



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2020-0040745
(43) 공개일자 2020년04월20일

- (51) 국제특허분류(Int. Cl.)
H04S 7/00 (2006.01) G06F 3/0346 (2013.01)
G10L 19/008 (2014.01) G10L 21/0272 (2013.01)
H04S 3/00 (2006.01) H04S 3/02 (2006.01)
- (52) CPC특허분류
H04S 7/304 (2013.01)
G06F 3/0346 (2013.01)
- (21) 출원번호 10-2020-7001183
- (22) 출원일자(국제) 2018년07월13일
심사청구일자 2020년01월14일
- (85) 번역문제출일자 2020년01월14일
- (86) 국제출원번호 PCT/EP2018/069140
- (87) 국제공개번호 WO 2019/012131
국제공개일자 2019년01월17일
- (30) 우선권주장
17181488.2 2017년07월14일
유럽특허청(EPO)(EP)

- (71) 출원인
프라온호퍼 게젤샤프트 쭈르 피르데룽 데어 안겐
반텐 포르슘 에. 베.
독일 80686 뮌헨 한자슈트라쎄 27 체
- (72) 발명자
헤레, 쥐르겐
라스버거 스트라쎄 10에이, 91054 에를랑겐 독일
하베츠, 엠마누엘
슈베덴스트라쎄 13, 91080 스파르도르프 독일
- (74) 대리인
한상수

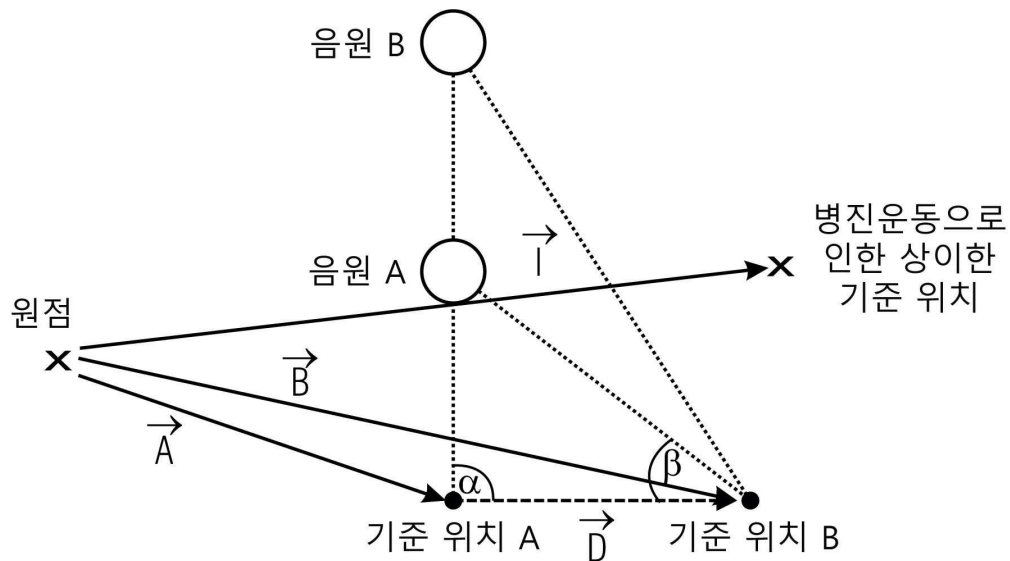
전체 청구항 수 : 총 50 항

(54) 발명의 명칭 다중-지점 음장 묘사를 이용하여 증강된 음장 묘사 또는 수정된 음장 묘사를 생성하기 위한 개념

(57) 요약

증강된 음장 묘사를 생성하기 위한 장치로서, 적어도 하나의 기준 위치에 대한 음장을 나타내는 적어도 하나의 음장 묘사를 생성하기 위한 음장 생성기(100, 250, 260); 및 음장의 공간적 정보와 관련된 메타데이터를 생성하기 위한 메타데이터 생성기(110)를 포함하고, 적어도 하나의 음장 묘사와 메타데이터가 증강된 음장 묘사를 구성한다. 적어도 하나의 음장 묘사는, 특정 경우에, 적어도 하나의 기준 위치와 관련된 제1 음장 묘사 및 기준 위치와 상이한 추가적인 기준 위치와 관련된 제2 음장 묘사를 포함하고, 공간적 정보와 관련된 메타데이터는 기준 위치 및 추가적인 기준 위치 또는 기준 위치 및 양 기준 위치 사이에서 연장되는 벡터 중 하나를 나타낸다.

대표도 - 도4c



(52) CPC특허분류

G10L 19/008 (2013.01)

G10L 21/0272 (2013.01)

H04S 3/008 (2013.01)

H04S 3/02 (2013.01)

H04S 7/306 (2013.01)

H04S 2400/11 (2013.01)

H04S 2400/15 (2013.01)

H04S 2420/01 (2013.01)

H04S 2420/03 (2013.01)

명세서

청구범위

청구항 1

증강된 음장 묘사를 생성하기 위한 장치에 있어서,
 적어도 하나의 기준 위치에 대한 음장을 나타내는 적어도 하나의 음장 묘사를 생성하기 위한 음장 생성기; 및
 상기 음장에 관한 공간적 정보와 관련된 메타데이터를 생성하기 위한 메타데이터 생성기를 포함하고,
 상기 적어도 하나의 음장 묘사와 상기 메타데이터가 증강된 음장 묘사를 구성하는 것을 특징으로 하는 증강된 음장 묘사를 생성하기 위한 장치.

청구항 2

제1항에 있어서,
 상기 음장 생성기는 상기 기준 위치와 관련된 제1 음장 묘사 및 추가적인 기준 위치와 관련된 제2 음장 묘사를 생성하도록 구성되고,
 상기 추가적인 기준 위치는 상기 기준 위치와 상이하며,
 상기 메타데이터 생성기는, 상기 기준 위치 및 상기 추가적인 기준 위치를 개별적으로 메타데이터로서 결정하도록, 또는 상기 기준 위치와 상기 추가적인 기준 위치 사이의 거리 또는 상기 기준 위치와 상기 추가적인 기준 위치 사이의 위치 벡터를 메타데이터로서 결정하도록 구성되는 것을 특징으로 하는 증강된 음장 묘사를 생성하기 위한 장치.

청구항 3

제2항에 있어서,
 상기 제1 음장 묘사는 제1 앰비소닉 묘사이고 상기 제2 음장 묘사는 제2 앰비소닉 묘사이거나, 또는
 상기 제1 음장 묘사와 상기 제2 음장 묘사가 앰비소닉 또는 DirAC 묘사 중 어느 하나인 것을 특징으로 하는 증강된 음장 묘사를 생성하기 위한 장치.

청구항 4

제1항 내지 제3항 중 어느 한 항에 있어서,
 제1 기하형태적 묘사가, 미리 결정된 원점으로부터 제1 음장 묘사의 기준 위치까지 지향된 제1 벡터에 관한 정보이고, 그리고
 제2 기하형태적 묘사가, 미리 결정된 원점으로부터 제2 음장 묘사의 추가적인 기준 위치까지 지향된 제2 벡터에 관한 정보인 것을 특징으로 하는 증강된 음장 묘사를 생성하기 위한 장치.

청구항 5

제1항 내지 제4항 중 어느 한 항에 있어서,
 제1 기하 형태적 묘사가, 미리 결정된 원점으로부터 제1 음장 묘사의 기준 위치까지 지향된 제1 벡터 및 미리 결정된 원점으로부터 제2 음장 묘사의 추가적인 기준 위치까지 지향된 제2 벡터 중 하나에 관한 정보, 그리고

상기 기준 위치와 상기 추가적인 기준 위치 사이의 벡터에 관한 정보인 것을 특징으로 하는 증강된 음장 묘사를 생성하기 위한 장치.

청구항 6

제1항 내지 제5항 중 어느 한 항에 있어서,

상기 기준 위치 및 상기 추가적인 기준 위치 중 하나가 미리 결정된 원점이고, 상기 메타데이터는 상기 미리 결정된 원점과 상기 기준 위치 및 상기 추가적인 기준 위치 중 다른 하나 사이의 벡터에 관한 정보를 포함하는 것을 특징으로 하는 증강된 음장 묘사를 생성하기 위한 장치.

청구항 7

제1항 내지 제6항 중 어느 한 항에 있어서,

상기 음장 생성기는, 실제 마이크로폰 디바이스를 이용하여 또는 가상 마이크로폰 기술을 이용하는 사운드 합성에 의해서, 상기 제1 음장 묘사 또는 상기 제2 음장 묘사를 생성하도록 구성되는 것을 특징으로 하는 증강된 음장 묘사를 생성하기 위한 장치.

청구항 8

음장 묘사 및 음장 묘사의 공간적 정보와 관련된 메타데이터로부터 수정된 음장 묘사를 생성하기 위한 장치에 있어서,

상기 공간적 정보, 상기 음장 묘사, 및 기준 위치로부터 상이한 기준 위치로의 병진 운동을 나타내는 병진 운동 정보를 이용하여 수정된 음장을 계산하기 위한 음장 계산기를 포함하는 것을 특징으로 하는 음장 묘사 및 음장 묘사의 공간적 정보와 관련된 메타데이터로부터 수정된 음장 묘사를 생성하기 위한 장치.

청구항 9

제8항에 있어서,

상기 음장 계산기는, 신호-기반 병진 운동 방법(signal-based translation method)을 이용하여, 수정된 음장 묘사를 계산하도록 구성되는 것을 특징으로 하는 음장 묘사 및 음장 묘사의 공간적 정보와 관련된 메타데이터로부터 수정된 음장 묘사를 생성하기 위한 장치.

청구항 10

제8항 또는 제9항에 있어서,

상기 음장 계산기가,

상기 제1 기준 위치와 관련된 제1 음장 묘사를 수신하고 상기 기준 위치와 상이한 추가적인 기준 위치와 관련된 제2 음장 묘사를 수신하고,

하나 이상의 추출된 객체의 제1 그룹을 획득하기 위한 제1 음장 묘사 및 하나 이상의 추출된 객체의 제2 그룹을 획득하기 위한 제2 음장 묘사로의 음원 분리를 실시하고,

추출된 객체에 대한 도달 방향(DoA) 데이터를 계산하고,

제1 음장 묘사의 추출된 객체의 제1 그룹 및 제2 음장 묘사의 추출된 객체의 제2 그룹을 매칭시키고,

하나 이상의 매칭된 객체의 위치를 추정하고, 그리고

상기 매칭된 객체의 추정된 위치 및 상기 상이한 기준 위치를 기초로 하나 이상의 매칭된 객체를 수정하도록 구성되는 것을 특징으로 하는 음장 묘사 및 음장 묘사의 공간적 정보와 관련된 메타데이터로부터 수정된 음장 묘사를 생성하기 위한 장치.

청구항 11

제10항에 있어서,

상기 음장 계산기가,

상기 매칭된 객체의 도달 방향 데이터(direction of arrival data)를 계산하고, 그리고

상기 상이한 기준 위치에 대한 계산된 도달 방향 데이터를 이용하여 각각의 매칭된 객체에 대한 음장 묘사를 결정하도록 구성되는 것을 특징으로 하는 음장 묘사 및 음장 묘사의 공간적 정보와 관련된 메타데이터로부터 수정된 음장 묘사를 생성하기 위한 장치.

청구항 12

제9항 내지 제11항 중 어느 한 항에 있어서,

상기 음장 계산기는, 비-매칭된 객체(non-matched objects)를 위해서 획득된 도달 방향 데이터를 이용하여 하나 이상의 비-매칭된 객체에 대한 음장 데이터를 계산하도록 구성되는 것을 특징으로 하는 음장 묘사 및 음장 묘사의 공간적 정보와 관련된 메타데이터로부터 수정된 음장 묘사를 생성하기 위한 장치.

청구항 13

제9항 내지 제12항 중 어느 한 항에 있어서,

상기 음장 계산기는, 상기 하나 이상의 매칭된 객체 및 상기 하나 이상의 비-매칭된 객체에 대한 음장 묘사 데이터를 조합하는 것에 의해서 상기 수정된 음장 묘사를 계산하도록 구성되는 것을 특징으로 하는 음장 묘사 및 음장 묘사의 공간적 정보와 관련된 메타데이터로부터 수정된 음장 묘사를 생성하기 위한 장치.

청구항 14

제8항에 있어서,

상기 음장 계산기는 파라메트릭-기반의 병진 운동 방법(parametric-based translation method)을 실시하도록 구성되는 것을 특징으로 하는 음장 묘사 및 음장 묘사의 공간적 정보와 관련된 메타데이터로부터 수정된 음장 묘사를 생성하기 위한 장치.

청구항 15

제8항 또는 제14항에 있어서,

상기 음장 계산기는,

음장 모델을 가정하고,

가정된 음장의 신호 성분 및/또는 매개변수를 추정하고,

상기 병진 운동 정보에 따라 또는 상기 상이한 기준 위치에 따라 상기 신호 성분 및/또는 매개변수를 수정하고, 그리고

상기 수정된 신호 성분 및/또는 상기 수정된 매개변수를 이용하여 상기 수정된 음장 묘사를 생성하도록 구성되

는 것을 특징으로 하는 음장 묘사 및 음장 묘사의 공간적 정보와 관련된 메타데이터로부터 수정된 음장 묘사를 생성하기 위한 장치.

청구항 16

제15항에 있어서,

상기 음장 계산기는, 상기 음장을 하나 이상의 다이렉트 사운드 성분 및 하나 이상의 확산 사운드 성분으로 분해하는 것을 이용하여 또는 상기 음장을 하나 이상의 다이렉트/주 사운드 성분 및 잔류 사운드 성분으로 분해하는 것을 이용하여 상기 신호 성분 및/또는 상기 매개변수를 추정하도록 구성되고, 상기 잔류 사운드 성분은 단일-채널 신호 또는 다중-채널 신호일 수 있는 것을 특징으로 하는 음장 묘사 및 음장 묘사의 공간적 정보와 관련된 메타데이터로부터 수정된 음장 묘사를 생성하기 위한 장치.

청구항 17

제8항 내지 제16항 중 어느 한 항에 있어서,

상기 음장 계산기는 상기 수정된 음장 묘사의 회전을 적용하도록 구성되는 것을 특징으로 하는 음장 묘사 및 음장 묘사의 공간적 정보와 관련된 메타데이터로부터 수정된 음장 묘사를 생성하기 위한 장치.

청구항 18

제8항 내지 제12항 중 어느 한 항에 있어서,

상기 수정된 음장 묘사가 앰비소닉 음장 묘사인 것을 특징으로 하는 음장 묘사 및 음장 묘사의 공간적 정보와 관련된 메타데이터로부터 수정된 음장 묘사를 생성하기 위한 장치.

청구항 19

제8항 내지 제18항 중 어느 한 항에 있어서,

상기 음장 계산기(420)는,

상기 음장 묘사로서, 제1 음장 묘사 및 제2 음장 묘사를 수신하고,

상기 제1 및 제2 음장 묘사에서 음원 분리를 수행하여 상기 제1 및 제2 음장 묘사의 음원 및 추출된 음원에 대한 도달 방향(DoA) 데이터를 추출하고,

각각의 추출된 음원에 대해서, 상기 DoA 데이터 및 상기 병진 운동 정보를 이용하여 상기 상이한 위치와 관련된 수정된 DoA 데이터를 계산하고,

상기 수정된 음장 묘사를 획득하기 위해서 상기 추출된 음원 및 상기 수정된 DoA 데이터를 프로세스하도록 구성되는 것을 특징으로 하는 음장 묘사 및 음장 묘사의 공간적 정보와 관련된 메타데이터로부터 수정된 음장 묘사를 생성하기 위한 장치.

청구항 20

제8항 내지 제19항 중 어느 한 항에 있어서,

상기 음장 계산기(420)는,

각각의 음장 묘사에 대해서 음원 분리를 개별적으로 실시하고 그리고 각각의 음장 묘사에 대해서 주변/확산/잔류 신호를 추출하도록 구성되는 것을 특징으로 하는 음장 묘사 및 음장 묘사의 공간적 정보와 관련된 메타데이터로부터 수정된 음장 묘사를 생성하기 위한 장치.

청구항 21

제8항 내지 제20항 중 어느 한 항에 있어서,

병진 운동 정보 또는 상기 수정된 음장에 대한 의도된 청취자의 회전을 나타내는 회전 정보를 제공하기 위한 병진 운동 인터페이스(410);

상기 메타데이터를 상기 음장 계산기(420)에 공급하기 위한 메타데이터 공급기(402, 400);

상기 음장 묘사를 상기 음장 계산기(420)에 공급하기 위한 음장 공급기(404, 400); 및

상기 수정된 음장 묘사 및 상기 병진 운동 정보를 이용하여 상기 메타데이터로부터 유도된 수정된 메타데이터를 포함하는 수정된 음장을 출력하기 위한, 또는 각각이 미리 규정된 라우드스피커 위치와 관련되는, 복수의 라우드스피커 채널을 출력하기 위한, 또는 상기 수정된 음장의 바이노럴 표상을 출력하기 위한 출력 인터페이스(421)를 더 포함하는 것을 특징으로 하는 음장 묘사 및 음장 묘사의 공간적 정보와 관련된 메타데이터로부터 수정된 음장 묘사를 생성하기 위한 장치.

청구항 22

제8항 내지 제21항 중 어느 한 항에 있어서,

상기 음장 묘사가 복수의 음장 성분을 포함하고, 상기 복수의 음장 성분은 전방향 성분 및 적어도 하나의 방향성 성분을 포함하며,

상기 음장 계산기는,

상이한 주파수 빈(different frequency bins)들에 대해서, 도달 방향 정보를 유도하기 위해서 상기 음장 성분을 분석하기 위한 음장 분석기(422);

거리 정보를 상기 주파수 빈에 의해서 표시된 음원에 연관시키는 깊이 맵을 포함하는 메타데이터와 상기 방향 정보를 이용하여 주파수 빈 마다 수정된 도달 방향 정보를 계산하기 위한 병진 운동 변환기(423)메타데이터; 및

상기 음원에 대한 깊이 맵에 의해서 제공되는 거리에 따라 달라지는 거리 보상 정보 및 상기 수정된 도달 방향 정보와 관련된 주파수 빈과 연관된 새로운 거리를 이용하여 상기 수정된 음장을 계산하기 위한 거리 보상을 포함하는 것을 특징으로 하는 음장 묘사 및 음장 묘사의 공간적 정보와 관련된 메타데이터로부터 수정된 음장 묘사를 생성하기 위한 장치.

청구항 23

제8항 내지 제22항 중 어느 한 항에 있어서,

하나 이상의 음원을 위한 상기 음장 계산기(420)는,

상기 음장의 분석(1102)에 의해서 얻어진 상기 기준 위치로부터 음원(510)으로 향하는 제1 벡터를 계산하고(1104);

상기 제1 벡터, 및 상기 기준 위치(522)로부터 상기 상이한 기준 위치(500)로의 병진 운동 벡터를 규정하는 상기 병진 운동 정보를 이용하여, 상이한 기준 위치(500)로부터 음원(510)으로 향하는 제2 벡터를 계산하고(1106); 그리고

상기 상이한 기준 위치(500), 상기 음원(510)의 위치, 및 상기 제2 벡터를 이용하여, 또는 상기 상이한 기준 위치(500)로부터 상기 음원(510)의 위치 및 상기 제2 벡터를 이용하여, 거리 수정 값을 계산하도록(1106) 구성되는 것을 특징으로 하는 음장 묘사 및 음장 묘사의 공간적 정보와 관련된 메타데이터로부터 수정된 음장 묘사를 생성하기 위한 장치.

청구항 24

제8항 내지 제23항 중 어느 한 항에 있어서,
 도달 방향 단위 벡터와 상기 메타데이터 내에 포함된 거리를 곱하는 것에 의해서 제1 벡터가 계산되거나,
 제1 벡터로부터 병진 운동 벡터를 차감함으로써 제2 벡터가 계산되거나,
 제2 벡터를 제1 벡터의 노름(norm)으로 나누는 것에 의해서 상기 거리 수정 값이 계산되는 것을 특징으로 하는
 음장 묘사 및 음장 묘사의 공간적 정보와 관련된 메타데이터로부터 수정된 음장 묘사를 생성하기 위한 장치.

청구항 25

제8항 내지 제24항 중 어느 한 항에 있어서,
 상기 음장 계산기(420)는, 상기 병진 운동 정보에 더하여, 회전 정보를 수신하도록 구성되고, 그리고
 상기 음장 계산기(420)는 상기 회전 정보를 이용하여 음장에 대한 도달 방향 데이터를 회전시키기 위해서 회전
 변환(424)을 실시하도록 구성되고,
 상기 도달 방향 데이터는 상기 음장 묘사의 음장 분석에 의해서 그리고 상기 병진 운동을 이용하여 얻어진 도달
 방향 데이터로부터 유도되는 것을 특징으로 하는 음장 묘사 및 음장 묘사의 공간적 정보와 관련된 메타데이터로
 부터 수정된 음장 묘사를 생성하기 위한 장치.

청구항 26

제8항 내지 제25항 중 어느 한 항에 있어서,
 상기 음장 계산기(420)는,
 음장 분석에 의해서 상기 음장 묘사 및 상기 음원에 대한 방향으로부터 음원을 결정하고(1200, 1240);
 음원에 대해서, 상기 메타데이터를 이용하여 상기 기준 위치로부터 상기 음원의 거리를 결정하고(1220);
 상기 음원에 대한 방향 및 상기 병진 운동 정보를 이용하여 상기 상이한 기준 위치와 관련된 상기 음원의 새로
 운 방향을 결정하고(1250);
 상기 상이한 기준 위치와 관련된 음원에 대한 새로운 거리 정보를 결정하고(1230); 그리고
 음원의 새로운 방향, 새로운 거리 정보, 및 음장 묘사 또는 음장 묘사로부터 유도된 음원에 상응하는 음원 신호
 를 이용하여 상기 수정된 음장을 생성하도록(1260) 구성되는 것을 특징으로 하는 음장 묘사 및 음장 묘사의 공
 간적 정보와 관련된 메타데이터로부터 수정된 음장 묘사를 생성하기 위한 장치.

청구항 27

제8항 내지 제26항 중 어느 한 항에 있어서,
 상기 음장 계산기는,
 사운드 분석에 의해서 상기 음장 묘사 및 상기 기준 위치와 관련된 음원 신호의 방향으로부터 음원 신호를 결정
 하고(1200);
 상기 병진 운동 정보를 이용하여 상기 상이한 기준 위치와 관련된 음원 신호의 새로운 방향을 계산하고(1230);
 상기 상이한 기준 위치와 관련된 음원에 대한 거리 정보를 계산하고(1230); 그리고
 상기 거리 정보, 상기 음원 신호 및 상기 새로운 방향을 이용하여 상기 수정된 음장을 합성하도록(1260) 구성되
 는 것을 특징으로 하는 음장 묘사 및 음장 묘사의 공간적 정보와 관련된 메타데이터로부터 수정된 음장 묘사를
 생성하기 위한 장치.

청구항 28

제27항에 있어서,

상기 음장 계산기(420)는,

리플레이 설정과 관련하여 새로운 방향에 의해서 주어진 방향에 대해서 음원 신호를 패닝하는 것에 의해서, 그리고 패닝 실시 전에 또는 패닝 실시 후에 거리 정보를 이용하여 음원 신호를 스케일링하는 것에 의해서 상기 수정된 음장을 합성하도록 구성되는 것을 특징으로 하는 음장 묘사 및 음장 묘사의 공간적 정보와 관련된 메타데이터로부터 수정된 음장 묘사를 생성하기 위한 장치.

청구항 29

제27항 또는 제28항에 있어서,

상기 음장 계산기(420)는 확산 신호를 상기 음원 신호의 다이렉트 부분에 추가하도록 구성되고, 상기 다이렉트 부분은 상기 확산 신호에 추가되기 전에 상기 거리 정보에 의해서 수정되는 것을 특징으로 하는 음장 묘사 및 음장 묘사의 공간적 정보와 관련된 메타데이터로부터 수정된 음장 묘사를 생성하기 위한 장치.

청구항 30

제27항 내지 제29항 중 어느 한 항에 있어서,

상기 음장 계산기(420)는,

상기 음장 묘사의 시간-주파수 변환을 실시하며 시간 프레임의 복수의 주파수 빈에 대해서 도달 방향을 계산하고(422),

각각의 주파수 빈에 대해서 새로운 방향을 계산하고(423, 424),

각각의 주파수 빈에 대해서 거리 정보를 계산하고, 그리고

각각의 주파수 빈에 대해서, 해당 주파수 빈에 대한 오디오 신호와, 해당 주파수 빈에 대한 새로운 방향으로부터 유도된 해당 주파수 빈에 대한 패닝 이득과, 그리고 상기 상응 음원에 대한 거리 정보로부터 유도된 해당 주파수 빈에 대한 스케일링 인자를 이용하여 다이렉트 합성을 실시하도록 구성되는 것을 특징으로 하는 음장 묘사 및 음장 묘사의 공간적 정보와 관련된 메타데이터로부터 수정된 음장 묘사를 생성하기 위한 장치.

청구항 31

제30항에 있어서,

상기 음장 계산기(420)는,

상기 주파수 빈에 대한 오디오 신호로부터 유도된 확산 오디오 신호를 이용하여 그리고 상기 주파수 빈에 대한 사운드 분석에 의해서 유도된 확산 매개변수를 이용하여 확산 합성을 수행하며 상기 주파수 빈에 대한 합성된 오디오 신호를 획득하기 위해서 상기 다이렉트 부분 및 상기 확산 부분을 조합하고; 그리고

상기 수정된 음장으로서 시간 도메인 합성된 오디오 신호를 획득하기 위해서 시간 프레임에 대해서 상기 주파수 빈에 대한 오디오 신호를 이용하여 주파수-시간 변환을 실시하도록 구성되는 것을 특징으로 하는 음장 묘사 및 음장 묘사의 공간적 정보와 관련된 메타데이터로부터 수정된 음장 묘사를 생성하기 위한 장치.

청구항 32

제27항 내지 제31항 중 어느 한 항에 있어서,

상기 음장 계산기(420)는, 각각의 음원에 대해서, 상기 상이한 기준 위치와 관련된 음장을 합성하도록 구성되고,

상기 합성은,

각각의 음원에 대해서, 상기 상이한 기준 위치와 관련된 음원 신호의 음장 묘사를 획득하기 위해서 상기 음원 신호에 대한 새로운 방향을 이용하여 음원 신호를 프로세싱하는 것(430);

상기 음원 신호를 프로세싱하기 전에 상기 음원 신호를 수정하는 것 또는 상기 방향 정보를 이용하여 상기 음장 묘사를 수정하는 것; 그리고

상기 상이한 기준 위치와 관련된 수정된 음장을 획득하기 위해서 상기 음원에 대한 상기 음장 묘사를 추가하는 것을 포함하는 것을 특징으로 하는 음장 묘사 및 음장 묘사의 공간적 정보와 관련된 메타데이터로부터 수정된 음장 묘사를 생성하기 위한 장치.

청구항 33

제27항 내지 제32항 중 어느 한 항에 있어서,

상기 사운드 분석(1200)은 음원 분리 알고리즘에 의해서 상기 음원 신호를 결정하고 그리고 상기 확산 신호를 획득하기 위해서 상기 음원 신호의 적어도 일부를 상기 음장 묘사로부터 차감하도록 구성되는 것을 특징으로 하는 음장 묘사 및 음장 묘사의 공간적 정보와 관련된 메타데이터로부터 수정된 음장 묘사를 생성하기 위한 장치.

청구항 34

제8항 내지 제33항 중 어느 한 항에 있어서,

상기 음장 계산기(420)는, 각각의 매칭된 음원에 대한 도달 방향 정보를 결정하되,

상기 도달 방향 정보 및 적어도 하나의 음장 묘사를 위한 메타데이터를 이용하여 새로운 기준 위치에 대한 음원의 거리를 결정하고,

상기 새로운 기준 위치에 대한 음원의 거리를 이용하여 스케일링 인자를 결정하도록 구성되는 것을 특징으로 하는 음장 묘사 및 음장 묘사의 공간적 정보와 관련된 메타데이터로부터 수정된 음장 묘사를 생성하기 위한 장치.

청구항 35

제8항 내지 제34항 중 어느 한 항에 있어서,

상기 음장 계산기(420)는,

상기 메타데이터를 이용하여 하나 이상의 추출된 객체의 제1 그룹 또는 하나 이상의 추출된 객체에 대한 도달 방향 정보를 획득하기 위해서 음장 묘사에 대해서 그리고 하나 이상의 추출된 객체의 제2 그룹 또는 하나 이상의 추출된 객체에 대한 도달 방향 정보를 획득하기 위해서 추가적인 기준 위치와 관련된 추가적인 음장 묘사에 대해서 사운드 분석을 수행하고(700, 740, 741);

적어도 하나의 매칭된 객체를 찾기 위해서 상기 그룹으로부터의 객체 신호 또는 상기 그룹으로부터 객체에 대한 도달 방향 정보를 이용하여 상기 제1 그룹 및 상기 제2 그룹을 매칭시키고(704);

상기 메타데이터 및 상기 제1 그룹을 획득하기 위해서 상기 사운드 분석을 실시하는 것에 의해서 상기 매칭된 객체에 대해서 얻어진 제1 도달 방향 정보 및 상기 제2 그룹을 획득하기 위해서 상기 사운드 분석을 실시하는 것에 의해서 상기 매칭된 객체에 대해서 얻어진 제2 도달 방향 정보를 이용하여 상기 매칭된 객체의 추정 위치를 계산하고(706, 746);

상기 추정된 위치 및 상기 병진 운동 정보를 기초로 거리 보상 프로세스를 상기 매칭된 객체에 적용하도록(710,

748) 구성되는 것을 특징으로 하는 음장 묘사 및 음장 묘사의 공간적 정보와 관련된 메타데이터로부터 수정된 음장 묘사를 생성하기 위한 장치.

청구항 36

제35항에 있어서,

상기 사운드 분석은 각각의 시간-주파수 빈에 대한 신호 및 각각의 시간-주파수 빈에 대한 도달 방향 값 및 선택적인 확산 값을 초래하는 DirAC 분석(740, 741)이고,

상기 매칭은 각각의 개별적인 시간-주파수 빈에 대해서 수행되어 적어도 하나의 빈을 상기 매칭된 객체로서 결정하고, 그리고

상기 적어도 하나의 빈에 대한 추정된 위치가 계산되는(746) 것을 특징으로 하는 음장 묘사 및 음장 묘사의 공간적 정보와 관련된 메타데이터로부터 수정된 음장 묘사를 생성하기 위한 장치.

청구항 37

제35항에 있어서,

상기 사운드 분석은 추출된 광대역 신호 및 추출된 광대역 신호에 대한 도달 방향 정보 그리고, 선택적으로, 확산 신호를 초래하는 광대역 음원 분석(700)이고,

상기 음원 매칭(704)은 제1 그룹으로부터의 추출된 광대역 신호 및 제2 그룹으로부터의 추출된 광대역 신호의 각각의 쌍에 대한 상관화 수단(correlation measure)을 이용하여 상기 추출된 광대역 신호에 대해서 수행되고, 그리고

상기 적어도 하나의 매칭된 광대역 객체에 대한 추정된 위치가 계산되는(706) 것을 특징으로 하는 음장 묘사 및 음장 묘사의 공간적 정보와 관련된 메타데이터로부터 수정된 음장 묘사를 생성하기 위한 장치.

청구항 38

제35항 내지 제37항 중 어느 한 항에 있어서,

상기 음장 계산기(420)는 상기 음장 묘사로부터 유도된 제1 도달 방향 정보 및 상기 추가적인 음장 묘사로부터 유도된 제2 도달 방향 정보, 그리고 상기 기준 위치와 상기 추가적인 기준 위치 사이의 벡터를 유도하는데 이용될 수 있는 정보를 포함하는 상기 메타데이터를 이용하여 상기 매칭된 객체의 위치의 계산을 실시하도록 구성되는 것을 특징으로 하는 음장 묘사 및 음장 묘사의 공간적 정보와 관련된 메타데이터로부터 수정된 음장 묘사를 생성하기 위한 장치.

청구항 39

제1항 내지 제38항 중 어느 한 항에 있어서,

상기 음장 계산기(420)는, 상기 상이한 기준 위치와 관련된 매칭된 객체에 대한 음장 묘사의 계산을 위해서, 병진 운동 정보에 의해서 결정된 청취자 위치에 대한 기준 위치 또는 추가적인 기준 위치의 거리에 따라 음장 묘사에 관련된 또는 추가적인 음장 묘사에 관련된 정보를 선택하도록 구성되는 것을 특징으로 하는 음장 묘사 및 음장 묘사의 공간적 정보와 관련된 메타데이터로부터 수정된 음장 묘사를 생성하기 위한 장치.

청구항 40

제39항에 있어서,

상기 음장 계산기는,

상기 기준 위치와 상기 새로운 청취자 위치 사이의 제1 거리를 계산하고(720);

상기 추가적인 기준 위치와 상기 새로운 청취자 위치 사이의 제2 거리를 계산하고(722);

상기 제1 및 제2 거리 중 더 작은 거리를 가지는 기준 위치를 갖는 음장 묘사를 선택하도록(724) 구성되는 것을 특징으로 하는 음장 묘사 및 음장 묘사의 공간적 정보와 관련된 메타데이터로부터 수정된 음장 묘사를 생성하기 위한 장치.

청구항 41

제35항 내지 제40항 중 어느 한 항에 있어서,

상기 음장 계산기는, 각각의 매칭된 객체에 대해서, 상기 음장 묘사 및 상기 추가적인 음장 묘사로부터 유도된 도달 방향 정보 중 하나를 이용하여 새로운 도달 방향 정보를 계산하도록 구성되는 것을 특징으로 하는 음장 묘사 및 음장 묘사의 공간적 정보와 관련된 메타데이터로부터 수정된 음장 묘사를 생성하기 위한 장치.

청구항 42

제40항 또는 제41항에 있어서,

상기 음장 계산기(420)는, 각각의 매칭된 객체에 대해서, 새로운 도달 방향 정보를 이용하여 음장 묘사를 계산하도록, 그리고 상기 매칭된 객체의 추정된 위치를 이용하여 상기 음원 신호 또는 상기 음원 신호의 음장 묘사를 스케일링하도록 구성되는 것을 특징으로 하는 음장 묘사 및 음장 묘사의 공간적 정보와 관련된 메타데이터로부터 수정된 음장 묘사를 생성하기 위한 장치.

청구항 43

제35항 내지 제42항 중 어느 한 항에 있어서,

상기 음장 계산기(420)는, 상응하는 도달 방향 정보를 이용하여, 상기 제1 그룹으로부터 각각의 비-매칭 객체에 대한 음장 묘사를 계산하도록 그리고 제2 그룹으로부터 각각의 비-매칭 객체에 대한 음장 묘사를 계산하도록 구성되는 것을 특징으로 하는 음장 묘사 및 음장 묘사의 공간적 정보와 관련된 메타데이터로부터 수정된 음장 묘사를 생성하기 위한 장치.

청구항 44

제43항에 있어서,

상기 음장 계산기(420)는, 상기 매칭된 객체의 음장 묘사 및 상기 비-매칭 객체의 음장 묘사를 조합하는 것(712)에 의해서 상기 수정된 음장을 계산하도록 구성되는 것을 특징으로 하는 음장 묘사 및 음장 묘사의 공간적 정보와 관련된 메타데이터로부터 수정된 음장 묘사를 생성하기 위한 장치.

청구항 45

제35항 내지 제44항 중 어느 한 항에 있어서,

상기 음장 계산기(420)는 상기 제1 그룹 및 상기 제2 그룹 중 적어도 하나에 대한 확산 신호에 관한 정보를 결정하도록 구성되고, 그리고

상기 음장 계산기는 확산 신호를 매칭된 객체의 음장 묘사 또는 비-매칭 객체의 음장 묘사에 부가하도록 구성되는 것을 특징으로 하는 음장 묘사 및 음장 묘사의 공간적 정보와 관련된 메타데이터로부터 수정된 음장 묘사를

생성하기 위한 장치.

청구항 46

증강된 음장 묘사를 생성하기 위한 방법에 있어서,

적어도 하나의 기준 위치에 대한 음장을 나타내는 적어도 하나의 음장 묘사를 생성하는 단계; 및

상기 음장의 공간 정보와 관련된 메타데이터를 생성하는 단계를 포함하고,

상기 적어도 하나의 음장 묘사와 상기 메타데이터가 증강된 음장 묘사를 구성하는 것을 특징으로 하는 증강된 음장 묘사를 생성하기 위한 방법.

청구항 47

음장 묘사 및 음장 묘사의 공간적 정보와 관련된 메타데이터로부터 수정된 음장 묘사를 생성하기 위한 방법에 있어서,

상기 공간적 정보, 상기 음장 묘사, 및 기준 위치로부터 상이한 기준 위치로의 병진 운동을 나타내는 병진 운동 정보를 이용하여 수정된 음장을 계산하는 단계를 포함하는 것을 특징으로 하는 음장 묘사 및 음장 묘사의 공간적 정보와 관련된 메타데이터로부터 수정된 음장 묘사를 생성하기 위한 방법.

청구항 48

컴퓨터 또는 프로세서 상에서 동작될 때, 제46항의 방법 또는 제47항의 방법을 실시하기 위한 컴퓨터 프로그램.

청구항 49

증강된 음장 묘사로서,

적어도 하나의 기준 위치에 대한 음장을 나타내는 적어도 하나의 음장 묘사 및 상기 음장의 공간 정보와 관련된 메타데이터를 포함하는 것을 특징으로 하는 증강된 음장 묘사.

청구항 50

제49항에 있어서,

상기 적어도 하나의 음장 묘사가 상기 기준 위치와 관련된 제1 음장 묘사 및 추가적인 기준 위치와 관련된 제2 음장 묘사를 포함하고, 상기 추가적인 기준 위치는 상기 기준 위치와 상이하며, 그리고

상기 메타데이터는, 개별적으로 상기 기준 위치 및 상기 추가적인 기준 위치 또는 상기 기준 위치와 상기 추가적인 기준 위치 사이의 거리 또는 상기 기준 위치와 상기 추가적인 기준 위치 사이의 위치 벡터 또는 상기 기준 위치 및 상기 추가적인 기준 위치와 관련된 임의의 정보를 포함하는 것을 특징으로 하는 증강된 음장 묘사.

발명의 설명

기술 분야

[0001] 본 발명은 오디오 프로세싱 그리고 특히, 마이크론 또는 가상 마이크론 위치와 같은 기준 위치에 대해서 규정되는 음장과 관련된 오디오 프로세싱에 관한 것이다.

배경 기술

- [0002] 앰비소닉 신호(ambisonics signals)는 음장(sound field)의 절두형 구면 조화 분해(truncated spherical harmonic decomposition)를 포함한다. 앰비소닉은 다양한 이유로 이용된다. 오늘날 '1차 앰비소닉'(FOA)으로 알려져 있고 제4 신호(즉, 하나의 전방향 신호 및 3개 이하의 8자 모양 방향 신호를 포함하는 '전통적인' 앰비소닉[31])이 있다. 보다 최근의 앰비소닉 변형에는 '고차 앰비소닉'(HOA)으로 알려져 있고, 더 많은 신호의 이송을 사용하여, 증강된 공간적 해상도 및 더 넓은 청취자 스위트-스팟 지역(sweet-spot area)을 제공한다. 일반적으로, 완전히 정의된 N차 HOA 표시는 $(N+1)^2$ 신호로 구성된다.
- [0003] 앰비소닉 아이디어와 관련하여, 방향성 오디오 코딩(DirAC) 표시는, FOA 또는 HOA 음향 장면(sound scene)을 더 밀집된 파라메트릭 스타일(parametric style)로 나타내는 것으로 인식되었다. 더 구체적으로, 공간적 음향 장면은, 각각의 시간-주파수(TF) 빈(time-frequency bin) 내의 음향 장면의 다운믹스 및 이와 연관된 방향 및 확산에 관한 부수적 정보를 나타내는 하나의(그 초과) 전송된 오디오 채널에 의해서 표시된다. DirAC에 관한 추가적인 정보를 [32] 및 [33]에서 확인할 수 있다.
- [0004] DirAC[32]는 상이한 마이크로폰 시스템 및 임의의 라우드스피커 설정과 함께 이용될 수 있다. DirAC 시스템의 목적은, 다중채널/3D 라우드스피커 시스템을 이용하여 가능한 한 정밀하게 기존 음향 환경의 공간적 인상을 재생하는 것이다. 선택된 환경에서, 응답(연속적 사운드 또는 임펄스 응답)은, 전방향 마이크로폰(W)으로 그리고 사운드의 도달-방향 및 사운드의 확산을 측정할 수 있는 마이크로폰의 세트로 측정된다. 일반적인 방법이, 데카르트 좌표 축과 정렬된 3개의 8자 모양의 마이크로폰(X, Y, Z)에 적용된다[34]. 이를 위한 방법은, 희망 응답 전부를 직접적으로 산출하는 음장 마이크로폰을 이용하는 것이다. W, X, Y, 및 Z 신호는 또한 구분된 전방향 마이크로폰의 세트로부터 연산될 수 있다.
- [0005] DirAC에서, 사운드 신호는 먼저 주파수 채널들로 분할된다. 사운드 방향 및 확산은 각각의 주파수 채널에서 시간에 따라 측정된다. 전송에서, 분석된 방향 및 확산 데이터와 함께, 하나 이상의 오디오 채널이 송신된다. 합성에서, 라우드스피커에 인가되는 오디오가 예를 들어 전방향 채널(W)일 수 있거나, 각각의 라우드스피커를 위한 사운드가 W, X, Y, 및 Z의 가중 합계로서 연산될 수 있고, 이는, 각각의 라우드스피커를 위한 특정 방향 특성을 갖는 신호를 형성한다. 각각의 오디오 채널이 주파수 채널로 분할되고, 이어서 주파수 채널은, 분석되는 확산에 따라, 확산 스트림 및 비-확산 스트림으로 선택적으로 분할된다. 확산 스트림은, 음향 장면의 확산 인식을 생성하는 기술, 예를 들어 Binaural Cue Coding[35-37]에서 이용되는 비상관화(decorrelation) 기술로 재생된다. 비확산 사운드는, 방향 데이터에 따라 점형 가상 음원을 생성하는 것을 목적으로 하는 기술(예를 들어, VBAP[38])로 재생된다.
- [0006] 제한된 자유도의 6DoF에서의 네비게이션을 위한 3개의 기술이 [39]에서 제시된다. 단일 앰비소닉 신호가 주어지면, 단일 앰비소닉 신호는 1) HOA 플레이백 및 가상 라우드스피커 어레이 내의 청취자 이동의 시뮬레이션, 2) 연산 및 평면-파동을 따른 해석, 및 3) 청취자를 중심으로 하는 음장의 재확장을 이용하여 연산된다.
- [0007] 또한, 예를 들어 간행물["Directional Audio Coding - Perception-Based Reproduction of Spatial Sound", V. Pulkki et al, International Workshop on the Principles and Applications of Spatial Hearing, November 11-13, 2009, Zao, Miyagi, Japan]에서 설명된 바와 같은 DirAC 기술을 참조한다. 이 참조 문헌은, 특히 공간적 오디오 프로세싱을 위한 인식적 동기부여 기술과 같은 기준 위치 관련 음장 프로세싱을 위한 예로서 방향성 오디오 코딩을 설명한다. 이는, 공간 신호의 캡처, 코딩 및 재-합성에서, 원격 회의에서, 방향성 필터링에서, 그리고 가상 청각 환경에서 적용된다.
- [0008] 음향 장면의 재생이 종종 라우드스피커 설정에 초점 맞춰졌는데, 이는, 그러한 재생이 전형적으로 사생활, 예를 들어 거실에서의, 그리고 전문적인 상황, 즉 극장에서의 재생이었기 때문이다. 여기에서, 재생 기하형태에 대한 장면의 관계는 정적인데, 이는, 그러한 것이, 청취자가 전방 방향을 바라보도록 강제하는 2차원적인 화상을 수반하기 때문이다. 이어서, 재생 시간(production time)에 사운드와 시각적 객체의 공간적 관계가 규정되어 확정된다.
- [0009] 가상 현실(VR)에서, 사용자가 장면 내에서 자유롭게 이동할 수 있게 하는 것에 의해서 몰입감이 분명하게 얻어진다. 그에 따라, 사용자의 이동을 추적하고 사용자의 위치에 맞춰 시각적 및 청각적 재생을 조정할 필요가 있다. 전형적으로, 사용자는 머리-장착형 디스플레이(HMD) 및 헤드폰을 착용한다. 헤드폰을 이용한 몰입 체험을 위해서, 오디오는 바이노럴화되어야 한다. 바이노럴화(binauralization)는 인간의 머리, 귀, 및 상부 몸통이 그 방향 및 거리에 따라 어떻게 음원의 사운드를 변화시키는지에 관한 시뮬레이션이다. 이는, 그 상대적인 방향에 대해서 헤드-관련 전달 함수(HRTF)로 신호를 콘볼루션(convolution)하는 것에 의해서 달성된다[1, 2]. 바이노럴

화는 또한, 사운드가 헤드의 내측으로부터가 아니라 장면으로부터 오는 것으로 생각되게 한다[3]. 이미 성공적으로 해결된 일반적인 시나리오는 360° 비디오 재생이다[4, 5]. 여기에서, 사용자는 HMD를 착용하거나 태블릿 또는 전화기를 그의 손에 들고 있다. 그의 머리 또는 장치를 이동시키는 것에 의해서, 사용자는 임의의 방향으로 둘러볼 수 있다. 이는 3 자유도(3DoF)의 시나리오인데, 이는 사용자가 3 이동도(피치, 요, 롤)를 가지기 때문이다. 시각적으로, 이는, 사용자 주위의 구체에 비디오를 투영하는 것에 의해서 실현된다. 오디오는 종종 비디오 카메라에 근접한, 공간적 마이크로폰, 예를 들어 1차 앰비소닉(FOA)으로 레코딩된다[6]. 앰비소닉 도메인 내에서, 사용자의 머리 회전은 직접적인 방식으로 적용된다[7]. 이어서, 오디오는 예를 들어 사용자 주위에 배치된 가상의 라우드스피커로 렌더링된다. 이러한 가상의 라우드스피커 신호가 이어서 바이노럴화(binauralized)된다.

[0010] 현대의 VR 적용에는 6 자유도(6DoF)를 허용한다. 머리 회전에 더하여, 사용자가 주변에서 이동하여, 3개의 공간적 차원에서 그의 위치의 병진 운동을 초래할 수 있다. 6DoF 재생은 걷는 지역의 전체 크기에 의해서 제한된다. 많은 경우에, 이러한 지역은 다소 작고, 예를 들어 통상적인 거실이다. 6DoF는 VR 게임에서 일반적으로 만나게 된다. 여기에서, 전체 장면은 컴퓨터-생성 이미지(CGI)와 합성된다. 오디오는 종종 객체-기반(object-based)의 렌더링을 이용하여 생성되고, 여기에서 각각의 오디오 객체는 트래킹 데이터를 기초로 하는 거리-의존적 이득(distance-dependent gain) 및 사용자로부터의 상대적인 방향으로 렌더링된다. 인공적인 반향 및 회절에 의해서 현실감이 향상될 수 있다[8, 9, 10].

[0011] 레코딩된 콘텐츠와 관련하여, 오디오-시각적 6DoF 재생을 확신하게 하는데 있어서 일부 분명한 해결과제가 있다. 공간적 병진 운동 도메인 내의 공간적 사운드 조작의 초기의 예는, '어쿠스틱 줌(acoustic zoom)' 기술의 공간적 사운드 조작이다[11, 12]. 여기에서, 청취자 위치는, 화상 내로의 주밍(zooming)하는 것과 유사하게, 레코딩된 시각적 장면 내로 가상으로 이동된다. 사용자는 하나의 방향 또는 화상 부분을 선택하고, 이어서 병진 운동된 지점으로부터 이를 청취할 수 있다. 이는, 모든 도달 방향(DoAs)이 원래의, 비-주밍(non-zoomed) 재생에 대해서 변화되는 것을 수반한다.

[0012] 공간적으로 분포된 레코딩 위치를 이용하는 레코딩된 콘텐츠의 6DoF 재생을 위한 방법이 제시되어 있다. 비디오에서, 명시적 렌더링을 생성하기 위해서 카메라의 어레이가 이용된다[13]. 오디오에서, 유사한 설정이, 분포된 마이크로폰 어레이 또는 앰비소닉 마이크로폰을 이용한다. 그러한 레코딩으로부터 임의의 위치에 배치된 '가상 마이크로폰'의 신호를 생성할 수 있다는 것을 보여주었다[14].

[0013] 그러한 공간적 사운드 수정을 기술적으로 편리한 방식으로 실현하기 위해서, 파라메트릭 사운드 프로세싱 또는 코딩 기술이 이용될 수 있다(개관을 위해 [15] 참조). 방향성 오디오 코딩(DirAC)[16]은 레코딩을 표상으로 변환하기 위한 대중적인 방법이며, 그러한 표상은 오디오 스펙트럼과, 사운드 방향 및 확산에 관한 파라메트릭 부수적 정보로 구성된다. 이는, 어쿠스틱 줌[11] 및 가상 마이크로폰[14] 적용예를 위해서 이용된다.

선행기술문헌

비특허문헌

[0014] (비특허문헌 0001) Liitola, T., Headphone sound externalization, Ph.D. thesis, Helsinki University of Technology. Department of Electrical and Communications Engineering Laboratory of Acoustics and Audio Signal Processing., 2006.

(비특허문헌 0002) Blauert, J., Spatial Hearing - Revised Edition: The Psychophysics of Human Sound Localization, The MIT Press, 1996, ISBN 0262024136.

(비특허문헌 0003) Zhang, W., Samarasinghe, P. N., Chen, H., and Abhayapala, T. D., "Surround by Sound: A Re-view of Spatial Audio Recording and Reproduction," Applied Sciences, 7(5), p. 532, 2017.

(비특허문헌 0004) Bates, E. and Boland, F., "Spatial Music, Virtual Reality, and 360 Media," in Audio Eng. Soc. Int. Conf. on Virtual and Augmented Reality, Los Angeles, CA, U.S.A., 2016.

(비특허문헌 0005) Anderson, R., Gallup, D., Barron, J. T., Kontkanen, J., Snavely, N., Esteban, C. H., Agarwal, S., and Seitz, S. M., "Jump: Virtual Reality Video," ACM Transactions on Graphics, 35(6), p. 198, 2016.

(비특허문헌 0006) Merimaa, J., Analysis, Synthesis, and Perception of Spatial Sound: Binaural

Localization Modeling and Multichannel Loudspeaker Reproduction, Ph.D. thesis, Helsinki University of Technology, 2006.

(비특허문헌 0007) Kronlachner, M. and Zotter, F., "Spatial Transformations for the Enhancement of Ambisonics Recordings," in 2nd International Conference on Spatial Audio, Erlangen, Germany, 2014.

(비특허문헌 0008) Tsingos, N., Gallo, E., and Drettakis, G., "Perceptual Audio Rendering of Complex Virtual Environments," ACM Transactions on Graphics, 23(3), pp. 249-258, 2004.

(비특허문헌 0009) Taylor, M., Chandak, A., Mo, Q., Lauterbach, C., Schissler, C., and Manocha, D., "Guided multi-view ray tracing for fast auralization," IEEE Trans. Visualization & Comp. Graphics, 18, pp. 1797- 1810, 2012.

(비특허문헌 0010) [10] Rungta, A., Schissler, C., Rewkowski, N., Mehra, R., and Manocha, D., "Diffraction Kernels for Interactive Sound Propagation in Dynamic Environments," IEEE Trans. Visualization & Comp. Graphics, 24(4), pp. 1613-1622, 2018.

(비특허문헌 0011) Thiergart. O., Kowalczyk, K., and Habets, E. A. P., "An Acoustical Zoom based on Informed Spatial Filtering," in Int. Workshop on Acoustic Signal Enhancement, pp. 109- 113, 2014.

(비특허문헌 0012) Khaddour, H., Schimmel, J., and Rund, F., "A Novel Combined System of Direction Estimation and Sound Zooming of Multiple Speakers," Radioengineering, 24(2), 2015.

(비특허문헌 0013) Ziegler, M., Keinert, J., Holzer, N., Wolf, T., Jaschke, T., op het Veld, R., Zakeri, F. S., and Foessel, S., "Immersive Virtual Reality for Live-Action Video using Camera Arrays," in IBC, Amsterdam, Netherlands, 2017.

(비특허문헌 0014) Thiergart. O., Galdo, G. D., Taseska, M., and Habets, E. A. P., "Geometry-Based Spatial Sound Acquisition using Distributed Microphone Arrays," IEEE Trans. Audio, Speech, Language Process., 21(12), pp. 2583-2594, 2013.

(비특허문헌 0015) Kowalczyk, K., Thiergart. O., Taseska, M., Del Galdo, G., Pulkki, V., and Habets, E. A. P., "Parametric Spatial Sound Processing: A Flexible and Efficient Solution to Sound Scene Acquisition, Modification, and Reproduction," IEEE Signal Process. Mag., 32(2), pp. 31-42, 2015.

(비특허문헌 0016) Pulkki, V., "Spatial Sound Reproduction with Directional Audio Coding," J. Audio Eng. Soc., 55(6), pp. 503-516, 2007.

(비특허문헌 0017) International Telecommunication Union, "ITU-R BS. 1534-3, Method for the subjective assessment of intermediate quality level of audio systems," 2015.

(비특허문헌 0018) Thiergart, O., Del Galdo, G., Kuech, F., and Prus, M., "Three-Dimensional Sound Field Analysis with Directional Audio Coding Based on Signal Adaptive Parameter Estimators," in Audio Eng. Soc. Conv. Spatial Audio: Sense the Sound of Space, 2010.

(비특허문헌 0019) Kuttruff, H., Room Acoustics, Taylor & Francis, 4 edition, 2000.

(비특허문헌 0020) BorB, C., "A polygon-based panning method for 3D loudspeaker setups," in Audio Eng. Soc. Conv., pp. 343-352, Los Angeles, CA, USA, 2014.

(비특허문헌 0021) Rummukainen. O., Schlecht, S., Plinge, A., and Habets, E. A. P., "Evaluating Binaural Reproduction Systems from Behavioral Patterns in a Virtual Reality - A Case Study with Impaired Binaural Cues and Tracking Latency," in Audio Eng. Soc. Conv. 143, New York, NY, USA, 2017.

(비특허문헌 0022) Engelke, U., Darcy, D. P., Mulliken, G. H., Bosse, S., Martini, M. G., Arndt, S., Antons, J.-N., Chan, K. Y., Ramzan, N., and Brunnstrom, K., "Psychophysiology-Based QoE Assessment: A Survey," IEEE Selected Topics in Signal Processing, 11(1), pp. 6-21, 2017.

(비특허문헌 0023) Schlecht, S. J. and Habets, E. A. P., "Sign-Agnostic Matrix Design for Spatial Artificial Reverberation with Feedback Delay Networks," in Proc. Audio Eng. Soc. Conf., pp. 1-10-

accepted, Tokyo, Japan, 2018.

(비특허문헌 0024) M. A. Gerzon, "Periphony: With-height sound reproduction," J. Acoust. Soc. Am., vol. 21,110, 1, pp. 2-10, 1973.

(비특허문헌 0025) V. Pulkki, "Directional audio coding in spatial sound reproduction and stereo upmixing," in Proc, of the 28th AES International Conference, 2006.

(비특허문헌 0026) -, "Spatial sound reproduction with directional audio coding," Journal Audio Eng. Soc., vol. 55, no. 6, pp. 503-516, Jun. 2007.

(비특허문헌 0027) C. G. and G. M., "Coincident microphone simulation covering three dimensional space and yielding various directional outputs," U.S. Patent 4042779, 1977.

(비특허문헌 0028) C. Faller and F. Baumgarte, "Binaural cue coding - part ii: Schemes and applications," IEEE Trans. Speech Audio Process., vol. 11, no. 6, Nov. 2003.

(비특허문헌 0029) C. Faller, "Parametric multichannel audio coding: Synthesis of coherence cues," IEEE Trans. Speech Audio Process., vol. 14, no. 1, Jan. 2006.

(비특허문헌 0030) H. P. J. E. E. Schuijers, J. Breebaart, "Low complexity parametric stereo coding," in Proc, of the 116th A ES Convention, Berlin, Germany, 2004.

(비특허문헌 0031) V. Pulkki, "Virtual sound source positioning using vector base amplitude panning," J. Acoust. Soc. A m., vol. 45, no. 6, pp. 456-466, Jun. 1997.

(비특허문헌 0032) J. G. Tylka and E. Y. Choueiri, "Comparison of techniques for binaural navigation of higher- order ambisonics sound fields," in Proc, of the AES International Conference on Audio for Virtual and Augmented Reality, New York, Sep. 2016.

발명의 내용

해결하려는 과제

- [0015] 본원에서 제시된 방법은, 단일 FOA 마이크로폰의 레코딩으로부터의 6DoF 재생을 가능하게 한다. 단일 공간적 위치로부터의 레코딩은 3DoF 재생 또는 어쿠스틱 줌을 위해서 이용되었다. 그러나, 본 발명자의 지식에 따르면, 그러한 데이터로부터의, 인터랙티브, 완전 6DoF 재생을 위한 방법은 아직까지 제시되지 않았다. 레코딩 내의 음원의 거리에 관한 정보를 통합함으로써 6DoF 재생을 실현한다. 이러한 거리 정보는 DirAC의 파라메트릭 표상 내로 통합되고, 그에 따라 청취자의 변화된 원근(perspective)이 정확하게 맵핑된다.
- [0016] 앰비소닉 음장 표상(일반적인 FOA 또는 HOA 앰비소닉과 같거나 또는 DirAC-스타일 파라메트릭 음장 표상과 같은 것임)은, 6DoF 적용예에서 요구되는 바와 같은 청취자 위치의 병진 운동적 이동을 가능하게 하는 충분한 정보를 제공하지 못하는데, 이는, 음향 장면 내의 객체 거리 및 절대 객체 위치 모두가 이러한 포맷으로 결정되지 않기 때문이다. 청취자 위치의 이동이 반대 방향을 따른 음향 장면의 동등한 이동으로 변환될 수 있다는 것을 주목하여야 한다.
- [0017] 6DoF에서 이동될 때의 전형적인 문제가 도 1b에 도시되어 있다. 음향 장면이 앰비소닉을 이용하여 지점 A에서 묘사되는 것으로 가정한다. 이러한 경우에, 음원 A 및 음원 B로부터의 사운드가 동일 방향으로부터 도달하고, 즉 이들은 동일한 도달-방향(direction-of-arrival, DOA)을 갖는다. 위치 B로 이동하는 경우에, 음원 A 및 음원 B의 DOA가 상이하다. 음장의 표준 앰비소닉 묘사, 즉 부가적인 정보가 없는 묘사를 이용할 때, 주어진 위치 A에서의 앰비소닉 신호로, 위치 B에서의 앰비소닉 신호를 연산할 수 없다.
- [0018] 본 발명의 목적은 한편으로는 증강된 음장 묘사를 그리고 다른 한편으로는 수정된 음장 묘사의 생성을 제공하는 것이고, 이는 개선된 또는 유연하거나 효과적인 프로세싱을 가능하게 한다.

과제의 해결 수단

- [0019] 이러한 목적은 제1항의 증강된 음장 묘사 생성 장치, 제8항의 수정된 음장 묘사 생성 장치, 제46항의 증강된 음

장 묘사 생성 방법, 제47항의 수정된 음장 묘사 생성 방법, 제48항의 컴퓨터 프로그램 또는 제49항의 증강된 음장 묘사에 의해서 달성된다.

[0020] 본 발명은, 해당 음장 묘사가 프로세스될 수 있도록, 그에 따라 원래의 기준 위치와 관련되지 않고 다른 기준 위치와 관련된 수정된 음장 묘사가 계산될 수 있도록, 기준 위치와 관련되는 전형적인 음장 묘사가 부가적인 정보를 필요로 한다는 발견을 기초로 한다. 이를 위해서, 이러한 음장의 공간적 정보와 관련된 메타데이터가 생성되고, 그러한 메타데이터는, 음장 묘사와 함께, 예를 들어 전달 또는 저장될 수 있는 증강된 음장 묘사에 상응한다. 음장 묘사 및 메타데이터 그리고, 구체적으로, 음장 묘사의 공간적 정보와 관련된 메타데이터로부터 수정된 음장 묘사를 생성하기 위해서, 수정된 음장이 이러한 공간적 정보, 음장 묘사, 및 기준 위치로부터 다른 기준 위치까지의 병진 운동을 나타내는 병진 운동 정보를 이용하여 계산된다. 따라서, 음장 묘사 및 음장 묘사의 기초가 되는 이러한 음장의 공간적 정보와 관련되는 메타데이터로 이루어진 증강된 음장 묘사를 프로세스하여, 예를 들어 디코더-측에 제공되거나 디코더-측에서 이용될 수 있는 부가적인 병진 운동 정보에 의해서 규정되는 상이한 기준 위치와 관련되는 수정된 음장 묘사를 획득한다.

[0021] 그러나, 본 발명은 인코더/디코더 시나리오와 관련된 뿐만 아니라, 증강된 음장 묘사의 생성 및 수정된 음장 묘사의 생성 모두가 기본적으로 하나의 그리고 동일한 위치에서 실시되는 적용예에서도 적용될 수 있다. 수정된 음장 묘사는, 예를 들어, 수정된 음장 자체 또는 실제로 채널 신호, 바이노럴 신호 내의 수정된 음장에 관한 묘사일 수 있거나, 다시 한번, 기준 위치-관련 신호 필드이나 이제는 원래의 기준 위치가 아니라 새로운 또는 다른 기준 위치와 관련된 기준 위치-관련 신호일 수 있다. 그러한 적용예는, 예를 들어, 음장 묘사가 메타데이터와 함께 존재하고 청취자가 음장이 주어지는 기준 위치로부터 이동하고 상이한 기준 위치로 이동하며, 이어서, 가상 지역 내의 주위에서 이동하는 청취자를 위한 음장이, 음장에 상응하게 그러나 이제 사용자가 이동한 상이한 기준 위치에서, 계산되는 가상 현실 시나리오일 수 있다.

[0022] 특정 실시예에서, 증강된 음장 묘사는 (제1) 기준 위치와 관련된 제1 음장 묘사, 및 (제1) 기준 위치와 상이한 추가적인 (제2) 기준 위치와 관련된 제2 음장 묘사를 가지고, 메타데이터는, 미리 결정된 원점으로부터 이러한 기준 위치들로 향하는 벡터와 같은, 기준 위치 및 추가적인 기준 위치에 관한 정보를 갖는다. 대안적으로, 메타데이터는 기준 위치 또는 추가적인 기준 위치를 향하는 단일 벡터, 및 2개의 상이한 음장 묘사들이 관련되는 2개의 기준 위치들 사이에서 연장되는 벡터일 수 있다.

[0023] 음장 묘사는, 1차 앰비소닉 또는 더 높은 차원의 앰비소닉 묘사와 같은, 비-파라메트릭 음장 묘사일 수 있다. 대안적으로 또는 부가적으로, 음장 묘사가 DirAC 묘사 또는 다른 파라메트릭 음장 묘사일 수 있거나, 하나의 음장 묘사가, 예를 들어, 파라메트릭 음장 묘사일 수 있고 다른 음장 묘사가, 예를 들어, 비-파라메트릭 음장 묘사일 수 있다.

[0024] 여기에서, 음장 묘사는, 각각의 음장 묘사를 위해서, 상이한 시간-주파수 빈들에 대한 하나 이상의 다운믹스 신호와 개별적인 방향 데이터와 그리고 선택적으로 확산 데이터를 가지는 음장의 DirAC 묘사를 생성할 수 있다. 이러한 맥락에서, 메타데이터 생성기는, 기준 위치 및 부가적인 기준 위치가 메타데이터로부터 식별될 수 있도록, 양 음장 묘사를 위한 기하형태적 메타데이터를 생성하도록 구성된다. 이어서, 개별적인 음원을 양 음장 묘사로부터 추출할 수 있을 것이고, 증강된 또는 수정된 음장 묘사를 생성하기 위한 부가적인 프로세싱을 실시할 수 있을 것이다.

[0025] 앰비소닉은 가상, 증강, 및 혼합 현실 적용예의 맥락에서 3D 오디오를 위해서 가장 일반적으로 사용되는 포맷 중 하나가 되었다. 앰비소닉 포맷의 출력 신호를 생성하는, 매우 다양한 오디오 획득 및 생산 도구가 개발되었다. 인터랙티브 가상 현실에서 앰비소닉 코딩된 콘텐츠를 제공하기 위해서, 앰비소닉 포맷이 재생을 위한 바이노럴 신호 또는 채널로 변환된다. 전술한 적용예에서, 청취자는 일반적으로, 3 자유도(3DoF, 즉 피치, 요, 및 롤) 그리고 여전히 적절한 사운드 품질의 경험할 수 있게 하면서 그들이 머리를 음향 장면 내에서 회전할 수 있는 범위까지, 그들의 배향을 제공된 장면 내에서 인터랙티브 방식으로 변화시킬 수 있다. 이는 머리 배향에 따른 렌더링 음향 장면을 회전시키는 것에 의해서 구현되고, 이는 덜 복잡한 연산으로 구현될 수 있고 앰비소닉 표상의 장점이다. 그러나, VR과 같은 새로운 적용예에서, 사용자가, 단순한 배향의 변화보다, 음향 장면 내에서 자유롭게 이동할 수 있게 하는 것이 요구된다(소위, '6 자유도' 또는 6DoF). 결과적으로, 음향 장면의 원근을 변화시키기 위한(즉, x-, y-, 또는 z-축을 따라 음향 장면 내에서 가상으로 이동시키기 위한) 신호 프로세싱이 요구된다. 그러나, 앰비소닉의 주요 단점은, 절차가 음향 장면 내의 단일 원근으로부터 음장을 묘사한다는 것이다. 구체적으로, 이는, 6DoF를 위해서 요구되는 바와 같이 음향 장면으로의 이동('병진 운동')을 허용할 수 있는 음향 장면 내의 음원의 실제 위치에 관한 정보를 포함하지 않는다. 본 발명의 설명은, 이러한 문제를 극복하

고 또한 병진 운동을 돕고 그에 따라 진정한 6DoF를 가능하게 하기 위한 앰비소닉에 대한 몇몇 확장을 제공한다.

- [0026] 1차 앰비소닉(FOA) 레코딩이 프로세스될 수 있고 헤드폰에서 재생될 수 있다. 이들은 청취자의 머리 배향을 고려하여 회전될 수 있다. 그러나, 가상 현실(VR) 시스템은, 청취자가 6 자유도(6DoF) 즉, 3개의 회전 자유도 더하기 3개의 병진 운동 자유도로 이동할 수 있게 한다. 여기에서, 음원들의 체감 각도(apparent angle) 및 거리는 청취자의 위치에 따라 달라진다. 6DoF를 돕기 위한 기술이 설명된다. 특히, 청취자의 위치와 음원까지의 거리에 관한 정보를 기초로 수정되는, 파라메트릭 모델을 이용하는 FOA 레코딩이 설명된다. 그러한 방법은, 청취자가 자유롭게 이동할 수 있는 합성 음향 장면의 상이한 바이노럴 렌더링을 비교하는, 청취 테스트에 의해서 평가된다.
- [0027] 추가적인 바람직한 실시예에서, 증강된 음향 묘사가 전달 또는 저장을 위한 출력 신호를 생성하기 위한 출력 인터페이스에 의해서 출력되고, 출력 신호는, 시간 프레임을 위해서, 음장으로부터 얻어진 하나 이상의 오디오 신호 및 시간 프레임을 위한 공간적 정보를 포함한다. 특히, 추가적인 실시예에서, 음향 생성기는 음장으로부터 방향 데이터를 유도하도록 구성되고, 방향 데이터는 주파수 빈 또는 기간 동안의 사운드의 도달 방향을 나타내고, 메타데이터 생성기는 거리 정보를 방향 데이터에 연관시키는 데이터 항목으로서 공간적 정보를 유도하도록 구성된다.
- [0028] 특히, 그러한 실시예에서, 시간 프레임을 위한 데이터 항목이 상이한 주파수 빈들에 대한 방향 데이터에 연계되도록, 출력 인터페이스가 출력 신호를 생성하도록 구성된다.
- [0029] 추가적인 실시예에서, 음향 생성기는 또한 음장의 시간 프레임의 복수의 주파수 빈에 대한 확산 정보를 생성하도록 구성되고, 메타데이터 생성기는 미리 결정된 값과 상이한 또는 무한대가 아닌 주파수 빈을 위한 거리 정보만을 생성하도록 또는, 확산 값이 미리 결정된 또는 적응형의 문턱값보다 낮을 때, 주파수 빈에 대한 거리 값을 조금이라도 생성하도록 구성된다. 따라서, 큰 확산을 가지는 시간/주파수 빈에서, 임의의 거리 값이 전혀 생성되지 않거나, 디코더에 의해서 특정 방식으로 해석되는 미리 결정된 거리 값이 생성된다. 따라서, 큰 확산을 가지는 시간/주파수 빈에서, 임의의 거리-관련 렌더링이 실시되지 않는 것이 확실한데, 이는, 큰 확산이, 그러한 시간/주파수 빈에서, 사운드가 특정 국소적 음원으로부터 오지 않고 임의의 방향으로부터 온다는 것, 그리고 그에 따라, 음장이 원래의 기준 위치에서 또는 상이하거나 새로운 기준 위치에서 인지되는지의 여부와 관계없이 동일하다는 것을 나타내기 때문이다.
- [0030] 음향 계산기와 관련하여, 바람직한 실시예는 의도된 청취자의 회전을 나타내는 병진 운동 정보 또는 회전 정보를 수정된 음장에 제공하기 위한 병진 운동 인터페이스, 메타데이터를 음향 계산기에 공급하기 위한 메타데이터 공급기 및 음향 묘사를 음향 계산기에 공급하기 위한 음향 공급기 그리고, 부가적으로, 수정된 음향 묘사 및 수정된 메타데이터를 포함하는 수정된 음장을 출력하기 위한 출력 인터페이스를 포함하고, 수정된 메타데이터는 병진 운동 정보를 이용하여 메타데이터로부터 유도되거나, 출력 인터페이스가 복수의 라우드스피커 채널을 출력하며, 각각의 라우드스피커 채널은 미리 규정된 라우드스피커 위치에 관련되거나, 출력 인터페이스가 수정된 음장의 바이노럴 표상을 출력한다.
- [0031] 실시예에서, 음향 묘사는 복수의 음향 성분을 포함한다. 복수의 음향 성분은 전방향 성분 및 적어도 하나의 방향 성분을 포함한다. 이러한 음향 묘사는, 예를 들어, 전방향 성분 및 3개의 방향 성분(X, Y, Z)을 가지는 1차 앰비소닉 음향 묘사이거나, 그러한 음장이, 전방향 성분, X, Y, 및 Z와 관련된 3개의 방향 성분, 그리고, 부가적으로, X, Y, Z 방향 이외의 다른 방향과 관련된 추가적인 방향 성분을 포함하는 더 높은 차수의 앰비소닉 묘사이다.
- [0032] 실시예에서, 장치는, 상이한 시간 또는 주파수 빈에 대해서, 도달 정보의 방향을 유도하기 위해서 음향 성분을 분석하기 위한 분석기를 포함한다. 장치는 부가적으로, DoA 정보 및 메타데이터를 이용하여 주파수 또는 시간 빈 마다의 수정된 DoA 정보를 계산하기 위한 병진 운동 변환기를 가지며, 메타데이터는, 예를 들어 2개의 상이한 기준 위치 및 거리/배치 또는 기준 위치와 관련된 2개의 각도를 이용한 삼각측정 프로세싱에 의해서 얻어지는 바와 같은 양 음향 묘사에 포함되는 음원에 거리를 연관시키는 깊이 맵과 관련된다. 이는, 전체 대역의 표상 또는 시간 프레임의 상이한 주파수 빈들에 적용될 수 있다.
- [0033] 또한, 음향 계산기는, 각각의 또는 일부의 시간/주파수 빈에서 상이한 음원의 각각의 주파수 또는 시간 빈에 대해서 동일한 메타데이터를 이용하여 계산된 거리에 따라, 그리고 시간 또는 주파수 빈과 연관된 새로운 거리에 따라 달라지는 거리 보상 정보를 이용하여, 수정된 음장을 계산하기 위한 거리 보상을 포함하고, 새로운 거리

는 수정된 DoA 정보와 관련된다.

- [0034] 실시예에서, 음장 계산기는 음장의 분석에 의해서 얻어진 기준 위치로부터 음원으로 향하는 제1 벡터를 계산한다. 또한, 음장 계산기는 상이한 기준 위치로부터 음원으로 향하는 제2 벡터를 계산하고, 이러한 계산은 제1 벡터 및 병진 운동 정보를 이용하여 이루어지고, 병진 운동 정보는 기준 위치로부터 상이한 기준 위치로의 병진 운동 벡터를 규정한다. 그리고, 이어서, 상이한 기준 위치로부터 음원까지의 거리는 제2 벡터를 이용하여 계산된다.
- [0035] 또한, 음장 계산기는, 병진 운동 정보에 더하여, 피치, 요 및 롤에 의해서 주어지는 3개의 회전 방향 중 하나를 따른 청취자의 머리의 회전을 나타내는 회전 정보를 수신하도록 구성된다. 이어서, 음장 계산기는 회전 정보를 이용하여 음장에 대한 수정된 도달 방향 데이터를 회전시키기 위해서 회전 변환을 실시하도록 구성되고, 수정된 도달 방향 데이터는 음장 묘사의 사운드 분석에 의해서 얻어진 도달 방향 데이터 및 병진 운동 정보로부터 유도된다.
- [0036] 실시예에서, 음장 계산기는, 음장 묘사 및 사운드 분석에 의한 기준 위치와 관련된 음원 신호의 방향으로부터 음원 신호를 결정하도록 구성된다.
- [0037] 이어서, 상이한 기준 위치와 관련된 음원의 새로운 방향이 계산되고, 이는 메타데이터를 이용하여 이루어지고, 이어서 상이한 기준 위치와 관련된 음원의 거리 정보가 계산되고, 이어서, 거리 정보 및 음원의 새로운 방향을 이용하여, 수정된 음장이 합성된다.
- [0038] 실시예에서, 음장 합성은, 재생 설정과 관련하여 새로운 방향 정보에 의해서 주어진 방향에 대해서 음원 신호를 패닝(panning)하는 것에 의해서 실시되고, 음원 신호의 스케일링(scaling)은, 패닝 동작을 실시하기 이전에 또는 패닝 동작의 실시 후에, 거리 정보를 이용하여 이루어진다.
- [0039] 추가적인 실시예에서, 음원 신호의 확산 부분이 음원 신호의 다이렉트 부분(direct part)에 부가되고, 다이렉트 부분은, 확산 부분에 부가되기 전에, 거리 정보에 의해서 수정된다.
- [0040] 특히, 음원 합성을 공간적 표상 내에서 실시하는 것이 바람직하고, 여기에서 새로운 방향 정보가 각각의 주파수 빈에 대해서 계산되고, 거리 정보가 각각의 주파수 빈에 대해서 계산되며, 해당 주파수 빈에 대한 오디오 신호를 이용하는 각각의 주파수 빈에 대한 다이렉트 합성이 해당 주파수 빈에 대한 오디오 신호와, 새로운 방향 정보로부터 유도된 해당 주파수 빈에 대한 패닝 이득과, 그리고 해당 주파수 빈에 대한 거리 정보로부터 유도된 주파수 빈에 대한 스케일링 인자를 이용하여 실시된다.
- [0041] 또한, 확산 합성이 해당 주파수 빈으로부터의 오디오 신호로부터 유도된 확산 오디오 신호를 이용하여 그리고 해당 주파수 빈에 대한 신호 분석에 의해서 유도된 확산 매개변수를 이용하여 실시되고, 이어서, 다이렉트 신호와 확산 신호를 조합하여 시간 또는 주파수 빈에 대한 합성된 오디오 신호를 획득하고, 이어서, 다른 시간/주파수 빈에 대한 오디오 신호를 이용하여 주파수-시간 변환을 실시하여 시간 도메인 합성된 오디오 신호를 수정된 음장으로서는 획득한다.
- [0042] 그에 따라, 일반적으로, 음장 계산기는, 예를 들어, 음원 신호에 대한 새로운 방향을 이용하여, 각각의 음원에 대해서, 음원 신호를 프로세스하여 상이한/새로운 기준 위치와 관련된 음원 신호의 음장 묘사를 획득하는 것에 의해서, 각각의 음원에 대해서, 상이한 기준 위치와 관련된 음장을 합성하도록 구성된다. 또한, 음원 신호는, 음원 신호를 프로세스하기 전에 또는 방향 정보를 이용하여 음원 신호를 프로세스하는 것에 후속하여, 수정된다. 그리고, 마지막으로, 음원에 대한 음장 묘사가 함께 부가되어, 상이한 기준 위치와 관련된 수정된 음장을 획득한다.
- [0043] 추가적인 실시예에서, 그리고 특히, 음장 묘사 및 음장 묘사의 공간적 정보와 관련된 메타데이터로부터 수정된 음장 묘사를 생성하기 위해서, 음장 계산기는, 제1 음장 묘사의 공간적 정보를 이용하여, 제2 음장 묘사의 공간적 정보를 이용하여, 그리고 상이한 기준 위치로의 기준 위치의 병진 운동을 나타내는 병진 운동 정보를 이용하여, 수정된 음장을 계산한다. 특히, 메타데이터는, 예를 들어, 음장 묘사의 기준 위치로 지향된 벡터 및 동일한 원점으로부터 제2 음장 묘사의 추가적인 기준 위치까지 지향된 다른 벡터일 수 있다.
- [0044] 병진 운동 정보를 어드레스하기 위해서, 음원 분리, 또는 빔 형성, 또는, 일반적으로, 임의의 종류의 음원 분석을 제1 및 제2 음장 묘사에 적용하는 것에 의해서, 객체가 생성된다. 이어서, 객체가 광대역 객체 또는 개별적인 시간/주파수 빈을 위한 객체인지의 여부와 관계없이, 모든 객체의 도달 방향 정보가 연산된다. 이어서, 적어도 하나의 매칭된 객체, 즉 제1 및 제2 음장 묘사 모두에서 발생되는 객체를 찾기 위해서, 상이한 음장 묘사로

부터 추출된 객체들을 서로 매칭시킨다. 이러한 매칭은, 예를 들어, 객체 신호 및/또는 도달 방향 정보 또는 다른 정보를 이용하는 상관 또는 가간섭성(coherence) 계산에 의해서 실시된다.

[0045] 따라서, 그러한 과정의 결과는, 매칭된 객체에 대한, 기준 위치와 관련된 제1 DoA 정보 및 추가적인 기준 위치와 관련된 제2 DoA 정보가 존재한다는 것이다. 이어서, 매칭된 객체의 위치 그리고, 특히, 기준 위치 또는 추가적인 기준 위치에 대한 매칭된 객체의 거리가, 기준 위치 또는 연관된 메타데이터에 포함된 기준 위치에 관한 정보를 이용하여 삼각측량을 기초로 계산된다.

[0046] 이러한 정보, 그리고, 특히, 이어서, 매칭된 객체에 대한 위치 정보를 이용하여, 추정된 위치 및 희망 위치를 기초로, 즉 거리 보상 프로세싱을 이용한 병진 운동 후에, 각각의 매칭된 객체를 수정한다. 새로운 청취자 위치에 대한 새로운 DoA 정보를 계산하기 위해서, 양 기준 위치로부터의 이전의 DoA 정보 및 병진 운동 정보가 이용된다. 기본적으로, 이러한 프로세싱은 양 개별적 음장 묘사를 위해서 실시될 수 있는데, 이는 각각의 매칭된 객체가 양 음장 묘사에서 발생되기 때문이다. 그러나, 바람직한 실시예에 따라, 병진 운동 후에 새로운 청취자 위치에 가장 가까운 기준 위치를 가지는 음장 묘사가 이용된다.

[0047] 이어서, 새로운 DoA를 이용하여, 상이한 기준 위치 즉, 사용자가 이동된 기준 위치와 관련된 매칭된 객체를 위한 새로운 음장 묘사를 계산한다. 이어서, 그리고 비-매칭 객체를 통합하기 위해서, 그러한 객체를 위한 음장 묘사가 또한 계산되나, 이때 이전의 DoA 정보가 이용된다. 그리고, 마지막으로, 수정된 음장은 모든 개별적인 음장 묘사들을 함께 더하는 것에 의해서 생성된다.

[0048] 단일 회전을 가상 앰비소닉 신호에 인가하는 것에 의해서, 배향의 임의의 변화가 실현될 수 있다.

[0049] 따라서, 메타데이터는 기준 위치에 대한 객체의 거리를 직접적으로 제공하기 위해서 이용되지 않는다. 그 대신, 메타데이터는 2개 이상의 음장 묘사의 각각의 기준 위치를 식별하기 위해서 제공되고, 기준 위치와 특정의 매칭된 객체 사이의 거리는, 예를 들어, 삼각측량 프로세싱 단계를 기초로 계산된다.

발명의 효과

[0050] 본원에서 제시된 방법은, 단일 FOA 마이크로폰의 레코딩으로부터의 6DoF 재생을 가능하게 한다. 단일 공간적 위치로부터의 레코딩은 3DoF 재생 또는 어쿠스틱 줌을 위해서 이용되었다. 그러나, 본 발명자의 지식에 따르면, 그러한 데이터로부터의, 인터랙티브, 완전 6DoF 재생을 위한 방법은 아직까지 제시되지 않았다. 레코딩 내의 음원의 거리에 관한 정보를 통합함으로써 6DoF 재생을 실현한다. 이러한 거리 정보는 DirAC의 파라메트릭 표상 내로 통합되고, 그에 따라 청취자의 변화된 원근(perspective)이 정확하게 맵핑된다.

[0051] 앰비소닉 음장 표상(일반적인 FOA 또는 HOA 앰비소닉과 같거나 또는 DirAC-스타일 파라메트릭 음장 표상과 같은 것임)은, 6DoF 적용예에서 요구되는 바와 같은 청취자 위치의 병진 운동적 이동을 가능하게 하는 충분한 정보를 제공하지 못하는데, 이는, 음향 장면 내의 객체 거리 및 절대 객체 위치 모두가 이러한 포맷으로 결정되지 않기 때문이다. 청취자 위치의 이동이 반대 방향을 따른 음향 장면의 동등한 이동으로 변환될 수 있다는 것을 주목하여야 한다.

도면의 간단한 설명

[0052] 이하에서, 첨부 도면과 관련하여 본 발명의 바람직한 실시예를 설명할 것이다.

도 1a는 증강된 음장 묘사를 생성하기 위한 장치의 바람직한 실시예이다.

도 1b는 본 발명의 기초가 되는 예시적인 문제를 설명하는 도면이다.

도 2는 증강된 음장 묘사를 생성하기 위한 장치의 바람직한 구현예이다.

도 3a는 오디오 데이터 및 오디오 데이터를 위한 부수적 정보를 포함하는 증강된 음장 묘사를 도시한다.

도 3b는 오디오 데이터 및 각각의 음장 묘사에 대한 기하형태적 정보와 같은 공간적 정보와 관련된 메타데이터를 포함하는 증강된 음장 묘사의 추가적인 도면을 도시한다.

도 4a는 수정된 음장 묘사를 생성하기 위한 장치의 구현예를 도시한다.

도 4b는 수정된 음장 묘사를 생성하기 위한 장치의 추가적인 구현예를 도시한다.

도 4c는 기준 배치/위치 A, 추가적인 기준 배치/위치 B, 및 병진 운동에 기인한 상이한 기준 위치를 갖는 시나

리오를 도시한다.

도 5는 일반적인 의미에서 공간적 오디오의 6DoF 재생을 나타낸다.

도 6a는 음장 계산기의 구현을 위한 바람직한 실시예를 도시한다.

도 6b는 새로운/상이한 기준 위치와 관련하여 음원의 새로운 DoA 및 새로운 거리를 계산하기 위한 바람직한 구현예를 도시한다.

도 6c는, 예를 들어, 각각의 개별적인 음장 묘사를 위한, 증강된 음장 묘사를 생성하기 위한 장치 및 매칭된 음원을 위한 수정된 음장 묘사를 생성하기 위한 장치를 포함하는 6DoF 재생의 바람직한 실시예를 도시한다.

도 7은 광대역 또는 협대역 객체에 대한 수정된 음장을 계산하기 위해서 제1 및 제2 음장 묘사 중 하나를 선택하기 위한 바람직한 실시예를 도시한다.

도 8은 모노-신호와 같은 오디오 신호 및 도달 방향 데이터로부터 음장 묘사를 생성하기 위한 예시적인 장치를 도시한다.

도 9는 음장 계산기의 추가적인 바람직한 실시예를 도시한다.

도 10은 수정된 음장 묘사를 생성하기 위한 장치의 바람직한 구현예를 도시한다.

도 11은 수정된 음장 묘사를 생성하기 위한 장치의 추가적인 바람직한 구현예를 도시한다.

도 12a는 종래 기술의 DirAC 분석 구현예를 도시한다.

도 12b는 종래 기술의 DirAC 합성 구현예를 도시한다.

발명을 실시하기 위한 구체적인 내용

- [0053] 언급한 앰비소닉/DirAC 표상을 위해서 6DoF를 적용할 수 있게 하기 위해서, 병진 운동 프로세싱(translational processing)을 위한 손실 정보(missing information)를 제공하는 방식으로 이러한 표상(representations)을 확장시킬 필요가 있다. 이러한 확장이 예를 들어 1) 객체의 거리 또는 위치를 기존 장면 표상에 부가할 수 있다는 것, 및/또는 2) 개별적인 객체들을 분리하는 프로세스를 도울 수 있는 정보를 부가하는 것일 수 있다는 것을 주목하여야 한다.
- [0054] 또한, 실시예의 목적은, 이하의 의미에서, 이러한 표상/시스템과의 역방향 양립성을 제공하기 위해서 기존 (비-파라메트릭 또는 파라메트릭) 앰비소닉 시스템의 구조를 보존/재사용하는 것이며,
- [0055] * 확장된 표상이 (예를 들어, 렌더링을 위한) 기존 비-확장 표상으로 변환될 수 있고, 그리고
- [0056] * 확장된 표상으로 작업할 때 기존 소프트웨어 및 하드웨어 구현예의 재사용을 가능하게 한다.
- [0057] 이하에서, 몇몇 접근 방식, 즉 하나의 제한된 (그러나 매우 단순한) 접근방식 및 6DoF를 가능하게 하기 위한 3개의 상이한 확장된 앰비소닉 포맷이 설명된다.
- [0058] 음향 장면은 둘 이상의 앰비소닉 신호를 이용하여 묘사되고, 각각의 앰비소닉 신호는 상이한 위치에서, 또는 달리 말해서 상이한 원근으로부터 음향 장면을 묘사한다. 상대적인 위치들이 알려져 있는 것으로 가정한다. 음향 장면 내의 희망 위치에서의 수정된 앰비소닉 신호가 입력 앰비소닉 신호로부터 생성된다. 신호-기반의 또는 파라메트릭-기반의 접근방식을 이용하여 희망 위치에서 가상 앰비소닉 신호를 생성할 수 있다.
- [0059] 다중-지점 앰비소닉 표상의 개념은 통상적 및 파라메트릭(DirAC-스타일) 앰비소닉 모두에 적용될 수 있다.
- [0060] 희망 위치에서의(즉, 병진 운동 후의) 가상의 앰비소닉 신호가 신호-기반의 병진 운동 실시예에서 이하의 단계를 이용하여 연산된다.
- [0061] 1. 음원 분리를 각각의 통상적인 앰비소닉 신호에 인가함으로써, 객체가 생성된다.
- [0062] 2. 모든 객체의 DoA가 각각의 통상적인 앰비소닉 신호에 대해서 연산된다.
- [0063] 3. 하나의 통상적인 앰비소닉 신호로부터 추출된 객체가 다른 통상적인 앰비소닉 신호로부터 추출된 객체에 대해서 매칭된다. 그러한 매칭은 (예를 들어, 상관/가간섭성에 의해서) 상응 DoA 및/또는 신호를 기초로 실시된다.

- [0064] 4. 매칭된 객체의 위치가 삼각측량을 기초로 추정된다.
- [0065] 5. 각각의 매칭된 객체(단일-채널 입력)가, 거리 보상 필터를 이용하여 추정된 위치 및 희망 위치(즉, 병진 운동 후)를 기초로 수정된다.
- [0066] 6. 희망 위치(즉, 병진 운동 이후)에서의 DoA가 각각의 매칭된 객체에 대해서 연산된다. 이러한 DoA는 DoA'에 의해서 표시된다.
- [0067] 7. 앰비소닉 객체 신호는 각각의 매칭된 객체에 대해서 연산된다. 매칭된 객체가 도착-방향 DoA'를 갖도록, 앰비소닉 객체 신호가 생성된다.
- [0068] 8. 앰비소닉 객체 신호는 각각의 비-매칭 객체에 대해서 연산된다. 비-매칭 객체가 도착-방향 DoA를 갖도록, 앰비소닉 객체 신호가 생성된다.
- [0069] 9. 모든 앰비소닉 객체 신호들을 함께 더하는 것에 의해서, 가상 앰비소닉 신호가 얻어진다.
- [0070] 희망 위치에서의(즉, 병진 운동 후의) 가상의 앰비소닉 신호가, 추가적인 실시예에 따른 파라메트릭-기반의 병진 운동 실시예에서 이하의 단계를 이용하여 연산된다.
- [0071] 1. 음장 모델이 가정된다. 음장은 하나 이상의 다이렉트 사운드 성분 및 확산 사운드 성분으로 분해될 수 있다. 다이렉트 사운드 성분은 신호 및 위치 정보(예를 들어, 극좌표 또는 데카르트 좌표)로 이루어진다. 대안적으로, 음장이 하나 이상의 다이렉트/원칙적 사운드 성분 및 나머지 사운드 성분(단일-채널 또는 다중-채널)으로 분해될 수 있다.
- [0072] 2. 가정된 음장 모델의 신호 성분 및 매개변수가 입력 앰비소닉 신호를 이용하여 추정된다.
- [0073] 3. 음향 장면에서, 희망 병진 운동, 또는 희망 위치에 따라 신호 성분 및/또는 매개변수가 수정된다.
- [0074] 4. 수정된 신호 성분 및 수정된 매개변수를 이용하여 가상 앰비소닉 신호가 생성된다.
- [0075] 다중-지점 앰비소닉 신호를 생성하는 것은, 컴퓨터로 생성되고 생산된 콘텐츠에서뿐만 아니라 마이크로폰 어레이 또는 공간적 마이크로폰(예를 들어, B-포맷 마이크로폰)을 통한 자연적인 레코딩의 맥락에서 단순하다. 실시예에서, 단계(2)에 후속하여 음원 매칭을 또는 단계 3 전에 삼각측량 계산을 실시하는 것이 바람직하다. 또한, 양 실시예 중 하나 이상의 단계가 또한 상응하는 다른 실시예에서 이용될 수 있다.
- [0076] 단일 회전을 가상 앰비소닉 신호에 인가하는 것에 의해서, 배향의 변화가 실현될 수 있다.
- [0077] 도 1a는 적어도 하나의 기준 위치와 관련된 음장을 나타내는 적어도 하나의 음장 묘사를 생성하기 위한 음장(묘사) 생성기(100)를 포함하는 증강된 음장 묘사를 생성하기 위한 장치를 도시한다. 또한, 장치는 음장의 공간 정보와 관련된 메타데이터를 생성하기 위한 메타데이터 생성기(110)를 포함한다. 메타데이터는, 입력으로서, 음장을, 또는 대안적으로 또는 부가적으로, 음원에 관한 분리된 정보를 수신한다.
- [0078] 음장 묘사 생성기(100)와 메타데이터 생성기(110)의 출력 모두가 증강된 음장 묘사를 구성한다. 실시예에서, 음장 묘사 생성기(100) 및 메타데이터 생성기(110)의 출력 모두를 조합기(120) 또는 출력 인터페이스(120) 내에서 조합하여, 메타데이터 생성기(110)에 의해서 생성되는 것과 같은 음장의 공간 정보 또는 공간 메타데이터를 포함하는 증강된 음장 묘사를 획득한다.
- [0079] 도 1b는 본 발명에 의해서 해결되는 상황을 도시한다. 예를 들어, 위치 A는 적어도 하나의 기준 위치이고, 음장은 음원 A 및 음원 B에 의해서 생성되고, 위치 A에 위치되는 특정의 실제의 또는, 예를 들어, 가상의 마이크로폰이 음원 A 및 음원 B로부터 사운드를 검출한다. 사운드는 사운드 방출 음원들로부터 오는 사운드의 중첩이다. 이는, 음장 묘사 생성기에 의해서 생성된 바와 같은 음장 묘사를 나타낸다.
- [0080] 또한, 메타데이터 생성기는, 특정 구현예에 의해서, 음원 A에 대한 공간적 정보 및 음원 B에 대한 다른 공간적 정보, 예를 들어 이러한 음원으로부터 위치 A와 같은 기준 위치까지의 거리를 유도할 수 있다.
- [0081] 당연하게, 기준 위치는, 대안적으로, 위치 B일 수 있다. 이어서, 실제 또는 가상 마이크로폰이 위치 B에 배치될 수 있고, 음장 묘사는 음장, 예를 들어 1차 앰비소닉 성분 또는 더 높은 차수의 앰비소닉 성분 또는 적어도 하나의 기준 위치 즉, 위치 B와 관련된 음장을 묘사하는 포텐셜을 갖는 임의의 다른 사운드 성분에 의해서 표시된, 음장일 수 있다.
- [0082] 이어서, 메타데이터 생성기는, 음원에 관한 정보로서, 위치 B에 대한 음원 A의 거리 또는 위치 B에 대한 음원 B

의 거리를 생성할 수 있다. 물론, 음원에 관한 대안적인 정보가 절대적 위치 또는 기준 위치에 대한 상대적 위치일 수 있다. 기준 위치는 일반적인 좌표계의 원점일 수 있거나, 일반적인 좌표계의 원점에 대한 규정된 관계로 위치될 수 있다.

- [0083] 다른 메타데이터는 하나의 음원의 절대 위치 및 제1 음원에 대한 다른 음원의 상대적인 위치 동일 수 있다.
- [0084] 도 2는 증강된 음장 묘사를 생성하기 위한 장치를 도시하고, 여기에서 음장 생성기는 제1 음장을 위한 음장 생성기(250), 제2 음장을 위한 음장 생성기(260), 및 제3, 제4 등의 음장과 같은 하나 이상의 음장을 위한 임의의 수의 음장 생성기를 포함한다. 또한, 메타데이터는 제1 음장 및 제2 음장에 관한 정보를 계산하고 조합기(120)에 전달하도록 구성된다. 모든 이러한 정보는, 증강된 음장 묘사를 생성하기 위해서 조합기(120)에 의해서 이용된다. 따라서, 조합기(120)는 또한 증강된 음장 묘사를 생성하기 위한 출력 인터페이스로서 구성된다.
- [0085] 도 3a는, 제1 음장 묘사(330), 제2 음장 묘사(340) 및, 그와 연관된, 제1 음장 묘사 및 제2 음장 묘사에 관한 정보를 포함하는 메타데이터(350)를 포함하는 데이터스트림으로서, 증강된 음장 묘사를 도시한다. 제1 음장 묘사는, 예를 들어, B-포맷 묘사 또는 더 높은 차수의 묘사 또는 전체-대역 표상으로 또는 주파수-선택된 표상으로 음원의 방향 분포를 결정할 수 있게 하는 임의의 다른 묘사일 수 있다. 따라서, 제1 음장 묘사(330) 및 제2 음장 묘사(340)은, 예를 들어, 상이한 시간/주파수 빈에 대한 다운믹스 신호 및 도달 방향 데이터를 가지는, 예를 들어, 상이한 기준 위치들에 대한 파라메트릭 음장 묘사일 수 있다.
- [0086] 그럼에도 불구하고, 제1 및 제2 음장 묘사에 대한 기하형태적 정보(350)는, 각각, 제1 음장 묘사(330)에 포함된 모든 음원에 대해서, 또는 제2 음장 묘사(340) 내의 음원에 대해서 동일하다. 따라서, 예시적으로, 제1 음장 묘사(330) 내에 3개의 음원 및 제1 음장 묘사에 관한 기하형태적 정보가 있는 경우에, 이러한 기하형태적 정보는 제1 음장 묘사 내의 3개의 음원에 대해서 동일하다. 유사하게, 예를 들어, 제2 음장 묘사 내에 5개의 음원이 존재할 때, 메타데이터(350) 내에 포함된 제2 음장에 대한 기하형태적 정보는 제2 음장 묘사 내의 모든 음원에 대해서 동일하다.
- [0087] 도 3b는 도 3a의 메타데이터(350)의 예시적인 구성을 도시한다. 실시예에서, 기준 위치(351)가 메타데이터 내에 포함될 수 있다. 그러나, 이는, 기준 위치 정보(351)가 또한 생략될 수 있는 경우에 필수적이지 않다.
- [0088] 제1 음장의 경우에, 예를 들어 원점으로부터, 제1 음장이 관련되는 기준 배치/위치 A로 향하는 도 4c에 도시된 벡터 A에 관한 정보일 수 있는 제1 기하형태적 정보가 주어진다.
- [0089] 제2 기하형태적 정보는, 예를 들어, 원점으로부터, 제2 음장 묘사가 관련되는 제2 기준 배치/위치 B로 향하는 벡터 B에 관한 정보일 수 있다.
- [0090] A 및 B는 양 음장 묘사에 대한 기준 위치 또는 레코딩 위치이다.
- [0091] 대안적인 기하형태적 정보는, 예를 들어, 기준 위치 A와 추가적인 기준 위치 B 및/또는 원점 사이에서 연장되는 벡터 D 그리고 원점으로부터 양 지점 중 하나로 향하는 벡터에 관한 정보일 수 있다. 따라서, 메타데이터 내에 포함되는 기하형태적 정보가 벡터 A 및 벡터 D일 수 있거나, 벡터 B 및 벡터 D를 포함할 수 있거나, 벡터 D가 없이 벡터 A 및 벡터 B를 포함할 수 있거나, 특정 3차원적 좌표계 내에서 기준 위치 A 및 기준 위치 B를 식별할 수 있게 하는 다른 정보를 포함할 수 있다. 그러나, 동일한 고려사항이 2차원적 사운드 묘사를 위해서 또한 적용될 뿐만 아니라, 특히 2차원적인 경우만을 도시하는 도 4c에 도시되어 있다.
- [0092] 도 4a는 음장 묘사 및 음장 묘사의 공간적 정보와 관련된 메타데이터로부터, 수정된 음장 묘사를 생성하기 위한 장치의 바람직한 구현예를 도시한다. 특히, 그러한 장치는, 메타데이터를 이용하여 수정된 음장을 생성하는 음장 계산기(420)를 포함하고, 음장 묘사 및 병진 운동 정보는 기준 위치로부터 다른 기준 위치로의 병진 운동을 나타낸다.
- [0093] 실시예에서, 음장 계산기(420)는, 예를 들어, 도 1a 또는 도 2와 관련하여 설명한 바와 같은 증강된 음장 묘사를 수신하기 위한 입력 인터페이스(400)에 연결되고, 이어서 입력 인터페이스(400)는 한편으로 음장 묘사 즉, 도 1a의 블록(100) 또는 도 2의 블록(210)에 의해서 생성된 것을 분리한다. 또한, 입력 인터페이스(400)는 메타데이터를 증강된 음장 묘사로부터, 즉 도 3a의 항목(350) 또는 도 3b의 선택적인 항목(351 및 352 내지 354)으로부터 분리한다.
- [0094] 또한, 병진 운동 인터페이스(410)는 청취자로부터 병진 운동 정보 및/또는 부가적인 또는 분리된 회전 정보를 획득한다. 병진 운동 인터페이스(410)의 구현예가, 가상 현실 환경에서 머리의 회전뿐만 아니라, 하나의 위치 즉, 도 1b의 위치 A로부터 다른 위치 즉, 도 1b의 위치 B로의 머리의 병진 운동을 추적하는 머리-추적 유닛일

수 있다.

- [0095] 도 4b는 도 1a와 유사한, 그러나 인코더/디코더 시나리오와 관련되지 않고, 메타데이터 공급기(402)에 의해서 표시된 메타데이터 공급, 음장 공급기(404)에 의해서 표시된 음장 공급이, 인코딩된 또는 증강된 음장 묘사를 분리하는 특정 입력 인터페이스가 없이 이루어지나, 예를 들어, 모두가, 예를 들어, 가상 현실 적용예에서 존재하는 실제 시나리오에서 이루어지는 일반적인 시나리오와 관련되는 다른 구현예를 도시한다. 그러나, 본 발명은 가상 현실 적용예로 제한되지 않고, 임의의 다른 적용예에서 또한 구현될 수 있고, 그러한 다른 적용예에서, 기준 위치와 관련된 음장의 공간적 오디오 프로세싱은, 제1 기준 위치와 관련된 음장을 다른 제2 기준 위치와 관련된 다른 음장으로 변환하는데 있어서 유용하다.
- [0096] 이어서, 음장 계산기(420)는 수정된 음장 묘사를 생성하거나, 대안적으로, (가상) 라우드스피커 표상을 생성하거나 헤드폰 재생을 위한 2-채널 표상과 같은 바이노럴 표상을 생성한다. 따라서, 음장 계산기(420)는, 수정된 음장으로서, 수정된 음장 묘사를 생성할 수 있고, 이는 기본적으로 원래의 음장 묘사와 동일하나, 이제는 새로운 기준 위치와 관련된다. 대안적인 실시예에서, 가상의 또는 실제의 라우드스피커 표상은 5.1 체계와 같은 미리 결정된 라우드스피커 설정 또는 더 많은 라우드스피커를 갖는 그리고, 특히, 2차원적 배열만이 아니라 3차원적인 라우드스피커의 배열, 즉 사용자 위치와 관련하여 상승된 라우드스피커를 가지는 라우드스피커 배열을 가지는 라우드스피커 설정을 위해서 생성될 수 있다. 가상 현실 적용예를 위해서 특히 유용한 다른 적용예는 바이노럴 재생을 위한 즉, 가상 현실 사용자의 머리에 적용될 수 있는 헤드폰을 위한 적용예이다.
- [0097] 예시적으로, 후술되는 도 6은, DirAC 합성기가 전방향적 또는 압력 성분과 같은 다운믹스 성분에서만 동작되는 한편, 도 12b와 관련하여 도시된 추가적인 대안적 실시예에서, DirAC 합성기는 전체 음장 데이터에서, 즉 도 12b의 이러한 실시예에서, 전방향 성분(w) 및 3개의 방향 성분(x, y, z)을 갖는 필드 묘사를 가지는 전체 성분 표상에서 동작된다.
- [0098] 도 4c는 본 발명의 바람직한 실시예의 기초가 되는 시나리오를 도시한다. 그러한 도면은 제1 기준 배치/위치 A, 제2 기준 배치/위치 B 및 2개의 상이한 음원 A 및 B, 그리고 병진 운동 벡터(I)를 도시한다.
- [0099] 음원 A 및 B 모두는 기준 위치 A와 관련된 음장 묘사 및 기준 위치 B에 관련된 제2 음장 묘사에 포함된다.
- [0100] 예를 들어, 제1 기준 위치 또는 제2 기준 위치까지의 음원 A의 거리를 계산하기 위해서, A 및 B와 관련된 상이한 2개의 음장 묘사 모두가 음원 분리 과정을 거치고, 이어서, 이러한 상이한 사운드 분리 과정에 의해서 얻어진 음원들의 매칭이 얻어진다. 이는, 예를 들어, 음원 A를 초래할 것이다. 음원 A는 제1 음장 묘사를 위한 그리고 또한 제2 음장 묘사를 위한 음원 분리 알고리즘에서 발견된다. 음원 A를 위한 도달 방향 정보는, 기준 위치 A와 관련된 제1 음장 묘사로부터 얻어질 때, 각도(α)가 될 것이다. 또한, 동일 음원 A에 대한 그러나 이제 추가적인 기준 위치 B와 관련된 제2 음장 묘사로부터 얻어진 도달 방향 정보가 각도(β)가 될 것이다.
- [0101] 이제, 예를 들어, 음장 묘사를 위한 메타데이터로부터 얻어질 수 있거나 계산될 수 있는 알고 있는 또는 계산 가능한 거리(D)를 기초로, 그리고 2개의 각도(α 및 β), 음원 A에 의해서 규정된 삼각형을 기초로, 기준 위치 A 및 기준 위치 B가 완전히 규정된다. 따라서, 음원 A로부터 기준 위치 A까지의 거리 또는 음원 A로부터 기준 위치 B까지의 거리 또는 음원 A의 일반 위치 즉, 원점으로부터 음원 A의 실제 위치를 향하는 벡터가, 예를 들어 삼각측량 프로세싱 동작에 의해서, 계산될 수 있다. 위치 또는 거리 모두는 거리 또는 위치에 관한 정보를 나타낸다.
- [0102] 이어서, 동일한 과정이 각각의 매칭된 음원에 대해서, 즉 또한 음원 B에 대해서 실시될 수 있다.
- [0103] 따라서, 각각의 매칭된 음원에 대한 거리/위치 정보가 계산되고, 이어서, 거리/위치가 완전히 알려진 것과 같이 또는, 예를 들어, 부가적인 메타데이터에 의해서 주어진 것과 같이, 각각의 매칭된 음원이 프로세스될 수 있다. 그러나, 각각의 개별적인 음원에 대한 임의의 거리/깊이 정보 대신, 제1 음장 묘사 및 제2 음장 묘사에 대한 기하학적 정보만이 필요하다.
- [0104] 도 8은 DirAC 합성기와 상이한 합성을 실시하기 위한 다른 구현예를 도시한다. 예를 들어, 음장 분석기가, 각각의 음원 신호에 대해서, 분리된 모노 신호(S) 및 원래의 도달 방향을 생성할 때 그리고, 병진 운동 정보에 따라서, 새로운 도달 방향이 계산될 때, 예를 들어, 도 8의 앰비소닉 신호 생성기(430)를 이용하여, 음원 신호, 즉 모노 신호(S)에 대한 그러나 수평 각도(θ) 또는 양각(Θ) 및 방위각(ϕ)으로 이루어진 새로운 도달 방향(DoA) 데이터에 대한 음장 묘사를 생성할 수 있다. 이어서, 도 4b의 음장 계산기(420)에 의해서 실시되는 과정이, 예를 들어, 새로운 도달 방향을 갖는 각각의 음원에 대한 1차 앰비소닉 음장 표상을 생성할 것이고, 이어서, 음원 마다의 추가적인 수정이 새로운 기준 위치에 대한 음장의 거리에 따라 달라지는 스케일링 인자를 이용하여 실시

되며, 이어서, 개별적인 음원으로부터의 모든 음장이 서로 중첩되어, 다시, 예를 들어, 특정의 새로운 기준 위치와 관련된 앰비소닉 표상에서, 수정된 음장을 최종적으로 획득할 수 있다

[0105] 도 6의 DirAC 분석기(422, 422a, 422b)에 의해서 프로세스되는 각각의 시간/주파수 빈이 특정 (제한된 대역폭의) 음원을 나타내는 것으로 해석할 때, 앰비소닉 신호 생성기(430)가, DirAC 합성기(425, 425a, 425b) 대신, 각각의 시간/주파수 빈에 대해서, 다운믹스 신호 또는 압력 신호 또는 이러한 시간/주파수 빈에 대한 전 방향 성분을 도 8의 "모노 신호(S)"로서 이용하여, 전체 앰비소닉 표상을 생성할 수 있다. 이어서, W, X, Y, Z 성분의 각각에 대한 주파수-시간 변환기에서의 개별적인 주파수-시간 변환이, 도 4c에 도시된 것과 상이한 음장 묘사를 초래할 것이다.

[0106] 그러한 장면은 마이크론의 시점(PoV)으로부터 레코딩되고, 그 위치는 기준 좌표계의 원점으로서 이용된다. 장면은, 6DoF에서 추적되는, 청취자의 PoV로부터 재생되어야 하고, 이에 대해서는 도 5를 참조한다. 단일 음원이 여기에서 예시를 위해서 도시되어 있고, 관계는 각각의 시간-주파수 빈에 대해서 유지된다.

[0107] 도 5는 공간적 오디오의 6DoF 재생을 도시한다. 음원은 마이크론 위치 및 배향(검은색 라인 및 원호)에 대한 거리(d_r)에서 DoA(r_r)을 갖춘 마이크론에 의해서 레코딩된다. 이는 DoA(r_r) 및 거리(d_r)(쇄선)를 가지고 이동 청취자에 대해서 재생되어야 한다. 이는, 청취자 병진 운동(l) 및 회전(o)(점선)을 고려하여야 한다. DOA는 음원을 향하는 단위 길이의 벡터로서 표시된다.

[0108] 좌표($d_r \in \mathbb{R}^3$)에서의 음원이 단위 벡터($r_r = d_r / \|d_r\|$)에 의해서 표현되는 도달 방향(DoA)으로부터 레코딩된다. 이러한 DoA는 레코딩의 분석으로부터 추정될 수 있다. 이는 거리($d_r = \|d_r\|$)로부터 온다. 이러한 정보가 각각의 음원에 대한 메타데이터로부터 또는, 일반적으로, 도 3b의 항목(352, 353, 354)으로부터 유도될 수 있고, 예를 들어 2개의 상이한 기준 위치들 및 거리/위치 또는 기준 위치와 관련된 2개의 각도를 이용하는 삼각측량 프로세싱에 의해서 얻어지는 바와 같은 레코딩 위치로부터 (예를 들어, 미터 등으로 주어진) 거리까지 임의의 방향(r)을 가지는 거리 묘사로서 표시될 수 있는 것으로 가정된다.

[0109] 청취자는 6DoF로 추적된다. 주어진 시간에, 청취자는 마이크론에 대해서 위치($l \in \mathbb{R}^3$)에 있고, 마이크론의 좌표계에 대해서 회전($o \in \mathbb{R}^3$)을 갖는다. 표기를 단순화하기 위해서, 레코딩 위치가 좌표계의 원점으로서 선택된다.

[0110] 따라서, 사운드는, 변화된 볼륨을 초래하는, 상이한 거리(d_1), 및 병진 운동 및 후속 회전 모두의 결과인 상이한 DoA(r_1)로 재생되어야 한다.

[0111] 이하의 섹션에서 설명되는 바와 같이, 파라메트릭 표상을 기초로 하는 전용 변환에 의해서 청취자 원근으로부터 가상 신호를 획득하기 위한 방법이 설명된다.

[0112] 제시된 방법은 파라메트릭 공간 사운드 인코딩을 위한 기본 DirAC 접근방식을 기초로 하고, 이에 대해서는 [16]를 참조한다. 분석 스펙트럼의 시간-주파수 순간마다 하나의 주요 다이렉트 음원이 있고, 이들이 독립적으로 처리될 수 있는 것으로 가정한다. 레코딩은 국소 푸리에 변환(STFT)을 이용하여 시간-주파수 표상으로 변환된다. 시간 프레임 인덱스는 n 으로 표시되고 주파수 인덱스는 k 로 표시된다. 이어서, 변환된 레코딩을 분석하여, 복소수 스펙트럼($P(k,n)$)의 각각의 시간-주파수 빈에 대한 방향($r_r(k,n)$) 및 확산($\psi(k,n)$)을 추정한다. 합성에서, 신호가 다이렉트 및 확산 부분으로 분할된다. 여기에서, 라우드스피커 신호는, 스피커 위치에 따라 다이렉트 부분을 패닝하는 것 그리고 확산 부분을 부가하는 것에 의해서 연산된다.

[0113] 6DoF에서 청취자 원근에 따라 FOA 신호를 변환하기 위한 방법이 5개의 단계로 분할될 수 있고, 이에 대해서 도 6c를 참조할 수 있다.

[0114] 도 6c는 6DoF 재생 방법을 도시한다. B-포맷의 레코딩된 FOA 신호가 DirAC 인코더에 의해서 프로세스되고, DirAC 인코더는 복소수 스펙트럼의 각각의 시간-주파수 빈에 대한 방향 및 확산을 연산한다. 이어서, 방향 벡터가 청취자의 추적된 위치에 의해서 그리고, 예를 들어, 삼각측량 계산에 의해서 유도된 각각의 음원에 대한 거리 맵에서 주어진 거리 정보에 따라 변환된다. 이어서, 머리 회전에 따라 결과적인 방향 벡터가 회전된다. 마치

막으로, 8+4개의 가상 라우드스피커 채널에 대한 신호가 DirAC 디코더에서 합성된다. 이들은 이어서 바이노럴화된다.

[0115] 실시예에서, 입력 신호가 DirAC 인코더(422)에서 분석되고, 거리 정보는 각각의 (매칭된) 음원에 대한 거리를 제공하는 거리 맵(m(r))으로부터 부가되고, 이어서 청취자의 추적된 병진 운동 및 회전이 신규 변환(423 및 424)에서 적용된다. DirAC 디코더(425)는 8+4개의 가상 라우드스피커를 위해서 신호들을 합성하고, 이는 다시 헤드폰 플레이백을 위해서 바이노럴화된다(427). 병진 운동 이후의 음향 장면의 회전이 독립적인 동작이기 때문에, 이러한 것이 바이노럴 렌더러 내에서 대안적으로 적용될 수 있다는 것을 주목하여야 한다. 6DoF을 위해서 변환된 매개변수만이 방향 벡터이다. 모델 정의에 의해서, 확산 부분이 등방적 및 균질한 것으로 가정되고, 그에 따라 변화되지 않고 유지된다.

[0116] DirAC 인코더에 대한 입력은 B-포맷 표상의 FOA 사운드 신호이다. 이는 4개의 채널, 즉 전방향 사운드 압력, 및 특정 가정하에서 입자 속도에 비례하는 3개의 1차 공간 구배로 구성된다. 이러한 신호가 파라메트릭 방식으로 인코딩되고, 이에 대해서는 [18]을 참조한다. 매개변수는 복소수 사운드 압력(P(k,n))으로부터 유도되고, 이는 변환된 전방향 신호 및 변환된 구배 신호에 상응하는 복소수 입자 속도 벡터 $U(k,n) = [U_x(k,n), U_y(k,n), U_z(k,n)]^T$ 이다.

[0117] DirAC 표상은 각각의 시간-주파수 빈에서의 사운드 파동의 신호(P(k,n)), 확산($\psi(k,n)$) 및 방향r(k,n)으로 구성된다. 후자를 유도하기 위해서, 첫 번째로, 활성 사운드 세기 벡터(I_a(k,n))가, 압력 벡터와 속도 벡터[18]의 ((·)^{*}에 의해서 표시된) 복소 켈레의 곱의 (Re(·)에 의해서 표시된) 실수 부분으로서 연산된다:

$$I_a(k,n) = \frac{1}{2} Re(P(k,n)U^*(k,n)) \quad (1)$$

[0118] 확산은 이러한 벡터[18]의 변이 계수로부터 추정된다.

$$\psi(k,n) = \sqrt{1 - \frac{\|E\{I_a(k,n)\}\|}{E\{\|I_a(k,n)\|\}}} \quad (2)$$

[0121] 여기에서 E는, 이동 평균으로서 구현되는, 시간 프레임을 따른 예상 연산자를 나타낸다.

[0122] 기준 위치에 대한 각각의 (매칭된) 음원의 거리를 가지는 방향-기반 거리 맵을 이용하여 사운드를 조작하도록 의도되기 때문에, 선택적인 실시예에서 방향 추정의 변경이 작아야 한다. 프레임이 전형적으로 짧기 때문에, 항상 그러한 것은 아니다. 그에 따라, 이동 평균이 적용되어 평탄화된 방향 추정($\bar{I}_a(k,n)$)을 획득한다. 이어서, 신호의 다이렉트 부분의 DoA가, 실시예에서, 반대 방향의 단위 길이 벡터로서 연산된다:

$$\mathbf{r}_r(k,n) = - \frac{\bar{I}_a(k,n)}{\|\bar{I}_a(k,n)\|} \quad (3)$$

[0124] 방향이 각각의 시간-주파수 빈에 대한 단위 길이의 3차원적 벡터로서 인코딩되기 때문에, 거리 정보를 쉽게 통합할 수 있다. 방향 벡터에 그 상응 맵 엔트리(map entry)를 곱하고, 그에 따라 벡터 길이는 상응 음원의 거리(d_r(k,n))를 나타낸다:

$$\begin{aligned} \mathbf{d}_r(k,n) &= \mathbf{r}_r(k,n)d_r(k,n) \\ &= \mathbf{r}_r(k,n)m(\mathbf{r}_r(k,n)) \end{aligned} \quad (4)$$

[0126] 여기에서, d_r(k,n)은 마이크로폰의 레코딩 위치로부터 시간(n) 및 주파수 빈(k)에서 활성적인 음원으로 향하는 벡터이다.

[0127] 청취자 위치는 현재 프로세싱 프레임에 대한 추적 시스템에 의해서 $l(n)$ 으로서 주어진다. 음원 위치의 벡터 표상으로, 길이가 $d_1(k, n) = \|\mathbf{d}_1(k, n)\|$ 인 새로운, 병진 운동된 방향 벡터($\mathbf{d}_1(k, n)$)를 얻기 위해서, 추적 위치 벡터($l(n)$)를 뺄 수 있다. 청취자의 PoV로부터 음원까지의 거리가 유도되고, DoA가 단일 단계에서 적용된다:

$$[0128] \quad \mathbf{d}_1(k, n) = \mathbf{d}_r(k, n) - l(n) \quad (5)$$

[0129] 현실적 재생의 중요 양태는 거리 감쇠이다. 그러한 감쇠는 음원 및 청취자 사이의 거리의 함수인 것으로 가정된다[19]. 거리 벡터의 길이는 재생을 위한 감쇠 또는 증폭을 인코딩하기 위한 것이다. 레코딩 위치에 대한 거리는 거리 탭에 따라 $d_r(k, n)$ 로 인코딩되고, 재생되는 거리는 $d_1(k, n)$ 로 인코딩된다. 벡터를 단위 길이로 정규화하고 이어서 이전의 그리고 새로운 거리의 비율을 곱하는 경우에, $d_1(k, n)$ 를 원래 벡터의 길이로 나누는 것에 의해서 필요 길이가 주어진다는 것을 알 수 있다:

$$[0130] \quad \mathbf{d}_v(k, n) = \frac{\mathbf{d}_1(k, n)}{\|\mathbf{d}_1(k, n)\|} \frac{\|\mathbf{d}_1(k, n)\|}{\|\mathbf{d}_r(k, n)\|} = \frac{\mathbf{d}_1(k, n)}{\|\mathbf{d}_r(k, n)\|} \quad (6)$$

[0131] 청취자의 배향에 대한 변화가 이하의 단계에서 적용된다. 추적에 의해서 주어진 배향은, 원점으로서의 레코딩 위치에 대한 피치, 요 및 롤로 구성된 벡터($\mathbf{o}(n) = [o_x(n), o_z(n), o_y(n)]^T$)로 작성될 수 있다. 음원 방향이 청취자 배향에 따라 회전되고, 이는 2D 회전 매트릭스를 이용하여 구현된다:

$$[0132] \quad \mathbf{d}_p(k, n) = \mathbf{R}_y(o_y(n))\mathbf{R}_z(o_z(n))\mathbf{R}_x(o_x(n))\mathbf{d}_v(k, n) \quad (7)$$

[0133] 이어서, 청취자에 대한 결과적인 DoA는 단위 길이로 정규화된 벡터에 의해서 주어진다:

$$[0134] \quad \mathbf{r}_p(k, n) = \frac{\mathbf{d}_p(k, n)}{\|\mathbf{d}_p(k, n)\|} \quad (8)$$

[0135] 변환된 방향 벡터, 확산, 및 복소수 스펙트럼을 이용하여, 균일하게 분포된 8+4개의 가상 라우드스피커 설정을 위한 신호를 합성한다. 8개의 가상 스피커가 청취자 평면(고도 0°)에서 45° 방위 단계로 위치되고, 4개는 45° 고도 초과에서 90° 의 십자가 형성으로 위치된다. 합성은 각각의 라우드스피커 채널 $1 \leq i \leq I$ 에 대해서 다이렉트 및 확산 부분으로 분할되고, 여기에서 $I = 12$ 는 라우드스피커의 수이다[16]:

$$[0136] \quad Y_i(k, n) = Y_{i,S}(k, n) + Y_{i,D}(k, n) \quad (9)$$

[0137] 다이렉트 부분을 위해서, 엣지 페이딩 진폭 페닝(edge fading amplitude panning)(EFAP)을 적용하여, 주어진 가상 라우드스피커 기하형태에 대한 직각 방향으로부터 사운드를 재생한다[20]. DoA 벡터($\mathbf{r}_p(k, n)$)가 주어지면, 이는 각각의 가상 라우드스피커 채널(i)에 대한 페닝 이득($G_i(r)$)을 제공한다. 각각의 DoA에 대한 거리-의존적 이득이 방향 벡터($\mathbf{d}_p(k, n)$)의 결과적인 길이로부터 유도된다. 채널(i)을 위한 다이렉트 합성이 이하가 되며:

$$[0138] \quad Y_{i,S}(k, n) = \sqrt{1 - \psi(k, n)}P(k, n) G_i(\mathbf{r}_p(k, n))(\|\mathbf{d}_p(k, n)\|)^{-\gamma} \quad (10)$$

[0139] 여기에서, 지수(γ)는, 전형적으로 약 1로 설정되는 튜닝 인자(tuning factor)이다[19]. $\gamma = 0$ 에서, 거리-의존적 이득이 턴 오프된다는 것을 주목하여야 한다.

[0140] 압력($P(k, n)$)은 I 개의 비상관화 신호($\tilde{P}_i(k, n)$)를 생성하기 위해서 이용된다. 이러한 비상관화된 신호는 확산 성분으로서 개별적인 라우드스피커 채널에 추가된다. 이는 표준 방법을 따른다[16]:

$$Y_{i,D}(k,n) = \sqrt{\psi(k,n)} \frac{1}{\sqrt{I}} \tilde{P}_i(k,n) \quad (11)$$

- [0141]
- [0142] 각각의 채널의 확산 및 다이렉트 부분이 함께 더해지고, 신호가 역 STFT에 의해서 시간 도메인으로 다시 변환된다. 이러한 채널 시간 도메인 신호는, 바이노럴화된 신호를 생성하기 위해서 라우드스피커 위치에 따라 좌측 또는 우측 귀를 위해서 HRTF로 컨볼브된다(convolved).
- [0143] 도 6a는 공간 정보, 그리고 제1 및 제2 음장 묘사 및, 예를 들어, 도 4c 또는 도 5의 벡터(1)에 대해서 설명한 바와 같이, 다른 기준 위치에 대한 기준 위치의 병진 운동을 나타내는 병진 운동 정보를 이용하여 수정된 음장을 계산하기 위한 추가적인 바람직한 실시예를 도시한다.
- [0144] 도 6a는, 사운드 분리 또는, 일반적으로, 사운드 분석 과정을 도 4c의 기준 위치 A와 관련된 제1 음장 묘사 및 도 4c의 기준 위치 B와 관련된 제2 음장 묘사에 적용하는 것을 나타내는 블록(700)을 도시한다.
- [0145] 이러한 과정은 하나 이상의 추출된 객체의 제1 그룹 및, 부가적으로, 하나 이상의 추출된 객체의 제2 그룹을 초래할 것이다.
- [0146] 이러한 그룹은 모든 분리된 음원에 대한, 즉 추출된 음원의 제1 그룹 및 하나 이상의 추출된 음원의 제2 그룹에 대한 도달 방향 정보를 계산하기 위해서 블록(702)에서 이용된다.
- [0147] 다른 실시예에서, 단계(700 및 702)가 단일 과정 내에 구현되어, 한편으로 음원을 위한 신호 및 다른 한편으로 음원을 위한 DoA 정보를 제공한다. 이는 또한 시간/주파수-선택적 과정과 같은 파라메트릭 과정에서 DirAC로서 마찬가지로 적용되고, 여기에서 음원 신호는 시간/주파수 빈에서의 B-포맷 표상의 신호 또는 시간/주파수 빈의 압력 신호 또는 전방향 신호 및 이러한 특정 빈에 대한 DoA 매개변수로서의 DoA 정보이다.
- [0148] 이어서, 단계(704)에서, 음원 매칭이 제1 그룹의 음원과 제2 그룹의 음원 사이에서 실시되고, 음원 매칭의 결과는 매칭된 음원이 된다.
- [0149] 이러한 매칭된 음원은, 블록(710)에 도시된 바와 같이, 새로운 DoA 및 새로운 거리를 이용하여 각각의 매칭된 객체에 대한 음장을 연산하기 위해서 이용된다. 또한, 예를 들어, 삼각측량 동작을 이용하여 매칭된 객체의 위치 또는, 대안적으로 또는 부가적으로, 매칭된 객체의 거리를 계산하기 위해서, 매칭된 객체의 도달 방향 정보 즉, 음원 A에 대한 도 4c의 α 및 β 와 같은 각각의 객체 당 2개가 블록(706)에서 이용된다.
- [0150] 블록(706)의 결과는, 예를 들어 도 4c에 도시된, 각각의 매칭된 객체의 위치 또는, 대안적으로 또는 부가적으로, 제1 또는 제2 기준 위치 A, B 중 하나에 대한 매칭된 객체의 거리이다.
- [0151] 또한, 매칭된 객체에 대한 새로운 도달 방향 정보 및 새로운 거리를 계산하기 위해서, 블록(708)에서 병진 운동 정보를 이용하는 것뿐만 아니라 회전 정보를 이용하는 것이 바람직하다.
- [0152] 비록 매칭된 객체의 위치가 단계(708) 내로 입력되는 것으로 설명되었지만, 매칭된 객체에 대한 새로운 도달 방향 정보만을 계산하기 위해서, 매칭된 객체에 대한 실제 위치 또는, 다시 말해서, 매칭된 객체의 거리가, 청취자가 이동한 새로운(상이한) 기준 위치에 대한 새로운 도달 방향을 계산하는데 필요치 않다는 것을 강조하며, 그러한 거리는 필수적이지 않다.
- [0153] 그러나, 그러한 거리는 음원 신호를 새로운 상황에 적응시키는데 있어서 필요하다. 따라서, 새로운 기준 위치에 대한 음원 또는 사운드 객체의 거리가 더 짧아질 때, 1 미만의 스케일링 인자가 계산될 것이다. 그러나, 거리가 더 길어질 때, 예를 들어, 도 6b와 관련하여 설명되는 바와 같이, 1 초과가 되도록 스케일링 인자가 계산된다. 따라서, 비록 도 6a가 실시예에 대해서 설명되었지만, 매칭된 객체의 명백한 위치 그리고 이어서 매칭된 객체의 거리가 계산되는 것, 그리고 이어서, 음장이 새로운 도달 방향 및 새로운 거리를 이용하여 각각의 매칭된 객체에 대해서 계산되는 것이 필수적인 것은 아니다. 그 대신, 2개의 기준 위치 중 하나의 기준 위치에 대한 매칭된 객체의 거리만으로 일반적으로 충분하고, 그리고 이어서, 각각의 매칭된 객체에 대한 음장이 새로운 DoA 및 새로운 거리를 이용하여 계산된다.
- [0154] 또한, 블록(714)은 블록(702)에 의해서 얻어진 이전의 DoA 정보를 이용하여 비-매칭 객체에 대한 음장의 계산을 도시한다. 또한, 블록(710)에서 얻어진 매칭된 객체 및 블록(714)에서 얻어진 비-매칭 객체에 대한 음장이 블록(712)에서 조합되고, 그에 따라, 예를 들어, 1차 앰비소닉 묘사, 더 높은 차수의 앰비소닉 묘사 또는, 대안적으로, 블록(710) 및 블록(714)에 대해서 동일한 특정 라우드스피커 설정과 관련된 라우드스피커 채널 묘사와 같은 앰비소닉 묘사일 수 있는 수정된 음장 묘사를 획득하며, 그에 따라 단순한 채널별 부가가 블록(712)에서 실시될

수 있다.

- [0155] 도 6b는 음장 계산기(420)의 바람직한 구현예를 도시한다. 블록(1102)에서, 각각의 음원에 대한 음원 분리 및 도달 방향 또는 일반적인 방향 정보 계산이 실시된다. 이어서, 블록(1104)에서, 도달 방향 벡터에 거리 정보 벡터, 즉 원래의 기준 위치로부터 음원으로서의 벡터, 즉 예를 들어 도 5의 항목(520)으로부터 항목(510)으로서의 벡터를 곱한다. 이어서, 블록(1106)에서, 병진 운동 정보, 즉 도 5의 항목(520)으로부터 항목(500)으로서의 벡터를 고려하여, 청취자 위치(500)로부터 음원 위치(510)까지의 벡터인 새로운 병진 운동된 방향 벡터를 계산한다. 이어서, d_r 로 표시된 정확한 길이를 갖는 새로운 도달 방향 벡터가 블록(1108)에서 계산된다. 이러한 벡터는 d_r 과 동일한 방향으로 지향되나, 상이한 길이를 가지는데, 이는 이러한 벡터의 길이가, 음원(510)이 특정 볼륨을 갖는 원래의 음장 내에 레코딩되고 그에 따라 d_r 의 길이는 라우드니스 변화를 다소간 나타낸다는 사실을 반영하기 때문이다. 이는, 벡터(d_r)를 레코딩 거리(d_r), 즉 마이크로폰(520)으로부터 음원(510)까지의 벡터(d_r)의 길이로 나누는 것에 의해서 얻어진다. 설명한 바와 같이, 마이크로폰(520)으로부터 음원(510)까지의 벡터(d_r)의 길이는 삼각측량 계산에 의해서 유도될 수 있다. 마이크로폰이 제1 음장 묘사의 기준 위치에 있을 때, 제1 음장 묘사의 기준 위치로부터 음원까지의 거리가 사용된다. 그러나, 마이크로폰이 제2 음장 묘사의 추가적인 기준 위치에 있을 때, 제2 음장 묘사의 추가적인 기준 위치로부터 음원까지의 거리가 사용된다.
- [0156] 도 5에서와 같이, 재생된 거리가 레코딩된 거리보다 멀 때, d_r 의 길이가 1(unity)보다 작을 것이다. 이는 새로운 청취자 위치에서의 재생을 위한 음원(510)의 감쇠를 초래할 것이다. 그러나, 재생된 거리(d_r)가 레코딩된 거리보다 짧을 때, 블록(1108)에 의해서 계산되는 바와 같은 d_r 의 길이는 1보다 클 것이고, 상응 스케일링 인자는 음원의 증폭을 초래할 것이다.
- [0157] 도 6a에서, 항목(710)은, 각각의 매칭된 객체에 대한 음장이 새로운 도달 방향 정보 및 새로운 거리를 이용하여 계산된다는 것을 나타낸다. 그러나, 기본적으로, 각각의 매칭된 객체에 대한 음장을 연산하기 위해서, 하나 이상의 추출된 음원의 제1 그룹 또는 하나 이상의 추출된 음원의 제2 그룹으로부터 얻어진 객체 신호가 일반적으로 이용될 수 있다. 그러나, 실시예에서, 블록(710)에서의 음장 연산을 실시하기 위해서 어떠한 음장 묘사를 이용할지를 결정하기 위해서, 도 7에 도시된 특정 선택이 실시된다. 블록(720)에서, 제1 음장 묘사의 제1 기준 위치에 대한 새로운 청취자 위치의 제1 거리가 결정된다. 도 4c와 관련하여, 이는 상이한 기준 위치와 기준 위치 A 사이의 거리이다.
- [0158] 또한, 단계(722)에서, 제2 음장 묘사의 제2 기준 위치에 대한 새로운 청취자 위치의 제2 거리가 결정된다. 도 4c의 이러한 실시예에서, 이는 (병진 운동으로 인한) 상이한 기준 위치와 기준 위치 B 사이의 거리일 수 있다.
- [0159] 상이한 기준 위치로부터 기준 위치 B까지의 거리가 상이한 기준 위치로부터 기준 위치 A까지의 차이보다 짧은 것으로 보인다. 따라서, 이러한 것이 단계(724)에서 결정될 수 있다. 그리고, 단계(726)에서, 객체 신호는 더 짧은 거리를 갖는 음장 묘사로부터 유도된 그룹으로부터 선택된다. 따라서, 도 4c의 도면에서 매칭된 음원에 상응하는 음원 A 및 B를 제공하기 위해서, 추가적인 기준 위치 B와 관련된 제2 음장 묘사로부터 유도된 음원 신호가 이용될 수 있다.
- [0160] 그러나, 병진 운동이 원점으로부터 상이한 기준 위치로, 예를 들어 도 4c 도면의 좌측으로 향하는 다른 실시예에서, 이러한 다른 기준 위치로부터 기준 위치 A까지 짧은 거리가 될 것이고, 이어서, 제1 음장 묘사는 도 6b의 블록(710)의 각각의 매칭된 객체에 대한 음장을 최종적으로 연산하기 위해서 이용될 수 있다. 다시, 선택이 도 7에 도시된 과정에 의해서 실시될 수 있다.
- [0161] 도 9는 추가적인 바람직한 실시예를 도시한다. 단계(740)에서, 제1 음장 묘사에 관한 음장 분석, 예를 들어 도 6c의 블록(422)에 도시된 DirAC 분석 형태의 파라메트릭 음장 분석이 실시된다.
- [0162] 이는, 예를 들어, 각각의 시간/주파수 빈에 대해서, 매개변수의 제1 세트를 초래하고, 매개변수의 각각의 세트는 DoA 매개변수 및, 선택적으로, 확산 매개변수를 포함한다.
- [0163] 단계(741)에서, 음장 분석이 제2 음장 묘사에서 실시되고, 다시 DirAC 분석이 블록(740)에서와 같이 그리고, 예를 들어, 도 6c의 블록(422)과 관련하여 설명한 바와 같이 실시된다.
- [0164] 이는, 예를 들어, 시간/주파수 빈에 대한, 매개변수의 제2 세트를 초래한다.
- [0165] 이어서, 블록(746)에서, 각각의 매개변수 쌍에 대한 위치가, 제1 시간/주파수 빈으로부터의 상응 DoA 매개변수

및 매개변수의 제2 세트로부터의 동일 시간/주파수 빈으로부터의 DoA 매개변수를 이용하여 결정될 수 있다. 이는, 각각의 매개변수 쌍에 대한 위치를 초래할 것이다. 그러나, 그러한 위치는, 매개변수의 제1 세트 및/또는 매개변수의 제2 세트 내에 상응 시간/주파수 빈에 대한 확산이 적을수록, 더 유용할 것이다.

- [0166] 따라서, 매개변수의 제1 및 제2 세트 모두에서 매우 작은 확산을 초래하는 시간/주파수 빈으로부터의 위치를 추가적으로 이용하는 것만이 바람직하다.
- [0167] 또한, 블록(740) 및 블록(741)에 의해서 또한 출력되는 바와 같은 시간/주파수 빈 내의 상응 신호에 대한 상관화를 또한 실시하는 것이 바람직하다.
- [0168] 따라서, 도 6a의 단계(704)의 "음원 매칭"을, 예를 들어, 완전히 피할 수 있고 확산 매개변수를 기초로 하는 매칭된 음원/매칭된 시간/주파수 빈의 결정에 의해서 대체될 수 있거나, 매칭이 부가적으로, 예를 들어, B-포맷 성분으로부터의 또는 도 6c의 블록(422)에 의해서 출력된 압력 신호 또는 객체 신호로부터의 시간/주파수 빈 내의 상응 신호를 이용하여 실시될 수 있다.
- [0169] 어떠한 경우에도, 블록(46)은, 도 6a의 블록(704)에서 발견되는 "매칭된 객체"에 상응하는 특정의 (선택된) 시간/주파수 빈에 대한 특정 위치를 초래할 것이다.
- [0170] 이어서, 블록(748)에서, 블록(746)에 의해서 얻어진 위치에 대한 수정된 매개변수 및/또는 신호 및/또는, 예를 들어, 햇 추적기(hat tracker)에 의해서 얻어진 바와 같은 상응 병진 운동/회전이 계산되고, 블록(748)의 출력은 상이한 시간/주파수 빈들에 대한 수정된 매개변수 및/또는 수정된 신호를 나타낸다.
- [0171] 따라서, 블록(748)은 수정된 매개변수의 계산을 위한 블록(424)의 병진 운동 변환(423) 및 회전 변환에 상응할 수 있고, 수정된 신호의 계산은, 예를 들어, 바람직하게 또한 상응 시간/주파수 빈에 대한 위치로부터 유도된 특정 스케일링 인자의 고려 하에서, 도 6c의 블록(425)에 의해서 실시될 수 있다.
- [0172] 마지막으로, 음장 묘사의 합성은 수정된 데이터를 이용하여 블록(750)에서 실시된다. 이는, 예를 들어, 제1 또는 제2 음장 묘사를 이용하여 DirAC 합성에 의해서 이루어질 수 있거나 블록(425)에 도시된 바와 같은 앰비소닉 신호 생성기에 의해서 실시될 수 있고, 결과는 전송/저장/렌더링을 위한 새로운 음장 묘사가 될 것이다.
- [0173] 도 10은 음장 계산기(420)의 추가적인 바람직한 구현예를 도시한다. 도 10에 도시된 과정의 적어도 일부가 각각의 매칭된 음원에 대해서 별도로 실시된다. 블록(1120)은, 예를 들어 삼각측량 계산에 의해서, 매칭된 음원에 대한 거리를 결정한다.
- [0174] 음장 묘사를 기초로, 전체 대역 도달 방향 또는 대역별 도달 방향이 1100에서 결정된다. 이러한 도달 방향 정보는 음장의 도달 방향 데이터를 나타낸다. 이러한 도달 방향 데이터를 기초로, 병진 운동 변환이 블록(1110)에서 실시된다. 이를 위해서, 블록(1120)은 각각의 매칭된 음원에 대한 거리를 계산한다. 데이터를 기초로, 블록(1110)은, 이러한 구현예에서, 기준 위치로부터 상이한 기준 위치로의 병진 운동에만 의존하는 음장에 대한 새로운 도달 방향 데이터를 생성한다. 이를 위해서, 블록(1110)은, 예를 들어, 가상 현실 구현예의 맥락의 추적에 의해서 생성된 병진 운동 정보를 수신한다.
- [0175] 바람직하게 또는 대안적으로, 회전 데이터가 또한 이용된다. 이를 위해서, 블록(1130)은 회전 정보를 이용하여 회전 변환을 실시한다. 병진 운동 및 회전 모두가 실시될 때, 병진 운동으로부터의 정보 및 블록(1120)으로부터의 음원 거리를 이미 포함하는 음장의 새로운 DoA의 계산에 이어서 회전 변환을 실시하는 것이 바람직하다.
- [0176] 이어서, 블록(1140)에서, 새로운 음장 묘사가 생성된다. 이를 위해서, 원래의 음장 묘사가 이용될 수 있거나, 대안적으로, 음원 분리 알고리즘에 의해서 음장 묘사로부터 분리된 음원 신호가 이용되거나 임의의 다른 적용예가 이용될 수 있다. 기본적으로, 새로운 음장 묘사는, 예를 들어, 앰비소닉 생성기(430)에 의해서 얻어진 것과 같은 또는 DirAC 합성기(425)에 의해서 생성된 바와 같은 방향성 음장 묘사일 수 있거나, 후속 바이노럴 렌더링에서 가상 스피커 표상으로부터 생성된 바이노럴 표상일 수 있다.
- [0177] 바람직하게, 도 10에 도시된 바와 같이, 도달 방향별 거리가 또한, 특정 음원의 볼륨 또는 라우드니스를 새로운 위치 즉, 새로운 또는 상이한 기준 위치에 적응시키기 위해서, 새로운 음장 묘사를 생성하는데 이용될 수 있다.
- [0178] 비록 도 10이, 회전 변환이 병진 운동 변환 후에 실시되는 상황을 도시하지만, 그 순서가 다를 수 있다는 것을 주목하여야 한다. 특히, 회전 변환이 블록(1100)에서 생성된 것과 같은 음장의 DoA에 적용될 수 있고, 이어서, 기준 위치로부터 상이한 기준 위치로의 대상의 병진 운동으로 인한 부가적인 병진 운동 변환이 적용될 수 있다.
- [0179] 음장의 DoA가 블록(1100)에 의해서 결정되자마자, 거리 정보가 블록(1120)을 이용하여 메타데이터로부터 획득되

고, 이러한 거리 정보는 이어서, 변화된 거리를 그리고, 그에 따라, 특정 기준 위치에 대한 특정 음원의 변화된 라우드니스를 고려하기 위해서 블록(1140)에서 새로운 음장 묘사를 생성하는 것에 의해서 이용된다. 기본적으로, 거리가 멀어지는 경우에, 특정 음원 신호가 감쇠되는 반면, 거리가 짧아질 때 음원 신호가 증폭된다고 할 수 있다. 당연히, 거리에 따른 특정 음원의 감쇠 또는 증폭이 거리 변화에 비례하여 만들어지나, 다른 실시예에서, 덜 복잡한 동작이, 상당히 큰 증분으로, 음원 신호의 이러한 증폭 또는 감쇠에 적용될 수 있다. 그러한 덜 복잡한 구현에는, 임의의 거리 변화가 완전히 무시되는 상황에 비해서, 더 우수한 결과를 제공한다.

- [0180] 도 11은 음장 계산기의 추가적인 바람직한 구현예를 도시한다. 블록(1200)에서, 음장으로부터의 개별적인 음원이, 예를 들어, 대역별로 또는 전체 대역에서 결정될 수 있다. 프레임 및 대역별 결정이 실시될 때, 이는 DirAC 분석에 의해서 이루어질 수 있다. 전체 대역 또는 하위대역 결정이 실시되는 경우에, 이는, 임의의 종류의 전체 대역 또는 하위대역 음원 분리 알고리즘에 의해서 이루어질 수 있다.
- [0181] 블록(1210)에서, 청취자의 병진 운동 및/또는 회전이, 예를 들어, 머리 추적에 의해서 결정된다.
- [0182] 블록(1220)에서, 각각의 음원에 대한 이전의 거리가 메타데이터의 이용에 의해서 그리고, 예를 들어, 삼각측량 계산을 위한 메타데이터의 이용에 의해서 결정된다. 따라서, 각각의 대역이 (확산이 특정 문턱값보다 작은 경우에) 특정 음원인 것으로 간주되고, 이어서, 작은 확산 값을 갖는 각각의 시간/주파수 빈에 대한 특정 거리가 결정된다.
- [0183] 이어서, 블록(1230)에서, 예를 들어, 도 6b의 맥락으로 설명된, 예를 들어, 대역마다의 벡터 계산에 의해서, 음원마다 새로운 거리가 얻어진다.
- [0184] 또한, 블록(1240)에서 도시된 바와 같이, 음원마다의 이전의 방향이, 예를 들어, DirAC 분석에서 얻어진 DoA 계산에 의해서 또는, 예를 들어, 음원 분리 알고리즘에서의 도달 방향 또는 방향 정보 분석에 의해서 결정된다.
- [0185] 이어서, 블록(1250)에서, 음원마다의 새로운 방향이, 예를 들어, 대역 마다 또는 전체 대역에서 벡터 계산을 실시하는 것에 의해서 결정된다.
- [0186] 이어서, 블록(1260)에서, 새로운 음장이 병진 운동되고 회전된 청취자에 대해서 생성된다. 이는, 예를 들어, DirAC 합성에서 채널마다 직접 부분을 스케일링하는 것에 의해서 이루어질 수 있다. 구체적인 구현예에 따라, 거리 수정은, 블록(1260)에서의 거리 수정 실시에 더하여 또는 대안적으로, 블록(1270a, 1270b 또는 1270c)에서 이루어질 수 있다.
- [0187] 예를 들어, 음장만이 단일 음원을 가지는 것으로 결정될 때, 거리 수정이 블록(1270a)에서 미리 실시될 수 있다.
- [0188] 대안적으로, 개별적인 음원 신호가 블록(1200)에 의해서 계산될 때, 실제 새로운 음장이 블록(1260)에서 생성되기 전에, 거리 수정이 블록(1270b)에서 개별적인 음원에 대해서 실시될 수 있다.
- [0189] 또한, 블록(1260)에서의 음장 생성이, 예를 들어, 라우드니스피커 설정 신호 또는 바이노럴 신호를 제공하지 않으나, 예를 들어 앰비소닉 인코더 또는 계산기(430)를 이용하여 다른 음장 묘사가 제공될 때, 거리 수정이 또한 블록(1260)에서의 생성에 이어서 실시될 수 있고, 이는 블록(1270c)에서의 실시를 의미한다. 구현예에 따라, 거리 수정이 또한 몇개의 수정기에 분배될 수 있고, 그에 따라, 마지막으로, 특정 음원이, 음원과 기준 위치 사이의 원래의 거리와, 음원과 상이한 기준 위치 사이의 새로운 거리 사이의 차이에 의해서 지향되는 특정 라우드니스로 있게 된다.
- [0190] 도 12a는, 예를 들어, 앞서 인용한 문헌["Directional Audio Coding" from IWPASH of 2009]에서 최초로 개시된 DirAC 분석기를 도시한다.
- [0191] DirAC 분석기는 대역 필터(1310)의 बैं크, 에너지 분석기(1320), 세기 분석기(1330), 일시적 평균화 블록(1340) 및 확산 계산기(1350) 그리고 방향 계산기(1360)를 포함한다.
- [0192] DirAC에서, 분석 및 합성 모두가 주파수 도메인에서 실시된다. 사운드를, 각각의 구분된 특성 내의 주파수 대역들로 분할하기 위한 몇 가지 방법이 있다. 가장 일반적으로 사용되는 주파수 변환은 국소 푸리에 변환(SIFT), 및 구적법 미러 필터 बैं크(Quadrature mirror filter bank)(QMF)를 포함한다. 이에 더하여, 임의의 구체적인 목적에 맞춰 최적화되는 임의의 필터를 갖는 필터 बैं크를 완전히 자유롭게 설계할 수 있다. 방향 분석의 목적은, 사운드가 하나의 또는 다수의 방향으로부터 동시에 도달하였는지를 평가하는 것과 함께, 각각의 주파수 대역에서 사운드의 도달 방향을 추정하는 것이다. 원칙적으로, 이는 많은 수의 기술로 실시될 수 있으나, 도

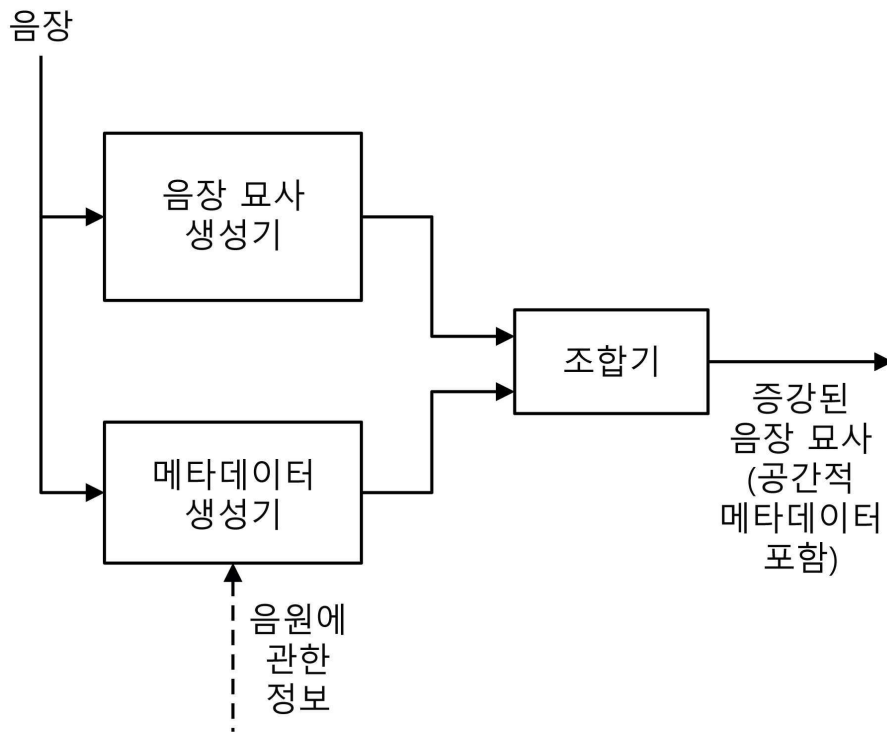
12a에 도시된, 음장의 에너지 분석이 적합한 것으로 확인되었다. 1차원, 2차원 또는 3차원의 압력 신호 및 속도 신호가 단일 위치로부터 캡처될 때, 에너지 분석이 실시될 수 있다. 1차 B-포맷 신호에서, 전방향 신호가 W-신호로 지칭되고, 이는 2의 평방 제곱근에 의해서 스케일 다운된 것이다. 사운드 압력은, STFT 도메인으로 표현된, $P = \sqrt{2} * W$ 로서 추정될 수 있다.

- [0193] X-, Y- 및 Z 채널은 데카르트 축을 따라 지향된 쌍극의 방향 패턴을 가지며, 이들은 벡터($U = [X, Y, Z]$)를 함께 형성한다. 벡터는 음장 속도 벡터를 추정하고, 또한 STFT 도메인으로 표현된다. 음장의 에너지(E)가 연산된다. B-포맷 신호의 캡처는, 방향성 마이크로폰들의 일치되는 배치로, 또는 전방향 마이크로폰들의 밀집-이격된 세트에 얻어질 수 있다. 일부 적용예에서, 마이크로폰 신호는 연산 도메인으로 형성될 수 있고, 즉 시플레이트 될 수 있다.
- [0194] 사운드의 방향은 세기 벡터(I)의 반대 방향으로 규정된다. 그러한 방향은 전달된 메타데이터 내의 각도 방위 및 고도 값에 상응하는 것으로 규정된다. 음장의 확산은 또한 세기 벡터 및 에너지의 예상 연산자를 이용하여 연산된다. 이러한 식은, 사운드 에너지가 단일 방향(확산이 0이다)으로부터 또는 모든 방향(확산이 1이다)으로부터 도달하는 경우를 특징으로 하는, 0과 1 사이의 실제-값의 수이다. 이러한 과정은, 전체 3D 또는 그 미만의 차수의 속도 정보가 이용 가능한 경우에 적합하다.
- [0195] 도 12b는, 다시 한번 대역 필터(1370), 가상 마이크로폰 블록(1400), 다이렉트/확산 합성 블록(1450), 및 특정 라우드스피커 설정 또는 가상의 의도된 라우드스피커 설정(1460)을 가지는, DirAC 합성을 도시한다. 또한, 확산-이득 변환기(1380), 벡터 기반의 증폭 패닝(VBAP) 이득 테이블 블록(1390), 마이크로폰 보상 블록(1420), 라우드스피커 이득 평균화 블록(1430) 및 다른 채널을 위한 분배기(1440)가 사용된다.
- [0196] 이러한 DirAC와 라우드스피커의 합성에서, 도 12b에 도시된 DirAC 합성의 고품질 버전이 모든 B-포맷 신호를 수신하고, 이를 위해서 가상 마이크로폰 신호가 라우드스피커 설정(1460)의 각각의 라우드스피커 방향에 대해서 연산된다. 이용되는 방향 패턴은 전형적으로 쌍극이다. 이어서, 메타데이터에 따라, 가상 마이크로폰 신호가 비-선형 방식으로 수정된다. DirAC의 낮은 비트레이트(bitrate) 버전이 도 12b에 도시되어 있지 않지만, 이러한 상황에서, 오디오의 단지 하나의 채널이 도 6에 도시된 바와 같이 전송된다. 프로세싱의 차이는, 모든 가상 마이크로폰 신호가 수신된 오디오의 단일 채널에 의해서 대체될 것이라는 것이다. 가상 마이크로폰 신호는 2개의 스트림: 분리되어 프로세스되는 확산 스트림 및 비-확산 스트림으로 분할된다.
- [0197] 비-확산 사운드는 벡터 기반 진폭 패닝(VBAP)을 이용함으로써 점 음원으로서 재생된다. 패닝에서, 마이크로폰 사운드 신호가 라우드스피커-특정 이득 인자와 곱해진 이후에 라우드스피커의 하위세트에 인가된다. 이득 인자는 라우드스피커 설정에 관한 정보 및 특정된 패닝 방향을 이용하여 연산된다. 저-비트-레이트 버전에서, 입력 신호는 메타데이터가 암시하는 방향으로 단순히 패닝된다. 고-품질 버전에서, 각각의 가상 마이크로폰 신호에 상응 이득 인자가 곱해지고, 이는 패닝과 동일한 효과를 생성하나, 임의의 비-선형 잡음을 생성할 가능성이 적다.
- [0198] 많은 경우에, 방향 메타데이터는 급격한 일시적 변화의 영향을 받는다. 잡음을 피하기 위해서, VBAP로 연산된 라우드스피커를 위한 이득 인자를, 각각의 대역에서 약 50 사이클 기간과 동일한 주파수-의존 시간 상수로 시간 적분하는 것에 의해서 평활화된다. 이는 잡음을 효과적으로 제거하나, 방향의 변화는, 대부분의 경우에 평균화가 없는 것보다 더 느린 것으로 인지되지 않는다.
- [0199] 확산 사운드의 합성의 목적은, 청취자를 둘러싸는 사운드의 인지를 생성하는 것이다. 저-비트-레이트 버전에서, 확산 스트림은 입력 신호를 비상관화하는 것 및 모든 라우드스피커로부터 이를 재생하는 것에 의해서 재생된다. 고품질 버전에서, 확산 스트림의 가상 마이크로폰 신호가 이미 어느 정도 비간섭적이고, 이들은 단지 약하게 비상관화될 필요가 있다. 이러한 접근방식은, 저-비트-레이트 버전보다 주위 반향 및 주변 사운드에 대해서 보다 양호한 공간적 품질을 제공한다.
- [0200] 헤드폰과의 DirAC 합성을 위해서, DirAC은 비-확산 스트림을 위한 청취자 주위의 특정 양의 가상 라우드스피커 및 확산 스트림을 위한 특정 수의 라우드스피커로 포물레이트된다. 가상 라우드스피커는 측정된 머리-관련 전달 함수(HRTF)를 이용한 입력 신호의 콘볼루션으로서 구현된다.
- [0201] 비록 장치의 맥락에서 몇몇 양태를 설명하였지만, 이러한 양태가 상응 방법의 설명을 또한 나타낸다는 것이 명확하고, 여기에서 블록 또는 디바이스는 방법 단계 또는 방법 단계의 특징에 상응한다. 마찬가지로, 방법 단계의 맥락으로 설명된 양태가 또한 상응 장치의 상응 블록 또는 항목 또는 특징에 관한 설명을 나타낸다.

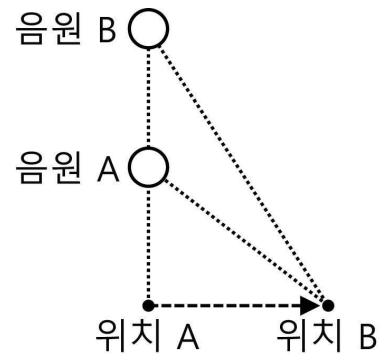
- [0202] 본 발명의 증강된 음장 묘사는 디지털 저장 매체 또는 비-일시적 저장 매체에 저장될 수 있거나, 무선 송신 매체 또는 유선 송신 매체, 예를 들어 인터넷과 같은 송신 매체에서 전송될 수 있다.
- [0203] 특정 구현예의 요건에 따라, 본 발명의 실시예가 하드웨어 또는 소프트웨어로 구현될 수 있다. 그러한 구현은, 각각의 방법이 실시되도록 프로그래밍 가능 컴퓨터 시스템과 협력하는(또는 협력할 수 있는), 전자적 관독 가능 제어 신호가 저장된, 디지털 저장 매체, 예를 들어 플로피 디스크, DVD, CD, ROM, PROM, EPROM, EEPROM 또는 FLASH 메모리를 이용하여 실시될 수 있다.
- [0204] 본 발명에 따른 일부 실시예는, 본원에서 설명된 방법 중 하나가 실시되도록, 프로그래밍 가능 컴퓨터 시스템과 협력할 수 있는, 전자 관독 가능 제어 신호를 갖는 비-일시적 데이터 캐리어를 포함한다.
- [0205] 일반적으로, 본 발명의 실시예는 프로그램 코드를 갖는 컴퓨터 프로그램 제품으로서 구현될 수 있고, 프로그램 코드는, 컴퓨터 프로그램 제품이 컴퓨터 상에서 작동될 때, 방법 중 하나를 실시하도록 동작된다. 프로그램 코드는 예를 들어 기계 관독 가능 캐리어에 저장될 수 있다.
- [0206] 다른 실시예는, 기계 관독 가능 캐리어에 저장된, 본원에서 설명된 방법 중 하나를 실시하기 위한 컴퓨터 프로그램을 포함한다.
- [0207] 다시 말해서, 본 발명의 방법의 실시예는, 그에 따라, 컴퓨터 프로그램이 컴퓨터에서 작동될 때, 본원에서 설명된 방법 중 하나를 실시하기 위한 프로그램 코드를 가지는 컴퓨터 프로그램이다.
- [0208] 그에 따라, 본 발명의 방법의 추가적인 실시예는, 본원에서 설명된 방법 중 하나를 실시하기 위한 컴퓨터 프로그램이 레코딩된 데이터 캐리어(또는 디지털 저장 매체, 또는 컴퓨터-관독 가능 매체)이다.
- [0209] 그에 따라, 본 발명의 방법의 추가적인 실시예는 본원에서 설명된 방법 중 하나를 실시하기 위한 컴퓨터 프로그램을 나타내는 신호의 시퀀스 또는 데이터 스트림이다. 데이터 스트림 또는 신호의 시퀀스는, 예를 들어, 데이터 통신 연결을 통해서, 예를 들어 인터넷을 통해서 전달되도록 구성될 수 있다.
- [0210] 추가적인 실시예는, 본원에서 설명된 방법 중 하나를 실시하도록 구성된 또는 적응된 프로세싱 수단, 예를 들어 컴퓨터 또는 프로그래밍 가능 로직 디바이스를 포함한다.
- [0211] 추가적인 실시예는 본원에서 설명된 방법 중 하나를 실시하기 위한 컴퓨터 프로그램이 설치된 컴퓨터를 포함한다.
- [0212] 일부 실시예에서, 프로그래밍 가능 로직 디바이스(예를 들어, 필드 프로그래밍 가능 게이트 어레이)를 이용하여 본원에서 설명된 방법의 일부 또는 모든 기능을 실시할 수 있다. 일부 실시예에서, 필드 프로그래밍 가능 게이트 어레이가 마이크로프로세서와 협력하여 본원에서 설명된 방법 중 하나를 실시할 수 있다. 일반적으로, 방법은 바람직하게 임의의 하드웨어 장치에 의해서 실시된다.
- [0213] 진술한 실시예는 본 발명의 원리에 대한 단순한 예시이다. 본원에서 설명된 배열 및 상세 내용에 대한 수정 및 변화가 다른 당업자에게 자명하다는 것을 이해하여야 한다. 그에 따라, 본 발명은 첨부된 특허 청구범위에 의해서만 제한될 것이고, 본원의 실시예에 관한 설명 및 기술 내용에 의해서 제공된 구체적인 상세 내용에 의해서 제한되지 않을 것이다.

도면

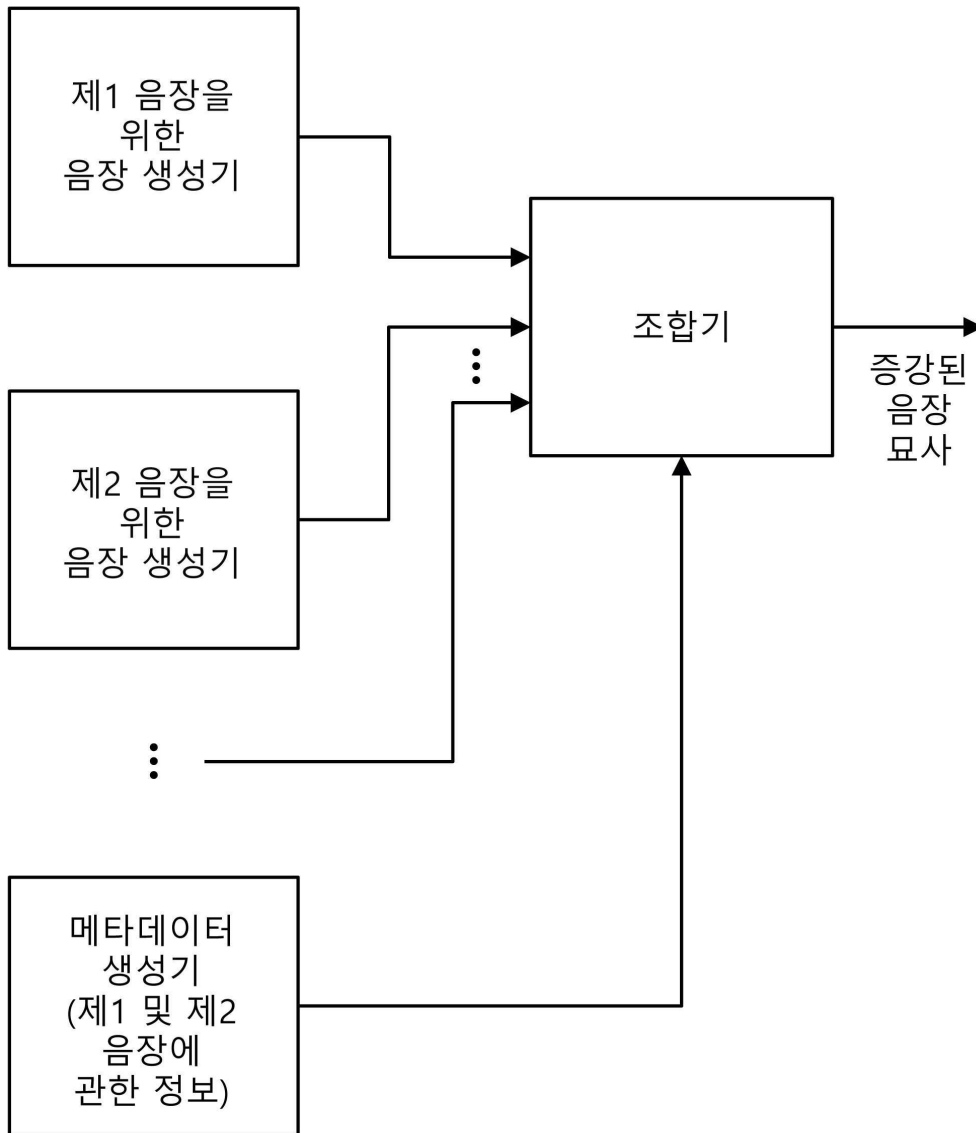
도면1a



도면1b



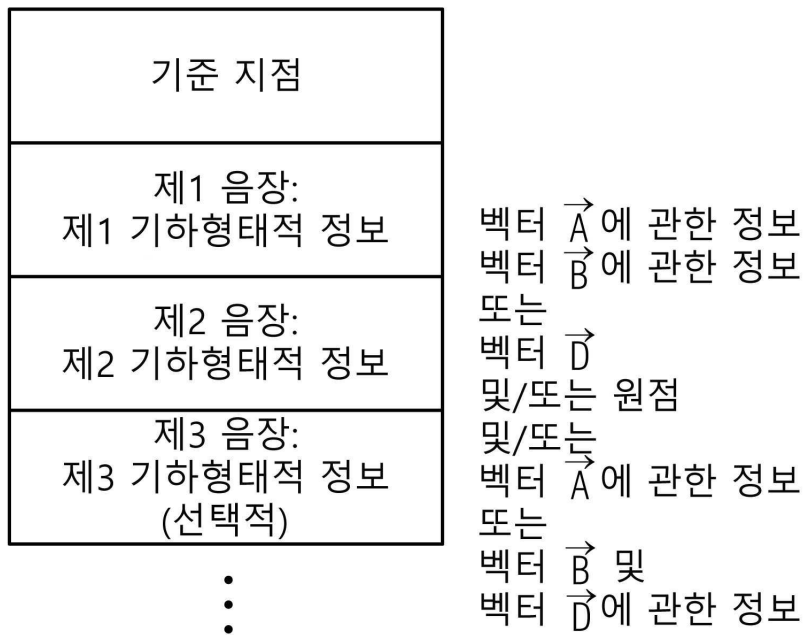
도면2



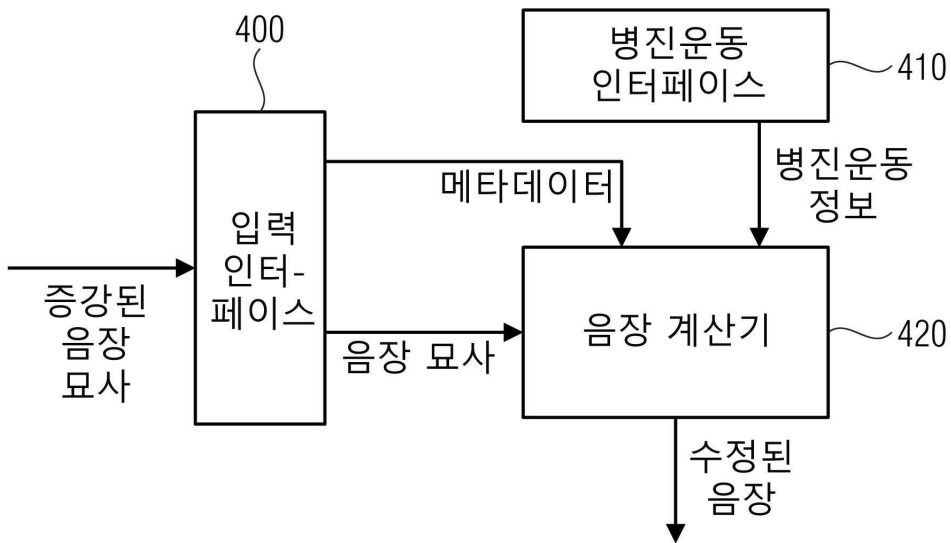
도면3a

제1 음장 묘사	제2 음장 묘사	제1 및 제2 음장에 관한 정보
----------	----------	----------------------

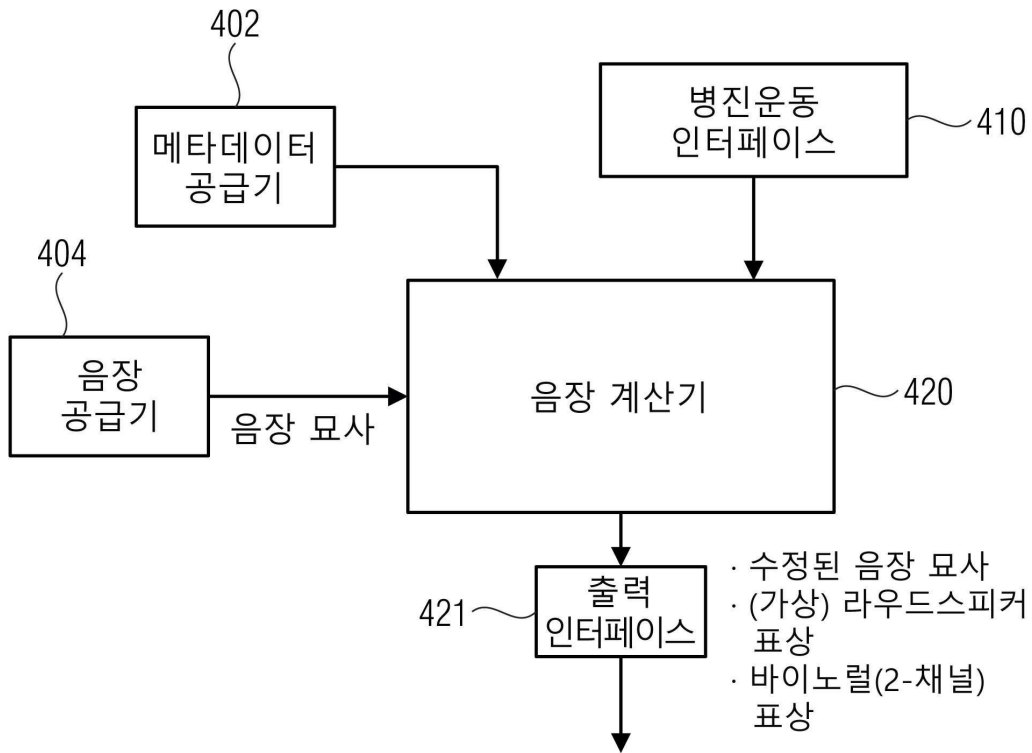
도면3b



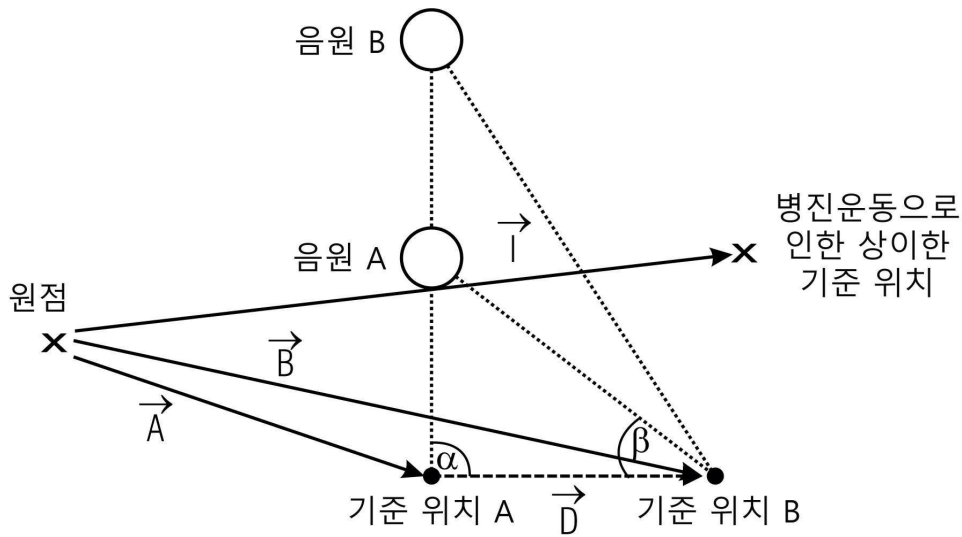
도면4a



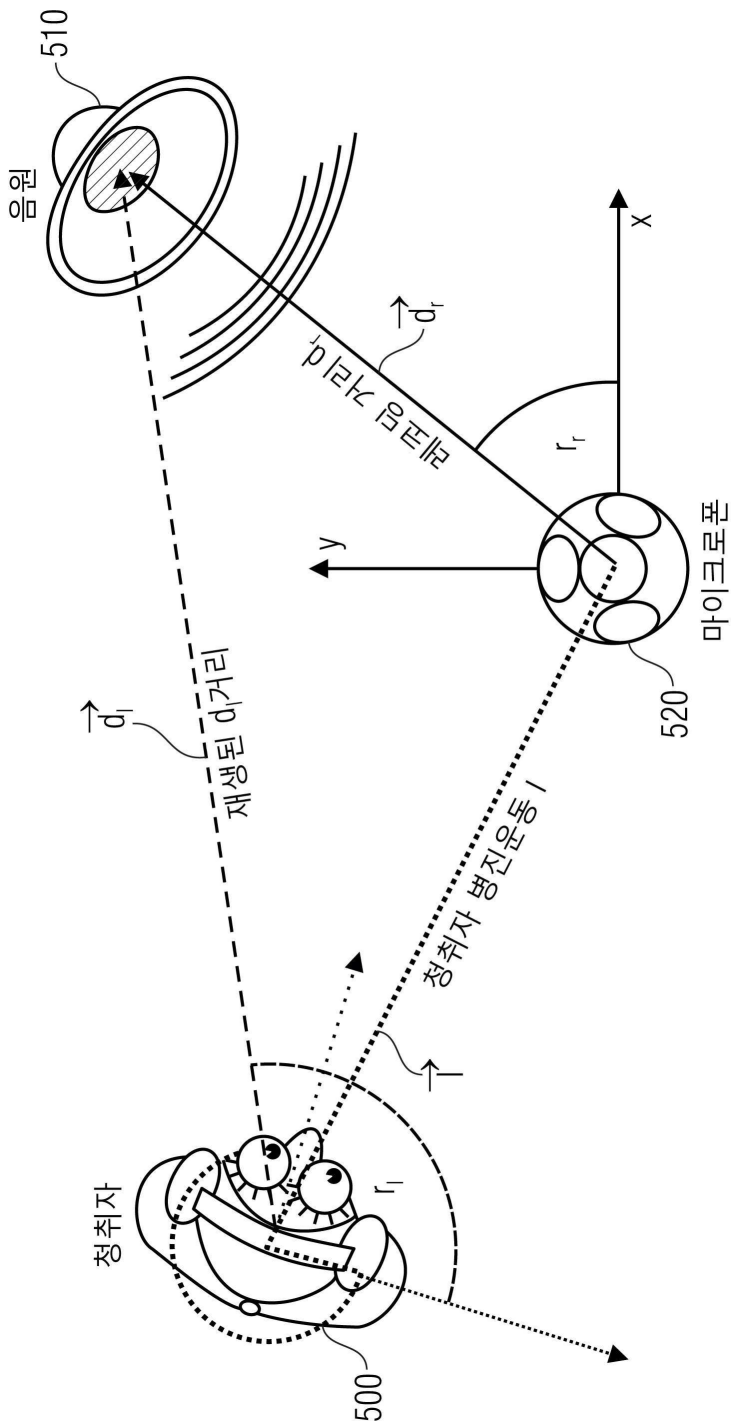
도면4b



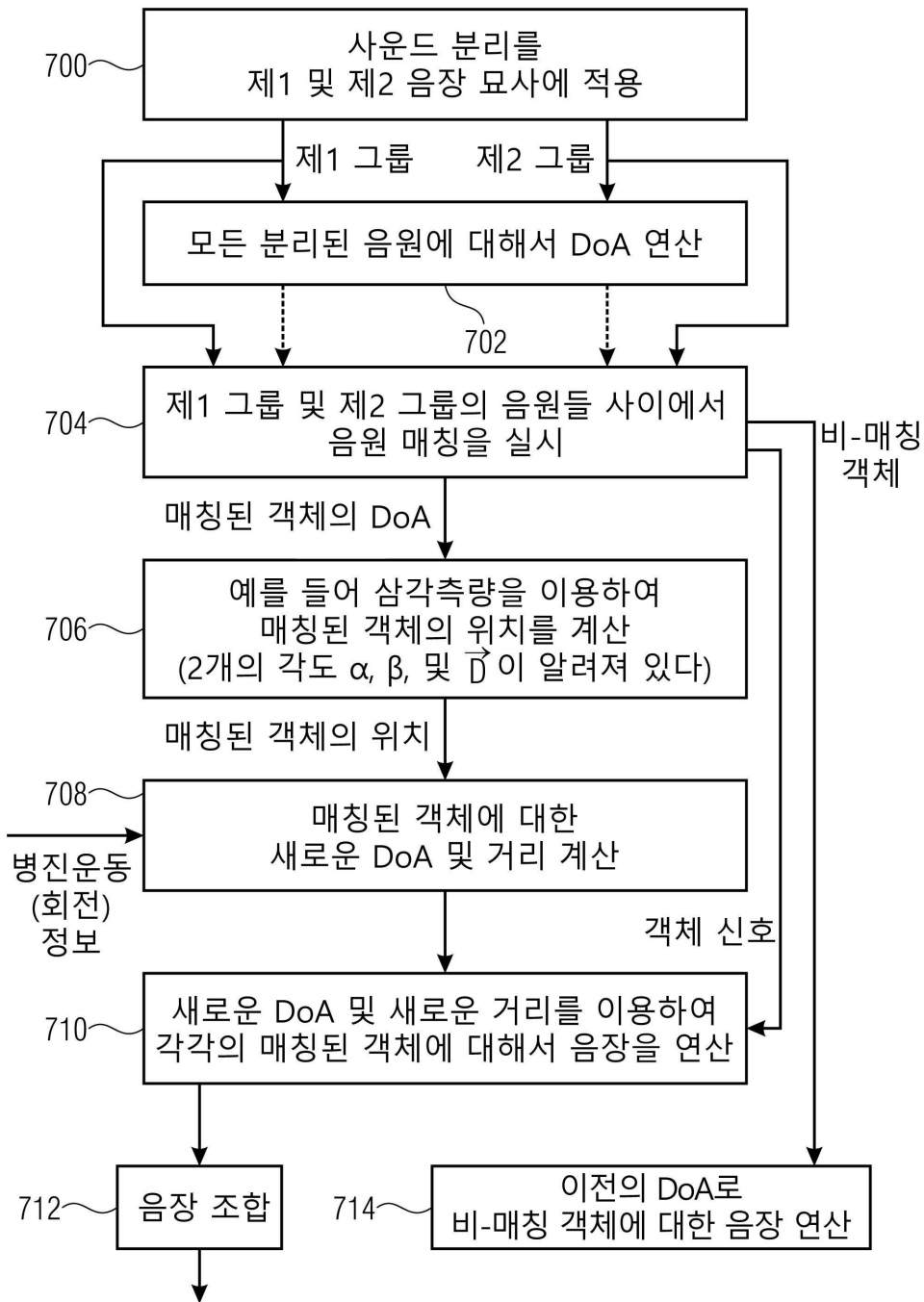
도면4c



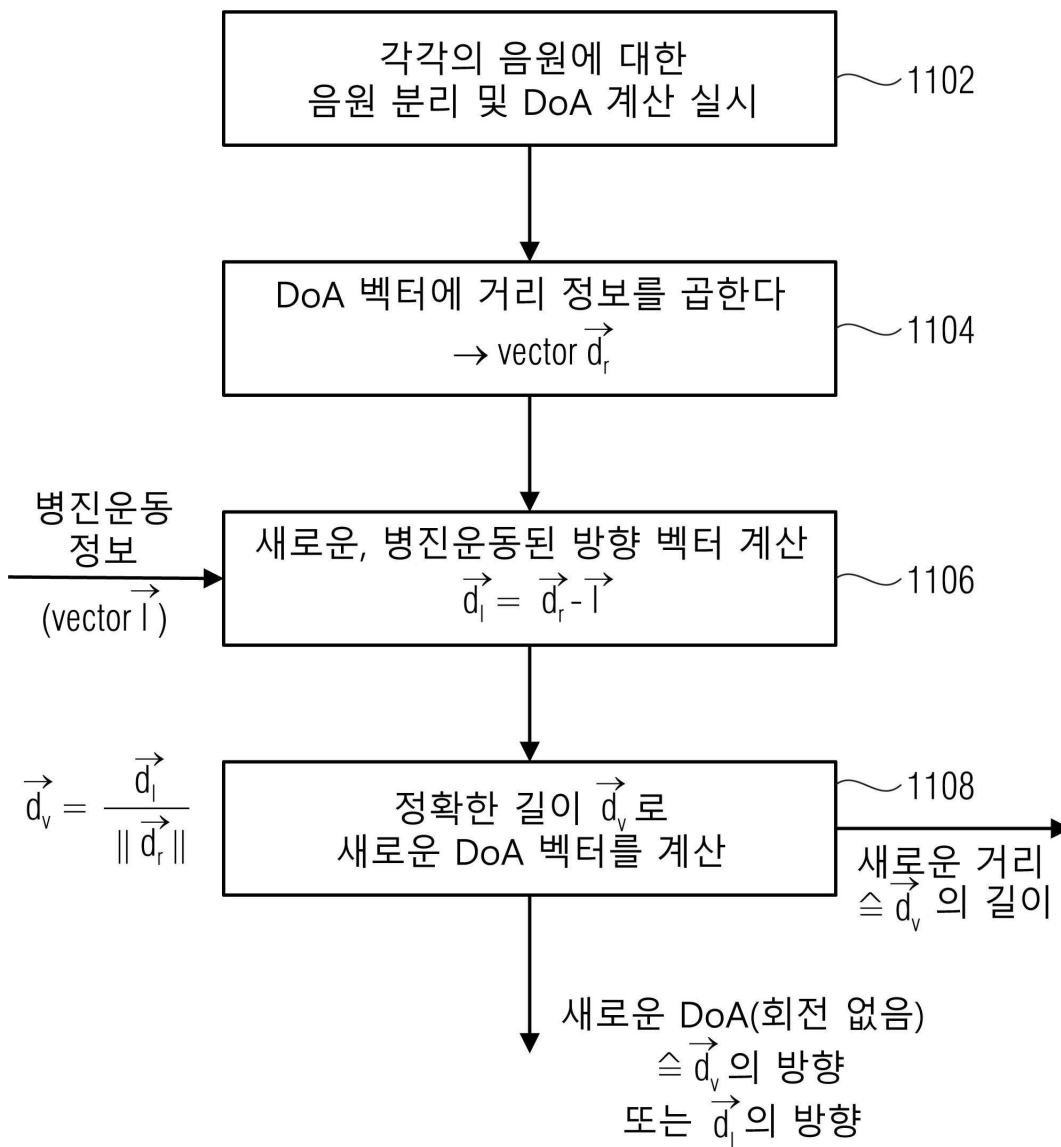
도면5



도면6a



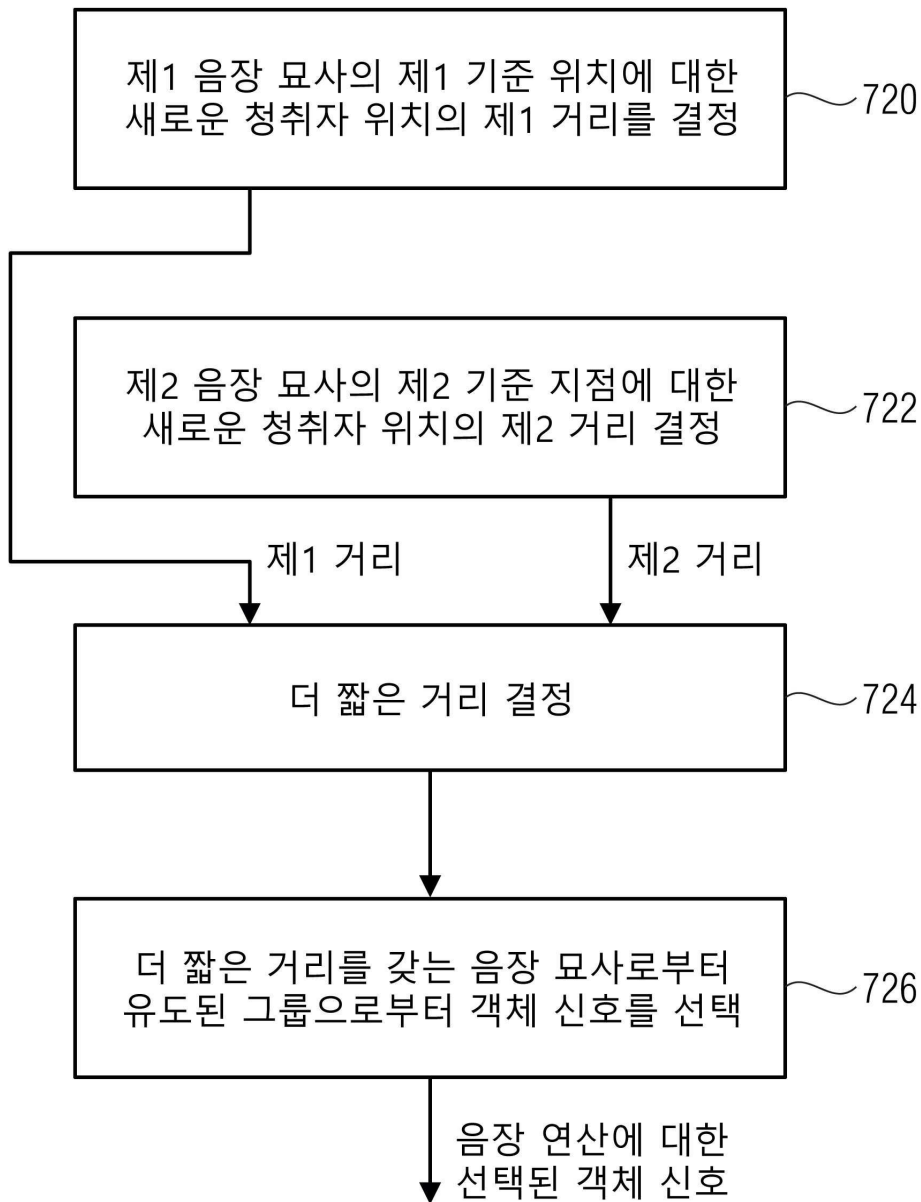
도면6b



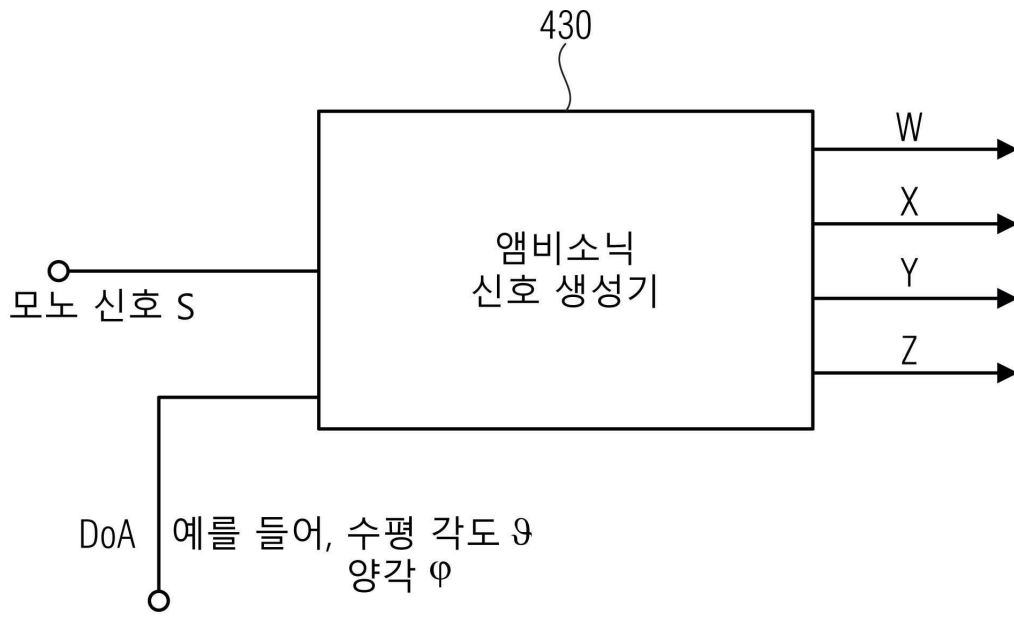
도면6c



도면7



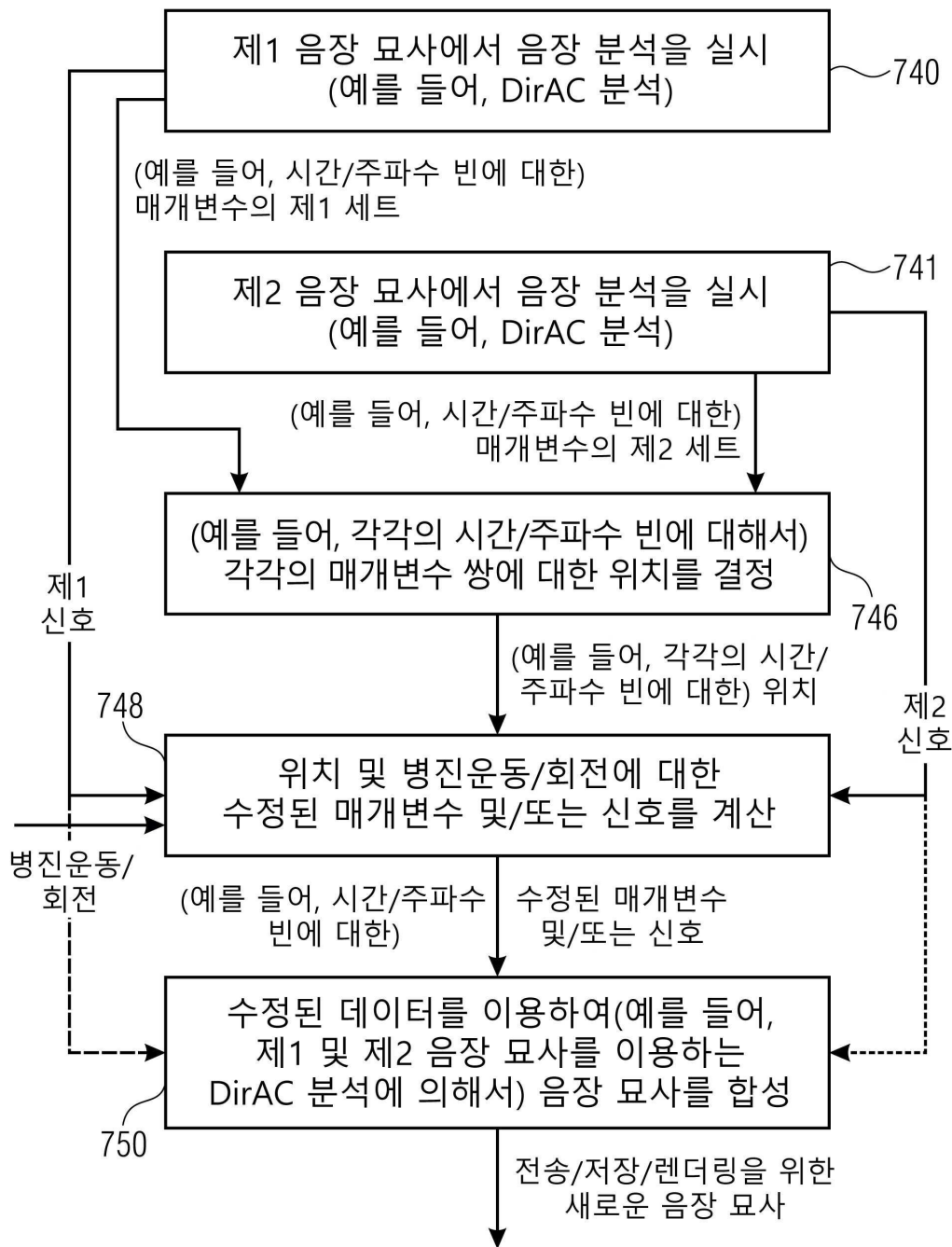
도면8



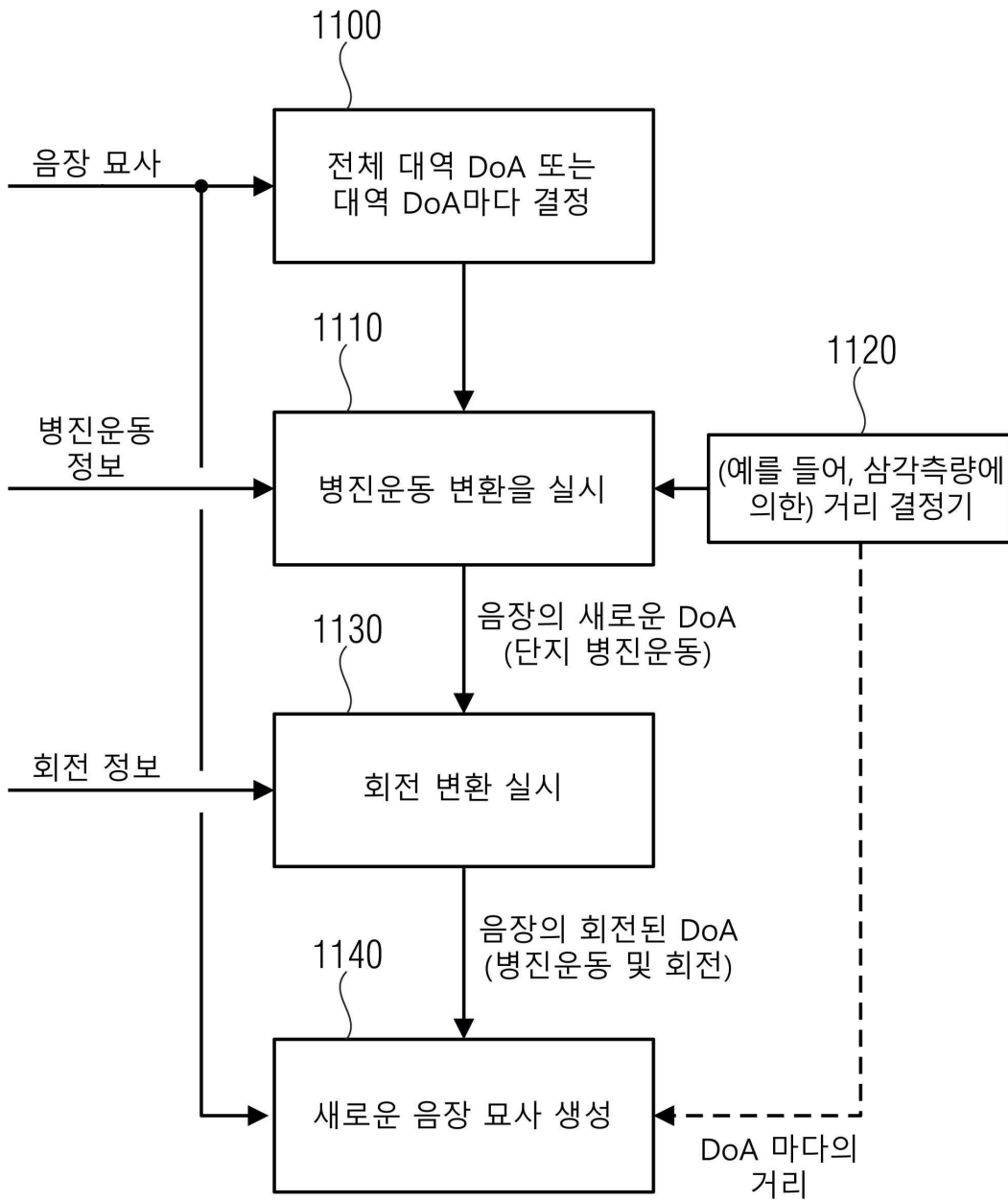
전방향 성분 $W = S \cdot \frac{1}{\sqrt{2}}$

방향 성분 $\begin{cases} X = S \cdot \cos \vartheta \cos \varphi \\ Y = S \cdot \sin \vartheta \cos \varphi \\ Z = S \cdot \sin \varphi \end{cases}$

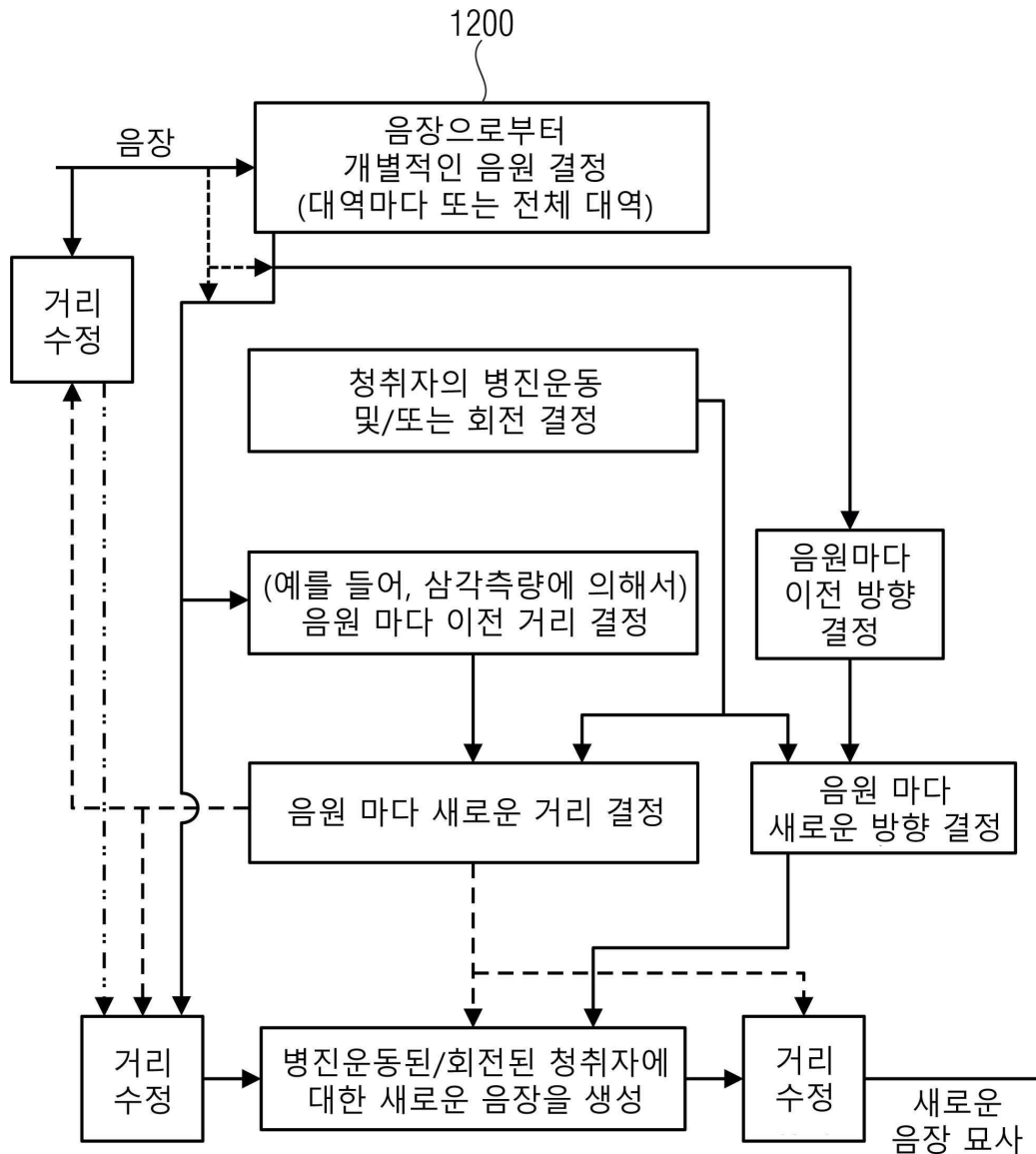
도면9



도면10

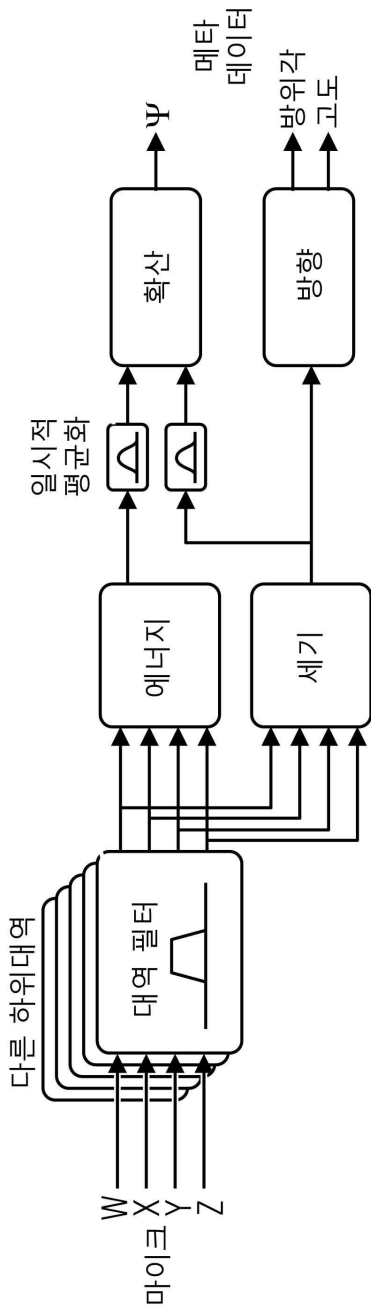


도면11



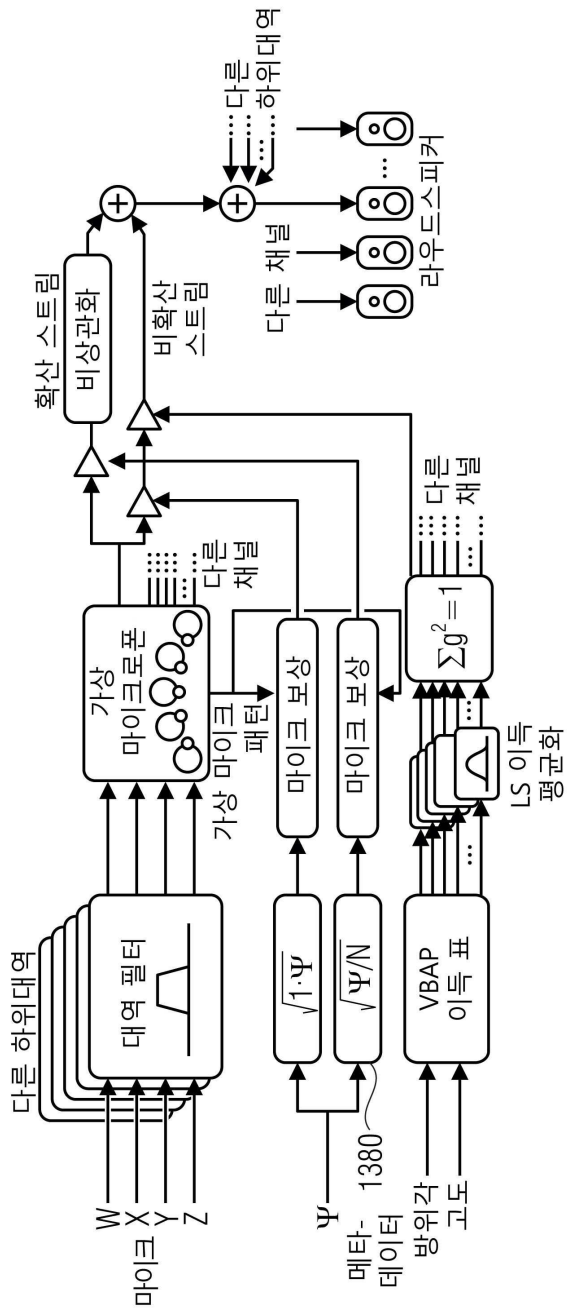
예를 들어, 원래의 음장에 대한 또는 제1 또는 제2 그룹으로부터의 분리된 음원에 대한, 또는 각각의 개별적인 새로운 앰비소닉 신호에 대한, 또는 모든 음원으로부터의 부가된 앰비소닉 신호에 대한 거리 수정

도면12a



DirAC 분석

도면12b



Dirac 합성