위치에서 상기 음장 내에 위치되어 있다면 청취자의 귀의 가능한 위치를 나타내는 지점에서 음장을 샘플하기 위하여 위치한 다수의 마이크로폰; 및 신호 프로세싱 유닛을 포함하되,

상기 신호 프로세싱 유닛은 음성 출력 장치에 연결되기 위한 출력을 가지고;

상기 신호 프로세싱 유닛은 헤드 트래킹 장치에 연결되기 위한 입력을 가지며;

상기 신호 프로세싱 유닛은 상기 마이크로폰 출력 신호를 처리하고 상기 헤드 트래킹 장치에 의하여 지시된 것처럼 상기 청취자의 머리의 방향에 대응하여 바이노럴 출력을 상기 음성 출력 장치로 나타내기 위한 수단을 포함하는 것을 특징으로 하는 동적인 바이노럴 음향 캡처 및 재생 장치.

### 청구항 26.

제 25항에 있어서,

상기 마이크로폰은 상기 청취자의 머리의 반경에 근접하는 반경을 가진 표면 주위에 원형 배열로 위치되어 있는 것을 특징으로 하는 장치.

### 청구항 27.

제 25항에 있어서,

상기 신호 프로세싱 유닛은,

상기 청취자의 머리가 상기 마이크로폰의 위치에서 상기 음장 내에 위치되어 있다면 상기 음장 내에서 청취자의 귀의 위치와 관련하여 가장 가까운 마이크로폰 및 다음으로 가까운 마이크로폰으로부터 출력을 나타내는 신호를 결합시키기 위한 수단을 더 포함하는 것을 특징으로 하는 장치.

## 청구항 28.

제 25항에 있어서,

상기 신호 프로세싱 유닛은 상기 마이크로폰 출력 신호의 각각과 연관된 저역 통과 필터; 및 청취자의 귀에 대하여 결합된 출력 신호를 만들기 위하여 상기 저역 통과 필터의 출력을 결합시키는 수단을 포함하며,

상기 결합된 출력 신호는 상기 청취자의 머리가 상기 마이크로폰의 위치에서 상기 음장 내에 위치되어 있다면 상기 음장 내에서 청취자의 귀의 위치와 관련하여 가장 가까운 마이크로폰 및 다음으로 가까운 마이크로폰으로부터의 출력을 나타내는 신호의 결합을 포함하는 것을 특징으로 하는 장치.

## 청구항 29.

제 28항에 있어서,

상기 음장 내에 위치한 보완적인 마이크로폰;

상기 보완적인 마이크로폰으로부터 출력 신호를 제공하기 위하여 배치된 고역 통과 필터; 및

상기 고역 통과 필터로부터 상기 출력 신호와 청취자의 귀에 대하여 상기 결합된 출력 신호를 결합시키기 위한 수단을 더 포함하고,

상기 저역 통과 필터에 의하여 제거된 고주파 내용물이 다시 끼워넣어 지는 것을 특징으로 하는 장치.

## 청구항 30.

제 29항에 있어서,

상기 보완적인 마이크로폰은,

상기 다수의 마이크로폰에서 상기 마이크로폰으로부터 분리된 마이크로폰, 상기 다수의 마이크로폰에서 상기 마이크로폰 중의 하나, 상기 다수의 마이크로폰에서 다수의 동적으로 변화된 마이크로폰으로부터의 신호로부터 만들어진 가상의 마이크로폰, 및 상기 다수의 마이크로폰에서 두개의 마이크로폰으로부터의 신호의 스펙트럼 내삽에 의하여 만들어진 가상의 마이크로폰으로 구성된 그룹으로부터 선택된 현실의 또는 가상의 마이크로폰을 포함하는 것을 특징으로 하는 장치.

## 요약

라이브 또는 녹음된 3차원 음향을 캡처하고 재생하는 새로운 접근이 기술된다. 이는 모션-트랙 바이노럴 즉, MTB라 불리우며, 몇개의 마이크로폰(14), 헤드 트래커(18) 및 마이크로폰(14)에 의하여 포착된 신호를 결합시키기 위한 특별한 신호 프로세싱 과정을 포함한다. MTB는 음향이 발생하는 장소에 청취자의 귀를 효과적으로 위치시키고, 청취자의 머리 운동과 동시에 가상의 귀를 움직이는 것에 의하여, 높은 현실감을 성취한다. MTB는 또한 공간 음향을 레코딩하기 위해서 보편적인 포맷을 제공한다.

## 대표도

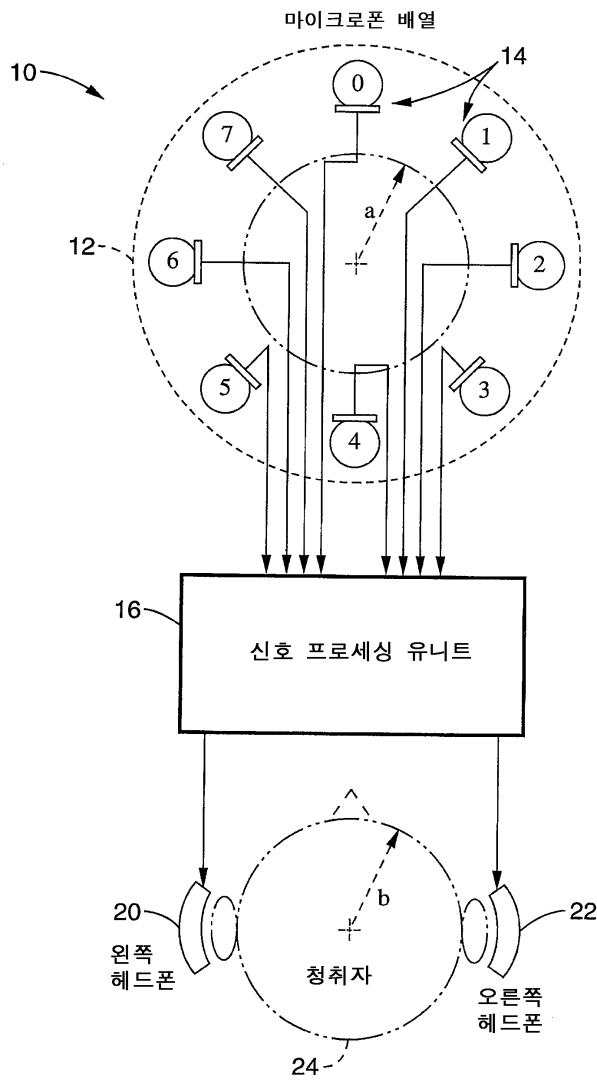
도 2

## 색인어

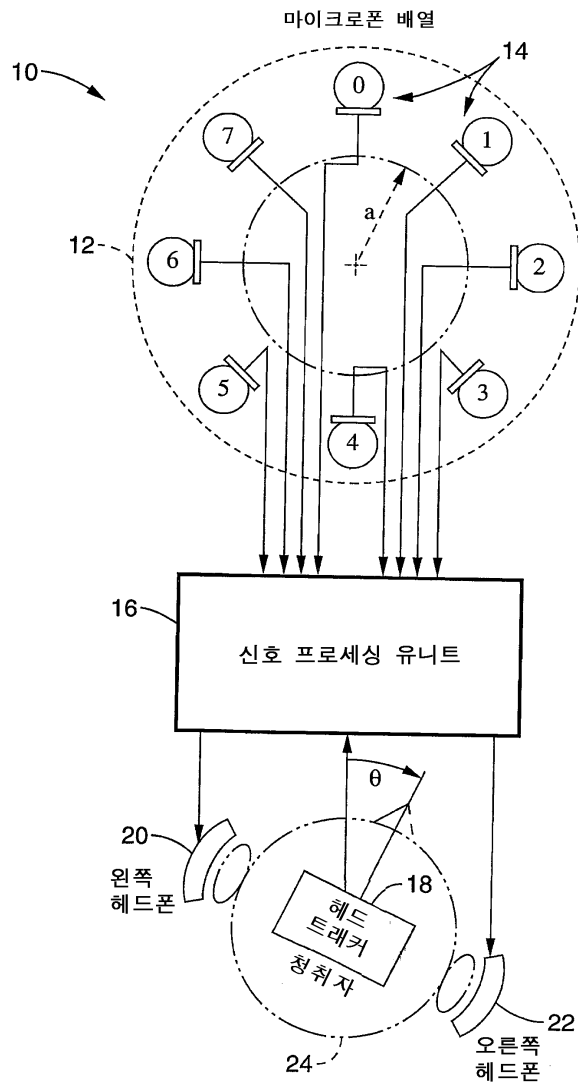
바이노럴 음향, 헤드 트래커.

도면

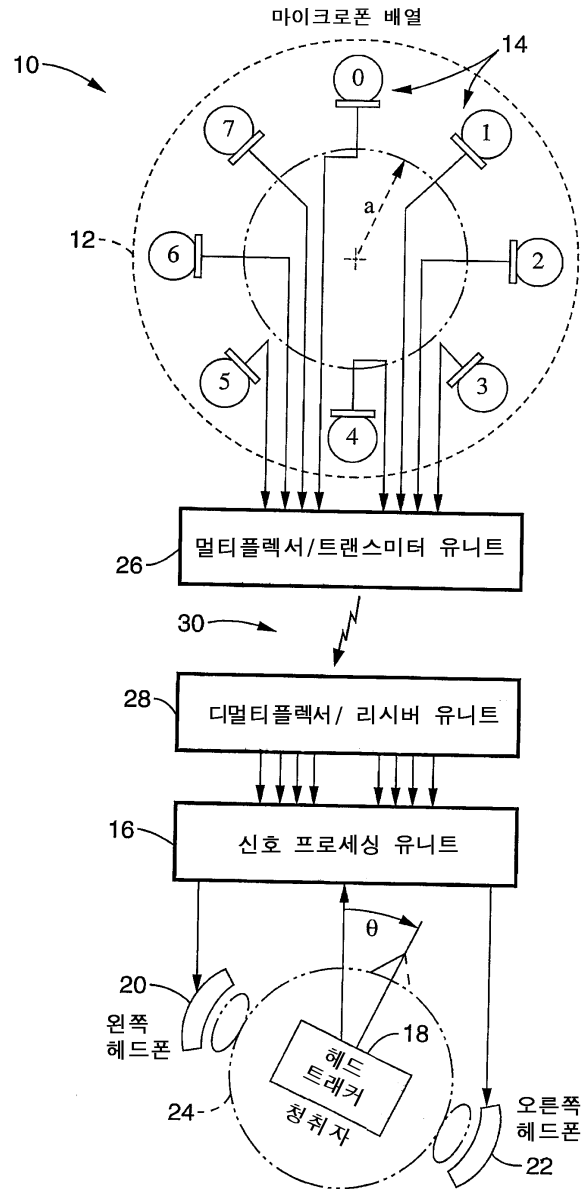
도면1



도면2

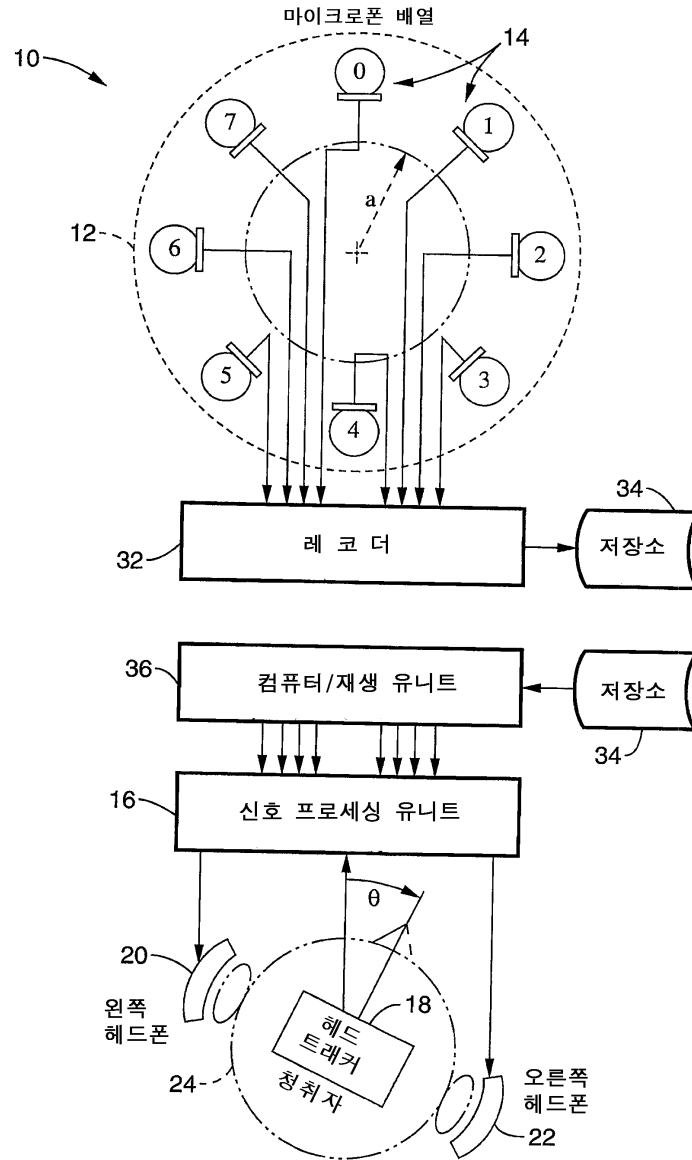


도면3

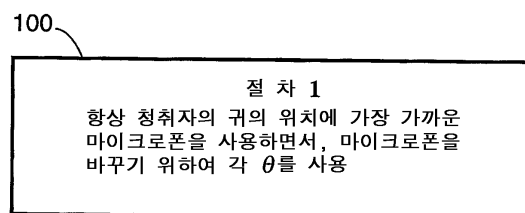




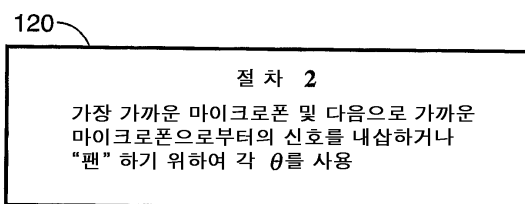
도면4



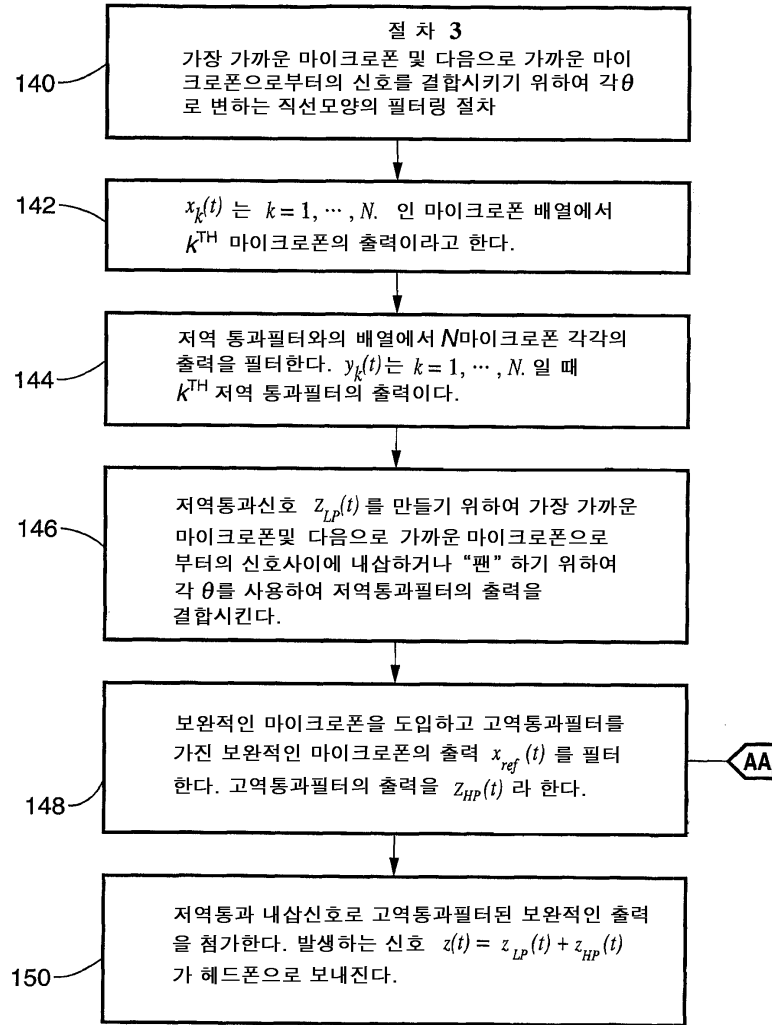
도면5



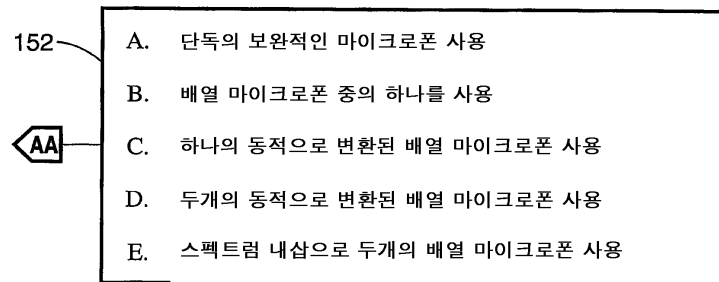
도면6



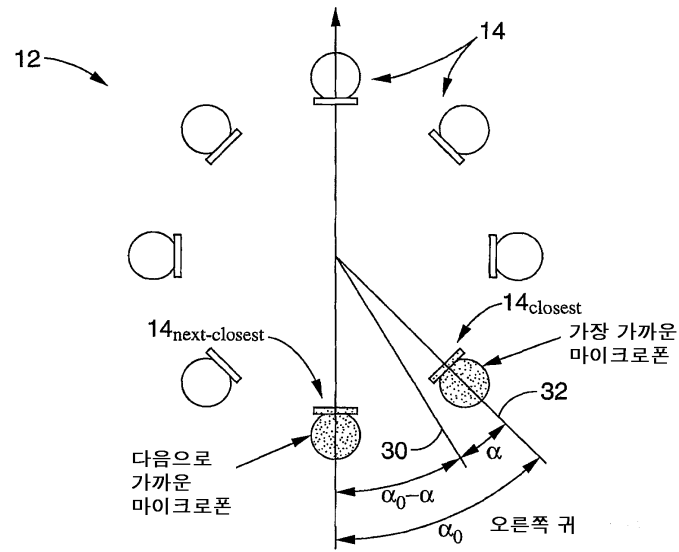
도면7A



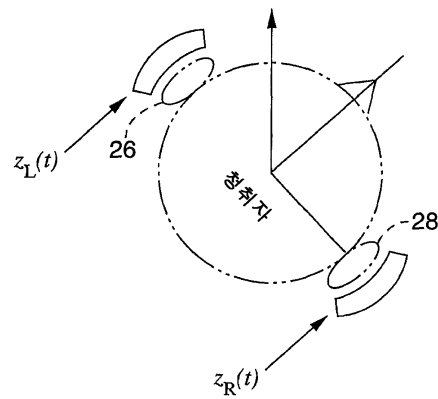
도면7B



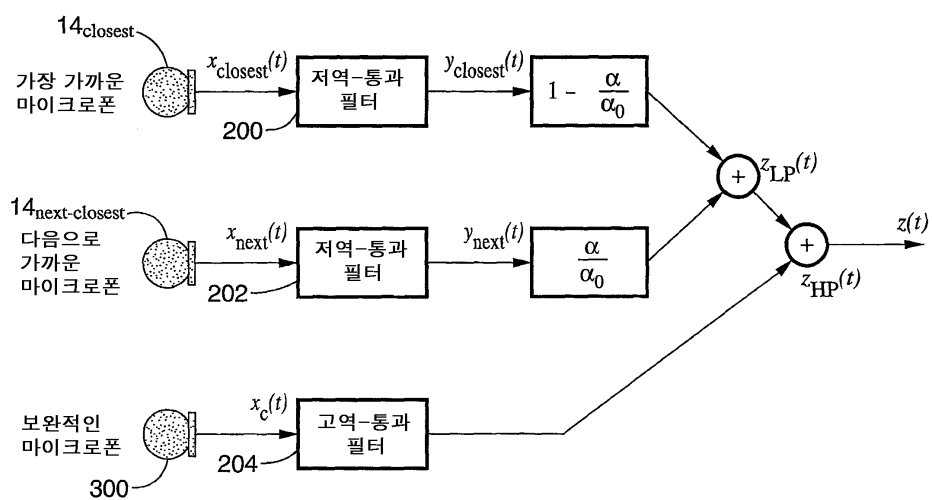
도면8A



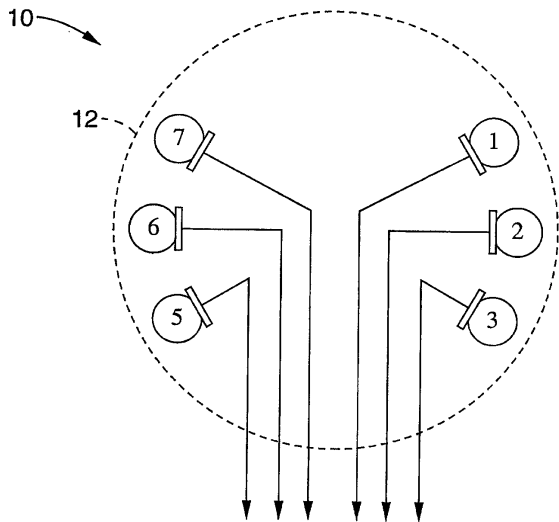
도면8B



도면9



도면10



도면11

방향,탐지를 위한 MTB 배열

