



(19) 대한민국특허청(KR)
 (12) 등록특허공보(B1)

(45) 공고일자 2017년04월18일
 (11) 등록번호 10-1728274
 (24) 등록일자 2017년04월12일

- (51) 국제특허분류(Int. Cl.)
H04S 7/00 (2006.01) *G10K 15/12* (2006.01)
G10L 19/008 (2014.01) *H04S 1/00* (2006.01)
H04S 3/00 (2006.01) *H04S 5/00* (2015.01)
- (52) CPC특허분류
H04S 7/305 (2013.01)
G10K 15/12 (2013.01)
- (21) 출원번호 10-2015-7036325
- (22) 출원일자(국제) 2014년05월28일
 심사청구일자 2016년09월27일
- (85) 번역문제출일자 2015년12월22일
- (65) 공개번호 10-2016-0015269
- (43) 공개일자 2016년02월12일
- (86) 국제출원번호 PCT/US2014/039863
- (87) 국제공개번호 WO 2014/194004
 국제공개일자 2014년12월04일
- (30) 우선권주장
 61/828,620 2013년05월29일 미국(US)
 (뒷면에 계속)
- (56) 선행기술조사문현
 US20060045275 A1

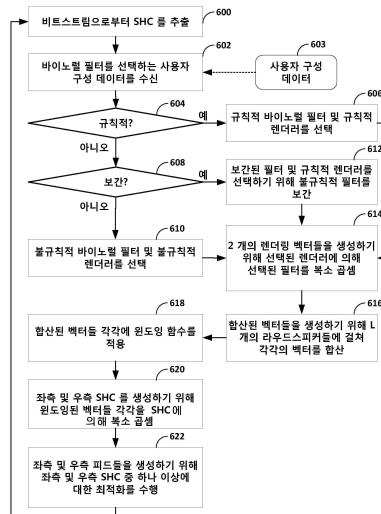
전체 청구항 수 : 총 31 항

심사관 : 권영학

(54) 발명의 명칭 구면 조화 계수들의 바이노럴 렌더링

(57) 요 약

디바이스는 사운드 필드를 렌더링하도록 3차원으로 사운드 필드를 나타내는 구면 조화 계수들에 바이노럴 킴 임 펠스 응답 필터를 적용하도록 구성되는 하나 이상의 프로세서들을 포함한다.

대 표 도 - 도15

(52) CPC특허분류

G10L 19/008 (2013.01)
H04S 1/002 (2013.01)
H04S 1/005 (2013.01)
H04S 3/004 (2013.01)
H04S 5/00 (2013.01)
H04S 7/306 (2013.01)
H04S 7/307 (2013.01)
H04S 2400/01 (2013.01)
H04S 2420/01 (2013.01)

(30) 우선권주장

61/847,543	2013년07월17일	미국(US)
61/886,593	2013년10월03일	미국(US)
61/886,620	2013년10월03일	미국(US)
14/288,276	2014년05월27일	미국(US)

명세서

청구범위

청구항 1

바이노럴 오디오 렌더링의 방법으로서,

복수의 스피커 피드들로서 사운드 필드를 렌더링하도록 고차 앰비소닉 계수들에 복수의 불규칙적 바이노럴 툴 임펄스 응답 (BRIR) 필터들을 적용하는 단계를 포함하고,

상기 고차 앰비소닉 계수들은 3차원으로 상기 사운드 필드를 나타내고,

상기 복수의 불규칙적 BRIR 필터들 중 각각의 개별적인 불규칙적 BRIR 필터는 복수의 가상 라우드스피커들 중 개별적인 가상 라우드스피커의 임펄스 위치에서 생성된 임펄스에 대한 응답을 나타내고,

상기 복수의 가상 라우드스피커들은 등간격으로 이격되어 있지 않은, 바이노럴 오디오 렌더링의 방법.

청구항 2

제 1 항에 있어서,

상기 고차 앰비소닉 계수들은 고차 앰비소닉 계수들의 제 1 세트이고, 상기 사운드 필드는 제 1 사운드 필드이고, 상기 복수의 가상 라우드스피커들은 제 1 복수의 가상 라우드스피커들이고,

상기 방법은:

복수의 규칙적 BRIR 필터들의 사용을 특정하는 사용자 구성 데이터를 수신하고 후속하여 복수의 불규칙적 BRIR 필터들을 상기 고차 앰비소닉 계수들의 제 1 세트에 적용하는 것에 응답하여, 제 2 사운드 필드를 렌더링하도록, 복수의 규칙적 BRIR 필터들을 고차 앰비소닉 계수들의 제 2 세트에 적용하는 단계를 더 포함하고,

상기 복수의 규칙적 BRIR 필터들 중 각각의 개별적인 규칙적 BRIR 필터는 제 2 복수의 가상 라우드스피커들 중 개별적인 가상 라우드스피커의 임펄스 위치에서 생성된 임펄스에 대한 응답을 나타내고,

상기 제 2 복수의 가상 라우드스피커들은 등간격으로 이격되어 있는, 바이노럴 오디오 렌더링의 방법.

청구항 3

제 1 항에 있어서,

상기 고차 앰비소닉 계수들에 복수의 불규칙적 BRIR 필터들을 적용하는 단계는, 좌측 및 우측 변경된 고차 앰비소닉 계수들을 생성하고, 상기 복수의 스피커 피드들은 제 1 주파수 도메인 스피커 피드 및 제 2 주파수 도메인 스피커 피드를 포함하고,

상기 방법은:

상기 제 1 주파수 도메인 스피커 피드를 생성하기 위해 상기 고차 앰비소닉 계수들과 연관된 차수들 및 하위 차수들의 수에 걸쳐 제 1 변경된 고차 앰비소닉 계수들을 합산하는 단계로서, 상기 제 1 변경된 고차 앰비소닉 계수들은 좌측 변경된 고차 앰비소닉 계수들 또는 우측 변경된 고차 앰비소닉 계수들을 포함하는, 상기 제 1 변경된 고차 앰비소닉 계수들을 합산하는 단계;

반전된 고차 앰비소닉 계수들을 생성하기 위해 상기 제 1 변경된 고차 앰비소닉 계수들 중, 음의 하위 차수와 연관된 고차 앰비소닉 계수들을 반전시키는 단계; 및

상기 제 2 주파수 도메인 스피커 피드를 생성하기 위해 차수들 및 하위 차수들의 수에 걸쳐 상기 반전된 고차 앰비소닉 계수들을 합산하는 단계를 더 포함하는, 바이노럴 오디오 렌더링의 방법.

청구항 4

제 1 항에 있어서,

상기 고차 앰비소닉 계수들이 대응하는 구면 기저 함수들의 차수는 1 보다 큰, 바이노럴 오디오 렌더링의 방법.

청구항 5

제 1 항에 있어서,

스피커들의 규칙적 배열에 대한 하나 이상의 규칙적 BRIR 필터들을 생성하기 위해 복수의 불규칙적 BRIR 필터들을 보간하는 단계를 더 포함하고,

상기 복수의 불규칙적 BRIR 필터들을 적용하는 단계는, 상기 사운드 필드를 렌더링하도록 상기 고차 앰비소닉 계수들에 복수의 규칙적 BRIR 필터들을 적용하는 단계를 포함하는, 바이노럴 오디오 렌더링의 방법.

청구항 6

제 1 항에 있어서,

윈도잉된 BRIR 필터를 생성하기 위해 상기 복수의 불규칙적 BRIR 필터들에 윈도잉 함수를 적용하는 단계를 더 포함하고,

상기 복수의 불규칙적 BRIR 필터들을 적용하는 단계는, 상기 사운드 필드를 렌더링하도록 상기 고차 앰비소닉 계수들에 상기 윈도잉된 BRIR 필터를 적용하는 단계를 포함하는, 바이노럴 오디오 렌더링의 방법.

청구항 7

제 1 항에 있어서,

변환된 불규칙적 BRIR 필터들을 생성하도록 상기 복수의 불규칙적 BRIR 필터들을 시간 도메인으로부터 주파수 도메인으로 변환하는 단계를 더 포함하고,

상기 복수의 불규칙적 BRIR 필터들을 적용하는 단계는, 상기 사운드 필드를 렌더링하도록 상기 고차 앰비소닉 계수들에 상기 변환된 불규칙적 BRIR 필터들을 적용하는 단계를 포함하는, 바이노럴 오디오 렌더링의 방법.

청구항 8

제 1 항에 있어서,

변환된 BRIR 필터들을 생성하도록 복수의 불규칙적 필터들을 시간 도메인으로부터 주파수 도메인으로 변환하는 단계; 및

변환된 고차 앰비소닉 계수들을 생성하도록 상기 고차 앰비소닉 계수들을 시간 도메인으로부터 주파수 도메인으로 변환하는 단계를 더 포함하고,

상기 복수의 불규칙적 BRIR 필터들을 적용하는 단계는, 상기 사운드 필드의 주파수 도메인 표현을 렌더링하도록, 상기 변환된 고차 앰비소닉 계수들에 변환된 상기 불규칙적 BRIR 필터들을 적용하는 단계를 포함하고,

상기 방법은 상기 사운드 필드를 렌더링하기 위해 상기 사운드 필드의 주파수 도메인 표현에 역 변환을 적용하는 단계를 더 포함하는, 바이노럴 오디오 렌더링의 방법.

청구항 9

제 1 항에 있어서,

상기 복수의 불규칙적 BRIR 필터들을 적용하는 단계는, 상기 고차 앰비소닉 계수들에 직접 상기 복수의 불규칙적 BRIR 필터들을 적용하는 단계를 포함하는, 바이노럴 오디오 렌더링의 방법.

청구항 10

제 1 항에 있어서,

상기 복수의 불규칙적 BRIR 필터들을 적용하는 단계는, 상기 고차 앰비소닉 계수들을 상기 불규칙적 BRIR 필터

들로 컨볼빙하는 단계를 포함하는, 바이노럴 오디오 렌더링의 방법.

청구항 11

제 10 항에 있어서,

상기 복수의 불규칙적 BRIR 필터들을 적용하는 단계는, 상기 스피커 피드들로서의 출력을 위하여 상기 사운드 필드를 렌더링하도록 컨볼루션들을 누적하는 단계를 더 포함하고,

상기 컨볼루션들은 상기 고차 앰비소닉 계수들을 상기 불규칙적 BRIR 필터들로 컨볼빙하는 것으로 인해 발생하는, 바이노럴 오디오 렌더링의 방법.

청구항 12

디바이스로서,

복수의 스피커 피드들로서 사운드 필드를 렌더링하도록 고차 앰비소닉 계수들에 복수의 불규칙적 바이노럴 툴 임펄스 응답 (BRIR) 필터들을 적용하도록 구성되는 하나 이상의 프로세서들을 포함하고,

상기 고차 앰비소닉 계수들은 3차원으로 상기 사운드 필드를 나타내고,

상기 복수의 불규칙적 BRIR 필터들 중 각각의 개별적인 불규칙적 BRIR 필터는 복수의 가상 라우드스피커들 중 개별적인 가상 라우드스피커의 임펄스 위치에서 생성된 임펄스에 대한 응답을 나타내고,

상기 복수의 가상 라우드스피커들은 등간격으로 이격되어 있지 않은, 디바이스.

청구항 13

제 12 항에 있어서,

상기 고차 앰비소닉 계수들은 고차 앰비소닉 계수들의 제 1 세트이고, 상기 사운드 필드는 제 1 사운드 필드이고, 상기 복수의 가상 라우드스피커들은 제 1 복수의 가상 라우드스피커들이고, 상기 하나 이상의 프로세서들은 또한, 스피커들의 규칙적 배열에 대한 복수의 규칙적 BRIR 필터들의 사용을 특정하는 사용자 구성 데이터를 수신하는 것에 응답하여, 제 2 사운드 필드를 렌더링하도록, 상기 복수의 규칙적 BRIR 필터들을 고차 앰비소닉 계수들의 제 2 세트에 적용하도록 구성되고,

상기 복수의 규칙적 BRIR 필터들 중 각각의 개별적인 규칙적 BRIR 필터는 제 2 복수의 가상 라우드스피커들 중 개별적인 가상 라우드스피커의 임펄스 위치에서 생성된 임펄스에 대한 응답을 나타내고,

상기 제 2 복수의 가상 라우드스피커들은 등간격으로 이격되어 있는, 디바이스.

청구항 14

제 12 항에 있어서,

상기 하나 이상의 프로세서들은 또한:

좌측 및 우측 변경된 고차 앰비소닉 계수들을 생성하기 위해 상기 고차 앰비소닉 계수들에 상기 복수의 불규칙적 BRIR 필터들을 적용하는 것으로서, 상기 복수의 스피커 피드들은 제 1 주파수 도메인 스피커 피드 및 제 2 주파수 도메인 스피커 피드를 포함하는, 상기 복수의 불규칙적 BRIR 필터들을 적용하고;

상기 제 1 주파수 도메인 스피커 피드를 생성하도록 상기 고차 앰비소닉 계수들과 연관된 차수들 및 하위 차수들의 수에 걸쳐 제 1 변경된 고차 앰비소닉 계수들을 합산하는 것으로서, 상기 제 1 변경된 고차 앰비소닉 계수들은 좌측 변경된 고차 앰비소닉 계수들 또는 우측 변경된 고차 앰비소닉 계수들을 포함하는, 상기 제 1 변경된 고차 앰비소닉 계수들을 합산하고;

반전된 고차 앰비소닉 계수들을 생성하기 위해 상기 제 1 변경된 고차 앰비소닉 계수들 중, 음의 하위 차수와 연관된 고차 앰비소닉 계수들을 반전시키고; 그리고

상기 제 2 주파수 도메인 스피커 피드를 생성하기 위해 차수들 및 하위 차수들의 수에 걸쳐 상기 반전된 고차 앰비소닉 계수들을 합산하도록 구성되는, 디바이스.

청구항 15

제 12 항에 있어서,

상기 고차 앰비소닉 계수들이 대응하는 구면 기저 함수들의 차수는 1 보다 큰, 디바이스.

청구항 16

제 12 항에 있어서,

상기 하나 이상의 프로세서들은 또한, 복수의 규칙적 BRIR 필터들을 생성하기 위해 상기 복수의 불규칙적 BRIR 필터들을 보간하도록 구성되고,

상기 규칙적 BRIR 필터들은 스피커들의 규칙적 배열에 대한 복수의 BRIR 필터들을 포함하며,

상기 하나 이상의 프로세서들은 또한, 상기 복수의 불규칙적 BRIR 필터들을 적용하기 위해, 상기 사운드 필드를 렌더링하도록 상기 고차 앰비소닉 계수들에 상기 복수의 규칙적 BRIR 필터들을 적용하도록 구성되는, 디바이스.

청구항 17

제 12 항에 있어서,

상기 하나 이상의 프로세서들은 또한, 원도잉된 BRIR 필터를 생성하기 위해 복수의 불규칙적 필터들에 원도잉 함수를 적용하도록 구성되고,

상기 하나 이상의 프로세서들은 또한, 상기 복수의 불규칙적 BRIR 필터들을 적용할 때, 상기 사운드 필드를 렌더링하도록 상기 고차 앰비소닉 계수들에 상기 원도잉된 BRIR 필터를 적용하도록 구성되는, 디바이스.

청구항 18

제 12 항에 있어서,

상기 하나 이상의 프로세서들은 또한, 변환된 불규칙적 BRIR 필터들을 생성하도록 상기 복수의 불규칙적 BRIR 필터들을 시간 도메인으로부터 주파수 도메인으로 변환하도록 구성되고,

상기 하나 이상의 프로세서들은 또한, 상기 복수의 불규칙적 BRIR 필터를 적용할 때, 상기 사운드 필드를 렌더링하도록 상기 고차 앰비소닉 계수들에 상기 변환된 불규칙적 BRIR 필터들을 적용하도록 구성되는, 디바이스.

청구항 19

제 12 항에 있어서,

상기 하나 이상의 프로세서들은 또한, 변환된 불규칙적 BRIR 필터들을 생성하도록 상기 복수의 불규칙적 BRIR 필터들을 시간 도메인으로부터 주파수 도메인으로 변환하고, 그리고 변환된 고차 앰비소닉 계수들을 생성하도록 상기 고차 앰비소닉 계수들을 시간 도메인으로부터 주파수 도메인으로 변환하도록 구성되고,

상기 하나 이상의 프로세서들은 또한, 상기 복수의 불규칙적 BRIR 필터들을 적용할 때, 상기 사운드 필드의 주파수 도메인 표현을 렌더링하도록, 상기 변환된 고차 앰비소닉 계수들에 상기 변환된 불규칙적 BRIR 필터들을 적용하도록 구성되고,

상기 하나 이상의 프로세서들은 또한, 상기 사운드 필드를 렌더링하기 위해 상기 사운드 필드의 주파수 도메인 표현에 역 변환을 적용하도록 구성되는, 디바이스.

청구항 20

제 12 항에 있어서,

상기 하나 이상의 프로세서들은 또한, 상기 복수의 불규칙적 BRIR 필터들을 적용할 때, 상기 고차 앰비소닉 계수들에 직접 상기 복수의 불규칙적 BRIR 필터들을 적용하도록 구성되는, 디바이스.

청구항 21

제 12 항에 있어서,

상기 하나 이상의 프로세서들은 상기 복수의 불규칙적 BRIR 필터들을 적용하는 부분으로서, 상기 하나 이상의 프로세서들이 상기 고차 앰비소닉 계수들을 상기 불규칙적 BRIR 필터들로 컨볼빙하게 하도록 구성되는, 디바이스.

청구항 22

제 21 항에 있어서,

상기 하나 이상의 프로세서들은 상기 복수의 불규칙적 BRIR 필터들을 적용하는 부분으로서, 상기 하나 이상의 프로세서들이 상기 스피커 피드들로서의 출력을 위하여 상기 사운드 필드를 렌더링하도록 컨볼루션들을 누적하게 하도록 구성되고,

상기 컨볼루션들은 상기 고차 앰비소닉 계수들을 상기 불규칙적 BRIR 필터들로 컨볼빙하는 것으로 인해 발생하는, 디바이스.

청구항 23

장치로서,

3차원으로 사운드 필드를 나타내는 고차 앰비소닉 계수들을 결정하는 수단; 및

복수의 스피커 피드들로서 사운드 필드를 렌더링하도록 상기 고차 앰비소닉 계수들에 복수의 불규칙적 바이노럴 룸 임펄스 응답 (BRIR; binaural room impulse response) 필터들을 적용하는 수단을 포함하고,

상기 복수의 불규칙적 BRIR 필터들 중 각각의 개별적인 불규칙적 BRIR 필터는 복수의 가상 라우드스피커들 중 개별적인 가상 라우드스피커의 임펄스 위치에서 생성된 임펄스에 대한 응답을 나타내고,

상기 복수의 가상 라우드스피커들은 등간격으로 이격되어 있지 않은, 장치.

청구항 24

제 23 항에 있어서,

상기 고차 앰비소닉 계수들은 고차 앰비소닉 계수들의 제 1 세트이고, 상기 사운드 필드는 제 1 사운드 필드이고, 상기 복수의 가상 라우드스피커들은 제 1 복수의 가상 라우드스피커들이고,

상기 장치는:

복수의 규칙적 BRIR 필터들의 사용을 특정하는 사용자 구성 데이터를 수신하는 수단; 및

제 2 사운드 필드를 렌더링하도록, 복수의 규칙적 BRIR 필터들을 고차 앰비소닉 계수들의 제 2 세트에 적용하는 수단을 더 포함하고,

상기 복수의 규칙적 BRIR 필터들 중 각각의 개별적인 규칙적 BRIR 필터는 제 2 복수의 가상 라우드스피커들 중 개별적인 가상 라우드스피커의 임펄스 위치에서 생성된 임펄스에 대한 응답을 나타내고,

상기 제 2 복수의 가상 라우드스피커들은 등간격으로 이격되어 있는, 장치.

청구항 25

제 23 항에 있어서,

상기 고차 앰비소닉 계수들에 복수의 불규칙적 BRIR 필터들을 적용하는 수단은, 좌측 및 우측 변경된 고차 앰비소닉 계수들을 생성하고, 상기 복수의 스피커 피드들은 제 1 주파수 도메인 스피커 피드 및 제 2 주파수 도메인 스피커 피드를 포함하고,

상기 장치는:

상기 제 1 주파수 도메인 스피커 피드를 생성하기 위해 상기 고차 앰비소닉 계수들과 연관된 차수들 및 하위 차수들의 수에 걸쳐 제 1 변경된 고차 앰비소닉 계수들을 합산하는 수단으로서, 상기 제 1 변경된 고차 앰비소닉 계수들은 좌측 변경된 고차 앰비소닉 계수들 또는 우측 변경된 고차 앰비소닉 계수들을 포함하는, 상기 제 1 변경된 고차 앰비소닉 계수들을 합산하는 수단;

반전된 고차 앰비소닉 계수들을 생성하기 위해 상기 제 1 변경된 고차 앰비소닉 계수들 중, 음의 하위 차수와 연관된 고차 앰비소닉 계수들을 반전시키는 수단; 및

상기 제 2 주파수 도메인 스피커 피드를 생성하기 위해 차수들 및 하위 차수들의 수에 걸쳐 상기 반전된 고차 앰비소닉 계수들을 합산하는 수단을 더 포함하는, 장치.

청구항 26

제 23 항에 있어서,

상기 고차 앰비소닉 계수들이 대응하는 구면 기저 함수들의 차수는 1 보다 큰, 장치.

청구항 27

제 23 항에 있어서,

복수의 규칙적 BRIR 필터들을 생성하기 위해 상기 복수의 불규칙적 BRIR 필터들을 보간하는 수단을 더 포함하고,

상기 규칙적 BRIR 필터들은 스피커들의 규칙적 배열에 대한 복수의 BRIR 필터들을 포함하며,

상기 복수의 불규칙적 BRIR 필터들을 적용하는 수단은, 상기 사운드 필드를 렌더링하도록 상기 고차 앰비소닉 계수들에 상기 복수의 규칙적 BRIR 필터들을 적용하는 수단을 포함하는, 장치.

청구항 28

제 23 항에 있어서,

원도잉된 BRIR 필터를 생성하기 위해 상기 복수의 불규칙적 BRIR 필터들에 원도잉 함수를 적용하는 수단을 더 포함하고,

상기 복수의 불규칙적 BRIR 필터들을 적용하는 수단은, 상기 사운드 필드를 렌더링하도록 상기 고차 앰비소닉 계수들에 상기 원도잉된 BRIR 필터를 적용하는 수단을 포함하는, 장치.

청구항 29

제 23 항에 있어서,

변환된 BRIR 필터들을 생성하도록 상기 복수의 불규칙적 BRIR 필터들을 시간 도메인으로부터 주파수 도메인으로 변환하는 수단을 더 포함하고,

상기 복수의 불규칙적 BRIR 필터들을 적용하는 수단은, 상기 사운드 필드를 렌더링하도록 상기 고차 앰비소닉 계수들에 변환된 상기 불규칙적 BRIR 필터들을 적용하는 수단을 포함하는, 장치.

청구항 30

제 23 항에 있어서,

변환된 불규칙적 BRIR 필터들을 생성하도록 상기 복수의 불규칙적 BRIR 필터들을 시간 도메인으로부터 주파수 도메인으로 변환하는 수단; 및

변환된 고차 앰비소닉 계수들을 생성하도록 상기 고차 앰비소닉 계수들을 시간 도메인으로부터 주파수 도메인으로 변환하는 수단을 더 포함하고,

상기 복수의 불규칙적 BRIR 필터들을 적용하는 수단은, 상기 사운드 필드의 주파수 도메인 표현을 렌더링하도록, 상기 변환된 고차 앰비소닉 계수들에 상기 변환된 불규칙적 BRIR 필터들을 적용하는 수단을 포함하고,

상기 장치는 상기 사운드 필드를 렌더링하기 위해 상기 사운드 필드의 주파수 도메인 표현에 역 변환을 적용하는 수단을 더 포함하는, 장치.

청구항 31

명령들을 저장하는 비일시적 컴퓨터 판독가능 저장 매체로서,
 상기 명령들은 실행될 때 하나 이상의 프로세서들로 하여금,
 복수의 스피커 피드들로서 사운드 필드를 렌더링하도록 고차 앰비소닉 계수들에 복수의 불규칙적 바이노럴 룸 임펄스 응답 (BRIR) 필터들을 적용하게 하고,
 상기 고차 앰비소닉 계수들은 3차원으로 상기 사운드 필드를 나타내고,
 상기 복수의 불규칙적 BRIR 필터들 중 각각의 개별적인 불규칙적 BRIR 필터는 복수의 가상 라우드스피커들 중 개별적인 가상 라우드스피커의 임펄스 위치에서 생성된 임펄스에 대한 응답을 나타내고,
 상기 복수의 가상 라우드스피커들은 등간격으로 이격되어 있지 않은, 명령들을 저장하는 비일시적 컴퓨터 판독 가능 저장 매체.

발명의 설명

기술 분야

[0001]

우선권 주장

[0002]

본 출원은 2013년 5월 29일 출원된 미국 특허 출원 번호 제 61/828,620호, 2013년 7월 17일 출원된 미국 특허 출원 번호 제61/847,543호, 2013년 10월 3일 출원된 미국 특허 출원 번호 제61/886,593호 및 2013년 10월 3일 출원된 미국 특허 출원번호 제61/886,620호의 이익을 우선권으로 주장한다.

[0003]

기술 분야

[0004]

본 개시물은 오디오 렌더링 및 보다 구체적으로 오디오 데이터의 바이노럴 렌더링에 관한 것이다.

발명의 내용

과제의 해결 수단

[0005]

일반적으로, 1 보다 큰 차수를 갖는 구면 조화 계수들 (고차 앰비소닉 (HOA; higher order ambisonics) 계수들이라 지칭될 수도 있음) 의 바이노럴 오디오 렌더링을 위한 기법들이 설명된다.

[0006]

일 예로서, 바이노럴 오디오 렌더링의 방법은 사운드 필드를 렌더링하도록 3차원으로 사운드 필드를 나타내는 구면 조화 계수들에 바이노럴 룸 임펄스 응답 필터를 적용하는 단계를 포함한다.

[0007]

다른 예에서, 디바이스는 사운드 필드를 렌더링하도록 3차원으로 사운드 필드를 나타내는 구면 조화 계수들에 바이노럴 룸 임펄스 응답 필터를 적용하도록 구성되는 하나 이상의 프로세서들을 포함한다.

[0008]

다른 예에서, 디바이스는 사운드 필드를 3 차원으로 표현하는 구면 조화 계수들을 결정하는 수단, 및 사운드 필드를 렌더링하도록 사운드 필드를 나타내는 구면 조화 계수들에 바이노럴 룸 임펄스 응답 필터를 적용하는 수단을 포함한다.

[0009]

다른 예에서, 명령들이 저장된 비일시적 컴퓨터 판독가능 저장 매체로서, 명령들은 실행될 때 하나 이상의 프로세서들로 하여금 사운드 필드를 렌더링하도록 3차원으로 사운드 필드를 나타내는 구면 조화 계수들에 바이노럴 룸 임펄스 응답 필터를 적용하게 한다.

[0010]

본 기법들의 하나 이상의 양태들의 세부사항들은 첨부되는 도면들 및 하기의 설명들에서 기술된다. 이를 기법들의 다른 특성들, 목적들 및 이점들은 상세한 설명 및 도면들로부터, 그리고 청구범위로부터 명확해질 것이다.

도면의 간단한 설명

[0011]

도 1 및 도 2 는 다양한 차수 (order) 를 및 하위-차수 (sub-order) 를의 구면 조화 기저 합수 (spherical harmonic basis function) 들을 예시하는 도면들이다.

도 3 은 오디오 신호 정보를 보다 효율적으로 렌더링하기 위해 본 개시물에 설명된 기법들을 수행할 수도 있는 시스템을 예시하는 다이어그램이다.

도 4 는 예시적인 바이노럴 툼 임펄스 응답 (BRIR) 을 예시하는 블록도이다.

도 5 는 툼에서 BRIR 을 생성하기 위한 예시적인 시스템 모델을 예시하는 블록도이다.

도 6 은 툼에서 BRIR 을 생성하기 위한 보다 세밀한 시스템 모델을 예시하는 블록도이다.

도 7 은 본 개시물에 설명된 바이노럴 오디오 렌더링 기법들의 여러 양태들을 수행할 수도 있는 오디오 재생 디바이스의 일 예를 예시하는 블록도이다.

도 8 은 본 개시물에 설명된 바이노럴 오디오 렌더링 기법들의 여러 양태들을 수행할 수도 있는 오디오 재생 디바이스의 일 예를 예시하는 블록도이다.

도 9 는 본 개시물에 설명된 기법들의 여러 양태들에 따라 구면 조화 계수들을 렌더링하기 위한 바이노럴 렌더링 디바이스에 대한 예시적인 동작 모드를 예시하는 흐름도이다.

도 10a 및 도 10b 는 본 개시물에 설명된 기법들의 여러 양태들에 따라 도 7 및 도 8 의 오디오 재생 디바이스들에 의해 수행될 수도 있는 대안의 동작 모드들을 예시하는 흐름도들이다.

도 11 은 본 개시물에 설명된 바이노럴 오디오 렌더링 기법들의 여러 양태들을 수행할 수도 있는 오디오 재생 디바이스의 일 예를 예시하는 블록도이다.

도 12 는 본 개시물에 설명된 기법들의 여러 양태들에 따라 도 11 의 오디오 재생 디바이스에 의해 수행될 수도 있는 프로세스를 예시하는 흐름도이다.

도 13 은 본 개시물에 설명된 바이노럴 오디오 렌더링 기법들의 여러 양태들을 수행할 수도 있는 오디오 재생 디바이스의 일 예를 예시하는 블록도이다.

도 14 는 본 개시물에 설명된 바이노럴 오디오 렌더링 기법들의 여러 양태들을 수행할 수도 있는 오디오 재생 디바이스의 일 예를 예시하는 블록도이다.

도 15 는 본 개시물에 설명된 기법들의 여러 양태들에 따라 구면 조화 계수들을 렌더링하기 위한 바이노럴 렌더링 디바이스에 대한 예시적인 동작 모드를 예시하는 흐름도이다.

도 16a, 도 16b 는 본 개시물에 설명된 기법들의 여러 양태들에 따라 도 13, 도 14 의 오디오 재생 디바이스들에 의해 수행될 수도 있는 개념적인 프로세스를 각각 예시하는 다이어그램들이다.

유사한 도면 문자들은 도면 및 텍스트 전반에 걸쳐 유사한 엘리먼트들을 표기한다.

발명을 실시하기 위한 구체적인 내용

[0012]

서라운드 사운드 (surround sound) 의 진화는 최근 엔터테인먼트를 위한 다수의 출력 포맷들을 이용가능하게 하였다. 이러한 서라운드 사운드 포맷들의 예들은 대중적인 5.1 포맷 (다음의 6 개의 채널들: 전방 좌측 (front left; FL), 전방 우측 (front right; FR), 중앙 또는 전방 중앙, 후방 좌측 또는 서라운드 좌측, 후방 우측 또는 서라운드 우측, 및 저주파수 효과들 (low frequency effects; LFE) 을 포함함), 성장하는 7.1 포맷, 및 (예컨대, 초고해상도 텔레비전 (Ultra High Definition Television) 표준과 함께 이용하기 위한) 당면한 22.2 포맷을 포함한다. 공간 오디오 포맷의 다른 예는 (또한 고차 앰비소닉으로서 알려진) 구면 조화 계수들이다.

[0013]

미래의 표준화된 오디오 인코더 (PCM 오디오 표현들을 비트스트림으로 변환하여 시간 샘플 당 요구되는 비트들의 수를 보존하는 디바이스) 에 대한 입력은 선택적으로 3 개의 가능한 포맷들 중 하나: (i) 사전-특정된 포지션들에서 라우드스피커들을 통해 플레이되도록 의도되는 전통적인 채널-기반 오디오; (ii) (다른 정보 중에서) 그 위치 좌표들을 포함하는 연관된 메타데이터를 갖는 단일 오디오 객체들에 대한 개별 펄스-코드-변조 (pulse-code-modulation; PCM) 데이터를 포함하는 객체-기반 오디오 (object-based audio); 및 (iii) 구면 조화 계수들 (SHC) - 여기에서 계수들은 구면 조화 기반 함수들의 선형 합산의 '가중치'를 나타냄 - 을 이용하여 사운드 필드를 나타내는 것을 포함하는 장면-기반 오디오 (scene-based audio) 일 수도 있다. 이 문맥에서, SHC 는 고차 앰비소닉 (HoA) 모델에 따라 HoA 신호들을 포함할 수도 있다. 구면 조화 계수들은 추가로 또는 대안으로 평면 모델들 및 구면 모델들을 포함할 수도 있다.

[0014]

시장에는 여러 '서라운드 사운드' 포맷이 존재한다. 이들은 예를 들어, (스테레오를 넘어서서 거실 내로 전출을 행하는 측면에서 가장 성공적이었던) 5.1 홈 시어터 시스템으로부터 NHK (일본 방송 협회 또는 Japan

Broadcasting Corporation)에 의해 개발된 22.2 시스템까지의 범위이다. 컨텐츠 크리에이터 (content creator) 들 (예컨대, 할리우드 스튜디오 (Hollywood studios))은 영화에 대한 사운드트랙을 한 번 생성하기를 원할 것이고, 각각의 스피커 구성에 대해 그것을 리믹스 (remix) 하기 위한 노력들을 소모하기를 원하지 않을 것이다. 최근, 표준 위원회들은 표준화된 비트스트림으로의 인코딩과, 렌더러 (renderer) 의 위치에서의 스피커 기하구조 및 음향 조건들에 대해 적응가능하고 불가지적 (agnostic) 인 추후의 디코딩을 제공하기 위한 방법들을 고려하고 있었다.

[0015] 컨텐츠 크리에이터들에 대해 이러한 유연성을 제공하기 위하여, 엘리먼트들의 계층적인 세트가 사운드 필드를 표현하기 위하여 이용될 수도 있다. 엘리먼트들의 계층적인 세트는, 저차 (lower-ordered) 엘리먼트들의 기본 세트가 모델링된 사운드 필드의 전체 표현을 제공하도록 엘리먼트들이 순서화되어 있는 엘리먼트들의 세트를 지칭할 수도 있다. 세트가 고차 엘리먼트들을 포함하도록 확장될수록, 표현은 더욱 상세하게 된다.

[0016] 엘리먼트들의 계층적인 세트의 하나의 예는 구면 조화 계수들 (SHC)의 세트이다. 다음의 표현식은 SHC 를 이용하여 사운드 필드의 설명 또는 표현을 입증한다:

$$p_i(t, r_r, \theta_r, \varphi_r) = \sum_{\omega=0}^{\infty} \left[4\pi \sum_{n=0}^{\infty} j_n(kr_r) \sum_{m=-n}^n A_n^m(k) Y_n^m(\theta_r, \varphi_r) \right] e^{j\omega t}$$

[0017] 이 표현식은 사운드 필드의 임의의 포인트 $\{r_r, \theta_r, \varphi_r\}$ (이는 이 예에서 사운드 필드를 캡쳐하는 마이크로폰에 대한 구면 좌표들로 표현됨)에서의 압력 p_i 이 $SHC A_n^m(k)$ 에 의해 고유하게 표현될 수 있음을 보여준다.

여기서, $k = \frac{\omega}{c}$ 이고, c 는 음의 속도 (~ 343 m/s)이고, $\{r_r, \theta_r, \varphi_r\}$ 는 기준 포인트 (또는 관찰 포인트)이고, $j_n(\cdot)$ 는 차수 n 의 구면 베셀 함수 (spherical Bessel function)이고, $Y_n^m(\theta_r, \varphi_r)$ 는 차수 n 및 하위 차수 m 의 구면 조화 기저 함수들이다. 대괄호 내의 항은, 이산 푸리에 변환 (discrete Fourier transform; DFT), 이산 코사인 변환 (discrete cosine transform; DCT), 또는 웨이블렛 변환 (wavelet transform)과 같은 다양한 시간-주파수 변환들에 의해 근사화될 수 있는 신호 (즉, $S(\omega, r_r, \theta_r, \varphi_r)$)의 주파수-도메인 표현인 것이 인식될 수 있다. 계층적인 세트들의 다른 예들은 웨이블렛 변환 계수들의 세트들, 및 멀티해상도 기저 함수들의 다른 세트들을 포함한다.

[0019] 도 1 은 제로 차수 ($n = 0$)로부터 제 4 차수 ($n = 4$) 까지의 구면 조화 기저 함수들을 예시하는 도면이다. 알 수 있는 바와 같이, 각각의 차수에 대하여, 예시 목적들의 용이함을 위하여 도 1 의 예에서 도시되어 있지만, 명시적으로 표기되지는 않은 하위 차수들 m 의 확장이 있다.

[0020] 도 2 는 제로 차수 ($n = 0$)로부터 제 4 차수 ($n = 4$) 까지의 구면 조화 기저 함수들을 예시하는 또 다른 도면이다. 도 2 에서, 구면 조화 기저 함수들은 차수 및 하위 차수 양쪽 모두가 도시되어 있는 3 차원 좌표 공간에서 도시되어 있다.

[0021] 어느 경우에도, $SHC A_n^m(k)$ 는 다양한 마이크로폰 어레이 구성들에 의해 물리적으로 취득 (예컨대, 레코딩) 될 수 있거나, 또는 대안적으로, 이들은 사운드 필드의 채널-기반 또는 객체-기반 설명들로부터 유도될 수 있다. SHC 는 장면 기반 오디오이다. 예를 들어, 4차 SHC 표현은 이 시간 샘플 당 $(1+4)^2 = 25$ 계수들을 수반한다.

[0022] 이 SHC 들이 객체-기반 설명으로부터 어떻게 유도될 수도 있는지를 예시하기 위하여, 다음의 수학식을 고려하자. 개별적인 오디오 객체에 대응하는 사운드 필드에 대한 계수들 $A_n^m(k)$ 은 다음과 같이 표현될 수도 있다:

$$A_n^m(k) = g(\omega)(-4\pi ik)h_n^{(2)}(kr_s)Y_n^m(\theta_s, \varphi_s)$$

[0023] 여기서, i 는 $\sqrt{-1}$ 이고, $h_n^{(2)}(\cdot)$ 는 차수 n 의 (제 2 종류의) 구면 한켈 함수 (spherical Hankel function)

이고, $\{r_s, \theta_s, \varphi_s\}$ 는 객체의 위치이다. (예를 들어, PCM 스트림에 대해 고속 푸리에 변환을 수행하는 것과 같은 시간-주파수 분석 기법들을 이용하여) 소스 에너지 $g(\omega)$ 를 주파수의 함수로서 인지하는 것은 우리가 각각의 PCM 객체 및 그 위치를 $\text{SHC } A_n^m(k)$ 로 변환하도록 한다. 또한, (상기한 것이 선형 및 직교 분해 이므로) 각각의 객체에 대한 $A_n^m(k)$ 계수들이 부가적인 것으로 보여질 수 있다. 이러한 방식으로, 다수의 PCM 객체들은 (예컨대, 개별적인 객체들에 대한 계수 벡터들의 합으로서) $A_n^m(k)$ 계수들에 의해 표현될 수 있다. 본질적으로, 이 계수들은 사운드 필드에 대한 정보 (3D 좌표들의 함수로서의 압력) 를 포함하고, 상기 한 것은 관찰 포인트 $\{r_r, \theta_r, \varphi_r\}$ 근처에서, 개별적인 객체들로부터 전체적인 사운드 필드의 표현으로의 변환을 나타낸다.

[0025]

SHC 들은 또한, 다음과 같이 마이크로폰-어레이로부터 유도될 수도 있다:

[0026]

$$a_n^m(t) = b_n(r_i, t) * \langle Y_n^m(\theta_i, \varphi_i), m_i(t) \rangle$$

[0027]

여기서, $a_n^m(t)$ 는 $A_n^m(k)$ (SHC) 의 시간-도메인 등가물이고, $*$ 는 컨볼루션 연산 (convolution operation) 을 나타내고, \langle , \rangle 는 내적 (inner product) 을 나타내고, $b_n(r_i, t)$ 는 r_i 에 의존적인 시간-도메인 필터 함수를 나타내고, $m_i(t)$ 는 i 번째 마이크로폰 신호이며, 여기서, i 번째 마이크로폰 트랜스듀서는 반경 r_i , 고도각 (elevation angle) θ_i 및 방위각 (azimuth angle) φ_i 에서 위치된다. 따라서, 마이크로폰 어레이에서 32 개의 트랜스듀서들이 있고 각각의 마이크로폰이 (mhAcoustics로부터의 Eigenmike EM32 상의 트랜스듀서들과 같아) $r_i = a$ 가 상수가 되도록 구 (sphere) 상에 위치될 경우, 25 개의 SHC 들은 다음과 같이 행렬 연산을 이용하여 유도될 수도 있다:

$$\begin{bmatrix} a_0^0(t) \\ a_1^{-1}(t) \\ \vdots \\ a_4^4(t) \end{bmatrix} = \begin{bmatrix} b_0(a, t) \\ b_1(a, t) \\ \vdots \\ b_4(a, t) \end{bmatrix} * \begin{bmatrix} Y_0^0(\theta_1, \varphi_1) & Y_0^0(\theta_2, \varphi_2) & \dots & Y_0^0(\theta_{32}, \varphi_{32}) \\ Y_1^{-1}(\theta_1, \varphi_1)Y_1^{-1}(\theta_2, \varphi_2) & \dots & Y_1^{-1}(\theta_{32}, \varphi_{32}) \\ \vdots & \ddots & \vdots \\ Y_4^4(\theta_1, \varphi_1) & Y_4^4(\theta_2, \varphi_2) & \dots & Y_4^4(\theta_{32}, \varphi_{32}) \end{bmatrix} \begin{bmatrix} m_1(a, t) \\ m_2(a, t) \\ \vdots \\ m_{32}(a, t) \end{bmatrix}$$

[0028]

상기 수학식에서의 행렬은 더욱 일반적으로 $E_s(\theta, \varphi)$ 로서 지칭될 수도 있으며, 여기서, 아래첨자 s 는 행렬이 어떤 트랜스듀서 기하구조-세트 s 에 대한 것임을 표시할 수도 있다. (* 에 의해 표시된) 상기 수학식에서의 컨볼루션은, 예를 들어, 출력 $a_0^0(t)$ 이 $b_0(a, t)$ 와, $E_s(\theta, \varphi)$ 행렬의 제 1 행 (row) 및 (벡터 곱셈의 결과가 시계열인 사실을 고려하여, 시간의 함수로서 변동되는) 마이크로폰 신호들의 열 (column) 의 벡터 곱셈 (vector multiplication) 으로부터 기인하는 시계열과의 사이의 컨볼루션의 결과가 되도록, 행-바이-행 (row-by-row) 에 기초하고 있다. 연산은 마이크로폰의 트랜스듀서 포지션들이 (Eigenmike 트랜스듀서 기하구조와 매우 근접한) 소위 T-설계 기하구조들 내에 있을 때에 가장 정확할 수도 있다. T-설계 기하구조의 하나의 특성은, 기하구조로부터 기인하는 $E_s(\theta, \varphi)$ 행렬이 매우 양호하게 거동된 역 (또는 의사 역 (pseudo inverse)) 을 가지는 것과, 또한, 역이 종종 행렬 $E_s(\theta, \varphi)$ 의 병치 (transpose) 에 의해 매우 양호하게 근사화될 수도 있다는 것일 수도 있다. $b_n(a, t)$ 에 의한 필터링 동작이 무시되어야 할 경우, 이 속성은 SHC (즉, 이 예에서 $[m_i(t)] = [E_s(\theta, \varphi)]^{-1}[\text{SHC}]$)로부터의 마이크로폰 신호들의 복구를 허용할 수도 있다. 나머지 도면들은 SHC-기반 오디오-코딩의 문맥에서 이하에서 설명된다.

[0030]

도 3 은 오디오 신호 정보를 보다 효율적으로 렌더링하기 위해 본 개시물에 설명된 기법들을 수행할 수도 있는 시스템 (20) 을 예시하는 디아이그램이다. 도 3 의 예에서 도시된 바와 같이, 시스템 (20) 은 컨텐츠 크리에이터 (22) 및 컨텐츠 소비자 (24) 를 포함한다. 컨텐츠 크리에이터 (22) 및 컨텐츠 소비자 (24) 의 문맥에서 설명되었지만, 기법들은 SHC 들 또는 사운드 필드의 계층적 표현을 정의하는 임의의 다른 계층적 엘리먼트

들을 이용하는 임의의 상황에서 구현될 수도 있다.

[0031] 컨텐츠 크리에이터 (22) 는 컨텐츠 소비자 (24) 와 같은 컨텐츠 소비자들에 의한 소비를 위한 멀티-채널 오디오 컨텐츠를 생성할 수도 있는 영화 스튜디오 또는 다른 엔티티를 나타낼 수도 있다. 종종, 이 컨텐츠 크리에이터는 비디오 컨텐츠와 함께 오디오 컨텐츠를 생성한다. 컨텐츠 소비자 (24) 는, 멀티-채널 오디오 컨텐츠로서의 재생 가능한 오디오 재생 시스템의 임의의 형태를 지정할 수도 있는 오디오 재생 시스템을 소유하거나 이 시스템에 액세스하는 개인을 나타낼 수도 있다. 도 3 의 예에서, 컨텐츠 소비자 (24) 는 사운드 필드의 계층적 표현을 정의하는 계층적 엘리먼트들을 렌더링하기 위한 오디오 재생 시스템 (32) 을 소유하거나 이 시스템에 액세스한다.

[0032] 컨텐츠 크리에이터 (22) 는 오디오 렌더러 (28) 및 오디오 편집 시스템 (30) 을 포함한다. 오디오 렌더러 (28) 는 스피커 피드들 (또한, "라우드스피커 피드들", "스피커 신호들", 또는 "라우드스피커 신호들" 로서 지칭될 수도 있음) 을 렌더링하거나 이와 다르게 생성하는 오디오 프로세싱 유닛을 나타낼 수도 있다. 각각의 스피커 피드는 멀티-채널 오디오 시스템의 특별한 채널에 대한 사운드를 재생하는 스피커 피드에 대응하거나, 또는 스피커 포지션에 매칭하는 헤드-관련 전달 함수 (HRTF) 필터들과의 컨볼루션에 의도되는 가상 라우드스피커 피드에 대응할 수도 있다. 각각의 스피커 피드는 방향성 사운드 필드를 나타내기 위해 SHC들의 다중 채널들을 이용하는 구면 조화 계수들의 채널에 대응할 수도 있다 (여기에서 채널은 구면 조화 계수들이 대응하는 관련 구면 기반 함수들의 차수 및/또는 하위 차수에 의해 표기될 수 있다).

[0033] 도 3 의 예에서, 렌더러 (28) 는 기존의 5.1, 7.1 또는 22.2 서라운드 사운드 포맷들에 대한 스피커 피드들을 렌더링할 수도 있어서, 5.1, 7.1 또는 22.2 서라운드 사운드 스피커 시스템들에서 5, 7 또는 22 개의 스피커들의 각각에 대한 스피커 피드를 생성할 수도 있다. 대안적으로, 렌더러 (28) 는 위에서 논의된 소스 구면 조화 계수들의 속성들이 주어질 시에, 임의의 수의 스피커들을 가지는 임의의 스피커 구성에 대한 소스 구면 조화 계수들로부터 스피커 피드들을 렌더링하도록 구성될 수도 있다. 이러한 방식으로, 오디오 렌더러 (28) 는 도 3 에서 스피커 피드들 (29) 로서 나타내어지는 다수의 스피커 피드들을 생성할 수도 있다.

[0034] 컨텐츠 크리에이터는 편집 프로세스 동안에, 구면 조화 계수들 (27) ("SHC (27)") 을 렌더링할 수도 있어서, 고충실도 (high fidelity) 를 가지지 않거나 설득력 있는 서라운드 사운드 경험을 제공하지 않는 사운드 필드의 양태들을 식별하기 위한 시도로 렌더링된 스피커 피드들을 청취할 수도 있다. 그 후, 컨텐츠 크리에이터 (22) 는 (소스 구면 조화 계수들이 위에서 설명된 방식으로 유도될 수도 있는 상이한 객체들의 조작을 통해 종종 간접적으로) 소스 구면 조화 계수들을 편집할 수도 있다. 컨텐츠 크리에이터 (22) 는 구면 조화 계수들 (27) 을 편집하기 위하여 오디오 편집 시스템 (30) 을 채용할 수도 있다. 오디오 편집 시스템 (30) 은 오디오 데이터를 편집할 수 있으며 이 오디오 데이터를 하나 이상의 소스 구면 조화 계수들로서 출력할 수 있는 임의의 시스템을 나타낸다.

[0035] 편집 프로세스가 완료될 때, 컨텐츠 크리에이터 (22) 는 구면 조화 계수들 (27) 에 기초하여 비트스트림 (31) 을 생성할 수도 있다. 즉, 컨텐츠 크리에이터 (22) 는 비트스트림 (31) 을 생성할 수 있는 임의의 디바이스를 나타낼 수도 있는 비트스트림 생성 디바이스 (36) 를 포함한다. 일부의 사례들에서, 비트스트림 생성 디바이스 (36) 는, (하나의 예로서 엔트로피 인코딩을 통해) 구면 조화 계수들 (27) 을 대역폭 압축하며 비트스트림 (31) 을 형성하기 위하여 수용된 포맷으로 구면 조화 계수들 (27) 의 엔트로피 인코딩된 버전을 배열하는 인코더를 나타낼 수도 있다. 다른 사례들에서, 비트스트림 생성 디바이스 (36) 는, 하나의 예로서, 멀티-채널 오디오 컨텐츠 또는 그 파생물 (derivative) 들을 압축하기 위하여 기존의 오디오 서라운드 사운드 인코딩 프로세스들의 것들과 유사한 프로세스들을 이용하여 멀티-채널 오디오 컨텐츠 (29) 를 인코딩하는 오디오 인코더 (아마도, MPEG 서라운드 또는 그 파생물과 같은 알려진 오디오 코딩 표준을 준수하는 것) 를 나타낼 수도 있다.

다음으로, 압축된 멀티-채널 오디오 컨텐츠 (29) 는 컨텐츠 (29) 를 대역폭 압축하기 위한 일부의 다른 방법으로 엔트로피 인코딩 또는 코딩될 수도 있고, 비트스트림 (31) 을 형성하기 위하여 합의된 포맷에 따라 배열될 수도 있다. 비트스트림 (31) 을 형성하기 위하여 직접적으로 압축되든, 또는 비트스트림 (31) 을 형성하기 위하여 렌더링되고, 그 다음에 압축되든 간에, 컨텐츠 크리에이터 (22) 는 비트스트림 (31) 을 컨텐츠 소비자 (24) 에게 송신할 수도 있다.

[0036] 도 3 에서, 컨텐츠 소비자 (24) 에게 직접적으로 송신되는 것으로 도시되어 있지만, 컨텐츠 크리에이터 (22) 는 비트스트림 (31) 을, 컨텐츠 크리에이터 (22) 및 컨텐츠 소비자 (24) 사이에 위치된 중간 디바이스로 출력할 수도 있다. 이 중간 디바이스는 이 비트스트림을 요청할 수도 있는 컨텐츠 소비자 (24) 로의 더 이후의 전달을 위해 비트스트림 (31) 을 저장할 수도 있다. 중간 디바이스는 파일 서버, 웹 서버, 데스크톱 컴퓨터, 랩

톱 컴퓨터, 태블릿 컴퓨터, 이동 전화, 스마트폰, 또는 오디오 디코더에 의해 더 이후의 추출 (retrieval) 을 위해 비트스트림 (31) 을 저장할 수 있는 임의의 다른 디바이스를 포함할 수도 있다. 이 중간 디바이스는 비트스트림 (31) 을 요청하는 컨텐츠 소비자 (24) 와 같은 가입자들에게 비트스트림 (31) 을 스트리밍 (그리고 아마도, 대응하는 비디오 데이터 비트스트림을 송신하는 것과 함께) 할 수 있는 컨텐츠 전달 네트워크에서 존재 할 수도 있다. 대안적으로, 컨텐츠 크리에이터 (22) 는 비트스트림 (31) 을, 컴팩트 디스크, 디지털 비디오 디스크, 고해상도 비디오 디스크 또는 다른 저장 매체들과 같은 저장 매체에 저장할 수도 있고, 이 매체들의 대부분은 컴퓨터에 의해 판독될 수 있고, 그러므로, 컴퓨터-판독가능한 저장 매체들 또는 비-일시적인 (non-transitory) 컴퓨터-판독가능한 저장 매체들로서 지칭될 수도 있다. 이 문맥에서, 송신 채널은 이 매체들에 저장된 컨텐츠가 송신되는 (그리고 소매점들 및 다른 점포-기반 전달 메커니즘을 포함할 수도 있는) 그러한 채널들을 지칭할 수도 있다. 어느 경우에도, 이 개시물의 기법들은 그러므로 도 3 의 예에 대해 이와 관련하여 제한되지 않아야 한다.

[0037] 도 3 의 예에서 추가로 도시된 바와 같이, 컨텐츠 소비자 (24) 는 오디오 재생 시스템 (32) 을 소유하거나 또는 달리 액세스한다. 오디오 재생 시스템 (32) 은 멀티-채널 오디오 데이터를 재생할 수 있는 임의의 오디오 재생 시스템을 나타낼 수도 있다. 오디오 재생 시스템 (32) 은 바이노럴 스피커 피드들 (35A-35B)(총괄적으로 "스피커 피드들 (35)" 로서 지칭됨) 로서의 출력을 위하여 SHCs (27') 을 렌더링하는 바이노럴 오디오 렌더러 (34) 를 포함한다. 바이노럴 오디오 렌더러 (34) 는 상이한 형태들의 렌더링, 이를 테면, 벡터 기반 진폭 폐닝 (VBAP) 을 수행하는 여러 방식들 중 하나 이상 및/또는 사운드 필드 합성을 수행하는 여러 방식들 중 하나 이상을 제공할 수도 있다.

[0038] 오디오 재생 시스템 (32) 은 추출 디바이스 (38) 를 더 포함할 수도 있다. 추출 디바이스 (38) 는 비트스트림 생성 디바이스 (36) 의 것과 일반적으로 역할 수도 있는 프로세스를 통하여 구면 조화 계수들 (27') ("SHCs (27')") 는 구면 조화 계수들 (27) 의 변경된 형태 또는 복제본을 표현할 수도 있음) 을 추출가능한 임의의 디바이스를 표현할 수도 있다. 어느 경우에도, 오디오 재생 시스템 (32) 은 구면 조화 계수들 (27') 을 수신할 수도 있고 바이노럴 오디오 렌더러 (34) 를 이용하여 구면 조화 계수들 (27') 을 렌더링하고 이에 의해 (설명의 용이성을 위하여 도 3 의 예에 도시되지 않은 오디오 재생 시스템 (32) 에 전기적으로 또는 가능하다면 무선으로 커플링된 복수의 라우드스피커들에 대응하는) 스피커 피드들 (35) 을 생성한다. 복수의 스피커 피드들 (35) 은 2 개일 수도 있고, 오디오 재생 시스템은 2 개의 대응하는 라우드스피커들을 포함하는 한 쌍의 헤드폰들에 무선으로 커플링할 수도 있다. 그러나, 여러 경우들에서, 바이노럴 오디오 렌더러 (34) 는 도 3 을 참조로 주로 설명되고 예시된 것보다 더 많거나 또는 더 적은 수의 스피커 피드들을 출력할 수도 있다.

[0039] 오디오 재생 시스템의 바이노럴 툼 임펄스 응답 (BRIR) 필터들 (37) 은 각각 소정 위치에서의 응답을 임펄스 위치에서 생성된 임펄스로 표현한다. BRIR 필터들 (37) 은 이들이 각각 그 위치에서 사람의 귀에 의해 경험되는 임펄스 응답을 나타내도록 생성된다는 점에서 "바이노럴"이다. 따라서, BRIR 필터들은 페어들로 렌더링 하는 사운드로 종종 생성되어 이용되며 여기에서 페어들 중 한 성분은 좌측 귀에 대한 것이고 페어들 중 다른 성분은 우측 귀에 대한 것이다. 예시된 예에서, 바이노럴 오디오 렌더러 (34) 는 좌측 BRIR 필터들 (33A) 및 우측 BRIR 필터들 (33B) 을 이용하여 각각의 바이노럴 오디오 출력들 (35A 및 35B) 을 렌더링한다.

[0040] 예를 들어, BRIR 필터들 (37) 은 임펄스 응답들 (IRs) 로서 측정된 헤드-관련 전달 함수들 (HRTFs) 로 사운드 소스 신호를 컨볼빙 (convolving) 함으로써 생성될 수도 있다. BRIR 필터들 (37) 각각에 대응하는 임펄스 위치는 가상 공간에서의 가상 라우드스피커의 포지션을 표현할 수도 있다. 일부 예들에서, 바이노럴 오디오 렌더러 (34) 는 SHCs (27') 를 가상 라우드스피커들에 대응하는 BRIR 필터들 (37) 로 컨볼빙한 다음, 스피커 피드들 (35) 로서 출력하기 위하여 SHCs (27') 에 의해 정의되는 사운드 필드를 렌더링하기 위해 결과적인 컨볼루션들을 누적 (즉, 합산) 한다. 여기에 설명된 바와 같이, 바이노럴 오디오 렌더러 (34) 는 SHCs (27') 를 스피커 피드들 (35) 로서 렌더링하면서 BRIR 필터들 (37) 를 조작함으로써 렌더링 연산을 감소시키기 위한 기법들을 적용할 수도 있다.

[0041] 일부 경우들에서, 기법들은 BRIR 필터들 (37) 을, 툴 내의 위치에서 임펄스 응답의 상이한 스테이지들을 표현하는 복수의 세그먼트들로 세그먼트하는 것을 포함한다. 이들 세그먼트들은 사운드 필드에 대한 임의의 지점에서의 압력 (또는 이것의 부재) 을 생성하는 상이한 물리적 현상에 대응한다. 예를 들어, BRIR 필터들 (37) 각각이 임펄스와 시간 일치되기 때문에, 첫번째 또는 "초기" 세그먼트는 임펄스 위치로부터의 압력파가, 임펄스 응답이 측정되는 위치에 도달할 때까지의 시간을 표현할 수도 있다. 타이밍 정보를 제외하고는, 각각의 초기 세그먼트들에 대한 BRIR 필터들 (37) 값들은 무의미할 수도 있고, 사운드 필드를 기술하는 계층적 엘리먼트들과의 컨볼루션으로부터 배제될 수도 있다. 이와 유사하게, BRIR 필터들 (37) 각각은 예를 들어, 지

정된 임계값 미만으로 감쇠되거나 또는 인간의 청각의 동작 범위 미만으로 감소되는 임펄스 응답 신호들을 포함하는 마지막 또는 "테일" 세그먼트를 포함할 수도 있다. 각각의 테일에 대한 BRIR 필터들 (37) 값들은 또한 무의미할 수도 있고 사운드 필드를 기술하는 계층적 엘리먼트들의 컨볼루션으로부터 배제될 수도 있다. 일부 예들에서, 기술들은 지정된 임계값과의 Schroeder 역진 적분 (Schroeder backward integration)을 수행하고 역진 적분이 지정된 임계값을 초과하는 경우 테일 세그먼트로부터 엘리먼트들을 폐기함으로써 테일 세그먼트를 결정하는 단계를 포함할 수도 있다. 일부 예들에서, 지정된 임계값은 반향 시간 (RT_{60})에 대해 -60 dB 이다.

[0042] BRIR 필터들 (37) 각각의 추가적인 세그먼트는 룸으로부터의 에코 효과들의 포함없이 임펄스 생성된 압력에 의해 야기되는 임펄스 응답을 표현할 수도 있다. 이를 세그먼트들은 BRIR 필터들 (37)에 대한 헤드-관련 전달 함수들 (HRTFs)로서 표현 및 기술될 수도 있고, 여기에서, HRTFs는 압력파가 고막을 향하여 진행할 때 헤드, 어깨/몸체 및 외이 주변의 압력 파들의 회절 및 반사로 인한 임펄스 응답을 캡처한다. HRTF 임펄스 응답들은 선형 및 시간 불변 시스템 (LTI)의 결과물이고, 최소 페이즈 필터들로서 모델링될 수도 있다. 일부 예들에서, 렌더링 동안에 HRTF 세그먼트 연산을 감소시키는 기법들은 오리지널 유한 임펄스 응답 (FIR) 필터의 차수 (예를 들어, HRTF 필터 세그먼트)를 감소시키기 위해 최소 페이즈 재구성 및 무한 임펄스 응답 (IIR) 필터들의 이용을 포함할 수도 있다.

[0043] IIR 필터들로서 구현되는 최소 페이즈 필터들은 감소된 필터 차수를 갖는 BRIR 필터들 (37)에 대해 HRTF 필터들을 근사화시키는데 이용될 수도 있다. 차수를 감소시키는 것은 주파수 도메인에서 시간 스텝에 대한 계산들의 수에서의 수반되는 감소로 귀결된다. 추가로, 최소 페이즈 필터들의 구성으로부터 야기되는 잔여/과다의 필터는 사운드 압력파가 소스로부터 각각의 귀로 진행하는 거리에 의해 야기되는 시간 또는 페이즈 거리를 표현하는 양이간 시간 차이 (ITD; interaural time difference)를 추정하는데 이용될 수도 있다. ITD는 하나 이상의 BRIR 필터들 (37)의, 사운드 필드를 기술하는 계층적 엘리먼트와의 컨볼루션을 연산 (즉, 바이노럴화를 결정) 한 후에 한쪽 귀 또는 양쪽 귀들에 대한 모델 사운드 국부화에 이용될 수도 있다.

[0044] BRIR 필터들 (37) 각각의 더 추가의 세그먼트는 HRTF 세그먼트에 후속하고, 임펄스 응답에 대한 룸의 효과들을 고려할 수도 있다. 이 룸 세그먼트는 조기 에코들 (또는 "조기 반사") 세그먼트 및 후기 반향 세그먼트로 추가로 분해될 수도 있다 (즉, 조기 에코들 및 후기 반향이 BRIR 필터들 (37) 각각의 별도의 세그먼트들에 의해 각각 표현될 수도 있다). HRTF 데이터가 BRIR 필터들 (37)에 이용가능한 경우, 조기 에코 세그먼트의 온셋 이, HRTF 와의 BRIR 필터들 (37)의 디컨볼루션에 의해 식별되어 HRTF 세그먼트가 식별될 수도 있다. 조기 에코 세그먼트가 HRTF 세그먼트에 후속한다. 잔여 룸 응답과 달리, HRTF 및 조기 에코 세그먼트들은 대응하는 가상 스피커의 위치가 유의적 관점에서 신호를 결정한다는 점에서 방향 의존성이다.

[0045] 일부 예들에서, 바이노럴 오디오 렌더러 (34)는 구면 조화 도메인 (θ, ϕ) 또는 사운드 필드를 기술하는 계층적 엘리먼트들에 대한 다른 도메인에 대해 준비되는 BRIR 필터들 (37)를 이용한다. 즉, BRIR 필터들 (37)은 바이노럴 오디오 렌더러 (34)가 BRIR 필터들 (37) (예를 들어, 좌측/우측)의 그리고 SHCs (27')의 대칭성을 포함하는 데이터 세트의 특정 특성들의 이점을 취하면서 고속 컨볼루션을 수행하도록 허용하기 위해, 변환된 BRIR 필터들 (37)로서 구면 조화 도메인 (SHD)에서 정의될 수도 있다. 이러한 예들에서, 변환된 BRIR 필터들 (37)은 SHC 렌더링 행렬과 오리지널 BRIR 필터들을 곱함으로써 (또는 시간 도메인에서 컨볼빙함으로써) 생성될 수도 있다. 수학적으로, 이는 다음 식 (1)-(5)에 따라 표현될 수 있다:

$$BRIR'_{(N+1)^2, L, left} = SHC_{(N+1)^2, L} * BRIR_{L, left} \quad (1)$$

$$BRIR'_{(N+1)^2, L, right} = SHC_{(N+1)^2, L} * BRIR_{L, right} \quad (2)$$

[0046] 또는

$$BRIR'_{(N+1)^2, L, right} = \begin{bmatrix} Y_0^0(\theta_1, \varphi_1) & Y_0^0(\theta_2, \varphi_2) & \dots & Y_0^0(\theta_L, \varphi_L) \\ Y_1^{-1}(\theta_1, \varphi_1) & Y_1^{-1}(\theta_2, \varphi_2) & \dots & Y_1^{-1}(\theta_L, \varphi_L) \\ \vdots & \vdots & \ddots & \vdots \\ Y_4^4(\theta_1, \varphi_1) & Y_4^4(\theta_2, \varphi_2) & \dots & Y_4^4(\theta_L, \varphi_L) \end{bmatrix} \begin{bmatrix} B_0 \\ B_1 \\ \vdots \\ B_L \end{bmatrix}^T \quad (3)$$

[0048]

$$BRIR''_{(N+1)^2, left} = \sum_{k=0}^{L-1} [BRIR'_{(N+1)^2, k, left}] \quad (4)$$

$$BRIR''_{(N+1)^2, right} = \sum_{k=0}^{L-1} [BRIR'_{(N+1)^2, k, right}] \quad (5)$$

[0049]

[0050] 여기에서 (3) 은 (1) 또는 (2) 를 4차 구면 조화 계수들에 대한 행렬 형태로 기술한다 (이는 4차 이하의 구면 기저 함수들과 연관된 구면 조화 계수들의 것들을 지칭하는 대안의 방식일 수도 있다). 식 (3) 은 물론, 고차 또는 저차 구면 조화 계수들에 대해 변경될 수도 있다. 식들 (4)-(5) 은 합산된 SHC-바이노럴 렌더링 행렬들 ($BRIR''$) 을 생성하기 위해 라우드스피커 차원 (L) 에 걸쳐 변환된 좌측 및 우측 BRIR 필터 (37) 의 합산을 묘사한다. 결합하여, 합산된 SHC-바이노럴 행렬들은 차원성 $[(N+1)^2, Length, 2]$ 을 갖고, 여기에서 $Length$ 는 식들 (1)-(5) 의 임의의 조합이 적용될 수도 있는 임펄스 응답 벡터들의 길이이다. 식들 (1) 및 (2) 의 일부 경우들에서, 렌더링 행렬 (SHC) 은 식 (1) 이 $BRIR'_{(N+1)^2, L, left} = SHC_{(N+1)^2, L, left} * BRIR_{L, left}$ 로 변경될 수도 있도록 그리고 식 (2) 가 $BRIR'_{(N+1)^2, L, right} = SHC_{(N+1)^2, L} * BRIR_{L, right}$ 로 변경될 수도 있도록 바이노럴화될 수도 있다.

[0051]

위의 식들 (1)-(3) 에서 제시된 SHC 렌더링 행렬 (SHC) 은 개별적인 SHC 채널을 효과적으로 정의하는 SHCs (27') 의 각각의 차수/하위 차수 조합에 대한 엘리먼트들을 포함하며, 여기에서 엘리먼트 값들은 구면 조화 도메인에서 스피커에 대한 포지션 (L) 에 대하여 설정된다. $BRIR_{L, left}$ 는 좌측 귀 또는 스피커에 대한 위치 (L) 에서 생성되는 임펄스에 대한 포지션에서의 BRIR 응답을 표현하며, 임펄스 응답 벡터들 B_i (여기에서 $\{i | i \in [0, L]\}$) 을 이용하여 식 (3) 으로 묘사된다. $BRIR'_{(N+1)^2, L, left}$ 는 구면 조화 도메인으로 변환된, "SHC-바이노럴 렌더링 행렬"의 1/2, 즉, 좌측 귀 또는 스피커들에 대한 포지션 (L) 에서 생성되는 임펄스에 대한 포지션에서의 SHC-바이노럴 렌더링 행렬을 표현한다. $BRIR'_{(N+1)^2, L, right}$ 는 SHC-바이노럴 렌더링 행렬의 다른 1/2 를 표현한다.

[0052]

일부 예들에서, 기법들은 SHC 렌더링 행렬을 HRTF 및 개별적인 오리지널 BRIR 필터들 (37) 의 조기 반사 세그먼트들에만 적용하여 변환된 BRIR 필터들 (37) 및 SHC-바이노럴 렌더링 행렬을 생성하는 것을 포함할 수도 있다. 이는 SHCs (27') 와의 컨볼루션들의 길이를 감소시킬 수도 있다.

[0053]

일부 예들에서, 식들 (4)-(5) 에서 묘사된 바와 같이, 구면 조화 도메인에서 여러 라우드스피커들을 통합하는 차원성을 갖는 SHC-바이노럴 렌더링 행렬들은 합산되어, SHC 렌더링 및 BRIR 렌더링/혼합을 결합한 $(N+1)^2 * Length * 2$ 필터 행렬이 생성될 수도 있다. 즉, L 개의 라우드스피커들 각각에 대한 SHC-바이노럴 렌더링 행렬들은 예를 들어, L 차원에 걸친 계수들을 합산함으로써 결합될 수도 있다. 길이 ($Length$) 의 SHC-바이노럴 렌더링 행렬에서, 이는 신호를 바이노럴화하기 위해 구면 조화 계수들의 오디오 신호에 적용될 수도 있는 $(N+1)^2 * Length * 2$ 합산된 SHC-바이노럴 렌더링 행렬을 생성한다. $Length$ 는 여기에 설명된 기법들에 따라 세그먼트된 BRIR 필터들의 길이일 수도 있다.

[0054]

모델 감소에 대한 기법들은 또한 변경된 렌더링 필터들에 적용될 수도 있고, 변경된 렌더링 필터들은 SHCs (27') (예를 들어, SHC 컨텐츠들) 가 새로운 필터 행렬 (합산된 SHC-바이노럴 렌더링 행렬) 로 직접 필터링되는 것을 허용한다. 그 후, 바이노럴 오디오 렌더러 (34) 는 바이노럴 출력 신호들 (35A, 35B) 을 획득하기 위해 필터링된 어레이들을 합산함으로서 바이노럴 오디오로 변환할 수도 있다.

[0055]

일부 예들에서, 오디오 재생 시스템 (32) 의 BRIR 필터들 (37) 은 상술한 기법들 중 임의의 하나 이상에 따라 이전에 연산된 구면 조화 도메인에서의 변환된 BRIR 필터들을 표현한다. 일부 예들에서, 오리지널 BRIR 필터들 (37) 의 변환은 런타임으로 수행될 수도 있다.

[0056]

일부 예들에서, BRIR 필터들 (37) 이 통상적으로 대칭적이기 때문에, 기법들은 좌측 또는 우측 귀에 대해 SHC-바이노럴 렌더링 행렬만을 이용함으로써 바이노럴 출력들 (35A, 35B) 의 연산의 추가적인 감소를 촉진할 수도 있다. 필터 행렬에 의해 필터링된 SHCs (27') 를 합산할 때, 바이노럴 오디오 렌더러 (34) 는 최종 출력을 렌더링할 때의 제 2 채널로서 출력 신호 (35A 또는 35B) 에 대한 조건적 결정들을 행할 수도 있다. 여기에 설명된 바와 같이, 좌측 또는 우측 귀에 대하여 기술되는 렌더링 행렬들을 변경하는 것 또는 컨텐츠를 프로세싱

하는 것에 대한 참조는 다른 귀에도 동일하게 적용가능함을 이해하여야 한다.

[0057] 이 방식으로 기법들은 다수의 채널들과 배제된 BRIR 필터 샘플들의 직접 컨볼루션을 가능성있게 회피하기 위하여 BRIR 필터들 (37)의 길이를 감소시키는 다수의 접근 방식들을 제공할 수도 있다. 그 결과, 바이노럴 오디오 웨더러 (34)는 SHCs (27')로부터의 바이노럴 출력 신호들 (35A, 35B)의 효율적인 웨더링을 제공할 수도 있다.

[0058] 도 4는 예시적인 바이노럴 룸 임펄스 응답 (BRIR)을 예시하는 블록도이다. BRIR (40)은 5개의 세그먼트들 (42A-42E)을 예시한다. 초기 세그먼트 (42A) 및 테일 세그먼트 (42E) 양쪽 모두는 무의미할 수도 있고 웨더링 연산으로부터 배제될 수도 있는 침묵 (quiet) 샘플들을 포함한다. 헤드-관련 전달 함수 (HRTF) 세그먼트 (42B)는 헤드-관련 전달로 인한 임펄스 응답을 포함하며, 여기에 설명된 기법들을 이용하여 식별될 수도 있다. 초기 에코들 (대안으로서, "초기 반사들") 세그먼트 (42C) 및 후기 룸 리버브 세그먼트 (42D)는 HRTF와 룸 효과들을 결합하며, 즉 초기 에코들 세그먼트 (42C)의 임펄스 응답은 룸의 초기 에코들과 후기 반향에 의해 필터링된 BRIR (40)의 HRTF의 것과 매칭한다. 그러나, 초기 에코들 세그먼트 (42C)는 후기 룸 리버브 세그먼트 (42D)에 비해 더 이산적인 에코들을 포함할 수도 있다. 혼합 시간은 초기 에코들 세그먼트 (42C)와 후기 룸 리버브 세그먼트 (42D) 사이의 시간이며 초기 에코들이 고밀도 리버브가 되는 시간을 표시한다. 혼합 시간은 HRTF로 대략 1.5×10^4 샘플들 또는 HRTF 세그먼트 (42B)의 온셋에서부터 대략 7.0×10^4 개의 샘플들에서 발생하는 것으로서 예시되어 있다. 일부 예들에서, 기법들은 룸 블루프로부터 통계 데이터 및 추정값을 이용하여 혼합 시간을 연산하는 것을 포함한다. 일부 예들에서, 50% 신뢰 구간 (t_{mp50})을 가진 지각적 혼합 시간은 대략 36 밀리초 (ms)이고 95% 신뢰 구간 (t_{mp95})을 가진 지각적 혼합 시간은 대략 80 밀리초이다. 일부 예들에서, BRIR (40)에 대응하는 필터의 후기 룸 리버브 세그먼트 (42D)는 코히어런스 매칭된 잡음 테일을 이용하여 합성될 수도 있다.

[0059] 도 5는 룸에서 BRIR, 이를 테면, 도 4의 BRIR (40)을 생성하기 위한 예시적인 시스템 모델 (50)을 예시하는 블록도이다. 모델은 케스케이드된 시스템들, 여기에서 룸 (52A) 및 HRTF (52B)를 포함한다. HRTF (52B)가 임펄스에 적용된 후, 임펄스 응답은 룸 (52A)의 초기 에코들에 의해 필터링된 HRTF의 것에 매칭한다.

[0060] 도 6은 룸에서 BRIR, 이를 테면, 도 4의 BRIR (40)을 생성하기 위한 보다 세밀한 시스템 모델 (60)을 예시하는 블록도이다. 이 모델 (60)은 케스케이드된 시스템들, 여기에서 HRTF (62A), 초기 에코들 (62B), 및 (HRTF 및 룸 에코들을 결합하는) 잔여 룸 (62C)을 포함한다. 모델 (60)은 룸 (52A)의, 초기 에코들 (62B) 및 잔여 룸 (62C)으로의 분해를 예시하며 각각의 시스템 (62A, 62B, 62C)을 선형 시간 불변으로서 처리한다.

[0061] 초기 에코들 (62B)은 잔여 룸 (62C)보다 더 이산적인 에코들을 포함한다. 따라서, 초기 에코들 (62B)은 가상 스피커 채널마다 변할 수도 있는 한편, 더 긴 테일을 갖는 잔여 룸 (62C)은 단일의 스테레오 카페본으로서 합성될 수도 있다. BRIR을 획득하는데 이용되는 일부 측정 모델들에서, HRTF 데이터는 무반향성 챕버에서 측정된 것으로서 이용가능할 수도 있다. 초기 에코들 (62B)은 ("반사들"로서 지칭될 수도 있는) 초기 에코들의 위치를 식별하기 위해 BRIR 및 HRTF 데이터를 디컨볼루션함으로써 결정될 수도 있다. 일부 예들에서, HRTF 데이터는 용이하게 이용가능하지 않고 초기 에코들 (62B)을 식별하기 위한 기법들은 블라인드 추정을 포함한다. 그러나, 간단한 접근 방식은 첫번째 수 밀리초 (예를 들어, 첫번째 5, 10, 15, 또는 20 ms)를 HRTF에 의해 필터링된 직접 임펄스로서 평가하는 것을 포함할 수도 있다. 위에 주지된 바와 같이, 기법들은 룸 블루프로부터 통계 데이터 및 추정값을 이용하여 혼합 시간을 연산하는 것을 포함할 수도 있다.

[0062] 일부 예들에서, 기법들은 잔여 룸 (62C)에 대한 하나 이상의 BRIR 필터들을 합성하는 것을 포함할 수도 있다. 혼합 시간 이후, BRIR 리버브 테일들 (도 6에서 시스템 잔여 룸 (62C)으로서 표현됨)은 지각적 부자극 없이 일부 예들에서 상호변경될 수 있다. 추가로, BRIR 리버브 테일들은 EDR (Energy Decay Relief) 및 FDIC (Frequency-Dependent Interaural Coherence)에 매칭하는 가우시안 백색 잡음으로 합성될 수 있다. 일부 예들에서, 공통 합성 BRIR 리버브 테일은 BRIR 필터들에 대해 생성될 수도 있다. 일부 예들에서, 공통 EDR은 모든 스피커들에 대한 EDRs의 평균일 수도 있거나 또는 평균 에너지에 매칭하는 에너지를 가진 프론트 제로도 EDR일 수도 있다. 일부 예들에서, FDIC는 모든 스피커들에 걸친 평균 FDIC일 수도 있거나 또는 공간감 (spaciousness)에 대한 최대 비상관화된 측정을 위해 모든 스피커들에 걸친 최소값일 수도 있다. 일부 예들에서, 리버브 테일들은 또한 피드백 자연 네트워크 (FDN)에 의한 인위적인 리버브로 시뮬레이션될 수 있다.

- [0063] 공통 리버브 테일에서, 대응하는 BRIR 필터의 나중 부분은 각각의 스피커 피드와의 별도의 컨볼루션으로부터 배제될 수도 있지만, 그 대신에 모든 스피커 피드들의 혼합에 대해 한번 적용될 수도 있다. 위에 설명된 바와 같이, 그리고 아래 보다 자세히 설명될 바와 같이, 모든 스피커 피드들의 혼합은 구면 조화 계수들의 신호 렌더링으로 추가로 간략화될 수 있다.
- [0064] 도 7 은 본 개시물에 설명된 바이노럴 오디오 렌더링 기법들의 여러 양태들을 수행할 수도 있는 오디오 재생 디바이스의 일 예를 예시하는 블록도이다. 도 7 의 예에서 단일 디바이스, 즉 오디오 재생 디바이스 (100) 로서 예시되어 있지만, 기법들은 하나 이상의 디바이스들에 의해 수행될 수도 있다. 따라서, 기법들은 이러한 면으로 제한되지 않아야 한다.
- [0065] 도 7 의 예에서 도시된 바와 같이, 오디오 재생 디바이스 (100) 는 바이노럴 렌더링 유닛 (102) 및 추출 유닛 (104) 을 포함할 수도 있다. 추출 유닛 (104) 은 비트스트림 (120) 으로부터 인코딩된 오디오 데이터를 출력하도록 구성되는 유닛을 나타낼 수도 있다. 추출 유닛 (104) 은 추출된 인코딩된 오디오 데이터를 구면 조화 계수들 (SHCs) (122) (이는 또한 SHCs (122) 가 1 보다 큰 차수와 연관된 적어도 하나의 계수를 포함할 수도 있다는 점에서 고차 엠비소닉 (HOA) 으로 지칭될 수도 있음) 의 형태로 바이노럴 렌더링 유닛 (146) 에 포워딩할 수도 있다.
- [0066] 일부 예들에서, 오디오 재생 디바이스 (100) 는 SHCs (122) 를 생성하도록 인코딩된 오디오 데이터를 디코딩하도록 구성된 오디오 디코딩 유닛을 포함한다. 오디오 디코딩 유닛은 일부 양태들에서, SHCs (122) 를 인코딩하는데 이용되는 오디오 인코딩 프로세스에 역이 되는 오디오 디코딩 프로세스를 수행할 수도 있다. 오디오 디코딩 유닛은 인코딩된 오디오 데이터의 SHCs 를 시간 도메인으로부터 주파수 도메인으로 변환하여 SHCs (122) 를 생성하도록 구성된 시간-주파수 분석 유닛을 포함할 수도 있다. 즉, 인코딩된 오디오 데이터가 시간 도메인으로부터 주파수 도메인으로 변환되지 않는 SHC (122) 의 압축된 형태를 표현할 때, 오디오 디코딩 유닛은 시간-주파수 분석 유닛을 인보크하여, (주파수 도메인에서 특정된) SHCs (122) 를 생성하도록 SHCs 들을 시간 도메인으로부터 주파수 도메인으로 변환할 수도 있다. 시간-주파수 분석 유닛은 몇몇 예를 제공하면 고속 푸리에 변환 (FFT), 이산 코사인 변환 (DCT), 수정된 이산 코사인 변환 (MDCT), 및 이산 사인 변환 (DST) 을 포함하는 임의의 형태의 푸리에 기반 변환을 적용하여, SHCs 를 시간 도메인으로부터 주파수 도메인에서의 SHCs (122) 로 변환할 수도 있다. 일부 경우들에서, SHCs (122) 는 비트스트림 (120) 으로 주파수 도메인에서 이미 특정될 수도 있다. 이들 경우들에서, 시간-주파수 분석 유닛은 변환을 적용하거나 수신된 SHCs (122) 를 달리 변환하는 것이 없이 바이노럴 렌더링 유닛 (102) 으로 SHCs (122) 를 패스시킬 수도 있다. 주파수 도메인에서 특정된 SHCs (122) 에 관하여 설명되어 있지만, 기법들은 시간 도메인에서 특정된 SHCs (122) 에 관하여 수행될 수도 있다.
- [0067] 바이노럴 렌더링 유닛 (102) 은 SHCs (122) 를 바이노럴화하도록 구성된 유닛을 나타낸다. 즉, 바이노럴 렌더링 유닛 (102) 은 SHCs (122) 이 기록되었던 룸에서의 청취자에 의해 좌측 및 우측 채널이 청취되는 방법을 모델링하기 위해 공간화를 피쳐링할 수도 있는, 좌측 및 우측 채널로 SHCs (122) 를 렌더링하도록 구성된 유닛을 표현할 수도 있다. 바이노럴 렌더링 유닛 (102) 은 헤드셋, 이를 테면, 헤드폰들을 통하여 재생에 적합한 좌측 채널 (136A) 및 우측 채널 (136B)(이들은 총괄적으로 "채널 (136)" 로서 지칭될 수도 있음) 을 생성하기 위해 SHCs (122) 를 렌더링할 수도 있다. 도 7 의 예에 도시된 바와 같이, 바이노럴 렌더링 유닛 (102) 은 BRIR 필터들 (108), BRIR 컨디셔닝 유닛 (106), 잔여 룸 응답 유닛 (110), BRIR SHC-도메인 변환 유닛 (112), 컨볼루션 유닛 (114) 및 결합 유닛 (116) 을 포함한다.
- [0068] BRIR 필터들 (108) 은 하나 이상의 BRIR 필터들을 포함하며, 도 3 의 BRIR 필터들 (37) 의 일 예를 나타낼 수도 있다. BRIR 필터들 (108) 은 개별적인 BRIRs 에 대한 좌측 및 우측 HRTF 의 효과를 표현하는 별도의 BRIR 필터들 (126A, 126B) 을 포함할 수도 있다.
- [0069] BRIR 컨디셔닝 유닛 (106) 은 가상 라우드스피커 (L) 에 대하여 하나씩, BRIR 필터들 (126A, 126B) 의 L 개의 인스턴스들을 수신하며 여기에서 각각의 BRIR 필터는 길이 (N) 를 갖는다. BRIR 필터들 (126A, 126B) 은 침묵 샘플들을 제거하도록 이미 컨디셔닝될 수도 있다. BRIR 컨디셔닝 유닛 (106) 은 위에 설명된 기법들을 적용하여 BRIR 필터들 (126A, 126B) 에 세그먼트하여, 각각의 HRTF, 조기 반사, 및 잔여 룸 세그먼트들을 식별 할 수도 있다. BRIR 컨디셔닝 유닛 (106) 은 사이즈 [a, L] 의 좌측 및 우측 행렬들을 표현하는 행렬들 (129A, 129B) 로서, HRTF 및 조기 반사 세그먼트들을 BRIR SHC-도메인 변환 유닛 (112) 에 제공하며, 여기에서, a 는 HRTF 및 조기 반사 세그먼트들의 연결의 길이이며, L 은 라우드스피커들 (가상 또는 실제) 의 수이다. BRIR 컨디셔닝 유닛 (106) 은 사이즈 [b, L] 의 좌측 및 우측 룸 행렬들 (128A, 128B) 로서, BRIR

필터들 (126A, 126B) 의 잔여 룸 세그먼트들을 잔여 룸 응답 유닛 (110) 에 제공하며, 여기에서, b 는 잔여 룸 세그먼트들의 길이이며, L 은 라우드스피커들 (가상 또는 실제) 의 수이다.

[0070] 잔여 룸 응답 유닛 (110) 은 도 7 에서 SHCs (122) 에 의해 나타낸 바와 같이, 사운드 필드를 기술하는 계층적 엘리먼트들 (예를 들어, 구면 조화 계수들) 의 적어도 일부분과의 컨볼루션을 위하여 좌측 및 우측 공통 잔여 룸 응답 세그먼트들을 연산 또는 달리 결정하기 위해 위에 설명된 기법들을 적용할 수도 있다. 즉, 잔여 룸 응답 유닛 (110) 은 좌측 및 우측 공통 잔여 룸 응답 세그먼트들을 생성하기 위해, 좌측 및 우측 잔여 룸 행렬들 (128A, 128B) 을 수신하고, L개에 걸쳐 각각의 좌측 및 우측 잔여 룸 행렬들 (128A, 128B) 을 결합할 수도 있다. 잔여 룸 응답 유닛 (110) 은 일부 경우들에서 L 개에 걸쳐 좌측 및 우측 잔여 룸 행렬들 (128A, 128B) 을 평균화함으로써 결합을 수행할 수도 있다.

[0071] 그 후, 잔여 룸 응답 유닛 (110) 은 도 7 에서 채널(들)(124B) 로서 예시된 SHCs (122) 의 적어도 하나의 채널과 좌측 및 우측 공통 잔여 룸 응답 세그먼트들의 고속 컨볼루션을 연산할 수도 있다. 일부 예들에서, 좌측 및 우측 공통 잔여 룸 응답 세그먼트들은 주변의 무지향성 사운드를 표현하기 때문에, 채널(들)(124B) 은 사운드 필드의 무지향성 부분을 인코딩하는 SHCs (122) 채널들 중 W 채널 (즉, 0 차수) 이다. 이러한 예들에서, 길이 (Length) 의 W 채널 샘플에 대해, 잔여 룸 응답 유닛 (110) 에 의한 좌측 및 우측 공통 잔여 룸 응답 세그먼트들과의 고속 컨볼루션은 길이 (Length) 의 좌측 및 우측 출력 신호들 (134A, 134B) 을 생성한다.

[0072] 여기에 이용된 바와 같이 용어 "고속 컨볼루션" 및 "컨볼루션" 은 주파수 도메인에서의 포인트 와이즈 곱셈 연산 뿐만 아니라 시간 도메인에서의 컨볼루션 연산을 지칭할 수도 있다. 즉, 신호 프로세싱의 당해 기술 분야의 당업자에게 잘 알려진 바와 같이, 시간 도메인에서의 컨볼루션은 주파수 도메인에서의 포인트 와이즈 곱셈과 등가이고, 여기에서, 시간 및 주파수 도메인들은 서로의 변환들이다. 출력 변환은 전달 함수와 입력 변환의 포인트 와이즈 곱이다. 따라서, 컨볼루션 및 포인트와이즈 곱셈 (또는 간단히 "곱셈") 은 각각의 도메인들 (여기에서는 시간 및 주파수) 에 대하여 행해진 개념적으로 유사한 연산들을 지칭할 수 있다. 컨볼루션 유닛들 (114, 214, 230); 잔여 룸 응답 유닛들 (210, 354); 필터들 (384) 및 리버브 (386) 는 대안적으로 주파수 도메인에서 곱셈을 적용할 수도 있고, 여기에서 이들 컴포넌트들로의 입력들은 시간 도메인이기보다는 주파수 도메인에서 제공된다. "고속 컨볼루션" 또는 "컨볼루션"으로서 여기에서 설명되는 다른 연산들은 주파수 도메인에서의 곱셈으로 간단하게 지칭될 수도 있고, 여기에서, 이들 연산들로의 입력들은 시간 도메인이기보다는 주파수 도메인에서 제공된다.

[0073] 일부 예들에서, 잔여 룸 응답 유닛 (110) 은 BRIR 컨디셔닝 유닛 (106) 으로부터, 공통 잔여 룸 응답 세그먼트들의 온셋 시간에 대한 값을 수신할 수도 있다. 잔여 룸 응답 유닛 (110) 은 BRIR 필터들 (108) 에 대한 조기 세그먼트들과의 결합을 예상하여 출력 신호들 (134A, 134B) 을 제로 패딩하거나 또는 달리 지연시킬 수도 있다.

[0074] BRIR SHC-도메인 변환 유닛 (112)(이하, "도메인 변환 유닛 (112)") 은 SHC 렌더링 행렬을 BRIR 행렬들에 적용하여 좌측 및 우측 BRIR 필터들 (126A, 126B) 을 구면 조화 도메인으로 가능성있게 변환시킨 다음, L 개에 걸쳐 필터들을 가능성있게 합산한다. 도메인 변환 유닛 (112) 은 변환 결과를 좌측 및 우측 SHC-바이노럴 렌더링 행렬들 (130A, 130B) 로서 각각 출력한다. 행렬들 (129A, 129B) 이 사이즈 [a, L] 로 된 경우, SHC-바이노럴 렌더링 행렬들 (130A, 130B) 각각은 L 개에 걸쳐 필터들을 합산한 후 사이즈 $[(N+1)^2, a]$ 로 된다 (예를 들어, 식들 (4)-(5) 을 참조한다). 일부 예들에서, SHC-바이노럴 렌더링 행렬들 (130A, 130B) 은 런타임 또는 세트업 시간에서 연산되는 것보다 오디오 재생 디바이스 (100) 에서 구성된다. 일부 예들에서, SHC-바이노럴 렌더링 행렬들 (130A, 130B) 의 다수의 인스턴스들은 오디오 재생 디바이스 (100) 로 구성되고, 오디오 재생 디바이스 (100) 는 SHCs (124A) 에 적용하기 위하여 다수의 인스턴스들의 좌측/우측 페어를 선택한다.

[0075] 컨볼루션 유닛 (114) 은 일부 예들에서 SHCs (122) 의 차수로부터 차례로 감소될 수도 있는 SHCs (124A) 와 좌측 및 우측 바이노럴 렌더링 행렬들 (130A, 130B) 을 컨볼브한다. 주파수 (예를 들어, SHC) 도메인에서의 SHCs (124A) 에서, 변환 유닛 (114) 은 SHCs (124A) 와 좌측 및 우측 바이노럴 렌더링 행렬들 (130A, 130B) 의 각각의 포인트 와이즈 곱셈들을 연산할 수도 있다. 길이 (Length) 의 SHC 신호에 대해, 컨볼루션은 사이즈 $[(N+1)^2]$ 의 좌측 및 우측 필터링된 SHC 채널들 (132A, 132B) 을 가져오며, 통상적으로, 구면 조화 도메인의 각각의 차수/하위 차수에 대한 각각의 출력 신호들 행렬에 대하여 로우가 존재한다.

[0076] 결합 유닛 (116) 은 바이노럴 출력 신호들 (136A, 136B) 을 생성하기 위해 출력 신호들 (134A, 134B) 과 좌측 및 우측 필터링된 채널들 (132A, 132B) 을 결합할 수도 있다. 결합 유닛 (116) 은 바이노럴 출력 신호들

(136A, 136B) 을 생성하기 위해 좌측 및 우측 출력 신호들 (134A, 134B) 과 좌측 및 우측 바이노럴 출력 신호들을 결합하기 전에 HRTF 및 조기 에코들 (반사) 세그먼트들에 대한 좌측 및 우측 바이노럴 출력 신호들을 생성하기 위해 L 개에 걸쳐 각각의 좌측 및 우측 필터링된 SHC 채널들 (132A, 132B) 을 개별적으로 합산할 수도 있다.

[0077] 도 8 은 본 개시물에 설명된 바이노럴 오디오 렌더링 기법들의 여러 양태들을 수행할 수도 있는 오디오 재생 디바이스의 일 예를 예시하는 블록도이다. 오디오 재생 디바이스 (200) 는 도 7 의 오디오 재생 디바이스 (100) 의 예시적인 경우를 보다 자세하게 표현할 수도 있다.

[0078] 오디오 재생 디바이스 (200) 는 SHCs (242) 의 차수를 감소시키기 위해 비트스트림 (240) 으로부터 되돌아오는 SHCs (242) 를 프로세싱하는 선택적 SHCs 차수 감소 유닛 (204) 을 포함할 수도 있다. 선택적 SHCs 차수 감소는 SHCs (242) 의 최고차 (예를 들어, 0차) 채널 (262)(예를 들어, W 채널) 을 잔여 룸 응답 유닛 (210) 에 제공하고 감소된 차수의 SHCs (242) 를 컨볼루션 유닛 (230) 에 제공한다. SHCs 차수 감소 유닛 (204) 이 SHCs (242) 의 차수를 감소시키지 않는 경우들에서, 컨볼루션 유닛 (230) 이 SHCs (242) 와 일치하는 SHCs (272) 를 수신한다. 어느 경우에도, SHCs (272) 는 차원들 $[Length, (N+1)^2]$ 을 가지며, 여기에서 N 은 SHCs (272) 의 차수이다.

[0079] BRIR 컨디셔닝 유닛 (206) 및 BRIR 필터들 (208) 은 도 7 의 BRIR 컨디셔닝 유닛 (106) 및 BRIR 필터들 (108) 의 예시적인 경우들을 표현할 수도 있다. 잔여 응답 유닛 (214) 의 컨볼루션 유닛 (214) 은 위에 설명된 기법들을 이용하여 BRIR 컨디셔닝 유닛 (206) 에 의해 컨디셔닝된 공통의 좌측 및 우측 잔여 룸 세그먼트들 (244A, 244B) 을 수신하고, 컨볼루션 유닛 (214) 은 좌측 및 우측 잔여 룸 신호들 (262A, 262B) 을 생성하기 위해 최고차 채널 (262) 과 공통의 좌측 및 우측 잔여 룸 세그먼트들 (244A, 244B) 을 컨볼빙한다. 지연 유닛 (216) 은 좌측 및 우측 잔여 룸 신호들 (262A, 262B) 을 공통의 좌측 및 우측 잔여 룸 세그먼트들 (244A, 244B) 에 대한 샘플들의 온셋 수로 제로 패딩하여 좌측 및 우측 잔여 룸 출력 신호들 (268A, 268B) 을 생성한다.

[0080] BRIR SHC-도메인 변환 유닛 (220)(이하, 도메인 변환 유닛 (220)) 은 도 7 의 도메인 변환 유닛의 예시적인 경우를 표현할 수도 있다. 예시된 예에서, 변환 유닛 (222) 은 $(N+1)^2$ 차원의 SHC 렌더링 행렬 (224) 을 사이즈 $[a, L]$ 의 좌측 및 우측 행렬을 표현하는 행렬들 (248A, 248B) 에 적용하며, 여기에서, a 는 HRTF 와 조기 반사 세그먼트들의 연결의 길이이며, L 은 라우드스피커들 (예를 들어, 가상 라우드스피커) 의 수이다. 변환 유닛 (222) 은 $[(N+1)^2, a, L]$ 차원들을 갖는 SHC-도메인에서 좌측 및 우측 행렬들 (252A, 252B) 을 출력한다.

합산 유닛 (226) 은 차원들 $[(N+1)^2, a]$ 을 갖는 좌측 및 우측 중간 SHC-렌더링 행렬들 (254A, 254B) 을 생성하기 위해 L 개에 걸쳐 좌측 및 우측 행렬들 (252A, 252B) 각각을 합산할 수도 있다. 감소 유닛 (228) 은 중간 SHC-렌더링 행렬들을 SHCs (272) 에 적용하는 연산 복잡도를 추가로 감소시키기 위해, 최소 페이즈 감소를 적용하였던 SHC-렌더링 행렬들 (254A, 254B) 의 개별적인 최소 페이즈 부분들의 주파수 응답을 근사화시키도록 IIR 필터들을 설계하는 밸런스드 모델 절단 방법들 (Balanced Model Truncation methods) 을 이용하여, 그리고 위에 설명된 기법들, 이를 테면, 최소 페이즈 감소를 적용할 수도 있다. 감소 유닛 (228) 은 좌측 및 우측 렌더링 행렬들 (256A, 256B) 을 출력한다.

[0081] 컨볼루션 유닛 (230) 은 합산 유닛 (232) 이 좌측 및 우측 신호들 (260A, 260B) 을 생성하도록 합산하는 중간 신호들 (258A, 258B) 을 생성하기 위해 SHCs (272) 의 형태로 SHC 컨텐츠들을 필터링한다. 결합 유닛 (234) 은 좌측 및 우측 잔여 룸 출력 신호들 (268A, 268B) 및 좌측 및 우측 신호들 (260A, 260B) 을 결합하여 좌측 및 우측 바이노럴 출력 신호들 (270A, 270B) 을 생성한다.

[0082] 일부 예들에서, 바이노럴 렌더링 유닛 (202) 은 변환 유닛 (222) 에 의해 생성되는 SHC-바이노럴 렌더링 행렬들 (252A, 252B) 중 하나만을 이용함으로써 연산에 대한 추가적인 감소를 구현할 수도 있다. 그 결과, 컨볼루션 유닛 (230) 은 좌측 또는 우측 신호들의 단지 하나에 대해서만 연산하여 컨볼루션 연산들을 반으로 감소시킬 수도 있다. 이러한 예들에서, 합산 유닛 (232) 은 출력들 (260A, 260B) 을 렌더링할 때 제 2 채널에 대한 조건적 결정들을 행한다.

[0083] 도 9 는 본 개시물에 설명된 기법들에 따라 구면 조화 계수들을 렌더링하기 위한 바이노럴 렌더링 디바이스에 대한 예시적인 동작 모드를 예시하는 흐름도이다. 예시 목적들을 위하여, 예시적인 동작 모드는 도 7 의 오디오 재생 디바이스 (200) 에 대하여 설명된다. 바이노럴 룸 임펄스 응답 (BRIR) 컨디셔닝 유닛 (206) 은 좌측 및 우측 BRIR 필터들 (246A, 246B) 로부터 방향 의존성 컴포넌트들/세그먼트들, 구체적으로 헤드-관련 전

달 함수 및 조기 에코 세그먼트들을 추출함으로써 좌측 및 우측 BRIR 필터들 (246A, 246B) 을 각각 컨디셔닝한다 (300). 좌측 및 우측 BRIR 필터들 (126A, 126B) 각각은 하나 이상의 대응하는 라우드스피커들에 대한 BRIR 필터들을 포함할 수도 있다. BRIR 컨디셔닝 유닛 (106) 은 BRIR SHC-도메인 변환 유닛 (220) 에 그 추출된 헤드-관련 전달 함수 및 조기 에코 세그먼트들의 연결을 좌측 및 우측 행렬들 (248A, 248B) 로서 제공한다.

[0084] BRIR SHC-도메인 변환 유닛 (220) 은 HOA 렌더링 행렬 (224) 을 적용하여 추출된 헤드-관련 전달 함수 및 조기 에코 세그먼트들을 포함하는 좌측 및 우측 필터 행렬들 (248A, 248B) 을 변환하여, 구면 조화 (예를 들어, HOA) 도메인에서 좌측 및 우측 필터 행렬들 (252A, 252B) 을 생성한다 (302). 일부 예들에서, 오디오 재생 디바이스 (200) 는 좌측 및 우측 필터 행렬들 (252A, 252B) 로 구성될 수도 있다. 예시적인 예들에서, 오디오 재생 디바이스 (200) 는 비트스트림 (240) 의 대역외 또는 대역내 신호에서 BRIR 필터들 (208) 을 수신하고, 이 경우에 오디오 재생 디바이스 (200) 는 좌측 및 우측 필터 행렬들 (252A, 252B) 을 생성한다. 합산 유닛 (226) 은 라우드스피커 차원에 걸쳐 개별적인 좌측 및 우측 필터 행렬들 (252A, 252B) 을 합산하여, 좌측 및 우측 중간 SHC-렌더링 행렬들 (254A, 254B) 을 포함하는 SHC-도메인에서 바이노럴 렌더링 행렬을 생성할 수도 있다 (304). 감소 유닛 (228) 은 좌측 및 우측 SHC-렌더링 행렬들 (254A, 254B) 을 생성하기 위해 중간 SHC-렌더링 행렬들 (254A, 254B) 을 추가로 감소시킬 수도 있다.

[0085] 바이노럴 렌더링 유닛 (202) 의 컨볼루션 유닛 (230) 은 좌측 및 우측 중간 SHC-렌더링 행렬들 (256A, 256B) 을 SHC 컨텐츠 (이를 테면, 구면 조화 계수들 (272)) 에 적용하여 좌측 및 우측 필터링된 SHC (예를 들어, HOA) 채널들 (258A, 258B) 을 생성한다 (306).

[0086] 합산 유닛 (232) 은 SHC 차원 $((N+1)^2)$ 에 걸쳐 좌측 및 우측 필터링된 SHC 채널들 (258A, 258B) 각각을 합산하여, 방향 의존성 세그먼트들에 대한 좌측 및 우측 신호들 (260A, 260B) 을 생성한다 (308). 그 후, 결합 유닛 (116) 은 좌측 및 우측 잔여 룸 출력 신호들 (268A, 268B) 과 좌측 및 우측 신호들 (260A, 260B) 을 결합하여, 좌측 및 우측 바이노럴 출력 신호들 (270A, 270B) 을 포함하는 바이노럴 출력 신호를 생성할 수도 있다.

[0087] 도 10a 는 본 개시물에 설명된 기법들의 여러 양태들에 따라 도 7 및 도 8 의 오디오 재생 디바이스들에 의해 수행될 수도 있는 예시적인 동작 모드 (310) 를 예시하는 다이어그램이다. 이후, 도 8 의 오디오 재생 디바이스 (200) 와 관련하여 동작 모드 (310) 가 여기에서 설명된다. 오디오 재생 디바이스 (200) 의 바이노럴 렌더링 유닛 (202) 은 BRIR 필터들 (208) 의 예시적인 인스턴스들일 수도 있는 BRIR 데이터 (312), 및 HOA 렌더링 행렬 (224) 의 예시적인 인스턴스일 수도 있는 HOA 렌더링 행렬 (314) 로 구성될 수도 있다. 오디오 재생 디바이스 (200) 는 BRIR 데이터 (312) 및 HOA 렌더링 행렬 (314) 을 비트스트림 (240) 의 존재에서 대역내 및 대역외 시그널링 채널에서 수신할 수도 있다. 이 예에서 BRIR 데이터 (312) 는 예를 들어, L 개의 실제 또는 가상의 라우드스피커들을 표현하는 L 개의 필터들을 가지며, L 필터들 각각은 길이 (K) 를 갖는다. L 개의 필터들 각각은 좌측 및 우측 컴포넌트들 ("x 2") 을 포함할 수도 있다. 일부 경우들에서, L 필터들 각각은 좌측 또는 우측에 대한 단일의 성분을 포함할 수도 있고 이는 그 우측 또는 좌측으로서 그 상대 부분과 대칭적이다. 이는 고속 컨볼루션의 비용을 감소시킬 수도 있다.

[0088] 오디오 재생 디바이스 (200) 의 BRIR 컨디셔닝 유닛 (206) 은 세그먼트화 및 결합 동작들을 적용함으로서 BRIR 데이터 (312) 를 컨디셔닝할 수도 있다. 구체적으로, 예시적인 동작 모드 (310) 에서, BRIR 컨디셔닝 유닛 (206) 은 여기에 설명된 기법들에 따라 L개의 필터들 각각을, HRTF 에, 결합된 길이 (a) 의 조기 에코 세그먼트들을 더한 것으로 세그먼트하여 행렬 (315) (차원성 [a, 2, L]) 을 생성하고 잔여 룸 응답 세그먼트들로 세그먼트하여 행렬 (339) (차원성 [b, 2, L]) 을 생성한다 (324). BRIR 데이터 (312) 의 L 필터들의 길이 (K) 는 대략 a 와 b 의 합이다. 변화 유닛 (222) 은 $(N+1)^2$ 차원성의 HOA/SHC 렌더링 행렬 (314) 을 행렬 (315) 의 L 필터들에 적용하여, (차원성 $[(N+1)^2, a, 2, L]$) 의 좌측 및 우측 행렬들 (252A, 252B) 의 결합의 예시적인 인스턴스일 수도 있는 행렬 (317) 을 생성할 수도 있다. 합산 유닛 (226) 은 차원성 $[(N+1)^2, a, 2]$ 를 갖는 중간 SHC-렌더링 행렬 (335) 을 생성하기 위해 L 개에 걸쳐 좌측 및 우측 행렬들 (252A, 252B) 각각을 합산 할 수도 있다 (제 3 차원이 좌측 및 우측 컴포넌트들을 표현하는 값 2 를 갖는 경우: 중간 SHC-렌더링 행렬 (335) 은 좌측 및 우측 중간 SHC-렌더링 행렬 (335) 양쪽 모두의 예시적인 인스턴스로서 표현할 수도 있다) (326). 일부 예들에서, 오디오 재생 디바이스 (200) 는 HOA 컨텐츠 (316) (또는 이들의 감소된 버전, 예를 들어, HOA 컨텐츠 (321)) 에 대한 적용을 위하여 중간 SHC-렌더링 행렬 (335) 로 구성될 수도 있다. 일부 예들에서, 감소 유닛 (228) 은 행렬 (317) 의 좌측 또는 우측 컴포넌트들 중 오직 하나만을 이용함으로써 연

산에 대한 추가적인 감산들을 적용할 수도 있다 (328).

[0089] 오디오 재생 디바이스 (200) 는 차수 (N_1) 와 길이 (*Length*) 의 HOA 컨텐츠 (316) 를 수신하고, 일부 양태들에서, 차수 감산 연산을 적용하여 그 안의 구면 조화 계수들 (SHCs) 의 차수를 N 으로 감소시킨다. N_1 는 입력 HOA 컨텐츠 (321) 의 차수를 표현한다. 차수 감산 연산 (330) 의 HOA 컨텐츠 (321) 는 HOA 컨텐츠 (316) 와 같이 SHC 도메인에 있다. 선택적인 차수 감산 연산은 또한 최고차 (예를 들어, 0차) 신호 (319) 를 고속 컨볼루션 연산 (338) 에 대한 잔여 응답 유닛 (210) 에 생성 및 제공한다. HOA 차수 감소 유닛 (204) 이 HOA 컨텐츠 (316) 의 차수를 감소시키지 않는 경우에, 고속 컨볼루션 연산 (332) 의 적용은 감소된 차수를 갖지 않는 입력에 대해 연산한다. 어느 경우에도, 고속 컨볼루션 연산 (332) 에 대한 HOA 컨텐츠 (321) 의 입력은 차원들 $[Length, (N+1)^2]$ 을 가지며, 여기에서 N 은 차수이다.

[0090] 오디오 재생 디바이스 (200) 는 행렬 (335) 을 이용한 HOA 컨텐츠 (321) 의 고속 컨볼루션을 적용하여 이에 따라 차원들 $([Length, (N+1)^2, 2])$ 을 가진 좌측 및 우측 컴포넌트들을 가진 HOA 신호 (323) 를 생성할 수도 있다 (332). 또한, 고속 컨볼루션은 주파수 도메인에서의 행렬 (335) 과 HOA 컨텐츠 (321) 의 포인트 와이즈 곱셈 또는 시간 도메인에서의 컨볼루션을 지칭할 수도 있다. 오디오 재생 디바이스 (200) 는 $(N+1)^2$ 에 걸쳐 HOA 신호 (323) 를 추가로 합산하여 차원들 $([Length, 2])$ 을 갖는 합산된 신호 (325) 를 생성할 수도 있다 (334).

[0091] 이하, 잔여 행렬 (339) 로 돌아가서, 오디오 재생 디바이스 (200) 는 여기에서 설명된 기법들에 따라 L 개의 잔여 룸 응답 세그먼트들을 결합하여 차원들 $([b, 2])$ 을 갖는 공통 잔여 룸 응답 행렬 (327) 을 생성할 수도 있다 (336). 오디오 재생 디바이스 (200) 는 공통 잔여 룸 응답 행렬 (327) 을 이용하여 0차 HOA 신호 (319) 의 고속 컨볼루션을 적용하여 차원들 $([Length, 2])$ 을 갖는 룸 응답 신호 (329) 를 생성할 수도 있다 (338).

잔여 행렬 (339) 의 L 개의 잔여 응답 룸 응답 세그먼트들을 생성하기 위해, 오디오 재생 디바이스 (200) 는 BRIR 데이터 (312) 의 L 개 필터들의 $(a+1)$ 번째 샘플들에서 시작하는 잔여 응답 룸 응답 세그먼트들을 획득하였기 때문에, 오디오 재생 디바이스 (200) 는 샘플들을 지연 (예를 들어, 패딩) 시킴으로써 초기 a 샘플들을 고려하여 차원들 $([Length, 2])$ 을 갖는 룸 응답 신호 (311) 를 생성한다 (340).

[0092] 오디오 재생 디바이스 (200) 는 차원들 $([Length, 2])$ 을 갖는 출력 신호 (318) 를 생성하기 위해 엘리먼트들을 추가함으로써 합산된 신호 (325) 를 룸 응답 신호 (311) 와 결합한다 (342). 이러한 방식으로, 오디오 재생 디바이스는 L 개의 잔여 응답 세그먼트들 각각에 대해 고속 컨볼루션을 적용하는 것을 회피할 수도 있다. 바이노럴 오디오 출력 신호로의 변환을 위한 22 개의 채널 입력에 대해, 이는 잔여 룸 응답을 생성하기 위한 고속 컨볼루션들의 수를 22 개에서 2 개로 감소시킬 수도 있다.

[0093] 도 10b 는 본 개시물에 설명된 기법들의 여러 양태들에 따라 도 7 및 도 8 의 오디오 재생 디바이스들에 의해 수행될 수도 있는 예시적인 동작 모드 (350) 를 예시하는 다이어그램이다. 여기에서 이후 동작 모드 (350) 는 도 8 의 오디오 재생 디바이스 (200) 에 대하여 설명되며, 동작 모드 (310) 와 유사하다. 그러나, 동작 모드 (350) 는 L개의 실제 또는 가상 라우드스피커들에 대하여 시간 도메인에서의 멀티채널 스피커 신호들로의 HOA 컨텐츠를 먼저 렌더링한 다음, 여기에 설명된 기법들에 따라 스피커 피드들 각각에 대하여 효율적인 BRIR 필터링을 적용하는 것을 포함한다. 그 목적을 위하여, 오디오 재생 디바이스 (200) 는 HOA 컨텐츠 (321) 를 차원들 $([Length, L])$ 을 갖는 멀티채널 오디오 신호 (333) 로 변환한다 (344). 추가로, 오디오 재생 디바이스는 BRIR 데이터 (312) 를 SHC 도메인으로 변환하지 않는다. 따라서, 오디오 재생 디바이스 (200) 에 의해 신호 (314) 에 감산을 적용하는 것은 차원들 $([a, 2, L])$ 을 갖는 행렬 (337) 을 생성한다 (328).

[0094] 그 후, 오디오 재생 디바이스 (200) 는 행렬 (337) 을 이용하여 멀티채널 오디오 신호 (333) 의 고속 컨볼루션 (332) 을 적용하여, (좌측 및 우측 성분들을 가진) 차원들 $([Length, L, 2])$ 을 갖는 멀티채널 오디오 신호 (341) 를 생성한다 (348). 그후, 오디오 재생 디바이스 (200) 는 L 개의 채널들/스피커들에 의한 멀티채널 오디오 신호 (341) 를 합산하여, 차원들 $([Length, 2])$ 을 갖는 신호 (325) 를 생성한다 (346).

[0095] 도 11 은 본 개시물에 설명된 바이노럴 오디오 렌더링 기법들의 여러 양태들을 수행할 수도 있는 오디오 재생 디바이스 (350) 의 일 예를 예시하는 블록도이다. 도 11 의 예에서 단일 디바이스, 즉 오디오 재생 디바이스 (350) 로서 예시되어 있지만, 기법들은 하나 이상의 디바이스들에 의해 수행될 수도 있다. 따라서, 기법들은 이러한 면으로 제한되지 않아야 한다.

[0096] 또한, 도 1 내지 도 10b 의 예들에 대하여 구면 조화 도메인에서 적용되고 있는 것으로서 위에 일반적으로 설명

되어 있지만, 기법들은 위에 주지된 서라운드 사운드 포맷들, 이를 테면, 5.1 서라운드 사운드 포맷, 7.1 서라운드 사운드 포맷, 및/또는 22.2 서라운드 사운드 포맷에 따르는 채널 기반 신호들을 포함하는 임의의 형태의 오디오 신호들에 대하여도 또한 구현될 수도 있다. 따라서, 기법들은 구면 조화 도메인에서 특정되는 오디오 신호들로 제한되지 않아야 하며 임의의 형태의 오디오 신호에 대하여 적용될 수도 있다.

[0097] 도 11 의 예에 도시된 바와 같이, 오디오 재생 디바이스 (350) 는 도 7 의 예에 도시된 오디오 재생 디바이스 (100) 와 유사할 수도 있다. 그러나, 오디오 재생 디바이스 (350) 는 일 예로서, 22.2 서라운드 사운드 포맷에 따르는 일반 채널 기반 오디오 신호들에 대하여 기법들을 동작시키거나 달리 수행할 수도 있다. 추출 유닛 (104) 은 오디오 채널들 (352) 을 추출할 수도 있으며, 오디오 채널들 (352) 은 일반적으로 "n" 개의 채널들을 포함할 수도 있으며, 이 예에서는 22.2 서라운드 사운드 포맷에 따르는 22 개의 채널들을 포함하는 것으로 본다. 이들 채널들 (352) 은 바이노럴 렌더링 유닛 (351) 의 잔여 룸 응답 유닛 (354) 및 채널당 절단된 필터 유닛 (356) 양쪽 모두에 제공될 수도 있다.

[0098] 위에 설명된 바와 같이, BRIR 필터들 (108) 은 하나 이상의 BRIR 필터들을 포함하고, 도 3 의 BRIR 필터들 (37) 의 예를 표현할 수도 있다. BRIR 필터들 (108) 은 개별적인 BRIRs 에 대한 좌측 및 우측 HRTF 의 효과를 표현하는 별도의 BRIR 필터들 (126A, 126B) 을 포함할 수도 있다.

[0099] BRIR 컨디셔닝 유닛 (106) 은 각각의 채널 (n) 에 대하여 하나씩, BRIR 필터들 (126A, 126B) 의 L 개의 인스턴스들을 수신하며 여기에서 각각의 BRIR 필터는 길이 (N) 를 갖는다. BRIR 필터들 (126A, 126B) 은 침묵 샘플들을 제거하도록 이미 컨디셔닝될 수도 있다. BRIR 컨디셔닝 유닛 (106) 은 위에 설명된 기법들을 적용하여 BRIR 필터들 (126A, 126B) 을 세그먼트하여, 각각의 HRTF, 조기 반사, 및 잔여 룸 세그먼트들을 식별할 수도 있다. BRIR 컨디셔닝 유닛 (106) 은 사이즈 [a, L] 의 좌측 및 우측 행렬들을 표현하는 행렬들 (129A, 129B) 로서, HRTF 및 조기 반사 세그먼트들을 채널당 절단된 필터 유닛 (356) 에 제공하며, 여기에서, a 는 HRTF 및 조기 반사 세그먼트들의 연결의 길이이며, n 은 라우드스피커들 (가상 또는 실제) 의 수이다. BRIR 컨디셔닝 유닛 (106) 은 사이즈 [b, L] 의 좌측 및 우측 룸 행렬들 (128A, 128B) 로서, BRIR 필터들 (126A, 126B) 의 잔여 룸 세그먼트들을 잔여 룸 응답 유닛 (354) 에 제공하며, 여기에서, b 는 잔여 룸 세그먼트들의 길이이며, n 은 라우드스피커들 (가상 또는 실제) 의 수이다.

[0100] 잔여 룸 응답 유닛 (354) 은 오디오 채널들 (352) 과의 컨볼루션을 위한 좌측 및 우측 공통 잔여 룸 응답 세그먼트들을 연산 또는 달리 결정하기 위해 위에 설명된 기법들을 적용할 수도 있다. 즉, 잔여 룸 응답 유닛 (110) 은 좌측 및 우측 공통 잔여 룸 응답 세그먼트들을 생성하기 위해, 좌측 및 우측 잔여 룸 행렬들 (128A, 128B) 을 수신하고, n개에 걸쳐 각각의 좌측 및 우측 잔여 룸 행렬들 (128A, 128B) 을 결합할 수도 있다. 잔여 룸 응답 유닛 (354) 은 일부 경우들에서 n 개에 걸쳐 좌측 및 우측 잔여 룸 행렬들 (128A, 128B) 을 평균화함으로써 결합을 수행할 수도 있다.

[0101] 그 후, 잔여 룸 응답 유닛 (354) 은 오디오 채널 (352) 중 적어도 하나와 좌측 및 우측 공통 잔여 룸 응답 세그먼트들의 고속 컨볼루션을 연산할 수도 있다. 일부 예들에서, 잔여 룸 응답 유닛 (352) 은 BRIR 컨디셔닝 유닛 (106) 으로부터, 공통 잔여 룸 응답 세그먼트들의 온셋 시간에 대한 값을 수신할 수도 있다. 잔여 룸 응답 유닛 (354) 은 BRIR 필터들 (108) 에 대한 조기 세그먼트들과의 결합을 예상하여 출력 신호들 (134A, 134B) 을 제로 패딩하거나 또는 달리 지연시킬 수도 있다. 출력 신호들 (134A) 은 좌측 오디오 신호들을 표현할 수도 있는 한편, 출력 신호들 (134B) 은 우측 오디오 신호들을 표현할 수도 있다.

[0102] 채널당 절단된 필터 유닛 (356)(이하, "절단된 필터 유닛 (356)") 은 HRTF 및 BRIR 필터들의 조기 반사 세그먼트들을 채널들 (352) 에 적용할 수도 있다. 보다 구체적으로, 채널당 절단된 필터 유닛 (356) 은 HRTF 및 BRIR 필터들의 조기 반사 세그먼트들을 표현하는 행렬들 (129A 및 129B) 을 채널들 (352) 의 각각의 하나에 적용할 수도 있다. 일부 경우들에서, 행렬들 (129A 및 129B) 이 결합되어 단일의 행렬 (129) 을 형성할 수도 있다. 또한, 통상적으로, HRTF 및 조기 반사 행렬들 (129A 및 129B) 각각의 좌측 하나와 HRTF 및 조기 반사 행렬들 (129A 및 129B) 각각의 우측 하나가 존재한다. 즉, 통상적으로 좌측 귀와 우측 귀에 대한 HRTF 및 조기 반사 행렬들이 존재한다. 채널당 방향 유닛 (356) 은 좌측 및 우측 필터링된 채널들 (358A 및 358B) 을 출력하기 위해 좌측 및 우측 행렬들 (129A, 129B) 각각을 적용할 수도 있다. 결합 유닛 (116) 은 출력 신호들 (134A) 과 우측 필터링된 채널들 (358A) 을 결합 (또는 즉, 혼합) 하는 한편, 출력 신호들 (134B) 과 우측 필터링된 채널들 (358B) 을 결합 (또는 즉, 혼합) 하여 바이노럴 출력 신호들 (136A, 136B) 을 생성할 수도 있다. 바이노럴 출력 신호 (136A) 는 좌측 오디오 채널에 대응할 수도 있고 바이노럴 출력 신호 (136B) 는 우측 오디오 채널과 대응할 수도 있다.

[0103] 일부 예들에서, 바이노럴 렌더링 유닛 (351) 은 잔여 툴 응답 유닛 (354) 이 채널당 절단된 필터 유닛 (356) 의 동작과 동시에 동작하도록, 잔여 툴 응답 유닛 (354) 및 채널당 절단된 필터 유닛 (356) 을 서로에 대해 동시에 인보크할 수도 있다. 즉, 일부 예들에서, 잔여 툴 응답 유닛 (354) 은 채널당 절단된 필터 유닛 (356) 과 병렬로 (그러나 종종 동시적이지 않은) 동작하여 바이노럴 출력 신호들 (136A, 136B) 이 생성될 수도 있는 속도를 종종 개선시킬 수도 있다. 여러 도면들에서 캐스케이드 방식으로 가능성있게 동작하는 것으로서 위에 도시되어 있지만, 기법들은 달리 특정되어 있지 않은 한, 본 개시물에 설명된 유닛들 또는 모듈들의 어느 것의 동시 또는 병렬 동작에 제공할 수도 있다.

[0104] 도 12 는 본 개시물에 설명된 기법들의 여러 양태들에 따라 도 11 의 오디오 재생 디바이스 (350) 에 의해 수행될 수도 있는 프로세스 (380) 를 예시하는 다이어그램이다. 프로세스 (380) 는 각각의 BRIR 의, 두개의 부분들: (a) 좌측 필터들 ($384A_L-384N_L$) 에 의해 그리고 우측 필터들 ($384A_R-384N_R$) (총괄적으로, "필터들 (384)") 에 의해 표현되는 HRTF 및 조기 반사들의 효과들을 통합하는 보다 적은 컴포넌트들, 및 (b) 좌측 리버브 필터 ($386L$) 및 우측 리버브 필터 ($386R$) (총괄적으로, "공통 필터들 (386)") 에 의해 표현되고 오리지널 BRIRs 의 테일들 모두의 특성들로부터 생성되는 공통의 '리버브 테일'로의 분해를 실현한다. 프로세스 (380) 에서 도시된 채널당 필터들 (384) 은 위에 주지된 부분 (a) 를 표현할 수도 있는 한편, 프로세스 (380) 에서 도시된 공통 필터들 (386) 은 위에 주지된 부분 (b) 을 표현할 수도 있다.

[0105] 프로세스 (380) 는 비가청 컴포넌트들을 제거하고 HRTF/조기 반사들 및 뒤의 반사들/확산으로 인한 성분들을 포함하는 성분들을 결정하기 위해 BRIRs 를 분석함으로서 이 분해를 수행한다. 이는 일 예로서, 부분 (a) 에 대해 2704 텁들인 길이의 FIR 필터, 및 다른 예로서, 부분 (b) 에 대해 15232 텁들인 길이의 FIR 필터를 가져온다. 프로세스 (380) 에 따르면, 오디오 재생 디바이스 (350) 는 동작 (396) 시 예시의 목적을 위하여 22 인 것으로 보는 개별적인 n 채널들 각각에 보다 짧은 FIR 필터들만을 적용할 수도 있다. 이 동작의 복잡도는 아래 재현된 식 (8) 에서 (4096 포인트 FFT 를 이용한) 연산의 제 1 부분에서 표현될 수도 있다. 프로세스 (380) 에서, 오디오 재생 디바이스 (350) 는 동작 (398) 에서 22 개의 채널 각각에 공통 '리버브 테일'을 적용하지 않고 이들 모두의 추가적인 혼합에 적용할 수도 있다. 이 복잡도는 또한 첨부된 부록에 도시되어 있는 식 (8) 에서의 복잡도 계산의 후반부에 표현된다.

[0106] 이러한 면에서, 프로세스 (380) 는 복수의 N 개의 채널들로부터의 혼합 오디오 컨텐츠에 기초하여 합성 오디오 신호를 생성하는 바이노럴 오디오 렌더링의 방법을 표현할 수도 있다. 추가로, 프로세스 (380) 는 N 개의 채널 필터들의 출력과 함께 지연에 의해 합성 오디오 신호를 추가로 정렬할 수도 있으며, 각각의 채널 필터가 절단된 BRIR 필터를 포함한다. 또한, 프로세스 (380) 에서, 그 후, 오디오 재생 디바이스 (350) 는 동작 (398) 에서 공통 합성 잔여 툴 임펄스 응답과 정렬된 합성 오디오 신호를 필터링할 수도 있고, 바이노럴 오디오 출력 ($388L$, $388R$) 의 좌측 및 우측 컴포넌트들에 대한 동작들 ($390L$ 및 $390R$) 에서 각각의 채널 필터의 출력과, 필터링된 정렬된 합성 오디오 신호를 혼합할 수도 있다.

[0107] 일부 예들에서, 절단된 BRIR 필터 및 공통 합성 잔여 임펄스 응답은 메모리에 사전 로딩된다.

[0108] 일부 예들에서, 정렬된 합성 오디오 신호의 필터링은 시간 주파수 도메인에서 수행된다.

[0109] 일부 예들에서, 정렬된 합성 오디오 신호의 필터링은 컨볼루션을 통하여 시간 도메인에서 수행된다.

[0110] 일부 예들에서, 절단된 BRIR 필터 및 공통 합성 잔여 임펄스 응답은 분해 분석에 기초한다.

[0111] 일부 예들에서, 분해 분석은 N 개의 툴 임펄스 응답들 각각에 대하여 수행되며, N 개의 절단된 툴 임펄스 응답들 및 N 개의 잔여 임펄스 응답들을 발생시킨다 (여기에서, N 은 n 또는 n 초과로서 표기될 수도 있다).

[0112] 일부 예들에서, 절단된 임펄스 응답은 각각의 툴 응답 임펄스 응답의 총 길이의 40% 미만을 표현한다.

[0113] 일부 예들에서, 절단된 임펄스 응답은 111 과 17,830 사이의 텁 범위를 포함한다.

[0114] 일부 예들에서, N 개의 잔여 임펄스 응답들 각각은 복잡도를 감소시키는 공통 합성 잔여 툴 응답에 결합된다.

[0115] 일부 예들에서, 각각의 채널 필터의 출력을 필터링된 정렬된 합성 오디오 신호와 혼합하는 것은 좌측 스피커 출력에 대한 제 1 세트의 혼합 및 우측 스피커 출력에 대한 제 2 세트의 혼합을 포함한다.

[0116] 여러 예들에서, 위에 설명된 프로세스 (380) 의 여러 예들의 방법, 또는 이들의 조합은 메모리 및 하나 이상의 프로세서들을 포함하는 디바이스, 방법의 각각의 단계를 수행하는 수단을 포함하는 장치, 및 비일시적 컴퓨터 관독가능 저장 매체 상에 저장된 명령들을 실행함으로써 방법의 각각의 단계를 수행하는 하나 이상의 프로세서

들에 의해 수행될 수도 있다.

[0117] 또한, 위에 설명된 임의의 예들에서 전술한 특정 피쳐들은 상술한 기법들의 유익한 예에 결합될 수도 있다. 즉, 특정 피쳐들 중 임의의 것이 일반적으로 기법들의 모든 예들에 적용가능하다. 기법들의 여러 예들이 설명되어 있다.

[0118] 본 개시물에 설명된 기법들은 일부 경우들에, 가정인 BRIR 세트를 따라 샘플들 111 내지 17830 만을 식별할 수도 있다. 그 후, 예시적인 룸의 볼륨으로부터 혼합 시간 (T_{mp95}) 을 계산함으로써, 기법들은 53.6 ms 후에 모든 BRIRs 을 공통 리버브 테일을 공유하게 할 수도 있고, 그 결과 15232 샘플 길이의 공통 리버브 테일 및 나머지 2704 샘플 HRTF + 반사 임펄스들을 발생시키며, 여기에서 이들 사이에 3ms 가 크로스페이드한다. 연산비용 분류의 관점에서, 다음에 도달될 수도 있다:

[0119] (a) 공통 리버브 테일 : $10 * 6 * \log_2(2 * 15232 / 10)$.

[0120] (b) 나머지 임펄스들 : 하나의 프레임에서 이를 행하기 위해 4096 FFT 를 이용하여, $22 * 6 * \log_2(2 * 4096)$.

[0121] (c) 추가적인 22 가산들.

[0122] 그 결과, 이에 따라 최종 성능 지수는 대략 $C_{mod} = \max(100 * (C_{conv} - C) / C_{conv}, 0) = 88.0$ 와 같을 수도 있고 여기에서:

$$C_{mod} = \max(100 * (C_{conv} - C) / C_{conv}, 0), \quad (6)$$

[0124] 여기에서 C_{conv} 는 최적화되지 않은 구현의 추정값이고:

$$C_{conv} = (22+2)*(10)*(6*\log_2(2*48000/10)), \quad (7)$$

[0126] C 는 일부 양태에서, 2 개의 추가적인 팩터들에 의해 결정될 수도 있다:

$$C = 22 * 6 * \log_2(2 * 4096) + 10 * 6 * \log_2(2 * \frac{15232}{10}). \quad (8)$$

[0128] 따라서, 일부 양태들에서, 성능 지수는 $C_{mod} = 87.35$ 이다.

[0129] $B_n(z)$ 로서 표기되는 BRIR 필터는 2 개의 함수들 $BT_n(z)$ 및 $BR_n(z)$ 로 분해될 수도 있고, 이들 함수는 절단된 BRIR 필터 및 리버브 BRIR 필터를 각각 표기한다. 위에 주지된 부분 (a) 는 이 절단된 BRIR 필터를 지칭할 수도 있는 한편, 위에서의 부분 (b) 는 리버브 BRIR 필터를 지칭할 수도 있다. 그 후, $B_n(z)$ 는 $BT_n(z) + (z^{-m} * |BR_n(z)|)$ 와 같을 수도 있고, 여기에서, m 은 지연을 표기한다. 이에 따라, 출력 신호 ($Y(z)$) 는 다음과 같이 연산될 수도 있다:

$$\sum_{n=0}^{N-1} [X_n(z) \cdot BT_n(z) + z^{-m} \cdot X_n(z) * BR_n(z)] \quad (9)$$

[0131] 프로세스 (380) 는 $BR_n(z)$ 를 분석하여, 공통 합성 리버브 테일 세그먼트를 유도할 수도 있고, 여기에서 이 공통 $BR(z)$ 가 채널 특정 $BR_n(z)$ 대신에 적용될 수도 있다. 이 공통 (또는 채널 일반) 합성 $BR(z)$ 가 이용될 때, $Y(z)$ 는 다음과 같이 연산될 수도 있다:

$$\sum_{n=0}^{N-1} [X_n(z) \cdot BT_n(z) + z^{-m} BR_n(z)] \cdot \sum_{n=0}^{N-1} X_n(z) \quad (10)$$

[0133] 도 13 은 본 개시물에 설명된 바이노럴 오디오 렌더링 기법들의 여러 양태들을 수행할 수도 있는 오디오 재생 디바이스의 일 예를 예시하는 블록도이다. 도 13 의 예에서 단일 디바이스, 즉 오디오 재생 디바이스 (400) 로서 예시되어 있지만, 기법들은 하나 이상의 디바이스들에 의해 수행될 수도 있다. 따라서, 기법들은 이러한 면으로 제한되지 않아야 한다. 또한, 오디오 재생 디바이스 (400) 는 오디오 재생 시스템 (62) 의 일 예를 나타낼 수도 있다.

[0134] 도 13 의 예에 도시된 바와 같이, 오디오 재생 디바이스 (400) 는 추출 유닛 (404), BRIR 선택 유닛 (424) 및

바이노럴 렌더링 유닛 (402) 을 포함할 수도 있다. 추출 유닛 (404) 은 비트스트림 (420) 으로부터 인코딩된 오디오 데이터를 출력하도록 구성되는 유닛을 나타낼 수도 있다. 추출 유닛 (404) 은 추출된 인코딩된 오디오 데이터를 구면 조화 계수들 (SHCs)(422) (이는 또한 SHCs (422) 가 1 보다 큰 차수와 연관된 적어도 하나의 계수를 포함할 수도 있다는 점에서 고차 엠피소닉 (HOA) 으로 지칭될 수도 있음) 의 형태로 바이노럴 렌더링 유닛 (146) 에 포워딩할 수도 있다. BRIR 선택 유닛 (424) 은 BRIRs 의 규칙적 또는 불규칙적 세트가 여기에 설명된 기술들에 따라 SHCs (422) 를 바이노럴화하는데 이용되는지의 여부를 선택하기 위해 사용자, 사용자 에이전트, 또는 다른 외부 엔티티가 사용자 입력 (425) 을 제공할 수도 있는 인터페이스를 나타낸다. BRIR 선택 유닛 (424) 은 커맨드 라인 또는 그래픽 사용자 인터페이스, 애플리케이션 프로그래밍 인터페이스, 네트워크 인터페이스, 애플리케이션 인터페이스, 이를 테면, 단순 오브젝트 액세스 프로토콜 (Simple Object Access Protocol), 원격 프로시저 호출 (Remote Procedure Call), 또는 외부 엔티티가 BRIRs 의 규칙적 또는 불규칙적 세트가 이용되는지의 여부를 구성할 수도 있는 임의의 다른 인터페이스를 포함할 수도 있다. 신호 (426) 는 SHCs (422) 를 바이노럴화하기 위하여 BRIRs 의 규칙적 또는 불규칙적 세트를 이용하도록 바이노럴 렌더링 유닛 (402) 에 지시 또는 구성하는 제어 신호 또는 사용자 구성 데이터를 나타낸다. 신호 (426) 는 플래그, 기능 파라미터, 신호, 또는 오디오 재생 디바이스 (400) 가 바이노럴 렌더링 유닛 (402) 에 SHCs (422) 를 바이노럴화하는데 규칙적 또는 불규칙적 세트를 선택하도록 지시할 수도 있는 임의의 다른 수단을 나타낼 수도 있다.

[0135]

일부 예들에서, 오디오 재생 디바이스 (400) 는 SHCs (422) 를 생성하도록 인코딩된 오디오 데이터를 디코딩하도록 구성된 오디오 디코딩 유닛을 포함한다. 오디오 디코딩 유닛은 일부 양태들에서, SHCs (422) 를 인코딩하는데 이용되는 오디오 인코딩 프로세스에 역이 되는 오디오 디코딩 프로세스를 수행할 수도 있다. 오디오 디코딩 유닛은 인코딩된 오디오 데이터의 SHCs 를 시간 도메인으로부터 주파수 도메인으로 변환하여 SHCs (422) 를 생성하도록 구성된 시간-주파수 분석 유닛을 포함할 수도 있다. 즉, 인코딩된 오디오 데이터가 시간 도메인으로부터 주파수 도메인으로 변환되지 않는 SHC (422) 의 압축된 형태를 표현할 때, 오디오 디코딩 유닛은 시간-주파수 분석 유닛을 인보크하여, (주파수 도메인에서 특정된) SHCs (422) 를 생성하도록 SHCs 를 시간 도메인으로부터 주파수 도메인으로 변환할 수도 있다.

[0136]

시간-주파수 분석 유닛은 몇몇 예를 제공하면 고속 푸리에 변환 (FFT), 이산 코사인 변환 (DCT), 수정된 이산 코사인 변환 (MDCT), 및 이산 사인 변환 (DST) 을 포함하는 임의의 형태의 푸리에 기반 변환을 적용하여, SHCs 를 시간 도메인으로부터 주파수 도메인에서의 SHCs (422) 로 변환할 수도 있다. 일부 경우들에서, SHCs (422) 는 비트스트림 (420) 으로 주파수 도메인에서 이미 특정될 수도 있다. 이들 경우들에서, 시간-주파수 분석 유닛은 변환을 적용하거나 수신된 SHCs (422) 를 달리 변환하는 것이 없이 바이노럴 렌더링 유닛 (402) 으로 SHCs (422) 를 패스시킬 수도 있다. 주파수 도메인에서 특정된 SHCs (422) 에 관하여 설명되어 있지만, 기법들은 시간 도메인에서 특정된 SHCs (422) 에 관하여 수행될 수도 있다.

[0137]

바이노럴 렌더링 유닛 (402) 은 SHCs (422) 를 바이노럴화하도록 구성된 유닛을 나타낸다. 즉, 바이노럴 렌더링 유닛 (402) 은 SHCs (422) 가 기록되었던 룸에서의 청취자에 의해 좌측 및 우측 채널이 청취되는 방법을 모델링하기 위해 공간화를 피쳐링할 수도 있는, 좌측 및 우측 채널로 SHCs (422) 를 렌더링하도록 구성된 유닛을 표현할 수도 있다. 바이노럴 렌더링 유닛 (402) 은 헤드셋, 이를 테면, 헤드폰들을 통하여 재생에 적합한 좌측 채널 (436A) 및 우측 채널 (436B)(이들은 총괄적으로 "채널 (436)" 로서 지칭될 수도 있음) 을 생성하기 위해 SHCs (422) 를 렌더링할 수도 있다. 도 13 의 예에 도시된 바와 같이, 바이노럴 렌더링 유닛 (402) 은 보간 유닛 (406), 시간 주파수 분석 유닛 (408), 복소 BRIR 유닛 (410), 합산 유닛 (442), 복소 곱셈 유닛 (414), 대칭 최적화 유닛 (416), 비대칭 최적화 유닛 (418) 및 역 시간 주파수 분석 유닛 (420) 을 포함한다.

[0138]

바이노럴 렌더링 유닛 (402) 은 보간된 규칙적 BRIR 필터들 (407C) 을 생성하도록 불규칙적 BRIR 필터들 (407A) 을 보간하는 보간 유닛 (406) 을 인보크할 수도 있고, 여기에서, BRIR 필터들의 문맥에서의 "규칙적" 또는 "불규칙적" 에 대한 언급은 서로에 대해 스피커들의 간격의 규칙성 또는 불규칙성을 표기할 수도 있다. 불규칙적 BRIR 필터들 (407A) 은 L x 2 와 동일한 사이즈일 수도 있다 (여기에서 L 은 라우드스피커들의 수를 표기한다). 규칙적 BRIR 필터들 (407A) 은 (페어들로서 규칙적으로 정렬되는 것으로 주어지면) L 개의 라우드스피커들 x 2 를 포함할 수도 있다. 오디오 재생 디바이스 (400) 의 사용자 또는 다른 오퍼레이터들은 불규칙적 BRIR 필터들 (407A) 또는 규칙적 필터들 (407B) 이 SHC (422) 의 바이노럴화 동안에 이용될지의 여부를 표시 또는 달리 구성할 수도 있다.

[0139]

또한, 오디오 재생 디바이스 (400) 의 사용자 또는 다른 오퍼레이터는 불규칙적 BRIR 필터들 (407A) 이 SHC (422) 의 바이노럴화 동안에 이용될 때, 규칙적 BRIR 필터들 (407C) 을 생성하기 위해 불규칙적 BRIR 필터들

(407A)에 보간이 수행되는지의 여부를 표시 또는 달리 구성할 수도 있다. 보간 유닛 (406)은 B 개의 라우드스피커 페어들을 형성하는 벡터 기반 진폭 패닝 또는 다른 패닝 기법들을 이용하여 규칙적 필터들 (407B)을 보간할 수도 있어, 이는 (또한 이것이 규칙적이고 이에 따라 축을 중심으로 하여 대칭인 것으로 주어지면) L x 2 사이즈를 갖는 규칙적 BRIR 필터들 (407C)을 발생시킨다. 도 13의 예에 도시되지 않았지만, 사용자 또는 다른 오퍼레이터는 불규칙적 BRIR 필터들 (407A), 규칙적 BRIR 필터들 (407B), 및/또는 규칙적 BRIR 필터들 (407C)이 SHC (422)를 바이노럴화할 때 이용될지의 여부를 선택하기 위해 (예를 들어, 일련의 버튼들 또는 다른 입력들로서) 물리적으로 제시되었는지 또는 그래픽 사용자 인터페이스를 통하여 그래픽적으로 제시되었는지의 여부를 사용자 인터페이스를 통하여 오디오 재생 디바이스 (400)와 인터페이스할 수도 있다.

[0140] 임의의 이벤트에서, BRIR 필터들이 (407A-407C) (어느 것이 SHC (422)를 바이노럴화하는데 선택되는지에 의존하여) 시간 도메인으로 제시될 때, 바이노럴 렌더링 유닛 (402)은 BRIR 필터들 (407A-407C) ("BRIR 필터들 (407)") 중 선택된 하나를 시간 도메인으로부터 주파수 도메인으로 변환하도록 시간 주파수 분석 유닛 (408)을 인보크할 수도 있고, 이는 변환된 BRIR 필터들 (409A-409C) ("BRIR 필터들 (409)")을 각각 발생시킨다. 복소 BRIR 유닛 (410)은 각각이 $L \times (N+1)^2$ 의 사이즈를 갖는 2개의 BRIR 렌더링 벡터들 (411A 및 411B)을 생성하기 위해 불규칙적 렌더러 (405A) ($L \times (N+1)^2$ 의 사이즈를 가짐) 또는 규칙적 렌더러 (405B) ($L \times (N+1)^2$ 의 사이즈를 가짐) 중 하나 및 하나 이상의 BRIR 필터 (409)에 대하여 엘리먼트-바이-엘리먼트 복소 곱셈 및 합산을 수행하도록 구성되는 유닛을 나타내며, 여기에서, N은 또한 SHC (422) 중 하나 이상이 대응하는 구면 기저 함수들의 최고 차수를 나타낸다.

[0141] BRIR 필터들 (407) 중 선택된 하나가 규칙적인지 또는 불규칙적인지의 여부에 의존하여, 복소 BRIR 유닛 (410)은 불규칙적 렌더러 (405A) 또는 규칙적 렌더러 (405B)를 선택할 수도 있다. 즉, 일 예로서, BRIR 필터들 (407) 중 선택된 하나가 규칙적이면 (예를 들어, BRIR 필터 (407B 또는 407C)), 복소 BRIR 유닛 (410)은 규칙적 렌더러 (405B)를 선택한다. BRIR 필터들 (407) 중 선택된 하나가 불규칙적이면 (예를 들어, BRIR 필터 (407A)), 복소 BRIR 유닛 (410)은 규칙적 렌더러 (405A)를 선택한다. 일부 예들에서, 오디오 재생 디바이스 (400)의 사용자 또는 다른 오퍼레이터는 불규칙적 렌더러 (405A) 또는 규칙적 렌더러 (405B)를 이용하는지의 여부를 표시 또는 달리 선택할 수도 있다. 일부 예들에서, 오디오 재생 디바이스 (400)의 사용자 또는 다른 오퍼레이터는 BRIR 필터들 (407) 중 하나를 사용하도록 선택하기보다는, 불규칙적 렌더러 (405A) 또는 규칙적 렌더러 (405B)를 사용할지의 여부를 표시 또는 달리 선택할 수도 있다 (렌더러 (405A 또는 405B)의 선택이 BRIR 필터들 (407)의 선택을 인에이블하는 경우, 규칙적 렌더러 (405B)를 선택하는 것은 BRIR 필터들 (407B 및/또는 407C)의 선택을 발생시키고, 불규칙적 렌더러 (405A)를 선택하는 것은 BRIR 필터들 (407A)의 선택을 발생시킨다).

[0142] 합산 유닛 (442)은 합산된 BRIR 렌더링 벡터들 (413A 및 413B)을 생성하기 위해 BRIR 렌더링 벡터들 (411A 및 411B) 각각을 합산하는 유닛을 나타낼 수도 있다. 윈도잉 유닛은 윈도잉된 BRIR 렌더링 벡터들 (415A 및 415B)을 생성하기 위해 합산된 BRIR 렌더링 벡터들 (413A 및 413B)에 윈도잉 함수를 적용하는 유닛을 나타낼 수도 있다. 윈도잉 함수들의 예들은, maxRE 윈도잉 함수, 동위상 윈도잉 함수 및 Kaiser 윈도잉 함수를 포함할 수도 있다. 복소 곱셈 유닛 (416)은 좌측 변경된 SHC (417A) 및 우측 변경된 SHC (417B)를 생성하기 위해 SHC (422)를 벡터들 (415A 및 415B) 각각으로 엘리먼트-바이-엘리먼트 복소 곱셈을 수행하는 유닛을 나타낸다.

[0143] 그 후, 바이노럴 렌더링 유닛 (402)은 오디오 재생 디바이스 (400)의 사용자 또는 다른 오퍼레이터에 의해 입력된 구성 데이터에 가능성있게 기초하여 대칭 최적화 유닛 (418) 또는 비대칭 최적화 유닛 (420)의 어느 것을 호출할 수도 있다. 즉, 불규칙적 BRIR 필터들 (407A)이 SHC (422)의 바이노럴화 동안에 이용될 것으로 사용자가 특정할 때, 바이노럴 렌더링 유닛 (402)은 불규칙적 BRIR 필터들 (407A)이 대칭 또는 비대칭인지의 여부를 결정할 수도 있다. 즉, 불규칙적 BRIR 필터들 (407A) 모두가 비대칭인 것은 아니며, 대칭일 수도 있다. 불규칙적 BRIR 필터들 (407A)이 대칭이지만, 규칙적으로 이간되어 있는 것은 아닐 때, 바이노럴 렌더링 유닛 (402)은 좌측 및 우측 변경된 SHC (417A 및 417B)의 렌더링을 최적화하기 위해 대칭 최적화 유닛 (418)을 인보크한다. 불규칙적 BRIR 필터들 (407A)이 비대칭일 때, 바이노럴 렌더링 유닛 (402)은 좌측 및 우측 변경된 SHC (417A 및 417B)의 렌더링을 최적화하기 위해 비대칭 최적화 유닛 (420)을 인보크한다.

규칙적 BRIR 필터들 (407B 또는 407C)이 선택될 때, 바이노럴 렌더링 유닛 (402)은 좌측 및 우측 변경된 SHC (417A 및 417B)의 렌더링을 최적화하기 위해 대칭 최적화 유닛 (420)을 인보크한다.

[0144] 대칭 최적화 유닛 (418)은 인보크될 때 n 개의 차수들 및 m 개의 하위 차수들에 걸쳐 좌측 또는 우측 변경된

SHC (417A 또는 417B) 중 하나만을 합산할 수도 있다. 즉, 대칭 최적화 유닛 (418)은 주파수 도메인 좌측 스피커 피드 (419A)를 생성하기 위해 n 개의 차수들 및 m 개의 하위 차수들에 걸쳐 좌측 변경된 SHC (417A)를 합산할 수도 있다. 그 후, 대칭 최적화 유닛 (418)은 주파수 도메인 우측 스피커 피드 (419B)를 생성하기 위해 음의 하위 차수를 갖는 구면 기저 함수와 연관된 SHC (417A)의 것들을 반전시킨 다음, n 개의 차수들 및 m 개의 하위 차수들에 걸쳐 SHC (417A)의 반전된 버전에 대해 합산할 수도 있다. 비대칭 최적화 유닛 (420)은 인보크될 때, 주파수 도메인 좌측 스피커 피드 (421A) 및 주파수 도메인 우측 스피커 피드 (421B)를 각각 생성하기 위해 n 개의 차수들 및 m 개의 하위 차수들에 걸쳐 좌측 변경된 SHC (417A) 및 우측 변경된 SHC (417B) 각각을 합산한다. 역 시간 주파수 분석 유닛 (422)은 좌측 스피커 피드 (436A) 및 우측 스피커 피드 (436B)를 생성하도록 주파수 도메인 좌측 스피커 피드 (419A 또는 421A)의 어느 것 및 대응하는 주파수 도메인 우측 스피커 피드 (419B 또는 421A)의 어느 것을 주파수 도메인으로부터 시간 도메인으로 변환하는 유닛을 나타낼 수도 있다.

[0145] 이러한 식으로, 기법들은 하나 이상의 프로세서들을 포함하는 디바이스 (400)로 하여금 사운드 필드를 렌더링하도록 3차원으로 사운드 필드를 나타내는 구면 조화 계수들에 바이노렐 룸 임펄스 응답 필터를 적용하게 한다.

[0146] 일부 예들에서, 하나 이상의 프로세서들은 또한, 바이노렐 룸 임펄스 응답 필터를 적용할 때, 사운드 필드를 렌더링하도록 구면 조화 계수들에 불규칙적 바이노렐 룸 임펄스 응답 필터를 적용하도록 구성되며, 여기에서, 불규칙적 바이노렐 룸 임펄스 응답 필터들은 스피커들의 불규칙적 배열에 대한 하나 이상의 바이노렐 룸 임펄스 필터들을 포함한다.

[0147] 일부 예들에서, 하나 이상의 프로세서들은 또한, 바이노렐 룸 임펄스 응답 필터를 적용할 때, 사운드 필드를 렌더링하도록 구면 조화 계수들에 규칙적 바이노렐 룸 임펄스 응답 필터를 적용하도록 구성되며, 여기에서, 규칙적 바이노렐 룸 임펄스 응답 필터들은 스피커들의 규칙적 배열에 대한 하나 이상의 바이노렐 룸 임펄스 필터들을 포함한다.

[0148] 일부 예들에서, 하나 이상의 프로세서들은 또한, 규칙적 바이노렐 룸 임펄스 응답 필터를 생성하기 위해 불규칙적 바이노렐 룸 임펄스 응답을 보간하도록 구성된다. 이들 및 다른 예들에서, 불규칙적 바이노렐 룸 임펄스 응답 필터들은 스피커들의 불규칙적 배열에 대한 하나 이상의 바이노렐 룸 임펄스 응답 필터들을 포함하고 규칙적 바이노렐 룸 임펄스 응답 필터들은 스피커들의 규칙적 배열에 대한 하나 이상의 바이노렐 룸 임펄스 응답 필터들을 포함한다. 이들 및 다른 예들에서, 하나 이상의 프로세서들은 또한, 바이노렐 룸 임펄스 응답 필터를 적용할 때, 사운드 필드를 렌더링하도록 구면 조화 계수들에 규칙적 바이노렐 룸 임펄스 응답 필터를 적용하도록 구성된다.

[0149] 일부 예들에서, 하나 이상의 프로세서들은 또한, 원도잉된 바이노렐 룸 임펄스 응답 필터를 생성하기 위해 바이노렐 룸 임펄스 응답 필터에 원도잉 함수를 적용하도록 구성된다. 이들 및 다른 예들에서, 하나 이상의 프로세서들은 또한, 바이노렐 룸 임펄스 응답 필터를 적용할 때, 사운드 필드를 렌더링하도록 구면 조화 계수들에 원도잉된 바이노렐 룸 임펄스 응답 필터를 적용하도록 구성된다.

[0150] 일부 예들에서, 하나 이상의 프로세서들은 또한, 변환된 바이노렐 룸 임펄스 응답 필터를 생성하도록 바이노렐 룸 임펄스 응답 필터를 시간 도메인으로부터 주파수 도메인으로 변환하도록 구성된다. 이들 및 다른 예들에서, 하나 이상의 프로세서들은 또한, 바이노렐 룸 임펄스 응답 필터를 적용할 때, 사운드 필드를 렌더링하도록 구면 조화 계수들에 변환된 바이노렐 룸 임펄스 응답 필터를 적용하도록 구성된다.

[0151] 일부 예들에서, 하나 이상의 프로세서들은 또한, 변환된 바이노렐 룸 임펄스 응답 필터를 생성하도록 바이노렐 룸 임펄스 응답 필터를 시간 도메인으로부터 주파수 도메인으로 변환하고 변환된 구면 조화 계수들을 생성하도록 구면 조화 계수들을 시간 도메인으로부터 주파수 도메인으로 변환하도록 구성된다. 이들 및 다른 예들에서, 하나 이상의 프로세서들은 또한, 바이노렐 룸 임펄스 응답 필터를 적용할 때, 사운드 필드의 주파수 도메인 표현을 렌더링하도록, 변환된 구면 조화 계수들에 변환된 바이노렐 룸 임펄스 응답 필터를 적용하도록 구성된다. 이들 및 다른 예들에서, 하나 이상의 프로세서들은 또한, 사운드 필드를 렌더링하기 위해 사운드 필드의 주파수 도메인 표현에 역 변환을 적용하도록 구성된다.

[0152] 도 14 는 본 개시물에 설명된 바이노렐 오디오 렌더링 기법들의 여러 양태들을 수행할 수도 있는 오디오 재생 디바이스의 일 예를 예시하는 블록도이다. 오디오 재생 디바이스 (500)는 도 1 의 오디오 재생 디바이스 (62)의 다른 예시적인 경우를 보다 자세하게 표현할 수도 있다. 오디오 재생 디바이스 (500)는, 도 13 의 오디오 재생 디바이스 (400)에 대하여 위에 설명된 것들과 유사한 동작들을 수행하는 오디오 재생 디바이스

(500) 가 추출 유닛 (404), BRIR 선택 유닛 (424), 및 바이노럴 렌더링 유닛 (402) 을 포함하는 점에서, 도 13 의 오디오 재생 디바이스 (400) 와 유사할 수도 있다.

[0153] 그러나, 오디오 재생 디바이스 (500) 는 또한, 차수 감소된 SHCs (502) 를 생성하기 위해 SHCs (422) 의 차수 또는 하위 차수를 감소시키도록 인바운드 SHCs (422) 를 프로세싱하는 차수 감소 유닛 (504) 을 포함할 수도 있다. 차수 감소 유닛 (504) 은 SHC (422) 로부터 하나 이상의 하위 차수들 (m) 또는 차수들 (n) 을 제거하기 위해 SHC (422) 의 분석, 이를 테면, 에너지 분석, 방향성 분석, 및 다른 형태의 분석들 또는 이들의 조합들에 기초하여 이 차수 감소를 수행할 수도 있다. 에너지 분석은 SHC (422) 에 대하여 특이 값 분해를 수행하는 것을 포함할 수도 있다. 방향성 분석은 또한 SHC (422) 에 대해 특이 값 분해를 수행하는 것을 포함할 수도 있다. 따라서, SHC (502) 는 SHC (422) 보다 적은 차수들 및/또는 하위 차수들을 포함할 수도 있다.

[0154] 차수 감소 유닛 (504) 은 SHC (502) 를 생성하기 위해 제거되었던 SHC (422) 의 차수들 및/또는 하위 차수들을 식별하는 차수 감소 데이터 (506) 를 또한 생성할 수도 있다. 차수 감소 유닛 (504) 은 바이노럴 렌더링 유닛 (402) 애, 이 차수 감소 데이터 (506) 및 차수 감소된 SHC (502) 를 제공할 수도 있다. 오디오 재생 디바이스 (500) 의 바이노럴 렌더링 유닛 (402) 은 오디오 재생 디바이스 (500) 의 바이노럴 렌더링 유닛 (402) 이 차수 감소된 SHC (502) 에 기초하여 렌더러들 (405) 중 여러 렌더러들을 변경할 수도 있으면서 또한 (비-차수 감소된 SHC (422) 가 아닌) 차수 감소된 SHC (502) 에 대하여 동작한다는 점을 제외하고는 오디오 재생 디바이스 (400) 의 바이노럴 렌더링 유닛 (402) 과 실질적으로 유사하게 기능할 수도 있다. 오디오 재생 디바이스 (500) 의 바이노럴 렌더링 유닛 (402) 은 SHC (422) 의 제거된 차수들 및/또는 하위 차수들을 렌더링하는 것을 담당하는 렌더러들 (405) 의 이들 부분들을 적어도 부분적으로 제거함으로써, 차수 감소 데이터 (506) 에 기초하여 렌더러들 (405) 을 수정, 변경 또는 결정할 수도 있다. 차수 감소를 수행하는 것은, 일반적으로 (의도된 사운드 필드의 재생을 달리 왜곡시키거나 현저한 아티팩트들을 도입하는 관점에서) 오디오 재생에 상당한 영향을 주지 않고 (프로세서 사이클들 및/또는 메모리 소모의 관점에서) SHC (422) 의 바이노럴화와 연관된 계산 복잡도를 감소시킬 수도 있다.

[0155] 도 13 및 도 14 의 예에서 도시되고 이 개시물에 설명된 기법들은 주파수 도메인에서 규칙적 또는 불규칙적 BRIRs 의 세트를 통하여 3D 사운드 필드들을 바이노럴화하는 효과적인 방식을 제공할 수도 있다. 예를 들어, BRIRs 의 불규칙적 세트 (407A) 가 SHCs (422) 를 렌더링하기 위해 바이노럴 렌더링 유닛 (402) 에 의해 이용되면, 일부 경우들에서 바이노럴 렌더링 유닛 (402) 은 BRIR 세트를 BRIRs 의 규칙적 이간 세트 (407C) 로 보간할 수도 있다. 이 보간은 선형 보간, 벡터 기반 진폭 패닝 (VBAP) 등을 통하여 수행될 수도 있다. 아직 주파수 도메인에 있지 않으면, 이용될 BRIR 세트 (또는 "선택된 BRIR 세트") 는 예를 들어, 고속 푸리에 변환 (FFT), 이산 푸리에 변환 (DFT), 이산 코사인 변환 (DCT), 변경된 DCT (MDCT), 및 데시메이트된 신호 대각화 (DSD) 를 이용하여 주파수 도메인으로 변환될 수도 있다. 그 후, 바이노럴 렌더링 유닛 (402) 은 규칙적 BRIR 필터들 (407B) 또는 불규칙적 BRIR 필터들 (407A) 중 어느 것의 이전 선택에 의존하여 규칙적 렌더러 (405B) 또는 불규칙적 렌더러 (405A) 에 이용될 BRIR 세트를 복소 곱셈할 수도 있다. 규칙적 렌더러 (405B) 또는 불규칙적 렌더러 (405A) 의 차수 (N) 는 $N \leq N_I$ (여기에서, N_I 는 인커밍 HOA 신호의 입력 차수 또는 전체 차수임) 이도록, 인커밍 HOA 신호 (예를 들어, SHCs (422)) 의 전체 차수를 이용하는 선택에 의해 결정될 수도 있다. 도 14 의 예에서 차수 감소 동작을 적용하는 차수 감소 유닛 (504) 은 또한 렌더러들 (405A, 406B) 양쪽 모두와 또한 BRIR 보간에 요구되는 라우드스피커들의 수 (L) 에 또한 영향을 줄 수도 있다. 그러나, BRIR 세트의 정규화가 선택되지 않으면, 이용될 BRIR 세트로부터의 L 의 값은 차수 감소 (504) 와 또한 렌더러 (405A, 406B) 로 퍼드 백워드될 수도 있다.

[0156] 이용될 BRIR 세트와 렌더러들 (405A, 406B) 중 적절한 렌더러의 복소 곱셈 이후에, 출력된 신호들 (411A, 411B) 은 바이노럴화 HOA 렌더러 신호들 (413A, 413B) 을 생성하도록 L 차원에 걸쳐 합산될 수도 있다. 렌더링을 추가로 강화하기 위해, 주파수에 대한 n, m (여기에서 m 은 HOA 하위 차수) 의 가중이 윈도잉 함수들, 이를 테면, maxRe, 동위상 또는 Kaiser 를 이용하여 변경될 수도 있도록, 윈도우 블록이 도입될 수도 있다. 이들 윈도우들은 음향 심리적 기준들을 충족시키도록 오브젝티브 대책을 제공하는 Gerzon 에 의해 설정된 통상적인 앰비소닉 기준들을 충족시키는 것을 도울 수도 있다. 이 동작적 윈도우 후에, 바이노럴 렌더링 유닛 (402) 은 바이노럴 HOA 신호들 (417A, 417B) (이들 신호들은 좌측, 우측 변경된 SHCs (417A, 417B) 로서 이 개시물 다른 곳에서 설명된 것의 예들임) 을 생성하기 위해 바이노럴화된 HOA 렌더러 신호들 (415A, 415B) 과 HOA 신호를 복소 곱셈한다. 기법들은 또한 일부 경우들에서, 대칭적 BRIR 최적화를 고려할 수도 있다. 바이노럴 렌더링 유닛 (402) 이 비대칭적 최적화를 적용하면, 바이노럴 렌더링 유닛 (402) 은 좌측 및 우측 채널들에 대한 n, m HOA 계수들을 합산한다. 그러나, 바이노럴 렌더링 유닛 (402) 이 대칭적 최적화를 적용하면, 바이노

럴 렌더링 유닛 (402) 은 좌측 채널에 대한 n, m HOA 계수들을 합산하여 출력한다. 그러나, 구면 조화 기저 함수들의 대칭성으로 인하여, $m < 0$ 에 대한 값들은 합산 전에 반전된다. 이 대칭성은 BRIR 세트의 좌측만을 결정하는 경우, 위에 설명된 기법들 전반에 걸쳐 역방향으로 적용될 수도 있다. 바이노럴 렌더링 유닛 (402) 은 바이노럴 출력 (436A, 436B) 에 대한 시간 도메인 (역 변환) 으로 좌측 및 우측 신호들을 다시 변환할 수도 있다.

[0157] 이러한 식으로, 기법들은 a) (단지 2D 가 아닌) 3D 의 포함, b) (단지 1차 앰비소닉이 아닌) 고차 앰비소닉의 바이노럴화, c) 규칙적 또는 불규칙적 BRIR 세트들의 적용, d) 불규칙적 또는 규칙적 BRIR 세트들로부터의 BRIRs 의 보간, e) 앰비소닉 재생 기준들에 보다 잘 매칭하는 BRIR 신호의 원도잉; 및 f) 시간 도메인 연산이 아닌 주파수 도메인 연산의 이점을 적어도 부분적으로 취함으로써 가능성있게 연산적으로 효율을 개선할 수도 있다.

[0158] 도 15 는 본 개시물에 설명된 기법들에 따라 구면 조화 계수들을 렌더링하기 위한 바이노럴 렌더링 디바이스에 대한 예시적인 동작 모드를 예시하는 흐름도이다. 예시 목적들을 위하여, 예시적인 동작 모드는 도 13 의 오디오 재생 디바이스 (400) 에 대하여 설명된다.

[0159] 추출 유닛 (404) 은 비트스트림 (420) 으로부터 인코딩된 오디오 데이터를 추출할 수도 있다. 추출 유닛 (404) 은 추출된 인코딩된 오디오 데이터를 구면 조화 계수들 (SHCs) (422) (이는 또한 SHCs (422) 가 1 보다 큰 차수와 연관된 적어도 하나의 계수를 포함할 수도 있다는 점에서 고차 앰비소닉 (HOA) 으로 지칭될 수도 있음) 의 형태로 바이노럴 렌더링 유닛 (146) 에 포워딩 할 수도 있다. SHCs (422) 가 비트스트림 (420) 에서 주파수 도메인으로 이미 특정되었다면, 시간-주파수 분석 유닛은 변환을 적용하거나 수신된 SHCs (422) 를 달리 변환하는 것이 없이 바이노럴 렌더링 유닛 (402) 으로 SHCs (422) 를 패스시킬 수도 있다. 주파수 도메인에서 특정된 SHCs (422) 에 관하여 설명되어 있지만, 기법들은 시간 도메인에서 특정된 SHCs (422) 에 관하여 수행될 수도 있다.

[0160] 즉, 어느 이벤트에도, 바이노럴 렌더링 유닛 (402) 은 SHCs (422) 가 기록되었던 룸에서의 청취자에 의해 좌측 및 우측 채널이 청취되는 방법을 모델링하기 위해 공간화를 피쳐링할 수도 있는, 좌측 및 우측 채널로 SHCs (422) 를 렌더링하도록 구성된 유닛을 표현할 수도 있다. 바이노럴 렌더링 유닛 (402) 은 헤드셋, 이를 테면, 헤드폰들을 통하여 재생에 적합한 좌측 채널 (436A) 및 우측 채널 (436B)(이들은 총괄적으로 "채널 (436)" 로서 지칭될 수도 있음) 을 생성하기 위해 SHCs (422) 를 렌더링 할 수도 있다.

[0161] 바이노럴 렌더링 유닛 (402) 은 불규칙적 BRIR 필터 (407A), 규칙적 BRIR 필터 (407B) 및/또는 보간된 BRIR 필터 (407C) 에 대하여 바이노럴 렌더링을 수행할지의 여부를 결정하도록 사용자 구성 데이터 (603) 를 수행할 수도 있다. 즉, 바이노럴 렌더링 유닛 (402) 은 SHC (422) 의 바이노럴화를 수행할 때 어느 필터들 (407) 이 이용되어야 하는지를 선택하는 사용자 구성 데이터 (603) 를 수신할 수도 있다 (602). 사용자 구성 데이터 (603) 는 도 13 및 도 14 의 신호 (426) 의 일 예를 나타낼 수도 있다. 사용자 구성 데이터 (603) 가, 규칙적 BRIR 필터 (407B) 가 이용됨 ("예" (604)) 을 특정할 때, 바이노럴 렌더링 유닛 (402) 은 규칙적 BRIR 필터 (407B) 및 규칙적 렌더러 (405B) 를 선택한다 (606). 사용자 구성 데이터 (603) 가, 불규칙적 BRIR 필터 (407A) 가 이용되고 ("아니오" (604)) 이 필터 (407A) 를 보간하지 않음 ("아니오" (608)) 을 표시할 때, 바이노럴 렌더링 유닛 (402) 은 불규칙적 BRIR 필터 (407A) 및 불규칙적 렌더러 (405A) 를 선택한다 (610). 사용자 구성 데이터 (603) 가, 불규칙적 BRIR 필터 (407A) 가 이용되지만 ("아니오" (604)) 이 BRIR 필터 (407A) 가 보간될 것임 ("예" (608)) 을 표시할 때, 바이노럴 렌더링 유닛 (402) 은 (필터 (407C) 를 생성하기 위해 선택된 필터 (407A) 를 보간하도록 보간 유닛 (406) 을 인보크한 후) 보간된 BRIR 필터 (407C) 및 규칙적 렌더러 (405B) 를 선택한다 (612).

[0162] 임의의 이벤트에서, BRIR 필터들이 (407A-407C) (어느 것이 SHC (422) 를 바이노럴화하는데 선택되는지에 의존하여) 시간 도메인으로 제시될 때, 바이노럴 렌더링 유닛 (402) 은 BRIR 필터들 (407A-407C) ("BRIR 필터들 (407)") 중 선택된 하나를 시간 도메인으로부터 주파수 도메인으로 변환하도록 시간 주파수 분석 유닛 (408) 을 인보크할 수도 있고, 이는 변환된 BRIR 필터들 (409A-409C) ("BRIR 필터들 (409)") 을 각각 발생시킨다. 복소 BRIR 유닛 (410) 은 2 개의 BRIR 렌더링 백터들 (411A 및 411B) 을 생성하기 위해 BRIR 필터 (409) 의 선택된 필터와 렌더러들 (405) 의 선택된 하나에 대하여 엘리먼트-바이-엘리먼트 복소 곱셈 및 합산을 수행할 수도 있다 (614).

[0163] 합산 유닛 (442) 은 합산된 BRIR 렌더링 백터들 (413A 및 413B) 을 생성하기 위해 BRIR 렌더링 백터들 (411A 및 411B) 각각을 합산할 수도 있다 (616). 원도잉 유닛은 원도잉된 BRIR 렌더링 백터들 (415A 및 415B) 을

생성하기 위해 합산된 BRIR 렌더링 벡터들 (413A 및 415B)에 윈도잉 함수를 적용할 수도 있다 (618). 그 후, 복소 곱셈 유닛 (416)은 좌측 변경된 SHC (417A) 및 우측 변경된 SHC (417B)를 생성하기 위해 SHC (422)를 벡터들 (415A 및 415B) 각각으로 엘리먼트-바이-엘리먼트 복소 곱셈을 수행할 수도 있다 (620).

[0164] 그 후, 바이노럴 렌더링 유닛 (402)은 위에 설명된 바와 같이, 오디오 재생 디바이스 (400)의 사용자 또는 다른 오퍼레이터에 의해 입력된 구성 데이터 (603)에 가능성있게 기초하여 대칭 최적화 유닛 (418) 또는 비대칭 최적화 유닛 (420)의 어느 것을 호출할 수도 있다.

[0165] 대칭 최적화 유닛 (418)은 인보크될 때 n 개의 차수들 및 m 개의 하위 차수들에 걸쳐 좌측 또는 우측 변경된 SHC (417A 또는 417B) 중 하나만을 합산할 수도 있다. 즉, 대칭 최적화 유닛 (418)은 주파수 도메인 좌측 스피커 피드 (419A)를 생성하기 위해 n 개의 차수들 및 m 개의 하위 차수들에 걸쳐 좌측 변경된 SHC (417A)를 합산할 수도 있다. 그 후, 대칭 최적화 유닛 (418)은 주파수 도메인 우측 스피커 피드 (419A)를 생성하기 위해 음의 하위 차수를 갖는 구면 기저 함수와 연관된 SHC (417A)의 것들을 반전시킨 다음, n 개의 차수들 및 m 개의 하위 차수들에 걸쳐 SHC (417A)의 이 버전에 대해 합산할 수도 있다.

[0166] 비대칭 최적화 유닛 (420)은 인보크될 때, 주파수 도메인 좌측 스피커 피드 (421A) 및 주파수 도메인 우측 스피커 피드 (421B)를 각각 생성하기 위해 n 개의 차수들 및 m 개의 하위 차수들에 걸쳐 좌측 변경된 SHC (417A) 및 우측 변경된 SHC (417B) 각각을 합산한다. 역 시간 주파수 분석 유닛 (422)은 좌측 스피커 피드 (436A) 및 우측 스피커 피드 (436B)를 생성하도록 주파수 도메인 좌측 스피커 피드 (419A 또는 421A)의 어느 것 및 대응하는 주파수 도메인 우측 스피커 피드 (419B 또는 421A)의 어느 것을 주파수 도메인으로부터 시간 도메인으로 변환하는 유닛을 나타낼 수도 있다. 이러한 방식으로, 바이노럴 렌더링 유닛 (402)은 좌측 및 우측 스피커 피드들 (436A 및 436B)을 생성하기 위해 좌측 및 우측 SHC (417A 및 417B) 중 하나 이상에 대하여 최적화를 수행할 수도 있다 (622). 오디오 재생 디바이스 (400)는 좌측 스피커 피드 (436A) 및 우측 스피커 피드 (436B)를 렌더링하기 위해 위에 설명된 방식으로 동작을 진행하고, SHC (422)를 추출 및 바이노럴화할 수도 있다 (600-622).

[0167] 도 16a 및 도 16b는 이 개시물에 설명된 기법들의 여러 양태들에 따라 도 13의 오디오 재생 디바이스 (400) 및 도 14의 오디오 재생 디바이스 (500)에 의해 수행될 수도 있는 개념적인 프로세스를 각각 예시하는 다이어그램을 표시한다. 고차 앰비소닉 (HOA) 계수들로 구성되는 공간 사운드 필드의 바이노럴화는 통상적으로, HOA 신호들을 라우드스피커 신호들에 렌더링한 다음, 라우드스피커 포지션에 대하여 취해진 BRIR의 좌측 및 우측 버전들로 라우드스피커 신호들을 컨볼빙하는 것을 포함한다. 이 통상적인 방법론은 HOA 계수들이 존재하는 것보다 더 많은 라우드스피커들을 갖는 경우에, (L 개의 라우드스피커들의) 라우드스피커 신호 당 2 개의 컨볼루션들을 일반적으로 요구하기 때문에 계산적으로 고비용일 수도 있다. 즉, 다중채널 (periphonic) 라우드스피커에 대하여 $L > (N+1)^2$ 이며 여기에서 N 앰비소닉 차수이다. 2 차원에 걸쳐 사운드 필드를 정의하는 고전 1차 앰비소닉에 대한 방법론은 1차 앰비소닉 컨텐츠를 재생하기 위하여 규칙적 (일부 경우에 규칙적의 의미는 등간격으로 이격된을 의미함) 가상 라우드스피커 배열들을 다룬다. 이 방법론이 최상의 경우의 시나리오이고 고차 앰비소닉 또는 3차원으로의 이것의 적용에 대한 정보가 제공되지 않는 것으로 보면 이 방법론은 지나치게 단순화된 것으로 고려될 수도 있다. 이 방법론은 또한 주파수 도메인의 언급없이 시간 도메인 내에서 컨볼루션에 의존하였다.

[0168] 도 8의 예에서 도시되고 이 개시물에 설명된 기법들은 주파수 도메인에서 규칙적 또는 불규칙적 BRIRs의 세트를 통하여 3D 사운드 필드들을 바이노럴화하는 효과적인 방식을 제공할 수도 있다. BRIRs의 불규칙적 세트가 이용되면, BRIR 세트를 BRIRs의 규칙적으로 이간된 세트로 보간하는 선택이 있을 수도 있다. 이 보간은 선형 보간, 벡터 기반 진폭 패닝 (VBAP) 등을 통하여 수행될 수도 있다. 도 16a에 도시된 바와 같이, 주파수 도메인에서 아직 있지 않으면, 이용될 BRIR 세트는 일부 예들에서, 몇몇 예들을 제공하기 위해 고속 푸리에 변환 (FFT), 이산 푸리에 변환 (DFT), 이산 코사인 변환 (DCT), MDCT 및 DSD을 이용하여 주파수 도메인으로 변환될 수도 있다. BRIR 세트는 이전의 규칙적/불규칙적 선택에 의존하여 규칙적 또는 불규칙적 렌더러와 복소 곱셈될 수도 있다. 규칙적 또는 불규칙적 렌더러의 차수 (N)는 $N \leq N_1$ 이도록, 인카밍 HOA 신호의 전체 차수를 이용하는 선택에 의해 조정될 수도 있다. 도 16a, 도 16b의 예에서의 '차수 감소' 블록은 또한 렌더러 양쪽 모두와 또한 BRIR 보간에 요구되는 라우드스피커들의 수 (L)에 영향을 줄 수도 있다. 그러나, BRIR 세트의 정규화가 선택되지 않으면, 이용될 BRIR 세트로부터의 L 의 값은 차소 감소와 또한 렌더러로 피드 백워드될 수도 있다.

[0169] 정확한 BRIR 신호 세트와 정확한 렌더러의 복소 곱셈 후에, 출력된 신호들은 바이노럴화된 HOA 렌더러 신호들을

생성하기 위해 L 차원에 걸쳐 합산될 수도 있다. 렌더링을 추가로 강화하기 위해, 주파수에 대한 n, m 의 가중이 원도잉 함수들, 이를 테면, maxRe, 동위상 또는 Kaiser 를 이용하여 변경될 수도 있도록, 원도우 블록이 도입될 수도 있다. 이들 원도우들은 음향 심리적 기준들을 충족시키도록 오브젝티브 대책을 제공하는 Gerzon 에 의해 설정된 통상적인 앰비소닉 기준들을 충족시키는 것을 도울 수도 있다. 이 선택된 원도우 후, (만약 도 16a 에 도시된 바와 같이 주파수 도메인에 있다면) HOA 는 바이노럴화된 HOA 렌더러 신호들과 복소 곱셈된다. HOA 가 시간 도메인에 있다면, HOA 는 도 16b 에 묘사된 바와 같이, 바이노럴화된 HOA 신호들과 고속 컨볼루션될 수도 있다.

[0170] 기법들은 또한 일부 경우들에서, 대칭적 BRIR 최적화를 고려할 수도 있다. 비최적화된 라우트가 수행되면, n, m HOA 계수들은 좌측 및 우측 채널들에 대하여 합산될 수도 있다. 대칭 경로가 선택되면, 좌측에 대해 출력된 신호는 n, m 값들의 합이지만, 구면 조화 기저 함수들의 대칭성으로 인하여, 합산 전에 $m < 0$ 의 값이 반전된다. 이 대칭성은 BRIR 세트의 좌측만을 결정하는 경우, 위에 설명된 기법들 전반에 걸쳐 역방향으로 적용될 수도 있다. 그 후, 좌측 및 우측 신호들은 바이노럴 출력을 위하여 시간 도메인으로 다시 변환 (역 변환) 될 수도 있다.

[0171] 기법들은 a) (단지 2D 가 아닌) 3D 의 포함, b) (단지 1차 앰비소닉이 아닌) 고차 앰비소닉의 바이노럴화, c) 규칙적 또는 불규칙적 BRIR 세트들의 적용, d) 불규칙적 또는 규칙적 BRIR 세트들로부터의 BRIRs 의 보간, e) 앰비소닉 재생 기준들에 보다 잘 매칭하는 BRIR 신호의 원도잉; 및 f) 시간 도메인 연산이 아닌 주파수 도메인 연산의 이점을 적어도 부분적으로 취함으로써 가능성있게 연산적으로 효율을 개선할 수도 있다.

[0172] 위의 것에 더하여, 또는 위의 것에 대안으로서, 다음 예들이 설명된다. 다음 예들 중 어느 것에서 설명된 피쳐들은 여기에 설명된 다른 예들의 어느 것과 함께 이용될 수도 있다.

[0173] 일 예는, 사운드 필드를 렌더링하도록 3차원으로 사운드 필드를 나타내는 구면 조화 계수들에 바이노럴 룸 임펄스 응답 필터를 적용하는 단계를 포함하는 바이노럴 오디오 렌더링의 방법에 대해 지시된다.

[0174] 일부 예들에서, 바이노럴 룸 임펄스 응답 필터를 적용하는 것은, 사운드 필드를 렌더링하도록 구면 조화 계수들에 불규칙적 바이노럴 룸 임펄스 응답 필터를 적용하는 것을 포함하며, 여기에서, 불규칙적 바이노럴 룸 임펄스 응답 필터들은 스피커들의 불규칙적 배열에 대한 하나 이상의 바이노럴 룸 임펄스 필터들을 포함한다.

[0175] 일부 예들에서, 바이노럴 룸 임펄스 응답 필터를 적용하는 것은, 사운드 필드를 렌더링하도록 구면 조화 계수들에 규칙적 바이노럴 룸 임펄스 응답 필터를 적용하는 것을 포함하며, 여기에서, 규칙적 바이노럴 룸 임펄스 응답 필터들은 스피커들의 규칙적 배열에 대한 하나 이상의 바이노럴 룸 임펄스 필터들을 포함한다.

[0176] 일부 예들에서, 구면 조화 계수들이 대응하는 구면 기저 함수들의 차수는 1 보다 크다.

[0177] 일부 예들에서, 본 방법은 규칙적 바이노럴 룸 임펄스 응답 필터를 생성하기 위해 불규칙적 바이노럴 룸 임펄스 응답 필터를 보간하는 단계를 더 포함하고, 불규칙적 바이노럴 룸 임펄스 응답 필터는 스피커들의 불규칙적 배열에 대한 하나 이상의 바이노럴 룸 임펄스 응답 필터들을 포함하고, 규칙적 바이노럴 룸 임펄스 응답 필터는 스피커들의 규칙적 배열에 대한 하나 이상의 바이노럴 룸 임펄스 응답 필터들을 포함하며, 그리고 바이노럴 룸 임펄스 응답 필터를 적용하는 것은 사운드 필드를 렌더링하도록 구면 조화 계수들에 규칙적 바이노럴 룸 임펄스 응답 필터를 적용하는 것을 포함한다.

[0178] 일부 예들에서, 본 방법은 원도잉된 바이노럴 룸 임펄스 응답 필터를 생성하기 위해 바이노럴 룸 임펄스 응답 필터에 원도잉 함수를 적용하는 단계를 더 포함하고, 바이노럴 룸 임펄스 응답 필터를 적용하는 것은 사운드 필드를 렌더링하도록 구면 조화 계수들에 원도잉된 바이노럴 룸 임펄스 응답 필터를 적용하는 것을 포함한다.

[0179] 일부 예들에서, 본 방법은 변환된 바이노럴 룸 임펄스 응답 필터를 생성하도록 바이노럴 룸 임펄스 응답 필터를 시간 도메인으로부터 주파수 도메인으로 변환하는 단계를 더 포함하고, 바이노럴 룸 임펄스 응답 필터를 적용하는 것은 사운드 필드를 렌더링하도록 구면 조화 계수들에 변환된 바이노럴 룸 임펄스 응답 필터를 적용하는 것을 포함한다.

[0180] 일부 예들에서, 본 방법은 변환된 바이노럴 룸 임펄스 응답 필터를 생성하도록 바이노럴 룸 임펄스 응답 필터를 시간 도메인으로부터 주파수 도메인으로 변환하는 단계; 및 변환된 구면 조화 계수들을 생성하도록 구면 조화 계수들을 시간 도메인으로부터 주파수 도메인으로 변환하는 단계를 더 포함하고, 바이노럴 룸 임펄스 응답 필터를 적용하는 것은 사운드 필드의 주파수 도메인 표현을 렌더링하도록 변환된 바이노럴 룸 임펄스 응답 필터를 변환된 구면 조화 함수들에 적용하는 것을 포함하고, 상기 방법은 사운드 필드를 렌더링하기 위해 사운드 필드

의 주파수 도메인 표현에 역 변환을 적용하는 단계를 더 포함한다.

[0181] 일 예는 사운드 필드를 렌더링하도록 3차원으로 사운드 필드를 나타내는 구면 조화 계수들에 바이노럴 룸 임펄스 응답 필터를 적용하도록 구성되는 하나 이상의 프로세서들을 포함하는 디바이스에 대해 지시된다.

[0182] 일부 예들에서, 하나 이상의 프로세서들은 또한, 바이노럴 룸 임펄스 응답 필터를 적용할 때, 사운드 필드를 렌더링하도록 구면 조화 계수들에 불규칙적 바이노럴 룸 임펄스 응답 필터를 적용하도록 구성되며, 여기에서, 불규칙적 바이노럴 룸 임펄스 응답 필터들은 스피커들의 불규칙적 배열에 대한 하나 이상의 바이노럴 룸 임펄스 필터들을 포함한다.

[0183] 일부 예들에서, 하나 이상의 프로세서들은 또한, 바이노럴 룸 임펄스 응답 필터를 적용할 때, 사운드 필드를 렌더링하도록 구면 조화 계수들에 규칙적 바이노럴 룸 임펄스 응답 필터를 적용하도록 구성되며, 여기에서, 규칙적 바이노럴 룸 임펄스 응답 필터들은 스피커들의 규칙적 배열에 대한 하나 이상의 바이노럴 룸 임펄스 필터들을 포함한다.

[0184] 일부 예들에서, 구면 조화 계수들이 대응하는 구면 기저 함수들의 차수는 1 보다 크다.

[0185] 일부 예들에서, 하나 이상의 프로세서들은 또한, 규칙적 바이노럴 룸 임펄스 응답 필터를 생성하기 위해 불규칙적 바이노럴 룸 임펄스 응답 필터를 보간하도록 구성되고, 불규칙적 바이노럴 룸 임펄스 응답 필터들은 스피커들의 불규칙적 배열에 대한 하나 이상의 바이노럴 룸 임펄스 응답 필터들을 포함하고, 규칙적 바이노럴 룸 임펄스 응답 필터들은 스피커들의 규칙적 배열에 대한 하나 이상의 바이노럴 룸 임펄스 응답 필터들을 포함하며, 그리고 하나 이상의 프로세서들은 또한, 바이노럴 룸 임펄스 응답 필터를 적용할 때, 사운드 필드를 렌더링하기 위해 구면 조화 계수들에 규칙적 바이노럴 룸 임펄스 응답 필터를 적용하도록 구성된다.

[0186] 일부 예들에서, 하나 이상의 프로세서들은 또한, 원도잉된 바이노럴 룸 임펄스 응답 필터를 생성하기 위해 바이노럴 룸 임펄스 응답 필터에 원도잉 함수를 적용하도록 구성되고, 하나 이상의 프로세서들은 바이노럴 룸 임펄스 응답 필터를 적용할 때, 사운드 필드를 렌더링하기 위해 구면 조화 계수들에 원도잉된 바이노럴 룸 임펄스 응답 필터를 적용하도록 구성된다.

[0187] 일부 예들에서, 하나 이상의 프로세서들은 또한, 변환된 바이노럴 룸 임펄스 응답 필터를 생성하기 위해 바이노럴 룸 임펄스 응답 필터를 시간 도메인으로부터 주파수 도메인으로 변환하도록 구성되고, 하나 이상의 프로세서들은 또한, 바이노럴 룸 임펄스 응답 필터를 적용할 때, 사운드 필드를 렌더링하기 위해 구면 조화 계수들에 변환된 바이노럴 룸 임펄스 응답 필터를 적용하도록 구성된다.

[0188] 일부 예들에서, 하나 이상의 프로세서들은 또한, 변환된 바이노럴 룸 임펄스 응답 필터를 생성하기 위해 바이노럴 룸 임펄스 응답 필터를 시간 도메인으로부터 주파수 도메인으로 변환하고; 그리고 변환된 구면 조화 계수들을 생성하기 위해 구면 조화 계수들을 시간 도메인으로부터 주파수 도메인으로 변환하도록 구성되고, 하나 이상의 프로세서들은 또한, 바이노럴 룸 임펄스 응답 필터를 적용할 때, 사운드 필드의 주파수 도메인 표현을 렌더링하기 위해 변환된 바이노럴 룸 임펄스 응답 필터를 변환된 구면 조화 함수들에 적용하도록 구성되고, 하나 이상의 프로세서들은 또한, 사운드 필드를 렌더링하기 위해 사운드 필드의 주파수 도메인 표현에 역 변환을 적용하도록 구성된다.

[0189] 일 예는 사운드 필드를 3 차원으로 표현하는 구면 조화 계수들을 결정하는 수단, 및 사운드 필드를 렌더링하도록 사운드 필드를 나타내는 구면 조화 계수들에 바이노럴 룸 임펄스 응답 필터를 적용하는 수단을 포함하는 디바이스에 대해 지시된다.

[0190] 일부 예들에서, 바이노럴 룸 임펄스 응답 필터를 적용하는 수단은, 사운드 필드를 렌더링하도록 구면 조화 계수들에 불규칙적 바이노럴 룸 임펄스 응답 필터를 적용하는 수단을 포함하며, 여기에서, 불규칙적 바이노럴 룸 임펄스 응답 필터들은 스피커들의 불규칙적 배열에 대한 하나 이상의 바이노럴 룸 임펄스 필터들을 포함한다.

[0191] 일부 예들에서, 바이노럴 룸 임펄스 응답 필터를 적용하는 수단은, 사운드 필드를 렌더링하도록 구면 조화 계수들에 규칙적 바이노럴 룸 임펄스 응답 필터를 적용하는 수단을 포함하며, 규칙적 바이노럴 룸 임펄스 응답 필터들은 스피커들의 규칙적 배열에 대한 하나 이상의 바이노럴 룸 임펄스 필터들을 포함한다.

[0192] 일부 예들에서, 구면 조화 계수들이 대응하는 구면 기저 함수들의 차수는 1 보다 크다.

[0193] 일부 예들에서, 본 디바이스는 규칙적 바이노럴 룸 임펄스 응답 필터를 생성하기 위해 불규칙적 바이노럴 룸 임펄스 응답 필터를 보간하는 수단을 더 포함하고, 불규칙적 바이노럴 룸 임펄스 응답 필터는 스피커들의 불규칙

적 배열에 대한 하나 이상의 바이노럴 룸 임펄스 응답 필터들을 포함하고, 규칙적 바이노럴 룸 임펄스 응답 필터는 스피커들의 규칙적 배열에 대한 하나 이상의 바이노럴 룸 임펄스 응답 필터들을 포함하며, 그리고 바이노럴 룸 임펄스 응답 필터를 적용하는 수단은 사운드 필드를 렌더링하도록 구면 조화 계수들에 규칙적 바이노럴 룸 임펄스 응답 필터를 적용하는 수단을 포함한다.

[0194] 일부 예들에서, 본 디바이스는 원도잉된 바이노럴 룸 임펄스 응답 필터를 생성하기 위해 바이노럴 룸 임펄스 응답 필터에 원도잉 함수를 적용하는 수단을 더 포함하고, 바이노럴 룸 임펄스 응답 필터를 적용하는 수단은 사운드 필드를 렌더링하도록 구면 조화 계수들에 원도잉된 바이노럴 룸 임펄스 응답 필터를 적용하는 수단을 포함한다.

[0195] 일부 예들에서, 본 디바이스는 변환된 바이노럴 룸 임펄스 응답 필터를 생성하도록 바이노럴 룸 임펄스 응답 필터를 시간 도메인으로부터 주파수 도메인으로 변환하는 단계를 더 포함하고, 바이노럴 룸 임펄스 응답 필터를 적용하는 수단은 사운드 필드를 렌더링하도록 구면 조화 계수들에 변환된 바이노럴 룸 임펄스 응답 필터를 적용하는 수단을 포함한다.

[0196] 일부 예들에서, 본 디바이스는 변환된 바이노럴 룸 임펄스 응답 필터를 생성하도록 바이노럴 룸 임펄스 응답 필터를 시간 도메인으로부터 주파수 도메인으로 변환하는 수단; 및 변환된 구면 조화 계수들을 생성하도록 구면 조화 계수들을 시간 도메인으로부터 주파수 도메인으로 변환하는 수단을 더 포함하고, 바이노럴 룸 임펄스 응답 필터를 적용하는 수단은 사운드 필드의 주파수 도메인 표현을 렌더링하도록 변환된 바이노럴 룸 임펄스 응답 필터를 변환된 구면 조화 함수들에 적용하는 수단을 포함하고, 상기 디바이스는 사운드 필드를 렌더링하기 위해 사운드 필드의 주파수 도메인 표현에 역 변환을 적용하는 수단을 더 포함한다.

[0197] 일 예는 명령들이 저장된 비일시적 컴퓨터 판독가능 저장 매체에 대하여 지시되며, 명령들은 실행될 때 하나 이상의 프로세서들로 하여금 사운드 필드를 렌더링하도록 3차원으로 사운드 필드를 나타내는 구면 조화 계수들에 바이노럴 룸 임펄스 응답 필터를 적용하게 한다.

[0198] 또한, 위에 설명된 임의의 예들에서 전술한 특정 피쳐들은 상술한 기법들의 유익한 예에 결합될 수도 있다. 즉, 특정 피쳐들 중 임의의 것이 일반적으로 본 발명의 모든 예들에 적용가능하다. 본 발명의 여러 예들이 기술되었다.

[0199] 예에 따라서는, 본원에서 설명된 방법들 중 임의의 것의 어떤 액트들 또는 이벤트들이 상이한 시퀀스에서 수행될 수 있거나, 추가될 수도 있거나, 병합될 수도 있거나, 또는 함께 배제 (예를 들어, 모든 설명된 액트들 또는 이벤트들이 방법의 실시를 위해 필요한 것은 아님) 될 수도 있다는 것을 이해해야 한다. 또한, 어떤 예들에서는, 액트들 또는 이벤트들이 순차적인 것이 아니라, 예를 들어, 멀티-스레딩된 (multi-threaded) 프로세싱, 인터럽트 프로세싱, 또는 다수의 프로세서들을 통해 동시에 수행될 수도 있다. 게다가, 이 개시물의 어떤 양태들은 명료함의 목적들을 위하여 단일 모듈 또는 유닛에 의해 수행되는 것으로 설명되어 있지만, 이 개시물의 기법들은 비디오 코더와 연관된 유닛들 또는 모듈들의 조합에 의해 수행될 수도 있는 것을 이해해야 한다.

[0200] 하나 이상의 예들에서는, 설명된 기능들이 하드웨어, 소프트웨어, 펌웨어, 또는 그 임의의 조합으로 구현될 수도 있다. 소프트웨어로 구현될 경우, 기능들은 하나 이상의 명령들 또는 코드로서, 컴퓨터-판독가능한 매체 상에 저장되거나 컴퓨터-판독가능한 매체를 통해 송신될 수도 있고, 하드웨어-기반 프로세싱 유닛에 의해 실행될 수도 있다. 컴퓨터-판독가능한 매체들은 데이터 저장 매체들과 같은 유형의 매체에 대응하는 컴퓨터-판독가능한 저장 매체들, 또는 예를 들어, 통신 프로토콜에 따라 하나의 위치로부터 또 다른 위치로 컴퓨터 프로그램의 전송을 용이하게 하는 임의의 매체를 포함하는 통신 매체들을 포함할 수도 있다.

[0201] 이러한 방식으로, 컴퓨터-판독가능한 매체들은 일반적으로 (1) 비-일시적인 유형의 컴퓨터-판독가능한 저장 매체들, 또는 (2) 신호 또는 반송파와 같은 통신 매체에 대응할 수도 있다. 데이터 저장 매체들은 이 개시물에서 설명된 기법들의 구현을 위한 명령들, 코드 및/또는 데이터 구조들을 취출하기 위해 하나 이상의 컴퓨터들 또는 하나 이상의 프로세서들에 의해 액세스될 수 있는 임의의 이용가능한 매체들일 수도 있다. 컴퓨터 프로그램 제품은 컴퓨터-판독가능한 매체를 포함할 수도 있다.

[0202] 제한이 아닌 예로서, 이러한 컴퓨터-판독가능한 저장 매체들은 RAM, ROM, EEPROM, CD-ROM 또는 다른 광학 디스크 저장, 자기 디스크 저장, 또는 다른 자기 저장 디바이스들, 플래시 메모리, 또는 명령들 또는 데이터 구조들의 형태로 희망하는 프로그램 코드를 저장하기 위해 이용될 수 있으며 컴퓨터에 의해 액세스될 수 있는 임의의 다른 매체를 포함할 수 있다. 또한, 임의의 접속은 컴퓨터-판독가능한 매체로 적절하게 지정된다. 예를 들어, 동축 케이블, 광섬유 케이블, 트위스트 페어 (twisted pair), 디지털 가입자 회선 (digital subscriber

line; DSL), 또는 무선 기술들 예컨대, 적외선, 라디오 (radio), 및 마이크로파 (microwave) 를 이용하여 웹사이트, 서버, 또는 다른 원격 소스로부터 명령들이 송신될 경우, 동축 케이블, 광섬유 케이블, 트위스트 페어, DSL, 또는 무선 기술들 예컨대, 적외선, 라디오, 및 마이크로파는 매체의 정의 내에 포함된다.

[0203] 그러나, 컴퓨터-관련 가능한 저장 매체들 및 데이터 저장 매체들은 접속들, 반송파들, 신호들, 또는 다른 순시적 매체 (transient medium) 들을 포함하는 것이 아니라, 그 대신에, 비-순시적인, 유형의 저장 매체들에 대해 지시되는 것이라는 것을 이해해야 한다. 본원에서 이용된 바와 같은 디스크 (disk) 및 디스크 (disc) 는 컴팩트 디스크 (compact disc; CD), 레이저 디스크 (laser disc), 광학 디스크 (optical disc), 디지털 다기능 디스크 (digital versatile disc; DVD), 플로피 디스크 (floppy disk) 및 블루레이 디스크 (blu-ray disc) 를 포함하고, 여기서, 디스크 (disk) 들은 통상 데이터를 자기적으로 재생하는 반면, 디스크 (disc) 들은 데이터를 레이저로 광학적으로 재생한다. 상기의 조합들은 컴퓨터-관련 가능한 매체들의 범위 내에 또한 포함되어야 한다.

[0204] 명령들은 하나 이상의 디지털 신호 프로세서 (digital signal processor; DSP) 들, 범용 마이크로프로세서들, 주문형 집적 회로 (application specific integrated circuit; ASIC) 들, 필드 프로그래밍가능한 로직 어레이 (field programmable logic array; FPGA) 들, 또는 다른 등가의 통합된 또는 개별 로직 회로부와 같은 하나 이상의 프로세서들에 의해 실행될 수도 있다. 따라서, 본원에서 이용된 바와 같은 용어 "프로세서" 는 상기한 구조 또는 본원에서 설명된 기술들의 구현을 위해 적당한 임의의 다른 구조 중 임의의 것을 지칭할 수도 있다.

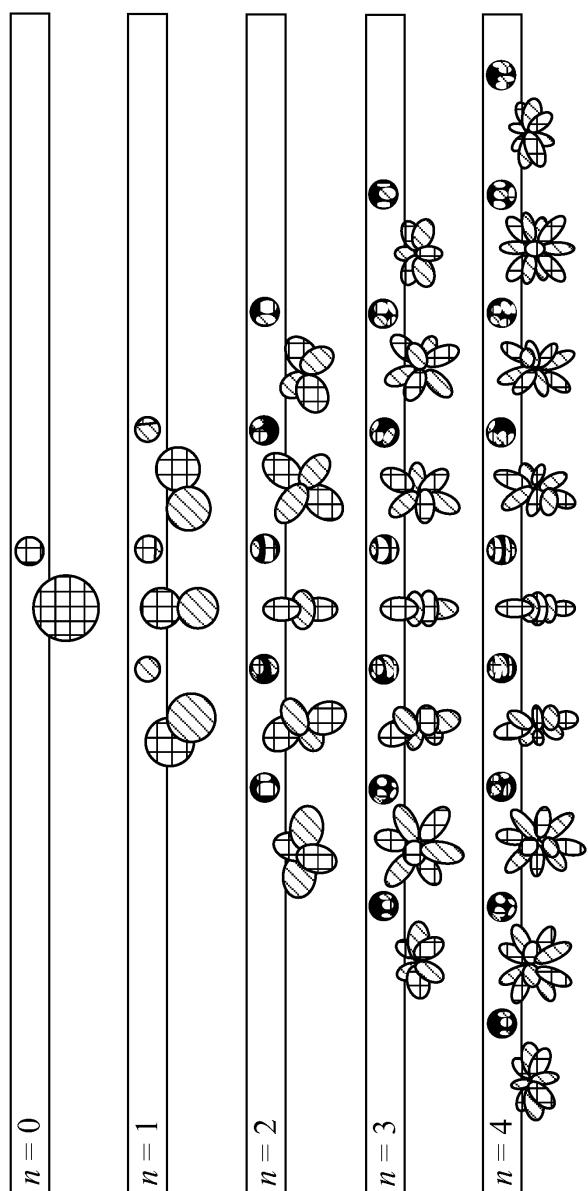
게다가, 일부의 양태들에서는, 본원에서 설명된 기능성이 인코딩 및 디코딩을 위해 구성되거나 조합된 코덱 (codec) 내에 통합되는 전용 하드웨어 및/또는 소프트웨어 모듈들 내에 제공될 수도 있다. 또한, 기법들은 하나 이상의 회로들 또는 로직 엘리먼트들에서 완전히 구현될 수 있다.

[0205] 이 개시물의 기법들은 무선 핸드셋, 집적 회로 (IC) 또는 IC 들의 세트 (예를 들어, 칩 세트) 를 포함하는 광범위한 디바이스들 또는 장치들에서 구현될 수도 있다. 다양한 컴포넌트들, 모듈들, 또는 유닛들은 개시된 기법들을 수행하도록 구성된 디바이스들의 기능적 양태들을 강조하기 위하여 이 개시물에서 설명되어 있지만, 상이한 하드웨어 유닛들에 의한 실현을 반드시 요구하지는 않는다. 오히려, 위에서 설명된 바와 같이, 다양한 유닛들은 코덱 하드웨어 유닛 내에 조합될 수도 있거나, 적당한 소프트웨어 및/또는 펌웨어와 함께, 위에서 설명된 바와 같은 하나 이상의 프로세서들을 포함하는 상호동작하는 하드웨어 유닛들의 집합에 의해 제공될 수도 있다.

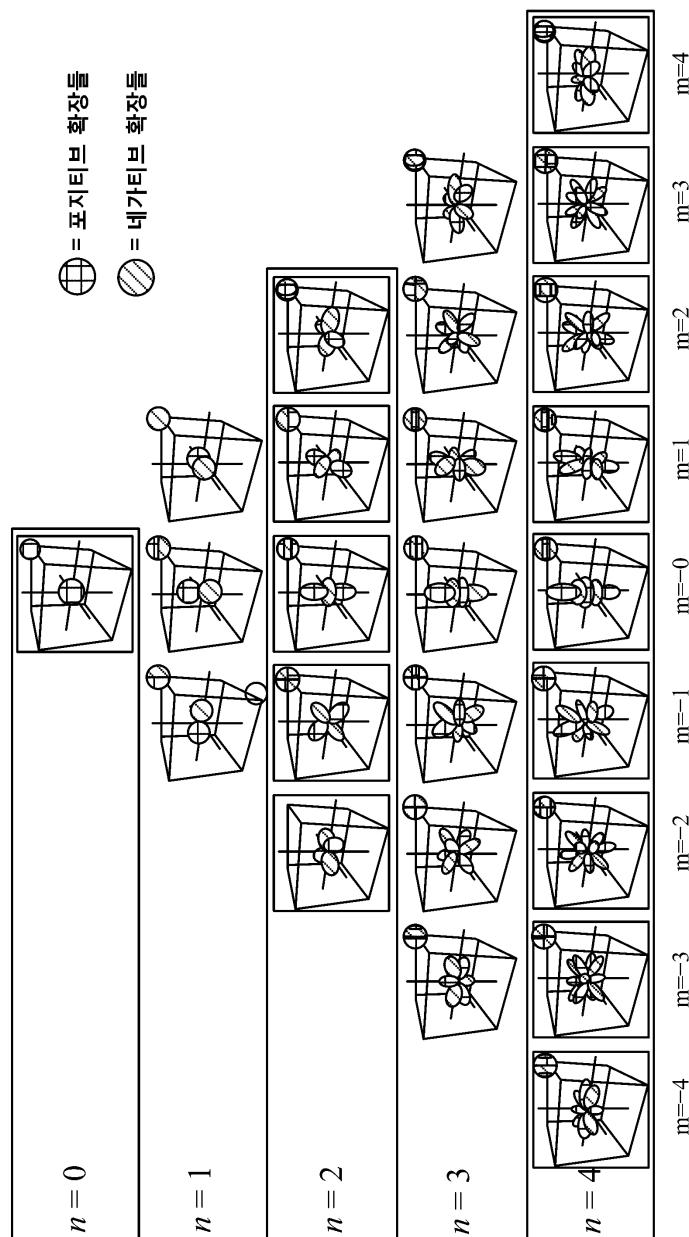
[0206] 개시물의 다양한 양태들이 설명되었다. 이러한 그리고 다른 양태들은 다음의 청구항들의 범위 내에 있다.

도면

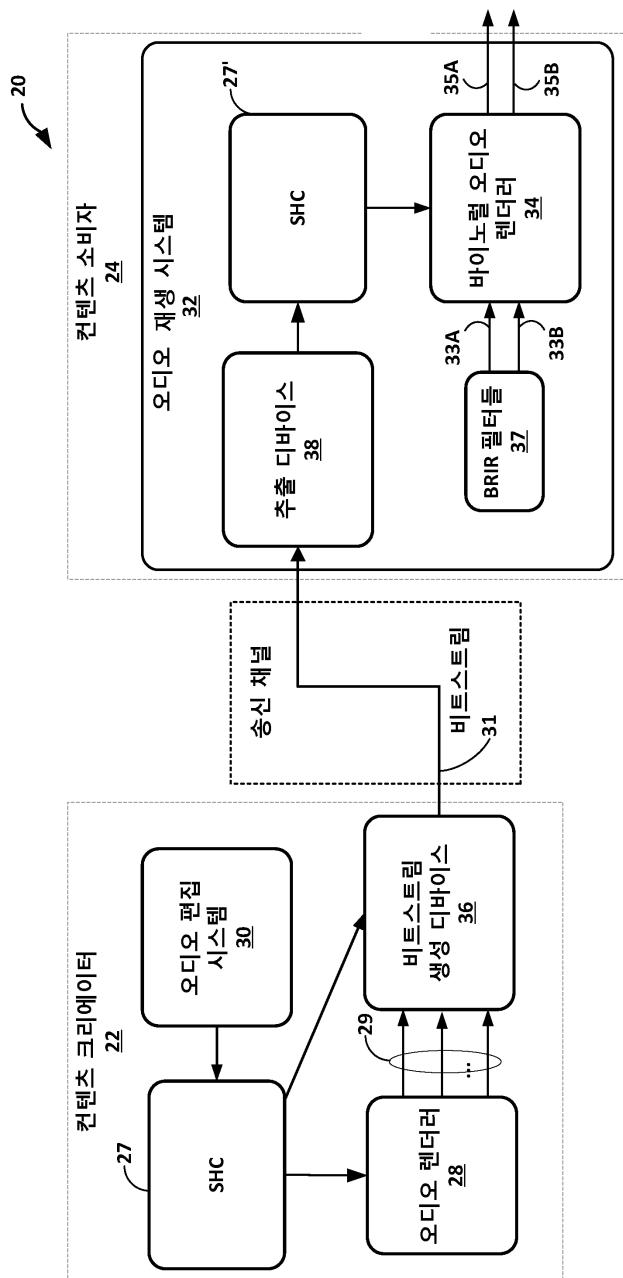
도면1



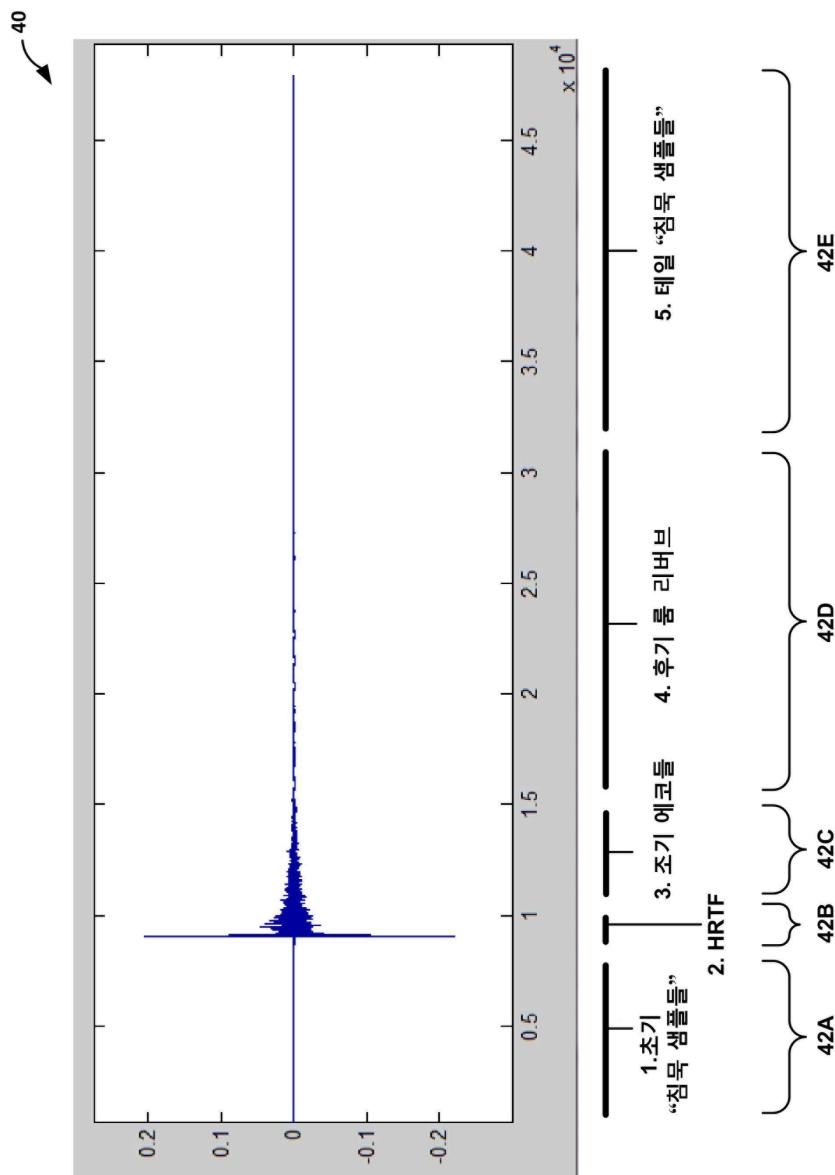
도면2



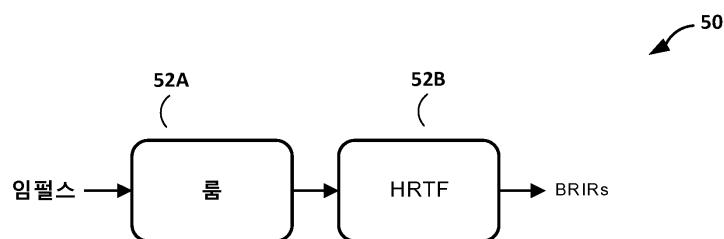
도면3



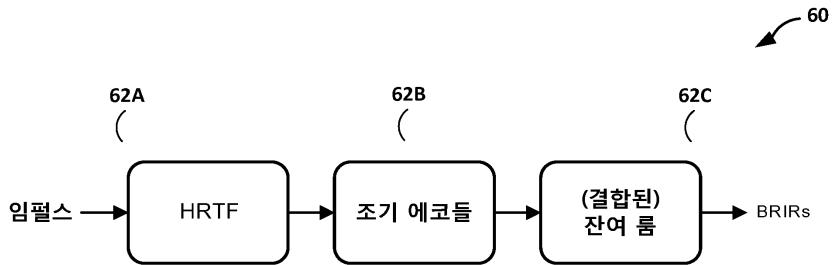
도면4



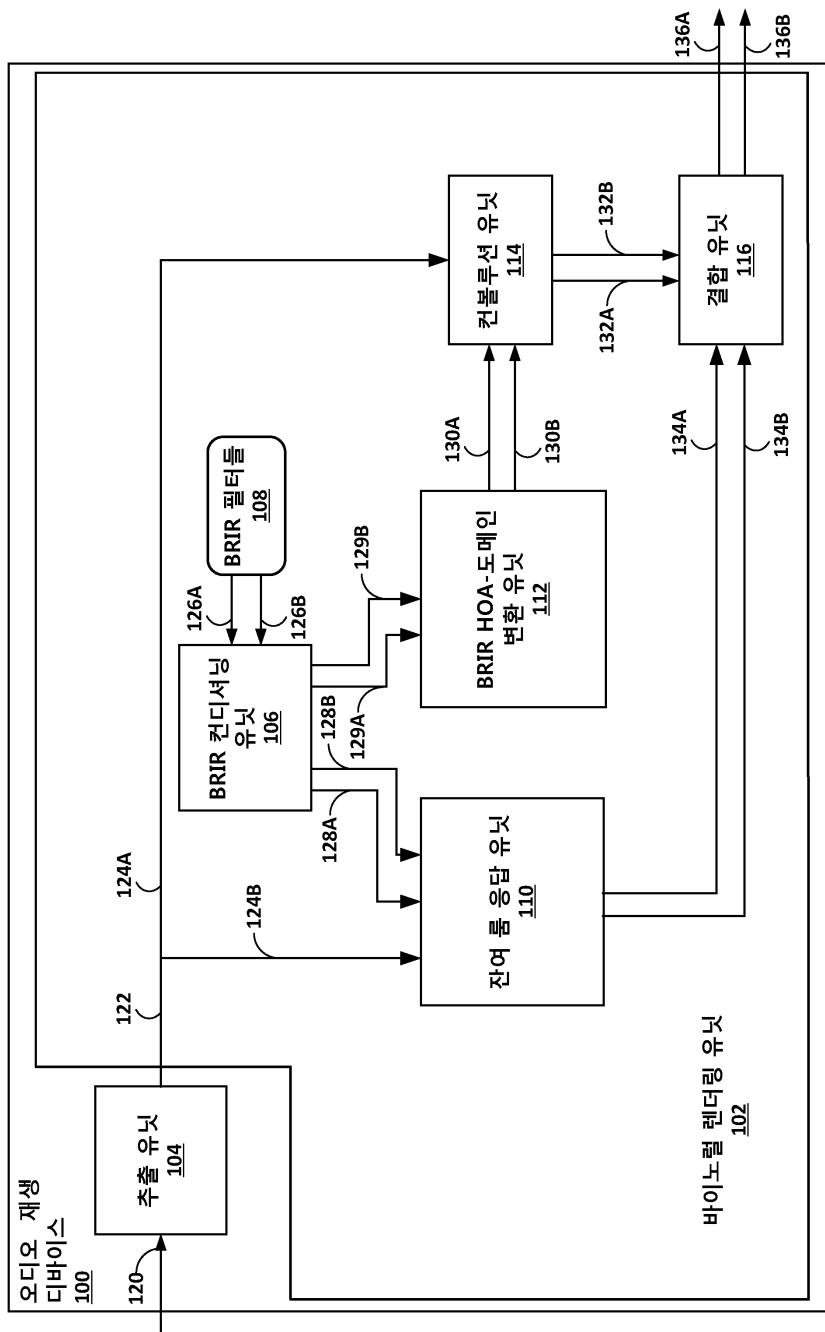
도면5



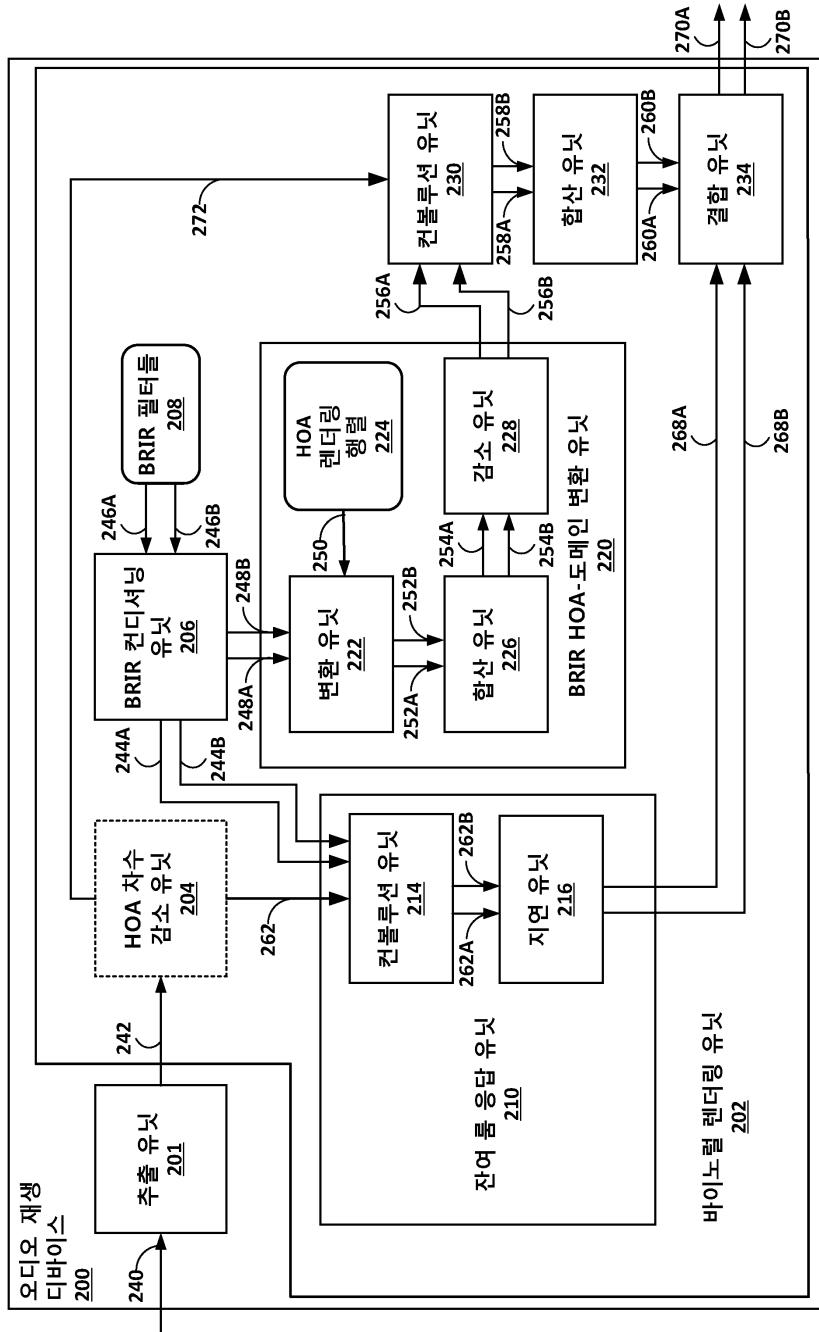
도면6



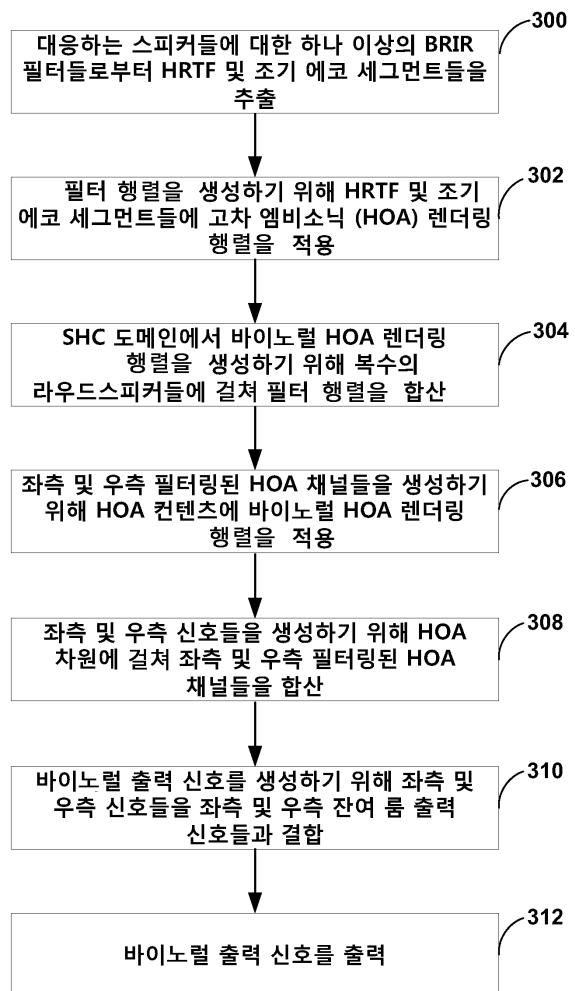
도면7



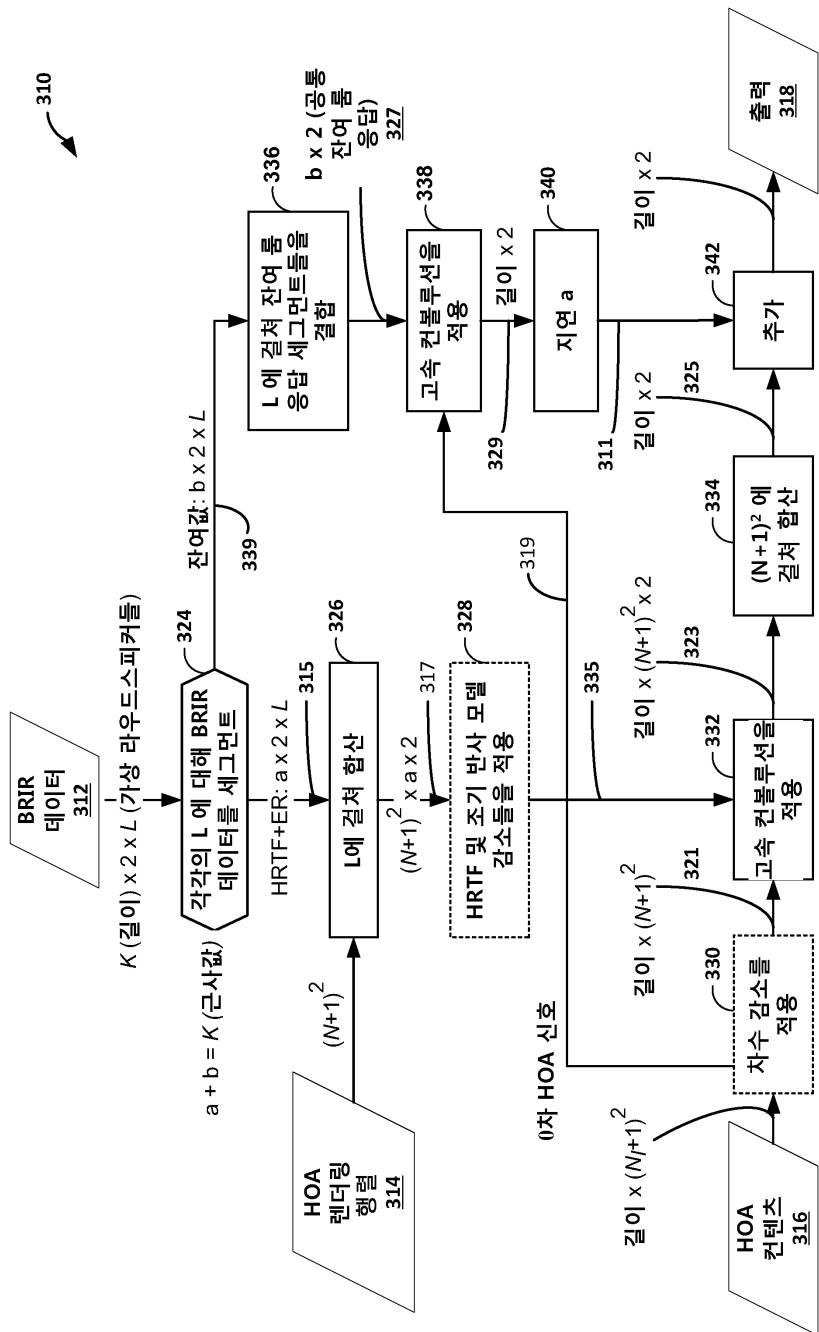
도면8



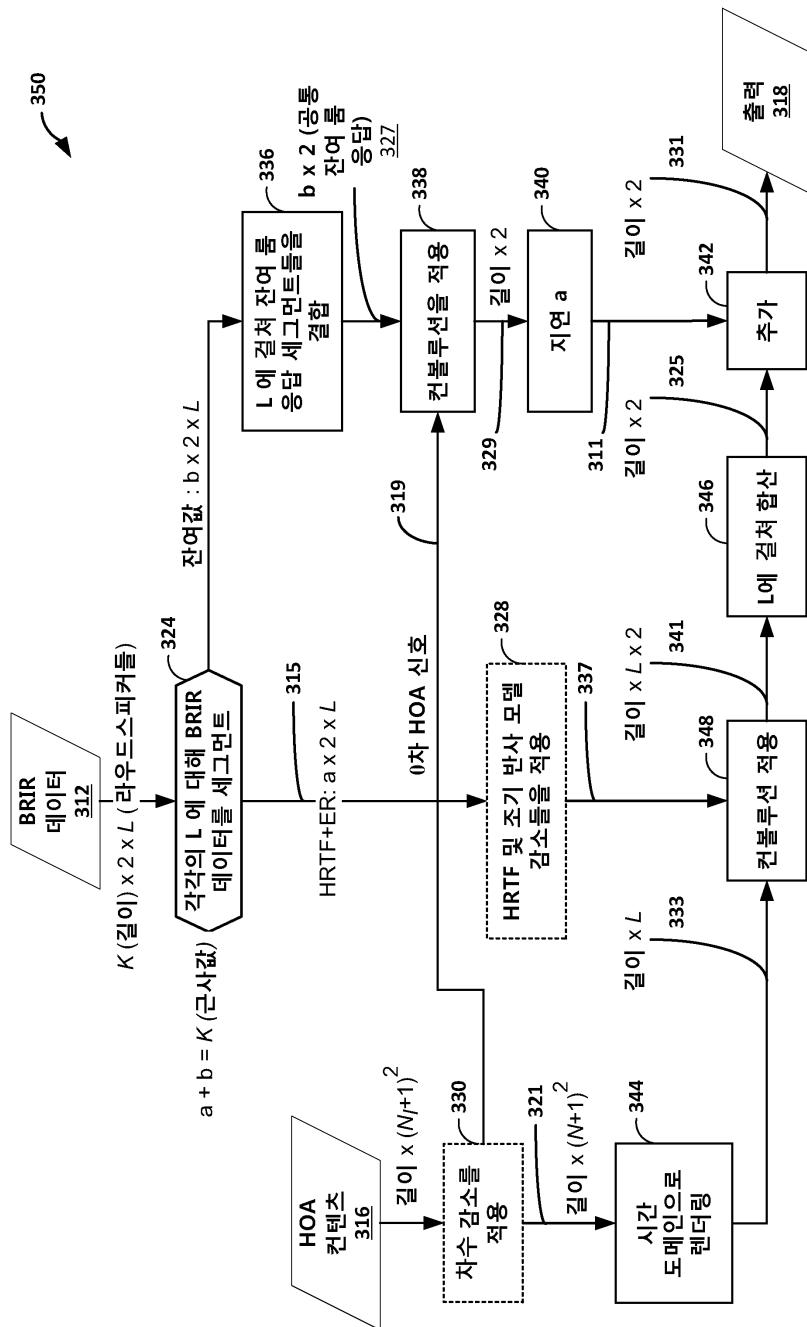
도면9



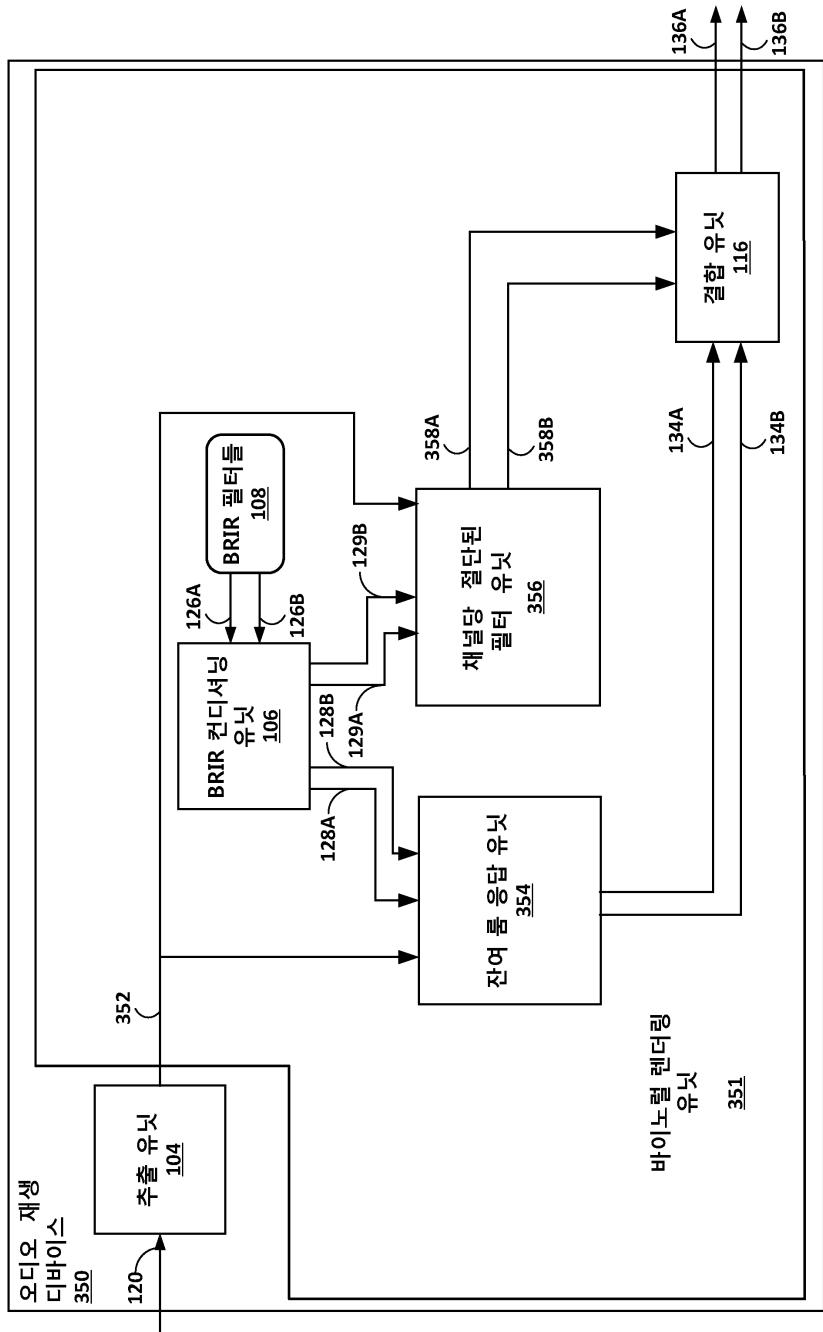
도면 10a



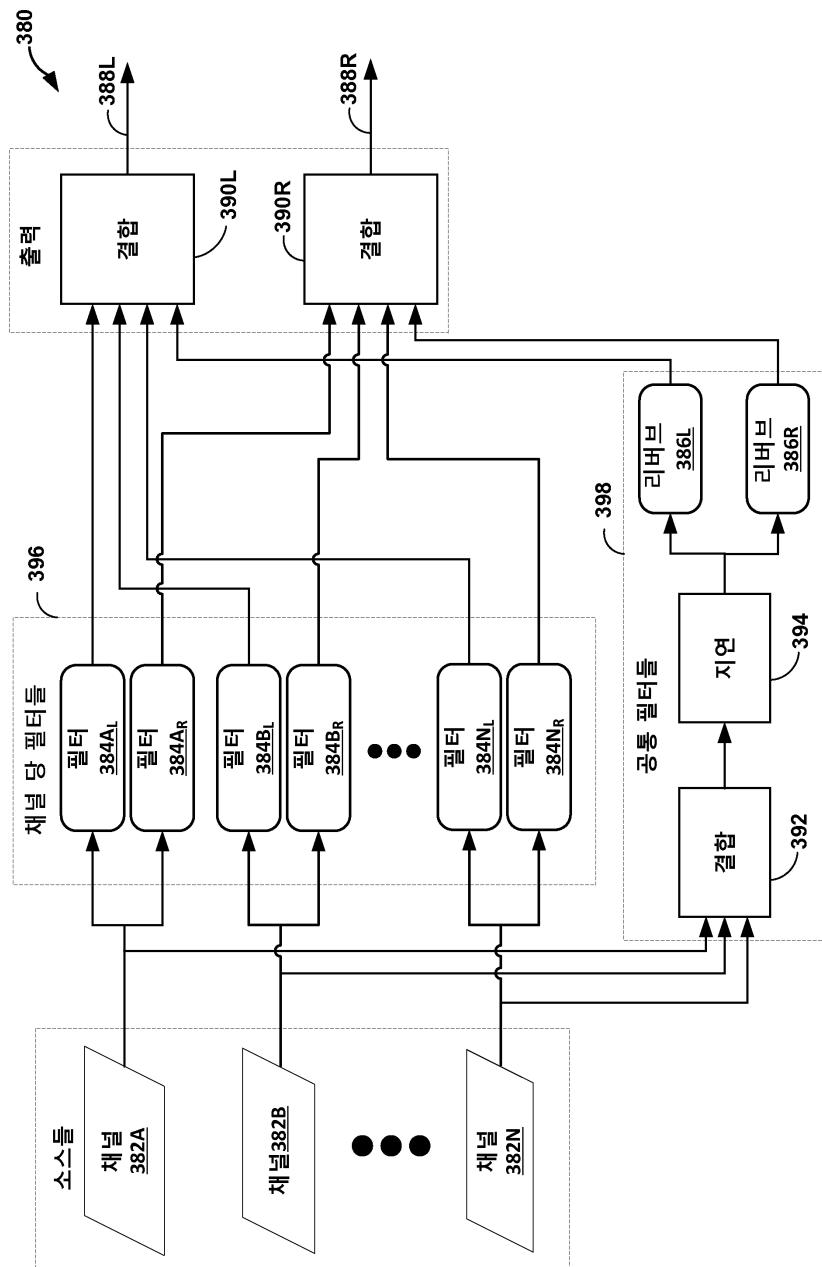
도면10b



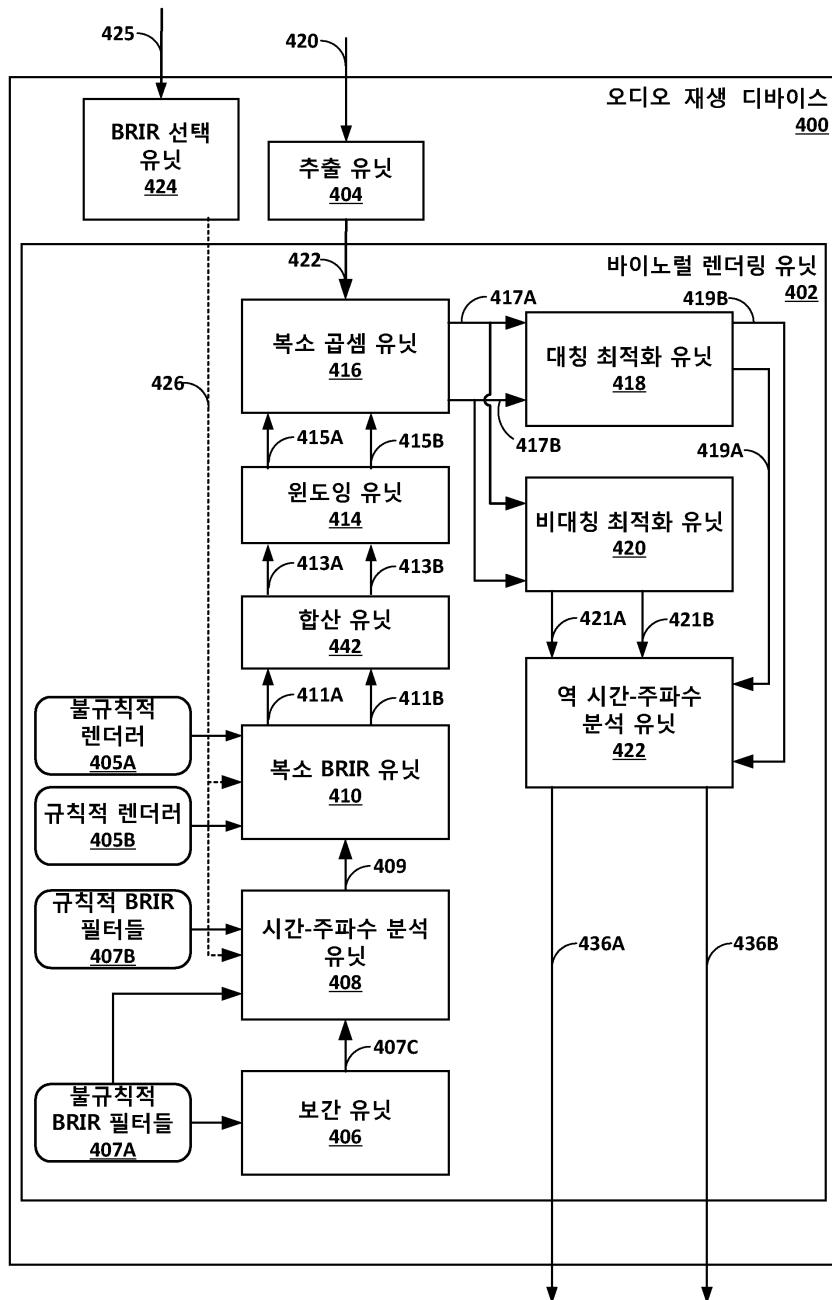
도면11



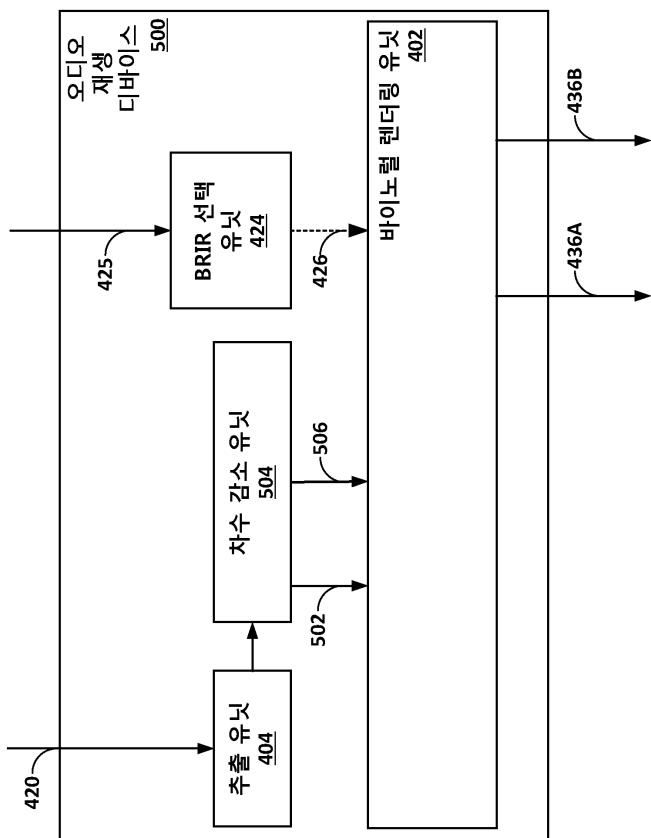
도면12



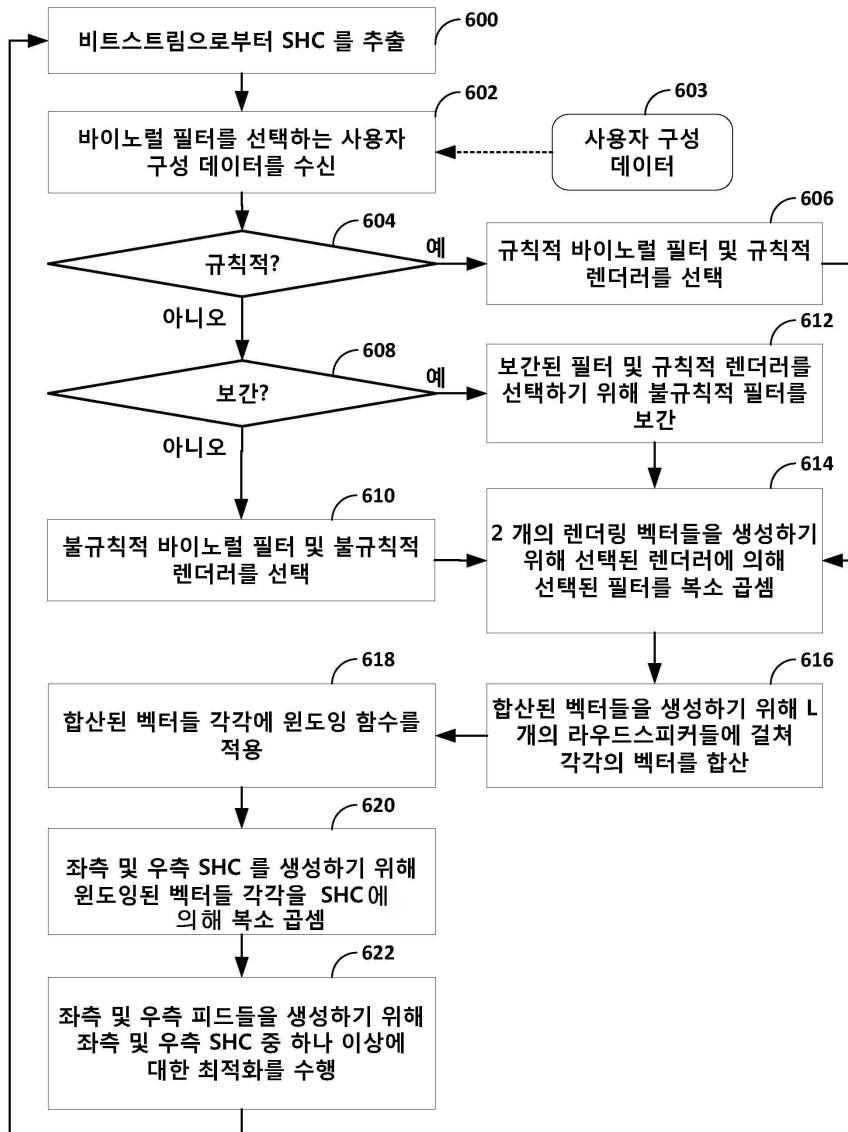
도면13



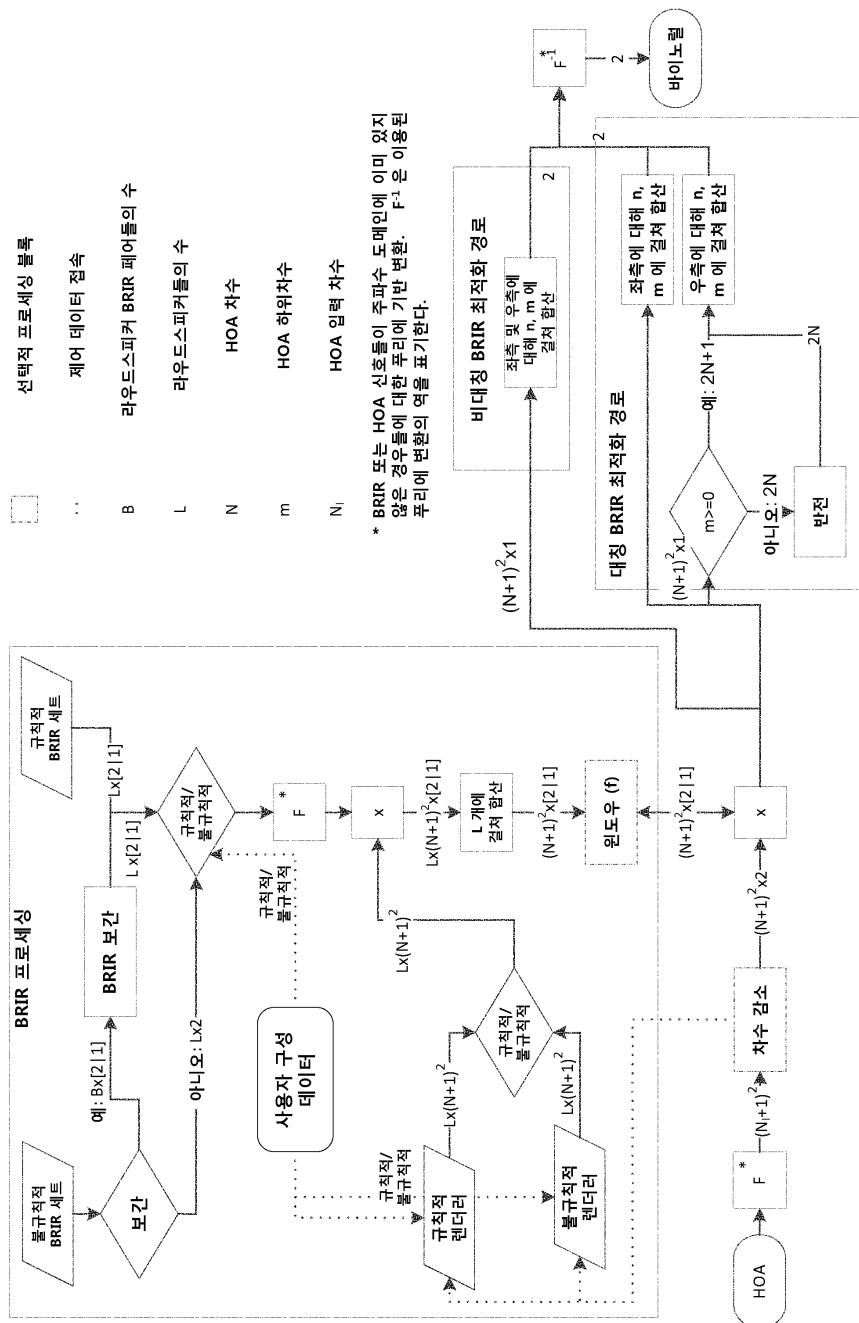
도면14



도면15



도면16a



도면 16b

