



등록특허 10-2235398



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2021년04월02일
(11) 등록번호 10-2235398
(24) 등록일자 2021년03월29일

- (51) 국제특허분류(Int. Cl.)
H04S 3/02 (2006.01) *H04S 7/00* (2006.01)
- (52) CPC특허분류
H04S 3/02 (2013.01)
H04S 7/30 (2020.05)
- (21) 출원번호 10-2016-7010383
- (22) 출원일자(국제) 2014년10월20일
심사청구일자 2019년10월14일
- (85) 번역문제출일자 2016년04월20일
- (65) 공개번호 10-2016-0074501
- (43) 공개일자 2016년06월28일
- (86) 국제출원번호 PCT/EP2014/072411
- (87) 국제공개번호 WO 2015/059081
국제공개일자 2015년04월30일
- (30) 우선권주장
13290255.2 2013년10월23일
유럽특허청(EPO)(EP)
- (56) 선행기술조사문헌
KR100964353 B1
WO2011117399 A1

(73) 특허권자
돌비 인터네셔널 에이비
네델란드 1101 씨엔 암스트레담 주이두스트 헤리
커퍼로그백 1-35 3이 아폴로 빌딩

(72) 발명자
케일러, 플로리안
독일 30625 하노버 칼 비헤르트 알레 74 도이치
톰슨 오에이치쥐 리서치 앤드 이노베이션
빔, 요하네스
독일 30625 하노버 칼 비헤르트 알레 74 도이치
톰슨 오에이치쥐 리서치 앤드 이노베이션

(74) 대리인
양영준, 백만기

전체 청구항 수 : 총 11 항

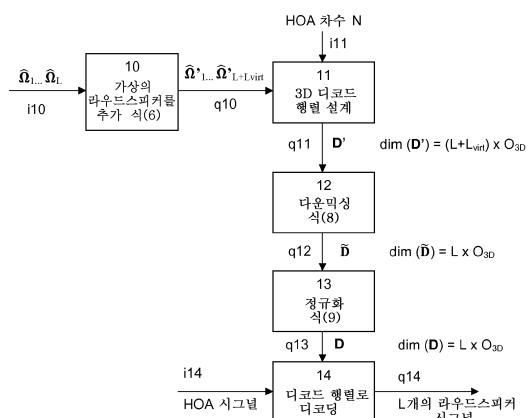
심사관 : 권영학

(54) 발명의 명칭 2D 셋업들을 이용하는 오디오 재생을 위해 앰비소닉스 오디오 음장 표현을 디코딩하기 위한 방법 및 장치

(57) 요약

3D의 사운드 션들은 자연 음장으로서 합성되거나 캡쳐될 수 있다. 디코딩하기 위해, 주어진 라우드스피커 셋업에 특정하고 알려진 라우드스피커 위치들을 이용하여 생성되는 디코드 행렬이 요구된다. 그러나, 몇몇 소스 방향들은 예를 들어 5.1 서라운드와 같은 2D 라우드스피커 셋업들에 대하여 감쇠된다. 알려진 위치들에서 L개의

(뒷면에 계속)

대 표 도 - 도1

라우드스피커를 위해 음장 포맷의 인코딩된 오디오 시그널을 디코딩하기 위한 개선된 방법은 L개의 라우드스피커의 위치들에 적어도 하나의 가상의 라우드스피커의 위치를 추가하는 단계(10), 3D 디코드 행렬(D')을 생성하는 단계(11) - L개의 라우드스피커의 위치들(공식 I) 및 적어도 하나의 가상의 위치(공식 II)가 이용됨 -, 3D 디코드 행렬(D')을 다운믹싱하는 단계(12), 및 다운스케일링된 3D 디코드 행렬(공식 III)을 이용하여, 인코딩된 오디오 시그널(i14)을 디코딩하는 단계(14)를 포함한다. 결과적으로, 복수의 디코딩된 라우드스피커 시그널(q14)이 획득된다.

(52) CPC특허분류

H04S 2420/11 (2013.01)

명세서

청구범위

청구항 1

알려진 위치들에 있는 L개의 라우드스피커(loudspeaker)를 위해 앰비소닉스 포맷(Ambisonics format)의 인코딩된 오디오 시그널을 디코딩하기 위한 방법으로서,

적어도 하나의 가상의 라우드스피커(virtual loudspeaker)의 적어도 하나의 가상의 위치를 상기 L개의 라우드스피커의 위치들에 추가하는 단계(10);

3D 디코드 행렬(\mathbf{D}')을 생성하는 단계(11) - 상기 L개의 라우드스피커의 위치들($\widehat{\Omega}_1 \dots \widehat{\Omega}_L$) 및 상기 적어도 하나의 가상의 라우드스피커의 적어도 하나의 가상의 위치($\widehat{\Omega}_{L+1}'$)가 이용되고, 상기 3D 디코드 행렬(\mathbf{D}')은 상기 L개의 라우드스피커 위치들 및 상기 가상의 라우드스피커 위치들을 위한 계수들을 가짐 -;

상기 3D 디코드 행렬(\mathbf{D}')을 다운믹싱(downmixing)하는 단계(12) - 상기 가상의 라우드스피커 위치들을 위한 상기 계수들은 상기 L개의 라우드스피커 위치들에 관한 계수들에 가중되고(weighted) 분산되며(distributed), 상기 L개의 라우드스피커 위치들을 위한 계수들을 갖는 다운스케일링된 3D 디코드 행렬($\tilde{\mathbf{D}}$)이 획득됨 -; 및

상기 다운스케일링된 3D 디코드 행렬($\tilde{\mathbf{D}}$)을 이용하여 상기 인코딩된 오디오 시그널(i14)을 디코딩하는 단계(14) - 복수의 디코딩된 라우드스피커 시그널(q14)이 획득됨 -

를 포함하는 방법.

청구항 2

◆ 청구항 2은(는) 설정등록료 납부시 포기되었습니다.◆

제1항에 있어서, 상기 가상의 라우드스피커 위치들을 위한 상기 계수들은 가중치 인자(weighting factor) $g = \frac{1}{\sqrt{L}}$ 로 가중되고, L은 라우드스피커들의 수인, 방법.

청구항 3

◆ 청구항 3은(는) 설정등록료 납부시 포기되었습니다.◆

제1항 또는 제2항에 있어서, 가상의 라우드스피커의 상기 적어도 하나의 가상의 위치($\widehat{\Omega}_{L+1}'$)는

$\widehat{\Omega}_{L+1}' = [0,0]^T$ 및 $\widehat{\Omega}_{L+1}' = [\pi, 0]^T$ 중 하나인, 방법.

청구항 4

◆ 청구항 4은(는) 설정등록료 납부시 포기되었습니다.◆

제1항에 있어서, 프로베니우스 놈(Frobenius norm)을 이용하여 상기 다운스케일링된 3D 디코드 행렬($\tilde{\mathbf{D}}$)을 정규화(normalizing)하는 단계(13)를 더 포함하고, 정규화되고 다운스케일링된 3D 디코드 행렬(\mathbf{D})이 획득되며, 상기 인코딩된 오디오 시그널을 디코딩하는 단계(14)는 상기 정규화되고 다운스케일링된 3D 디코드 행렬(\mathbf{D})을 이용하는, 방법.

청구항 5

$$\mathbf{D} = \frac{\tilde{\mathbf{D}}}{\sqrt{\sum_{l=1}^L \sum_{q=1}^{O_{3D}} |\tilde{d}_{l,q}|^2}}$$

제4항에 있어서, 상기 정규화하는 단계는 제4항에 있어서, 상기 정규화하는 단계는 \mathbf{D} 에 따라 수행되고, $\tilde{d}_{l,q}$ 는 상기 다운스케일링된 3D 디코드 행렬($\tilde{\mathbf{D}}$)의 계수이고, $\tilde{d}_{l,q} = d'_{l,q} + g \cdot d'_{L+1,q} + g \cdot d'_{L+2,q}$ for $l = 1, \dots, L$ and $q = 1, \dots, O_{3D}$ 에 의해 정의되는, 방법.

청구항 6

제1항에 있어서,

상기 L개의 라우드스피커의 위치들($\widehat{\Omega}_1 \dots \widehat{\Omega}_L$) 및 상기 인코딩된 오디오 시그널의 계수들의 차수 N을 결정하는 단계(101);

상기 위치들로부터, 상기 L개의 라우드스피커가 2D 평면에 있음을 결정하는 단계(102); 및

가상의 라우드스피커의 적어도 하나의 가상의 위치($\widehat{\Omega}'_{L+1}$)를 생성하는 단계(103)

를 더 포함하는 방법.

청구항 7

제1항에 있어서, 대역 통과 필터(band pass filter)들을 이용하여 상기 인코딩된 오디오 시그널을 복수의 주파수 대역으로 분리(separating)하는 단계를 더 포함하고, 각각의 주파수 대역에 대해 하나씩 복수의 분리된 3D 디코드 행렬(\mathbf{D}_b')이 생성되고(711b), 각각의 3D 디코드 행렬(\mathbf{D}_b')은 다운믹싱되고(712b) 선택적으로는 (optionally) 별개로 정규화되고(713b), 상기 인코딩된 오디오 시그널(i14)을 디코딩하는 단계(714b)는 각각의 주파수 대역을 위해 별개로 수행되는, 방법.

청구항 8

제1항에 있어서, 상기 알려진 L개의 라우드스피커의 위치들은 10° 를 넘지 않는 양각(elevation)들을 갖고서 하나의 2D 평면 내에 있는, 방법.

청구항 9

알려진 위치들에 있는 L개의 라우드스피커를 위해 앰비조닉스 포맷의 인코딩된 오디오 시그널을 디코딩하기 위한 장치로서,

적어도 하나의 가상의 라우드스피커의 적어도 하나의 가상의 위치를 상기 L개의 라우드스피커의 위치들에 추가하기 위한 추가 유닛(adder unit)(410);

3D 디코드 행렬(\mathbf{D}')을 생성하기 위한 디코드 행렬 생성기 유닛(decode matrix generator unit)(411) - 상기 L개의 라우드스피커의 위치들($\widehat{\Omega}_1 \dots \widehat{\Omega}_L$) 및 상기 적어도 하나의 가상의 라우드스피커의 적어도 하나의 가상의 위치($\widehat{\Omega}'_{L+1}$)가 이용되고, 상기 3D 디코드 행렬(\mathbf{D}')은 상기 L개의 라우드스피커 위치들 및 상기 가상의 라우드스피커 위치들을 위한 계수들을 가짐 -;

상기 3D 디코드 행렬(\mathbf{D}')을 다운믹싱하기 위한 행렬 다운믹싱 유닛(412) - 상기 가상의 라우드스피커 위치들을 위한 상기 계수들은 상기 L개의 라우드스피커 위치들에 관한 계수들에 가중되고 분산되며, 상기 L개의 라우드스피커 위치들을 위한 계수들을 갖는 다운스케일링된 3D 디코드 행렬($\tilde{\mathbf{D}}$)이 획득됨 -; 및

상기 다운스케일링된 3D 디코드 행렬($\tilde{\mathbf{D}}$)을 이용하여 상기 인코딩된 오디오 시그널(i14)을 디코딩하기 위한 디코딩 유닛(414) - 복수의 디코딩된 라우드스피커 시그널(q14)이 획득됨 -

을 포함하는 장치.

청구항 10

제9항에 있어서, 프로베니우스 놈을 이용하여 상기 다운스케일링된 3D 디코드 행렬($\tilde{\mathbf{D}}$)을 정규화하기 위한 정규화 유닛(413)을 더 포함하고, 정규화되고 다운스케일링된 3D 디코드 행렬(\mathbf{D})이 획득되고, 상기 디코딩 유닛(414)은 상기 정규화되고 다운스케일링된 3D 디코드 행렬(\mathbf{D})을 이용하는, 장치.

청구항 11

제9항 또는 제10항에 있어서,

상기 L개의 라우드스피커의 위치들($\widehat{\mathbf{\Omega}}_1 \dots \widehat{\mathbf{\Omega}}_L$) 및 상기 인코딩된 오디오 시그널의 계수들의 차수 N을 결정하기 위한 제1 결정 유닛(101);

상기 위치들로부터, 상기 L개의 라우드스피커가 2D 평면에 있음을 결정하기 위한 제2 결정 유닛(102); 및

가상의 라우드스피커의 적어도 하나의 가상의 위치($\widehat{\mathbf{\Omega}}_{L+1}'$)를 생성하기 위한 가상 라우드스피커 위치 생성 유닛(103)

을 더 포함하는 장치.

청구항 12

제9항에 있어서, 상기 인코딩된 오디오 시그널을 복수의 주파수 대역으로 분리하기 위한 복수의 대역 통과 필터(715b)를 더 포함하고, 각각의 주파수 대역에 대해 하나씩 복수의 분리된 3D 디코드 행렬(\mathbf{D}_b')이 생성되고(711b), 각각의 3D 디코드 행렬(\mathbf{D}_b')은 다운믹싱되고(712b) 선택적으로는 별개로 정규화되고, 상기 디코딩 유닛은 각각의 주파수 대역을 별개로 디코딩하는, 장치.

청구항 13

컴퓨터가 알려진 위치들에 있는 L개의 라우드스피커를 위해 앰비소닉스 포맷의 인코딩된 오디오 시그널을 디코딩하기 위한 방법을 수행하게 하는 실행 가능한 명령어들이 저장되어 있는 컴퓨터 판독 가능 저장 매체로서, 상기 방법은

적어도 하나의 가상의 라우드스피커의 적어도 하나의 가상의 위치를 상기 L개의 라우드스피커의 위치들에 추가하는 단계(10);

3D 디코드 행렬(\mathbf{D}')을 생성하는 단계(11) - 상기 L개의 라우드스피커의 위치들($\widehat{\mathbf{\Omega}}_1 \dots \widehat{\mathbf{\Omega}}_L$ 및 상기 적어도 하나의 가상의 라우드스피커의 적어도 하나의 가상의 위치($\widehat{\mathbf{\Omega}}_{L+1}'$)가 이용되고, 상기 3D 디코드 행렬(\mathbf{D}')은 상기 L개의 라우드스피커 위치들 및 상기 가상의 라우드스피커 위치들을 위한 계수들을 가짐 -;

상기 3D 디코드 행렬(\mathbf{D}')을 다운믹싱하는 단계(12) - 상기 가상의 라우드스피커 위치들을 위한 상기 계수들은 상기 L개의 라우드스피커 위치들에 관한 계수들에 가중되고 분산되며, 상기 L개의 라우드스피커 위치들을 위한 계수들을 갖는 다운스케일링된 3D 디코드 행렬($\tilde{\mathbf{D}}$)이 획득됨 -; 및

상기 다운스케일링된 3D 디코드 행렬($\tilde{\mathbf{D}}$)을 이용하여 상기 인코딩된 오디오 시그널(i14)을 디코딩하는 단계(14) - 복수의 디코딩된 라우드스피커 시그널(q14)이 획득됨 -

를 포함하는, 컴퓨터 판독 가능 저장 매체.

청구항 14

◆청구항 14은(는) 설정등록료 납부시 포기되었습니다.◆

제13항에 있어서, 상기 가상의 라우드스피커 위치들을 위한 상기 계수들은 가중치 인자 $g = \frac{1}{\sqrt{L}}$ 로 가중되고, L은 라우드스피커들의 수인, 컴퓨터 판독 가능 저장 매체.

청구항 15

제13항 또는 제14항에 있어서, 가상의 라우드스피커의 상기 적어도 하나의 가상의 위치($\widehat{\Omega}'_{L+1}$)는 $\widehat{\Omega}'_{L+1} = [0,0]^T$ 및 $\widehat{\Omega}'_{L+1} = [\pi,0]^T$ 중 하나인, 컴퓨터 판독 가능 저장 매체.

발명의 설명

기술 분야

[0001] 본 발명은 2D 또는 2D에 가까운(near-2D) 셋업을 이용하는 오디오 재생을 위해 오디오 음장 표현(audio soundfield representation), 특히 앰비소닉스 포맷의 오디오 표현(Ambisonics formatted audio representation)을 디코딩하기 위한 방법 및 장치에 관한 것이다.

배경 기술

[0002] 정확한 정위(location)는 임의의 공간 오디오 재생 시스템(spatial audio reproduction system)에 주된 목표이다. 그러한 재생 시스템들은 3D 사운드로부터 이득을 얻는 회의 시스템들, 게임들, 또는 다른 가상 환경들에 대해 고도로 적용될 수 있다. 3D의 사운드 씬(sound scene)들은 자연 음장(natural sound field)으로서 합성되거나 캡쳐될 수 있다. 예컨대, 앰비소닉스와 같은 음장 시그널들은 원하는 음장의 표현을 운반한다. 음장 표현으로부터 개별 라우드스피커(loudspeaker) 시그널들을 획득하기 위해 디코딩 프로세스가 요구된다. 앰비소닉스 포맷의 시그널을 디코딩하는 것은 또한 "렌더링(rendering)"으로 지칭된다. 오디오 씬들을 합성하기 위해, 공간적 라우드스피커 배열을 참조하는 패닝 기능들은 주어진 사운드 소스의 공간적 정위를 획득하기 위해 요구된다. 자연 음장을 레코딩(recording)하기 위해, 마이크로폰 어레이들은 공간적 정보를 캡처하도록 요구된다. 앰비소닉스 접근은 이를 성취하기 위한 매우 적합한 틀이다. 앰비소닉스 포맷의 시그널들은 음장의 구면 조화 분해(spherical harmonic decomposition)에 기반하는 원하는 음장의 표현을 운반한다. 기본 앰비소닉스 포맷이나 B-포맷은 0차 및 1차의 구면 조화 함수(spherical harmonic)들을 이용하는 반면, 소위 고차 앰비소닉스(HOA: Higher Order Ambisonics)는 적어도 2차의 추가 구면 조화 함수들도 이용한다. 라우드스피커들의 공간적 배열은 라우드스피커 셋업으로서 지정된다. 디코딩 프로세스에 대하여, 디코드 행렬(또한 렌더링 행렬로도 지칭됨)이 요구되고, 이는 주어진 라우드스피커 셋업에 특정하고, 알려진 라우드스피커 위치들을 이용하여 생성된다.

[0003] 보통 이용되는 라우드스피커 셋업들은 두 개의 라우드스피커를 채용하는 스테레오 셋업, 다섯 개의 라우드스피커를 이용하는 표준 서라운드 셋업(standard surround setup), 및 다섯 개보다 더 많은 라우드스피커를 이용하는 서라운드 셋업의 확장들이다. 그러나, 이러한 잘 알려진 셋업들은 2차원(2D)에 제한되고, 예컨대, 어떠한 높이 정보도 재생되지 않는다. 높이 정보를 재생할 수 있는 알려진 라우드스피커 셋업들을 위해 렌더링하는 것은 사운드 정위 및 배색(coloration)에서 단점을 갖는데, 즉, 공간적 수직의 팬(pan)들이 매우 불균등한(uneven) 음량(loudness)으로 인지되거나, 라우드스피커 시그널들이 강한 사이드 로브(side lobe)들을 가지며, 이는 특히 중심에서 벗어난(off-center) 리스닝 위치들에 불리하다. 따라서, 소위 에너지-보존 렌더링 설계는 라우드스피커들에 HOA 음장 서술을 렌더링하는 경우 선호된다. 이것은 단일 사운드 소스의 렌더링이 소스의 방향에 독립하여, 일정한 에너지의 라우드스피커 시그널들을 야기하는 것을 의미한다. 다시 말해서, 앰비소닉스 표현에 의해 운반되는 입력 에너지는 라우드스피커 렌더러(renderer)에 의해 보존된다. 본 발명자들의 국제 특허 공보 WO2014/012945A1[1]은 3D 라우드스피커 셋업들을 위한 양호한 에너지 보존 및 정위 속성을 갖는 HOA 렌더러 설계를 설명한다. 그러나, 이 접근은 모든 방향을 포함하는 3D 라우드스피커 셋업들에 대해 꽤 잘 동작 하지만, 몇몇 소스 방향들은 2D 라우드스피커 셋업들(예컨대, 5.1 서라운드 같은 것)을 위해 감쇠된다(attenuated). 이는 특히 어떠한 라우드스피커들도 위치하지 않는 방향들을 위해, 예를 들어 최상부로부터의 방향에 적용된다.

[0004] F.Zotter 및 M.Frank의 "올라운드 앰비소닉 패닝 및 디코딩(All-Round Ambisonic Panning and Decoding)"[2]에서, 라우드스피커들에 의해 구축된(built) 볼록 꼽질(convex hull)에서 홀이 있는 경우 "가상의(imaginary)" 라

우드스피커가 추가된다. 그러나, 그러한 가상의 라우드스피커를 위한 결과적인 시그널은 실제 라우드스피커 상에서의 재생을 위해서는 생략된다. 따라서, 그 방향(예컨대, 어떠한 실제 라우드스피커도 위치하지 않는 방향)으로부터 소스 시그널은 여전히 감쇠될 것이다. 추가로, 이 논문은 VBAP(vector base amplitude panning)와의 이용만을 위해 가상의 라우드스피커의 이용을 보여준다.

발명의 내용

[0005] 따라서, 2D (2차원) 라우드스피커 셋업들을 위해, 어떠한 라우드스피커들도 위치하지 않는 방향으로부터 사운드 소스들이 덜 감쇠되거나 전혀 감쇠되지 않는 에너지-보존 앰비소닉스 렌더리들을 설계하는 문제가 남아있다. 2D 라우드스피커 셋업들은 라우드스피커들의 앙각(elevation angle)들이 정의된 작은 범위(예컨대, $<10^\circ$) 내에 있어서 그 라우드스피커들이 수평면에 가까운 것으로서 분류될 수 있다.

[0006] 본 명세서는 정규의 또는 비정규의 공간적 라우드스피커 분포들을 위해 앰비소닉스 포맷의 오디오 음장 표현을 렌더링/디코딩하기 위한 솔루션을 설명하고, 여기서 렌더링/디코딩은 고도로 개선된 정위 및 배색 속성들을 제공하고 에너지 보존적이며, 어떠한 라우드스피커도 이용 가능하지 않는 방향들로부터의 사운드조차 렌더링된다. 유리하게도, 어떠한 라우드스피커도 이용 가능하지 않은 방향들로부터의 사운드는 라우드스피커가 각각의 방향에서 이용 가능했더라면 가졌을 것과 실질적으로 동일한 에너지 및 인지된 음량으로 렌더링된다. 물론, 이러한 사운드 소스들의 정확한 정위는 그것의 방향에서 어떠한 라우드스피커도 이용 가능하지 않기 때문에 가능하지 않다.

[0007] 특히, 적어도 몇몇 설명된 실시예들은 HOA 포맷에서 음장 데이터를 디코딩하기 위해 디코드 행렬을 획득하는 새로운 방식을 제공한다. 적어도 HOA 포맷은 라우드스피커 위치들에 직접적으로 관련되지 않는 음장을 설명하고, 획득되는 라우드스피커 시그널들은 반드시 채널 기반 오디오 포맷으로 되어 있으므로, HOA 시그널들의 디코딩은 언제나 오디오 시그널을 렌더링하는 것에 엄격하게 관련된다. 원칙적으로, 다른 오디오 음장 포맷들에도 동일하게 적용된다. 따라서, 본 개시내용은 음장 관련 오디오 포맷들을 디코딩하고 렌더링하는 것 모두에 관한 것이다. 디코드 행렬 및 렌더링 행렬이라는 용어는 동의어로서 이용된다.

[0008] 양호한 에너지 보존 속성들과 함께 주어진 셋업을 위한 디코드 행렬을 획득하기 위해, 하나 이상의 가상의 라우드스피커(virtual loudspeaker)는 어떠한 라우드스피커도 이용 가능하지 않는 위치들에 추가된다. 예를 들어, 2D 셋업을 위한 개선된 디코드 행렬을 획득하기 위해, 두 개의 가상의 라우드스피커는 상단 및 하단에 추가된다 (이것은 $+90^\circ$ 및 -90° 앙각에 대응하며, 2D 라우드스피커들은 대략 0° 의 앙각에 위치함). 이 가상의 3D 라우드스피커 셋업에 대하여, 디코드 행렬은 에너지 보존 속성을 만족시키도록 설계된다. 최종적으로, 가상의 라우드스피커들을 위한 디코드 행렬로부터의 가중치 인자(weighting factor)들은 2D 셋업의 실제 라우드스피커들에 대한 일정한 이득들과 맥락된다.

[0009] 일 실시예에 따르면, 라우드스피커들의 주어진 세트에 앰비소닉스 포맷의 오디오 시그널을 렌더링하거나 디코딩하기 위한 디코드 행렬(또는 렌더링 행렬)은 종래의 방법 및 수정된 라우드스피커 위치들을 이용하여 제1 예비(preliminary) 디코드 행렬을 생성하고 - 수정된 라우드스피커 위치들은 라우드스피커들의 주어진 세트의 라우드스피커 위치들 및 적어도 하나의 추가적인 가상의 라우드스피커 위치를 포함함 -, 제1 예비 디코드 행렬을 다운믹싱(downmixing)하는 것에 의해 생성되고, 적어도 하나의 추가적인 가상의 라우드스피커에 관한 계수들은 제거되고 라우드스피커들의 주어진 세트의 라우드스피커들에 관한 계수들에 분산된다(distributed). 일 실시예에서, 디코드 행렬을 정규화하는 후속 단계가 이어진다. 결과적인 디코드 행렬은 라우드스피커들의 주어진 세트에 앰비소닉스 시그널을 렌더링하거나 디코딩하는 것에 적절하고, 어떠한 라우드스피커도 존재하지 않는 위치들로부터의 사운드조차 올바른(correct) 시그널 에너지와 함께 재생된다. 이것은 개선된 디코드 행렬의 구조(construction) 때문이다. 바람직하게는, 제1 예비 디코드 행렬은 에너지 보존적이다.

[0010] 일 실시예에서, 디코드 행렬은 L개의 행 및 O_{3D} 개의 열을 가진다. 행들의 수는 2D 라우드스피커 셋업에서 라우드스피커들의 수에 대응하고, 열들의 수는 $O_{3D} = (N+1)^2$ 에 따른 HOA 차수 N에 의존하는 앰비소닉스 계수들 O_{3D} 의 수에 대응한다. 2D 라우드스피커 셋업을 위한 디코드 행렬의 각각의 계수들은 적어도 제1 중간 계수(intermediate coefficient) 및 제2 중간 계수의 합이다. 제1 중간 계수는 2D 라우드스피커 셋업의 현재의 라우드스피커 위치를 위한 에너지-보존 3D 행렬 설계 방법에 의해 획득되고, 에너지-보존 3D 행렬 설계 방법은 적어도 하나의 가상의 라우드스피커 위치를 이용한다. 제2 중간 계수는 적어도 하나의 가상의 라우드스피커 위치를 위한 상기 에너지-보존 3D 행렬 설계 방법으로부터 획득되는 계수에 가중치 인자 G 를 곱함으로써 획득된다.

일 실시예에서, 가중치 인자 $g = \frac{1}{\sqrt{L}}$ 에 따라 계산되고, L은 2D 라우드스피커 셋업의 라우드스피커들의 수이다.

[0011] 일 실시예에서, 본 발명은 컴퓨터가 청구항들에서 또는 앞서 개시된 방법의 단계들을 포함하는 방법을 수행하게 하는 실행 가능한 명령어들이 저장되어 있는 컴퓨터 판독 가능 저장 매체에 관한 것이다. 이 방법을 활용하는 장치는 청구항 9에 개시된다.

[0012] 유리한 실시예들은 종속 청구항들, 다음의 설명 및 도면들에 개시된다.

도면의 간단한 설명

[0013] 본 발명의 대표적인 실시예들은 첨부 도면들을 참조하여 설명된다.

도 1은 일 실시예에 따른 방법의 흐름도이다.

도 2는 다운믹싱된 HOA 디코드 행렬의 대표적인 구조이다.

도 3은 라우드스피커 위치들을 획득하고 수정하는 것에 대한 흐름도이다.

도 4는 일 실시예에 따른 장치의 블록 다이어그램이다.

도 5는 종래의 디코드 행렬로부터 야기되는 에너지 분포이다.

도 6은 실시예들에 따른 디코드 행렬로부터 야기되는 에너지 분포이다.

도 7은 상이한 주파수 대역들에서 별개로 최적화된 디코드 행렬들의 이용이다.

발명을 실시하기 위한 구체적인 내용

[0014] 도 1은 일 실시예에 따른 오디오 시그널, 특히 음장 시그널을 디코딩하기 위한 방법의 흐름도를 보여준다. 음장 시그널들을 디코딩하는 것은 일반적으로 오디오 시그널이 렌더링될 라우드스피커들의 위치들을 요구한다. L 개의 라우드스피커를 위한 그러한 라우드스피커 위치들 $\widehat{\Omega}_1 \dots \widehat{\Omega}_L$ 은 프로세스에 대한 입력(i10)이다. 위치들이 언급되는 경우, 본 명세서에서 실제로는 공간적 방향들을 의미한다는 것에 유의해야 하며, 즉, 라우드스피커들의 위치들은 그것의 경사각(inclination angle)들 θ_l 및 방위각(azimuth angle)들 ϕ_l 에 의해 정의되며, 그들은 벡터 $\widehat{\Omega}_l = [\theta_l, \phi_l]^T$ 로 조합된다. 다음으로, 가상의 라우드스피커의 적어도 하나의 위치가 추가된다(i10). 일 실시예에서, 프로세스(i10)에 입력되는 모든 라우드스피커 위치는 실질적으로 동일한 평면에 있고, 따라서 그들은 2D 셋업을 구성하고, 추가되는 적어도 하나의 가상의 라우드스피커는 이 평면 밖에 있다. 특히 유리한 일 실시예에서, 프로세스(i10)에 입력되는 모든 라우드스피커 위치는 실질적으로 동일한 평면에 있고 두 개의 가상의 라우드스피커의 위치들은 단계(10)에서 추가된다. 두 개의 가상의 라우드스피커의 유리한 위치들은 이하에서 설명된다. 일 실시예에서, 추가는 이하의 식(6)에 따라 수행된다. 추가하는 단계(10)는 q10에서 라우드스피커 각도들의 수정된 세트 $\widehat{\Omega}'_1 \dots \widehat{\Omega}'_{L+L_{virt}}$ 을 야기한다. L_{virt} 은 가상의 라우드스피커들의 수이다. 라우드스피커 각도들의 수정된 세트는 3D 디코드 행렬 설계 단계(11)에서 이용된다. 또한, HOA 차수 N(일반적으로 음장 시그널의 계수들의 차수)은 단계(11)에서 제공되는 것(i11)이 필요하다.

[0015] 3D 디코드 행렬 설계 단계(11)는 3D 디코드 행렬을 생성하기 위한 임의의 알려진 방법을 수행한다. 바람직하게는, 3D 디코드 행렬은 에너지-보존 타입의 디코딩/렌더링에 적절하다. 예를 들어, PCT/EP2013/065034에서 설명된 방법이 이용될 수 있다. 3D 디코드 행렬 설계 단계(11)는 $L' = L + L_{virt}$ 개의 라우드스피커 시그널을 렌더링하는 것에 적절한 디코드 행렬 또는 렌더링 행렬 D' 을 야기하고, L_{virt} 은 "가상의 라우드스피커 위치 추가" 단계(10)에서 추가된 가상의 라우드스피커 위치들의 수이다.

[0016] L개의 라우드스피커만이 물리적으로 이용 가능하므로, 3D 디코드 행렬 설계 단계(11)로부터 야기되는 디코드 행렬 D' 은 다운믹싱하는 단계(12)에서 L개의 라우드스피커에 맞춰 조정되는(adapted) 것이 필요하다. 이 단계는 디코드 행렬 D' 의 다운믹싱을 수행하고, 가상의 라우드스피커들에 관한 계수들은 존재하는 라우드스피커들에 관한 계수들에 가중되고 분산된다. 바람직하게는, 임의의 특정한 HOA 차수(즉, 디코드 행렬 D' 의 열)의 계수

들은 동일한 HOA 차수(즉, 디코드 행렬 \mathbf{D}' 의 동일한 열)의 계수들에 가중되고 추가된다. 하나의 예는 이하의 식(8)에 따른 다운믹싱이다. 다운믹싱하는 단계(12)는 L개의 행을 갖는, 즉, 디코드 행렬 \mathbf{D}' 보다 더 적은 행들을 갖지만 디코드 행렬 \mathbf{D}' 과 동일한 수의 열들을 갖는 다운믹싱된 3D 디코드 행렬 $\tilde{\mathbf{D}}$ 를 야기한다. 다시 말해서, 디코드 행렬 \mathbf{D}' 의 차원은 $(L+L_{\text{virt}}) \times O_{3D}$ 이고, 다운믹싱된 3D 디코드 행렬 $\tilde{\mathbf{D}}$ 의 차원은 $L \times O_{3D}$ 이다.

[0017] 도 2는 HOA 디코드 행렬 \mathbf{D}' 으로부터 다운믹싱된 HOA 디코드 행렬 $\tilde{\mathbf{D}}$ 의 대표적인 구조를 보여준다. HOA 디코드 행렬 \mathbf{D}' 은 2개의 가상의 라우드스피커 위치가 L개의 이용 가능한 라우드스피커 위치에 추가되는 것을 의미하는 $L+2$ 개의 행 및 O_{3D} 개의 열을 가지며, $O_{3D}=(N+1)^2$ 이고 N은 HOA 차수이다. 다운믹싱하는 단계(12)에서, HOA 디코드 행렬 \mathbf{D}' 의 행 $L+1$ 및 $L+2$ 의 계수들은 그들 각각의 열의 계수들에 가중되고 분산되며, 행 $L+1$ 및 $L+2$ 는 제거된다. 예를 들어, 행 $L+1$ 및 $L+2$ 각각의 제1 계수들 $d'_{L+1,1}$ 및 $d'_{L+2,1}$ 은 $d'_{1,1}$ 과 같은 각각의 남은 행의 제1 계수들에 가중되고 추가된다. 다운믹싱된 HOA 디코드 행렬 $\tilde{\mathbf{D}}$ 의 결과적인 계수 $\tilde{d}_{1,1}$ 은 $d'_{1,1}$, $d'_{L+1,1}$, $d'_{L+2,1}$ 및 가중치 인자 \mathbf{g} 의 함수이다. 동일한 방식으로, 예컨대, 다운믹싱된 HOA 디코드 행렬 $\tilde{\mathbf{D}}$ 의 결과적인 계수 $\tilde{d}_{2,1}$ 은 $d'_{2,1}$, $d'_{L+1,1}$, $d'_{L+2,1}$ 및 가중치 인자 \mathbf{g} 의 함수이고, 다운믹싱된 HOA 디코드 행렬 $\tilde{\mathbf{D}}$ 의 결과적인 계수 $\tilde{d}_{1,2}$ 는 $d'_{1,2}$, $d'_{L+1,2}$, $d'_{L+2,2}$ 및 가중치 인자 \mathbf{g} 의 함수이다.

[0018] 보통, 다운믹싱된 HOA 디코드 행렬 $\tilde{\mathbf{D}}$ 는 정규화 단계(13)에서 정규화될 것이다. 그러나, 이 단계(13)는 선택적인데, 왜냐하면 비정규화된 디코드 행렬 또한 음장 시그널을 디코딩하기 위해 이용될 수 있기 때문이다. 일 실시예에서, 다운믹싱된 HOA 디코드 행렬 $\tilde{\mathbf{D}}$ 는 이하의 식(9)에 따라 정규화된다. 정규화 단계(13)는 정규화되고 다운믹싱된 HOA 디코드 행렬 \mathbf{D} 를 야기하고, 이는 다운믹싱된 HOA 디코드 행렬 $\tilde{\mathbf{D}}$ 와 동일한 차원 $L \times O_{3D}$ 를 가진다.

[0019] 다음으로, 정규화되고 다운믹싱된 HOA 디코드 행렬 \mathbf{D} 는 음장 디코딩 단계(14)에서 이용될 수 있고, 입력 음장 시그널(i14)은 L개의 라우드스피커 시그널(q14)로 디코딩된다. 보통, 정규화되고 다운믹싱된 HOA 디코드 행렬 \mathbf{D} 는 라우드스피커 셋업이 수정될 때까지는 수정될 필요가 없다. 따라서, 일 실시예에서, 정규화되고 다운믹싱된 HOA 디코드 행렬 \mathbf{D} 는 디코드 행렬 저장소에 저장된다.

[0020] 도 3은 실시예에서 라우드스피커 위치들이 어떻게 획득되고 수정되는지 상세하게 보여준다. 이 실시예는 L개의 라우드스피커의 위치들 $\widehat{\mathbf{\Omega}}_1 \dots \widehat{\mathbf{\Omega}}_L$ 및 음장 시그널의 계수들의 차수 N을 결정하는 단계(101), 위치들로부터 L개의 라우드스피커가 실질적으로 2D 평면에 있는 것을 결정하는 단계(102), 및 가상의 라우드스피커의 적어도 하나의 가상의 위치 $\widehat{\mathbf{\Omega}}'_{L+1}$ 을 생성하는 단계(103)를 포함한다. 일 실시예에서, 적어도 하나의 가상의 위치 $\widehat{\mathbf{\Omega}}'_{L+1}$ 은 $\widehat{\mathbf{\Omega}}'_{L+1} = [0,0]^T$ 및 $\widehat{\mathbf{\Omega}}'_{L+1} = [\pi, 0]^T$ 중 하나이다. 일 실시예에서, 두 개의 가상의 라우드스피커에 대응하는 두 개의 가상의 위치 $\widehat{\mathbf{\Omega}}'_{L+1}$ 및 $\widehat{\mathbf{\Omega}}'_{L+2}$ 가 생성되며(103), $\widehat{\mathbf{\Omega}}'_{L+1} = [0,0]^T$ 및 $\widehat{\mathbf{\Omega}}'_{L+2} = [\pi, 0]^T$ 이다.

[0021] 일 실시예에 따르면, 알려진 위치들에서의 L개의 라우드스피커를 위해, 인코딩된 오디오 시그널을 디코딩하기 위한 방법은 L개의 라우드스피커의 위치들 $\widehat{\mathbf{\Omega}}_1 \dots \widehat{\mathbf{\Omega}}_L$ 및 음장 시그널의 계수들의 차수 N을 결정하는 단계(101), 위치들로부터 L개의 라우드스피커가 실질적으로 2D 평면에 있는 것을 결정하는 단계(102), 가상의 라우드스피커의 적어도 하나의 가상의 위치 $\widehat{\mathbf{\Omega}}'_{L+1}$ 을 생성하는 단계(103), 3D 디코드 행렬 \mathbf{D}' 를 생성하는 단계(11) - L개의 라우드스피커의 결정된 위치들 $\widehat{\mathbf{\Omega}}_1 \dots \widehat{\mathbf{\Omega}}_L$ 및 적어도 하나의 가상의 위치 $\widehat{\mathbf{\Omega}}'_{L+1}$ 이용되고, 3D 디

코드 행렬 \mathbf{D}' 은 상기 결정된 라우드스피커 위치들 및 가상의 라우드스피커 위치들을 위한 계수들을 가짐 -, 3D 디코드 행렬 \mathbf{D}' 을 다운믹싱하는 단계(12) - 가상의 라우드스피커 위치들을 위한 계수들은 결정된 라우드스피커 위치들에 관한 계수들에 가중되고 분산되며, 결정된 라우드스피커 위치들을 위한 계수들을 갖는 다운스케일링된 3D 디코드 행렬 $\tilde{\mathbf{D}}$ 가 획득됨 -, 및 다운스케일링된 3D 디코드 행렬 $\tilde{\mathbf{D}}$ 를 이용하여, 인코딩된 오디오 시그널(i14)을 디코딩하는 단계(14) - 복수의 디코딩된 라우드스피커 시그널(q14)이 획득됨 - 를 포함한다.

[0022] 일 실시예에서, 인코딩된 오디오 시그널은 예컨대, HOA 포맷의 음장 시그널이다. 일 실시예에서, 가상의 라우드스피커의 적어도 하나의 가상의 위치 $\widehat{\Omega}'_{L+1}$ 은 $\widehat{\Omega}'_{L+1} = [0, 0]^T$ 및 $\widehat{\Omega}'_{L+1} = [\pi, 0]^T$ 중 하나이다. 일 실시예에서, 가상의 라우드스피커 위치들을 위한 계수들은 가중치 요소 $g = \frac{1}{\sqrt{L}}$ 로 가중된다. 일 실시예에서, 방법은 다운스케일링된 3D 디코드 행렬 $\tilde{\mathbf{D}}$ 를 정규화하는 추가적인 단계를 가지고, 정규화되고 다운스케일링된 3D 디코드 행렬 \mathbf{D} 가 획득되며, 인코딩된 오디오 시그널(i14)을 디코딩하는 단계(14)는 정규화되고 다운스케일링된 3D 디코드 행렬 \mathbf{D} 를 이용한다. 일 실시예에서, 방법은 디코드 행렬 저장소에 다운스케일링된 3D 디코드 행렬 $\tilde{\mathbf{D}}$ 또는 정규화되고 다운믹싱된 HOA 디코드 행렬 \mathbf{D} 를 저장하는 추가적인 단계를 가진다.

[0023] 일 실시예에 따르면, 라우드스피커들의 주어진 세트에 음장 시그널을 렌더링하거나 디코딩하기 위한 디코드 행렬은 종래의 방법을 이용하고 수정된 라우드스피커 위치들을 이용하여 제1 예비 디코드 행렬을 생성하고 - 수정된 라우드스피커 위치들은 라우드스피커들의 주어진 세트의 라우드스피커 위치들 및 적어도 하나의 추가적인 가상의 라우드스피커 위치를 포함함 -, 제1 예비 디코드 행렬을 다운믹싱하는 것에 의해 생성되고, 적어도 하나의 추가적인 가상의 라우드스피커에 관한 계수들은 제거되고 라우드스피커들의 주어진 세트의 라우드스피커들에 관한 계수들에 분산된다. 일 실시예에서, 디코드 행렬을 정규화하는 후속 단계가 이어진다. 결과적인 디코드 행렬은 라우드스피커들의 주어진 세트에 음장 시그널을 렌더링하거나 디코딩하는 것에 적절하고, 어떠한 라우드스피커도 존재하지 않는 위치들로부터의 사운드조차 을바른 시그널 에너지와 함께 재생된다. 이것은 개선된 디코드 행렬의 구조 때문이다. 바람직하게는, 제1 예비 디코드 행렬은 에너지 보존적이다.

[0024] 도 4의 a)는 일 실시예에 따른 장치의 블록 다이어그램을 보여준다. 알려진 위치들에 있는 L개의 라우드스피커를 위해 음장 포맷의 인코딩된 오디오 시그널을 디코딩하기 위한 장치(400)는 적어도 하나의 가상의 라우드스피커의 적어도 하나의 위치를 L개의 라우드스피커의 위치들에 추가하기 위한 추가 유닛(addition unit)(410), 3D 디코드 행렬 \mathbf{D}' 을 생성하기 위한 디코드 행렬 생성기 유닛(decode matrix generator unit)(411) - L개의 라우드스피커의 위치들 $\widehat{\Omega}_1 \dots \widehat{\Omega}_L$ 및 적어도 하나의 가상의 위치 $\widehat{\Omega}'_{L+1}$ 이 이용되고, 3D 디코드 행렬 \mathbf{D}' 은 상기 결정된 라우드스피커 위치들 및 가상의 라우드스피커 위치들을 위한 계수들을 가짐 -, 3D 디코드 행렬 \mathbf{D}' 을 다운믹싱하기 위한 행렬 다운믹싱 유닛(412) - 가상의 라우드스피커 위치들을 위한 계수들은 결정된 라우드스피커 위치들에 관한 계수들에 가중되고 분산되며, 결정된 라우드스피커 위치들을 위한 계수들을 갖는 다운스케일링된 3D 디코드 행렬 $\tilde{\mathbf{D}}$ 가 획득됨 -, 및 다운스케일링된 3D 디코드 행렬 $\tilde{\mathbf{D}}$ 를 이용하여, 인코딩된 오디오 시그널을 디코딩하기 위한 디코딩 유닛(414) - 복수의 디코딩된 라우드스피커 시그널이 획득됨 - 을 포함한다.

[0025] 일 실시예에서, 장치는 다운스케일링된 3D 디코드 행렬 $\tilde{\mathbf{D}}$ 를 정규화하기 위한 정규화 유닛(413)을 더 포함하고, 정규화되고 다운스케일링된 3D 디코드 행렬 \mathbf{D} 가 획득되며, 디코딩 유닛(414)은 정규화되고 다운스케일링된 3D 디코드 행렬 \mathbf{D} 를 이용한다. 도 4의 b)에서 보여진 일 실시예에서, 장치는 L개의 라우드스피커의 위치들(Ω_L) 및 음장 시그널의 계수들의 차수 N을 결정하기 위한 제1 결정 유닛(4101), 위치들로부터 L개의 라우드스피커가 실질적으로 2D 평면에 있는 것을 결정하기 위한 제2 결정 유닛(4102), 및 가상의 라우드스피커의 적어도 하나의 가상의 위치($\widehat{\Omega}'_{L+1}$)를 생성하기 위한 가상 라우드스피커 위치 생성 유닛(4103)을 더 포함한다. 일 실시예에서, 장치는 인코딩된 오디오 시그널을 복수의 주파수 대역으로 분리(separating)하기 위한 복수의 대역 통과 필터(band pass filter)(715b)를 더 포함하고, 각각의 주파수 대역에 대해 하나씩 복수의 분리된 3D 디코드 행렬

\mathbf{D}_b' 이 생성되고(711b), 각각의 3D 디코드 행렬 \mathbf{D}_b' 은 다운믹싱되고(712b) 선택적으로는(optionally) 별개로 정규화되고, 디코딩 유닛(714b)은 각각의 주파수 대역을 별개로 디코딩한다. 이 실시예에서, 장치는 각각의 라우드스피커에 대해 하나씩 복수의 추가 유닛(716b)을 더 포함한다. 각각의 추가 유닛은 각각의 라우드스피커에 관한 주파수 대역들을 합산한다(add up).

[0026] 추가 유닛(410), 디코드 행렬 생성기 유닛(411), 행렬 다운믹싱 유닛(412), 정규화 유닛(413), 디코딩 유닛(414), 제1 결정 유닛(4101), 제2 결정 유닛(4102) 및 가상 라우드스피커 위치 생성기 유닛(4103) 중 각각은 하나 이상의 프로세서에 의해 구현될 수 있고, 이러한 유닛들 중 각각은 이러한 또는 다른 유닛들 중 임의의 다른 것과 동일한 프로세서를 공유할 수 있다.

[0027] 도 7은 입력 시그널의 상이한 주파수 대역들에서 별개로 최적화된 디코드 행렬들을 이용하는 실시예를 보여준다. 이 실시예에서, 디코딩 방법은 대역 통과 필터들을 이용하여, 인코딩된 오디오 시그널을 복수의 주파수 대역으로 분리하는 단계를 포함한다. 각각의 주파수 대역에 대해 하나씩 복수의 분리된 3D 디코드 행렬 \mathbf{D}_b' 이 생성되고(711b), 각각의 3D 디코드 행렬 \mathbf{D}_b' 은 다운믹싱되고(712b) 선택적으로는 별개로 정규화된다. 인코딩된 오디오 시그널을 디코딩하는 것(714b)은 각각의 주파수 대역에서 별개로 수행된다. 이는 인간의 인지(human perception)에서의 주파수-의존 차이점들이 고려될 수 있고, 상이한 주파수 대역들에 대해 상이한 디코드 행렬들을 야기할 수 있다는 장점을 가진다. 일 실시예에서, 디코드 행렬 중 단 하나 또는 그 이상(전부는 아님)만이 가상의 라우드스피커 위치들을 추가하고, 다음으로 앞서 설명한 것처럼 존재하는 라우드스피커 위치들을 위한 계수들에 그것들의 계수들을 가중하고 분산시키는 것에 의해 생성된다. 다른 실시예에서, 각각의 디코드 행렬들은 가상의 라우드스피커 위치들을 추가하고, 다음으로 앞서 설명한 것처럼 존재하는 라우드스피커 위치들을 위한 계수들에 그것들의 계수들을 가중하고 분산시키는 것에 의해 생성된다. 최종적으로, 동일한 라우드스피커에 관한 모든 주파수 대역은 주파수 대역 분할(frequency band splitting)의 반대의 동작으로, 라우드스피커 당 하나의 주파수 대역 추가 유닛(716b)에서 합산된다.

[0028] 추가 유닛(410), 디코드 행렬 생성기 유닛(711b), 행렬 다운믹싱 유닛(712b), 정규화 유닛(713b), 디코딩 유닛(714b), 주파수 대역 추가 유닛(716b) 및 대역 통과 필터 유닛(715b) 중 각각은 하나 이상의 프로세서에 의해 구현될 수 있고, 이러한 유닛들 중 각각은 이러한 또는 다른 유닛들 중 임의의 다른 것과 동일한 프로세서를 공유할 수 있다.

[0029] 본 개시내용의 일 양태는 양호한 에너지 보존 속성들과 함께 2D 셋업을 위한 렌더링 행렬을 획득하는 것이다. 일 실시예에서, 두 개의 가상의 라우드스피커는 상단 및 하단에 추가된다(이것은 $+90^\circ$ 및 -90° 양각에 대응하며, 2D 라우드스피커들은 대략 0° 의 양각에 위치함). 이 가상의 3D 라우드스피커 셋업에 대하여, 렌더링 행렬은 에너지 보존 속성을 만족시키도록 설계된다. 최종적으로, 가상의 라우드스피커들을 위한 렌더링 행렬로부터의 가중치 인자들은 2D 셋업의 실제 라우드스피커들에 대한 일정한 이득들과 맥심된다. 이하에서는, 앰비소닉스(특히, HOA) 렌더링이 설명된다. 앰비소닉스 렌더링은 앰비소닉스 음장 설명으로부터의 라우드스피커 시그널들의 계산(computation) 프로세스이다. 종종, 이는 앰비소닉스 디코딩으로도 지칭된다. N 차수의 3D 앰비소닉스 음장 표현이 고려되는데, 계수들의 수는 다음과 같다.

$$O_{3D} = (N+1)^2 \quad (1)$$

[0031] 시간 샘플 t 에 대한 계수들은 O_{3D} 개의 요소와 함께 벡터 $\mathbf{b}(t) \in \mathbb{C}^{O_{3D} \times 1}$ 에 의해 나타내어진다. 렌더링 행렬 $\mathbf{D} \in \mathbb{C}^{L \times O_{3D}}$ 에서, 시간 샘플 t 에 대한 라우드스피커 시그널들은

$$\mathbf{w}(t) = \mathbf{D} \mathbf{b}(t) \quad (2)$$

[0033] 예 의해 계산되고, $\mathbf{D} \in \mathbb{C}^{L \times O_{3D}}$ 이고 $\mathbf{w} \in \mathbb{R}^{L \times 1}$ 이며, L은 라우드스피커들의 수이다.

[0034] 라우드스피커들의 위치들은 $l = 1, \dots, L$ 에 대하여, 그것의 경사각 θ_l 및 방위각 ϕ_l 에 의해 정의되며, 이들은 벡터 $\hat{\Omega}_l = [\theta_l, \phi_l]^T$ 로 조합된다. 리스닝 위치로부터의 상이한 라우드스피커 거리들은 라우드스피커 채널들에 대하여 개별 지연들을 이용하는 것에 의해 보상된다. HOA 도메인에서 시그널 에너지는

[0035] $E = \mathbf{b}^H \mathbf{b}$ (3)

[0036] 예 의해 주어지고, \cdot^H 는 {켤레 복소(conjugate complex)} 전치됨(transposed)을 의미한다. 라우드스피커 시 그널들의 대응하는 에너지는 다음에 의해 계산된다.

[0037] $\hat{E} = \mathbf{w}^H \mathbf{w} = \mathbf{b}^H \mathbf{D}^H \mathbf{D} \mathbf{b}$ (4)

[0038] 에너지-보존 디코딩/렌더링을 달성하기 위해, 에너지 보존 디코드/렌더링 행렬에 대한 비율 \hat{E}/E 가 일정해야만 한다.

[0039] 원칙적으로, 개선된 2D 렌더링을 위한 다음의 확장이 제안된다. 2D 라우드스피커 셋업들을 위한 렌더링 행렬들의 설계를 위해, 하나 이상의 가상의 라우드스피커가 추가된다. 2D 셋업들은 라우드스피커들의 양각들이 정의된 작은 범위 내에 있어서 그것들이 수평면에 가까운 것으로서 이해된다. 이는 다음에 의해 표현될 수 있다.

[0040] $|\theta_l - \frac{\pi}{2}| \leq \theta_{thres2d}; l = 1, \dots, L$ (5)

[0041] 임계값 $\theta_{thres2d}$ 는 일 실시예에서 보통 5° 내지 10° 의 범위 내의 값에 대응하도록 선택된다.

[0042] 렌더링 설계에 대하여, 라우드스피커 각도들 $\widehat{\Omega}'_l$ 의 수정된 세트가 정의된다. 마지막 라우드스피커 위치들(이 예에서는 두 개)은 극 좌표계의 북점(north pole) 및 남점(south pole)(수직 방향으로, 즉, 상단 및 하단)에서 두 개의 가상의 라우드스피커의 위치들이다.

[0043] $\widehat{\Omega}'_l = \widehat{\Omega}_l; l = 1, \dots, L$

[0044] $\widehat{\Omega}'_{L+1} = [0, 0]^T$ (6)

[0045] $\widehat{\Omega}'_{L+2} = [\pi, 0]^T$

[0046] 따라서, 렌더링 설계를 위해 이용되는 라우드스피커의 새로운 수는 $L' = L + 2$ 이다. 이러한 수정된 라우드스피커 위치들로부터, 렌더링 행렬 $\mathbf{D}' \in \mathbb{C}^{(L+2) \times O_{3D}}$ 은 에너지 보존 접근으로 설계된다. 예를 들어, [1]에서 설명된 설계 방법이 이용될 수 있다. 이제 본래의 라우드스피커 셋업을 위한 최종적인 렌더링 행렬은 \mathbf{D}' 로부터 도출된다. 하나의 아이디어는 실제의 라우드스피커들에 행렬 \mathbf{D}' 에서 정의된 가상의 라우드스피커를 위한 가중치 인자들을 맥심하는 것이다. 고정된 이득 인자는 다음으로 선택되어 이용된다.

[0047] $g = \frac{1}{\sqrt{L}}$ (7)

[0048] 중간 행렬(intermediate matrix) $\tilde{\mathbf{D}} \in \mathbb{C}^{L \times O_{3D}}$ (본 명세서에서 다운스케일링된 3D 디코드 행렬로도 지칭됨)의 계수들은

[0049] $l = 1, \dots, L$ 및 $q = 1, \dots, O_{3D}$ 에 대해 $\tilde{d}_{l,q} = d'_{l,q} + g \cdot d'_{L+1,q} + g \cdot d'_{L+2,q}$ (8)

[0050] 예 의해 정의되고, $\tilde{d}_{l,q}$ 는 l 번째 행 및 q 번째 열에서의 $\tilde{\mathbf{D}}$ 의 행렬 요소이다. 선택적인 최종의 단계에서, 중간 행렬(다운스케일링된 3D 디코드 행렬)은 프로베니우스 놈(Frobenius norm)을 이용하여 정규화된다.

[0051]
$$\mathbf{D} = \frac{\tilde{\mathbf{D}}}{\sqrt{\sum_{l=1}^L \sum_{q=1}^{O_{3D}} |\tilde{d}_{l,q}|^2}}$$
 (9)

[0052] 도 5 및 도 6은 5.0 서라운드 라우드스피커 셋업을 위한 에너지 분포들을 보여준다. 두 도면들에서, 에너지 값들은 그레이스케일(greyscale)들로서 보여지고, 원들은 라우드스피커 위치들을 나타낸다. 개시된 방법에서, 특

히 상단에서의 감쇠는(여기서 보여지지 않지만, 하단에서도) 분명히 감소된다.

[0053] 도 5는 종래의 디코드 행렬로부터 야기되는 에너지 분포를 보여준다. $z=0$ 평면 근처의 작은 원들은 라우드스피커 위치들을 나타낸다. 볼 수 있는 것처럼, 에너지 범위 [-3.9, ..., 2.1] dB이 포괄되고, 이는 6 dB의 에너지 차이들을 야기한다. 추가로, 단위 구(unit sphere)의 상단으로부터의(보여지지 않지만, 하단에서도) 시그널들은 매우 낮은 에너지로, 즉, 들을 수 없게 재생되는데, 왜냐하면 여기서는 어떠한 라우드스피커들도 이용 가능하지 않기 때문이다.

[0054] 도 6은 하나 이상의 실시예에 따른 디코드 행렬로부터 야기되는 에너지 분포를 보여주며, 도 5에서와 동일한 수량의 라우드스피커들이 동일한 위치들에 있다. 적어도 다음의 장점들이 제공된다: 첫째로, 더 작은 에너지 범위 [-1.6, ..., 0.8] dB이 포괄되고, 이는 더 작은 2.4 dB의 에너지 차이들만을 야기한다. 둘째로, 단위 구의 모든 방향으로부터의 시그널들은 이용 가능한 라우드스피커들이 거기에 없더라도 그것들의 올바른 에너지와 함께 재생된다. 이러한 시그널들이 이용 가능한 라우드스피커들을 통하여 재생되므로, 그것들의 정위는 올바르지는 않지만, 시그널들은 올바른 음량으로 들을 수 있다. 이 예에서, 상단으로부터의 시그널들과 하단에서의(보이지 않음) 시그널들은 개선된 디코드 행렬을 이용한 디코딩으로 인하여 들을 수 있게 된다.

[0055] 실시예에서, 알려진 위치들에 있는 L개의 라우드스피커를 위해 앰비소닉스 포맷의 인코딩된 오디오 시그널을 디코딩하기 위한 방법은 L개의 라우드스피커의 위치들에 적어도 하나의 가상의 라우드스피커의 적어도 하나의 위치를 추가하는 단계, 3D 디코드 행렬 \mathbf{D}' 을 생성하는 단계 - L개의 라우드스피커의 위치들 $\Omega_1 \dots \Omega_L$ 및 적어도 하나의 가상의 위치 Ω'_{L+1} 이 이용되고, 3D 디코드 행렬 \mathbf{D}' 은 상기 결정된 라우드스피커 위치들 및 가상의 라우드스피커 위치들을 위한 계수들을 가짐 -, 3D 디코드 행렬 \mathbf{D}' 을 다운믹싱하는 단계 - 가상의 라우드스피커 위치들을 위한 계수들은 결정된 라우드스피커 위치들에 관한 계수들에 가중되고 분산되며, 결정된 라우드스피커 위치들을 위한 계수들을 갖는 다운스케일링된 3D 디코드 행렬 $\tilde{\mathbf{D}}$ 가 획득됨 -, 및 다운스케일링된 3D 디코드 행렬 $\tilde{\mathbf{D}}$ 를 이용하여, 인코딩된 오디오 시그널을 디코딩하는 단계 - 복수의 디코딩된 라우드스피커 시그널이 획득됨 - 를 포함한다.

[0056] 다른 실시예에서, 알려진 위치들에 있는 L개의 라우드스피커를 위해 앰비소닉스 포맷의 인코딩된 오디오 시그널을 디코딩하기 위한 장치는 L개의 라우드스피커의 위치들에 적어도 하나의 가상의 라우드스피커의 적어도 하나의 위치를 추가하기 위한 추가 유닛(410), 3D 디코드 행렬 \mathbf{D}' 을 생성하기 위한 디코드 행렬 생성기 유닛(411) - L개의 라우드스피커의 위치들 $\Omega_1 \dots \Omega_L$ 및 적어도 하나의 가상의 위치 Ω'_{L+1} 이 이용되고, 3D 디코드 행렬 \mathbf{D}' 은 상기 결정된 라우드스피커 위치들 및 가상의 라우드스피커 위치들을 위한 계수들을 가짐 -, 3D 디코드 행렬 \mathbf{D}' 을 다운믹싱하기 위한 행렬 다운믹싱 유닛(412) - 가상의 라우드스피커 위치들을 위한 계수들은 결정된 라우드스피커 위치들에 관한 계수들에 가중되고 분산되며, 결정된 라우드스피커 위치들을 위한 계수들을 갖는 다운스케일링된 3D 디코드 행렬 $\tilde{\mathbf{D}}$ 가 획득됨 -, 및 다운스케일링된 3D 디코드 행렬 $\tilde{\mathbf{D}}$ 를 이용하여, 인코딩된 오디오 시그널을 디코딩하기 위한 디코딩 유닛(414) - 복수의 디코딩된 라우드스피커 시그널이 획득됨 - 을 포함한다.

[0057] 또 다른 실시예에서, 알려진 위치들에 있는 L개의 라우드스피커를 위해 앰비소닉스 포맷의 인코딩된 오디오 시그널을 디코딩하기 위한 장치는 적어도 하나의 프로세서 및 적어도 하나의 메모리를 포함하며, 메모리는 프로세서 상에서 실행될 때 L개의 라우드스피커의 위치들에 적어도 하나의 가상의 라우드스피커의 적어도 하나의 위치를 추가하기 위한 추가 유닛(410), 3D 디코드 행렬 \mathbf{D}' 을 생성하기 위한 디코드 행렬 생성기 유닛(411) - L개의 라우드스피커의 위치들 $\Omega_1 \dots \Omega_L$ 및 적어도 하나의 가상의 위치 Ω'_{L+1} 이 이용되고, 3D 디코드 행렬 \mathbf{D}' 은 상기 결정된 라우드스피커 위치들 및 가상의 라우드스피커 위치들을 위한 계수들을 가짐 -, 3D 디코드 행렬 \mathbf{D}' 을 다운믹싱하기 위한 행렬 다운믹싱 유닛(412) - 가상의 라우드스피커 위치들을 위한 계수들은 결정된 라우드스피커 위치들에 관한 계수들에 가중되고 분산되며, 결정된 라우드스피커 위치들을 위한 계수들을 갖는 다운스케일링된 3D 디코드 행렬 $\tilde{\mathbf{D}}$ 가 획득됨 -, 및 다운스케일링된 3D 디코드 행렬 $\tilde{\mathbf{D}}$ 를 이용하여, 인코딩된 오디오 시그널을 디코딩하기 위한 디코딩 유닛(414) - 복수의 디코딩된 라우드스피커 시그널이 획득됨 - 을 구현하는

저장된 명령어들을 가진다.

[0058] 또 다른 실시예에서, 컴퓨터 관독 가능 저장 매체는 컴퓨터가 알려진 위치들에 있는 L개의 라우드스피커를 위해 앰비소닉스 포맷의 인코딩된 오디오 시그널을 디코딩하기 위한 방법을 수행하게 하는 저장된 실행 가능한 명령어들을 가지고, 방법은 L개의 라우드스피커의 위치들에 적어도 하나의 가상의 라우드스피커의 적어도 하나의 위치를 추가하는 단계, 3D 디코드 행렬 \mathbf{D}' 을 생성하는 단계 - L개의 라우드스피커의 위치들 $\hat{\Omega}_1 \dots \hat{\Omega}_L$ 및 적어도 하나의 가상의 위치 $\hat{\Omega}'_{L+1}$ 이 이용되고, 3D 디코드 행렬 \mathbf{D}' 은 상기 결정된 라우드스피커 위치들 및 가상의 라우드스피커 위치들을 위한 계수들을 가짐 -, 3D 디코드 행렬 \mathbf{D}' 을 다운믹싱하는 단계 - 가상의 라우드스피커 위치들을 위한 계수들은 결정된 라우드스피커 위치들에 관한 계수들에 가중되고 분산되며, 결정된 라우드스피커 위치들을 위한 계수들을 갖는 다운스케일링된 3D 디코드 행렬 $\tilde{\mathbf{D}}$ 가 획득됨 -, 및 다운스케일링된 3D 디코드 행렬 $\tilde{\mathbf{D}}$ 를 이용하여, 인코딩된 오디오 시그널을 디코딩하는 단계 - 복수의 디코딩된 라우드스피커 시그널이 획득됨 - 를 포함한다. 추가로, 컴퓨터 관독 가능 저장 매체의 실시예들은 앞서 설명된 임의의 특징들, 특히 청구항 1을 다시 참조하는 종속 청구항들에서 개시된 특징들을 포함할 수 있다.

[0059] 본 발명은 순수하게 예시의 방식으로만 설명되었고, 세부 사항의 수정들은 본 발명의 범위를 벗어나지 않고 만들어질 수 있다고 이해될 것이다. 예를 들어, HOA에 관하여만 설명되었다 하더라도, 본 발명은 또한 다른 음장 오디오 포맷들에 적용될 수 있다. 본 설명 및 (적당한) 청구항들 및 도면들에 개시된 각각의 특징은 독립적으로 또는 임의의 적절한 조합으로 제공될 수 있다. 특징들은 하드웨어, 소프트웨어, 또는 두 개의 조합으로 적절하게 구현될 수 있다. 청구항들에 나타나는 참조 번호들은 오직 예시이고 청구항들의 범위를 한정하는 효과를 갖지 않는다.

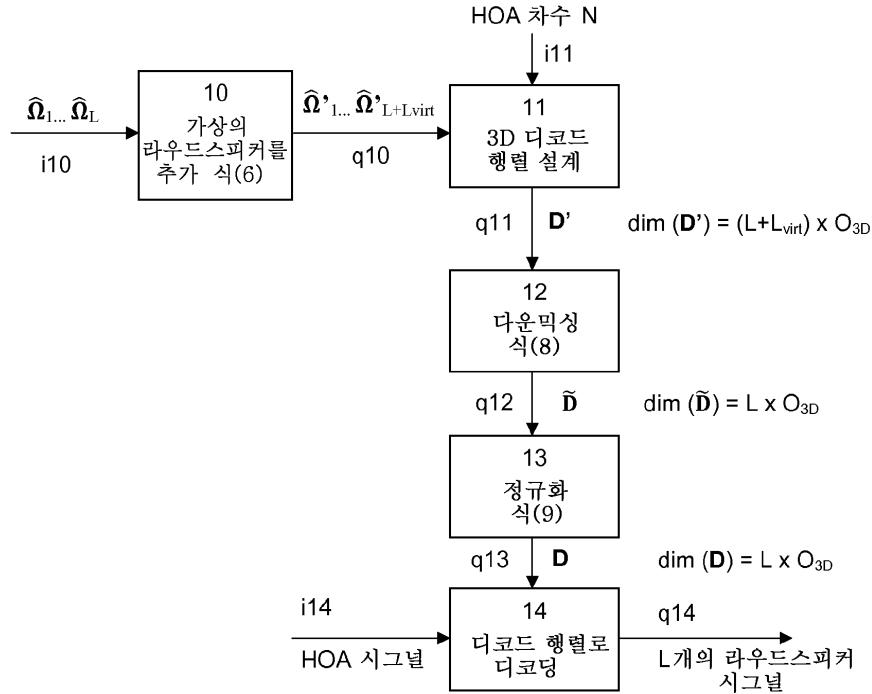
[0060] 다음의 참조들은 앞서 인용되었다.

[0061] [1] 국제 특허 공보 No.WO2014/012945A1 (PD120032)

[0062] [2] F.Zotter 및 M.Frank, "올라운드 앰비소닉 패닝 및 디코딩(All-Round Ambisonic Panning and Decoding)", J.Audio Eng. Soc., 2012, Vol. 60, pp. 807-820

도면

도면1

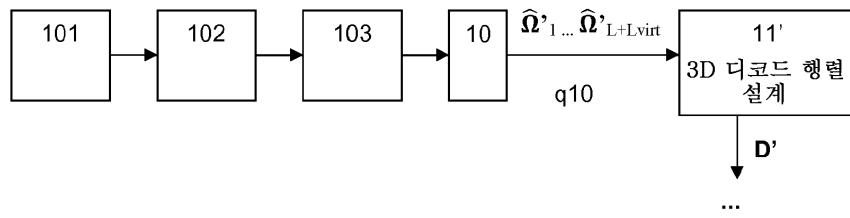


도면2

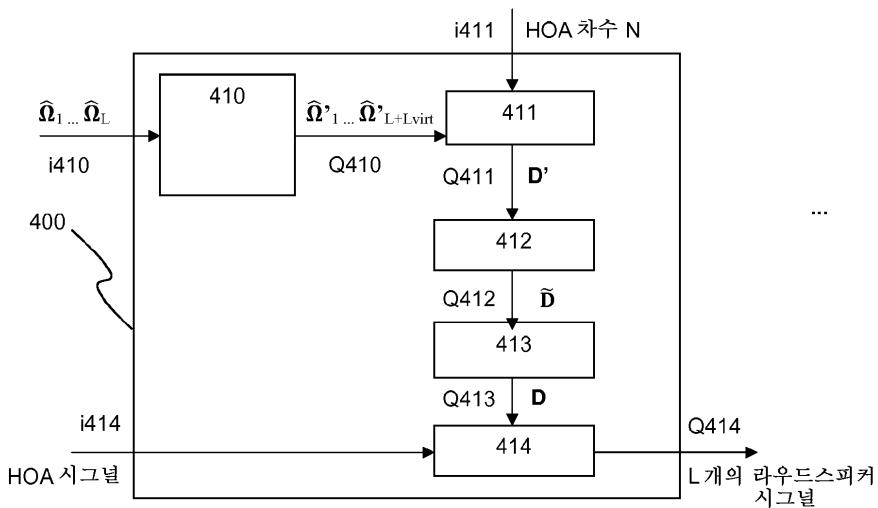
$$\mathbf{D}' = \begin{bmatrix} d'_{1,1} & d'_{1,2} & \dots & d'_{1,O3D} \\ \dots & \dots & \dots & \dots \\ d'_{L,1} & d'_{L,2} & \dots & d'_{L,O3D} \\ d'_{L+1,1} & d'_{L+1,2} & \dots & d'_{L+1,O3D} \\ d'_{L+2,1} & d'_{L+2,2} & \dots & d'_{L+2,O3D} \end{bmatrix} \rightarrow \tilde{\mathbf{D}} = \begin{bmatrix} \tilde{d}_{1,1} & \tilde{d}_{1,2} & \dots & \tilde{d}_{1,O3D} \\ \dots & \dots & \dots & \dots \\ \tilde{d}_{L,1} & \tilde{d}_{L,2} & \dots & \tilde{d}_{L,O3D} \end{bmatrix}$$

$$\mathbf{D}' = \begin{bmatrix} d'_{1,1} & d'_{1,2} & \dots & d'_{1,O3D} \\ \dots & \dots & \dots & \dots \\ d'_{L,1} & d'_{L,2} & \dots & d'_{L,O3D} \\ d'_{L+1,1} & d'_{L+1,2} & \dots & d'_{L+1,O3D} \\ d'_{L+2,1} & d'_{L+2,2} & \dots & d'_{L+2,O3D} \end{bmatrix} \rightarrow \tilde{\mathbf{D}} = \begin{bmatrix} \tilde{d}_{1,1} & \tilde{d}_{1,2} & \dots & \tilde{d}_{1,O3D} \\ \dots & \dots & \dots & \dots \\ \tilde{d}_{L,1} & \tilde{d}_{L,2} & \dots & \tilde{d}_{L,O3D} \end{bmatrix}$$

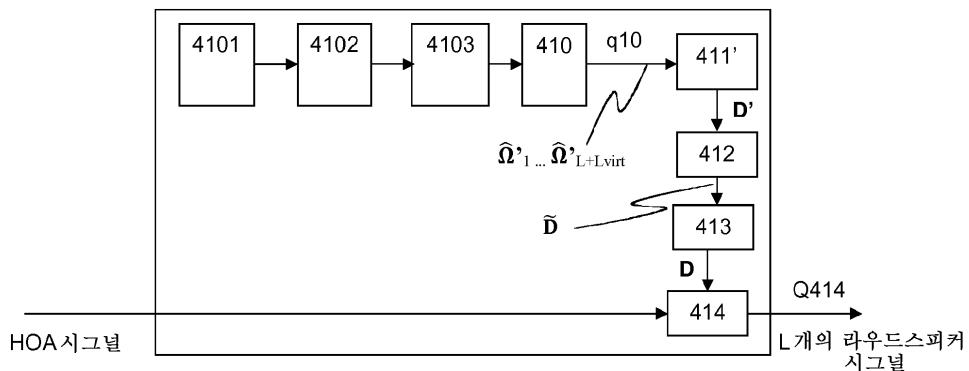
도면3



도면4

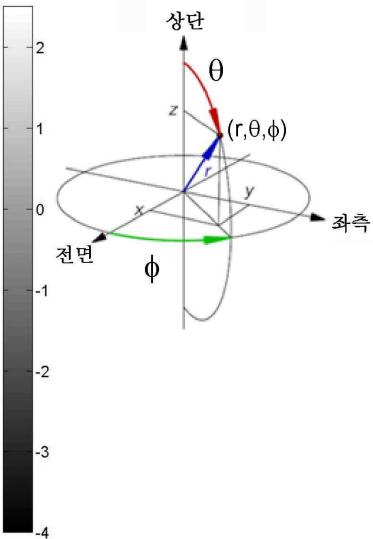
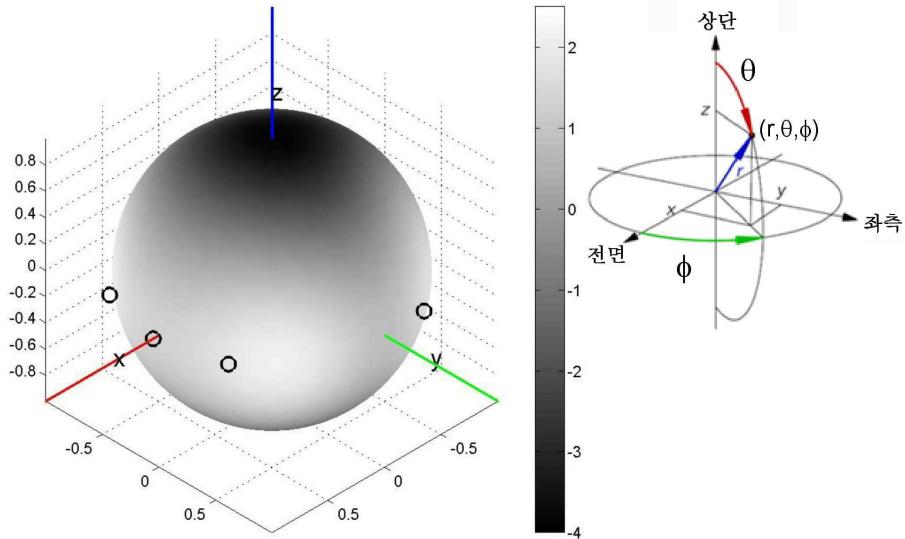


a)

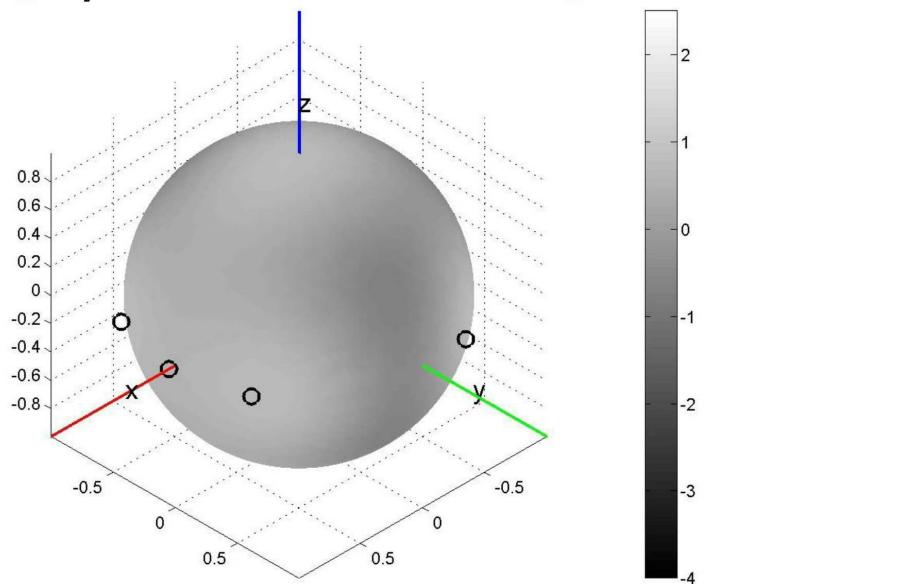


b)

도면5

 S_2 상에서의 dB 단위의 에너지, [-3.9, 2.1] dB

도면6

 S_2 상에서의 dB 단위의 에너지, [-1.6, 0.8] dB

도면7

