



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2017-0084335
(43) 공개일자 2017년07월19일

- (51) 국제특허분류(Int. Cl.)
H04S 3/02 (2006.01) G10L 19/008 (2014.01)
H04S 7/00 (2006.01)
- (52) CPC특허분류
H04S 3/02 (2013.01)
G10L 19/008 (2013.01)
- (21) 출원번호 10-2017-7018317(분할)
- (22) 출원일자(국제) 2011년03월25일
심사청구일자 없음
- (62) 원출원 특허 10-2012-7025099
원출원일자(국제) 2011년03월25일
심사청구일자 2016년03월24일
- (85) 번역문제출일자 2017년07월03일
- (86) 국제출원번호 PCT/EP2011/054644
- (87) 국제공개번호 WO 2011/117399
국제공개일자 2011년09월29일
- (30) 우선권주장
10305316.1 2010년03월26일
유럽특허청(EPO)(EP)

- (71) 출원인
돌비 인터네셔널 에이비
네덜란드 1101 씨엔 암스트레담 주이두스트 헤리
커베르그백 1-35 3이 아폴로 빌딩
- (72) 발명자
마트케, 요한-마커스
독일 30625 하노버 칼 비헤르트 알레 74 리서치
앤드 이노베이션 도이치 톰슨 오에이치취
케일러, 플로리안
독일 30625 하노버 칼 비헤르트 알레 74 리서치
앤드 이노베이션 도이치 톰슨 오에이치취
보엠, 요하네스
독일 30625 하노버 칼 비헤르트 알레 74 리서치
앤드 이노베이션 도이치 톰슨 오에이치취
- (74) 대리인
양영준, 전경석, 백만기

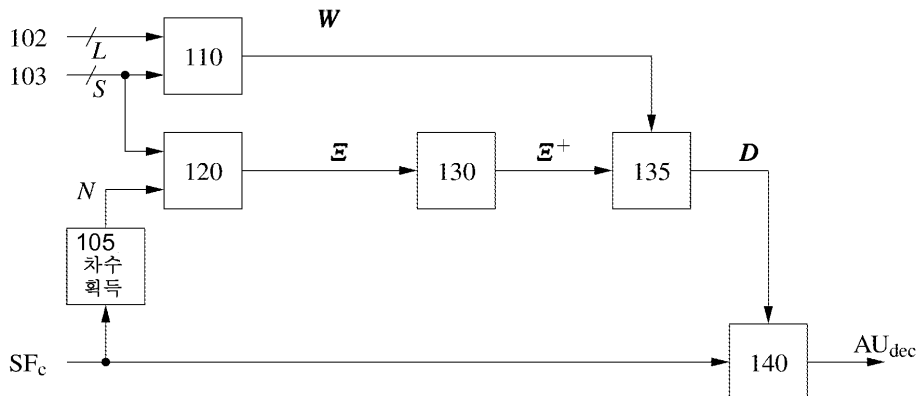
전체 청구항 수 : 총 13 항

(54) 발명의 명칭 오디오 재생을 위한 오디오 사운드필드 표현을 디코딩하는 방법 및 장치

(57) 요약

예컨대 앰비소닉스와 같은 사운드필드 신호는 원하는 사운드필드의 표현을 갖고 있다. 앰비소닉스 포맷은 사운드필드의 구면 고조파 분해에 바탕을 두고 있으며, 고차 앰비소닉스(HOA)는 적어도 2차의 구면 고조파를 이용한다. 그러나, 일반적으로 이용되는 확성기 설정은 불규칙적이며 디코더 설계 시에 문제가 된다. 오디오 재생을 위한 오디오 사운드필드 표현을 디코딩하는 방법은 복수의 확성기 위치와 복수의 소스 방향에 기초한 기하학적 방법을 이용하여 패닝 함수(W)를 산출하는 단계(110), 상기 확성기 위치들로부터 모드 행렬(Ξ)을 산출하는 단계(120), 의사 역모드 행렬(Ξ^+)을 산출하는 단계(130), 및 상기 오디오 사운드필드 표현을 디코딩하는 단계(140)를 포함한다. 상기 디코딩은 상기 패닝 함수(W)와 상기 의사 역모드 행렬(Ξ^+)로부터 구한 디코드 행렬(D)에 기초한다.

대표도



(52) CPC특허분류

H04S 7/308 (2013.01)

H04S 2400/13 (2013.01)

H04S 2420/11 (2013.01)

명세서

청구범위

청구항 1

복수의 확성기를 통한 재생을 위하여 앰비소닉스(ambisonics) 오디오 사운드필드 표현을 디코딩하는 방법에 있어서,

복수의 확성기 각각에 대하여, 상기 확성기들의 위치들과 복수의 소스 방향에 기초한 기하학적 방법을 사용하여 패닝 함수(panning function)를 획득하는 단계;

상기 소스 방향들로부터 모드 행렬(mode matrix)과, 상기 앰비소닉스 오디오 사운드필드 표현의 차수(order)를 획득하는 단계;

상기 모드 행렬로부터 베이스 행렬을 획득하는 단계; 및

상기 앰비소닉스 오디오 사운드필드 표현을 디코딩 행렬로 디코딩하는 단계 - 상기 디코딩 행렬은 상기 패닝 함수 및 상기 베이스 행렬에 기초하고, 상기 소스 방향들은 단위 구면에 고르게 분포되고, 상기 소스 방향들의 수는 S이고, 상기 앰비소닉스 오디오 사운드필드 표현의 상기 차수는 N이고, $S \geq (N+1)^2$ 임 -

를 포함하는 방법.

청구항 2

제1항에 있어서,

상기 패닝 함수를 획득하는 단계에서 사용되는 상기 기하학적 방법은 벡터 베이스 진폭 패닝(VBAP; Vector Base Amplitude Panning)에 기초하는, 방법.

청구항 3

제1항에 있어서,

상기 앰비소닉스 오디오 사운드필드 표현은 적어도 제2 차수인, 방법.

청구항 4

제1항에 있어서,

상기 베이스 행렬은 상기 모드 행렬 및 전치 행렬의 곱에 기초하는, 방법.

청구항 5

제1항에 있어서,

상기 패닝 함수는 행렬로서 표현되고, 상기 베이스 행렬은 상기 모드 행렬의 규칙화(regularization)인, 방법.

청구항 6

제1항에 있어서,

상기 패닝 함수는 이득값들로서 표현되는, 방법.

청구항 7

복수의 확성기를 통한 재생을 위하여 앰비소닉스 오디오 사운드필드 표현을 디코딩하기 위한 장치로서,

복수의 확성기 각각에 대하여, 상기 확성기들의 위치들과 복수의 소스 방향에 기초한 기하학적 방법을 사용하여 패닝 함수를 획득하기 위한 수단;

상기 소스 방향들로부터 모드 행렬과, 상기 앰비소닉스 오디오 사운드필드 표현의 차수를 획득하기 위한 수단;

상기 모드 행렬로부터 베이스 행렬을 획득하기 위한 수단; 및

상기 앰비소닉스 오디오 사운드필드 표현을 디코딩 행렬로 디코딩하기 위한 수단 - 상기 디코딩 행렬은 상기 패닝 함수 및 상기 베이스 행렬에 기초하고, 상기 소스 방향들은 단위 구면에 고르게 분포되고, 상기 소스 방향들의 수는 S이고, 상기 앰비소닉스 오디오 사운드필드 표현의 상기 차수는 N이고, $S \geq (N+1)^2$ 임 -

을 포함하는 장치.

청구항 8

제7항에 있어서,

상기 패닝 함수를 획득하는 데 사용되는 상기 기하학적 방법은 벡터 베이스 진폭 패닝(VBAP)에 기초하는, 장치.

청구항 9

제7항에 있어서,

상기 앰비소닉스 오디오 사운드필드 표현은 적어도 제2 차수인, 장치.

청구항 10

제7항에 있어서,

상기 베이스 행렬은 상기 모드 행렬 및 전치 행렬의 곱에 기초하는, 장치.

청구항 11

제7항에 있어서,

상기 패닝 함수는 행렬로서 표현되고, 상기 베이스 행렬은 상기 모드 행렬의 규칙화인, 장치.

청구항 12

제7항에 있어서,

상기 패닝 함수는 이득값들로서 표현되는, 장치.

청구항 13

컴퓨터로 하여금, 오디오 재생을 위하여 앰비소닉스 오디오 사운드필드 표현을 디코딩하는 방법을 수행하도록 하는 실행가능한 명령어들을 저장하는 컴퓨터 판독가능 매체로서,

상기 방법은,

복수의 확성기 각각에 대하여, 상기 확성기들의 위치들과 복수의 소스 방향에 기초한 기하학적 방법을 사용하여 패닝 함수(panning function)를 획득하는 단계;

상기 소스 방향들로부터 모드 행렬과, 상기 앰비소닉스 오디오 사운드필드 표현의 차수(order)를 획득하는 단계;

상기 모드 행렬로부터 베이스 행렬을 획득하는 단계; 및

상기 앰비소닉스 오디오 사운드필드 표현을 디코딩 행렬로 디코딩하는 단계 - 상기 디코딩 행렬은 상기 패닝 함수 및 상기 베이스 행렬에 기초하고, 상기 소스 방향들은 단위 구면에 고르게 분포되고, 상기 소스 방향들의 수는 S이고, 상기 앰비소닉스 오디오 사운드필드 표현의 상기 차수는 N이고, $S \geq (N+1)^2$ 임 -

를 포함하는, 컴퓨터 판독가능 매체.

발명의 설명

기술 분야

본 발명은 오디오 재생을 위한 오디오 사운드필드 표현, 특히 앰비소닉스 포맷 오디오 표현을 디코딩하는 방법

및 장치에 관한 것이다.

배경 기술

- [0002] 이 절에서는 하기에 설명 및/또는 권리청구되는 본 발명의 여러 가지 양상에 관련될 수 있는 기술의 여러 가지 양상을 독자에게 소개하고자 한다. 이 설명은 본 발명의 여러 가지 양상을 더 잘 이해할 수 있도록 하는 배경 정보를 독자에게 제공하는 데 도움이 될 것으로 생각한다. 따라서, 이 설명은 소스가 명시적으로 언급되지 않는 한, 이러한 견지에서 파악되어야 하며, 종래 기술을 인정하는 것으로 이해되어서는 안 된다는 것을 알아야 한다.
- [0003] 정확한 로컬라이제이션(localization)은 임의의 공간적 오디오 재생 시스템의 핵심적인 목표이다. 그와 같은 재생 시스템은 화상회의 시스템, 게임, 또는 기타 여러 가지 3D 사운드로 특을 보는 가상 환경에 매우 잘 적용될 수 있다. 3D 사운드 장면은 자연스러운 사운드필드로서 합성되거나 캡처될 수 있다. 예컨대 앰비소닉스와 같은 사운드필드 신호는 원하는 사운드필드의 표현을 갖고 있다. 앰비소닉스 포맷은 사운드필드의 구면 고조파 분해(spherical harmonic decomposition)에 바탕을 두고 있다. 기본 앰비소닉스 포맷, 즉 B-포맷은 차수가 제로 또는 1인 구면 고조파를 이용하지만, 소위 고차 앰비소닉스(Higher Order Ambisonics: HOA)는 또한 적어도 2차의 구면 고조파를 더 이용한다. 개별적인 확성기 신호를 얻기 위해서는 디코딩 프로세스가 필요하다. 오디오 장면을 합성하려면, 특정 사운드 소스의 공간적 로컬라이제이션을 얻기 위해서 공간적 확성기 배치를 지시하는 패닝(panning) 함수가 필요하다. 자연스러운 사운드필드가 녹화되기 위해서는 공간 정보를 캡처하는 마이크로폰 어레이가 필요하다. 공지의 앰비소닉스 방식은 이를 달성할 수 있는 매우 적합한 도구이다. 앰비소닉스 포맷 신호는 원하는 사운드필드의 표현을 갖고 있다. 그와 같은 앰비소닉스 포맷 신호로부터 개별적인 확성기 신호를 얻기 위해서는 디코딩 프로세스가 필요하다. 이 경우에도 패닝 함수가 디코딩 함수로부터 유도될 수 있기 때문에, 패닝 함수는 공간적 로컬라이제이션 업무를 기술하는 데 있어 핵심적인 문제이다. 확성기의 공간적 배치는 여기서는 확성기 설정이라고 한다.
- [0004] 일반적으로 이용되는 확성기 설정은 2개의 확성기를 이용하는 스테레오 설정, 5개의 확성기를 이용하는 표준 서라운드 설정, 5개 초과와 확성기를 이용하는 확장된 서라운드 설정이다. 이들 설정은 공지되어 있다. 그러나, 이들은 2차원(2D)에 한정된다. 예컨대, 높이 정보는 재생되지 않는다.
- [0005] 3차원(3D) 재생을 위한 확성기 설정은 예컨대 「"Wide listening area with exceptional spatial sound quality of a 22.2 multichannel sound system", K. Hamasaki, T. Nishiguchi, R. Okumaura, and Y. Nakayama in Audio Engineering Society Preprints, Vienna, Austria, May 2007」(이것은 22.2 format, 즉 2+2+2 arrangement of Dabringhaus(mdg-musikproduktion dabringhaus und grimm, www.mdg.de)를 가진 NHK 초고해상도 TV에 대한 제안서임)과, 「10.2 setup in "Sound for Film and Television", T. Holman in 2nd ed. Boston: Focal Press, 2002」에 기재되어 있다. 공간 재생과 패닝 전략을 지시하는 몇 가지 공지의 시스템들 중 하나는 「"Virtual sound source positioning using vector base amplitude panning," Journal of Audio Engineering Society, vol. 45, no. 6, pp. 456-466, June 1997」(여기에서, Pulkki라 함)에서의 벡터 베이스 진폭 패닝(VBAP) 방식이다. Pulkki는 VBAP(Vector Base Amplitude Panning)를 이용하여 임의의 확성기 설정을 가지고 가상 음향 소스를 재생하였다. 가상 소스를 2D 평면에 배치하기 위해서는 한쌍의 확성기가 필요한 반면에, 3D 경우에는 3중의 확성기가 필요하다. 각 가상 소스에 있어서, (가상 소스의 위치에 따라서) 이득이 서로 다른 모노포닉 신호가 전체 설정 중에서 선택된 확성기에 공급된다. 그런 다음에, 모든 가상 소스에 대한 확성기 신호가 합산된다. VBAP는 확성기들 간의 패닝을 위해 확성기 신호의 이득을 산출하기 위해 기하학적 방식을 적용한다.
- [0006] 여기서 고려되고 새로이 제시되는 예시적인 3D 확성기 설정에는 도 2에 도시된 바와 같이 위치하는 16개의 확성기를 갖고 있다. 각각이 3개의 확성기를 가진 4개의 기둥을 가지고 이들 기둥 사이에 추가 확성기가 있는 실제 고려 사항 때문에 이러한 위치 설정이 선택되었다. 더 구체적으로 설명하면, 청취자의 머리 주위에는 8개의 확성기가 원형으로 45도 각도로 동등하게 분포되어 있다. 추가적인 4개의 확성기가 상단과 하단에 90도의 방위각으로 배치되어 있다. 앰비소닉스에 관해서는, 「"An ambisonics format for flexible playback layouts," by H. Pomberger and F. Zotter in Proceedings of the 1st Ambisonis Symposium, Graz, Austria, July 2009」에서 언급한 바와 같이, 이 설정은 불규칙적이며 디코더 설계시에 문제가 된다.
- [0007] 「"Three-dimensional surround sound systems based on spherical harmonics" by M. Poletti in J. Audio Eng. Soc., vol. 53, no. 11, pp. 1004-1025, Nov. 2005」에 기재되어 있는 바와 같이, 종래의 앰비소닉스 디

코딩은 일반적으로 알려져 있는 모드 매칭 프로세스를 이용한다. 모드는 명확한 입사 방향에 대한 구면 고조파의 값을 포함하는 모드 벡터에 의해 기술된다. 개별 확성기에 의해 주어지는 모든 방향을 조합하면 확성기 설정의 모드 행렬이 되며, 이 모드 행렬은 확성기 위치를 나타낸다. 명확한 소스 신호의 모드를 재생하기 위해서는 개별적인 확성기의 중첩된 모드들이 원하는 모드로 합산되는 식으로 확성기 모드들이 가중된다. 필요한 가중치를 얻기 위해서는 확성기 모드 행렬의 역행렬 표현이 산출될 필요가 있다. 신호 디코딩 면에서 보면, 가중치는 확성기의 구동 신호를 형성하며, 역 확성기 모드 행렬은 앰비소닉스 포맷 신호 표현을 디코딩하는 데 적용되는 "디코딩 행렬"이라 한다. 특히, 많은 확성기 설정에 있어서, 예컨대 도 2에 도시된 설정에 있어서는 모드 행렬의 역을 구하기가 어렵다.

[0008] 전술한 바와 같이, 일반적으로 이용되는 확성기 설정은 2D에 한정된다. 즉, 높이 정보는 재생되지 않는다. 수학적으로 규칙적이지 않는 공간 분포를 갖는 확성기 설정에 대한 사운드필드 표현을 디코딩하면 일반적으로 알려진 기법으로는 로컬라이제이션과 음색 변화 문제가 생긴다. 앰비소닉스 신호를 디코딩하기 위해서는 디코딩 행렬(즉, 디코딩 계수 행렬)이 이용된다. 종래의 앰비소닉스 신호, 특히 HOA 신호 디코딩에서는 적어도 2가지 문제가 발생한다. 첫째, 올바른 디코딩을 위해서는 디코딩 행렬을 얻기 위해 신호 소스 방향을 알아야 한다. 둘째, 수학적으로 올바른 디코딩의 결과 양의 확성기 진폭뿐만 아니라 일부 음의 확성기 진폭도 생길 것이라는 수학적 문제 때문에 기존의 확성기 설정에의 맵핑은 조직적으로 틀린다. 그러나, 이들은 양의 신호로서 틀리게 재생되고, 따라서 전술한 문제가 생기게 된다.

발명의 내용

[0009] 본 발명은 개선된 로컬라이제이션과 음색 변화 특성을 가진 비규칙적 공간 분포에 대한 사운드필드 표현을 디코딩하는 방법을 제공한다. 본 발명은 예컨대 앰비소닉스 포맷으로 된 사운드필드 데이터에 대한 디코딩 행렬을 구하는 다른 방법을 대표하며, 프로세스를 시스템 평가 방식으로 이용한다. 가능한 입사 방향 세트를 고려하여, 원하는 확성기와 관련된 패닝 함수가 산출된다. 패닝 함수는 앰비소닉스 디코딩 프로세스의 출력으로서 취해진다. 필요한 입력 신호는 모든 고려되는 방향의 모드 행렬이다. 그러므로, 후술하는 바와 같이, 디코딩 행렬은 다중 행렬에 입력 신호의 모드 행렬의 역행렬을 직접적으로 곱함으로써 구해진다.

[0010] 전술한 두 번째 문제와 관련하여, 확성기 위치를 나타내는 소위 모드 행렬의 역과 위치 종속 가중 함수("패닝 함수") W 로부터 디코딩 행렬을 구하는 것도 가능함을 알았다. 본 발명의 일 양상은 이들 패닝 함수 W 가 일반적으로 이용되는 것과 다른 방법을 이용하여 도출될 수 있다는 것이다. 양호하게도 간단한 기하학적 방법이 이용된다. 그와 같은 방법은 신호 소스 방향을 몰라도 되며, 따라서 전술한 첫 번째 문제를 해결할 수 있다. 그와 같은 하나의 방법은 VBAP(Vector-Based Amplitude Panning)로 알려져 있다. 본 발명에 따라서, VBAP는 필요한 패닝 함수를 산출하는 데 이용되며, 이 패닝 함수는 앰비소닉스 디코딩 행렬을 산출하는 데 이용된다. (확성기 설정을 표현하는) 모드 행렬의 역행렬이 필요하다는 점에서 다른 문제가 발생한다. 그러나, 정확한 역행렬을 구하기는 어렵고, 이 또한 오디오 재생을 틀리게 한다. 따라서, 추가적인 양상은 디코딩 행렬을 구하기 위해 구하기 훨씬 쉬운 의사 역모드 행렬이 산출된다는 것이다.

[0011] 본 발명은 2단계 방식을 이용한다. 제1 단계는 재생을 위해 이용된 확성기 설정에 의존하는 패닝 함수를 도출하는 것이다. 제2 단계에서는 모든 확성기에 대한 패닝 함수로부터 앰비소닉스 디코딩 행렬이 계산된다.

[0012] 본 발명의 이점은 사운드 소스의 파라메트릭 기술(description)이 필요치 않고, 대신에 앰비소닉스와 같은 사운드필드 기술이 이용될 수 있다는 것이다.

[0013] 본 발명에 따라서, 오디오 재생을 위한 오디오 사운드필드 표현을 디코딩하는 방법은, 복수의 확성기 각각에 대해, 확성기들의 위치와 복수의 소스 방향에 기초한 기하학적 방법을 이용하여 패닝 함수를 산출하는 단계, 상기 소스 방향으로부터 모드 행렬을 산출하는 단계, 상기 모드 행렬의 의사 역모드 행렬을 산출하는 단계, 및 상기 오디오 사운드필드 표현을 디코딩하는 단계를 포함하며, 상기 디코딩은 적어도 상기 패닝 함수와 상기 의사 역모드 행렬로부터 구한 디코딩 행렬에 기초한다.

[0014] 다른 양상에 따라서, 오디오 재생을 위한 오디오 사운드필드 표현을 디코딩하는 장치는, 복수의 확성기 각각에 대해, 확성기들의 위치와 복수의 소스 방향에 기초한 기하학적 방법을 이용하여 패닝 함수를 산출하기 위한 제1 산출 수단, 상기 소스 방향으로부터 모드 행렬을 산출하기 위한 제2 산출 수단, 상기 모드 행렬의 의사 역모드 행렬을 산출하기 위한 제3 산출 수단, 및 상기 사운드필드 표현을 디코딩하기 위한 디코더 수단을 포함하고, 상기 디코딩은 디코딩 행렬에 기초하고, 상기 디코더 수단은 적어도 상기 패닝 함수와 상기 의사 역모드 행렬을 이용하여 상기 디코딩 행렬을 구한다. 상기 제1, 제2 및 제3 산출 수단은 단일 프로세서 또는 2개 이상의 독립

적인 프로세서일 수 있다.

[0015] 또 다른 양상에 따라서, 컴퓨터 판독 가능 매체는, 컴퓨터에게, 복수의 확성기 각각에 대해, 확성기들의 위치와 복수의 소스 방향에 기초한 기하학적 방법을 이용하여 패닝 함수를 산출하는 단계, 상기 소스 방향으로부터 모드 행렬을 산출하는 단계, 상기 모드 행렬의 의사 역모드 행렬을 산출하는 단계, 및 상기 오디오 사운드필드 표현을 디코딩하는 단계를 포함하며, 상기 디코딩은 적어도 상기 패닝 함수와 상기 의사 역모드 행렬로부터 구한 디코드 행렬에 기초하는, 오디오 재생을 위한 오디오 사운드필드 표현을 디코딩하는 방법을 수행하게 하는 실행 가능 명령어를 저장한다.

[0016] 본 발명의 바람직한 실시예들은 종속항, 하기의 상세한 설명 및 도면에 개시된다.

도면의 간단한 설명

[0017] 첨부 도면을 참조로 본 발명의 예시적 실시예들에 대해 설명한다.

도 1은 본 방법의 플로우차트.

도 2는 16개의 확성기를 가진 예시적 3D 설정을 보여주는 도.

도 3은 비규칙화(non-regularized) 모드 매칭을 이용하여 디코딩한 결과로서 생긴 빔 패턴을 보여주는 도.

도 4는 규칙화 모드 행렬을 이용하여 디코딩한 결과로서 생긴 빔 패턴을 보여주는 도.

도 5는 VBAP로부터 도출된 디코딩 행렬을 이용하여 디코딩한 결과로서 생긴 빔 패턴을 보여주는 도.

도 6은 듣기 평가의 결과를 보여주는 도.

도 7은 장치의 블록도를 도시한 도.

발명을 실시하기 위한 구체적인 내용

[0018] 도 1에 도시된 바와 같이, 오디오 재생을 위한 오디오 사운드필드 표현(SF_c)을 디코딩하는 방법은, 복수의 확성기 각각에 대해, 확성기들의 위치(102)(L은 확성기 수)와 복수의 소스 방향(103)(S는 소스 방향 수)에 기초한 기하학적 방법을 이용하여 패닝 함수(W)를 산출하는 단계(110), 상기 소스 방향과 상기 사운드필드 표현의 주어진 차수(N)로부터 모드 행렬 Ξ 을 산출하는 단계(120), 모드 행렬 Ξ 의 의사 역모드 행렬 Ξ^+ 을 산출하는 단계(130), 및 디코딩된 사운드 데이터(AU_{dec})를 얻도록 상기 오디오 사운드필드 표현(SF_c)을 디코딩하는 단계(135, 140)를 포함한다. 이 디코딩은 적어도 패닝 함수 W와 의사 역모드 행렬 Ξ^+ 로부터 구한(135) 디코드 행렬 D에 기초한다. 일 실시예에서, 의사 역모드 행렬은 $\Xi^+ = \Xi^H [\Xi \Xi^H]^{-1}$ 에 따라서 구해진다. 사운드필드 표현의 차수(N)는 미리 정의되거나 입력 신호(SF_c)로부터 추출될(105) 수 있다.

[0019] 도 7에 도시된 바와 같이, 오디오 재생을 위한 오디오 사운드필드 표현을 디코딩하는 장치는, 복수의 확성기 각각에 대해, 확성기들의 위치(102)와 복수의 소스 방향(103)에 기초한 기하학적 방법을 이용하여 패닝 함수(W)를 산출하기 위한 제1 산출 수단(210), 상기 소스 방향으로부터 모드 행렬 Ξ 을 산출하기 위한 제2 산출 수단(220), 모드 행렬 Ξ 의 의사 역모드 행렬 Ξ^+ 을 산출하기 위한 제3 산출 수단(230), 및 상기 사운드필드 표현을 디코딩하기 위한 디코더 수단(240)을 포함한다. 이 디코딩은 디코드 행렬 산출 수단(235)(예컨대 곱셈기)에 의해 적어도 패닝 함수 W와 의사 역모드 행렬 Ξ^+ 로부터 구한 디코드 행렬 D에 기초한다. 디코더 수단(240)은 디코드 행렬 D를 이용하여 디코딩된 오디오 신호(AU_{dec})를 얻는다. 제1, 제2 및 제3 산출 수단(210, 220, 230)은 단일 프로세서 또는 2 이상의 독립적인 프로세서일 수 있다. 사운드필드 표현의 차수(N)는 미리 정의되거나 입력 신호(SF_c)로부터 그 차수를 추출하기 위한 수단(205)에 의해 구해질 수 있다.

[0020] 특히 유용한 3D 확성기 설정은 16개의 확성기를 갖고 있다. 도 2에 도시된 바와 같이, 각각이 3개의 확성기를 가진 4개의 기둥이 있으며, 이들 기둥 사이에 추가 확성기가 있다. 청취자의 머리 주위에는 8개의 확성기가 원형으로 45도 각도로 동등하게 분포되어 있다. 추가적인 4개의 확성기가 상단과 하단에 90도의 방위각으로 배치되어 있다. 앰비소닉스에 관해서는 이 설정은 불규칙적이며, 보통은 디코더 설계 시에 문제가 된다.

[0021] 이하에서는 VBAP(Vector Base Amplitude Panning)에 대해 상세히 설명한다. 일 실시예에서, VBAP는 확성기들이 청취 위치로부터 동일한 거리에 있다고 가정한 임의의 확성기 설정으로 가상 음향 소스들을 배치하는 데 이용된다. VBAP는 3개의 확성기를 이용하여 3D 공간에 가상 소스를 배치한다. 각 가상 소스에 대해서는, 사용될 확성기들에 이득이 서로 다른 모노포닉(monophonic) 신호가 공급된다. 이들 서로 다른 확성기의 이득은 가상 소스의 위치에 따라 달라진다. VBAP는 확성기들 간의 패닝(panning)을 위한 확성기 신호들의 이득을 산출하는 기하학적 방식이다. 3D 경우에는, 삼각형으로 배치된 3개의 확성기가 벡터 베이스를 구축한다. 각 벡터 베이스는 확성기 번호 k, m, n 으로 식별되며, 확성기 위치 벡터 l_k, l_m, l_n 은 단위 길이로 정규화된 직교 좌표로 주어진다. 확성기(k, m, n)에 대한 벡터 베이스는 다음과 같이 정의된다.

수학식 1

[0022] $L_{kmn} = \{l_k, l_m, l_n\}$

[0023] 가상 소스의 원하는 방향 $\Omega = (\theta, \phi)$ 는 방위각 ϕ 와 경사각 θ 로 주어져야 한다. 그러므로 직교 좌표에서 가상 소스의 단위 길이 위치 벡터 $p(\Omega)$ 는 다음과 같이 정의된다.

수학식 2

[0024] $p(\Omega) = \{\cos\phi \sin\theta, \sin\phi \sin\theta, \cos\theta\}^T$

[0025] 가상 소스 위치는 상기 벡터 베이스와 이득 계수 $g(\Omega) = (\tilde{g}_k, \tilde{g}_m, \tilde{g}_n)^T$ 를 가지고 다음과 같이 표현될 수 있다.

수학식 3

[0026] $p(\Omega) = L_{kmn} g(\Omega) = \tilde{g}_k l_k + \tilde{g}_m l_m + \tilde{g}_n l_n$

[0027] 벡터 베이스 행렬을 역변환(invert)함으로써 필요한 이득 계수는 다음과 같이 계산될 수 있다.

수학식 4

[0028] $g(\Omega) = L_{kmn}^{-1} p(\Omega)$

[0029] 사용될 벡터 베이스는 Pulkki의 논문에 따라서 결정된다. 먼저, 모든 벡터 베이스에 대해 Pulkki에 따라서 이득이 산출된다. 그 다음, 각 벡터 베이스에 대해 이득 계수의 최소치가 $\tilde{g}_{min} = \min\{\tilde{g}_k, \tilde{g}_m, \tilde{g}_n\}$ 에 따라서 구해진다. 마지막으로 \tilde{g}_{min} 가 최고치를 갖는 벡터 베이스가 이용된다. 이렇게 해서 도출되는 이득 계수는 음수이어서는 안 된다. 청취방 음향에 따라서는 이득 계수는 에너지 보존을 위해 정규화될 수 있다.

[0030] 이하에서는 사운드필드 포맷의 예로서 앰비소닉스 포맷에 대해서 설명한다. 앰비소닉스 표현은 일 위치에서의 사운드필드의 수학적 근사를 이용하는 사운드필드 기술 방법이다. 구면 좌표계를 이용하여 공간 내 지점 $r = (r, \theta, \phi)$ 에서의 압력은 구면 푸리에 변환에 의해 다음과 같이 기술된다.

수학식 5

$$p(\mathbf{r}, k) = \sum_{n=0}^{\infty} \sum_{m=-n}^n A_n^m(k) j_n(kr) Y_n^m(\theta, \phi)$$

[0031]

[0032]

여기서, k 는 파수(wave number)이다. 통상적으로 n 은 유한 차수 M 에 이른다. 이 급수의 계수 $A_n^m(k)$ 는 (소스가 유효 영역밖에 있다고 가정하면) 사운드필드를 기술하며, $j_n(kr)$ 은 제1종 구면 베셀 함수이고, $Y_n^m(\theta, \phi)$ 는 구면 고조파를 나타낸다. 이와 관련하여 계수 $A_n^m(k)$ 는 앰비소닉스 계수로 간주된다. 구면 고조파 $Y_n^m(\theta, \phi)$ 는 경사각과 방위각에만 종속되며 단위 구면 상의 함수를 기술한다.

[0033]

단순화하기 위해 사운드필드 재생에 종종 평면파가 가정된다. 평면파를 방향 Ω_s 로부터의 음향 소스로서 기술하는 앰비소닉스 계수는 다음과 같다.

수학식 6

$$A_{n,\text{plane}}^m(\Omega_s) = 4\pi i^n Y_n^m(\Omega_s)^*$$

[0034]

[0035]

이 계수의 파수 k 에의 종속성은 이러한 특수한 경우에는 순수한 방향 종속성으로 감소한다. 한정된 차수 M 에 대해서는 계수는 $0 = (M+1)^2$ 원소를 유지하면서 다음과 같이 배열될 수 있는 벡터 A 를 형성한다.

수학식 7

$$A(\Omega_s) = [A_0^0 \ A_1^{-1} \ A_1^0 \ A_1^1 \ \dots \ A_M^M]^T$$

[0036]

[0037]

이 배열은 벡터 $Y(\Omega_s)^* = [Y_0^0 \ Y_1^{-1} \ Y_1^0 \ Y_1^1 \ \dots \ A_M^M]^H$ 를 산출하는 구면 고조파 계수에 이용된다. 윗 첨자 H 는 복소 공액 전치를 나타낸다.

[0038]

사운드필드의 앰비소닉스 표현으로부터 확성기 신호를 산출하는 데는 일반적으로 모드 매칭이 이용된다. 기본 개념은 특정 앰비소닉스 사운드필드 기술(description) $A(\Omega_s)$ 를 확성기들의 사운드필드 기술 $A(\Omega_i)$ 의 가중합으로 표현하는 것이다.

수학식 8

$$A(\Omega_s) = \sum_{i=1}^L w_i A(\Omega_i)$$

[0039]

[0040]

여기서, Ω_i 은 확성기 방향을 나타내며, w_i 은 가중치이고, L 은 확성기 수이다. 수학식 8로부터 페닝 함수를 유도하기 위해서 입사 방향 Ω_s 는 이미 알고 있는 것으로 가정한다. 소스와 스피커 사운드필드 모두 평면파라면 계수 $4\pi i^n$ (수학식 6 참조)은 뺄 수 있고, 수학식 8은 "모드"라고도 하는 구면 고조파 벡터의 복소 공액에만 의존한다. 이는 행렬식으로는 다음과 같다.

수학식 9

$$Y(\Omega_s)^* = \Psi w(\Omega_s)$$

[0041]

여기서, Ψ 는 $0 \times L$ 개의 원소를 가진 다음과 같은 확성기 설정의 모드 행렬이다.

수학식 10

$$\Psi = [Y(\Omega_1)^*, Y(\Omega_2)^*, \dots, Y(\Omega_L)^*]$$

[0043]

원하는 가중 벡터 w 를 얻기 위해, 이를 달성하는 여러 가지 전략이 알려져 있다. 만일 $M=3$ 이 선택되면, Ψ 는 정방형(square)이며 역변환가능(invertible)하다. 그렇지만 불규칙적인 확성기 설정으로 인해 이 행렬은 확장성이 나쁘다. 그와 같은 경우에, 대개는 의사 역행렬이 선택되며, 하기 수학식은 $L \times 0$ 디코딩 행렬 D 를 산출한다.

수학식 11

$$D = [\Psi^H \Psi]^{-1} \Psi^H$$

[0045]

최종적으로 다음과 같은 수학식이 성립할 수 있다.

수학식 12

$$w(\Omega_s) = DY(\Omega_s)^*$$

[0047]

여기서, 가중치 $w(\Omega_s)$ 는 수학식 9에 대한 최소 에너지 해이다. 이하, 의사 역행렬을 이용하여 얻은 결과에 대해 설명한다.

[0048]

이하에서는 패닝 함수와 앰비소닉스 디코딩 행렬 간의 연계에 대해 설명한다. 앰비소닉스부터 시작하여, 개별 확성기에 대한 패닝 함수는 수학식 12를 이용하여 산출될 수 있다.

[0049]

수학식 13

$$\Xi = [Y(\Omega_1)^*, Y(\Omega_2)^*, \dots, Y(\Omega_S)^*]$$

[0050]

이 S 개의 입력 신호 방향의 모드 행렬, 예컨대, 1° 에서 180° 까지 1 도씩 단계적으로 증가하는 경사각과 1° 에서 360° 까지의 방위각을 가진 구면 그리드(spherical grid)라고 하자. 이 모드 행렬은 $0 \times S$ 개의 원소를 갖고 있다. 수학식 12를 이용하여 도출된 행렬 W 는 $L \times S$ 개의 원소를 갖고 있으며, 행 l 은 각자의 확성기에 대해 S 개의 패닝 가중치를 갖고 있다.

[0051]

수학식 14

$$W = D\Xi$$

[0052]

[0053] 대표적인 예로서, 단일 확성기(2)의 패닝 함수는 도 3에서 빔 패턴으로서 나타나 있다. 이 예에서 디코드 행렬 D의 차수 M=3이다. 도시된 바와 같이, 패닝 함수값은 확성기의 물리적 위치를 나타내지 않는다. 이는 선택된 차수에 대한 공간 샘플링 방식으로서 충분한 낮은 확성기의 수학적 불규칙적 위치 설정에 기인한다. 그러므로 디코드 행렬은 비규칙화 모드 행렬이라고 한다. 이 문제는 수학식 11에서 확성기 모드 행렬 Ψ 의 규칙화에 의해 극복될 수 있다. 이 해는 디코딩 행렬의 공간 분해능을 희생하는 것이며, 따라서 보다 낮은 앰비소닉스 차수로 표현될 수 있다. 도 4는 규칙화 모드 행렬을 이용하여, 특히 규칙화를 위한 모드 행렬의 고유값들(eigenvalues)의 평균을 이용하여 디코딩한 결과로서 생긴 예시적인 빔 패턴을 보여준다. 도 3과 비교해서, 다루어진 확성기의 방향은 이제 명확히 인식된다.

[0054] 배경 기술 부분에서 설명한 바와 같이, 패닝 함수가 이미 알려져 있는 경우에는 앰비소닉스 신호의 재생을 위한 디코딩 행렬 D를 얻는 다른 방법도 가능하다. 패닝 함수 W는 가상 소스 방향 Ω 세트 상에서 정의된 원하는 신호로 간주되며, 이들 방향의 모드 행렬 Ξ 은 입력 신호로서 이용된다. 그러면, 디코딩 행렬은 하기 수학식을 이용하여 산출될 수 있다.

수학식 15

[0055]
$$D = W \Xi^H [\Xi \Xi^H]^{-1} = W \Xi^+$$

[0056] 여기서, $\Xi^H [\Xi \Xi^H]^{-1}$ 또는 간단히 Ξ^+ 는 모드 행렬 Ξ 의 의사 역행렬이다. 이 새로운 방식에서는 VBAP로부터 W 패닝 함수를 취하고 이로부터 앰비소닉스 디코딩 행렬을 산출한다.

[0057] W 패닝 함수는 다시 수학식 4를 이용해 산출된 이득값 $g(\Omega)$ 로서 취해지며, Ω 는 수학식 13에 따라서 선택된다. 수학식 15를 이용하는 최종적인 디코드 행렬은 VBAP 패닝 함수를 용이하게 하는 앰비소닉스 디코딩 행렬이다. VBAP로부터 도출된 디코딩 행렬을 이용하여 디코딩한 결과로서 생긴 빔 패턴을 보여주는 도 5에 일례가 도시되어 있다. 양호하게도, 사이드로브 SL은 도 4의 규칙화 모드 매칭 결과의 사이드로브 SL_{reg} 보다 훨씬 작다. 더욱이, 개별적인 확성기에 대한 VBAP 도출 빔 패턴은, VBAP 패닝 함수가 다루어진 방향의 벡터 베이스에 의존함에 따라, 확성기 설정의 기하학적 형태를 따른다. 결과적으로, 본 발명에 따른 새로운 방식은 확성기 설정의 모든 방향에 대해 더 양호한 결과를 준다.

[0058] 소스 방향(103)은 상당히 자유롭게 정의될 수 있다. 소스 방향 S의 수에 대한 조건은 이것이 적어도 $(N+1)^2$ 이어야 한다는 것이다. 따라서, 사운드필드 신호 SF_c 의 특정 차수 N을 갖는다면, $S \geq (N+1)^2$ 에 따라서 S를 정의하고, S 소스 방향을 단위 구면에 고르게 분포시키는 것이 가능하다. 전술한 바와 같이, 결과는 1°에서 180°까지 x(예컨대 x=1...5 또는 x=10, 20 등)도씩 단계적으로 증가하는 경사각 θ 와 1°에서 360°까지의 방위각 ϕ 를 가진 구면 그리드일 수 있으며, 각 소스 방향 $\Omega=(\theta, \phi)$ 는 방위각 ϕ 와 경사각 θ 로 주어질 수 있다.

[0059] 듣기 평가에서 양호한 효과가 확인되었다. 단일 소스의 로컬라이제이션의 평가를 위해, 가상 소스가 기준으로서의 실제 소스와 비교된다. 실제 소스에 대해서는, 원하는 위치에 있는 확성기가 사용된다. 사용된 재생 방법은 VBAP, 앰비소닉스 모드 매칭 디코딩, 그리고, 본 발명에 따라 VBAP 패닝 함수를 이용하는 새로이 제시된 앰비소닉스 디코딩이다. 후자의 2가지 방법에 있어서는 각 평가 위치와 각 평가 입력 신호에 대해 3차(third order)의 앰비소닉스 신호가 생성된다. 그 후, 이 합성 앰비소닉스 신호는 대응하는 디코딩 행렬을 이용하여 디코딩된다. 사용된 평가 신호는 광대역 핑크 노이즈와 남성 음성 신호이다. 평가 위치는 다음과 같은 방향을 가진 정면 영역에 위치한다.

수학식 16

[0060]
$$\Omega_1 = (76.1^\circ, -23.2^\circ), \Omega_2 = (63.3^\circ, -4.3^\circ)$$

[0061] 듣기 평가는 대략 0.2초의 평균 잔향 시간을 가진 음향실에서 실시되었다. 이 듣기 평가에는 9명의 사람이 참

여했다. 평가 대상자에게 기준과 비교한 모든 재생 방법의 공간 재생 성능에 등급을 매기도록 하였다. 단일 등급값은 가상 소스의 로컬라이제이션과 음색 변화를 나타내는 것이어야 했다. 도 5는 듣기 평가의 결과를 보여준다.

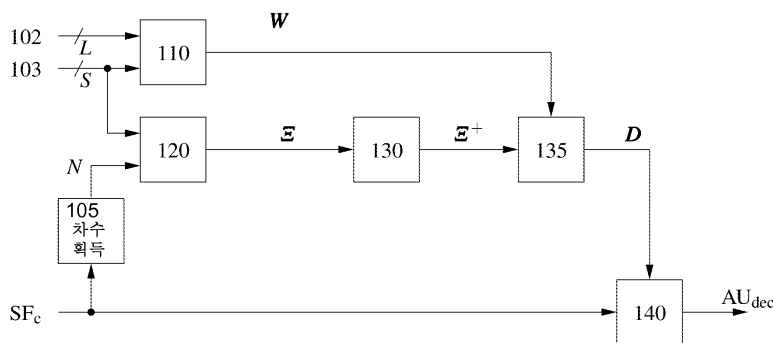
[0062] 결과가 보여주듯이, 비규칙화 앰비소닉스 모드 매칭 디코딩은 평가한 다른 방법보다 지각적으로 더 나쁜 등급이 매겨져 있다. 이 결과는 도 3에 대응한다. 앰비소닉스 모드 매칭 방법은 이 듣기 평가에서 앵커(anchor) 역할을 한다. 다른 이점은 잡음 신호에 대한 신뢰 구간이 다른 방법에서보다 VBAP에서 더 크다는 것이다. 평균값은 VBAP 패닝 함수를 이용한 앰비소닉스 디코딩에서 최고치를 보여준다. 따라서, 이용된 앰비소닉스 차수로 인해 공간 분해능이 감소되기는 하나, 이 방법은 파라메트릭 VBAP 방식에 비해 이점을 보여준다. VBAP에 비해, 강건 패닝 함수와 VBAP 패닝 함수를 가진 양쪽 앰비소닉스 디코딩은 가상 소스를 렌더링하는 데 3개의 확성기만 사용되는 것은 아니라는 이점을 갖고 있다. VBAP에서는, 가상 소스 위치가 확성기의 물리적 위치들 중 하나에 가까이 있다면 단일 확성기들이 우세할 수 있다. 대부분의 평가 대상자들은 직접 적용 VBAP에서보다는 앰비소닉스 구동 VBAP에서 음색 변화가 적었다고 했다. VBAP에서의 음색 변화 문제는 Pulkki에서 이미 알려져 있다. VBAP와는 달리, 상기 새로이 제시된 방법은 가상 소스의 재생을 위해 3개 초과 확성기를 사용하지만 놀랍게도 음색 변화가 더 적다.

[0063] 결론으로서, VBAP 패닝 함수로부터 앰비소닉스 디코딩 행렬을 얻는 새로운 방법이 개시된다. 이 방법은 여러 가지 서로 다른 확성기 설정에 있어서는 모드 매칭 방식의 행렬에 비해 유리하다. 이들 디코딩 행렬의 특성과 결과에 대해서는 전술하였다. 요약하면, VBAP 패닝 함수를 이용한 새로이 제시된 앰비소닉스 디코딩은 공지 모드 매칭 방법의 통상적인 문제를 방지한다. 듣기 평가는 VBAP 도출 앰비소닉스 디코딩이 VBAP의 직접적인 이용보다도 공간 재생 품질이 더 양호할 수 있다는 것을 보여주었다. VBAP는 렌더링될 가상 소스의 파라메트릭 기술(description)을 필요로 하지만, 이 제시된 방법은 사운드필드 기술만을 필요로 한다.

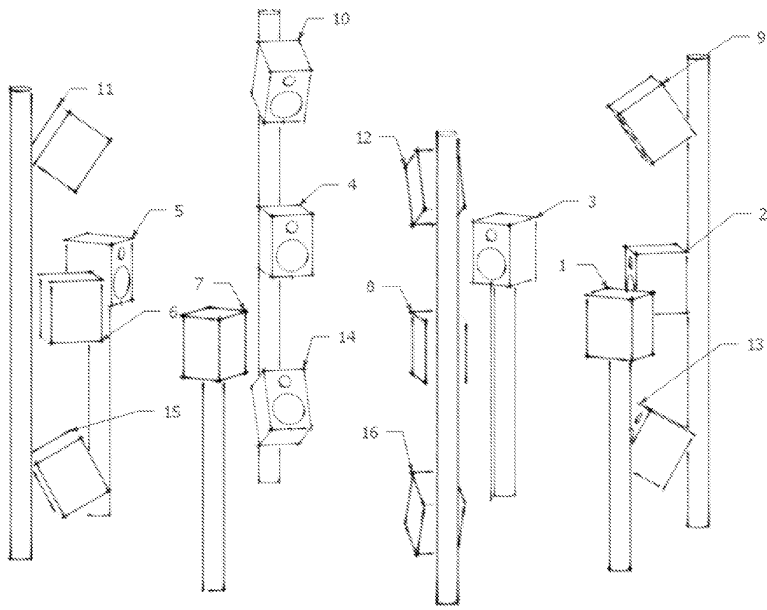
[0064] 지금까지 바람직한 실시예에 적용된 본 발명의 기본적인 신규한 특징들을 도시하고, 설명하고, 지적하였지만, 당업자라면 본 발명의 본질로부터 벗어남이 없이, 설명된 장치와 방법, 개시된 디바이스의 형태와 세부 사항, 그리고 그들의 동작에 있어 여러 가지 생략, 치환, 및 수정이 가능함을 잘 알 것이다. 동일한 결과를 달성하기 위해 실질적으로 동일한 기능을 실질적으로 동일한 방식으로 수행하는 구성요소들의 모든 조합이 본 발명의 범위 내에 속하는 것이다. 설명된 실시예들 간의 구성요소 치환도 충분히 의도하고 고려할 수 있다. 본 발명의 범위로부터 벗어남이 없이 세부 사항의 변경이 가능함을 알아야 한다. 상세한 설명 및 (적당한 경우) 청구범위 및 도면에 개시된 각 특징은 서로 독립적으로 또는 임의의 적당한 조합으로 제공될 수 있다. 특징들은 적당한 경우 하드웨어, 소프트웨어, 또는 이 둘의 조합으로 구현될 수 있다. 청구범위에 나타난 도면부호는 단지 예시적인 것이며 청구범위를 한정하는 것이 아니다.

도면

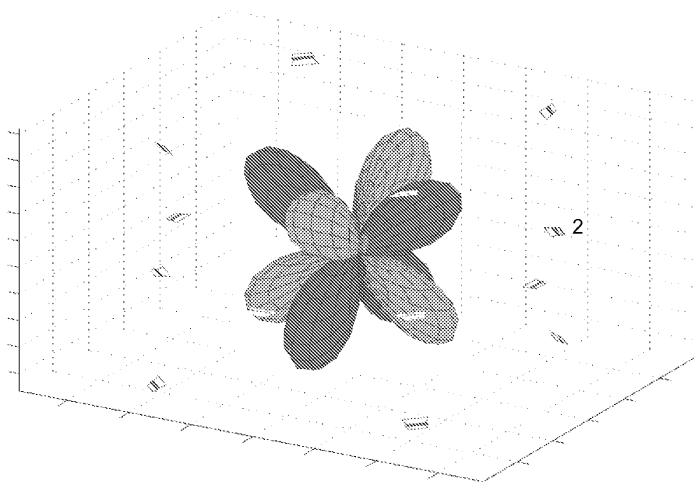
도면1



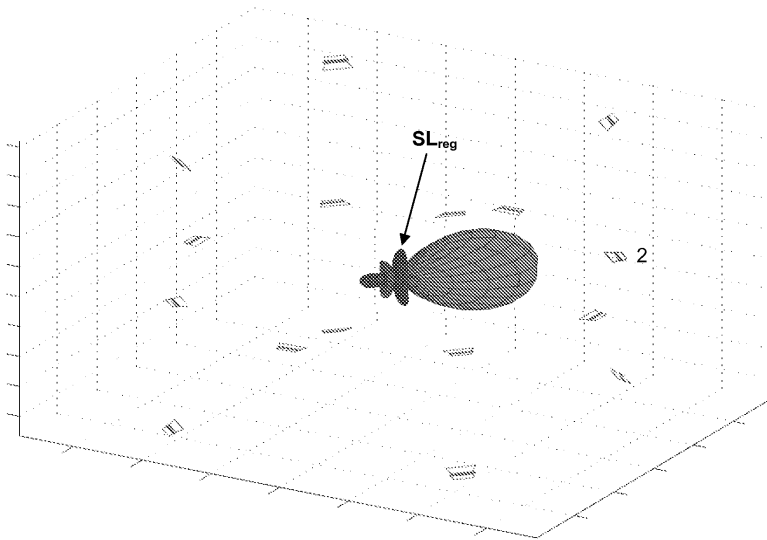
도면2



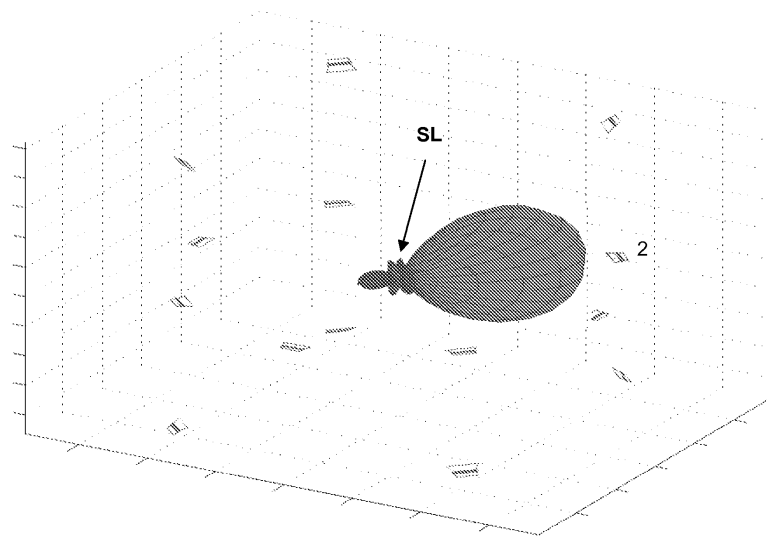
도면3



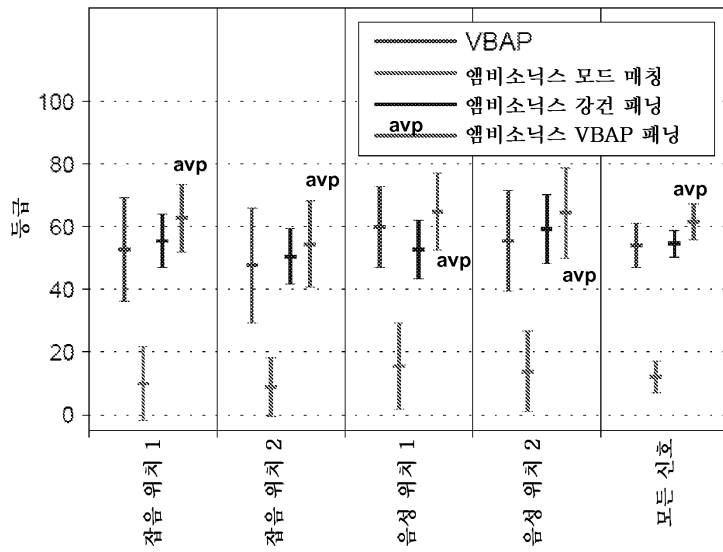
도면4



도면5



도면6



도면7

