(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2018년12월13일
(11) 등록번호 10-1929092
(24) 등록일자 2018년12월07일

(51) 국제특허분류(Int. Cl.)
H04S 7/00 (2006.01) *G10L 19/002* (2013.01)
G10L 19/008 (2014.01) *G10L 19/02* (2006.01)
G10L 19/038 (2013.01) *G10L 19/06* (2006.01)
G10L 25/18 (2013.01) *H04S 5/00* (2015.01)
(52) CPC특허분류
H04S 7/30 (2013.01)
G10L 19/002 (2013.01)
(21) 출원번호 10-2015-7036200
(22) 출원일자(국제) 2014년05월29일
심사청구일자 2018년01월29일
(85) 번역문제출일자 2015년12월22일
(65) 공개번호 10-2016-0015264
(43) 공개일자 2016년02월12일
(86) 국제출원번호 PCT/US2014/040035
(87) 국제공개번호 WO 2014/194099
국제공개일자 2014년12월04일
(30) 우선권주장
61/828,445 2013년05월29일 미국(US)
(뒷면에 계속)
(56) 선행기술조사문헌
KR1020120070521 A
WO2011117399 A1

(73) 특허권자
퀄컴 인코포레이티드
미국 92121-1714 캘리포니아주 샌 디에고 모어하우스 드라이브 5775
(72) 발명자
센 디판잔
미국 92121-1714 캘리포니아주 샌디에고 모어하우스 드라이브 5775
페터스 닐스 귄터
미국 92121-1714 캘리포니아주 샌디에고 모어하우스 드라이브 5775
(74) 대리인
특허법인코리아나

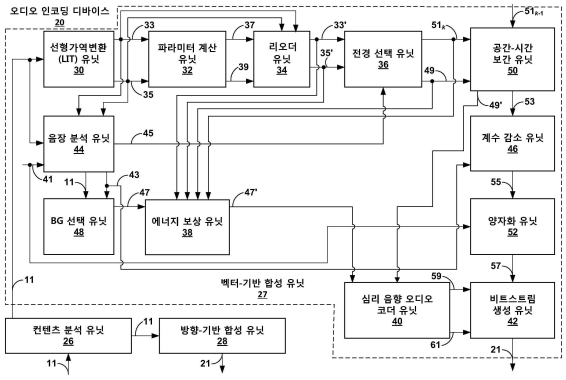
전체 청구항 수 : 총 50 항 심사관 : 권영학

(54) 발명의 명칭 **음장의 분해된 표현들에 대한 보간**

(57) 요 약

일반적으로, 기법들은 음장의 분해된 버전들에 대해 보간을 수행하기 위해 설명된다. 하나 이상의 프로세서들을 포함하는 디바이스가 그 기법들을 수행하도록 구성될 수도 있다. 프로세서들은 적어도 부분적으로, 제 1 복수의 구면 조화 계수들의 제 1 분해 및 제 2 복수의 구면 조화 계수들의 제 2 분해에 대해 보간을 수행함으로써, 시간 세그먼트에 대한 분해된 보간된 구면 조화 계수들을 획득하도록 구성될 수도 있다.

대 표 도



(52) CPC특허분류
G10L 19/008 (2013.01)
G10L 19/0204 (2013.01)
G10L 19/038 (2013.01)
G10L 19/06 (2013.01)
G10L 25/18 (2013.01)
H04S 5/005 (2013.01)
H04S 7/304 (2013.01)
H04S 7/40 (2013.01)
H04S 2400/15 (2013.01)

(30) 우선권주장
61/828,615 2013년05월29일 미국(US)
61/829,174 2013년05월30일 미국(US)
61/829,155 2013년05월30일 미국(US)
61/829,182 2013년05월30일 미국(US)
61/829,846 2013년05월31일 미국(US)
61/829,791 2013년05월31일 미국(US)
61/886,617 2013년10월03일 미국(US)
61/886,605 2013년10월03일 미국(US)
61/899,041 2013년11월01일 미국(US)
61/899,034 2013년11월01일 미국(US)
61/925,074 2014년01월08일 미국(US)
61/925,112 2014년01월08일 미국(US)
61/925,126 2014년01월08일 미국(US)
61/925,158 2014년01월08일 미국(US)
61/933,706 2014년01월30일 미국(US)
61/933,721 2014년01월30일 미국(US)
62/003,515 2014년05월27일 미국(US)
14/289,539 2014년05월28일 미국(US)

명 세 서

청구범위

청구항 1

오디오 디코딩 디바이스에 의해, 압축된 오디오 데이터를 나타내는 비트스트림의 제 1 프레임으로부터, 제 1 복수의 구면 조화 계수들의 제 1 부분의 제 1 분해를 나타내는 데이터를 획득하는 단계;

상기 오디오 디코딩 디바이스에 의해, 상기 비트스트림의 제 2 프레임으로부터, 제 2 복수의 구면 조화 계수들의 제 2 부분의 제 2 분해를 나타내는 데이터를 획득하는 단계로서, 상기 제 1 복수의 구면 조화 계수들의 상기 제 1 분해 및 상기 제 2 복수의 구면 조화 계수들의 상기 제 2 분해는 구면 조화 도메인에서 각각 정의되고 상기 제 1 복수의 구면 조화 계수들 및 상기 제 2 복수의 구면 조화 계수들에 의해 나타낸 음장에 존재하는 우월한 신호 (predominant signal) 의 형상 및 폭을 각각 나타내는, 상기 제 2 분해를 나타내는 데이터를 획득하는 단계;

상기 오디오 디코딩 디바이스에 의해, 분해된 보간된 구면 조화 계수들을 획득하기 위해 상기 제 1 분해 및 상기 제 2 분해에 대하여 보간을 수행하는 단계;

상기 오디오 디코딩 디바이스에 의해, 상기 비트스트림으로부터 상기 분해된 보간된 구면 조화 계수들에 대응하는 우월한 신호를 획득하는 단계;

상기 오디오 디코딩 디바이스에 의해, 상기 분해된 보간된 구면 조화 계수들 및 상기 대응하는 우월한 신호에 기초하여 하나 이상의 스피커 피드들을 렌더링하는 단계; 및

상기 오디오 디코딩 디바이스에 의해, 상기 하나 이상의 스피커 피드들을 하나 이상의 스피커들에 출력하는 단계를 포함하는, 방법.

청구항 2

제 1 항에 있어서,

상기 제 1 분해는 제 1 V 벡터를 포함하는, 방법.

청구항 3

제 1 항에 있어서,

상기 제 2 분해는 제 2 V 벡터를 포함하는, 방법.

청구항 4

제 1 항에 있어서,

상기 보간을 수행하는 단계는 상기 우월한 신호에 대응하는 보간된 V 벡터를 획득하기 위해 제 1 V 벡터 및 제 2 V 벡터에 대하여 상기 보간을 수행하는 단계를 포함하는, 방법.

청구항 5

제 1 항에 있어서,

시간 세그먼트는 상기 제 1 프레임의 서브-프레임을 포함하는, 방법.

청구항 6

제 1 항에 있어서,

시간 세그먼트는 상기 제 1 프레임의 시간 샘플을 포함하는, 방법.

청구항 7

제 1 항에 있어서,

제 1 인공 시간 성분 및 제 2 인공 시간 성분을 수신하는 단계; 및

보간된 분해들의 역들을, 상기 제 1 인공 시간 성분에 적용하여 제 1 시간 성분을 복원하고 상기 제 2 인공 시간 성분에 적용하여 제 2 시간 성분을 복원하는 단계를 더 포함하는, 방법.

청구항 8

제 1 항에 있어서,

시간 세그먼트에 대한 상기 분해된 보간된 구면 조화 계수들을 획득하는 것은 상기 제 1 복수의 구면 조화 계수들의 제 1 공간 성분 및 상기 제 2 복수의 구면 조화 계수들의 제 2 공간 성분을 보간하는 것을 포함하는, 방법.

청구항 9

제 8 항에 있어서,

상기 제 1 공간 성분은 상기 제 1 복수의 구면 조화 계수들에 대한 구면 조화 계수들의 M 개의 시간 세그먼트들을 나타내고 상기 제 2 공간 성분은 상기 제 2 복수의 구면 조화 계수들에 대한 구면 조화 계수들의 M 개의 시간 세그먼트들을 나타내는, 방법.

청구항 10

제 8 항에 있어서,

상기 제 1 공간 성분은 상기 제 1 복수의 구면 조화 계수들에 대한 구면 조화 계수들의 M 개의 시간 세그먼트들을 나타내고 상기 제 2 공간 성분은 상기 제 2 복수의 구면 조화 계수들에 대한 구면 조화 계수들의 M 개의 시간 세그먼트들을 나타내며,

상기 시간 세그먼트에 대한 상기 분해된 보간된 구면 조화 계수들을 획득하는 것은 상기 제 1 공간 성분의 마지막 N 개의 엘리먼트들 및 상기 제 2 공간 성분의 처음 N 개의 엘리먼트들을 보간하는 것을 포함하는, 방법.

청구항 11

제 1 항에 있어서,

상기 제 2 복수의 구면 조화 계수들은 시간 도메인에서 상기 제 1 복수의 구면 조화 계수들에 후속하는, 방법.

청구항 12

제 1 항에 있어서,

상기 제 1 복수의 구면 조화 계수들 및 상기 제 2 복수의 구면 조화 계수들 각각은 상기 음장의 평면파 표현을 나타내는, 방법.

청구항 13

제 1 항에 있어서,

상기 제 1 복수의 구면 조화 계수들 및 상기 제 2 복수의 구면 조화 계수들 각각은 함께 믹싱된 하나 이상의 모노-오디오 오브젝트들을 나타내는, 방법.

청구항 14

제 1 항에 있어서,

상기 제 1 복수의 구면 조화 계수들 및 상기 제 2 복수의 구면 조화 계수들 각각은 3 차원 음장을 나타내는 각각의 제 1 및 제 2 구면 조화 계수들을 포함하는, 방법.

청구항 15

제 1 항에 있어서,

상기 제 1 복수의 구면 조화 계수들 및 상기 제 2 복수의 구면 조화 계수들은 각각 1 보다 더 큰 오더를 갖는 적어도 하나의 구면 기저 함수와 연관되는, 방법.

청구항 16

제 1 항에 있어서,

상기 제 1 복수의 구면 조화 계수들 및 상기 제 2 복수의 구면 조화 계수들은 각각 4 와 동일한 오더를 갖는 적어도 하나의 구면 기저 함수와 연관되는, 방법.

청구항 17

제 1 항에 있어서,

상기 보간은 상기 제 1 분해 및 상기 제 2 분해의 가중된 보간이며, 상기 제 1 분해에 적용된 상기 가중된 보간의 가중치들은 상기 제 1 분해 및 상기 제 2 분해의 벡터들에 의해 나타낸 시간에 반비례하고 상기 제 2 분해에 적용된 상기 가중된 보간의 가중치들은 상기 제 1 분해 및 상기 제 2 분해의 벡터들에 의해 나타낸 시간에 비례하는, 방법.

청구항 18

제 1 항에 있어서,

상기 분해된 보간된 구면 조화 계수들은 상기 제 1 복수의 구면 조화 계수들 및 상기 제 2 복수의 구면 조화 계수들의 공간 성분들 및 시간 성분들 중 적어도 하나를 평활화하는, 방법.

청구항 19

제 1 항에 있어서,

(1) 시간 세그먼트에 대한 상기 분해된 보간된 구면 조화 계수들의 표현; 및

(2) 상기 보간의 타입의 표시

를 포함하는 상기 비트스트림을 획득하는 단계를 더 포함하는, 방법.

청구항 20

제 19 항에 있어서,

상기 표시는 상기 보간의 타입에 맵핑하는 하나 이상의 비트들을 포함하는, 방법.

청구항 21

제 1 항에 있어서,

상기 하나 이상의 스피커들에 의해, 상기 스피커 피드들에 기초하여 상기 분해된 보간된 구면 조화 계수들에 의해 나타낸 음장을 재생하는 단계를 더 포함하는, 방법.

청구항 22

제 1 항에 있어서,

상기 오디오 디코딩 디바이스에 의해, 상기 분해된 보간된 구면 조화 계수들 및 상기 우월한 신호에 기초하여 상기 구면 조화 계수들을 재구성하는 단계를 더 포함하고,

상기 하나 이상의 스피커 피드들을 렌더링하는 단계는, 재구성된 상기 구면 조화 계수들에 기초하여, 상기 하나 이상의 스피커 피드들을 렌더링하는 단계를 포함하는, 방법.

청구항 23

제 1 항에 있어서,

상기 하나 이상의 스피커 피드들을 렌더링하는 단계는, 상기 분해된 보간된 구면 조화 계수들에 기초하여, 하나 이상의 라우드스피커 피드들을 렌더링하는 단계를 포함하고,

상기 하나 이상의 스피커들은 하나 이상의 라우드스피커들을 포함하는, 방법

청구항 24

제 1 항에 있어서,

상기 하나 이상의 스피커 피드들을 렌더링하는 단계는, 상기 분해된 보간된 구면 조화 계수들에 기초하여, 하나 이상의 바이노럴 (binaural) 오디오 헤드폰 피드들을 렌더링하는 단계를 포함하고,

상기 하나 이상의 스피커들은 하나 이상의 헤드폰 스피커들을 포함하는, 방법.

청구항 25

제 1 항에 있어서,

상기 제 1 분해를 획득하기 위해 상기 제 1 분해를 나타내는 상기 데이터에 대하여 역양자화를 수행하는 단계; 및

상기 제 2 분해를 획득하기 위해 상기 제 2 분해를 나타내는 상기 데이터에 대하여 역양자화를 수행하는 단계를 더 포함하는, 방법.

청구항 26

하나 이상의 프로세서들로서,

압축된 오디오 데이터를 나타내는 비트스트림의 제 1 프레임으로부터, 제 1 복수의 구면 조화 계수들의 제 1 부분의 제 1 분해를 나타내는 데이터를 획득하고;

상기 비트스트림의 제 2 프레임으로부터, 제 2 복수의 구면 조화 계수들의 제 2 부분의 제 2 분해를 나타내는 데이터를 획득하는 것으로서, 상기 제 1 복수의 구면 조화 계수들의 상기 제 1 분해 및 상기 제 2 복수의 구면 조화 계수들의 상기 제 2 분해는 구면 조화 도메인에서 각각 정의되고 상기 제 1 복수의 구면 조화 계수들 및 상기 제 2 복수의 구면 조화 계수들에 의해 나타낸 음장에 존재하는 우월한 신호의 형상 및 폭을 각각 나타내는, 상기 제 2 분해를 나타내는 데이터를 획득하고;

분해된 보간된 구면 조화 계수들을 획득하기 위해 상기 제 1 분해 및 상기 제 2 분해에 대하여 보간을 수행하고;

상기 비트스트림으로부터, 분해된 보간된 구면 조화 계수들에 대응하는 우월한 신호를 획득하고;

상기 분해된 보간된 구면 조화 계수들 및 상기 대응하는 우월한 신호에 기초하여 하나 이상의 스피커 피드들을 렌더링하고; 그리고

상기 하나 이상의 스피커 피드들을 하나 이상의 스피커들에 출력하도록

구성된, 상기 하나 이상의 프로세서들; 및

상기 하나 이상의 프로세서들에 커플링되고, 상기 스피커 피드들을 저장하도록 구성된 메모리를 포함하는, 디바이스.

청구항 27

제 26 항에 있어서,

상기 제 1 분해는 제 1 V 벡터를 포함하는, 디바이스.

청구항 28

제 26 항에 있어서,

상기 제 2 분해는 제 2 V 벡터를 포함하는, 디바이스.

청구항 29

제 26 항에 있어서,

상기 하나 이상의 프로세서들은 상기 우월한 신호에 대응하는 보간된 V 벡터를 획득하기 위해 제 1 V 행렬 및 제 2 V 행렬에 대하여 상기 보간을 수행하도록 구성되는, 디바이스.

청구항 30

제 26 항에 있어서,

시간 세그먼트는 상기 제 1 프레임의 시간 샘플을 포함하는, 디바이스.

청구항 31

제 26 항에 있어서,

상기 하나 이상의 프로세서들은 또한,

제 1 인공 시간 성분 및 제 2 인공 시간 성분을 수신하고; 그리고

보간된 분해들의 역들을, 상기 제 1 인공 시간 성분에 적용하여 제 1 시간 성분을 복원하고 상기 제 2 인공 시간 성분에 적용하여 제 2 시간 성분을 복원하도록 구성되는, 디바이스.

청구항 32

제 26 항에 있어서,

상기 하나 이상의 프로세서들은 상기 제 1 복수의 구면 조화 계수들의 제 1 공간 성분 및 상기 제 2 복수의 구면 조화 계수들의 제 2 공간 성분을 보간하도록 구성되는, 디바이스.

청구항 33

제 32 항에 있어서,

상기 제 1 공간 성분은 상기 제 1 복수의 구면 조화 계수들의 좌-특이 벡터들을 나타내는 제 1 U 행렬을 포함하는, 디바이스.

청구항 34

제 32 항에 있어서,

상기 제 2 공간 성분은 상기 제 2 복수의 구면 조화 계수들의 좌-특이 벡터들을 나타내는 제 2 U 행렬을 포함하는, 디바이스.

청구항 35

제 32 항에 있어서,

상기 제 1 공간 성분은 상기 제 1 복수의 구면 조화 계수들에 대한 구면 조화 계수들의 M 개의 시간 세그먼트들을 나타내고 상기 제 2 공간 성분은 상기 제 2 복수의 구면 조화 계수들에 대한 구면 조화 계수들의 M 개의 시간 세그먼트들을 나타내는, 디바이스.

청구항 36

제 32 항에 있어서,

상기 제 1 공간 성분은 상기 제 1 복수의 구면 조화 계수들에 대한 구면 조화 계수들의 M 개의 시간 세그먼트들을 나타내고 상기 제 2 공간 성분은 상기 제 2 복수의 구면 조화 계수들에 대한 구면 조화 계수들의 M 개의 시간 세그먼트들을 나타내며,

상기 하나 이상의 프로세서들은 상기 제 1 공간 성분의 마지막 N 개의 엘리먼트들 및 상기 제 2 공간 성분의 처음 N 개의 엘리먼트들을 보간하도록 구성되는, 디바이스.

청구항 37

제 26 항에 있어서,

상기 제 2 복수의 구면 조화 계수들은 시간 도메인에서 상기 제 1 복수의 구면 조화 계수들에 후속하는, 디바이스.

청구항 38

제 26 항에 있어서,

상기 제 1 복수의 구면 조화 계수들 및 상기 제 2 복수의 구면 조화 계수들 각각은 상기 음장의 평면파 표현을 나타내는, 디바이스.

청구항 39

제 26 항에 있어서,

상기 제 1 복수의 구면 조화 계수들 및 상기 제 2 복수의 구면 조화 계수들 각각은 함께 믹싱된 하나 이상의 모노-오디오 오브젝트들을 나타내는, 디바이스.

청구항 40

제 26 항에 있어서,

상기 제 1 복수의 구면 조화 계수들 및 상기 제 2 복수의 구면 조화 계수들은 각각 1 보다 더 큰 오더를 갖는 적어도 하나의 구면 기저 함수와 연관되는, 디바이스.

청구항 41

제 26 항에 있어서,

상기 제 1 복수의 구면 조화 계수들 및 상기 제 2 복수의 구면 조화 계수들은 각각 4 와 동일한 오더를 갖는 적어도 하나의 구면 기저 함수와 연관되는, 디바이스.

청구항 42

제 26 항에 있어서,

상기 하나 이상의 프로세서들은 또한, 시간 세그먼트에 대한 상기 분해된 보간된 구면 조화 계수들의 표현, 및 상기 보간의 타입의 표시를 포함하는 상기 비트스트림을 획득하도록 구성되는, 디바이스.

청구항 43

제 42 항에 있어서,

상기 표시는 상기 보간의 타입에 맵핑하는 하나 이상의 비트들을 포함하는, 디바이스.

청구항 44

제 26 항에 있어서,

상기 하나 이상의 스피커들을 더 포함하고,

상기 하나 이상의 스피커들은, 상기 스피커 피드들에 기초하여, 상기 분해된 보간된 구면 조화 계수들에 의해 나타낸 음장을 재생하도록 구성된, 디바이스.

청구항 45

제 26 항에 있어서,

상기 하나 이상의 프로세서들은 또한, 상기 분해된 보간된 구면 조화 계수들 및 상기 우월한 신호에 기초하여, 상기 구면 조화 계수들을 재구성하도록 구성되고,

상기 하나 이상의 프로세서들은 재구성된 상기 구면 조화 계수들에 기초하여, 상기 하나 이상의 스피커 피드들을 렌더링하도록 구성되는, 디바이스.

청구항 46

제 26 항에 있어서,

상기 하나 이상의 프로세서들은 상기 분해된 보간된 구면 조화 계수들에 기초하여, 하나 이상의 라우드스피커 피드들을 렌더링하도록 구성되고,

상기 하나 이상의 스피커들은 하나 이상의 라우드스피커들을 포함하는, 디바이스.

청구항 47

제 26 항에 있어서,

상기 하나 이상의 프로세서들은 상기 분해된 보간된 구면 조화 계수들에 기초하여, 하나 이상의 바이노럴 오디오 헤드폰 피드들을 렌더링하도록 구성되고,

상기 하나 이상의 스피커들은 하나 이상의 헤드폰 스피커들을 포함하는, 디바이스.

청구항 48

제 26 항에 있어서,

상기 하나 이상의 프로세서들은 또한,

상기 제 1 분해를 획득하기 위해 상기 제 1 분해를 나타내는 상기 데이터에 대하여 역양자화를 수행하고; 그리고

상기 제 2 분해를 획득하기 위해 상기 제 2 분해를 나타내는 상기 데이터에 대하여 역양자화를 수행하도록 구성되는, 디바이스.

청구항 49

오디오 디코딩 디바이스에 의해, 압축된 오디오 데이터를 나타내는 비트스트림의 제 1 프레임으로부터, 제 1 복수의 구면 조화 계수들의 제 1 부분의 제 1 분해를 획득하는 수단;

상기 오디오 디코딩 디바이스에 의해, 상기 비트스트림의 제 2 프레임으로부터, 제 2 복수의 구면 조화 계수들의 제 2 부분의 제 2 분해를 획득하는 수단으로서, 상기 제 1 복수의 구면 조화 계수들의 상기 제 1 분해 및 상기 제 2 복수의 구면 조화 계수들의 상기 제 2 분해는 구면 조화 도메인에서 각각 정의되고 상기 제 1 복수의 구면 조화 계수들 및 상기 제 2 복수의 구면 조화 계수들에 의해 나타낸 음장에 존재하는 우월한 신호의 형상 및 폭을 각각 나타내는, 상기 제 2 분해를 획득하는 수단;

분해된 보간된 구면 조화 계수들을 획득하기 위해 상기 제 1 분해 및 상기 제 2 분해에 대하여 보간을 수행하는 수단;

상기 비트스트림으로부터, 분해된 보간된 구면 조화 계수들에 대응하는 우월한 신호를 획득하는 수단;

상기 분해된 보간된 구면 조화 계수들 및 상기 대응하는 우월한 신호에 기초하여 하나 이상의 스피커 피드들을 렌더링하는 수단; 및

상기 하나 이상의 스피커 피드들을 하나 이상의 스피커들에 출력하는 수단을 포함하는, 디바이스.

청구항 50

명령들을 저장하고 있는 비-일시적 컴퓨터 판독가능 저장 매체로서,

상기 명령들은, 실행될 때, 하나 이상의 프로세서들로 하여금 :

압축된 오디오 데이터를 나타내는 비트스트림의 제 1 프레임으로부터, 제 1 복수의 구면 조화 계수들의 제 1 부분의 제 1 분해를 획득하게 하고;

상기 비트스트림의 제 2 프레임으로부터, 제 2 복수의 구면 조화 계수들의 제 2 부분의 제 2 분해를 획득하게

하는 것으로서, 상기 제 1 복수의 구면 조화 계수들의 상기 제 1 분해 및 상기 제 2 복수의 구면 조화 계수들의 상기 제 2 분해는 구면 조화 도메인에서 각각 정의되고 상기 제 1 복수의 구면 조화 계수들 및 상기 제 2 복수의 구면 조화 계수들에 의해 나타낸 음장에 존재하는 우월한 신호의 형상 및 폭을 각각 나타내는, 상기 제 2 분해를 획득하게 하고;

분해된 보간된 구면 조화 계수들을 획득하기 위해 상기 제 1 분해 및 상기 제 2 분해에 대하여 보간을 수행하게 하고;

상기 비트스트림으로부터, 분해된 보간된 구면 조화 계수들에 대응하는 우월한 신호를 획득하게 하고;

상기 분해된 보간된 구면 조화 계수들 및 상기 대응하는 우월한 신호에 기초하여 하나 이상의 스피커 피드들을 렌더링하게 하고; 그리고

상기 하나 이상의 스피커 피드들을 하나 이상의 스피커들에 출력하게 하는, 비-일시적 컴퓨터 판독가능 저장 매체.

청구항 51

삭제

청구항 52

삭제

청구항 53

삭제

청구항 54

삭제

청구항 55

삭제

청구항 56

삭제

청구항 57

삭제

청구항 58

삭제

발명의 설명

기 술 분 야

[0001] 본 출원은 2013 년 5 월 29 일에 출원된 미국 가출원 제 61/828,445 호, 2013 년 5 월 31 일에 출원된 미국 가출원 제 61/829,791 호, 2013 년 11 월 1 일에 출원된 미국 가출원 제 61/899,034 호, 2013 년 11 월 1 일에 출원된 미국 가출원 제 61/899,041 호, 2013 년 5 월 30 일에 출원된 미국 가출원 제 61/829,182 호, 2013 년 5 월 30 일에 출원된 미국 가출원 제 61/829,174 호, 2013 년 5 월 30 일에 출원된 미국 가출원 제 61/829,155 호, 2014 년 1 월 30 일에 출원된 미국 가출원 제 61/933,706 호, 2013 년 5 월 31 일에 출원된 미국 가출원 제 61/829,846 호, 2013 년 10 월 3 일에 출원된 미국 가출원 제 61/886,605 호, 2013 년 10 월 3 일에 출원된 미국 가출원 제 61/886,617 호, 2014 년 1 월 8 일에 출원된 미국 가출원 제 61/925,158 호, 2014 년 1 월 30 일에 출원된 미국 가출원 제 61/933,721 호, 2014 년 1 월 8 일에 출원된 미국 가출원 제 61/925,074 호, 2014 년 1 월 8 일에 출원된 미국 가출원 제 61/925,112 호, 2014 년 1 월 8 일에 출원된 미국 가출원 제

61/925,126 호, 2014 년 5 월 27 일에 출원된 미국 가출원 제 62/003,515 호, 및 2013 년 5 월 29 일에 출원된 미국 가출원 제 61/828,615 호의 이익을 청구하며, 이들 각각의 전체 내용은 본 명세서에서 참조로서 통합된다.

[0002] 본 개시물은 오디오 데이터에 관한 것이며, 더 구체적으로 오디오 데이터의 압축에 관한 것이다.

배 경 기 술

[0003] (종종 복수의 구면 조화 계수들 (SHC) 또는 다른 계층 엘리먼트들로 표현되는) 고차 앰비소닉스 (HOA) 신호는 음장의 3 차원 표현이다. 이러한 HOA 또는 SHC 표현은 상기 SHC 신호로부터 렌더링된 다중-채널 오디오 신호를 재생하는데 사용된 로컬 스피커 지오메트리에 독립적인 방식으로 상기 음장을 표현할 수도 있다. 이러한 SHC 신호는 또한, 이러한 SHC 신호가 5.1 오디오 채널 포맷 또는 7.1 오디오 채널 포맷과 같은 잘 알려지고 고도로 채택된 다중-채널 포맷들로 렌더링될 수도 있기 때문에, 역호환성을 가능하게 할 수도 있다. 그러므로, SHC 표현은 역호환성을 또한 수용하는 음장의 더 양호한 표현을 가능하게 할 수도 있다.

발명의 내용

해결하려는 과제

과제의 해결 수단

[0004] 일반적으로, 기술들은 고차 앰비소닉 오디오 데이터의 압축 및 압축해제를 위해 설명된다.

[0005] 일 양태에서, 일 방법은 음장의 구별되는 성분들을 설명하는 하나 이상의 제 1 벡터들 및 음장의 배경 성분들을 설명하는 하나 이상의 제 2 벡터들을 획득하는 단계를 포함하며, 하나 이상의 제 1 벡터들과 하나 이상의 제 2 벡터들 양자는 적어도 복수의 구면 조화 계수들에 대해 변환을 수행함으로써 생성된다.

[0006] 다른 양태에서, 일 디바이스는 음장의 구별되는 성분들을 설명하는 하나 이상의 제 1 벡터들 및 음장의 배경 성분들을 설명하는 하나 이상의 제 2 벡터들을 결정하도록 구성된 하나 이상의 프로세서들을 포함하며, 하나 이상의 제 1 벡터들과 하나 이상의 제 2 벡터들 양자는 적어도 복수의 구면 조화 계수들에 대해 변환을 수행함으로써 생성된다.

[0007] 또 다른 양태에서, 일 디바이스는 음장의 구별되는 성분들을 설명하는 하나 이상의 제 1 벡터들 및 음장의 배경 성분들을 설명하는 하나 이상의 제 2 벡터들을 획득하는 수단으로서, 하나 이상의 제 1 벡터들과 하나 이상의 제 2 벡터들 양자는 적어도 복수의 구면 조화 계수들에 대해 변환을 수행함으로써 생성되는, 상기 획득하는 수단 및 하나 이상의 제 1 벡터들을 저장하는 수단을 포함한다.

[0008] 또 다른 양태에서, 비-일시적 컴퓨터 판독가능 저장 매체는 저장된 명령들을 가지고, 그 명령들은 실행될 경우, 하나 이상의 프로세서들로 하여금, 음장의 구별되는 성분들을 설명하는 하나 이상의 제 1 벡터들 및 음장의 배경 성분들을 설명하는 하나 이상의 제 2 벡터들을 획득하게 하며, 하나 이상의 제 1 벡터들과 하나 이상의 제 2 벡터들 양자는 적어도 복수의 구면 조화 계수들에 대해 변환을 수행함으로써 생성된다.

[0009] 또 다른 양태에서, 일 방법은 음장을 나타내는 구면 조화 계수들의 압축 버전이 합성 오디오 오브젝트로부터 생성되는지 여부의 표시에 기초하여 복수의 압축해제 방식들 중 하나를 선택하는 단계, 및 복수의 압축해제 방식들 중 선택된 하나를 사용하여 구면 조화 계수들의 압축 버전을 압축해제하는 단계를 포함한다.

[0010] 또 다른 양태에서, 일 디바이스는 음장을 나타내는 구면 조화 계수들의 압축 버전이 합성 오디오 오브젝트로부터 생성되는지 여부의 표시에 기초하여 복수의 압축해제 방식들 중 하나를 선택하도록 구성된 하나 이상의 프로세서들을 포함하고, 복수의 압축해제 방식들 중 선택된 하나를 사용하여 구면 조화 계수들의 압축 버전을 압축해제한다.

[0011] 또 다른 양태에서, 일 디바이스는 음장을 나타내는 구면 조화 계수들의 압축 버전이 합성 오디오 오브젝트로부터 생성되는지 여부의 표시에 기초하여 복수의 압축해제 방식들 중 하나를 선택하는 수단, 및 복수의 압축해제 방식들 중 선택된 하나를 사용하여 구면 조화 계수들의 압축 버전을 압축해제하는 수단을 포함한다.

[0012] 또 다른 양태에서, 비-일시적 컴퓨터 판독가능 저장 매체는 저장된 명령들을 가지고, 그 명령들은 실행될 경우, 통합된 디코딩 디바이스의 하나 이상의 프로세서들로 하여금, 음장을 나타내는 구면 조화 계수들의 압축 버전이 합성 오디오 오브젝트로부터 생성되는지 여부의 표시에 기초하여 복수의 압축해제 방식들 중 하나를 선택하게

하고, 복수의 압축해제 방식들 중 선택된 하나를 사용하여 구면 조화 계수들의 압축 버전을 압축해제하게 한다.

[0013] 또 다른 양태에서, 일 방법은 음장을 나타내는 구면 조화 계수들이 합성 오디오 오브젝트로부터 생성되는지 여부의 표시를 획득하는 단계를 포함한다.

[0014] 또 다른 양태에서, 일 디바이스는 음장을 나타내는 구면 조화 계수들이 합성 오디오 오브젝트로부터 생성되는지 여부의 표시를 획득하도록 구성된 하나 이상의 프로세서들을 포함한다.

[0015] 또 다른 양태에서, 일 디바이스는 음장을 나타내는 구면 조화 계수들을 저장하는 수단 및 구면 조화 계수들이 합성 오디오 오브젝트로부터 생성되는지 여부의 표시를 획득하는 수단을 포함한다.

[0016] 또 다른 양태에서, 비-일시적 컴퓨터 판독가능 저장 매체는 저장된 명령들을 가지고, 그 명령들은 실행될 경우, 하나 이상의 프로세서들로 하여금, 음장을 나타내는 구면 조화 계수들이 합성 오디오 오브젝트로부터 생성되는지 여부의 표시를 획득하게 한다.

[0017] 또 다른 양태에서, 일 방법은 음장의 하나 이상의 성분들을 나타내는 하나 이상의 제 1 벡터들을 양자화하는 단계, 및 음장의 동일한 하나 이상의 성분들을 또한 나타내는 하나 이상의 제 2 벡터들에서 하나 이상의 제 1 벡터들의 양자화로 인해 도입된 에러에 대하여 보상하는 단계를 포함한다.

[0018] 또 다른 양태에서, 일 디바이스는 음장의 하나 이상의 성분들을 나타내는 하나 이상의 제 1 벡터들을 양자화하고, 음장의 동일한 하나 이상의 성분들을 또한 나타내는 하나 이상의 제 2 벡터들에서 하나 이상의 제 1 벡터들의 양자화로 인해 도입된 에러에 대하여 보상하도록 구성된 하나 이상의 프로세서들을 포함한다.

[0019] 또 다른 양태에서, 일 디바이스는 음장의 하나 이상의 성분들을 나타내는 하나 이상의 제 1 벡터들을 양자화하는 수단, 및 음장의 동일한 하나 이상의 성분들을 또한 나타내는 하나 이상의 제 2 벡터들에서 하나 이상의 제 1 벡터들의 양자화로 인해 도입된 에러에 대하여 보상하는 수단을 포함한다.

[0020] 또 다른 양태에서, 비-일시적 컴퓨터 판독가능 저장 매체는 저장된 명령들을 가지고, 그 명령들은 실행될 경우, 하나 이상의 프로세서들로 하여금, 음장의 하나 이상의 성분들을 나타내는 하나 이상의 제 1 벡터들을 양자화하게 하고, 음장의 동일한 하나 이상의 성분들을 또한 나타내는 하나 이상의 제 2 벡터들에서 하나 이상의 제 1 벡터들의 양자화로 인해 도입된 에러에 대하여 보상하게 한다.

[0021] 또 다른 양태에서, 일 방법은 타겟 비트레이트에 기초하여, 복수의 구면 조화 계수들 또는 그의 분해들에 대해 오더 감소를 수행하여 감소된 구면 조화 계수들 또는 감소된 분해들을 생성하는 단계를 포함하며, 복수의 구면 조화 계수들은 음장을 표현한다.

[0022] 또 다른 양태에서, 일 디바이스는 타겟 비트레이트에 기초하여, 복수의 구면 조화 계수들 또는 그의 분해들에 대해 오더 감소를 수행하여 감소된 구면 조화 계수들 또는 감소된 분해들을 생성하도록 구성된 하나 이상의 프로세서들을 포함하며, 복수의 구면 조화 계수들은 음장을 표현한다.

[0023] 또 다른 양태에서, 일 디바이스는 복수의 구면 조화 계수들 또는 그의 분해들을 저장하는 수단, 및 타겟 비트레이트에 기초하여, 복수의 구면 조화 계수들 또는 그의 분해들에 대해 오더 감소를 수행하여 감소된 구면 조화 계수들 또는 감소된 분해들을 생성하는 수단을 포함하며, 복수의 구면 조화 계수들은 음장을 표현한다.

[0024] 또 다른 양태에서, 비-일시적 컴퓨터 판독가능 저장 매체는 저장된 명령들을 가지고, 그 명령들은 실행될 경우, 하나 이상의 프로세서들로 하여금, 타겟 비트레이트에 기초하여, 복수의 구면 조화 계수들 또는 그의 분해들에 대해 오더 감소를 수행하여 감소된 구면 조화 계수들 또는 감소된 분해들을 생성하게 하고, 복수의 구면 조화 계수들은 음장을 표현한다.

[0025] 또 다른 양태에서, 일 방법은 음장의 구별되는 성분을 나타내는 벡터의 계수들의 제 1 비-제로 세트를 획득하는 단계를 포함하며, 벡터는 음장을 설명하는 복수의 구면 조화 계수들로부터 분해된다.

[0026] 또 다른 양태에서, 일 디바이스는 음장의 구별되는 성분을 나타내는 벡터의 계수들의 제 1 비-제로 세트를 획득하도록 구성된 하나 이상의 프로세서들을 포함하며, 벡터는 음장을 설명하는 복수의 구면 조화 계수들로부터 분해된다.

[0027] 또 다른 양태에서, 일 디바이스는 음장의 구별되는 성분을 나타내는 벡터의 계수들의 제 1 비-제로 세트를 획득하는 수단으로서, 벡터는 음장을 설명하는 복수의 구면 조화 계수들로부터 분해되는, 상기 계수들의 제 1 비-제로 세트를 획득하는 수단, 및 계수들의 제 1 비-제로 세트를 저장하는 수단을 포함한다.

[0028] 또 다른 양태에서, 비-일시적 컴퓨터 판독가능 저장 매체는 저장된 명령들을 가지고, 그 명령들은 실행될 경우, 하나 이상의 프로세서들로 하여금, 음장의 구별되는 성분을 나타내는 벡터의 계수들의 제 1 비-제로 세트를 결정하게 하고, 벡터는 음장을 설명하는 복수의 구면 조화 계수들로부터 분해된다.

[0029] 또 다른 양태에서, 일 방법은 비트스트림으로부터, 배경 구면 조화 계수들과 재결합된 구면 조화 계수들로부터 분해된 하나 이상의 벡터들 중 적어도 하나를 획득하는 단계를 포함하며, 구면 조화 계수들은 음장을 설명하고, 배경 구면 조화 계수들은 동일한 음장의 하나 이상의 배경 성분들을 설명한다.

[0030] 또 다른 양태에서, 일 디바이스는 비트스트림으로부터, 배경 구면 조화 계수들과 재결합된 구면 조화 계수들로부터 분해된 하나 이상의 벡터들 중 적어도 하나를 결정하도록 구성된 하나 이상의 프로세서들을 포함하며, 구면 조화 계수들은 음장을 설명하고, 배경 구면 조화 계수들은 동일한 음장의 하나 이상의 배경 성분들을 설명한다.

[0031] 또 다른 양태에서, 일 디바이스는 비트스트림으로부터, 배경 구면 조화 계수들과 재결합된 구면 조화 계수들로부터 분해된 하나 이상의 벡터들 중 적어도 하나를 획득하는 수단을 포함하며, 구면 조화 계수들은 음장을 설명하고, 배경 구면 조화 계수들은 동일한 음장의 하나 이상의 배경 성분들을 설명한다.

[0032] 또 다른 양태에서, 비-일시적 컴퓨터 판독가능 저장 매체는 저장된 명령들을 가지고, 그 명령들은 실행될 경우, 하나 이상의 프로세서들로 하여금, 비트스트림으로부터, 배경 구면 조화 계수들과 재결합된 구면 조화 계수들로부터 분해된 하나 이상의 벡터들 중 적어도 하나를 획득하게 하며, 구면 조화 계수들은 음장을 설명하고, 배경 구면 조화 계수들은 동일한 음장의 하나 이상의 배경 성분들을 설명한다.

[0033] 또 다른 양태에서, 일 방법은 하나 이상의 오디오 오브젝트들에 대하여 결정된 방향성에 기초하여 오디오 오브젝트들과 연관된 하나 이상의 구면 조화 계수들 (SHC) 로부터 하나 이상의 구별되는 오디오 오브젝트들을 식별하는 단계를 포함한다.

[0034] 또 다른 양태에서, 일 디바이스는 하나 이상의 오디오 오브젝트들에 대하여 결정된 방향성에 기초하여 오디오 오브젝트들과 연관된 하나 이상의 구면 조화 계수들 (SHC) 로부터 하나 이상의 구별되는 오디오 오브젝트들을 식별하도록 구성된 하나 이상의 프로세서들을 포함한다.

[0035] 또 다른 양태에서, 일 디바이스는 하나 이상의 구면 조화 계수들 (SHC) 을 저장하는 수단, 및 하나 이상의 오디오 오브젝트들에 대하여 결정된 방향성에 기초하여 오디오 오브젝트들과 연관된 하나 이상의 구면 조화 계수들 (SHC) 로부터 하나 이상의 구별되는 오디오 오브젝트들을 식별하는 수단을 포함한다.

[0036] 또 다른 양태에서, 비-일시적 컴퓨터 판독가능 저장 매체는 저장된 명령들을 가지고, 그 명령들은 실행될 경우, 하나 이상의 프로세서들로 하여금, 하나 이상의 오디오 오브젝트들에 대하여 결정된 방향성에 기초하여 오디오 오브젝트들과 연관된 하나 이상의 구면 조화 계수들 (SHC) 로부터 하나 이상의 구별되는 오디오 오브젝트들을 식별하게 한다.

[0037] 또 다른 양태에서, 일 방법은 복수의 구면 조화 계수들에 대해 벡터-기반 합성을 수행하여 하나 이상의 오디오 오브젝트들을 나타내는 복수의 구면 조화 계수들의 분해된 표현들 및 대응하는 방향 정보를 생성하는 단계로서, 구면 조화 계수들은 오더와 연관되고 음장을 설명하는, 상기 벡터-기반 합성을 수행하는 단계, 방향 정보로부터 구별되는 배경 방향 정보를 결정하는 단계, 배경 오디오 오브젝트들과 연관된 방향 정보의 오더를 감소시켜 변환된 배경 방향 정보를 생성하는 단계, 변환된 방향 정보의 값들을 증가시키기 위해 보상을 적용하여 음장의 전체 에너지를 보존하는 단계를 포함한다.

[0038] 또 다른 양태에서, 일 디바이스는 복수의 구면 조화 계수들에 대해 벡터-기반 합성을 수행하여 하나 이상의 오디오 오브젝트들을 나타내는 복수의 구면 조화 계수들의 분해된 표현들 및 대응하는 방향 정보를 생성하는 것으로서, 구면 조화 계수들은 오더와 연관되고 음장을 설명하는, 상기 벡터-기반 합성을 수행하고, 방향 정보로부터 구별되는 배경 방향 정보를 결정하고, 배경 오디오 오브젝트들과 연관된 방향 정보의 오더를 감소시켜 변환된 배경 방향 정보를 생성하고, 변환된 방향 정보의 값들을 증가시키기 위해 보상을 적용하여 음장의 전체 에너지를 보존하도록 구성된 하나 이상의 프로세서들을 포함한다.

[0039] 또 다른 양태에서, 일 디바이스는 복수의 구면 조화 계수들에 대해 벡터-기반 합성을 수행하여 하나 이상의 오디오 오브젝트들을 나타내는 복수의 구면 조화 계수들의 분해된 표현들 및 대응하는 방향 정보를 생성하는 수단으로서, 구면 조화 계수들은 오더와 연관되고 음장을 설명하는, 상기 벡터-기반 합성을 수행하는 수단, 방향 정보로부터 구별되는 배경 방향 정보를 결정하는 수단, 배경 오디오 오브젝트들과 연관된 방향 정보의 오더를 감

소시켜 변환된 배경 방향 정보를 생성하는 수단, 및 변환된 방향 정보의 값들을 증가시키기 위해 보상을 적용하여 음장의 전체 에너지를 보존하는 수단을 포함한다.

[0040] 또 다른 양태에서, 비-일시적 컴퓨터 판독가능 저장 매체는 저장된 명령들을 가지고, 그 명령들은 실행될 경우, 하나 이상의 프로세서들로 하여금, 복수의 구면 조화 계수들에 대해 벡터-기반 합성을 수행하여 하나 이상의 오디오 오브젝트들을 나타내는 복수의 구면 조화 계수들의 분해된 표현들 및 대응하는 방향 정보를 생성하게 하는 것으로서, 구면 조화 계수들은 오더와 연관되고 음장을 설명하는, 상기 벡터-기반 합성을 수행하게 하고, 방향 정보로부터 구별되는 배경 방향 정보를 결정하게 하고, 배경 오디오 오브젝트들과 연관된 방향 정보의 오더를 감소시켜 변환된 배경 방향 정보를 생성하게 하고, 변환된 방향 정보의 값들을 증가시키기 위해 보상을 적용하여 음장의 전체 에너지를 보존하게 한다.

[0041] 또 다른 양태에서, 일 방법은 제 1 복수의 구면 조화 계수들의 제 1 분해 및 제 2 복수의 구면 조화 계수들의 제 2 분해에 대해 적어도 부분적으로 보간을 수행함으로써, 시간 세그먼트에 대하여 분해된 보간된 구면 조화 계수들을 획득하는 단계를 포함한다.

[0042] 또 다른 양태에서, 일 디바이스는 제 1 복수의 구면 조화 계수들의 제 1 분해 및 제 2 복수의 구면 조화 계수들의 제 2 분해에 대해 적어도 부분적으로 보간을 수행함으로써, 시간 세그먼트에 대하여 분해된 보간된 구면 조화 계수들을 획득하도록 구성된 하나 이상의 프로세서들을 포함한다.

[0043] 또 다른 양태에서, 일 디바이스는 제 1 복수의 구면 조화 계수들 및 제 2 복수의 구면 조화 계수들을 저장하는 수단, 및 제 1 복수의 구면 조화 계수들의 제 1 분해 및 제 2 복수의 구면 조화 계수들의 제 2 분해에 대해 적어도 부분적으로 보간을 수행함으로써, 시간 세그먼트에 대하여 분해된 보간된 구면 조화 계수들을 획득하는 수단을 포함한다.

[0044] 또 다른 양태에서, 또 다른 양태에서, 비-일시적 컴퓨터 판독가능 저장 매체는 저장된 명령들을 가지고, 그 명령들은 실행될 경우, 하나 이상의 프로세서들로 하여금, 제 1 복수의 구면 조화 계수들의 제 1 분해 및 제 2 복수의 구면 조화 계수들의 제 2 분해에 대해 적어도 부분적으로 보간을 수행함으로써, 시간 세그먼트에 대하여 분해된 보간된 구면 조화 계수들을 획득하게 한다.

[0045] 또 다른 양태에서, 일 방법은 음장의 공간 성분의 압축 버전을 포함하는 비트스트림을 획득하는 단계를 포함하며, 공간 성분은 복수의 구면 조화 계수들에 대해 벡터 기반 합성을 수행함으로써 생성된다.

[0046] 또 다른 양태에서, 일 디바이스는 음장의 공간 성분의 압축 버전을 포함하는 비트스트림을 획득하도록 구성된 하나 이상의 프로세서들을 포함하며, 공간 성분은 복수의 구면 조화 계수들에 대해 벡터 기반 합성을 수행함으로써 생성된다.

[0047] 또 다른 양태에서, 일 디바이스는 음장의 공간 성분의 압축 버전을 포함하는 비트스트림을 획득하는 수단으로서, 공간 성분은 복수의 구면 조화 계수들에 대해 벡터 기반 합성을 수행함으로써 생성되는, 상기 비트스트림을 획득하는 수단, 및 비트스트림을 저장하는 수단을 포함한다.

[0048] 또 다른 양태에서, 비-일시적 컴퓨터 판독가능 저장 매체는 저장된 명령들을 가지고, 그 명령들은 실행될 경우, 하나 이상의 프로세서들로 하여금, 음장의 공간 성분의 압축 버전을 포함하는 비트스트림을 획득하게 하며, 공간 성분은 복수의 구면 조화 계수들에 대해 벡터 기반 합성을 수행함으로써 생성된다.

[0049] 또 다른 양태에서, 일 방법은 음장의 공간 성분의 압축 버전을 포함하는 비트스트림을 생성하는 단계를 포함하며, 공간 성분은 복수의 구면 조화 계수들에 대해 벡터 기반 합성을 수행함으로써 생성된다.

[0050] 또 다른 양태에서, 일 디바이스는 음장의 공간 성분의 압축 버전을 포함하는 비트스트림을 생성하도록 구성된 하나 이상의 프로세서들을 포함하며, 공간 성분은 복수의 구면 조화 계수들에 대해 벡터 기반 합성을 수행함으로써 생성된다.

[0051] 또 다른 양태에서, 일 디바이스는 음장의 공간 성분의 압축 버전을 포함하는 비트스트림을 생성하는 수단으로서, 공간 성분은 복수의 구면 조화 계수들에 대해 벡터 기반 합성을 수행함으로써 생성되는, 상기 비트스트림을 생성하는 수단, 및 비트스트림을 저장하는 수단을 포함한다.

[0052] 또 다른 양태에서, 비-일시적 컴퓨터 판독가능 저장 매체는 저장된 명령들을 가지고, 그 명령들은 실행될 경우, 하나 이상의 프로세서들로 하여금, 음장의 공간 성분의 압축 버전을 포함하는 비트스트림을 생성하게 하며, 공간 성분은 복수의 구면 조화 계수들에 대해 벡터 기반 합성을 수행함으로써 생성된다.

[0053] 또 다른 양태에서, 일 방법은 복수의 압축된 공간 성분들 중 나머지 공간 성분들에 대한 공간 성분의 압축 버전의 오더에 기초하여 복수의 압축된 공간 성분들 중의 공간 성분의 압축된 버전을 압축해제할 때 사용할 허프만(Huffman) 코드북을 식별하는 단계를 포함하며, 공간 성분은 복수의 구면 조화 계수들에 대해 벡터 기반 합성을 수행함으로써 생성된다.

[0054] 또 다른 양태에서, 일 디바이스는 복수의 압축된 공간 성분들 중 나머지 공간 성분들에 대한 공간 성분의 압축 버전의 오더에 기초하여 복수의 압축된 공간 성분들 중의 공간 성분의 압축된 버전을 압축해제할 때 사용할 허프만 코드북을 식별하도록 구성된 하나 이상의 프로세서들을 포함하며, 공간 성분은 복수의 구면 조화 계수들에 대해 벡터 기반 합성을 수행함으로써 생성된다.

[0055] 또 다른 양태에서, 일 디바이스는 복수의 압축된 공간 성분들 중 나머지 공간 성분들에 대한 공간 성분의 압축 버전의 오더에 기초하여 복수의 압축된 공간 성분들 중의 공간 성분의 압축된 버전을 압축해제할 때 사용할 허프만 코드북을 식별하는 수단으로서, 공간 성분은 복수의 구면 조화 계수들에 대해 벡터 기반 합성을 수행함으로써 생성되는, 상기 허프만 코드북을 식별하는 수단, 및 복수의 압축된 공간 성분들을 저장하는 수단을 포함한다.

[0056] 또 다른 양태에서, 비-일시적 컴퓨터 판독가능 저장 매체는 저장된 명령들을 가지고, 그 명령들은 실행될 경우, 하나 이상의 프로세서들로 하여금, 복수의 압축된 공간 성분들 중 나머지 공간 성분들에 대한 공간 성분의 압축 버전의 오더에 기초하여 복수의 압축된 공간 성분들 중의 공간 성분의 압축된 버전을 압축해제할 때 사용할 허프만 코드북을 식별하게 하며, 공간 성분은 복수의 구면 조화 계수들에 대해 벡터 기반 합성을 수행함으로써 생성된다.

[0057] 또 다른 양태에서, 일 방법은 복수의 공간 성분들 중 나머지 공간 성분들에 대한 공간 성분의 오더에 기초하여 복수의 공간 성분들 중의 공간 성분을 압축할 때 사용할 허프만 코드북을 식별하는 단계를 포함하며, 공간 성분은 복수의 구면 조화 계수들에 대해 벡터 기반 합성을 수행함으로써 생성된다.

[0058] 또 다른 양태에서, 일 디바이스는 복수의 공간 성분들 중 나머지 공간 성분들에 대한 공간 성분의 오더에 기초하여 복수의 공간 성분들 중의 공간 성분을 압축할 때 사용할 허프만 코드북을 식별하도록 구성된 하나 이상의 프로세서들을 포함하며, 공간 성분은 복수의 구면 조화 계수들에 대해 벡터 기반 합성을 수행함으로써 생성된다.

[0059] 또 다른 양태에서, 일 디바이스는 허프만 코드북을 저장하는 수단, 및 복수의 공간 성분들 중 나머지 공간 성분들에 대한 공간 성분의 오더에 기초하여 복수의 공간 성분들 중의 공간 성분을 압축할 때 사용할 허프만 코드북을 식별하는 수단을 포함하며, 공간 성분은 복수의 구면 조화 계수들에 대해 벡터 기반 합성을 수행함으로써 생성된다.

[0060] 또 다른 양태에서, 비-일시적 컴퓨터 판독가능 저장 매체는 저장된 명령들을 가지고, 그 명령들은 실행될 경우, 하나 이상의 프로세서들로 하여금, 복수의 공간 성분들 중 나머지 공간 성분들에 대한 공간 성분의 오더에 기초하여 복수의 공간 성분들 중의 공간 성분을 압축할 때 사용할 허프만 코드북을 식별하게 하며, 공간 성분은 복수의 구면 조화 계수들에 대해 벡터 기반 합성을 수행함으로써 생성된다.

[0061] 또 다른 양태에서, 일 방법은 음장의 공간 성분을 압축할 때 사용될 양자화 스텝 사이즈를 결정하는 단계를 포함하며, 공간 성분은 복수의 구면 조화 계수들에 대해 벡터 기반 합성을 수행함으로써 생성된다.

[0062] 또 다른 양태에서, 일 디바이스는 음장의 공간 성분을 압축할 때 사용될 양자화 스텝 사이즈를 결정하도록 구성된 하나 이상의 프로세서들을 포함하며, 공간 성분은 복수의 구면 조화 계수들에 대해 벡터 기반 합성을 수행함으로써 생성된다.

[0063] 또 다른 양태에서, 일 디바이스는 음장의 공간 성분을 압축할 때 사용될 양자화 스텝 사이즈를 결정하는 수단으로서, 공간 성분은 복수의 구면 조화 계수들에 대해 벡터 기반 합성을 수행함으로써 생성되는, 상기 양자화 스텝 사이즈를 결정하는 수단, 및 양자화 스텝 사이즈를 저장하는 수단을 포함한다.

[0064] 또 다른 양태에서, 비-일시적 컴퓨터 판독가능 저장 매체는 저장된 명령들을 가지고, 그 명령들은 실행될 경우, 하나 이상의 프로세서들로 하여금, 음장의 공간 성분을 압축할 때 사용될 양자화 스텝 사이즈를 결정하게 하며, 공간 성분은 복수의 구면 조화 계수들에 대해 벡터 기반 합성을 수행함으로써 생성된다.

[0065] 기술들의 하나 이상의 양태들의 세부사항들은 첨부된 도면들 및 이하 설명에서 기술된다. 이들 기술들의 다

른 특징들, 목적들, 및 장점들은 그 설명 및 도면들로부터, 및 청구범위들로부터 명백할 것이다.

도면의 간단한 설명

[0066] 도 1 및 도 2 는 다양한 오더들 및 서브-오더들의 구면 조화 기저 함수들을 예시하는 다이어그램들이다.

도 3 은 본 개시물에서 설명되는 기술들의 다양한 양태들을 수행할 수도 있는 시스템을 예시하는 다이어그램이다.

도 4 는 본 개시물에서 설명되는 기술들의 다양한 양태들을 수행할 수도 있는 도 3 의 예에 도시된 오디오 인코딩 디바이스의 일 예를 더 상세하게 예시하는 블록 다이어그램이다.

도 5 는 도 3 의 오디오 디코딩 디바이스를 더 상세하게 예시하는 블록 다이어그램이다.

도 6 은 본 개시물에서 설명되는 기술들의 다양한 양태들을 수행할 시 오디오 인코딩 디바이스의 컨텐츠 분석 유닛의 예시적인 동작을 예시하는 플로우차트이다.

도 7 은 본 개시물에서 설명되는 벡터-기반 합성 기술들의 다양한 양태들을 수행할 시 오디오 인코딩 디바이스의 예시적인 동작을 예시하는 플로우차트이다.

도 8 은 본 개시물에서 설명되는 기술들의 다양한 양태들을 수행할 시 오디오 인코딩 디바이스의 예시적인 동작을 예시하는 플로우차트이다.

도 9a 내지 도 9l 은 도 4 의 예의 오디오 인코딩 디바이스의 다양한 양태들을 더 상세하게 예시하는 블록 다이어그램들이다.

도 10a 내지 도 10o(ii) 는 압축된 공간 성분들을 더 상세하게 명시할 수도 있는 비트스트림 또는 측면 채널 정보의 일부분을 예시하는 다이어그램들이다.

도 11a 내지 도 11g 는 도 5 의 예에 도시된 오디오 디코딩 디바이스의 다양한 유닛들을 더 상세하게 예시하는 블록 다이어그램들이다.

도 12 는 본 개시물에서 설명되는 기술들의 다양한 양태들을 수행할 수도 있는 예시적인 오디오 에코시스템을 예시하는 다이어그램이다.

도 13 은 도 12 의 오디오 에코시스템의 일 예를 더 상세하게 예시하는 다이어그램이다.

도 14 는 도 12 의 오디오 에코시스템의 일 예를 더 상세하게 예시하는 다이어그램이다.

도 15a 및 도 15b 는 도 12 의 오디오 에코시스템의 다른 예들을 더 상세하게 예시하는 다이어그램들이다.

도 16 은 본 개시물에서 설명되는 기술들의 다양한 양태들을 수행할 수도 있는 예시적인 오디오 인코딩 디바이스를 예시하는 다이어그램이다.

도 17 은 도 16 의 오디오 인코딩 디바이스의 일 예를 더 상세하게 예시하는 다이어그램이다.

도 18 은 본 개시물에서 설명되는 기술들의 다양한 양태들을 수행할 수도 있는 예시적인 오디오 디코딩 디바이스를 예시하는 다이어그램이다.

도 19 는 도 18 의 오디오 디코딩 디바이스의 일 예를 더 상세하게 예시하는 다이어그램이다.

도 20a 내지 도 20g 는 본 개시물에서 설명되는 기술들의 다양한 양태들을 수행할 수도 있는 예시적인 오디오 포착 디바이스들을 예시하는 다이어그램들이다.

도 21a 내지 도 21e 는 본 개시물에서 설명되는 기술들의 다양한 양태들을 수행할 수도 있는 예시적인 오디오 재생 디바이스들을 예시하는 다이어그램들이다.

도 22a 내지 도 22h 는 본 개시물에서 설명되는 하나 이상의 기술들에 따른 예시적인 오디오 재생 환경들을 예시하는 다이어그램들이다.

도 23 은 본 개시물에서 설명된 하나 이상의 기술들에 따라 사용자가 헤드폰들을 착용하면서 스포츠 게임의 3D 음장을 경험할 수도 있는 예시적인 사용 케이스를 예시하는 다이어그램이다.

도 24 는 본 개시물에서 설명된 하나 이상의 기술들에 따라 3D 음장이 레코딩될 수도 있는 스포츠 스타디움을

예시하는 다이어그램이다.

도 25 는 본 개시물에서 설명된 하나 이상의 기술들에 따라 로컬 오디오 랜드스케이프에 기초하여 3D 음장을 렌더링하는 기술을 예시하는 플로우 다이어그램이다.

도 26 은 본 개시물에서 설명되는 하나 이상의 기술들에 따른 예시적인 게임 스튜디오를 예시하는 다이어그램이다.

도 27 은 본 개시물에서 설명되는 하나 이상의 기술들에 따른 렌더링 엔진들을 포함하는 복수의 게임 시스템들을 예시하는 다이어그램이다.

도 28 은 본 개시물에서 설명되는 하나 이상의 기술들에 따라 헤드폰들에 의해 시뮬레이팅될 수도 있는 스피커 구성을 예시하는 다이어그램이다.

도 29 는 본 개시물에서 설명되는 하나 이상의 기술들에 따라 3D 음장을 포착 및/또는 편집하는데 사용될 수도 있는 복수의 모바일 디바이스들을 예시하는 다이어그램이다.

도 30 은 본 개시물에서 설명된 하나 이상의 기술들에 따라 프로세싱될 수도 있는 3D 음장과 연관된 비디오 프레임을 예시하는 다이어그램이다.

도 31a 내지 도 31m 은 본 개시물에서 설명된 기술들의 다양한 양태들에 따라 음장의 합성의 또는 레코딩된 카테고리화를 수행한 다양한 시뮬레이션 결과들을 보여주는 그래프들을 예시하는 다이어그램들이다.

도 32 는 본 개시물에서 설명되는 기술들에 따라 고차 앰비소닉 계수들로부터 분해된 S 행렬로부터의 특이값들의 그래프를 예시하는 다이어그램이다.

도 33a 및 도 33b 는 본 개시물에서 설명된 기술들에 따라 음장의 전경 성분들을 설명하는 벡터들을 인코딩할 때 리오더링이 갖는 잠정적인 영향을 보여주는 개별 그래프들을 예시하는 다이어그램들이다.

도 34 및 도 35 는 본 개시물에 따라, 구별되는 오디오 오브젝트들의 오직 에너지-기반 식별과 방향-기반 식별 간의 차이들을 예시하는 개념 다이어그램들이다.

도 36a 내지 도 36g 는 본 개시물에서 설명된 기술들의 다양한 양태들에 따라 보간을 수행하도록, 구면 조화 계수들의 분해된 버전의 적어도 일부분의 공간 도메인으로의 투영들을 예시하는 다이어그램들이다.

도 37 은 본 명세서에 설명된 것과 같이 공간-시간 보간을 획득하는 기술들의 표현을 예시한다.

도 38 은 본 명세서에 설명된 기술들에 따른 다차원 신호의 순차적인 SVD 블록들에 대한 인공적인 US 행렬들, US1 및 US2 을 예시하는 블록 다이어그램이다.

도 39 는 본 개시물에서 설명되는 기술들에 따라 공간-시간 성분들의 특이값 분해 및 평활화를 사용하는 고차 앰비소닉스 (HOA) 의 순차적인 프레임들의 분해를 예시하는 블록 다이어그램이다.

도 40a 내지 도 40j 는 각각, 2 또는 3 차원 음장들을 설명하는 구면 조화 계수들을 압축하기 위해 본 개시물에서 설명되는 기술들의 다양한 양태들을 수행할 수도 있는 예시적인 오디오 인코딩 디바이스들을 예시하는 블록 다이어그램이다.

도 41a 내지 도 40d 는 각각, 2 또는 3 차원 음장들을 설명하는 구면 조화 계수들을 디코딩하기 위해 본 개시물에서 설명되는 기술들의 다양한 양태들을 수행할 수도 있는 예시적인 오디오 디코딩 디바이스들을 예시하는 블록 다이어그램들이다.

도 42a 내지 도 42c 는 각각, 도 40b 내지 도 40j 의 예들에 도시된 오더 감소 유닛을 더 상세하게 예시하는 블록 다이어그램들이다.

도 43 은 도 40i 에 도시된 V 압축 유닛을 예시하는 다이어그램이다.

도 44 는 본 개시물에서 설명되는 기술들의 다양한 양태들에 따라 양자화 에러에 대하여 보상하기 위해 오디오 인코딩 디바이스에 의해 수행되는 예시적인 동작들을 예시하는 다이어그램이다.

도 45a 및 도 45b 는 본 개시물에서 설명된 기술들의 다양한 양태들에 따라 2 개의 프레임들의 부분들로부터의 서브-프레임들의 보간을 예시하는 다이어그램들이다.

도 46a 내지 도 46e 는 본 개시물에서 설명된 기술들에 따라 보간된 복수의 구면 조화 계수들의 분해된 버전의

하나 이상의 벡터들의 투영의 횡단면을 예시하는 다이어그램들이다.

도 47 은 도 41a 내지 도 41d 의 예들에 도시된 오디오 디코딩 디바이스들의 추출 유닛을 더 상세하게 예시하는 블록 다이어그램들이다.

도 48 은 도 41a 내지 도 41d 의 예들에 도시된 오디오 디코딩 디바이스의 오디오 렌더링 유닛을 더 상세하게 예시하는 블록 다이어그램이다.

도 49a 내지 도 49e(ii) 는 본 개시물에서 설명되는 기술들의 다양한 양태들을 구현할 수도 있는 개별 오디오 코딩 시스템들을 예시하는 다이어그램들이다.

도 50a 및 도 50b 는 본 개시물에서 설명된 기술들에 따라 배경 컨텐츠의 오더를 잠재적으로 감소시키는 2 개의 상이한 접근법들 중 하나를 각각 예시하는 블록 다이어그램들이다.

도 51 은 구면 조화 계수들을 압축하기 위해 본 개시물에서 설명되는 기술들의 다양한 양태들을 구현할 수도 있는 오디오 인코딩 디바이스의 구별되는 성분 압축 경로의 예들을 예시하는 블록 다이어그램이다.

도 52 는 구면 조화 계수들 (SHC) 을 복원하거나 거의 복원하기 위해 본 개시물에서 설명되는 기술들의 다양한 양태들을 구현할 수도 있는 오디오 디코딩 디바이스의 다른 예를 예시하는 블록 다이어그램이다.

도 53 은 본 개시물에서 설명되는 기술들의 다양한 양태들을 수행할 수도 있는 오디오 인코딩 디바이스의 다른 예를 예시하는 블록 다이어그램이다.

도 54 는 도 53 의 예에 도시된 오디오 인코딩 디바이스의 예시적인 구현을 더 상세하게 예시하는 블록 다이어그램이다.

도 55a 및 도 55b 는 음장을 회전시키기 위해 본 개시물에 설명된 기술들의 다양한 양태들을 수행하는 예를 예시하는 다이어그램들이다.

도 56 은 제 2 참조 프레임 면에서 음장을 표현하기 위해 본 개시물에서 설명되는 기술들에 따라 회전되는 제 1 참조 프레임에 따라 캡처된 예시적인 음장을 예시하는 다이어그램이다.

도 57a 내지 도 57e 는 각각, 본 개시물에서 설명되는 기술들에 따라 형성된 비트스트림들을 예시하는 다이어그램이다.

도 58 은 본 개시물에서 설명되는 기술들의 회전 양태들을 구현할 때 도 53 의 예에 도시된 오디오 인코딩 디바이스의 예시적인 동작을 예시하는 플로우차트이다.

도 59 는 본 개시물에서 설명되는 기술들의 변환 양태들을 수행할 때 도 53 의 예에 도시된 오디오 인코딩 디바이스의 예시적인 동작을 예시하는 플로우차트이다.

발명을 실시하기 위한 구체적인 내용

[0067] 서라운드 사운드의 진화는, 요즘, 엔터테인먼트를 위한 다수의 출력 포맷들을 이용가능하게 하였다. 그러한 소비자 서라운드 사운드 포맷들의 예들은, 특정 기하학적 좌표들에서의 라우드스피커들로의 피드들을 암시적으로 명시한다는 점에 있어서 대부분 '채널' 기반이다. 이들은 대중적인 5.1 포맷 (이는 다음의 6개 채널들: 즉, 전방 좌측 (FL), 전방 우측 (FR), 중앙 또는 전방 중앙, 후방 좌측 또는 서라운드 좌측, 후방 우측 또는 서라운드 우측, 및 저주파수 효과들 (LFE) 을 포함함), 및 성장하는 7.1 포맷, (예를 들어, 울트라 고선명 텔레비전 표준으로 사용하기 위한) 22.2 포맷 및 7.1.4 포맷과 같은 높이 스피커들을 포함하는 다양한 포맷들을 포함한다. 비-소비자 포맷들은, 종종 '서라운드 어레이들' 로 명명되는 (대칭 및 비대칭 지오메트리들에 있어서) 임의의 수의 스피커들에 걸칠 수 있다. 그러한 어레이의 일 예는 절단된 20면체의 코너들 상의 좌표들에 위치된 32개의 라우드스피커들을 포함한다.

[0068] 장래의 MPEG 인코더로의 입력은 옵션적으로, 3가지 가능한 포맷들 중 하나이다: 즉, (i) 미리 명시된 포지션들에서의 라우드스피커들을 통해 재생되도록 의도된 (상기 논의된 바와 같은) 종래의 채널 기반 오디오; (ii) (다른 정보 중에서) 그 위치 좌표를 포함하는 메타데이터와 연관된 단일 오디오 오브젝트들에 대한 이산 펄스 코드 변조 (PCM) 데이터를 오브젝트 기반 오디오; 및 (iii) 구면 조화 기저 함수들의 계수들 (또는 "구면 조화 계수들" 또는 SHC, "고차 앰비소닉스" 또는 HOA, 및 "HOA 계수들" 로도 지칭됨) 을 사용하여 음장을 나타낸 것을 수반하는 장면 기반 오디오. 이러한 장래의 MPEG 인코더는 스위스 제네바에서 2013년 1월 공개되고 http://mpeg.chiariglione.org/sites/default/files/files/standards/parts/docs/w13411.zip 에서 입수가능한

국제 표준화 기구/국제 전자기술 위원회 (ISO)/(IEC) JTC1/SC29/WG11/N13411 에 의한 "Call for Proposals for 3D Audio" 의 명칭의 문헌에 더 상세하게 기술될 수도 있다.

[0069] 다양한 '서라운드 사운드' 채널 기반 포맷들이 시장에 존재한다. 이들은, 예를 들어, 5.1 홈 씨어터 시스템 (스테레오를 넘어 거실로의 잠식을 행하는 것의 관점에서 가장 성공적이었음) 으로부터 NHK (Nippon Hoso Kyokai 또는 일본 방송사) 에 의해 개발된 22.2 시스템까지에 이른다. 컨텐츠 창작자 (예를 들어, 헐리우드 스튜디오) 는 영화용 사운드트랙을 한번 제작하고 싶어 하며, 각각의 스피커 구성을 위해 그것을 리믹스하려는 노력을 쏟지는 않는다. 최근, 표준 개발 기구들은 표준화된 비트스트림으로의 인코딩, 및 재생 (렌더러 수반) 의 위치에서 스피커 지오메트리 (및 개수) 및 음향 조건들에 대해 적응가능하고 불가지한 후속 디코딩을 제공하는 방식들을 고려하고 있었다.

[0070] 컨텐츠 창작자들에 대한 그러한 유연성을 제공하기 위해, 엘리먼트들의 계위적 세트가 음장을 나타내기 위해 사용될 수도 있다. 엘리먼트들의 계위적 세트는, 하위 오더링된 엘리먼트들의 기본 세트가 모델링된 음장의 전체 표현을 제공하도록 엘리먼트들이 오더링되는 엘리먼트들의 세트를 지칭할 수도 있다. 그 세트가 상위 오더 엘리먼트들을 포함하도록 확장됨에 따라, 그 표현은 더 상세화되어, 해상도를 증가시킨다.

[0071] 엘리먼트들의 계위적 세트의 일 예는 구면 조화 계수들 (SHC) 의 세트이다. 다음의 식은 SHC 를 사용하여 음장의 서술 또는 표현을 나타낸다:

[0072]
$$p_i(t, r_r, \theta_r, \varphi_r) = \sum_{\omega=0}^{\infty} \left[4\pi \sum_{n=0}^{\infty} j_n(kr_r) \sum_{m=-n}^{n} A_n^m(k)\, Y_n^m(\theta_r, \varphi_r) \right] e^{j\omega t}$$

[0073] 이 식은 시간 t 에서 음장의 임의의 포인트 $\{r_r, \theta_r, \varphi_r\}$ 에서의 압력 (p_i) 이 SHC ($A_n^m(k)$) 에 의해 고유하게 표현될 수 있음을 나타낸다. 여기서, $k = \frac{\omega}{c}$ 이고, c 는 사운드의 속도 (~343 m/s) 이고, $\{r_r, \theta_r, \varphi_r\}$은 참조의 포인트 (또는 관측 포인트) 이고, $j_n(\cdot)$ 는 오더 n 의 곡면 베셀 함수이고, $Y_n^m(\theta_r, \varphi_r)$는 오더 n 및 서브오더 m 의 구면 조화 기저 함수들이다. 대괄호의 용어는 신호의 주파수 도메인 표현 (즉, $S(\omega, r_r, \theta_r, \varphi_r)$) 이고, 이는 이산 푸리에 변환 (DFT), 이산 코사인 변환 (DCT), 또는 웨이블릿 변환과 같은 다양한 시간-주파수 변환들에 의해 근사화될 수 있음이 인식될 수 있다. 계위적 세트들의 다른 예들은 웨이블릿 변환 계수들의 세트들 및 다해상도 기저 함수들의 계수들의 다른 세트들을 포함한다.

[0074] 도 1 은 제로 오더 (n = 0) 로부터 제 4 오더 (n = 4) 까지의 구면 조화 기저 함수들을 예시하는 다이어그램이다. 볼 수 있는 바와 같이, 각각의 오더에 대해, 예시 목적들의 용이를 위해 도 1 의 예에 도시되지만 명시적으로 노트되지 않는 서브오더들 (m) 의 확장이 존재한다.

[0075] 도 2 는 제로 오더 (n = 0) 로부터 제 4 오더 (n = 4) 까지의 구면 조화 기저 함수들을 도시한 다른 다이어그램이다. 도 2 에 있어서, 구면 조화 기저 함수들이 오더 및 서브오더 양자가 도시된 2 차원 좌표 공간에 도시된다.

[0076] SHC ($A_n^m(k)$) 는 다양한 마이크로폰 어레이 구성들에 의해 물리적으로 획득 (예를 들어, 레코딩) 될 수 있거나, 대안적으로, 음장의 채널 기반 또는 오브젝트 기반 디스크립션들로부터 도출될 수 있다. SHC 는 장면 기반 오디오를 나타내며, 여기서, SHC 는, 더 효율적인 송신 또는 저장을 촉진할 수도 있는 인코딩된 SHC 를 획득하기 위한 오디오 인코더에 입력될 수도 있다. 예를 들어, $(1+4)^2$ (25, 및 따라서 제 4 오더) 계수들을 수반하는 제 4 오더 표현이 사용될 수도 있다.

[0077] 상기 노트된 바와 같이, SHC 는 마이크로폰을 사용한 마이크로폰 레코딩으로부터 도출될 수도 있다. SHC 가 마이크로폰 어레이들로부터 어떻게 도출될 수 있는지의 다양한 예들은 Poletti, M., 의 "Three-Dimensional Surround Sound Systems Based on Spherical Harmonics", J.Audio Eng.Soc., Vol. 53, No. 11, 2005년 11월자, pp. 1004-1025 에 기술된다.

[0078] 이들 SHC들이 오브젝트 기반 디스크립션으로부터 어떻게 도출될 수 있는지를 예시하기 위해, 다음의 수학식을 고려한다. 개별 오디오 오브젝트에 대응하는 음장에 대한 계수들 ($A_n^m(k)$) 은

[0079] $$A_n^m(k) = g(\omega)(-4\pi ik)h_n^{(2)}(kr_s)Y_n^{m^*}(\theta_s,\varphi_s)$$

[0080] 와 같이 표현될 수도 있으며, 여기서, i 는 $\sqrt{-1}$ 이고, $h_n^{(2)}(\cdot)$ 는 오더 n 의 (제 2 종의) 구면 핸켈 함수이고, $\{r_s,\theta_s,\varphi_s\}$ 는 오브젝트의 위치이다. (예를 들어, PCM 스트림에 대한 고속 푸리에 변환을 수행하는 것과 같은 시간-주파수 분석 기술을 사용하여) 주파수의 함수로서 오브젝트 소스 에너지 (g(ω)) 를 아는 것은 각각의 PCM 오브젝트 및 그 위치를 SHC ($A_n^m(k)$) 로 변환하게 한다. 추가로, 이는 (상기가 선형 및 직교 분해이기 때문에) 각각의 오브젝트에 대한 $A_n^m(k)$ 계수들이 가산적임을 보일 수 있다. 이러한 방식으로, 다수의 PCM 오브젝트들은 (예를 들어, 개별 오브젝트들에 대한 계수 벡터들의 합으로서) $A_n^m(k)$ 계수들에 의해 표현될 수 있다. 본질적으로, 이들 계수들은 음장에 관한 정보 (3D 좌표의 함수로서의 압력) 를 포함하고, 상기는 관측 포인트 $\{r_r,\theta_r,\varphi_r\}$ 의 근방에서, 개별 오브젝트로부터 전체 음장의 표현으로의 변환을 나타낸다. 나머지 도면들은 오브젝트 기반 및 SHC 기반 오디오 코딩의 컨텍스트에 있어서 하기에서 설명된다.

[0081] 도 3 은 본 개시물에서 설명된 기술들의 다양한 양태들을 수행할 수도 있는 시스템 (10) 을 예시하는 다이어그램이다. 도 3 의 예에 도시된 바와 같이, 시스템 (10) 은 컨텐츠 창작자 (12) 및 컨텐츠 소비자 (14) 를 포함한다. 컨텐츠 창작자 (12) 및 컨텐츠 소비자 (14) 의 컨텍스트에서 설명되지만, 그 기술들은 음장의 (HOA 계수들로서 또한 지칭될 수도 있는) SHC들 또는 임의의 다른 계위적 표현이 오디오 데이터의 비트스트림 표현을 형성하도록 인코딩되는 임의의 컨텍스트에서 구현될 수도 있다. 더욱이, 컨텐츠 창작자 (12) 는, 몇몇 예들을 제공하기 위해 핸드셋 (또는 셀룰러 전화기), 태블릿 컴퓨터, 스마트 폰, 또는 데스크탑 컴퓨터를 포함하여 본 개시물에서 설명된 기술들을 구현 가능한 임의의 형태의 컴퓨팅 디바이스를 나타낼 수도 있다. 유사하게, 컨텐츠 소비자 (14) 는, 몇몇 예들을 제공하기 위해 핸드셋 (또는 셀룰러 전화기), 태블릿 컴퓨터, 스마트 폰, 셋탑 박스, 또는 데스크탑 컴퓨터를 포함하여 본 개시물에서 설명된 기술들을 구현 가능한 임의의 형태의 컴퓨팅 디바이스를 나타낼 수도 있다.

[0082] 컨텐츠 창작자 (12) 는, 컨텐츠 소비자 (14) 와 같은 컨텐츠 소비자들에 의한 소비를 위해 다중 채널 오디오 컨텐츠를 생성할 수도 있는 영화 스튜디오 또는 다른 엔터티를 나타낼 수도 있다. 일부 예들에 있어서, 컨텐츠 창작자 (12) 는, HOA 계수들 (11) 을 압축하고자 하는 개별 사용자를 나타낼 수도 있다. 종종, 이러한 컨텐츠 창작자는 비디오 컨텐츠와 함께 오디오 컨텐츠를 생성한다. 컨텐츠 소비자 (14) 는 오디오 재생 시스템을 소유하거나 오디오 재생 시스템에 액세스하는 개인을 나타내며, 이 오디오 재생 시스템은 다중 채널 오디오 컨텐츠로서의 재생을 위해 SHC 를 렌더링가능한 임의의 형태의 오디오 재생 시스템을 지칭할 수도 있다.
도 3 의 예에 있어서, 컨텐츠 소비자 (14) 는 오디오 재생 시스템 (16) 을 포함한다.

[0083] 컨텐츠 창작자 (12) 는 오디오 편집 시스템 (18) 을 포함한다. 컨텐츠 창작자 (12) 는 컨텐츠 창작자 (12) 가 오디오 편집 시스템 (18) 을 사용하여 편집할 수도 있는, 다양한 포맷들 (HOA 계수들로서 직접 포함) 의 라이브 레코딩들 (7) 및 오디오 오브젝트들 (9) 을 획득한다. 컨텐츠 창작자는, 편집 프로세스 동안, 오디오 오브젝트들 (9) 로부터 HOA 계수들 (11) 을 렌더링하여, 추가로 편집할 것을 요구하는 음장의 다양한 양태들을 식별하기 위한 시도 시 렌더링된 스피커 피드들을 리스닝할 수도 있다. 그 후, 컨텐츠 창작자 (12) 는 (소스 HOA 계수들이 상기 설명된 방식으로 도출될 수도 있는 오디오 오브젝트들 (9) 중 상이한 오디오 오브젝트들의 조작을 통해 잠재적으로 간접적으로) HOA 계수들 (11) 을 편집할 수도 있다. 컨텐츠 창작자 (12) 는 오디오 편집 시스템 (18) 을 채용하여 HOA 계수들 (11) 을 생성할 수도 있다. 오디오 편집 시스템 (18) 은 오디오 데이터를 편집하고 그리고 이 오디오 데이터를 하나 이상의 소스 구면 조화 계수들로서 출력 가능한 임의의 시스템을 나타낸다.

[0084] 편집 프로세스가 완료할 경우, 컨텐츠 창작자 (12) 는 HOA 계수들 (11) 에 기초하여 비트스트림 (21) 을 생성할 수도 있다. 즉, 컨텐츠 창작자 (12) 는 비트스트림 (21) 을 생성하기 위해 본 개시물에서 설명된 기술들의 다양한 양태들에 따라 HOA 계수들 (11) 을 인코딩하거나 그렇지 않으면 압축하도록 구성된 디바이스를 나타내는 오디오 인코딩 디바이스 (20) 를 포함한다. 오디오 인코딩 디바이스 (20) 는, 일 예로서, 유선 또는 무선 채널, 데이터 저장 디바이스 등일 수도 있는 송신 채널 상으로의 송신을 위해 비트스트림 (21) 을 생성할 수도 있다. 비트스트림 (21) 은 HOA 계수들 (11) 의 인코딩된 버전을 나타낼 수도 있고, 프라이머리 비트스트림 또는 다른 사이드 비트스트림을 포함할 수도 있으며, 이 사이드 비트스트림은 사이드 채널 정보로서 지칭될 수

도 있다.

[0085] 비록 하기에서 더 상세하게 설명되더라도, 오디오 인코딩 디바이스 (20) 는 벡터 기반 합성 또는 방향 기반 합성에 기초하여 HOA 계수들 (11) 을 인코딩하도록 구성될 수도 있다. 벡터 기반 합성 방법을 수행할지 또는 방향 기반 합성 방법을 수행할지를 결정하기 위해, 오디오 인코딩 디바이스 (20) 는, HOA 계수들 (11) 에 적어도 부분적으로 기초하여, HOA 계수들 (11) 이 음장의 자연적 레코딩 (예를 들어, 라이브 레코딩 (7)) 을 통해 생성되었는지 또는 일 예로서 PCM 오브젝트와 같은 오디오 오브젝트들 (9) 로부터 인공적으로 (즉, 합성적으로) 생성되었는지 여부를 결정할 수도 있다. HOA 계수들 (11) 이 오디오 오브젝트들 (9) 로부터 생성되었을 경우, 오디오 인코딩 디바이스 (20) 는 방향 기반 합성 방법을 사용하여 HOA 계수들 (11) 을 인코딩할 수도 있다.

HOA 계수들 (11) 이 예를 들어 아이겐마이크를 사용하여 라이브로 캡처되었을 경우, 오디오 인코딩 디바이스 (20) 는 벡터 기반 합성 방법을 사용하여 HOA 계수들 (11) 을 인코딩할 수도 있다. 상기 구별은 벡터 기반 또는 방향 기반 합성 방법이 사용될 수도 있는 경우의 일 예를 나타낸다. 이들 중 어느 하나 또는 그 양자는 자연적 레코딩들, 인공적으로 생성된 컨텐츠 또는 이 둘의 혼합 (하이브리드 컨텐츠) 을 위해 유용할 수도 있는 다른 경우들이 존재할 수도 있다. 더욱이, HOA 계수들의 단일 시간 프레임을 코딩하기 위해 양자의 방법들을 동시에 사용하는 것이 또한 가능하다.

[0086] 예시의 목적을 위해, HOA 계수들 (11) 이 라이브로 캡처되었거나 또는 그렇지 않으면 라이브 레코딩 (7) 과 같은 라이브 레코딩들을 나타냄을 오디오 인코딩 디바이스 (20) 가 결정한다고 가정하면, 오디오 인코딩 디바이스 (20) 는 선형 가역 변환 (LIT) 의 애플리케이션을 수반하는 벡터 기반 합성 방법을 사용하여 HOA 계수들 (11) 을 인코딩하도록 구성될 수도 있다. 선형 가역 변환의 일 예는 "특이값 분해" (또는 "SVD") 로서 지칭된다.

이 예에 있어서, 오디오 인코딩 디바이스 (20) 는 SVD 를 HOA 계수들 (11) 에 적용하여 HOA 계수들 (11) 의 분해된 버전을 결정할 수도 있다. 그 후, 오디오 인코딩 디바이스 (20) 는 HOA 계수들 (11) 의 분해된 버전을 분석하여 다양한 파라미터들을 식별할 수도 있으며, 이 파라미터들은 HOA 계수들 (11) 의 분해된 버전의 리오더링을 용이하게 할 수도 있다. 그 후, 오디오 인코딩 디바이스 (20) 는 식별된 파라미터들에 기초하여 HOA 계수들 (11) 의 분해된 버전을 리오더링할 수도 있으며, 여기서, 하기에서 더 상세하게 설명되는 바와 같은 그러한 리오더링은, 변환이 HOA 계수들 (11) 의 프레임들에 걸쳐 HOA 계수들을 리오더링할 수도 있다면 코딩 효율을 개선시킬 수도 있다 (여기서, 프레임은 일반적으로 HOA 계수들 (11) 의 M개 샘플들을 포함하고 M 은 일부 예들에 있어서 1024 로 설정된다). HOA 계수들 (11) 의 분해된 버전을 리오더링한 이후, 오디오 인코딩 디바이스 (20) 는 HOA 계수들 (11) 의 분해된 버전 중, 음장의 전경 (즉, 구별되는, 우월한 또는 돌출하는) 성분들을 나타내는 HOA 계수들을 선택할 수도 있다. 오디오 인코딩 디바이스 (20) 는 전경 성분들을 나타내는 HOA 계수들 (11) 의 분해된 버전을 오디오 오브젝트 및 관련 방향 정보로서 명시할 수도 있다.

[0087] 오디오 인코딩 디바이스 (20) 는 또한, HOA 계수들 (11) 의 분해된 버전 중, 음장의 하나 이상의 배경 (즉, 주변) 성분들을 나타내는 HOA 계수들을 적어도 부분적으로 식별하기 위해 HOA 계수들 (11) 에 관한 음장 분석을 수행할 수도 있다. 오디오 인코딩 디바이스 (20) 는, 일부 예들에 있어서, 배경 성분들이 오직 (예를 들어, 제 2 또는 상위 오더 구면 기저 함수들에 대응하는 HOA 계수들이 아닌 제로 및 제 1 오더 구면 기저 함수들에 대응하는 HOA 계수들과 같은) HOA 계수들 (11) 의 임의의 소정의 샘플의 서브세트만을 포함할 수도 있다고 주어지면 배경 성분들에 관한 에너지 보상을 수행할 수도 있다. 즉, 오더 감소가 수행될 경우, 오디오 인코딩 디바이스 (20) 는, 오더 감소를 수행하는 것으로부터 기인하는 전체 에너지에서의 변화를 보상하기 위해 HOA 계수들 (11) 의 나머지 배경 HOA 계수들을 증강 (예를 들어, 에너지를 나머지 배경 HOA 계수들에 가산/에너지를 나머지 배경 HOA 계수들로부터 감산) 할 수도 있다.

[0088] 오디오 인코딩 디바이스 (20) 는, 다음으로, 전경 오디오 오브젝트들의 각각 및 배경 성분들을 나타내는 HOA 계수들 (11) 각각에 관한 (MPEG 서라운드, MPEG-AAC, MPEG-USAC, 또는 다른 공지된 형태들의 심리 음향 인코딩과 같은) 일 형태의 심리 음향 인코딩을 수행할 수도 있다. 오디오 인코딩 디바이스 (20) 는 전경 방향 정보에 관한 일 형태의 보간을 수행하고, 그 후, 보간된 전경 방향 정보에 관한 오더 감소를 수행하여 오더 감소된 전경 방향 정보를 생성할 수도 있다. 오디오 인코딩 디바이스 (20) 는 추가로, 일부 예들에 있어서, 오더 감소된 전경 방향 정보에 관한 양자화를 수행하여, 코딩된 전경 방향 정보를 출력할 수도 있다. 일부 예들에 있어서, 이러한 양자화는 스칼라/엔트로피 양자화를 포함할 수도 있다. 그 후, 오디오 인코딩 디바이스 (20) 는 인코딩된 배경 성분들, 인코딩된 전경 오디오 오브젝트들, 및 양자화된 방향 정보를 포함하도록 비트스트림 (21) 을 형성할 수도 있다. 그 후, 오디오 인코딩 디바이스 (20) 는 비트스트림 (21) 을 컨텐츠 소비자 (14) 에 송신하거나 그렇지 않으면 출력할 수도 있다.

[0089] 컨텐츠 소비자 (14) 에 직접 송신되는 것으로서 도 3 에 도시되지만, 컨텐츠 창작자 (12) 는 컨텐츠 창작자

(12) 와 컨텐츠 소비자 (14) 사이에 위치된 중간 디바이스에 비트스트림 (21) 을 출력할 수도 있다. 이러한 중간 디바이스는, 이러한 비트스트림을 요청할 수도 있는 컨텐츠 소비자 (14) 로의 추후 전달을 위해 비트스트림 (21) 을 저장할 수도 있다. 중간 디바이스는 파일 서버, 웹 서버, 데스크탑 컴퓨터, 랩톱 컴퓨터, 태블릿 컴퓨터, 모바일 전화기, 스마트 폰, 또는 오디오 디코더에 의한 추후 취출을 위해 비트스트림 (21) 을 저장 가능한 임의의 다른 디바이스를 포함할 수도 있다. 이러한 중간 디바이스는 비트스트림 (21) 을, 비트스트림 (21) 을 요청하는 컨텐츠 소비자 (14) 와 같은 가입자들에게 스트리밍하는 것 (및 가능하게는, 대응하는 비디오 데이터 비트스트림을 송신하는 것과 함께) 이 가능한 컨텐츠 전달 네트워크에 상주할 수도 있다.

[0090] 대안적으로, 컨텐츠 창작자 (12) 는 컴팩트 디스크, 디지털 비디오 디스크, 고선명 비디오 디스크 또는 다른 저장 매체와 같은 저장 매체에 비트스트림 (21) 을 저장할 수도 있으며, 이들 중 대부분은 컴퓨터에 의해 판독가능하고, 따라서, 컴퓨터 판독가능 저장 매체 또는 비-일시적인 컴퓨터 판독가능 저장 매체로서 지칭될 수도 있다. 이러한 컨텍스트에서, 송신 채널은 이들 매체들에 저장된 컨텐츠가 송신되는 그러한 채널들을 지칭할 수도 있다 (그리고, 소매 스토어들 및 다른 스토어 기반 전달 메커니즘을 포함할 수도 있음). 따라서, 어떤 경우든, 본 개시의 기술들은 이러한 점에서 도 3 의 예로 한정되지 않아야 한다.

[0091] 도 3 의 예에서 추가로 도시된 바와 같이, 컨텐츠 소비자 (14) 는 오디오 재생 시스템 (16) 을 포함한다. 오디오 재생 시스템 (16) 은 다중 채널 오디오 데이터를 재생 가능한 임의의 오디오 재생 시스템을 나타낼 수도 있다. 오디오 재생 시스템 (16) 은 다수의 상이한 렌더러들 (22) 을 포함할 수도 있다. 렌더러들 (22) 은 각각 상이한 형태의 렌더링을 제공할 수도 있으며, 여기서, 상이한 형태들의 렌더링은 벡터 기반 진폭 플래닝 (VBAP) 을 수행하는 다양한 방식들 중 하나 이상, 및/또는 음장 합성을 수행하는 다양한 방식들 중 하나 이상을 포함할 수도 있다. 본 명세서에서 사용된 바와 같이, "A 및/또는 B"는 "A 또는 B" 또는 "A 및 B" 양자를 의미한다.

[0092] 오디오 재생 시스템 (16) 은 오디오 디코딩 디바이스 (24) 를 더 포함할 수도 있다. 오디오 디코딩 디바이스 (24) 는 비트스트림 (21) 으로부터의 HOA 계수들 (11') 을 디코딩하도록 구성된 디바이스를 나타낼 수도 있으며, 여기서, HOA 계수들 (11') 은 HOA 계수들 (11) 과 유사하지만 손실성 동작들 (예를 들어, 양자화) 및/또는 송신 채널을 통한 송신으로 인해 상이할 수도 있다. 즉, 오디오 디코딩 디바이스 (24) 는 비트스트림 (21) 에 명시된 전경 방향 정보를 역양자화하면서 또한 비트스트림 (21) 에 명시된 전경 오디오 오브젝트들 및 배경 성분들을 나타내는 인코딩된 HOA 계수들에 관한 심리 음향 디코딩을 수행할 수도 있다. 오디오 디코딩 디바이스 (24) 는 추가로, 디코딩된 전경 방향 정보에 관한 보간을 수행하고, 그 후, 디코딩된 전경 오디오 오브젝트들 및 보간된 전경 방향 정보에 기초하여 전경 성분들을 나타내는 HOA 계수들을 결정할 수도 있다. 그 후, 오디오 디코딩 디바이스 (24) 는 전경 성분들을 나타내는 결정된 HOA 계수들 및 배경 성분들을 나타내는 디코딩된 HOA 계수들에 기초하여 HOA 계수들 (11') 을 결정할 수도 있다.

[0093] 오디오 재생 시스템 (16) 은 비트스트림 (21) 을 디코딩한 이후 HOA 계수들 (11') 을 획득하고 HOA 계수들 (11') 을 렌더링하여 라우드스피커 피드들 (25) 을 출력할 수도 있다. 라우드스피커 피드들 (25) 은 (예시 목적들의 용이를 위해 도 3 의 예에 도시되지 않은) 하나 이상의 라우드스피커들을 구동할 수도 있다.

[0094] 적절한 렌더러를 선택하기 위해 또는 일부 경우들에 있어서 적절한 렌더러를 생성하기 위해, 오디오 재생 시스템 (16) 은 라우드스피커들의 개수 및/또는 라우드스피커들의 공간 지오메트리를 나타내는 라우드스피커 정보 (13) 를 획득할 수도 있다. 일부 경우들에 있어서, 오디오 재생 시스템 (16) 은 라우드스피커 정보 (13) 를 동적으로 결정하도록 하는 그러한 방식으로 라우드스피커들을 구동하고 그리고 참조 마이크로폰을 이용하여 라우드스피커 정보 (13) 를 획득할 수도 있다. 다른 경우들에 있어서 또는 라우드스피커 정보 (13) 의 동적 결정과 함께, 오디오 재생 시스템 (16) 은 오디오 재생 시스템 (16) 과 상호작용하고 그리고 라우드스피커 정보 (16) 를 입력하게 하도록 사용자를 프롬프트할 수도 있다.

[0095] 그 후, 오디오 재생 시스템 (16) 은 라우드스피커 정보 (13) 에 기초하여 오디오 렌더러들 (22) 중 하나를 선택할 수도 있다. 일부 경우들에 있어서, 오디오 재생 시스템 (16) 은, 어떠한 오디오 렌더러들 (22) 도 라우드스피커 정보 (13) 에 명시된 일부 임계 유사도 측정치 (라우드스피커 지오메트리별) 내에 있지 않을 경우, 오디오 재생 시스템 (16) 은 라우드스피커 정보 (13) 에 기초하여 오디오 렌더러들 (22) 중 하나를 생성할 수도 있다. 오디오 재생 시스템 (16) 은, 일부 경우들에 있어서, 오디오 렌더러들 (22) 중 기존의 하나를 선택하려는 처음 시도없이 라우드스피커 정보 (13) 에 기초하여 오디오 렌더러들 (22) 중 하나를 생성할 수도 있다.

[0096] 도 4 는, 본 개시물에서 설명된 기술들의 다양한 양태들을 수행할 수도 있는 도 3 의 예에 도시된 오디오 인코딩 디바이스 (20) 의 일 예를 더 상세하게 예시하는 블록 다이어그램이다. 오디오 인코딩 디바이스 (20) 는

컨텐츠 분석 유닛 (26), 벡터 기반 합성 방법 유닛 (27), 및 방향 기반 합성 방법 유닛 (28) 을 포함한다.

[0097] 컨텐츠 분석 유닛 (26) 은, HOA 계수들 (11) 이 라이브 레코딩 또는 오디오 오브젝트로부터 생성된 컨텐츠를 나타내는지 여부를 식별하기 위해 HOA 계수들 (11) 의 컨텐츠를 분석하도록 구성된 유닛을 나타낸다. 컨텐츠 분석 유닛 (26) 은 HOA 계수들 (11) 이 실제 음장의 레코딩으로부터 생성되었는지 또는 인공적인 오디오 오브젝트로부터 생성되었는지 여부를 결정할 수도 있다. 컨텐츠 분석 유닛 (26) 은 다양한 방식들로 이러한 결정을 행할 수도 있다. 예를 들어, 컨텐츠 분석 유닛 (26) 은 $(N+1)^2$-1개 채널들을 코딩하고 (벡터로서 표현될 수도 있는) 최종 나머지 채널을 예측할 수도 있다. 컨텐츠 분석 유닛 (26) 은 $(N+1)^2$-1개 채널들 중 적어도 일부에 스칼라들을 적용하고 결과적인 값들을 가산하여 최종 나머지 채널을 결정할 수도 있다. 더욱이, 이 예에 있어서, 컨텐츠 분석 유닛 (26) 은 예측된 채널의 정확도를 결정할 수도 있다. 이 예에 있어서, 예측된 채널의 정확도가 상대적으로 높으면 (예를 들어, 정확도가 특정 임계치를 초과하면), HOA 계수들 (11) 은 합성 오디오 오브젝트로부터 생성될 가능성이 있다. 대조적으로, 예측된 채널의 정확도가 상대적으로 낮으면 (예를 들어, 정확도가 특정 임계치 이하면), HOA 계수들 (11) 은 레코딩된 음장을 나타낼 가능성이 있다. 예를 들어, 이 예에 있어서, 예측된 채널의 신호대 노이즈 비 (SNR) 이 100 데시벨 (dbs) 초과이면, HOA 계수들 (11) 은 합성 오디오 오브젝트로부터 생성된 음장을 나타낼 더 많은 가능성이 있다. 대조적으로, 아이겐 마이크로폰을 이용하여 레코딩된 음장의 SNR 은 5 내지 20 dbs 일 수도 있다. 따라서, 실제 직접 레코딩으로부터 그리고 합성 오디오 오브젝트로부터 생성된 HOA 계수들 (11) 에 표현된 음장 간의 SNR 비에서의 명백한 구분이 존재할 수도 있다.

[0098] 더 상세하게, 컨텐츠 분석 유닛 (26) 은, 음장을 나타내는 HOA 계수들 (11) 이 합성 오디오 오브젝트로부터 생성되는지 여부를 결정할 경우, HOA 계수들 중 프레임화된 HOA 계수를 획득할 수도 있으며, 이 프레임화된 HOA 계수는 제 4 오더 표현 (즉, N = 4) 에 대해 25 × 1024 의 사이즈일 수도 있다. 프레임화된 HOA 계수들을 획득한 이후 (이는 또한 프레임화된 SHC 행렬 (11) 로서 본 명세서에서 표기될 수도 있으며, 후속적인 프레임화된 SHC 행렬들은 프레임화된 SHC 행렬들 (27B, 27C 등) 으로서 표기될 수도 있음). 그 후, 컨텐츠 분석 유닛 (26) 은 프레임화된 HOA 계수들 (11) 의 제 1 벡터를 배제하여 감소된 프레임화된 HOA 계수들을 생성할 수도 있다. 일부 예들에 있어서, 프레임화된 HOA 계수들 (11) 로부터 배제된 이러한 제 1 벡터는 제로-오더, 제로-서브오더 구면 조화 기저 함수와 연관된 HOA 계수들 (11) 의 계수들에 대응할 수도 있다.

[0099] 그 후, 컨텐츠 분석 유닛 (26) 은 감소된 프레임화된 HOA 계수들의 제 1 비-제로 벡터를 감소된 프레임화된 HOA 계수들의 나머지 벡터들로부터 예측할 수도 있다. 제 1 비-제로 벡터는 제 1 오더 (및 오더 의존형 서브오더들 각각을 고려) 로부터, 제로 이외의 값들을 갖는 제 4 오더 (및 오더 의존형 서브오더들 각각을 고려) 로 가는 제 1 벡터를 지칭할 수도 있다. 일부 예들에 있어서, 감소된 프레임화된 HOA 계수들의 제 1 비-제로 벡터는 제 1 오더, 제로-서브오더 구면 조화 기저 함수와 연관된 HOA 계수들 (11) 의 계수들을 지칭한다. 제 1 비-제로 벡터에 관하여 설명되지만, 그 기술들은 감소된 프레임화된 HOA 계수들의 다른 벡터들을 감소된 프레임화된 HOA 계수들의 나머지 벡터들로부터 예측할 수도 있다. 예를 들어, 컨텐츠 분석 유닛 (26) 은 감소된 프레임화된 HOA 계수들 중, 제 1 오더, 제 1 서브오더 구면 조화 기저 함수 또는 제 1 오더 네거티브 제 1 오더 구면 조화 기저 함수와 연관된 HOA 계수들을 예측할 수도 있다. 또 다른 예들로서, 컨텐츠 분석 유닛 (26) 은 감소된 프레임화된 HOA 계수들 중, 제 2 오더, 제로-오더 구면 조화 기저 함수와 연관된 HOA 계수들을 예측할 수도 있다.

[0100] 제 1 비-제로 벡터를 예측하기 위해, 컨텐츠 분석 유닛 (26) 은 다음의 수학식에 따라 동작할 수도 있으며:

[0101] $$\sum_i (\alpha_i v_i)$$,

[0102] 여기서, i 는 1 내지 $(N+1)^2$-2 (제 4 오더 표현에 대해 23 임) 이고, α_i 는 i 번째 벡터에 대한 어떤 상수를 나타내고, v_i 는 i 번째 벡터를 지칭한다. 제 1 비-제로 벡터를 예측한 이후, 컨텐츠 분석 유닛 (26) 은 예측된 제 1 비-제로 벡터 및 실제 비-제로 벡터에 기초하여 에러를 획득할 수도 있다. 일부 예들에 있어서, 컨텐츠 분석 유닛 (26) 은 예측된 제 1 비-제로 벡터를 실제 제 1 비-제로 벡터로부터 감산하여 에러를 도출한다. 컨텐츠 분석 유닛 (26) 은 예측된 제 1 비-제로 벡터 및 실제 제 1 비-제로 벡터에서의 각각의 엔트리 간의 차이들의 절대값의 합을 컴퓨팅할 수도 있다.

[0103] 일단 에러가 획득되면, 컨텐츠 분석 유닛 (26) 은 실제 제 1 비-제로 벡터의 에너지 및 에러에 기초하여 비율을

컴퓨팅할 수도 있다. 컨텐츠 분석 유닛 (26) 은 제 1 비-제로 벡터의 각각의 엔트리를 제곱하고 제곱된 엔트리들을 서로 가산함으로써 이 에너지를 결정할 수도 있다. 그 후, 컨텐츠 분석 유닛 (26) 은 이 비율을 임계치와 비교할 수도 있다. 그 비율이 임계치를 초과하지 않을 경우, 컨텐츠 분석 유닛 (26) 은 프레임화된 HOA 계수들 (11) 이 레코딩으로부터 생성됨을 결정하고, HOA 계수들 (11) 의 대응하는 코딩된 표현이 레코딩으로부터 생성되었음을 비트스트림에 표시할 수도 있다. 그 비율이 임계치를 초과할 경우, 컨텐츠 분석 유닛 (26) 은 프레임화된 HOA 계수들 (11) 이 합성 오디오 오브젝트로부터 생성됨을 결정하고, 프레임화된 HOA 계수들 (11) 의 대응하는 코딩된 표현이 합성 오디오 오브젝트로부터 생성되었음을 비트스트림에 표시할 수도 있다.

[0104] 프레임화된 HOA 계수들 (11) 이 레코딩으로부터 생성되었는지 또는 합성 오디오 오브젝트로부터 생성되었는지 여부의 표시는 각각의 프레임에 대한 단일 비트를 포함할 수도 있다. 그 단일 비트는, 상이한 인코딩들이 대응하는 프레임을 인코딩하기 위한 상이한 방식들 사이에서 효과적으로 토글링하는 각각의 프레임에 대해 사용되었음을 표시할 수도 있다. 일부 경우들에 있어서, 프레임화된 HOA 계수들 (11) 이 레코딩으로부터 생성되었을 경우, 컨텐츠 분석 유닛 (26) 은 HOA 계수들 (11) 을 벡터 기반 합성 유닛 (27) 으로 전달한다. 일부 경우들에 있어서, 프레임화된 HOA 계수들 (11) 이 합성 오디오 오브젝트로부터 생성되었을 경우, 컨텐츠 분석 유닛 (26) 은 HOA 계수들 (11) 을 방향 기반 합성 유닛 (28) 으로 전달한다. 방향 기반 합성 유닛 (28) 은 방향 기반 비트스트림 (21) 을 생성하기 위해 HOA 계수들 (11) 의 방향 기반 합성을 수행하도록 구성된 유닛을 나타낼 수도 있다.

[0105] 즉, 그 기술들은 프론트-엔드 분류자를 사용하여 HOA 계수들을 코딩하는 것에 기초한다. 그 분류자는 다음과 같이 작동할 수도 있다:

[0106] 프레임화된 SH 행렬로 시작함 (즉, 제 4 오더, 1024 의 프레임 사이즈, 이는 또한 프레임화된 HOA 계수들로서 또는 HOA 계수들로서 지칭할 수도 있음) - 여기서, 사이즈 25 × 1024 의 행렬이 획득됨.

[0107] 제 1 벡터 (제로 오더 SH) 를 배제함 - 그래서 사이즈 24 × 1024 의 행렬이 존재함.

[0108] 행렬의 벡터들 중 나머지 (사이즈 1 × 1024 의 23개 벡터들) 로부터 행렬에서 제 1 비-제로 벡터 (1 × 1024 사이즈 벡터) 를 예측함.

[0109] 예측은 다음과 같음: 예측된 벡터 = i 에 걸친 합 [알파-i × 벡터-I] (여기서, I 에 걸친 합은 23개 인덱스들에 걸쳐 수행됨, i=1…23).

[0110] 그 후, 에러를 체크함: 실제 벡터 - 예측된 벡터 = 에러.

[0111] 벡터의 에너지/에러의 비율이 크면 (즉, 에러가 작으면), (그 프레임에서의) 기본 음장은 희박/합성적임. 그렇지 않으면, 기본 음장은 (소위 마이크 어레이를 사용한) 레코딩된 음장임.

[0112] 레코딩 대 합성 판정에 의존하여, 상이한 방식들로 인코딩/디코딩을 수행함 (이는 대역폭 압축으로 지칭할 수도 있음). 그 판정은 1비트 판정임, 이는 각각의 프레임에 대해 비트스트림 상으로 전송됨.

[0113] 도 4 의 예에서 도시된 바와 같이, 벡터 기반 합성 유닛 (27) 은 선형 가역 변환 (LIT) 유닛 (30), 파라미터 계산 유닛 (32), 리오더 유닛 (34), 전경 선택 유닛 (36), 에너지 보상 유닛 (38), 심리 음향 오디오 코더 유닛 (40), 비트스트림 생성 유닛 (42), 음장 분석 유닛 (44), 계수 감소 유닛 (46), 배경 (BG) 선택 유닛 (48), 공간-시간 보간 유닛 (50), 및 양자화 유닛 (52) 을 포함할 수도 있다.

[0114] 선형 가역 변환 (LIT) 유닛 (30) 은 HOA 채널들의 형태로 HOA 계수들 (11) 을 수신하고, 각각의 채널은 구면 기저 함수들의 소정의 오더, 서브오더와 연관된 계수의 블록 또는 프레임을 나타낸다 (이는 HOA[k] 로서 표기될 수도 있으며, 여기서, k 는 샘플들의 현재 프레임 또는 블록을 나타낼 수도 있음). HOA 계수들 (11) 의 행렬은 디멘전들 D: M × $(N+1)^2$ 을 가질 수도 있다.

[0115] 즉, LIT 유닛 (30) 은 특이값 분해로 지칭되는 일 형태의 분석을 수행하도록 구성된 유닛을 나타낼 수도 있다. SVD 에 관하여 설명되지만, 본 개시물에서 설명된 기술들은, 선형적으로 미상관된 에너지 집중형 출력의 세트들을 제공하는 임의의 유사한 변환 또는 분해에 관하여 수행될 수도 있다. 또한, 본 개시물에서의 "세트들" 의 참조는, 구체적으로 반하여 서술되지 않으면 비-제로 세트들을 지칭하도록 일반적으로 의도되고, 소위 "공집합" 을 포함하는 세트들의 고전적인 수학적 정의를 지칭하도록 의도되지 않는다.

[0116] 대안적인 변환은 "PCA" 로서 종종 지칭되는 주요 성분 분석을 포함할 수도 있다. PCA 는, 가능하게 상관된

변수들의 관측들의 세트를 주요 성분들로서 지칭된 선형적으로 미상관된 변수들의 세트로 변환하기 위해 직교 변환을 채용하는 수학적 절차를 지칭한다. 선형적으로 미상관된 변수들은 서로 선형 통계적 관계 (또는 의존성) 을 갖지 않는 변수를 나타낸다. 이들 주요 성분들은 서로 작은 통계적 상관도를 갖는 것으로서 설명될 수도 있다. 어떤 경우든, 소위 주요 성분들의 수는 원래의 변수들의 수보다 작거나 같다. 일부 예들에 있어서, 변환은, 제 1 주요 성분이 최대 가능한 분산을 갖고 (즉, 가능한 많은 데이터의 가변성을 설명함) 그리고 차례로 각각의 다음 성분은 이러한 후속 성분이 이전 성분들에 직교해야 하는 (이전 성분들과 미상관됨으로써 재서술될 수도 있음) 제약 하에서 가능한 최고 분산을 갖는 그러한 방식으로 정의된다. PCA 는 일 형태의 오더 감소를 수행할 수도 있으며, 이는 HOA 계수들 (11) 이 관점에서 HOA 계수들 (11) 의 압축을 발생시킬 수도 있다. 컨텍스트에 의존하여, PCA 는, 몇몇 예들만 들자면, 이산 카루넨-루베 변환, 호텔링 변환, 적절한 직교 분해 (POD), 및 고유값 분해 (EVD) 와 같은 다수의 상이한 이름들에 의해 지칭될 수도 있다. 오디오 데이터를 압축하는 기본 목적에 도움이 되는 그러한 동작들의 특성들은 다중 채널 오디오 데이터의 '에너지 집중' 및 '탈상관' 이다.

[0117] 어떤 경우든, LIT 유닛 (30) 은 특이값 분해 (이는, 다시, "SVD" 로서 지칭될 수도 있음) 를 수행하여, HOA 계수들 (11) 을 변환된 HOA 계수들의 2 이상의 세트들로 변환한다. 변환된 HOA 계수들의 이들 "세트들" 은 변환된 HOA 계수들의 벡터들을 포함할 수도 있다. 도 4 의 예에 있어서, LIT 유닛 (30) 은 HOA 계수들 (11) 에 관하여 SVD 를 수행하여 소위 V 행렬, S 행렬, 및 U 행렬을 생성할 수도 있다. 선형 대수에서의 SVD 는 y-바이-z 실수 또는 복소 행렬 (X) (여기서, X 는 HOA 계수들 (11) 과 같은 다중 채널 오디오 데이터를 나타냄) 의 인수분해를 다음의 형태로 나타낼 수도 있다:

[0118] X = USV*

[0119] U 는 y-바이-y 실수 또는 복소 유니터리 행렬을 나타낼 수도 있으며, 여기서, U 의 y 컬럼들은 일반적으로, 다중 채널 오디오 데이터의 좌-특이 벡터들로서 공지된다. S 는 대각선으로 비-네거티브 실수를 갖는 y-바이-z 직사각형 대각 행렬을 나타낼 수도 있으며, 여기서, S 의 대각선 값들은 일반적으로, 다중 채널 오디오 데이터의 특이값들로서 공지된다. (V 의 컨주게이트 트랜스포즈를 나타낼 수도 있는) V* 는 z-바이-z 실수 또는 복소 유니터리 행렬을 나타낼 수도 있으며, 여기서, V* 의 z 컬럼들은 일반적으로, 다중 채널 오디오 데이터의 우-특이 벡터들로서 공지된다.

[0120] HOA 계수들 (11) 을 포함하는 다중 채널 오디오 데이터에 적용되는 것으로서 본 개시물에서 설명되지만, 그 기술들은 임의의 형태의 다중 채널 오디오 데이터에 적용될 수도 있다. 이러한 방식으로, 오디오 인코딩 디바이스 (20) 는 음장의 적어도 일부분을 나타내는 다중 채널 오디오 데이터에 관하여 특이값 분해를 수행하여, 다중 채널 오디오 데이터의 좌-특이 벡터들을 나타내는 U 행렬, 다중 채널 오디오 데이터의 특이값들을 나타내는 S 행렬, 및 다중 채널 오디오 데이터의 우-특이 벡터들을 나타내는 V 행렬을 생성하고, U 행렬, S 행렬, 및 V 행렬 중 하나 이상의 적어도 일부분의 함수로서 다중 채널 오디오 데이터를 나타낼 수도 있다.

[0121] 일부 예들에 있어서, 상기에서 참조된 SVD 수학식에서의 V* 행렬은, SVD 가 복소수들을 포함하는 행렬들에 적용될 수도 있음을 반영하기 위해 V 행렬의 컨주게이트 트랜스포즈로서 표기된다. 오직 실수들만을 포함하는 행렬들에 적용될 경우, V 행렬의 복소 컨주게이트 (즉, V* 행렬) 는 V 행렬의 트랜스포즈로 고려될 수도 있다. 이하, 예시 목적들의 용이를 위해, HOA 계수들 (11) 은, V* 행렬보다는 V 행렬이 SVD 를 통해 출력되는 결과로 실수들을 포함함이 가정된다. 더욱이, 본 개시물에서 V 행렬로서 표기되지만, V 행렬의 참조는 적절할 경우 V 행렬의 트랜스포즈로 지칭하도록 이해되어야 한다. V 행렬로 가정되지만, 그 기술들은 복소 계수들을 갖는 HOA 계수들 (11) 과 유사한 방식으로 적용될 수도 있으며, 여기서, SVD 의 출력은 V* 행렬이다. 이에 따라, 그 기술들은 이러한 점에서 V 행렬을 생성하기 위해 오직 SVD 의 적용만을 제공하도록 제한되어야 하지만, V* 행렬을 생성하기 위해 복소 성분들을 갖는 HOA 계수들 (11) 로의 SVD 의 적용을 포함할 수도 있다.

[0122] 어떤 경우든, LIT 유닛 (30) 은 고차 앰비소닉스 (HOA) 오디오 데이터 (여기서, 이 앰비소닉스 오디오 데이터는 HOA 계수들 (11) 의 블록들 또는 샘플들 또는 임의의 다른 형태의 다중 채널 오디오 데이터를 포함함) 의 각각의 블록 (이는 프레임으로 지칭될 수도 있음) 에 관하여 블록별 형태의 SVD 를 수행할 수도 있다. 상기 노트된 바와 같이, 변수 M 은 샘플들에서의 오디오 프레임의 길이를 나타내는데 사용될 수도 있다. 예를 들어, 오디오 샘플이 1024개 오디오 샘플들을 포함할 경우, M 은 1024 와 동일하다. 비록 M 에 대한 이러한 전형적인 값에 관하여 설명되더라도, 본 개시의 기술들은 M 에 대한 이러한 전형적인 값으로 한정되지 않아야 한다. 따라서, LIT 유닛 (30) 은 M-바이-$(N+1)^2$ HOA 계수들을 갖는 HOA 계수들 (11) 블록에 관하여 블록별 SVD 를 수행할 수도 있으며, 여기서, N 은 다시 HOA 오디오 데이터의 오더를 나타낸다. LIT 유닛 (30) 은,

이러한 SVD 를 수행하는 것을 통해, V 행렬, S 행렬, 및 U 행렬을 생성할 수도 있으며, 여기서, 행렬들 각각은 상기 설명된 개별 V, S 및 U 행렬들을 나타낼 수도 있다. 이러한 방식으로, 선형 가역 변환 유닛 (30) 은 HOA 계수들 (11) 에 관하여 SVD 를 수행하여, 디멘전들 D: M × $(N+1)^2$ 를 갖는 US[k] 벡터들 (33) (이는 S 벡터들과 U 벡터들의 결합된 버전을 나타낼 수도 있음), 및 디멘전들 D: $(N+1)^2$ × $(N+1)^2$ 를 갖는 V[k] 벡터들 (35) 을 출력할 수도 있다. US[k] 행렬에서의 개별 벡터 엘리먼트들이 또한 $X_{PS}(k)$ 로 명명될 수도 있는 한편, V[k] 행렬에서의 개별 벡터들이 또한 v(k) 로 명명될 수도 있다.

[0123] U, S 및 V 행렬들의 분석은 이들 행렬들이 X 에 의해 상기에 표현된 기본 음장의 공간 및 시간 특성들을 포함 또는 표현함을 드러낼 수도 있다. (길이 M 샘플들의) U 에서의 N 벡터들 각각은, 서로 직교하고 그리고 임의의 공간 특성들 (이는 또한 방향 정보로서 지칭될 수도 있음) 로부터 디커플링되었던 정규화된 분리된 오디오 신호들을 (M개 샘플들에 의해 표현된 기간 주기에 대한) 시간의 함수로서 표현할 수도 있다. 공간 형상 및 포지션 (r, 세타, 파이) 폭을 나타내는 공간 특성들은, V 행렬 (각각 길이 $(N+1)^2$) 에서 개별 i번째 벡터들 ($v^{(i)}(k)$) 에 의해 대신 표현될 수도 있다. U 행렬 및 V 행렬에서의 벡터들 양자는 그 제곱-평균-제곱근 에너지들이 1 과 동일하도록 정규화된다. 따라서, U 에서의 오디오 신호들의 에너지는 S 에서의 대각선 엘리먼트들에 의해 표현된다. 따라서, U 와 S 를 곱하여 (개별 벡터 엘리먼트들 ($X_{PS}(k)$) 를 갖는) US[k] 를 형성하는 것은 진정한 에너지들을 갖는 오디오 신호를 표현한다. (U 에서의) 오디오 시간 신호들, (S 에서의) 그 에너지들, 및 (V 에서의) 그 공간 특성들을 디커플링하기 위한 SVD 분해의 능력은 본 개시물에서 설명된 기술들의 다양한 양태들을 지원할 수도 있다. 추가로, 기본 HOA[k] 계수들 (X) 을 US[k] 와 V[k] 의 벡터 곱셈에 의해 합성하는 이러한 모델은, 본 문헌 전반에 걸쳐 사용되는 용어 "벡터 기반 합성 방법" 을 발생시킨다.

[0124] 비록 HOA 계수들 (11) 에 관하여 직접 수행되는 것으로서 설명되더라도, LIT 유닛 (30) 은 선형 가역 변환을 HOA 계수들 (11) 의 도함수들에 적용할 수도 있다. 예를 들어, LIT 유닛 (30) 은 HOA 계수들 (11) 로부터 도출된 전력 스펙트럼 밀도 행렬에 관하여 SVD 를 적용할 수도 있다. 전력 스펙트럼 밀도 행렬은 PSD 로서 표기될 수도 있으며, 하기에 이어지는 의사 코드에서 나타낸 바와 같이, hoaFrame 의 트랜스포즈의 hoaFrame 과의 행렬 곱셈을 통해 획득될 수도 있다. hoaFrame 표기술은 HOA 계수들 (11) 의 프레임을 지칭한다.

[0125] LIT 유닛 (30) 은, SVD (svd) 를 PSD 에 적용한 이후, $S[k]^2$ 행렬 (S_squared) 및 V[k] 행렬을 획득할 수도 있다. $S[k]^2$ 행렬은 제곱된 S[k] 행렬을 나타낼 수도 있으며, 그 결과, LIT 유닛 (30) 은 $S[k]^2$ 행렬에 제곱근 연산을 적용하여 S[k] 행렬을 획득할 수도 있다. LIT 유닛 (30) 은, 일부 경우들에 있어서, V[k] 행렬에 관하여 양자화를 수행하여 양자화된 V[k] 행렬 (이는 V[k]' 행렬로서 표기될 수도 있음) 을 획득할 수도 있다.

LIT 유닛 (30) 은, 먼저, S[k] 행렬을 양자화된 V[k]' 행렬로 곱하여 SV[k]' 행렬을 획득함으로써 U[k] 행렬을 획득할 수도 있다. 다음으로, LIT 유닛 (30) 은 SV[k]' 행렬의 의사-역 (pinv) 을 획득하고, 그 후, HOA 계수들 (11) 을 SV[k]' 행렬의 의사-역을 곱하여 U[k] 행렬을 획득할 수도 있다. 전술한 바는 다음의 의사-코드에 의해 표현될 수도 있다:

[0126] PSD = hoaFrame' * hoaFrame;

[0127] [V, S_squared] = svd(PSD,' econ');

[0128] S = sqrt(S_squared);

[0129] U = hoaFrame * pinv(S*V');

[0130] 계수들 자체보다는 HOA 계수들의 전력 스펙트럼 밀도 (PSD) 에 관하여 SVD 를 수행함으로써, LIT 유닛 (30) 은 프로세서 사이클들 및 저장 공간 중 하나 이상의 관점에서 SVD 를 수행하는 계산 복잡도를 잠재적으로 감소시키면서, SVD 가 HOA 계수들에 직접 적용되는 것처럼 동일한 소스 오디오 인코딩 효율을 달성할 수도 있다. 즉, 상기 설명된 PSD 타입 SVD 는, SVD 가 F*F 행렬 (F 는 HOA 계수들의 수임) 에 대해 수행되기 때문에 잠재적으로 계산상 덜 요구될 수도 있다. M 을 갖는 M * F 행렬과 비교되는 것은 프레임 길이, 즉, 1024개 이상의 샘플들이다. SVD 의 복잡도는 이제, HOA 계수들 (11) 보다는 PSD 로의 적용을 통해, HOA 계수들 (11) 에 적용될 경우 O(M*L^2) 에 비해 대략 O(L^3) 일 수도 있다 (여기서, O(*) 는 컴퓨터 과학 기술들에서 일반적인 계산 복잡도의 빅-O 표기술을 나타냄).

[0131] 파라미터 계산 유닛 (32) 은 상관 파라미터 (R), 방향 특성 파라미터들 (θ, φ, r) 및 에너지 특성 (e) 과 같

은 다양한 파라미터들을 계산하도록 구성된 유닛을 나타낸다. 현재 프레임에 대한 이들 파라미터들 각각은 R[k], Θ[k], φ[k], r[k] 및 e[k] 로서 표시될 수도 있다. 파라미터 계산 유닛 (32) 은 US[k] 벡터들 (33) 에 관하여 에너지 분석 및/또는 상관 (또는 소위 상호-상관) 을 수행하여 이들 파라미터들을 식별할 수도 있다. 파라미터 계산 유닛 (32) 은 또한 이들 파라미터들을 이전 프레임에 대해 결정할 수도 있으며, 여기서, 이전 프레임 파라미터들은 US[k-1] 벡터 및 V[k-1] 벡터들의 이전 프레임에 기초하여 R[k-1], Θ[k-1], φ[k-1], r[k-1] 및 e[k-1] 로 표기될 수도 있다. 파라미터 계산 유닛 (32) 은 현재 파라미터들 (37) 및 이전 파라미터들 (39) 을 리오더 유닛 (34) 으로 출력할 수도 있다.

[0132] 즉, 파라미터 계산 유닛 (32) 은 제 1 시간에 대응하는 L개의 제 1 US[k] 벡터들 (33) 의 각각 및 제 2 시간에 대응하는 제 2 US[k-1] 벡터들 (33) 의 각각에 관하여 에너지 분석을 수행하여, 제 1 오디오 프레임의 적어도 일부분 (종종 그 전부) 및 제 2 오디오 프레임의 일부분 (종종 그 전부) 에 대한 제곱 평균 제곱근화된 에너지를 컴퓨팅하고 이에 의해 2L개의 에너지들, 즉, 제 1 오디오 프레임의 L개의 제 1 US[k] 벡터들 (33) 의 각각에 대해 하나 및 제 2 오디오 프레임의 제 2 US[k-1] 벡터들 (33) 의 각각에 대해 하나를 생성할 수도 있다.

[0133] 다른 예들에 있어서, 파라미터 계산 유닛 (32) 은 제 1 US[k] 벡터들 (33) 의 각각 및 제 2 US[k-1] 벡터들 (33) 의 각각에 대한 샘플들의 세트의 (전부가 아니면) 일부분 사이에서 상호-상관을 수행할 수도 있다. 상호-상관은 신호 프로세싱 기술들에서 이해되는 바와 같은 상호-상관을 지칭할 수도 있다. 즉, 상호-상관은 2개 파형들 (이 경우에 있어서, M개 샘플들의 이산 세트로서 정의됨) 간의 유사도의, 그들 중 하나에 적용된 시간-지연의 함수로서의 측정치를 지칭할 수도 있다. 일부 예들에 있어서, 상호-상관을 수행하기 위해, 파라미터 계산 유닛 (32) 은 각각의 제 1 US[k] 벡터들 (27) 의 마지막 L개 샘플들을, 순번별로, 제 2 US[k-1] 벡터들 (33) 의 나머지 벡터들 각각의 제 1 의 L개 샘플들과 비교하여 상관 파라미터를 결정한다. 본 명세서에서 사용된 바와 같이, "순번별 (turn-wise)" 동작은 엘리먼트들의 제 1 세트 및 엘리먼트들의 제 2 세트에 관하여 행해진 엘리먼트별 동작을 지칭하며, 여기서, 그 동작은 세트들의 오더링에 따라 엘리먼트의 제 1 및 제 2 세트들 각각으로부터 일 엘리먼트를 "차례로" 인출한다.

[0134] 파라미터 계산 유닛 (32) 은 또한, V[k] 및/또는 V[k-1] 벡터들 (35) 을 분석하여 방향 특성 파라미터들을 결정할 수도 있다. 이들 방향 특성 파라미터들은 대응하는 US[k] 및/또는 US[k-1] 벡터들 (33) 에 의해 표현되는 오디오 오브젝트의 움직임 및 위치의 표시를 제공할 수도 있다. 파라미터 계산 유닛 (32) 은 (US[k] 벡터들 (33) 및/또는 V[k] 벡터들 (35) 에 관하여 결정된) 전술한 현재 파라미터들 (37) 의 임의의 조합 및 (US[k-1] 벡터들 (33) 및/또는 V[k-1] 벡터들 (35) 에 관하여 결정된) 이전 파라미터들 (39) 의 임의의 조합을 리오더 유닛 (34) 에 제공할 수도 있다.

[0135] SVD 분해는, US[k-1] 벡터들 (33) 에서의 p번째 벡터에 의해 표현된 오디오 신호/오브젝트 (이는 US[k-1][p] 벡터로서 (또는 대안적으로, $X_{PS}^{(p)}(k-1)$ 로서) 표기될 수도 있음) 가 US[k] 벡터들 (33) 에서의 p번째 벡터에 의해 표현된 (시간적으로 진행된) 동일한 오디오 신호/오브젝트 (이는 US[k][p] 벡터들 (33) 로서 (또는 대안적으로, $X_{PS}^{(p)}(k)$ 로서) 표기될 수도 있음) 일 것임을 보장하지 않는다. 파라미터 계산 유닛 (32) 에 의해 계산된 파라미터들은, 오디오 오브젝트들을, 시간에 걸친 그 자연적 평가 또는 연속성을 표현하도록 리오더링하기 위해 리오더 유닛 (34) 에 의해 사용될 수도 있다.

[0136] 즉, 그 후, 리오더 유닛 (34) 은 제 1 US[k] 벡터들 (33) 로부터의 파라미터들 (37) 각각을, 순번별로, 제 2 US[k-1] 벡터들 (33) 에 대한 파라미터들 (39) 각각에 대해 비교할 수도 있다. 리오더 유닛 (34) 은 현재 파라미터들 (37) 및 이전 파라미터들 (39) 에 기초하여 US[k] 행렬 (33) 및 V[k] 행렬 (35) 내의 다양한 벡터들을 (일 예로서, 헝가리안 알고리즘을 사용하여) 리오더링하여, 리오더링된 US[k] 행렬 (33') (이는 수학적으로 $\overline{US}[k]$ 로서 표기될 수도 있음) 및 리오더링된 V[k] 행렬 (35') (이는 수학적으로 $\overline{V}[k]$ 로서 표기될 수도 있음) 를 전경 사운드 (또는 우월한 사운드 - PS) 선택 유닛 (36) ("전경 선택 유닛 (36)") 및 에너지 보상 유닛 (38) 에 출력할 수도 있다.

[0137] 즉, 리오더 유닛 (34) 은 리오더링된 US[k] 행렬 (33') 을 생성하기 위해 US[k] 행렬 (33) 내의 벡터들을 리오더링하도록 구성된 유닛을 나타낼 수도 있다. 리오더 유닛 (34) 은, US[k] 벡터들 (33) 의 오더 (여기서, $X_{PS}^{(p)}(k)$ 로서 다시 대안적으로 표기될 수도 있는 US[k] 벡터들 (33) 의 각각의 벡터는 다시, 음장에 존재하는 하나 이상의 구별되는 (즉, 우월한) 모노 오디오 오브젝트를 나타낼 수도 있음) 가 오디오 데이터의 부분들로부

터 변할 수도 있기 때문에, US[k] 행렬 (33) 을 리오더링할 수도 있다. 즉, 오디오 인코딩 디바이스 (12) 가 일부 예들에 있어서 오디오 프레임들로서 일반적으로 지칭되는 오디오 데이터의 이들 부분들 상에서 동작한다고 주어지면, 도출됨에 따라 US[k] 행렬 (33) 에서 표현된 바와 같은 이들 구별되는 모노 오디오 오브젝트들에 대응하는 벡터들의 포지션은, 프레임별의 각각의 오디오 오브젝트의 가변하는 돌출성 및 프레임들로의 SVD 의 적용으로 인해 오디오 프레임별로 별할 수도 있다.

[0138] US[k] 행렬 (33) 내의 벡터들을 오디오 프레임별로 리오더링할 것없이 US[k] 행렬 (33) 내의 벡터들을 심리 음향 오디오 코더 유닛 (40) 으로 직접 전달하는 것은, 모노 오디오 오브젝트들이 오디오 프레임들에 걸쳐 (채널별로, (이는, 이 예에 있어서, 서로에 대해 US[k] 행렬 (33) 내의 벡터들의 위치 오더에 의해 정의됨)) 연속적일 경우에 더 우수하게 수행하는 레거시 압축 방식들과 같은 일부 압축 방식들에 대해 달성가능한 압축의 정도를 감소시킬 수도 있다. 더욱이, 리오더링되지 않을 경우, US[k] 행렬 (33) 내의 벡터들의 인코딩은, 디코딩될 경우 오디오 데이터의 품질을 감소시킬 수도 있다. 예를 들어, 심리 음향 오디오 코더 유닛 (40) 에 의해 도 3 의 예에서 표현될 수도 있는 AAC 인코더들은, US[k] 행렬 (33) 내의 벡터들을 프레임별로 직접 인코딩할 경우에 달성된 압축에 비하여 US[k] 행렬 (33') 내의 리오더링된 하나 이상의 벡터들을 프레임별로 더 효율적으로 압축할 수도 있다. AAC 인코더들에 관하여 상기 설명되지만, 그 기술들은, 모노 오디오 오브젝트들이 특정 오더 또는 포지션 (채널별) 에서의 프레임들에 걸쳐 명시될 경우 더 우수한 압축을 제공하는 임의의 인코더에 관하여 수행될 수도 있다.

[0139] 그 기술들의 다양한 양태들은, 이러한 방식으로, 오디오 인코딩 디바이스 (12) 로 하여금 하나 이상의 벡터들 (예를 들어, 리오더링된 US[k] 행렬 (33') 내의 리오더링된 하나 이상의 벡터들을 생성하고 그에 의해 심리 음향 오디오 코더 유닛 (40) 과 같은 레거시 오디오 인코더에 의해 US[k] 행렬 (33) 내의 벡터들의 압축을 용이하게 하는 US[k] 행렬 (33) 내의 벡터들) 을 리오더링하게 할 수도 있다.

[0140] 예를 들어, 리오더 유닛 (34) 은, US[k-1] 행렬 (33) 내의 하나 이상의 제 2 벡터들이 현재 파라미터들 (37) 및 이전 파라미터들 (39) 에 기초하여 대응하는 제 2 프레임에 시간적으로 후속적인 제 1 오디오 프레임으로부터, US[k] 행렬 (33) 내의 하나 이상의 벡터들을 리오더링할 수도 있다. 제 2 오디오 프레임에 시간적으로 후속적인 제 1 오디오 프레임의 컨텍스트에서 설명되지만, 제 1 오디오 프레임은 제 2 오디오 프레임에 시간적으로 선행할 수도 있다. 이에 따라, 그 기술들은 본 개시물에서 설명된 예로 한정되지 않아야 한다.

[0141] 예시하기 위해, US[k] 행렬 (33) 내의 p개 벡터들 각각이 US[k][p] 로서 표기되는 다음의 표 1 을 고려하며, 여기서, k 는 대응하는 벡터가 k번째 프레임으로부터 비롯되는지 또는 이전 (k-1)번째 프레임으로부터 비롯되는지 여부를 나타내고, p 는 동일한 오디오 프레임의 벡터들에 대한 벡터의 로우를 나타낸다 (여기서, US[k] 행렬은 그러한 $(N+1)^2$ 벡터들을 가짐). 상기 노트된 바와 같이, N 이 1 인 것으로 결정된다고 가정하면, p 는 벡터일 (1) 내지 벡터 (4) 를 나타낼 수도 있다.

[0142] [표 1]

고려 중인 에너지	비교 대상
US[*k*-1][1]	US[*k*][1], US[*k*][2], US[*k*][3], US[*k*][4]
US[*k*-1][2]	US[*k*][1], US[*k*][2], US[*k*][3], US[*k*][4]
US[*k*-1][3]	US[*k*][1], US[*k*][2], US[*k*][3], US[*k*][4]
US[*k*-1][4]	US[*k*][1], US[*k*][2], US[*k*][3], US[*k*][4]

[0143]

[0144] 상기 표 1 에 있어서, 리오더 유닛 (34) 은 US[k-1][1] 에 대해 컴퓨팅된 에너지를 US[k][1], US[k][2], US[k][3], US[k][4] 의 각각에 대해 컴퓨팅된 에너지와 비교하고, US[k-1][2] 에 대해 컴퓨팅된 에너지를 US[k][1], US[k][2], US[k][3], US[k][4] 의 각각에 대해 컴퓨팅된 에너지와 비교하는 등이다. 그 후, 리오더 유닛 (34) 은 제 2 선행 오디오 프레임 (시간별) 의 제 2 US[k-1] 벡터들 (33) 중 하나 이상을 폐기할 수도 있다. 예시하기 위해, 나머지 제 2 US[k-1] 벡터들 (33) 을 나타내는 다음의 표 2 를 고려한다:

[0145] [표 2]

고려 중인 벡터	고려 중인 나머지
US[*k*-1][1]	US[*k*][1], US[*k*][2]
US[*k*-1][2]	US[*k*][1], US[*k*][2]
US[*k*-1][3]	US[*k*][3], US[*k*][4]
US[*k*-1][4]	US[*k*][3], US[*k*][4]

[0146]

[0147] 상기 표 2 에 있어서, 리오더 유닛 (34) 은, US[k-1][1] 에 대해 컴퓨팅된 에너지가 US[k][1] 및 US[k][2] 의 각각에 대해 컴퓨팅된 에너지와 유사하고 US[k-1][2] 에 대해 컴퓨팅된 에너지가 US[k][1] 및 US[k][2] 의 각각에 대해 컴퓨팅된 에너지와 유사하고 US[k-1][3] 에 대해 컴퓨팅된 에너지가 US[k][3] 및 US[k][4] 의 각각에 대해 컴퓨팅된 에너지와 유사하고 US[k-1][4] 에 대해 컴퓨팅된 에너지가 US[k][3] 및 US[k][4] 의 각각에 대해 컴퓨팅된 에너지와 유사하다는 에너지 비교에 기초하여 결정될 수도 있다. 일부 예들에 있어서, 리오더 유닛 (34) 은 US[k] 행렬 (33) 의 제 1 벡터들 각각과 US[k-1] 행렬 (33) 의 제 2 벡터들 각각 사이의 유사도를 식별하기 위해 추가적인 에너지 분석을 수행할 수도 있다.

[0148] 다른 예들에 있어서, 리오더 유닛 (32) 은 상호-상관에 관한 현재 파라미터들 (37) 및 이전 파라미터들 (39) 에 기초하여 벡터들을 리오더링할 수도 있다. 이들 예들에 있어서, 상기 표 2 를 다시 참조하면, 리오더 유닛 (34) 은 이들 상호-상관 파라미터들에 기초하여 표 3 에 표현된 다음의 예시적인 상관을 결정할 수도 있다:

[0149] [표 3]

고려 중인 벡터	상관 대상
US[*k*-1][1]	US[*k*][2]
US[*k*-1][2]	US[*k*][1]
US[*k*-1][3]	US[*k*][3]
US[*k*-1][4]	US[*k*][4]

[0150]

[0151] 상기 표 3 으로부터, 리오더 유닛 (34) 은, 일 예로서, US[k-1][1] 벡터가 상이하게 위치된 US[k][2] 벡터에 상관하고 US[k-1][2] 벡터가 상이하게 위치된 US[k][1] 벡터에 상관하고 US[k-1][3] 벡터가 유사하게 위치된 US[k][3] 벡터에 상관하고 US[k-1][4] 벡터가 유사하게 위치된 US[k][4] 벡터에 상관함을 결정한다. 즉, 리오더 유닛 (34) 은, US[k][2] 벡터가 US[k] 행렬 (33) 의 제 1 벡터들의 제 1 로우에 재위치되고 US[k][1] 벡터가 제 1 US[k] 벡터들 (33) 의 제 2 로우에 재위치되도록 US[k] 행렬 (33) 의 제 1 벡터들을 리오더링하는 방법을 기술하는 리오더 정보로서 지칭될 수도 있는 것을 결정한다. 그 후, 리오더 유닛 (34) 은 이러한 리오더 정보에 기초하여 US[k] 행렬 (33) 의 제 1 벡터들을 리오더링하여 리오더링된 US[k] 행렬 (33') 을 생성할 수도 있다.

[0152] 부가적으로, 리오더 유닛 (34) 은, 비록 도 4 의 예에 도시되지 않더라도, 이러한 리오더 정보를 비트스트림 생성 디바이스 (42) 에 제공할 수도 있으며, 이 비트스트림 생성 디바이스 (42) 는, 도 3 및 도 5 의 예에 도시된 오디오 디코딩 디바이스 (24) 와 같은 오디오 디코딩 디바이스가 US[k] 행렬 (33') 의 리오더링된 벡터들을 리오더링하는 방법을 결정하여 US[k] 행렬 (33) 의 벡터들을 복원할 수도 있도록 이러한 리오더 정보를 포함하기 위해 비트스트림 (21) 을 생성할 수도 있다.

[0153] 먼저 에너지 특정 파라미터들 및 그 후 상호-상관 파라미터들에 기초한 분석을 수반하는 2단계 프로세스를 수행하는 것으로서 상기 설명되었지만, 리오더 유닛 (32) 은 리오더 정보를 결정하기 위해 오직 에너지 파라미터들에 관해서만 오직 이러한 분석만을 수행하거나, 리오더 정보를 결정하기 위해 오직 상호-상관 파라미터들에 관해서만 이러한 분석을 수행하거나, 또는 상기에서 설명된 방식으로 에너지 파라미터들 및 상호-상관 파라미터들 양자에 관하여 분석을 수행할 수도 있다. 부가적으로, 그 기술들은, 에너지 비교 및/또는 상호-상관 중 하나 또는 그 양자를 수행하는 것을 수반하지 않는 상관을 결정하기 위한 다른 타입들의 프로세스들을 채용할 수도 있다. 부가적으로, 그 기술들은 이러한 점에서 상기 설명된 예들로 한정되지 않아야 한다. 더욱이, (V 벡터들로부터 도출된 공간 포지션 파라미터들 또는 V[k] 및 V[k-1] 에서의 벡터들의 상관과 같은) 파라미터

계산 유닛 (32) 으로부터 획득된 다른 파라미터들은 또한 US 에서의 벡터들의 정확한 오더링을 결정하기 위해 US[k] 및 US[k-1] 로부터 획득된 에너지 및 상호-상관 파라미터들과 함께 (동시에/공동으로 또는 순차적으로) 사용될 수 있다.

[0154] V 행렬에서의 벡터들의 상관을 사용하는 일 예로서, 파라미터 계산 유닛 (34) 은, V[k] 행렬 (35) 의 벡터들이 다음의 표 4 에서 명시되는 바와 같이 상관됨을 결정할 수도 있다:

[0155] [표 4]

[0156]

고려 중인 벡터	상관 대상
V[*k*-1][1]	V[*k*][2]
V[*k*-1][2]	V[*k*][1]
V[*k*-1][3]	V[*k*][3]
V[*k*-1][4]	V[*k*][4]

[0157] 상기 표 4 로부터, 리오더 유닛 (1) 은, 일 예로서, V[k-1][1] 벡터가 상이하게 위치된 V[k][2] 벡터에 상관하고 V[k-1][2] 벡터가 상이하게 위치된 V[k][1] 벡터에 상관하고 V[k-1][3] 벡터가 유사하게 위치된 V[k][3] 벡터에 상관하고 V[k-1][4] 벡터가 유사하게 위치된 V[k][4] 벡터에 상관함을 결정한다. 리오더 유닛 (34) 은 V[k] 행렬 (35) 의 벡터들의 리오더링된 버전을 리오더링된 V[k] 행렬 (35') 로서 출력할 수도 있다.

[0158] 일부 예들에 있어서, US 행렬에서의 벡터들에 적용되는 동일한 리오더링이 또한 V 행렬에서의 벡터들에 적용된다. 즉, V 벡터들을 리오더링함에 있어서 사용된 임의의 분석은 US 벡터들을 리오더링하는데 사용된 임의의 분석과 함께 사용될 수도 있다. 리오더 정보가 US[k] 벡터들 (35) 에 대하여 에너지 파라미터들 및/또는 상호-상관 파라미터들에 관하여 결정될 뿐 만은 아닌 예를 예시하기 위해, 리오더 유닛 (34) 은 또한 V[k] 벡터들 (35) 에 관하여 상기 설명된 바와 유사한 방식으로 상호-상관 파라미터들 및 에너지 파라미터들에 기초하여 V[k] 벡터들 (35) 에 관하여 이러한 분석을 수행할 수도 있다. 더욱이, US[k] 벡터들 (33) 이 임의의 방향 특성들을 갖지 않지만, V[k] 벡터들 (35) 은 대응하는 US[k] 벡터들 (33) 의 방향성에 관한 정보를 제공할 수도 있다. 이러한 의미에서, 리오더 유닛 (34) 은 대응하는 방향 특성 파라미터들의 분석에 기초하여 V[k] 벡터들 (35) 과 V[k-1] 벡터들 (35) 간의 상관들을 식별할 수도 있다. 즉, 일부 예들에 있어서, 오디오 오브젝트는 움직이고 있을 때 연속적인 방식으로 음장 내에서 움직이거나 그것은 상대적으로 안정적 위치에 머무른다. 그에 따라, 리오더 유닛 (34) 은, 일부 공지된 물리적으로 현실적인 모션을 나타내거나 또는 이러한 방향 특성 상관에 기초하여 US[k] 벡터들 (33) 및 V[k] 벡터들 (35) 를 리오더링하는 상관될 때의 음장 내에서 정지식으로 머무르는 V[k] 행렬 (35) 및 V[k-1] 행렬 (35) 의 그 벡터들을 식별할 수도 있다. 어떤 경우든, 리오더 유닛 (34) 은 리오더링된 US[k] 벡터들 (33') 및 리오더링된 V[k] 벡터들 (35') 을 전경 선택 유닛 (36) 에 출력할 수도 있다.

[0159] 부가적으로, 그 기술들은, 에너지 비교 및/또는 상호-상관 중 하나 또는 그 양자를 수행하는 것을 수반하지 않는 정확한 오더를 결정하기 위한 다른 타입들의 프로세스들을 채용할 수도 있다. 부가적으로, 그 기술들은 이러한 점에서 상기 설명된 예들로 한정되지 않아야 한다.

[0160] 비록 US 행렬의 벡터들의 리오더링을 미러링하기 위해 V 행렬의 벡터들을 리오더링하는 것으로서 상기 설명되더라도, 특정 경우들에 있어서, V 벡터들은 US 벡터들과는 상이하게 리오더링될 수도 있으며, 여기서, 별도의 신택스 엘리먼트들은 US 벡터들의 리오더링 및 V 벡터들의 리오더링을 나타내기 위해 생성될 수도 있다. 일부 경우들에 있어서, V 벡터들이 심리 음향적으로 인코딩되지 않을 수도 있다고 주어지면, V 벡터들은 리오더링되지 않을 수도 있고 오직 US 벡터들만이 리오더링될 수도 있다.

[0161] V 행렬의 벡터들 및 US 행렬의 벡터들의 리오더링이 상이한 실시형태는 그 의도가 오디오 오브젝트들을 공간에서 교환하려는 것일 때이며, 즉, 원래의 레코딩된 위치 (기본 음장이 자연적 리코딩이었을 경우) 또는 예술적으로 의도된 포지션 (기본 음장이 오브젝트들의 인공적인 혼합일 경우) 으로부터 이격되게 이동할 때이다. 예로서, 2개의 오디오 소스들 (A 및 B) 가 존재하고 A 는 음장의 "좌측"부로부터 나오는 고양이 "야옹" 의 사운드일 수도 있고 B 는 음장의 "우측"부로부터 나오는 개 "으르렁" 의 사운드일 수도 있다고 가정한다. V 및 US 의 리오더링이 상이할 경우, 2개 사운드 소스들의 포지션이 교환된다. 음장의 우측부로부터 나오는 A ("야옹") 를 교환한 이후, B ("으르렁") 는 음장의 좌측부로부터 나온다.

[0162] 음장 분석 유닛 (44) 은 타겟 비트레이트 (41) 를 잠재적으로 달성하도록 HOA 계수들 (11) 에 관하여 음장 분석을 수행하도록 구성된 유닛을 나타낼 수도 있다. 음장 분석 유닛 (44) 은, 이러한 분석에 및/또는 수신된 타겟 비트레이트 (41) 에 기초하여, 심리 음향 코더 인스턴스화의 총 수 (이는 주변 또는 배경 채널들 (BG_{TOT}) 의 총 수의 함수일 수도 있음) 및 전경 채널들 즉 우월한 채널들의 수를 결정할 수도 있다. 심리 음향 코더 인스턴스화들의 총 수는 numHOATransportChannels 로서 표기될 수 있다. 음장 분석 유닛 (44) 은 또한, 다시 타겟 비트레이트 (41) 을 잠재적으로 달성하기 위해, 전경 채널들의 총 수 (nFG) (45), 배경 (즉, 주변) 음장의 최소 오더 (N_{BG} 또는 대안적으로, MinAmbHoaOrder), 배경 음장의 최소 오더를 나타내는 실제 채널들의 대응하는 수 (nBGa = $(MinAmbHoaOrder+1)^2$), 및 전송할 추가적인 BG HOA 채널들의 인덱스들 (i) (이는 도 4 의 예에 있어서 배경 채널 정보 (43) 로서 일괄적으로 표기될 수도 있음) 을 결정할 수도 있다. 배경 채널 정보 (42) 는 또한 주변 채널 정보 (43) 으로 지칭될 수도 있다. numHOATransportChannels - nBGa 로부터 남겨지는 채널들 각각은 "부가적인 배경/주변 채널", "활성 벡터 기반 우월한 채널", "활성 방향 기반 우월한 신호" 또는 "완전히 비활성" 일 수도 있다. 일 실시형태에 있어서, 이들 채널 타입들은 2개 비트들에 의해 신택스 엘리먼트를 ("ChannelType" 으로서) 표시받을 수도 있다 (예를 들어, 00: 부가적인 배경 채널; 01: 벡터 기반 우월한 신호; 10: 비활성 신호; 11: 방향 기반 신호). 배경 또는 주변 신호들의 총 수 (nBGa) 는 $(MinAmbHoaOrder +1)^2$ + (상기 예에 있어서) 인덱스 00 이 그 프레임에 대한 비트스트림에서 채널 타입으로서 나타나는 횟수에 의해 주어질 수도 있다.

[0163] 어떤 경우든, 음장 분석 유닛 (44) 은 타겟 비트레이트 (41) 에 기초하여 배경 (즉, 주변) 채널들의 수 및 전경 (즉, 우월한) 채널들의 수를 선택하여, 타겟 비트레이트 (41) 이 상대적으로 더 높을 경우 (예를 들어, 타겟 비트레이트 (41) 가 512 Kbps 이상일 경우) 더 많은 배경 및/또는 전경 채널들을 선택할 수도 있다. 일 실시형태에 있어서, 비트스트림의 헤더 선택에 있어서 MinAmbHoaOrder 가 1 로 설정될 수도 있는 동안 numHOATransportChannels 는 8 로 설정될 수도 있다 (이는 도 10 내지 도 10o(ii) 에 대하여 더 상세하게 설명됨). 이 시나리오에 있어서, 모든 프레임에서, 4개 채널들이 음장의 배경 또는 주변부를 나타내는데 전용될 수도 있는 한편, 다른 4개 채널들은 채널의 타입에 대해 프레임별 기반으로 변할 수 있으며, 예를 들어, 부가적인 배경/주변 채널 또는 전경/우월한 채널로서 사용될 수 있다. 전경/우월한 신호들은, 상기 설명된 바와 같이, 벡터 기반 또는 방향 기반 신호들 중 하나일 수 있다.

[0164] 일부 경우들에 있어서, 프레임에 대한 벡터 기반 우월한 신호들의 총 수는, 상기 예에 있어서, 그 프레임의 비트스트림에서 ChannelType 인덱스가 01 인 횟수에 의해 주어질 수도 있다. 상기 실시형태에 있어서, 모든 부가적인 배경/주변 채널 (예를 들어, 00 의 ChannelType 에 대응) 에 대해, (첫번째 4개 초과의) 가능한 HOA 계수들 중 어느 계수의 대응하는 정보가 그 채널에서 표현될 수도 있다. 제 4 오더 HOA 컨텐츠에 대한 이 정보는 5 내지 25 사이를 나타내기 위한 인덱스일 수도 있다 (첫번째 4개 (1-4) 는 minAmbHoaOrder 가 1 로 설정될 시간에 모두 전송될 수도 있으며, 따라서, 오직 5 내지 25 사이 중 하나를 나타낼 필요만이 있음). 따라서, 이 정보는, "CodedAmbCoeffIdx" 로서 표기될 수도 있는 (제 4 오더 컨텐츠에 대한) 5 비트 신택스 엘리먼트를 사용하여 전송될 수 있다.

[0165] 제 2 실시형태에 있어서, 전경/우월한 신호들 모두는 벡터 기반 신호들이다. 이 제 2 실시형태에 있어서, 전경/우월한 신호들의 총 수는 nFG = numHOATransportChannels - [$(MinAmbHoaOrder +1)^2$ + 인덱스 00 의 횟수] 에 의해 주어질 수도 있다.

[0166] 음장 분석 유닛 (44) 은 배경 채널 정보 (43) 및 HOA 계수들 (11) 을 배경 (BG) 선택 유닛 (46) 에 출력하고, 배경 채널 정보 (43) 을 계수 감소 유닛 (46) 및 비트스트림 생성 유닛 (42) 에 출력하며, nFG (45) 를 전경 선택 유닛 (36) 에 출력한다.

[0167] 일부 예들에 있어서, 음장 분석 유닛 (44) 은, US[k] 행렬 (33) 의 벡터들 및 타겟 비트레이트 (41) 의 분석에 기초하여, 최대 값을 갖는 이들 성분들의 변수 nFG 수를 선택할 수도 있다. 즉, 음장 분석 유닛 (44) 은 S[k] 행렬 (33) 의 벡터들의 내림차순의 대각선 값들에 의해 생성된 곡선의 기울기를 분석함으로써 2개 부분공간을 분리하는 변수 A (이는 N_{BG} 와 유사하거나 실질적으로 유사할 수도 있음) 에 대한 값을 결정할 수도 있으며, 여기서, 큰 특이값들은 전경 또는 구별되는 사운드들을 나타내고 낮은 특이값들은 음장의 배경 성분들을 나타낸다. 즉, 변수 A 는 전체 음장을 전경 부분공간 및 배경 부분공간으로 분할할 수도 있다.

[0168] 일부 예들에 있어서, 음장 분석 유닛 (44) 은 특이값 곡선의 제 1 및 제 2 도함수를 이용할 수도 있다. 음

장 분석 유닛 (44) 은 또한, 변수 A 에 대한 값을 1 과 5 사이로 제한할 수도 있다. 다른 예로서, 음장 분석 유닛 (44) 은 변수 A 에 대한 값을 1 과 $(N+1)^2$ 사이로 제한할 수도 있다. 대안적으로, 음장 분석 유닛 (44) 은 변수 A 에 대한 값을, 예컨대 4 의 값으로 미리정의할 수도 있다. 어떤 경우든, A 의 값에 기초하여, 음장 분석 유닛 (44) 은 전경 채널들의 총 수 (nFG) (45), 배경 음장의 오더 (N_{BG}) 및 수 (nBGa), 그리고 전송할 추가적인 BG HOA 채널들의 인덱스들 (i) 을 결정한다.

[0169] 더욱이, 음장 분석 유닛 (44) 은 V[k] 행렬 (35) 에서의 벡터들의 에너지를 벡터당 기반으로 결정할 수도 있다. 음장 분석 유닛 (44) 은 V[k] 행렬 (35) 에서의 벡터들의 각각에 대한 에너지를 결정하고, 고 에너지를 갖는 벡터들을 전경 성분들로서 식별할 수도 있다.

[0170] 더욱이, 음장 분석 유닛 (44) 은 공간 에너지 분석, 공간 마스킹 분석, 확산 분석 또는 다른 형태들의 청각 분석을 포함하여 HOA 계수들 (11) 에 관하여 다양한 다른 분석들을 수행할 수도 있다. 음장 분석 유닛 (44) 은 HOA 계수들 (11) 의 공간 도메인으로의 변환을 통해 공간 에너지 분석을 수행하고, 보존되어야 하는 음장의 방향 성분들을 나타내는 고 에너지의 영역들을 식별할 수도 있다. 음장 분석 유닛 (44) 은, 상위 에너지 사운드들을 공간적으로 근사화함으로써 마스킹되는 공간 영역들을 음장 분석 유닛 (44) 이 식별할 수도 있는 것을 제외하면, 공간 에너지 분석의 것과 유사한 방식으로 지각 공간 마스킹 분석을 수행할 수도 있다. 그 후, 음장 분석 유닛 (44) 은 일부 경우들에 있어서, 지각적으로 마스킹된 영역들에 기초하여, 더 적은 전경 성분들을 식별할 수도 있다. 음장 분석 유닛 (44) 은 추가로, HOA 계수들 (11) 에 관하여 확산 분석을 수행하여, 음장의 배경 성분들을 나타낼 수도 있는 확산 에너지의 영역들을 식별할 수도 있다.

[0171] 음장 분석 유닛 (44) 은 또한, 오디오 데이터와 연관된 방향성 기반 정보를 이용하여, 음장을 나타내는 오디오 데이터의 돌출성, 구별성 또는 우월성을 결정하도록 구성된 유닛을 나타낼 수도 있다. 에너지 기반 결정들이 음장의 구별되는 오디오 성분들을 식별하기 위해 SVD 에 의해 분해된 음장의 렌더링을 개선할 수도 있지만, 배경 오디오 성분들이 고 에너지 레벨을 나타낼 경우, 에너지 기반 결정들은 또한 디바이스로 하여금 배경 오디오 성분들을 구별되는 오디오 성분들로서 잘못 식별하게 할 수도 있다. 즉, 구별되는 및 배경의 오디오 성분들의 유일한 에너지 기반 구분은 강인하지 않을 수도 있는데, 왜냐하면 왕성한 (예를 들어, 소리가 더 큰) 배경 오디오 성분들이 구별되는 오디오 성분들로서 잘못 식별될 수도 있기 때문이다. 음장의 구별되는 및 배경의 오디오 성분들 간을 더 강인하게 구별하기 위해, 본 개시물에서 설명된 기술들의 다양한 양태들은 음장 분석 유닛 (44) 으로 하여금 HOA 계수들 (11) 의 분해된 버전으로부터 전경 및 주변 오디오 성분들을 분리하도록 HOA 계수들 (11) 의 방향성 기반 분석을 수행하게 할 수도 있다.

[0172] 이러한 점에서, 음장 분석 유닛 (44) 은 US[k] 행렬 (33) 에서의 벡터들 및 V[k] 행렬 (35) 에서의 벡터들 중 하나 이상에 포함된 배경 엘리먼트들로부터 구별되는 (또는 전경) 엘리먼트들을 식별하도록 구성되거나 그렇지 않으면 동작가능한 유닛을 나타낼 수도 있다. 일부 SVD 기반 기술들에 따르면, 가장 왕성한 성분들 (예를 들어, US[k] 행렬 (33) 및 V[k] 행렬 (35) 중 하나 이상의 제 1 의 몇몇 벡터들, 또는 그로부터 도출된 벡터들) 이 구별되는 성분들로서 처리될 수도 있다. 하지만, US[k] 행렬 (33) 에서의 벡터들 및 V[k] 행렬 (35) 에서의 벡터들 중 하나 이상의 가장 왕성한 성분들 (이는 벡터들에 의해 표현됨) 은, 모든 시나리오들에 있어서, 가장 방향성인 성분들/신호들을 나타내지 않을 수도 있다.

[0173] 음장 분석 유닛 (44) 은 US[k] 행렬 (33) 에서의 벡터들 및 V[k] 행렬 (35) 에서의 벡터들 중 하나 이상의 벡터들 또는 그로부터 도출된 벡터들의 방향성에 기초하여 전경/직접/우월한 엘리먼트들을 식별하기 위해 본 명세서에서 설명된 기술들의 하나 이상의 양태들을 구현할 수도 있다. 일부 예들에 있어서, 음장 분석 유닛 (44) 은 벡터들의 에너지 및 방향성 양자에 기초하여 하나 이상의 벡터들을, 구별되는 오디오 성분들 (여기서, 성분들은 또한 "오브젝트들" 로서 지칭될 수도 있음) 로서 식별하거나 선택할 수도 있다. 예를 들어, 음장 분석 유닛 (44) 은, 고 에너지 및 고 방향성 (예를 들어, 방향성 지수로서 표현됨) 을 구별되는 오디오 성분들로서 디스플레이하는 US[k] 행렬 (33) 에서의 벡터들 및 V[k] 행렬 (35) 에서의 벡터들 중 하나 이상의 그 벡터들 (또는 그로부터 도출된 벡터들) 을 식별할 수도 있다. 결과적으로, US[k] 행렬 (33) 에서의 벡터들 및 V[k] 행렬 (35) 에서의 벡터들 중 하나 이상의 다른 벡터들 (또는 그로부터 도출된 벡터들) 과 비교할 때 특정 벡터가 상대적으로 덜 방향성이라고 음장 분석 유닛 (44) 이 결정하면, 특정 벡터와 연관된 에너지 레벨과 무관하게, 음장 분석 유닛 (44) 은 특정 벡터가 HOA 계수들 (11) 에 의해 표현된 음장의 배경 (또는 주변) 오디오 성분들을 나타낸다고 결정할 수도 있다.

[0174] 일부 예들에 있어서, 음장 분석 유닛 (44) 은, 다음의 동작들을 수행함으로써, 방향성에 기초하여 구별되는 오디오 오브젝트들 (이는 또한, 상기 노트된 바와 같이, "성분들" 로서 지칭될 수도 있음) 을 식별할 수도 있다.

음장 분석 유닛 (44) 은 S[k] 행렬 (이는 US[k] 벡터들 (33) 로부터 도출될 수도 있거나 또는 비록 도 4 의 예에 도시되지 않지만 LIT 유닛 (30) 에 의해 별도로 출력됨) 에서의 벡터들을 V[k] 행렬 (35) 에서의 벡터들에 의해 (예를 들어, 하나 이상의 행렬 곱셈 프로세스들을 이용하여) 곱할 수도 있다. V[k] 행렬 (35) 과 S[k] 벡터들을 곱함으로써, 음장 분석 유닛 (44) 은 VS[k] 행렬을 획득할 수도 있다. 부가적으로, 음장 분석 유닛 (44) 은 VS[k] 행렬에서의 벡터들 각각의 엔트리들 중 적어도 일부를 제곱할 수도 있다 (즉, 2 의 거듭제곱으로 거듭제곱함). 일부 경우들에 있어서, 음장 분석 유닛 (44) 은, 1 보다 큰 오더와 연관되는 각각의 벡터의 그 제곱된 엔트리들을 합산할 수도 있다.

[0175] 일 예로서, VS[k] 행렬의 각 벡터가 25개 엔트리들을 포함하면, 음장 분석 유닛 (44) 은, 각각의 벡터에 관하여, 제 5 엔트리에서 시작하고 제 25 엔트리에서 종료하는 각각의 벡터의 엔트리들을 제곱하고 제곱된 엔트리들을 합산하여 방향성 지수 (또는 방향성 표시자) 를 결정할 수도 있다. 각각의 합산 연산은 대응하는 벡터에 대한 방향성 지수를 발생시킬 수도 있다. 이 예에 있어서, 음장 분석 유닛 (44) 은, 1 보다 작거나 같은 오더와 연관되는 각각의 로우의 그 엔트리들, 즉, 제 1 내지 제 4 엔트리들이 에너지의 양에 더 일반적으로 지향되고 그들 엔트리들의 방향성에 덜 지향됨을 결정할 수도 있다. 즉, 제로 또는 1 의 오더와 연관된 저차 앰비소닉스는, 도 1 및 도 2 에 도시된 바와 같이, 압력파의 방향의 관점에서 더 많이 제공하지 않지만 일부 볼륨 (이는 에너지로 나타냄) 을 제공하는 구면 기저 함수들에 대응한다.

[0176] 상기 예에서 설명된 동작들은 또한, 다음의 의사-코드에 따라 표현될 수도 있다. 하기의 의사-코드는 (따옴표없이) 문자 스트링들 “/*” 과 “*/” 의 연속적인 인스턴스들 내에 포함된 코멘트 스테이트먼트들의 형태로 주석들을 포함한다.

[0177] [U,S,V] = svd(audioframe,'ecom');

[0178] VS = V*S;

[0179] /* 다음 라인은 각각의 로우를 독립적으로 분석하고 (일 예로서) 제 1 로우에서의 값들을 제 5 엔트리로부터 제 25 엔트리까지 합산하여 대응하는 벡터의 방향성 지수 또는 방향성 메트릭을 결정하도록 지향됨. 엔트리들을 합산하기 전에 제곱함. 1 보다 큰 오더와 연관되는 각 로우에서의 엔트리들은 고차 앰비소닉스와 연관되고, 따라서, 방향성일 가능성이 더 많음.*/

[0180] sumVS = sum(VS(5:end,:).^2,1);

[0181] /* 다음 라인은 생성된 VS 행렬에 대한 제곱의 합을 분류하고 최대 값들의 세트 (예를 들어, 최대 값들 중 3 또는 4) 를 선택하도록 지향됨 */

[0182] [~,idxVS] = sort(sumVS,'descend');

[0183] U = U(:,idxVS);

[0184] V = V(:,idxVS);

[0185] S = S(idxVS,idxVS);

[0186] 즉, 상기 의사-코드에 따르면, 음장 분석 유닛 (44) 은, 1 보다 큰 오더를 갖는 구면 기저 함수에 대응하는 HOA 계수들 (11) 의 계수들로부터 분해된 VS[k] 행렬의 각각의 벡터의 엔트리들을 선택할 수도 있다. 그 후, 음장 분석 유닛 (44) 은 VS[k] 행렬의 각각의 벡터에 대한 이들 엔트리들을 제곱하고 제곱된 엔트리들을 합산하여, VS[k] 행렬의 각각의 벡터에 대한 방향성 메트릭 또는 지수를 식별하거나 컴퓨팅하거나 그렇지 않으면 결정할 수도 있다. 다음으로, 음장 분석 유닛 (44) 은 벡터들 각각의 개별 방향성 메트릭들에 기초하여 VS[k] 행렬의 벡터들을 분류할 수도 있다. 음장 분석 유닛 (44) 은, 최고의 대응하는 방향성을 갖는 그 벡터들이 먼저이고 최저의 대응하는 방향성을 갖는 그 벡터들이 마지막이도록 이들 벡터들을 방향성 메트릭들의 내림차순으로 분류할 수도 있다. 그 후, 음장 분석 유닛 (44) 은 최고의 상대적인 방향성 메트릭을 갖는 벡터들의 비-제로 서브세트를 선택할 수도 있다.

[0187] 음장 분석 유닛 (44) 은 전경 분석들의 임의의 조합을 수행하여, 심리 음향 코더 인스턴스화의 총 수 (이는 주변 또는 배경 채널들 (BG_{TOT}) 의 총 수의 함수일 수도 있음) 및 전경 채널들의 수를 결정할 수도 있다. 음장 분석 유닛 (44) 은, 전경 분석들의 임의의 조합에 기초하여, 전경 채널들의 총 수 (nFG) (45), 배경 음장의 오

더 (N_{BG}) 및 수 (nBGa), 그리고 전송할 추가적인 BG HOA 채널들의 인덱스들 (i) (이는, 도 4 의 예에서 배경 채널 정보 (43) 으로서 일괄적으로 표기될 수도 있음) 을 결정할 수도 있다.

[0188] 일부 예들에 있어서, 음장 분석 유닛 (44) 은 이러한 분석을 매 M개 샘플들마다 수행할 수도 있으며, 이는 프레임별 기반으로서 재서술될 수도 있다. 이러한 점에서, A 에 대한 값은 프레임별로 변할 수도 있다. 판정이 매 M개 샘플들마다 행해지는 비트스트림의 예가 도 10 내지 도 10o(ii) 에 도시된다. 다른 예들에 있어서, 음장 분석 유닛 (44) 은 이러한 분석을 프레임당 1회 초과로 수행하여, 프레임의 2 이상의 부분들을 분석할 수도 있다. 이에 따라, 그 기술들은 이러한 점에서 본 개시물에서 설명된 예들로 한정되지 않아야 한다.

[0189] 배경 선택 유닛 (48) 은 배경 채널 정보 (예를 들어, 배경 음장 (N_{BG}) 및 수 (nBGa), 그리고 전송할 추가적인 BG HOA 채널들의 인덱스들 (i)) 에 기초하여 배경 또는 주변 HOA 계수들 (47) 을 결정하도록 구성된 유닛을 나타낼 수도 있다. 예를 들어, N_{BG} 가 1 과 동일할 경우, 배경 선택 유닛 (48) 은 1 보다 작거나 같은 오더를 갖는 오디오 프레임의 각각의 샘플에 대한 HOA 계수들 (11) 을 선택할 수도 있다. 그 후, 배경 선택 유닛 (48) 은, 이 예에 있어서, 추가적인 BG HOA 계수들로서 인덱스들 (i) 중 하나에 의해 식별된 인덱스를 갖는 HOA 계수들 (11) 을 선택할 수도 있으며, 여기서, nBGa 는, 도 3 의 예에 도시된 오디오 디코딩 디바이스 (24) 와 같은 오디오 디코딩 디바이스로 하여금 비트스트림 (21) 로부터 BG HOA 계수들 (47) 을 파싱할 수 있게 하도록 비트스트림 (21) 에 명시되도록 비트스트림 생성 유닛 (42) 에 제공된다. 그 후, 배경 선택 유닛 (48) 은 주변 HOA 계수들 (47) 을 에너지 보상 유닛 (38) 에 출력할 수도 있다. 주변 HOA 계수들 (47) 은 디멘전들 D: $M \times [(N_{BG}+1)^2 + nBGa]$ 을 가질 수도 있다.

[0190] 전경 선택 유닛 (36) 은 (이들 전경 벡터들을 식별하는 하나 이상의 인덱스들을 나타낼 수도 있는) nFG (45) 에 기초하여 음장의 전경 또는 구별되는 성분들을 나타내는 리오더링된 US[k] 행렬 (33') 및 리오더링된 V[k] 행렬 (35') 의 성분들을 선택하도록 구성된 유닛을 나타낼 수도 있다. 전경 선택 유닛 (36) 은 nFG 신호들 (49) (이는 리오더링된 $US[k]_{1,\ldots,nFG}$ (49), $FG_{1,\ldots,nfG}[k]$ (49), 또는 $X_{PS}^{(1..nFG)}(k)$ (49) 로서 표기될 수도 있음) 를 심리 음향 오디오 코더 유닛 (40) 에 출력할 수도 있으며, 여기서, nFG 신호들 (49) 은 디멘전들 D: $M \times nFG$ 을 가질 수도 있고 각각은 모노 오디오 오브젝트들을 나타낼 수도 있다. 전경 선택 유닛 (36) 은 또한, 음장의 전경 성분들에 대응하는 리오더링된 V[k] 행렬 (35') (또는 $v^{(1..nFG)}(k)$ (35')) 을 공간-시간 보간 유닛 (50) 에 출력할 수도 있으며, 여기서, 전경 성분들에 대응하는 리오더링된 V[k] 행렬 (35') 의 성분들은 디멘전들 D: $(N+1)^2 \times nFG$ 을 갖는 전경 V[k] 행렬 (51_k) 로서 표기될 수도 있다 (이는 수학적으로 $\overline{V}_{1,\ldots,nFG}[k]$로서 표기될 수도 있음).

[0191] 에너지 보상 유닛 (38) 은 배경 선택 유닛 (48) 에 의한 HOA 채널들 중 다양한 채널들의 제거로 인한 에너지 손실을 보상하기 위해 주변 HOA 계수들 (47) 에 관하여 에너지 보상을 수행하도록 구성된 유닛을 나타낼 수도 있다. 에너지 보상 유닛 (38) 은 리오더링된 US[k] 행렬 (33'), 리오더링된 V[k] 행렬 (35'), nFG 신호들 (49), 전경 V[k] 벡터들 (51_k) 및 주변 HOA 계수들 (47) 중 하나 이상에 관하여 에너지 분석을 수행하고, 그 후, 이러한 에너지 분석에 기초하여 에너지 보상을 수행하여 에너지 보상된 주변 HOA 계수들 (47') 을 생성할 수도 있다. 에너지 보상 유닛 (38) 은 에너지 보상된 주변 HOA 계수들 (47') 을 심리 음향 오디오 코더 유닛 (40) 에 출력할 수도 있다.

[0192] 효과적으로, 에너지 보상 유닛 (38) 은 오더 감소된 주변 HOA 계수들 (47) (이는, 일부 예들에 있어서, 다음의 오더들/서브오더들: $[(N_{BG}+1)^2 + nBGa]$ 를 갖는 구면 기저 함수들에 대응하는 오직 포함된 계수들만의 관점에서 N 미만의 오더를 가짐) 을 생성하기 위해 HOA 계수들 (11) 에 의해 기술된 음장의 주변 성분들의 오더를 감소함으로써 야기된 음장의 배경 사운드 성분들의 전체 에너지에서의 가능한 감소들을 보상하는데 사용될 수도 있다.

일부 예들에 있어서, 에너지 보상 유닛 (38) 은, 주변 HOA 계수들 (47) 을 심리 음향 오디오 코더 유닛 (40) 에 출력하기 이전, 주변 HOA 계수들 (47) 의 제곱 평균 제곱근 (RMS) 에너지를 HOA 계수들 (11) 의 RMS (리오더링된 US[k] 행렬 (33'), 리오더링된 V[k] 행렬 (35'), nFG 신호들 (49), 전경 V[k] 벡터들 (51_k), 및 오더 감소된 주변 HOA 계수들 (47) 중 하나 이상의 총 에너지 분석을 통해 결정됨) 와 동일하거나 적어도 더 가까이 근사화시키도록 증가시키기 위하여 주변 HOA 계수들 (47) 의 $[(N_{BG}+1)^2 - nBGa]$ 컬럼들 각각에 적용하기 위한 증폭

값들의 형태로 보상 이득을 보상함으로써 에너지의 이러한 손실을 보상한다.

[0193] 일부 경우들에 있어서, 에너지 보상 유닛 (38) 은 리오더링된 US[k] 행렬 (33') 및 리오더링된 V[k] 행렬 (35') 중 하나 이상의 각각의 로우 및/또는 컬럼에 대한 RMS 를 식별할 수도 있다. 에너지 보상 유닛 (38) 은 또한, nFG 신호들 (49), 전경 V[k] 벡터들 (51_k), 및 오더 감소된 주변 HOA 계수들 (47) 을 포함할 수도 있는 선택된 전경 채널들 중 하나 이상의 각각의 로우 및/또는 컬럼에 대한 RMS 를 식별할 수도 있다. 리오더링된 US[k] 행렬 (33') 및 리오더링된 V[k] 행렬 (35') 중 하나 이상의 각각의 로우 및/또는 컬럼에 대한 RMS 는 RMS_{FULL} 로 표기된 벡터에 저장될 수도 있지만, nFG 신호들 (49), 전경 V[k] 벡터들 (51_k), 및 오더 감소된 주변 HOA 계수들 (47) 중 하나 이상의 각각의 로우 및/또는 컬럼에 대한 RMS 는 $RMS_{REDUCED}$ 로 표기된 벡터들에 저장될 수도 있다. 그 후, 에너지 보상 유닛 (38) 은 증폭 값 벡터 (Z) 를 다음의 수학식에 따라 컴퓨팅할 수도 있다: $Z = RMS_{FULL}/RMS_{REDUCED}$. 그 후, 에너지 보상 유닛 (38) 은 이 증폭 값 벡터 (Z) 또는 그 다양한 부분들을, nFG 신호들 (49), 전경 V[k] 벡터들 (51_k), 및 오더 감소된 주변 HOA 계수들 (47) 중 하나 이상에 적용할 수도 있다. 일부 경우들에 있어서, 증폭 값 벡터 (Z) 는 오직 다음의 수학식 $HOA_{BG-RED}' = HOA_{BG-RED}Z^T$ 마다 오더 감소된 주변 HOA 계수들 (47) 에만 적용되며, 여기서, HOA_{BG-RED} 는 오더 감소된 주변 HOA 계수들 (47) 을 나타내고, HOA_{BG-RED}' 는 에너지 보상된 감소된 주변 HOA 계수들 (47') 을 나타내며, Z^T 는 Z 벡터의 트랜스포즈를 나타낸다.

[0194] 일부 예들에 있어서, 리오더링된 US[k] 행렬 (33'), 리오더링된 V[k] 행렬 (35'), nFG 신호들 (49), 전경 V[k] 벡터들 (51_k) 및 오더 감소된 주변 HOA 계수들 (47) 중 하나 이상의 개별 로우들 및/또는 컬럼들의 각각의 RMS 를 결정하기 위해, 에너지 보상 유닛 (38) 은 먼저, 참조 구면 조화 계수들 (SHC) 렌더러를 컬럼들에 적용할 수도 있다. 에너지 보상 유닛 (38) 에 의한 참조 SHC 렌더러의 적용은 또한, 하기에서 더 상세하게 설명되는 바와 같이, 리오더링된 US[k] 행렬 (33'), 리오더링된 V[k] 행렬 (35'), nFG 신호들 (49), 전경 V[k] 벡터들 (51_k) 및 오더 감소된 주변 HOA 계수들 (47) 중 하나 이상의 로우들 및/또는 컬럼들에 의해 표현된 프레임의 각각의 로우 및/또는 컬럼에 의해 기술된 전체 음장의 에너지를 결정하기 위한 SHC 도메인에서의 RMS 의 결정을 허용한다.

[0195] 공간-시간 보간 유닛 (50) 은 k'번째 프레임에 대한 전경 V[k] 벡터들 (51_k) 및 이전 프레임 (이에 따라, k-1 표기) 에 대한 전경 V[k-1] 벡터들 (51_{k-1}) 을 수신하고 그리고 공간-시간 보간을 수행하여 보간된 전경 V[k] 벡터들을 생성하도록 구성된 유닛을 나타낼 수도 있다. 공간-시간 보간 유닛 (50) 은 nFG 신호들 (49) 을 전경 V[k] 벡터들 (51_k) 과 재결합하여 리오더링된 전경 HOA 계수들을 복원할 수도 있다. 그 후, 공간-시간 보간 유닛 (50) 은 리오더링된 전경 HOA 계수들을 보간된 V[k] 벡터들에 의해 제산하여 보간된 nFG 신호들 (49') 을 생성할 수도 있다. 공간-시간 보간 유닛 (50) 은 또한, 보간된 전경 V[k] 벡터들을 생성하는데 사용되었던 전경 V[k] 벡터들 (51_k) 의 벡터들을 출력하여, 오디오 디코딩 디바이스 (24) 와 같은 오디오 디코딩 디바이스가 보간된 전경 V[k] 벡터들을 생성하고 그에 의해 전경 V[k] 벡터들 (51_k) 을 복원하게 할 수도 있다.

보간된 전경 V[k] 벡터들을 생성하는데 사용된 전경 V[k] 벡터들 (51_k) 의 벡터들은 나머지 전경 V[k] 벡터들 (53) 로서 표기된다. 동일한 V[k] 및 V[k-1] 이 (보간된 벡터들 (V[k]) 을 생성하기 위해) 인코더 및 디코더에서 사용됨을 보장하기 위하여, 이들의 양자화된/역양자화된 버전들이 인코더 및 디코더에서 사용될 수도 있다.

[0196] 이러한 점에서, 공간-시간 보간 유닛 (50) 은 제 1 오디오 프레임 및 제 2 의 시간적으로 후속하는 또는 선행하는 오디오 프레임의 일부 다른 부분들로부터 제 1 오디오 프레임의 제 1 부분을 보간하는 유닛을 나타낼 수도 있다. 일부 예들에 있어서, 그 부분들은 서브-프레임들로서 표기될 수도 있으며, 여기서, 서브-프레임들에 관하여 수행된 바와 같은 보간은 도 45 내지 도 46E 에 관하여 하기에서 더 상세하게 설명된다. 다른 예들에 있어서, 공간-시간 보간 유닛 (50) 은, 도 37 내지 도 39 에서 설명된 바와 같이, 이전 프레임의 일부 최종 수의 샘플들 및 후속 프레임의 일부 제 1 수의 샘플들에 관하여 동작할 수도 있다. 공간-시간 보간 유닛 (50) 은, 이러한 보간을 수행함에 있어서, 비트스트림 (21) 에 명시되도록 요구되는 전경 V[k] 벡터들 (51_k) 의 샘플들의 수를 감소시킬 수도 있는데, 왜냐하면 보간된 전경 V[k] 벡터들을 생성하는데 사용된 전경 V[k] 벡터들 (51_k) 의 벡터들은 전경 V[k] 벡터들 (51_k) 의 서브세트를 나타내기 때문이다. 즉, (비트스트림 (21) 에

명시되는 전경 V[k] 벡터들 (51_k) 의 수를 감소시킴으로써) HOA 계수들 (11) 의 압축을 잠재적으로 더 효율적으로 하게 하기 위하여, 본 개시물에서 설명된 기술들의 다양한 양태들이 제 1 오디오 프레임의 하나 이상의 부분들의 보간을 제공할 수도 있으며, 여기서, 그 부분들 각각은 HOA 계수들 (11) 의 분해된 버전들을 나타낼 수도 있다.

[0197] 공간-시간 보간은 다수의 이점들을 발생시킬 수도 있다. 먼저, nFG 신호들 (49) 은, SVD 또는 다른 LIT 가 수행되는 블록별 특성으로 인해 프레임에 따라 연속적이지 않을 수도 있다. 즉, LIT 유닛 (30) 이 SVD 를 프레임별 기반으로 적용한다고 주어지면, 예를 들어, US[k] 행렬 (33) 및 V[k] 행렬 (35) 의 오더링되지 않은 특성에 의한 증거로서, 특정 불연속성들이 결과적인 변환된 HOA 계수들에 존재할 수도 있다. 이러한 보간을 수행함으로써, 보간이 프레임 경계들 (즉, HOA 계수들 (11) 의 프레임들로의 구분) 로 인해 도입된 임의의 아티팩트들을 잠재적으로 감소하는 평활화 효과를 가질 수도 있다고 주어지면, 불연속성은 감소될 수도 있다. 전경 V[k] 벡터들 (51_k) 을 사용하여 이러한 보간을 수행하고, 그 후, 복원된 리오더링된 HOA 계수들로부터의 보간된 전경 V[k] 벡터들 (51_k) 에 기초하여 보간된 nFG 신호들 (49') 을 생성하는 것은 프레임별 동작으로 인한 뿐만 아니라 nFG 신호들 (49) 의 리오더링으로 인한 적어도 일부 효과들을 평활화할 수도 있다.

[0198] 동작 시, 공간-시간 보간 유닛 (50) 은 제 1 프레임에 포함된 제 1 복수의 HOA 계수들 (11) 의 부분의 제 1 분해, 예를 들어, 전경 V[k] 벡터들 (51_k), 및 제 2 프레임에 포함된 제 2 복수의 HOA 계수들 (11) 의 부분의 제 2 분해, 예를 들어, 전경 V[k] 벡터들 (51_{k-1}) 로부터 제 1 오디오 프레임의 하나 이상의 서브-프레임들을 보간하여, 하나 이상의 서브-프레임들에 대한 분해된 보간된 구면 조화 계수들을 생성할 수도 있다.

[0199] 일부 예들에 있어서, 제 1 분해는 HOA 계수들 (11) 의 부분의 우-특이 벡터들을 나타내는 제 1 전경 V[k] 벡터들 (51_k) 을 포함한다. 유사하게, 일부 예들에 있어서, 제 2 분해는 HOA 계수들 (11) 의 부분의 우-특이 벡터들을 나타내는 제 2 전경 V[k] 벡터들 (51_k) 을 포함한다.

[0200] 즉, 구면 조화 기반 3D 오디오는 구면 상의 직교 기저 함수들의 관점에서 3D 압력 필드를 나타내는 파라메트릭 표현일 수도 있다. 표현의 오더 (N) 이 높아질수록, 공간 해상도도 잠재적으로 높아지고, 종종 (총 $(N+1)^2$ 계수들에 대한) 구면 조화 (SH) 계수들의 수도 커진다. 다수의 애플리케이션들에 대해, 계수들의 대역폭 압축이 계수들을 효율적으로 송신 및 저장할 수 있도록 요구될 수도 있다. 본 개시물에서 안내된 이 기술들은 특이값 분해 (SVD) 를 사용한 프레임 기반 차원수 감소 프로세스를 제공할 수도 있다. SVD 분석은 계수들의 각각의 프레임을 3개의 행렬들 (U, S 및 V) 로 분해할 수도 있다. 일부 예들에 있어서, 그 기술들은 US[k] 행렬에서의 벡터들 중 일부를 기본 음장의 전경 성분들로서 처리할 수도 있다. 하지만, 이러한 방식으로 처리될 경우, (U S[k] 행렬에서의) 이들 벡터들은, 심지어 동일의 구별되는 오디오 성분을 나타내더라도, 프레임별로 불연속이다. 이들 불연속성들은, 성분들이 변환 오디오 코더들을 통해 제공될 경우에 현저한 아티팩트들을 안내할 수도 있다.

[0201] 본 개시물에서 설명된 기술들은 이러한 불연속성을 다룰 수도 있다. 즉, 그 기술들은, V 행렬이 구면 조화 도메인에서 직교 공간 축들로서 해석될 수 있다는 관측에 기초할 수도 있다. U[k] 행렬은 구면 조화 (HOA) 데이터의 프로젝션을 그 기저 함수들의 관점에서 표현할 수도 있으며, 여기서, 불연속성은 모든 프레임을 변경하는 직교 공간 축 (V[k]) 에 기인될 수 있으며 따라서 그 자체가 불연속이다. 이는 푸리에 변환과 같은 유사한 분해와 다르며, 여기서, 기저 함수들은, 일부 예들에 있어서 프레임별로 일정하다. 이들 용어들에 있어서, SVD 는 매칭 추구 알고리즘으로 고려될 수도 있다. 본 개시물에서 설명된 기술들은 공간-시간 보간 유닛 (50) 으로 하여금 기저 함수들 (V[k]) 간의 연속성을, 그들 사이의 보간에 의해 프레임별로 유지하게 할 수도 있다.

[0202] 전술된 바와 같이, 보간이 샘플들에 대해 수행될 수도 있다. 이 경우는 서브프레임들이 샘플들의 단일 세트를 포함할 때 상기 설명에서 일반화된다. 샘플들에 걸친 그리고 서브프레임들에 걸친 보간의 경우 양쪽 모두에서, 보간 동작은 다음 식의 형태를 취할 수도 있다:

[0203] $$\overline{v(l)} = w(l)v(k) + \big(1 - w(l)\big)v(k-1).$$

[0204] 이 상기 식에서, 보간은 단일 V-벡터 $v(k-1)$ 로부터 단일 V-벡터 $v(k)$ 에 대해 수행될 수도 있고, 이것은 하나의 실시형태에서 인접 프레임들 *k* 와 *k-1* 로부터의 V-벡터들을 나타낼 수 있다. 상기 식에서, *l* 은 보

간이 실행되고 있는 레졸루션을 나타내고, 여기서 l 은 정수 샘플을 표시할 수도 있고, $l = 1, \ldots, T$ 이다 (여기서 T 는 보간이 실행되고 있고 출력 보간된 벡터들 $\overline{v(l)}$ 이 요구되는 샘플들의 길이이고, 또한 이 프로세스의 출력이 이들 벡터들의 l 을 생성함을 표시한다). 대안적으로, l 은 다수의 샘플들로 이루어진 서브프레임들을 표시할 수 있다. 예를 들어, 프레임이 4 개의 서브프레임들로 분할될 때, l 은, 각각의 하나의 서브프레임들에 대해, 1, 2, 3 및 4 의 값들을 포함할 수도 있다. l 의 값은 비트스트림을 통해 "CodedSpatialInterpolationTime" 이라고 불리는 필드로서 시그널링될 수도 있다 - 그래서 보간 동작이 디코더에서 복제될 수도 있도록 한다. $w(l)$ 은 보간 가중치들의 값들을 포함할 수도 있다. 보간이 선형적일 때, $w(l)$ 은, l 의 함수로서, 0 과 1 사이에서 선형적으로 그리고 단조적으로 변할 수도 있다. 다른 경우들에서, $w(l)$ 은 l 의 함수로서 0 과 1 사이에서 비-선형적으로 그러나 단조적 방식 (이를 테면 상향 코사인의 1/4 사이클) 으로 변할 수도 있다. 함수 $w(l)$ 은 함수들의 몇몇 상이한 가능성들 사이에서 인덱싱되고 비트스트림에서 "SpatialInterpolationMethod" 라고 불리는 필드로서 시그널링될 수도 있어서, 동일한 보간 동작이 디코더에 의해 복제될 수도 있도록 한다. $w(l)$ 이 0 에 가까운 값일 때, 출력 $\overline{v(l)}$ 은 $v(k-1)$ 에 의해 고도로 가중되거나 영향을 받을 수도 있다. 그에 반해 $w(l)$ 이 1 에 가까운 값일 때, 이것은 출력 $\overline{v(l)}$ 이 $v(k-1)$ 에 고도로 가중되거나 영향을 받는다는 것을 보장한다.

[0205] 계수 감소 유닛 (46) 은 감소된 전경 V[k] 벡터들 (55) 을 양자화 유닛 (52) 에 출력하기 위해 배경 채널 정보 (43) 에 기초하여 나머지 전경 V[k] 벡터들 (53) 에 대해 계수 감소를 수행하도록 구성된 유닛을 나타낼 수도 있다. 감소된 전경 V[k] 벡터들 (55) 은 디멘전들 $\mathrm{D}: [(N+1)^2 - (N_{BG}+1)^2\text{-nBGa}] \times \mathrm{nFG}$ 을 가질 수도 있다.

[0206] 계수 감소 유닛 (46) 은, 이와 관련하여, 나머지 전경 V[k] 벡터들 (53) 의 계수들의 수를 감소시키도록 구성된 유닛을 나타낼 수도 있다. 다시 말해서, 계수 감소 유닛 (46) 은 어떠한 방향 정보에 대해서도 거의 갖고 있지 않은 (나머지 전경 V[k] 벡터들 (53) 을 형성하는) 전경 V[k] 벡터들의 이들 계수들을 제거하도록 구성된 유닛을 나타낼 수도 있다. 전술된 바와 같이, 일부 예들에서, (N_{BG} 로서 표기될 수도 있는) 1 차 그리고 0 차 기저 함수들에 대응하는 구별되는 또는, 다시 말해서, 전경 V[k] 벡터들의 이들 계수들은 거의 방향 정보를 제공하지 않아서, ("계수 감소" 라고 지칭될 수도 있는 프로세스를 통해) 전경 V 벡터들로부터 삭제될 수 있다.

이 예에서, N_{BG} 에 대응하는 이들 계수들을 식별할 뿐만 아니라 $[(N_{BG}+1)^2+1, (N+1)^2]$ 의 세트로부터 (변수 TotalOfAddAmbHOAChan 에 의해 표기될 수도 있는) 부가적 HOA 채널들을 식별하기 위해 가장 큰 유연성이 제공될 수도 있다. 음장 분석 유닛 (44) 은 HOA 계수들 (11) 을 분석하여 BG_{TOT} 를 결정할 수도 있고, 이 BG_{TOT} 는 $(N_{BG}+1)^2$ 뿐만 아니라 TotalOfAddAmbHOAChan 을 식별할 수도 있고, 이는 배경 채널 정보 (43) 라고 총칭될 수도 있다. 계수 감소 유닛 (46) 은 그 후에 나머지 전경 V[k] 벡터들 (53) 로부터 $(N_{BG}+1)^2$ 및 TotalOfAddAmbHOAChan 에 대응하는 이들 계수들을 삭제하여 사이즈 $((N+1)^2 - (BG_{TOT}) \times nFG$ 의 보다 작은 디멘전의 V[k] 행렬 (55) 을 생성할 수도 있고, 이 행렬은 또한 감소된 전경 V[k] 벡터들 (55) 이라고 지칭될 수도 있다.

[0207] 양자화 유닛 (52) 은 임의의 형태의 양자화를 수행하여 감소된 전경 V[k] 벡터들 (55) 을 압축하여 코딩된 전경 V[k] 벡터들 (57) 을 생성하여, 이들 코딩된 전경 V[k] 벡터들 (57) 을 비트스트림 생성 유닛 (42) 에 출력하도록 구성된 유닛을 나타낼 수도 있다. 동작시, 양자화 유닛 (52) 은 음장의 공간 성분, 즉, 이 예에서 감소된 전경 V[k] 벡터들 (55) 중 하나 이상을 압축하도록 구성된 유닛을 나타낼 수도 있다. 예의 목적들을 위해, 감소된 전경 V[k] 벡터들 (55) 은, 계수 감소의 결과로서, (음장의 4 차 HOA 표현을 내포하는) 25 개 미만의 엘리먼트들을 각각 갖는 2 개의 로우 벡터들을 포함하는 것으로 가정한다. 2 개의 로우 벡터들에 대해

설명하였지만, 임의의 수의 벡터들이 감소된 전경 V[k] 벡터들 (55) 에 $(n+1)^2$ 까지 포함될 수도 있고, 여기서 n 은 음장의 HOA 표현의 오더를 표기한다. 게다가, 스칼라 및/또는 엔트로피 양자화를 수행하는 것으로서 아래에 설명하였지만, 양자화 유닛 (52) 은 감소된 전경 V[k] 벡터들 (55) 의 압축시 발생하는 임의의 형태의 양자화를 수행할 수도 있다.

[0208] 양자화 유닛 (52) 은 감소된 전경 V[k] 벡터들 (55) 을 수신하고 압축 방식을 수행하여 코딩된 전경 V[k] 벡터들 (57) 을 생성할 수도 있다. 이 압축 방식은 일반적으로 벡터 또는 데이터의 엘리먼트들을 압축하기 위한 임의의 상상가능한 압축 방식을 수반할 수도 있고, 더 상세하게 아래에 설명되는 예로 제한되어서는 안된다. 양자화 유닛 (52) 은, 예로서, 감소된 전경 V[k] 벡터들 (55) 의 각 엘리먼트의 부동 소수점 표현을 감소된 전경 V[k] 벡터들 (55) 의 각 엘리먼트의 정수 표현으로의 변환, 감소된 전경 V[k] 벡터들 (55) 의 정수 표현들의 균일한 양자화 및 나머지 전경 V[k] 벡터들 (55) 의 양자화된 정수 표현들의 카테고리화 및 코딩 중 하나 이상을 포함하는 압축 방식을 수행할 수도 있다.

[0209] 일부 예들에서, 이 압축 방식의 다양한 하나 이상의 프로세스들은, 하나의 예로서, 결과적인 비트스트림 (21) 에 대한 타겟 비트레이트를 달성하거나 거의 달성하기 위해 파라미터들에 의해 동적으로 제어될 수도 있다. 감소된 전경 V[k] 벡터들 (55) 각각이 서로 정규직교한다고 가정하면, 감소된 전경 V[k] 벡터들 (55) 각각은 독립적으로 코딩될 수도 있다. 일부 예들에서, 아래에 더 상세하게 설명된 바와 같이, 각각의 감소된 전경 V[k] 벡터들 (55) 의 각각의 엘리먼트는 (다양한 서브-모드들에 의해 정의된) 동일한 코딩 모드를 사용하여 코딩될 수도 있다.

[0210] 어떤 경우든, 전술된 바와 같이, 이 코딩 방식은 감소된 전경 V[k] 벡터들 (55) 각각의 (일부 예들에서, 32-비트 부동 소수점 수인) 각 엘리먼트의 부동 소수점 표현들을 16-비트 정수 표현으로 변환하는 것을 우선 수반할 수도 있다. 양자화 유닛 (52) 은, 일부 예들에서, 15 만큼의 우측 시프트에 의해 수행되는, 감소된 전경 V[k] 벡터들 (55) 중 주어진 하나의 감소된 전경 V[k] 벡터의 각각의 엘리먼트를 2^{15} 로 곱하는 것에 의해 이 부동 소수점-정수-변환을 수행할 수도 있다.

[0211] 양자화 유닛 (52) 은 그 후에 감소된 전경 V[k] 벡터들 (55) 중 주어진 하나의 감소된 전경 V[k] 벡터의 엘리먼트들 모두에 대해 균일한 양자화를 수행할 수도 있다. 양자화 유닛 (52) 은 $nbits$ 파라미터로서 표시될 수도 있는 값에 기초하여 양자화 스텝 사이즈를 식별할 수도 있다. 양자화 유닛 (52) 은 타겟 비트레이트 (41) 에 기초하여 이 $nbits$ 파라미터를 동적으로 결정할 수도 있다. 양자화 유닛 (52) 은 이 $nbits$ 파라미터의 함수로서 양자화 스텝 사이즈를 결정할 수도 있다. 하나의 예로서, 양자화 유닛 (52) 은 $2^{16-nbits}$ 과 동일한 양자화 스텝 사이즈 (본 개시물에서는 "델타" 또는 "Δ" 로서 표기됨) 를 결정할 수도 있다. 이 예에서, $nbits$ 가 6 과 동일한 경우, 델타는 2^{10} 과 동일하고 2^6 개의 양자화 레벨들이 존재한다. 이와 관련하여, 벡터 엘리먼트 v 에 대해, 양자화된 벡터 엘리먼트 v_q 는 [v/Δ] 와 동일하고 $-2^{nbits-1} < v_q < 2^{nbits-1}$ 이다.

[0212] 양자화 유닛 (52) 은 그 후에 양자화된 벡터 엘리먼트들의 카테고리화 및 잔차 코딩을 수행할 수도 있다. 하나의 예로서, 양자화 유닛 (52) 은, 주어진 양자화된 벡터 엘리먼트 v_q 에 대해 다음 식을 사용하여 이 엘리먼트가 대응하는 카테고리를 (카테고리 식별자 cid 를 결정함으로써) 식별할 수도 있다:

[0213] $$cid = \begin{cases} 0, & v_q = 0 \text{ 인 경우} \\ \lfloor log_2|v_q| \rfloor + 1, & v_q \neq 0 \text{ 인 경우} \end{cases}$$

[0214] 양자화 유닛 (52) 은 그 후에 v_q 가 포지티브 값인지 또는 네거티브 값인지 여부를 표시하는 부호 비트를 또한 식별하는 동안, 이 카테고리 인덱스 cid 를 허프만 코딩할 수도 있다. 양자화 유닛 (52) 은 그 다음에 이 카테고리에서 잔차를 식별할 수도 있다. 하나의 예로서, 양자화 유닛 (52) 은 다음 식에 따라 이 잔차를 결정할 수도 있다:

[0215] $$\text{잔차} = |v_q| - 2^{cid-1}$$

[0216] 양자화 유닛 (52) 은 그 후에 cid-1 비트들로 이 잔차를 블록 코딩할 수도 있다.

[0217] 다음 예는 이 카테고리화 및 잔차 코딩 프로세스의 단순화된 예를 예시한다. 우선, *nbits* 가 6 과 동일하여 $v_q \in [-31,31]$ 이도록 가정한다. 그 다음에, 다음을 가정한다:

cid	vq	cid 에 대한 허프만 코드
0	0	'1'
1	-1, 1	'01'
2	-3,-2, 2,3	'000'
3	-7,-6,-5,-4, 4,5,6,7	'0010'
4	-15,-14,...,-8, 8,...,14,15	'00110'
5	-31,-30,...,-16, 16,...,30,31	'00111'

[0218]

[0219] 또한, 다음을 가정한다:

cid	잔차에 대한 블록 코드
0	N/A
1	0, 1
2	01,00, 10,11
3	011,010,001,000, 100,101,110,111
4	0111, 0110...,0000, 1000,...,1110,1111
5	01111, ... ,00000, 10000, ... ,11111

[0220]

[0221] 따라서, v_q = [6, -17, 0, 0, 3] 에 대해, 다음이 결정될 수도 있다:

[0222] » cid = 3,5,0,0,2

[0223] » 부호 = 1,0,x,x,1

[0224] » 잔차 = 2,1,x,x,1

[0225] » 6 에 대한 비트들 = '0010' + '1' + '10'

[0226] » -17 에 대한 비트들 = '00111' + '0' + '0001'

[0227] » 0 에 대한 비트들 = '0'

[0228] » 0 에 대한 비트들 = '0'

[0229] » 3 에 대한 비트들 = '000' + '1' + '1'

[0230] » 총 비트들 = 7+10+1+1+5 = 24

[0231] » 평균 비트들 = 24/5 = 4.8

[0232] 전술한 단순화된 예에서 보여지지 않았지만, 양자화 유닛 (52) 은 *cid* 를 코딩할 때 *nbits* 의 상이한 값들에 대한 상이한 허프만 코드 북들을 선택할 수도 있다. 일부 예들에서, 양자화 유닛 (52) 은 *nbits* 값들 6, ..., 15 에 대한 상이한 허프만 코딩 테이블을 제공할 수도 있다. 게다가, 양자화 유닛 (52) 은 총 50 개의 허프만 코드 북들에 대해 6, ..., 15 로부터의 범위에 있는 상이한 *nbits* 값들 각각에 대한 5 개의 상이한 허프만 코드 북들을 포함할 수도 있다. 이와 관련하여, 양자화 유닛 (52) 은 다수의 상이한 통계적 컨텍스트들에서 *cid* 의 코딩을 수용하기 위해 복수의 상이한 허프만 코드 북들을 포함할 수도 있다.

[0233] 예시하기 위해, 양자화 유닛 (52) 은, *nbits* 값들 각각에 대해, 벡터 엘리먼트들 1 내지 4 를 코딩하기 위한 제 1 허프만 코드 북, 벡터 엘리먼트들 5 내지 9 를 코딩하기 위한 제 2 허프만 코드 북, 벡터 엘리먼트들 9 이상을 코딩하기 위한 제 3 허프만 코드 북을 포함할 수도 있다. 압축될 감소된 전경 V[*k*] 벡터들 (55) 중 하나가 감소된 전경 V[*k*] 벡터들 (55) 중 시간적으로 후속하는 대응하는 하나의 감소된 전경 V[*k*] 벡터로부터 예측되지 않고 합성 오디오 오브젝트 (예를 들어, 펄스 코드 변조된 (PCM) 오디오 오브젝트에 의해 원래 정의된 것) 의 공간 정보를 나타내지 않을 때 이들 처음 3 개의 허프만 코드 북들이 사용될 수도 있다. 양자화 유닛 (52) 은, *nbits* 값들 각각에 대해, 감소된 전경 V[*k*] 벡터들 (55) 중 하나가 감소된 전경 V[*k*] 벡터들 (55) 중 시간적으로 후속하는 대응하는 하나의 감소된 전경 V[*k*] 벡터로부터 예측될 때 이 감소된 전경 V[*k*] 벡터들 (55) 중 하나를 코딩하기 위한 제 4 허프만 코드 북을 부가적으로 포함할 수도 있다. 양자화 유닛 (52) 은 또한, *nbits* 값들 각각에 대해, 감소된 전경 V[*k*] 벡터들 (55) 중 하나가 합성 오디오 오브젝트를 나타낼 때 이 감소된 전경 V[*k*] 벡터들 (55) 중 하나를 코딩하기 위한 제 5 허프만 코드 북을 포함할 수도 있다. 이들 상이한 통계적 컨텍스트들 각각에 대해, 즉, 이 예에서 비-예측된 그리고 비-합성 컨텍스트, 예측된 컨텍스트 및 합성 컨텍스트에 대해 다양한 허프만 코드 북들이 개발될 수도 있다.

[0234] 다음 테이블은 허프만 테이블 선택 및 압축해제 유닛이 적절한 허프만 테이블을 선택할 수 있게 하기 위해 비트스트림에 특정될 비트들을 예시한다:

Pred mode	**HT info**	**HT 테이블**
0	0	HT5
0	1	HT{1,2,3}
1	0	HT4
1	1	HT5

[0235]

[0236] 전술한 테이블에서, 예측 모드 ("Pred mode") 는 예측이 현재 벡터에 대해 수행되었는지 여부를 표시하는 한편, 허프만 테이블 ("HT info") 은 허프만 테이블들 1 내지 5 중 하나를 선택하는데 사용되는 부가적인 허프만 코드 북 (또는 테이블) 정보를 표시한다.

[0237] 다음 테이블은 다양한 통계적 컨텍스트들 또는 시나리오들이 주어진 이 허프만 테이블 선택 프로세스를 추가로 예시한다.

	레코딩	합성
예측 없음	HT{1,2,3}	HT5
예측 있음	HT4	HT5

[0238]

[0239] 전술한 테이블에서, "레코딩" 컬럼은 벡터가 레코딩된 오디오 오브젝트를 나타낼 때의 코딩 컨텍스트를 표시하는 한편, "합성" 컬럼은 벡터가 합성 오디오 오브젝트를 나타낼 때의 코딩 컨텍스트를 표시한다. "예측 없음" 로우는 예측이 벡터 엘리먼트들에 대해 수행되지 않을 때의 코딩 컨텍스트를 표시하는 한편, "예측 있음" 로우는 예측이 벡터 엘리먼트들에 대해 수행될 때의 코딩 컨텍스트를 표시한다. 이 테이블에 도시된 바와 같이, 양자화 유닛 (52) 은 벡터가 레코딩된 오디오 오브젝트를 나타내고 예측이 벡터 엘리먼트들에 대해 수행되지 않을 때 HT{1, 2, 3} 을 선택한다. 양자화 유닛 (52) 은 오디오 오브젝트가 합성 오디오 오브젝트를 나타내고 예측이 벡터 엘리먼트들에 대해 수행되지 않을 때 HT5 를 선택한다. 양자화 유닛 (52) 은 벡터가 레코딩된 오디오 오브젝트를 나타내고 예측이 벡터 엘리먼트들에 대해 수행될 때 HT4 를 선택한다. 양자화 유닛 (52) 은 오디오 오브젝트가 합성 오디오 오브젝트를 나타내고 예측이 벡터 엘리먼트들에 대해 수행될 때 HT5 를 선택한다.

[0240] 이와 관련하여, 양자화 유닛 (52) 은 전술된 스칼라 양자화 및/또는 허프만 인코딩을 수행하여 감소된 전경 V[k] 벡터들 (55) 을 압축하여, 사이드 채널 정보 (57) 라고 지칭될 수도 있는 코딩된 전경 V[k] 벡터들 (57) 을 출력할 수도 있다. 이 사이드 채널 정보 (57) 는 나머지 전경 V[k] 벡터들 (55) 을 코딩하는데 사용되는 신택스 엘리먼트들을 포함할 수도 있다. 양자화 유닛 (52) 은 도 10b 및 도 10c 중 하나의 예에서 도시된 것과 유사한 방식으로 사이드 채널 정보 (57) 를 출력할 수도 있다.

[0241] 전술된 바와 같이, 양자화 유닛 (52) 은 사이드 채널 정보 (57) 에 대한 신택스 엘리먼트들을 생성할 수도 있다. 예를 들어, 양자화 유닛 (52) 은 복수의 구성 모드들 중 어떤 것이 선택되었는지를 표기한 (하나 이상의 프레임들을 포함할 수도 있는) 액세스 유닛의 헤더에서의 신택스 엘리먼트를 특정할 수도 있다. 액세스 유닛 기반으로 특정되는 것으로 설명되었지만, 양자화 유닛 (52) 은 프레임 기반으로 또는 임의의 다른 주기적 기반 또는 비-주기적 기반 (이를 테면 전체 비트스트림에 대해 한번) 으로 이 신택스 엘리먼트를 특정할 수도 있다. 어떤 경우든, 이 신택스 엘리먼트는 이 구별되는 성분의 방향 양태들을 나타내도록 감소된 전경 V[k] 벡터들 (55) 의 계수들의 비-제로 세트를 명시하기 위해 4 개의 구성 모드들 중 어떤 것이 선택되었는지를 표시하는 2 비트들을 포함할 수도 있다. 신택스 엘리먼트는 "codedVVecLength" 로서 표기될 수도 있다. 이러한 방식으로, 양자화 유닛 (52) 은 4 개의 구성 모드들 중 어떤 것이 비트스트림에서 코딩된 전경 V[k] 벡터들 (57) 을 명시하는데 사용되었는지를 비트스트림에서 시그널링 또는 그렇지 않으면 특정할 수도 있다. 4 개의 구성 모드들에 대해 설명되었지만, 이 기술들은 단일 구성 모드 또는 복수의 구성 모드들을 포함하는 임의의 수의 구성 모드들이 아닌 4 개의 구성 모드들로 제한되어서는 안된다. 스칼라/엔트로피 양자화 유닛 (53) 은 또한 사이드 채널 정보 (57) 에서 다른 신택스 엘리먼트로서 플래그 (63) 를 특정할 수도 있다.

[0242] 오디오 인코딩 디바이스 (20) 내에 포함된 심리 음향 오디오 코더 유닛 (40) 은 심리 음향 오디오 코더의 다수의 인스턴스들을 나타낼 수도 있고, 그 각각은 에너지 보상된 주변 HOA 계수들 (47') 및 보간된 nFG 신호들 (49') 각각의 상이한 오디오 오브젝트 또는 HOA 채널을 인코딩하여 인코딩된 주변 HOA 계수들 (59) 및 인코딩된 nFG 신호들 (61) 을 생성하는데 사용된다. 심리 음향 오디오 코더 유닛 (40) 은 인코딩된 주변 HOA 계수들 (59) 및 인코딩된 nFG 신호들 (61) 을 비트스트림 생성 유닛 (42) 에 출력할 수도 있다.

[0243] 일부 경우들에서, 이 심리 음향 오디오 코더 유닛 (40) 은 AAC (advanced audio coding) 인코딩 유닛의 하나 이상의 인스턴스들을 나타낼 수도 있다. 심리 음향 오디오 코더 유닛 (40) 은 에너지 보상된 주변 HOA 계수들 (47') 및 보간된 nFG 신호들 (49') 의 각각의 컬럼 또는 로우를 인코딩할 수도 있다. 종종, 심리 음향 오디오 코더 유닛 (40) 은 에너지 보상된 주변 HOA 계수들 (47') 및 보간된 nFG 신호들 (49') 에 남아있는 오더/서브-오더 조합들 각각에 대해 AAC 인코딩 유닛의 인스턴스를 인보크할 수도 있다. 배경 구면 조화 계수들 (31) 이 AAC 인코딩 유닛을 사용하여 어떻게 인코딩될 수도 있는지에 관한 더 많은 정보는, 124 차 컨벤션, 2008년 5월 17-20일자로 제시되고 http://ro.uow.edu.au/cgi/viewcontent.cgi?article=8025&context=engpapers 에서 입수가능한, "Encoding

Higher Order Ambisonics with AAC" 라는 명칭의, Eric Hellerud 등에 의한 컨벤션 문서에서 발견될 수 있다.

일부 경우들에서, 오디오 인코딩 유닛 (14) 은 보간된 nFG 신호들 (49') 을 인코딩하는데 사용되는 것보다 더 낮은 타겟 비트레이트를 사용하여 에너지 보상된 주변 HOA 계수들 (47') 을 오디오 인코딩할 수도 있어서, 보간된 nFG 신호들 (49') 에 비해 에너지 보상된 주변 HOA 계수들 (47') 을 더 많이 잠재적으로 압축한다.

[0244] 오디오 인코딩 디바이스 (20) 내에 포함된 비트스트림 생성 유닛 (42) 은 (디코딩 디바이스에 의해 공지된 포맷으로 지칭할 수도 있는) 공지된 포맷을 준수하도록 데이터를 포맷팅하여, 벡터-기반 비트스트림 (21) 을 생성하는 유닛을 나타낸다. 비트스트림 생성 유닛 (42) 은, 일부 예들에서, 코딩된 전경 V[*k*] 벡터들 (57), 인코딩된 주변 HOA 계수들 (59), 인코딩된 nFG 신호들 (61) 및 배경 채널 정보 (43) 를 수신할 수도 있는 멀티플렉서를 나타낼 수도 있다. 비트스트림 생성 유닛 (42) 은 그 후에 코딩된 전경 V[*k*] 벡터들 (57), 인코딩된 주변 HOA 계수들 (59), 인코딩된 nFG 신호들 (61) 및 배경 채널 정보 (43) 에 기초하여 비트스트림 (21) 을 생성할 수도 있다. 비트스트림 (21) 은 프라이머리 또는 메인 비트스트림 및 하나 이상의 사이드 채널 비트스트림들을 포함할 수도 있다.

[0245] 도 4 의 예에서 도시되지 않았지만, 오디오 인코딩 디바이스 (20) 는 또한 현재 프레임이 방향-기반 합성 또는 벡터-기반 합성을 사용하여 인코딩되어야 하는지 여부에 기초하여 (예를 들어, 방향-기반 비트스트림 (21) 과 벡터-기반 비트스트림 (21) 사이에서) 오디오 인코딩 디바이스 (20) 로부터 출력된 비트스트림을 스위칭하는 비트스트림 출력 유닛을 포함할 수도 있다. 이 비트스트림 출력 유닛은, (HOA 계수들 (11) 이 합성 오디오 오브젝트로부터 생성되었다는 검출 결과로서) 방향-기반 합성이 수행되었는지 또는 (HOA 계수들이 레코딩되었다는 검출 결과로서) 벡터-기반 합성이 수행되었는지 여부를 표시하는, 컨텐츠 분석 유닛 (26) 에 의해 출력된 신택스 엘리먼트에 기초하여 이 스위치를 수행할 수도 있다. 비트스트림 출력 유닛은 비트스트림들 (21) 중의 각각의 하나의 비트스트림과 함께 현재 프레임에 대해 사용된 현재 인코딩 또는 이 스위치를 표시하기 위해 올바른 헤더 신택스를 특정할 수도 있다.

[0246] 일부 경우들에서, 기술들의 다양한 양태들은 또한 오디오 인코딩 디바이스 (20) 를 인에이블시켜서 HOA 계수들 (11) 이 합성 오디오 오브젝트로부터 생성되는지 여부를 결정할 수도 있다. 기술들의 이들 양태들은 오디오 인코딩 디바이스 (20) 를 인에이블시켜서 음장을 나타내는 구면 조화 계수들이 합성 오디오 오브젝트로부터 생성되는지 여부의 표시를 획득하도록 구성될 수도 있다.

[0247] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 또한 구면 조화 계수들이 합성 오디오 오브젝트로부터 생성되는지 여부를 결정하도록 구성된다.

[0248] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 감소된 프레임화된 구면 조화 계수 행렬을 획득하기 위해 음장을 나타내는 구면 조화 계수들의 적어도 일부분을 저장한 프레임화된 구면 조화 계수 행렬로부터 제 1 벡터를 제외하도록 구성된다.

[0249] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 감소된 프레임화된 구면 조화 계수 행렬을 획득하기 위해 음장을 나타내는 구면 조화 계수들의 적어도 일부분을 저장한 프레임화된 구면 조화 계수 행렬로부터 제 1 벡터를 제외하고, 감소된 프레임화된 구면 조화 계수 행렬의 나머지 벡터들에 기초하여 감소된 프레임화된 구면 조화 계수 행렬의 벡터를 예측하도록 구성된다.

[0250] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 감소된 프레임화된 구면 조화 계수 행렬을 획득하기 위해 음장을 나타내는 구면 조화 계수들의 적어도 일부분을 저장한 프레임화된 구면 조화 계수 행렬로부터 제 1 벡터를 제외하고, 감소된 프레임화된 구면 조화 계수 행렬의 나머지 벡터들의 합산에 적어도 부분적으로 기초하여 감소된 프레임화된 구면 조화 계수 행렬의 벡터를 예측하도록 구성된다.

[0251] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 프레임화된 구면 조화 계수 행렬의 나머지 벡터들의 합산에 적어도 부분적으로 기초하여 구면 조화 계수들의 적어도 일부분을 저장한 프레임화된 구면 조화 계수 행렬의 벡터를 예측하도록 구성된다.

[0252] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 프레임화된 구면 조화 계수 행렬의 나머지 벡터들의 합산에 적어도 부분적으로 기초하여 구면 조화 계수들의 적어도 일부분을 저장한 프레임화된 구면 조화 계수 행렬의 벡터를 예측하고, 예측된 벡터에 기초하여 에러를 컴퓨팅하도록 또한 구성되도록 구성된다.

[0253] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 프레임화된 구면 조화 계수 행렬의 나머지 벡터들의 합산에 적어도 부분적으로 기초하여 구면 조화 계수들의 적어도 일부분을 저장한 프레임화된 구면 조화 계수 행렬의 벡터를 예측하고, 예측된 벡터 및 프레임화된 구면 조화 계수 행렬의 대응 벡터에 기초하여 에러를 컴퓨팅

하도록 구성되도록 구성된다.

[0254] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 프레임화된 구면 조화 계수 행렬의 나머지 벡터들의 합산에 적어도 부분적으로 기초하여 구면 조화 계수들의 적어도 일부분을 저장한 프레임화된 구면 조화 계수 행렬의 벡터를 예측하고, 예측된 벡터 및 프레임화된 구면 조화 계수 행렬의 대응 벡터의 차이의 절대 값의 합산으로서 에러를 컴퓨팅하도록 구성되도록 구성된다.

[0255] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 프레임화된 구면 조화 계수 행렬의 나머지 벡터들의 합산에 적어도 부분적으로 기초하여 구면 조화 계수들의 적어도 일부분을 저장한 프레임화된 구면 조화 계수 행렬의 벡터를 예측하고, 예측된 벡터 및 프레임화된 구면 조화 계수 행렬의 대응 벡터에 기초하여 에러를 컴퓨팅하고, 에러 및 프레임화된 구면 조화 계수 행렬의 대응 벡터의 에너지에 기초하여 비율을 컴퓨팅하며, 이 비율을 임계치와 비교하여 음장을 나타내는 구면 조화 계수들이 합성 오디오 오브젝트로부터 생성되는지 여부를 결정하도록 구성되도록 구성된다.

[0256] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 구면 조화 계수들의 압축된 버전을 저장한 비트스트림 (21) 에서의 표시를 명시하도록 구성되도록 구성된다.

[0257] 일부 경우들에서, 다양한 기술들은 오디오 인코딩 디바이스 (20) 를 인에이블시켜서 HOA 계수들 (11) 에 대한 변환을 수행할 수도 있다. 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 음장의 구별되는 성분들을 설명하는 하나 이상의 제 1 벡터들 및 음장의 배경 성분들을 설명하는 하나 이상의 제 2 벡터들을 획득하도록 구성될 수도 있고, 하나 이상의 제 1 벡터들 및 하나 이상의 제 2 벡터들 양쪽 모두는 적어도 복수의 구면 조화 계수들 (11) 에 대한 변환을 수행함으로써 생성된다.

[0258] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 에서, 여기서 변환은 복수의 구면 조화 계수들의 좌-특이 벡터들을 나타내는 U 행렬, 복수의 구면 조화 계수들의 특이값들을 나타내는 S 행렬 및 복수의 구면 조화 계수들 (11) 의 우-특이 벡터들을 나타내는 V 행렬을 생성하는 특이값 분해를 포함한다.

[0259] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 에서, 여기서 하나 이상의 제 1 벡터들은, 오디오 인코딩 이전에, U 행렬의 하나 이상의 인코딩된 U_{DIST} 벡터들을 S 행렬의 하나 이상의 S_{DIST} 벡터들에 의해 곱함으로써 생성된 하나 이상의 오디오 인코딩된 $U_{DIST} * S_{DIST}$ 벡터들을 포함하고, 여기서 U 행렬 및 S 행렬은 적어도 복수의 구면 조화 계수들에 대해 특이값 분해를 수행함으로써 생성된다.

[0260] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 에서, 여기서 하나 이상의 제 1 벡터들은, 오디오 인코딩 이전에, U 행렬의 하나 이상의 인코딩된 U_{DIST} 벡터들을 S 행렬의 하나 이상의 S_{DIST} 벡터들에 의해 곱함으로써 생성된 하나 이상의 오디오 인코딩된 $U_{DIST} * S_{DIST}$ 벡터들, 및 V 행렬의 트랜스포즈의 하나 이상의 V^{T}_{DIST} 벡터들을 포함하고, 여기서 U 행렬 및 S 행렬 및 V 행렬은 적어도 복수의 구면 조화 계수들 (11) 에 대해 특이값 분해를 수행함으로써 생성된다.

[0261] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 에서, 여기서 하나 이상의 제 1 벡터들은, 오디오 인코딩 이전에, U 행렬의 하나 이상의 인코딩된 U_{DIST} 벡터들을 S 행렬의 하나 이상의 S_{DIST} 벡터들에 의해 곱함으로써 생성된 하나 이상의 $U_{DIST} * S_{DIST}$ 벡터들, 및 V 행렬의 트랜스포즈의 하나 이상의 V^{T}_{DIST} 벡터들을 포함하고, 여기서 U 행렬, S 행렬 및 V 행렬은 적어도 복수의 구면 조화 계수들에 대해 특이값 분해를 수행함으로써 생성되었고, 여기서 오디오 인코딩 디바이스 (20) 는 또한 하나 이상의 $U_{DIST} * S_{DIST}$ 벡터들 및 하나 이상의 V^{T}_{DIST} 벡터들을 형성하기 위해 비트스트림으로부터 추출될 벡터들의 수를 표시한 값 D 를 획득하도록 구성된다.

[0262] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 에서, 여기서 하나 이상의 제 1 벡터들은, 오디오 인코딩 이전에, U 행렬의 하나 이상의 인코딩된 U_{DIST} 벡터들을 S 행렬의 하나 이상의 S_{DIST} 벡터들에 의해 곱함으로써 생성된 하나 이상의 $U_{DIST} * S_{DIST}$ 벡터들, 및 V 행렬의 트랜스포즈의 하나 이상의 V^{T}_{DIST} 벡터들을 포함하고, 여기서 U 행렬, S 행렬 및 V 행렬은 적어도 복수의 구면 조화 계수들에 대해 특이값 분해를 수행함으로써 생성되었고, 여기서 오디오 인코딩 디바이스 (20) 는 또한 하나 이상의 $U_{DIST} * S_{DIST}$ 벡터들 및 하나 이상의 V^{T}_{DIST} 벡터들을 형성하기 위해 비트스트림으로부터 추출될 벡터들의 수를 표시한 값 D 를 오디오-프레임간 기반으로 획득하

도록 구성된다.

[0263] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 에서, 여기서 변환은 음장의 구별되는 성분들 및 음장의 배경 성분들을 식별하기 위한 주요 성분 분석을 포함한다.

[0264] 본 개시물에서 설명된 기술들의 다양한 양태들은 양자화 에러를 보상하도록 구성된 오디오 인코딩 디바이스 (20) 를 제공할 수도 있다.

[0265] 일부 경우들에서, 오디오 인코딩 디바이스 (20) 는 음장의 하나 이상의 성분들을 나타내는 하나 이상의 제 1 벡터들을 양자화하고, 음장의 동일한 하나 이상의 성분들을 또한 나타내는 하나 이상의 제 2 벡터들에 하나 이상의 제 1 벡터들의 양자화로 인해 도입된 에러를 보상하도록 구성될 수도 있다.

[0266] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스는 적어도 부분적으로 음장을 설명하는 복수의 구면 조화 계수들에 대해 특이값 분해를 수행함으로써 생성된 V 행렬의 트랜스포즈로부터 하나 이상의 벡터들을 양자화하도록 구성된다.

[0267] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스는 또한 음장을 나타내는 복수의 구면 조화 계수들에 대해 특이값 분해를 수행하여 복수의 구면 조화 계수들의 좌-특이 벡터들을 나타내는 U 행렬, 복수의 구면 조화 계수들의 특이값들을 나타내는 S 행렬 및 복수의 구면 조화 계수들의 우-특이 벡터들을 나타내는 V 행렬을 생성하도록 구성되고, V 행렬의 트랜스포즈로부터 하나 이상의 벡터들을 양자화하도록 구성된다.

[0268] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스는 또한 음장을 나타내는 복수의 구면 조화 계수들에 대해 특이값 분해를 수행하여 복수의 구면 조화 계수들의 좌-특이 벡터들을 나타내는 U 행렬, 복수의 구면 조화 계수들의 특이값들을 나타내는 S 행렬 및 복수의 구면 조화 계수들의 우-특이 벡터들을 나타내는 V 행렬을 생성하도록 구성되고, V 행렬의 트랜스포즈로부터 하나 이상의 벡터들을 양자화하도록 구성되며, U 행렬의 하나 이상의 U 벡터들을 S 행렬의 하나 이상의 S 벡터들에 의해 곱함으로써 컴퓨팅된 하나 이상의 U * S 벡터들에 양자화로 인해 도입된 에러를 보상하도록 구성된다.

[0269] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스는 또한 음장을 나타내는 복수의 구면 조화 계수들에 대해 특이값 분해를 수행하여 복수의 구면 조화 계수들의 좌-특이 벡터들을 나타내는 U 행렬, 복수의 구면 조화 계수들의 특이값들을 나타내는 S 행렬 및 복수의 구면 조화 계수들의 우-특이 벡터들을 나타내는 V 행렬을 생성하고, U 행렬의 하나 이상의 U_{DIST} 벡터들을 결정하는 것으로서, 그 각각은 음장의 구별되는 성분에 대응하는, 그 U 행렬의 하나 이상의 U_{DIST} 벡터들을 결정하는 것을 행하고, S 행렬의 하나 이상의 S_{DIST} 벡터들을 결정하는 것으로서, 그 각각은 음장의 동일한 구별되는 성분에 대응하는, 그 S 행렬의 하나 이상의 S_{DIST} 벡터들을 결정하는 것을 행하며, V 행렬의 트랜스포즈의 하나 이상의 V^{T}_{DIST} 벡터들을 결정하는 것으로서, 그 각각은 음장의 동일한 구별되는 성분에 대응하는, 그 V 행렬의 트랜스포즈의 하나 이상의 V^{T}_{DIST} 벡터들을 결정하는 것을 행하도록 구성되고, 하나 이상의 V^{T}_{DIST} 벡터들을 양자화하여 하나 이상의 $V^{T}_{Q_DIST}$ 벡터들을 생성하도록 구성되며, 하나 이상의 에러 보상된 $U_{DIST} * S_{DIST}$ 벡터들을 생성하도록 U 행렬의 하나 이상의 U_{DIST} 벡터들을 S 행렬의 하나 이상의 S_{DIST} 벡터들에 의해 곱함으로써 컴퓨팅된 하나 이상의 $U_{DIST} * S_{DIST}$ 벡터들에 양자화로 인해 도입된 에러를 보상하도록 구성된다.

[0270] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스는 하나 이상의 U_{DIST} 벡터들, 하나 이상의 S_{DIST} 벡터들 및 하나 이상의 V^{T}_{DIST} 벡터들에 기초하여 구별되는 구면 조화 계수들을 결정하고, 하나 이상의 $V^{T}_{Q_DIST}$ 벡터들에 대해 의사 역을 수행하여 구별되는 구면 조화 계수들을 하나 이상의 $V^{T}_{Q_DIST}$ 벡터들에 의해 나누어서 V^{T}_{DIST} 벡터들의 양자화를 통해 도입된 에러를 적어도 부분적으로 보상하는 에러 보상된 하나 이상의 $U_{C_DIST} * S_{C_DIST}$ 벡터들을 생성하도록 구성된다.

[0271] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스는 또한 음장을 나타내는 복수의 구면 조화 계수들에 대해 특이값 분해를 수행하여 복수의 구면 조화 계수들의 좌-특이 벡터들을 나타내는 U 행렬, 복수의 구면 조화 계수들의 특이값들을 나타내는 S 행렬 및 복수의 구면 조화 계수들의 우-특이 벡터들을 나타내는 V 행렬을 생성하고, 음장의 하나 이상의 배경 성분을 설명하는 U 행렬의 하나 이상의 U_{BG} 벡터들 및 음장의 하나 이상의 구별되는 성

분들을 설명하는 U 행렬의 하나 이상의 U_{DIST} 벡터들을 결정하고, 음장의 하나 이상의 배경 성분을 설명하는 S 행렬의 하나 이상의 S_{BG} 벡터들 및 음장의 하나 이상의 구별되는 성분들을 설명하는 S 행렬의 하나 이상의 S_{DIST} 벡터들을 결정하며, 하나 이상의 V^{T}_{DIST} 벡터들 및 V 행렬의 트랜스포즈의 하나 이상의 V^{T}_{BG} 벡터들을 결정하도록 구성되고, 여기서 V^{T}_{DIST} 벡터들은 음장의 하나 이상의 구별되는 성분들을 설명하고, V^{T}_{BG} 는 음장의 하나 이상의 배경 성분들을 설명하고, 하나 이상의 V^{T}_{DIST} 벡터들을 양자화하여 하나 이상의 $V^{T}_{Q_DIST}$ 벡터들을 생성하도록 구성되며, 에러 보상된 배경 구면 조화 계수들을 생성하도록 하나 이상의 U_{BG} 벡터들을 하나 이상의 S_{BG} 벡터들에 의해 그리고 그 후에 하나 이상의 V^{T}_{BG} 벡터들에 의해 곱함으로써 형성된 배경 구면 조화 계수들에 양자화로 인해 도입된 에러를 보상하도록 구성된다.

[0272] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스는 U_{DIST} 벡터들 및 U_{DIST} 벡터들을 S_{DIST} 벡터들에 의해 곱함으로써 형성된 하나 이상의 $U_{DIST} * S_{DIST}$ 벡터들에 기초하여 에러를 결정하고 그 결정된 에러를 배경 구면 조화 계수들에 가산하여 에러 보상된 배경 구면 조화 계수들을 생성하도록 구성된다.

[0273] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스는 하나 이상의 에러 보상된 제 2 벡터들을 생성하기 위해 음장의 동일한 하나 이상의 성분들을 또한 나타내는 하나 이상의 제 2 벡터들에 하나 이상의 제 1 벡터들의 양자화로 인해 도입된 에러를 보상하도록 구성되고, 또한 하나 이상의 에러 보상된 제 2 벡터들 및 양자화된 하나 이상의 제 1 벡터들을 포함시키기 위해 비트스트림을 생성하도록 구성된다.

[0274] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스는 하나 이상의 에러 보상된 제 2 벡터들을 생성하기 위해 음장의 동일한 하나 이상의 성분들을 또한 나타내는 하나 이상의 제 2 벡터들에 하나 이상의 제 1 벡터들의 양자화로 인해 도입된 에러를 보상하도록 구성되고, 또한 하나 이상의 에러 보상된 제 2 벡터들을 오디오 인코딩하고, 오디오 인코딩된 하나 이상의 에러 보상된 제 2 벡터들 및 양자화된 하나 이상의 제 1 벡터들을 포함시키기 위해 비트스트림을 생성하도록 구성된다.

[0275] 기술들의 다양한 양태들은 또한 오디오 인코딩 디바이스 (20) 를 인에이블시켜서 감소된 구면 조화 계수들 또는 그의 분해들을 생성한다. 일부 경우들에서, 오디오 인코딩 디바이스 (20) 는, 타겟 비트레이트에 기초하여, 복수의 구면 조화 계수들 또는 그의 분해들에 대해 오더 감소를 수행하여 감소된 구면 조화 계수들 또는 감소된 그의 분해들을 생성하도록 구성될 수도 있고, 여기서 복수의 구면 조화 계수들은 음장을 나타낸다.

[0276] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 또한, 오더 감소를 수행하기 이전에, 복수의 구면 조화 계수들에 대해 특이값 분해를 수행하여 음장의 구별되는 성분들을 설명하는 하나 이상의 제 1 벡터들 및 음장의 배경 성분들을 식별하는 하나 이상의 제 2 벡터들을 식별하도록 구성되고, 하나 이상의 제 1 벡터들, 하나 이상의 제 2 벡터들 또는 하나 이상의 제 1 벡터들과 하나 이상의 제 2 벡터들 양쪽 모두에 대해 오더 감소를 수행하도록 구성된다.

[0277] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 또한 복수의 구면 조화 계수들 또는 그의 분해들에 대해 컨텐츠 분석을 수행하도록 구성되고, 타겟 비트레이트 및 컨텐츠 분석에 기초하여, 복수의 구면 조화 계수들 또는 그의 분해들에 대해 오더 감소를 수행하여 감소된 구면 조화 계수들 또는 감소된 그의 분해들을 생성하도록 구성된다.

[0278] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 복수의 구면 조화 계수들 또는 그의 분해들에 대해 공간 분석을 수행하도록 구성된다.

[0279] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 복수의 구면 조화 계수들 또는 그의 분해들에 대해 확산 분석을 수행하도록 구성된다.

[0280] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 하나 이상의 프로세서들이 복수의 구면 조화 계수들 또는 그의 분해들에 대해 공간 분석 및 확산 분석을 수행하도록 구성된다.

[0281] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 또한 감소된 구면 조화 계수들 또는 감소된 그의 분해들 중 어떤 것들이 감소된 구면 조화 계수들 또는 감소된 그의 분해들을 포함하는 비트스트림에 대응하는지에 대해 구면 기저 함수들의 하나 이상의 오더들 및/또는 하나 이상의 서브-오더들을 명시하도록 구성된다.

[0282] 이들 및 다른 경우들에서, 감소된 구면 조화 계수들 또는 감소된 그의 분해들은 복수의 구면 조화 계수들 또는 그의 분해들보다 더 작은 값들을 갖는다.

[0283] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 감소된 구면 조화 계수들 또는 감소된 그의 분해들을 생성하기 위해 특정된 오더 및/또는 서브-오더를 갖는 복수의 구면 조화 계수들의 것들 또는 그의 분해들의 벡터들을 삭제하도록 구성된다.

[0284] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 감소된 구면 조화 계수들 또는 감소된 그의 분해들을 생성하기 위해 특정된 오더 및/또는 서브-오더를 갖는 복수의 구면 조화 계수들의 것들 또는 그의 분해들의 이들 벡터들을 제로 아웃하도록 구성된다.

[0285] 기술들의 다양한 양태들은 또한 오디오 인코딩 디바이스 (20) 가 음장의 구별되는 성분들을 나타내도록 구성되게 할 수도 있다. 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 음장의 구별되는 성분을 나타내는데 사용될 벡터의 계수들의 제 1 비-제로 세트를 획득하도록 구성되고, 여기서 벡터는 음장을 설명하는 복수의 구면 조화 계수들로부터 분해된다.

[0286] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 모든 계수들을 포함하도록 벡터의 계수들의 제 1 비-제로 세트를 결정하도록 구성된다.

[0287] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 복수의 구면 조화 계수들 중 하나 이상이 대응하는 기저 함수의 오더보다 더 큰 오더에 대응하는 계수들의 것들로서 계수들의 제 1 비-제로 세트를 결정하도록 구성된다.

[0288] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 복수의 구면 조화 계수들 중 하나 이상이 대응하는 기저 함수의 오더보다 더 큰 오더에 대응하는 계수들의 것들을 포함하도록 그리고 복수의 구면 조화 계수들 중 하나 이상이 대응하는 기저 함수의 오더보다 더 큰 오더에 대응하는 계수들 중 적어도 하나를 제외하도록 계수들의 제 1 비-제로 세트를 결정하도록 구성된다.

[0289] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 복수의 구면 조화 계수들 중 하나 이상이 대응하는 기저 함수의 오더보다 더 큰 오더에 대응하는 계수들 중 적어도 하나를 제외하고는 모든 계수들을 포함하도록 계수들의 제 1 비-제로 세트를 결정하도록 구성된다.

[0290] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 또한 사이드 채널 정보에서 벡터의 계수들의 제 1 비-제로 세트를 명시하도록 구성된다.

[0291] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 또한 벡터의 계수들의 제 1 비-제로 세트의 오디오 인코딩 없이 사이드 채널 정보에서 벡터의 계수들의 제 1 비-제로 세트를 명시하도록 구성된다.

[0292] 이들 및 다른 경우들에서, 벡터는 벡터 기반 합성을 사용하여 복수의 구면 조화 계수들로부터 분해된 벡터를 포함한다.

[0293] 이들 및 다른 경우들에서, 벡터 기반 합성은 특이값 분해를 포함한다.

[0294] 이들 및 다른 경우들에서, 벡터는 특이값 분해를 사용하여 복수의 구면 조화 계수들로부터 분해된 V 벡터를 포함한다.

[0295] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 또한 벡터의 계수들의 비-제로 세트를 명시하게 하는 복수의 구성 모드들 중 하나를 선택하고, 복수의 구성 모드들 중 선택된 하나에 기초하여 벡터의 계수들의 비-제로 세트를 명시하도록 구성된다.

[0296] 이들 및 다른 경우들에서, 복수의 구성 모드들 중 하나는 계수들의 비-제로 세트가 모든 계수들을 포함함을 표시한다.

[0297] 이들 및 다른 경우들에서, 복수의 구성 모드들 중 하나는, 계수들의 비-제로 세트가, 복수의 구면 조화 계수들 중 하나 이상이 대응하는 기저 함수의 오더보다 더 큰 오더에 대응하는 계수들의 것들을 포함함을 표시한다.

[0298] 이들 및 다른 경우들에서, 복수의 구성 모드들 중 하나는, 계수들의 비-제로 세트가, 복수의 구면 조화 계수들 중 하나 이상이 대응하는 기저 함수의 오더보다 더 큰 오더에 대응하는 계수들의 것들을 포함함 그리고 복수의 구면 조화 계수들 중 하나 이상이 대응하는 기저 함수의 오더보다 더 큰 오더에 대응하는 계수들 중 적어도 하나를 제외함을 표시한다.

[0299] 이들 및 다른 경우들에서, 복수의 구성 모드들 중 하나는, 계수들의 비-제로 세트가, 계수들 중 적어도 하나를 제외하고는 모든 계수들을 포함함을 표시한다.

[0300] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 또한 비트스트림에서 복수의 구성 모드들 중 선택된 하나를 명시하도록 구성된다.

[0301] 본 개시물에서 설명된 기술들의 다양한 양태들은 또한 오디오 인코딩 디바이스 (20) 가 다양한 방법으로 음장의 구별되는 성분을 나타내도록 구성되게 할 수도 있다. 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 음장의 구별되는 성분을 나타내는 벡터의 계수들의 제 1 비-제로 세트를 획득하도록 구성되고, 이 벡터는 음장을 설명하는 복수의 구면 조화 계수들로부터 분해되었다.

[0302] 이들 및 다른 경우들에서, 계수들의 제 1 비-제로 세트는 벡터의 계수들 모두를 포함한다.

[0303] 이들 및 다른 경우들에서, 계수들의 제 1 비-제로 세트는 복수의 구면 조화 계수들 중 하나 이상이 대응하는 기저 함수의 오더보다 더 큰 오더에 대응하는 계수들의 것들을 포함한다.

[0304] 이들 및 다른 경우들에서, 계수들의 제 1 비-제로 세트는 복수의 구면 조화 계수들 중 하나 이상이 대응하는 기저 함수의 오더보다 더 큰 오더에 대응하는 계수들의 것들을 포함하고 복수의 구면 조화 계수들 중 하나 이상이 대응하는 기저 함수의 오더보다 더 큰 오더에 대응하는 계수들 중 적어도 하나를 제외한다.

[0305] 이들 및 다른 경우들에서, 계수들의 제 1 비-제로 세트는 충분한 방향 정보를 갖지 않는 것으로서 식별된 계수들 중 적어도 하나를 제외하고는 모든 계수들을 포함한다.

[0306] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 또한 벡터의 제 1 부분으로서 계수들의 제 1 비-제로 세트를 추출하도록 구성된다.

[0307] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 또한 사이드 채널 정보로부터 벡터의 제 1 비-제로 세트를 추출하고, 벡터의 계수들의 제 1 비-제로 세트에 기초하여 복수의 구면 조화 계수들의 재구성된 버전을 획득하도록 구성된다.

[0308] 이들 및 다른 경우들에서, 벡터는 벡터 기반 합성을 사용하여 복수의 구면 조화 계수들로부터 분해된 벡터를 포함한다.

[0309] 이들 및 다른 경우들에서, 벡터 기반 합성은 특이값 분해를 포함한다.

[0310] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 또한 복수의 구성 모드들 중 하나를 결정하여 그 복수의 구성 모드들 중 하나에 따라 벡터의 계수들의 비-제로 세트를 추출하고, 복수의 구성 모드들 중 획득된 하나에 기초하여 벡터의 계수들의 비-제로 세트를 추출하도록 구성된다.

[0311] 이들 및 다른 경우들에서, 복수의 구성 모드들 중 하나는 계수들의 비-제로 세트가 모든 계수들을 포함함을 표시한다.

[0312] 이들 및 다른 경우들에서, 복수의 구성 모드들 중 하나는, 계수들의 비-제로 세트가, 복수의 구면 조화 계수들 중 하나 이상이 대응하는 기저 함수의 오더보다 더 큰 오더에 대응하는 계수들의 것들을 포함함을 표시한다.

[0313] 이들 및 다른 경우들에서, 복수의 구성 모드들 중 하나는, 계수들의 비-제로 세트가, 복수의 구면 조화 계수들 중 하나 이상이 대응하는 기저 함수의 오더보다 더 큰 오더에 대응하는 계수들의 것들을 포함함 그리고 복수의 구면 조화 계수들 중 하나 이상이 대응하는 기저 함수의 오더보다 더 큰 오더에 대응하는 계수들 중 적어도 하나를 제외함을 표시한다.

[0314] 이들 및 다른 경우들에서, 복수의 구성 모드들 중 하나는, 계수들의 비-제로 세트가, 계수들 중 적어도 하나를 제외하고는 모든 계수들을 포함함을 표시한다.

[0315] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 비트스트림에서 시그널링된 값에 기초하여 복수의 구성 모드들 중 하나를 결정하도록 구성된다.

[0316] 기술들의 다양한 양태들은 또한, 일부 경우들에서, 오디오 인코딩 디바이스 (20) 를 인에이블시켜서 하나 이상의 구별되는 오디오 오브젝트들 (또는, 다시 말해서, 우월한 오디오 오브젝트들) 을 식별할 수도 있다. 일부 경우들에서, 오디오 인코딩 디바이스 (20) 는 오디오 오브젝트들 중 하나 이상에 대해 결정된 방향성에 기초하여 오디오 오브젝트들과 연관된 하나 이상의 구면 조화 계수들 (SHC) 로부터 하나 이상의 구별되는 오디오 오

브젝트들을 식별하도록 구성될 수도 있다.

[0317] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 또한 오디오 오브젝트들과 연관된 구면 조화 계수들에 기초하여 하나 이상의 오디오 오브젝트들의 방향성을 결정하도록 구성된다.

[0318] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 또한 구면 조화 계수들에 대해 특이값 분해를 수행하여 복수의 구면 조화 계수들의 좌-특이 벡터들을 나타내는 U 행렬, 복수의 구면 조화 계수들의 특이값들을 나타내는 S 행렬 및 복수의 구면 조화 계수들의 우-특이 벡터들을 나타내는 V 행렬을 생성하고, U 행렬, S 행렬 및 V 행렬 중 하나 이상의 행렬의 적어도 일부분의 행렬의 함수로서 복수의 구면 조화 계수들을 나타내도록 구성되고, 여기서 오디오 인코딩 디바이스 (20) 는 하나 이상의 오디오 오브젝트들의 각각의 방향성이 V 행렬에 적어도 부분적으로 기초하는 것을 결정하도록 구성된다.

[0319] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 또한 V 행렬의 하나 이상의 벡터들을 리오더링하여 리오더링된 V 행렬에서 보다 큰 방향성 지수를 갖는 벡터들이 보다 작은 방향성 지수를 갖는 벡터들보다 위에 위치되도록 구성된다.

[0320] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 또한 보다 큰 방향성 지수를 갖는 벡터들이 보다 작은 방향성 지수를 갖는 벡터들보다 더 큰 방향 정보를 포함하는 것을 결정하도록 구성된다.

[0321] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 또한 V 행렬을 S 행렬에 의해 곱하여 VS 행렬을 생성하도록 구성되고, 이 VS 행렬은 하나 이상의 벡터들을 포함한다.

[0322] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 또한 14 보다 큰 오더와 연관된 VS 행렬의 각각의 로우의 엔트리들을 선택하고, 선택된 엔트리들 각각을 제곱하여 대응하는 제곱한 엔트리들을 형성하며, VS 행렬의 각각의 로우에 대해, 제곱한 엔트리들 모두를 합산하여 대응하는 벡터에 대한 방향성 지수를 결정하도록 구성된다.

[0323] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는, 14 보다 큰 오더와 연관된 VS 행렬의 각각의 로우의 엔트리들을 선택하는 것이, VS 행렬의 각각의 로우의 18 번째 엔트리에서 시작하고 VS 행렬의 각각의 로우의 38 번째 엔트리에서 끝나는 모든 엔트리들을 선택하는 것을 포함하도록 구성된다.

[0324] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 또한 구별되는 오디오 오브젝트들을 나타내도록 VS 행렬의 벡터들의 서브세트를 선택하도록 구성된다. 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 VS 행렬의 4 개의 벡터들을 선택하도록 구성되고, 여기서 선택된 4 개의 벡터들은 VS 행렬의 벡터들 모두의 4 개의 최대 방향성 지수들을 갖는다.

[0325] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 구별되는 오디오 오브젝트들을 나타내도록 선택된 벡터들의 서브세트가 각각의 벡터의 에너지 및 방향성 양쪽에 기초하는 것을 결정하도록 구성된다.

[0326] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 또한 구별되는 오디오 오브젝트들을 나타내는 하나 이상의 제 1 벡터들과 하나 이상의 제 2 벡터들 사이의 에너지 비교를 수행하여 리오더링된 하나 이상의 제 1 벡터들을 결정하도록 구성되고, 여기서 하나 이상의 제 1 벡터들은 오디오 데이터의 제 1 부분에서 구별되는 오디오 오브젝트들을 설명하고 하나 이상의 제 2 벡터들은 오디오 데이터의 제 2 부분에서 구별되는 오디오 오브젝트들을 설명한다.

[0327] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 또한 구별되는 오디오 오브젝트들을 나타내는 하나 이상의 제 1 벡터들과 하나 이상의 제 2 벡터들 사이의 상호 상관을 수행하여 리오더링된 하나 이상의 제 1 벡터들을 결정하도록 구성되고, 여기서 하나 이상의 제 1 벡터들은 오디오 데이터의 제 1 부분에서 구별되는 오디오 오브젝트들을 설명하고 하나 이상의 제 2 벡터들은 오디오 데이터의 제 2 부분에서 구별되는 오디오 오브젝트들을 설명한다.

[0328] 기술들의 다양한 양태들은 또한, 일부 경우들에서, 오디오 인코딩 디바이스 (20) 를 인에이블시켜서 HOA 계수들 (11) 의 분해들에 대해 에너지 보상을 수행하도록 구성될 수도 있다. 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 복수의 구면 조화 계수들에 대해 벡터-기반 합성을 수행하여 하나 이상의 오디오 오브젝트들 및 대응하는 방향 정보를 나타내는 복수의 구면 조화 계수들의 분해된 표현들을 생성하도록 구성될 수도 있고, 여기서 구면 조화 계수들은 오더와 연관되고 음장을 설명하고, 방향 정보로부터 구별되는 그리고 배경 방향 정보를 결정하고, 배경 오디오 오브젝트들과 연관된 방향 정보의 오더를 감소시켜서 변환된 배경 방향 정보

를 생성하고, 변환된 방향 정보의 값들을 증가시키기 위한 보상을 적용하여 음장의 전체 에너지를 보존한다.

[0329] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 복수의 구면 조화 계수들에 대해 특이값 분해를 수행하여 오디오 오브젝트들을 나타내는 U 행렬 및 S 행렬 및 방향 정보를 나타내는 V 행렬을 생성하고, V 행렬의 구별되는 컬럼 벡터들 및 V 행렬의 배경 컬럼 벡터들을 결정하고, V 행렬의 배경 컬럼 벡터들의 오더를 감소시켜서 V 행렬의 변환된 배경 컬럼 벡터들을 생성하며, V 행렬의 변환된 배경 컬럼 벡터들의 값들을 증가시키기 위한 보상을 적용하여 음장의 전체 에너지를 보존하도록 구성될 수도 있다.

[0330] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 또한 S 행렬의 돌출 특이값들의 수를 결정하도록 구성되고, 여기서 V 행렬의 구별되는 컬럼 벡터들의 수는 S 행렬의 돌출 특이값들의 수이다.

[0331] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 구면 조화 계수들에 대한 감소된 오더, 및 감소된 오더보다 더 큰 오더와 연관된 V 행렬의 배경 컬럼 벡터들의 로우들에 대한 제로 값들을 결정하도록 구성된다.

[0332] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 또한 U 행렬의 배경 컬럼들, S 행렬의 배경 컬럼들, 및 V 행렬의 변환된 배경 컬럼들의 트랜스포즈를 조합하여 변경된 구면 조화 계수들을 생성하도록 구성된다.

[0333] 이들 및 다른 경우들에서, 변경된 구면 조화 계수들은 음장의 하나 이상의 배경 성분들을 설명한다.

[0334] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 V 행렬의 배경 컬럼 벡터들의 벡터의 제 1 에너지 및 V 행렬의 변환된 배경 컬럼 벡터들의 벡터의 제 2 에너지를 결정하고, V 행렬의 변환된 배경 컬럼 벡터들의 벡터의 각각의 엘리먼트에 증폭 값을 적용하도록 구성되고, 여기서 증폭 값은 제 1 에너지 대 제 2 에너지의 비율을 포함한다.

[0335] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 V 행렬의 배경 컬럼 벡터들의 벡터의 제 1 의 평균 제곱한 에너지 및 V 행렬의 변환된 배경 컬럼 벡터들의 벡터의 제 2 의 평균 제곱한 에너지를 결정하고, V 행렬의 변환된 배경 컬럼 벡터들의 벡터의 각각의 엘리먼트에 증폭 값을 적용하도록 구성되고, 여기서 증폭 값은 제 1 에너지 대 제 2 에너지의 비율을 포함한다.

[0336] 본 개시물에서 설명된 기술들의 다양한 양태들은 또한 오디오 인코딩 디바이스 (20) 를 인에이블시켜서 HOA 계수들 (11) 의 분해된 버전들에 대해 보간을 수행할 수도 있다. 일부 경우들에서, 오디오 인코딩 디바이스 (20) 는, 적어도 부분적으로, 복수의 제 1 구면 조화 계수들의 제 1 분해 및 복수의 제 2 구면 조화 계수들의 제 2 분해에 대해 보간을 수행함으로써 시간 세그먼트에 대해 분해된 보간된 구면 조화 계수들을 획득하도록 구성될 수도 있다.

[0337] 이들 및 다른 경우들에서, 제 1 분해는 복수의 제 1 구면 조화 계수들의 우-특이 벡터들을 나타내는 제 1 V 행렬을 포함한다.

[0338] 이들 및 다른 경우들에서, 제 2 분해는 복수의 제 2 구면 조화 계수들의 우-특이 벡터들을 나타내는 제 2 V 행렬을 포함한다.

[0339] 이들 및 다른 경우들에서, 제 1 분해는 복수의 제 1 구면 조화 계수들의 우-특이 벡터들을 나타내는 제 1 V 행렬을 포함하고, 제 2 분해는 복수의 제 2 구면 조화 계수들의 우-특이 벡터들을 나타내는 제 2 V 행렬을 포함한다.

[0340] 이들 및 다른 경우들에서, 시간 세그먼트는 오디오 프레임의 서브-프레임을 포함한다.

[0341] 이들 및 다른 경우들에서, 시간 세그먼트는 오디오 프레임의 시간 샘플을 포함한다.

[0342] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 복수의 제 1 구면 조화 계수들의 구면 조화 계수에 대한 제 1 분해 및 제 2 분해의 보간된 분해를 획득하도록 구성된다.

[0343] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 제 1 프레임에 포함된 복수의 제 1 구면 조화 계수들의 제 1 부분에 대한 제 1 분해 및 제 2 프레임에 포함된 복수의 제 2 구면 조화 계수들의 제 2 부분에 대한 제 2 분해의 보간된 분해들을 획득하도록 구성되고, 오디오 인코딩 디바이스 (20) 는 또한 제 1 프레임에 포함된 복수의 제 1 구면 조화 계수들의 제 1 부분의 제 1 시간 성분에 보간된 분해들을 적용하여 복수의 제 1 구면 조화 계수들의 제 1 인공 시간 성분을 생성하고, 제 2 프레임에 포함된 복수의 제 2 구면 조화 계수들의 제 2 부분의 제 2 시간 성분에 각각의 보간된 분해들을 적용하여 포함된 복수의 제 2 구면 조화 계수들의 제 2 인공 시간 성분을 생성하도록 구성된다.

[0344] 이들 및 다른 경우들에서, 제 1 시간 성분은 복수의 제 1 구면 조화 계수들에 대해 벡터-기반 합성을 수행함으로써 생성된다.

[0345] 이들 및 다른 경우들에서, 제 2 시간 성분은 복수의 제 2 구면 조화 계수들에 대해 벡터-기반 합성을 수행함으로써 생성된다.

[0346] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 또한 제 1 인공 시간 성분 및 제 2 인공 시간 성분을 수신하고, 복수의 제 1 구면 조화 계수들의 제 1 부분에 대한 제 1 분해 및 복수의 제 2 구면 조화 계수들의 제 2 부분에 대한 제 2 분해의 보간된 분해들을 컴퓨팅하며, 보간된 분해들의 역들을 제 1 인공 시간 성분에 적용하여 제 1 시간 성분을 복원하고 제 2 인공 시간 성분에 적용하여 제 2 시간 성분을 복원하도록 구성된다.

[0347] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 복수의 제 1 구면 조화 계수들의 제 1 공간 성분 및 복수의 제 2 구면 조화 계수들의 제 2 공간 성분을 보간하도록 구성된다.

[0348] 이들 및 다른 경우들에서, 제 1 공간 성분은 복수의 제 1 구면 조화 계수들의 좌-특이 벡터들을 나타내는 제 1 U 행렬을 포함한다.

[0349] 이들 및 다른 경우들에서, 제 2 공간 성분은 복수의 제 2 구면 조화 계수들의 좌-특이 벡터들을 나타내는 제 2 U 행렬을 포함한다.

[0350] 이들 및 다른 경우들에서, 제 1 공간 성분은 복수의 제 1 구면 조화 계수들에 대한 구면 조화 계수들의 M 개의 시간 세그먼트들을 나타내고 제 2 공간 성분은 복수의 제 2 구면 조화 계수들에 대한 구면 조화 계수들의 M 개의 시간 세그먼트들을 나타낸다.

[0351] 이들 및 다른 경우들에서, 제 1 공간 성분은 복수의 제 1 구면 조화 계수들에 대한 구면 조화 계수들의 M 개의 시간 세그먼트들을 나타내고 제 2 공간 성분은 복수의 제 2 구면 조화 계수들에 대한 구면 조화 계수들의 M 개의 시간 세그먼트들을 나타내며, 오디오 인코딩 디바이스 (20) 는 제 1 공간 성분의 마지막 N 개의 엘리먼트들 및 제 2 공간 성분의 처음 N 개의 엘리먼트들을 보간하도록 구성된다.

[0352] 이들 및 다른 경우들에서, 복수의 제 2 구면 조화 계수들은 시간 도메인에서 복수의 제 1 구면 조화 계수들에 후속한다.

[0353] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 또한 복수의 제 1 구면 조화 계수들을 분해하여 복수의 제 1 구면 조화 계수들의 제 1 분해를 생성하도록 구성된다.

[0354] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 또한 복수의 제 2 구면 조화 계수들을 분해하여 복수의 제 2 구면 조화 계수들의 제 2 분해를 생성하도록 구성된다.

[0355] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 또한 복수의 제 1 구면 조화 계수들에 대해 특이값 분해를 수행하여 복수의 제 1 구면 조화 계수들의 좌-특이 벡터들을 나타내는 U 행렬, 복수의 제 1 구면 조화 계수들의 특이값들을 나타내는 S 행렬 및 복수의 제 1 구면 조화 계수들의 우-특이 벡터들을 나타내는 V 행렬을 생성하도록 구성된다.

[0356] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 또한 복수의 제 2 구면 조화 계수들에 대해 특이값 분해를 수행하여 복수의 제 2 구면 조화 계수들의 좌-특이 벡터들을 나타내는 U 행렬, 복수의 제 2 구면 조화 계수들의 특이값들을 나타내는 S 행렬 및 복수의 제 2 구면 조화 계수들의 우-특이 벡터들을 나타내는 V 행렬을 생성하도록 구성된다.

[0357] 이들 및 다른 경우들에서, 복수의 제 1 및 제 2 구면 조화 계수들 각각은 음장의 평면파 표현을 나타낸다.

[0358] 이들 및 다른 경우들에서, 복수의 제 1 및 제 2 구면 조화 계수들 각각은 함께 믹싱된 하나 이상의 모노-오디오 오브젝트들을 나타낸다.

[0359] 이들 및 다른 경우들에서, 복수의 제 1 및 제 2 구면 조화 계수들 각각은 3 차원 음장을 나타내는 각각이 제 1 및 제 2 구면 조화 계수들을 포함한다.

[0360] 이들 및 다른 경우들에서, 복수의 제 1 및 제 2 구면 조화 계수들은 각각 1 보다 더 큰 오더를 갖는 적어도 하나의 구면 기저 함수와 연관된다.

[0361] 이들 및 다른 경우들에서, 복수의 제 1 및 제 2 구면 조화 계수들은 각각 4 와 동일한 오더를 갖는 적어도 하나

의 구면 기저 함수와 연관된다.

[0362] 이들 및 다른 경우들에서, 보간은 제 1 분해 및 제 2 분해의 가중된 보간이고, 여기서 제 1 분해에 적용된 가중된 보간의 가중치들은 제 1 및 제 2 분해의 벡터들에 의해 나타낸 시간에 반비례하고 여기서 제 2 분해에 적용된 가중된 보간의 가중치들은 제 1 및 제 2 분해의 벡터들에 의해 나타낸 시간에 비례한다.

[0363] 이들 및 다른 경우들에서, 분해된 보간된 구면 조화 계수들은 복수의 제 1 구면 조화 계수들 및 복수의 제 2 구면 조화 계수들의 공간 성분들 및 시간 성분들 중 적어도 하나를 평활화한다.

[0364] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 Us[n] = HOA(n) * (V_vec[n])-1 을 컴퓨팅하여 스칼라를 획득하도록 구성된다.

[0365] 이들 및 다른 경우들에서, 보간은 선형 보간을 포함한다. 이들 및 다른 경우들에서, 보간은 비-선형 보간을 포함한다. 이들 및 다른 경우들에서, 보간은 코사인 보간을 포함한다. 이들 및 다른 경우들에서, 보간은 가중된 코사인 보간을 포함한다. 이들 및 다른 경우들에서, 보간은 큐빅 보간을 포함한다. 이들 및 다른 경우들에서, 보간은 적응적 스플라인 보간을 포함한다. 이들 및 다른 경우들에서, 보간은 최소 곡률 보간을 포함한다.

[0366] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 또한 시간 세그먼트에 대한 분해된 보간된 구면 조화 계수들의 표현, 및 보간의 타입의 표시를 포함하는 비트스트림을 생성하도록 구성된다.

[0367] 이들 및 다른 경우들에서, 이 표시는 보간의 타입에 맵핑하는 하나 이상의 비트들을 포함한다.

[0368] 이러한 방법으로, 본 개시물에서 설명된 기술들의 다양한 양태들은 오디오 인코딩 디바이스 (20) 를 인에이블시켜서 시간 세그먼트에 대한 분해된 보간된 구면 조화 계수들의 표현, 및 보간의 타입의 표시를 포함하는 비트스트림을 획득하도록 구성될 수도 있다.

[0369] 이들 및 다른 경우들에서, 이 표시는 보간의 타입에 맵핑하는 하나 이상의 비트들을 포함한다.

[0370] 이와 관련하여, 오디오 인코딩 디바이스 (20) 는, 오디오 인코딩 디바이스 (20) 가, 일부 경우들에서, 음장의 공간 성분의 압축된 버전을 포함하는 비트스트림을 생성하도록 구성될 수도 있는 기술들의 하나의 실시형태를 나타낼 수도 있고, 공간 성분은 복수의 구면 조화 계수들에 대해 벡터 기반 합성을 수행함으로써 생성된다.

[0371] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 또한 공간 성분을 압축할 때 예측 모드를 명시하는 필드를 포함하도록 비트스트림을 생성하도록 구성된다.

[0372] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 공간 성분을 압축할 때 허프만 테이블을 명시하는 허프만 테이블 정보를 포함하도록 비트스트림을 생성하도록 구성된다.

[0373] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 공간 성분을 압축할 때 양자화 스텝 사이즈 또는 그의 변수를 표현하는 값을 표시하는 필드를 포함하도록 비트스트림을 생성하도록 구성된다.

[0374] 이들 및 다른 경우들에서, 값은 *nbits* 값을 포함한다.

[0375] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 공간 성분의 압축된 버전이 포함된 음장의 복수의 공간 성분들의 압축된 버전을 포함하도록 비트스트림을 생성하도록 구성되고, 여기서 값은 복수의 공간 성분들을 압축할 때 사용된 양자화 스텝 사이즈 또는 그의 변수를 나타낸다.

[0376] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 또한 공간 성분이 대응하는 압축 카테고리를 식별하는 카테고리 식별자를 나타내기 위한 허프만 코드를 포함하도록 비트스트림을 생성하도록 구성된다.

[0377] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 공간 성분이 포지티브 값인지 또는 네거티브 값인지 여부를 식별하는 부호 비트를 포함하도록 비트스트림을 생성하도록 구성된다.

[0378] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 공간 성분의 잔차 값을 나타내기 위한 허프만 코드를 포함하도록 비트스트림을 생성하도록 구성된다.

[0379] 이들 및 다른 경우들에서, 벡터 기반 합성은 특이값 분해를 포함한다.

[0380] 이와 관련하여, 오디오 인코딩 디바이스 (20) 는 또한, 오디오 인코딩 디바이스 (20) 가, 일부 경우들에서, 복수의 공간 성분들 중 나머지 것들에 대한 공간 성분의 오더에 기초하여 복수의 공간 성분들 중의 공간 성분을 압축할 때 사용할 허프만 코드북을 식별하도록 구성될 수도 있는 기술들의 다양한 양태들을 시행할 수도 있고,

공간 성분은 복수의 구면 조화 계수들에 대해 벡터 기반 합성을 수행함으로써 생성된다.

[0381] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 공간 성분을 압축할 때 사용된 예측 모드에 기초하여 허프만 코드북을 식별하도록 구성된다.

[0382] 이들 및 다른 경우들에서, 공간 성분의 압축된 버전은 허프만 코드북을 식별하는 허프만 테이블 정보를 적어도 부분적으로 사용하여 비트스트림에서 나타낸다.

[0383] 이들 및 다른 경우들에서, 공간 성분의 압축된 버전은 공간 성분을 압축할 때 사용된 양자화 스텝 사이즈 또는 그의 변수를 표현하는 값을 표시하는 필드를 적어도 부분적으로 사용하여 비트스트림에서 나타낸다.

[0384] 이들 및 다른 경우들에서, 값은 *nbits* 값을 포함한다.

[0385] 이들 및 다른 경우들에서, 비트스트림은 공간 성분의 압축된 버전이 포함된 음장의 복수의 공간 성분들의 압축된 버전을 포함하고, 이 값은 복수의 공간 성분들을 압축할 때 사용된 양자화 스텝 사이즈 또는 그의 변수를 나타낸다.

[0386] 이들 및 다른 경우들에서, 공간 성분의 압축된 버전은 공간 성분이 대응하는 압축 카테고리를 식별하는 카테고리 식별자를 나타내기 위해 식별된 허프만 코드북으로부터 선택된 허프만 코드를 적어도 부분적으로 사용하여 비트스트림에서 나타낸다.

[0387] 이들 및 다른 경우들에서, 공간 성분의 압축된 버전은 공간 성분이 포지티브 값인지 또는 네거티브 값인지 여부를 식별하는 부호 비트를 적어도 부분적으로 사용하여 비트스트림에서 나타낸다.

[0388] 이들 및 다른 경우들에서, 공간 성분의 압축된 버전은 공간 성분의 잔차 값을 나타내기 위해 식별된 허프만 코드북으로부터 선택된 허프만 코드를 적어도 부분적으로 사용하여 비트스트림에서 나타낸다.

[0389] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 또한 식별된 허프만 코드북에 기초하여 공간 성분을 압축하여 공간 성분의 압축된 버전을 생성하고, 공간 성분의 압축된 버전을 포함하도록 비트스트림을 생성하도록 구성된다.

[0390] 게다가, 오디오 인코딩 디바이스 (20) 는, 일부 경우들에서, 오디오 인코딩 디바이스 (20) 가 음장의 공간 성분을 압축할 때 사용될 양자화 스텝 사이즈를 결정하도록 구성될 수도 있는 기술들의 다양한 양태들을 시행할 수도 있고, 공간 성분은 복수의 구면 조화 계수들에 대해 벡터 기반 합성을 수행함으로써 생성된다.

[0391] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 또한 타겟 비트 레이트에 기초하여 양자화 스텝 사이즈를 결정하도록 구성된다.

[0392] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 공간 성분을 나타내는데 사용되는 비트들의 수의 추정치를 결정하고, 그 추정치와 타겟 비트 레이트 사이의 차이에 기초하여 양자화 스텝 사이즈를 결정하도록 구성된다.

[0393] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 공간 성분을 나타내는데 사용되는 비트들의 수의 추정치를 결정하고, 그 추정치와 타겟 비트 레이트 사이의 차이를 결정하며, 타겟 비트 레이트에 차이를 가산함으로써 양자화 스텝 사이즈를 결정하도록 구성된다.

[0394] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 타겟 비트 레이트에 대응하는 코드 북이 주어진 공간 성분에 대해 발생되어야 하는 비트들의 수의 추정치를 계산하도록 구성된다.

[0395] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 공간 성분을 압축할 때 사용된 코딩 모드가 주어진 공간 성분에 대해 발생되어야 하는 비트들의 수의 추정치를 계산하도록 구성된다.

[0396] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 공간 성분을 압축할 때 사용될 제 1 코딩 모드가 주어진 공간 성분에 대해 발생되어야 하는 비트들의 수의 제 1 추정치를 계산하고, 공간 성분을 압축할 때 사용될 제 2 코딩 모드가 주어진 공간 성분에 대해 발생되어야 하는 비트들의 수의 제 2 추정치를 계산하며, 비트들의 수의 결정된 추정치로서 사용될 비트들의 최소 수를 갖는, 제 1 추정치와 제 2 추정치 중 하나를 선택하도록 구성된다.

[0397] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 공간 성분이 대응하는 카테고리를 식별하는 카테고리 식별자를 식별하고, 카테고리에 대응하는 공간 성분을 압축할 때 발생하는 공간 성분에 대한 잔차 값의 비트 길이를 식별하며, 적어도 부분적으로, 잔차 값의 비트 길이에 카테고리 식별자를 나타내는데 사용된 비트들의

수를 가산함으로써 비트들의 수의 추정치를 결정하도록 구성된다.

[0398] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 또한 공간 성분을 압축할 때 사용될 복수의 코드 북들 중 하나를 선택하도록 구성된다.

[0399] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 또한 복수의 코드 북들 각각을 사용하여 공간 성분을 나타내는데 사용된 비트들의 수의 추정치를 결정하고, 비트들의 최소 수를 갖는 결정된 추정치를 발생시킨 복수의 코드 북들 중 하나를 선택하도록 구성된다.

[0400] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 또한 복수의 코드 북들 중 하나 이상을 사용하여 공간 성분을 나타내는데 사용된 비트들의 수의 추정치를 결정하도록 구성되고, 복수의 코드 북들 중 하나 이상은 공간 성분의 다른 엘리먼트에 대해 압축될 공간 성분의 엘리먼트들의 오더에 기초하여 선택된다.

[0401] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 또한 공간 성분이 후속 공간 성분로부터 예측되지 않을 때 사용되도록 설계된 복수의 코드 북들 중 하나를 사용하여 공간 성분을 나타내는데 사용된 비트들의 수의 추정치를 결정하도록 구성된다.

[0402] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 또한 공간 성분이 후속 공간 성분로부터 예측될 때 사용되도록 설계된 복수의 코드 북들 중 하나를 사용하여 공간 성분을 나타내는데 사용된 비트들의 수의 추정치를 결정하도록 구성된다.

[0403] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 또한 공간 성분이 음장에서 합성 오디오 오브젝트를 나타낼 때 사용되도록 설계된 복수의 코드 북들 중 하나를 사용하여 공간 성분을 나타내는데 사용된 비트들의 수의 추정치를 결정하도록 구성된다.

[0404] 이들 및 다른 경우들에서, 합성 오디오 오브젝트는 펄스 코드 변조된 (PCM) 오디오 오브젝트를 포함한다.

[0405] 이들 및 다른 경우들에서, 오디오 인코딩 디바이스 (20) 는 또한 공간 성분이 음장에서 레코딩된 오디오 오브젝트를 나타낼 때 사용되도록 설계된 복수의 코드 북들 중 하나를 사용하여 공간 성분을 나타내는데 사용된 비트들의 수의 추정치를 결정하도록 구성된다.

[0406] 전술된 다양한 경우들 각각에서, 오디오 인코딩 디바이스 (20) 는 오디오 인코딩 디바이스 (20) 가 수행하도록 구성된 방법을 수행할 수도 있고 또는 그렇지 않으면 방법의 각각의 단계를 수행하는 수단을 포함할 수도 있다는 것이 이해되어야 한다. 일부 경우들에서, 이들 수단은 하나 이상의 프로세서들을 포함할 수도 있다. 일부 경우들에서, 하나 이상의 프로세서들은 비-일시적 컴퓨터 판독가능 저장 매체에 저장된 명령들로 구성된 특수 목적 프로세서를 나타낼 수도 있다. 다시 말해서, 인코딩 예들의 세트들 각각에서의 기술들의 다양한 양태들은, 실행될 때, 오디오 인코딩 디바이스 (20) 가 수행하도록 구성된 방법을 하나 이상의 프로세서들이 수행하게 하는 명령들을 저장한 비-일시적 컴퓨터 판독가능 저장 매체를 제공할 수도 있다.

[0407] 도 5 는 도 3 의 오디오 디코딩 디바이스 (24) 를 더 상세하게 예시한 블록 다이어그램이다. 도 5 의 예에 도시된 바와 같이, 오디오 디코딩 디바이스 (24) 는 추출 유닛 (72), 방향-기반 복원 유닛 (90) 및 벡터-기반 복원 유닛 (92) 을 포함할 수도 있다.

[0408] 추출 유닛 (72) 은 비트스트림 (21) 을 수신하고 HOA 계수들 (11) 의 다양한 인코딩된 버전들 (예를 들어, 방향-기반 인코딩된 버전 또는 벡터-기반 인코딩된 버전) 을 추출하도록 구성된 유닛을 나타낼 수도 있다. 추출 유닛 (72) 은 HOA 계수들 (11) 이 다양한 버전들을 통해 인코딩되었는지 여부를 전술된 신택스 엘리먼트 (예를 들어, 도 10e 및 도 10h(i) 내지 도 10o(ii) 의 예에서 도시된 채널 타입 신택스 엘리먼트) 로부터 결정할 수도 있다. 방향-기반 인코딩이 수행되었을 때, 추출 유닛 (72) 은 HOA 계수들 (11) 의 방향-기반 버전 및 이 인코딩된 버전과 연관된 신택스 엘리먼트들 (도 5 의 예에서 방향-기반 정보 (91) 로서 표기됨) 을 추출하여, 이 방향 기반 정보 (91) 를 방향-기반 복원 유닛 (90) 에 전달할 수도 있다. 이 방향-기반 복원 유닛 (90) 은 방향-기반 정보 (91) 에 기초하여 HOA 계수들 (11') 의 형태로 HOA 계수들을 복원하도록 구성된 유닛을 나타낼 수도 있다. 비트스트림 및 이 비트스트림 내의 신택스 엘리먼트들의 배열은 도 10 내지 도 10o(ii) 및 도 11 의 예와 관련하여 더 상세하게 아래에 설명된다.

[0409] 신택스 엘리먼트가 HOA 계수들 (11) 이 벡터-기반 합성을 사용하여 인코딩되었음을 표시할 때, 추출 유닛 (72) 은 코딩된 전경 V[k] 벡터들 (57), 인코딩된 주변 HOA 계수들 (59) 및 인코딩된 nFG 신호들 (59) 을 추출할 수도 있다. 추출 유닛 (72) 은 코딩된 전경 V[k] 벡터들 (57) 을 양자화 유닛 (74) 에 전달하고 인코딩된 nFG

신호들 (61) 과 함께 인코딩된 주변 HOA 계수들 (59) 을 심리 음향 디코딩 유닛 (80) 에 전달할 수도 있다.

[0410] 코딩된 전경 V[k] 벡터들 (57), 인코딩된 주변 HOA 계수들 (59) 및 인코딩된 nFG 신호들 (59) 을 추출하기 위해, 추출 유닛 (72) 은 codedVVecLength 로 표기된 신택스 엘리먼트를 포함하는 사이드 채널 정보 (57) 를 획득할 수도 있다. 추출 유닛 (72) 은 사이드 채널 정보 (57) 로부터의 codedVVecLength 를 파싱할 수도 있다. 추출 유닛 (72) 은 codedVVecLength 신택스 엘리먼트에 기초하여 전술된 구성 모드들 중 임의의 하나에서 동작하도록 구성될 수도 있다.

[0411] 추출 유닛 (72) 은 그 후에 사이드 채널 정보 (57) 로부터 감소된 전경 V[k] 벡터들 (55_k) 의 압축된 형태를 파싱하기 위해 구성 모드들 중 임의의 하나에 따라 동작할 수도 있다. 추출 유닛 (72) 은 VVectorData 에 대한 다음 신택스 테이블에 제시된 신택스를 갖는 다음 의사-코드에 제시된 스위치 스테이트먼트에 따라 동작할 수도 있다.

[0412]
```
switch CodedVVecLength{
       case 0:
              VVecLength = NumOfHoaCoeffs;
              for (m=0; m<VVecLength; ++m){
                     VVecCoeffId[m] = m;

              }
              break;
        case 1:
              VVecLength = NumOfHoaCoeffs - MinNumOfCoeffsForAmbHOA -
        NumOfContAddHoaChans;
        n = 0;
        for(m=MinNumOfCoeffsForAmbHOA; m<NumOfHoaCoeffs; ++m){
                     CoeffIdx = m+1;
              if(CoeffIdx isNotMemberOf ContAddHoaCoeff){
                     VVecCoeffId[n] = CoeffIdx-1;
                     n++;
              }
        }
              break;
  case 2:
              VVecLength = NumOfHoaCoeffs - MinNumOfCoeffsForAmbHOA;
              for (m=0; m< VVecLength; ++m){
                     VVecCoeffId[m] = m + MinNumOfCoeffsForAmbHOA;
              }
        break;
        case 3:
        VVecLength = NumOfHoaCoeffs - NumOfContAddHoaChans;
        n = 0;
        for(m=0; m<NumOfHoaCoeffs; ++m){
                     c = m+1;
              if(c isNotMemberOf ContAddHoaCoeff){
                     VVecCoeffId[n] = c-1;
                     n++;
              }
        }
  }
```

[0413]

신택스	비트들의 수	연상기호

VVectorData(i)
{
if (NbitsQ(k)[i] == 5){
for (m=0; m< VVecLength; ++m){
VVec[i][VVecCoeffId[m]](k) = (**VecVal** / 128.0) – 1.0;
}
elseif(NbitsQ(k)[i] >= 6){
for (m=0; m< VVecLength; ++m){
huffIdx = *huffSelect*(VVecCoeffId[m], PFlag[i], CbFlag[i]);
cid = *huffDecode*(NbitsQ[i], huffIdx, **huffVal**);
aVal[i][m] = 0.0;
if (cid > 0) {
aVal[i][m] = sgn = (**sgnVal** * 2) - 1;
if (cid > 1) {
aVal[i][m] = sgn * (2.0^(cid -1) + **intAddVal**);
}
}
VVec[i][VVecCoeffId[m]](k) = aVal[i][m] *(2^(16 – NbitsQ(k)[i])*aVal[i][m])/2^15;
if (PFlag(k)[i] == 1) {
VVec[i][VVecCoeffId[m]](k)+= VVec[i][VVecCoeffId[m]](k-1)
}
}
}
	8
	동적
	1
	cid-1
		uimsbf
		huffDecode
		bslbf
		uimsbf

[0414]

[0415] 상기 신택스 테이블에서, 4 개의 케이스들 (케이스 0-3) 을 갖는 제 1 스위치 스테이트먼트는 수의 관점에서 V^{T}_{DIST} 벡터 길이 (VVecLength) 및 계수들의 인덱스들 (VVecCoeffId) 을 결정하는 방식을 제공한다. 제 1 케이스, 케이스 0 은, V^{T}_{DIST} 벡터들에 대한 계수들 (NumOfHoaCoeffs) 전부가 지정되는 것을 나타낸다. 제 2 케이스, 케이스 1 은, MinNumOfCoeffsForAmbHOA 보다 큰 수에 대응하는 V^{T}_{DIST} 벡터의 이들 계수들 만이 지정된다는 것을 나타내고, 이것은 상기에서 $(N_{DIST}+1)^2 - (N_{BG}+1)^2$ 로서 지칭되는 것이 무엇인지를 가리킬 수도 있다.

추가로, ContAddAmbHoaChan 에서 식별된 이들 NumOfContAddAmbHoaChan 계수들은 차감된다. 리스트 ContAddAmbHoaChan 은 오더 MinAmbHoaOrder 를 초과하는 오더에 대응하는 추가의 채널들 (여기서, "채널들" 은 소정의 오더, 서브-오더 조합에 대응하는 특정 계수를 지칭함) 을 지정한다. 제 3 케이스, 케이스 2 는, MinNumOfCoeffsForAmbHOA 보다 큰 수에 대응하는 V^{T}_{DIST} 벡터의 이들 계수들이 지정된다는 것을 나타내고, 이것은 상기에서 $(N_{DIST}+1)^2 - (N_{BG}+1)^2$ 으로서 지칭되는 것이 무엇인지를 가리킬 수도 있다. 제 4 케이스, 케이스 3 은, NumOfContAddAmbHoaChan 에 의해 식별된 계수들을 제거한 후에 남겨진 V^{T}_{DIST} 벡터의 이들 계수들이 지정된다는 것을 나타낸다. VVecLength 뿐만 아니라 VVecCoeffId 리스트 양자 모두는 HOAFrame 상 내의 모든 VVectors 에 대해 유효하다.

[0416] 이 스위치 스테이트먼트 후에, 균일한 역양자화를 수행할지 여부의 결정은 NbitsQ (또는, 상기에서 *nbits* 로 표기됨) 에 의해 제어될 수도 있고, 이것이 5 와 동일하다면, 균일한 8 비트 스칼라 역양자화가 수행된다. 반대로, 6 이상의 NbitsQ 값은 허프만 디코딩의 적용을 초래할 수도 있다. 상기에서 지칭된 *cid* 값은 NbitsQ 값의 2 개의 최하위 비트들에 동일할 수도 있다. 전술된 예측 모드는 상기 신택스 테이블에서 PFlag 로서 표기되는 한편, HT info 비트는 상기 신택스 테이블에서 CbFlag 로서 표기된다. 나머지 신택스는, 디코딩이 전술된 것과 실질적으로 유사한 방식으로 어떻게 발행했는지를 지정한다. 전술된 다양한 케이스들 각각에 따르는 비트스트림 (21) 의 다양한 예들은 도 10h(i)-10o(ii) 에 대하여 이하에서 더 상세하게 설명된다.

[0417] 벡터-기반 복원 유닛 (92) 은 벡터-기반 합성 유닛 (27) 에 대하여 전술된 것들에 상반된 동작들을 수행하여

HOA 계수들 (11') 을 복원하도록 구성된 유닛을 나타낸다. 벡터 기반 복원 유닛 (92) 은 양자화 유닛 (74), 공간-시간 보간 유닛 (76), 전경 공식화 유닛 (78), 심리 음향 디코딩 유닛 (80), HOA 계수 공식화 유닛 (82) 및 리오더 유닛 (84) 을 포함할 수도 있다.

[0418] 양자화 유닛 (74) 은 코딩된 전경 V[k] 벡터들 (57) 을 역양자화하고 이에 의해 감소된 전경 V[k] 벡터들 (55_k) 을 생성하도록, 도 4 의 예에 도시된 양자화 유닛 (52) 에 상반된 방식으로 동작하도록 구성된 유닛을 나타낼 수도 있다. 역양자화 유닛 (74) 은, 일부 예들에서 양자화 유닛 (52) 에 대하여 전술된 것과 상반된 방식으로 스칼라 역양자화 및 엔트로피 디코딩의 형태를 수행할 수도 있다. 역양자화 유닛 (74) 은 감소된 V[k] 벡터들 (55_k) 을 리오더 유닛 (84) 으로 포워딩할 수도 있다.

[0419] 심리 음향 디코딩 유닛 (80) 은, 인코딩된 주변 HOA 계수들 (59) 및 인코딩된 nFG 신호들 (61) 을 디코딩하고 이에 의해 에너지 보상된 주변 HOA 계수들 (47') 및 보간된 nFG 신호들 (49')(이것은 또한, 보간된 nFG 오디오 오브젝트들 (49') 로서 지칭될 수도 있음) 을 생성하도록, 도 4 의 예에 도시된 심리 음향 오디오 코딩 유닛 (40) 에 상반된 방식으로 동작할 수도 있다. 심리 음향 디코딩 유닛 (80) 은 에너지 보상된 주변 HOA 계수들 (47') 을 HOA 계수 공식화 유닛 (82) 으로 그리고 nFG 신호들 (49') 을 리오더 유닛 (84) 으로 패스할 수도 있다.

[0420] 리오더 유닛 (84) 은 리오더 유닛 (34) 에 대하여 전술된 것과 유사한 상반된 방식으로 동작하도록 구성된 유닛을 나타낼 수도 있다. 리오더 유닛 (84) 은 HOA 계수들 (11) 의 전경 성분들의 원래 오더를 나타내는 신택스 엘리먼트들을 수신할 수도 있다. 리오더 유닛 (84) 은, 이들 리오더 신택스 엘리먼트들에 기초하여 보간된 nFG 신호들 (49') 및 감소된 전경 V[k] 벡터들 (55_k) 을 리오더링하여 리오더링된 nFG 신호들 (49") 및 리오더링된 전경 V[k] 벡터들 ($55_k'$) 을 생성할 수도 있다. 리오더 유닛 (84) 은 리오더링된 nFG 신호들 (49") 을 전경 공식화 유닛 (78) 으로 그리고 리오더링된 전경 V[k] 벡터들 ($55_k'$) 을 공간-시간 보간 유닛 (76) 으로 출력할 수도 있다.

[0421] 공간-시간 보간 유닛 (76) 은 공간-시간 보간 유닛 (50) 에 대하여 전술된 것과 유사한 방식으로 동작할 수도 있다. 공간-시간 보간 유닛 (76) 은 리오더링된 전경 V[k] 벡터들 ($55_k'$) 을 수신하고, 리오더링된 전경 V[k] 벡터들 ($55_k'$) 및 리오더링된 전경 V[k-1] 벡터들 ($55_{k-1}'$) 에 대하여 공간-시간 보간을 수행하여 보간된 전경 V[k] 벡터들 ($55_k"$) 을 생성할 수도 있다. 공간-시간 보간 유닛 (76) 은 보간된 전경 V[k] 벡터들 ($55_k"$) 을 전경 공식화 유닛 (78) 으로 포워딩할 수도 있다.

[0422] 전경 공식화 유닛 (78) 은 보간된 전경 V[k] 벡터들 ($55_k"$) 및 리오더링된 nFG 신호들 (49") 에 대하여 행렬 곱셈을 수행하여 전경 HOA 계수들 (65) 을 생성하도록 구성된 유닛을 나타낼 수도 있다. 전경 공식화 유닛 (78) 은 보간된 전경 V[k] 벡터들 ($55_k"$) 에 의한 리오더링된 nFG 신호들 (49") 의 행렬 곱셈을 수행할 수도 있다.

[0423] HOA 계수 공식화 유닛 (82) 은 전경 HOA 계수들 (65) 을 주변 HOA 채널들 (47') 에 추가하여 HOA 계수들 (11') 을 획득하도록 구성된 유닛을 나타낼 수도 있고, 여기서 주된 표기 (notation) 는, 이들 HOA 계수들 (11') 이 HOA 계수들 (11) 과 유사하지만 동일하지는 않을 수도 있다는 것을 반영한다. HOA 계수들 (11 과 11') 간의 차이들은 손실있는 송신 매체, 양자화 또는 다른 손실있는 동작들을 통한 송신으로 인한 손실을 초래할 수도 있다.

[0424] 이 방식에서, 기술들은 오디오 디코딩 디바이스, 예컨대 오디오 디코딩 디바이스 (24) 로 하여금 비트스트림으로부터, 양자화된 방향 정보, 인코딩된 전경 오디오 오브젝트, 및 인코딩된 주변 고차 앰비소닉스 (higher order ambisonic; HOA) 계수들을 결정하게 할 수도 있고, 여기서 양자화된 방향 정보 및 인코딩된 전경 오디오 오브젝트는 음장의 전경 성분을 설명하는 전경 HOA 계수들을 나타내고, 인코딩된 주변 HOA 계수들은 음장의 주변 성분을 설명하고, 양자화된 방향 정보를 역양자화하여 방향 정보를 생성하고, 방향 정보에 대하여 공간-시간 보간을 수행하여 보간된 방향 정보를 생성하고, 인코딩된 전경 오디오 오브젝트를 오디오 디코딩하여 전경 오디오 오브젝트를 생성하고 인코딩된 주변 HOA 계수들을 오디오 디코딩하여 주변 HOA 계수들을 생성하고, 보간된 방향 정보 및 전경 오디오 오브젝트의 함수로서 전경 HOA 계수들을 결정하며, 전경 HOA 계수들 및 주변 HOA 계수들의 함수로서 HOA 계수들을 결정한다.

[0425] 이 방식에서, 기술들의 다양한 양태들은 통일된 오디오 디코딩 디바이스 (24) 로 하여금, 2 개의 상이한 분해 방식들 간에 스위칭하게 할 수도 있다. 일부 경우들에서, 오디오 디코딩 디바이스 (24) 는, 합성 오디오 오브젝트로부터 음장을 나타내는 구면 조화 계수들의 압축된 버전이 생성되는지 여부의 표시에 기초하여 복수의 분해 방식들 중 하나를 선택하고, 복수의 분해 방식들 중 선택된 방식을 사용하여 구면 조화 계수들의 압축된 버전을 분해하도록 구성될 수도 있다. 이들 및 다른 경우들에서, 오디오 디코딩 디바이스 (24) 는 통합된 디코더를 포함한다.

[0426] 일부 경우들에서, 오디오 디코딩 디바이스 (24) 는, 음장을 나타내는 구면 조화 계수들이 합성 오디오 오브젝트로부터 생성되는지 여부의 표시를 획득하도록 구성될 수도 있다.

[0427] 이들 및 다른 경우들에서, 오디오 디코딩 디바이스 (24) 는 구면 조화 계수들의 압축된 버전을 저장하는 비트스트림으로부터 이 표시를 획득하도록 구성된다.

[0428] 이 방식에서, 기술들의 다양한 양태들은 오디오 디코딩 디바이스 (24) 로 하여금, 음장의 구별되는 배경 성분들을 설명하는 벡터들을 획득하게 할 수도 있다. 일부 경우들에서, 오디오 디코딩 디바이스 (24) 는 음장의 구별되는 성분들을 설명하는 하나 이상의 제 1 벡터들 및 음장의 배경 성분들을 설명하는 하나 이상의 제 2 벡터들을 결정하도록 구성될 수도 있고, 하나 이상의 제 1 벡터들 및 하나 이상의 제 2 벡터들은 복수의 구면 조화 계수들에 대하여 적어도 변환을 수행함으로써 생성된다.

[0429] 이들 및 다른 경우들에서, 오디오 디코딩 디바이스 (24) 에서, 변환은 복수의 구면 조화 계수들의 좌-특이 벡터들을 나타내는 U 행렬, 복수의 구면 조화 계수들의 특이값들을 나타내는 S 행렬 및 복수의 구면 조화 계수들의 우-특이 벡터들을 나타내는 V 행렬을 생성하는 특이값 분해를 포함한다.

[0430] 이들 및 다른 경우들에서, 오디오 디코딩 디바이스 (24) 에서, 하나 이상의 제 1 벡터들은, 오디오 인코딩 전에, S 행렬의 하나 이상의 S_{DIST} 벡터들을 U 행렬의 하나 이상의 오디오 인코딩된 U_{DIST} 벡터들에 곱함으로써 생성된 하나 이상의 오디오 인코딩된 $U_{DIST} * S_{DIST}$ 벡터들을 포함하고, 여기서 U 행렬 및 S 행렬은 복수의 구면 조화 계수들에 대하여 적어도 특이값 분해를 수행함으로써 생성된다.

[0431] 이들 및 다른 경우들에서, 오디오 디코딩 디바이스 (24) 는 또한, 하나 이상의 오디오 인코딩된 $U_{DIST} * S_{DIST}$ 벡터들을 오디오 디코딩하여 하나 이상의 오디오 인코딩된 $U_{DIST} * S_{DIST}$ 벡터들의 오디오 디코딩된 버전을 생성한다.

[0432] 이들 및 다른 경우들에서, 오디오 디코딩 디바이스 (24) 에서, 하나 이상의 제 1 벡터들은, 오디오 인코딩 전에, S 행렬의 하나 이상의 S_{DIST} 벡터들을 U 행렬의 하나 이상의 오디오 인코딩된 U_{DIST} 벡터들에 곱함으로써 생성된 하나 이상의 오디오 인코딩된 $U_{DIST} * S_{DIST}$ 벡터들, 및 V 행렬의 트랜스포즈의 하나 이상의 V^{T}_{DIST} 벡터들을 포함하고, 여기서 U 행렬 및 S 행렬 및 V 행렬은 복수의 구면 조화 계수들에 대하여 적어도 특이값 분해를 수행함으로써 생성된다.

[0433] 이들 및 다른 경우들에서, 오디오 디코딩 디바이스 (24) 는 또한, 하나 이상의 오디오 인코딩된 $U_{DIST} * S_{DIST}$ 벡터들을 오디오 디코딩하여 하나 이상의 오디오 인코딩된 $U_{DIST} * S_{DIST}$ 벡터들의 오디오 디코딩된 버전을 생성한다.

[0434] 이들 및 다른 경우들에서, 오디오 디코딩 디바이스 (24) 는 또한, $U_{DIST} * S_{DIST}$ 벡터들에 V^{T}_{DIST} 벡터들을 곱하여 음장의 구별되는 성분들을 나타내는 복수의 구면 조화함수들의 것들을 복구하도록 구성된다.

[0435] 이들 및 다른 경우들에서, 오디오 디코딩 디바이스 (24) 에서, 하나 이상의 제 2 벡터들은, 오디오 인코딩 전에, S 행렬 내에 포함된 SBG 벡터들을 그 후 V 행렬의 트랜스포즈 내에 포함된 V^{T}_{BG} 벡터들을 U 행렬 내에 포함된 U_{BG} 벡터들에 곱함으로써 생성된 하나 이상의 오디오 인코딩된 $U_{BG} * S_{BG} * V^{T}_{BG}$ 벡터들을 포함하고, 여기서 S 행렬, U 행렬 및 V 행렬 각각은 복수의 구면 조화 계수들에 대하여 적어도 특이값 분해를 수행함으로써 생성된다.

[0436] 이들 및 다른 경우들에서, 오디오 디코딩 디바이스 (24) 에서, 하나 이상의 제 2 벡터들은, 오디오 인코딩

전에, S 행렬 내에 포함된 S_{BG} 벡터들을 그 후 V 행렬의 트랜스포즈 내에 포함된 V^{T}_{BG} 벡터들을 U 행렬 내에 포함된 U_{BG} 벡터들에 곱함으로써 생성된 하나 이상의 오디오 인코딩된 $U_{BG} * S_{BG} * V^{T}_{BG}$ 벡터들을 포함하고, 여기서 S 행렬, U 행렬 및 V 행렬 각각은 복수의 구면 조화 계수들에 대하여 적어도 특이값 분해를 수행함으로써 생성되었고, 오디오 디코딩 디바이스 (24) 는 또한, 하나 이상의 오디오 인코딩된 $U_{BG} * S_{BG} * V^{T}_{BG}$ 벡터들을 오디오 디코딩하여 하나 이상의 오디오 디코딩된 $U_{BG} * S_{BG} * V^{T}_{BG}$ 벡터들을 생성하도록 구성된다.

[0437] 이들 및 다른 경우들에서, 오디오 디코딩 디바이스 (24) 에서, 하나 이상의 제 1 벡터들은, 오디오 인코딩 전에, S 행렬의 하나 이상의 S_{DIST} 벡터들을 U 행렬의 하나 이상의 오디오 인코딩된 U_{DIST} 벡터들에 곱함으로써 생성된 하나 이상의 오디오 인코딩된 $U_{DIST} * S_{DIST}$ 벡터들, 및 V 행렬의 트랜스포즈의 하나 이상의 V^{T}_{DIST} 벡터들을 포함하고, 여기서 U 행렬, S 행렬 및 V 행렬은 복수의 구면 조화 계수들에 대하여 적어도 특이값 분해를 수행함으로써 생성되었고, 여기서 오디오 디코딩 디바이스 (24) 는 또한, 하나 이상의 오디오 인코딩된 $U_{DIST} * S_{DIST}$ 벡터들을 오디오 디코딩하여 하나 이상의 $U_{DIST} * S_{DIST}$ 벡터들을 생성하고, $U_{DIST} * S_{DIST}$ 벡터들에 V^{T}_{DIST} 벡터들을 곱하여 음장의 구별되는 성분들을 설명하는 복수의 구면 조화 계수들의 것들을 복원하도록 구성되며, 여기서 하나 이상의 제 2 벡터들은, 오디오 인코딩 전에, S 행렬 내에 포함된 S_{BG} 벡터들을 그 후 V 행렬의 트랜스포즈 내에 포함된 V^{T}_{BG} 벡터들을 U 행렬 내에 포함된 U_{BG} 벡터들에 곱함으로써 생성되었고, 오디오 디코딩 디바이스 (24) 는 또한, 하나 이상의 오디오 인코딩된 $U_{BG} * S_{BG} * V^{T}_{BG}$ 벡터들을 오디오 디코딩하여 음장의 배경 성분들을 설명하는 복수의 구면 조화 계수들의 적어도 일부분을 복원하고, 음장의 구별되는 성분들을 설명하는 복수의 구면 조화 계수들을 음장의 배경 성분들을 설명하는 복수의 구면 조화 계수들의 적어도 일부분에 추가하여 복수의 구면 조화 계수들의 복원된 버전을 생성한다.

[0438] 이들 및 다른 경우들에서, 오디오 디코딩 디바이스 (24) 에서, 하나 이상의 제 1 벡터들은, 오디오 인코딩 전에, S 행렬의 하나 이상의 S_{DIST} 벡터들을 U 행렬의 하나 이상의 오디오 인코딩된 U_{DIST} 벡터들에 곱함으로써 생성된 하나 이상의 $U_{DIST} * S_{DIST}$ 벡터들, 및 V 행렬의 트랜스포즈의 하나 이상의 V^{T}_{DIST} 벡터들을 포함하고, 여기서 U 행렬, S 행렬 및 V 행렬은 복수의 구면 조화 계수들에 대하여 적어도 특이값 분해를 수행함으로써 생성되었고, 오디오 디코딩 디바이스 (20) 는 또한, 비트스트림으로부터 추출될 벡터들의 수를 나타내는 값 D 를 획득하여 하나 이상의 $U_{DIST} * S_{DIST}$ 벡터들 및 하나 이상의 V^{T}_{DIST} 벡터들을 형성하도록 구성된다.

[0439] 이들 및 다른 경우들에서, 오디오 디코딩 디바이스 (24) 에서, 하나 이상의 제 1 벡터들은, 오디오 인코딩 전에, S 행렬의 하나 이상의 S_{DIST} 벡터들을 U 행렬의 하나 이상의 오디오 인코딩된 U_{DIST} 벡터들에 곱함으로써 생성된 하나 이상의 $U_{DIST} * S_{DIST}$ 벡터들, 및 V 행렬의 트랜스포즈의 하나 이상의 V^{T}_{DIST} 벡터들을 포함하고, 여기서 U 행렬, S 행렬 및 V 행렬은 복수의 구면 조화 계수들에 대하여 적어도 특이값 분해를 수행함으로써 생성되었고, 오디오 디코딩 디바이스 (24) 는 또한, 비트스트림으로부터 추출될 벡터들의 수를 나타내는 오디오-프레임 단위 상의 값 D 를 획득하여 하나 이상의 $U_{DIST} * S_{DIST}$ 벡터들 및 하나 이상의 V^{T}_{DIST} 벡터들을 형성하도록 구성된다.

[0440] 이들 및 다른 경우들에서, 오디오 디코딩 디바이스 (24) 에서, 변환은 음장의 구별되는 성분들 및 음장의 배경 성분들을 식별하기 위한 주요 성분 분석을 포함한다.

[0441] 본 개시물에 설명된 기술들의 다양한 양태들은 또한, 오디오 인코딩 디바이스 (24) 로 하여금 HOA 계수들의 분해된 버전들에 대하여 보간을 수행하게 할 수도 있다. 일부 경우들에서, 오디오 디코딩 디바이스 (24) 는, 제 1 복수의 구면 조화 계수들의 제 1 분해 및 제 2 복수의 구면 조화 계수들의 제 2 분해에 대하여 적어도 부분적으로 보간을 수행함으로써, 시간 세그먼트에 대한 분해된 보간 구면 조화 계수들을 획득하도록 구성될 수도 있다.

[0442] 이들 및 다른 경우들에서, 제 1 분해는 제 1 복수의 구면 조화 계수들의 우-특이 벡터들을 나타내는 제 1 V 행

렬을 포함한다.

[0443] 이들 및 다른 예들에서, 제 2 분해는 제 2 복수의 구면 조화 계수들의 우-특이 벡터들을 나타내는 제 2 V 행렬을 포함한다.

[0444] 이들 및 다른 경우들에서, 제 1 분해는 제 1 복수의 구면 조화 계수들의 우-특이 벡터들을 나타내는 제 1 V 행렬을 포함하고, 제 2 분해는 제 2 복수의 구면 조화 계수들의 우-특이 벡터들을 나타내는 제 2 V 행렬을 포함한다.

[0445] 이들 및 다른 경우들에서, 시간 세그먼트는 오디오 프레임의 서브-프레임을 포함한다.

[0446] 이들 및 다른 경우들에서, 시간 세그먼트는 오디오 프레임의 시간 샘플을 포함한다.

[0447] 이들 및 다른 경우들에서, 오디오 디코딩 디바이스 (24) 는 제 1 복수의 구면 조화 계수들의 구면 조화 계수들에 대한 제 2 분해 및 제 1 분해의 보간된 분해를 획득하도록 구성된다.

[0448] 이들 및 다른 경우들에서, 오디오 디코딩 디바이스 (24) 는 제 1 프레임에 포함된 제 1 복수의 구면 조화 계수들의 제 1 부분에 대한 제 1 분해 및 제 2 프레임에 포함된 제 2 복수의 구면 조화 계수들의 제 2 부분에 대한 제 2 분해의 보간된 분해들을 획득하도록 구성되고, 오디오 디코딩 디바이스 (24) 는 또한, 제 1 프레임에 포함된 제 1 복수의 구면 조화 계수들의 제 1 부분의 제 1 시간 성분에 보간된 분해들을 적용하여 제 1 복수의 구면 조화 계수들의 제 1 인공 시간 성분을 생성하고, 제 2 프레임에 포함된 제 2 복수의 구면 조화 계수들의 제 2 부분의 제 2 시간 성분에 각각의 보간된 분해들을 적용하여 포함된 제 2 복수의 구면 조화 계수들의 제 2 인공 시간 성분을 생성하도록 구성된다.

[0449] 이들 및 다른 경우들에서, 제 1 시간 성분은 제 1 복수의 구면 조화 계수들에 대하여 벡터-기반 합성을 수행함으로써 제 1 시간 성분이 생성된다.

[0450] 이들 및 다른 경우들에서, 제 2 시간 성분은 제 2 복수의 구면 조화 계수들에 대하여 벡터-기반 합성을 수행함으로써 제 2 시간 성분이 생성된다.

[0451] 이들 및 다른 경우들에서, 오디오 디코딩 디바이스 (24) 는 또한, 제 1 인공 시간 성분 및 제 2 인공 시간 성분을 수신하고, 제 1 복수의 구면 조화 계수들의 제 1 부분에 대한 제 1 분해 및 제 2 복수의 구면 조화 계수들의 제 2 부분에 대한 제 2 분해의 보간된 분해들을 컴퓨팅하며, 제 1 인공 시간 성분에 보간된 분해들의 역 (inverse) 들을 적용하여 제 1 시간 성분을 복원하고 제 2 인공 시간 성분에 보간된 분해들의 역들을 적용하여 제 2 시간 성분을 복원하도록 구성된다.

[0452] 이들 및 다른 경우들에서, 오디오 디코딩 디바이스 (24) 는 제 1 복수의 구면 조화 계수들의 제 1 공간 성분 및 제 2 복수의 구면 조화 계수들의 제 2 공간 성분을 보간하도록 구성된다.

[0453] 이들 및 다른 경우들에서, 제 1 공간 성분은 제 1 복수의 구면 조화 계수들의 좌-특이 벡터들을 나타내는 제 1 U 행렬을 포함한다.

[0454] 이들 및 다른 경우들에서, 제 2 공간 성분은 제 2 복수의 구면 조화 계수들의 좌-특이 벡터들을 나타내는 제 2 U 행렬을 포함한다.

[0455] 이들 및 다른 경우들에서, 제 1 공간 성분은 제 1 복수의 구면 조화 계수들에 대한 구면 조화 계수들의 M 시간 세그먼트들을 나타내고, 제 2 공간 성분은 제 2 복수의 구면 조화 계수들에 대한 구면 조화 계수들의 M 시간 세그먼트들을 나타낸다.

[0456] 이들 및 다른 경우들에서, 제 1 공간 성분은 제 1 복수의 구면 조화 계수들에 대한 구면 조화 계수들의 M 시간 세그먼트들을 나타내고, 제 2 공간 성분은 제 2 복수의 구면 조화 계수들에 대한 구면 조화 계수들의 M 시간 세그먼트들을 나타내며, 오디오 디코딩 디바이스 (24) 는 제 1 공간 성분의 최종 N 개의 엘리먼트들 및 제 2 공간 성분의 제 1 N 개의 성분들을 보간하도록 구성된다.

[0457] 이들 및 다른 경우들에서, 제 2 복수의 구면 조화 계수들은 시간 도메인에서 제 1 복수의 구면 조화 계수들에 후속된다.

[0458] 이들 및 다른 경우들에서, 오디오 디코딩 디바이스 (24) 는 또한, 제 1 복수의 구면 조화 계수들을 분해하여 제 1 복수의 구면 조화 계수들의 제 1 분해를 생성하도록 구성된다.

[0459] 이들 및 다른 경우들에서, 오디오 디코딩 디바이스 (24) 는 또한, 제 2 복수의 구면 조화 계수들을 분해하여 제

2 복수의 구면 조화 계수들의 제 2 분해를 생성하도록 구성된다.

[0460] 이들 및 다른 경우들에서, 오디오 디코딩 디바이스 (24) 는 또한, 제 1 복수의 구면 조화 계수들에 대하여 특이값 분해를 수행하여 제 1 복수의 구면 조화 계수들의 좌-특이 벡터들을 나타내는 U 행렬, 제 1 복수의 구면 조화 계수들의 특이값들을 나타내는 S 행렬 및 제 1 복수의 구면 조화 계수들의 우-특이 벡터들을 나타내는 V 행렬을 생성하도록 구성된다.

[0461] 이들 및 다른 경우들에서, 오디오 디코딩 디바이스 (24) 는 또한, 제 2 복수의 구면 조화 계수들에 대하여 특이값 분해를 수행하여 제 2 복수의 구면 조화 계수들의 좌-특이 벡터들을 나타내는 U 행렬, 제 2 복수의 구면 조화 계수들의 특이값들을 나타내는 S 행렬 및 제 2 복수의 구면 조화 계수들의 우-특이 벡터들을 나타내는 V 행렬을 생성하도록 구성된다.

[0462] 이들 및 다른 경우들에서, 제 1 및 제 2 복수의 구면 조화 계수들 각각은, 음장의 평면파 표현을 나타낸다.

[0463] 이들 및 다른 경우들에서, 제 1 및 제 2 복수의 구면 조화 계수들 각각은 함께 혼합된 하나 이상의 모노-오디오 오브젝트들을 나타낸다.

[0464] 이들 및 다른 경우들에서, 제 1 및 제 2 복수의 구면 조화 계수들 각각은, 3 차원 음장을 나타내는 각각의 제 1 및 제 2 구면 조화 계수들을 포함한다.

[0465] 이들 및 다른 경우들에서, 제 1 및 제 2 복수의 구면 조화 계수들 각각은, 1 보다 큰 오더를 갖는 적어도 하나의 구면 기저 함수와 연관된다.

[0466] 이들 및 다른 경우들에서, 제 1 및 제 2 복수의 구면 조화 계수들 각각은, 4 와 동일한 오더를 갖는 적어도 하나의 구면 기저 함수와 연관된다.

[0467] 이들 및 다른 경우들에서, 보간은 제 1 분해 및 제 2 분해의 가중된 보간이고, 여기서 제 1 분해에 적용된 가중된 보간의 가중치들은 제 1 및 제 2 분해의 벡터들로 표현된 시간에 역비례하며, 제 2 분해에 적용된 가중화된 보간의 가중치들은 제 1 및 제 2 분해의 벡터들로 표현된 시간에 비례한다.

[0468] 이들 및 다른 경우들에서, 분해된 보간 구면 조화 계수들은 제 1 복수의 구면 조화 계수들 및 제 2 복수의 구면 조화 계수들의 시간 성분들 및 공간 성분들 중 적어도 하나를 평활화한다.

[0469] 이들 및 다른 경우들에서, 오디오 디코딩 디바이스 (24) 는 Us[n] = HOA(n) * (V_vec[n])-1 을 컴퓨팅하여 스칼라를 획득하도록 구성된다.

[0470] 이들 및 다른 경우들에서, 보간은 선형 보간을 포함한다. 이들 및 다른 경우들에서, 보간은 비-선형 보간을 포함한다. 이들 및 다른 경우들에서, 보간은 코사인 보간을 포함한다. 이들 및 다른 경우들에서, 보간은 가중된 코사인 보간을 포함한다. 이들 및 다른 경우들에서, 보간은 큐빅 보간을 포함한다. 이들 및 다른 경우들에서, 보간은 적응적 스플라인 (Adaptive Spline) 보간을 포함한다. 이들 및 다른 경우들에서, 보간은 최소 곡률 보간을 포함한다.

[0471] 이들 및 다른 경우들에서, 오디오 디코딩 디바이스 (24) 는 또한, 시간 세그먼트에 대한 분해된 보간 구면 조화 계수들의 표현 및 보간의 유형의 표시를 포함하는 비트스트림을 생성하도록 구성된다.

[0472] 이들 및 다른 경우들에서, 이 표시는 보간의 유형으로 맵핑하는 하나 이상의 비트들을 포함한다.

[0473] 이들 및 다른 경우들에서, 오디오 디코딩 디바이스 (24) 는 또한, 시간 세그먼트에 대한 분해된 보간 구면 조화 계수들의 표현 및 보간의 유형의 표시를 포함하는 비트스트림을 획득하도록 구성된다.

[0474] 이들 및 다른 경우들에서, 이 표시는 보간의 유형으로 맵핑하는 하나 이상의 비트들을 포함한다.

[0475] 기술들의 다양한 양태들은, 일부 경우들에서 또한, 오디오 디코딩 디바이스 (24) 로 하여금 음장의 공간 성분의 압축된 버전을 포함하는 비트스트림을 획득하게 할 수도 있고, 이 공간 성분은 복수의 구면 조화 계수들에 대하여 벡터 기반 합성을 수행함으로써 생성된다.

[0476] 이들 및 다른 경우들에서, 공간 성분의 압축된 버전은, 공간 성분을 압축할 때 사용된 예측 모드를 지정하는 필드를 적어도 부분적으로 사용하여 비트스트림에서 표현된다.

[0477] 이들 및 다른 경우들에서, 공간 성분의 압축된 버전은, 공간 성분을 압축할 때 사용된 허프만 테이블을 지정하는 허프만 테이블 정보를 적어도 부분적으로 사용하여 비트스트림에서 표현된다.

[0478] 이들 및 다른 경우들에서, 공간 성분의 압축된 버전은, 공간 성분을 압축할 때 사용된 양자화 스텝 사이즈 또는 그의 변수를 표현하는 값을 나타내는 필드를 적어도 부분적으로 사용하여 비트스트림에서 표현된다.

[0479] 이들 및 다른 경우들에서, 값은 n비트들 값을 포함한다.

[0480] 이들 및 다른 경우들에서, 비트스트림은 공간 성분의 압축된 버전이 포함되는 음장의 복수의 공간 성분들의 압축된 버전을 포함하고, 값은 복수의 공간 성분들을 압축할 때 사용된 양자화 스텝 사이즈 또는 그의 변수를 표현한다.

[0481] 이들 및 다른 경우들에서, 공간 성분의 압축된 버전은, 공간 성분이 대응하는 압축 카테고리를 식별하는 카테고리 식별자를 표현하기 위해 허프만 코드를 적어도 부분적으로 사용하여 비트스트림에서 표현된다.

[0482] 이들 및 다른 경우들에서, 공간 성분의 압축된 버전은, 이 공간 성분이 포지티브 값인지 또는 네거티브 값인지 여부를 식별하는 사인 비트를 적어도 부분적으로 사용하여 비트스트림에서 표현된다.

[0483] 이들 및 다른 경우들에서, 공간 성분의 압축된 버전은 공간 성분의 잔차 값을 나타내기 위해 허프만 코드를 적어도 부분적으로 사용하여 비트스트림에서 표현된다.

[0484] 이들 및 다른 경우들에서, 디바이스는 오디오 디코딩 디바이스를 포함한다.

[0485] 기술들의 다양한 양태들은 또한, 오디오 디코딩 디바이스 (24) 로 하여금, 복수의 압축된 공간 성분들 중 나머지 성분들에 대한 공간 성분의 압축된 버전의 오더에 기초하여 복수의 압축된 공간 성분들 중 공간 성분의 압축된 버전을 분해하는 경우 사용할 허프만 코드북을 식별하게 할 수도 있고, 이 공간 성분은 복수의 구면 조화 계수들에 대하여 벡터 기반 합성을 수행함으로써 생성된다.

[0486] 이들 및 다른 경우들에서, 오디오 디코딩 디바이스 (24) 는 음장의 공간 성분의 압축된 버전을 포함하는 비트스트림을 획득하고, 식별된 허프만 코드북을 적어도 부분적으로 사용하여 공간 성분의 압축된 버전을 분해하여 공간 성분을 획득하도록 구성된다.

[0487] 이들 및 다른 경우들에서, 공간 성분의 압축된 버전은 공간 성분을 압축할 때 사용된 예측 모드를 지정하는 필드를 적어도 부분적으로 사용하여 비트스트림에서 표현되고, 오디오 디코딩 디바이스 (24) 는 예측 모드에 적어도 부분적으로 기초하여 공간 성분의 압축된 버전을 분해하여 공간 성분을 획득하도록 구성된다.

[0488] 이들 및 다른 경우들에서, 공간 성분의 압축된 버전은 공간 성분을 압축할 때 사용된 허프만 테이블을 지정하는 허프만 테이블 정보를 적어도 부분적으로 사용하여 비트스트림에서 표현되고, 오디오 디코딩 디바이스 (24) 는 허프만 테이블 정보에 적어도 부분적으로 기초하여 공간 성분의 압축된 버전을 분해하도록 구성된다.

[0489] 이들 및 다른 경우들에서, 공간 성분의 압축된 버전은 공간 성분을 압축할 때 사용된 양자화 스텝 사이즈 또는 그의 변수를 표현하는 값을 나타내는 필드를 적어도 부분적으로 사용하여 비트스트림에서 표현되고, 오디오 디코딩 디바이스 (24) 는 그 값에 적어도 부분적으로 기초하여 공간 성분의 압축된 버전을 분해하도록 구성된다.

[0490] 이들 및 다른 경우들에서, 값은 *nbits* 값을 포함한다.

[0491] 이들 및 다른 경우들에서, 비트스트림은, 공간 성분의 압축된 버전이 포함되는 음장의 복수의 공간 성분들의 압축된 버전을 포함하고, 값은 복수의 공간 성분들을 압축할 때 사용된 양자화 스텝 사이즈 또는 그의 변수를 표현하며, 오디오 디코딩 디바이스 (24) 는 그 값에 적어도 부분적으로 기초하여 공간 성분의 복수의 압축된 버전을 분해하도록 구성된다.

[0492] 이들 및 다른 경우들에서, 공간 성분의 압축된 버전은 공간 성분이 대응하는 압축 카테고리를 식별하는 카테고리 식별자를 나타내기 위해 허프만 코드를 적어도 부분적으로 사용하여 비트스트림에서 표현되고, 오디오 디코딩 디바이스 (24) 는 그 허프만 코드에 적어도 부분적으로 기초하여 공간 성분의 압축된 버전을 분해하도록 구성된다.

[0493] 이들 및 다른 경우들에서, 공간 성분의 압축된 버전은, 이 공간 성분이 포지티브 값인지 또는 네거티브 값인지 여부를 식별하는 사인 비트를 적어도 부분적으로 사용하여 비트스트림에서 표현되고, 오디오 디코딩 디바이스 (24) 는 그 사인 비트에 적어도 부분적으로 기초하여 공간 성분의 압축된 버전을 분해하도록 구성된다.

[0494] 이들 및 다른 경우들에서, 공간 성분의 압축된 버전은 공간 성분의 잔차 값을 나타내기 위해 허프만 코드를 적어도 부분적으로 사용하여 비트스트림에서 표현되고, 오디오 디코딩 디바이스 (24) 는 식별된 허프만 코드북에 포함된 허프만 코드에 적어도 부분적으로 기초하여 공간 성분의 압축된 버전을 분해하도록 구성된다.

[0495] 전술된 다양한 경우들 각각에서, 오디오 디코딩 디바이스 (24) 는, 오디오 디코딩 디바이스 (24) 가 수행하도록 구성되는 방법을 수행하거나 다르게는 그 방법의 각 단계를 수행하기 위한 수단을 포함할 수도 있음이 이해되어야 한다. 일부 경우들에서, 이들 수단은 하나 이상의 프로세서들을 포함할 수도 있다. 일부 경우들에서, 하나 이상의 프로세서들은 비-일시적 컴퓨터 판독가능 저장 매체에 저장된 명령들의 방식에 의해 구성된 특수 용도의 프로세서를 나타낼 수도 있다. 다시 말해, 인코딩 예들의 세트들 각각에서 기술들의 다양한 양태들은 명령들이 저장되어 있는 비-일시적 컴퓨터 판독가능 저장 매체에 대해 제공할 수도 있고, 이 명령들은 실행되는 경우 하나 이상의 프로세서들로 하여금 오디오 디코딩 디바이스 (24) 가 수행하도록 구성된 방법을 수행하게 한다.

[0496] 도 6 은 본 개시물에 설명된 기술들의 다양한 양태들을 수행하는데 있어서, 도 4 의 예에 도시된 컨텐츠 분석 유닛 (26) 과 같은 오디오 인코딩 디바이스의 컨텐츠 분석 유닛의 예시적인 동작을 예시하는 플로우차트이다.

[0497] 컨텐츠 분석 유닛 (26) 은, 음장을 나타내는 HOA 계수들 (11) 이 합성 오디오 오브젝트로부터 생성되는지 여부를 결정하는 경우, 네 번째 오더의 표현 (즉, N=4) 에 대해 25 × 1024 사이즈일 수도 있는 프레임화된 HOA 계수들을 획득할 수도 있다 (93). 프레임화된 HOA 계수들 (이것은 또한, 프레임화된 SHC 행렬 (11) 로서 본원에서 표기될 수도 있고, 후속의 프레임화된 SHC 행렬들은 프레임화된 SHC 행렬들 (27B, 27C 등) 으로서 표기될 수도 있음) 을 획득한 후에, 컨텐츠 분석 유닛 (26) 은 그 후, 감소된 프레임화된 HOA 계수들을 생성하기 위해 프레임화된 HOA 계수들 (11) 의 제 1 벡터를 제외시킬 수도 있다 (94).

[0498] 컨텐츠 분석 유닛 (26) 은 그 후, 감소된 프레임화된 HOA 계수들의 나머지 벡터들로부터 감소된 프레임화된 HOA 계수들의 제 1 비-제로 벡터를 예측할 수도 있다 (95). 제 1 비-제로 벡터를 예측한 후에, 컨텐츠 분석 유닛 (26) 은 예측된 제 1 비-제로 벡터 및 실제 비-제로 벡터에 기초하여 에러를 획득할 수도 있다 (96). 일단 이 에러가 획득되면, 컨텐츠 분석 유닛 (26) 은 실제 제 1 비-제로 벡터 및 에러의 에너지에 기초하여 비율을 컴퓨팅할 수도 있다 (97). 컨텐츠 분석 유닛 (26) 은 그 후, 이 비율을 임계치에 비교할 수도 있다 (98). 비율이 임계치를 초과하지 않는 경우 (98 에서 "아니오"), 컨텐츠 분석 유닛 (26) 은, 프레임화된 SHC 행렬 (11) 이 레코딩으로부터 생성된다는 것을 결정하고, SHC 행렬 (11) 의 대응하는 코딩된 표현이 레코딩으로부터 생성되었다는 것을 비트스트림에 표시할 수도 있다 (100, 101). 비율이 임계치를 초과하는 경우 (98 에서 "예"), 컨텐츠 분석 유닛 (26) 은, 프레임화된 SHC 행렬 (11) 이 합성 오디오 오브젝트로부터 생성된다는 것을 결정하고, SHC 행렬 (11) 의 대응하는 코딩된 표현이 합성 오디오 오브젝트로부터 생성되었다는 것을 비트스트림에 표시할 수도 있다 (102, 103). 일부 경우들에서, 프레임화된 SHC 행렬 (11) 이 레코딩으로부터 생성된 경우, 컨텐츠 분석 유닛 (26) 은 프레임화된 SHC 행렬 (11) 을 벡터-기반 합성 유닛 (27) 으로 패스한다 (101). 일부 경우들에서, 프레임화된 SHC 행렬 (11) 이 합성 오디오 오브젝트로부터 생성된 경우, 컨텐츠 분석 유닛 (26) 은 프레임화된 SHC 행렬 (11) 을 방향-기반 합성 유닛 (28) 으로 패스한다 (104).

[0499] 도 7 은 본 개시물에 설명된 벡터-기반 합성 기술들의 다양한 양태들을 수행하는데 있어서, 도 4 의 예에 도시된 오디오 인코딩 디바이스 (20) 와 같은 오디오 인코딩 디바이스의 예시적인 동작을 예시하는 플로우차트이다. 먼저, 오디오 인코딩 디바이스 (20) 는 HOA 계수들 (11) 을 수신한다 (106). 오디오 인코딩 디바이스 (20) 는 LIT 유닛 (30) 을 인보크할 수도 있고, 이것은 HOA 계수들에 대하여 LIT 를 적용하여 변환된 HOA 계수들 (예를 들어, SVD 의 경우에서, 변환된 HOA 계수들은 US[*k*] 벡터들 (33) 및 V[*k*] 벡터들 (35) 을 포함할 수도 있음) 을 출력할 수도 있다 (107).

[0500] 오디오 인코딩 디바이스 (20) 는 다음으로, US[*k*] 벡터들 (33), US[*k-1*] 벡터들 (33), V[*k*] 및/또는 V[*k-1*] 벡터들 (35) 의 임의의 조합에 대하여 전술된 분석을 수행하여 전술된 방식으로 다양한 파라미터들을 식별하도록, 파라미터 계산 유닛 (32) 을 인보크할 수도 있다. 즉, 파라미터 계산 유닛 (32) 은 변환된 HOA 계수들 (33/35) 의 분석에 기초하여 적어도 하나의 파라미터를 결정할 수도 있다 (108).

[0501] 오디오 인코딩 디바이스 (20) 는 그 후, 전술된 바와 같이, 리오더 유닛 (34) 을 인보크할 수도 있고, 이것은 파라미터에 기초하여 변환된 HOA 계수들 (이것은, 다시 SVD 의 컨텍스트에서, US[*k*] 벡터들 (33) 및 V[*k*] 벡터들 (35) 을 지칭할 수도 있음) 을 리오더링하여, 리오더링된 변환된 HOA 계수들 (33'/35')(또는, 다시 말해, US[*k*] 벡터들 (33') 및 V[*k*] 벡터들 (35')) 을 생성할 수도 있다 (109). 오디오 인코딩 디바이스 (20) 는, 앞서의 동작들 또는 후속의 동작들 중 어느 하나 동안, 또한 음장 분석 유닛 (44) 을 인보크할 수도 있다. 음장 분석 유닛 (44) 은, 전술된 바와 같이 HOA 계수들 (11) 및/또는 변환된 HOA 계수들 (33/35) 에 대하여 음장 분석을 수행하여 전경 채널들 (nFG)(45) 의 총 수, 배경 음장 (N_{BG}) 의 오더 및 수 (nBGa), 및 (도 4 의 예에서 배경 채널 정보 (43) 로서 총괄하여 표기될 수도 있는) 전송할 추가의 BG HOA 채널들의 인덱스들 (i) 을

결정할 수도 있다 (109).

[0502] 오디오 인코딩 디바이스 (20) 는 또한, 배경 선택 유닛 (48) 을 인보크할 수도 있다. 배경 선택 유닛 (48) 은 배경 채널 정보 (43) 에 기초하여 배경 또는 주변 HOA 계수들 (47) 을 결정할 수도 있다 (110). 오디오 인코딩 디바이스 (20) 는 또한, 전경 선택 유닛 (36) 을 인보크할 수도 있고, 이것은 (이들 전경 벡터들을 식별하는 하나 이상의 인덱스들을 나타낼 수도 있는) nFG (45) 에 기초하여 음장의 구별되는 성분들 또는 전경을 나타내는 리오더링된 V[k] 벡터들 (35') 및 리오더링된 US[*k*] 벡터들 (33') 의 것들을 선택할 수도 있다 (112).

[0503] 오디오 인코딩 디바이스 (20) 는 에너지 보상 유닛 (38) 을 인보크할 수도 있다. 에너지 보상 유닛 (38) 은 주변 HOA 계수들 (47) 에 대하여 에너지 보상을 수행하여 배경 선택 유닛에 의해 HOA 채널들 중 다양한 채널들의 제거로 인한 에너지 손실을 보상하고 (114), 이에 의해 에너지 보상된 주변 HOA 계수들 (47') 을 생성할 수도 있다.

[0504] 오디오 인코딩 디바이스 (20) 는 또한, 공간-시간 보간 유닛 (50) 을 인보크한다. 공간-시간 보간 유닛 (50) 은 리오더링된 변환된 HOA 계수들 (33'/35') 에 대하여 공간-시간 보간을 수행하여, ("보간된 nFG 신호들 (49')" 로서 지칭될 수도 있는) 보간된 전경 신호들 (49') 및 ("V[*k*] 벡터들 (53)" 로서 지칭될 수도 있는) 나머지 전경 방향 정보 (53) 를 획득할 수도 있다 (116). 오디오 인코딩 디바이스 (20) 는 그 후, 계수 감소 유닛 (46) 을 인보크할 수도 있다. 계수 감소 유닛 (46) 은 배경 채널 정보 (43) 에 기초하여 나머지 전경 V[k] 벡터들 (53) 에 대하여 계수 감소를 수행하여, (감소된 전경 V[k] 벡터들 (55) 로서 지칭될 수도 있는) 감소된 전경 방향 정보 (55) 를 획득할 수도 있다 (118).

[0505] 오디오 인코딩 디바이스 (20) 는 그 후, 감소된 전경 V[k] 벡터들 (55) 을 전술된 방식으로 압축하고 코딩된 전경 V[k] 벡터들 (57) 을 생성하도록, 양자화 유닛 (52) 을 인보크할 수도 있다 (120).

[0506] 오디오 인코딩 디바이스 (20) 는 또한, 심리 음향 오디오 코더 유닛 (40) 을 인보크할 수도 있다. 심리 음향 오디오 코더 유닛 (40) 은 에너지 보상된 주변 HOA 계수들 (47') 및 보간된 nFG 신호들 (49') 의 각각의 벡터를 심리 음향 코딩하여, 인코딩된 주변 HOA 계수들 (59) 및 인코딩된 nFG 신호들 (61) 을 생성할 수도 있다. 오디오 인코딩 디바이스는 그 후, 비트스트림 생성 유닛 (42) 을 인보크할 수도 있다. 비트스트림 생성 유닛 (42) 은 코딩된 전경 방향 정보 (57), 코딩된 주변 HOA 계수들 (59), 코딩된 nFG 신호들 (61) 및 배경 채널 정보 (43) 에 기초하여 비트스트림 (21) 을 생성할 수도 있다.

[0507] 도 8 은 본 개시물에 설명된 기술들의 다양한 양태들을 수행하는데 있어서, 도 5 에 도시된 오디오 디코딩 디바이스 (24) 와 같은 오디오 디코딩 디바이스의 예시적인 동작을 나타내는 플로우차트이다. 먼저, 오디오 디코딩 디바이스 (24) 는 비트스트림 (21) 을 수신할 수도 있다 (130). 비트스트림의 수신 시에, 오디오 디코딩 디바이스 (24) 는 추출 유닛 (72) 을 인보크할 수도 있다. 논의의 목적을 위해 비트스트림 (21) 이, 벡터-기반 복원이 수행된다는 것을 나타낸다고 가정하면, 추출 디바이스 (72) 는 이 정보를 벡터-기반 복원 유닛 (92) 으로 패스하는, 상기 주목된 정보를 취출하도록 이 비트스트림을 파싱할 수도 있다.

[0508] 다시 말해, 추출 유닛 (72) 은 전술된 방식으로 비트스트림 (21) 으로부터, 코딩된 전경 방향 정보 (57)(이것은, 다시 또한 코딩된 전경 V[k] 벡터들 (57) 로서 지칭될 수도 있음), 코딩된 주변 HOA 계수들 (59) 및 코딩된 전경 신호들 (이것은 또한, 코딩된 전경 nFG 신호들 (59) 또는 코딩된 전경 오디오 오브젝트들 (59) 로서 지칭될 수도 있음) 을 추출할 수도 있다 (132).

[0509] 오디오 디코딩 디바이스 (24) 는 또한, 양자화 유닛 (74) 을 인보크할 수도 있다. 양자화 유닛 (74) 은 코딩된 전경 방향 정보 (57) 를 엔트로피 디코딩 및 역양자화하여, 감소된 전경 방향 정보 (55_k) 를 획득할 수도 있다 (136). 오디오 디코딩 디바이스 (24) 는 또한, 심리 음향 디코딩 유닛 (80) 을 인보크할 수도 있다. 심리 음향 오디오 코딩 유닛 (80) 은 인코딩된 주변 HOA 계수들 (59) 및 인코딩된 전경 신호들 (61) 을 디코딩하여, 에너지 보상된 주변 HOA 계수들 (47') 및 보간된 전경 신호들 (49') 을 획득할 수도 있다 (138). 심리 음향 디코딩 유닛 (80) 은 에너지 보상된 주변 HOA 계수들 (47') 을 HOA 계수 공식화 유닛 (82) 으로 그리고 nFG 신호들 (49') 을 리오더 유닛 (84) 으로 패스할 수도 있다.

[0510] 리오더 유닛 (84) 은 HOA 계수들 (11) 의 전경 성분들의 원래 오더를 나타내는 신택스 엘리먼트들을 수신할 수도 있다. 리오더 유닛 (84) 은, 이들 리오더 신택스 엘리먼트들에 기초하여 보간된 nFG 신호들 (49') 및 감소된 전경 V[k] 벡터들 (55_k) 을 리오더링하여, 리오더링된 nFG 신호들 (49') 및 리오더링된 전경 V[k] 벡터들 (55_k') 을 생성할 수도 있다 (140). 리오더 유닛 (84) 은 리오더링된 nFG 신호들 (49") 을 전경 공식화 유

닛 (78) 으로 그리고 리오더링된 전경 V[k] 벡터들 (55_k') 을 공간-시간 보간 유닛 (76) 으로 출력할 수도 있다.

[0511] 오디오 디코딩 디바이스 (24) 는 다음으로, 공간-시간 보간 유닛 (76) 을 인보크할 수도 있다. 공간-시간 보간 유닛 (76) 은 리오더링된 전경 방향 정보 (55_k') 를 수신하고, 감소된 전경 방향 정보 ($55_k/55_{k-1}$) 에 대하여 공간-시간 보간을 수행하여, 보간된 전경 방향 정보 (55_k") 를 생성할 수도 있다 (142). 공간-시간 보간 유닛 (76) 은 보간된 전경 V[k] 벡터들 (55_k") 을 전경 공식화 유닛 (718) 으로 포워딩할 수도 있다.

[0512] 오디오 디코딩 디바이스 (24) 는 전경 공식화 유닛 (78) 을 인보크할 수도 있다. 전경 공식화 유닛 (78) 은 보간된 전경 방향 정보 (55_k") 에 의한 보간된 전경 신호들 (49") 의 행렬 곱셈을 수행하여 전경 HOA 계수들 (65) 을 획득할 수도 있다 (144). 오디오 디코딩 디바이스 (24) 는 또한, HOA 계수 공식화 유닛 (82) 을 인보크할 수도 있다. HOA 계수 공식화 유닛 (82) 은 전경 HOA 계수들 (65) 을 주변 HOA 채널들 (47') 에 합산하여, HOA 계수들 (11') 을 획득할 수도 있다 (146).

[0513] 도 9a 내지 도 91 은 도 4 의 예의 오디오 인코딩 디바이스 (20) 의 다양한 양태들을 더 상세하게 예시하는 블록도들이다. 도 9a 는 오디오 인코딩 디바이스 (20) 의 LIT 유닛 (30) 을 더 상세하게 예시하는 블록 다이어그램이다. 도 9a 의 예에 도시된 바와 같이, LIT 유닛 (30) 은 다수의 상이한 선형 도치 변환들 (200-200N) 을 포함할 수도 있다. LIT 유닛 (30) 은, 몇몇 예들을 제공하기 위해 특이값 분해 (singular value decomposition; SVD) 변환 (200A)("SVD 200A"), 주요 성분 분석 (principle component analysis; PCA) 변환 (200B)("PCA 200B"), 카루넨-루베 변환 (KLT)(200C)("KLT 200C"), 고속 푸리에 변환 (FFT)(200D)("FFT 200D") 및 이산 코사인 변환 (DCT)(200N)("DCT 200N") 을 포함할 수도 있다. LIT 유닛 (30) 은 이들 선형 도치 변환들 (200) 중 어느 하나를 인보크하여, HOA 계수들 (11) 에 대하여 각각의 변환을 적용하고 각각의 변환된 HOA 계수들 (33/35) 을 생성할 수도 있다.

[0514] HOA 계수들 (11) 에 대하여 직접 수행된 것으로 설명되었으나, LIT 유닛 (30) 은 HOA 계수들 (11) 의 파생물들에 선형 도치 변환들 (200) 을 적용할 수도 있다. 예를 들어, LIT 유닛 (30) 은 HOA 계수들 (11) 로부터 도출된 전력 스펙트럼 밀도 행렬에 대하여 SVD (200) 를 적용할 수도 있다. 전력 스펙트럼 밀도 행렬은 PSD 로서 표기되고, 이하에서 이어지는 의사-코드에서 요약된 바와 같이 hoaFrame 의 hoaFrame 로의 트랜스포즈의 행렬 곱셈을 통해 획득될 수도 있다. hoaFrame 표기는 HOA 계수들 (11) 의 프레임을 지칭한다.

[0515] LIT 유닛 (30) 은, SVD (200)(svd) 를 PSD 에 적용한 후에, $S[k]^2$ 행렬 (S_squared) 및 V[k] 행렬을 획득할 수도 있다. $S[k]^2$ 행렬은 제곱 S[k] 행렬을 가리킬 수도 있고, 그 결과 LIT (30) 유닛 (30)(또는, 다르게는 일 예로서 SVD 유닛 (200)) 은 제곱근 연산을 $S[k]^2$ 에 적용하여 S[k] 행렬을 획득할 수도 있다. SVD 유닛 (200) 은, 일부 경우들에서 V[k] 행렬에 대하여 양자화를 수행하여, (V[k]' 로서 표기될 수도 있는) 양자화된 V[k] 행렬을 획득할 수도 있다. LIT 유닛 (30) 은 먼저, S[k] 행렬에 양자화된 V[k]' 를 곱함으로써 U[k] 행렬을 획득하여, SV[k]' 행렬을 획득할 수도 있다. LIT 유닛 (30) 은 다음으로, SV[k]' 행렬의 의사-역 (pinv) 을 획득하고, 그 후 HOA 계수들 (11) 에 SV[k]' 행렬의 의사-역을 곱하여 U[k] 행렬을 획득할 수도 있다.

[0516] 상기는 다음의 의사-코드로 표현될 수도 있다:

[0517] PSD = hoaFrame'*hoaFrame;

[0518] [V, S_squared] = svd(PSD,'econ');

[0519] S = sqrt(S_squared);

[0520] U = hoaFrame * pinv(S*V).

[0521] 계수들 자체들 보다는 HOA 계수들의 전력 스펙트럼 밀도 (PSD) 에 대하여 SVD 를 수행함으로써, LIT 유닛 (30) 은 잠재적으로 하나 이상의 프로세서 사이클들 및 저장 공간의 관점에서 SVD 를 수행하는 컴퓨팅 복잡성을 감소시키면서, SVD 가 HOA 계수들에 직접 적용된 것과 같은 동일한 소스 오디오 인코딩 효율성을 달성할 수도 있다.
즉, 전술된 PSD-유형 SVD 는 잠재적으로, SVD 가 F*F 행렬 (F 는 HOA 계수들의 수) 상에 행해지기 때문에 컴퓨팅 부담이 덜할 수도 있다. M * F 행렬에 비해, M 은 프레임길이, 즉 1024 이상의 샘플들이다. SVD 의 복잡성은 이제, HOA 계수들 (11) 보다는 PSD 에의 적용을 통해, HOA 계수들 (11) 에 적용된 경우 O(M*L^2)

에 비해 대략 O(L^3) 일 수도 있다 (여기서, O(*) 는 컴퓨터-과학 분야들에 공통인 컴퓨팅 복잡성의 big-O 표기를 가리킨다).

[0522] 도 9b 는 오디오 인코딩 디바이스 (20) 의 파라미터 계산 유닛 (32) 을 더 상세하게 예시하는 블록 다이어그램이다. 파라미터 계산 유닛 (32) 은 에너지 분석 유닛 (202) 및 상호-상관 유닛 (204) 을 포함할 수도 있다.

에너지 분석 유닛 (202) 은 US[k] 벡터들 (33) 및 V[k] 벡터들 (35) 중 하나 이상에 대하여 전술된 에너지 분석을 수행하여, 현재 프레임 (k) 또는 이전 프레임 (k-1) 중 하나 이상에 대한 상관 파라미터 (R), 방향 특성 파라미터들 (θ, φ, r), 및 에너지 특성 (e) 중 하나 이상을 생성할 수도 있다. 유사하게, 상호-상관 유닛 (204) 은 US[k] 벡터들 (33) 및 V[k] 벡터들 (35) 중 하나 이상에 대하여 전술된 상호-상관을 수행하여, 현재 프레임 (k) 또는 이전 프레임 (k-1) 중 하나 이상에 대한 상관 파라미터 (R), 방향 특성 파라미터들 (θ, φ, r), 및 에너지 특성 (e) 중 하나 이상을 생성할 수도 있다. 파라미터 계산 유닛 (32) 은 현재 프레임 파라미터들 (37) 및 이전 프레임 파라미터들 (39) 을 출력할 수도 있다.

[0523] 도 9c 는 오디오 인코딩 디바이스 (20) 의 리오더 유닛 (34) 을 더 상세하게 예시하는 블록 다이어그램이다.

리오더 유닛 (34) 은 파라미터 평가 유닛 (206) 및 벡터 리오더 유닛 (208) 을 포함한다. 파라미터 평가 유닛 (206) 은 전술된 방식으로 이전 프레임 파라미터들 (39) 및 현재 프레임 파라미터들 (37) 을 평가하여 리오더 인덱스들을 (205) 을 생성하도록 구성된 유닛을 나타낸다. 리오더 인덱스들 (205) 은, US[k] 벡터들 (33) 의 벡터들 및 V[k] 벡터들 (35) 의 벡터들이 (예를 들어, 현재 벡터 로케이션의 인덱스를 식별하는 제 1 인덱스 쌍 및 벡터의 리오더링된 로케이션을 식별하는 제 2 인덱스 쌍을 갖는 인덱스 쌍들에 의해) 어떻게 리오더링되는지를 식별하는 인덱스들을 포함한다. 벡터 리오더 유닛 (208) 은 리오더 인덱스들 (205) 에 따라 US[k] 벡터들 (33) 및 V[k] 벡터들 (35) 을 리오더링하도록 구성된 유닛을 나타낸다. 리오더 유닛 (34) 은 리오더링된 US[k] 벡터들 (33') 및 리오더링된 V[k] 벡터들 (35') 을 출력할 수도 있는 한편, 또한 하나 이상의 신택스 엘리먼트들과 같은 리오더 인덱스들 (205) 을 비트스트림 생성 유닛 (42) 으로 패스한다.

[0524] 도 9d 는 오디오 인코딩 디바이스 (20) 의 음장 분석 유닛 (44) 을 더 상세하게 예시하는 블록 다이어그램이다.

도 9d 의 예에 도시된 바와 같이, 음장 분석 유닛 (44) 은 특이값 분석 유닛 (210A), 에너지 분석 유닛 (210B), 공간 분석 유닛 (210C), 공간 마스킹 분석 유닛 (210d), 확산 분석 유닛 (210e) 및 방향 분석 유닛 (210F) 을 포함할 수도 있다. 전술된 바와 같이, 특이값 분석 유닛 (210A) 은 (US[k] 벡터들 (33) 의 부분을 형성하는) S 벡터들의 내림차순 사선 값들에 의해 생성된 커브의 슬로프를 분석하도록 구성된 유닛을 나타낼 수도 있고, 여기서 큰 특이값들은 전경 또는 구별되는 사운드들을 나타내고 낮은 특이값들은 음장의 배경 성분들을 나타낸다. 에너지 분석 유닛 (210B) 은 벡터 단위 당 V[k] 벡터들 (35) 의 에너지를 결정하도록 구성된 유닛을 나타낼 수도 있다.

[0525] 공간 분석 유닛 (210C) 은 보존되어야 하는 음장의 방향 성분들을 나타내는 고 에너지의 영역들을 식별하고 공간 도메인으로의 HOA 계수들 (11) 의 변환을 통해 전술된 공간 에너지 분석을 수행하도록 구성된 유닛을 나타낼 수도 있다. 공간 마스킹 분석 유닛 (210d) 은, 공간 마스킹 분석 유닛 (210d) 이 공간적으로 근접한 더 높은 에너지 사운드들에 의해 마스킹되는 공간 영역을 식별할 수도 있는 것을 제외하고, 공간 에너지 분석의 방식과 유사한 방식으로 공간 마스킹 분석을 수행하도록 구성된 유닛을 나타낼 수도 있다. 확산 분석 유닛 (210e) 은 HOA 계수들 (11) 에 대하여 전술된 확산 분석을 수행하여 음장의 배경 성분들을 나타낼 수도 있는 확산 에너지의 영역들을 식별하도록 구성된 유닛을 나타낼 수도 있다. 방향 분석 유닛 (210F) 은 VS[k] 벡터들을 컴퓨팅하고, 이들 VS[k] 벡터들 각각의 각 엔트리를 제곱 및 합산하여 방향성 지수 (directionality quotient) 를 식별하는 것을 수반하는 전술된 방향 분석을 수행하도록 구성된 유닛을 나타낼 수도 있다. 방향 분석 유닛 (210F) 은 VS[k] 벡터들 각각에 대한 이 방향성 지수를 배경/전경 (BG/FG) 식별 (ID) 유닛 (212) 에 제공할 수도 있다.

[0526] 음장 분석 유닛 (44) 은 또한, BG/FG ID 유닛 (212) 을 포함할 수도 있고, 이 유닛은 전경 채널들 (nFG)(45) 의 총 수, 배경 음장 (N_{BG}) 의 오더 및 추가의 BG HOA 채널들의 인덱스들 (i) 및 수 (nBGa) 를 결정하여, 분석 유닛들 (210-210F) 의 임의의 조합에 의한 분석 출력의 임의의 조합에 기초하여 전송 (도 4 의 예에서 배경 채널 정보 (43) 로서 총괄하여 표기될 수도 있음) 하도록 구성된 유닛을 나타낼 수도 있다. BG/FG ID 유닛 (212) 은 타겟 비트레이트 (41) 를 달성하도록 배경 채널 정보 (43) 및 nFG (45) 를 결정할 수도 있다.

[0527] 도 9e 는 오디오 인코딩 디바이스 (20) 의 전경 선택 유닛 (36) 을 더 상세하게 예시하는 블록 다이어그램이다.

전경 선택 유닛 (36) 은 리오더링된 US[k] 벡터들 (33') 및 리오더링된 V[k] 벡터들 (35') 로부터 nFG 신택

스 엘리먼트 (45) 에 의해 식별된 전경 US[k] 벡터들 (49) 및 전경 V[k] 벡터들 (51_k) 을 파싱 또는 다르게는 추출할 수도 있는 벡터 파싱 유닛 (214) 을 포함한다. 벡터 파싱 유닛 (214) 은 음장 분석 유닛 (44) 에 의해 식별되고 nFG 신택스 엘리먼트 (45)(이것은 또한, 전경 채널 정보 (45) 로서 지칭될 수도 있음) 에 의해 지정된 음장의 전경 성분들을 나타내는 다양한 벡터들을 파싱할 수도 있다. 도 9e 의 예에서 도시된 바와 같이, 벡터 파싱 유닛 (214) 은 일부 경우들에서, 전경 US[k] 벡터들 (49) 및 전경 V[k] 벡터들 (51_k) 내의 비-연속적 벡터들을 선택하여, 음장의 전경 성분들을 나타낼 수도 있다. 더욱이, 벡터 파싱 유닛 (214) 은, 일부 경우들에서 전경 US[k] 벡터들 (49) 및 전경 V[k] 벡터들 (51_k) 의 동일한 벡터들 (포지션-와이즈) 을 선택하여, 음장의 전경 성분들을 나타낼 수도 있다.

[0528] 도 9f 는 오디오 인코딩 디바이스 (20) 의 배경 선택 유닛 (48) 을 더 상세하게 예시하는 블록 다이어그램이다. 배경 선택 유닛 (48) 은 배경 채널 정보 (예를 들어, 배경 음장 (N_{BG}) 및 전송할 추가의 BG HOA 채널들의 인덱스들 (i) 및 수 (nBGa)) 에 기초하여 배경 또는 주변 HOA 계수들 (47) 을 결정할 수도 있다. 예를 들어, N_{BG} 가 1 과 동일한 경우, 배경 선택 유닛 (48) 은 1 이하의 오더를 갖는 오디오 프레임의 각 샘플에 대한 HOA 계수들 (11) 을 선택할 수도 있다. 배경 선택 유닛 (48) 은, 이 예에서 그 후 추가의 BG HOA 계수들과 같은 인덱스들 (i) 중 하나에 의해 식별된 인덱스를 갖는 HOA 계수들 (11) 을 선택할 수도 있고, 여기서 nBGa 는 비트스트림 (21) 에서 지정되도록 비트스트림 생성 유닛 (42) 에 제공되어 오디오 디코딩 디바이스, 예컨대 도 5 의 예에 도시된 오디오 디코딩 디바이스 (24) 가 비트스트림 (21) 으로부터 BG HOA 계수들 (47) 을 파싱하게 한다. 배경 선택 유닛 (48) 은 그 후, 주변 HOA 계수들 (47) 을 에너지 보상 유닛 (38) 으로 출력할 수도 있다. 주변 HOA 계수들 (47) 은 디멘전들 D: $M \times [(N_{BG}+1)^2 + nBGa]$ 을 가질 수도 있다.

[0529] 도 9g 는 오디오 인코딩 디바이스 (20) 의 에너지 보상 유닛 (38) 을 더 상세하게 예시하는 블록 다이어그램이다. 에너지 보상 유닛 (38) 은 주변 HOA 계수들 (47) 에 대하여 에너지 보상을 수행하여 배경 선택 유닛 (48) 에 의해 HOA 채널들 중 다양한 채널들의 제거로 인한 에너지 손실을 보상하도록 구성된 유닛을 나타낼 수도 있다. 에너지 보상 유닛 (38) 은 에너지 결정 유닛 (218), 에너지 분석 유닛 (220) 및 에너지 증폭 유닛 (222) 을 포함할 수도 있다.

[0530] 에너지 결정 유닛 (218) 은 리오더링된 US[k] 행렬 (33') 및 리오더링된 V[k] 행렬 (35') 중 하나 이상의 각각의 행 및/또는 열에 대한 RMS 를 식별하도록 구성된 유닛을 나타낼 수도 있다. 에너지 결정 유닛 (38) 은 또한, nFG 신호들 (49) 및 전경 V[k] 벡터들 (51_k) 을 포함할 수도 있는 선택된 전경 채널들, 및 오더-감소된 주변 HOA 계수들 (47) 중 하나 이상의 각각의 행 및/또는 열에 대한 RMS 를 식별할 수도 있다. 리오더링된 US[k] 행렬 (33') 및 리오더링된 V[k] 행렬 (35') 의 하나 이상의 각각의 행 및/또는 열에 대한 RMS 는 RMS_{FULL} 로 표기된 벡터에 저장될 수도 있는 한편, nFG 신호들 (49), 전경 V[k] 벡터들 (51_k), 및 오더-감소된 주변 HOA 계수들 (47) 중 하나 이상의 각각의 행 및/또는 열에 대한 RMS 는 벡터 $RMS_{REDUCED}$ 로 표기된 벡터에 저장될 수도 있다.

[0531] 일부 예들에서, 리오더링된 US[k] 행렬 (33'), 리오더링된 V[k] 행렬 (35'), nFG 신호들 (49), 전경 V[k] 벡터들 (51_k), 및 오더-감소된 주변 HOA 계수들 (47) 중 하나 이상의 각각의 행들 및/또는 열들의 각각의 RMS 를 결정하기 위해, 에너지 결정 유닛 (218) 은 먼저, 참조 구면 조화 계수들 (SHC) 렌더러를 그 열들에 적용할 수도 있다. 에너지 결정 유닛 (218) 에 의한 참조 SHC 렌더러의 적용은, SHC 도메인에서 RMS 의 결정을 허용하여, 리오더링된 US[k] 행렬 (33'), 리오더링된 V[k] 행렬 (35'), nFG 신호들 (49), 전경 V[k] 벡터들 (51_k), 및 오더-감소된 주변 HOA 계수들 (47) 중 하나 이상의 행들 및/또는 열들로 표현된 프레임의 각각의 행 및/또는 열에 의해 설명된 전체 음장의 에너지를 결정하는 것을 허용한다. 에너지 결정 유닛 (38) 은 이 RMS_{FULL} 및 $RMS_{REDUCED}$ 벡터들을 에너지 분석 유닛 (220) 으로 패스할 수도 있다.

[0532] 에너지 분석 유닛 (220) 은 다음의 식: $Z = RMS_{FULL}/RMS_{REDUCED}$ 에 따라 증폭 값 벡터 (Z) 를 컴퓨팅하도록 구성된 유닛을 나타낼 수도 있다. 에너지 분석 유닛 (220) 은 그 후, 이 증폭 값 벡터 (Z) 를 에너지 증폭 유닛 (222) 으로 패스할 수도 있다. 에너지 증폭 유닛 (222) 은 nFG 신호들 (49), 전경 V[k] 벡터들 (51_k), 및 오더-감소된 주변 HOA 계수들 (47) 중 하나 이상에 이 증폭 값 벡터 (Z) 또는 그 다양한 부분들을 적용하도록 구성된 유닛을 나타낼 수도 있다. 일부 경우들에서, 증폭 값 벡터 (Z) 는 다음의 식 $HOA_{BG_RED}' = HOA_{BG_RED}Z^T$

마다 단지 오더-감소된 주변 HOA 계수들 (47) 에 적용되고, 여기서 HOA_{BG-RED} 는 오더-감소된 주변 HOA 계수들 (47) 을 가리키고, HOA_{BG-RED}' 는 에너지 보상된, 감소된 주변 HOA 계수들 (47') 을 가리키며 Z^T 은 Z 벡터의 트랜스포즈를 가리킨다.

[0533] 도 9h 는 도 4 의 예에 도시된 오디오 인코딩 디바이스 (20) 의 공간-시간 보간 유닛 (50) 을 더 상세하게 예시하는 블록 다이어그램이다. 공간-시간 보간 유닛 (50) 은 k'번째 프레임에 대한 전경 V[k] 벡터들 (51_k) 및 이전 프레임 (따라서, k-1 표기) 에 대한 전경 V[k-1] 벡터들 (5_{k-1}) 을 수신하고, 공간-시간 보간을 수행하여 보간된 전경 V[k] 벡터들을 생성하도록 구성된 유닛을 나타낼 수도 있다. 공간-시간 보간 유닛 (50) 은 V 보간 유닛 (224) 및 전경 적응화 유닛 (226) 을 포함할 수도 있다.

[0534] V 보간 유닛 (224) 은 현재 전경 V[k] 벡터들 (51_k) 및 이전 전경 V[k-1] 벡터들 (51_{k-1}) 의 나머지 부분들에 기초하여 보간하도록 현재 전경 V[k] 벡터들 (51_k) 의 부분을 선택할 수도 있다. V 보간 유닛 (224) 은 전술된 서브-프레임들 중 하나 이상인 부분 또는 프레임 단위로 변할 수도 있는 단일의 정의되지 않은 부분만을 선택할 수도 있다. V 보간 유닛 (224) 은, 일부 경우들에서 현재 전경 V[k] 벡터들 (51_k) 의 1024 개의 샘플들 중 단일의 128 개의 샘플 부분을 선택하여, 보간할 수도 있다. V 보간 유닛 (224) 은 그 후, 현재 전경 V[k] 벡터들 (51_k) 및 이전의 전경 V[k-1] 벡터들 (51_{k-1}) 에서의 벡터들 각각을, (T-설계 행렬과 같은 투영 행렬을 사용하는) 구체 (sphere) 위로 벡터들을 투영함으로써 구별되는 공간 맵들로 컨버팅한다. V 보간 유닛 (224) 은 그 후, 구체 상의 형상들로서 V 에서의 벡터들을 해석할 수도 있다. 256 개의 샘플 부분에 대한 V 행렬들을 보간하기 위해, V 보간 유닛 (224) 은 그 후, 이들 공간 형상들을 보간하고, 그 후 그들을 투영 행렬의 역을 통해 구면 조화 도메인 벡터들로 다시 변환할 수도 있다. 본 개시물의 기술들은, 이 방식으로 V 행렬들 간의 평활한 트랜지션을 제공할 수도 있다. V 보간 유닛 (224) 은 그 후, 전경 V[k] 벡터들 (51_k) 의 보간된 부분을 제거하도록 변경된 후에 전경 V[k] 벡터들 (51_k) 을 나타내는, 나머지 V[k] 벡터들 (53) 을 생성할 수도 있다. V 보간 유닛 (224) 은 그 후, 보간된 전경 V[k] 벡터들 ($51_k'$) 을 nFG 적응 유닛 (226) 으로 패스할 수도 있다.

[0535] 보간할 단일의 부분을 선택하는 경우, V 보간 유닛 (224) 은 CodedSpatialInterpolationTime (254) 로 표기된 신택스 엘리먼트를 생성할 수도 있고, 이것은 지속기간, 또는 다시 말해 (예를 들어, 다수의 샘플들의 관점에서) 보간의 시간을 식별한다. 서브-프레임 보간을 수행하는 것의 단일 부분을 선택하는 경우, V 보간 유닛 (224) 은 또한, SpatialInterpolationMethod (255) 로 표기된 다른 신택스 엘리먼트를 생성할 수도 있고, 이것은 수행된 보간의 유형 (또는, 일부 경우들에서 보간이 수행되었는지 아닌지의 여부) 을 식별할 수도 있다. 공간-시간 보간 유닛 (50) 은 이들 신택스 엘리먼트들 (254 및 255) 을 비트스트림 생성 유닛 (42) 으로 출력할 수도 있다.

[0536] nFG 적응 유닛 (226) 은 적응된 nFG 신호들 (49') 을 생성하도록 구성된 유닛을 나타낼 수도 있다. nFG 적응 유닛 (226) 은 전경 V[k] 벡터들 (51_k) 에 의한 nFG 신호들 (49) 의 곱셈을 통해 전경 HOA 계수들을 먼저 획득함으로써, 적응된 nFG 신호들 (49') 을 생성할 수도 있다. 전경 HOA 계수들을 획득한 후에, nFG 적응 유닛 (226) 은 전경 HOA 계수들을 보간된 전경 V[k] 벡터들 (53) 로 나누어, 적응된 nFG 신호들 (49')(이것은, 보간된 전경 V[k] 벡터들 ($51_k'$) 로부터 이들 신호들이 도출되는 경우 보간된 nFG 신호들 (49') 로서 지칭될 수도 있음) 을 획득할 수도 있다.

[0537] 도 9i 는 도 4 의 예에 도시된 오디오 인코딩 디바이스 (20) 의 계수 감소 유닛 (46) 을 더 상세하게 예시하는 블록 다이어그램이다. 계수 감소 유닛 (46) 은 배경 채널 정보 (43) 에 기초하여 나머지 전경 V[k] 벡터들 (53) 에 대하여 계수 감소를 수행하여, 감소된 전경 V[k] 벡터들 (55) 을 양자화 유닛 (52) 으로 출력하도록 구성된 유닛을 나타낼 수도 있다. 감소된 전경 V[k] 벡터들 (55) 은 디멘전들 D: $[(N+1)^2 - (N_{BG}+1)^2\text{-nBGa}] \times$ nFG 을 가질 수도 있다.

[0538] 계수 감소 유닛 (46) 은 계수 최소화 유닛 (228) 을 포함할 수도 있고, 이것은 (배경 채널 정보 (43) 에 의해 식별된 바와 같은) 배경 HOA 계수들 (47) 에 대해 고려되는 임의의 계수들을 제거함으로써 나머지 전경 V[k] 벡터들 (53) 각각의 사이즈를 감소 또는 다르게는 최소화하도록 구성된 유닛을 나타낼 수도 있다. 계수 최소화 유닛 (228) 은 배경 채널 정보 (43) 에 의해 식별된 이들 계수들을 제거하여 감소된 전경 V[k] 벡터들 (55)

을 획득할 수도 있다.

[0539] 도 9j 는 도 4 의 예에 도시된 오디오 인코딩 디바이스 (20) 의 심리 음향 오디오 코더 유닛 (40) 을 더 상세하게 예시하는 블록 다이어그램이다. 심리 음향 오디오 코더 유닛 (40) 은 에너지 보상된 배경 HOA 계수들 (47') 및 보간된 nFG 신호들 (49') 에 대하여 심리 음향 인코딩을 수행하도록 구성된 유닛을 나타낼 수도 있다.

도 9h 의 예에 도시된 바와 같이, 심리 음향 오디오 코더 유닛 (40) 은, 에너지 보상된 배경 HOA 계수들 (47') 의 채널들 (여기서, 채널은 이 컨텍스트에서 특정 오더/서브-오더 구면 기저 함수에 대응하는 프레임에서의 샘플들의 전부에 대한 계수들을 지칭함) 각각 및 보간된 nFG 신호들 (49') 의 각각의 신호를 오디오 인코딩하도록 심리 음향 오디오 인코더들 (40A-40N) 의 다수의 경우들을 인보크할 수도 있다. 일부 예들에서, 심리 음향 오디오 코더 유닛 (40) 은 에너지 보상된 배경 HOA 계수들 (47')(또는 nBGa 플러스 인덱스들 (i) 의 총 수) 의 각각의 채널 및 총 nBGa 에 대한 보간된 nFG 신호들 (49')(또는 nFG) 플러스 추가의 주변 HOA 채널들의 인덱스들 (i) 의 총 수 플러스 nFG 의 각 신호를 별개로 인코딩하기에 충분한 수의 오디오 인코더들 (40A-40N) 을 (하드웨어에서 구현되는 경우) 개시하거나 다르게는 포함한다. 오디오 인코더들 (40A-40N) 은 인코딩된 배경 HOA 계수들 (59) 및 인코딩된 nFG 신호들 (61) 을 출력할 수도 있다.

[0540] 도 9k 는 도 4 의 예에 도시된 오디오 인코딩 디바이스 (20) 의 양자화 유닛 (52) 을 더 상세하게 예시하는 블록 다이어그램이다. 도 9k 의 예에서, 양자화 유닛 (52) 은 균일 양자화 유닛 (230), n비트들 유닛 (232), 예측 유닛 (234), 예측 모드 유닛 (236)("Pred Mode Unit 236"), 카테고리 및 잔차 코딩 유닛 (238), 및 허프만 테이블 선택 유닛 (240) 을 포함한다. 균일 양자화 유닛 (230) 은 (감소된 전경 V[k] 벡터들 (55) 중 어느 하나를 나타낼 수도 있는) 공간 성분들 중 하나에 대하여 전술된 균일한 양자화를 수행하도록 구성된 유닛을 나타낸다. n비트들 유닛 (232) 은 *nbits* 파라미터 또는 값을 결정하도록 구성된 유닛을 나타낸다.

[0541] 예측 유닛 (234) 은 양자화된 공간 성분에 대하여 예측을 수행하도록 구성된 유닛을 나타낸다. 예측 유닛 (234) 은 (감소된 전경 V[k-1] 벡터들 (55) 로서 표기될 수도 있는) 감소된 전경 V[k] 벡터들 (55) 중 시간적으로 후속의 대응하는 벡터에 의한, 감소된 전경 V[k] 벡터들 (55) 중 현재 벡터의 엘리먼트-와이즈 차감을 수행함으로써 예측을 수행할 수도 있다. 이 예측의 결과는 예측된 공간 성분로서 지칭될 수도 있다.

[0542] 예측 모드 유닛 (236) 은 예측 모드를 선택하도록 구성된 유닛을 나타낼 수도 있다. 허프만 테이블 선택 유닛 (240) 은 *cid* 의 코딩을 위해 적합한 허프만 테이블을 선택하도록 구성된 유닛을 나타낼 수도 있다. 예측 모드 유닛 (236) 및 허프만 테이블 선택 유닛 (240) 은 다음의 의사-코드에 따라 일 예로서 동작할 수도 있다:

[0543] 소정의 n비트들에 대해, n비트들을 갖는 허프만 테이블들 전부를 취출한다

[0544] B00 = 0; B01 = 0; B10 = 0; B11 = 0; // 코딩 모드 마다 예상된 비트들을 컴퓨팅하도록 초기화

[0545] m=1 에 대해 : (벡터에서의 엘리먼트들의 #)

[0546] // 벡터 엘리먼트에 대한 비트들의 예상된 수를 계산

[0547] // 예측 없이 그리고 허프만 테이블 5 를 사용

[0548] B00 = B00 + calculate_bits(v(m), HT5);

[0549] // 예측 없이 그리고 허프만 테이블 {1,2,3} 을 사용

[0550] B01 = B01 + calculate_bits(v(m), HTq); q in {1,2,3}

[0551] // 예측 잔차 e(m) 에 대한 비트들의 예상된 수를 계산

[0552] e(m) = v(m) - vp(m); // vp(m): 이전 프레임 벡터 엘리먼트

[0553] // 예측으로 그리고 허프만 테이블 4 를 사용

[0554] B10 = B10 + calculate_bits(e(m), HT4);

[0555] // 예측으로 그리고 허프만 테이블 5 를 사용

[0556] B11 = B11 + calculate_bits(e(m), HT5);

[0557] end

[0558] // 최소 비트들을 산출하는 허프만 테이블 및 최선의 예측 모드 찾기

[0559] // 최선의 예측 모드 및 허프만 테이블은 pflag 및 Htflag 각각에 의해 플래그됨

[0560] [Be, id] = min([B00 B01 B10 B11]);

[0561] Switch id

[0562] case 1 : pflag = 0; HTflag = 0;

[0563] case 2: pflag = 0; HTflag = 1;

[0564] case 3: pflag = 1; HTflag = 0;

[0565] case 4: pflag = 1; HTflag = 1;

[0566] end

[0567] 카테고리 및 잔차 코딩 유닛 (238) 은 상기에서 더 상세하게 설명된 방식으로 예측된 공간 성분 또는 (예측이 디스에이블되는 경우) 양자화된 공간 성분의 카테고리화 및 잔차 코딩을 수행하도록 구성된 유닛을 나타낼 수도 있다.

[0568] 도 9k 의 예에 도시된 바와 같이, 양자화 유닛 (52) 은 (비트스트림 (21) 으로부터 분리된 비트스트림 그 자체일 수도 있는) 사이드 정보 또는 비트스트림 (21) 중 어느 하나의 포함을 위해 다양한 파라미터들 또는 값들을 출력할 수도 있다. 사이드 채널 정보에서 정보가 지정된다고 가정하면, 스칼라/엔트로피 양자화 유닛 (50) 은 *nbits* 값을 *nbits* 값 (233) 으로서, 예측 모드를 예측 모드 (237) 로서 그리고 허프만 테이블 정보를 허프만 테이블 정보 (241) 로서, 이 예에서 *cid*, 사인 비트, 및 블록 코딩된 잔차를 인코딩하기 위해 선택된 허프만 코드로서 지칭될 수도 있는, (도 4 의 예에서 코딩된 전경 V[k] 벡터들 (57) 로서 도시된) 공간 성분의 압축된 버전과 함께 비트스트림 생성 유닛 (42) 으로 출력할 수도 있다. *nbits* 값은 코딩된 전경 V[k] 벡터들 (57) 의 전부에 대해 사이드 채널 정보에서 일단 지정될 수도 있는 한편, 예측 모드 및 허프만 테이블 정보는 코딩된 전경 V[k] 벡터들 (57) 의 각 벡터에 대해 지정될 수도 있다. 공간 성분의 압축된 버전을 지정하는 비트스트림의 부분은 도 10b 및/또는 도 10c 의 예에서 더 도시된다.

[0569] 도 9l 는 도 4 의 예에 도시된 오디오 인코딩 디바이스 (20) 의 비트스트림 생성 유닛 (42) 을 더 상세하게 예시하는 블록 다이어그램이다. 비트스트림 생성 유닛 (42) 은 메인 채널 정보 생성 유닛 (242) 및 사이드 채널 정보 생성 유닛 (244) 을 포함할 수도 있다. 메인 채널 정보 생성 유닛 (242) 은 리오더와 인덱스들 (205), CodedSpatialInterpolationTime 신택스 엘리먼트 (254), SpatialInterpolationMethod 신택스 엘리먼트 (255), 인코딩된 배경 HOA 계수들 (59), 및 인코딩된 nFG 신호들 (61) 중 전부는 아니더라도, 하나 이상을 포함하는 메인 비트스트림 (21) 을 생성할 수도 있다. 사이드 채널 정보 생성 유닛 (244) 은 *nbits* 값 (233), 예측 모드 (237), 허프만 테이블 정보 (241) 및 코딩된 전경 V[k] 벡터들 (57) 중, 전부는 아니더라도 하나 이상을 포함할 수도 있는 사이드 채널 비트스트림 (21B) 을 생성하도록 구성된 유닛을 나타낼 수도 있다. 비트스트림들 (21 및 21B) 은 비트스트림 (21) 으로서 총괄하여 지칭될 수도 있다. 일부 컨텍스트들에서, 비트스트림 (21) 은 단지 메인 채널 비트스트림 (21) 을 지칭할 수도 있는 한편, 비트스트림 (21B) 은 사이드 채널 정보 (21B) 로서 지칭될 수도 있다.

[0570] 도 10a 내지 도 10o(ii) 는 압축된 공간 성분들을 지정할 수도 있는 비트스트림 또는 사이드 채널 정보의 부분들을 더 상세하게 예시하는 다이어그램들이다. 도 10a 의 예에서, 부분 (250) 은 렌더러 식별자 ("renderer ID") 필드 (251) 및 HOADecoderConfig 필드 (252) 를 포함한다. 렌더러 ID 필드 (251) 는 HOA 컨텐츠의 혼합을 위해 사용되었던 렌더러의 ID 를 저장하는 필드를 나타낼 수도 있다. HOADecoderConfig 필드 (252) 는 HOA 공간 디코더를 초기화하기 위한 정보를 저장하도록 구성된 필드를 나타낼 수도 있다.

[0571] HOADecoderConfig 필드 (252) 는 방향 정보 ("direction info") 필드 (253), CodedSpatialInterpolationTime 필드 (254), SpatialInterpolationMethod 필드 (255), CodedVVecLength 필드 (256) 및 이득 정보 필드 (257) 를 더 포함한다. 방향 정보 필드 (253) 는 방향-기반 합성 디코더를 구성하기 위한 정보를 저장하는 필드를

나타낼 수도 있다. CodedSpatialInterpolationTime 필드 (254) 는 벡터-기반 신호들의 공간-시간 보간의 시간을 저장하는 필드를 나타낼 수도 있다. SpatialInterpolationMethod 필드 (255) 는 벡터-기반 신호들의 공간-시간 보간 동안 적용된 보간 유형의 표시 (indication) 를 저장하는 필드를 나타낼 수도 있다. CodedVVecLength 필드 (256) 는 벡터-기반 신호들을 합성하는데 사용된 송신된 데이터 벡터의 길이를 저장하는 필드를 나타낼 수도 있다. 이득 정보 필드 (257) 는 신호들에 적용된 이득 보정을 나타내는 정보를 저장하는 필드를 나타낸다.

[0572] 도 10b 의 예에서, 부분 (258A) 은 사이드-정보 채널의 부분을 나타내고, 여기서 부분 (258A) 은 다수의 바이트 필드 (260) 및 n비트들 필드 (261) 를 포함하는 프레임 헤더 (259) 를 포함한다. 다수의 바이트 필드 (260) 는 바이트 정렬 필드 (264) 를 위한 0 들을 포함하는 공간 성분들 (v1 내지 vn) 을 지정하기 위해 프레임에 포함된 바이트들의 수를 표현하기 위한 필드를 나타낼 수도 있다. n비트들 필드 (261) 는 공간 성분들 (v1 내지 vn) 을 분해하는데 사용하기 위해 식별된 n비트들 값을 지정할 수도 있는 필드를 나타낸다.

[0573] 도 10b 의 예에 추가로 도시된 바와 같이, 부분 (258A) 은 예측 모드 필드 (262), 허프만 테이블 정보 필드 (263) 및 압축된 공간 성분들 (v1 내지 vn) 중 대응하는 성분을 포함하는 v1 내지 vn 각각에 대한 서브-비트스트림들을 포함할 수도 있다. 예측 모드 필드 (262) 는 압축된 공간 성분들 (v1 내지 vn) 중 대응하는 성분에 대하여 예측이 수행되었는지 여부의 표시를 저장하는 필드를 나타낼 수도 있다. 허프만 테이블 정보 필드 (263) 는, 압축된 공간 성분들 (v1 내지 vn) 중 대응하는 성분의 다양한 양태들을 디코딩하는데 어느 허프만 테이블이 사용될지를 적어도 부분적으로 표시하기 위한 필드를 나타낸다.

[0574] 이 점에 있어서, 기술들은 오디오 인코딩 디바이스 (20) 로 하여금 음장의 공간 성분의 압축된 버전을 포함하는 비트스트림을 획득하게 할 수도 있고, 이 공간 성분은 복수의 구면 조화 계수들에 대하여 벡터 기반 합성을 수행함으로써 생성된다.

[0575] 도 10c 는 압축된 공간 성분들을 더 상세하게 지정할 수도 있는 사이드 채널 정보의 부분 (258B) 의 대안의 예를 예시하는 다이어그램이다. 도 10c 의 예에서, 부분 (258B) 은 N비트들 필드 (261) 를 포함하는 프레임 헤더 (259) 를 포함한다. N비트들 필드 (261) 는 공간 성분들 (v1 내지 vn) 을 분해하는데 사용하기 위해 식별된 n비트들 값을 지정할 수도 있는 필드를 나타낸다.

[0576] 도 10c 의 예에 추가로 도시된 바와 같이, 부분 (258B) 은, 예측 모드 필드 (262), 허프만 테이블 정보 필드 (263) 및 압축된 공간 성분들 (v1 내지 vn) 중 대응하는 성분을 포함하는 v1 내지 vn 각각에 대한 서브-비트스트림들을 포함할 수도 있다. 예측 모드 필드 (262) 는 압축된 공간 성분들 (v1 내지 vn) 중 대응하는 성분에 대하여 예측이 수행되었는지 여부의 표시를 저장하는 필드를 나타낼 수도 있다. 허프만 테이블 정보 필드 (263) 는, 압축된 공간 성분들 (v1 내지 vn) 중 대응하는 성분의 다양한 양태들을 디코딩하는데 어느 허프만 테이블이 사용될지를 적어도 부분적으로 표시하기 위한 필드를 나타낸다.

[0577] 예시된 예의 N비트들 필드 (261) 는 서브필드들 A (265), B (266), 및 C (267) 를 포함한다. 이 예에서, A (265) 및 B (266) 는 각각 1 비트 서브-필드들인 한편, C (267) 는 2 비트 서브-필드이다. 다른 예들은 상이하게-크기가 정해진 서브-필드들 (265, 266, 및 267) 을 포함할 수도 있다. A 필드 (265) 및 B 필드 (266) 는 N비트들 필드 (261) 의 제 1 및 제 2 최상위 비트들을 저장하는 필드들을 나타낼 수도 있는 한편, C 필드 (267) 는 N비트들 필드 (261) 의 최하위 비트들을 저장하는 필드를 나타낼 수도 있다.

[0578] 부분 (258B) 은 또한, AddAmbHoaInfoChannel 필드 (268) 를 포함할 수도 있다. AddAmbHoaInfoChannel 필드 (268) 는 추가의 주변 HOA 계수들에 대한 정보를 저장하는 필드를 나타낼 수도 있다. 도 10c 의 예에 도시된 바와 같이, AddAmbHoaInfoChannel (268) 는 CodedAmbCoeffIdx 필드 (246), AmbCoeffIdxTransition 필드 (247) 를 포함한다. CodedAmbCoeffIdx 필드 (246) 는 추가의 주변 HOA 계수의 인덱스를 저장하는 필드를 나타낼 수도 있다. AmbCoeffIdxTransition 필드 (247) 는, 이 프레임에서 추가의 주변 HOA 계수가 페이드 인되는지 또는 페이드 아웃되는지 여부를 나타내는 데이터를 저장하도록 구성된 필드를 나타낼 수도 있다.

[0579] 도 10c(i) 는 압축된 공간 성분들을 지정할 수도 있는 사이드 채널 정보의 부분 (258B') 의 대안의 예를 더 상세하게 예시하는 다이어그램이다. 도 10c(i) 의 예에서, 부분 (258B') 은 N비트들 필드 (261) 를 포함하는 프레임 헤더 (259) 를 포함한다. N비트들 필드 (261) 는 공간 성분들 (v1 내지 vn) 을 분해하는데 사용하기 위해 식별된 n비트들 값을 지정할 수도 있는 필드를 나타낸다.

[0580] 도 10c(i) 의 예에 추가로 도시된 바와 같이, 부분 (258B') 은 v1 내지 vn 에 대한 서브-비트스트림들을 포함할 수도 있고, 서브-비트스트림들 각각은 예측 모드 필드 (262) 를 포함하지 않고 압축된 공간 성분들 (v1 내지

vn) 중 대응하는 성분 및 허프만 테이블 정보 필드 (263) 를 포함한다. 모든 다른 양태들에서, 부분 (258B') 은 부분 (258B) 에 유사할 수도 있다.

[0581] 도 10d 는 비트스트림 (21) 의 부분 (258C) 을 더 상세하게 예시하는 다이어그램이다. 부분 (258C) 은, 프레임 헤더 (259) 및 제로 바이트 정렬 (264) 이 제거된 한편, N비트들 (261) 필드가 도 10d 의 예에 도시된 바와 같이 v1-vn 에 대한 비트스트림들 각각 전에 추가된 것을 제외하고, 부분 (258) 과 유사하다.

[0582] 도 10d(i) 는 비트스트림 (21) 의 부분 (258C') 을 더 상세하게 예시하는 다이어그램이다. 부분 (258C') 은, 부분 (258C') 이 V 벡터들 (v1-vn) 각각에 대한 예측 모드 필드 (262) 를 포함하지 않는 것을 제외하고 부분 (258C) 과 유사하다.

[0583] 도 10e 는 비트스트림 (21) 의 부분 (258D) 을 더 상세하게 예시하는 다이어그램이다. 부분 (258D) 은, 프레임 헤더 (259) 및 제로 바이트 정렬 (264) 이 제거된 한편, N비트들 (261) 필드가 도 10e 의 예에 도시된 바와 같이 v1-vn 에 대한 비트스트림들 각각 전에 추가된 것을 제외하고, 부분 (258B) 과 유사하다.

[0584] 도 10e(i) 는 비트스트림 (21) 의 부분 (258D') 을 더 상세하게 예시하는 다이어그램이다. 부분 (258D') 은, 부분 (258D') 이 V 벡터들 (v1-vn) 각각에 대한 예측 모드 필드 (262) 를 포함하지 않는 것을 제외하고 부분 (258D) 과 유사하다. 이 점에 있어서, 오디오 인코딩 디바이스 (20) 는 도 10c(i), 10d(i) 및 10e(i) 에 대하여 예시된 바와 같이, 각각의 압축된 V 벡터에 대한 예측 모드 필드 (262) 를 포함하지 않는 비트스트림 (21) 을 생성할 수도 있다.

[0585] 도 10f 는 도 10a 의 예에 도시된 비트스트림 (21) 의 부분 (250) 을 상이한 방식으로 예시하는 다이어그램이다. 도 10d 의 예에 도시된 부분 (250) 은 HOAOrder 필드 (이것은 예시의 목적을 위해 도 10f 의 예에 도시되지 않음), MinAmbHoaOrder 필드 (이것은 다시, 예시의 목적을 위해 도 10 의 예에 도시되지 않음), 방향 정보 필드 (253), CodedSpatialInterpolationTime 필드 (254), SpatialInterpolationMethod 필드 (255), CodedVVecLength 필드 (256) 및 이득 정보 필드 (257) 를 포함한다. 도 10f 의 예에 도시된 바와 같이, CodedSpatialInterpolationTime 필드 (254) 는 3 개의 비트 필드를 포함할 수도 있고, SpatialInterpolationMethod 필드 (255) 는 1 개의 비트 필드를 포함할 수도 있으며, CodedVVecLength 필드 (256) 는 2 개의 비트 필드를 포함할 수도 있다.

[0586] 도 10g 는 비트스트림 (21) 의 부분 (248) 을 더 상세하게 예시하는 다이어그램이다. 부분 (248) 은 (사이드대역 정보, 사이드 채널 정보, 또는 사이드 채널 비트스트림으로서 또한, 표기될 수도 있는) HOAframe 필드 (249) 를 포함하는 통일된 스피치/오디오 코더 (USAC) 3 차원 (3D) 페이로드를 나타낸다. 도 10e 의 예에 도시된 바와 같이, HOAFrame 필드 (249) 의 확대도는 도 10c 의 예에 도시된 비트스트림 (21) 의 부분 (258B) 과 유사할 수도 있다. "ChannelSideInfoData" 는 예시의 목적을 위해 도 10c 의 예에 도시되지 않은 ChannelType (채널타입) 필드 (269), 도 10e 의 예에서 "ba" 로서 표기된 A 필드 (265), 도 10e 의 예에서 "bb" 로서 표기된 B 필드 (266) 및 도 10e 의 예에서 "unitC" 로서 표기된 C 필드 (267) 를 포함한다. ChannelType 필드는, 그 채널이 방향-기반 신호, 벡터-기반 신호 또는 추가의 주변 HOA 계수인지 여부를 나타낸다. 상이한 ChannelSideInfoData 사이에는, 그레이로 표기된 상이한 V 벡터 비트스트림들 (예를 들어, "v1 에 대한 비트스트림" 및 "v2 에 대한 비트스트림") 을 갖는 AddAmbHoaInfoChannel 필드들 (268) 이 존재한다.

[0587] 도 10h 내지 도 10o(ii) 는 첨부한 HOAconfig 부분들 (250H-250O) 과 함께 다른 다양한 예시의 부분들 (248H-248O) 을 더 상세하게 예시하는 다이어그램들이다. 도 10h(i) 및 10h(ii) 는 상기 의사-코드에서 케이스 0 에 대응하도록 생성되고 있는 첨부한 HOAconfig 부분 (250H) 및 제 1 예시의 비트스트림 (248H) 을 예시한다.

도 10h(i) 의 예에서, HOAconfig 부분 (250H) 은 V 벡터의 모든 엘리먼트들, 예를 들어 모든 16 개의 V 벡터 엘리먼트들이 코딩된 것을 나타내도록 설정된 CodedVVecLength 신택스 엘리먼트 (256) 를 포함한다. HOAconfig 부분 (250H) 은 또한, 공간-시간 보간의 보간 함수가 상향 코사인 (raised cosine) 인 것을 나타내도록 설정된 SpatialInterpolationMethod 신택스 엘리먼트 (255) 를 포함한다. HOAconfig 부분 (250H) 은 또한, 256 의 보간된 샘플 지속기간을 나타내도록 설정된 CodedSpatialInterpolationTime (254) 을 포함한다.

HOAconfig 부분 (250H) 은 주변 HOA 컨텐츠의 MinimumHOA 오더가 1 인 것을 나타내도록 설정된 MinAmbHoaOrder 신택스 엘리먼트 (150) 를 더 포함하고, 여기서 오디오 디코딩 디바이스 (24) 는 $(1+1)^2$ 또는 4 와 동일하도록 MinNumofCoeffsForAmbHOA 신택스 엘리먼트를 도출할 수도 있다. HOAconfig 부분 (250H) 은 컨텐츠의 HOA 오더가 3 (또는 다시 말해, N = 3) 과 동일한 것을 나타내도록 설정된 HoaOrder 신택스 엘리먼트

(152) 를 포함하고, 여기서 오디오 디코딩 디바이스 (24) 는 $(N+1)^2$ 또는 16 과 동일한 것으로 NumOfHoaCoeffs 을 도출할 수도 있다.

[0588] 또한 도 10h(i) 의 예에 보여진 바처럼, 부분 (248H) 은 통일된 스피치 및 오디오 코딩 (USAC) 3 차원 (USAC-3D) 오디오 프레임을 포함하고, 여기서 스펙트럼 대역 복사 (SBR) 가 가능해질 때 하나의 USAC-3D 프레임 내에 2 개의 오디오 프레임들이 저장되는 것을 고려하면 2 개의 HOA 프레임들 (249A 및 249B) 이 USAC 확장 페이로드에 저장된다. 오디오 디코딩 디바이스 (24) 는 numHOATransportChannels 신택스 엘리먼트 및 MinNumOfCoeffsForAmbHOA 신택스 엘리먼트의 함수로서 플렉서블 전송 채널들의 수를 도출할 수도 있다. 다음의 예들에서, numHOATransportChannels 신택스 엘리먼트는 7 과 동일하고 MinNumOfCoeffsForAmbHOA 신택스 엘리먼트가 4 와 동일한 것으로 가정되고, 여기서 플렉서블 전송 채널들의 수는 numHOATransportChannels 신택스 엘리먼트 마이너스 MinNumOfCoeffsForAmbHOA 신택스 엘리먼트 (또는 3) 과 동일하다.

[0589] 도 10h(ii) 는 프레임들 (249A 및 249B) 를 더 상세하게 예시한다. 도 10h(ii) 의 예에 보여진 바처럼, 프레임 (249A) 는 CSID (ChannelSideInfoData) 필드 (154-154C), HOAGCD (HOAGainCorrectionData) 필드, VVectorData 필드들 (156 및 156B) 및 HOAPredictionInfo 필드들을 포함한다. CSID 필드 (154) 는 ChannelType (269) 과 함께 unitC (267), bb (266) 및 ba (265) 를 포함하고, 이들의 각각은 도 10h(i) 의 예에 보여진 대응하는 값들 01, 1, 0 및 01 로 설정된다. CSID 필드 (154B) 는 ChannelType (269) 과 함께 unitC (267), bb (266) 및 ba (265) 를 포함하고, 이들의 각각은 도 10h(ii) 의 예에 보여진 대응하는 값들 01, 1, 0 및 01 로 설정된다. CSID 필드 (154C) 는 3 의 값을 갖는 ChannelType 필드 (269) 를 포함한다.

CSID 필드들 (154-154C) 의 각각은 전송 채널들 (1, 2 및 3) 의 각각의 하나에 대응한다. 사실상, 각각의 CSID 필드 (154-154C) 는, (대응하는 ChannelType 이 0 과 같을 때) 대응하는 페이로드 (156 및 156B) 가 방향 기반 신호들인지, (대응하는 ChannelType 이 1 과 같을 때) 벡터 기반 신호들인지, (대응하는 ChannelType 이 2와 같을 때) 추가적인 주변 HOA 계수인지, 또는 (ChannelType 이 3 과 같을 때) 공 (empty) 인지를 표시한다.

[0590] 도 10h(ii) 의 예에서, 프레임 (249A) 은 (CSID 필드들 (154 및 154B) 에서 ChannelType (269) 이 1 과 같음을 고려하면) 2 개의 벡터 기반 신호들 및 (CSID 필드들 (154C) 에서 ChannelType (269) 이 3 과 같음을 고려하면) 공을 포함한다. 앞서 말한 HOAconfig 부분 (250H) 을 고려하면, 오디오 디코딩 디바이스 (24) 는 모든 16 개 V 벡터 엘리먼트들이 인코딩되는 것을 결정할 수도 있다. 그러므로, VVectorData (156 및 156B) 각각은 모든 16 개 벡터 엘리먼트들을 포함하고, 이들의 각각은 8 비트로 균일하게 양자화된다. 각주 1 에 의해 의해 언급된 바처럼, 코딩된 VVectorData 엘리먼트들의 수 및 인덱스들은 파라미터 CodedVVecLength=0 에 의해 명시된다. 또한, 단일 별표 (*) 에 의해 언급된 바처럼, 코딩 방식은 대응하는 전송 채널에 대해 CSID 필드에서 NbitsQ = 5 에 의해 시그널링된다.

[0591] 프레임 (249B) 에서, CSID 필드 (154 및 154B) 는 프레임 (249) 에 있는 것과 동일한 한편, 프레임 (249B) 의 CSID 필드 (154C) 는 1 의 ChannelType 으로 스위칭된다. 그러므로, 프레임 (249B) 의 CSID 필드 (154C) 는 Cbflag (267), (허프만 인코딩을 표시하는) Pflag (267) 및 (12와 같은) Nbits (261) 를 포함한다. 결과적으로, 프레임 (249B) 는 16 개 V 벡터 엘리먼트들을 포함하는 제 3 VVectorData 필드 (156C) 를 포함하고, 그들의 각각은 12 비트로 균일하게 양자화되고 허프만 코딩된다. 위에 언급된 바처럼, 코딩된 VVectorData 엘리먼트들의 수 및 인덱스들은 파라미터 CodedVVecLength = 0 에 의해 명시되는 한편, 허프만 코딩 방식은 이 특정 전송 채널 (예를 들어, 전송 채널 번호 3) 에 대해 CSID 필드 (154C) 에서 NbitsQ = 12, CbFlag = 0 및 Pflag = 0 에 의해 시그널링된다.

[0592] 도 10i(i) 및 도 10i(ii) 의 예들은 제 2 예시적 비트스트림 (248I) 및 수반하는 HOA config 부분 (250I) 이 위의 의사 코드에서 위에 있는 케이스 0 과 대응하도록 생성된 것을 예시한다. 도 10i(i) 의 예에서, HOAconfig 부분 (250I) 은, V 벡터의 모든 엘리먼트들, 예를 들어 모든 16 개 V 벡터 엘리먼트들이 코딩된다는 것을 표시하도록 설정된 CodedVVecLength 신택스 엘리먼트 (256) 를 포함한다. HOAconfig 부분 (250I) 은 또한, 공간-시간 보간의 보간 함수가 상향 코사인 (raised cosine) 이라는 것을 표시하도록 설정된 SpatialInterpolationMethod 신택스 엘리먼트 (255) 를 포함한다. HOAconfig 부분 (250I) 은 더욱이, 256 의 보간된 샘플 지속시간을 표시하도록 설정된 odedSpatialInterpolationTime (254) 을 포함한다.

[0593] HOAconfig 부분 (250I) 은 주변 HOA 컨텐츠의 MinimumHOA 오더가 1 이라는 것을 표시하도록 설정된 MinAmbHoaOrder 신택스 엘리먼트 (150) 를 더 포함하고, 여기서 오디오 디코딩 디바이스 (24) 는 $(1+1)^2$ 또는 4 와 동일한 것으로 MinNumofCoeffsForAmbHOA 신택스 엘리먼트를 도출할 수도 있다. 오디오 디코딩 디바이스

(24) 는 또한, 이 예에서 16-4 또는 12 와 같은 것으로 가정되는 MinNumOfCoeffsForAmbHOA 와 NumOfHoaCoeff 신택스 엘리먼트 사이의 차이로 설정된 것으로 MaxNoofAddActiveAmbCoeffs 신택스 엘리먼트를 도출할 수도 있다. 오디오 디코딩 디바이스 (24) 는 또한, ceil(log2(MaxNoOfAddActiveAmbCoeffs)) = ceil(log2(12)) = 4 으로 설정된 것으로 AmbAsignmBits 신택스 엘리먼트를 도출할 수도 있다. HOAconfig 부분 (250H) 은, 3 과 동일한 것으로 컨텐츠의 HOA 오더 (또는 다른 말로, N = 3) 를 표시하도록 설정된 HoaOrder 신택스 엘리먼트 (152) 를 포함하고, 여기서 오디오 디코딩 디바이스 (24) 는 $(N + 1)^2$ 또는 16 와 동일한 것으로 NumOfHoaCoeffs 을 도출할 수도 있다.

[0594] 또한 도 10i(i) 의 예에 보여진 바처럼, 부분 (248H) 은 USAC-3D 오디오 프레임을 포함하고, 여기서 스펙트럼 대역 복사 (SBR) 가 가능해질 때 하나의 USAC-3D 프레임 내에 2 개의 오디오 프레임들이 저장되는 것을 고려하면 2 개의 HOA 프레임들 (249C 및 249D) 이 USAC 확장 페이로드에 저장된다. 오디오 디코딩 디바이스 (24) 는 numHOATransportChannels 신택스 엘리먼트 및 MinNumOfCoeffsForAmbHOA 신택스 엘리먼트의 함수로서 플렉서블 전송 채널들의 수를 도출할 수도 있다. 다음의 예들에서, numHOATransportChannels 신택스 엘리먼트는 7 과 동일하고 MinNumOfCoeffsForAmbHOA 신택스 엘리먼트가 4 와 동일한 것으로 가정되고, 여기서 플렉서블 전송 채널들의 수는 numHOATransportChannels 신택스 엘리먼트 마이너스 MinNumOfCoeffsForAmbHOA 신택스 엘리먼트 (또는 3) 과 동일하다.

[0595] 도 10i(ii) 는 프레임들 (249C 및 249D) 를 더 상세하게 예시한다. 도 10i(ii) 의 예에 보여진 바처럼, 프레임 (249C) 은 CSID 필드들 (154-154C) 및 VVectorData 필드들 (156) 을 포함한다. CSID 필드 (154) 는 CodedAmbCoeffIdx (246), AmbCoeffIdxTransition (247) (여기서 이중 별표 (**) 는, 플렉서블 전송 채널 번호 1 에 대하여, 디코더의 내부 상태가 여기서 AmbCoeffIdxTransitionState = 2 인 것으로 가정되고, 이는 CodedAmbCoeffIdx 비트필드가 시그널링되거나 또는 그렇지 않으면 비트스트림에서 명시되는 것에 귀결된다), 및 ChannelType (269) (이는 2 와 같으며, 대응하는 페이로드가 추가적인 주변 HOA 계수라는 것을 시그널링한다) 를 포함한다. 오디오 디코딩 디바이스 (24) 는 CodedAmbCoeffIdx+1+MinNumOfCoeffsForAmbHOA 와 같은 것으로서 또는 이 예에서 5 로서 AmbCoeffIdx 를 도출할 수도 있다. CSID 필드 (154B) 는 ChannelType (269) 과 함께 unitC (267), bb (266) 및 ba (265) 를 포함하고, 이들의 각각은 도 10i(ii) 의 예에 보여진 대응하는 값들 01, 1, 0 및 01 로 설정된다. CSID 필드 (154C) 는 3 의 값을 갖는 ChannelType 필드 (269) 를 포함한다.

[0596] 도 10i(ii) 의 예에서, 프레임 (249C) 은 (CSID 필드들 (154B) 에서 ChannelType (269) 이 1 과 같음을 고려하면) 특이 벡터 기반 신호 및 (CSID 필드들 (154C) 에서 ChannelType (269) 이 3 과 같음을 고려하면) 공을 포함한다. 앞서 말한 HOAconfig 부분 (250I) 을 고려하면, 오디오 디코딩 디바이스 (24) 는 모든 16 개 V 벡터 엘리먼트들이 인코딩되는 것을 결정할 수도 있다. 그러므로, VVectorData (156) 는 모든 16 개 벡터 엘리먼트들을 포함하고, 이들의 각각은 8 비트로 균일하게 양자화된다. 각주 1 에 의해 의해 언급된 바처럼, 코딩된 VVectorData 엘리먼트들의 수 및 인덱스들은 파라미터 CodedVVecLength=0 에 의해 명시된다. 더욱이, 각주 2 에 의해 언급된 바처럼, 코딩 방식은 대응하는 전송 채널에 대해 CSID 필드에서 NbitsQ = 5 에 의해 시그널링된다.

[0597] 프레임 (249D) 에서, CSID 필드 (154) 는, 천이가 일어나지 않았다는 것을 표시하는 AmbCoeffIdxTransition (247) 를 포함하고 따라서 CodedAmbCoeffIdx (246) 이 이전 프레임으로부터 암시될 수도 있고 시그널링되거나 또는 그렇지 않으면 다시 명시될 필요가 없다. 프레임 (249D) 의 CSID 필드 (154B 및 154C) 는 프레임 (249C) 에 대한 것과 동일하고 따라서, 프레임 (249C) 과 같이, 프레임 (249D) 은 단일 VVectorData 필드 (156) 를 포함하고, 이는 모든 16 개의 벡터 엘리먼트들을 포함하며, 이들의 각각은 8 비트로 균일하게 양자화된다.

[0598] 도 10j(i) 및 도 10j(ii) 는 제 1 예의 비트스트림 (248J) 및 수반하는 HOA config 부분 (250J) 이 위의 의사 코드에서 케이스 1 과 대응하도록 생성된 것을 예시한다. 도 10j(i) 의 예에서, HOAconfig 부분 (250J) 은, 엘리먼트들 1 내지 MinNumOfCoeffsForAmbHOA 신택스 엘리먼트들 및 (이 예에서 0 으로 가정된) ContAddAmbHoaChan 신택스 엘리먼트에 명시된 그러한 엘리먼트들을 제외하고, V 벡터의 모든 엘리먼트들이 코딩되는 것을 나타내도록 설정된 CodedVVecLength 신택스 엘리먼트 (256) 를 포함한다. HOAconfig 부분 (250J) 은 또한, 공간-시간 보간의 보간 함수가 상향 코사인이라는 것을 표시하도록 설정된 SpatialInterpolationMethod 신택스 엘리먼트 (255) 를 포함한다. HOAconfig 부분 (250J) 은 더욱이, 256 의 보간된 샘플 지속시간을 표시하도록 설정된 CodedSpatialInterpolationTime (254) 을 포함한다.

HOAconfig 부분 (250J) 은 주변 HOA 컨텐츠의 MinimumHOA 오더가 1 이라는 것을 표시하도록 설정된 MinAmbHoaOrder 신택스 엘리먼트 (150) 를 더 포함하고, 여기서 오디오 디코딩 디바이스 (24) 는 $(1+1)^2$ 또는 4 와 동일한 것으로 MinNumofCoeffsForAmbHOA 신택스 엘리먼트를 도출할 수도 있다. HOAconfig 부분 (250J) 는, 3 과 동일한 것으로 컨텐츠의 HOA 오더 (또는 다른 말로, N = 3) 를 표시하도록 설정된 HoaOrder 신택스 엘리먼트 (152) 를 포함하고, 여기서 오디오 디코딩 디바이스 (24) 는 $(N + 1)^2$ 또는 16 와 동일한 것으로 NumOfHoaCoeffs 을 도출할 수도 있다.

[0599] 또한 도 10j(i) 의 예에 보여진 바처럼, 부분 (248J) 은 USAC-3D 오디오 프레임을 포함하고, 여기서 스펙트럼 대역 복사 (SBR) 가 가능해질 때 하나의 USAC-3D 프레임 내에 2 개의 오디오 프레임들이 저장되는 것을 고려하면 2 개의 HOA 프레임들 (249E 및 249F) 이 USAC 확장 페이로드에 저장된다. 오디오 디코딩 디바이스 (24) 는 numHOATransportChannels 신택스 엘리먼트 및 MinNumOfCoeffsForAmbHOA 신택스 엘리먼트의 함수로서 플렉서블 전송 채널들의 수를 도출할 수도 있다. 다음의 예들에서, numHOATransportChannels 신택스 엘리먼트가 7 과 동일하고 MinNumOfCoeffsForAmbHOA 신택스 엘리먼트가 4 와 동일한 것으로 가정되고, 여기서 플렉서블 전송 채널들의 수는 numHOATransportChannels 신택스 엘리먼트 마이너스 MinNumOfCoeffsForAmbHOA 신택스 엘리먼트 (또는 3) 과 동일하다.

[0600] 도 10j(ii) 는 프레임들 (249E 및 249F) 을 더 상세하게 예시한다. 도 10j(ii) 의 예에 보여진 바처럼, 프레임 (249E) 은 CSID 필드들 (154-154C) 및 VVectorData 필드들 (156 및 156B) 을 포함한다. CSID 필드 (154) 는 ChannelType (269) 과 함께 unitC (267), bb (266) 및 ba (265) 를 포함하고, 이들의 각각은 도 10j(i) 의 예에 보여진 대응하는 값들 01, 1, 0 및 01 로 설정된다. CSID 필드 (154B) 는 ChannelType (269) 과 함께 unitC (267), bb (266) 및 ba (265) 를 포함하고, 이들의 각각은 도 10j(ii) 의 예에 보여진 대응하는 값들 01, 1, 0 및 01 로 설정된다. CSID 필드 (154C) 는 3 의 값을 갖는 ChannelType 필드 (269) 를 포함한다. CSID 필드들 (154-154C) 의 각각은 전송 채널들 (1, 2 및 3) 의 각각의 하나에 대응한다.

[0601] 도 10j(ii) 의 예에서, 프레임 (249E) 은 (CSID 필드들 (154 및 154B) 에서 ChannelType (269) 이 1 과 같음을 고려하면) 2 개의 벡터 기반 신호들 및 (CSID 필드들 (154C) 에서 ChannelType (269) 이 3 과 같음을 고려하면) 공 (empty) 을 포함한다. 앞서 말한 HOAconfig 부분 (250H) 을 고려하면, 오디오 디코딩 디바이스 (24) 는, 모든 12 개 V 벡터 엘리먼트들이 인코딩되는 것을 결정할 수도 있다 (여기서 12는 $(HOAOrder + 1)^2 - (MinNumOfCoeffsForAmbHOA) - (ContAddAmbHoaChan) = 16-4-0 = 12$ 로서 도출된다). 그러므로, VVectorData (156 및 156B) 각각은 모든 12 개 벡터 엘리먼트들을 포함하고, 이들의 각각은 8 비트로 균일하게 양자화된다. 각주 1 에 의해 의해 언급된 바처럼, 코딩된 VVectorData 엘리먼트들의 수 및 인덱스들은 파라미터 CodedVVecLength=0 에 의해 명시된다. 또한, 단일 별표 (*) 에 의해 언급된 바처럼, 코딩 방식은 대응하는 전송 채널에 대해 CSID 필드에서 NbitsQ = 5 에 의해 시그널링된다.

[0602] 프레임 (249F) 에서, CSID 필드 (154 및 154B) 는 프레임 (249E) 에 있는 것과 동일한 한편, 프레임 (249F) 의 CSID 필드 (154C) 는 1 의 ChannelType 으로 스위칭된다. 그러므로, 프레임 (249B) 의 CSID 필드 (154C) 는 Cbflag (267), (허프만 인코딩을 표시하는) Pflag (267) 및 (12와 같은) Nbits (261) 를 포함한다. 결과적으로, 프레임 (249F) 는 12 개 V 벡터 엘리먼트들을 포함하는 제 3 VVectorData 필드 (156C) 를 포함하고, 그들의 각각은 12 비트로 균일하게 양자화되고 허프만 코딩된다. 위에 언급된 바처럼, 코딩된 VVectorData 엘리먼트들의 수 및 인덱스들은 파라미터 CodedVVecLength = 0 에 의해 명시되는 한편, 허프만 코딩 방식은 이 특정 전송 채널 (예를 들어, 전송 채널 번호 3) 에 대해 CSID 필드 (154C) 에서 NbitsQ = 12, CbFlag = 0 및 Pflag = 0 에 의해 시그널링된다.

[0603] 도 10k(i) 및 도 10k(ii) 의 예는 제 2 예의 비트스트림 (248K) 및 수반하는 HOA config 부분 (250K) 이 위의 의사 코드에서 케이스 1 과 대응하도록 생성된 것을 예시한다. 도 10K(i) 의 예에서, HOAconfig 부분들 (250K) 은, 엘리먼트들 1 내지 MinNumOfCoeffsForAmbHOA 신택스 엘리먼트들 및 (이 예에서 1 로 가정되는) ContAddAmbHoaChan 신택스 엘리먼트에 명시된 그러한 엘리먼트들을 제외하고, V 벡터의 모든 엘리먼트들이 코딩되는 것을 나타내도록 설정된 CodedVVecLength 신택스 엘리먼트 (256) 를 포함한다. HOAconfig 부분 (250K) 은 또한, 공간-시간 보간의 보간 함수가 상향 코사인이라는 것을 표시하도록 설정된 SpatialInterpolationMethod 신택스 엘리먼트 (255) 를 포함한다. HOAconfig 부분 (250K) 은 더욱이, 256 의 보간된 샘플 지속시간을 표시하도록 설정된 CodedSpatialInterpolationTime (254) 을 포함한다.

[0604] HOAconfig 부분 (250K) 은 주변 HOA 컨텐츠의 MinimumHOA 오더가 1 이라는 것을 표시하도록 설정된

MinAmbHoaOrder 신택스 엘리먼트 (150) 를 더 포함하고, 여기서 오디오 디코딩 디바이스 (24) 는 $(1+1)^2$ 또는 4 와 동일한 것으로 MinNumofCoeffsForAmbHOA 신택스 엘리먼트를 도출할 수도 있다. 오디오 디코딩 디바이스 (24) 는 또한, 이 예에서 16-4 또는 12 인 것으로 가정되는 MinNumOfCoeffsForAmbHOA 와 NumOfHoaCoeff 신택스 엘리먼트 사이의 차이로 설정된 것으로서 MaxNoOfAddActiveAmbCoeffs 신택스 엘리먼트를 도출할 수도 있다. 오디오 디코딩 디바이스 (24) 는 또한, ceil(log2(MaxNoOfAddActiveAmbCoeffs)) = ceil(log2(12)) = 4 으로 설정된 것으로서 AmbAsignmBits 신택스 엘리먼트를 도출할 수도 있다. HOAconfig 부분 (250K) 는, 3 과 동일한 것으로 컨텐츠의 HOA 오더 (또는 다른 말로, N = 3) 를 표시하도록 설정된 HoaOrder 신택스 엘리먼트 (152) 를 포함하고, 여기서 오디오 디코딩 디바이스 (24) 는 $(N + 1)^2$ 또는 16 와 동일한 것으로 NumOfHoaCoeffs 을 도출할 수도 있다.

[0605] 또한 도 10k(i) 의 예에 보여진 바처럼, 부분 (248K) 은 USAC-3D 오디오 프레임을 포함하고, 여기서 스펙트럼 대역 복사 (SBR) 가 가능해질 때 하나의 USAC-3D 프레임 내에 2 개의 오디오 프레임들이 저장되는 것을 고려하면 2 개의 HOA 프레임들 (249G 및 249H) 이 USAC 확장 페이로드에 저장된다. 오디오 디코딩 디바이스 (24) 는 numHOATransportChannels 신택스 엘리먼트 및 MinNumOfCoeffsForAmbHOA 신택스 엘리먼트의 함수로서 플렉서블 전송 채널들의 수를 도출할 수도 있다. 다음의 예들에서, numHOATransportChannels 신택스 엘리먼트가 7 과 동일하고 MinNumOfCoeffsForAmbHOA 신택스 엘리먼트가 4 와 동일한 것으로 가정되고, 여기서 플렉서블 전송 채널들의 수는 numHOATransportChannels 신택스 엘리먼트 마이너스 MinNumOfCoeffsForAmbHOA 신택스 엘리먼트 (또는 3) 과 동일하다.

[0606] 도 10k(ii) 는 프레임들 (249G 및 249H) 을 더 상세하게 예시한다. 도 10k(ii) 의 예에 보여진 바처럼, 프레임 (249G) 는 CSID 필드들 (154-154C) 및 VVectorData 필드들 (156) 을 포함한다. CSID 필드 (154) 는 CodedAmbCoeffIdx (246), AmbCoeffIdxTransition (247) (여기서 이중 별표 (**) 는, 플렉서블 전송 채널 번호 1 에 대하여, 디코더의 내부 상태가 여기서 AmbCoeffIdxTransitionState = 2 인 것으로 가정되고, 이는 CodedAmbCoeffIdx 비트필드가 시그널링되거나 또는 그렇지 않으면 비트스트림에서 명시되는 것에 귀결된다), 및 ChannelType (269) (이는 2 와 같으며, 대응하는 페이로드가 추가적인 주변 HOA 계수라는 것을 시그널링한다) 를 포함한다. 오디오 디코딩 디바이스 (24) 는 CodedAmbCoeffIdx+1+MinNumOfCoeffsForAmbHOA 와 같은 것으로서 또는 이 예에서 5 로서 AmbCoeffIdx 를 도출할 수도 있다. CSID 필드 (154B) 는 ChannelType (269) 과 함께 unitC (267), bb (266) 및 ba (265) 를 포함하고, 이들의 각각은 도 10k(ii) 의 예에 보여진 대응하는 값들 01, 1, 0 및 01 로 설정된다. CSID 필드 (154C) 는 3 의 값을 갖는 ChannelType 필드 (269) 를 포함한다.

[0607] 도 10k(ii) 의 예에서, 프레임 (249G) 은 (CSID 필드들 (154B) 에서 ChannelType (269) 이 1 과 같음을 고려하면) 특이 벡터 기반 신호 및 (CSID 필드들 (154C) 에서 ChannelType (269) 이 3 과 같음을 고려하면) 공을 포함한다. 앞서 말한 HOAconfig 부분 (250K) 을 고려하면, 오디오 디코딩 디바이스 (24) 는, 11 개 V 벡터 엘리먼트들이 인코딩된다고 결정할 수도 있다 (여기서 12는 $(HOAOrder + 1)^2$ - (MinNumOfCoeffsForAmbHOA) - (ContAddAmbHoaChan) = 16-4-1 = 11 로서 도출된다). 그러므로, VVectorData (156) 는 모든 11개 벡터 엘리먼트들을 포함하고, 이들의 각각은 8 비트로 균일하게 양자화된다. 각주 1 에 의해 의해 언급된 바처럼, 코딩된 VVectorData 엘리먼트들의 수 및 인덱스들은 파라미터 CodedVVecLength=0 에 의해 명시된다. 더욱이, 각주 2 에 의해 언급된 바처럼, 코딩 방식은 대응하는 전송 채널에 대해 CSID 필드에서 NbitsQ = 5 에 의해 시그널링된다.

[0608] 프레임 (249H) 에서, CSID 필드 (154) 는, 천이가 일어나지 않았다는 것을 표시하는 AmbCoeffIdxTransition (247) 를 포함하고 따라서 CodedAmbCoeffIdx (246) 이 이전 프레임으로부터 암시될 수도 있고 시그널링되거나 또는 그렇지 않으면 다시 명시될 필요가 없다. 프레임 (249H) 의 CSID 필드 (154B 및 154C) 는 프레임 (249G) 에 대한 것과 동일하고 따라서, 프레임 (249G) 과 같이, 프레임 (249H) 은 단일 VVectorData 필드 (156) 를 포함하고, 이는 11 개의 벡터 엘리먼트들을 포함하며, 이들의 각각은 8 비트로 균일하게 양자화된다.

[0609] 도 10l(i) 및 도 10l(ii) 는 제 1 예의 비트스트림 (248L) 및 수반하는 HOA config 부분 (250L) 이 위의 의사 코드에서 케이스 2 와 대응하도록 생성된 것을 예시한다. 도 10l(i) 의 예에서, HOAconfig 부분 (250L) 은, 제 0 의 오더에서 (이 예에서 $(HoaOrder + 1)^2$ - $(MinAmbHoaOrder + 1)^2$ = 16 - 4 = 12 와 같은) MinAmbHoaOrder 신택스 엘리먼트 (150) 에 의해 명시되는 오더에 이르기 까지의 엘리먼트들을 제외하고, V 벡터의 모든 엘리먼트들이 코딩된다는 것을 나타내도록 설정된 CodedVVecLength 신택스 엘리먼트 (256) 를

포함한다. HOAconfig 부분 (250L) 은 또한, 공간-시간 보간의 보간 함수가 상향 코사인이라는 것을 표시하도록 설정된 SpatialInterpolationMethod 신택스 엘리먼트 (255) 를 포함한다. HOAconfig 부분 (250L) 은 더욱이, 256 의 보간된 샘플 지속시간을 표시하도록 설정된 CodedSpatialInterpolationTime (254) 을 포함한다. HOAconfig 부분 (250L) 은 주변 HOA 컨텐츠의 MinimumHOA 오더가 1 이라는 것을 표시하도록 설정된 MinAmbHoaOrder 신택스 엘리먼트 (150) 를 더 포함하고, 여기서 오디오 디코딩 디바이스 (24) 는 $(1+1)^2$ 또는 4 와 동일한 것으로 MinNumofCoeffsForAmbHOA 신택스 엘리먼트를 도출할 수도 있다. HOAconfig 부분 (250L) 은, 3 과 동일한 것으로 컨텐츠의 HOA 오더 (또는 다른 말로, N = 3) 를 표시하도록 설정된 HoaOrder 신택스 엘리먼트 (152) 를 포함하고, 오디오 디코딩 디바이스 (24) 는 $(N + 1)^2$ 또는 16 와 동일한 것으로 NumOfHoaCoeffs 을 도출할 수도 있다.

[0610] 또한 도 10l(i) 의 예에 보여진 바처럼, 부분 (248L) 은 USAC-3D 오디오 프레임을 포함하고, 여기서 스펙트럼 대역 복사 (SBR) 가 가능해질 때 하나의 USAC-3D 프레임 내에 2 개의 오디오 프레임들이 저장되는 것을 고려하면 2 개의 HOA 프레임들 (249I 및 249I) 이 USAC 확장 페이로드에 저장된다. 오디오 디코딩 디바이스 (24) 는 numHOATransportChannels 신택스 엘리먼트 및 MinNumOfCoeffsForAmbHOA 신택스 엘리먼트의 함수로서 플렉서블 전송 채널들의 수를 도출할 수도 있다. 다음의 예들에서, numHOATransportChannels 신택스 엘리먼트가 7 과 동일하고 MinNumOfCoeffsForAmbHOA 신택스 엘리먼트가 4 와 동일한 것으로 가정되고, 여기서 플렉서블 전송 채널들의 수는 numHOATransportChannels 신택스 엘리먼트 마이너스 MinNumOfCoeffsForAmbHOA 신택스 엘리먼트 (또는 3) 과 동일하다.

[0611] 도 10l(ii) 은 프레임들 (249I 및 249J) 를 더 상세하게 예시한다. 도 10l(ii) 의 예에 보여진 바처럼, 프레임 (249I) 는 CSID 필드들 (154-154C) 및 VVectorData 필드들 (156 및 156B) 을 포함한다. CSID 필드 (154) 는 ChannelType (269) 과 함께 unitC (267), bb (266) 및 ba (265) 를 포함하고, 이들의 각각은 도 10j(i) 의 예에 보여진 대응하는 값들 01, 1, 0 및 01 로 설정된다. CSID 필드 (154B) 는 ChannelType (269) 과 함께 unitC (267), bb (266) 및 ba (265) 를 포함하고, 이들의 각각은 도 10l(ii) 의 예에 보여진 대응하는 값들 01, 1, 0 및 01 로 설정된다. CSID 필드 (154C) 는 3 의 값을 갖는 ChannelType 필드 (269) 를 포함한다. CSID 필드들 (154-154C) 의 각각은 전송 채널들 (1, 2 및 3) 의 각각의 하나에 대응한다.

[0612] 도 10l(ii) 의 예에서, 프레임 (249I) 은 (CSID 필드들 (154 및 154B) 에서 ChannelType (269) 이 1 과 같음을 고려하면) 2 개의 벡터 기반 신호들 및 (CSID 필드들 (154C) 에서 ChannelType (269) 이 3 과 같음을 고려하면) 공을 포함한다. 앞서 말한 HOAconfig 부분 (250H) 을 고려하면, 오디오 디코딩 디바이스 (24) 는 12 개 V 벡터 엘리먼트들이 인코딩되는 것을 결정할 수도 있다. 그러므로, VVectorData (156 및 156B) 각각은 12 개 벡터 엘리먼트들을 포함하고, 이들의 각각은 8 비트로 균일하게 양자화된다. 각주 1 에 의해 의해 언급된 바처럼, 코딩된 VVectorData 엘리먼트들의 수 및 인덱스들은 파라미터 CodedVVecLength=0 에 의해 명시된다. 또한, 단일 별표 (*) 에 의해 언급된 바처럼, 코딩 방식은 대응하는 전송 채널에 대해 CSID 필드에서 NbitsQ = 5 에 의해 시그널링된다.

[0613] 프레임 (249J) 에서, CSID 필드 (154 및 154B) 는 프레임 (249I) 에 있는 것과 동일한 한편, 프레임 (249F) 의 CSID 필드 (154C) 는 1 의 ChannelType 으로 스위칭된다. 그러므로, 프레임 (249B) 의 CSID 필드 (154C) 는 Cbflag (267), (허프만 인코딩을 표시하는) Pflag (267) 및 (12와 같은) Nbits (261) 를 포함한다. 결과적으로, 프레임 (249F) 은 12 개 V 벡터 엘리먼트들을 포함하는 제 3 VVectorData 필드 (156C) 를 포함하고, 그들의 각각은 12 비트로 균일하게 양자화되고 허프만 코딩된다. 위에 언급된 바처럼, 코딩된 VVectorData 엘리먼트들의 수 및 인덱스들은 파라미터 CodedVVecLength = 0 에 의해 명시되는 한편, 허프만 코딩 방식은 이 특정 전송 채널 (예를 들어, 전송 채널 번호 3) 에 대해 CSID 필드 (154C) 에서 NbitsQ = 12, CbFlag = 0 및 Pflag = 0 에 의해 시그널링된다.

[0614] 도 10m(i) 및 도 10m(ii) 의 예는 제 2 예의 비트스트림 (248M) 및 수반하는 HOA config 부분 (250M) 이 위의 의사 코드에서 케이스 2 와 대응하도록 생성된 것을 예시한다. 도 10m(i) 의 예에서, HOAconfig 부분 (250M) 은, 제 0 의 오더에서 (이 예에서 $(HoaOrder + 1)^2 - (MinAmbHoaOrder + 1)^2 = 16 - 4 = 12$ 와 같은) MinAmbHoaOrder 신택스 엘리먼트 (150) 에 의해 명시되는 오더에 이르기 까지의 엘리먼트들을 제외하고, V 벡터의 모든 엘리먼트들이 코딩된다는 것을 나타내도록 설정된 CodedVVecLength 신택스 엘리먼트 (256) 를 포함한다. HOAconfig 부분 (250M) 은 또한, 공간-시간 보간의 보간 함수가 상향 코사인이라는 것을 표시하도록 설정된 SpatialInterpolationMethod 신택스 엘리먼트 (255) 를 포함한다. HOAconfig 부분 (250M) 은 더욱이, 256 의 보간된 샘플 지속시간을 표시하도록 설정된 CodedSpatialInterpolationTime (254) 을

포함한다.

[0615] HOAconfig 부분 (250M) 은 주변 HOA 컨텐츠의 MinimumHOA 오더가 1 이라는 것을 표시하도록 설정된 MinAmbHoaOrder 신택스 엘리먼트 (150) 를 더 포함하고, 여기서 오디오 디코딩 디바이스 (24) 는 $(1+1)^2$ 또는 4 와 동일한 것으로 MinNumofCoeffsForAmbHOA 신택스 엘리먼트를 도출할 수도 있다. 오디오 디코딩 디바이스 (24) 는 또한, 이 예에서 16-4 또는 12 과 같은 것으로 가정되는 MinNumOfCoeffsForAmbHOA 와 NumOfHoaCoeff 신택스 엘리먼트 사이의 차이로 설정된 것으로서 MaxNoOfAddActiveAmbCoeffs 신택스 엘리먼트를 도출할 수도 있다. 오디오 디코딩 디바이스 (24) 는 또한, ceil(log2(MaxNoOfAddActiveAmbCoeffs)) = ceil(log2(12)) = 4 로 설정된 것으로서 AmbAsignmBits 신택스 엘리먼트를 도출할 수도 있다. HOAconfig 부분 (250M) 은, 3 과 동일한 것으로 컨텐츠의 HOA 오더 (또는 다른 말로, N = 3) 를 표시하도록 설정된 HoaOrder 신택스 엘리먼트 (152) 를 포함하고, 오디오 디코딩 디바이스 (24) 는 $(N + 1)^2$ 또는 16 와 동일한 것으로 NumOfHoaCoeffs 을 도출할 수도 있다.

[0616] 또한 도 10m(i) 의 예에 보여진 바처럼, 부분 (248M) 은 USAC-3D 오디오 프레임을 포함하고, 여기서 스펙트럼 대역 복사 (SBR) 가 가능해질 때 하나의 USAC-3D 프레임 내에 2 개의 오디오 프레임들이 저장되는 것을 고려하면 2 개의 HOA 프레임들 (249K 및 249L) 이 USAC 확장 페이로드에 저장된다. 오디오 디코딩 디바이스 (24) 는 numHOATransportChannels 신택스 엘리먼트 및 MinNumOfCoeffsForAmbHOA 신택스 엘리먼트의 함수로서 플렉서블 전송 채널들의 수를 도출할 수도 있다. 다음의 예들에서, numHOATransportChannels 신택스 엘리먼트가 7 과 동일하고 MinNumOfCoeffsForAmbHOA 신택스 엘리먼트가 4 와 동일한 것으로 가정되고, 여기서 플렉서블 전송 채널들의 수는 numHOATransportChannels 신택스 엘리먼트 마이너스 MinNumOfCoeffsForAmbHOA 신택스 엘리먼트 (또는 3) 과 동일하다.

[0617] 도 10m(ii) 은 프레임들 (249K 및 249L) 을 더 상세하게 예시한다. 도 10m(ii) 의 예에 보여진 바처럼, 프레임 (249K) 는 CSID 필드들 (154-154C) 및 VVectorData 필드 (156) 을 포함한다. CSID 필드 (154) 는 CodedAmbCoeffIdx (246), AmbCoeffIdxTransition (247) (여기서 이중 별표 (**) 는, 플렉서블 전송 채널 번호 1 에 대하여, 디코더의 내부 상태가 여기서 AmbCoeffIdxTransitionState = 2 인 것으로 가정되고, 이는 CodedAmbCoeffIdx 비트필드가 시그널링되거나 또는 그렇지 않으면 비트스트림에서 명시되는 것에 귀결된다), 및 ChannelType (269) (이는 2 와 같으며, 대응하는 페이로드가 추가적인 주변 HOA 계수라는 것을 시그널링한다) 를 포함한다. 오디오 디코딩 디바이스 (24) 는 CodedAmbCoeffIdx+1+MinNumOfCoeffsForAmbHOA 와 같은 것으로 또는 이 예에서 5 로 AmbCoeffIdx 를 도출할 수도 있다. CSID 필드 (154B) 는 ChannelType (269) 과 함께 unitC (267), bb (266) 및 ba (265) 를 포함하고, 이들의 각각은 도 10m(ii) 의 예에 보여진 대응하는 값들 01, 1, 0 및 01 로 설정된다. CSID 필드 (154C) 는 3 의 값을 갖는 ChannelType 필드 (269) 를 포함한다.

[0618] 도 10m(ii) 의 예에서, 프레임 (249K) 은 (CSID 필드들 (154B) 에서 ChannelType (269) 이 1 과 같음을 고려하면) 특이 벡터 기반 신호 및 (CSID 필드들 (154C) 에서 ChannelType (269) 이 3 과 같음을 고려하면) 공을 포함한다. 앞서 말한 HOAconfig 부분 (250M) 을 고려하면, 오디오 디코딩 디바이스 (24) 는 12 개 V 벡터 엘리먼트들이 인코딩되는 것을 결정할 수도 있다. 그러므로, VVectorData (156) 는 12 개 벡터 엘리먼트들을 포함하고, 이들의 각각은 8 비트로 균일하게 양자화된다. 각주 1 에 의해 의해 언급된 바처럼, 코딩된 VVectorData 엘리먼트들의 수 및 인덱스들은 파라미터 CodedVVecLength=0 에 의해 명시된다. 더욱이, 각주 2 에 의해 언급된 바처럼, 코딩 방식은 대응하는 전송 채널에 대해 CSID 필드에서 NbitsQ = 5 에 의해 시그널링된다.

[0619] 프레임 (249L) 에서, CSID 필드 (154) 는, 천이가 일어나지 않았다는 것을 표시하는 AmbCoeffIdxTransition (247) 를 포함하고 따라서 CodedAmbCoeffIdx (246) 이 이전 프레임으로부터 암시될 수도 있고 시그널링되거나 또는 그렇지 않으면 다시 명시될 필요가 없다. 프레임 (249L) 의 CSID 필드 (154B 및 154C) 는 프레임 (249K) 에 대한 것과 동일하고 따라서, 프레임 (249K) 과 같이, 프레임 (249L) 은 단일 VVectorData 필드 (156) 를 포함하고, 이는 12 개의 벡터 엘리먼트들을 포함하며, 이들의 각각은 8 비트로 균일하게 양자화된다.

[0620] 도 10n(i) 및 도 10n(ii) 는 제 1 예의 비트스트림 (248N) 및 수반하는 HOA config 부분 (250N) 이 위의 의사 코드에서 케이스 3 과 대응하도록 생성된 것을 예시한다. 도 10n(i) 의 예에서, HOAconfig 부분 (250N) 은, (이 예에서 0 인 것으로 가정되는) ContAddAmbHoaChan 신택스 엘리먼트에 명시된 그러한 엘리먼트들을 제외하고, V 벡터의 모든 엘리먼트들이 코딩되는 것을 나타내도록 설정된 CodedVVecLength 신택스 엘리먼트 (256) 를 포함한다. HOAconfig 부분 (250N) 은 또한, 공간-시간 보간의 보간 함수가 상향 코사인이라는 것을 표시하도록 설정된 SpatialInterpolationMethod 신택스 엘리먼트 (255) 를 포함한다. HOAconfig 부분

(250N) 은 더욱이, 256 의 보간된 샘플 지속시간을 표시하도록 설정된 CodedSpatialInterpolationTime (254) 을 포함한다. HOAconfig 부분 (250N) 은 주변 HOA 컨텐츠의 MinimumHOA 오더가 1 이라는 것을 표시하도록 설정된 MinAmbHoaOrder 신택스 엘리먼트 (150) 를 더 포함하고, 여기서 오디오 디코딩 디바이스 (24) 는 $(1+1)^2$ 또는 4 와 동일한 것으로 MinNumofCoeffsForAmbHOA 신택스 엘리먼트를 도출할 수도 있다. HOAconfig 부분 (250N) 는, 3 과 동일한 것으로 컨텐츠의 HOA 오더 (또는 다른 말로, N = 3) 를 표시하도록 설정된 HoaOrder 신택스 엘리먼트 (152) 를 포함하고, 여기서 오디오 디코딩 디바이스 (24) 는 $(N + 1)^2$ 또는 16 와 동일한 것으로 NumOfHoaCoeffs 을 도출할 수도 있다.

[0621] 또한 도 10n(i) 의 예에 보여진 바처럼, 부분 (248N) 은 USAC-3D 오디오 프레임을 포함하고, 여기서 스펙트럼 대역 복사 (SBR) 가 가능해질 때 하나의 USAC-3D 프레임 내에 2 개의 오디오 프레임들이 저장되는 것을 고려하면 2 개의 HOA 프레임들 (249M 및 249N) 이 USAC 확장 페이로드에 저장된다. 오디오 디코딩 디바이스 (24) 는 numHOATransportChannels 신택스 엘리먼트 및 MinNumOfCoeffsForAmbHOA 신택스 엘리먼트의 함수로서 플렉서블 전송 채널들의 수를 도출할 수도 있다. 다음의 예들에서, numHOATransportChannels 신택스 엘리먼트는 7 과 동일하고 MinNumOfCoeffsForAmbHOA 신택스 엘리먼트가 4 와 동일한 것으로 가정되고, 여기서 플렉서블 전송 채널들의 수는 numHOATransportChannels 신택스 엘리먼트 마이너스 MinNumOfCoeffsForAmbHOA 신택스 엘리먼트 (또는 3) 과 동일하다.

[0622] 도 10n(ii) 은 프레임들 (249M 및 249N) 을 더 상세하게 예시한다. 도 10n(ii) 의 예에 보여진 바처럼, 프레임 (249M) 은 CSID 필드들 (154-154C) 및 VVectorData 필드들 (156 및 156B) 을 포함한다. CSID 필드 (154) 는 ChannelType (269) 과 함께 unitC (267), bb (266) 및 ba (265) 를 포함하고, 이들의 각각은 도 10j(i) 의 예에 보여진 대응하는 값들 01, 1, 0 및 01 로 설정된다. CSID 필드 (154B) 는 ChannelType (269) 과 함께 unitC (267), bb (266) 및 ba (265) 를 포함하고, 이들의 각각은 도 10n(ii) 의 예에 보여진 대응하는 값들 01, 1, 0 및 01 로 설정된다. CSID 필드 (154C) 는 3 의 값을 갖는 ChannelType 필드 (269) 를 포함한다. CSID 필드들 (154-154C) 의 각각은 전송 채널들 (1, 2 및 3) 의 각각의 하나에 대응한다.

[0623] 도 10n(ii) 의 예에서, 프레임 (249M) 은 (CSID 필드들 (154 및 154B) 에서 ChannelType (269) 이 1 과 같음을 고려하면) 2 개의 벡터 기반 신호들 및 (CSID 필드들 (154C) 에서 ChannelType (269) 이 3 과 같음을 고려하면) 공을 포함한다. 앞서 말한 HOAconfig 부분 (250M) 을 고려하면, 오디오 디코딩 디바이스 (24) 는 16 개 V 벡터 엘리먼트들이 인코딩되는 것을 결정할 수도 있다. 그러므로, VVectorData (156 및 156B) 각각은 16 개 벡터 엘리먼트들을 포함하고, 이들의 각각은 8 비트로 균일하게 양자화된다. 각주 1 에 의해 의해 언급된 바처럼, 코딩된 VVectorData 엘리먼트들의 수 및 인덱스들은 파라미터 CodedVVecLength=0 에 의해 명시된다. 또한, 단일 별표 (*) 에 의해 언급된 바처럼, 코딩 방식은 대응하는 전송 채널에 대해 CSID 필드에서 NbitsQ = 5 에 의해 시그널링된다.

[0624] 프레임 (249N) 에서, CSID 필드 (154 및 154B) 는 프레임 (249M) 에 있는 것과 동일한 한편, 프레임 (249F) 의 CSID 필드 (154C) 는 1 의 ChannelType 으로 스위칭된다. 그러므로, 프레임 (249B) 의 CSID 필드 (154C) 는 Cbflag (267), (허프만 인코딩을 표시하는) Pflag (267) 및 (12와 같은) Nbits (261) 를 포함한다. 결과적으로, 프레임 (249F) 는 16 개 V 벡터 엘리먼트들을 포함하는 제 3 VVectorData 필드 (156C) 를 포함하고, 그들의 각각은 12 비트로 균일하게 양자화되고 허프만 코딩된다. 위에 언급된 바처럼, 코딩된 VVectorData 엘리먼트들의 수 및 인덱스들은 파라미터 CodedVVecLength = 0 에 의해 명시되는 한편, 허프만 코딩 방식은 이 특정 전송 채널 (예를 들어, 전송 채널 번호 3) 에 대해 CSID 필드 (154C) 에서 NbitsQ = 12, CbFlag = 0 및 Pflag = 0 에 의해 시그널링된다.

[0625] 도 10o(i) 및 도 10o(ii) 의 예는 제 2 예의 비트스트림 (248O) 및 수반하는 HOA config 부분 (250O) 이 위의 의사 코드에서 케이스 3 과 대응하도록 생성된 것을 예시한다. 도 10o(i) 의 예에서, HOAconfig 부분 (2500) 은, (이 예에서 1 로 가정되는) ContAddAmbHoaChan 신택스 엘리먼트에 명시된 그러한 엘리먼트들을 제외하고, V 벡터의 모든 엘리먼트들이 코딩되는 것을 나타내도록 설정된 CodedVVecLength 신택스 엘리먼트 (256) 를 포함한다. HOAconfig 부분 (2500) 은 또한, 공간-시간 보간의 보간 함수가 상승 코사인이라는 것을 표시하도록 설정된 SpatialInterpolationMethod 신택스 엘리먼트 (255) 를 포함한다. HOAconfig 부분 (2500) 은 더욱이, 256 의 보간된 샘플 지속시간을 표시하도록 설정된 CodedSpatialInterpolationTime (254) 을 포함한다.

[0626] HOAconfig 부분 (2500) 은 주변 HOA 컨텐츠의 MinimumHOA 오더가 1 이라는 것을 표시하도록 설정된

MinAmbHoaOrder 신택스 엘리먼트 (150) 를 더 포함하고, 여기서 오디오 디코딩 디바이스 (24) 는 $(1+1)^2$ 또는 4 와 동일한 것으로 MinNumofCoeffsForAmbHOA 신택스 엘리먼트를 도출할 수도 있다. 오디오 디코딩 디바이스 (24) 는 또한, 이 예에서 16-4 또는 12 와 같은 것으로 가정되는 MinNumOfCoeffsForAmbHOA 와 NumOfHoaCoeff 신택스 엘리먼트 사이의 차이로 설정된 것으로서 MaxNoOfAddActiveAmbCoeffs 신택스 엘리먼트를 도출할 수도 있다. 오디오 디코딩 디바이스 (24) 는 또한, ceil(log2(MaxNoOfAddActiveAmbCoeffs)) = ceil(log2(12)) = 4 으로 설정된 것으로서 AmbAsignmBits 신택스 엘리먼트를 도출할 수도 있다. HOAconfig 부분 (2500) 은, 3 과 동일한 것으로 컨텐츠의 HOA 오더 (또는 다른 말로, N = 3) 를 표시하도록 설정된 HoaOrder 신택스 엘리먼트 (152) 를 포함하고, 오디오 디코딩 디바이스 (24) 는 $(N + 1)^2$ 또는 16 와 동일한 것으로 NumOfHoaCoeffs 을 도출할 수도 있다.

[0627] 또한 도 10o(i) 의 예에 보여진 바처럼, 부분 (2480) 은 USAC-3D 오디오 프레임을 포함하고, 여기서 스펙트럼 대역 복사 (SBR) 가 가능해질 때 하나의 USAC-3D 프레임 내에 2 개의 오디오 프레임들이 저장되는 것을 고려하면 2 개의 HOA 프레임들 (2490 및 249P) 이 USAC 확장 페이로드에 저장된다. 오디오 디코딩 디바이스 (24) 는 numHOATransportChannels 신택스 엘리먼트 및 MinNumOfCoeffsForAmbHOA 신택스 엘리먼트의 함수로서 플렉서블 전송 채널들의 수를 도출할 수도 있다. 다음의 예들에서, numHOATransportChannels 신택스 엘리먼트가 7 과 동일하고 MinNumOfCoeffsForAmbHOA 신택스 엘리먼트가 4 와 동일한 것으로 가정되고, 여기서 플렉서블 전송 채널들의 수는 numHOATransportChannels 신택스 엘리먼트 마이너스 MinNumOfCoeffsForAmbHOA 신택스 엘리먼트 (또는 3) 과 동일하다.

[0628] 도 10o(ii) 는 프레임들 (2490 및 249P) 을 더 상세하게 예시한다. 도 10o(ii) 의 예에 보여진 바처럼, 프레임 (2490) 는 CSID 필드들 (154-154C) 및 VVectorData 필드 (156) 을 포함한다. CSID 필드 (154) 는 CodedAmbCoeffIdx (246), AmbCoeffIdxTransition (247) (여기서 이중 별표 (**) 는, 플렉서블 전송 채널 번호 1 에 대하여, 디코더의 내부 상태가 여기서 AmbCoeffIdxTransitionState = 2 인 것으로 가정되고, 이는 CodedAmbCoeffIdx 비트필드가 시그널링되거나 또는 그렇지 않으면 비트스트림에서 명시되는 것에 귀결된다), 및 ChannelType (269) (이는 2 와 같으며, 대응하는 페이로드가 추가적인 주변 HOA 계수라는 것을 시그널링한다) 를 포함한다. 오디오 디코딩 디바이스 (24) 는 CodedAmbCoeffIdx+1+MinNumOfCoeffsForAmbHOA 와 같은 것으로서 또는 이 예에서 5 로서 AmbCoeffIdx 를 도출할 수도 있다. CSID 필드 (154B) 는 ChannelType (269) 과 함께 unitC (267), bb (266) 및 ba (265) 를 포함하고, 이들의 각각은 도 10o(ii) 의 예에 보여진 대응하는 값들 01, 1, 0 및 01 로 설정된다. CSID 필드 (154C) 는 3 의 값을 갖는 ChannelType 필드 (269) 를 포함한다.

[0629] 도 10o(ii) 의 예에서, 프레임 (2490) 은 (CSID 필드들 (154B) 에서 ChannelType (269) 이 1 과 같음을 고려하면) 특이 벡터 기반 신호 및 (CSID 필드들 (154C) 에서 ChannelType (269) 이 3 과 같음을 고려하면) 공을 포함한다. 앞서 말한 HOAconfig 부분 (2500) 을 고려하면, 오디오 디코딩 디바이스 (24) 는 16 마이너스 ContAddAmbHoaChan 신택스 엘리먼트에 의해 명시된 하나 (예를들어, 6의 인덱스와 연관된 벡터 엘리먼트는 ContAddAmbHoaChan 신택스 엘리먼트로서 명시된다) 또는 15개 V 벡터 엘리먼트들이 인코딩된다고 결정할 수도 있다. 그러므로, VVectorData (156) 는 15개 벡터 엘리먼트들을 포함하고, 이들의 각각은 8 비트로 균일하게 양자화된다. 각주 1 에 의해 의해 언급된 바처럼, 코딩된 VVectorData 엘리먼트들의 수 및 인덱스들은 파라미터 CodedVVecLength=0 에 의해 명시된다. 더욱이, 각주 2 에 의해 언급된 바처럼, 코딩 방식은 대응하는 전송 채널에 대해 CSID 필드에서 NbitsQ = 5 에 의해 시그널링된다.

[0630] 프레임 (249P) 에서, CSID 필드 (154) 는, 천이가 일어나지 않았다는 것을 표시하는 AmbCoeffIdxTransition (247) 를 포함하고 따라서 CodedAmbCoeffIdx (246) 이 이전 프레임으로부터 암시될 수도 있고 시그널링되거나 또는 그렇지 않으면 다시 명시될 필요가 없다. 프레임 (249P) 의 CSID 필드 (154B 및 154C) 는 프레임 (2490) 에 대한 것과 동일하고 따라서, 프레임 (2490) 과 같이, 프레임 (249P) 은 단일 VVectorData 필드 (156) 를 포함하고, 이는 15 개의 벡터 엘리먼트들을 포함하며, 이들의 각각은 8 비트로 균일하게 양자화된다.

[0631] 도 11a 내지 도 11g 는 도 5 의 예에 보여진 오디오 디코딩 디바이스 (24) 의 다양한 유닛들을, 더 상세하게, 예시하는 블록 다이어그램들이다. 도 11a 는 오디오 디코딩 디바이스 (24) 의 추출 유닛 (72) 을, 더 상세하게 예시하는 블록 다이어그램이다. 도 11a 의 예에 보여진 바처럼, 추출 유닛 (72) 은 모드 파싱 유닛 (270), 모드 구성 유닛 (272) ("모드 config 유닛 (272)"), 및 구성가능 추출 유닛 (274) 을 포함할 수도 있다.

[0632] 모드 파싱 유닛 (270) 은, 비트스트림 (21) 을 형성하도록 HOA 계수들 (11) 을 인코딩하는데 사용된 코딩 모드

를 나타내는 위에 언급된 신택스 엘리먼트 (도 10e 의 예에 보여진 ChannelType 신택스 엘리먼트) 를 파싱하도록 구성된 유닛을 나타낼 수도 있다. 모드 파싱 유닛 (270) 은 결정된 신택스 엘리먼트를 모드 구성 유닛 (272) 으로 보낼 수도 있다. 모드 구성 유닛 (272) 은, 파싱된 신택스 엘리먼트에 기초하여 구성가능 추출 유닛 (274) 을 구성하도록 구성된 유닛을 나타낼 수도 있다. 모드 구성 유닛 (272) 은, 비트스트림 (21) 으로부터 HOA 계수들 (11) 의 방향 기반 코딩된 표현을 추출하거나, 또는 파싱된 신택스 엘리먼트에 기초하여 비트스트림 (21) 으로부터 HOA 계수들 (11) 의 벡터 기반 코딩된 표현을 추출하기 위하여 구성가능 추출 유닛 (274) 을 구성할 수도 있다.

[0633] 방향 기반 인코딩이 수행되었을 때, 구성가능 추출 유닛 (274) 은 HOA 계수들 (11) 의 방향 기반 버전 및 (도 11a 의 예에서 방향 기반 정보 (91) 로 표기되는) 이 인코딩된 버전과 연관된 신택스 엘리먼트들을 추출할 수도 있다. 이 방향 기반 정보 (91) 는, 도 10d 의 예에서 나타낸 방향 정보 (253) 및 0 과 동일한 ChannelType 으로서 정의된 도 10e 의 예에 나타낸 SideChannelInfoData 를 포함할 수도 있다.

[0634] 신택스 엘리먼트가 HOA 계수들 (11) 이 벡터 기반 합성을 이용하여 인코딩되었다는 것을 표시할 때 (예를 들어, ChannelType 신택스 엘리먼트가 1 과 동일할 때), 구성가능 추출 유닛 (274) 은 코딩된 전경 V[k] 벡터들 (57), 인코딩된 주변 HOA 계수들 (59) 및 인코딩된 nFG 신호들 (59) 을 추출할 수도 있다. 구성가능 추출 유닛 (274) 은, 또한, 신택스 엘리먼트가 HOA 계수들 (11) 이 벡터 기반 합성을 이용하여 인코딩되었다는 것을 나타낸다고 결정할 시에, 비트스트림 (21) 으로부터 CodedSpatialInterpolationTime 신택스 엘리먼트 (254) 및 SpatialInterpolationMethod 신택스 엘리먼트 (255) 를 추출하며, 이들 신택스 엘리먼트들 (254 및 255) 을 공간-시간 보간 유닛 (76) 으로 보낸다.

[0635] 도 11b 는 도 5 의 예에 보여진 오디오 디코딩 디바이스 (24) 의 양자화 유닛 (74) 을, 더 상세하게, 예시하는 블록 다이어그램이다. 양자화 유닛 (74) 은, 코딩된 전경 V[k] 벡터들 (57) 을 엔트로피 디코딩 및 역양자화함으로써 감소된 전경 V[k] 벡터들 (55_k) 을 생성하기 위하여 도 4의 예에 보여진 양자화 유닛 (52) 에 상보적인 방식으로 동작하도록 구성된 유닛을 표현할 수도 있다. 스칼라/엔트로피 역양자화 유닛 (984) 은 카테고리/잔차 디코딩 유닛 (276), 예측 유닛 (278) 및 균일 역양자화 유닛 (280) 을 포함할 수도 있다.

[0636] 카테고리/잔차 디코딩 유닛 (276) 은, (위에 언급된 바처럼, 비트스트림 (21) 에서 신택스 엘리먼트로서 표현되는) 허프만 테이블 정보 (241) 에 의해 식별되는 허프만 테이블을 이용하여 코딩된 전경 V[k] 벡터들 (57) 에 관하여 허프만 디코딩을 수행하도록 구성된 유닛을 나타낼 수도 있다. 카테고리/잔차 디코딩 유닛 (276) 은 예측 유닛 (278) 으로 양자화 전경 V[k] 벡터들을 출력할 수도 있다. 예측 유닛 (278) 은, 예측 모드 (237) 에 기초하여 양자화 전경 V[k] 벡터들에 관하여 예측을 수행하여, 균일 역양자화 유닛 (280) 으로 증강된 양자화 전경 V[k] 벡터들을 출력하도록 구성된 유닛을 나타낼 수도 있다. 균일 역양자화 유닛 (280) 은, nbits 값 (233) 에 기초하여 증강된 양자화 전경 V[k] 벡터들에 관하여 역양자화를 수행하여, 감소된 전경 V[k] 벡터들 (55_k) 을 출력하도록 구성된 유닛을 나타낼 수도 있다.

[0637] 도 11c 는 도 5 의 예에 보여진 오디오 디코딩 디바이스 (24) 의 심리 음향 디코딩 유닛 (80) 을, 더 상세하게, 예시하는 블록 다이어그램이다. 위에 언급된 바처럼, 심리 음향 디코딩 유닛 (80) 은, 인코딩된 주변 HOA 계수들 (59) 및 인코딩된 nFG 신호들 (61) 을 디코딩하고 이에 의해 에너지 보상된 주변 HOA 계수들 (47') 및 (보간된 nFG 오디오 오브젝트들 (49') 로도 지칭될 수도 있는) 보간된 nFG 신호들 (49') 을 생성하기 위하여 도 4의 예에 보여진 심리 음향 오디오 코딩 유닛 (40) 에 상반되는 방식으로 동작할 수도 있다. 심리 음향 디코딩 유닛 (80) 은 에너지 보상된 주변 HOA 계수들 (47') 을 HOA 계수 공식화 유닛 (82) 으로 그리고 nFG 신호들 (49') 을 리오더 (84) 로 보낼 수도 있다. 심리 음향 디코딩 유닛 (80) 은, 심리 음향 오디오 코딩 유닛 (40) 과 유사한 복수의 오디오 디코더들 (80-80N) 을 포함할 수도 있다. 오디오 디코더들 (80-80N) 은, 위에 언급된 바처럼, 배경 HOA 계수들 (47') 의 각각의 채널 및 nFG 신호들 (49') 의 각각의 신호의 동시 디코딩을, 지원하기에 충분한 양에서 심리 음향 오디오 코딩 유닛 (40) 에 의해 예시될 수도 있거나 또는 그렇지 않으면 내부에 포함될 수도 있다.

[0638] 도 11d 는 도 5 의 예에 보여진 오디오 디코딩 디바이스 (24) 의 리오더 유닛 (84) 을, 더 상세하게, 예시하는 블록 다이어그램이다. 리오더 유닛 (84) 은, 리오더 유닛 (34) 에 관하여 위에 설명된 것과 상호 유사한 방식으로 동작하도록 구성된 유닛을 나타낼 수도 있다. 리오더 유닛 (84) 은, HOA 계수들 (11) 의 전경 성분들의 원래 오더를 표시하는 신택스 엘리먼트들 (205) 을 수신하도록 구성된 유닛을 나타낼 수도 있는 벡터 리오더 유닛 (282) 을 포함할 수도 있다. 추출 유닛 (72) 은 비트스트림 (21) 으로부터 이들 신택스 엘리먼트들 (205) 을 파싱하고 신택스 엘리먼트 (205) 를 리오더 유닛 (84) 으로 보낼 수도 있다. 벡터 리오더 유닛

(282) 은, 이들 리오더 신택스 엘리먼트들 (205) 에 기초하여, 보간된 nFG 신호들 (49') 및 감소된 전경 V[k] 벡터들 (55_k) 을 리오더링하여 리오더링된 nFG 신호들 (49'') 및 리오더링된 전경 V[k] 벡터들 ($55_k'$) 을 생성할 수도 있다. 리오더 유닛 (84) 은, 리오더링된 nFG 신호들 (49'') 을 전경 공식화 유닛 (78) 으로 그리고 리오더링된 전경 V[k] 벡터들 ($55_k'$) 을 공간-시간 보간 유닛 (76) 으로 출력할 수도 있다.

[0639] 도 11e 는 도 5 의 예에 보여진 오디오 디코딩 디바이스 (24) 의 공간-시간 보간 유닛 (76) 을, 더 상세하게, 예시하는 블록 다이어그램이다. 공간-시간 보간 유닛 (76) 은, 공간-시간 보간 유닛 (50) 에 관하여 위에 설명된 것과 유사한 방식으로 동작할 수도 있다. 공간-시간 보간 유닛 (76) 은 V 보간 유닛 (284) 을 포함할 수도 있고, 이 V 보간 유닛 (284) 은 리오더링된 전경 V[k] 벡터들 ($55_k'$) 을 수신하고 리오더링된 전경 V[k] 벡터들 ($55_k'$) 및 리오더링된 전경 V[k-1] 벡터들 ($55_{k-1}'$) 에 관하여 공간-시간 보간을 수행하여 보간된 전경 V[k] 벡터들 ($55_k''$) 을 생성하도록 구성된 유닛을 나타낼 수도 있다. V 보간 유닛 (284) 은, CodedSpatialInterpolationTime 신택스 엘리먼트 (254) 및 SpatialInterpolationMethod 신택스 엘리먼트 (255) 에 기초하여 보간을 수행할 수도 있다. 일부 실시형태들에서, V 보간 유닛 (285) 은, SpatialInterpolationMethod 신택스 엘리먼트 (255) 에 의해 식별되는 보간의 타입을 사용하여 CodedSpatialInterpolationTime 신택스 엘리먼트 (254) 에 의해 명시되는 지속시간에 대해 V 벡터들을 보간할 수도 있다. 공간-시간 보간 유닛 (76) 은 보간된 전경 V[k] 벡터들 ($55_k''$) 을 전경 공식화 유닛 (78) 으로 포워딩할 수도 있다.

[0640] 도 11f 는 도 5 의 예에 보여진 오디오 디코딩 디바이스 (24) 의 전경 공식화 유닛 (78) 을, 더 상세하게, 예시하는 블록 다이어그램이다. 전경 공식화 유닛 (78) 은 곱셈 유닛 (286) 을 포함할 수도 있고, 이 곱셈 유닛 (286) 은 전경 HOA 계수들 (65) 을 생성하기 위하여 보간된 전경 V[k] 벡터들 ($55_k''$) 및 리오더링된 nFG 신호들 (49'') 에 대해 행렬 곱셈을 수행하도록 구성된 유닛을 나타낼 수도 있다.

[0641] 도 11g 는 도 5 의 예에 보여진 오디오 디코딩 디바이스 (24) 의 HOA 계수 공식화 유닛 (82) 을, 더 상세하게, 예시하는 블록 다이어그램이다. HOA 계수 공식화 유닛 (82) 은 가산 유닛 (288) 을 포함할 수도 있고, 이 가산 유닛 (288) 은, HOA 계수들 (11') 을 획득하기 위하여 주변 HOA 채널들 (47') 에 전경 HOA 계수들 (65) 을 가산하도록 구성된 유닛을 나타낼 수도 있다.

[0642] 도 12 는 본 개시물에 기재된 기술들의 다양한 양태들을 수행할 수도 있는 일 예의 오디오 에코시스템을 예시하는 다이어그램이다. 도 12에 예시된 바처럼, 오디오 에코시스템 (300) 은 포착 (301), 편집 (302), 코딩 (303), 송신 (304) 및 재생 (305) 을 포함할 수도 있다.

[0643] 포착 (301) 은 오디오 컨텐츠가 포착되는 오디오 에코시스템 (300) 의 기술들을 나타낼 수도 있다. 포착 (301) 의 예들은, 사운드 (예를 들어, 라이브 사운드) 레코딩, 오디오 생성 (예를 들어, 오디오 오브젝트, 폴리 생성, 사운드 합성, 시뮬레이션) 등을 포함하지만 이에 한정되지는 않는다. 일부 예들에서, 사운드는 콘서트, 스포츠 이벤트 및 감시를 행할 때 레코딩될 수도 있다. 일부 예들에서, 오디오는 시뮬레이션, 및 저작/믹싱 (예를 들어, 영화, 게임) 을 수행할 때 생성될 수도 있다. 오디오 오브젝트들은 헐리우드 (예를 들어, IMAX 스튜디오들) 에서 사용되는 바와 같을 수도 있다. 일부 예들에서, 포착 (301) 은 도 3 의 컨텐츠 창작자 (12) 와 같은 컨텐츠 창작자에 의해 수행될 수도 있다.

[0644] 편집 (302) 은 오디오 컨텐츠가 편집 및/또는 수정되는 오디오 에코시스템 (300) 의 기술들을 나타낼 수도 있다. 일 예로서, 오디오 컨텐츠는, 오디오 컨텐츠의 다수의 유닛들을 오디오 컨텐츠의 단일 유닛으로 조합하는 것에 의해 편집될 수도 있다. 또 다른 예로서, 오디오 컨텐츠는 실제 오디오 컨텐츠를 조정 (예를 들어, 오디오 컨텐츠의 하나 이상의 주파수 성분들의 레벨들을 조정) 하는 것에 의해 편집될 수도 있다. 일부 예들에서, 편집 (302) 은 도 3 의 오디오 편집 시스템 (18) 과 같은 오디오 편집 시스템에 의해 수행될 수도 있다. 일부 예들에서, 편집 (302) 은 도 29 에 예시된 모바일 디바이스들의 하나 이상과 같은 모바일 디바이스 상에서 수행될 수도 있다.

[0645] 코딩 (303) 은 오디오 컨텐츠가 오디오 컨텐츠의 표현으로 코딩되는 오디오 에코시스템 (300) 의 기술들을 나타낼 수도 있다. 일부 실시형태들에서, 오디오 컨텐츠의 표현은 도 3 의 비트스트림 (21) 과 같은 비트스트림일 수도 있다. 일부 예들에서, 코딩 (302) 은 도 3 의 오디오 인코딩 디바이스 (20) 와 같은 오디오 인코딩 디바이스에 의해 수행될 수도 있다.

[0646] 송신 (304) 은 오디오 컨텐츠가 컨텐츠 창작자로부터 컨텐츠 소비자로 전송되는 오디오 에코시스템 (300) 의 엘

리먼트들을 나타낼 수도 있다. 일부 실시형태들에서, 오디오 컨텐츠는 실시간 또는 근 실시간으로 전송될 수도 있다. 가령, 오디오 컨텐츠는 컨텐츠 소비자에게 스트리밍될 수도 있다. 일부 예들에서, 오디오 컨텐츠는 컴퓨터 판독가능 저장 매체와 같은 매체 상에 오디오 컨텐츠를 코딩함으로써 옮겨질 수도 있다. 가령, 오디오 컨텐츠는 디스크, 드라이브 등에 저장될 수도 있다 (예를 들어, 블루레이 디스크, 메모리 카드, 하드 드라이브 등).

[0647] 재생 (305) 은 오디오 컨텐츠가 컨텐츠 소비자에게 렌더링되고 재생되는 오디오 에코시스템 (300) 의 기술들을 나타낼 수도 있다. 일부 예들에서, 재생 (305) 은 재생 환경의 하나 이상의 양태들에 기초하여 3D 음장을 렌더링할 수도 있다. 즉, 재생 (305) 은 로컬 음향 랜드스케이프 (local acoustic landscape) 에 기초할 수도 있다.

[0648] 도 13 은 도 12 의 오디오 에코시스템의 일 예를 더 상세하게 예시하는 다이어그램이다. 도 13 에 예시된 바처럼, 오디오 에코시스템 (300) 은 오디오 컨텐츠 (308), 영화 스튜디오들 (310), 음악 스튜디오들 (311), 게이밍 오디오 스튜디오들 (312), 채널 기반 오디오 컨텐츠 (313), 코딩 엔진들 (314), 게임 오디오 스템들 (315), 게임 오디오 코딩/렌더링 엔진들 (316) 및 전달 시스템들 (317) 을 포함할 수도 있다. 일 예의 게이밍 오디오 스튜디오 (312) 가 도 26 에 예시되어 있다. 일부 예의 게임 오디오 코딩/렌더링 엔진들 (316) 이 도 27 에 예시되어 있다.

[0649] 도 13 에 의해 예시된 바처럼, 영화 스튜디오들 (310), 음악 스튜디오들 (311) 및 게이밍 오디오 스튜디오들 (312) 은 오디오 컨텐츠 (308) 를 수신할 수도 있다. 일부 예에서, 오디오 컨텐츠 (308) 는 도 12 의 포착 (301) 의 출력을 나타낼 수도 있다. 영화 스튜디오들 (310) 은, 이를테면 디지털 오디오 워크스테이션 (DAW) 을 이용하여, (예를 들어, 2.0, 5.1, 및 7.1 에서) 채널 기반 오디오 컨텐츠 (313) 를 출력할 수도 있다.

음악 스튜디오들 (310) 은, 이를테면 DAW 을 이용하여, (예를 들어, 2.0, 및 5.1 에서) 채널 기반 오디오 컨텐츠 (313) 를 출력할 수도 있다. 어떤 경우든, 코딩 엔진들 (314) 은, 전달 시스템들 (317) 에 의한 출력을 위해 하나 이상의 코덱들 (예를 들어, AAC, AC3, Dolby True HD, Dolby Digital Plus, 및 DTS Master Audio) 에 기초하여 채널 기반 오디오 컨텐츠 (313) 를 수신 및 인코딩할 수도 있다. 이런 식으로, 코딩 엔진들 (314) 은 도 12 의 코딩 (303) 의 일 예일 수도 있다. 게이밍 오디오 스튜디오들 (312) 은, 이를테면 DAW 를 이용하여, 하나 이상의 게임 오디오 스템들 (315) 을 출력할 수도 있다. 게임 오디오 코딩/렌더링 엔진들 (316) 은, 전달 시스템들 (317) 에 의한 출력을 위해 채널 기반 오디오 컨텐츠로 오디오 스템들 (315) 을 코딩 및 또는 렌더링할 수도 있다. 일부 예들에서, 영화 스튜디오들 (310), 음악 스튜디오들 (311) 및 게이밍 오디오 스튜디오들 (312) 의 출력은 도 12 의 편집 (302) 의 출력을 나타낼 수도 있다. 일부 예들에서, 코딩 엔진들 (314) 및/또는 게임 오디오 코딩/렌더링 엔진들 (316) 의 출력은 도 12 의 송신 (304) 의 기술들을 통해 전달 시스템들 (317) 로 전송될 수도 있다.

[0650] 도 14 는 도 12 의 오디오 에코시스템의 또 다른 예를 더 상세하게 예시하는 다이어그램이다. 도 14 에 예시된 바처럼, 오디오 에코시스템 (300B) 은, 브로드캐스트 레코딩 오디오 오브젝트들 (319), 프로페셔널 오디오 시스템들 (320), 소비자 온 디바이스 캡처 (322), HOA 오디오 포맷 (323), 온 디바이스 렌더링 (324), 소비자 오디오, TV, 및 액세서리들 (325) 및 카 오디오 시스템들 (326) 을 포함할 수도 있다.

[0651] 도 14 에 예시된 바처럼, 브로드캐스트 레코딩 오디오 오브젝트들 (319), 프로페셔널 오디오 시스템들 (320) 및 소비자 온 디바이스 캡처 (322) 는 모두 그들의 출력을 HOA 오디오 포맷 (323) 을 이용하여 코딩할 수도 있다.

이런 식으로, 오디오 컨텐츠는, 온 디바이스 렌더링 (324), 소비자 오디오, TV 및 액세서리들 (325), 및 카 오디오 시스템들 (326) 을 이용하여 재생될 수도 있는 단일 표현으로 HOA 오디오 포맷 (323) 을 이용하여 코딩될 수도 있다. 즉, 오디오 컨텐츠의 단일 표현은 일반 오디오 재생 시스템에서 (즉, 5.1, 7.1 등과 같은 특정 구성을 필요로 하는 것과 반대로) 재생될 수도 있다.

[0652] 도 15a 및 도 15b는 도 12 의 오디오 에코시스템의 다른 예를 더 상세하게 예시하는 다이어그램들이다. 도 15a 에 예시된 바처럼, 오디오 에코시스템 (300C) 은 포착 엘리먼트들 (331), 및 재생 엘리먼트들 (336) 을 포함할 수도 있다. 포착 엘리먼트들 (331) 은, 유선 및/또는 무선 포착 디바이스들 (332) (예를 들어, 아이겐 마이크로폰들), 온 디바이스 서라운드 사운드 캡처 (334), 및 모바일 디바이스들 (335) (예를 들어, 스마트폰 및 태블릿들) 을 포함할 수도 있다. 일부 예들에서, 유선 및/또는 무선 포착 디바이스들 (332) 은 유선 및/또는 무선 통신 채널(들) (333) 을 통해 모바일 디바이스 (335) 에 연결될 수도 있다.

[0653] 본 개시의 하나 이상의 기술들에 따르면, 모바일 디바이스 (335) 는 음장을 포착하는데 사용될 수도 있다. 가령, 모바일 디바이스 (335) 는 유선 및/또는 무선 포착 디바이스들 (332) 및/또는 온 디바이스 서라운드 사운

드 캡처 (334) (예를 들어, 모바일 디바이스 (335) 내에 통합되는 복수의 마이크로폰들) 을 통해 음장을 포착할 수도 있다. 다음으로, 모바일 디바이스 (335) 는 하나 이상의 재생 엘리먼트들 (336) 에 의한 재생을 위해 포착된 음장을 HOA들 (337) 로 코딩할 수도 있다. 가령, 모바일 디바이스 (335) 의 사용자는 라이브 이벤트 (예를 들어, 미팅, 컨퍼런스, 연극, 콘서트 등) 을 레코딩 (라이브 이벤트의 음장을 포착) 하고 그 레코딩을 HOA들로 코딩할 수도 있다.

[0654] 모바일 디바이스 (335) 는 또한, HOA 코딩된 음장을 재생하기 위하여 재생 엘리먼트들 (336) 중의 하나 이상을 이용할 수도 있다. 가령, 모바일 디바이스 (335) 는 HOA 코딩된 음장을 디코딩하고, 재생 엘리먼트들 (336) 중의 하나 이상으로 하여금 음장을 재생성하게 하는 신호를 재생 엘리먼트들 (336) 중의 하나 이상으로 출력할 수도 있다. 일 예로서, 모바일 디바이스 (335) 는, 신호를 하나 이상의 스피커들 (예를 들어, 스피커 어레이, 사운드 바 등) 로 출력하기 위하여 무선 및/또는 무선 통신 채널들 (338) 을 이용할 수도 있다. 또 다른 예로서, 모바일 디바이스 (335) 는 하나 이상의 도킹 스테이션들 및/또는 하나 이상의 도킹된 스피커들 (예를 들어, 스마트 카 및/또는 가정에 있는 사운드 시스템들) 로 신호를 출력하기 위하여 도킹 솔루션들 (339) 을 이용할 수도 있다. 또 다른 예로서, 모바일 디바이스 (335) 는, 예를 들어, 현실적인 바이오럴 사운드를 생성하기 위하여, 헤드폰들의 세트로 신호를 출력하는데 헤드폰 렌더링 (340) 을 이용할 수도 있다.

[0655] 일부 예들에서, 특정 모바일 디바이스 (335) 는 3D 음장을 포착할 뿐만 아니라 나중에 동일한 3D 음장을 재생할 수도 있다. 일부 예들에서, 모바일 디바이스 (335) 는 3D 음장을 포착하고, 3D 음장을 HOA 로 인코딩하고, 재생을 위해 인코딩된 3D 음장을 하나 이상의 다른 디바이스들 (예를 들어, 다른 모바일 디바이스들 및/또는 다른 비모바일 디바이스들) 로 송신할 수도 있다.

[0656] 도 15b 에 예시된 바처럼, 오디오 에코시스템 (300D) 은 오디오 컨텐츠 (343), 게임 스튜디오들 (344), 코딩된 오디오 컨텐츠 (345), 렌더링 엔진들 (346), 및 전달 시스템들 (347) 을 포함할 수도 있다. 일부 예들에서, 게임 스튜디오들 (344) 은, HOA 신호들의 편집을 지원할 수도 있는 하나 이상의 DAW 들을 포함할 수도 있다. 가령, 하나 이상의 DAW 들은, 하나 이상의 게임 오디오 시스템들과 동작 (예를 들어, 작동) 하도록 구성될 수도 있는 HOA 플러그인들 및/또는 툴들을 포함할 수도 있다. 일부 예들에서, 게임 스튜디오들 (344) 은 HOA 를 지원하는 새로운 스템 포맷들을 출력할 수도 있다. 어느 경우든, 게임 스튜디오들 (344) 은, 전달 시스템들 (347) 에 의한 음장을 렌더링할 수도 있는 렌더링 엔진들 (346) 로 코딩된 오디오 컨텐츠 (345) 를 출력할 수도 있다.

[0657] 도 16 는 본 개시물에 기재된 기술들의 다양한 양태들을 수행할 수도 있는 일 예의 오디오 인코딩 디바이스를 예시하는 다이어그램이다. 도 16 에 예시된 바처럼, 오디오 에코시스템 (300E) 은 원래의 3D 오디오 컨텐츠 (351), 인코더 (352), 비트스트림 (353), 디코더 (354), 렌더러 (355) 및 재생 엘리먼트들 (356) 을 포함할 수도 있다. 또한, 도 16 에 의해 예시된 바처럼, 인코더 (352) 는 음장 분석 및 분해 (357), 배경 추출 (358), 배경 돌출 결정 (359), 오디오 코딩 (360), 전경/구별되는 오디오 추출 (361), 및 오디오 코딩 (362) 을 포함할 수도 있다. 일부 실시형태들에서, 인코더 (352) 는 도 3 및 도 4의 오디오 인코딩 디바이스 (20) 에 유사한 동작들을 수행하도록 구성될 수도 있다. 일부 예들에서, 음장 분석 및 분해 (357) 는 도 4의 음장 분석 유닛 (44) 에 유사한 동작들을 수행하도록 구성될 수도 있다. 일부 예들에서, 배경 추출 (358) 및 배경 돌출 결정 (359) 은 도 4의 BG 선택 유닛 (48) 에 유사한 동작들을 수행하도록 구성될 수도 있다. 일부 예들에서, 오디오 코딩 (360) 및 오디오 코딩 (362) 은 도 4의 심리 음향 오디오 코더 유닛 (40) 에 유사한 동작들을 수행하도록 구성될 수도 있다. 일부 예들에서, 전경/구별되는 오디오 추출 (361) 은 도 4의 전경 선택 유닛 (36) 에 유사한 동작들을 수행하도록 구성될 수도 있다.

[0658] 일부 예들에서, 전경/구별되는 오디오 추출 (361) 은 도 33 의 비디오 프레임 (390) 에 대응하는 오디오 컨텐츠를 분석할 수도 있다. 가령, 전경/구별되는 오디오 추출 (361) 은, 영역들 (391A - 391C) 에 대응하는 오디오 컨텐츠가 전경 오디오라는 것을 결정할 수도 있다.

[0659] 도 16 에 예시된 바처럼, 인코더 (352) 는, 25 - 75 Mbps 의 비트레이트를 가질 수도 있는 원래 컨텐츠 (351) 를 256kbps - 1.2Mbps 의 비트레이트를 가질 수도 있는 비트스트림 (353) 으로 인코딩하도록 구성될 수도 있다.
도 17은 도 16의 오디오 인코딩 디바이스의 일 예를 더 상세하게 예시하는 다이어그램이다.

[0660] 도 18 은 본 개시물에 기재된 기술들의 다양한 양태들을 수행할 수도 있는 일 예의 오디오 디코딩 디바이스를 예시하는 다이어그램이다. 도 18 에 예시된 바처럼, 오디오 에코시스템 (300E) 은 원래 3D 오디오 컨텐츠 (351), 인코더 (352), 비트스트림 (353), 디코더 (354), 렌더러 (355) 및 재생 엘리먼트들 (356) 을 포함할 수도 있다. 또한 도 16 에 예시된 바처럼, 디코더 (354) 는 오디오 디코더 (363), 오디오 디코더 (364), 전경

복원부 (365) 및 믹싱부 (366) 를 포함할 수도 있다. 일부 예들에서, 디코더 (354) 는 도 3 및 도 5 의 오디오 디코딩 디바이스 (24) 에 유사한 동작들을 수행하도록 구성될 수도 있다. 일부 예들에서, 오디오 디코더 (363), 오디오 디코더 (364) 는 도 5 의 심리 음향 디코딩 유닛 (80) 에 유사한 동작들을 수행하도록 구성될 수도 있다. 일부 예들에서, 전경 복원부 (365) 는 도 5 의 전경 공식화 유닛 (78) 에 유사한 동작들을 수행하도록 구성될 수도 있다.

[0661] 도 16 에 예시된 바처럼, 디코더 (354) 는 비트스트림 (353) 을 수신 및 디코딩하고 결과적인 복원된 3D 음장을 렌더러 (355) 에 출력하도록 구성될 수도 있고, 다음으로 이 렌더러 (355) 는 재생 엘리먼트들 (356) 중의 하나 이상으로 하여금 원래 3D 컨텐츠 (351) 의 표현을 출력하게 할 수도 있다. 도 19 는 도 18 의 오디오 디코딩 디바이스의 일 예를 더 상세하게 예시하는 다이어그램이다.

[0662] 도 20a 내지 도 20g 는 본 개시물에 기재된 기술들의 다양한 양태들을 수행할 수도 있는 예시적인 오디오 포착 디바이스들을 예시하는 다이어그램들이다. 도 20a 는, 3D 음장을 레코딩하도록 집합적으로 구성되는 복수의 마이크로폰들을 포함할 수도 있는 아이겐 마이크로폰 (370) 을 예시한다. 일부 예들에서, 아이겐 마이크로폰 (370) 의 복수의 마이크로폰들은, 대략 4cm 의 반경을 갖는 실질적으로 구형 볼의 표면에 위치될 수도 있다. 일부 예들에서, 오디오 인코딩 디바이스 (20) 는 마이크로폰 (370) 으로부터 직접 비트스트림 (17) 을 출력하기 위하여 아이겐 마이크로폰내에 통합될 수도 있다.

[0663] 도 20b 는, 하나 이상의 아이겐 마이크로폰들 (370) 과 같은 하나 이상의 마이크로폰들로부터 신호를 수신하도록 구성될 수도 있는 제조 트럭 (production truck; 372) 을 예시한다. 제조 트럭 (372) 은 또한, 도 3 의 오디오 인코더 (20) 와 같은 오디오 인코더를 포함할 수도 있다.

[0664] 도 20c 내지 도 20e 는, 3D 음장을 레코딩하도록 집합적으로 구성되는 복수의 마이크로폰들을 포함할 수도 있는 모바일 디바이스 (374) 를 예시한다. 다른 말로, 복수의 마이크로폰은 X, Y, Z 다이버시티 (diversity) 를 가질 수도 있다. 일부 예들에서, 모바일 디바이스 (374) 는 마이크로폰 (376) 을 포함할 수도 있고, 이 마이크로폰 (376) 은 모바일 디바이스 (374) 의 하나 이상의 다른 마이크로폰들에 관하여 X, Y, Z 다이버시티를 제공하기 위하여 회전될 수도 있다. 모바일 디바이스 (374) 는 또한, 도 3 의 오디오 인코더 (20) 와 같은 오디오 인코더를 포함할 수도 있다.

[0665] 도 20f 는, 3D 음장을 기록하도록 구성될 수도 있는 러기다이즈드 (ruggedized) 비디오 캡처 디바이스 (378) 를 예시한다. 일부 예들에서, 러기다이즈드 비디오 캡처 디바이스 (378) 는 활동에 관여된 사용자의 헬멧에 부착될 수도 있다. 가령, 러기다이즈드 비디오 캡처 디바이스 (378) 는, 급류 래프팅 사용자의 헬멧에 부착될 수도 있다. 이런 식으로, 러기다이즈드 비디오 캡처 디바이스 (378) 는, (예를 들어, 사용자 뒤쪽의 물 충돌, 사용자 앞쪽에서 말하고 있는 또 다른 래프터 등) 사용자의 모든 주변의 액션을 나타내는 3D 음장을 캡처할 수도 있다.

[0666] 도 20g 는, 3D 음장을 레코딩하도록 구성될 수도 있는 액세서리 향상 모바일 디바이스 (380) 를 예시한다. 일부 예들에서, 모바일 디바이스 (380) 는, 하나 이상의 액세서리들이 추가되는, 도 15 의 모바일 디바이스 (335) 와 유사할 수도 있다. 가령, 아이겐 마이크로폰이 도 15 의 모바일 디바이스 (335) 에 부착되어 액세서리 향상 모바일 디바이스 (380) 를 형성할 수도 있다. 이런 식으로, 액세서리 향상 모바일 디바이스 (380) 는, 액세서리 향상 모바일 디바이스 (380) 에 내장된 사운드 캡처 성분들만을 이용하는 것보다 더 높은 품질 버전의 3D 음장을 캡처할 수도 있다.

[0667] 도 21a 내지 도 21e 는 본 개시물에 기재된 기술들의 다양한 양태들을 수행할 수도 있는 예시적인 오디오 재생 디바이스들을 예시하는 다이어그램들이다. 도 21a 내지 도 21b 는 복수의 스피커들 (382) 및 사운드 바들 (384) 을 예시한다. 본 개시의 하나 이상의 기술들에 따르면, 스피커들 (382) 및/또는 사운드 바들 (384) 은, 여전히 3D 음장을 재생하면서 어느 임의의 구성으로도 배열될 수도 있다. 도 21c 내지 도 21e 는 복수의 헤드폰 재생 디바이스들 (386 - 386C) 을 예시한다. 헤드폰 재생 디바이스들 (386 - 386C) 은, 유선 또는 무선 접속 중 어느 하나를 통해 디코더에 연결될 수도 있다. 본 개시의 하나 이상의 기술들에 따르면, 음장의 단일 일반적 표현이 스피커들 (382), 사운드 바들 (384), 및 헤드폰 재생 디바이스들 (386 - 386C) 의 임의의 조합 상에 음장을 렌더링하는데 이용될 수도 있다.

[0668] 도 22a 내지 도 22h 는 본 개시물에 기재된 하나 이상의 기술들에 따른 예시적인 오디오 재생 환경들을 예시하는 다이어그램들이다. 가령, 도 22a 는 5.1 스피커 재생 환경을 예시하고, 도 22b 는 2.0 (예를 들어, 스테레오) 스피커 재생 환경을 예시하고, 도 22c 는 풀 하이트 전면 라우드스피커를 갖는 9.1 스피커 재생 환경을

예시하고, 도 22d 및 도 22e 각각은 22.2 스피커 재생 환경을 예시하고, 도 22f 는 16.0 스피커 재생 환경을 예시하고, 도 22g 는 자동차 스피커 재생 환경을 예시하고, 도 22h 는 이어 버드 (ear bud) 재생 환경을 갖는 모바일 디바이스를 예시한다.

[0669] 본 개시의 하나 이상의 기술들에 따르면, 음장의 단일 일반적 표현이 도 22a 내지 도 22h 에 예시된 재생 환경들 중의 어느 것에 상에 음장을 렌더링하는데 이용될 수도 있다. 추가적으로, 본 개시의 기술들은, 렌더러로 하여금, 도 22a 내지 도 22h 에 예시된 것들 외의 재생 환경들 상에 재생을 위한 일반 표현으로부터의 음장을 렌더링하는 것을 가능하게 한다. 예를 들어, 설계 고려사항들이 7.1 스피커 재생 환경에 따른 스피커들의 적당한 배치를 금하는 경우 (예를 들어, 우측 서라운드 스피커를 배치하는 것이 불가능한 경우), 본 개시의 기술들은, 렌더러로 하여금, 6.1 스피커 재생 환경 상에서 재생이 달성될 수도 있도록 다른 6개의 스피커들로 보상하는 것을 가능하게 한다.

[0670] 도 23 에 예시된 바처럼, 사용자는 헤드폰들 (386) 을 착용하고 있으면서 스포츠 게임을 볼 수도 있다. 본 개시의 하나 이상의 기술들에 따르면, 스포츠 게임의 3D 음장이 포착될 수도 있으며 (예를 들어, 하나 이상의 아이겐 마이크로폰들이 도 24 에 예시된 야구 스타디움에 및/또는 주변에 배치될 수도 있다), 3D 음장에 대응하는 HOA 계수들이 획득되고 디코더로 송신될 수도 있고, 디코더는 HOA 계수들에 기초하여 3D 음장을 복원하고 복원된 3D 음장을 렌더러로 출력하는것을 결정할 수도 있고, 렌더러는 그 재생 환경 (예를 들어, 헤드폰들) 의 유형에 대한 표시를 획득할 수도 있고, 헤드폰들로 하여금 스포츠 게임의 3D 음장의 표현을 출력하게 하는 신호들로 복원된 3D 음장을 렌더링할 수도 있다. 일부 예들에서, 렌더러는 도 25 의 기술들에 따라 재생 환경의 유형에 대한 표시를 획득할 수도 있다. 이런 식으로, 렌더러는 다양한 스피커 위치들, 수 유형, 크기를 위해 "적응" (adapt) 될 수도 있고 또한 이상적으로 로컬 환경을 위해 등화될 수도 있다.

[0671] 도 28 은, 본 개시물에 기재된 하나 이상의 기술들에 따라 헤드폰들에 의해 시뮬레이션될 수도 있는 스피커 구성을 예시하는 다이어그램이다. 도 28 에 의해 예시된 바처럼, 본 개시의 기술들은, 헤드폰들 (389) 을 착용하고 있는 사용자가, 음장을, 마치 그 음장이 스피커들 (388) 에 의해 재생된 것처럼, 체험하는 것을 가능하게 할 수도 있다. 이런 식으로, 사용자는 큰 영역에 사운드가 출력됨이 없이 3D 음장을 청취할 수도 있다.

[0672] 도 30 은, 본 개시물에 기재된 하나 이상의 기술들에 따라 프로세싱될 수도 있는 3D 음장과 연관된 비디오 프레임을 예시하는 다이어그램이다.

[0673] 도 31a 내지 도 31m 은 본 개시물에 기재된 기술들의 다양한 양태들에 따른 음장의 합성 또는 레코딩된 카테고리화를 수행하는 다양한 시뮬레이션 결과들을 보여주는 그래프들 (400A-400M) 을 예시하는 다이어그램들이다.

도 31a 내지 도 31m 의 예들에서, 그래프들 (400A-400M) 의 각각은, 점선에 의해 표기되는 임계치 (402) 및 파선에 의해 표기되는 각각의 오디오 오브젝트 (404A-404M) (총괄적으로, "오디오 오브젝트들 (404)") 을 포함한다.

[0674] 컨텐츠 분석 유닛 (26) 에 관하여 위에서 설명된 분석을 통하여 오디오 오브젝트들 (404) 이 임계치 (402) 아래에 있는 것으로 결정될 때, 컨텐츠 분석 유닛 (26) 은, 오디오 오브젝트들 (404) 중의 대응하는 하나가 레코딩된 오디오 오브젝트를 나타낸다는 것을 결정한다. 도 31b, 도 31d - 도 31h 및 도 31j - 도 31l 중의 예들에 보여진 바처럼, 컨텐츠 분석 유닛 (26) 은 오디오 오브젝트들 (404B, 404D-404H, 404J-404L) 이 임계치 (402) (시간의 적어도 +90% 및 종종 시간의 100%) 밑에 있고 따라서 레코딩된 오디오 오브젝트들을 나타낸다고 결정한다. 도 31a, 도 31c 및 도 31l 의 예들에 보여진 바처럼, 컨텐츠 분석 유닛 (26) 은, 오디오 오브젝트들 (404A, 404C 및 404I) 이 임계치 (402) 를 초과하고 따라서 합성 오디오 오브젝트들을 나타낸다고 결정한다.

[0675] 도 31m 의 예에서, 오디오 오브젝트 (404M) 는, (예를 들어, 임계치 (402) 더 높은) 일부 합성 부분들 및 (예를 들어, 임계치 (402) 보다 낮은) 일부 합성 부분들을 갖는, 믹싱된 합성/레코딩된 오디오 오브젝트를 나타낸다.

이 예에서 컨텐츠 분석 유닛 (26) 은 오디오 오브젝트 (404M) 의 합성 및 레코딩된 부분들을 식별하고 그 결과 오디오 인코딩 디바이스 (20) 는 방향성 기반 인코딩된 오디오 데이터 및 벡터 기반 인코딩된 오디오 데이터 양자 모두를 포함할 비트스트림 (21) 을 생성한다.

[0676] 도 32 는 본 개시물에 기재된 기술들에 따른 고차 앰비소닉 계수들로부터 분해된 S 행렬로부터 특이값들의 그래프 (406) 를 예시하는 다이어그램이다. 도 32에 도시된 바처럼, 큰 값들을 갖는 비-제로 특이값들은 적다.

도 4 의 음장 분석 유닛 (44) 은, 리오더링된 US[k] 벡터들 (33') 및 리오더링된 V[k] 벡터들 (35') 의 nFG 전경 (또는, 다른 말로, 우월한) 성분들 (종종, 벡터들에 의해 표시됨) 을 결정하기 위하여 이들 특이값들을 분

석할 수도 있다.

[0677] 도 33a 및 도 33b 는 본 개시물에 기재된 기술들에 따른 음장의 전경 성분들을 기술하는 벡터들을 인코딩할 때 리오더링이 갖는 잠재적인 영향을 보여주는 각각의 그래프들 (410A 및 410B) 을 예시하는 다이어그램들이다. 그래프 (410A) 는 오더링되지 않은 (또는, 다른 말로, 원래의) US[k] 벡터들 (33) 중의 적어도 일부를 인코딩하는 결과를 보여주는 한편, 그래프 (410B) 는 오더링된 US[k] 벡터들 (33') 중의 대응하는 것들을 인코딩하는 결과를 보여준다. 그래프들 (410A 및 410B) 의 각각에 있는 상단 플롯은 인코딩에 있어서 에러를 보여주고, 여기서 그래프 (410B) 에서 뚜렷한 에러는 프레임 경계들에서만 있을 가능성이 있다. 따라서, 본 개시물에 기재된 리오더링 기술들은 레거시 오디오 코더를 이용하여 모노-오디오 오브젝트들의 코딩을 용이하게 하거나 또는 그렇지 않으면 증진시킬 수도 있다.

[0678] 도 34 및 도 35는 본 개시물에 따른 구별되는 오디오 오브젝트들의 오직 에너지 기반 및 방향성 기반 식별 사이의 차이들을 예시하는 개념도들이다. 도 34의 예에서, 보다 큰 에너지를 나타내는 벡터들은, 방향성에 관계없이, 구별되는 오디오 오브젝트들인 것으로 식별된다. 도 34 에 도시된 바처럼, (y-축 상에 플롯화된) 더 높은 에너지 값들에 따라 배치되는 오디오 오브젝트들은, (예를 들어, x-축 상에 플롯화된 방향성 지수들에 의해 표현되는) 방향성에 관계 없이, "전경 내" 에 있는 것으로 결정된다.

[0679] 도 35 는 방향성 및 에너지 양자 모두에 기초하여, 이를테면 도 4 의 음장 분석 유닛 (44) 에 의해 구현되는 기술들에 따른, 구별되는 오디오 오브젝트들의 식별을 예시한다. 도 35에 도시된 바처럼, 더 큰 방향성 지수들이 x-축의 좌측쪽으로 플롯되고, 더 큰 에너지 레벨들이 y-축의 상단쪽으로 플롯화된다. 이 예에서, 음장 분석 유닛 (44) 은, (예를 들어, "전경 내" 에 있는) 구별되는 오디오 오브젝트들이 그래프의 상단 좌측쪽으로 상대적으로 플롯화된 벡터 데이터와 연관된다고 결정할 수도 있다. 일 예로서, 음장 분석 유닛 (44) 은 그래프의 상단 좌측 사분면에 플롯화된 그러한 벡터들이 구별되는 오디오 오브젝트들과 연관된다고 결정할 수도 있다.

[0680] 도 36a 내지 도 36f 는, 본 개시물에 기재된 기술들의 다양한 양태들에 따라 보간을 수행하기 위하여 공간 도메인으로의 구면 조화 계수의 분해된 버전의 적어도 일부의 투영들을 예시하는 다이어그램들이다. 도 36a 는 구체 (412) 상으로의 V[k] 벡터들 (35) 중의 하나 이상의 투영을 예시하는 다이어그램이다. 도 36a 의 예에서, 각각의 수는 (가능하게는 V 행렬 (19') 의 하나의 행 및/또는 열과 연관된) 구체 상으로 투영된 상이한 구면 조화 계수를 식별한다. 상이한 컬러들은 구별되는 오디오 성분들의 방향을 시사하고, 여기서 더 밝은 (그리고 점진적으로 더 어두워지는) 컬러는 구별되는 성분의 1차 방향을 표시한다. 도 4의 예에 보여진 오디오 인코딩 디바이스 (20) 의 공간-시간 보간 유닛 (50) 은 도 36a 의 예에 보여진 구체를 생성하기 위하여 적색점들의 각각 사이의 공간-시간 보간을 수행할 수도 있다.

[0681] 도 36b 는 빔 상으로의 V[k] 벡터들 (35) 중의 하나 이상의 투영들을 예시하는 다이어그램이다. 공간-시간 보간 유닛 (50) 은 V[k] 벡터들 (35) 의 하나의 행 및/또는 열 또는 V[k] 벡터들 (35) 의 다수의 행들 및/또는 열들을 투영하여 도 36b 의 예에 보여진 빔 (414) 을 생성할 수도 있다.

[0682] 도 36c 는 도 36 의 예에 보여진 구체 (412) 와 같은 구체 상으로의 V[k] 벡터들 (35) 중의 하나 이상의 하나 이상의 벡터들의 투영의 단면을 예시하는 다이어그램이다.

[0683] 도 36d 내지 도 36g 는, 상이한 사운드 소스들 (벌, 헬리콥터, 전자 음악, 및 스타디움에 있는 사람들) 이 3 차원 공간에 예시될 수도 있을 때 (약 20 밀리초의 1 프레임에 대한) 시간의 스냅샷들의 예들이다.

[0684] 본 개시물에 기재된 기술들은 이들 상이한 사운드 소스들의 표현이 단일 US[k] 벡터 및 단일 V[k] 벡터를 이용하여 식별되고 나타내어지는 것을 허용한다. 사운드 소스들의 시간 가변성은 US[k] 벡터에 나타나 있는 한편, 각각의 사운드 소스의 공간 분포는 단일 V[k] 벡터에 의해 나타내어진다. 하나의 V[k] 벡터는 사운드 소스의 폭, 위치 및 크기를 나타낼 수도 있다. 게다가, 단일 V[k] 벡터는 구면 조화 기저 함수들의 선형 조합으로서 나타내어질 수도 있다. 도 36d 내지 도 36g 의 플롯들에서, 사운드 소스들의 표현은, 공간 좌표 시스템으로의 단일 V 벡터들을 변환하는 것에 기초할 수도 있다. 사운드 소스들을 예시하는 유사한 방법들이 도 36 내지 도 36c 에서 사용된다.

[0685] 도 37 은 여기에 기재된 바처럼 공간-시간 보간을 획득하기 위한 기술들의 표현을 예시한다. 도 4 의 예에 보여진 오디오 인코딩 디바이스 (20) 의 공간-시간 보간 유닛 (50) 은 더 상세하게 아래에 기재된 바처럼 공간-시간 보간을 수행할 수도 있다. 공간-시간 보간은 공간 및 시간 차원들 양자 모두에서 더 높은 해상도 공간 성분들을 획득하는 것을 포함할 수도 있다. 공간 성분들은 고차 앰비소닉 (higher-order ambisonic; HOA)

계수들로 구성된 다차원 신호의 직교 분해에 기초할 수도 있다 (또는, 여기서 HOA 계수들은 "구면 조화 계수들"로도 지칭될 수도 있다).

[0686] 예시된 그래프에서, 벡터들 V_1 및 V_2 은 다차원 신호의 2 개 상이한 공간 성분들의 대응하는 벡터들을 나타낸다.

공간 성분들은 다차원 신호의 블록 방식 (block-wise) 분해에 의해 획득될 수도 있다. 일부 예들에서, 공간 성분들은, 고차 앰비소닉스 (HOA) 오디오 데이터의 (프레임을 지칭할 수도 있는) 각각의 블록에 관하여 SVD 의 블록 방식 형태를 수행하는 것으로부터 비롯된다 (여기서 이 앰비소닉스 오디오 데이터는 블록들, 샘플들, 또는 임의의 다른 형태의 다중 채널 오디오 데이터를 포함한다). 변수 M 은 샘플들에서 오디오 프레임의 길이를 표시하는데 사용될 수도 있다.

[0687] 따라서, V_1 및 V_2 는 HOA 계수들 (11) 의 순차적인 블록들에 대해 전경 V[k] 벡터들 (51_k) 및 전경 V[k-1] 벡터들 (51_{k-1}) 의 대응하는 벡터들을 나타낼 수도 있다. V_1 는, 가령, 제 1 프레임 (k-1) 에 대해 전경 V[k-1] 벡터들 (51_{k-1}) 의 제 1 벡터를 나타내는 한편, V_2 는 제 2 그리고 후속 프레임 (k) 에 대해 전경 V[k] 벡터들 (51_k) 의 제 1 벡터를 나타낼 수도 있다. V_1 및 V_2 는 다차원 신호에 포함된 단일 오디오 오브젝트를 위한 공간 성분을 나타낼 수도 있다.

[0688] 각각의 x 에 대해 보간된 벡터들 V_x 은, 보간된 벡터들 V_x 이 시간 (및, 따라서, 일부 경우들에서는 공간) 성분을 평활화하기 위해 적용될 수도 있는 다차원 신호의 시간 성분에 대한 시간 세그먼트들 또는 "시간 샘플들"의 수, x 에 따라 V_1 및 V_2 를 가중함으로써 획득된다. 위에 기재된 바처럼, SVD 합성을 가정하면, nFG 신호들 (49) 을 평활화하는 것은, 대응하는 보간된 V_x 로 각각의 시간 샘플 벡터 (예를 들어, HOA 계수들 (11) 의 샘플) 의 벡터 분할 (vector division) 을 행함으로써 획득될 수도 있다. 즉, $US[n] = HOA[n] * V_x[n]^{-1}$, 여기서 이것은 컬럼 벡터가 곱해진 로우 벡터를 나타내며, 따라서 US 에 대해 스칼라 엘리먼트를 산출한다. $V_x[n]^{-1}$ 은 $V_x[n]$ 의 의사역 (pseudoinverse) 으로서 획득될 수도 있다.

[0689] V_1 및 V_2 의 가중에 관하여, V_1 은, V_2 가 V_1에 시간적으로 후속하여 일어나는 것에 기인하여, 시간 치원을 따라 비례적으로 낮게 가중된다. 즉, 전경 V[k-1] 벡터들 (51_{k-1}) 이 분해의 공간 성분들이지만, 시간적으로 순차적인 전경 V[k] 벡터들 (51_k) 은 시간이 흐름에 따라 공간 성분의 상이한 값을 나타낸다. 따라서, x 가 t 를 따라 증가함에 따라, V_1 의 가중치는 감소되는 반면, V_2 의 가중치는 커진다. 여기서 d_1 및 d_2 는 가중치들을 나타낸다.

[0690] 도 38 은 여기에 기재된 기술들에 따른 다차원 신호를 위한 순차적인 SVD 블록들을 위한 인공 (artificial) US 행렬들, 즉 US1 및 US2 를 예시하는 블록 다이어그램이다. 보간된 V-벡터들은 원래 다차원 신호를 복구하기 위하여 인공 US 행렬들의 로우 벡터들에 적용될 수도 있다. 보다 구체적으로, 공간-시간 보간 유닛 (50) 은, U2 행렬의 도 38의 예에 보여진 바처럼 제 1 K/2 샘플들로서 nFG 신호들의 K/2 샘플들 대신에 사용될 수도 있는, K/2 보간된 샘플들을 획득하기 위하여 (전경 HOA 계수들로 표시될 수도 있는) 전경 V[k] 벡터들 (51_k) 에 의해 nFG 신호들 (49) 을 곱한 결과에 보간된 전경 V[k] 벡터들 (53) 의 의사역을 곱할 수도 있다.

[0691] 도 39 는 본 개시물에 기재된 기술들에 따라 특이값 분해 및 공간-시간 성분들의 평활화를 사용하여 고차 앰비소닉스 (HOA) 신호의 후속 프레임들의 분해를 예시하는 블록 다이어그램이다. (또한 프레임 n 및 프레임 n+1 으로서 표기될 수도 있는) 프레임 n-1 및 프레임 n 은 시간적으로 후속하는 프레임들을 나타내며, 각각의 프레임은 1024 시간 세그먼트들을 포함하고 HOA 오더가 4 이며, $(4+1)^2 = 25$ 계수들을 낳는다. 프레임 n-1 및 프레임 n 에서 인공적으로 평활화된 U 행렬들인 US 행렬들은 예시된 바처럼 보간된 V 벡터들의 적용에 의해 획득될 수도 있다. 각각의 회색 로우 또는 컬럼 벡터들은 하나의 오디오 오브젝트를 나타낸다.

[0692] *액티브 벡터 기반 신호들의 HOA 표현을 컴퓨팅*

[0693] 인스턴스 CVECk 는, XVECk 에 나타낸 벡터 기반 신호들의 각각을 취하고 이를 그의 대응하는 (역양자화된) 공간 벡터, VVECk 와 곱함으로써 생성된다. 각각의 VVECk 는 MVECk 에서 나타내어진다. 따라서, 오더 L HOA 신호, 그리고 M 벡터 기반 신호들에 대해, M 개 벡터 기반 신호들이 있을 것이며, 이들의 각각은 프레임 길이 P 에 의해 주어지는 디멘전을 가질 것이다. 따라서, 이들 신호들은 XVECkmn, n=0,..P-1; m=0,..M-1 로서 나

타내어질 수 있다. 대응하여, 디멘전 $(L+1)^2$ 의 M 개 공간 벡터들 VVECk 이 있을 것이다. 이것들은 MVECkml, l=0,..,$(L+1)^{2-1}$;m=0,..,M-1 로서 나타내어질 수도 있다. 각각의 벡터 기반 신호를 위한 HOA 표현 CVECkm 는 다음에 의해 주어지는 행렬 벡터 곱셈이다:

[0694] $$\boldsymbol{C}\mathrm{VEC}km=(\boldsymbol{X}\mathrm{VEC}km(\boldsymbol{M}\mathrm{VEC}km)T)T$$

[0695] 이는 $(L+1)^{2\times}$ P 의 행렬을 산출한다. 완전한 HOA 표현은 다음과 같이 각각의 벡터 기반 신호의 기여를 합산함으로써 주어진다:

[0696] $$\boldsymbol{C}\mathrm{VEC}k=m=0M-1\boldsymbol{C}\mathrm{VEC}k[m]$$

[0697] *V-벡터들의 공간-시간 보간*

[0698] 하지만, 평활한 공간-시간 연속성을 유지하기 위하여, 위의 연산은 프레임 길이 P-B 의 부분에 대해서만 수행된다. 그 대신에 HOA 행렬의 제 1 B 샘플들은, 현재 MVECkm 및 이전 값들 MVECk-1m 로부터 도출된, 보간된 세트의 MVECkml, m=0,..,M-1;l=0,..,$(L+1)^2$ 를 이용하여 수행된다. 이것은 다음과 같이 각각의 시간 샘플 p 에 대한 벡터를 도출할 때 더 높은 시간 밀도 공간 벡터에 귀결된다:

[0699] $$\boldsymbol{M}\mathrm{VEC}kmp=pB-1\boldsymbol{M}\mathrm{VEC}km+B-1-pB-1\boldsymbol{M}\mathrm{VEC}k-1m,\ p=0,..,B-1.$$

[0700] 각각의 시간 샘플 p 에 대해, $(L+1)^2$ 디멘전의 새로운 HOA 벡터가 다음과 같이 컴퓨팅된다:

[0701] $$\boldsymbol{C}\mathrm{VEC}kp=(\boldsymbol{X}\mathrm{VEC}kmp)\boldsymbol{M}\mathrm{VEC}kmp,\ p=0,..,B-1$$

[0702] 이들, 제 1 B 샘플들은 이전 섹션의 P-B 샘플들로 증강되어, m번째 벡터 기반 신호의 완전한 HOA 표현, CVECkm 에 귀결된다.

[0703] 디코더 (예를 들어, 도 5 의 예에 나타낸 오디오 디코딩 디바이스 (24)) 에서, 어떤 구별되는, 전경 또는 벡터 기반 우월 사운드에 대해, 이전 프레임으로부터의 V-벡터 및 현재 프레임으로부터의 V 벡터가 특정 시간 세그먼트에 대해 고 해상도 (시간적으로) 보간된 V-벡터를 산출하기 위하여 선형 (또는 비선형) 보간을 이용하여 보간될 수도 있다. 공간-시간 보간 유닛 (76) 은 이 보간을 수행할 수도 있고, 여기서 공간-시간 보간 유닛 (76) 은 그 특정 시간 세그먼트에 대해 HOA 행렬을 산출하기 위하여 고 해상도 보간된 V-벡터와 현재 프레임에 있는 US 벡터를 곱할 수도 있다.

[0704] 대안적으로, 공간-시간 보간 유닛 (76) 은 제 1 HOA 행렬을 생성하기 위하여 현재 프레임의 V-벡터와 US 벡터를 곱할 수도 있다. 디코더는 추가적으로, 제 2 HOA 행렬을 생성하기 위하여 이전 프레임으로부터 V-벡터와 US 벡터를 곱할 수도 있다. 다음으로 공간-시간 보간 유닛 (76) 은, 특정 시간 세그먼트에 대해 제 1 및 제 2 HOA 행렬들에 선형 (또는 비선형) 보간을 적용할 수도 있다. 이 보간의 출력은 보간된 V-벡터와 US 벡터의 곱셈의 것에 매칭될 수도 있으며, 공통 입력 행렬들/벡터들을 제공한다.

[0705] 이와 관련하여, 그 기술들은 오디오 인코딩 디바이스 (20) 및/또는 오디오 디코딩 디바이스 (24) 로 하여금 하기 절들에 따라 동작하도록 구성될 수 있게 할 수도 있다.

[0706] 조항 135054-1C. 오디오 인코딩 디바이스 (20) 또는 오디오 디코딩 디바이스 (24) 와 같은 디바이스로서, 공간 및 시간 양자 모두에서 복수의 고 해상도 공간 성분들을 획득하도록 구성된 하나 이상의 프로세서들을 포함하고, 그 공간 성분들은 구면 조화 계수들로 구성되는 다차원 신호의 직교 분해에 기초하는, 디바이스.

[0707] 조항 135054-1D. 오디오 인코딩 디바이스 (20) 또는 오디오 디코딩 디바이스 (24) 와 같은 디바이스로서, 제 1 복수의 구면 조화 계수들 및 제 2 복수의 구면 조화 계수들의 공간 성분 및 시간 성분들 중의 적어도 하나를 평활화하도록 구성된 하나 이상의 프로세서들을 포함하는, 디바이스.

[0708] 조항 135054-1E. 오디오 인코딩 디바이스 (20) 또는 오디오 디코딩 디바이스 (24) 와 같은 디바이스로서, 공간 및 시간 양자 모두에서 복수의 고 해상도 공간 성분들을 획득하도록 구성된 하나 이상의 프로세서들을 포함하고, 그 공간 성분들은 구면 조화 계수들로 구성되는 다차원 신호의 직교 분해에 기초하는, 디바이스.

[0709] 조항 135054-1G. 오디오 인코딩 디바이스 (20) 또는 오디오 디코딩 디바이스 (24) 와 같은 디바이스로서,

제 1 복수의 구면 조화 계수들의 제 1 분해 및 제 2 복수의 구면 조화 계수들의 제 2 분해에 관하여 해상도를 적어도 부분적으로 증가시키는 것에 의해 시간 세그먼트에 대해 분해된 증가 해상도 구면 조화 계수들을 획득하도록 구성된 하나 이상의 프로세서들을 포함하는, 디바이스.

[0710] 조항 135054-2G. 조항 135054-1G 에 있어서, 제 1 분해는 제 1 복수의 구면 조화 계수들의 우-특이 벡터들을 표시하는 제 1 V 행렬을 포함하는, 디바이스.

[0711] 조항 135054-3G. 조항 135054-1G 에 있어서, 제 2 분해는 제 2 복수의 구면 조화 계수들의 우-특이 벡터들을 표시하는 제 2 V 행렬을 포함하는, 디바이스.

[0712] 조항 135054-4G. 조항 135054-1G 에 있어서, 제 1 분해는 제 1 복수의 구면 조화 계수들의 우-특이 벡터들을 표시하는 제 1 V 행렬을 포함하고, 제 2 분해는 제 2 복수의 구면 조화 계수들의 우-특이 벡터들을 표시하는 제 2 V 행렬을 포함하는, 디바이스.

[0713] 조항 135054-5G. 조항 135054-1G 에 있어서, 시간 세그먼트는 오디오 프레임의 서브 프레임을 포함하는, 디바이스.

[0714] 조항 135054-6G. 조항 135054-1G 에 있어서, 시간 세그먼트는 오디오 프레임의 시간 샘플을 포함하는, 디바이스.

[0715] 조항 135054-7G. 조항 135054-1G 에 있어서, 하나 이상의 프로세서들은 제 1 복수의 구면 조화 계수들의 구면 조화 계수를 위해 제 1 분해 및 제 2 분해의 보간된 분해를 획득하도록 구성되는, 디바이스.

[0716] 조항 135054-8G. 조항 135054-1G 에 있어서, 하나 이상의 프로세서들은, 제 1 프레임에 포함된 제 1 복수의 구면 조화 계수들의 제 1 부분을 위한 제 1 분해 및 제 2 프레임에 포함된 제 2 복수의 구면 조화 계수들의 제 2 부분을 위한 제 2 분해의 보간된 분해들을 획득하도록 구성되고, 하나 이상의 프로세서들은 또한, 제 1 복수의 구면 조화 계수들의 제 1 인공 시간 성분을 생성하기 위해 제 1 프레임에 포함된 제 1 복수의 구면 조화 계수들의 제 1 부분의 제 1 시간 성분에 보간된 분해들을 적용하고, 포함된 제 2 복수의 구면 조화 계수들의 제 2 인공 시간 성분을 생성하기 위해 제 2 프레임에 포함된 제 2 복수의 구면 조화 계수들의 제 2 부분의 제 2 시간 성분에 각각의 보간된 분해들을 적용하도록 구성되는, 디바이스.

[0717] 조항 135054-9G. 조항 135054-8G 에 있어서, 제 1 시간 성분은 제 1 복수의 구면 조화 계수들에 관하여 벡터 기반 합성을 수행하는 것에 의해 생성되는, 디바이스.

[0718] 조항 135054-10G. 제 2 시간 성분이 제 2 복수의 구면 조화 계수들에 대해 벡터 기반 합성을 수행하는 것에 의해 생성되는, 조항 135054-8G 의 디바이스.

[0719] 조항 135054-11G. 하나 이상의 프로세서들이 또한, 제 1 인공 시간 성분 및 제 2 인공 시간 성분을 수신하고, 제 1 복수의 구면 조화 계수들의 제 1 부분을 위한 제 1 분해 및 제 2 복수의 구면 조화 계수들의 제 2 부분을 위한 제 2 분해의 보간 분해들을 컴퓨팅하고, 보간된 분해들의 역들을 제 1 인공 시간 성분에 적용하여 제 1 시간 성분을 복구하고 또한 제 2 인공 시간 성분에 적용하여 제 2 시간 성분을 복구하도록 구성되는, 조항 135054-8G 의 디바이스.

[0720] 조항 135054-12G. 하나 이상의 프로세서들이 제 1 복수의 구면 조화 계수들의 제 1 공간 성분 및 제 2 복수의 구면 조화 계수들의 제 2 공간 성분을 보간하도록 구성되는, 조항 135054-1G 의 디바이스.

[0721] 조항 135054-13G. 제 1 공간 성분이 제 1 복수의 구면 조화 계수들의 좌-특이 벡터들을 나타내는 제 1 U 행렬을 포함하는, 조항 135054-12G 의 디바이스.

[0722] 조항 135054-14G. 제 2 공간 성분이 제 2 복수의 구면 조화 계수들의 좌-특이 벡터들을 나타내는 제 2 U 행렬을 포함하는, 조항 135054-12G 의 디바이스.

[0723] 조항 135054-15G. 제 1 공간 성분이 제 1 복수의 구면 조화 계수들에 대한 구면 조화 계수들의 M 개의 시간 세그먼트들을 나타내고, 제 2 공간 성분이 제 2 복수의 구면 조화 계수들에 대한 구면 조화 계수들의 M 개의 시간 세그먼트들을 나타내는, 조항 135054-12G 의 디바이스.

[0724] 조항 135054-16G. 제 1 공간 성분이 제 1 복수의 구면 조화 계수들에 대한 구면 조화 계수들의 M 개의 시간 세그먼트들을 나타내고, 제 2 공간 성분이 제 2 복수의 구면 조화 계수들에 대한 구면 조화 계수들의 M 개의 시간 세그먼트들을 나타내며, 그리고 하나 이상의 프로세서들이 시간 세그먼트에 대해 분해된 보간된 공간 조화

계수들을 획득하도록 구성되고, 제 1 공간 성분의 마지막 N 엘리먼트들 및 제 2 공간 성분의 제 1 N 엘리먼트들을 보간하는 것을 포함하는, 조항 135054-12G 의 디바이스.

[0725] 조항 135054-17G. 제 2 복수의 구면 조화 계수들이 시간 도메인에서 제 1 복수의 구면 조화 계수들에 후속하는, 조항 135054-1G 의 디바이스.

[0726] 조항 135054-18G. 하나 이상의 프로세서들이 또한, 제 1 복수의 구면 조화 계수들을 분해하여 제 1 복수의 구면 조화 계수들의 제 1 분해를 생성하도록 구성되는, 조항 135054-1G 의 디바이스.

[0727] 조항 135054-19G. 하나 이상의 프로세서들이 또한, 제 2 복수의 구면 조화 계수들을 분해하여 제 2 복수의 구면 조화 계수들의 제 2 분해를 생성하도록 구성되는, 조항 135054-1G 의 디바이스.

[0728] 조항 135054-20G. 하나 이상의 프로세서들이 또한, 제 1 복수의 구면 조화 계수들에 대해 특이값 분해를 수행하여 제 1 복수의 구면 조화 계수들의 좌-특이 벡터들을 나타내는 U 행렬, 제 1 복수의 구면 조화 계수들의 특이값들을 나타내는 S 행렬, 및 제 1 복수의 구면 조화 계수들의 우-특이 벡터들을 나타내는 V 행렬을 생성하도록 구성되는, 조항 135054-1G 의 디바이스.

[0729] 조항 135054-21G. 하나 이상의 프로세서들이 또한, 제 2 복수의 구면 조화 계수들에 대해 특이값 분해를 수행하여 제 2 복수의 구면 조화 계수들의 좌-특이 벡터들을 나타내는 U 행렬, 제 2 복수의 구면 조화 계수들의 특이값들을 나타내는 S 행렬, 및 제 2 복수의 구면 조화 계수들의 우-특이 벡터들을 나타내는 V 행렬을 생성하도록 구성된, 조항 135054-1G 의 디바이스.

[0730] 조항 135054-22G. 제 1 및 제 2 복수의 구면 조화 계수들이 각각 음장의 평면파 표현을 나타내는, 조항 135054-1G 의 디바이스.

[0731] 조항 135054-23G. 제 1 및 제 2 복수의 구면 조화 계수들이 각각 함께 혼합된 하나 이상의 모노 오디오 오브젝트들을 나타내는, 조항 135054-1G 의 디바이스.

[0732] 조항 135054-24G. 제 1 및 제 2 복수의 구면 조화 계수들이 각각 3 차원 음장을 나타내는 개별 제 1 및 제 2 구면 조화 계수들을 포함하는, 조항 135054-1G 의 디바이스.

[0733] 조항 135054-25G. 제 1 및 제 2 복수의 구면 조화 계수들이 1 보다 큰 오더를 갖는 적어도 하나의 구면 기저 함수와 각각 연관되는, 조항 135054-1G 의 디바이스.

[0734] 조항 135054-26G. 제 1 및 제 2 복수의 구면 조화 계수들이 각각 4 인 오더를 갖는 적어도 하나의 구면 기저 함수와 각각 연관되는, 조항 135054-1G 의 디바이스.

[0735] 조항 135054-27G. 보간이 제 1 분해 및 제 2 분해의 가중화된 보간이고, 제 1 분해에 적용된 가중화된 보간의 가중치들은 제 1 및 제 2 분해의 벡터로 나타낸 시간에 반비례하고, 제 2 분해에 적용된 가중화된 보간의 가중치들이 제 1 및 제 2 분해의 벡터들에 의해 나타낸 시간에 비례하는, 조항 135054-1G 의 디바이스.

[0736] 조항 135054-28G. 분해된 보간된 구면 조화 계수들은 제 1 복수의 구면 조화 계수들 및 제 2 복수의 구면 조화 계수들의 공간 성분들 및 시간 성분들 중 적어도 하나를 평활화하는, 조항 135054-1G 의 디바이스.

[0737] 도 40a 내지 도 40j 는 2 차원 또는 3 차원 음장들을 기술하는 구면 조화 계수들을 압축하기 위해 본 개시물에 기재된 기술들의 다양한 양태를 수행할 수도 있는 예시의 오디오 인코딩 디바이스들 (510A-510J) 를 예시하는 각각의 블록 다이어그램이다. 도 40a 내지 도 40j 의 예들의 각각에서, 오디오 인코딩 디바이스들 (510A 및 510B) 는 각각, 일부 예들에서, 오디오 데이터를 인코딩할 수도 있는 임의의 디바이스, 예컨대 데스크탑 컴퓨터, 랩탑 컴퓨터, 워크스테이션, 태블릿 또는 슬레이트 컴퓨터, 전용 오디오 레코딩 디바이스, 셀룰러 폰 (소위 "스마트 폰들" 을 포함), 개인용 미디어 플레이어 디바이스, 개인용 게이밍 디바이스, 또는 오디오 데이터를 인코딩할 수 있는 디바이스의 임의의 다른 타입을 나타낸다.

[0738] 단일 디바이스로서, 즉 도 40a 내지 도 40j 에서 디바이스들 (510A-510J) 로서 나타나 있지만, 디바이스들 (510A-510J) 내에 포함되는 것으로서 하기에서 지칭되는 다양한 성분들 또는 유닛들은 실제로 디바이스들 (510A-510J) 외부에 있는 별도의 디바이스들을 형성할 수도 있다. 즉, 본 개시물에서는 단일 디바이스에 의해, 즉 도 40a 내지 도 40j 의 예들에서의 디바이스들 (510A-510J) 에 의해 수행되는 것으로 기재되어 있지만, 기술들은 다중 디바이스들을 포함하는 시스템에 의해 구현되거나 또는 그렇지 않으면 수행될 수도 있으며, 여기서 이들 디바이스들은 각각 하기에서 더 상세하게 기재되는 다양한 성분들 또는 유닛들 중 하나 이상을 포함할 수도 있다. 따라서, 기술들은 도 40a 내지 도 40j 의 예들에 한정되지 않아야 한다.

[0739] 일부 예들에서, 오디오 인코딩 디바이스들 (510A-510J) 은 도 3 및 도 4 의 예들에 대해 전술한 것에 대한 대안의 오디오 인코딩 디바이스들을 나타낸다. 오디오 인코딩 디바이스들 (510A-510J) 의 하기의 논의 전체에 걸쳐 동작에 관한 다양한 유사성들은 도 4 에 대해 전술한 오디오 인코딩 디바이스 (20) 의 다양한 유닛들 (30-52) 에 대해 주목된다. 많은 양태들에서, 오디오 인코딩 디바이스들 (510A-510J) 은, 하기에 기재되는 바와 같이, 약간의 도출들 또는 수정들을 갖지만 오디오 인코딩 디바이스 (20) 와 실질적으로 유사한 방식으로 동작할 수도 있다.

[0740] 도 40a 의 예에 나타낸 바와 같이, 오디오 인코딩 디바이스 (510A) 는 오디오 압축 유닛 (512), 오디오 인코딩 유닛 (514) 및 비트스트림 생성 유닛 (516) 을 포함한다. 오디오 압축 유닛 (512) 은 구면 조화 계수들 (SHC)(511)("SHC (511)") 을 압축하는 유닛을 나타낼 수도 있으며, 이는 또한 고차 앰비소닉스 (HOA) 계수들 (511) 로서 지칭될 수도 있다. 일부 경우들에서, 오디오 압축 유닛 (512) 은 SHC (511) 에 대해 무손실로 압축하거나 로시 (lossy) 압축을 수행할 수도 있는 유닛을 나타낸다. SCH (511) 은 복수의 SHC들을 나타낼 수도 있으며, 여기서 복수의 SHC 중 적어도 하나는 위에서 더 상세하게 기재된 바와 같이, 1 보다 큰 오더를 갖는 구면 기저 함수에 대응한다 (여기서, 이러한 다양성의 SHC 는 일 예가 소위 "B-포맷" 인 저차 앰비소닉스와 구별하도록 고차 앰비소닉스로서 지칭된다). 오디오 압축 유닛 (512) 이 SHC (511) 을 무손실로 압축할 수도 있지만, 일부 예들에서, 오디오 압축 유닛 (512) 은 (일부는 인간의 청각 체계에 의해 들을 수 없을 수도 있다는 점에서) 재생될 때 음장을 기술하는데 있어서 돌출 또는 관련이 없는 SHC (511) 의 것들을 제거한다. 이러한 관점에서, 이러한 압축의 로시 성질은 SHC (511) 의 압축된 버전으로부터 재생될 때 음장의 감지된 품질에 과도하게 영향을 미치지 않을 수도 있다.

[0741] 도 40a 의 예에서, 오디오 압축 유닛은 분해 유닛 (518) 및 음장 성분 추출 유닛 (520) 을 포함한다. 분해 유닛 (518) 은 오디오 인코딩 디바이스 (20) 의 선형 가역 변환 유닛 (30) 과 유사할 수도 있다. 즉, 분해 유닛 (518) 은 특이값 분해로서 지칭되는 분석의 형태를 수행하도록 구성되는 유닛을 나타낼 수도 있다. SVD 에 대해 기재되었지만, 기술들은 선형으로 보정되지 않은 데이터의 세트들을 제공하는 임의의 유사한 변환 또는 분해에 대해 수행될 수도 있다. 또한, 본 개시물에서 "세트들" 에 대한 언급은 상반되는 것으로 특별히 정해지지 않으면 "0 이 아닌 것" 을 지칭하는 것으로 의도되고 소위 "공 세트" 를 포함하는 세트들의 전형적인 수학적 정의를 지칭하는 것으로 의도되지 않는다.

[0742] 어떤 경우든, 분해 유닛 (518) 은 특이값 분해 (다시, 그 두문자어로 "SVD" 로 지칭될 수도 있음) 를 수행하여 구면 조화 계수들 (511) 을 변환된 구면 조화 계수들의 2 이상의 세트들로 변환한다. 도 40 의 예에서, 분해 유닛 (518) 은 SHC (511) 에 대해 SVD 를 수행하여 소위 V 행렬 (519), S 행렬 (519B), 및 U 행렬 (519C) 을 생성할 수도 있다. 도 40 의 예에서, 분해 유닛 (518) 은 선형 가역 변환 유닛 (30) 에 대해 전술한 바와 같이 결합된 형태에서 US[k] 벡터들을 출력하기 보다는 별도로 별도로 행렬들의 각각을 출력한다.

[0743] 위에서 지적된 바와 같이, 위에서 언급된 SVD 수학적 표현에서 V* 행렬은, SVD 가 복소수들을 포함하는 행렬들에 적용될 수도 있는 것을 반영하기 위해 V 행렬의 컨주게이트 트랜스포즈로서 지칭된다. 단지 실수들만을 포함하는 행렬들에 적용될 때, V 행렬의 복소 컨주게이트 (또는, 다시 말해서, V* 행렬) 는 V 행렬과 동등한 것으로 고려될 수도 있다. 예시를 용이하게 하기 위해, 하기에서는 V 행렬이 V* 행렬 보다는 SVD 를 통해 출력되는 결과를 갖는 실수들을 SHC (511) 가 포함한다는 것이 가정된다. V 행렬인 것이 가정되지만, 기술들은 복소 계수들을 갖는 SHC (511) 와 유사한 방식으로 적용될 수도 있으며, 여기서 SVC 의 출력은 V* 행렬이다. 따라서, 기술들은 이러한 V 행렬을 생성하기 위해 SVD 의 적용을 제공하는 것에만 제한되어야 하는 것이 아니라, V* 행렬을 생성하기 위해 복소 성분들을 갖는 SHC (511) 로의 SVD 의 적용을 포함할 수도 있다.

[0744] 어떤 경우든, 분해 유닛 (518) 은 고차 앰비소닉스 (HOA) 오디오 데이터 (여기서 이 앰비소닉스 오디오 데이터는 SHC (511) 의 블록들 또는 샘플들 또는 다중 채널 오디오 데이터의 임의의 다른 형태를 포함한다) 의 각각의 블록 (프레임을 지칭할 수도 있음) 에 대해 SVD 의 블록 방식 형태를 수행할 수도 있다. 변수 M 은 샘플들에서 오디오 프레임의 길이를 지칭하기 위해 사용될 수도 있다. 예를 들어, 오디오 프레임이 1024 오디오 샘플들을 포함할 때, M 은 1024 이다. 이에 따라, 분해 유닛 (518) 은 블록에 대해 블록 방식 SVD 를 수행할 수도 있고, SHC (511) 은 M-바이-$(N+1)^2$ SHC 를 가지며, 여기서 N 은 다시 HOA 오디오 데이터의 오더를 지칭한다. 분해 유닛 (518) 은 이러한 SVD 를 수행하는 것을 통해, V 행렬 (519), S 행렬 (519B), 및 U 행렬 (519C) 을 생성할 수도 있으며, 행렬들 (519-519C)("행렬들 (519)") 의 각각은 위에서 더 상세하게 기재된 개별 V, S, 및 U 행렬을 나타낸다. 분해 유닛 (518) 은 이들 행렬들 (519A) 을 음장 성분 추출 유닛 (520) 로 패스하거나 출력할 수도 있다. V 행렬 (519A) 은 사이즈가 $(N+1)^2$-바이-$(N+1)^2$ 일 수도 있고, S 행렬

(519B) 은 사이즈가 $(N+1)^2$-바이-$(N+1)^2$ 일 수도 있으며, U 행렬은 사이즈가 M-바이-$(N+1)^2$ 일 수도 있으며, 여기서 M 은 오디오 프레임에서의 샘플들의 수를 지칭한다. M 에 대한 통상적인 값은 1024 이지만, 본 개시물의 기술들은 이러한 M 에 대한 통상적인 값에 제한되지 않아야 한다.

[0745] 음장 성분 추출 유닛 (520) 은 음장의 구별되는 성분들 및 음장의 배경 성분들을 결정한 다음 추출하여, 음장의 배경 성분들로부터 음장의 구별되는 성분들을 효과적으로 분리하도록 구성되는 유닛을 나타낼 수도 있다. 이러한 관점에서, 음장 성분 유닛 (520) 은 도 4 의 예에 나타낸 오디오 인코딩 유닛 (20) 의 음장 분석 유닛 (44), 배경 선택 유닛 (48), 및 전경 선택 유닛 (36) 에 대해 전술한 많은 동작들을 수행할 수도 있다. 일부 예들에서, 음장의 구별되는 성분들이 이들 성분들의 구별되는 성질을 정확하게 나타내기 위해 상위 오더 (음장의 배경 성분들에 비해) 기저 함수들 (그리고 이에 따른 많은 SHC) 를 필요로 한다고 하면, 배경 성분들로부터 구별되는 성분들을 분리하는 것은 많은 비트들이 구별되는 성분들에 할당되는 것을 가능하게 하고 적은 비트들 (상대적으로, 스피킹) 이 배경 성분들에 할당되는 것을 가능하게 할 수도 있다. 따라서, 이러한 변환 (SVD 의 형태로 또는 PCA 를 포함하는 변환의 다른 형태로) 의 적용을 통해, 본 개시물에 기재된 기술들은 다양한 SHC 로의 비트들의 할당 그리고 이에 의한 SHC (511) 의 압축을 용이하게 할 수도 있다.

[0746] 게다가, 도 40b 에 대해 하기에서 더 상세하게 기재되는 바와 같이, 기술들은 또한, 일부 예들에서, 상위 오더 기저 함수들이, 이러한 성분들의 확산 또는 배경 성질이 주어지면 음장의 이러한 배경 부분들을 나타내기 위해 필요하지 않다고 하면, 음장의 배경 성분들의 오더 감소를 가능하게 할 수도 있다. 이에 따라, 기술들은 음장의 SHC (511) 로의 SVD 의 적용을 통해 음장의 돌출하는 구별되는 성분들 또는 양태들을 보존하면서 음장의 배경 양태들 또는 확산의 압축을 가능하게 할 수도 있다.

[0747] 도 40 의 예에 추가로 나타낸 바와 같이, 음장 성분 추출 유닛 (520) 은 트랜스포즈 유닛 (522), 돌출 성분 분석 유닛 (524) 및 계산 유닛 (526) 을 포함한다. 트랜스포즈 유닛 (522) 은 "V^T 행렬 (523)" 로서 지칭되는, V 행렬 (519) 의 트랜스포즈를 생성하기 위해 V 행렬 (519A) 을 트랜스포즈하도록 구성된 유닛을 나타낸다. 트랜스포즈 유닛 (522) 은 이러한 V^T 행렬 (523) 을 계산 유닛 (526) 에 출력할 수도 있다. V^T 행렬 (523) 은 사이즈가 $(N+1)^2$-바이-$(N+1)^2$ 일 수도 있다.

[0748] 돌출 성분 분석 유닛 (524) 은 S 행렬 (519B) 에 대해 돌출 분석을 수행하도록 구성된 유닛을 나타낸다. 이러한 관점에서, 돌출 성분 분석 유닛 (524) 은 도 4 의 예에 나타낸 오디오 인코딩 디바이스 (20) 의 음장 분석 유닛 (44) 에 대해 전술한 것과 유사한 동작들을 수행할 수도 있다. 돌출 성분 분석 유닛 (524) 은 S 행렬 (519B) 의 대각선 값들을 분석하여, 최대 값을 갖는 이들 성분들의 변수 (D) 를 선택할 수도 있다. 즉, 돌출 성분 분석 유닛 (524) 은 S 의 하향 대각선 값들에 의해 생성된 곡선의 기울기를 분석하는 것에 의해, 2 개의 서브스페이스들 (예를 들어, 전경 또는 우월한 서브스페이스 및 배경 또는 주변 서브스페이스) 를 분리하는, 값 (D) 를 결정할 수도 있으며, 여기서 큰 특이값들은 전경 또는 구별되는 사운드들을 나타내고 낮은 특이값들은 음장의 배경 성분들을 나타낸다. 일부 예들에서, 돌출 성분 분석 유닛 (524) 은 특이값 곡선의 제 1 및 제 2 도함수를 사용할 수도 있다. 돌출 성분 분석 유닛 (524) 은 또한 수 (D) 를 1 과 5 사이로 제한할 수도 있다. 또 다른 예에서, 돌출 성분 분석 유닛 (524) 은 수 (D) 를 1 과 $(N+1)^2$ 로 제한할 수도 있다. 대안으로, 돌출 성분 분석 유닛 (524) 은 수 (D) 를, 예컨대 4 의 값으로 미리 정의 할 수도 있다. 어떤 경우든, 수 (D) 가 추정되면, 돌출 성분 분석 유닛 (524) 은 행렬들 (U, V 및 S) 로부터 전경 및 배경 공간을 추출한다.

[0749] 일부 예들에서, 돌출 성분 분석 유닛 (524) 은 이러한 분석을 M-샘플들 마다 수행할 수도 있으며, 이는 프레임 마다에 기초하는 것으로 다시 언급될 수도 있는 있다. 이러한 관점에서, D 는 프레임에서 프레임으로 변화할 수도 있다. 다른 예들에서, 돌출 성분 분석 유닛 (524) 은 프레임 당 2 번 이상 이러한 분석을 수행하여, 프레임의 2 이상의 부분들을 분석할 수도 있다. 따라서, 기술들은 이러한 관점에서 본 개시물에 기재된 예들에 한정되지 않아야 한다.

[0750] 사실상, 돌출 성분 분석 유닛 (524) 은, 도 40 의 예에서 S 행렬 (519B) 로서 지칭되는, 대각 행렬의 특이값들을 분석하여, 대각 S 행렬 (519B) 의 다른 값들 보다 큰 상대값을 갖는 그러한 값들을 식별할 수도 있다. 돌출 성분 분석 유닛 (524) 은 D 값들을 식별하여, S_{DIST} 행렬 (525A) 및 S_{BG} 행렬 (525B) 를 생성하기 위해 이들 값들을 추출할 수도 있다. S_{DIST} 행렬 (525A) 는 원래 S 행렬 (519B) 의 $(N+1)^2$ 를 갖는 D 컬럼들을 포함하

는 대각 행렬을 나타낼 수도 있다. 일부 경우들에서, S_{BG} 행렬 (525B) 는 각각이 원래 S 행렬 (519B) 의 $(N+1)^2$ 변환된 구면 조화 계수들을 포함하는, $(N+1)^2$-D 컬럼들을 갖는 행렬을 나타낼 수도 있다. 원래 S 행렬 (519B) 의 $(N+1)^2$ 값들을 갖는 D 컬럼들을 포함하는 행렬을 나타내는 S_{DIST} 행렬로서 기재되어 있지만, 돌출 성분 분석 유닛 (524) 은, S 행렬 (519B) 가 대각 행렬이고, 각각의 컬럼에서 D 번째 값 이후의 D 컬럼들의 $(N+1)^2$ 값들이 종종 0 의 값이라고 하면, 원래 S 행렬 (519B) 의 D 값들을 갖는 D 컬럼들을 갖는 S_{DIST} 행렬을 생성하기 위해 이 행렬을 절단 (truncate) 할 수도 있다. 전체 S_{DIST} 행렬 (525A) 및 전체 S_{BG} 행렬 (525B) 에 대해 기재되었지만, 기술들은 이들 S_{DIST} 행렬 (525A) 의 절단된 버전 및 이 S_{BG} 행렬 (525B) 의 절단된 버전에 대해 구현될 수도 있다. 따라서, 본 개시물의 기술들은 이러한 관점에서 제한되지 않아야 한다.

[0751] 다시 말해서, S_{DIST} 행렬 (525A) 는 사이즈가 D-바이-$(N+1)^2$ 일 수도 있는 한편, S_{BG} 행렬 (525B) 는 사이즈가 $(N+1)^2$-D-바이-$(N+1)^2$ 일 수도 있다. S_{DIST} 행렬 (525A) 는 그러한 원리의 성분들, 또는 다시 말해서, 음장의 구별되는 (DIST) 오디오 성분들에 관하여 핵심적이라고 결정되는 특이값들을 포함할 수도 있는 한편, S_{BG} 행렬 (525B) 는 배경 (BG) 또는 다시 말해서, 음장의 주변의 또는 불구별되는 오디오 성분들이라고 결정되는 그러한 특이값들을 포함할 수도 있다. 도 40 의 예에서 별도의 행렬들 (525A 및 525B) 로 나타나 있지만, 행렬들 (525A 및 5225B) 은 S_{DIST} 행렬 (525) 를 나타내는 이러한 단일 행렬의 컬럼들의 수 (좌측에서 우측으로) 를 나타내기 위해 변수 (D) 를 사용하여 단일 행렬로서 특정될 수도 있다. 일부 예들에서, 변수 (D) 는 4 로 설정될 수도 있다.

[0752] 돌출 성분 분석 유닛 (525) 은 또한 U 행렬 (519C) 를 분석하여 U_{DIST} 행렬 (525C) 및 U_{BG} 행렬 (525D) 를 생성할 수도 있다. 종종, 돌출 성분 분석 유닛 (524) 은 S 행렬 (519B) 를 분석하여 변수 (D) 를 식별하여, 변수 D 에 기초한 U_{DIST} 행렬 (525C) 및 U_{BG} 행렬 (525B) 를 생성할 수도 있다. 즉, 돌출하는 S 행렬 (519B) 의 D 컬럼들을 식별한 후, 돌출 성분 분석 유닛 (524) 은 결정된 변수 (D) 에 기초하여 U 행렬 (519C) 를 스플릿할 수도 있다. 이 경우, 돌출 성분 분석 유닛 (524) 은 원래 U 행렬 (519C) 의 $(N+1)^2$ 변환된 구면 조화 계수들의 D 컬럼들 (좌측에서 우측으로) 을 포함하기 위해 U_{DIST} 행렬 (525C) 를 생성할 수도 있고, 원래 U 행렬 (519C) 의 $(N+1)^2$ 변환된 구면 조화 계수들의 나머지 $(N+1)^2$-D 를 포함하기 위해 U_{BG} 행렬 (525D) 를 생성할 수도 있다. U_{DIST} 행렬 (525C) 는 사이즈가 M 바이 D 일 수도 있는 한편, U_{BG} 행렬 (525D) 는 사이즈가 M 바이 $(N+1)^2$-D 일 수도 있다. 도 40 의 예에서 별도의 행렬들 (525C 및 525D) 인 것으로 도시되어 있지만, 행렬들 (525C 및 525D) 은 U_{DIST} 행렬 (525B) 를 나타내는 이러한 단일 행렬의 컬럼들의 수 (좌측에서 우측으로) 를 나타내기 위해 변수 (D) 를 사용하여 단일 행렬로서 특정될 수도 있다.

[0753] 돌출 성분 분석 유닛 (524) 은 또한 V^T 행렬 (523) 을 분석하여 V^T_{DIST} 행렬 (525E) 및 V^T_{BG} 행렬 (525F) 를 생성할 수도 있다. 종종, 돌출 성분 분석 유닛 (524) 은 S 행렬 (519B) 를 분석하여 변수 (D) 를 식별하여, 변수 (D) 에 기초한 V^T_{DIST} 행렬 (525E) 및 V_{BG} 행렬 (525F) 를 생성할 수도 있다. 즉, 돌출하는 S 행렬 (519B) 의 D 컬럼들을 식별한 후, 돌출 성분 분석 유닛 (254) 는 이러한 결정된 변수 D 에 기초하여 V 행렬 (519A) 를 스플릿할 수도 있다. 이러한 경우, 돌출 성분 분석 유닛 (524) 는 원래 V^T 행렬 (523) 의 D 값들의 $(N+1)^2$ 로우들 (상부에서 하부로) 를 포함하기 위해 V^T_{DIST} 행렬 (525E) 를 생성하고 원래 V^T 행렬 (523) 의 $(N+1)^2$-D 의 나머지 $(N+1)^2$ 로우들을 포함하기 위해 V^T_{BG} 행렬 (525F) 를 생성할 수도 있다. V^T_{DIST} 행렬 (525E) 는 사이즈가 $(N+1)^2$-바이-D 일 수도 있는 한편, V^T_{BG} 행렬 (525D) 는 사이즈가 $(N+1)^2$-바이-$(N+1)^2$-D 일 수도 있다. 도 40 의 예에서는 별도의 행렬들 (525E 및 525F) 인 것으로 나타나 있지만, 행렬들 (525E 및 525F) 은 V_{DIST} 행렬 (525E) 를 나타내는 이러한 단일 행렬의 컬럼들의 수 (좌측에서 우측으로) 를 나타내기 위해 변수 (D) 를 사용하여 단일 행렬로서 특정될 수도 있다. V^T_{DIST} 행렬 (525E) 를 비트스트림 생성 유닛 (516) 에 또한 출력하

면서, 돌출 성분 분석 유닛 (524) 은 S_{DIST} 행렬 (525), S_{BG} 행렬 (525B), U_{DIST} 행렬 (525C), U_{BG} 행렬 (525D) 및 V^{T}_{BG} 행렬 (525F) 를 계산 유닛 (526) 에 출력할 수도 있다.

[0754] 계산 유닛 (526) 은 행렬 곱셈들 또는 하나 이상의 행렬들 (또는 벡터들) 에 대해 수행될 수 있는 임의의 다른 수학적 연산을 수행하도록 구성된 유닛을 나타낼 수도 있다. 보다 구체적으로, 도 40 의 예에 나타낸 바와 같이, 계산 유닛 (526) 은 행렬 곱셈을 수행하여 S_{DIST} 행렬 (525A) 로 U_{DIST} 행렬 (525C) 를 곱하여 사이즈가 M 바이 D 인 $U_{DIST} * S_{DIST}$ 벡터들 (527) 을 생성하도록 구성된 유닛을 나타낼 수도 있다. 행렬 계산 유닛 (526) 은 또한 행렬 곱셈을 수행하여 S_{BG} 행렬 (525B) 로 U_{BG} 행렬 (525D) 를 곱하고 그 후 V^{T}_{BG} 행렬 (525F) 로 곱하여 $U_{BG} * S_{BG} * V^{T}_{BG}$ 행렬 (525F) 를 생성하여 사이즈가 M 바이 $(N+1)^{2}$ (음장의 배경 성분들을 나타내는 구면 조화 계수들 (511) 의 것들을 나타낼 수도 있음) 인 배경 구면 조화 계수들 (531) 을 생성할 수도 있다. 계산 유닛 (526) 은 $U_{DIST} * S_{DIST}$ 벡터들 (527) 및 배경 구면 조화 계수들 (531) 을 오디오 인코딩 유닛 (514) 에 출력할 수도 있다.

[0755] 이에 따라, 오디오 인코딩 디바이스 (510) 는, 인코딩 프로세스의 종료 시 행렬 곱셈을 통해 오디오 인코딩 디바이스 (510) 가 $U_{DIST} * S_{DIST}$ 벡터들 (527) 및 배경 구면 조화 계수들 (531) 을 생성하도록 구성된 이러한 계산 유닛 (526) 을 포함한다는 점에서 오디오 인코딩 디바이스 (20) 와 상이하다. 오디오 인코딩 디바이스 (20) 의 선형 가역 변환 유닛 (30) 은 U 및 S 행렬들의 곱셈을 수행하여도 40 의 예에는 나타내지 않은 리오더와 같은 이후의 동작들을 용이하게 할 수도 있는, 인코딩 프로세스의 상대적 시작에서 US[*k*] 벡터들 (33) 을 출력한다. 게다가, 오디오 인코딩 디바이스 (20) 는 인코딩 프로세스의 종료에서 배경 SHC (531) 을 복구하기 보다는, HOA 계수들 (11) 로부터 직접 배경 HOA 계수들 (47) 을 선택함으로써, 행렬 곱셈들을 잠재적으로 회피하여 배경 SHC (531) 을 복구한다.

[0756] 오디오 인코딩 유닛 (514) 은 인코딩의 형태를 수행하여 $U_{DIST} * S_{DIST}$ 벡터들 (527) 및 배경 구면 조화 계수들 (531) 을 추가로 압축하는 유닛을 나타낼 수도 있다. 오디오 인코딩 유닛 (514) 은 도 4 의 예에 나타낸 오디오 인코딩 디바이스 (20) 의 심리 음향 오디오 코더 유닛 (40) 과 실질적으로 유사한 방식으로 동작할 수도 있다. 일부 경우들에서, 이러한 오디오 인코딩 유닛 (514) 은 어드밴스드 오디오 코딩 (AAC) 인코딩 유닛의 하나 이상의 경우들을 나타낼 수도 있다. 오디오 인코딩 유닛 (514) 은 $U_{DIST} * S_{DIST}$ 벡터들 (527) 의 각각의 컬럼 또는 로우를 인코딩할 수도 있다. 종종, 오디오 인코딩 유닛 (514) 은 배경 구면 조화 계수들 (531) 에서 유지하고 있는 오더/서브 오더 조합들의 각각에 대해 AAC 인코딩 유닛의 경우를 작동시킬 수도 있다. 배경 구면 조화 계수들 (531) 에 관한 정보가 AAC 인코딩 유닛을 사용하여 어떻게 인코딩될 수 있는지는, 2008 년 5 월 12 일부터 20 일, 124 회 컨벤션에서 제시된 명칭이 "Encoding Higher Order Ambisonics with AAC" 이고, http://ro.uow.edu.au/cgi/viewcontent.cgi?article=8025&context=engpapers 에서 입수가능한, Eric Hellerud 등에 의한 컨벤션 논문에서 알 수 있다. 오디오 인코딩 유닛 (514) 은 $U_{DIST} * S_{DIST}$ 벡터들 (527) 의 인코딩된 버전 ("인코딩된 $U_{DIST} * S_{DIST}$ 벡터들 (515)" 로 지칭됨) 및 배경 구면 조화 계수들 (531) 의 인코딩된 버전 ("인코딩된 배경 구면 조화 계수들 (515B)" 로 지칭됨) 을 비트스트림 생성 유닛 (516) 에 출력할 수도 있다. 일부 경우들에서, 오디오 인코딩 유닛 (514) 은 $U_{DIST} * S_{DIST}$ 벡터들 (527) 을 인코딩하기 위해 사용되는 것 보다 낮은 타겟 비트레이트를 사용하여 배경 구면 조화 계수들 (531) 을 오디오 인코딩함으로써, $U_{DIST} * S_{DIST}$ 벡터들 (527) 에 비해 많이 배경 구면 조화 계수들 (531) 을 잠재적으로 압축할 수도 있다.

[0757] 비트스트림 생성 유닛 (516) 은 알려진 포맷 (디코딩 디바이스에 의해 알려진 포맷을 지칭할 수도 있음) 에 따르도록 데이터를 포맷함으로써, 비트스트림 (517) 을 생성하는 유닛을 나타낸다. 비트스트림 생성 유닛 (42) 은 도 4 의 예에 나타낸 오디오 인코딩 유닛 (24) 의 비트스트림 생성 유닛 (42) 에 대해 전술한 것과 실질적으로 유사한 방식으로 동작할 수도 있다. 비트스트림 생성 유닛 (516) 은 인코딩된 $U_{DIST} * S_{DIST}$ 벡터들 (515), 인코딩된 배경 구면 조화 계수들 (515B) 및 V^{T}_{DIST} 행렬 (525E) 를 멀티플렉싱하는 멀티플렉서를 포함할 수도 있다.

[0758] 도 40b 는 2 차원 또는 3 차원 음장들을 기술하는 구면 조화 계수들을 압축하기 위해 본 개시물에 기재된 기술들의 다양한 양태들을 수행할 수도 있는 일 예의 오디오 인코딩 디바이스 (510B) 를 예시하는 블록 다이어그램

이다. 오디오 인코딩 디바이스 (510B) 는, 오디오 인코딩 디바이스 (510B) 가 오디오 압축 유닛 (512), 오디오 인코딩 유닛 (514) 및 비트스트림 생성 유닛 (516) 을 포함한다는 점에서 오디오 인코딩 디바이스 (510) 와 유사할 수도 있다. 게다가, 오디오 인코딩 디바이스 (510B) 의 오디오 압축 유닛 (512) 은 오디오 압축 유닛 (512) 이 분해 유닛 (518) 을 포함한다는 점에서 오디오 인코딩 디바이스 (510) 의 것과 유사할 수도 있다. 오디오 인코딩 디바이스 (510B) 의 오디오 압축 유닛 (512) 은, 음장 성분 추출 유닛 (520) 이 오더 감소 유닛 (528A)("오더 감소 유닛 (528)") 으로서 지칭되는 부가 유닛을 포함한다는 점에서, 오디오 인코딩 디바이스 (510) 의 오디오 압축 유닛 (512) 과 상이할 수도 있다. 이러한 이유로, 오디오 인코딩 디바이스 (510B) 의 음장 성분 추출 유닛 (520) 은 "음장 성분 추출 유닛 (520B)" 로서 지칭된다.

[0759] 오더 감소 유닛 (528A) 은 배경 구면 조화 계수들 (531) 의 부가적인 오더 감소를 수행하도록 구성된 유닛을 나타낸다. 일부 경우들에서, 오더 감소 유닛 (528A) 은 음장을 나타내는데 필요한 배경 구면 조화 계수들 (531) 의 수를 감소시키기 위해 배경 구면 조화 계수들 (531) 을 나타내었던 음장을 회전할 수도 있다. 일부 경우들에서, 배경 구면 조화 계수들 (531) 이 음장의 배경 성분들을 나타낸다고 하면, 오더 감소 유닛 (528A)은 고차 구면 기저 함수들에 대응하는 배경 구면 조화 계수들 (531) 의 것들을 (종종 삭감하는 것에 의해) 제거, 소거 또는 그렇지 않으면 삭제할 수도 있다. 이러한 관점에서, 오더 감소 유닛 (528A) 은 도 4 의 예에 나타낸 오디오 인코딩 디바이스 (20) 의 배경 선택 유닛 (48) 과 유사한 동작들을 수행할 수도 있다.

오더 감소 유닛 (528A) 은 배경 구면 조화 계수들 (531) 의 감소된 버전 ("감소된 배경 구면 조화 계수들 (529)" 로서 지칭됨) 을, 감소된 구면 조화 계수들 (529) 를 인코딩하기 위해 전술한 방식으로 오디오 인코딩을 수행하고, 이에 의해 인코딩된 감소된 배경 구면 조화 계수들 (515B) 를 생성할 수도 있는, 오디오 인코딩 유닛 (514) 에 출력할 수도 있다.

[0760] 하기에 열거되는 다양한 조항들은 본 개시물에 기재된 기술들의 다양한 양태들을 제시할 수도 있다.

[0761] 조항 132567-1. 복수의 구면 조화 계수들의 좌-특이 벡터들을 나타내는 U 행렬, 복수의 구면 조화 계수들의 특이값들을 나타내는 S 행렬, 및 복수의 구면 조화 계수들의 우-특이 벡터들을 나타내는 V 행렬을 생성하기 위해 음장을 나타내는 복수의 구면 조화 계수들에 대해 특이값 분해를 수행하고, 그리고 U 행렬, S 행렬, 및 V 행렬의 하나 이상의 적어도 일부의 함수로서 복수의 구면 조화 계수들을 나타내도록 구성된 하나 이상의 프로세서를 포함하는, 오디오 인코딩 디바이스 (510) 또는 오디오 인코딩 디바이스 (510B) 와 같은, 디바이스.

[0762] 조항 132567-2. 하나 이상의 프로세서들이 또한, U 행렬, S 행렬 및 V 행렬의 조합들 또는 그 유도들을 포함하는 U 행렬, S 행렬 및 V 행렬의 하나 이상의 벡터들로서 복수의 구면 조화 계수들의 표현을 포함하기 위해 비트스트림을 생성하도록 구성되는, 조항 132567-1 의 디바이스.

[0763] 조항 132567-3. 하나 이상의 프로세서들이 또한, 복수의 구면 조화 계수들을 나타낼 때, 음장의 구별되는 성분들을 기술하는 U 행렬 내에 포함된 하나 이상의 U_{DIST} 벡터들을 결정하도록 구성되는, 조항 132567-1 의 디바이스.

[0764] 조항 132567-4. 복수의 구면 조화 계수들을 나타낼 때, 음장의 구별되는 성분들을 기술하는 U 행렬 내에 포함된 하나 이상의 U_{DIST} 벡터들을 결정하고, 음장의 구별되는 성분들을 또한 기술하는 S 행렬 내에 포함된 하나 이상의 S_{DIST} 벡터들을 결정하며, 그리고 하나 이상의 U_{DIST} 벡터들 및 하나 이상의 S_{DIST} 벡터들을 곱하여 $U_{DIST} * S_{DIST}$ 벡터들을 생성하도록 구성되는, 조항 132567-1 의 디바이스.

[0765] 조항 132567-5. 하나 이상의 프로세서들은 또한, 복수의 구면 조화 계수들을 나타낼 때, 음장의 구별되는 성분들을 기술하는 U 행렬 내에 포함된 하나 이상의 U_{DIST} 벡터들을 결정하고, 음장의 구별되는 성분들을 또한 기술하는 S 행렬 내에 포함된 하나 이상의 S_{DIST} 벡터들을 결정하며, 그리고 하나 이상의 U_{DIST} 벡터들 및 하나 이상의 S_{DIST} 벡터들을 곱하여 하나 이상의 $U_{DIST} * S_{DIST}$ 벡터들을 생성하도록 구성되고, 하나 이상의 프로세서들이 또한 하나 이상의 $U_{DIST} * S_{DIST}$ 벡터들을 오디오 인코딩하여 하나 이상의 $U_{DIST} * S_{DIST}$ 벡터들의 오디오 인코딩된 버전을 생성하도록 구성되는, 조항 132567-1 의 디바이스.

[0766] 조항 132567-6. 하나 이상의 프로세서들이 또한, 복수의 구면 조화 계수들을 나타낼 때, U 행렬 내에 포함된 하나 이상의 U_{BG} 벡터들을 결정하도록 구성되는, 조항 132567-1 의 디바이스.

[0767] 조항 132567-7. 하나 이상의 프로세서들이 또한, 복수의 구면 조화 계수들을 나타낼 때, S 행렬을 분석하여

음장의 구별되는 그리고 배경 성분들을 식별하도록 구성되는, 조항 132567-1 의 디바이스.

[0768] 조항 132567-8. 하나 이상의 프로세서들이 또한, 복수의 구면 조화 계수들을 나타낼 때, S 행렬을 분석하여 음장의 구별되는 그리고 배경 성분들을 식별하고, 그리고 S 행렬의 분석에 기초하여, 음장의 성분들을 기술하는 U 행렬의 하나 이상의 U_{DIST} 벡터들, 및 음장의 배경 성분들을 기술하는 U 행렬의 하나 이상의 U_{BG} 벡터들을 결정하도록 구성되는, 조항 132567-1 의 디바이스.

[0769] 조항 132567-9. 하나 이상의 프로세서들이 또한, 복수의 구면 조화 계수들을 나타낼 때, S 행렬을 분석하여 오디오 프레임 마다에 기초하여 음장의 구별되는 그리고 배경 성분들을 식별하고, 그리고 S 행렬의 오디오 프레임 마다에 기초하여, 음장의 구별되는 성분들을 기술하는 U 행렬의 하나 이상의 U_{DIST} 벡터들 및 음장의 배경 성분들을 기술하는 U 행렬의 하나 이상의 U_{BG} 벡터들을 결정하도록 구성되는, 조항 132567-1 의 디바이스.

[0770] 조항 132567-10. 하나 이상의 프로세서들이 또한, 복수의 구면 조화 계수들을 나타낼 때, S 행렬을 분석하여 음장의 구별되는 그리고 배경 성분들을 식별하고, S 행렬의 분석에 기초하여, 음장의 구별되는 성분들을 기술하는 U 행렬의 하나 이상의 U_{DIST} 벡터들 및 음장의 배경 성분을 기술하는 U 행렬의 하나 이상의 U_{BG} 벡터들을 결정하도록 구성되고, S 행렬의 분석에 기초하여 S 행렬의 하나 이상의 S_{DIST} 벡터들 및 하나 이상의 S_{BG} 벡터들을 결정하는 것은, 하나 이상의 U_{DIST} 벡터들 및 하나 이상의 U_{BG} 벡터들에 대응하며, 그리고 S 행렬의 분석에 기초하여, 하나 이상의 U_{DIST} 벡터들 및 하나 이상의 U_{BG} 벡터들에 대응하는 V 행렬의 트랜스포즈의 하나 이상의 V^{T}_{BG} 벡터들 및 하나 이상의 V^{T}_{DIST} 벡터들을 결정하도록 구성되는, 조항 132567-1 의 디바이스.

[0771] 조항 132567-11. 하나 이상의 프로세서들이 또한, 복수의 구면 조화 계수들을 나타낼 때, 추가로 하나 이상의 S_{BG} 벡터들로 하나 이상의 U_{BG} 벡터들을 곱한 후 하나 이상의 V^{T}_{BG} 벡터들로 곱하여 하나 이상의 $U_{BG} * S_{BG} * V^{T}_{BG}$ 벡터들을 생성하도록 구성되고, 하나 이상의 프로세서들이 또한, $U_{BG} * S_{BG} * V^{T}_{BG}$ 벡터들을 오디오 인코딩하여 $U_{BG} * S_{BG} * V^{T}_{BG}$ 벡터들의 오디오 인코딩된 버전을 생성하도록 구성되는, 조항 132567-10 의 디바이스.

[0772] 조항 132567-12. 하나 이상의 프로세서들이 또한, 복수의 구면 조화 계수들을 나타낼 때, 하나 이상의 S_{BG} 벡터들로 하나 이상의 U_{BG} 벡터들을 곱한 후 하나 이상의 V^{T}_{BG} 벡터들로 곱하여 하나 이상의 $U_{BG} * S_{BG} * V^{T}_{BG}$ 벡터들을 생성하고, 그리고 오더 감소 프로세스를 수행하여 하나 이상의 구면 조화 기저 함수들과 연관된 하나 이상의 $U_{BG} * S_{BG} * V^{T}_{BG}$ 벡터들의 계수들의 것들을 제거하고 이에 의해 하나 이상의 $U_{BG} * S_{BG} * V^{T}_{BG}$ 벡터들의 오더 감소된 버전을 생성하도록 구성되는, 조항 132567-10 의 디바이스.

[0773] 조항 132567-13. 하나 이상의 프로세서들이 또한, 복수의 구면 조화 계수들을 나타낼 때, 하나 이상의 S_{BG} 벡터들로 하나 이상의 U_{BG} 벡터들을 곱한 후 하나 이상의 V^{T}_{BG} 벡터들로 곱하여 하나 이상의 $U_{BG} * S_{BG} * V^{T}_{BG}$ 벡터들을 생성하고, 그리고 오더 감소 프로세스를 수행하여 구면 조화 기저 함수들의 하나 이상의 오더들과 연관된 하나 이상의 $U_{BG} * S_{BG} * V^{T}_{BG}$ 벡터들의 계수들의 것들을 소거하고 이에 의해 하나 이상의 $U_{BG} * S_{BG} * V^{T}_{BG}$ 벡터들의 오더 감소된 버전을 생성하도록 구성되며, 하나 이상의 프로세서들은 또한, 하나 이상의 $U_{BG} * S_{BG} * V^{T}_{BG}$ 벡터들의 오더 감소된 버전을 오디오 인코딩하여 오더 감소된 하나 이상의 $U_{BG} * S_{BG} * V^{T}_{BG}$ 벡터들의 오디오 인코딩된 버전을 생성하도록 구성되는, 조항 132567-10 의 디바이스.

[0774] 조항 132567-14. 하나 이상의 프로세서들은 또한, 복수의 구면 조화 계수들을 나타낼 때, 하나 이상의 S_{BG} 벡터들로 하나 이상의 U_{BG} 벡터들을 곱한 후 하나 이상의 V^{T}_{BG} 벡터들을 곱하여 하나 이상의 $U_{BG} * S_{BG} * V^{T}_{BG}$ 벡터들을 생성하고, 오더 감소 프로세스를 수행하여 구면 조화 기저 함수들 중 하나 보다 큰 하나 이상의 오더들과 연관된 하나 이상의 $U_{BG} * S_{BG} * V^{T}_{BG}$ 벡터들의 계수의 것들을 소거하고, 이에 의해 하나 이상의 $U_{BG} * S_{BG} * V^{T}_{BG}$ 벡터들의 오더 감소된 버전을 생성하며, 하나 이상의 $U_{BG} * S_{BG} * V^{T}_{BG}$ 벡터들의 오더 감소된 버전을 오디오 인코

딩하여 오더 감소된 하나 이상의 $U_{BG} * S_{BG} * V^{T}_{BG}$ 벡터들의 오디오 인코딩된 버전을 생성하도록 구성되는, 조항 132567-10 의 디바이스.

[0775] 조항 132567-15. 하나 이상의 프로세서들은 또한, 하나 이상의 V^{T}_{DIST} 벡터들을 포함하기 위해 비트스트림을 생성하도록 구성되는, 조항 132567-10 의 디바이스.

[0776] 조항 132567-16. 하나 이상의 프로세서들은 또한, 하나 이상의 V^{T}_{DIST} 벡터들을 오디오 인코딩하지 않으면서 하나 이상의 V^{T}_{DIST} 벡터들을 포함하기 위해 비트스트림을 생성하도록 구성되는, 조항 132567-10 의 디바이스.

[0777] 조항 132567-1F. 음장의 적어도 일부를 나타내는 다중 채널 오디오 데이터에 대해 특이값 분해를 수행하여 다중 채널 오디오 데이터의 좌-특이 벡터들을 나타내는 U 행렬, 다중 채널 오디오 데이터의 특이값들을 나타내는 S 행렬, 및 다중 채널 오디오 데이터의 우-특이 벡터들을 나타내는 V 행렬을 생성하고, 그리고 U 행렬, S 행렬, 및 V 행렬의 하나 이상의 적어도 일부의 함수로서 다중 채널 오디오 데이터를 나타내도록 구성된 하나 이상의 프로세서들을 포함하는, 오디오 인코딩 디바이스 (510 또는 510B) 와 같은 디바이스.

[0778] 조항 132567-2F. 다중 채널 오디오 데이터가 복수의 구면 조화 계수들을 포함하는, 조항 132567-1F 의 디바이스.

[0779] 조항 132567-3F. 하나 이상의 프로세서들이 또한, 조항 132567-2 내지 132567-16 의 임의의 조합에 의해 기재된 바와 같이 수행하도록 구성되는, 조항 132567-2F 의 디바이스.

[0780] 전술한 다양한 조항들의 각각으로부터, 오디오 인코딩 디바이스들 (510A-510J) 중 어느 것이 오디오 인코딩 디바이스 (510A-510J) 가 수행하도록 구성되는 방법을 수행하거나 그렇지 않으면 그 방법의 각각의 단계를 수행하기 위한 수단을 포함할 수도 있다는 것을 이해해야 한다. 일부 경우들에서, 이들 수단은 하나 이상의 프로세서들을 포함할 수도 있다. 일부 경우들에서, 하나 이상의 프로세서들은 비-일시적 컴퓨터 판독가능 저장 매체에 저장된 명령들에 의해 구성된 특수 목적 프로세서를 나타낼 수도 있다. 즉, 인코딩 예들의 세트들의 각각에서 기술들의 다양한 양태들은, 실행될 때, 하나 이상의 프로세서들로 하여금 오디오 인코딩 디바이스 (510A-510J) 가 수행하도록 구성된 방법을 수행하게 하는 명령들이 저장된 비-일시적 컴퓨터 판독가능 저장 매체를 제공할 수도 있다.

[0781] 예를 들어, 조항 132567-17 은, 음장을 나타내는 복수의 구면 조화 계수들에 대해 특이값 분해를 수행하여 복수의 구면 조화 계수들의 좌-특이 벡터들을 나타내는 U 행렬, 복수의 구면 조화 계수들의 특이값들을 나타내는 S 행렬, 및 복수의 구면 조화 계수들의 우-특이 벡터들을 나타내는 V 행렬을 생성하는 단계, 및 U 행렬, S 행렬, 및 V 행렬의 하나 이상의 적어도 일부의 함수로써 복수의 구면 조화 계수들을 나타내는 단계를 포함하는 방법인 상기 조항 132567-1 로부터 도출될 수도 있다.

[0782] 또 다른 예로써, 조항 132567-18 은, 복수의 구면 조화 계수들의 좌-특이 벡터들을 나타내는 U 행렬, 복수의 구면 조화 계수들의 특이값들을 나타내는 S 행렬, 및 복수의 구면 조화 계수들의 우-특이 벡터들을 나타내는 V 행렬을 생성하기 위해 음장을 나타내는 복수의 구면 조화 계수들에 대해 특이값 분해를 수행하는 수단, 및 U 행렬, S 행렬, 및 V 행렬의 하나 이상의 적어도 일부의 함수로서 복수의 구면 조화 계수들을 나타내는 수단을 포함하는, 오디오 인코딩 디바이스 (510B) 와 같은, 디바이스인 상기 조항 132567-1 로부터 도출될 수도 있다.

[0783] 또 다른 예로써, 조항 132567-18 은, 실행될 때, 하나 이상의 프로세서로 하여금 음장을 나타내는 복수의 구면 조화 계수들에 대해 특이값 분해를 수행하여 복수의 구면 조화 계수들의 좌-특이 벡터들을 나타내는 U 행렬, 복수의 구면 조화 계수들의 특이값들을 나타내는 S 행렬, 및 복수의 구면 조화 계수들의 우-특이 벡터들을 나타내는 V 행렬을 생성하게 하고, U 행렬, S 행렬, 및 V 행렬의 하나 이상의 적어도 일부의 함수로서 복수의 구면 조화 계수들을 나타내게 하는, 명령들이 저장되는 비-일시적 컴퓨터 판독가능 저장 매체인, 상기 조항 132567-1 로부터 도출될 수도 있다.

[0784] 마찬가지로 다양한 조항들이 위에서 예시된 바와 같이 도출되는 다양한 디바이스들, 방법들 및 비-일시적 컴퓨터 판독가능 저장 매체에 대해 도출될 수도 있다. 본 개시물 전체에 걸쳐 열거되는 다양한 다른 조항들에 대해서도 동일하게 수행될 수도 있다.

[0785] 도 40c 는 2 차원 또는 3 차원 음장들을 기술하는 구면 조화 계수들을 압축하기 위해 본 개시물에 기재된 기술들의 다양한 양태들을 수행할 수도 있는 예시의 오디오 인코딩 디바이스들 (510C) 을 예시하는 블록 다이어그램

이다. 오디오 인코딩 디바이스 (510C) 는, 오디오 인코딩 디바이스 (510C) 가 오디오 압축 유닛 (512), 오디오 인코딩 유닛 (514), 및 비트스트림 생성 유닛 (516) 을 포함한다는 점에서 오디오 인코딩 디바이스 (510B) 와 유사할 수도 있다. 게다가, 오디오 인코딩 디바이스 (510C) 의 오디오 압축 유닛 (512) 은 오디오 압축 유닛 (512) 이 분해 유닛 (518) 을 포함한다는 점에서 오디오 인코딩 디바이스 (510B) 와 유사할 수도 있다.

[0786] 하지만, 오디오 인코딩 디바이스 (510C) 의 오디오 압축 유닛 (512) 은, 음장 성분 추출 유닛 (520) 이 벡터 리오더 유닛 (532) 로 지칭된, 부가 유닛을 포함한다는 점에서 오디오 인코딩 디바이스 (510B) 의 오디오 압축 유닛 (512) 와 상이할 수도 있다. 이러한 이유로, 오디오 인코딩 디바이스 (510C) 의 음장 성분 추출 유닛 (520) 은 "음장 성분 추출 유닛 (520C)" 으로서 지칭된다.

[0787] 벡터 리오더 유닛 (532) 은 리오더링된 하나 이상의 $U_{DIST} * S_{DIST}$ 벡터들 (533) 을 생성하기 위해 $U_{DIST} * S_{DIST}$ 벡터들 (527) 을 리오더링하도록 구성된 유닛을 나타낼 수도 있다. 이러한 관점에서, 벡터 리오더 유닛 (532) 은 도 4 의 예에 나타낸 오디오 인코딩 디바이스 (20) 의 리오더 유닛 (34) 에 대해 전술한 것과 유사한 방식으로 동작할 수도 있다. 음장 성분 추출 유닛 (520C) 은 $U_{DIST} * S_{DIST}$ 벡터들 (527) 을 리오더링하기 위해 벡터 리오더 유닛 (532) 을 작동시킬 수도 있는데, 이는 $U_{DIST} * S_{DIST}$ 벡터들 (527)(여기서 $U_{DIST} * S_{DIST}$ 벡터들 (527) 의 각각의 벡터는 음장에 존재하는 하나 이상의 구별되는 모노 오디오 오브젝트를 나타낼 수도 있음) 의 오더가 위에서 언급된 이유로 오디오 데이터의 부분들로부터 달라질 수도 있다. 즉, 일부 예들에서, 오디오 압축 유닛 (512) 이 오디오 프레임들 (구면 조화 계수들 (511) 의 M 개의 샘플들을 가질 수도 있으며, 여기서 M 은 일부 예들에서 1024 로 설정된다) 로서 일반적으로 지칭되는 오디오 데이터의 이러한 부분들 상에서 동작한다고 하면, $U_{DIST} * S_{DIST}$벡터들 (527) 이 도출되는 U 행렬 (519C) 에서 나타낸 바와 같이 이들 구별되는 모노 오디오 오브젝트들에 대응하는 벡터들의 부분이 오디오 프레임에서 오디오 프레임으로 달라질 수도 있다.

[0788] 오디오 프레임에서 오디오 프레임으로 이들 $U_{DIST} * S_{DIST}$ 벡터들 (527) 을 리오더링하지 않으면서 이들 $U_{DIST} * S_{DIST}$ 벡터들 (527) 이 직접 오디오 인코딩 유닛 (514) 을 패스하면, 오디오 프레임들을 통해 모노 오디오 오브젝트들이 상관할 때 (서로에 대해 $U_{DIST} * S_{DIST}$ 벡터들 (527) 의 오더에 의해 본 예에서 정의되는, 채널 방식) 더 양호하게 수행하는 레거시 압축 방식들과 같은, 일부 압축 방식들에 대해 달성가능한 압축의 범위를 감소시킬 수도 있다. 게다가, 리오더링되지 않을 때, $U_{DIST} * S_{DIST}$ 벡터들 (527) 의 인코딩은 복구될 때 오디오 데이터의 품질을 감소시킬 수도 있다. 예를 들어, 오디오 인코딩 유닛 (514) 에 의해 도 40c 의 예에 나타낼 수도 있는, AAC 인코더들은, 프레임에서 프레임으로 $U_{DIST} * S_{DIST}$ 벡터들 (527) 을 직접 인코딩할 때 달성되는 압축에 비해 프레임에서 프레임으로 리오더링된 하나 이상의 $U_{DIST} * S_{DIST}$ 벡터들 (527) 을 보다 효율적으로 압축할 수도 있다. 위에서는 AAC 인코더들에 대해 기재하였지만, 기술들은 특정 오더 또는 포지션 (채널 방식) 에서 모노 오디오 오브젝트들이 특정될 때 더 양호한 압축을 제공하는 임의의 인코더에 대해 수행될 수도 있다.

[0789] 하기에서 더 상세하게 기재되는 바와 같이, 기술들은 오디오 인코딩 디바이스 (510C) 가 하나 이상의 벡터들 (즉, $U_{DIST} * S_{DIST}$ 벡터들 (527)) 를 리오더링하는 것을 가능하게 하여 리오더링된 하나 이상의 벡터들 $U_{DIST} * S_{DIST}$ 벡터들 (527) 를 생성하고 이에 의해 오디오 인코딩 유닛 (514) 과 같은 레거시 오디오 인코더에 의해 $U_{DIST} * S_{DIST}$ 벡터들 (527) 의 압축을 용이하게 할 수도 있다. 오디오 인코딩 디바이스 (510C) 는 또한, 오디오 인코딩 유닛 (514) 을 사용하여 리오더링된 하나 이상의 $U_{DIST} * S_{DIST}$ 벡터들 (527) 을 오디오 인코딩하여 리오더링된 하나 이상의 $U_{DIST} * S_{DIST}$ 벡터들 (527) 의 인코딩된 버전 (515A) 을 생성하기 위해 본 개시물에 기재된 기술들을 수행할 수도 있다.

[0790] 예를 들어, 음장 성분 추출 유닛 (520C) 은 벡터 리오더 유닛 (532) 을 작동시켜 시간에서 후속하는 제 1 오디오 프레임에서 하나 이상의 제 2 $U_{DIST} * S_{DIST}$ 벡터들 (527) 이 대응하는 제 2 프레임으로 하나 이상의 제 1 벡터 리오더 유닛 (532) 을 리오더링할 수도 있다. 제 2 오디오 프레임에 대해 시간에서 후속인 제 1 오디오 프레임의 컨텍스트에서 기재하였지만, 제 1 오디오 프레임이 제 2 오디오 프레임을 선행할 수도 있다. 따라서, 기술들은 본 개시물에 기재된 예에 한정되지 않아야 한다.

[0791] 벡터 리오더 유닛 (532) 은 먼저 제 1 $U_{DIST} * S_{DIST}$ 벡터들 (527) 및 제 2 $U_{DIST} * S_{DIST}$ 벡터들 (527) 의 각각에 대해 에너지 분석을 수행하여, 제 1 오디오 프레임의 적어도 일부 (하지만 종종 전체) 에 대해 실효값 (root mean squared) 에너지를 계산하고, 이에 의해 (D 가 4 인 것으로 가정하여) 8 개의 에너지들, 제 1 오디오 프레

임의 제 1 U_{DIST} * S_{DIST} 벡터들 (527) 의 각각에 대한 하나 및 제 2 오디오 프레임의 제 2 U_{DIST} * S_{DIST} 벡터들 (527) 의 각각에 대한 하나를 생성할 수도 있다. 벡터 리오더 유닛 (532) 은 그 후 표 1 내지 표 4 에 대해 위에서 기재한 바와 같이 제 1 U_{DIST} * S_{DIST} 벡터들 (527) 로부터의 각각의 에너지를 제 2 U_{DIST} * S_{DIST} 벡터들 (527) 의 각각에 대해 턴방식으로 제 1 U_{DIST} * S_{DIST} 벡터들 (527) 을 비교할 수도 있다.

[0792] 즉, HoA 신호들 상의 프레임 기반 SVD (또는 KLT & PCA 와 같은 관련 방법들) 분해를 사용할 때, 프레임에서 프레임으로의 벡터들의 리오더링이 일치되도록 보장되지 않을 수도 있다. 예를 들어, 하부에 놓인 음장에 2 개의 오브젝트들이 있다면, 분해 (적절하게 수행될 때가 "이상적인 분해" 로서 지칭될 수도 있음) 는 U 행렬에서 하나의 벡터가 하나의 오브젝트를 나타내게 되도록 2 개의 오브젝트들의 분리를 초래할 수도 있다. 하지만, 분해가 "이상적인 분해" 로서 지칭될 수도 있을 때에도, 벡터들은 프레임에서 프레임으로 U 행렬에서의 (그리고 대응하여 S 및 V 행렬에서의) 포지션에서 변경할 수도 있다. 또한, 위상차가 상당히 있을 수도 있으며, 여기서 벡터 리오더 유닛 (532) 은 (마이너스 또는 네거티브 1 로 반전된 벡터의 각각의 엘리먼트를 벡터곱 (dot multiplying) 하는 것에 의해) 반전된 벡터의 위상 반전을 사용하여 위상을 반전할 수도 있다. 이들 벡터들을 피드하기 위해서, 동일한 "AAC/오디오 코딩 엔진" 으로의 프레임 마다, 식별되는 오더 (또는, 즉 매칭되는 신호들), 정류되는 위상, 및 적용되는 프레임 경계들에서의 세심한 보간을 필요로 할 수도 있다. 이러한 것 없이도, 하부에 놓인 오디오 코덱은 '시간적 스미어링 (temporal smearing)' 또는 '프리-에코' 로서 알려진 것들을 포함하는 극도의 하시 아티팩트들 (harsh artifacts) 을 재생할 수도 있다.

[0793] 본 개시물에 기재된 기술들의 다양한 양태들에 따라, 오디오 인코딩 디바이스 (510C) 는 벡터들의 프레임 경계들에서 에너지 및 상호 상관을 사용하여, 벡터들을 식별/매칭하기 위해 다중 방법론들을 적용할 수도 있다. 오디오 인코딩 디바이스 (510C) 는 또한 - 프레임 경계들에서 종종 발생하는 - 180 도의 위상 변화가 보정되는 것을 보장할 수도 있다. 벡터 리오더 유닛 (532) 은 프레임들 사이의 평활한 트랜지션을 보장하기 위해 벡터들 사이에서 페이드-인/페이드-아웃 보간 윈도우의 형태를 적용한다.

[0794] 이러한 식으로, 오디오 인코딩 디바이스 (530C) 는 하나 이상의 벡터들을 리오더링하여 리오더링된 하나 이상의 제 1 벡터들을 생성하고 이에 의해 레거시 오디오 인코더에 의해 인코딩을 용이하게 할 수도 있으며, 하나 이상의 벡터들은 음장의 구별되는 성분들을 기술하고 나타내며, 레거시 오디오 인코더를 사용하여 리오더링된 하나 이상의 벡터들을 오디오 인코딩하여 리오더링된 하나 이상의 벡터들의 인코딩된 버전을 생성한다.

[0795] 본 개시물에 기재된 기술들의 다양한 양태들은 오디오 인코딩 디바이스 (510C) 가 다음의 조항들에 따라 동작하는 것을 가능하게 할 수도 있다.

[0796] 조항 133143-1A. 하나 이상의 제 1 벡터들과 하나 이상의 제 2 벡터들 사이의 에너지 비교를 수행하여 리오더링된 하나 이상의 제 1 벡터들을 결정하고 하나 이상의 제 1 벡터들 및 하나 이상의 제 2 벡터들 중 하나 또는 양자의 추출을 용이하게 하도록 구성된 하나 이상의 프로세서들을 포함하고, 하나 이상의 벡터들은 오디오 데이터의 제 1 부분에서 음장의 구별되는 성분들을 기술하고, 하나 이상의 제 2 벡터들은 오디오 데이터의 제 2 부분에서 음장의 구별되는 성분들을 기술하는, 오디오 인코딩 디바이스 (510C) 와 같은 디바이스.

[0797] 조항 133143-2A. 하나 이상의 제 1 벡터들이 오디오 데이터의 제 1 부분에서 음장의 배경 성분들을 나타내지 않고, 하나 이상의 제 2 벡터들이 오디오 데이터의 제 2 부분에서 음장의 배경 성분들을 나타내지 않는, 조항 133143-1A 의 디바이스.

[0798] 조항 133143-3A. 하나 이상의 프로세서들은 또한, 에너지 비교를 수행한 후, 하나 이상의 제 1 벡터들과 하나 이상의 제 2 벡터들 사이의 상호 상관을 수행하여 하나 이상의 제 2 벡터들과 상관되는 하나 이상의 제 1 벡터들을 식별하도록 구성되는, 조항 133143-1A 의 디바이스.

[0799] 조항 133143-4A. 하나 이상의 프로세서들은 또한, 에너지 비교에 기초하여 제 2 벡터들의 하나 이상을 폐기하여 하나 이상의 제 2 벡터들보다 적은 벡터들을 갖는 감소된 하나 이상의 제 2 벡터들을 생성하고, 하나 이상의 제 1 벡터들 중 적어도 하나와 감소된 하나 이상의 제 2 벡터들 사이의 상호 상관을 수행하여 하나 이상의 제 1 벡터들 중 적어도 하나와 상관하는 감소된 하나 이상의 제 2 벡터들 중 하나를 식별하며, 그리고 상호 상관에 기초하여 하나 이상의 제 1 벡터들 중 적어도 하나를 리오더링하여 리오더링된 하나 이상의 제 1 벡터들을 생성하도록 구성되는, 조항 133143-1A 의 디바이스.

[0800] 조항 133143-5A. 하나 이상의 프로세서들은 또한, 에너지 비교에 기초하여 제 2 벡터들의 하나 이상을 폐기하여 하나 이상의 제 2 벡터들 보다 적은 벡터들을 갖는 감소된 하나 이상의 제 2 벡터들을 생성하고, 하나 이

상의 제 1 벡터들 중 적어도 하나와 감소된 하나 이상의 제 2 벡터들 사이의 상호 상관을 수행하여 하나 이상의 제 1 벡터들의 적어도 하나와 상관하는 감소된 하나 이상의 제 2 벡터들 중 하나를 식별하고, 상호 상관에 기초하여 하나 이상의 제 1 벡터들의 적어도 하나를 리오더링하여 리오더링된 하나 이상의 제 1 벡터들을 생성하며, 그리고 리오더링된 하나 이상의 제 1 벡터들을 인코딩하여 리오더링된 하나 이상의 제 1 벡터들의 오디오 인코딩된 버전을 생성하도록 구성되는, 조항 133143-1A 의 디바이스.

[0801] 조항 133143-6A. 하나 이상의 프로세서들은 또한, 에너지 비교에 기초하여 제 2 벡터들의 하나 이상을 폐기하여 하나 이상의 제 2 벡터들보다 적은 감소된 하나 이상의 제 2 벡터들을 생성하고, 하나 이상의 제 1 벡터들 중 적어도 하나와 감소된 하나 이상이 제 2 벡터들 사이의 상호 상관을 수행하여 하나 이상의 제 1 벡터들 중 적어도 하나와 상관하는 감소된 하나 이상의 제 2 벡터들 중 하나를 식별하고, 상호 상관에 기초하여 하나 이상의 제 1 벡터들 중 적어도 하나를 리오더링하여 리오더링된 하나 이상의 제 1 벡터들을 생성하고, 리오더링된 하나 이상의 제 1 벡터들을 인코딩하여 리오더링된 하나 이상의 제 1 벡터들의 오디오 인코딩된 버전을 생성하며, 그리고 리오더링된 하나 이상의 제 1 벡터들의 인코딩된 버전을 포함하기 위해 비트스트림을 생성하도록 구성되는, 조항 133143-1A 의 디바이스.

[0802] 조항 133143-7A. 오디오 데이터의 제 1 부분은 M 개의 샘플들을 갖는 제 1 오디오 프레임을 포함하고, 오디오 데이터의 제 2 부분은 M 개의 동일한 수의 샘플들을 갖는 제 2 오디오를 포함하고, 하나 이상의 프로세서들은 또한, 상호 상관을 수행할 때, 하나 이상의 제 1 벡터들 중 적어도 하나의 마지막 M-Z 값들 및 감소된 하나 이상의 제 2 벡터들의 각각의 처음 M-Z 값들에 대해 상호 상관을 수행하여 하나 이상의 제 1 벡터들 중 적어도 하나와 상관하는 감소된 하나 이상의 제 2 벡터들 중 하나를 식별하도록 구성되며, 여기서 Z 는 M 보다 작은, 청구항들 3A-6A 의 디바이스.

[0803] 조항 133143-8A. 오디오 데이터의 제 1 부분은 M 개의 샘플들을 갖는 제 1 오디오 프레임을 포함하고, 오디오 데이터의 제 2 부분은 M 개의 동일한 수의 샘플들을 갖는 제 2 오디오 프레임을 포함하고, 하나 이상의 프로세서들은 또한, 상호 상관을 수행할 때, 하나 이상의 제 1 벡터들 중 적어도 하나의 마지막 M-Y 값들 및 감소된 하나 이상의 제 2 벡터들의 각각의 처음 M-Z 값들에 대해 상호 상관을 수행하여 하나 이상의 제 1 벡터들 중 적어도 하나와 상관하는 감소된 하나 이상의 제 2 벡터들 중 하나를 식별하도록 구성되고, 여기서 Z 및 Y 는 모두 M 보다 작은, 청구항들 3A-6A 의 디바이스.

[0804] 조항 133143-9A. 하나 이상의 프로세서들은 또한, 상호 상관을 수행할 때, 하나 이상의 제 1 벡터들 중 적어도 하나 및 하나 이상의 제 1 벡터들을 반전하도록 구성되는, 청구항 3A-6A 의 디바이스.

[0805] 조항 133143-10A. 하나 이상의 프로세서들은 또한, 음장을 나타내는 복수의 구면 조화 계수들에 대해 특이값 분해를 수행하여 하나 이상의 제 2 벡터들 및 하나 이상의 제 2 벡터들을 생성하도록 구성되는, 조항 133143-1A 의 디바이스.

[0806] 조항 133143-11A. 하나 이상의 프로세서들은 또한, 음장을 나타내는 복수의 구면 조화 계수들에 대해 특이값 분해를 수행하여 복수의 구면 조화 계수들의 좌-특이 벡터들을 나타내는 U 행렬, 복수의 구면 조화 계수들이 특이값들을 나타내는 S 행렬, 및 복수의 구면 조화 계수들의 우-특이 벡터들을 나타내는 V 행렬을 생성하고, U 행렬, S 행렬, 및 V 행렬의 하나 이상의 함수로서 하나 이상의 제 1 벡터들 및 하나 이상의 제 2 벡터들을 생성하도록 구성되는, 조항 133143-1A 의 디바이스.

[0807] 조항 133143-12A. 하나 이상의 프로세서들은 또한, 음장을 나타내는 복수의 구면 조화 계수들에 대해 특이값 분해를 수행하여 복수의 구면 조화 계수들의 좌-특이 벡터들을 나타내는 U 행렬, 복수의 구면 조화 계수들의 특이값들을 나타내는 S 행렬, 및 복수의 구면 조화 계수들의 우-특이 벡터들을 나타내는 V 행렬을 생성하고, S 행렬에 대해 돌출 분석을 수행하여 U 행렬의 하나 이상의 U_{DIST} 벡터들 및 S 행렬의 하나 이상의 S_{DIST} 벡터들을 식별하고, 그리고 하나 이상의 S_{DIST} 벡터들로 하나 이상의 U_{DIST} 벡터들을 적어도 부분적으로 곱하는 것에 의해 하나 이상의 제 1 벡터들 및 하나 이상의 제 2 벡터들을 결정하도록 구성되는, 조항 133143-1A 의 디바이스.

[0808] 조항 133143-13A. 오디오 데이터의 제 1 부분은 시간에 있어서 오디오 데이터의 제 2 부분 전에 발생하는, 조항 133143-1A 의 디바이스.

[0809] 조항 133143-14A. 오디오 데이터의 제 1 부분은 시간에 있어서 오디오 데이터의 제 2 부분 후에 발생하는, 조항 133143-1A 의 디바이스.

[0810] 조항 133143-15A. 하나 이상의 프로세서들은 또한, 에너지 비교를 수행할 때, 하나 이상의 제 1 벡터들 및

하나 이상의 제 2 벡터들의 각각에 대해 실효값 에너지를 계산하고, 그리고 하나 이상의 제 2 벡터들의 각각에 대해 계산된 실효값 에너지에 하나 이상의 제 1 벡터들 중 적어도 하나에 대해 계산된 실효값 에너지를 비교하도록 구성되는, 조항 133143-1A 의 디바이스.

[0811] 조항 133143-16A. 하나 이상의 프로세서들은 또한, 에너지 비교에 기초하여 하나 이상의 제 1 벡터들 중 적어도 하나를 리오더링하여 리오더링된 하나 이상의 제 1 벡터들을 생성하도록 구성되고, 하나 이상의 프로세서들이 또한, 제 1 벡터들을 리오더링할 때, 하나 이상의 제 1 벡터들 사이에서 페이드-인/페이드-아웃 보간 윈도우를 적용하여 리오더링된 하나 이상의 벡터들을 생성할 때 평활한 트랜지션을 보장하도록 구성되는, 조항 133143-1A 의 디바이스.

[0812] 조항 133143-17A. 하나 이상의 프로세서들은 또한, 적어도 에너지 비교에 기초하여 하나 이상의 제 1 벡터들을 리오더링하여 리오더링된 하나 이상의 제 1 벡터들을 생성하고, 리오더링된 하나 이상의 제 1 벡터들 및 리오더링된 하나 이상의 제 1 벡터들의 인코딩된 버전을 포함하도록 비트스트림을 생성하며, 그리고 하나 이상의 제 1 벡터들이 어떻게 리오더링되었는지를 기술하는 리오더 정보를 비트스트림에서 명시하도록 구성되는, 조항 133143-1A 의 디바이스.

[0813] 조항 133143-18A. 에너지 비교가 하나 이상의 제 1 벡터들 및 하나 이상의 제 2 벡터들 중 하나 또는 양자의 오디오 인코딩을 촉진하기 위해서 하나 이상의 제 1 벡터들 및 하나 이상의 제 2 벡터들 중 하나 또는 양자의 추출을 용이하게 하는, 조항 133143-1A 의 디바이스.

[0814] 조항 133143-1B. 하나 이상의 제 1 벡터들 및 하나 이상의 제 2 벡터들에 대해 상호 상관을 수행하여 리오더링된 하나 이상의 제 1 벡터들을 결정하고 하나 이상의 제 1 벡터들 및 하나 이상의 제 2 벡터들 중 하나 또는 양자의 추출을 용이하게 하도록 구성된 하나 이상의 프로세서를 포함하고, 하나 이상의 제 1 벡터들은 오디오 데이터의 제 1 부분에서 음장의 구별되는 성분들을 기술하고, 하나 이상의 제 2 벡터들은 오디오 데이터의 제 2 부분에서 음장의 구별되는 성분들을 기술하는, 오디오 인코딩 디바이스 (510C) 와 같은 디바이스.

[0815] 조항 133143-2B. 하나 이상의 제 1 벡터들은 오디오 데이터의 제 1 부분에서 음장의 배경 성분을 나타내지 않고, 하나 이상의 제 2 벡터들은 오디오 데이터의 제 2 부분에서 음장의 배경 성분들을 나타내지 않는, 조항 133143-1B 의 디바이스.

[0816] 조항 133143-3B. 하나 이상의 프로세서들은 또한, 상호 상관을 수행하기 전에, 하나 이상의 제 1 벡터들과 하나 이상의 제 2 벡터들 사이에서 에너지 비교를 수행하여 하나 이상의 제 2 벡터들보다 적은 벡터들을 갖는 감소된 하나 이상의 제 2 벡터들을 생성하도록 구성되고, 하나 이상이 프로세서들은 또한, 상호 상관을 수행할 때, 하나 이상의 제 1 벡터들과 감소된 하나 이상의 제 2 벡터들 사이에서 상호 상관을 수행하여 하나 이상의 제 1 벡터들 및 하나 이상의 제 2 벡터들 중 하나 또는 양자의 오디오 인코딩을 용이하게 하도록 구성되는, 조항 133143-1B 의 디바이스.

[0817] 조항 133143-4B. 하나 이상의 프로세서들은 또한, 에너지 비교를 수행할 때, 하나 이상의 제 1 벡터들 및 하나 이상의 제 2 벡터들의 각각에 대해 실효값 에너지를 계산하고, 하나 이상의 제 2 벡터들의 각각에 대해 계산된 실효값 에너지에 하나 이상의 제 1 벡터들 중 적어도 하나에 대해 계산된 실효값 에너지를 비교하도록 구성되는, 조항 133143-3B 의 디바이스.

[0818] 조항 133143-5B. 하나 이상의 프로세서들은 또한, 에너지 비교에 기초하여 제 2 벡터들 중 하나 이상을 폐기하여 하나 이상의 제 2 벡터들 보다 적은 벡터들을 갖는 감소된 하나 이상의 제 2 벡터들을 생성하도록 구성되고, 하나 이상의 프로세서들은 또한, 상호 상관을 수행할 때, 하나 이상의 제 1 벡터들 중 적어도 하나와 감소된 하나 이상의 제 2 벡터들 사이에서 상호 상관을 수행하여 하나 이상의 제 1 벡터들 중 적어도 하나와 상관하는 감소된 하나 이상의 벡터들 중 하나를 식별하도록 구성되며, 하나 이상의 프로세서들은 또한, 상호 상관에 기초하여 하나 이상의 제 1 벡터들 중 적어도 하나를 리오더링하여 리오더링된 하나 이상의 제 1 벡터들을 생성하도록 구성되는, 조항 133143-3B 의 디바이스.

[0819] 조항 133143-6B. 하나 이상의 프로세서들은 또한, 에너지 비교에 기초하여 제 2 벡터들 중 하나 이상을 폐기하여 하나 이상의 제 2 벡터들 보다 적은 벡터들을 갖는 감소된 하나 이상이 제 2 벡터들을 생성하도록 구성되고, 하나 이상의 프로세서들은 또한, 상호 상관을 수행할 때, 하나 이상의 제 1 벡터들 중 적어도 하나와 감소된 하나 이상의 제 2 벡터들 사이에서 상호 상관을 수행하여 하나 이상의 제 1 벡터들 중 적어도 하나와 상관하는 감소된 하나 이상의 제 2 벡터들 중 하나를 식별하도록 구성되며, 하나 이상의 프로세서들은 또한, 상호 상관에 기초하여 하나 이상의 벡터들 중 적어도 하나를 리오더링하여 리오더링된 하나 이상의 제 1 벡터들을 생

성하고, 그리고 리오더링된 하나 이상의 제 1 벡터들을 인코딩하여 리오더링된 하나 이상의 제 1 벡터들의 오디오 인코딩된 버전을 생성하도록 구성되는, 조항 133143-3B 의 디바이스.

[0820] 조항 133143-7B. 하나 이상의 프로세서들은 또한, 에너지 비교에 기초하여 제 2 벡터들 중 하나 이상을 폐기하여 하나 이상의 제 2 벡터들보다 적은 벡터들을 갖는 감소된 하나 이상의 제 2 벡터들을 생성하도록 구성되고, 하나 이상의 프로세서들은 또한, 상호 상관을 수행할 때, 하나 이상의 제 1 벡터들 중 적어도 하나와 감소된 하나 이상의 제 2 벡터들 사이의 상호 상관을 수행하여 하나 이상의 제 1 벡터들 중 적어도 하나와 상관하는 감소된 하나 이상의 제 2 벡터들 중 하나를 식별하도록 구성되며, 하나 이상의 프로세서들은 또한, 상호 상관에 기초하여 하나 이상의 제 1 벡터들 중 적어도 하나를 리오더링하여 리오더링된 하나 이상의 제 1 벡터들을 생성하고, 리오더링된 하나 이상의 제 1 벡터들을 인코딩하여 리오더링된 하나 이상의 제 1 벡터들의 오디오 인코딩된 버전을 생성하며, 그리고 리오더링된 하나 이상의 제 1 벡터들의 인코딩된 버전을 포함하도록 비트스트림을 생성하도록 구성되는, 조항 133143-3B 의 디바이스.

[0821] 조항 133143-8B. 오디오 데이터의 제 1 부분은 M 개의 샘플들을 갖는 제 1 오디오 프레임을 포함하고, 오디오 데이터의 제 2 부분은 M 개의 동일한 수의 샘플들을 갖는 제 2 오디오 프레임을 포함하고, 하나 이상의 프로세서들은 또한, 상호 상관을 수행할 때, 하나 이상의 제 1 벡터들 중 적어도 하나의 마지막 M-Z 값들 및 감소된 하나 이상의 제 2 벡터들 중 처음 M-Z 값들에 대해 상호 상관을 수행하여 하나 이상의 제 1 벡터들 중 적어도 하나와 상관하는 감소된 하나 이상의 제 2 벡터들 중 하나를 식별하도록 구성되며, 여기서 Z 는 M 보다 작은, 청구항들 3B-7B 의 디바이스.

[0822] 조항 133143-9B. 오디오 데이터의 제 1 부분은 M 개의 샘플들을 갖는 제 1 오디오 프레임을 포함하고, 오디오 데이터의 제 2 부분은 M 개의 동일한 수의 샘플들을 갖는 제 2 오디오 프레임을 포함하고, 하나 이상의 프로세서들은 또한, 상호 상관을 수행할 때, 하나 이상의 제 1 벡터들 중 적어도 하나의 마지막 M-Y 값들 및 감소된 하나 이상의 제 2 벡터들의 각각의 처음 M-Z 값들에 대해 상호 상관을 수행하여 하나 이상의 제 1 벡터들 중 적어도 하나와 상관하는 감소된 하나 이상의 제 2 벡터들 중 하나를 식별하도록 구성되며, 여기서 Z 및 Y 는 모두 M 보다 작은, 청구항들 3B-7B 의 디바이스.

[0823] 조항 133143-10B. 하나 이상의 프로세서들은 또한, 상호 상관을 수행할 때, 하나 이상의 제 1 벡터들 및 하나 이상의 제 2 벡터들 중 적어도 하나를 반전하도록 구성되는, 청구항 1B 의 디바이스.

[0824] 조항 133143-11B. 하나 이상의 프로세서들은 또한, 음장을 나타내는 복수의 구면 조화 계수들에 대해 특이값 분해를 수행하여 하나 이상의 제 1 벡터들 및 하나 이상의 제 2 벡터들을 생성하도록 구성되는, 조항 133143-1B 의 디바이스.

[0825] 조항 133143-12B. 하나 이상의 프로세서들은 또한 음장을 나타내는 복수의 구면 조화 계수들에 대해 특이값 분해를 수행하여 복수의 구면 조화 계수들의 좌-특이 벡터들을 나타내는 U 행렬, 복수의 구면 조화 계수들의 특이값들을 나타내는 S 행렬, 및 복수의 구면 조화 계수들의 우-특이 벡터들을 나타내는 V 행렬을 생성하고, 그리고 U 행렬, S 행렬, 및 V 행렬의 하나 이상의 함수로서 하나 이상의 제 1 벡터들 및 하나 이상의 제 2 벡터들을 생성하도록 구성되는, 조항 133143-1B 의 디바이스.

[0826] 조항 133143-13B. 하나 이상의 프로세서들은 또한, 음장을 나타내는 복수의 구면 조화 계수들에 대해 특이값 분해를 수행하여 복수의 구면 조화 계수들의 좌-특이 벡터들을 나타내는 U 행렬, 복수의 구면 조화 계수들의 특이값들을 나타내는 S 행렬, 및 복수의 구면 조화 계수들의 우-특이 벡터들을 나타내는 V 행렬을 생성하고, S 행렬에 대해 핵심 분석을 수행하여 U 행렬의 하나 이상의 U_{DIST} 벡터들 및 S 행렬의 하나 이상의 S_{DIST} 벡터들을 식별하며, 그리고 하나 이상의 S_{DSIST} 벡터들로 하나 이상의 U_{DIST} 벡터들을 적어도 부분적으로 곱하는 것에 의해 하나 이상의 제 1 벡터들 및 하나 이상의 제 2 벡터들을 결정하도록 구성되는, 조항 133143-1B 의 디바이스.

[0827] 조항 133143-14B. 하나 이상의 프로세서들은 또한, 음장을 나타내는 복수의 구면 조화 계수들에 대해 특이값 분해를 수행하여 복수의 구면 조화 계수들의 좌-특이 벡터들을 나타내는 U 행렬, 복수의 구면 조화 계수들의 특이값들을 나타내는 S 행렬, 및 복수의 구면 조화 계수들의 우-특이 벡터들을 나타내는 V 행렬을 생성하고, 그리고 하나 이상의 제 1 벡터들 및 하나 이상의 제 2 벡터들을 결정할 때, S 행렬에 대해 핵심 분석을 수행하여 하나 이상의 제 1 벡터들 및 하나 이상의 제 2 벡터들 중 적어도 하나로서 V 행렬의 하나 이상의 V_{DIST} 벡터들을 식별하도록 구성되는, 조항 133143-1B 의 디바이스.

[0828] 조항 133143-15B. 오디오 데이터의 제 1 부분은 시간에 있어서 오디오 데이터의 제 2 부분 전에 발생하는,

조항 133143-1B 의 디바이스.

[0829] 조항 133143-16B. 오디오 데이터의 제 1 부분은 시간에 있어서 오디오 데이터의 제 2 부분 후에 발생하는, 조항 133143-1B 의 디바이스.

[0830] 조항 133143-17B. 하나 이상의 프로세서들은 또한, 상호 상관에 기초하여 하나 이상의 제 1 벡터들 중 적어도 하나를 리오더링하여 리오더링된 하나 이상의 제 1 벡터들을 생성하고, 그리고 제 1 벡터들을 리오더링할 때, 하나 이상의 제 1 벡터들 사이에 페이드-인/페이드-아웃 보간 윈도우를 적용하여 리오더링된 하나 이상의 제 1 벡터들을 생성할 때 평활한 트랜지션을 보장하도록 구성되는, 조항 133143-1B 의 디바이스.

[0831] 조항 133143-18B. 하나 이상의 프로세서들은 또한, 상호 상관에 기초하여 하나 이상의 제 1 벡터들을 리오더링하여 리오더링된 하나 이상의 제 1 벡터들을 생성하고, 리오더링된 하나 이상의 제 1 벡터들 또는 리오더링된 하나 이상의 제 1 벡터들의 인코딩된 버전을 포함하도록 비트스트림을 생성하며, 그리고 하나 이상의 제 1 벡터들이 어떻게 리오더링되었는지를 비트스트림에 명시하도록 구성되는, 조항 133143-1B 의 디바이스.

[0832] 조항 133143-19B. 하나 이상의 제 1 벡터들 및 하나 이상의 제 2 벡터들 중 하나 또는 양자의 오디오 인코딩을 촉진하기 위해서, 상호 상관이 하나 이상의 제 1 벡터들 및 하나 이상의 제 2 벡터들 중 하나 또는 양자의 추출을 용이하게 하는, 조항 133143-1B 의 디바이스.

[0833] 도 40d 는 2 차원 또는 3 차원 음장들을 기술하는 구면 조화 계수들을 압축하기 위해 본 개시물에 기재된 기술들의 다양한 양태들을 수행할 수도 있는 일 예의 오디오 인코딩 디바이스 (510D) 를 예시하는 블록 다이어그램이다. 오디오 인코딩 디바이스 (510D) 는, 오디오 인코딩 디바이스 (510D) 가 오디오 압축 유닛 (512), 오디오 인코딩 유닛 (514), 및 비트스트림 생성 유닛 (516) 을 포함한다는 점에서 오디오 인코딩 디바이스 (510C) 와 유사할 수도 있다. 게다가, 오디오 인코딩 디바이스 (510D) 의 오디오 압축 유닛 (512) 은, 오디오 압축 유닛 (512) 이 분해 유닛 (518) 을 포함한다는 점에서 오디오 인코딩 디바이스 (510C) 와 유사할 수도 있다.

[0834] 하지만, 오디오 인코딩 디바이스 (510D) 의 오디오 압축 유닛 (512) 은, 음장 성분 추출 유닛 (520) 이 양자화 유닛 (534)("양자 유닛 (534)") 으로서 지칭되는 부가 유닛을 포함한다는 점에서 오디오 인코딩 디바이스 (510C) 의 오디오 압축 유닛 (512) 과 상이할 수도 있다. 이러한 이유로, 오디오 인코딩 디바이스 (510D) 의 음장 성분 추출 유닛 (520) 은 "음장 성분 추출 유닛 (520D)" 으로서 지칭된다.

[0835] 양자화 유닛 (534) 은 하나 이상의 V_{DIST} 벡터들 (525E) 및/또는 하나 이상의 V^{T}_{BG} 벡터들 (525F) 를 양자화하여 대응하는 하나 이상의 $V^{T}_{Q_DIST}$ 벡터들 (525G) 및/또는 하나 이상의 $V^{T}_{Q_BG}$ 벡터들 (525H) 을 생성하도록 구성되는 유닛을 나타낸다. 양자화 유닛 (534) 은 비트스트림 (517) 에서 하나 이상의 V_{DIST} 벡터들 (525E) 을 나타내기 위해 사용되는 비트들의 수를 감소시키도록 하나 이상의 V^{T}_{DIST} 벡터들 (525E) 을 양자화 (값을 나타내기 위해 사용되는 비트들의 소거를 통한 수학적 라운딩을 위한 신호 프로세싱 용어임) 할 수도 있다. 일부 예들에서, 양자화 유닛 (534) 은 하나 이상의 V_{DIST} 벡터들 (525E) 의 32 비트 값들을 양자화하여, 이들 32 비트 값들을 라운딩된 16 비트 값들로 대체하여 하나 이상의 $V^{T}_{Q_DIST}$ 벡터들 (525G) 을 생성할 수도 있다. 이러한 관점에서, 양자화 유닛 (534) 은 도 4 이 예에 나타낸 오디오 인코딩 디바이스 (20) 의 양자화 유닛 (52) 에 대해 전술한 것과 유사한 방식으로 동작할 수도 있다.

[0836] 이러한 성질의 양자화는 양자화의 조악함에 따라 변화하는 음장들의 표현에 에러를 도입할 수도 있다. 즉, V^{T}_{DIST} 벡터들 (525E) 을 나타내기 위해 사용되는 많은 비트들은 보다 적은 양자화 에러를 유발할 수도 있다. V^{T}_{DIST} Q벡터들 (525E) 의 양자화에 기인하는 양자화 에러 ("E_{DIST}" 로 지칭될 수도 있음) 는 하나 이상의 $V^{T}_{Q_DIST}$ 벡터들 (525G) 로부터 하나 이상의 V^{T}_{DIST} 벡터들 (525E) 을 감산하는 것에 의해 결정될 수도 있다.

[0837] 본 개시물에 기재된 기술들에 따라, 오디오 인코딩 디바이스 (510D) 는 E_{DIST} 에러를 프로젝트하거나 그렇지 않으면 $U_{DIST} * S_{DIST}$ 벡터들 (527) 의 하나 이상을 수정하는 것에 의해 E_{DIST} 양자화 에러들의 하나 이상을 보상하거나, 하나 이상의 U_{BG} 벡터들 (525D) 을 하나 이상의 S_{BG} 벡터들 (525B) 로 곱한 후 하나 이상의 V^{T}_{BG} 벡터들 (525F) 로 곱하는 것에 의해 생성된 배경 구면 조화 계수들 (531) 을 보상할 수도 있다. 일부 예들에

서, 오디오 인코딩 디바이스 (510D) 는 단지 U_{DIST} * S_{DIST} 벡터들 (527) 에서 E_{DIST} 에러만을 보상할 수도 있다.

다른 예들에서, 오디오 인코딩 디바이스 (510D) 는 단지 배경 구면 조화 계수들에서 E_{BG} 에러들만을 보상할 수도 있다. 또 다른 예들에서, 오디오 인코딩 디바이스 (510D) 는 U_{DIST} * S_{DIST} 벡터들 (527) 및 배경 구면 조화 계수들 모두에서 E_{DIST} 에러를 보상할 수도 있다.

[0838] 동작에 있어서, 돌출 성분 분석 유닛 (524) 은 하나 이상의 S_{DIST} 벡터들 (525), 하나 이상의 S_{BG} 벡터들 (525B), 하나 이상의 U_{DIST} 벡터들 (525C), 하나 이상의 U_{BG} 벡터들 (525D), 하나 이상의 V_{DIST} 벡터들 (525E), 및 하나 이상의 V^{T}_{BG} 벡터들 (525F) 을 계산 유닛 (526) 에 출력하도록 구성될 수도 있다. 돌출 성분 분석 유닛 (524) 은 또한 하나 이상의 V_{DIST} 벡터들 (525E) 을 양자화 유닛 (534) 에 출력할 수도 있다. 양자화 유닛 (534) 은 하나 이상의 V^{T}_{DIST} 벡터들 (525E) 을 양자화하여 하나 이상의 $V^{T}_{Q_DIST}$ 벡터들 (525G) 을 생성할 수도 있다.

양자화 유닛 (534) 은 하나 이상의 $V^{T}_{Q_DIST}$ 벡터들 (525G) 를 벡터 리오더링 유닛 (535) (전술한 바와 같음) 에 또한 제공하면서, 하나 이상의 $V^{T}_{Q_DIST}$ 벡터들 (525G) 을 계산 유닛 (526) 에 제공한다. 벡터 리오더 유닛 (532) 은 V^{T}_{DIST} 벡터들 (525E) 에 대해 전술한 것과 유사한 방식으로 하나 이상의 $V^{T}_{Q_DIST}$ 벡터들 (525G) 에 대해 동작할 수도 있다.

[0839] 이들 벡터들 (525-525G)("벡터들 (525)") 을 수신하면, 계산 유닛 (526) 은 먼저 음장의 구별되는 성분들을 기술하는 구별되는 구면 조화 계수들 및 음장의 배경 성분들을 기술했던 배경 구면 조화 계수들을 결정할 수도 있다. 행렬 계산 유닛 (526) 은 하나 이상의 U_{DIST} (525C) 벡터들을 하나 이상의 S_{DIST} 벡터들 (525A) 로 곱한 후 하나 이상의 V_{DIST} 벡터들 (525E) 로 곱하는 것에 의해 구별되는 구면 조화 계수들을 결정하도록 구성될 수도 있다. 계산 유닛 (526) 은 하나 이상의 U_{BG} (525D) 벡터들을 하나 이상의 S_{BG} 벡터들 (525A) 로 곱한 후 하나 이상의 V^{T}_{BG} 벡터들 (525E) 로 곱하는 것에 의해 배경 구면 조화 계수들을 결정하도록 구성될 수도 있다.

[0840] 계산 유닛 (526) 은 그 후 하나 이상의 $V^{T}_{Q_DIST}$ 벡터들 (525G) 에 대해 의사 역 동작을 수행하고 그 후 하나 이상의 $V^{T}_{Q_DIST}$ 벡터들 (525G) 의 의사 역으로 구별되는 구면 조화들을 곱하는 것에 의해 하나 이상의 보상된 U_{DIST} * S_{DIST} 벡터들 (527')(이들 벡터들이 E_{DIST} 에러를 보상하기 위한 값들을 포함하는 것을 제외하고 U_{DIST} * S_{DIST} 벡터들 (527) 과 유사할 수도 있음) 을 결정할 수도 있다. 벡터 리오더 유닛 (532) 은 리오더링된 벡터들 (527') 을 생성하기 위해 전술한 방식으로 동작할 수도 있으며, 리오더링된 벡터들 (527') 은 그 후 또한 전술한 바와 같이 오디오 인코딩 유닛 (515A) 에 의해 오디오 인코딩되어 오디오 인코딩된 리오더링된 벡터들 (515') 을 생성한다.

[0841] 계산 유닛 (526) 은 다음에 E_{DIST} 에러를 배경 구면 조화 계수들에 프로젝트할 수도 있다. 계산 유닛 (526) 은, 이러한 프로젝션을 수행하기 위해서, 구별되는 구면 조화 계수들을 배경 구면 조화 계수들에 더하는 것에 의해 원래 구면 조화 계수들 (511) 을 결정하거나 그렇지 않으면 복구할 수도 있다. 계산 유닛 (526) 은 그 후 양자화된 구별되는 구면 조화 계수들 (U_{DIST} 벡터들 (525C) 을 S_{DIST} 벡터들 (525A) 로 곱한 후 $V^{T}_{Q_DIST}$ 벡터들 (525G) 로 곱하는 것에 의해 생성될 수도 있음) 및 배경 구면 조화 계수들을 구면 조화 계수들 (511) 로부터 빼서 V^{T}_{DIST} 벡터들 (519) 의 양자화에 기인하는 나머지 에러를 결정할 수도 있다. 계산 유닛 (526) 은 그 후 이러한 에러를 양자화된 배경 구면 조화 계수들에 부가하여 보상된 양자화된 구면 조화 계수들 (531') 을 생성할 수도 있다.

[0842] 어떤 경우든, 오더 감소 유닛 (528A) 은 보상된 양자화된 배경 구면 조화 계수들 (531') 을, 오디오 인코딩된 감소된 배경 구면 조화 계수들 (515B') 을 생성하기 위해 전술한 방식으로 오디오 인코딩 유닛 (514) 에 의해 오디오 인코딩될 수도 있는 감소된 배경 구면 조화 계수들 (529') 로 감소시키기 위해 전술한 바와 같이 수행할 수도 있다.

[0843] 이러한 방식으로, 기술들은 오디오 인코딩 디바이스 (510D) 가 하나 이상의 벡터들, 예컨대 음장의 하나 이상의

성분들을 나타내는 V^{T}_{DIST} 벡터들 (525E) 을 양자화 하는 것을 가능하게 하고, 또한 음장의 동일한 하나 이상의 성분들을 나타내는, 배경 조화 계수들 (531) 의 벡터들 및/또는 $U_{DIST} * S_{DIST}$ 벡터들 (527) 과 같은, 하나 이상의 제 2 벡터들에서 하나 이상의 제 1 벡터들의 양자화에 기인하여 도입된 에러를 보상할 수도 있다.

[0844] 게다가, 기술들은 다음의 조항들에 따라 이러한 양자화 에러 보상을 제공할 수도 있다.

[0845] 조항 133146-1B. 음장의 하나 이상의 구별되는 성분들을 나타내는 하나 이상의 제 1 벡터들을 양자화하고, 그리고 음장의 동일한 하나 이상의 구별되는 성분들을 또한 나타내는 하나 이상의 제 2 벡터들에서 하나 이상의 제 1 벡터들의 양자화에 기인하여 도입되는 에러를 보상하도록 구성되는 하나 이상의 프로세서들을 포함하는, 오디오 인코딩 디바이스 (510D) 와 같은, 디바이스.

[0846] 조항 133146-2B. 하나 이상의 프로세서들은 음장을 기술하는 복수의 구면 조화 계수들에 대해 특이값 분해를 적어도 부분적으로 수행하는 것에 의해 생성된 V 행렬의 트랜스포즈로부터 하나 이상의 벡터들을 양자화하도록 구성되는, 조항 133146-1B 의 디바이스.

[0847] 조항 133146-3B. 하나 이상의 프로세서들은 또한, 음장을 나타내는 복수의 구면 조화 계수들에 대해 특이값 분해를 수행하여 복수의 구면 조화 계수들의 좌-특이 벡터들을 나타내는 U 행렬, 복수의 구면 조화 계수들의 특이값들을 나타내는 S 행렬, 및 복수의 구면 조화 계수들의 우-특이 벡터들을 나타내는 V 행렬을 생성하도록 구성되고, 하나 이상의 프로세서들은 V 행렬의 트랜스포즈로부터 하나 이상의 벡터들을 양자화하도록 구성되는, 조항 133146-1B 의 디바이스.

[0848] 조항 133146-4B. 하나 이상의 프로세서들은 음장의 복수의 구면 조화 계수들에 대해 특이값 분해를 수행하여 복수의 구면 조화 계수들의 좌-특이 벡터들을 나타내는 U 행렬, 복수의 구면 조화 계수들의 특이값들을 나타내는 S 행렬, 및 복수의 구면 조화 계수들의 우-특이 벡터들을 나타내는 V 행렬을 생성하도록 구성되고, 하나 이상의 프로세서들은 V 행렬의 트랜스포즈로부터 하나 이상의 벡터들을 양자화하도록 구성되며, 하나 이상의 프로세서들은 S 행렬의 하나 이상의 S 벡터들로 U 행렬의 하나 이상의 U 벡터들을 곱한 것에 의해 컴퓨팅된 하나 이상의 U * S 벡터들에서의 양자화에 기인하여 도입되는 에러를 보상하도록 구성되는, 조항 133146-1B 의 디바이스.

[0849] 조항 133146-5B. 음장을 나타내는 복수의 구면 조화 계수들에 대해 특이값 분해를 수행하여 복수의 구면 조화 계수들의 좌-특이 벡터들을 나타내는 U 행렬, 복수의 구면 조화 계수들의 특이값들을 나타내는 S 행렬, 및 복수의 구면 조화 계수들의 우-특이 벡터들을 나타내는 V 행렬을 생성하고, 각각이 음장의 구별되는 성분들 중 하나에 대응하는, U 행렬의 하나 이상의 U_{DIST} 벡터들을 결정하고, 각각이 음장의 구별되는 성분들의 동일한 것에 대응하는, S 행렬의 하나 이상의 S_{DIST} 벡터들을 결정하며, 각각이 음장의 구별되는 성분들의 동일한 것에 대응하는 V 행렬의 트랜스포즈의 하나 이상의 V^{T}_{DIST} 벡터들을 결정하도록 구성되고,

[0850] 하나 이상의 프로세서들은 또한 하나 이상이 V^{T}_{DIST} 벡터들을 양자화하여 하나 이상의 $V^{T}_{Q_DIST}$ 벡터들을 생성하도록 구성되고, 하나 이상의 프로세서들은 하나 이상의 에러 보상된 $U_{DIST} * S_{DIST}$ 벡터들을 생성하도록 S 행렬의 하나 이상의 S_{DIST} 벡터들로 U 행렬의 하나 이상의 U_{DIST} 벡터들을 곱하는 것에 의해 컴퓨팅된 하나 이상의 $U_{DIST} * S_{DIST}$ 벡터들에서의 양자화에 기인하여 도입되는 에러를 보상하도록 구성되는, 조항 133146-1B 의 디바이스.

[0851] 조항 133146-6B. 하나 이상의 프로세서들은 하나 이상의 U_{DIST} 벡터들, 하나 이상의 S_{DIST} 벡터들 및 하나 이상의 V^{T}_{DIST} 벡터들에 기초하여 구별되는 구면 조화 계수들을 결정하고, 그리고 $V^{T}_{Q_DIST}$ 벡터들에 대해 의사 역을 수행하여 하나 이상의 $V^{T}_{Q_DIST}$ 벡터들로 구별되는 구면 조화 계수들을 나누고 이에 의해 V^{T}_{DIST} 벡터들의 양자화를 통해 도입되는 에러를 적어도 부분적으로 보상하는 에러 보상된 하나 이상의 $U_{C_DIST} * S_{C_DIST}$ 벡터들을 생성하도록 구성되는, 조항 133146-5B 의 디바이스.

[0852] 조항 133146-7B. 하나 이상의 프로세서들은 또한 하나 이상의 보상된 $U_{DIST} * S_{DIST}$ 벡터들을 오디오 인코딩하도록 구성되는, 조항 133146-5B 의 디바이스.

[0853] 조항 133146-8B. 조항 133146-1B 의 디바이스에 있어서, 하나 이상의 프로세서들은 또한, 음장을 나타내는

복수의 구면 조화 계수들에 대하여 특이값 분해를 수행하여 복수의 구면 조화 계수들의 좌-특이 벡터들을 나타내는 U 행렬, 복수의 구면 조화 계수들의 특이값들을 나타내는 S 행렬 및 복수의 구면 조화 계수들의 우-특이 벡터들을 나타내는 V 행렬을 생성하고, 음장의 하나 이상의 배경 성분들을 기술하는 U 행렬의 하나 이상의 U_{BG} 벡터들 및 음장의 하나 이상의 구별되는 성분들을 기술하는 U 행렬의 하나 이상의 U_{DIST} 벡터들을 결정하고, 음장의 하나 이상의 배경 성분들을 기술하는 S 행렬의 하나 이상의 S_{BG} 벡터들 및 음장의 하나 이상의 구별되는 성분들을 기술하는 S 행렬의 하나 이상의 S_{DIST} 벡터들을 결정하며, V 행렬의 트랜스포즈의 하나 이상의 V^T_{DIST} 벡터들 및 하나 이상의 V^T_{BG} 벡터들을 결정하도록 구성되며, V^T_{DIST} 벡터들은 음장의 하나 이상의 구별되는 성분들을 기술하며 V^T_{BG} 는 음장의 하나 이상의 배경 성분들을 기술하며, 하나 이상의 프로세서들은 하나 이상의 V^T_{DIST} 벡터들을 양자화하여 하나 이상의 $V^T_{Q_DIST}$ 벡터들을 생성하도록 구성되며, 하나 이상의 프로세서들은 또한, 하나 이상의 U_{BG} 벡터들에 하나 이상의 S_{BG} 벡터들을 곱하고 그 후 하나 이상의 V^T_{BG} 벡터들을 곱함으로써 형성된 배경 구면 조화 계수들에 있어서 양자화로 인해 도입된 에러의 적어도 부분을 보상하여 에러 보상된 배경 구면 조화 계수들을 생성하도록 구성되는, 디바이스.

[0854] 조항 133146-9B. 조항 133146-8B 의 디바이스에 있어서, 하나 이상의 프로세서들은 V^T_{DIST} 벡터들 및 U_{DIST} 벡터들에 S_{DIST} 벡터들을 곱함으로써 형성된 하나 이상의 $U_{DIST} * S_{DIST}$ 벡터들에 기초하여 에러를 결정하며, 배경 구면 조화 계수들에 결정된 에러를 더하여 에러 보상된 배경 구면 조화 계수들을 생성하도록 구성되는, 디바이스.

[0855] 조항 133146-10B. 조항 133146-8B 의 디바이스에 있어서, 하나 이상의 프로세서들은 또한, 에러 보상된 배경 구면 조화 계수들을 오디오 인코딩하도록 구성되는, 디바이스.

[0856] 조항 133146-11B. 조항 133146-1B 의 디바이스에 있어서, 하나 이상의 프로세서들은 음장의 동일한 하나 이상의 성분들을 또한 나타내는 하나 이상의 제 2 벡터들에 있어서 하나 이상의 제 1 벡터들의 양자화로 인해 도입된 에러를 보상하여 하나 이상의 에러 보상된 제 2 벡터들을 생성하도록 구성되며, 하나 이상의 프로세서들은 또한, 하나 이상의 에러 보상된 제 2 벡터들 및 양자화된 하나 이상의 제 1 벡터들을 포함하기 위한 비트스트림을 생성하도록 구성되는, 디바이스.

[0857] 조항 133146-12B. 조항 133146-1B 의 디바이스에 있어서, 하나 이상의 프로세서들은 음장의 동일한 하나 이상의 성분들을 또한 나타내는 하나 이상의 제 2 벡터들에 있어서 하나 이상의 제 1 벡터들의 양자화로 인해 도입된 에러를 보상하여 하나 이상의 에러 보상된 제 2 벡터들을 생성하도록 구성되며, 하나 이상의 프로세서들은 또한, 하나 이상의 에러 보상된 제 2 벡터들을 오디오 인코딩하고 오디오 인코딩된 하나 이상의 에러 보상된 제 2 벡터들 및 양자화된 하나 이상의 제 1 벡터들을 포함하기 위한 비트스트림을 생성하도록 구성되는, 디바이스.

[0858] 조항 133146-1C. 오디오 인코딩 디바이스 (510D) 와 같은 디바이스로서, 음장의 하나 이상의 구별되는 성분들을 나타내는 하나 이상의 제 1 벡터들을 양자화하며, 음장의 하나 이상의 배경 성분들을 나타내는 하나 이상의 제 2 벡터들에 있어서 하나 이상의 제 1 벡터들의 양자화로 인해 도입된 에러를 보상하도록 구성된 하나 이상의 프로세서들을 포함하는, 디바이스.

[0859] 조항 133146-2C. 조항 133146-1C 의 디바이스에 있어서, 하나 이상의 프로세서들은 적어도 부분적으로, 음장을 기술하는 복수의 구면 조화 계수들에 대하여 특이값 분해를 수행함으로써 생성된 V 행렬의 트랜스포즈로부터 하나 이상의 벡터들을 양자화하도록 구성되는, 디바이스.

[0860] 조항 133146-3C. 조항 133146-1C 의 디바이스에 있어서, 하나 이상의 프로세서들은 또한, 음장을 나타내는 복수의 구면 조화 계수들에 대하여 특이값 분해를 수행하여 복수의 구면 조화 계수들의 좌-특이 벡터들을 나타내는 U 행렬, 복수의 구면 조화 계수들의 특이값들을 나타내는 S 행렬 및 복수의 구면 조화 계수들의 우-특이 벡터들을 나타내는 V 행렬을 생성하도록 구성되며, 하나 이상의 프로세서들은 V 행렬의 트랜스포즈로부터 하나 이상의 벡터들을 양자화하도록 구성되는, 디바이스.

[0861] 조항 133146-4C. 조항 133146-1C 의 디바이스에 있어서, 하나 이상의 프로세서들은 또한, 음장을 나타내는 복수의 구면 조화 계수들에 대하여 특이값 분해를 수행하여 복수의 구면 조화 계수들의 좌-특이 벡터들을 나타내는 U 행렬, 복수의 구면 조화 계수들의 특이값들을 나타내는 S 행렬 및 복수의 구면 조화 계수들의 우-특이 벡터들을 나타내는 V 행렬을 생성하고, 각각이 음장의 구별되는 성분들 중 하나의 성분에 대응하는 U 행렬의 하

나 이상의 U_{DIST} 벡터들을 결정하고, 각각이 음장의 구별되는 성분들 중 동일한 하나의 성분에 대응하는 S 행렬의 하나 이상의 S_{DIST} 벡터들을 결정하며, 각각이 음장의 구별되는 성분들 중 동일한 하나의 성분에 대응하는 V 행렬의 트랜스포즈의 하나 이상의 V^{T}_{DIST} 벡터들을 결정하도록 구성되며, 하나 이상의 프로세서들은 하나 이상의 V^{T}_{DIST} 벡터들을 양자화하여 하나 이상의 $V^{T}_{Q_DIST}$ 벡터들을 생성하며, U 행렬의 하나 이상의 U_{DIST} 벡터들에 S 행렬의 하나 이상의 S_{DIST} 벡터들을 곱함으로써 컴퓨팅된 하나 이상의 $U_{DIST} * S_{DIST}$ 벡터들에 있어서 양자화로 인해 도입된 에러의 적어도 부분을 보상하여 하나 이상의 에러 보상된 $U_{DIST} * S_{DIST}$ 벡터들을 생성하도록 구성되는, 디바이스.

[0862] 조항 133146-5C. 조항 133146-4C 의 디바이스에 있어서, 하나 이상의 프로세서들은 하나 이상의 U_{DIST} 벡터들, 하나 이상의 S_{DIST} 벡터들 및 하나 이상의 V^{T}_{DIST} 벡터들에 기초하여 구별되는 구면 조화 계수들을 결정하며, $V^{T}_{Q_DIST}$ 벡터들에 대하여 의사역행렬을 수행하여 구별되는 구면 조화 계수들을 하나 이상의 $V^{T}_{Q_DIST}$ 벡터들로 나누고 이로써 V^{T}_{DIST} 벡터들의 양자화를 통하여 도입된 에러를 적어도 부분적으로 보상하는 하나 이상의 $U_{C_DIST} * S_{C_DIST}$ 벡터들을 생성하도록 구성되는, 디바이스.

[0863] 조항 133146-6C. 조항 133146-4C 의 디바이스에 있어서, 하나 이상의 프로세서들은 또한, 하나 이상의 에러 보상된 $U_{DIST} * S_{DIST}$ 벡터들을 오디오 인코딩하도록 구성되는, 디바이스.

[0864] 조항 133146-7C. 조항 133146-1C 의 디바이스에 있어서, 하나 이상의 프로세서들은 또한, 음장을 나타내는 복수의 구면 조화 계수들에 대하여 특이값 분해를 수행하여 복수의 구면 조화 계수들의 좌-특이 벡터들을 나타내는 U 행렬, 복수의 구면 조화 계수들의 특이값들을 나타내는 S 행렬 및 복수의 구면 조화 계수들의 우-특이 벡터들을 나타내는 V 행렬을 생성하고, 음장의 하나 이상의 배경 성분들을 기술하는 U 행렬의 하나 이상의 U_{BG} 벡터들 및 음장의 하나 이상의 구별되는 성분들을 기술하는 U 행렬의 하나 이상의 U_{DIST} 벡터들을 결정하고, 음장의 하나 이상의 배경 성분들을 기술하는 S 행렬의 하나 이상의 S_{BG} 벡터들 및 음장의 하나 이상의 구별되는 성분들을 기술하는 S 행렬의 하나 이상의 S_{DIST} 벡터들을 결정하며, V 행렬의 트랜스포즈의 하나 이상의 V^{T}_{DIST} 벡터들 및 하나 이상의 V^{T}_{BG} 벡터들을 결정하도록 구성되며, V^{T}_{DIST} 벡터들은 음장의 하나 이상의 구별되는 성분들을 기술하며 V^{T}_{BG} 는 음장의 하나 이상의 배경 성분들을 기술하며, 하나 이상의 프로세서들은 하나 이상의 V^{T}_{DIST} 벡터들을 양자화하여 하나 이상의 $V^{T}_{Q_DIST}$ 벡터들을 생성하도록 구성되며, 하나 이상의 프로세서들은 하나 이상의 U_{BG} 벡터들에 하나 이상의 S_{BG} 벡터들을 곱하고 그 후 하나 이상의 V^{T}_{BG} 벡터들을 곱함으로써 형성된 배경 구면 조화 계수들에 있어서 양자화로 인해 도입된 에러를 보상하여 에러 보상된 배경 구면 조화 계수들을 생성하도록 구성되는, 디바이스.

[0865] 조항 133146-8C. 조항 133146-7C 의 디바이스에 있어서, 하나 이상의 프로세서들은 V^{T}_{DIST} 벡터들 및 U_{DIST} 벡터들에 S_{DIST} 벡터들을 곱함으로써 형성된 하나 이상의 $U_{DIST} * S_{DIST}$ 벡터들에 기초하여 에러를 결정하며, 배경 구면 조화 계수들에 결정된 에러를 더하여 에러 보상된 배경 구면 조화 계수들을 생성하도록 구성되는, 디바이스.

[0866] 조항 133146-9C. 조항 133146-7C 의 디바이스에 있어서, 하나 이상의 프로세서들은 또한, 에러 보상된 배경 구면 조화 계수들을 오디오 인코딩하도록 구성되는, 디바이스.

[0867] 조항 133146-10C. 조항 133146-1C 의 디바이스에 있어서, 하나 이상의 프로세서들은 또한, 음장의 동일한 하나 이상의 성분들을 또한 나타내는 하나 이상의 제 2 벡터들에 있어서 하나 이상의 제 1 벡터들의 양자화로 인해 도입된 에러를 보상하여 하나 이상의 에러 보상된 제 2 벡터들을 생성하며, 하나 이상의 에러 보상된 제 2 벡터들 및 양자화된 하나 이상의 제 1 벡터들을 포함하기 위한 비트스트림을 생성하도록 구성되는, 디바이스.

[0868] 조항 133146-11C. 조항 133146-1C 의 디바이스에 있어서, 하나 이상의 프로세서들은 또한, 음장의 동일한 하나 이상의 성분들을 또한 나타내는 하나 이상의 제 2 벡터들에 있어서 하나 이상의 제 1 벡터들의 양자화로 인해 도입된 에러를 보상하여 하나 이상의 에러 보상된 제 2 벡터들을 생성하고, 하나 이상의 에러 보상된 제 2

벡터들을 오디오 인코딩하며, 오디오 인코딩된 하나 이상의 에러 보상된 제 2 벡터들 및 양자화된 하나 이상의 제 1 벡터들을 포함하기 위한 비트스트림을 생성하도록 구성되는, 디바이스.

[0869] 즉, 대역폭 감소를 목적으로 HoA 신호들에 대해 프레임 기반 SVD (또는 KLT & PCA 와 같은 관련 방법들) 분해를 이용할 때, 본 개시물에서 설명된 기술들은 오디오 인코딩 디바이스 (10D) 가 (S 행렬의 대응하는 특이값들을 곱한) U 행렬의 처음의 몇몇 벡터들은 물론 V 벡터의 대응하는 벡터들을 양자화하는 것을 가능하게 할 수도 있다. 이것은 음장의 '전경' 또는 '구별되는' 성분들을 포함할 것이다. 기술들은 그 후, 오디오 인코딩 디바이스 (510D) 가 AAC 인코더와 같은 '블랙-박스' 오디오-코딩 엔진을 이용하여 U*S 벡터들을 코딩하는 것을 가능하게 할 수도 있다. V 벡터는 스칼라 양자화되거나 벡터 양자화되거나 둘 중 어느 하나일 수도 있다.

[0870] 또한, U 행렬에서의 나머지 벡터들 중 일부는 S 행렬 및 V 행렬의 대응하는 특이값들과 곱해지고 또한 '블랙-박스' 오디오-코딩 엔진을 이용하여 코딩될 수도 있다. 이들은 음장의 '배경' 성분들을 포함할 것이다. V 벡터들의 단순 16 비트 스칼라 양자화는 대략적으로 제 4 오더 (25 개의 계수들) 에 대해 80kbps 오버헤드 및 제 6 오더 (49 개의 계수들) 에 대해 160kbps 를 초래할 수도 있다. 더 개략적인 (coarse) 양자화는 더 큰 양자화 에러들을 초래할 수도 있다. 본 개시물에서 설명된 기술들은 V 벡터들의 양자화 에러를 - V 벡터의 양자화 에러를 전경 및 배경 성분들로 '투영하는 것' 에 의해 보상할 수도 있다.

[0871] 본 개시물에서의 기술들은 실제 V 벡터의 양자화된 버전을 계산하는 것을 포함할 수도 있다. 이 양자화된 V 벡터는 V' (여기서 V'=V+e) 로 불릴 수도 있다. 기술들이 재현하기 위해 시도중인 - 전경 성분들에 대한 - 근본적인 HoA 신호는 H_f=USV 에 의해 주어지며, 여기서 U, S 및 V 는 단지 전경 엘리먼트들만을 포함한다. 이 논의의 목적을 위해, US 는 단일 세트의 벡터들 U 에 의해 대체될 것이다. 따라서, H_f=UV 이다. 잘못된 V' 를 갖는다는 것을 고려해 볼 때, 기술들은 H_f 를 가능한 한 가깝게 재현하려고 시도하고 있다. 따라서, 기술들은 오디오 인코딩 디바이스 (10D) 가 H_f=U'V' 가 되도록 U' 를 발견하는 것을 가능하게 할 수도 있다. 오디오 인코딩 디바이스 (10D) 는 U' = H_f [V']^(-1) 을 허용하는 의사역행렬 방법론을 이용할 수도 있다. U' 를 코딩하기 위해 소위 '블랙박스' 오디오-코딩 엔진을 이용하면, 기술들은 잘못된 V' 벡터로 지칭될 수도 있는 것에 의해 야기된, H 의 에러를 최소화할 수도 있다.

[0872] 유사한 방식으로, 기술들은 또한, 오디오 인코딩 디바이스가 V 를 배경 엘리먼트들로 양자화하는 것으로 인해 에러를 투영하는 것을 가능하게 할 수도 있다. 오디오 인코딩 디바이스 (510D) 는 전경 및 배경 HoA 신호들의 결합, 즉 H=H_f + H_b 인 총 HoA 신호를 재현하도록 구성될 수도 있다. 이것은 다시 V' 의 양자화 에러로 인해 H = H_f + e + H_b 로서 모델링될 수 있다. 이 방식으로, '블랙-박스 오디오-코더' 를 통하여 H_b 를 넣는 대신에, 오디오-코더를 통하여 (e + H_b) 를 넣어, 사실상 V' 의 에러를 보상한다. 실제로, 이것은 배경 엘리먼트들을 청하기 위해 오디오 인코딩 디바이스 (510D) 에 의해 결정된 오더까지만 에러를 보상한다.

[0873] 도 40e 는 2 차원 또는 3 차원 음장들을 기술하는 구면 조화 계수들을 압축하기 위해 본 개시물에서 설명된 기술들의 다양한 양태들을 수행할 수도 있는 예시적인 오디오 인코딩 디바이스 (510E) 를 예시하는 블록 다이어그램이다. 오디오 인코딩 디바이스 (510E) 는 오디오 인코딩 디바이스 (510E) 가 오디오 압축 유닛 (512), 오디오 인코딩 유닛 (514) 및 비트스트림 생성 유닛 (516) 을 포함한다는 점에서 오디오 인코딩 디바이스 (510D) 와 유사할 수도 있다. 더욱이, 오디오 인코딩 디바이스 (510E) 의 오디오 압축 유닛 (512) 은 오디오 압축 유닛 (512) 이 분해 유닛 (518) 을 포함한다는 점에서 오디오 인코딩 디바이스 (510D) 의 것과 유사할 수도 있다.

[0874] 그러나, 오디오 인코딩 디바이스 (510E) 의 오디오 압축 유닛 (512) 은 음장 성분 추출 유닛 (520) 의 계산 유닛 (526) 이 비트스트림 (517) 에 V 행렬 (519A) 의 트랜스포즈의 감소된 버전을 포함시키기 이전에 V 행렬 (519A) 을 더 감소시키기 위해 본 개시물에서 설명된 기술들의 추가적인 양태들을 수행한다는 점에서 오디오 인코딩 디바이스 (510D) 의 오디오 압축 유닛 (512) 과는 상이할 수도 있다. 이런 이유로, 오디오 인코딩 디바이스 (510E) 의 음장 성분 추출 유닛 (520) 은 "음장 성분 추출 유닛 (520E)" 으로 표기된다.

[0875] 도 40e 의 예에서, 오더 감소 유닛 (528) 은, 감소된 배경 구면 조화 계수들 (529') 을 오디오 인코딩 유닛 (514) 으로 포워딩하기 보다는, 감소된 배경 구면 조화 계수들 (529') 을 계산 유닛 (526) 으로 반환한다. 상기 언급한 바와 같이, 이들 감소된 배경 구면 조화 계수들 (529') 은 계수들 중 하나 이상의 식별된 오더들 및/또는 서브-오더들을 갖는 구면 기저 함수 (spherical basis function) 들에 대응하는 그 계수들을 제거함으로써 감소되었을 수도 있다. 감소된 배경 구면 조화 계수들 (529') 의 감소된 오더는 변수 N_{BG} 로 표기될 수도 있다.

[0876] 음장 성분 추출 유닛 (520E) 이 리오더링된 하나 이상의 $U_{DIST} * S_{DIST}$ 벡터들 (533') 에 대하여 오더 감소를 수행하지 않을 수도 있다는 것을 고려해 볼 때, (변수 N_{DIST} 로 표기될 수도 있는) 음장의 구별되는 성분들을 기술하는 이 구면 조화 계수들의 분해의 오더는 배경 오더 N_{BG} 보다 더 클 수도 있다. 즉, N_{BG} 는 통상 N_{DIST} 보다 더 작을 수도 있다. N_{BG} 가 N_{DIST} 보다 더 작을 수도 있는 하나의 가능한 이유는 배경 성분들이 많은 방향성을 갖지 않아 상위 오더의 구면 기저 함수들이 요구되지 않으며, 이로써 오더 감소를 가능하게 하고 그 결과 N_{BG} 가 N_{DIST} 보다 더 작게 되는 것으로 가정된다는 것이다.

[0877] 리오더링된 하나 이상의 $V^T_{Q_DIST}$ 벡터들 (539) 이 도 40a 내지 도 40d 의 예들에 도시한 바와 같이, 비트스트림 (517) 에서의 이들 벡터들 (539) 을 오디오 인코딩하지 않고, 이전에 공개적으로 전송되었다는 것을 고려해 볼 때, 리오더링된 하나 이상의 $V^T_{Q_DIST}$ 벡터들 (539) 은 상당한 대역폭을 소비할 수도 있다. 하나의 예로서, 리오더링된 하나 이상의 $V^T_{Q_DIST}$ 벡터들 (539) 각각은, 16-비트 스칼라 값들로 양자화될 때, 대략적으로, 제 4 오더 앰비소닉스 오디오 데이터 (여기서 각각의 벡터는 25 개의 계수들을 갖는다) 에 대해 20Kbps 및 제 6 오더 앰비소닉스 오디오 데이터 (여기서 각각의 벡터는 49 개의 계수들을 갖는다) 에 대해 40Kbps 를 소비할 수도 있다.

[0878] 본 개시물에서 설명된 기술들의 다양한 양태들에 따르면, 음장 성분 추출 유닛 (520E) 은 구면 조화 계수들 또는 이들의 분해들, 이를 테면 리오더링된 하나 이상의 $V^T_{Q_DIST}$ 벡터들 (539) 에 대해 특정되어야 하는 비트들의 양을 감소시킬 수도 있다. 일부 예들에서, 계산 유닛 (526) 은 오더 감소된 구면 조화 계수들 (529') 에 기초하여, 리오더링된 $V^T_{Q_DIST}$ 벡터들 (539) 중 오더 감소된 구면 조화 계수들 (529') 과 제거 및 재결합될 그 벡터들 및 리오더링된 $V^T_{Q_DIST}$ 벡터들 (539) 중 V^T_{SMALL} 벡터들 (521) 을 형성할 그 벡터들을 결정할 수도 있다. 즉, 계산 유닛 (526) 은 오더 감소된 구면 조화 계수들 (529') 의 오더를 결정할 수도 있으며, 여기서 이 오더는 N_{BG} 로 표기될 수도 있다. 리오더링된 $V^T_{Q_DIST}$ 벡터들 (539) 은 변수 N_{DIST} 로 표기된 오더가 될 수도 있으며, 여기서 N_{DIST} 는 오더 N_{BG} 보다 더 크다.

[0879] 계산 유닛 (526) 은 그 후 리오더링된 $V^T_{Q_DIST}$ 벡터들 (539) 의 처음의 N_{BG} 오더들을 파싱하여, N_{BG} 이상인 오더를 갖는 구면 기저 함수들에 대응하는 분해된 구면 조화 계수들을 명시하는 그 벡터들을 제거할 수도 있다. 이들 제거된 리오더링된 $V^T_{Q_DIST}$ 벡터들 (539) 은 그 후 중간 구별되는 구면 조화 계수들을 형성하기 위해 구면 조화 계수들 (511) 의 분해된 버전들을 나타내는 리오더링된 $U_{DIST} * S_{DIST}$ 벡터들 (533') 중 N_{BG} 이하인 오더를 갖는 구면 기저 함수들에 대응하는 그 벡터들에 제거된 리오더링된 $V^T_{Q_DIST}$ 벡터들 (539) 을 곱함으로써 중간 구면 조화 계수들을 형성하는데 이용될 수도 있다. 계산 유닛 (526) 은 그 후 오더 감소된 구면 조화 계수들 (529') 에 중간 구별되는 구면 조화 계수들을 합산함으로써 수정된 배경 구면 조화 계수들 (537) 을 생성할 수도 있다. 계산 유닛 (526) 은 그 후 이 수정된 배경 구면 조화 계수들 (537) 을 오디오 인코딩 유닛 (514) 으로 전달할 수도 있으며, 오디오 인코딩 유닛 (514) 은 오디오 인코딩된 수정된 배경 구면 조화 계수들 (515B') 을 형성하기 위해 이들 계수들 (537) 을 오디오 인코딩한다.

[0880] 계산 유닛 (526) 은 그 후 N_{BG} 보다 더 크고 N_{DIST} 이하인 오더를 갖는 구면 기저 함수들에 대응하는 구면 조화 계수들 (511) 의 분해된 형태를 나타내는 그 벡터들 (539) 을 표현할 수도 있는 하나 이상의 V^T_{SMALL} 벡터들 (521) 을 전달할 수도 있다. 이 점에 있어서, 계산 유닛 (526) 은 도 4 의 예에 도시된 오디오 인코딩 디바이스 (20) 의 계수 감소 유닛 (46) 과 유사한 동작들을 수행할 수도 있다. 계산 유닛 (526) 은 하나 이상의 V^T_{SMALL} 벡터들 (521) 을 비트스트림 생성 유닛 (516) 으로 전달할 수도 있으며, 비트스트림 생성 유닛 (516) 은 V^T_{SMALL} 벡터들 (521) 을 종종 그들의 원래의 비-오디오 인코딩된 형태로 포함하기 위한 비트스트림 (517) 을 생성할 수도 있다. V^T_{SMALL} 벡터들 (521) 이 리오더링된 $V^T_{Q_DIST}$ 벡터들 (539) 보다 더 적은 벡터들을 포함한다

는 것을 고려해 볼 때, 기술들은 단지 비트스트림 (517) 에서 V^{T}_{SMALL} 벡터들 (521) 을 명시하는 것에 의해 리오더링된 $V^{T}_{Q_DIST}$ 벡터들 (539) 에 더 적은 비트들의 할당을 가능하게 할 수도 있다.

[0881] 양자화되고 있지 않은 것으로 도시하였지만, 일부 경우들에서, 오디오 인코딩 디바이스 (510E) 는 V^{T}_{BG} 벡터들 (525F) 을 양자화할 수도 있다. 오디오 인코딩 유닛 (514) 이 배경 구면 조화 계수들을 압축하기 위해 이용되지 않을 때와 같은 일부 경우들에서는, 오디오 인코딩 디바이스 (510E) 는 V^{T}_{BG} 벡터들 (525F) 을 양자화할 수도 있다.

[0882] 이 방식으로, 기술들은 오디오 인코딩 디바이스 (510E) 가 비트스트림에서 하나 이상의 벡터들에 할당되도록 요구되는 비트들의 양을 감소시키기 위해 배경 구면 조화 계수들과 재결합될 구면 조화 계수들로부터 분해된 하나 이상의 벡터들 중의 적어도 하나의 벡터를 결정하는 것을 가능하게 할 수도 있으며, 여기서 구면 조화 계수들은 음장을 기술하며, 배경 구면 조화 계수들은 동일한 음장의 하나 이상의 배경 성분들을 기술하였다.

[0883] 즉, 기술들은 오디오 인코딩 디바이스 (510E) 가 다음의 조항들에 의해 나타내진 방식으로 구성되는 것을 가능하게 할 수도 있다.

[0884] 조항 133149-1A. 오디오 인코딩 디바이스 (510E) 와 같은 디바이스로서, 비트스트림에서 하나 이상의 벡터들에 할당되도록 요구되는 비트들의 양을 감소시키기 위해 배경 구면 조화 계수들과 재결합될 구면 조화 계수들로부터 분해된 하나 이상의 벡터들 중의 적어도 하나의 벡터를 결정하도록 구성된 하나 이상의 프로세서들을 포함하며, 구면 조화 계수들은 음장을 기술하며, 배경 구면 조화 계수들은 동일한 음장의 하나 이상의 배경 성분들을 기술하는, 디바이스.

[0885] 조항 133149-2A. 조항 133149-1A 의 디바이스에 있어서, 하나 이상의 프로세서들은 또한, 하나 이상의 벡터들로부터 하나 이상의 벡터들 중의 결정된 적어도 하나의 벡터를 제거함으로써 하나 이상의 벡터들의 감소된 세트를 생성하도록 구성되는, 디바이스.

[0886] 조항 133149-3A. 조항 133149-1A 의 디바이스에 있어서, 하나 이상의 프로세서들은 또한, 하나 이상의 벡터들로부터 하나 이상의 벡터들 중의 결정된 적어도 하나의 벡터를 제거함으로써 하나 이상의 벡터들의 감소된 세트를 생성하고, 하나 이상의 벡터들 중의 제거된 적어도 하나의 벡터를 배경 구면 조화 계수들과 재결합하여 수정된 배경 구면 조화 계수들을 생성하며, 하나 이상의 벡터들의 감소된 세트 및 수정된 배경 구면 조화 계수들을 포함하기 위한 비트스트림을 생성하도록 구성되는, 디바이스.

[0887] 조항 133149-4A. 조항 133149-3A 의 디바이스에 있어서, 하나 이상의 벡터들의 감소된 세트는 먼저 오디오 인코딩되지 않고 비트스트림에 포함되는, 디바이스.

[0888] 조항 133149-5A. 조항 133149-1A 의 디바이스에 있어서, 하나 이상의 프로세서들은 또한, 하나 이상의 벡터들로부터 하나 이상의 벡터들 중의 결정된 적어도 하나의 벡터를 제거함으로써 하나 이상의 벡터들의 감소된 세트를 생성하고, 하나 이상의 벡터들 중의 제거된 적어도 하나의 벡터를 배경 구면 조화 계수들과 재결합하여 수정된 배경 구면 조화 계수들을 생성하고, 수정된 배경 구면 조화 계수들을 오디오 인코딩하며, 하나 이상의 벡터들의 감소된 세트 및 오디오 인코딩된 수정된 배경 구면 조화 계수들을 포함하기 위한 비트스트림을 생성하도록 구성되는, 디바이스.

[0889] 조항 133149-6A. 조항 133149-1A 의 디바이스에 있어서, 하나 이상의 벡터들은 음장의 하나 이상의 구별되는 성분들의 적어도 일부 양태를 나타내는 벡터들을 포함하는, 디바이스.

[0890] 조항 133149-7A. 조항 133149-1A 의 디바이스에 있어서, 하나 이상의 벡터들은 적어도 부분적으로, 음장을 기술하는 복수의 구면 조화 계수들에 대하여 특이값 분해를 수행함으로써 생성된 V 행렬의 트랜스포즈로부터의 하나 이상의 벡터들을 포함하는, 디바이스.

[0891] 조항 133149-8A. 조항 133149-1A 의 디바이스에 있어서, 하나 이상의 프로세서들은 또한, 복수의 구면 조화 계수들에 대하여 특이값 분해를 수행하여 복수의 구면 조화 계수들의 좌-특이 벡터들을 나타내는 U 행렬, 복수의 구면 조화 계수들의 특이값들을 나타내는 S 행렬 및 복수의 구면 조화 계수들의 우-특이 벡터들을 나타내는 V 행렬을 생성하도록 구성되며, 하나 이상의 벡터들은 V 행렬의 트랜스포즈로부터 하나 이상의 벡터들을 포함하는, 디바이스.

[0892] 조항 133149-9A. 조항 133149-1A 의 디바이스에 있어서, 하나 이상의 프로세서들은 또한, 배경 구면 조화 계수들 중 식별된 오더 및/또는 서브-오더를 갖는 구면 기저 함수들에 대응하는 그 계수들을 제거하기 위하여 배경 구면 조화 계수들에 대한 오더 감소를 수행하도록 구성되며, 배경 구면 조화 계수들은 오더 N_{BG} 에 대응하는, 디바이스.

[0893] 조항 133149-10A. 조항 133149-1A 의 디바이스에 있어서, 하나 이상의 프로세서들은 또한, 배경 구면 조화 계수들 중 식별된 오더 및/또는 서브-오더를 갖는 구면 기저 함수들에 대응하는 그 계수들을 제거하기 위하여 배경 구면 조화 계수들에 대하여 오더 감소를 수행하도록 구성되며, 배경 구면 조화 계수들은 구별되는 구면 조화 계수들 N_{DIST} 의 오더보다 더 작은 오더 N_{BG} 에 대응하며, 구별되는 구면 조화 계수들은 음장의 구별되는 성분들을 나타내는. 디바이스.

[0894] 조항 133149-11A. 조항 133149-1A 의 디바이스에 있어서, 하나 이상의 프로세서들은 또한, 배경 구면 조화 계수들 중 식별된 오더 및/또는 서브-오더를 갖는 구면 기저 함수들에 대응하는 그 계수들을 제거하기 위하여 배경 구면 조화 계수들에 대하여 오더 감소를 수행하도록 구성되며, 배경 구면 조화 계수들은 구별되는 구면 조화 계수들 N_{DIST} 의 오더보다 더 작은 오더 N_{BG} 에 대응하며, 구별되는 구면 조화 계수들은 음장의 구별되는 성분들을 표현하며 오더 감소되지 않는, 디바이스.

[0895] 조항 133149-12A. 조항 133149-1A 의 디바이스에 있어서, 하나 이상의 프로세서들은 또한, 복수의 구면 조화 계수들에 대하여 특이값 분해를 수행하여 복수의 구면 조화 계수들의 좌-특이 벡터들을 나타내는 U 행렬, 복수의 구면 조화 계수들의 특이값들을 나타내는 S 행렬 및 복수의 구면 조화 계수들의 우-특이 벡터들을 나타내는 V 행렬을 생성하며, V 행렬의 트랜스포즈의 하나 이상의 V^T_{DIST} 벡터들 및 하나 이상의 V^T_{BG} 를 결정하도록 구성되며, 하나 이상의 V^T_{DIST} 벡터들은 음장의 하나 이상의 구별되는 성분들을 기술하며 하나 이상의 V^T_{BG} 벡터들은 음장의 하나 이상의 배경 성분들을 기술하며, 하나 이상의 벡터들은 하나 이상의 V^T_{DIST} 벡터들을 포함하는, 디바이스.

[0896] 조항 133149-13A. 조항 133149-1A 의 디바이스에 있어서, 하나 이상의 프로세서들은 또한, 복수의 구면 조화 계수들에 대하여 특이값 분해를 수행하여 복수의 구면 조화 계수들의 좌-특이 벡터들을 나타내는 U 행렬, 복수의 구면 조화 계수들의 특이값들을 나타내는 S 행렬 및 복수의 구면 조화 계수들의 우-특이 벡터들을 나타내는 V 행렬을 생성하고, V 행렬의 트랜스포즈의 하나 이상의 V^T_{DIST} 벡터들 및 하나 이상의 V^T_{BG} 를 결정하는 것으로서, 하나 이상의 V_{DIST} 벡터들은 음장의 하나 이상의 구별되는 성분들을 기술하며 하나 이상의 V_{BG} 벡터들은 음장의 하나 이상의 배경 성분들을 기술하는, 상기 하나 이상의 V^T_{DIST} 벡터들 및 하나 이상의 V^T_{BG} 를 결정하며, 하나 이상의 V^T_{DIST} 벡터들을 양자화하여 하나 이상의 $V^T_{Q_DIST}$ 벡터들을 생성하도록 구성되며, 하나 이상의 벡터들은 하나 이상의 $V^T_{Q_DIST}$ 벡터들을 포함하는, 디바이스.

[0897] 조항 133149-14A. 조항 133149-12A 또는 조항 133149-13A 중 어느 하나의 디바이스에 있어서, 하나 이상의 프로세서들은 또한, U 행렬의 하나 이상의 U_{DIST} 벡터들 및 하나 이상의 U_{BG} 벡터들을 결정하는 것으로서, 하나 이상의 U_{DIST} 벡터들은 음장의 하나 이상의 구별되는 성분들을 기술하며 하나 이상의 U_{BG} 벡터들은 음장의 하나 이상의 배경 성분들을 기술하는, 상기 하나 이상의 U_{DIST} 벡터들 및 하나 이상의 U_{BG} 벡터들을 결정하며, S 행렬의 하나 이상의 S_{DIST} 벡터들 및 하나 이상의 S_{BG} 벡터들을 결정하는 것으로서, 하나 이상의 S_{DIST} 벡터들은 음장의 하나 이상의 구별되는 성분들을 기술하며 하나 이상의 S_{BG} 벡터들은 음장의 하나 이상의 배경 성분들을 기술하는, 상기 하나 이상의 S_{DIST} 벡터들 및 하나 이상의 S_{BG} 벡터들을 결정하도록 구성되는, 디바이스.

[0898] 조항 133149-15A. 조항 133149-14A 의 디바이스에 있어서, 하나 이상의 프로세서들은 또한, 배경 구면 조화 계수들을 하나 이상의 U_{BG} 벡터들, 하나 이상의 S_{BG} 벡터들, 및 하나 이상의 V^T_{BG} 의 함수로서 결정하고, 배경 구면 조화 계수들에 대하여 오더 감소를 수행하여 N_{BG} 와 동일한 오더를 갖는 감소된 배경 구면 조화 계수들을 생성하고, 하나 이상의 U_{DIST} 에 하나 이상의 S_{DIST} 벡터들을 곱하여 하나 이상의 $U_{DIST} * S_{DIST}$ 벡터들을 생성하고, 하

나 이상의 벡터들로부터 하나 이상의 벡터들 중의 결정된 적어도 하나의 벡터를 제거하여 하나 이상의 벡터들의 감소된 세트를 생성하고, 하나 이상의 $U_{DIST} * S_{DIST}$ 벡터들에 하나 이상의 V^T_{DIST} 벡터들 또는 하나 이상의 $V^T_{Q_DIST}$ 벡터들 중의 제거된 적어도 하나의 벡터를 곱하여 중간 구별되는 구면 조화 계수들을 생성하며, 배경 구면 조화 계수에 중간 구별되는 구면 조화 계수들을 더하여 하나 이상의 V^T_{DIST} 벡터들 또는 하나 이상의 $V^T_{Q_DIST}$ 벡터들 중의 제거된 적어도 하나의 벡터를 배경 구면 조화 계수들과 재결합하도록 구성되는, 디바이스.

[0899] 조항 133149-16A. 조항 133149-14A 의 디바이스에 있어서, 하나 이상의 프로세서들은 또한, 배경 구면 조화 계수들을 하나 이상의 U_{BG} 벡터들, 하나 이상의 S_{BG} 벡터들, 및 하나 이상의 V^T_{BG} 의 함수로서 결정하고, 배경 구면 조화 계수들에 대하여 오더 감소를 수행하여 N_{BG} 와 동일한 오더를 갖는 감소된 배경 구면 조화 계수들을 생성하고, 하나 이상의 U_{DIST} 에 하나 이상의 S_{DIST} 벡터들을 곱하여 하나 이상의 $U_{DIST} * S_{DIST}$ 벡터들을 생성하고, 하나 이상의 $U_{DIST} * S_{DIST}$ 벡터들을 리오더링하여 리오더링된 하나 이상의 $U_{DIST} * S_{DIST}$ 벡터들을 생성하고, 하나 이상의 벡터들로부터 하나 이상의 벡터들 중의 결정된 적어도 하나의 벡터를 제거하여 하나 이상의 벡터들의 감소된 세트를 생성하고, 리오더링된 하나 이상의 $U_{DIST} * S_{DIST}$ 벡터들에 하나 이상의 V^T_{DIST} 벡터들 또는 하나 이상의 $V^T_{Q_DIST}$ 벡터들 중의 제거된 적어도 하나의 벡터를 곱하여 중간 구별되는 구면 조화 계수들을 생성하며, 배경 구면 조화 계수에 중간 구별되는 구면 조화 계수들을 더하여 하나 이상의 V^T_{DIST} 벡터들 또는 하나 이상의 $V^T_{Q_DIST}$ 벡터들 중의 제거된 적어도 하나의 벡터를 배경 구면 조화 계수들과 재결합하도록 구성되는, 디바이스.

[0900] 조항 133149-17A. 조항 133149-15A 또는 조항 133149-16A 중 어느 하나의 디바이스에 있어서, 하나 이상의 프로세서들은 또한, 배경 구면 조화 계수들에 중간 구별되는 구면 조화 계수들을 더한 후 배경 구면 조화 계수들을 오디오 인코딩하며, 오디오 인코딩된 배경 구면 조화 계수들을 포함하기 위한 비트스트림을 생성하도록 구성되는, 디바이스.

[0901] 조항 133149-18A. 조항 133149-1A 의 디바이스에 있어서, 하나 이상의 프로세서들은 또한, 복수의 구면 조화 계수들에 대하여 특이값 분해를 수행하여 복수의 구면 조화 계수들의 좌-특이 벡터들을 나타내는 U 행렬, 복수의 구면 조화 계수들의 특이값들을 나타내는 S 행렬 및 복수의 구면 조화 계수들의 우-특이 벡터들을 나타내는 V 행렬을 생성하고, V 행렬의 트랜스포즈의 하나 이상의 V^T_{DIST} 벡터들 및 하나 이상의 V^T_{BG} 를 결정하는 것으로서, 하나 이상의 V_{DIST} 벡터들은 음장의 하나 이상의 구별되는 성분들을 기술하며 하나 이상의 V_{BG} 벡터들은 음장의 하나 이상의 배경 성분들을 기술하는, 상기 하나 이상의 V^T_{DIST} 벡터들 및 하나 이상의 V^T_{BG} 를 결정하고, 하나 이상의 V^T_{DIST} 벡터들을 양자화하여 하나 이상의 $V^T_{Q_DIST}$ 벡터들을 생성하며, 하나 이상의 $V^T_{Q_DIST}$ 벡터들을 리오더링하여 리오더링된 하나 이상의 $V^T_{Q_DIST}$ 벡터들을 생성하도록 구성되며, 하나 이상의 벡터들은 리오더링된 하나 이상의 $V^T_{Q_DIST}$ 벡터들을 포함하는, 디바이스.

[0902] 도 40f 는 2 차원 또는 3 차원 음장들을 기술하는 구면 조화 계수들을 압축하기 위해 본 개시물에서 설명된 기술들의 다양한 양태들을 수행할 수도 있는 예시적인 오디오 인코딩 디바이스 (510F) 를 예시하는 블록 다이어그램이다. 오디오 인코딩 디바이스 (510F) 는 오디오 인코딩 디바이스 (510F) 가 오디오 압축 유닛 (512), 오디오 인코딩 유닛 (514) 및 비트스트림 생성 유닛 (516) 을 포함한다는 점에서 오디오 인코딩 디바이스 (510C) 와 유사할 수도 있다. 더욱이, 오디오 인코딩 디바이스 (510F) 의 오디오 압축 유닛 (512) 은 오디오 압축 유닛 (512) 이 오디오 인코딩 디바이스 (510C) 의 동일한 유닛들과 유사하게 동작할 수도 있는 분해 유닛 (518) 및 벡터 리오더 유닛 (532) 을 포함한다는 점에서 오디오 인코딩 디바이스 (510C) 의 것과 유사할 수도 있다.

일부 예들에서, 오디오 인코딩 디바이스 (510F) 는 도 40d 및 도 40e 에 대하여 설명한 바와 같이, 양자화 유닛 (534) 을 포함하여, U_{DIST} 벡터들 (525C), U_{BG} 벡터들 (525D), V^T_{DIST} 벡터들 (525E), 및 V^T_{BG} 벡터들 (525J) 중 임의의 하나 이상의 벡터들을 양자화할 수도 있다.

[0903] 그러나, 오디오 인코딩 디바이스 (510F) 의 오디오 압축 유닛 (512) 은 음장 성분 추출 유닛 (520) 의 돌출 성분 분석 유닛 (524) 이 도 40a 내지 도 40j 의 맥락에서 D 로서 표기된 전경 성분들의 수를 선택하기 위해 컨텐

츠 분석을 수행할 수도 있다는 점에서 오디오 인코딩 디바이스 (510C) 의 오디오 압축 유닛 (512) 과는 상이할 수도 있다. 즉, 돌출 성분 분석 유닛 (524) 은 구면 조화 계수들의 분해된 버전들이 합성 오디오 오브젝트들로부터 또는 마이크로폰으로의 내추럴 레코딩으로부터 생성되었는지의 여부를 식별하기 위해 U, S 및 V 행렬들 (519) 에 대하여 상기 설명된 방식으로 동작할 수도 있다. 돌출 성분 분석 유닛 (524) 은 그 후 이 합성 결정에 기초하여 D 를 결정할 수도 있다.

[0904] 더욱이, 오디오 인코딩 디바이스 (510F) 의 오디오 압축 유닛 (512) 은 음장 성분 추출 유닛 (520) 이 추가적인 유닛, 오더 감소 및 에너지 보존 유닛 (528F) ("ORDER RED. AND ENERGY PRSV. UNIT (528F)" 으로서 예시됨) 을 포함할 수도 있다는 점에서 오디오 인코딩 디바이스 (510C) 의 오디오 압축 유닛 (512) 과는 상이할 수도 있다. 이런 이유들로, 오디오 인코딩 디바이스 (510F) 의 음장 성분 추출 유닛 (520) 은 "음장 성분 추출 유닛 (520F)" 으로서 표기된다.

[0905] 오더 감소 및 에너지 보존 유닛 (528F) 은 풀 V_{BG} 행렬 (525H) 에 의해 부분적으로 기술된 음장의 전체 에너지 (및 수반되는 음압) 를 보존하면서 복수의 구면 조화 계수들 (511) 의 우-특이 벡터들을 나타내는 V_{BG} 행렬 (525H) 의 배경 성분들의 오더 감소를 수행하도록 구성된 유닛을 표현한다. 이 점에 있어서, 오더 감소 및 에너지 보존 유닛 (528F) 은 도 4 의 예에 도시된 오디오 인코딩 디바이스 (20) 의 배경 선택 유닛 (48) 및 에너지 보상 유닛 (38) 에 대하여 상기 설명된 동작들과 유사한 동작들을 수행할 수도 있다.

[0906] 풀 V_{BG} 행렬 (525H) 은 차원성 $(N+1)^{2} \times (N+1)^{2} - D$ 를 갖고, 여기서 D 는 음장의 구별되는 오디오 성분들인 관점에서 돌출인 것으로 결정되는 주성분들, 또는 다른 말로 특이값들의 수를 표현한다. 즉, 풀 V_{BG} 행렬 (525H) 은 음장의 배경 (BG) 또는 다른 말로, 주변 또는 비구별되는 오디오 성분들인 것으로 결정되는 그 특이값들을 포함한다.

[0907] 예를 들어, 도 40b 내지 도 40e 의 오더 감소 유닛 (524) 에 대하여 상기 설명한 바와 같이, 오더 감소 및 에너지 보존 유닛 (528F) 은 V_{BG} 행렬 (525H) 의 배경 특이값들 중 상위 오더의 구면 기저 함수들에 대응하는 그 것들을 (종종 제로 아웃함으로써) 제거하거나, 없애거나 또는 다르게는 삭제할 수도 있다. 오더 감소 및 에너지 보존 유닛 (528F) 은 V_{BG} 행렬 (525H) 의 감소된 버전 ("V_{BG}' 행렬 (525I)" 로 표기되고 이하 "감소된 V_{BG}' 행렬 (525I)" 로 지칭됨) 을 트랜스포즈 유닛 (522) 에 출력할 수도 있다. 감소된 V_{BG}' 행렬 (525I) 은 차원성 $(\tilde{\eta}+1)^{2} \times (N+1)^{2} - D$ 을 가질 수도 있으며, 여기서 $\tilde{\eta} < N$ 이다. 트랜스포즈 유닛 (522) 은 감소된 V_{BG}' 행렬 (525I) 에 트랜스포즈 동작을 적용하여 트랜스포즈된 감소된 V^{T}_{BG}' 행렬 (525J) 을 생성하여 계산 유닛 (526) 에 출력하고, 계산 유닛 (526) 은 U_{BG} 행렬 (525D), S_{BG} 행렬 (525B), 및 트랜스포즈된 감소된 V^{T}_{BG}' 행렬 (525J) 을 이용하여 $U_{BG}*S_{BG}*V^{T}_{BG}$ 를 컴퓨팅함으로써 음장의 배경 사운드 성분들을 복원하도록 동작할 수도 있다.

[0908] 본 개시물에서 설명된 기술들에 따르면, 오더 감소 및 에너지 보존 유닛 (528F) 은 또한, 풀 V_{BG} 행렬 (525H) 의 오더를 감소시켜 감소된 V_{BG}' 행렬 (525I) 을 생성함으로써 야기된 음장의 배경 사운드 성분들의 전체 에너지에 있어서의 가능한 감소들을 보상하도록 구성된다. 일부 예들에서, 오더 감소 및 에너지 보존 유닛 (528F) 은 감소된 V_{BG}' 행렬 (525I) 을 트랜스포즈 유닛 (522) 에 출력하기 이전에, 풀 V_{BG} 행렬 (525H) 의 RMS (root mean-squared) 와 동일하거나 적어도 더 가깝게 근사하도록 감소된 V_{BG}' 행렬 (525I) 의 RMS 에너지를 증가시키기 위하여 감소된 V_{BG}' 행렬 (525I) 의 $(N+1)^{2} - D$ 컬럼들 각각에 적용할 증폭 값들의 형태의 보상 이득을 결정함으로써 보상한다.

[0909] 일부 경우들에서, 오더 감소 및 에너지 보존 유닛 (528F) 은 풀 V_{BG} 행렬 (525H) 의 각각의 컬럼의 RMS 에너지 및 감소된 V_{BG}' 행렬 (525I) 의 각각의 컬럼의 RMS 에너지를 결정한 후, 다음의 식에서 나타낸 바와 같이, 컬럼에 대한 증폭 값을 전자 대 후자의 비율로서 결정할 수도 있다 :

[0910] $$\propto = v_{BG} / v_{BG}'$$

[0911] 여기서 $\propto$ 는 컬럼에 대한 증폭 값이고, v_{BG} 는 V_{BG} 행렬 (525H) 의 단일 컬럼을 표현하며, v_{BG}' 는 V_{BG}' 행렬 (525I) 의 대응하는 단일 컬럼을 표현한다. 이것은 다음과 같이 행렬 표현으로 표현될 수도 있으며 :

$$A = V_{BG}{}^{RMS} / V_{BG}{}'^{RMS},$$

$$A = [\propto_1 \ldots \propto_{(N+1)^2 - D}]$$

[0912]

[0913] 여기서 $V_{BG}{}^{RMS}$ 는 V_{BG} 행렬 (525H) 의 각각의 컬럼의 RMS 를 표기하는 엘리먼트들을 갖는 RMS 벡터이고, $V_{BG}{}'^{RMS}$ 는 감소된 V_{BG}' 행렬 (525I) 의 각각의 컬럼의 RMS 를 표기하는 엘리먼트들을 갖는 RMS 벡터이며, A 는 V_{BG} 행렬 (525H) 의 각각의 컬럼에 대한 엘리먼트들을 갖는 증폭 값 벡터이다. 오더 감소 및 에너지 보존 유닛 (528F) 은 대응하는 증폭 값 $\propto$ 를 이용하여, 또는 벡터 형태 :

$$V_{BG}'' = V_{BG}' A^T$$

[0914]

[0915] 로 감소된 V_{BG} 행렬 (525I) 의 각각의 컬럼에 스칼라 곱을 적용하며, 여기서 V_{BG}'' 는 에너지 보상을 포함하는 감소된 V_{BG}' 행렬 (525I) 을 표현한다. 오더 감소 및 에너지 보존 유닛 (528F) 은 감소된 V_{BG}' 행렬 (525I) 의 RMS 를 풀 V_{BG} 행렬 (525H) 의 RMS 와 등화 (또는 거의 등화) 시키기 위해 트랜스포즈 유닛 (522) 에 에너지 보상을 포함하는 감소된 V_{BG}' 행렬 (525I) 을 출력할 수도 있다. 에너지 보상을 포함하는 감소된 V_{BG}' 행렬 (525I) 의 출력 차원성은 $(\tilde{\eta}+1)^2 \mathrm{x}\ (N+1)^2 - D$ 일 수도 있다.

[0916] 일부 예들에서, 감소된 V_{BG}' 행렬 (525I) 및 풀 V_{BG} 행렬 (525H) 의 개별의 컬럼들의 각각의 RMS 를 결정하기 위해, 오더 감소 및 에너지 보존 유닛 (528F) 은 먼저 컬럼들에 참조 구면 조화 계수들 (SHC) 렌더러를 적용할 수도 있다. 오더 감소 및 에너지 보존 유닛 (528F) 에 의한 참조 SHC 렌더러의 적용은 감소된 V_{BG}' 행렬 (525I) 및 풀 V_{BG} 행렬 (525H) 에 의해 표현된 프레임의 각각의 컬럼에 의해 기술된 전체 음장의 에너지를 결정하기 위해 SHC 도메인에서의 RMS 의 결정을 허용한다. 따라서, 이러한 예들에서, 오더 감소 및 에너지 보존 유닛 (528F) 은 참조 SHC 렌더러를 풀 V_{BG} 행렬 (525H) 의 각각의 컬럼에 그리고 감소된 V_{BG}' 행렬 (525I) 의 각각의 감소된 컬럼에 적용하고, 컬럼 및 감소된 컬럼에 대한 개별의 RMS 값들을 결정하며, 컬럼에 대한 증폭 값을 컬럼에 대한 RMS 값 대 감소된 컬럼에 대한 RMS 값의 비율로서 결정할 수도 있다. 일부 예들에서, 감소된 V_{BG}' 행렬 (525I) 에 대한 오더 감소는 에너지 보존에 부합하는 컬럼-방향으로 진행된다. 이것은 다음과 같이 의사코드로 표현될 수도 있다 :

```
R = ReferenceRenderer;
for m = numDist+1 : numChannels
   fullV = V(:,m); //takes one column of V => fullV
   reducedV =[fullV(1:numBG); zeros(numChannels-numBG,1)];
   alpha=sqrt( sum((fullV'*R).^2)/sum((reducedV'*R).^2) );
   if isnan(alpha) || isinf(alpha), alpha = 1; end;
   V_out(:,m) = reducedV * alpha;
end
```

[0917]

[0918] 상기 의사코드에서, numChannels 는 $(N+1)^2$ - D 를 표현할 수도 있고, numBG 는 $(\tilde{\eta}+1)^2$ 를 표현할 수도 있고, V 는 V_{BG} 행렬 (525H) 을 표현할 수도 있고, V_out 은 감소된 V_{BG}' 행렬 (525I) 을 표현할 수도 있으며, R 은 오더 감소 및 에너지 보존 유닛 (528F) 의 참조 SHC 렌더러를 표현할 수도 있다. V 의 차원성은 $(N+1)^2 \times (N+1)^2$ - D 일 수도 있고, V_out 의 차원성은 $(\tilde{\eta}+1)^2 \text{ x } (N+1)^2 - D$ 일 수도 있다.

[0919] 그 결과, 오디오 인코딩 디바이스 (510F) 는, 복수의 구면 조화 계수들 (511) 을 표현할 때, 오더 감소 프로세스에 대한 결과로서 손실될 수도 있는 에너지에 대한 보상을 포함하는 오더-감소된 V_{BG}' 행렬 (525I) 을 이용하여 배경 사운드 성분들을 복원할 수도 있다.

[0920] 도 40g 는 2 차원 또는 3 차원 음장들을 기술하는 구면 조화 계수들을 압축하기 위해 본 개시물에서 설명된 기술들의 다양한 양태들을 수행할 수도 있는 예시적인 오디오 인코딩 디바이스 (510G) 를 예시하는 블록 다이어그램이다. 도 40g 의 예에서, 오디오 인코딩 디바이스 (510G) 는 음장 성분 추출 유닛 (520F) 을 포함한다. 결국, 음장 성분 추출 유닛 (520F) 은 돌출 성분 분석 유닛 (524G) 을 포함한다.

[0921] 그러나, 오디오 인코딩 디바이스 (510G) 의 오디오 압축 유닛 (512) 은 오디오 인코딩 디바이스 (510G) 의 오디오 압축 유닛 (512) 이 돌출 성분 분석 유닛 (524G) 을 포함한다는 점에서 오디오 인코딩 디바이스 (10F) 의 오디오 압축 유닛 (512) 과는 상이할 수도 있다. 돌출 성분 분석 유닛 (524G) 은 음장을 나타내는 오디오 데이터와 연관된 방향-기반 정보를 이용하여, 그 오디오 데이터의 돌출성 (saliency) 또는 구별성 (distinctness) 을 결정하도록 구성된 유닛을 표현할 수도 있다.

[0922] 에너지 기반 결정들은 음장의 구별되는 오디오 성분들을 식별하기 위해 SVD 에 의해 분해된 음장의 렌더링을 개선시킬 수도 있지만, 에너지 기반 결정들은 또한, 디바이스로 하여금, 배경 오디오 성분들을 구별되는 오디오 성분들로서 잘못 식별하게 할 수도 있으며, 이 경우들에서 배경 오디오 성분들은 높은 에너지 레벨을 보인다. 즉, 에너제틱한 (예를 들어, 소리가 더 큰 (louder)) 배경 오디오 성분들이 구별되는 오디오 성분들인 것으로 부정확하게 식별될 수도 있기 때문에, 구별되는 및 배경 오디오 성분들의 단독 에너지 기반 분리는 강인하지 않을 수도 있다. 음장의 구별되는 및 배경 오디오 성분들 간을 더 강인하게 구별하기 위해, 본 개시물에서 설명된 기술들의 다양한 양태들은 돌출 성분 분석 유닛 (524G) 이 SHC (511) 의 방향-기반 분석을 수행하여 SHC (511) 의 분해된 버전들로부터 구별되는 및 배경 오디오 성분들을 분리하는 것을 가능하게 할 수도 있다.

[0923] 돌출 성분 분석 유닛 (524G) 은 도 40h 의 예에서, 이전에 설명된 오디오 인코딩 디바이스들 (510 내지 510F) 의 돌출 성분 분석 유닛들 (524) 과 유사하게, V 행렬 (519), S 행렬 (518B), 및 U 행렬 (519C) 중 하나 이상에 포함된 배경 엘리먼트들로부터 구별되는 (또는 전경) 엘리먼트들을 분리하도록 구성되거나 또는 다르게는 동작가능한 유닛을 표현할 수도 있다. 일부 SVD 기반 기술들에 따르면, 가장 에너지틱한 성분들 (예를 들어, V, S 및 U 행렬들 (519 내지 519C) 중 하나 이상의 행렬들, 또는 그로부터 유도된 행렬의 처음의 몇몇 벡터들) 은 구별되는 성분들로서 취급될 수도 있다. 그러나, 하나 이상의 행렬들 (519 내지 519C) 중 가장 에너제틱한 성분들 (이는 벡터들로 표현된다) 은 모든 시나리오들에서, 가장 방향성인 성분들/신호들을 표현하지 않을 수도 있다.

[0924] 이전에 설명된 돌출 성분 분석 유닛들 (524) 과 달리, 돌출 성분 분석 유닛 (524G) 은 하나 이상의 행렬들 (519 내지 519C) 또는 그로부터 유도된 행렬의 벡터들의 방향성에 기초하여 전경 엘리먼트들을 식별하기 위해 본 개시물에서 설명된 기술들의 하나 이상의 양태들을 구현할 수도 있다. 일부 예들에서, 돌출 성분 분석 유닛 (524G) 은 벡터들의 에너지 및 방향성 양자에 기초하여 하나 이상의 벡터들을 구별되는 오디오 성분들 (그 성분들은 또한 "오브젝트들" 로 지칭될 수도 있다) 로서 식별 또는 선택할 수도 있다. 예를 들어, 돌출 성분 분석 유닛 (524G) 은 고 에너지 및 고 방향성 (예를 들어, 방향성 지수 (directionality quotient) 로서 표현됨) 을 구별되는 오디오 성분들로서 디스플레이하는 하나 이상의 행렬들 (519 내지 519C) 의 그 벡터들을 식별할 수도 있다. 그 결과, 돌출 성분 분석 유닛 (524G) 이 특정 벡터가 하나 이상의 행렬들 (519 내지 519C) (또는 그로부터 유도된 행렬) 의 다른 벡터들과 비교할 때 상대적으로 덜 방향성이라고 결정한다면, 그 특정 벡터와 연관된 에너지 레벨에 상관없이, 돌출 성분 분석 유닛 (524G) 은 특정 벡터가 SHC (511) 에 의해 표현된 음장의 배경 (또는 주변) 오디오 성분들을 표현한다고 결정할 수도 있다. 이 점에 있어서, 돌출 성분 분석 유닛 (524G) 은 도 4 의 예에 도시된 오디오 인코딩 디바이스 (20) 의 음장 분석 유닛 (44) 에 대하여 상기 설명된 동작들과 유사한 동작들을 수행할 수도 있다.

[0925] 일부 구현들에서, 돌출 성분 분석 유닛 (524G) 은 다음의 동작들을 수행함으로써, 방향성에 기초하여 구별되는 오디오 오브젝트들 (상기 언급한 바와 같이, "성분들" 로도 지칭될 수도 있다) 을 식별할 수도 있다. 돌출 성분 분석 유닛 (524G) 은 (예를 들어 하나 이상의 행렬 곱셈 프로세스들을 이용하여) V 행렬 (519A) 에 S 행렬 (519B) 를 곱할 수도 있다. V 행렬 (519A) 과 S 행렬 (519B) 을 곱함으로써, 돌출 성분 분석 유닛 (524G) 은 VS 행렬을 획득할 수도 있다. 추가적으로, 돌출 성분 분석 유닛 (524G) 은 VS 행렬의 (로우일 수도 있는) 벡터들 각각의 엔트리들 중 적어도 일부를 제곱 (즉, 2 의 거듭제곱에 의한 지수효과) 할 수도 있다. 일부 경우들에서, 돌출 성분 분석 유닛 (524G) 은 1 보다 더 큰 오더와 연관되는 각각의 벡터의 그 제곱 엔트리들을 합산할 수도 있다. 하나의 예로서, 행렬의 각각의 벡터가 25 개의 엔트리들을 포함한다면, 돌출 성분 분석 유닛 (524G) 은 각각의 벡터에 대하여, 제 5 엔트리에서 시작하고 제 25 엔트리에서 끝나는 각각의 벡터의 엔트리들을 제곱하여, 제곱 엔트리들을 합산하여 방향성 지수 (또는 방향성 표시자) 를 결정할 수도 있다. 각각의 합산 연산은 대응하는 벡터에 대한 방향성 지수를 초래할 수도 있다. 이 예에서, 돌출 성분 분석 유닛 (524G) 은 1 이하인 오더와 연관되는 각각의 로우의 그 엔트리들, 즉 제 1 엔트리 내지 제 4 엔트리가 에너지의 양에 더 일반적으로 관련되고 그 엔트리들의 방향성에 덜 관련된다고 결정할 수도 있다. 즉, 0 또는 1 의 오더와 연관된 하위 오더의 앰비소닉스는 도 1 및 도 2 에 예시한 바와 같이, 압력파의 방향의 관점에서 많이 제공하지 않고, 오히려 (에너지를 나타내는) 일부 볼륨을 제공하는 구면 기저 함수들에 대응한다.

[0926] 상기 예에서 설명된 동작들은 또한 다음의 의사-코드에 따라 표현될 수도 있다. 이하의 의사-코드는 문자열들 "/*" 및 "*/" (인용구들 없음) 의 연속적인 인스턴스들 내에 포함되는 주석문 형태의 주석들을 포함한다.

```
[U,S,V] = svd(audioframe,'ecom');
VS = V*S;
```

[0927]

[0928] /*다음 라인은 각각의 로우를 독립적으로 분석하는 것, 및 대응하는 벡터에 대한 방향성 지수 또는 방향성 메트릭을 결정하기 위해 제 5 엔트리로부터 제 25 엔트리까지의 제 1 (하나의 예로서) 로우의 값들을 합산하는 것과 관련된다. 합산하기 전에는 그 엔트리들을 제곱한다. 1 보다 더 큰 오더와 연관되는 각각의 로우에서의 엔트리들은 고차 앰비소닉스와 연관되고, 따라서 방향성이 있을 가능성이 더 크다.*/

```
sumVS = sum(VS(5:end,:).^2,1);
```

[0929]

[0930] /* 다음 라인은 생성된 VS 행렬에 대한 제곱들의 합을 분류하는 것, 및 가장 큰 값들의 세트 (예를 들어, 3 개 또는 4 개의 가장 큰 값들) 를 선택하는 것과 관련된다*/

```
[~,idxVS] = sort(sumVS,'descend');
U = U(:,idxVS);
V = V(:,idxVS);
S = S(idxVS,idxVS);
```

[0931]

[0932] 즉, 상기 의사-코드에 따르면, 돌출 성분 분석 유닛 (524G) 은 SHC (511) 중 1 보다 더 큰 오더를 갖는 구면 기저 함수에 대응하는 그 것들로부터 분해된 VS 행렬의 각각의 벡터의 엔트리들을 선택할 수도 있다. 돌출 성분 분석 유닛 (524G) 은 그 후 VS 행렬의 각각의 벡터에 대한 이들 엔트리들을 제곱하여, 제곱 엔트리들을 합산하여, VS 행렬의 각각의 벡터에 대한 방향성 메트릭 또는 지수를 식별, 컴퓨팅 또는 다르게는 결정할 수도 있다. 다음에, 돌출 성분 분석 유닛 (524G) 은 벡터들 각각의 개별의 방향성 메트릭들에 기초하여 VS 행렬의 벡터들을 분류할 수도 있다. 돌출 성분 분석 유닛 (524G) 은 최고의 대응하는 방향성을 가진 그 벡터들이 처음이고 최저의 대응하는 방향성을 가진 그 벡터들이 마지막이 되도록 이들 벡터들을 방향성 메트릭들의 내림차순으로 분류할 수도 있다. 돌출 성분 분석 유닛 (524G) 은 그 후 최고의 상대적 방향성 메트릭을 갖는 벡터들의 비-제로 서브세트를 선택할 수도 있다.

[0933] 본 개시물에서 설명된 기술들의 일부 양태들에 따르면, 오디오 인코딩 디바이스 (510G), 또는 이들의 하나 이상의 성분들은 VS 행렬의 미리 결정된 수의 벡터들을 구별되는 오디오 성분들로서 식별 또는 다르게는 이용할 수도 있다. 예를 들어, VS 행렬의 각각의 로우의 엔트리 5 내지 엔트리 25 를 선택하고 선택된 엔트리들을 제

곱 및 합산하여 각 개별의 벡터에 대한 상대적 방향성 메트릭을 결정한 후, 돌출 성분 분석 유닛 (524G) 은 구별되는 오디오 성분들을 나타내는 벡터들을 식별하기 위해 벡터들 중에서의 추가 선택을 구현할 수도 있다. 일부 예들에서, 돌출 성분 분석 유닛 (524G) 은 벡터들의 방향성 지수들을 비교함으로써, VS 행렬의 미리 결정된 수의 벡터들을 선택할 수도 있다. 하나의 예로서, 돌출 성분 분석 유닛 (524G) 은 4 개의 최고의 방향성 지수들을 갖는 (그리고 분류된 VS 행렬의 처음의 4 개의 벡터들인) VS 행렬에서 표현된 4 개의 벡터들을 선택할 수도 있다. 결국, 돌출 성분 분석 유닛 (524G) 은 4 개의 선택된 벡터들이 음장의 대응하는 SHC 표현과 연관된 4 개의 가장 구별되는 오디오 오브젝트들을 표현한다고 결정할 수도 있다.

[0934] 일부 예들에서, 돌출 성분 분석 유닛 (524G) 은 상기 설명한 바와 같이, 4 개의 선택된 벡터들의 구별성을 반영하기 위해 VS 행렬로부터 유도된 벡터들을 리오더링할 수도 있다. 하나의 예에서, 돌출 성분 분석 유닛 (524G) 은, 4 개의 선택된 엔트리들이 VS 행렬의 상부로 리로케이팅되도록 벡터들을 리오더링할 수도 있다. 예를 들어, 돌출 성분 분석 유닛 (524G) 은 4 개의 선택된 엔트리들 모두가 결과의 리오더링된 VS 행렬의 제 1 (또는 가장 높은) 로우에 포지셔닝되도록 VS 행렬을 수정할 수도 있다. 돌출 성분 분석 유닛 (524G) 에 대하여 본 명세서에서 설명하였지만, 다양한 구현들에서는, 오디오 인코딩 디바이스 (510G) 의 다른 성분들, 이를테면 벡터 리오더 유닛 (532) 이 리오더링을 수행할 수도 있다.

[0935] 돌출 성분 분석 유닛 (524G) 은 결과의 행렬 (즉, VS 행렬, 이는 경우에 따라 리오더링되거나 리오더링되지 않음) 을 비트스트림 생성 유닛 (516) 에 통신할 수도 있다. 결국, 비트스트림 생성 유닛 (516) 은 VS 행렬 (525K) 을 이용하여 비트스트림 (517) 을 생성할 수도 있다. 예를 들어, 돌출 성분 분석 유닛 (524G) 이 VS 행렬 (525K) 을 리오더링했다면, 비트스트림 생성 유닛 (516) 은 VS 행렬 (525K) 의 리오더링된 버전의 나머지 벡터들을 양자화 또는 폐기하는 것에 의해서와 같이, VS 행렬 (525K) 의 리오더링된 버전의 상부 로우를 구별되는 오디오 오브젝트들로서 이용할 수도 있다. VS 행렬 (525K) 의 리오더링된 버전의 나머지 벡터들을 양자화함으로써, 비트스트림 생성 유닛 (516) 은 나머지 벡터들을 주변 또는 배경 오디오 데이터로서 취급할 수도 있다.

[0936] 돌출 성분 분석 유닛 (524G) 이 VS 행렬 (525K) 을 리오더링하지 않았다는 예들에서, 비트스트림 생성 유닛 (516) 은 돌출 성분 분석 유닛 (524G) 에 의해 선택한 바와 같이, VS 행렬 (525K) 의 각각의 로우의 특정 엔트리들 (예를 들어, 제 5 엔트리 내지 제 25 엔트리) 에 기초하여, 구별되는 오디오 데이터를 배경 오디오 데이터와 구별할 수도 있다. 예를 들어, 비트스트림 생성 유닛 (516) 은 VS 행렬 (525K) 의 각각의 로우의 처음의 4 개의 엔트리들을 양자화 또는 폐기하는 것에 의해 비트스트림 (517) 을 생성할 수도 있다.

[0937] 이 방식으로, 오디오 인코딩 디바이스 (510G) 및/또는 그의 성분들, 이를테면 돌출 성분 분석 유닛 (524G) 은 음장을 나타내는 배경 오디오 데이터와 구별되는 오디오 오브젝트들 간을 구별하기 위하여, 오디오 데이터의 더 높은 및 더 낮은 계수들의 에너지들의 비율들을 결정 또는 다르게는 활용하기 위해 본 개시물의 기술들을 구현할 수도 있다. 예를 들어, 설명한 바와 같이, 돌출 성분 분석 유닛 (524G) 은 돌출 성분 분석 유닛 (524H) 에 의해 생성된 VS 행렬 (525K) 의 다양한 엔트리들의 값들에 기초하여 에너지 비율들을 활용할 수도 있다. V 행렬 (519A) 및 S 행렬 (519B) 에 의해 제공된 데이터를 결합함으로서, 돌출 성분 분석 유닛 (524G) 은 벡터들 및 관련 데이터 (예를 들어, 방향성 지수들) 의 형태로, 오디오 데이터의 다양한 성분들의 전체 에너지 및 방향성 양자에 대한 정보를 제공하기 위해 VS 행렬 (525K) 을 생성할 수도 있다. 보다 구체적으로, V 행렬 (519A) 은 방향성 결정들과 관련된 정보를 제공할 수도 있는 한편, S 행렬 (519B) 은 오디오 데이터의 성분들에 대한 전체 에너지 결정들과 관련된 정보를 제공할 수도 있다.

[0938] 다른 예들에서, 돌출 성분 분석 유닛 (524G) 은 리오더링된 V^{T}_{DIST} 벡터들 (539) 을 이용하여 VS 행렬 (525K) 을 생성할 수도 있다. 이들 예들에서, 돌출 성분 분석 유닛 (524G) 은 S 행렬 (519B) 에 기초한 임의의 수정 이전에, V 행렬 (519) 에 기초한 구별성을 결정할 수도 있다. 즉, 이들 예들에 따르면, 돌출 성분 분석 유닛 (524G) 은 VS 행렬 (525K) 을 생성하는 단계를 수행하지 않고, 단지 V 행렬 (519) 만을 이용하여 방향성을 결정할 수도 있다. 보다 구체적으로, V 행렬 (519A) 은 오디오 데이터의 성분들 (예를 들어, V 행렬 (519) 의 벡터들) 이 혼합되는 방식에 대한 정보, 및 잠재적으로는, 벡터들에 의해 전달된 데이터의 다양한 시너지 효과들에 대한 정보를 제공할 수도 있다. 예를 들어, V 행렬 (519A) 은 EigenMike® 에 의해 오디오 인코딩 디바이스 (510G) 로 중계된 바와 같이, 각각의 오디오 성분의 도달 방향과 같은, 벡터들에 의해 표현되는 다양한 오디오 성분들의 "도달 방향" 에 대한 정보를 제공할 수도 있다. 본 명세서에서 사용한 바와 같이, 용어 "오디오 데이터의 성분" 은 행렬들 (519) 또는 그로부터 유도된 임의의 행렬들 중 임의의 것의 "엔트리" 로 상

호교환가능하게 사용될 수도 있다.

[0939] 본 개시물의 기술들의 일부 구현들에 따르면, 돌출 성분 분석 유닛 (524G) 은 본 명세서에서 설명된 다양한 결정들을 행하기 위해 SHC 표현들을 외부 정보로 보충 또는 보강할 수도 있다. 하나의 예로서, 돌출 성분 분석 유닛 (524G) 은 행렬들 (519 내지 519C) 에서 표현된 다양한 오디오 성분들의 돌출성을 결정하기 위하여 SHC 를 외부 정보로 보강할 수도 있다. 다른 예로서, 돌출 성분 분석 유닛 (524G) 및/또는 벡터 리오더 유닛 (532) 은 구별되는 오디오 오브젝트들 및 배경 오디오 데이터 간을 구별하기 위해 HOA 를 외부 데이터로 보강할 수도 있다.

[0940] 일부 예들에서, 돌출 성분 분석 유닛 (524G) 은 오디오 데이터의 부분들 (예를 들어, 구별되는 오디오 오브젝트들) 이 Keynesian 에너지를 디스플레이한다는 것을 검출할 수도 있다. 이러한 구별되는 오브젝트들의 예는 변조되는 인간 음성과 연관될 수도 있다. 변조되는 음성-기반 오디오 데이터의 경우에, 돌출 성분 분석 유닛 (524G) 은 변조 데이터의 에너지가 나머지 성분들의 에너지들에 대한 비율로서, 시간의 경과에 따라 계속 대략 일정 (예를 들어, 임계 범위 내에 일정) 하거나 대략 변하지 않는다고 결정할 수도 있다. 종래에, Keynesian 에너지를 가진 구별되는 오디오 성분들의 에너지 특성들 (예를 들어, 변조 음성과 연관된 것들) 이 하나의 오디오 프레임으로부터 다른 것으로 바뀌면, 디바이스는 일련의 오디오 성분들을 단일 신호로서 식별가능하지 않을 수도 있다. 그러나, 돌출 성분 분석 유닛 (524G) 은 다양한 행렬들에서 벡터로서 표현된 구별되는 오브젝트의 애퍼처 또는 방향성을 결정하기 위해 본 개시물의 기술들을 구현할 수도 있다.

[0941] 보다 구체적으로, 돌출 성분 분석 유닛 (524G) 은 방향성 및/또는 애퍼처와 같은 특성들이 실질적으로 오디오 프레임에 걸쳐 바뀔 가능성이 없다고 결정할 수도 있다. 본 명세서에서 사용한 바와 같이, 애퍼처는 오디오 데이터 내에서, 상위 오더의 계수들 대 하위 오더의 계수들의 비율을 표현한다. V 행렬 (519A) 의 각각의 로우는 특정 SHC 에 대응하는 벡터들을 포함할 수도 있다. 돌출 성분 분석 유닛 (524G) 은 하위 오더의 SHC (예를 들어, 1 이하인 오더와 연관됨) 가 주변 데이터를 나타내는 경향이 있는 한편, 상위 오더의 엔트리들이 구별되는 데이터를 나타내는 경향이 있다고 결정할 수도 있다. 추가적으로, 돌출 성분 분석 유닛 (524G) 은 많은 경우들에서, 상위 오더의 SHC (예를 들어, 1 보다 더 큰 오더와 연관됨) 가 더 큰 에너지를 디스플레이하고, 상위 오더 대 하위 오더의 SHC 의 에너지 비율이 오디오 프레임 간에 실질적으로 유사하게 (또는 대략적으로 일정하게) 유지된다고 결정할 수도 있다.

[0942] 돌출 성분 분석 유닛 (524G) 의 하나 이상의 성분들은 V 행렬 (519) 을 이용하여 방향성 및 애퍼처와 같은 오디오 데이터의 특성들을 결정할 수도 있다. 이 방식으로, 돌출 성분 분석 유닛 (524G) 과 같은 오디오 인코딩 디바이스 (510G) 의 성분들은 방향-기반 정보를 이용하여, 배경 오디오로부터 구별되는 오디오 오브젝트들을 구별하고/하거나 돌출성을 결정하기 위해 본 명세서에서 설명된 기술들을 구현할 수도 있다. 돌출성 및/또는 구별성을 결정하기 위해 방향성을 이용함으로써, 돌출 성분 분석 유닛 (524G) 은 단지 에너지-기반 데이터만을 이용하여 돌출성 및/또는 구별성을 결정하도록 구성된 디바이스의 경우들에서보다 더 강인한 결정들에 이를 수도 있다. 돌출성 및/또는 구별성의 방향-기반 결정들에 대하여 위에서 설명하였지만, 돌출 성분 분석 유닛 (524G) 은 하나 이상의 행렬들 (519 내지 519C) (또는 그로부터 유도된 임의의 행렬) 의 벡터들로 표현한 바와 같이, 오디오 데이터의 특정 성분들의 돌출성 및/또는 구별성을 결정하기 위해, 방향성을 에너지와 같은 다른 특성들에 더하여 이용하기 위해 본 개시물의 기술들을 구현할 수도 있다.

[0943] 일부 예들에서, 방법은 하나 이상의 오디오 오브젝트들에 대해 결정된 방향성에 기초하여 오디오 오브젝트들과 연관된 하나 이상의 구면 조화 계수들 (SHC) 로부터 하나 이상의 구별되는 오디오 오브젝트들을 식별하는 단계를 포함한다. 하나의 예에서, 방법은 오디오 오브젝트들과 연관된 구면 조화 계수들에 기초하여 하나 이상의 오디오 오브젝트들의 방향성을 결정하는 단계를 더 포함한다. 일부 예들에서, 방법은 구면 조화 계수들에 대하여 특이값 분해를 수행하여 복수의 구면 조화 계수들의 좌-특이 벡터들을 나타내는 U 행렬, 복수의 구면 조화 계수들의 특이값들을 나타내는 S 행렬 및 복수의 구면 조화 계수들의 우-특이 벡터들을 나타내는 V 행렬을 생성하는 단계; 및 복수의 구면 조화 계수들을 U 행렬, S 행렬 및 V 행렬 중 하나 이상의 행렬의 적어도 부분의 함수로서 나타내는 단계를 더 포함하며, 하나 이상의 오디오 오브젝트들의 개별의 방향성을 결정하는 단계는 V 행렬에 적어도 부분적으로 기초한다.

[0944] 하나의 예에서, 방법은 더 큰 방향성 지수를 갖는 벡터들이 리오더링된 V 행렬에서 더 적은 방향성 지수를 갖는 벡터들 상방에 포지셔닝되도록 V 행렬의 하나 이상의 벡터들을 리오더링하는 단계를 더 포함한다. 하나의 예에서, 방법은 더 큰 방향성 지수를 갖는 벡터들이 더 적은 방향성 지수를 갖는 벡터들보다 더 큰 방향성 정보를 포함한다고 결정하는 단계를 더 포함한다. 하나의 예에서, 방법은 V 행렬에 S 행렬을 곱하여 VS 행렬을

생성하는 단계를 더 포함하며, VS 행렬은 하나 이상의 벡터들을 포함한다. 하나의 예에서, 방법은 1 보다 더 큰 오더와 연관되는 VS 행렬의 각각의 로우의 엔트리들을 선택하는 단계, 선택된 엔트리들 각각을 제곱하여 대응하는 제곱 엔트리들을 형성하는 단계, 및 VS 행렬의 각각의 로우에 대해, 제곱 엔트리들 모두를 합산하여 대응하는 벡터에 대한 방향성 지수를 결정하는 단계를 더 포함한다.

[0945] 일부 예들에서, VS 행렬의 각각의 로우는 25 개의 엔트리들을 포함한다. 하나의 예에서, 1 보다 더 큰 오더와 연관되는 VS 행렬의 각각의 로우의 엔트리들을 선택하는 단계는 VS 행렬의 각각의 로우의 제 5 엔트리에서 시작하고 VS 행렬의 각각의 로우의 제 25 엔트리에서 끝나는 모든 엔트리들을 선택하는 단계를 포함한다. 하나의 예에서, 방법은 구별되는 오디오 오브젝트들을 표현하기 위해 VS 행렬의 벡터들의 서브세트를 선택하는 단계를 더 포함한다. 일부 예들에서, 서브세트를 선택하는 단계는 VS 행렬의 4 개의 벡터들을 선택하는 단계를 포함하며, 선택된 4 개의 벡터들은 VS 행렬의 벡터들 모두의 4 개의 가장 큰 방향성 지수들을 갖는다. 하나의 예에서, 선택된 벡터들의 서브세트가 구별되는 오디오 오브젝트들을 표현한다고 결정하는 단계는 각각의 벡터의 방향성 및 에너지 양자에 기초한다.

[0946] 일부 예들에서, 방법은 하나 이상의 오디오 오브젝트들에 대해 결정된 방향성 및 에너지에 기초하여, 오디오 오브젝트들과 연관된 하나 이상의 구면 조화 계수들로부터 하나 이상의 구별되는 오디오 오브젝트들을 식별하는 단계를 포함한다. 하나의 예에서, 방법은 오디오 오브젝트들과 연관된 구면 조화 계수들에 기초하여 하나 이상의 오디오 오브젝트들의 방향성 및 에너지 중 하나 또는 양자를 결정하는 단계를 더 포함한다. 일부 예들에서, 방법은 음장을 나타내는 구면 조화 계수들에 대하여 특이값 분해를 수행하여 복수의 구면 조화 계수들의 좌-특이 벡터들을 나타내는 U 행렬, 복수의 구면 조화 계수들의 특이값들을 나타내는 S 행렬 및 복수의 구면 조화 계수들의 우-특이 벡터들을 나타내는 V 행렬을 생성하는 단계, 및 복수의 구면 조화 계수들을 U 행렬, S 행렬 및 V 행렬 중 하나 이상의 행렬의 적어도 부분의 함수로서 나타내는 단계를 더 포함하며, 하나 이상의 오디오 오브젝트들의 개별의 방향성을 결정하는 단계는 V 행렬에 적어도 부분적으로 기초하며, 하나 이상의 오디오 오브젝트들의 개별의 에너지를 결정하는 단계는 S 행렬에 적어도 부분적으로 기초한다.

[0947] 하나의 예에서, 방법은 V 행렬에 S 행렬을 곱하여 VS 행렬을 생성하는 단계를 더 포함하며, VS 행렬은 하나 이상의 벡터들을 포함한다. 일부 예들에서, 방법은 1 보다 더 큰 오더와 연관되는 VS 행렬의 각각의 로우의 엔트리들을 선택하는 단계, 선택된 엔트리들 각각을 제곱하여 대응하는 제곱 엔트리들을 형성하는 단계, 및 VS 행렬의 각각의 로우에 대해, 제곱 엔트리들 모두를 합산하여 VS 행렬의 대응하는 벡터에 대한 방향성 지수를 생성하는 단계를 더 포함한다. 일부 예들에서, VS 행렬의 각각의 로우는 25 개의 엔트리들을 포함한다. 하나의 예에서, 1 보다 더 큰 오더와 연관되는 VS 행렬의 각각의 로우의 엔트리들을 선택하는 단계는 VS 행렬의 각각의 로우의 제 5 엔트리에서 시작하고 VS 행렬의 각각의 로우의 제 25 엔트리에서 끝나는 모든 엔트리들을 선택하는 단계를 포함한다. 일부 예들에서, 방법은 구별되는 오디오 오브젝트들을 표현하기 위해 벡터들의 서브세트를 선택하는 단계를 더 포함한다. 하나의 예에서, 서브세트를 선택하는 단계는 VS 행렬의 4 개의 벡터들을 선택하는 단계를 포함하며, 선택된 4 개의 벡터들은 VS 행렬의 벡터들 모두의 4 개의 가장 큰 방향성 지수들을 갖는다. 일부 예들에서, 선택된 벡터들의 서브세트가 구별되는 오디오 오브젝트들을 표현한다고 결정하는 단계는 각각의 벡터의 방향성 및 에너지 양자에 기초한다.

[0948] 일부 예들에서, 방법은 방향-기반 정보를 이용하여, 음장의 구별되는 성분들을 기술하는 하나 이상의 제 1 벡터들 및 음장의 배경 성분들을 기술하는 하나 이상의 벡터들을 결정하는 단계를 포함하며, 하나 이상의 제 1 벡터들 및 하나 이상의 제 2 벡터들 양자는 적어도 복수의 구면 조화 계수들에 대하여 변환을 수행함으로써 생성되었다. 하나의 예에서, 변환은 복수의 구면 조화 계수들의 좌-특이 벡터들을 나타내는 U 행렬, 복수의 구면 조화 계수들의 특이값들을 나타내는 S 행렬 및 복수의 구면 조화 계수들의 우-특이 벡터들을 나타내는 V 행렬을 생성하는 특이값 분해를 포함한다. 하나의 예에서, 변환은 음장의 구별되는 성분들 및 음장의 배경 성분들을 식별하기 위해 주성분 분석을 포함한다.

[0949] 일부 예들에서, 디바이스는 본 명세서에서 설명된 기술들 또는 그 기술들의 임의의 조합 중 임의의 것을 수행하도록 구성되거나 다르게는 동작가능하다. 일부 예들에서, 컴퓨터 판독가능 저장 매체는 실행될 때, 하나 이상의 프로세서들로 하여금, 본 명세서에서 설명된 기술들 또는 그 기술들의 임의의 조합 중 임의의 것을 수행하게 하는 명령들로 인코딩된다. 일부 예들에서, 디바이스는 본 명세서에서 설명된 기술들 또는 그 기술들의 임의의 조합 중 임의의 것을 수행하기 위한 수단을 포함한다.

[0950] 즉, 기술들의 전술한 양태들은 오디오 인코딩 디바이스 (510G) 가 다음의 조항들에 따라 동작하도록 구성되는 것을 가능하게 할 수도 있다.

[0951] 조항 134954-1B. 오디오 인코딩 디바이스 (510G) 와 같은 디바이스로서, 하나 이상의 오디오 오브젝트들에 대해 결정된 방향성 및 에너지에 기초하여, 오디오 오브젝트들과 연관된 하나 이상의 구면 조화 계수들로부터 하나 이상의 구별되는 오디오 오브젝트들을 식별하도록 구성된 하나 이상의 프로세서들을 포함하는, 디바이스.

[0952] 조항 134954-2B. 조항 134954-1B 의 디바이스에 있어서, 하나 이상의 프로세서들은 또한, 오디오 오브젝트들과 연관된 구면 조화 계수들에 기초하여 하나 이상의 오디오 오브젝트들의 방향성 및 에너지 중 하나 또는 양자를 결정하도록 구성되는, 디바이스.

[0953] 조항 134954-3B. 청구항 1B 또는 청구항 2B 또는 이들의 조합 중 임의의 것의 디바이스에 있어서, 하나 이상의 프로세서들은 또한, 음장을 나타내는 구면 조화 계수들에 대하여 특이값 분해를 수행하여 복수의 구면 조화 계수들의 좌-특이 벡터들을 나타내는 U 행렬, 복수의 구면 조화 계수들의 특이값들을 나타내는 S 행렬 및 복수의 구면 조화 계수들의 우-특이 벡터들을 나타내는 V 행렬을 생성하며, 복수의 구면 조화 계수들을 U 행렬, S 행렬 및 V 행렬 중 하나 이상의 행렬의 적어도 부분의 함수로서 표현하도록 구성되며, 하나 이상의 프로세서들은 V 행렬에 적어도 부분적으로 기초하여 하나 이상의 오디오 오브젝트들의 개별의 방향성을 결정하도록 구성되며, 하나 이상의 프로세서들은 하나 이상의 오디오 오브젝트들의 개별의 에너지가 S 행렬에 적어도 부분적으로 기초한다고 결정하도록 구성되는, 디바이스.

[0954] 조항 134954-4B. 조항 134954-3B 의 디바이스에 있어서, 하나 이상의 프로세서들은 또한, V 행렬에 S 행렬을 곱하여 VS 행렬을 생성하도록 구성되며, VS 행렬은 하나 이상의 벡터들을 포함하는, 디바이스.

[0955] 조항 134954-5B. 조항 134954-4B 의 디바이스에 있어서, 하나 이상의 프로세서들은 또한, 1 보다 더 큰 오더와 연관되는 VS 행렬의 각각의 로우의 엔트리들을 선택하고, 선택된 엔트리들 각각을 제곱하여 대응하는 제곱 엔트리들을 형성하며, VS 행렬의 각각의 로우에 대해, 제곱 엔트리들 모두를 합산하여 VS 행렬의 대응하는 벡터에 대한 방향성 지수를 생성하도록 구성되는, 디바이스.

[0956] 조항 134954-6B. 청구항 5B 및 청구항 6B 또는 이들의 조합의 임의의 것의 디바이스에 있어서, VS 행렬의 각각의 로우는 25 개의 엔트리들을 포함하는, 디바이스.

[0957] 조항 134954-7B. 조항 134954-6B 의 디바이스에 있어서, 하나 이상의 프로세서들은 VS 행렬의 각각의 로우의 제 5 엔트리에서 시작하고 VS 행렬의 각각의 로우의 제 25 엔트리에서 끝나는 모든 엔트리들을 선택하도록 구성되는, 디바이스.

[0958] 조항 134954-8B. 조항 134954-6B 및 조항 134954-7B 또는 이들의 조합 중 임의의 것의 디바이스에 있어서, 하나 이상의 프로세서들은 또한, 구별되는 오디오 오브젝트들을 표현하기 위해 벡터들의 서브세트를 선택하도록 구성되는, 디바이스.

[0959] 조항 134954-9B. 조항 134954-8B 의 디바이스에 있어서, 하나 이상의 프로세서들은 VS 행렬의 4 개의 벡터들을 선택하도록 구성되며, 선택된 4 개의 벡터들은 VS 행렬의 벡터들 모두의 4 개의 가장 큰 방향성 지수들을 갖는, 디바이스.

[0960] 조항 134954-10B. 조항 134954-8B 및 조항 134954-9B 또는 이들의 조합 중 임의의 것의 디바이스에 있어서, 하나 이상의 프로세서들은 또한, 각각의 벡터의 방향성 및 에너지 양자에 기초하여, 선택된 벡터들의 서브세트가 구별되는 오디오 오브젝트들을 표현한다고 결정하도록 구성되는, 디바이스.

[0961] 조항 134954-1C. 오디오 인코딩 디바이스 (510G) 와 같은 디바이스로서, 방향-기반 정보를 이용하여, 음장의 구별되는 성분들을 기술하는 하나 이상의 제 1 벡터들 및 음장의 배경 성분들을 기술하는 하나 이상의 제 2 벡터들을 결정하도록 구성된 하나 이상의 프로세서들을 포함하며, 하나 이상의 제 1 벡터들 및 하나 이상의 제 2 벡터들 양자는 적어도 복수의 구면 조화 계수들에 대하여 변환을 수행함으로써 생성된, 디바이스.

[0962] 조항 134954-2C. 조항 134954-1C 의 방법에 있어서, 변환은 복수의 구면 조화 계수들의 좌-특이 벡터들을 나타내는 U 행렬, 복수의 구면 조화 계수들의 특이값들을 나타내는 S 행렬 및 복수의 구면 조화 계수들의 우-특이 벡터들을 나타내는 V 행렬을 생성하는 특이값 분해를 포함하는, 방법.

[0963] 조항 134954-3C. 조항 134954-2C 의 방법에 있어서, 조항 134954-1A 내지 조항 134954-12A 및 조항 134954-1B 내지 조항 134954-9B 의 임의의 조합에 의해 인용된 동작들을 더 포함하는, 방법.

[0964] 조항 134954-4C. 조항 134954-1C 의 방법에 있어서, 변환은 음장의 구별되는 성분들 및 음장의 배경 성분들을 식별하기 위해 주요 성분 분석을 포함하는, 방법.

[0965] 도 40h 는 2 차원 또는 3 차원 음장들을 기술하는 구면 조화 계수들을 압축하기 위해 본 개시물에서 설명된 기술들의 다양한 양태들을 수행할 수도 있는 예시적인 오디오 인코딩 디바이스 (510H) 를 예시하는 블록 다이어그램이다. 오디오 인코딩 디바이스 (510H) 는 오디오 인코딩 디바이스 (510H) 가 오디오 압축 유닛 (512), 오디오 인코딩 유닛 (514) 및 비트스트림 생성 유닛 (516) 을 포함한다는 점에서 오디오 인코딩 디바이스 (510G) 와 유사할 수도 있다. 더욱이, 오디오 인코딩 디바이스 (510H) 의 오디오 압축 유닛 (512) 은 오디오 압축 유닛 (512) 이 오디오 인코딩 디바이스 (510G) 의 동일한 유닛들과 유사하게 동작할 수도 있는 분해 유닛 (518) 및 음장 성분 추출 유닛 (520G) 을 포함한다는 점에서 오디오 인코딩 디바이스 (510G) 의 것과 유사할 수도 있다. 일부 예들에서, 오디오 인코딩 디바이스 (510H) 는 도 40d 및 도 40e 에 대하여 설명한 바와 같이, 양자화 유닛 (534) 을 포함하여 U_{DIST} 벡터들 (525C), U_{BG} 벡터들 (525D), V^{T}_{DIST} 벡터들 (525E), 및 V^{T}_{BG} 벡터들 (525J) 중 임의의 하나 이상의 벡터들을 양자화할 수도 있다.

[0966] 그러나, 오디오 인코딩 디바이스 (510H) 의 오디오 압축 유닛 (512) 은 오디오 인코딩 디바이스 (510H) 의 오디오 압축 유닛 (512) 이 보간 유닛 (550) 으로서 표기된 추가적인 유닛을 포함한다는 점에서 오디오 인코딩 디바이스 (510G) 의 오디오 압축 유닛 (512) 과는 상이할 수도 있다. 보간 유닛 (550) 은 도 45 및 도 45b 에 대하여 이하 더 상세하게 설명한 바와 같이, 제 1 오디오 프레임의 서브-프레임들을 제 1 오디오 프레임 및 제 2 시간적으로 후속 또는 선행하는 오디오 프레임의 서브-프레임들로부터 보간하는 유닛을 표현할 수도 있다. 보간 유닛 (550) 은 이 보간을 수행하는데 있어서, 분해 유닛 (518) 이 SHC (511) 를 분해하기 위해 요구되는 정도를 잠재적으로 감소시킴으로써 (프로세싱 사이클들 및/또는 메모리 소비의 관점에서) 컴퓨테이션 복잡성을 감소시킬 수도 있다. 이 점에 있어서, 보간 유닛 (550) 은 도 4 의 예에 도시된 오디오 인코딩 디바이스 (24) 의 공간-시간 보간 유닛 (50) 에 대하여 상기 설명된 동작들과 유사한 동작들을 수행할 수도 있다.

[0967] 즉, 분해 유닛 (518) 에 의해 수행된 특이값 분해는 잠재적으로는 매우 프로세서 및/또는 메모리 집약적인 동시에, 또한, 일부 예들에서는, 특히 SHC (511) 의 오더가 증가할 때 SHC (511) 를 분해하는데 광범위한 시간량이 든다. 그 시간량을 감소시키고 SHC (511) 의 압축을 (프로세싱 사이클들 및/또는 메모리 소비의 관점에서) 더 효율적이게 만들기 위하여, 본 개시물에서 설명된 기술들은 제 1 오디오 프레임의 하나 이상의 서브-프레임들의 보간을 허용할 수도 있으며, 서브-프레임들 각각은 SHC (511) 의 분해된 버전들을 표현할 수도 있다. 전체 프레임에 대하여 SVD 를 수행하기 보다는, 그 기술들은 분해 유닛 (518) 이 제 1 오디오 프레임의 제 1 서브-프레임을 분해하여, V 행렬 (519') 을 생성하는 것을 가능하게 할 수도 있다.

[0968] 분해 유닛 (518) 은 또한, 제 2 오디오 프레임의 제 2 서브-프레임을 분해할 수도 있으며, 여기서 이 제 2 오디오 프레임은 제 1 오디오 프레임에 시간적으로 후속하거나 또는 시간적으로 선행할 수도 있다. 분해 유닛 (518) 은 제 2 오디오 프레임의 이 서브-프레임에 대해 V 행렬 (519') 을 출력할 수도 있다. 보간 유닛 (550) 은 그 후 제 1 및 제 2 서브-프레임들로부터 분해된 V 행렬들 (519') 에 기초하여 제 1 오디오 프레임의 나머지 서브-프레임들을 보간하여, V 행렬 (519), S 행렬 (519B) 및 U 행렬 (519C) 을 출력할 수도 있으며, 여기서 나머지 서브-프레임들에 대한 분해들은 SHC (511), 제 1 오디오 프레임에 대한 V 행렬 (519A) 및 제 1 오디오 프레임의 나머지 서브-프레임들에 대한 보간된 V 행렬들 (519) 에 기초하여 컴퓨팅될 수도 있다. 보간은 따라서 제 1 오디오 프레임의 나머지 서브-프레임들에 대한 분해들의 컴퓨테이션을 회피할 수도 있다.

[0969] 더욱이, 상기 언급한 바와 같이, U 행렬 (519C) 은 프레임 간 연속적이지 않을 수도 있으며, SHC (511) 의 제 1 오디오 프레임으로부터 분해된 U 행렬 (519C) 의 구별되는 성분들은 SHC (511) 의 제 2 오디오 프레임으로부터 분해된 U 행렬 (519C) 에서와는 상이한 로우들 및/또는 컬럼들에서 특정될 수도 있다. 이 보간을 수행함으로써, 불연속성은, 선형 보간이 프레임 경계들로 인해 도입된 임의의 아티팩트들 (또는 다른 말로, SHC (511) 의 프레임들로의 세그먼트화) 감소시킬 수도 있는 평활화 효과를 가질 수도 있다는 것을 고려해 볼 때 감소될 수도 있다. 이 보간을 수행하기 위해 V 행렬 (519') 을 이용한 후 SHC (511) 로부터 보간된 V 행렬 (519') 에 기초하여 U 행렬들 (519C) 을 복구하는 것은 U 행렬 (519C) 을 리오더링하는 것으로부터의 임의의 효과들을 평활화할 수도 있다.

[0970] 동작에 있어서, 보간 유닛 (550) 은 제 1 프레임에 포함된 제 1 복수의 구면 조화 계수들 (511) 의 부분의 제 1 분해, 예를 들어, V 행렬 (519') 및 제 2 프레임에 포함된 제 2 복수의 구면 조화 계수들 (511) 의 부분의 제 2 분해, 예를 들어, V 행렬 (519') 로부터 제 1 오디오 프레임의 하나 이상의 서브-프레임들을 보간하여 하나 이상의 서브-프레임들에 대한 분해된 보간된 구면 조화 계수들을 생성할 수도 있다.

[0971] 일부 예들에서, 제 1 분해는 제 1 복수의 구면 조화 계수들 (511) 의 부분의 우-특이 벡터들을 나타내는 제 1 V 행렬 (519') 을 포함한다. 마찬가지로, 일부 예들에서, 제 2 분해는 제 2 복수의 구면 조화 계수들의 부분

의 우-특이 벡터들을 나타내는 제 2 V 행렬 (519') 을 포함한다.

[0972] 보간 유닛 (550) 은 제 1 V 행렬 (519') 및 제 2 V 행렬 (519') 에 기초하여 하나 이상의 서브-프레임들에 대하여 시간적 보간을 수행할 수도 있다. 즉, 보간 유닛 (550) 은 예를 들어, 제 1 오디오 프레임의 제 1 서브-프레임으로부터 분해된 V 행렬 (519') 및 제 2 오디오 프레임의 제 1 서브-프레임으로부터 분해된 V 행렬 (519') 에 기초하여 제 1 오디오 프레임에 대한 4 개의 총 서브-프레임들 중에서 예를 들어, 제 2, 제 3 및 제 4 서브-프레임들을 시간적으로 보간할 수도 있다. 일부 예들에서, 이 시간적 보간은 선형 시간적 보간이며, 제 1 오디오 프레임의 제 1 서브-프레임으로부터 분해된 V 행렬 (519') 은 제 1 오디오 프레임의 제 4 서브-프레임을 보간할 때보다 제 1 오디오 프레임의 제 2 서브-프레임을 보간할 때 더 과중하게 가중된다. 제 3 서브-프레임을 보간할 때, V 행렬들 (519') 은 균등하게 가중될 수도 있다. 제 4 서브-프레임을 보간할 때, 제 2 오디오 프레임의 제 1 서브-프레임으로부터 분해된 V 행렬 (519') 은 제 1 오디오 프레임의 제 1 서브-프레임으로부터 분해된 V 행렬 (519') 보다 더 과중하게 가중될 수도 있다.

[0973] 즉, 선형 시간적 보간은 보간될 제 1 오디오 프레임의 서브-프레임들 중 하나의 서브-프레임의 근접성을 고려하여 V 행렬들 (519') 을 가중할 수도 있다. 보간될 제 2 서브-프레임의 경우, 제 1 오디오 프레임의 제 1 서브-프레임으로부터 분해된 V 행렬 (519') 은 보간될 제 2 서브-프레임에 대한 그것의 근접성을 고려해 볼 때 제 2 오디오 프레임의 제 1 서브-프레임으로부터 분해된 V 행렬 (519') 보다 더 과중하게 가중된다. 가중치들은 이런 이유로 V 행렬들 (519') 에 기초하여 제 3 서브-프레임을 보간할 때 등가일 수도 있다. 제 2 오디오 프레임의 제 1 서브-프레임으로부터 분해된 V 행렬 (519') 에 적용된 가중치는, 보간될 제 4 서브-프레임이 제 1 오디오 프레임의 제 1 서브-프레임보다 제 2 오디오 프레임의 제 1 서브-프레임에 더 근접한다는 것을 고려하면 제 1 오디오 프레임의 제 1 서브-프레임으로부터 분해된 V 행렬 (519') 에 적용된 것보다 더 클 수도 있다.

[0974] 일부 예들에서, 단지 각각의 오디오 프레임의 제 1 서브-프레임만이 보간을 수행하기 위해 사용되지만, 복수의 구면 조화 계수들의 부분은 제 1 복수의 구면 조화 계수들 (511) 의 4 개의 서브-프레임들 중 2 개를 포함할 수도 있다. 이들 및 다른 예들에서, 제 2 복수의 구면 조화 계수들 (511) 의 부분은 제 2 복수의 구면 조화 계수들 (511) 의 4 개의 서브-프레임들 중 2 개를 포함한다.

[0975] 상기 언급한 바와 같이, 단일 디바이스, 예를 들어, 오디오 인코딩 디바이스 (510H) 는 제 1 복수의 구면 조화 계수들의 부분의 제 1 분해들을 생성하기 위해 제 1 복수의 구면 조화 계수들의 부분을 또한 분해하면서 보간을 수행할 수도 있다. 이들 및 다른 예들에서, 분해 유닛 (518) 은 제 2 복수의 구면 조화 계수들의 부분의 제 2 분해들을 생성하기 위해 제 2 복수의 구면 조화 계수들의 부분을 분해할 수도 있다. 단일 디바이스에 대하여 설명하였지만, 2 개 이상의 디바이스들은 본 개시물에서 설명된 기술들을 수행할 수도 있고, 2 개의 디바이스들 중 하나는 분해를 수행하고 디바이스들 중 다른 하나는 본 개시물에서 설명된 기술들에 따른 보간을 수행한다.

[0976] 즉, 구면 조화-기반 3D 오디오는 구 상의 직교 기저 함수들의 관점에서의 3D 압력장 (pressure field) 의 파라메트릭 표현일 수도 있다. 표현의 오더 N 이 높을수록, 공간 해상도는 잠재적으로 더 높고 종종 구면 조화 (SH) 계수들의 수가 커진다 (총 $(N+1)^2$ 계수들). 많은 애플리케이션들을 위해, 계수들의 대역폭 압축은 계수들을 효율적으로 송신 및 저장가능하기 위해 요구될 수도 있다. 본 개시물에 관련된 이 기술들은 특이값 분해 (SVD) 를 이용하여 프레임-기반 차원성 감소 프로세스를 제공할 수도 있다. SVD 분석은 3 개의 행렬들 U, S 및 V 로 계수들의 각각의 프레임을 분해할 수도 있다. 일부 예들에서, 기술들은 U 의 벡터들의 일부를 근본적인 음장의 방향성 성분들로서 취급할 수도 있다. 그러나, 이 방식으로 취급될 때, 이들 (U 의) 벡터들은 - 그들이 동일한 구별되는 오디오 성분을 표현하더라도 프레임 간 불연속적이다. 이들 불연속성은 성분들이 변환-오디오-코더들을 통하여 피딩될 때 상당한 아티팩트들을 야기할 수도 있다.

[0977] 본 개시물에서 설명된 기술들은 이 불연속성을 다룰 수도 있다. 즉, 그 기술들은 V 행렬이 구면 조화 도메인에서 직교 공간 축들로서 해석될 수 있다는 관찰에 기초할 수도 있다. U 행렬은 그 기저 함수들의 관점에서 구면 조화 (HOA) 데이터의 프로젝션을 표현할 수도 있으며, 불연속성은 매 프레임마다 변하며 - 따라서 그들 스스로 불연속적인 기저 함수들 (V) 에 기여될 수 있다. 이것은 다른 유사한 분해, 이를 테면 푸리에 변환이며, 여기서 기저 함수들은 일부 예들에서 프레임 간 일정하다. 이런 면에서, SVD 는 정합 추적 알고리즘인 것으로 간주될 수도 있다. 본 개시물에서 설명된 기술들은 보간 유닛 (550) 이 프레임들 간 - 그들 사이에 보간함으로써 기저 함수들 (V) 사이의 연속성을 유지하는 것을 가능하게 할 수도 있다.

[0978] 일부 예들에서, 기술들은 보간 유닛 (550) 이 상기 설명되고 도 45 및 도 45b 에 대하여 이하 추가 설명된 바와

같이, SH 의 프레임을 4 개의 서브프레임들로 분할하는 것을 가능하게 한다. 보간 유닛 (550) 은 그 후 제 1 서브-프레임에 대한 SVD 를 컴퓨팅할 수도 있다. 유사하게 제 2 프레임의 제 1 서브-프레임에 대해 SVD 를 컴퓨팅한다. 제 1 프레임 및 제 2 프레임 각각에 대해, 보간 유닛 (550) 은 V 의 벡터들을, (예를 들어, T-설계 행렬과 같은 투영 행렬을 이용하여) 구 상으로 그 벡터들을 투영함으로써 공간 맵으로 변환할 수도 있다. 보간 유닛 (550) 은 그 후 V 의 벡터들을 구 상의 형상들로서 해석할 수도 있다. 3 개의 서브-프레임들에 대한 V 행렬들을 제 1 프레임의 제 1 서브-프레임과 다음 프레임의 제 1 서브-프레임 사이에 보간하기 위해, 보간 유닛 (550) 은 그 후 이들 공간 형상들을 보간할 수도 있고 - 그 후 그들을 다시 SH 벡터들로 투영 행렬의 역을 통해 변환할 수도 있다. 본 개시물의 기술들은 이 방식으로 V 행렬들 사이의 평활한 천이를 제공할 수도 있다.

[0979] 이 방식으로, 오디오 인코딩 디바이스 (510H) 는 다음의 조항들에 대하여 이하 기재된 기술들의 다양한 양태들을 수행하도록 구성될 수도 있다.

[0980] 조항 135054-1A. 오디오 인코딩 디바이스 (510H) 와 같은 디바이스로서, 제 1 프레임에 포함된 제 1 복수의 구면 조화 계수들의 부분의 제 1 분해 및 제 2 프레임에 포함된 제 2 복수의 구면 조화 계수들의 부분의 제 2 분해로부터 제 1 프레임의 하나 이상의 서브-프레임들을 보간하여 하나 이상의 서브-프레임들에 대한 분해된 보간된 구면 조화 계수들을 생성하도록 구성된 하나 이상의 프로세서들을 포함하는, 디바이스.

[0981] 조항 135054-2A. 조항 135054-1A 의 디바이스에 있어서, 제 1 분해는 제 1 복수의 구면 조화 계수들의 부분의 우-특이 벡터들을 나타내는 제 1 V 행렬을 포함하는, 디바이스.

[0982] 조항 135054-3A. 조항 135054-1A 의 디바이스에 있어서, 제 2 분해는 제 2 복수의 구면 조화 계수들의 부분의 우-특이 벡터들을 나타내는 제 2 V 행렬을 포함하는, 디바이스.

[0983] 조항 135054-4A. 조항 135054-1A 의 디바이스에 있어서, 제 1 분해는 제 1 복수의 구면 조화 계수들의 부분의 우-특이 벡터들을 나타내는 제 1 V 행렬을 포함하며, 제 2 분해는 제 2 복수의 구면 조화 계수들의 부분의 우-특이 벡터들을 나타내는 제 2 V 행렬을 포함하는, 디바이스.

[0984] 조항 135054-5A. 조항 135054-1A 의 디바이스에 있어서, 하나 이상의 프로세서들은 또한, 하나 이상의 서브-프레임들을 보간할 때, 제 1 분해 및 제 2 분해에 기초하여 하나 이상의 서브-프레임들을 시간적으로 보간하도록 구성된다.

[0985] 조항 135054-6A. 조항 135054-1A 의 디바이스, 여기서 하나 이상의 프로세서들은 또한, 하나 이상의 서브프레임들을 보간할 때, 제 1 투영된 분해들을 생성하기 위해 공간 도메인 내로 제 1 분해를 투영하고, 제 2 투영된 분해들을 생성하기 위해 공간 도메인 내로 제 2 분해를 투영하며, 제 1 공간적으로 보간된 투영된 분해 및 제 2 공간적으로 보간된 투영된 분해를 생성하기 위해 제 1 투영된 분해들 및 제 2 투영된 분해들을 공간적으로 보간하고, 제 1 공간적으로 보간된 투영된 분해 및 제 2 공간적으로 보간된 투영된 분해에 기초하여 하나 이상의 서브프레임들을 시간적으로 보간하도록 구성된다.

[0986] 조항 135054-7A. 조항 135054-6A 의 디바이스, 여기서 하나 이상의 프로세서들은 또한, 구면 조화 도메인으로 하나 이상의 서브프레임들을 다시 보간하는 것으로부터 초래된 시간적으로 보간된 구면 조화 계수들을 투영하도록 구성된다.

[0987] 조항 135054-8A. 조항 135054-1A 의 디바이스, 여기서 제 1 복수의 구면 조화 계수들의 일부는 제 1 복수의 구면 조화 계수들의 단일의 서브프레임을 포함한다.

[0988] 조항 135054-9A. 조항 135054-1A 의 디바이스, 여기서 제 2 복수의 구면 조화 계수들의 일부는 제 2 복수의 구면 조화 계수들의 단일의 서브프레임을 포함한다.

[0989] 조항 135054-10A. 조항 135054-1A 의 디바이스,

[0990] 여기서 제 1 프레임은 4 개의 서브프레임들로 분할되고,

[0991] 여기서 제 1 복수의 구면 조화 계수들의 일부는 제 1 복수의 구면 조화 계수들의 제 1 서브프레임만을 포함한다.

[0992] 조항 135054-11A. 조항 135054-1A 의 디바이스,

[0993] 여기서 제 2 프레임은 4 개의 서브프레임들로 분할되고,

[0994] 여기서 제 2 복수의 구면 조화 계수들의 일부는 제 2 복수의 구면 조화 계수들의 제 1 서브프레임만을 포함한다.

[0995] 조항 135054-12A. 조항 135054-1A 의 디바이스, 여기서 제 1 복수의 구면 조화 계수들의 일부는 제 1 복수의 구면 조화 계수들의 4 개의 서브프레임들 중 2 개를 포함한다.

[0996] 조항 135054-13A. 조항 135054-1A 의 디바이스, 여기서 제 2 복수의 구면 조화 계수들의 일부는 제 2 복수의 구면 조화 계수들의 4 개의 서브프레임들 중 2 개를 포함한다.

[0997] 조항 135054-14A. 조항 135054-1A 의 디바이스, 여기서 하나 이상의 프로세서들은 또한 제 1 복수의 구면 조화 계수들의 일부의 제 1 분해들을 생성하기 위해 제 1 복수의 구면 조화 계수들의 일부를 분해하도록 구성된다.

[0998] 조항 135054-15A. 조항 135054-1A 의 디바이스, 여기서 하나 이상의 프로세서들은 또한 제 2 복수의 구면 조화 계수들의 일부의 제 2 분해들을 생성하기 위해 제 2 복수의 구면 조화 계수들의 일부를 분해하도록 구성된다.

[0999] 조항 135054-16A. 조항 135054-1A 의 디바이스, 여기서 하나 이상의 프로세서들은 또한, 제 1 복수의 구면 조화 계수들의 좌-특이 벡터들을 나타내는 U 행렬, 제 1 복수의 구면 조화 계수들의 특이값들을 나타내는 S 행렬 및 제 1 복수의 구면 조화 계수들의 우-특이 벡터들을 나타내는 V 행렬을 생성하기 위해 제 1 복수의 구면 조화 계수들의 일부에 대해 특이값 분해를 수행하도록 구성된다.

[1000] 조항 135054-17A. 조항 135054-1A 의 디바이스, 여기서 하나 이상의 프로세서들은 또한, 제 2 복수의 구면 조화 계수들의 좌-특이 벡터들을 나타내는 U 행렬, 제 2 복수의 구면 조화 계수들의 특이값들을 나타내는 S 행렬 및 제 2 복수의 구면 조화 계수들의 우-특이 벡터들을 나타내는 V 행렬을 생성하기 위해 제 2 복수의 구면 조화 계수들의 일부에 대해 특이값 분해를 수행하도록 구성된다.

[1001] 조항 135054-18A. 조항 135054-1A 의 디바이스, 여기서 제 1 및 제 2 복수의 구면 조화 계수들은 각각 음장의 평면파 표현을 나타낸다.

[1002] 조항 135054-19A. 조항 135054-1A 의 디바이스, 여기서 제 1 및 제 2 복수의 구면 조화 계수들은 각각 함께 혼합된 하나 이상의 모노-오디오 오브젝트들을 나타낸다.

[1003] 조항 135054-20A. 조항 135054-1A 의 디바이스, 여기서 제 1 및 제 2 복수의 구면 조화 계수들은 각각 3 차원 음장을 나타내는 각각의 제 1 및 제 2 구면 조화 계수들을 포함한다.

[1004] 조항 135054-21A. 조항 135054-1A 의 디바이스, 여기서 제 1 및 제 2 복수의 구면 조화 계수들은 1 보다 큰 오더를 갖는 적어도 하나의 구면 기저 함수와 각각 연관된다.

[1005] 조항 135054-22A. 조항 135054-1A 의 디바이스, 여기서 제 1 및 제 2 복수의 구면 조화 계수들은 4 보다 큰 오더를 갖는 적어도 하나의 구면 기저 함수와 각각 연관된다.

[1006] 오디오 인코딩 디바이스 (510H) 에 의해 수행되는 것으로서 전술되었지만, 여러 오디오 디코딩 디바이스들 (24 및 540) 은 또한 조항들 135054-1A 내지 135054-22A 에 대해 전술된 기술들의 임의의 여러 양태들을 수행할 수도 있다.

[1007] 도 40i 는 2 또는 3 차원 음장들을 기술하는 구면 조화 계수들을 압축하기 위해 본 개시물에 기술된 기술들의 여러 양태들을 수행할 수도 있는 예시의 오디오 인코딩 디바이스 (510I) 를 예시하는 블록 다이어그램이다. 오디오 인코딩 디바이스 (510I) 는 오디오 인코딩 디바이스 (510I) 가 오디오 압축 유닛 (512), 오디오 인코딩 유닛 (514) 및 비트스트림 생성 유닛 (516) 을 포함한다는 점에서 오디오 인코딩 디바이스 (510H) 와 유사할 수도 있다. 게다가, 오디오 인코딩 디바이스 (510I) 의 오디오 압축 유닛 (512) 은 오디오 압축 유닛 (512) 이 오디오 인코딩 디바이스 (510H) 의 유사한 유닛들과 유사하게 동작할 수도 있는 분해 유닛 (518) 및 음장 성분 추출 유닛 (520) 을 포함한다는 점에서 오디오 인코딩 디바이스 (510H) 의 그것과 유사할 수도 있다. 일부 예들에서, 오디오 인코딩 디바이스 (10I) 는 U_{DIST} (25C), U_{BG} (25D), V^{T}_{DIST} (25E), 및 V^{T}_{BG} (25J) 의 임의의 것의 하나 이상의 벡터들을 양자화하기 위해, 도 3D 내지 도 3E 에 대해 기술된, 양자화 유닛 (34) 을 포함할 수도 있다.

[1008] 그러나, 오디오 인코딩 디바이스 (510I) 의 오디오 압축 유닛 (512) 및 오디오 인코딩 디바이스 (10H) 의 오디

오 압축 유닛 (512) 의 양자 모두는 음장 성분 추출 유닛을 포함하는 반면, 오디오 인코딩 디바이스 (510I) 의 음장 성분 추출 유닛 (520I) 은 V 압축 유닛 (552) 로서 지칭되는 추가적인 모듈을 포함할 수도 있다. V 압축 유닛 (552) 은 음장의 공간 성분, 즉 이 예시에서 하나 이상의 V^{T}_{DIST} 벡터들 (539) 을 압축하도록 구성된 유닛을 나타낼 수도 있다. 즉, SHC 에 대해 수행된 특이값 분해는 (음장을 나타내는) SHC 를 S 행렬의 벡터들에 의해 표현되는 에너지 성분들, U 행렬에 의해 표현되는 시간 성분들 및 V 행렬에 의해 표현되는 공간 성분들로 분해할 수도 있다. V 압축 유닛 (552) 은 양자화 유닛 (52) 에 대해 전술된 것들과 유사한 동작들을 수행할 수도 있다.

[1009] 예시의 목적으로, V^{T}_{DIST} 벡터들 (539) 은 (음장의 제 4 차 HOA 표현을 암시하는) 25 개의 엘리먼트들 각각을 갖는 2 개의 로우 벡터들을 포함하는 것으로 가정된다. 2 개의 로우 벡터들에 대해 기술되지만, 임의의 수의 벡터들이 최대 $(n+1)^{2}$ 까지 V^{T}_{DIST} 벡터들 (539) 에 포함될 수도 있으며, 여기서 n 은 음장의 HOA 표현의 오더를 나타낸다.

[1010] V 압축 유닛 (552) 은 V^{T}_{DIST} 벡터들 (539) 을 수신하고 압축된 V^{T}_{DIST} 벡터 표현들 (539') 을 생성하기 위해 압축 방식을 수행할 수도 있다. 이러한 압축 방식은 일반적으로 벡터 또는 데이터의 엘리먼트들을 압축하는 임의의 생각할 수 있는 압축 방식을 수반할 수도 있고, 이하에 더 상세하게 기술되는 예시에 제한되지 않아야 한다.

[1011] V 압축 유닛 (552) 은 일예로서, V^{T}_{DIST} 벡터들 (539) 의 각 엘리먼트의 부동소수점 표현들을 V^{T}_{DIST} 벡터들 (539) 의 각 엘리먼트의 정수 표현들로 변환하는 것, V^{T}_{DIST} 벡터들 (539) 의 정수 표현들의 균일한 양자화 및 V^{T}_{DIST} 벡터들 (539) 의 양자화된 정수 표현들의 카테고리화 및 코딩 중 하나 이상을 포함하는 압축 방식을 수행할 수도 있다. 이러한 압축 방식의 여러가지 하나 이상의 프로세스들은, 일 예로서, 결과의 비트스트림 (517) 에 대한 목표 비트레이트를 달성하거나 거의 달성하기 위해 파라미터들에 의해 동적으로 제어될 수도 있다.

[1012] V^{T}_{DIST} 벡터들 (539) 각각이 서로에 대해 정규직교인 경우, V^{T}_{DIST} 벡터들 (539) 각각은 독립적으로 코딩될 수도 있다. 일부 예들에서, 이하에 더 상세하게 기술되는 바와 같이, V^{T}_{DIST} 벡터 (539) 의 각각의 엘리먼트는 (여러 서브모드들에 의해 정의된) 동일한 코딩 모드를 사용하여 코딩될 수도 있다.

[1013] 임의의 경우에, 전술된 바와 같이, 이러한 코딩 방식은 먼저 V^{T}_{DIST} 벡터들 (539) 각각의 (일부 예들에서, 32-비트 부동소수점수인) 각 엘리먼트의 부동소수점 표현들을 16-비트 정수 표현들으로 변환하는 것을 수반할 수도 있다. V 압축 유닛 (552) 은, 일부 예시들에서 15 만큼의 우측 시프트에 의해 수행되는, V^{T}_{DIST} 벡터들 (539) 중 주어진 것의 각 엘리먼트를 2^{15} 로 곱하는 것에 의해 이러한 부동소수점-대-정수 변환을 수행할 수도 있다.

[1014] V 압축 유닛 (552) 은 그 후 V^{T}_{DIST} 벡터들 (539) 의 주어진 것의 엘리먼트들 모두에 대해 균일한 양자화를 수행할 수도 있다. V 압축 유닛 (552) 은 nbits 파라미터로서 표시될 수도 있는 값에 기초하여 양자화 스텝 사이즈를 식별할 수도 있다. V 압축 유닛 (552) 은 목표 비트 레이트에 기초하여 이러한 nbits 파라미터를 동적으로 결정할 수도 있다. V 압축 유닛 (552) 은 이러한 nbits 파라미터의 함수로서 양자화 스텝 사이즈를 결정할 수도 있다. 하나의 예로서, V 압축 유닛 (552) 은 $2^{16-nbits}$ 와 동일한 (본 개시물에서 "델타" 또는 "Δ" 로서 표시된) 양자화 스텝 사이즈를 결정할 수도 있다. 이러한 예에서, nbits 가 6 과 동일한 경우, 델타는 2^{10} 과 같고, 2^{6} 개의 양자화 레벨들이 존재한다. 이점에 있어서, 벡터 엘리먼트 v 의 경우, 양자화된 벡터 엘리먼트 v_{q} 는 [v/Δ] 와 같고, $-2^{nbits-1} < v_{q} < 2^{nbits-1}$ 이다.

[1015] V 압축 유닛 (552) 은 그 후 양자화된 벡터 엘리먼트들의 카테고리화 및 잔차 코딩을 수행할 수도 있다. 일 예시로서, V 압축 유닛 (552) 은, 주어진 양자화된 벡터 엘리먼트 v_{q} 에 대해, 다음의 방정식을 사용하여 이러한 엘리먼트가 대응하는 (카테고리 식별자 cid 를 결정함으로써) 카테고리를 식별할 수도 있다:

$$cid = \begin{cases} 0, & if\ v_q = 0 \\ \lfloor log_2|v_q| \rfloor + 1, & if\ v_q \neq 0 \end{cases}$$

[1016]

[1017] V 압축 유닛 (552) 은 그 후, 또한 v_q 가 포지티브 값인지 네거티브 값인지 여부를 나타내는 부호 비트를 식별하면서, 이러한 카테고리 인덱스 cid 를 허프만 코딩할 수도 있다. V 압축 유닛 (552) 은 다음에 이러한 카테고리에서 잔차를 식별할 수도 있다. 하나의 예로서, V 압축 유닛 (552) 은 다음의 방정식에 따라 이러한 잔차를 결정할 수도 있다:

[1018] 잔차 = $|v_q| - 2^{cid-1}$

[1019] V 압축 유닛 (552) 은 그 후 cid-1 비트들로 이러한 잔차를 블록 코딩할 수도 있다.

[1020] 다음의 예는 이러한 카테고리화 및 잔차 코딩 프로세스의 단순화된 예를 도시한다. 먼저, $v_q \in [-31,31]$ 이도록 nbits 가 6 과 동일하다고 가정하라. 다음에, 다음을 가정하라:

cid	vq	cid 에 대한 허프만 코드
0	0	'1'
1	-1, 1	'01'
2	-3,-2, 2,3	'000'
3	-7,-6,-5,-4, 4,5,6,7	'0010'
4	-15,-14,...,-8, 8,...,14,15	'00110'
5	-31,-30,...,-16, 16,...,30,31	'00111'

[1021]

[1022] 또한, 다음을 가정하라:

cid	잔차에 대한 블록 코드
0	N/A
1	0, 1
2	01,00, 10,11
3	011,010,001,000, 100,101,110,111
4	0111, 0110...,0000, 1000,...,1110,1111
5	01111, ... ,00000, 10000, ... ,11111

[1023]

[1024] 따라서, v_q = [6, -17, 0, 0, 3] 에 대해, 다음이 결정될 수도 있다:

[1025] >> cid = 3,5,0,0,2

[1026] >> 부호=1,0,x,x,1

[1027] >> 잔차 = 2,1,x,x,1

[1028] >> 6 에 대한 비트들 = '0010' + '1' + '10'

[1029] >> -17 에 대한 비트들 = '00111' + '0' + '0001'

[1030] >> 0 에 대한 비트들 = '0'

[1031] >> 0 에 대한 비트들 = '0'

[1032] >> 3 에 대한 비트들 = '000' + '1' + '1'

[1033] >> 총 비트들 = 7+10+1+1+5 = 24

[1034] >> 평균 비트들 = 24/5 = 4.8

[1035] 다음의 단순화된 예에서 도시되지 않지만, V 압축 유닛 (552) 은 cid 를 코딩할 때 nbits 의 상이한 값들에 대해 상이한 허프만 코드 북들을 선택할 수도 있다. 일부 예들에서, V 압축 유닛 (552) 은 nbits 값들 6, ..., 15 에 대해 상이한 허프만 코딩 테이블을 제공할 수도 있다. 게다가, V 압축 유닛 (552) 은 총 50 개의 허프만 코드 북들에 대해 6, ..., 15 의 범위의 상이한 nbits 값들 각각에 대한 5 개의 상이한 허프만 코드 북들을 포함할 수도 있다. 이러한 점에서, V 압축 유닛 (552) 은 다수의 상이한 통계적 컨텍스트들에서 cid 의 코딩을 수용하기 위해 복수의 상이한 허프만 코드 북들을 포함할 수도 있다.

[1036] 설명을 위해, V 압축 유닛 (552) 은, nbits 값들 각각에 대해, 벡터 엘리먼트들 1 내지 4 를 코딩하기 위한 제 1 허프만 코드 북, 벡터 엘리먼트들 5 내지 9 를 코딩하기 위한 제 2 허프만 코드 북, 벡터 엘리먼트들 9 이상을 코딩하기 위한 제 3 허프만 코드 북을 포함할 수도 있다. 이들 제 1 의 3 개의 허프만 코드 북들은, 압축될 V^{T}_{DIST} 벡터들 (539) 중 하나가 V^{T}_{DIST} 벡터들 (539) 중 시간적으로 후속하는 대응하는 하나로부터 예측되지 않고 합성 오디오 오브젝트 (예를 들어, 원래 펄스 코드 변조된 (PCM) 오디오 오브젝트에 의해 정의된 것) 의 공간적 정보를 나타내지 않는 경우 사용될 수도 있다. V 압축 유닛 (552) 은, V^{T}_{DIST} 벡터들 (539) 중 이러한 하나가 V^{T}_{DIST} 벡터들 (539) 중 시간적으로 후속하는 대응하는 하나로부터 예측되는 경우, nbits 값들 각각에 대해, V^{T}_{DIST} 벡터들 (539) 중 그 하나를 코딩하기 위해 제 4 허프만 코드 북을 추가로 포함할 수도 있다. V 압축 유닛 (552) 은 또한, V^{T}_{DIST} 벡터들 (539) 중 이러한 하나가 합성 오디오 오브젝트를 나타내는 경우, nbits 값들 각각에 대해, V^{T}_{DIST} 벡터들 (539) 중 그 하나를 코딩하기 위해 제 5 허프만 코드 북을 추가로 포함할 수도 있다. 여러 허프만 코드 북들이 이들 상이한 통계적 컨텍스트들, 즉 이 예에서 비예측 및 비합성 컨텍스트, 예측 컨텍스트 및 합성 컨텍스트의 각각에 대해 전개될 수도 있다.

[1037] 다음의 표는 분해 유닛이 적절한 허프만 테이블을 선택하는 것을 가능하게 하기 위해 비트스트림에서 특정될 비트들 및 허프만 테이블 선택을 도시한다:

Pred mode	**HT info**	**HT** 테이블
0	0	HT5
0	1	HT{1,2,3}
1	0	HT4
1	1	HT5

[1038]

[1039] 전술한 표에서, 예측 모드 ("Pred mode") 는 예측이 현재의 벡터에 대해 수행되는지 여부를 나타내는 반면, 허프만 테이블 ("HT info") 은 허프만 테이블들 1 내지 5 중 하나를 선택하기 위해 사용되는 추가적인 허프만 코드 북 (또는 테이블) 정보를 나타낸다.

[1040] 다음의 표는 또한 여러 통계적 컨텍스트들 또는 시나리오들이 주어질 때 이러한 허프만 테이블 선택 프로세스를 도시한다.

[1041]

	레코딩	합성
예측 없음	HT{1,2,3}	HT5
예측 있음	HT4	HT5

[1042] 상기 표에서, "레코딩" 컬럼은 벡터가 레코딩된 오디오 오브젝트를 나타내는 경우의 코딩 컨텍스트를 나타내는 반면, "합성" 컬럼은 벡터가 합성 오디오 오브젝트를 나타내는 경우에 대해 코딩 컨텍스트를 나타낸다. "예측 없음" 로우는 예측이 벡터 엘리먼트들에 대해 수행되지 않는 경우의 코딩 컨텍스트를 나타내는 반면, "예측 있음" 로우는 예측이 벡터 엘리먼트들에 대해 수행되는 경우의 코딩 컨텍스트를 나타낸다. 이러한 표에서 보여주는 바와 같이, V 압축 유닛 (552) 은 벡터가 레코딩된 오디오 오브젝트를 나타내고 예측이 벡터 엘리먼트들에 대해 수행되지 않는 경우 HT{1, 2, 3} 을 선택한다. V 압축 유닛 (552) 은 오디오 오브젝트가 합성 오디오 오브젝트를 나타내고 예측이 벡터 엘리먼트들에 대해 수행되지 않는 경우 HT5 를 선택한다. V 압축 유닛 (552) 은 벡터가 레코딩된 오디오 오브젝트를 나타내고 예측이 벡터 엘리먼트들에 대해 수행되는 경우 HT4 를 선택한다. V 압축 유닛 (552) 은 오디오 오브젝트가 합성 오디오 오브젝트를 나타내고 예측이 벡터 엘리먼트들에 대해 수행되는 경우 HT5 를 선택한다.

[1043] 이러한 방식으로, 그 기술들은 오디오 압축 디바이스가 음장의 공간 성분을 압축하는 것을 가능하게 할 수도 있으며, 여기서 공간 성분은 복수의 구면 조화 계수들에 대해 벡터 기반 합성을 수행함으로써 생성된다.

[1044] 도 43 은 도 40i 에 도시된 V 압축 유닛 (552) 을 더 상세하게 예시하는 다이어그램이다. 도 43 의 예에서, V 압축 유닛 (552) 은 균일 양자화 유닛 (600), n 비트들 유닛 (602), 예측 유닛 (604), 예측 모드 유닛 (606) ("Pred Mode Unit 606"), 카테고리 및 잔차 코딩 유닛 (608), 허프만 테이블 선택 유닛 (610) 을 포함한다. 균일 양자화 유닛 (600) 은 (V^{T}_{DIST} 벡터들 (539) 중 임의의 하나를 나타낼 수도 있는) 도 43 의 예에서 v 로서 표시되는 공간 성분들 중 하나에 대해 전술된 균일 양자화를 수행하도록 구성된 유닛을 나타낸다. n 비트들 유닛 (602) 은 nbits 파라미터 또는 값을 결정하도록 구성된 유닛을 나타낸다.

[1045] 예측 유닛 (604) 은 도 43 의 예에서 v_q 로서 표시된 양자화된 공간 성분에 대한 예측을 수행하도록 구성된 유닛을 나타낸다. 예측 유닛 (604) 은 V^{T}_{DIST} 벡터들 (539) 중 시간적으로 후속하는 대응하는 벡터에 의한 V^{T}_{DIST} 벡터들 (539) 중 현재의 벡터의 엘리먼트 기반 감산을 수행함으로써 예측을 수행할 수도 있다. 이러한 예측의 결과는 예측된 공간 성분으로서 지칭될 수도 있다.

[1046] 예측 모드 유닛 (606) 은 예측 모드를 선택하도록 구성된 유닛을 나타낼 수도 있다. 허프만 테이블 선택 유닛 (610) 은 cid 의 코딩을 위한 적절한 허프만 테이블을 선택하도록 구성된 유닛을 나타낼 수도 있다. 예측 모드 유닛 (606) 및 허프만 테이블 선택 유닛 (610) 은, 하나의 예로서, 다음의 의사-코드에 따라 동작할 수도 있다:

[1047] 주어진 nbits 에 대해, nbits 들을 갖는 모든 허프만 테이블들을 취출

[1048] B00 = 0; B01 = 0; B10 = 0; B11 = 0; // 코딩 모드 당 예상된 비트들을 컴퓨팅하기 위해 초기화

[1049] for m = 1:(벡터 내의 # elements)

[1050] // 벡터 엘리먼트 v(m) 에 대한 비트들의 예상된 수를 계산

[1051] // 예측 없고 허프만 테이블 5 를 사용

[1052] B00 = B00 + calculate_bits(v(m), HT5);

[1053] // 예측 없고 허프만 테이블 {1,2,3} 을 사용

[1054] B01 = B01 + calculate_bits(v(m), HTq); q in {1,2,3}

[1055]

[1056] // 예측 잔차 e(m) 에 대한 비트들의 예상된 수를 계산

[1057] e(m) = v(m) - vp(m); // vp(m): 이전 프레임 벡터 엘리먼트

[1058] // 예측 있고 허프만 테이블 4 를 사용

[1059] B10 = B10 + calculate_bits(e(m), HT4);

[1060] // 예측 있고 허프만 테이블 5 를 사용

[1061] B11 = B11 + calculate_bits(e(m), HT5);

[1062] end

[1063] // 최소 비트들을 산출하는 최선의 예측 모드 및 허프만 테이블을 발견

[1064] // 최선의 예측 모드 및 허프만 데이블은 각각 pflag 및 Htflag 에 의해 플래깅된다

[1065] [Be, id] = min([B00 B01 B10 B11]);

[1066] Switch id

[1067] case 1: pflag = 0; HTflag = 0;

[1068] case 2: pflag = 0; HTflag = 1;

[1069] case 3: pflag = 1; HTflag = 0;

[1070] case 4: pflag = 1; HTflag = 1;

[1071] end

[1072] 카테고리 및 잔차 코딩 유닛 (608) 은 위에서 더 상세하게 기술된 방식으로 (예측이 가능하지 않은 경우) 예측된 공간 성분 또는 양자화된 공간 성분의 카테고리화 및 잔차 코딩을 수행하도록 구성된 유닛을 나타낼 수도 있다.

[1073] 도 43 의 예에서 도시된 바와 같이, V 압축 유닛 (552)은 (그 자체가 비트스트림 (517) 과는 별개인 비트스트림일 수도 있는) 비트스트림 (517) 또는 사이드 정보에의 포함을 위한 여러 파라미터들 또는 값들을 출력할 수도 있다. 정보가 비트스트림 (517) 에서 특정된다고 가정하면, V 압축 유닛 (552) 은, 이러한 예에서 cid, 부호 비트, 및 블록 코딩된 잔차를 인코딩하기 위해 선택된 허프만 코드로서 지칭될 수도 있는 (도 40i 의 예에서 압축된 공간 성분 (539') 으로서 도시된) 공간 성분의 압축된 버전과 함께 비트스트림 생성 유닛 (516) 으로 nbits 값, 예측 모드 및 허프만 테이블 정보를 출력할 수도 있다. nbits 값은 V^{T}_{DIST} 벡터들 (539) 의 모두에 대해 비트스트림 (517) 에서 한번 특정될 수도 있는 반면, 예측 모드 및 허프만 테이블 정보는 V^{T}_{DIST} 벡터들 (539) 중 각각의 벡터에 대해 특정될 수도 있다. 공간 성분의 압축된 버전을 명시하는 비트스트림의 일부가도 10b 및 도 10c 의 예에서 도시된다.

[1074] 이러한 방식으로, 오디오 인코딩 디바이스 (510H) 는 다음의 조항들에 대해 이하에 진술되는 기술들의 여러 양태들을 수행할 수도 있다.

[1075] 조항 141541-1A. 오디오 인코딩 디바이스 (510H) 와 같은 디바이스는: 음장의 공간 성분의 압축된 버전을

포함하는 비트스트림을 획득하도록 구성된 하나 이상의 프로세서들을 포함하고, 그 공간 성분은 복수의 구면 조화 계수들에 대해 벡터 기반 합성을 수행함으로써 생성된다.

[1076] 조항 141541-2A. 조항 141541-1A 의 디바이스, 여기서 공간 성분의 압축된 버전은 적어도 부분적으로 공간 성분을 압축할 때 사용되는 예측 모드를 명시하는 필드를 사용하여 비트스트림에서 표현된다.

[1077] 조항 141541-3A. 조항 141541-1A 및 조항 141541-2A 의 임의의 조합의 디바이스, 여기서 공간 성분의 압축된 버전은 적어도 부분적으로 공간 성분을 압축할 때 사용되는 허프만 테이블을 명시하는 허프만 테이블 정보를 사용하여 비트스트림에서 표현된다.

[1078] 조항 141541-4A. 조항 141541-1A 내지 조항 141541-3A 의 임의의 조합의 디바이스, 여기서 공간 성분의 압축된 버전은 적어도 부분적으로 공간 성분을 압축할 때 사용되는 양자화 스텝 사이즈 또는 이들의 변수를 표현하는 값을 나타내는 필드를 사용하여 비트스트림에서 표현된다.

[1079] 조항 141541-5A. 조항 141541-4A 의 디바이스, 여기서 그 값은 nbits 값을 포함한다.

[1080] 조항 141541-6A. 조항 141541-4A 및 조항 141541-5A 의 임의의 조합의 디바이스, 여기서 비트스트림은 그의 공간 성분의 압축된 버전이 포함되는 음장의 복수의 공간 성분들의 압축된 버전을 포함하고, 여기서 그 값은 복수의 공간 성분들을 압축할 때 사용되는 양자화 스텝 사이즈 및 그의 변수를 표현한다.

[1081] 조항 141541-7A. 조항 141541-1A 내지 조항 141541-6A 의 임의의 조합의 디바이스, 여기서 공간 성분의 압축된 버전은 적어도 부분적으로 공간 성분이 대응하는 압축 카테고리를 식별하는 카테고리 식별자를 표현하기 위해 허프만 코드를 사용하여 비트스트림에서 표현된다.

[1082] 조항 141541-8A. 조항 141541-1A 내지 조항 141541-7A 의 임의의 조합의 디바이스, 여기서 공간 성분의 압축된 버전은 적어도 부분적으로 공간 성분이 포지티브 값인지 또는 네거티브 값인지 여부를 식별하는 부호 비트를 사용하여 비트스트림에서 표현된다.

[1083] 조항 141541-9A. 조항 141541-1A 내지 조항 141541-8A 의 임의의 조합의 디바이스, 여기서 공간 성분의 압축된 버전은 적어도 부분적으로 공간 성분의 잔차 값을 표현하기 위해 허프만 코드를 사용하여 비트스트림에서 표현된다.

[1084] 조항 141541-10A. 조항 141541-1A 내지 조항 141541-9A 의 임의의 조합의 디바이스, 여기서 디바이스는 오디오 인코딩 디바이스 및 비트스트림 생성 디바이스를 포함한다.

[1085] 조항 141541-12A. 조항 141541-1A 내지 조항 141541-11A 의 임의의 조합의 디바이스, 여기서 벡터 기반 합성은 특이값 분해를 포함한다.

[1086] 오디오 인코딩 디바이스 (510H) 에 의해 수행되는 것으로서 기술되지만, 그 기술들은 또한 오디오 디코딩 디바이스들 (24 및/또는 540) 중 임의의 것에 의해 수행될 수도 있다.

[1087] 이러한 방식으로, 오디오 인코딩 디바이스 (510H) 는 다음의 조항들에 대해 아래에서 진술된 기술들의 여러 양태들을 추가적으로 수행할 수도 있다.

[1088] 조항 141541-1D. 오디오 인코딩 디바이스 (510H) 와 같은 디바이스는: 음장의 공간 성분의 압축된 버전을 포함하는 비트스트림을 생성하도록 구성된 하나 이상의 프로세서들을 포함하고, 그 공간 성분은 복수의 구면 조화 계수들에 대해 벡터 기반 합성을 수행함으로써 생성된다.

[1089] 조항 141541-2D. 조항 141541-1D 의 디바이스, 여기서 하나 이상의 프로세서들은 또한, 비트스트림을 생성할 때, 공간 성분을 압축할 때 사용되는 예측 모드를 명시하는 필드를 포함하는 비트스트림을 생성하도록 구성된다.

[1090] 조항 141541-3D. 조항 141541-1D 및 조항 141541-2D 의 임의의 조합의 디바이스, 여기서, 하나 이상의 프로세서들은 또한, 비트스트림을 생성할 때, 공간 성분을 압축할 때 사용되는 허프만 테이블을 명시하는 허프만 테이블 정보를 포함하는 비트스트림을 생성하도록 구성된다.

[1091] 조항 141541-4D. 조항 141541-1D 내지 조항 141541-3D 의 임의의 조합의 디바이스, 여기서, 하나 이상의 프로세서들은 또한, 비트스트림을 생성할 때, 공간 성분을 압축할 때 사용되는 양자화 스텝 사이즈 또는 그의 변수를 표현하는 값을 나타내는 필드를 포함하는 비트스트림을 생성하도록 구성된다.

[1092] 조항 141541-5D. 조항 141541-4D 의 디바이스, 여기서 그 값은 nbits 값을 포함한다.

[1093] 조항 141541-6D. 조항 141541-4D 및 조항 141541-5D 의 임의의 조합의 디바이스, 여기서 하나 이상의 프로세서들은 또한, 비트스트림을 생성할 때, 그의 공간 성분의 압축된 버전이 포함되는 음장의 복수의 공간 성분들의 압축된 버전을 포함하는 비트스트림을 생성하도록 더 구성되며, 여기서 그 값은 복수의 공간 성분들을 압축할 때 사용되는 양자화 스텝 사이즈 또는 그의 변수를 표현한다.

[1094] 조항 141541-7D. 조항 141541-1D 내지 조항 141541-6D 의 임의의 조합의 디바이스, 여기서 하나 이상의 프로세서들은 또한, 비트스트림을 생성할 때, 공간 성분이 대응하는 압축 카테고리를 식별하는 카테고리 식별자를 나타내는 허프만 코드를 포함하는 비트스트림을 생성하도록 구성된다.

[1095] 조항 141541-8D. 조항 141541-1D 내지 조항 141541-7D 의 임의의 조합의 디바이스, 여기서 하나 이상의 프로세서들은 또한, 비트스트림을 생성할 때, 공간 성분이 포지티브 값인지 또는 네거티브 값인지 여부를 식별하는 부호 비트를 포함하는 비트스트림을 생성하도록 구성된다.

[1096] 조항 141541-9D. 조항 141541-1D 내지 조항 141541-8D 의 임의의 조합의 디바이스, 여기서 하나 이상의 프로세서들은 또한, 비트스트림을 생성할 때, 공간 성분의 잔차 값을 표혐하는 허프만 코드를 포함하는 비트스트림을 생성하도록 구성된다.

[1097] 조항 141541-10D. 조항 141541-1D 내지 조항 141541-10D 의 임의의 조합의 디바이스, 여기서 벡터 기반 합성은 특이값 분해를 포함한다.

[1098] 오디오 인코딩 디바이스 (510H) 는 또한 다음의 조항들에서 진술된 기술들의 여러 양태들을 구현하도록 구성될 수도 있다.

[1099] 조항 141541-1E. 오디오 인코딩 디바이스 (510H) 와 같은 디바이스는: 음장의 공간 성분을 압축하도록 구성된 하나 이상의 프로세서들을 포함하며, 그 공간 성분은 복수의 구면 조화 계수들에 대해 벡터 기반 합성을 수행함으로써 생성된다.

[1100] 조항 141541-2E. 조항 141541-1E 의 디바이스, 여기서 하나 이상의 프로세서들은 또한, 공간 성분을 압축할 때, 공간 성분을 부동소수점 표현으로부터 정수 표현으로 변환하도록 구성된다.

[1101] 조항 141541-3E. 조항 141541-1E 및 조항 141541-2E 의 임의의 조합의 디바이스, 여기서 하나 이상의 프로세서들은 또한, 공간 성분을 압축할 때, 양자화 스텝 사이즈를 표현하는 값을 동적으로 결정하고, 양자화된 공간 성분을 생성하기 위해 그 값에 기초하여 공간 성분을 양자화하도록 구성된다.

[1102] 조항 141541-4E. 조항 141541-1E 내지 조항 141541-3E 의 임의의 조합의 디바이스, 여기서 하나 이상의 프로세서들은 또한, 공간 성분을 압축할 때, 공간 성분이 대응하는 카테고리를 식별하도록 구성된다.

[1103] 조항 141541-5E. 조항 141541-1E 내지 조항 141541-4E 의 임의의 조합의 디바이스, 여기서 하나 이상의 프로세서들은 또한, 공간 성분을 압축할 때, 공간 성분에 대한 잔차 값을 식별하도록 구성된다.

[1104] 조항 141541-6E. 조항 141541-1E 내지 조항 141541-5E 의 임의의 조합의 디바이스, 여기서 하나 이상의 프로세서들은 또한, 공간 성분을 압축할 때, 예측된 공간 성분을 생성하기 위해 그 공간 성분 및 후속하는 공간 성분에 대해 예측을 수행하도록 구성된다.

[1105] 조항 141541-7E. 조항 141541-1E 의 임의의 조합의 디바이스, 여기서 하나 이상의 프로세서들은 또한, 공간 성분을 압축할 때, 공간 성분을 부동소수점 표현으로부터 정수 표현으로 변환하고, 양자화 스텝 사이즈를 표현하는 값을 동적으로 결정하며, 양자화된 공간 성분을 생성하기 위해 그 값에 기초하여 공간 성분의 정수 표현을 양자화하고, 카테고리 식별자를 생성하기 위해 양자화된 공간 성분에 기초하여 공간 성분이 대응하는 카테고리를 식별하며, 공간 성분의 부호를 결정하고, 양자화된 공간 성분 및 카테고리 식별자에 기초하여 공간 성분에 대한 잔차 값을 식별하며, 카테고리 식별자, 부호 및 잔차 값에 기초하여 공간 성분의 압축된 버전을 생성하도록 구성된다.

[1106] 조항 141541-8E. 조항 141541-1E 의 임의의 조합의 디바이스, 여기서 하나 이상의 프로세서들은 또한, 공간 성분을 압축할 때, 공간 성분을 부동소수점 표현으로부터 정수 표현으로 변환하고, 양자화 스텝 사이즈를 표현하는 값을 동적으로 결정하며, 양자화된 공간 성분을 생성하기 위해 그 값에 기초하여 공간 성분의 정수 표현을 양자화하고, 예측된 공간 성분을 생성하기 위해 그 공간 성분 및 후속하는 공간 성분에 대해 예측을 수행하며, 카테고리 식별자를 생성하기 위해 양자화된 공간 성분에 기초하여 예측된 공간 성분이 대응하는 카테고리를 식

별하며, 공간 성분의 부호를 결정하고, 양자화된 공간 성분 및 카테고리 식별자에 기초하여 공간 성분에 대한 잔차 값을 식별하며, 카테고리 식별자, 부호 및 잔차 값에 기초하여 공간 성분의 압축된 버전을 생성하도록 구성된다.

[1107] 조항 141541-9E. 조항 141541-1E 내지 조항 141541-8E 의 임의의 조합의 디바이스, 여기서 벡터 기반 합성은 특이값 분해를 포함한다.

[1108] 기술들의 여러 양태들이 또한 오디오 인코딩 디바이스 (510H) 가 다음의 조항들에서 진술된 바와 같이 동작하도록 구성되는 것을 가능하게 할 수도 있다.

[1109] 조항 141541-1F. 오디오 인코딩 디바이스 (510H) 와 같은 디바이스는: 복수의 공간 성분들 중 남아있는 것들에 대한 현재의 공간 성분의 오더에 기초하여 복수의 공간 성분들 중 현재의 공간 성분을 압축할 때 사용할 허프만 코드북을 식별하도록 구성된 하나 이상의 프로세서들을 포함하며, 공간 성분은 복수의 구면 조화 계수들에 대해 벡터 기반 합성을 수행함으로써 생성된다.

[1110] 조항 141541-2F. 조항 141541-3F 의 디바이스, 여기서, 하나 이상의 프로세서들은 또한 조항 141541-1A 내지 조항 141541-12A, 조항 141541-1B 내지 조항 141541-10B, 및 조항 141541-1C 내지 조항 141541-9C 에 기재된 단계들의 임의의 조합을 수행하도록 구성된다.

[1111] 기술들의 여러 양태들이 또한 오디오 인코딩 디바이스 (510H) 가 다음의 조항들에서 진술된 바와 같이 동작하도록 구성되는 것을 가능하게 할 수도 있다.

[1112] 조항 141541-1H. 오디오 인코딩 디바이스 (510H) 와 같은 디바이스는: 음장의 공간 성분을 압축할 때 사용될 양자화 스텝 사이즈를 결정하도록 구성된 하나 이상의 프로세서들을 포함하며, 공간 성분은 복수의 구면 조화 계수들에 대해 벡터 기반 합성을 수행함으로써 생성된다.

[1113] 조항 141541-2H. 조항 141541-1H 의 디바이스, 여기서 하나 이상의 프로세서들은 또한, 양자화 스텝 사이즈를 결정할 때, 목표 비트 레이트에 기초하여 양자화 스텝 사이즈를 결정하도록 구성된다.

[1114] 조항 141541-3H. 조항 141541-1H 의 디바이스, 여기서 하나 이상의 프로세서들은 또한, 복수의 양자화 스텝 사이즈들 중 하나를 선택할 때, 공간 성분을 표현하기 위해 사용되는 비트들의 수의 추정을 결정하고, 그 추정 및 목표 비트 레이트 사이의 차이에 기초하여 양자화 스텝 사이즈를 결정하도록 구성된다.

[1115] 조항 141541-4H. 조항 141541-1H 의 디바이스, 여기서 하나 이상의 프로세서들은 또한, 복수의 양자화 스텝 사이즈들 중 하나를 선택할 때, 공간 성분을 표현하기 위해 사용되는 비트들의 수의 추정을 결정하고, 그 추정 및 목표 비트 레이트 사이의 차이를 결정하며, 목표 비트 레이트에 그 차이를 가산함으로써 양자화 스텝 사이즈를 결정하도록 구성된다.

[1116] 조항 141541-5H. 조항 141541-3H 또는 조항 141541-4H 의 디바이스, 여기서 하나 이상의 프로세서들은 또한, 비트들의 수의 추정을 결정할 때, 목표 비트 레이트에 대응하는 코드 북이 주어지면 공간 성분에 대해 생성되어야 하는 비트들의 수의 추정을 계산하도록 구성된다.

[1117] 조항 141541-6H. 조항 141541-3H 또는 조항 141541-4H 의 디바이스, 여기서 하나 이상의 프로세서들은 또한, 비트들의 수의 추정을 결정할 때, 공간 성분을 압축할 때 사용되는 코딩 모드가 주어지면 공간 성분에 대해 생성되어야 하는 비트들의 수의 추정을 계산하도록 구성된다.

[1118] 조항 141541-7H. 조항 141541-3H 또는 조항 141541-4H 의 디바이스, 여기서 하나 이상의 프로세서들은 또한, 비트들의 수의 추정을 결정할 때, 공간 성분을 압축할 때 사용될 제 1 코딩 모드가 주어지면 공간 성분에 대해 생성되어야 하는 비트들의 수의 제 1 추정을 계산하고, 공간 성분을 압축할 때 사용될 제 2 코딩 모드가 주어지면 공간 성분에 대해 생성되어야 하는 비트들의 수의 제 2 추정을 계산하며, 비트들의 수의 결정된 추정으로서 사용될 비트들의 최소 수를 갖는 제 1 추정 및 제 2 추정 중 하나를 선택하도록 구성된다.

[1119] 조항 141541-8H. 조항 141541-3H 또는 조항 141541-4H 의 디바이스, 여기서 하나 이상의 프로세서들은 또한, 비트들의 수의 추정을 결정할 때, 공간 성분이 대응하는 카테고리를 식별하는 카테고리 식별자를 식별하고, 카테고리에 대응하는 공간 성분을 압축할 때 초래할 공간 성분에 대한 잔차 값의 비트 길이를 식별하며, 적어도 부분적으로 그 잔차 값의 비트 길이에 카테고리 식별자를 나타내는데 사용되는 비트들의 수를 가산함으로써 비트들의 수의 추정을 결정하도록 구성된다.

[1120] 조항 141541-9H. 조항 141541-1H 내지 조항 141541-8H 의 임의의 조합의 디바이스, 여기서 벡터 기반 합성

은 특이값 분해를 포함한다.

[1121] 비록 오디오 인코딩 디바이스 (510H) 에 의해 수행되는 것으로서 기술되지만, 전술된 조항들인 조항 141541-1H 내지 조항 141541-9H 에서 진술된 기술들은 또한 오디오 디코딩 디바이스 (540D) 에 의해 수행될 수도 있다.

[1122] 추가적으로, 기술들의 여러 양태들은 오디오 인코딩 디바이스 (510H) 가 다음의 조항들에서 진술된 바와 같이 동작하도록 구성되는 것을 가능하게 할 수도 있다.

[1123] 조항 141541-1J. 오디오 인코딩 디바이스 (510J) 와 같은 디바이스는: 음장의 공간 성분을 압축할 때 사용될 복수의 코드북들 중 하나를 선택하도록 구성된 하나 이상의 프로세서들을 포함하며, 공간 성분은 복수의 구면 조화 계수들에 대해 벡터 기반 합성을 수행함으로써 생성된다.

[1124] 조항 141541-2J. 조항 141541-1J 의 디바이스, 여기서 하나 이상의 프로세서들은 또한, 복수의 코드 북들 중 하나를 선택할 때, 복수의 코드 북들 각각을 사용하여 공간 성분을 표현하기 위해 사용되는 비트들의 수의 추정을 결정하고, 비트들의 최소 수를 갖는 결정된 추정을 야기하는 복수의 코드북들 중 하나를 선택하도록 구성된다.

[1125] 조항 141541-3J. 조항 141541-1J 의 디바이스, 여기서 하나 이상의 프로세서들은 또한, 복수의 코드 북들 중 하나를 선택할 때, 복수의 코드 북들 중 하나 이상을 사용하여 공간 성분을 나타내는데 사용되는 비트들의 수의 추정을 결정하도록 구성되고, 복수의 코드 북들 중 하나 이상은 공간 성분의 다른 엘리먼트들에 대해 압축될 공간 성분의 엘리먼트들의 오더에 기초하여 선택된다.

[1126] 조항 141541-4J. 조항 141541-1J 의 디바이스, 여기서 하나 이상의 프로세서들은 또한, 복수의 코드 북들 중 하나를 선택할 때, 공간 성분이 후속하는 공간 성분으로부터 예측되지 않을 때 사용되도록 설계된 복수의 코드 북들 중 하나를 사용하여 공간 성분을 표현하기 위해 사용되는 비트들의 수의 추정을 결정하도록 구성된다.

[1127] 조항 141541-5J. 조항 141541-1J 의 디바이스, 여기서 하나 이상의 프로세서들은 또한, 복수의 코드 북들 중 하나를 선택할 때, 공간 성분이 후속하는 공간 성분으로부터 예측될 때 사용되도록 설계된 복수의 코드 북들 중 하나를 사용하여 공간 성분을 표현하기 위해 사용되는 비트들의 수의 추정을 결정하도록 구성된다.

[1128] 조항 141541-6J. 조항 141541-1J 의 디바이스, 여기서 하나 이상의 프로세서들은 또한, 복수의 코드 북들 중 하나를 선택할 때, 공간 성분이 음장 내의 합성 오디오 오브젝트를 나타낼 때 사용되도록 설계된 복수의 코드 북들 중 하나를 사용하여 공간 성분을 표현하기 위해 사용되는 비트들의 수의 추정을 결정하도록 구성된다.

[1129] 조항 141541-7J. 조항 141541-1J 의 디바이스, 여기서 합성 오디오 오브젝트는 펄스 코드 변조된 (PCM) 오디오 오브젝트를 포함한다.

[1130] 조항 141541-8J. 조항 141541-1J 의 디바이스, 여기서 하나 이상의 프로세서들은 또한, 복수의 코드 북들 중 하나를 선택할 때, 공간 성분이 음장 내의 레코딩된 오디오 오브젝트를 나타낼 때 사용되도록 설계된 복수의 코드 북들 중 하나를 사용하여 공간 성분을 표현하기 위해 사용되는 비트들의 수의 추정을 결정하도록 구성된다.

[1131] 조항 141541-9J. 조항 141541-1J 내지 조항 141541-8J 의 임의의 조합의 디바이스, 여기서 벡터 기반 합성은 특이값 분해를 포함한다.

[1132] 전술된 여러 예들 각각에서, 오디오 인코딩 디바이스 (510) 는 방법을 수행하거나 또는 다르게는 오디오 인코딩 디바이스 (510) 가 수행하도록 구성되는 방법의 각 단계를 수행하는 수단을 포함할 수도 있다는 것이 이해되어야 한다. 일부 예들에서, 이들 수단은 하나 이상의 프로세서들을 포함할 수도 있다. 일부 예들에서, 하나 이상의 프로세서들은 비-일시적 컴퓨터 판독가능 저장 매체에 저장된 명령들에 의해 구성된 특수 목적 프로세서를 나타낼 수도 있다. 즉, 인코딩 예들의 세트들 각각에서의 기술들의 여러 양태들은, 실행될 때 하나 이상의 프로세서들로 하여금 오디오 인코딩 디바이스 (510) 가 수행하도록 구성된 방법을 수행하게 하는 명령들을 저장한 비-일시적 컴퓨터 판독가능 저장 매체를 제공할 수도 있다.

[1133] 도 40j 는 2 또는 3 차원 음장들을 기술하는 구면 조화 계수들을 압축하기 위해 본 개시물에서 기술된 기술들의 여러 양태들을 수행할 수도 있는 예시의 오디오 인코딩 디바이스 (510J) 를 예시하는 블록 다이어그램이다. 오디오 인코딩 디바이스 (510J) 는 오디오 인코딩 디바이스 (510J) 가 오디오 압축 유닛 (512), 오디오 인코딩

유닛 (514) 및 비트스트림 생성 유닛 (516) 을 포함한다는 점에서 오디오 인코딩 디바이스 (510G) 와 유사할 수도 있다. 게다가, 오디오 인코딩 디바이스 (510J) 의 오디오 압축 유닛 (512) 은 오디오 압축 유닛 (512) 이 오디오 인코딩 디바이스 (510I) 의 유사한 유닛들과 유사하게 동작할 수도 있는 분해 유닛 (518) 및 음장 성분 추출 유닛 (520) 을 포함한다는 점에서 오디오 인코딩 디바이스 (510G) 의 그것과 유사할 수도 있다. 일부 예들에서, 오디오 인코딩 디바이스 (510J) 는 U_{DIST} 벡터들 (525C), U_{BG} 벡터들 (525D), V^{T}_{DIST} 벡터들 (525E), 및 V^{T}_{BG} 벡터들 (525J) 의 임의의 것의 하나 이상의 벡터들을 양자화하기 위해, 도 40D 내지 도 40E 에 대해 기술된, 양자화 유닛 (534) 을 포함할 수도 있다.

[1134] 오디오 인코딩 디바이스 (510J) 의 오디오 압축 유닛 (512) 은, 그러나 오디오 인코딩 디바이스 (510J) 의 오디오 압축 유닛 (512) 이 보간 유닛 (550) 으로서 표시된 추가적인 유닛을 포함한다는 점에서 오디오 인코딩 디바이스 (510G) 의 오디오 압축 유닛 (512) 과 상이할 수도 있다. 보간 유닛 (550) 은 도 45 및 도 45b 에 대해 이하에 더 상세하게 기술되는 바와 같이, 제 1 오디오 프레임 및 제 2 의 시간적으로 후속하거나 선행하는 오디오 프레임의 서브 프레임들로부터 제 1 오디오 프레임의 서브프레임들을 보간하는 유닛을 나타낼 수도 있다. 보간 유닛 (550) 은, 이러한 보간을 수행함에 있어서, 분해 유닛 (518) 이 SHC (511) 를 분해하는데 요구되는 정도를 잠재적으로 감소시킴으로써 (프로세싱 사이클들 및/또는 메모리 소비 면에서) 컴퓨테이션 복잡성을 감소시킬 수도 있다. 보간 유닛 (550) 은 도 40h 및 도 40i 의 예들에서 도시된 오디오 인코딩 디바이스들 (510H 및 510I) 의 보간 유닛 (550) 에 대해 전술된 것과 유사한 방식으로 동작할 수도 있다.

[1135] 동작에 있어서, 보간 유닛 (200) 은 하나 이상의 서브프레임들에 대해 분해된 보간된 구면 조화 계수들을 생성하기 위해 제 1 프레임에 포함된 제 1 복수의 구면 조화 계수들 (11) 의 일부의 제 1 분해, 예를 들어 V 행렬 (19') 및 제 2 프레임에 포함된 제 2 복수의 구면 조화 계수들 (11) 의 일부의 제 2 분해, 예를 들어 V 행렬 (19') 로부터 제 1 오디오 프레임의 하나 이상의 서브프레임들을 보간할 수도 있다.

[1136] 보간 유닛 (550) 은 제 1 복수의 구면 조화 계수들의 제 1 분해 및 제 2 복수의 구면 조화 계수들의 제 2 분해에 대해 보간을, 적어도 부분적으로, 수행함으로써 시간 세그먼트에 대한 분해된 보간된 구면 조화 계수들을 획득할 수도 있다. 평활화 유닛 (554) 은 제 1 복수의 구면 조화 계수들 및 제 2 복수의 구면 조화 계수들의 공간 성분들 및 시간 성분들 중 적어도 하나를 평활화하기 위해 분해된 보간된 구면 조화 계수들을 적용할 수도 있다. 평활화 유닛 (554) 은 도 37 내지 도 39 에 대해 전술된 바와 같이 평활화된 U_{DIST} 행렬들 (525C') 를 생성할 수도 있다. 제 1 및 제 2 분해들은 도 40j 에서 V_1^{T} (556), 및 V_2^{T} (556B) 를 지칭할 수도 있다.

[1137] 일부 경우들에서, V^{T} 또는 다른 V-벡터들 또는 V-행렬들은 보간을 위해 양자화된 버전으로 출력될 수도 있다. 이러한 방식으로, 보간을 위한 V 벡터들은, 예를 들어 다차원 신호를 복구하기 위해, V 벡터 보간을 또한 수행하는 디코더에서의 V 벡터들과 동일할 수도 있다.

[1138] 일부 예들에서, 제 1 분해는 제 1 복수의 구면 조화 계수들 (511) 의 일부의 우-특이 벡터들을 나타내는 제 1 V 행렬 (519') 을 포함한다. 마찬가지로, 일부 예들에서, 제 2 분해는 제 2 복수의 구면 조화 계수들의 일부의 우-특이 벡터들을 나타내는 제 2 V 행렬 (519') 을 포함한다.

[1139] 보간 유닛 (550) 은 제 1 V 행렬 (519') 및 제 2 V 행렬 (519') 에 기초하여 하나 이상의 서브프레임들에 대해 시간적 보간을 수행할 수도 있다. 즉, 보간 유닛 (550) 은 제 1 오디오 프레임의 제 1 서브프레임으로부터 분해된 V 행렬 (519') 및 제 2 오디오 프레임의 제 1 서브프레임으로부터 분해된 V 행렬 (519') 에 기초하여 제 1 오디오 프레임에 대한 4 개의 총 서브프레임들로부터 예를 들어 제 2, 제 3 및 제 4 서브프레임들을 시간적으로 보간할 수도 있다. 일부 예들에서, 이러한 시간적 보간은 선형 시간적 보간이며, 여기서 제 1 오디오 프레임의 제 1 서브프레임으로부터 분해된 V 행렬 (519') 은 제 1 오디오 프레임의 제 4 서브프레임을 보간할 때 보다 제 1 오디오 프레임의 제 2 서브프레임을 보간할 때 더 무겁게 가중된다. 제 3 서브프레임을 보간할 때, V 행렬들 (519') 은 고르게 가중될 수도 있다. 제 4 서브프레임을 보간할 때, 제 2 오디오 프레임의 제 1 서브프레임으로부터 분해된 V 행렬 (519') 은 제 1 오디오 프레임의 제 1 서브프레임으로부터 분해된 V 행렬 (519') 보다 더 무겁게 가중될 수도 있다.

[1140] 즉, 선형 시간적 보간은 보간될 제 1 오디오 프레임의 서브프레임들 중 하나의 근접성이 주어지면 V 행렬들 (519') 을 가중할 수도 있다. 보간될 제 2 서브프레임의 경우, 제 1 오디오 프레임의 제 1 서브프레임으로

부터 분해된 V 행렬 (519') 은 제 2 오디오 프레임의 제 1 서브프레임으로부터 분해된 V 행렬 (519') 보다 보간될 제 2 서브프레임에 대한 그것의 근접성이 주어지면 더 무겁게 가중된다. 가중치들은 V 행렬들 (519') 에 기초하여 제 3 서브프레임을 보간할 때 이러한 이유로 등가일 수도 있다. 제 2 오디오 프레임의 제 1 서브프레임으로부터 분해된 V 행렬 (519') 에 인가된 가중치는, 보간될 제 4 서브프레임이 제 1 오디오 프레임의 제 1 서브프레임보다 제 2 오디오 프레임의 제 1 서브프레임에 더 근접하면 제 1 오디오 프레임의 제 1 서브프레임으로부터 분해된 V 행렬 (519') 에 인가된 것보다 더 클 수도 있다.

[1141] 일부 예들에서, 보간 유닛 (550) 은 제 1 투영된 분해들을 생성하기 위해 제 1 오디오 프레임의 제 1 서브프레임으로부터 분해된 제 1 V 행렬 (519') 을 공간 도메인으로 투영할 수도 있다. 일부 예들에서, 이러한 투영은 (예를 들어, T-설계 행렬과 같은 투영 행렬을 사용하여) 구면 (sphere) 으로의 투영을 포함한다. 보간 유닛 (550) 은 그 후 제 2 투영된 분해들을 생성하기 위해 제 2 오디오 프레임의 제 1 서브프레임으로부터 분해된 제 2 V 행렬 (519') 을 공간 도메인으로 투영할 수도 있다. 보간 유닛 (550) 은 그 후 제 1 공간적으로 보간된 투영된 분해 및 제 2 공간적으로 보간된 투영된 분해를 생성하기 위해 제 1 투영된 분해들 및 제 2 투영된 분해들을 공간적으로 보간할 수도 있다 (이것은 다시 선형 보간일 수도 있다). 보간 유닛 (550) 은 그 후 제 1 공간적으로 보간된 투영된 분해 및 제 2 공간적으로 보간된 투영된 분해에 기초하여 하나 이상의 서브프레임들을 시간적으로 보간할 수도 있다.

[1142] 보간 유닛 (550) 이 V 행렬 (519') 을 공간적으로 그리고 그 후 시간적으로 투영하는 이들 예들에서, 보간 유닛 (550) 은 구면 조화 도메인으로 하나 이상의 서브프레임들을 다시 보간하는 것으로부터 야기되는 시간적으로 보간된 구면 조화 계수들을 투영할 수도 있어, V 행렬 (519), S 행렬 (519B) 및 U 행렬 (519C) 을 생성한다.

[1143] 일부 예들에서, 제 1 복수의 구면 조화 계수들의 일부는 제 1 복수의 구면 조화 계수들 (511) 의 단일의 서브프레임을 포함한다. 일부 예들에서, 제 2 복수의 구면 조화 계수들의 일부는 제 2 복수의 구면 조화 계수들 (511) 의 단일의 서브프레임을 포함한다. 일부 예들에서, V 행렬들 (19') 이 분해되는 이러한 단일의 서브프레임은 제 1 서브프레임이다.

[1144] 일부 예들에서, 제 1 프레임은 4 개의 서브프레임들로 분할된다. 이들 및 다른 예들에서, 제 1 복수의 구면 조화 계수들의 일부는 제 1 복수의 구면 조화 계수들 (511) 의 제 1 서브프레임만을 포함한다. 이들 및 다른 예들에서, 제 2 프레임은 4 개의 서브프레임들로 분할되고, 제 2 복수의 구면 조화 계수들 (511) 의 일부는 제 2 복수의 구면 조화 계수들 (511) 의 제 1 서브프레임만을 포함한다.

[1145] 비록, 일부 예들에서, 각 오디오 프레임의 제 1 서브프레임만이 보간을 수행하는데 사용되지만, 제 1 복수의 구면 조화 계수들의 일부는 제 1 복수의 구면 조화 계수들 (511) 의 4 개의 서브프레임들 중 2 개를 포함할 수도 있다. 이들 및 다른 예들에서, 제 2 복수의 구면 조화 계수들 (511) 은 제 2 복수의 구면 조화 계수들 (511) 의 4 개의 서브프레임들 중 2 개를 포함한다.

[1146] 전술된 바와 같이, 단일의 디바이스, 예를 들어 오디오 인코딩 디바이스 (510J) 는 제 1 복수의 구면 조화 계수들의 일부의 제 1 분해들을 생성하기 위해 제 1 복수의 구면 조화 계수들의 일부를 또한 분해하면서 보간을 수행할 수도 있다. 이들 및 다른 예들에서, 분해 유닛 (518) 은 제 2 복수의 구면 조화 계수들의 일부의 제 2 분해들을 생성하기 위해 제 2 복수의 구면 조화 계수들의 일부를 분해할 수도 있다. 단일의 디바이스에 대해 기술되지만, 2 이상의 디바이스들이 본 개시물에 기술된 기술들을 수행할 수도 있고, 여기서 본 개시물에 기술된 기술들에 따라 2 개의 디바이스들 중 하나가 분해를 수행하고, 그 디바이스들 중 다른 하나는 보간을 수행한다.

[1147] 일부 예들에서, 분해 유닛 (518) 은 제 1 복수의 구면 조화 계수들 (511) 의 우-특이 벡터들을 나타내는 V 행렬 (519') (뿐아니라 설명의 편의 목적으로 도시되지 않은 S 행렬 (519B') 및 U 행렬 (519C')) 을 생성하기 위해 제 1 복수의 구면 조화 계수들 (511) 의 일부에 대해 특이값 분해를 수행할 수도 있다. 이들 및 다른 예들에서, 분해 유닛 (518) 은 제 2 복수의 구면 조화 계수들의 우-특이 벡터들을 나타내는 V 행렬 (519') (뿐아니라 설명의 편의 목적으로 도시되지 않은 S 행렬 (519B') 및 U 행렬 (519C')) 을 생성하기 위해 제 2 복수의 구면 조화 계수들 (511) 의 일부에 대해 특이값 분해를 수행할 수도 있다.

[1148] 일부 예들에서, 전술된 바와 같이, 제 1 및 제 2 복수의 구면 조화 계수들 각각은 음장의 평면파 표현을 나타낸다. 이들 및 다른 예들에서, 제 1 및 제 2 복수의 구면 조화 계수들 (511) 각각은 함께 혼합된 하나 이상의 모노-오디오 오브젝트들을 나타낸다.

[1149] 즉, 구면 조화 함수 기반 3D 오디오는 구면에 대한 직교 기저 함수들의 견지에서 3D 압력장의 파라미터 표현일

수도 있다. 그 표현의 오더 (N) 가 더 높을 수록, 공간 해상도는 잠재적으로 더 높고, 종종 (총 $(N+1)^2$ 계수들에 대해) 구면 조화 (SH) 계수들의 수가 더 크다. 많은 애플리케이션들의 경우, 계수들의 대역폭 압축이 계수들을 효율적으로 송신 및 저장할 수 있기 위해 요구될 수도 있다. 본 개시물에 지향된 이러한 기술들은 특이값 분해 (SVD) 를 사용하는 프레임 기반 차원수 감소 프로세스를 제공할 수도 있다. SVD 분석은 계수들의 각 프레임을 3 개의 행렬들 U, S 및 V 로 분해할 수도 있다. 일부 예들에서, 기술들은 U 에서의 벡터들의 일부를 근원적인 음장의 방향성 성분들로서 핸들링할 수도 있다. 그러나, 이러한 방식으로 핸들링되는 경우, (U 에서의) 이들 벡터들은 그들이 동일한 구별되는 오디오 성분을 나타낼지라도 프레임에서 프레임으로 불연속이다. 이들 불연속성들은 성분들이 변환-오디오-코더들을 통해 피딩되는 경우 상당한 아티팩트들을 초래할 수도 있다.

[1150] 이러한 개시물에 기술된 기술들은 이러한 불연속성을 다룰 수도 있다. 즉, 그 기술들은 V 행렬이 구면 조화 함수 도메인에서 직교 공간 축들로서 해석될 수 있다는 관찰에 기초할 수도 있다. U 행렬은 이들 기저 함수들의 견지에서 구면 조화 함수 (HOA) 데이터의 투영을 표현할 수도 있으며, 여기서 불연속성은 매 프레임마다 변하고, 따라서 그것들 자체가 불연속인 기저 함수들 (V) 에 귀착될 수 있다. 이것은 푸리에 변환과 같은 유사한 분해와 다르며, 여기서 기저 함수들은, 일부 예들에서 프레임에서 프레임으로 일정하다. 이들 조건에서, SVD 는 매칭 추구 알고리즘으로서 고려될 수도 있다. 본 개시물에 기술된 기술들은 보간 유닛 (550) 이 프레임에서 프레임으로 기저 함수들 (V) 사이의 연속성을 그들 사이에 보간함으로써 유지하는 것을 가능하게 할 수도 있다.

[1151] 일부 예들에서, 그 기술들은 보간 유닛 (550) 이 전술된 그리고 도 45 및 도 45b 에 대해 이하에 더 기술된 바와 같이 SH 데이터의 프레임을 4 개의 서브프레임들로 분할하는 것을 가능하게 한다. 보간 유닛 (550) 은 그 후 제 1 서브프레임에 대해 SVD 를 컴퓨팅할 수도 있다. 유사하게 우리는 제 2 프레임의 제 1 서브프레임에 대해 SVD 를 컴퓨팅한다. 제 1 프레임 및 제 2 프레임 각각에 대해, 보간 유닛 (550) 은 (T-설계 행렬과 같은 투영 행렬을 사용하여) 벡터들을 구면상으로 투영함으로써 V 에서의 벡터들을 공간 맵으로 변환할 수도 있다. 보간 유닛 (550) 은 그 후 V 에서의 벡터들을 구면상의 형상들로서 해석할 수도 있다. 제 1 프레임의 제 1 서브프레임과 다음 프레임의 제 1 서브프레임 사이에 3 개의 서브 프레임들에 대한 V 행렬들을 보간하기 위해, 보간 유닛 (550) 은 그 후 이들 공간 형상들을 보간하고, 그 후 투영 행렬의 역을 통해 그들을 다시 SH 벡터들로 변환할 수도 있다. 이러한 개시의 기술들은, 이러한 방식으로 V 행렬들 사이의 평활한 천이를 제공할 수도 있다.

[1152] 도 41 내지 도 41d 는 2 또는 3 차원 음장들을 기술하는 구면 조화 계수들을 디코딩하기 위해 본 개시물에서 기술된 기술들의 여러 양태들을 수행할 수도 있는 예시의 오디오 디코딩 디바이스 (540A-540D) 를 각각 예시하는 블록도들이다. 오디오 디코딩 디바이스 (540A) 는 데스크톱 컴퓨터, 랩톱 컴퓨터, 워크스테이션, 태블릿 또는 슬레이트 컴퓨터, 전용 오디오 레코딩 디바이스, (소위 "스마트 폰들" 을 포함하는) 셀룰러 폰, 개인용 미디어 플레이어 디바이스, 개인용 게이밍 디바이스, 또는 오디오 데이터를 디코딩할 수 있는 임의의 다른 타입의 디바이스와 같은 오디오 데이터를 디코딩할 수 있는 임의의 디바이스를 표현할 수도 있다.

[1153] 일부 예들에서, 오디오 디코딩 디바이스 (540A) 는, 일부 예들에서 외래의 무관계한 데이터의 제거를 용이하게 하기 위해 오디오 인코딩 디바이스들 (510B-510J) 에 의해 사용되는 (도 40b 내지 도 40j 의 예들에 대해 전술된 바와 같은) 오더 감소를 수행하는 것을 제외하고 오디오 인코딩 디바이스들 (510 또는 510B) 중 임의의 것에 의해 수행되는 오디오 인코딩 프로세스에 대해 역인 오디오 디코딩 프로세스를 수행한다.

[1154] 단일의 디바이스, 즉 도 41 의 예에서 디바이스 (540A) 로서 도시되지만, 디바이스 (540A) 내에 포함되는 것으로서 이하에 참조되는 여러 성분들 또는 유닛들은 디바이스 (540A) 로부터 외부에 있는 구별되는 디바이스들을 형성할 수도 있다. 다시 말해서, 단일의 디바이스, 즉 도 41 의 예에서 디바이스 (540A) 에 의해 수행되는 것으로서 본 개시물에서 기술되지만, 그 기술들은 다수의 디바이스들을 포함하는 시스템에 의해 구현되거나 또는 다르게는 수행될 수도 있으며, 여기서 이들 디바이스들 각각은 이하에 더 상세하게 기술되는 여러 성분들 또는 유닛들의 하나 이상을 각각 포함할 수도 있다. 이에 따라, 기술들은 이러한 점에서 도 41 의 예로 제한되지 않아야 한다.

[1155] 도 41 의 예에서 도시된 바와 같이, 오디오 디코딩 디바이스 (540A) 는 추출 유닛 (542), 오디오 디코딩 유닛 (544), 계산 유닛 (546), 및 오디오 렌더링 유닛 (548) 을 포함한다. 추출 유닛 (542) 은 비트스트림 (517) 으로부터 인코딩된 감소된 배경 구면 조화 계수들 (515B), 인코딩된 $U_{DIST} * S_{DIST}$ 벡터들 (515A) 및 V^T_{DIST}

벡터들 (525E) 를 추출하도록 구성된 유닛을 나타낸다. 추출 유닛 (542) 은 오디오 디코딩 유닛 (544) 으로 인코딩된 감소된 배경 구면 조화 계수들 (515B) 및 인코딩된 $U_{DIST} * S_{DIST}$ 벡터들 (515A) 를 출력하는 한편, 계산 유닛 (546) 으로 V^{T}_{DIST} 행렬 (525E) 을 또한 출력한다. 이러한 점에서, 추출 유닛 (542) 은 도 5 의 예에서 도시된 오디오 디코딩 디바이스 (24) 의 추출 유닛 (72) 과 유사한 방식으로 동작할 수도 있다.

[1156] 오디오 디코딩 유닛 (544) 은 $U_{DIST} * S_{DIST}$ 벡터들 (527) 및 감소된 배경 구면 조화 계수들 (529) 을 복구하기 위해 (종종 AAC 디코딩 방식과 같은 역의 오디오 디코딩 방식에 따라) 인코딩된 오디오 데이터를 디코딩하는 유닛을 나타낸다. 오디오 디코딩 유닛 (544) 은 계산 유닛 (546) 으로 $U_{DIST} * S_{DIST}$ 벡터들 (527) 및 감소된 배경 구면 조화 계수들 (529) 을 출력한다. 이러한 점에서, 오디오 디코딩 유닛 (544) 은 도 5 의 예에서 도시된 오디오 디코딩 디바이스 (24) 의 심리 음향 디코딩 유닛 (80) 과 유사한 방식으로 동작할 수도 있다.

[1157] 계산 유닛 (546) 은 행렬 곱셈 및 덧셈 (뿐아니라, 일부 예들에서, 임의의 다른 행렬 수학 동작) 을 수행하도록 구성된 유닛을 표현할 수도 있다. 계산 유닛 (546) 은 먼저 V^{T}_{DIST} 행렬 (525E) 에 의한 $U_{DIST} * S_{DIST}$ 벡터들 (527) 의 행렬 곱셈을 수행할 수도 있다. 계산 유닛 (546) 은 그 후 복구된 구면 조화 계수들 (547) 로서 표시되는 원래의 구면 조화 계수들 (11) 의 감소된 버전을 생성하기 위해 V^{T}_{DIST} 행렬 (525E) 에 의한 $U_{DIST} * S_{DIST}$ 벡터들 (527) 의 행렬 곱셈의 결과에 (다시 S_{BG} 행렬 (525B) 에 의한 그리고 그 후 V^{T}_{BG} 행렬 (525F) 에 의한 U_{BG} 행렬 (525D) 의 곱셈의 결과로 지칭될 수도 있는) 감소된 배경 구면 조화 계수들 (529) 에 의해 V^{T}_{DIST} 행렬 (525E) 에 의한 $U_{DIST} * S_{DIST}$ 벡터들 (527) 의 곱셈의 결과를 가산할 수도 있다. 계산 유닛 (546) 은 오디오 렌더링 유닛 (548) 으로 복구된 구면 조화 계수들 (547) 을 출력할 수도 있다. 이러한 점에서, 계산 유닛 (546) 은 도 5 의 예에서 도시된 오디오 디코딩 디바이스 (24) 의 전경 공식화 유닛 (78) 및 HOA 계수 공식화 유닛 (82) 과 유사한 방식으로 동작할 수도 있다.

[1158] 오디오 렌더링 유닛 (548) 은 채널들 (549A-549N) ("다중 채널 오디오 데이터 (549)" 로서 또는 "라우드스피커 피드들 (549)" 로서 또한 일반적으로 지칭될 수도 있는 "채널들 (549)") 을 렌더링하도록 구성된 유닛을 표현할 수도 있다. 오디오 렌더링 유닛 (548) 은 복구된 구면 조화 계수들 (547) 에 (종종 행렬의 형태로 표현되는) 변환을 적용할 수도 있다. 복구된 구면 조화 계수들 (547) 이 3 차원에서 음장을 기술하기 때문에, 복구된 구면 조화 계수들 (547) 은 (다중 채널 오디오 데이터 (549) 를 재생할 스피커들의 지오메트리를 지칭할 수도 있는) 대부분의 디코더-로컬 스피커 지오메트리들을 수용할 수 있는 방식으로 다중채널 오디오 데이터 (549A) 의 렌더링을 용이하게 하는 오디오 포맷을 표현한다. 다중채널 오디오 데이터 (549A) 의 렌더링에 관한 더 많은 정보는 도 48 에 대해 위에서 기술되어 있다.

[1159] 서라운드 사운드 다중채널 오디오 데이터 (549) 인 다중채널 오디오 데이터 (549A) 의 컨텍스트에서 기술되는 반면, 오디오 렌더링 유닛 (48) 은 또한 복구된 구면 조화 계수들 (549A) 을 바이노럴 (binaural) 화하기 위해 바이노럴화의 형성을 수행하여, 2 개의 바이노럴적으로 렌더링된 채널들 (549) 을 생성할 수도 있다. 이에 따라, 그 기술들은 다중 채널 오디오 데이터의 서라운드 사운드 형태들에 제한되는 것이 아니라, 바이노럴화된 다중 채널 오디오 데이터를 포함할 수도 있다.

[1160] 이하에 리스트된 여러 조항들은 본 개시물에 기술된 기술들의 여러 양태들을 제시할 수도 있다.

[1161] 조항 132567-1B. 오디오 디코딩 디바이스 (540) 와 같은 디바이스는: 음장의 구별되는 성분들을 기술하는 하나 이상의 제 1 벡터들 및 음장의 배경 성분들을 기술하는 하나 이상의 제 2 벡터들을 결정하도록 구성된 하나 이상의 프로세서들을 포함하고, 하나 이상의 제 1 벡터들 및 하나 이상의 제 2 벡터들 양자 모두는 적어도 복수의 구면 조화 계수들에 대해 특이값 분해를 수행함으로써 생성된다.

[1162] 조항 132567-2B. 조항 132567-1B 의 디바이스, 여기서 하나 이상의 제 1 벡터들은, 오디오 인코딩 이전에, S 행렬의 하나 이상의 S_{DIST} 벡터들에 의해 U 행렬의 하나 이상의 오디오 인코딩된 U_{DIST} 벡터들을 곱함으로써 생성되었던 하나 이상의 오디오 인코딩된 $U_{DIST} * S_{DIST}$ 벡터들을 포함하고, 여기서 U 행렬 및 S 행렬은 적어도 복수의 구면 조화 계수들에 대해 특이값 분해를 수행함으로써 생성되며, 여기서 하나 이상의 프로세서들은 또한 하나 이상의 오디오 인코딩된 $U_{DIST} * S_{DIST}$ 벡터들의 오디오 디코딩된 버전을 생성하기 위해 하나 이상의 오디오 인

코딩된 $U_{DIST} * S_{DIST}$ 벡터들을 오디오 디코딩하도록 구성된다.

[1163] 조항 132567-3B. 조항 132567-1B 의 디바이스, 여기서 하나 이상의 제 1 벡터들은, 오디오 인코딩 이전에, S 행렬의 하나 이상의 S_{DIST} 벡터들에 의해 U 행렬의 하나 이상의 오디오 인코딩된 U_{DIST} 벡터들을 곱함으로써 생성되었던 하나 이상의 오디오 인코딩된 $U_{DIST} * S_{DIST}$ 벡터들, 및 V 행렬의 트랜스포즈의 하나 이상의 V^{T}_{DIST} 벡터들을 포함하고, 여기서 U 행렬 및 S 행렬 및 V 행렬은 적어도 복수의 구면 조화 계수들에 대해 특이값 분해를 수행함으로써 생성되며, 여기서 하나 이상의 프로세서들은 또한 하나 이상의 오디오 인코딩된 $U_{DIST} * S_{DIST}$ 벡터들의 오디오 디코딩된 버전을 생성하기 위해 하나 이상의 오디오 인코딩된 $U_{DIST} * S_{DIST}$ 벡터들을 오디오 디코딩하도록 구성된다.

[1164] 조항 132567-4B. 조항 132567-1B 의 디바이스, 여기서 하나 이상의 프로세서들은 또한 음장의 구별되는 성분들을 나타내는 복수의 구면 조화 계수들의 성분들을 복구하기 위해 V^{T}_{DIST} 벡터들에 의해 $U_{DIST} * S_{DIST}$ 벡터들을 곱하도록 구성된다.

[1165] 조항 132567-5B. 조항 132567-1B 의 디바이스, 여기서 하나 이상의 제 2 벡터들은, 오디오 인코딩 이전에, S 행렬 내에 포함된 S_{BG} 벡터들에 의해 그리고 그 후 V 행렬의 트랜스포즈 내에 포함된 V^{T}_{BG} 벡터들에 의해 U 행렬 내에 포함된 U_{BG} 벡터들을 곱함으로써 생성되었던 하나 이상의 오디오 인코딩된 $U_{BG} * S_{BG} * V^{T}_{BG}$ 벡터들을 포함하고, 여기서 S 행렬, U 행렬 및 V 행렬은 적어도 복수의 구면 조화 계수들에 대해 특이값 분해를 수행함으로써 각각 생성되었다.

[1166] 조항 132567-6B. 조항 132567-1B 의 디바이스, 여기서 하나 이상의 제 2 벡터들은, 오디오 인코딩 이전에, S 행렬 내에 포함된 S_{BG} 벡터들에 의해 그리고 그 후 V 행렬의 컨주게이트 내에 포함된 V^{T}_{BG} 벡터들에 의해 U 행렬 내에 포함된 U_{BG} 벡터들을 곱함으로써 생성되었던 하나 이상의 오디오 인코딩된 $U_{BG} * S_{BG} * V^{T}_{BG}$ 벡터들을 포함하고, 여기서 S 행렬, U 행렬 및 V 행렬은 적어도 복수의 구면 조화 계수들에 대해 특이값 분해를 수행함으로써 각각 생성되었으며, 여기서 하나 이상의 프로세서들은 또한 하나 이상의 오디오 디코딩된 $U_{BG} * S_{BG} * V^{T}_{BG}$ 벡터들을 생성하기 위해 하나 이상의 오디오 인코딩된 $U_{BG} * S_{BG} * V^{T}_{BG}$ 벡터들을 오디오 디코딩하도록 구성된다.

[1167] 조항 132567-7B. 조항 132567-1B 의 디바이스, 여기서 하나 이상의 제 1 벡터들은, 오디오 인코딩 이전에, S 행렬의 하나 이상의 S_{DIST} 벡터들에 의해 U 행렬의 하나 이상의 오디오 인코딩된 U_{DIST} 벡터들을 곱함으로써 생성되었던 하나 이상의 오디오 인코딩된 $U_{DIST} * S_{DIST}$ 벡터들, 및 V 행렬의 트랜스포즈의 하나 이상의 V^{T}_{DIST} 벡터들을 포함하고, 여기서 S 행렬, U 행렬 및 V 행렬은 적어도 복수의 구면 조화 계수들에 대해 특이값 분해를 수행함으로써 각각 생성되었으며, 여기서 하나 이상의 프로세서들은 또한 하나 이상의 $U_{DIST} * S_{DIST}$ 벡터들을 생성하기 위해 하나 이상의 오디오 인코딩된 $U_{DIST} * S_{DIST}$ 벡터들을 오디오 디코딩하고, 음장의 구별되는 성분들을 기술하는 복수의 구면 조화 계수들의 성분들을 복구하기 위해 V^{T}_{DIST} 벡터들에 의해 $U_{DIST} * S_{DIST}$ 벡터들을 곱하도록 구성되며, 여기서 하나 이상의 제 2 벡터들은, 오디오 인코딩 이전에, S 행렬 내에 포함된 S_{BG} 벡터들에 의해 그리고 그 후 V 행렬의 트랜스포즈 내에 포함된 V^{T}_{BG} 벡터들에 의해 U 행렬 내에 포함된 U_{BG} 벡터들을 곱함으로써 생성되었던 하나 이상의 오디오 인코딩된 $U_{BG} * S_{BG} * V^{T}_{BG}$ 벡터들을 포함하고, 여기서 하나 이상의 프로세서들은 또한 음장의 배경 성분들을 기술하는 복수의 구면 조화 계수들의 적어도 일부를 복구하기 위해 하나 이상의 오디오 인코딩된 $U_{BG} * S_{BG} * V^{T}_{BG}$ 벡터들을 오디오 디코딩하고, 복수의 구면 조화 계수들의 재구성된 버전을 생성하기 위해 음장의 배경 성분들을 기술하는 복수의 구면 조화 계수들의 그 적어도 일부에 음장의 구별되는 성분들을 기술하는 복수의 구면 조화 계수들을 가산하도록 구성된다.

[1168] 조항 132567-8B. 조항 132567-1B 의 디바이스, 여기서 하나 이상의 제 1 벡터들은, 오디오 인코딩 이전에, S 행렬의 하나 이상의 S_{DIST} 벡터들에 의해 U 행렬의 하나 이상의 오디오 인코딩된 U_{DIST} 벡터들을 곱함으로써 생

성되었던 하나 이상의 $U_{DIST} * S_{DIST}$ 벡터들, 및 V 행렬의 트랜스포즈의 하나 이상의 V^{T}_{DIST} 벡터들을 포함하고, 여기서 U 행렬, S 행렬 및 V 행렬은 적어도 복수의 구면 조화 계수들에 대해 특이값 분해를 수행함으로써 생성되었으며, 여기서 하나 이상의 프로세서들은 또한 하나 이상의 $U_{DIST} * S_{DIST}$ 벡터들 및 하나 이상의 V^{T}_{DIST} 벡터들을 형성하기 위해 비트스트림으로부터 추출될 벡터들의 수를 나타내는 값 (D) 을 결정하도록 구성된다.

[1169] 조항 132567-9B. 조항 132567-1B 의 디바이스, 여기서 하나 이상의 제 1 벡터들은, 오디오 인코딩 이전에, S 행렬의 하나 이상의 S_{DIST} 벡터들에 의해 U 행렬의 하나 이상의 오디오 인코딩된 U_{DIST} 벡터들을 곱함으로써 생성되었던 하나 이상의 $U_{DIST} * S_{DIST}$ 벡터들, 및 V 행렬의 트랜스포즈의 하나 이상의 V^{T}_{DIST} 벡터들을 포함하고, 여기서 U 행렬, S 행렬 및 V 행렬은 적어도 복수의 구면 조화 계수들에 대해 특이값 분해를 수행함으로써 생성되었으며, 여기서 하나 이상의 프로세서들은 또한 하나 이상의 $U_{DIST} * S_{DIST}$ 벡터들 및 하나 이상의 V^{T}_{DIST} 벡터들을 형성하기 위해 비트스트림으로부터 추출될 벡터들의 수를 나타내는, 오디오-프레임-바이-오디오-프레임 기반의 값 (D) 을 결정하도록 구성된다.

[1170] 조항 132567-1G. 오디오 디코딩 디바이스 (540) 와 같은 디바이스는: 음장의 구별되는 성분들을 기술하는 하나 이상의 제 1 벡터들 및 음장의 배경 성분들을 기술하는 하나 이상의 제 2 벡터들을 결정하도록 구성된 하나 이상의 프로세서들을 포함하고, 하나 이상의 제 1 벡터들 및 하나 이상의 제 2 벡터들 양자 모두는 적어도 음장의 적어도 일부를 나타내는 다중 채널 오디오 데이터에 대해 특이값 분해를 수행함으로써 생성된다.

[1171] 조항 132567-2G. 조항 132567-1G 의 디바이스, 여기서 다중 채널 오디오 데이터는 복수의 구면 조화 계수들을 포함한다.

[1172] 조항 132567-3G. 조항 132567-2G 의 디바이스, 여기서 하나 이상의 프로세서들은 또한 조항 132567-2B 내지 조항 132567-9B 의 임의의 조합을 수행하도록 구성된다.

[1173] 전술된 여러 조항들 각각으로부터, 오디오 디코딩 디바이스들 (540A-540D) 중 임의의 것은 오디오 디코딩 디바이스들 (540A-540D) 이 수행하도록 구성되는 방법을 수행하거나 다르게는 그 방법의 각 단계를 수행하는 수단을 포함할 수도 있다는 것이 이해되어야 한다. 일부 예들에서, 이들 수단은 하나 이상의 프로세서들을 포함할 수도 있다. 일부 예들에서, 하나 이상의 프로세서들은 비-일시적 컴퓨터 판독가능 저장 매체에 저장된 명령들에 의해 구성된 특수 목적 프로세서를 표현할 수도 있다. 즉, 인코딩 예들의 세트들 각각에서의 기술들의 여러 양태들은, 실행될 때, 하나 이상의 프로세서들로 하여금 오디오 디코딩 디바이스들 (540A-540D) 이 수행하도록 구성되었던 방법을 수행하게 하는 명령들을 저장한 비-일시적 컴퓨터 판독가능 저장 매체를 제공할 수도 있다.

[1174] 예를 들어, 조항 132567-10B 는 음장의 구별되는 성분들을 기술하는 하나 이상의 제 1 벡터들 및 음장의 배경 성분들을 기술하는 하나 이상의 제 2 벡터들을 결정하는 단계를 포함하는 A 방법을 포함하는 방법이도록 상기 조항 132567-1B 로부터 도출될 수도 있고, 하나 이상의 제 1 벡터들 및 하나 이상의 제 2 벡터들 양자 모두는 적어도 음장을 나타내는 복수의 구면 조화 계수들에 대해 특이값 분해를 수행함으로써 생성된다.

[1175] 다른 예로서, 조항 132567-11B 는 음장의 구별되는 성분들을 기술하는 하나 이상의 제 1 벡터들 및 음장의 배경 성분들을 기술하는 하나 이상의 제 2 벡터들을 결정하는 수단으로서, 하나 이상의 제 1 벡터들 및 하나 이상의 제 2 벡터들 양자 모두는 적어도 복수의 구면 조화 계수들에 대해 특이값 분해를 수행함으로써 생성되는, 상기 결정하는 수단; 및 하나 이상의 제 1 벡터들 및 하나 이상의 제 2 벡터들을 저장하는 수단을 포함하는, 오디오 디코딩 디바이스 (540) 와 같은 디바이스이도록 상기 조항 132567-1B 로부터 도출될 수도 있다.

[1176] 또 다른 예로서, 조항 132567-12B 는, 실행될 때, 하나 이상의 프로세서로 하여금 음장의 구별되는 성분들을 기술하는 하나 이상의 제 1 벡터들 및 음장의 배경 성분들을 기술하는 하나 이상의 제 2 벡터들을 결정하게 하는 명령들을 저장한 비-일시적 컴퓨터 판독가능 저장 매체이도록 조항 132567-1B 로부터 도출될 수도 있고, 하나 이상의 제 1 벡터들 및 하나 이상의 제 2 벡터들 양자 모두는 적어도 음장을 기술하는 고차 앰비소닉스 오디오 데이터 내에 포함된 복수의 구면 조화 계수들에 대해 특이값 분해를 수행함으로써 생성된다.

[1177] 여러 조항들은 마찬가지로 위에서 예시된 바와 같이 도출되는 여러 디바이스들, 방법들 및 비-일시적 컴퓨터 판독가능 저장 매체들에 대해 조항 132567-2B 내지 조항 132567-9B 로부터 도출될 수도 있다. 본 개시물에 걸쳐 리스트된 여러 다른 조항들에 대해 동일한 것이 수행될 수도 있다.

[1178] 도 41b 는 2 또는 3 차원 음장들을 기술하는 구면 조화 계수들을 디코딩하기 위해 본 개시물에 기술된 기술들의 여러 양태들을 수행할 수도 있는 예시의 오디오 디코딩 디바이스 (540B) 를 예시하는 블록 다이어그램이다. 오디오 디코딩 디바이스 (540B) 는, 일부 예들에서, 추출 유닛 (542) 이 V^{T}_{DIST} 벡터들 (525E) 보다는 오히려 레코딩된 V^{T}_{DIST} 벡터들 (539) 을 추출할 수도 있다는 것을 제외하고, 오디오 디코딩 디바이스 (540) 와 유사할 수도 있다. 다른 예들에서, 추출 유닛 (542) 은 V^{T}_{DIST} 벡터들 (525E) 을 추출하고, 그 후 리오더링된 V^{T}_{DIST} 벡터들 (539) 를 결정하기 위해 (다른 벡터들의 분석을 통해) 추론되거나 비트스트림에서 특정된 리오더 (reorder) 정보에 기초하여 이들 V^{T}_{DIST} 벡터들 (525E) 을 리오더링할 수도 있다. 이러한 점에서, 추출 유닛 (542) 은 도 5 의 예에서 도시된 오디오 디코딩 디바이스 (24) 의 추출 유닛 (72) 과 유사한 방식으로 동작할 수도 있다. 어떤 경우든, 추출 유닛 (542) 은 계산 유닛 (546) 으로 리오더링된 V^{T}_{DIST} 벡터들 (539) 을 출력할 수도 있고, 여기서 구면 조화 계수들을 복구하는 것에 대해 전술된 프로세스는 이들 리오더링된 V^{T}_{DIST} 벡터들 (539) 에 대해 수행될 수도 있다.

[1179] 이러한 방식으로, 그 기술들은 오디오 디코딩 디바이스 (540B) 가 음장의 구별되는 성분들을 나타내는 리오더링된 하나 이상의 벡터들을 오디오 디코딩하는 것을 가능하게 할 수도 있고, 그 리오더링된 하나 이상의 벡터들은 하나 이상의 벡터들을 압축하는 것을 용이하게 하기 위해 리오더링되었다. 이들 및 다른 예들에서, 오디오 디코딩 디바이스 (540B) 는 음장의 구별되는 성분들을 나타내는 구면 조화 계수들을 복구하기 위해 그 리오더링된 하나 이상의 벡터들을 리오더링된 하나 이상의 추가적인 벡터들과 재결합할 수도 있다. 이들 및 다른 예들에서, 오디오 디코딩 디바이스 (540B) 는 그 후 음장의 구별되는 성분들을 나타내는 구면 조화 계수들 및 음장의 배경 성분들을 나타내는 구면 조화 계수들에 기초하여 복수의 구면 조화 계수들을 복구할 수도 있다.

[1180] 즉, 본 기술들의 다양한 양태들은, 다음과 같은 조항들에 따라 리오더링된 하나 이상의 벡터들을 디코딩하도록 구성되는 오디오 디코딩 디바이스 (540B) 를 제공할 수도 있다.

[1181] 조항 133146-1F. 음장에서의 성분들에 대응하는 다수의 벡터들을 결정하도록 구성된 하나 이상의 프로세서들을 포함하는, 오디오 인코딩 디바이스 (540B) 와 같은 디바이스.

[1182] 조항 133146-2F. 상기 하나 이상의 프로세서들은, 상기 설명된 경우들의 임의의 조합에 따라 오더 감소를 수행한 후에 상기 다수의 벡터들을 결정하도록 구성되는, 조항 133146-1F 의 디바이스.

[1183] 조항 133146-3F. 상기 하나 이상의 프로세서들은, 상기 설명된 경우들의 임의의 조합에 따라 오더 감소를 수행하도록 더 구성되는, 조항 133146-1F 의 디바이스.

[1184] 조항 133146-4F. 상기 하나 이상의 프로세서들은, 비트스트림에서 특정된 값으로부터 상기 다수의 벡터들을 결정하도록 구성되고, 상기 하나 이상의 프로세서들은, 음장에서의 구별되는 성분들을 나타내는 비트스트림에서의 하나 이상의 벡터들을 식별하기 위해, 결정된 수의 벡터들에 기초하여 비트스트림을 파싱하도록 더 구성되는, 조항 133146-1F 의 디바이스.

[1185] 조항 133146-5F. 상기 하나 이상의 프로세서들은, 비트스트림에서 특정된 값으로부터 상기 다수의 벡터들을 결정하도록 구성되고, 상기 하나 이상의 프로세서들은, 음장에서의 배경 성분들을 나타내는 비트스트림에서의 하나 이상의 벡터들을 식별하기 위해, 결정된 수의 벡터들에 기초하여 비트스트림을 파싱하도록 더 구성되는, 조항 133146-1F 의 디바이스.

[1186] 조항 133143-1C. 음장에서의 구별되는 성분들을 나타내는 리오더링된 하나 이상의 벡터들을 리오더링하도록 구성된 하나 이상의 프로세서들을 포함하는, 오디오 디코딩 디바이스 (540B) 와 같은 디바이스.

[1187] 조항 133143-2C. 상기 하나 이상의 프로세서들은, 리오더링된 하나 이상의 벡터들을 결정하고, 그 리오더링된 하나 이상의 벡터들이 어떻게 리오더링되었는지를 기술하는 리오더 정보를 결정하도록 더 구성되고, 상기 하나 이상의 프로세서들은, 리오더링된 하나 이상의 벡터들을 리오더링할 때, 결정된 리오더 정보에 기초하여, 리오더링된 하나 이상의 벡터들을 리오더링하도록 더 구성되는, 조항 133143-1C 의 디바이스.

[1188] 조항 133143-3C. 리오더링된 하나 이상의 벡터들은 청구항 1A-18A 의 임의의 조합 또는 청구항 1B-19B 의 임의의 조합에 의해 인용된 하나 이상의 리오더링된 제 1 벡터들을 포함하고, 상기 하나 이상의 제 1 벡터들은 청구항 1A-18A 의 임의의 조합 또는 청구항 1B-19B 의 임의의 조합에 의해 인용된 방법에 따라 결정되는, 1C 의

디바이스.

[1189] 조항 133143-4D. 음장에서의 구별되는 성분들을 나타내는 리오더링된 하나 이상의 벡터들을 오디오 디코딩하도록 구성된 하나 이상의 프로세서들을 포함하고, 그 리오더링된 하나 이상의 벡터들은 그 하나 이상의 벡터들을 압축하는 것을 용이하게 하기 위해 리오더링된 것인, 오디오 디코딩 디바이스 (540B) 와 같은 디바이스.

[1190] 조항 133143-5D. 상기 하나 이상의 프로세서들은, 음장에서의 구별되는 성분들을 나타내는 구면 조화 계수들을 복원하기 위해, 리오더링된 하나 이상의 벡터들을 리오더링된 하나 이상의 추가적인 벡터들과 재결합하도록 더 구성되는, 조항 133143-4D 의 디바이스.

[1191] 조항 133143-6D. 상기 하나 이상의 프로세서들은, 음장에서의 구별되는 성분들을 나타내는 구면 조화 계수들 및 음장의 배경 성분들을 나타내는 구면 조화 계수들에 기초하여 복수의 구면 조화 계수들을 복원하도록 더 구성되는, 조항 133143-5D 의 디바이스.

[1192] 조항 133143-1E. 하나 이상의 벡터들을 리오더링하여 리오더링된 하나 이상의 제 1 벡터들을 생성하여 레거시 오디오 인코더에 의한 인코딩을 용이하게 하도록 구성된 하나 이상의 프로세서들을 포함하는 오디오 디코딩 디바이스 (540B) 와 같은 디바이스로서, 상기 하나 이상의 벡터들은 음장의 구별되는 성분들을 기술하고 나타내며, 상기 하나 이상의 프로세서들은, 상기 리오더링된 하나 이상의 벡터들의 인코딩된 버전을 생성하기 위해 레거시 오디오 인코더를 이용하여 상기 리오더링된 하나 이상의 벡터들을 오디오 인코딩하도록 구성되는, 오디오 디코딩 디바이스 (540B) 와 같은 디바이스.

[1193] 조항 133143-2E. 리오더링된 하나 이상의 벡터들은 청구항 1A-18A 의 임의의 조합 또는 청구항 1B-19B 의 임의의 조합에 의해 인용된 하나 이상의 리오더링된 제 1 벡터들을 포함하고, 상기 하나 이상의 제 1 벡터들은 청구항 1A-18A 의 임의의 조합 또는 청구항 1B-19B 의 임의의 조합에 의해 인용된 방법에 따라 결정되는, 1E 의 디바이스.

[1194] 도 41c 는 또 다른 예시적인 오디오 인코딩 디바이스 (540C) 를 나타내는 블록 다이어그램이다. 오디오 디코딩 디바이스 (540C) 는, 데스크톱 컴퓨터, 랩톱 컴퓨터, 워크스테이션, 태블릿 또는 슬레이트 컴퓨터, 전용 오디오 레코딩 디바이스, (소위 "스마트 폰들" 을 포함하는) 셀룰러 폰, 퍼스널 미디어 플레이어 디바이스, 퍼스널 게이밍 디바이스, 또는 오디오 데이터를 디코딩할 수 있는 임의의 다른 유형의 디바이스와 같은, 오디오 데이터를 디코딩할 수 있는 임의의 디바이스를 나타낼 수도 있다.

[1195] 도 41c 의 예에서, 오디오 디코딩 디바이스 (540C) 는, 일부 예들에서, 비본질적인 무관한 데이터의 제거를 용이하게 하기 위해 오디오 인코딩 디바이스 (510B-510J) 에 의해 사용되는, (도 40b 내지 도 40j 의 예들과 관련하여 상기 설명된 바와 같이) 오더 감소를 수행하는 것을 제외하고는 오디오 인코딩 디바이스들 (510B-510E) 중 어느 것에 의해 수행되는 오디오 인코딩 프로세스에 대해 상호적으로 역의 관계를 갖는 오디오 디코딩 프로세스를 수행한다.

[1196] 도 41c 의 예에서는 단일의 디바이스, 즉, 디바이스 (540C) 로서 도시되지만, 디바이스 (540C) 내에 포함되는 것으로서 이하 언급되는 다양한 성분들 또는 유닛들은 디바이스 (540C) 로부터 외부에 있는 분리된 디바이스들을 형성할 수도 있다. 다르게 말하면, 단일의 디바이스, 즉, 도 41c 의 예에서의 디바이스 (540C) 에 의해 수행되는 것으로서 이 개시물에서 설명되지만, 본 기술들은 다수의 디바이스들을 포함하는 시스템에 의해 구현 또는 그 외에 수행될 수도 있으며, 여기서, 이들 디바이스들의 각각은 이하 더 상세하게 설명되는 다양한 성분들 또는 유닛들 중 하나 이상을 각각 포함할 수도 있다. 따라서, 본 기술들은 이러한 점에서 도 41c 의 예에 대한 것으로 제한되어서는 아니된다.

[1197] 또한, 오디오 인코딩 디바이스 (540C) 는 오디오 인코딩 디바이스 (540B) 와 유사할 수도 있다. 하지만, 추출 유닛 (542) 은 (도 40 의 오디오 인코딩 디바이스 (510) 와 관련하여 설명된 경우에서와 같이) 리오더링된 $V^T_{Q_DIST}$ 벡터들 (539) 또는 V^T_{DIST} 벡터들 (525E) 보다는 비트스트림 (517) 으로부터 하나 이상의 V^T_{SMALL} 벡터들 (521) 을 결정할 수도 있다. 그 결과로서, 추출 유닛 (542) 은 V^T_{SMALL} 벡터들 (521) 을 계산 유닛 (546) 에 패스할 수도 있다.

[1198] 또한, 추출 유닛 (542) 은 비트스트림 (517) 으로부터 오디오 인코딩된 변형된 배경 구면 조화 계수들 (515B') 을 결정하고, 이들 계수들 (515B') 을 오디오 디코딩 유닛 (544) 에 패스할 수도 있고, 이 오디오 디코딩 유닛 (544) 은 인코딩된 변형된 배경 구면 조화 계수들 (515B) 을 오디오 디코딩하여 변형된 배경 구면 조화 계수들

(537) 을 복원할 수도 있다. 오디오 디코딩 유닛 (544) 은 이들 변형된 배경 구면 조화 계수들 (537) 을 계산 유닛 (546) 에 패스할 수도 있다.

[1199] 계산 유닛 (546) 은 그 후, 오디오 디코딩된 (그리고 가능하게는 오더링되지 않은) $U_{DIST} * S_{DIST}$ 벡터들 (527') 을 하나 이상의 V^{T}_{SMALL} 벡터들 (521) 에 의해 곱하여 상위 오더의 구별되는 구면 조화 계수들을 복원할 수도 있다. 계산 유닛 (546) 은 그 후, 변형된 배경 구면 조화 계수들 (537) 에 그 상위 오더의 구별되는 구면 조화 계수들을 부가하여 복수의 구면 조화 계수들 (511) 또는 (인코더 유닛 (510E) 에서 수행된 오더 감소로 인해 미분일 수도 있는) 그것의 몇몇 미분을 복원할 수도 있다.

[1200] 이러한 방식으로, 본 기술들은, 오디오 디코딩 디바이스 (540C) 로 하여금, 비트스트림에서 하나 이상의 벡터들에 할당될 것이 요구되는 비트들의 양을 감소시키기 위해 배경 구면 조화 계수들과 재결합된 구면 조화 계수들로부터 분해된 하나 이상의 벡터들 중의 적어도 하나를 비트스트림으로부터 결정하는 것을 가능하게 할 수도 있고, 여기서, 구면 조화 계수들은 음장을 기술하고, 배경 구면 조화 계수들은 동일한 음장의 하나 이상의 배경 성분들을 기술하였다.

[1201] 본 기술들의 다양한 양태들은 이러한 점에서, 오디오 디코딩 디바이스 (540C) 로 하여금, 일부 경우들에서, 배경 구면 조화 계수들과 재결합된 구면 조화 계수들로부터 분해된 하나 이상의 벡터들 중의 적어도 하나를 비트스트림으로부터 결정하도록 구성되는 것을 가능하게 할 수도 있고, 여기서, 구면 조화 계수들은 음장을 기술하고, 배경 구면 조화 계수들은 동일한 음장의 하나 이상의 배경 성분들을 기술하였다.

[1202] 이들 및 다른 경우들에서, 오디오 디코딩 디바이스 (540C) 는, N_{BG} 와 동일한 오더를 갖는 구면 조화 계수들의 제 1 부분을 비트스트림으로부터 획득하도록 구성된다.

[1203] 이들 및 다른 경우들에서, 오디오 디코딩 디바이스 (540C) 는 또한, N_{BG} 와 동일한 오더를 갖는 구면 조화 계수들의 제 1 오디오 인코딩된 부분을 비트스트림으로부터 획득하고, 그 구면 조화 계수들의 오디오 인코딩된 제 1 부분을 오디오 디코딩하여 구면 조화 계수들의 제 1 부분을 생성하도록 구성된다.

[1204] 이들 및 다른 경우들에서, 하나 이상의 벡터들 중의 적어도 하나는, 하나 이상의 V^{T}_{SMALL} 벡터들을 포함하고, 이 하나 이상의 V^{T}_{SMALL} 벡터들은, 복수의 구면 조화 계수들에 대해 특이값 분해를 수행함으로써 생성된 V 행렬의 트랜스포즈로부터 결정된 것이다.

[1205] 이들 및 다른 경우들에서, 하나 이상의 벡터들 중의 적어도 하나는, 하나 이상의 V^{T}_{SMALL} 벡터들을 포함하고, 이 하나 이상의 V^{T}_{SMALL} 벡터들은, 복수의 구면 조화 계수들에 대해 특이값 분해를 수행함으로써 생성된 V 행렬의 트랜스포즈로부터 결정된 것이고, 오디오 디코딩 디바이스 (540C) 는 또한, 복수의 구면 조화 계수들에 대해 특이값 분해를 수행함으로써 생성된 U 행렬 및 S 행렬로부터 도출된 것인 하나 이상의 $U_{DIST} * S_{DIST}$ 벡터들을 비트스트림으로부터 획득하고, 그 $U_{DIST} * S_{DIST}$ 벡터들을 V^{T}_{SMALL} 벡터들에 의해 곱하도록 구성된다.

[1206] 이들 및 다른 경우들에서, 하나 이상의 벡터들 중의 적어도 하나는, 하나 이상의 V^{T}_{SMALL} 벡터들을 포함하고, 이 하나 이상의 V^{T}_{SMALL} 벡터들은, 복수의 구면 조화 계수들에 대해 특이값 분해를 수행함으로써 생성된 V 행렬의 트랜스포즈로부터 결정된 것이고, 오디오 디코딩 디바이스 (540C) 는 또한, 복수의 구면 조화 계수들에 대해 특이값 분해를 수행함으로써 생성된 U 행렬 및 S 행렬로부터 도출된 것인 하나 이상의 $U_{DIST} * S_{DIST}$ 벡터들을 비트스트림으로부터 획득하고, 그 $U_{DIST} * S_{DIST}$ 벡터들을 V^{T}_{SMALL} 벡터들에 의해 곱하여 상위 오더의 구별되는 배경 구면 조화 계수들을 복원하며, 그 상위 오더의 구별되는 배경 구면 조화 계수들에 하위 오더의 구별되는 배경 구면 조화 계수들을 포함하는 배경 구면 조화 계수들을 부가하여 복수의 구면 조화 계수들을 적어도 부분적으로 복원하도록 구성된다.

[1207] 이들 및 다른 경우들에서, 하나 이상의 벡터들 중의 적어도 하나는, 하나 이상의 V^{T}_{SMALL} 벡터들을 포함하고, 이 하나 이상의 V^{T}_{SMALL} 벡터들은, 복수의 구면 조화 계수들에 대해 특이값 분해를 수행함으로써 생성된 V 행렬의 트

랜스포즈로부터 결정된 것이고, 오디오 디코딩 디바이스 (540C) 는 또한, 복수의 구면 조화 계수들에 대해 특이값 분해를 수행함으로써 생성된 U 행렬 및 S 행렬로부터 도출된 것인 하나 이상의 $U_{DIST} * S_{DIST}$ 벡터들을 비트스트림으로부터 획득하고, 그 $U_{DIST} * S_{DIST}$ 벡터들을 V^{T}_{SMALL} 벡터들에 의해 곱하여 상위 오더의 구별되는 배경 구면 조화 계수들을 복원하며, 그 상위 오더의 구별되는 배경 구면 조화 계수들에 하위 오더의 구별되는 배경 구면 조화 계수들을 포함하는 배경 구면 조화 계수들을 부가하여 복수의 구면 조화 계수들을 적어도 부분적으로 복원하며, 그 복원된 복수의 구면 조화 계수들을 렌더링하도록 구성된다.

[1208] 도 41d 는 또 다른 예시적인 오디오 인코딩 디바이스 (540D) 를 나타내는 블록 다이어그램이다. 오디오 디코딩 디바이스 (540D) 는, 데스크톱 컴퓨터, 랩톱 컴퓨터, 워크스테이션, 태블릿 또는 슬레이트 컴퓨터, 전용 오디오 레코딩 디바이스, (소위 "스마트 폰들" 을 포함하는) 셀룰러 폰, 퍼스널 미디어 플레이어 디바이스, 퍼스널 게이밍 디바이스, 또는 오디오 데이터를 디코딩할 수 있는 임의의 다른 유형의 디바이스와 같은, 오디오 데이터를 디코딩할 수 있는 임의의 디바이스를 나타낼 수도 있다.

[1209] 도 41d 의 예에서, 오디오 디코딩 디바이스 (540D) 는, 일부 예들에서, 비본질적인 무관한 데이터의 제거를 용이하게 하기 위해 오디오 인코딩 디바이스들 (510B-510J) 에 의해 사용되는, (도 40b 내지 도 40j 의 예들과 관련하여 상기 설명된 바와 같이) 오더 감소를 수행하는 것을 제외하고는 오디오 인코딩 디바이스들 (510B-510J) 중 어느 것에 의해 수행되는 오디오 인코딩 프로세스에 대해 상호적으로 역의 관계를 갖는 오디오 디코딩 프로세스를 수행한다.

[1210] 도 41d 의 예에서는 단일의 디바이스, 즉, 디바이스 (540D) 로서 도시되지만, 디바이스 (540D) 내에 포함되는 것으로서 이하 언급되는 다양한 성분들 또는 유닛들은 디바이스 (540D) 로부터 외부에 있는 분리된 디바이스들을 형성할 수도 있다. 다르게 말하면, 단일의 디바이스, 즉, 도 41d 의 예에서의 디바이스 (540D) 에 의해 수행되는 것으로서 이 개시물에서 설명되지만, 본 기술들은 다수의 디바이스들을 포함하는 시스템에 의해 구현 또는 그 외에 수행될 수도 있으며, 여기서, 이들 디바이스들의 각각은 이하 더 상세하게 설명되는 다양한 성분들 또는 유닛들 중 하나 이상을 각각 포함할 수도 있다. 따라서, 본 기술들은 이러한 점에서 도 41d 의 예에 대한 것으로 제한되어서는 아니된다.

[1211] 또한, 오디오 디코딩 디바이스 (540D) 는, 오디오 디코딩 디바이스 (540D) 가 도 40i 와 관련하여 상기 설명된 V 압축 유닛 (552) 에 의해 수행되는 압축에 대해 일반적으로 상호적으로 역의 관계에 있는 추가적인 V 압축해제를 수행하는 점을 제외하고는, 오디오 디코딩 디바이스 (540B) 와 유사할 수도 있다. 도 41d 의 예에서, 추출 유닛 (542) 은, 비트스트림 (517) 에 포함된 (그리고 도 10b 및 도 10c 중의 하나에서 나타난 예에 따라 일반적으로 특정된) 압축된 공간적 성분들 (539') 의 이러한 V 압축해제를 수행하는 V 압축해제 유닛 (555) 을 포함한다. V 압축해제 유닛 (555) 은

[1212]
$$\hat{v}_q = \begin{cases} 0, & if\ cid = 0 \\ sgn * (2^{cid-1} + residual), & if\ cid \neq 0 \end{cases}$$

[1213] 와 같은 식에 기초하여 V^{T}_{DIST} 벡터들 (539) 을 압축해제할 수도 있다. 다르게 말하면, V 압축해제 유닛 (555) 은 먼저 비트스트림 (517) 으로부터 *nbits* 값을 파싱하고, *cid* 를 나타내는 허프만 코드를 디코딩할 때 이용하기 위해 5 개의 허프만 코드 테이블들의 적절한 세트를 식별할 수도 있다. 비트스트림 (517) 및 가능하게는 공간 성분의 다른 엘리먼트들에 대한 공간 성분의 엘리먼트의 오더에서 특정된 허프만 코딩 정보 및 예측 모드에 기초하여, V 압축해제 유닛 (555) 은 파싱된 *nbits* 값에 대해 정의된 5 개의 허프만 테이블들 중 정확한 하나를 식별할 수도 있다. 이 허프만 테이블을 이용하여, V 압축해제 유닛 (555) 은 허프만 코드로부터 *cid* 값을 디코딩할 수도 있다. V 압축해제 유닛 (555) 은 그 후, 사인 비트 및 잔차 블록 코드를 파싱하고, 잔차를 식별하기 위해 잔차 블록 코드를 디코딩할 수도 있다. 상기 식에 따라서, V 압축해제 유닛 (555) 은 V^{T}_{DIST} 벡터들 (539) 중 하나를 디코딩할 수도 있다.

[1214] 전술한 바는 다음과 같은 신택스 테이블에서 요약될 수도 있다:

[1215] **테이블 - 디코딩된 벡터들**

신택스	비트들의 수	연상 기호

```
decodeVVec(i)
{
 switch codedVVecLength {
    case 0: //벡터 완성
      VVecLength = NumOfHoaCoeffs;
      for (m=0; m< VVecLength; ++m){ VecCoeff[m] = m+1; }
      break;
    case 1: //하위 오더들은 제거됨
      VVecLength = NumOfHoaCoeffs –
MinNumOfCoeffsForAmbHOA;

       for (m=0; m< VVecLength; ++m){
         VecCoeff[m] = m + MinNumOfCoeffsForAmbHOA +
1;
       }
       break;
    case 2:
      VVecLength = NumOfHoaCoeffs –
MinNumOfCoeffsForAmbHOA - NumOfAddAmbHoaChan;
     n = 0;
     for(m=0;m<NumOfHoaCoeffs–
MinNumOfCoeffsForAmbHOA; ++m){
        c = m + MinNumOfCoeffsForAmbHOA + 1;
        if ( ismember(c, AmbCoeffIdx) == 0){
          VecCoeff[n] = c;
           n++;
        }
     }
      break;
    case 3:
      VVecLength = NumOfHoaCoeffs -
NumOfAddAmbHoaChan;
     n = 0;
     for(m=0; m<NumOfHoaCoeffs; ++m){
        c = m + 1;
```

[1216]

```
        if ( ismember(c, AmbCoeffIdx) == 0){
          VecCoeff[n] = c;
          n++;
        }
      }
  }

  if (NbitsQ[i] == 5) { /* 균일한 양자화기 */
     for (m=0; m< VVecLength; ++m){
        VVec(k)[i][m] = (VecValue  / 128.0) – 1.0;                    8          uimsbf
      }
  }
  else { /* 허프만 디코딩 */
     for (m=0; m< VVecLength; ++m){
       Idx = 5;
       If (CbFlag[i] == 1) {
          idx = (min(3, max(1, ceil(sqrt(VecCoeff[m]) - 1)));
       }
       else if (PFlag[i] == 1) {idx = 4;}
         cid =                                                        동적       huffDe
huffDecode(huffmannTable[NbitsQ].codebook[idx]; huffVal);                        code
        if ( cid > 0 ) {
          aVal = sgn = (sgnVal * 2) - 1;                              1          bslbf
          if (cid > 1) {
            aVal =  sgn * (2.0^(cid -1 ) + intAddVal);                cid-1      uimsbf
          }
        } else {aVal = 0.0;}
     }
  }
}
```

주의:

- 균일한 양자화기를 위한 인코더 함수는 min(255, round((x + 1.0) * 128.0))
- 연상 기호 huffDecode 에 대한 비트들의 수는 동적이다

[1217]

[1218] 전술한 신택스 테이블에서, 4 가지 경우들 (경우 0-3) 을 갖는 제 1 스위치 구문은 계수들의 수의 면에서 V^{T}_{DIST} 벡터 길이를 결정하기 위한 방식을 제공한다. 제 1 경우, 경우 0 은 V^{T}_{DIST} 벡터들에 대한 계수들의 전부가 특정되는 것을 나타낸다. 제 2 경우, 경우 1 은 위에서 (N_{DIST} +1) - (N_{BG} + 1) 로서 지칭되는 것을 표시할 수도 있는, MinNumOfCoeffsForAmbHOA 보다 더 큰 오더에 대응하는 V^{T}_{DIST} 벡터의 그들 계수들만이 특정되는 것을 나타낸다. 제 3 경우, 경우 2 는 제 2 경우와 유사하지만, 오더 N_{BG} 를 초과하는 오더에 대응하는 추가적인 채널들 (여기서, "채널들" 은 소정 오더, 서브-오더 조합에 대응하는 특정 계수를 지칭한다) 을 명시하기 위한 변수를 표시하는, NumOfAddAmbHoaChan 에 의해 식별되는 계수들을 더 뺀다. 제 4 경우, 경우 3 은 NumOfAddAmbHoaChan 에 의해 식별된 계수들을 제거한 후에 남겨진 V^{T}_{DIST} 벡터의 그들 계수들만이 특정되는 것을 나타낸다.

[1219] 이 스위치 구문 후에, 통일된 양자화해제를 수행할지 여부에 관한 결정이 NbitsQ (또는, 위에서 *nbits* 로서 표시됨) 에 의해 제어되고, 이는, 5 와 동일하지 않은 경우에 허프만 디코딩의 적용을 초래한다. 상기 언급된 *cid* 값은 NbitsQ 값의 2 개의 최하위 비트들과 동일하다. 상기 논의된 예측 모드는 상기 신택스 테이블에서 PFlag 로서 표시되는 한편, HT 정보 비트는 상기 신택스 테이블에서 CbFlag 로서 표시된다. 나머지 신택스는 어떻게 디코딩이 상기 설명된 것과 실질적으로 유사한 방식으로 일어나는지를 명시한다.

[1220] 이러한 방식으로, 이 개시의 기술들은 오디오 디코딩 디바이스 (540D) 로 하여금, 복수의 구면 조화 계수들에 대해 벡터 기반 합성을 수행함으로써 생성된, 음장의 공간 성분의 압축된 버전을 포함하는 비트스트림을 획득하고, 그 공간 성분을 획득하기 위해 공간 성분의 압축된 버전을 압축해제하는 것을 가능하게 할 수도 있다.

[1221] 또한, 본 기술들은 오디오 디코딩 디바이스 (540D) 로 하여금, 복수의 구면 조화 계수들에 대해 벡터 기반 합성을 수행함으로써 생성된, 음장의 공간 성분의 압축된 버전을 압축해제하는 것을 가능하게 할 수도 있다.

[1222] 이러한 방식으로, 오디오 인코딩 디바이스 (540D) 는 다음과 같은 조항들에 대해 이하 전개되는 본 기술들의 다양한 양태들을 수행할 수도 있다.

[1223] 조항 141541-1B. 하나 이상의 프로세서들을 포함하는 디바이스로서, 상기 하나 이상의 프로세서들은, 복수의 구면 조화 계수들에 대해 벡터 기반 합성을 수행함으로써 생성된, 음장의 공간 성분의 압축된 버전을 포함하는 비트스트림을 획득하고, 그 공간 성분을 획득하기 위해 공간 성분의 압축된 버전을 압축해제하도록 구성되는, 디바이스.

[1224] 조항 141541-2B. 조항 141541-1B 의 디바이스에 있어서, 공간 성분의 압축된 버전은, 공간 성분을 압축할 때 사용된 예측 모드를 명시하는 필드를 적어도 부분적으로 이용하여 비트스트림에서 표현되고, 상기 하나 이상의 프로세서들은, 공간 성분의 압축된 버전을 압축해제할 때, 공간 성분을 획득하기 위해 예측 모드에 적어도 부분적으로 기초하여 공간 성분의 압축된 버전을 압축해제하도록 더 구성되는, 디바이스.

[1225] 조항 141541-3B. 조항 141541-1B 및 조항 141541-2B 의 임의의 조합의 디바이스에 있어서, 공간 성분의 압축된 버전은, 공간 성분을 압축할 때 사용되는 허프만 테이블을 명시하는 허프만 테이블 정보를 적어도 부분적으로 이용하여 비트스트림에서 표현되고, 상기 하나 이상의 프로세서들은 또한, 공간 성분의 압축된 버전을 압축해제할 때, 허프만 테이블 정보에 적어도 부분적으로 기초하여 공간 성분의 압축된 버전을 압축해제하도록 구성되는, 디바이스.

[1226] 조항 141541-4B. 조항 141541-1B 내지 조항 141541-3B 의 임의의 조합의 디바이스에 있어서, 공간 성분의 압축된 버전은, 공간 성분을 압축할 때 사용되는 양자화 스텝 사이즈 또는 그의 변수를 표현하는 값을 나타내는 필드를 적어도 부분적으로 이용하여 비트스트림에서 표현되고, 상기 하나 이상의 프로세서들은 또한, 공간 성분의 압축된 버전을 압축해제할 때, 상기 값에 적어도 부분적으로 기초하여 공간 성분의 압축된 버전을 압축해제하도록 구성되는, 디바이스.

[1227] 조항 141541-5B. 조항 141541-4B 에 있어서, 상기 값은 *nbits* 값을 포함하는, 디바이스.

[1228] 조항 141541-6B. 조항 141541-4B 및 조항 141541-5B 의 임의의 조합의 디바이스에 있어서, 비트스트림은 공간 성분의 압축된 버전이 포함되는 음장의 복수의 공간 성분들의 압축된 버전을 포함하고, 상기 값은 복수의 공간 성분들을 압축할 때 사용되는 양자화 스텝 사이즈 또는 그의 변수를 표현하고, 상기 하나 이상의 프로세서들은 또한, 공간 성분의 압축된 버전을 압축해제할 때, 상기 값에 적어도 부분적으로 기초하여 복수의 공간 성분의 압축된 버전을 압축해제하도록 더 구성되는, 디바이스.

[1229] 조항 141541-7B. 조항 141541-1B 내지 조항 141541-6B 의 임의의 조합의 디바이스에 있어서, 공간 성분의 압축된 버전은, 공간 성분이 대응하는 압축 카테고리를 식별하는 카테고리 식별자를 표현하기 위한 허프만 코드를 적어도 부분적으로 이용하여 비트스트림에서 표현되고, 상기 하나 이상의 프로세서들은 또한, 공간 성분의 압축된 버전을 압축해제할 때, 허프만 코드에 적어도 부분적으로 기초하여 공간 성분의 압축된 버전을 압축해제하도록 구성되는, 디바이스.

[1230] 조항 141541-8B. 조항 141541-1B 내지 조항 141541-7B 의 임의의 조합의 디바이스에 있어서, 공간 성분의 압축된 버전은, 공간 성분이 포지티브 값인지 또는 네거티브 값인지 여부를 식별하는 사인 비트를 적어도 부분적으로 이용하여 비트스트림에서 표현되고, 상기 하나 이상의 프로세서들은 또한, 공간 성분의 압축된 버전을 압축해제할 때, 그 사인 비트에 적어도 부분적으로 기초하여 공간 성분의 압축된 버전을 압축해제하도록 구성되는, 디바이스.

[1231] 조항 141541-9B. 조항 141541-1B 내지 조항 141541-8B 의 임의의 조합의 디바이스에 있어서, 공간 성분의 압축된 버전은, 공간 성분의 잔차 값을 나타내기 위한 허프만 코드를 적어도 부분적으로 이용하여 비트스트림에서 표현되고, 상기 하나 이상의 프로세서들은 또한, 공간 성분의 압축된 버전을 압축해제할 때, 허프만 코드에 적어도 부분적으로 기초하여 공간 성분의 압축된 버전을 압축해제하도록 구성되는, 디바이스.

[1232] 조항 141541-10B. 조항 141541-1B 내지 조항 141541-10B 의 임의의 조합의 디바이스에 있어서, 벡터 기반

합성은 특이값 분해를 포함하는, 디바이스.

[1233] 또한, 오디오 디코딩 디바이스 (540D) 는 다음과 같은 조항들에 대해 이하 전개되는 기술들의 다양한 양태들을 수행하도록 구성될 수도 있다.

[1234] 조항 141541-1C. 하나 이상의 프로세서들을 포함하는, 오디오 디코딩 디바이스 (540D) 와 같은 디바이스로서, 상기 하나 이상의 프로세서들은, 복수의 구면 조화 계수들에 대해 벡터 기반 합성을 수행함으로써 생성된, 음장의 공간 성분의 압축된 버전을 압축해제하도록 구성되는, 디바이스.

[1235] 조항 141541-2C. 조항 141541-1C 및 조항 141541-2C 의 임의의 조합의 디바이스에 있어서, 상기 하나 이상의 프로세서들은 또한, 공간 성분의 압축된 버전을 압축해제할 때, 공간 성분이 압축될 때 카테고리화되었던 카테고리를 식별하는 카테고리 식별자를 획득하고, 공간 성분이 포지티브 값인지 또는 네거티브 값인지 여부를 식별하는 사인을 획득하며, 공간 성분의 압축된 버전과 연관된 잔차 값을 획득하고, 그 카테고리 식별자, 사인 및 잔차 값에 기초하여 공간 성분의 압축된 버전을 압축해제하도록 구성되는, 디바이스.

[1236] 조항 141541-3C. 조항 141541-2C 의 디바이스에 있어서, 상기 하나 이상의 프로세서들은 또한, 카테고리 식별자를 획득할 때, 그 카테고리 식별자를 나타내는 허프만 코드를 획득하고, 그 카테고리 식별자를 획득하기 위해 허프만 코드를 디코딩하도록 구성되는, 디바이스.

[1237] 조항 141541-4C. 조항 141541-3C 의 디바이스에 있어서, 상기 하나 이상의 프로세서들은 또한, 상기 허프만 코드를 디코딩할 때, 복수의 공간 성분들을 명시하는 벡터에서의 공간 성분의 상대적인 포지션에 적어도 부분적으로 기초하여 허프만 코드를 디코딩하기 위해 사용된 허프만 테이블을 식별하도록 구성되는, 디바이스.

[1238] 조항 141541-5C. 조항 141541-3C 및 조항 141541-4C 의 임의의 조합의 디바이스에 있어서, 상기 하나 이상의 프로세서들은 또한, 상기 허프만 코드를 디코딩할 때, 공간 성분을 압축할 때 사용된 예측 모드에 적어도 부분적으로 기초하여 허프만 코드를 디코딩하기 위해 사용된 허프만 테이블을 식별하도록 구성되는, 디바이스.

[1239] 조항 141541-6C. 조항 141541-3C 내지 조항 141541-5C 의 임의의 조합의 디바이스에 있어서, 상기 하나 이상의 프로세서들은 또한, 상기 허프만 코드를 디코딩할 때, 공간 성분의 압축된 버전과 연관된 허프만 테이블 정보에 적어도 부분적으로 기초하여 허프만 코드를 디코딩하기 위해 사용된 허프만 테이블을 식별하도록 구성되는, 디바이스.

[1240] 조항 141541-7C. 조항 141541-3C 의 디바이스에 있어서, 상기 하나 이상의 프로세서들은 또한, 상기 허프만 코드를 디코딩할 때, 복수의 공간 성분들을 명시하는 벡터에서의 공간 성분의 상대적인 포지션에 적어도 부분적으로 기초하여 허프만 코드를 디코딩하기 위해 사용된 허프만 테이블, 공간 성분을 압축할 때 사용된 예측 모드, 및 공간 성분의 압축된 버전과 연관된 허프만 테이블 정보를 식별하도록 구성되는, 디바이스.

[1241] 조항 141541-8C. 조항 141541-2C 의 디바이스에 있어서, 상기 하나 이상의 프로세서들은 또한, 잔차 값을 획득할 때, 그 잔차 값을 획득하기 위해 잔차 값을 나타내는 블록 코드를 디코딩하도록 구성되는, 디바이스.

[1242] 조항 141541-9C. 조항 141541-1C 내지 조항 141541-8C 의 임의의 조합의 디바이스에 있어서, 벡터 기반 합성은 특이값 분해를 포함하는, 디바이스.

[1243] 또한, 오디오 디코딩 디바이스 (540D) 는 다음과 같은 조항들에 대해 이하 전개되는 기술들의 다양한 양태들을 수행하도록 구성될 수도 있다.

[1244] 조항 141541-1G. 하나 이상의 프로세서들을 포함하는, 오디오 디코딩 디바이스 (540D) 와 같은 디바이스로서, 상기 하나 이상의 프로세서들은, 복수의 압축된 공간 성분들의 나머지 것들에 대한 현재의 공간 성분의 압축된 버전의 오더에 기초하여 복수의 압축된 공간 성분들의 현재의 공간 성분의 압축된 버전을 압축해제할 때 사용하기 위해 허프만 코드북을 식별하도록 구성되고, 공간 성분은 복수의 구면 조화 계수들에 대해 벡터 기반 합성을 수행함으로써 생성되는, 디바이스.

[1245] 조항 141541-2G. 조항 141541-1G 의 디바이스에 있어서, 상기 하나 이상의 프로세서들은, 조항 141541-1D 내지 조항 141541-10D, 및 조항 141541-1E 내지 조항 141541-9E 에 인용된 단계들의 임의의 조합을 수행하도록 더 구성되는, 디바이스.

[1246] 도 42 내지 도 42c 는 각각, 도 40b 내지 도 40j 의 예들에서 나타난 오더 감소 유닛 (528A) 을 나타내는 블록 다이어그램이다. 도 42 는 도 40b 내지 도 40j 의 오더 감소 유닛 (528A) 의 하나의 예를 나타낼 수도 있는, 오더 감소 유닛 (528) 을 나타내는 블록 다이어그램이다. 오더 감소 유닛 (528A) 은 타겟 비트레이트

(535) 를 수신 또는 그 외에 결정하고, 이 타겟 비트레이트 (535) 에만 기초하여 배경 구면 조화 계수들 (531) 에 대해 오더 감소를 수행할 수도 있다. 일부 예들에서, 오더 감소 유닛 (528A) 은, 감소된 배경 구면 조화 계수들 (529) 을 생성하기 위해 배경 구면 조화 계수들 (531) 로부터 제거되어야 하는 그들 오더들 및/또는 서브오더들을 식별하기 위해 타겟 비트레이트 (535) 를 이용하여 테이블 또는 다른 데이터 구조를 액세스할 수도 있다.

[1247] 이러한 방식으로, 본 기술들은 오디오 인코딩 디바이스 (510B-410J) 와 같은 오디오 인코딩 디바이스로 하여금, 감소된 구면 조화 계수들 (529) 또는 그것의 감소된 분해들을 생성하기 위해, 배경 구면 조화 계수들 (531) 과 같은, 복수의 구면 조화 계수들 또는 그것의 분해들에 대해 오더 감소를 수행하는 것을 가능하게 할 수도 있고, 여기서, 복수의 구면 조화 계수들은 음장을 나타낸다.

[1248] 상기 설명된 다양한 경우들의 각각에서, 오디오 디코딩 디바이스 (540) 는, 오디오 디코딩 디바이스 (540) 가 수행하도록 구성되는 방법을 수행하거나 그 외에 그 방법의 각 단계를 수행하기 위한 수단을 포함할 수도 있음을 이해하여야 한다. 일부 경우들에서, 이들 수단은 하나 이상의 프로세서들을 포함할 수도 있다. 일부 경우들에서, 하나 이상의 프로세서들은 비-일시적 컴퓨터 판독가능 저장 매체에 저장된 명령들에 의해 구성되는 특수 목적 프로세서를 나타낼 수도 있다. 다르게 말하면, 인코딩 예들의 세트들의 각각에서의 기술들의 다양한 양태들은, 실행될 때 하나 이상의 프로세서들로 하여금 오디오 디코딩 디바이스 (540) 가 수행하도록 구성된 방법을 수행하게 하는 명령들을 저장한 비-일시적 컴퓨터 판독가능 저장 매체를 제공할 수도 있다.

[1249] 도 42b 는, 도 40b 내지 도 40j 의 오더 감소 유닛 (528A) 의 하나의 예를 나타낼 수도 있는 오더 감소 유닛 (528B) 을 나타내는 블록 다이어그램이다. 도 42b 의 예에서, 타겟 비트레이트 (535) 에만 기초하여 오더 감소를 수행하기 보다는, 오더 감소 유닛 (528B) 은 배경 구면 조화 계수들 (531) 의 컨텐츠 분석에 기초하여 오더 감소를 수행할 수도 있다. 오더 감소 유닛 (528B) 은 이 컨텐츠 분석을 수행하는 컨텐츠 분석 유닛 (536A) 을 포함할 수도 있다.

[1250] 일부 예들에서, 컨텐츠 분석 유닛 (536A) 은 공간 분석으로 지칭되는 컨텐츠 분석의 형태를 수행하는 공간 분석 유닛 (536A) 을 포함할 수도 있다. 공간 분석은 음장의 배경 성분들의 형상 또는 다른 공간적 특성들을 기술하는 공간 정보를 식별하기 위해 배경 구면 조화 계수들 (531) 을 분석하는 것을 수반할 수도 있다. 이 공간 정보에 기초하여, 오더 감소 유닛 (528B) 은 배경 구면 조화 계수들 (529) 을 생성하기 위해 배경 구면 조화 계수들 (531) 로부터 제거되어야 하는 그들 오더들 및/또는 서브오더들을 식별할 수도 있다.

[1251] 일부 예들에서, 컨텐츠 분석 유닛 (536A) 은 확산 분석으로 지칭되는 컨텐츠 분석의 형태를 수행하는 확산 분석 유닛 (536B) 을 포함할 수도 있다. 확산 분석은, 음장의 배경 성분들의 확산성을 기술하는 확산 정보를 식별하기 위해 배경 구면 조화 계수들 (531) 을 분석하는 것을 수반할 수도 있다. 이 확산 정보에 기초하여, 오더 감소 유닛 (528B) 은 배경 구면 조화 계수들 (529) 을 생성하기 위해 배경 구면 조화 계수들 (531) 로부터 제거되어야 하는 그들 오더들 및/또는 서브오더들을 식별할 수도 있다.

[1252] 공간 분석 유닛 (536A) 및 확산 분석 유닛 (36B) 양자 모두를 포함하는 것으로 도시되었지만, 컨텐츠 분석 유닛 (536A) 은 오직 공간 분석 유닛 (536) 만을, 오직 확산 분석 유닛 (536B) 만을, 또는 공간 분석 유닛 (536A) 및 확산 분석 유닛 (536B) 양자 모두를 포함할 수도 있다. 일부 예들에서, 컨텐츠 분석 유닛 (536A) 은 공간 분석 및 확산 분석 중의 일방 또는 양방에 추가하여 또는 그에 대한 대안으로서 컨텐츠 분석의 다른 형태들을 수행할 수도 있다. 따라서, 이 개시물에 설명된 기술들은 이 점에 한정되어서는 안된다.

[1253] 이러한 식으로, 본 기술들은, 오디오 인코딩 디바이스들 (510B-510J) 과 같은 오디오 인코딩 디바이스로 하여금, 음장을 기술하는 복수의 구면 조화 계수들 또는 그것의 분해들의 컨텐츠 분석에 기초하여, 감소된 구면 조화 계수들 또는 그것의 감소된 분해들을 생성하기 위해 복수의 구면 조화 계수들 또는 그것의 분해들에 대해 오더 감소를 수행하는 것을 가능하게 할 수도 있다.

[1254] 다르게 말하면, 본 기술들은 오디오 인코딩 디바이스들 (510B-510J) 과 같은 디바이스로 하여금, 다음과 같은 조항들에 따라 구성되도록 하는 것을 가능하게 할 수도 있다.

[1255] 조항 133146-1E. 하나 이상의 프로세서들을 포함하는, 오디오 인코딩 디바이스들 (510B-510J) 과 같은 디바이스에 있어서, 상기 하나 이상의 프로세서들은, 음장을 기술하는 복수의 구면 조화 계수들 또는 그것의 분해들의 컨텐츠 분석에 기초하여, 감소된 구면 조화 계수들 또는 그것의 감소된 분해들을 생성하기 위해 복수의 구면 조화 계수들 또는 그것의 분해들에 대해 오더 감소를 수행하도록 구성되는, 디바이스.

[1256] 조항 133146-2E. 조항 133146-1E 의 디바이스에 있어서, 상기 하나 이상의 프로세서들은 또한, 오더 감소를

수행하기 이전에, 음장의 구별되는 성분들을 기술하는 하나 이상의 제 1 벡터들 및 음장의 배경 성분들을 식별하는 하나 이상의 제 2 벡터들을 식별하기 위해 복수의 구면 조화 계수들에 대해 특이값 분해를 수행하도록 구성되고, 상기 하나 이상의 프로세서들은, 상기 하나 이상의 제 1 벡터들, 상기 하나 이상의 제 2 벡터들, 또는 상기 하나 이상의 제 1 벡터들 및 상기 하나 이상의 제 2 벡터들 양자 모두에 대해 오더 감소를 수행하도록 구성되는, 디바이스.

[1257] 조항 133146-3E. 조항 133146-1E 의 디바이스에 있어서, 상기 하나 이상의 프로세서들은 또한, 복수의 구면 조화 계수들 또는 그것의 분해들에 대해 컨텐츠 분석을 수행하도록 구성되는, 디바이스.

[1258] 조항 133146-4E. 조항 133146-3E 의 디바이스에 있어서, 상기 하나 이상의 프로세서들은, 복수의 구면 조화 계수들 또는 그것의 분해들에 대해 공간 분석을 수행하도록 구성되는, 디바이스.

[1259] 조항 133146-5E. 조항 133146-3E 의 디바이스에 있어서, 컨텐츠 분석을 수행하는 것은, 복수의 구면 조화 계수들 또는 그것의 분해들에 대해 확산 분석을 수행하는 것을 포함하는, 디바이스.

[1260] 조항 133146-6E. 조항 133146-3E 의 디바이스에 있어서, 상기 하나 이상의 프로세서들은, 복수의 구면 조화 계수들 또는 그것의 분해들에 대해 공간 분석 및 확산 분석을 수행하도록 구성되는, 디바이스.

[1261] 조항 133146-7E. 청구항 1 의 디바이스에 있어서, 하나 이상의 프로세서들은, 복수의 구면 조화 계수들 또는 그것의 분해들의 컨텐츠 분석 및 타겟 비트레이트에 기초하여, 감소된 구면 조화 계수들 또는 그것의 감소된 분해들을 생성하기 위해 복수의 구면 조화 계수들 또는 그것의 분해들에 대해 오더 감소를 수행하도록 구성되는, 디바이스.

[1262] 조항 133146-8E. 조항 133146-1E 의 디바이스에 있어서, 상기 하나 이상의 프로세서들은 또한, 감소된 구면 조화 계수들 또는 그것의 분해들을 오디오 인코딩하도록 구성되는, 디바이스.

[1263] 조항 133146-9E. 조항 133146-1E 의 디바이스에 있어서, 상기 하나 이상의 프로세서들은 또한, 감소된 구면 조화 계수들 또는 그것의 감소된 분해들을 오디오 인코딩하고, 감소된 구면 조화 계수들 또는 그것의 감소된 분해들을 포함하도록 비트스트림을 생성하도록 구성되는, 디바이스.

[1264] 조항 133146-10E. 조항 133146-1E 의 디바이스에 있어서, 상기 하나 이상의 프로세서들은 또한, 감소된 구면 조화 계수들 또는 그것의 감소된 분해들이 그 감소된 구면 조화 계수들 또는 그것의 감소된 분해들을 포함하는 비트스트림에서 대응하는, 구면 기초 함수들의 하나 이상의 오더들 및/또는 하나 이상의 서브-오더들을 명시하도록 구성되는, 디바이스.

[1265] 조항 133146-11E. 조항 133146-1E 의 디바이스에 있어서, 감소된 구면 조화 계수들 또는 그것의 감소된 분해들은 복수의 구면 조화 계수들 또는 그것의 분해들보다 더 적은 값들을 갖는, 디바이스.

[1266] 조항 133146-12E. 조항 133146-1E 의 디바이스에 있어서, 상기 하나 이상의 프로세서들은 또한, 감소된 구면 조화 계수들 또는 그것의 감소된 분해들을 생성하기 위해 복수의 구면 조화 계수들 또는 그것의 분해들의 벡터들 중 명시된 오더 및/또는 서브-오더를 갖는 것들을 제거하도록 구성되는, 디바이스.

[1267] 조항 133146-13E. 조항 133146-1E 의 디바이스에 있어서, 상기 하나 이상의 프로세서들은, 감소된 구면 조화 계수들 또는 그것의 감소된 분해들을 생성하기 위해 복수의 구면 조화 계수들 중 또는 그것의 분해들의 벡터들 중 명시된 오더 및/또는 서브-오더를 갖는 것들을 제로 아웃하도록 구성되는, 디바이스.

[1268] 도 42c 는, 도 40b 내지 도 40j 의 오더 감소 유닛 (528A) 의 하나의 예를 나타낼 수도 있는 오더 감소 유닛 (528C) 을 나타내는 블록 다이어그램이다. 도 42b 의 오더 감소 유닛 (528C) 은 오더 감소 유닛 (528B) 과 실질적으로 동일하지만, 도 42b 의 오더 감소 유닛 (528B) 에 대해 상기 설명된 방식으로 컨텐츠 분석을 또한 수행하는 한편, 도 42 의 오더 감소 유닛 (528A) 에 대해 상기 설명된 방식으로 타겟 비트레이트 (535) 를 수신 또는 그 외에 결정할 수도 있다. 오더 감소 유닛 (528C) 은 그 다음, 이 타겟 비트레이트 (535) 및 컨텐츠 분석에 기초하여 배경 구면 조화 계수들 (531) 에 대해 오더 감소를 수행할 수도 있다.

[1269] 이러한 방식으로, 본 기술들은 오디오 인코딩 디바이스들 (510B-510J) 과 같은 오디오 인코딩 디바이스로 하여금, 복수의 구면 조화 계수들 또는 그것의 분해들에 대해 컨텐츠 분석을 수행하는 것을 가능하게 할 수도 있다.

오더 감소를 수행할 때, 오디오 인코딩 디바이스 (510B-510J) 는, 타겟 비트레이트 (535) 및 컨텐츠 분석에 기초하여, 감소된 구면 조화 계수들 또는 그것의 감소된 분해들을 생성하기 위해 복수의 구면 조화 계수들 또는 그것의 분해들에 대해 오더 감소를 수행할 수도 있다.

[1270] 하나 이상의 벡터들이 제거되면, 오디오 인코딩 디바이스들 (510B-510J) 은 제어 데이터로서 비트스트림에서의 벡터들의 수를 명시할 수도 있다. 오디오 인코딩 디바이스들 (510B-510J) 은 오디오 디코딩 디바이스에 의한 비트스트림으로부터의 벡터들의 추출을 용이하게 하기 위해 비트스트림에서의 이 벡터들의 수를 명시할 수도 있다.

[1271] 도 44 는 이 개시물에서 설명되는 기술들의 다양한 양태들에 따라 양자화 에러를 보상하기 위해 오디오 인코딩 디바이스 (410D) 에 의해 수행되는 예시적인 동작들을 나타내는 다이어그램이다. 도 44 의 예에서, 오디오 인코딩 디바이스 (510D) 의 계산 유닛 (526) 은 수학적 연산들이 오디오 디코딩 디바이스 (510D) 의 계산 유닛 (526) 에 의해 수행될 수도 있음을 나타내기 위해 점선의 블록으로서 도시된다.

[1272] 도 44 의 예에서 도시된 바와 같이, 계산 유닛 (526) 은 먼저, ("H_{DIST} 벡터들 (630)" 로서 표시된) 구별되는 구면 조화 계수들을 생성하기 위해 U_{DIST} * S_{DIST} 벡터들 (527) 을 V^{T}_{DIST} 벡터들 (525E) 에 의해 곱할 수도 있다. 계산 유닛 (526) 은 그 다음, ("$V^{T}_{Q_DIST}$ 벡터들 (525G)" 로서 다시 표시되는) V^{T}_{DIST} 벡터들 (525E) 의 양자된 버전에 의해 H_{DIST} 벡터들 (630) 을 나눌 수도 있다. 계산 유닛 (526) 은, $V^{T}_{Q_DIST}$ 벡터들 (525G) 의 의사 역을 결정한 다음, H_{DIST} 벡터들을 $V^{T}_{Q_DIST}$ 벡터들 (525G) 의 의사 역에 의해 곱함으로써 이 나누기를 수행하고, ("US_{DIST}" 또는 "US_{DIST} 벡터들"로서 약칭될 수도 있는) U_{DIST} * S_{DIST} 의 에러 보상된 버전을 출력할 수도 있다. US_{DIST} 의 에러 보상된 버전은 도 44 의 예에서 US^{*}_{DIST} 벡터들 (527') 로서 표시될 수도 있다. 이러한 방식으로, 본 기술들은 US_{DIST} 벡터들 (527) 에 적어도 부분적으로 양자화 에러를 유효하게 투영하여 US^{*}_{DIST} 벡터들 (527') 을 생성할 수도 있다.

[1273] 계산 유닛 (526) 은 그 다음, (U_{DIST} * S_{DIST} 벡터들 (527) 내로 투영된 양자화로 인한 에러의 적어도 부분을 나타낼 수도 있는) US_{ERR} 벡터들 (634) 을 결정하기 위해 U_{DIST} * S_{DIST} 벡터들 (527) 로부터 US^{*}_{DIST} 벡터들 (527') 을 뺄 수도 있다. 계산 유닛 (526) 은 그 다음, H_{ERR} 벡터들 (636) 을 결정하기 위해 US_{ERR} 벡터들 (634) 을 $V^{T}_{Q_DIST}$ 벡터들 (525G) 에 의해 곱할 수도 있다. 수학적으로, H_{ERR} 벡터들 (636) 은 US_{DIST} 벡터들 (527) - US^{*}_{DIST} 벡터들 (527') 의 결과에, 그 다음, V^{T}_{DIST} 벡터들 (525E) 에 의해 곱해진 것과 동일할 수도 있다. 계산 유닛 (526) 은 그 다음, U_{BG} 벡터들 (525D) 을 S_{BG} 벡터들 (525B) 에 의해 곱한 후 V^{T}_{BG} 벡터들 (525F) 에 의해 곱함으로써 계산된 (도 44 의 예에서 H_{BG} 벡터들 (531) 로서 표시된) 배경 구면 조화 계수들 (531) 에 H_{ERR} 벡터들 (636) 을 더할 수도 있다. 계산 유닛 (526) 은, H_{BG} 벡터들 (531) 에 H_{ERR} 벡터들 (636) 을 부가하여, H_{BG} 벡터들 (531) 에 양자화 에러의 적어도 부분을 유효하게 반영하여, 보상된 H_{BG} 벡터들 (531') 을 생성할 수도 있다. 이러한 방식으로, 본 기술들은 H_{BG} 벡터들 (531) 내로 양자화 에러의 적어도 부분을 투영할 수도 있다.

[1274] 도 45 및 도 45b 는 이 개시물에서 설명된 기술들의 다양한 양태들에 따른, 2 개의 프레임들의 부분들로부터 서브-프레임들의 보간을 나타내는 다이어그램이다. 도 45 의 예에서, 제 1 프레임 (650) 및 제 2 프레임 (652) 이 도시된다. 제 1 프레임 (650) 은 U[1], S[1] 및 V'[1] 행렬들로 분해될 수도 있는 구면 조화 계수들 ("SH[1]") 을 포함할 수도 있다. 제 2 프레임 (652) 은 구면 조화 계수들 ("SH[2]") 을 포함할 수도 있다. SH[1] 및 SH[2] 는 상기 설명된 SHC (511) 의 상이한 프레임들을 식별할 수도 있다.

[1275] 도 45b 의 예에서, 도 40h 의 예에서 도시된 오디오 인코딩 디바이스 (510H) 의 분해 유닛 (518) 은, 프레임들 (650 및 652) 의 각각을 4 개의 각각의 서브-프레임들 (651A-651D 및 653A-653D) 로 분리할 수도 있다. 분해 유닛 (518) 은 그 다음, 프레임 (650) 의 ("SH[1,1]" 로서 표시된) 제 1 서브-프레임 (651A) 을 U[1, 1], S[1, 1] 및 V[1, 1] 행렬들로 분해하고, V[1, 1] 행렬 (519') 을 보간 유닛 (550) 으로 출력할 수도 있다. 분해 유닛 (518) 은 그 다음, 프레임 (652) 의 ("SH[2,1]" 로서 표시된) 제 2 서브-프레임 (653A) 을 U[1, 1], S[1, 1] 및 V[1, 1] 행렬들로 분해하고, V[2, 1] 행렬 (519') 을 보간 유닛 (550) 으로 출력할 수도 있다. 분해 유닛 (518) 은 또한, SHC 11 의 SH[1, 1], SH[1, 2], SH[1, 3] 및 SH[1, 4] 및 SHC (511) 의

SH[2, 1], SH[2, 2], SH[2, 3] 및 SH[2, 4] 를 보간 유닛 (550) 으로 출력할 수도 있다.

[1276] 보간 유닛 (550) 은 그 다음, 도 45b 에 나타난 도시의 바닥에서 식별된 보간들을 수행할 수도 있다. 즉, 보간 유닛 (550) 은 V'[1, 1] 및 V'[2, 1] 에 기초하여 V'[1, 2] 를 보간할 수도 있다. 보간 유닛 (550) 은 또한, V'[1, 1] 및 V'[2, 1] 에 기초하여 V'[1, 3] 를 보간할 수도 있다. 추가적으로, 보간 유닛 (550) 은 또한, V'[1, 1] 및 V'[2, 1] 에 기초하여 V'[1, 4] 를 보간할 수도 있다. 이들 보간들은, 도 46 내지 도 46e 의 예들에서 나타난 바와 같이 V'[1, 1] 및 V'[2, 1] 의 공간 도메인 내로의 투영에 이어, 시간적 보간, 그리고 그 후, 구면 조화 도메인 내로의 다시 역으로의 투영을 수반할 수도 있다.

[1277] 보간 유닛 (550) 은 다음으로, SH[1, 2] 를 $(V'[1, 2])^{-1}$ 에 의해 곱함으로써 U[1, 2]S[1, 2] 를 도출하고, SH[1, 3] 를 $(V'[1, 3])^{-1}$ 에 의해 곱함으로써 U[1, 3]S[1, 3] 를 도출하며, SH[1, 4] 를 $(V'[1, 4])^{-1}$ 에 의해 곱함으로써 U[1, 4]S[1, 4] 를 도출할 수도 있다. 보간 유닛 (550) 은 그 다음, V 행렬 (519), S 행렬 (519B) 및 U 행렬 (519C) 을 출력하는 분해된 형태로 프레임을 바꿀 수도 있다.

[1278] 도 46a 내지 도 46e 는 이 개시물에서 설명된 기술들에 따라, 보간된 복수의 구면 조화 계수들의 분해된 버전의 하나 이상의 벡터들의 투영의 단면을 나타내는 다이어그램이다. 도 46a 는 SVD 프로세스를 통해 제 1 프레임으로부터의 제 1 서브-프레임의 SHC (511) 로부터 분해된 제 1 V 행렬 (19') 의 하나 이상의 제 1 벡터들의 투영의 단면을 나타낸다. 도 46b 는 SVD 프로세스를 통해 제 2 프레임으로부터의 제 1 서브-프레임의 SHC (511) 로부터 분해된 제 2 V 행렬 (519') 의 하나 이상의 제 2 벡터들의 투영의 단면을 나타낸다.

[1279] 도 46c 는 제 1 프레임으로부터의 제 2 서브-프레임을 나타내는 V 행렬 (519A) 에 대한 하나 이상의 보간된 벡터들의 투영의 단면을 나타내고, 이들 벡터들은, SHC (511) 의 제 1 프레임의 제 1 서브-프레임으로부터 분해된 V 행렬 (519') 로부터 이 개시물에서 기술된 기술들에 따라 보간된 것 (즉, 이 예에서 도 46 의 예에서 나타난 V 행렬 (519') 의 하나 이상의 벡터들) 및 SHC (511) 의 제 2 프레임의 제 1 서브-프레임으로부터 분해된 V 행렬 (519') 로부터 이 개시물에서 기술된 기술들에 따라 보간된 것 (즉, 이 예에서 도 46b 의 예에서 나타난 V 행렬 (519') 의 하나 이상의 벡터들) 이다.

[1280] 도 46d 는 제 1 프레임으로부터의 제 3 서브-프레임을 나타내는 V 행렬 (519A) 에 대한 하나 이상의 보간된 벡터들의 투영의 단면을 나타내고, 이들 벡터들은, SHC (511) 의 제 1 프레임의 제 1 서브-프레임으로부터 분해된 V 행렬 (519') 로부터 이 개시물에서 기술된 기술들에 따라 보간된 것 (즉, 이 예에서 도 46 의 예에서 나타난 V 행렬 (519') 의 하나 이상의 벡터들) 및 SHC (511) 의 제 2 프레임의 제 1 서브-프레임으로부터 분해된 V 행렬 (519') 로부터 이 개시물에서 기술된 기술들에 따라 보간된 것 (즉, 이 예에서 도 46b 의 예에서 나타난 V 행렬 (519') 의 하나 이상의 벡터들) 이다.

[1281] 도 46e 는 제 1 프레임으로부터의 제 4 서브-프레임을 나타내는 V 행렬 (519A) 에 대한 하나 이상의 보간된 벡터들의 투영의 단면을 나타내고, 이들 벡터들은, SHC (511) 의 제 1 프레임의 제 1 서브-프레임으로부터 분해된 V 행렬 (519') 로부터 이 개시물에서 기술된 기술들에 따라 보간된 것 (즉, 이 예에서 도 46 의 예에서 나타난 V 행렬 (519') 의 하나 이상의 벡터들) 및 SHC (511) 의 제 2 프레임의 제 1 서브-프레임으로부터 분해된 V 행렬 (519') 로부터 이 개시물에서 기술된 기술들에 따라 보간된 것 (즉, 이 예에서 도 46b 의 예에서 나타난 V 행렬 (519') 의 하나 이상의 벡터들) 이다.

[1282] 도 47 은 예들 도 41 내지 도 41d 에서 도시된 오디오 디코딩 디바이스들 (540A-540D) 의 추출 유닛 (542) 을 더 상세하게 나타내는 블록 다이어그램이다. 일부 예들에서, 추출 유닛 (542) 은 2 이상의 디코딩 방식들을 수행할 수도 있는 "통합된 디코더" (여기서, 이들 2 이상의 방식들을 수행함으로써, 디코더는 2 이상의 방식들을 "통합" 하는 것으로 고려될 수도 있다) 로서 지칭될 수도 있는 것에 대한 프론트 엔드를 나타낼 수도 있다.

도 44 에서 도시된 바와 같이, 추출 유닛 (542) 은 멀티플렉서 (620) 및 추출 서브-유닛들 (622A 및 622B) ("추출 서브-유닛들 (622)") 을 포함한다. 멀티플렉서 (620) 는, 연관된 인코딩된 프레임화된 SHC 행렬들 (547-547N) 이 합성 오디오 오브젝트 또는 레코딩으로부터 생성되는지 여부에 관한 대응하는 표시에 기초하여, 인코딩된 프레임화된 SHC 행렬들 (547-547N) 중 추출 서브-유닛 (622A) 및 추출 서브-유닛 (622B) 에 전송될 것들을 식별한다. 추출 서브-유닛들 (622A) 의 각각은, 일부 예들에서 합성 오디오 오브젝트로부터 생성된 SHC 또는 레코딩으로부터 생성된 SHC 중 어느 일방에 테일러링되는 상이한 디코딩 (이는 "압축해제" 로서 지칭될 수도 있다) 방식을 수행할 수도 있다. 추출 서브-유닛들 (622A) 의 각각은, SHC (547) 로 출력되는 SHC (547) 의 프레임들을 생성하기 위해 이들 압축해제 방식들 중 각각의 하나를 수행할 수도 있다.

[1283] 예를 들어, 추출 유닛 (622A) 은 다음과 같은 식:

[1284] HOA = DirV × PS

[1285] 을 이용하여 우월한 신호 (predominant signal; PS) 로부터 SA 를 재구성하기 위해 압축해제 방식을 수행할 수도 있고, 여기서, DirV 는 (다양한 방향들 및 폭들을 나타내는) 방향-벡터이고, 이는 사이드 채널을 통해 송신될 수도 있다. 추출 유닛 (622B) 은, 이 예에서, 다음과 같은 식:

[1286] HOA = sqrt(4π) * Ynm (theta, phi) * PS

[1287] 을 이용하여 PS 로부터 HOA 행렬을 재구성하는 압축해제 방식을 수행할 수도 있고, 여기서, Ynm 은 구면 조화 함수이고, 세타 (theta) 및 파이 (phi) 정보는 사이드 채널을 통해 전송될 수도 있다.

[1288] 이러한 점에서, 본 기술들은 추출 유닛 (538) 으로 하여금, 음장을 나타내는 구면 조화 계수들의 압축된 버전이 합성 오디오 오브젝트로부터 생성되는지 여부에 관한 표시에 기초하여 복수의 압축해제 방식들 중 하나를 선택하고, 복수의 압축해제 방식들 중의 선택된 하나를 이용하여 구면 조화 계수들의 압축된 버전을 압축해제하는 것을 가능하게 한다. 일부 예들에서, 디바이스는 통합된 디코더를 포함한다.

[1289] 도 48 은 도 41a 내지 도 41d 의 예들에서 도시된 오디오 디코딩 디바이스 (540A-540D) 의 오디오 렌더링 유닛 (48) 을 더 상세하게 나타내는 블록 다이어그램이다. 도 48 은, 복원된 구면 조화 계수들 (547) 로부터 다중 채널 오디오 데이터 (549A) 로의 변환을 나타내고, 이는 디코더-로컬 스피커 지오메트리와 양립될 수 있는 것이다. 몇몇 로컬 스피커 지오메트리들 (이는 다시 디코더에서의 스피커 지오메트리로 지칭될 수도 있다) 에 대해, 전환가능성을 보장하는 몇몇 변환들은 바람직한 오디오-이미지 품질보다 더 적은 품질을 초래할 수도 있다. 즉, 사운드 재생은 캡처되는 오디오와 비교될 때 사운드들의 정확한 로컬화를 항상 초래하지 못할 수도 있다. 이 바람직한 품질보다 적은 품질을 정정하기 위해, 본 기술들은, "가상 스피커들" 로서 지칭될 수도 있는 개념을 도입하도록 더 증강될 수도 있다.

[1290] 하나 이상의 라우드스피커들이 상기 언급된 ITU-R BS.775-1 과 같은 표준에 의해 명시되는 소정의 각도 공차들을 갖는 공간의 특정 또는 정의된 구역들에서 재포지셔닝 또는 포지셔닝되는 것을 요구하기보다는, 상기 프레임워크는 벡터 베이스 진폭 패닝 (vector base amplitude panning; VBAP), 거리 기반 진폭 패닝, 또는 다른 형태들의 패닝과 같은 몇몇 형태의 패닝을 포함하도록 수정될 수도 있다. 예시의 목적을 위해 VBAP 에 초점을 두어, VBAP 는 "가상 스피커들" 로서 특징지어질 수도 있는 것을 유효하게 도입할 수도 있다. VBAP 는, 가상 스피커를 지원하는 그 하나 이상의 라우드스피커들의 로케이션 및/또는 각도 중 적어도 하나와 상이한 로케이션 및 각도 중 하나 이상에서 가상 스피커로부터 발생되어 나오는 것처럼 보이는 사운드를 하나 이상의 라우드스피커들이 출력하도록 이들 하나 이상의 라우드스피커들에 대한 피드 (feed) 를 수정할 수도 있다.

[1291] 예시를 위해, SHC 의 면에서 라우드스피커 피드들을 결정하기 위한 식은 다음과 같다:

$$\begin{bmatrix} A_0^0(\omega) \\ A_1^1(\omega) \\ A_1^{-1}(\omega) \\ \dots \\ A_{(Order+1)(Order+1)}^{-(Order+1)(Order+1)}(\omega) \end{bmatrix} = -ik \begin{bmatrix} VBAP \\ MATRIX \\ MxN \end{bmatrix} \begin{bmatrix} D \\ Nx(Order+1)^2 \end{bmatrix} \begin{bmatrix} g_1(\omega) \\ g_2(\omega) \\ g_3(\omega) \\ \dots \\ g_M(\omega) \end{bmatrix}.$$

[1292]

[1293] 상기 식에서, VBAP 행렬은 M 행에 N 열들의 사이즈의 것이고, 여기서, M 은 스피커들의 수를 나타내고 (그리고 상기 식에서는 5 와 동일할 것이다) N 은 가상 스피커들의 수를 나타낸다. VBAP 행렬은 스피커들의 포지션들의 각각에 대한 청취자의 정의된 로케이션으로부터의 벡터들 및 가상 스피커들의 포지션들의 각각에 대한 청취자의 정의된 로케이션으로부터의 벡터들의 함수로서 컴퓨팅될 수도 있다. 상기 식에서의 D 행렬은 N 행들에 $(order+1)^2$ 열들의 사이즈의 것일 수도 있고, 여기서, 오더 (order) 는 SH 함수들의 오더를 지칭할 수도 있다. D 행렬은 다음을 나타낼 수도 있다:

[1294] 행렬 : $$\begin{bmatrix} h_0^{(2)}(kr_1)Y_0^{0^*}(\theta_1,\varphi_1) & h_0^{(2)}(kr_2)Y_0^{0^*}(\theta_2,\varphi_2) & . & . & . \\ h_1^{(2)}(kr_1)Y_1^{1^*}(\theta_1,\varphi_1). & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \end{bmatrix}.$$

[1295] $\boldsymbol{g}$ 행렬 (또는 오직 단일 열만이 존재하면 벡터) 은 디코더-로컬 지오메트리에서 배열된 스피커들에 대한 스피커 피드들에 대한 이득을 나타낼 수도 있다. 식에서, $\boldsymbol{g}$ 행렬은 M 사이즈의 것이다. $\boldsymbol{A}$ 행렬 (또는 오직 단일 열만이 존재하면 벡터) 은 SHC (520) 를 나타낼 수도 있고, $(Order+1)^2$ 으로서 또한 표시될 수도 있는 (Order+1)(Order+1) 의 사이즈의 것이다.

[1296] 사실상, VBAP 행렬은 스피커들의 로케이션 및 가상 스피커들의 포지션을 요인으로 포함하는 "이득 조정" 으로서 지칭될 수도 있는 것을 제공하는 MxN 행렬이다. 이러한 방식으로 패닝을 도입하는 것은, 로컬 스피커 지오메트리에 의해 재생되는 더 양호한 품질의 이미지를 초래하는 다중 채널 오디오의 더 양호한 재생을 초래할 수도 있다. 또한, VBAP 를 이 식에 통합함으로써, 본 기술들은 다양한 표준들에서 명시된 것들과 정렬되지 않는 열악한 스피커 지오메트리들을 극복할 수도 있다.

[1297] 실제로, 식은, 이 개시물에서 디코더-로컬 지오메트리로서 다시 지칭될 수도 있는 라우드스피커들의 특정 지오메트리 또는 구성에 대해 SHC 를 다시 역으로 다중 채널 피드들로 변환하기 위해 인버트되고 채용될 수도 있다. 즉, 식은 $\boldsymbol{g}$ 행렬에 대해 풀기 위해 인버트될 수도 있다. 인버트된 식은 다음과 같을 수도 있다:

[1298] $$\begin{bmatrix} g_1(\omega) \\ g_2(\omega) \\ g_3(\omega) \\ \dots \\ g_M(\omega) \end{bmatrix} = -ik \begin{bmatrix} VBAP \\ MATRIX^{-1} \\ MxN \end{bmatrix} \begin{bmatrix} D^{-1} \\ Nx(Order+1)^2 \end{bmatrix} \begin{bmatrix} A_0^0(\omega) \\ A_1^1(\omega) \\ A_1^{-1}(\omega) \\ \dots \\ A_{(Order+1)(Order+1)}^{-(Order+1)(Order+1)}(\omega) \end{bmatrix}.$$

[1299] $\boldsymbol{g}$ 행렬은, 이 예에서 5.1 스피커 구성에서의 5 개의 라우드스피커들의 각각에 대한 스피커 이득을 나타낼 수도 있다. 이 구성에서 사용된 가상 스피커들 로케이션들은 5.1 멀티채널 포맷 사양 또는 표준에서 정의된 로케이션들에 대응할 수도 있다. 이들 가상 스피커들의 각각을 지원할 수도 있는 라우드스피커들의 로케이션은 임의의 수의 알려진 오디오 로컬화 기술들을 이용하여 결정될 수도 있고, 이들 알려진 오디오 로컬화 기술들의 많은 것들은 (오디오/비디오 수신기 (A/V 수신기), 텔레비젼, 게이밍 시스템, 디지털 비디오 디스크 시스템, 또는 다른 유형들의 헤드엔드 시스템들과 같은) 헤드엔드 유닛에 대한 각각의 라우드스피커의 로케이션을 결정하기 위해 특정 주파수를 갖는 톤을 플레이하는 것을 수반한다. 대안적으로, 헤드엔드 유닛의 사용자는 라우드스피커들의 각각의 로케이션을 수동으로 특정할 수도 있다. 어느 경우에도, 이들 알려진 로케이션들 및 가능한 각도들이 주어지면, 헤드엔드 유닛은 VBAP 를 통해 가상 라우드스피커들의 이상적인 구성을 가정하여 이득들에 대해 풀 수도 있다.

[1300] 이러한 점에서, 디바이스 또는 장치는, 상이한 로컬 지오메트리에서 구성된 가상 스피커들로부터 유래되는 것처럼 보이는 사운드들을 방출하기 위해 디코더-로컬 지오메트리에서 스피커들을 구동하는 복수의 채널들을 생성하기 위해 복수의 가상 채널들에 대해 벡터 베이스 진폭 패닝 또는 다른 형태의 패닝을 수행할 수도 있다. 본 기술들은, 따라서, 오디오 디코딩 디바이스 (40) 로 하여금, 복원된 구면 조화 계수들 (47) 과 같은 복수의 구면 조화 계수들에 대해 복수의 채널들을 생성하기 위한 변환을 수행하는 것을 가능하게 할 수도 있다. 복수의 채널들의 각각은 공간의 대응하는 상이한 구역과 연관될 수도 있다. 더욱이, 복수의 채널들의 각각은 복수의 가상 채널들을 포함할 수도 있고, 여기서, 복수의 가상 채널들은 대응하는 공간의 상이한 구역과 연관될 수도 있다. 디바이스는, 따라서, 다중 채널 오디오 데이터 (49) 의 복수의 채널을 생성하기 위해 가상 채널들에 대해 벡터 베이스 진폭 패닝을 수행할 수도 있다.

[1301] 도 49a 내지 도 49e(ii) 는 이 개시물에서 기술된 기술들의 다양한 양태들을 구현할 수도 있는 각각의 오디오 코딩 시스템들 (560A-560C, 567D, 569D, 571E 및 573E) 을 나타내는 도들이다. 도 49a 의 예에서 나타난 바와 같이, 오디오 코딩 시스템 (560A) 은 오디오 인코딩 디바이스 (562) 및 오디오 디코딩 디바이스 (564) 를 포함할 수도 있다. 오디오 인코딩 디바이스 (562) 는 도 4 및 도 40a 내지 도 40d 의 예에서 각각 나타난 오디오 인코딩 디바이스들 (20 및 510A-510D) 중 임의의 것과 유사할 수도 있다. 오디오 디코딩 디바이스

(564) 는 도 5 및 도 41 의 예에서 나타난 오디오 디코딩 디바이스 (24 및 40) 와 유사할 수도 있다.

[1302] 상기 기술된 바와 같이, 고차 앰비소닉스 (higher-order ambisonics; HOA) 는 공간 퓨리에 변환에 기초하여 음장의 모든 방향성 정보를 기술하는 방식이다. 일부 예들에서, 앰비소닉스 오더 N 이 더 높을 수록, 공간 분해능이 더 높아지고, 구면 조화 (spherical harmonics; SH) 계수들 $(N+1)^2$ 의 수가 더 커진다. 따라서, 일부 예들에서, 더 높은 앰비소닉스 오더 N 은 계수들을 송신 및 저장하기 위한 더 큰 대역폭 요건들을 초래한다.

HOA 의 대역폭 요건들은 예를 들어 5.1 또는 7.1 서라운드 사운드 오디오 데이터에 비해 오히려 높기 때문에, 대역폭 감소가 많은 애플리케이션들에 대해 소망될 수도 있다.

[1303] 이 개시물에서 설명된 기술들에 따르면, 오디오 코딩 시스템 (560A) 은 공간 사운드 장면에서 불구별되는 (non-distinct) (배경 또는 환경) 엘리먼트들로부터 구별되는 (distinct) (전경) 엘리먼트를 분리하는 것에 기초한 방법을 수행할 수도 있다. 이러한 분리는 오디오 코딩 시스템 (560A) 으로 하여금, 전경 및 배경 엘리먼트들을 서로 독립적으로 프로세싱하는 것을 허용할 수도 있다. 이 예에서, 오디오 코딩 시스템 (560A) 은, 전경 엘리먼트들이 (청취자에 의한) 보다 많은 주의를 끌 수도 있고 배경 엘리먼트들에 비해 (다시 청취자에 의해) 로컬화하기가 더 쉬울 수도 있는 특성을 이용한다. 그 결과로서, 오디오 코딩 시스템 (560A) 은 HOA 컨텐츠를 보다 효율적으로 저장 또는 송신할 수도 있다.

[1304] 일부 예들에서, 오디오 코딩 시스템 (560A) 은 특이값 분해 (SVD) 프로세스를 채용함으로써 이러한 분리를 달성할 수도 있다. SVD 프로세스는 HOA 계수들의 프레임을 3 개의 행렬들 (U, S, V) 로 분리할 수도 있다. 행렬 U 는 좌-특이 벡터들을 포함하고, V 행렬은 우-특이 벡터들을 포함한다. 대각 행렬 S 는 비-네거티브, 분류된 특이값들을 그것의 대각선에 포함한다. HOA 계수들의 일반적으로 양호한 (또는, 일부 경우들에서, HOA 계수들을 나타냄에 있어서 무제한의 정확도를 가정하는 완벽한) 재구성은, U*S*V' 에 의해 주어질 것이다.

오직 D 최대 특이값들: U(:,1:D)*S(1:D,:)*V' 로 부공간을 재구성함으로써, 오디오 코딩 시스템 (560A) 은 이 HOA 프레임, 즉, 전경 사운드 엘리먼트들로부터 가장 핵심적인 공간 정보 (및 아마도 몇몇 강한 조기 실내 반향들) 를 추출할 수도 있다. 나머지 U(:,D+1:end)*S(D+1:end,:)*V' 는 배경 엘리먼트들 및 컨텐츠로부터의 잔향을 재구성할 수도 있다.

[1305] 오디오 코딩 시스템 (560A) 은, S 의 하향 대각선 값들에 의해 형성된 곡선의 기울기를 분석함으로써 2 개의 부공간들을 분리하는 값 D 를 결정할 수도 있고, 여기서, 큰 특이값들은 전경 사운드들을 나타내고, 낮은 특이값들은 배경 값들을 나타낸다. 오디오 코딩 시스템 (560A) 은 특이값 곡선의 제 1 및 제 2 미분을 이용할 수도 있다. 오디오 코딩 시스템 (560A) 은 또한 넘버 D 를 1 과 5 사이에 있도록 제한할 수도 있다. 대안적으로, 오디오 코딩 시스템 (560A) 은 4 의 값에 대해서와 같이 넘버 D 를 사전-정의할 수도 있다. 어느 경우에도, 일단 넘버 D 가 추정되면, 오디오 코딩 시스템 (560A) 은 행렬들 U, 및 S 로부터 전경 및 배경 부공간을 추출한다.

[1306] 오디오 코딩 시스템 (560A) 은 그 다음, HOA 계수들의 $(N+1)^2$ 채널들을 초래하는, U(:,D+1:end)*S(D+1:end,:)*V' 를 통해 배경 장면의 HOA 계수들을 재구성할 수도 있다. 배경 엘리먼트들은, 일부 예들에서, 현저하지 않고 전경 엘리먼트들에 대해 로컬화가능하지 않은 것으로 알려져 있기 때문에, 오디오 코딩 시스템 (560A) 은 HOA 채널들의 오더를 절단하여 줄일 수도 있다. 또한, 오디오 코딩 시스템 (560A) 은, AAC 와 같은 손실있는 또는 무손실의 오디오 코덱들로, 또는 선택적으로 돌출하는 전경 엘리먼트들을 압축하기 위해 사용되는 것에 비해 보다 공격적인 오디오 코덱으로, 이들 채널들을 압축할 수도 있다.

일부 경우들에서, 대역폭을 절약하기 위해, 오디오 코딩 시스템 (560A) 은 전경 엘리먼트들을 상이하게 송신할 수도 있다. 즉, 오디오 코딩 시스템은, (AAC 와 같은) 손실있는 또는 무손실의 오디오 코덱들로 압축된 후에 좌-특이 벡터들 U(:,1:D) 을 송신할 수도 있고, 재구성 행렬 R=S(1:D,:)*V' 과 함께 이들 압축된 좌-특이 값들을 송신할 수도 있다. R 은 $D \times (N+1)^2$ 행렬을 나타낼 수도 있고, 이는 프레임들에 걸쳐 상이할 수도 있다.

[1307] 오디오 코딩 시스템 (560) 의 수신기 측에서, 오디오 코딩 시스템은 이들 2 개의 행렬들을 곱하여 $(N+1)^2$ HOA 채널들의 프레임을 재구성할 수도 있다. 일단 배경 및 전경 HOA 채널들이 함께 합산되면, 오디오 코딩 시스템 (560A) 은 임의의 적절한 앰비소닉스 렌더러를 이용하여 임의의 라우드스피커 셋업에 렌더링할 수도 있다.

본 기술들은 배경 엘리먼트들로부터 전경 엘리먼트들 (다이렉트 또는 구분 사운드) 의 분리를 제공하기 때문에, 청취력 약화된 사람은 명료도를 증가시키기 위해 배경 엘리먼트들에 대한 전경의 혼합을 제어할 수 있을 것이다. 또한, 예컨대 단지 전경 엘리먼트들에 대한 동적 압축기와 같은 다른 오디오 효과들이 또한 적용가능

할 수도 있다.

[1308] 도 49b 는 오디오 인코딩 시스템 (560B) 을 더 상세하게 나타내는 블록 다이어그램이다. 도 49b 의 예에서 나타난 바와 같이, 오디오 코딩 시스템 (560B) 은 오디오 인코딩 디바이스 (566) 및 오디오 디코딩 디바이스 (568) 를 포함할 수도 있다. 오디오 인코딩 디바이스 (566) 는 도 4 및 도 40e 의 예에서 나타난 오디오 인코딩 디바이스들 (24 및 510E) 과 유사할 수도 있다. 오디오 디코딩 디바이스 (568) 는 도 5 및 도 41b 의 예에서 나타난 오디오 디코딩 디바이스 (24 및 540B) 와 유사할 수도 있다.

[1309] 이 개시물에서 기술된 본 기술들에 따라, HoA 신호들에 대한 프레임 기반 SVD (또는 KLT & PCA 와 같은 관련된 방법들) 분해를 이용할 때, 대역폭 감소의 목적을 위해, 오디오 인코딩 디바이스 (66) 는 (S 행렬의 대응하는 특이값들에 의해 곱해진) U 행렬의 처음 소수의 벡터들 및 V^T 벡터의 대응하는 벡터들을 양자화할 수도 있다. 이것은 음장의 '전경' 성분들을 포함할 것이다. 본 기술들은 오디오 인코딩 디바이스 (566) 로 하여금 '블랙-박스' 오디오-코딩 엔진을 이용하여 $U_{DIST} * S_{DIST}$ 벡터를 코딩하는 것을 가능하게 할 수도 있다. V 벡터는 스칼라 또는 벡터 양자화 중 어느 일방으로 양자화될 수도 있다. 또한, U 행렬의 나머지 벡터들의 일부 또는 전부는 S 행렬 및 V 행렬의 대응하는 특이값들과 곱해지고 또한 '블랙-박스' 오디오-코딩 엔진을 이용하여 코딩될 수도 있다. 이들은 음장의 "배경' 성분들을 포함할 것이다.

[1310] 소리가 가장 큰 청각 성분들은 '전경 성분들' 로 분해되기 때문에, 오디오 인코딩 디바이스 (566) 는, 배경이 중요한 로컬화가능한 컨텐츠를 포함하지 않기 때문에 (우리가 그렇게 가정한다), 오디오 인코딩 디바이스 (566) 는 '블랙-박스' 오디오-코딩 엔진을 이용하기 이전에 '배경' 성분들의 앰비소닉스 오더를 감소시킬 수도 있다. 전경 성분들의 앰비소닉스 오더에 의존하여, 오디오 인코딩 유닛 (566) 은 약간 클 수도 있는 대응하는 V-벡터(들)를 송신할 수도 있다. 예를 들어, V 벡터들의 단순한 16 비트 스칼라 양자화는 대략적으로 전경 성분당 4 차 오더에 대해 20kbps 오버헤드 (25 계수들) 및 6 차 오더에 대해 40kbps (49 계수들) 를 초래할 것이다. 이 개시물에서 설명된 기술들은 이 V-벡터의 이 오버헤드를 감소시키기 위한 방법을 제공할 수도 있다.

[1311] 예시를 위해, 전술한 바와 같이, 전경 엘리먼트들의 앰비소닉스 오더는 N_{DIST} 이고, 배경 엘리먼트들의 앰비소닉스 오더는 N_{BG} 이다. 오디오 인코딩 디바이스 (566) 는 전술한 바와 같이 배경 엘리먼트들의 앰비소닉스 오더를 감소시킬 수도 있기 때문에, N_{BG} 는 N_{DIST} 보다 더 적을 수도 있다. 수신기 측에서 전경 엘리먼트들을 재구성하기 위해 송신될 필요성이 있는 전경 V-벡터의 길이는 전경 엘리먼트 당 $(N_{DIST}+1)^2$ 의 길이를 가지는 반면, 제 1 $((N_{DIST}+1)^2)$ - $((N_{BG}+1)^2)$ 계수들은 오더 N_{BG} 까지 전경 또는 구별되는 성분들을 재구성하기 위해 이용될 수도 있다. 이 개시물에서 설명될 기술들을 이용하여, 오디오 인코딩 디바이스 (566) 는 오더 N_{BG} 까지 전경을 재구성하고, 결과적인 $(N_{BG}+1)^2$ 채널들을 배경 채널들과 병합할 수도 있고, 오더 N_{BG} 까지 완전한 음장을 발생시킬 수도 있다. 오디오 인코딩 디바이스 (566) 는 그 다음, V-벡터를 송신을 위해 $(N_{BG}+1)^2$ 보다 더 높은 인덱스를 갖는 그들 계수들로 감소시킬 수도 있다 (여기서, 이들 벡터들은 "V^T_{SMALL}" 로서 지칭될 수도 있다).

수신기 측에서, 오디오 디코딩 유닛 (568) 은 전경 엘리먼트들을 V^T_{SMALL} 벡터들에 의해 곱함으로써 N_{BG} 보다 더 큰 앰비소닉스 오더에 대해 전경 오디오-채널들을 재구성할 수도 있다.

[1312] 도 49c 는 오디오 인코딩 시스템 (560C) 을 더 상세하게 나타내는 블록 다이어그램이다. 도 49c 의 예에서 나타난 바와 같이, 오디오 코딩 시스템 (560B) 은 오디오 인코딩 디바이스 (567) 및 오디오 디코딩 디바이스 (569) 를 포함할 수도 있다. 오디오 인코딩 디바이스 (567) 는 도 4 및 도 40f 의 예에서 나타난 오디오 인코딩 디바이스들 (20 및 510F) 과 유사할 수도 있다. 오디오 디코딩 디바이스 (569) 는 도 5 및 도 41b 의 예에서 나타난 오디오 디코딩 디바이스들 (24 및 540B) 과 유사할 수도 있다.

[1313] 이 개시물에서 기술된 본 기술들에 따라, HoA 신호들에 대한 프레임 기반 SVD (또는 KLT & PCA 와 같은 관련된 방법들) 분해를 이용할 때, 대역폭 감소의 목적을 위해, 오디오 인코딩 디바이스 (567) 는 (S 행렬의 대응하는 특이값들에 의해 곱해진) U 행렬의 처음 소수의 벡터들 및 V^T 벡터의 대응하는 벡터들을 양자화할 수도 있다. 이것은 음장의 '전경' 성분들을 포함할 것이다. 본 기술들은 오디오 인코딩 디바이스 (567) 로 하여금 '블랙-박스' 오디오-코딩 엔진을 이용하여 $U_{DIST} * S_{DIST}$ 벡터를 코딩하는 것을 가능하게 할 수도 있다. V 벡터

는 스칼라 또는 벡터 양자화 중 어느 일방으로 양자화될 수도 있다. 또한, U 행렬의 나머지 벡터들의 일부 또는 전부는 S 행렬 및 V 행렬의 대응하는 특이값들과 곱해지고 또한 '블랙-박스' 오디오-코딩 엔진을 이용하여 코딩될 수도 있다. 이들은 음장의 '배경' 성분들을 포함할 것이다.

[1314] 소리가 가장 큰 청각 성분들은 '전경 성분들' 로 분해되기 때문에, 오디오 인코딩 디바이스 (567) 는, 배경이 중요한 로컬화가능한 컨텐츠를 포함하지 않기 때문에 (우리가 그렇게 가정한다), 오디오 인코딩 디바이스 (566) 는 '블랙-박스' 오디오-코딩 엔진을 이용하기 이전에 '배경' 성분들의 앰비소닉스 오더를 감소시킬 수도 있다.

오디오 인코딩 디바이스 (567) 는 본 명세서에서 설명된 기술들에 따라 음장의 전체 에너지를 보존하는 그러한 방식으로 오더를 감소시킬 수도 있다. 전경 성분들의 앰비소닉스 오더에 의존하여, 오디오 인코딩 유닛 (567) 은 약간 클 수도 있는 대응하는 V-벡터(들)를 송신할 수도 있다. 예를 들어, V 벡터들의 단순한 16 비트 스칼라 양자화는 대략적으로 전경 성분 당 4 차 오더에 대해 20kbps 오버헤드 (25 계수들) 및 6 차 오더에 대해 40kbps (49 계수들) 를 초래할 것이다. 이 개시물에서 설명된 기술들은 이 V-벡터(들)의 이 오버헤드를 감소시키기 위한 방법을 제공할 수도 있다.

[1315] 예시를 위해, 전경 엘리먼트들의 그리고 배경 엘리먼트들의 앰비소닉스 오더는 N 이라고 가정한다. 오디오 인코딩 디바이스 (567) 는 $\tilde{\eta} < N$ 이도록 N 에서부터 $\tilde{\eta}$ 으로 V-벡터(들)의 배경 엘리먼트들의 앰비소닉스 오더를 감소시킬 수도 있다. 오디오 인코딩 디바이스 (67) 는 SHC 들에 의해 기술되는 음장의 전체 에너지를 보존하기 위해 V-벡터(들)의 배경 엘리먼트들의 값들을 증가시키기 위해 보상을 더 적용한다. 보상을 적용하는 예시적인 기술들은 도 40f 와 관련하여 상기 설명되었다. 수신기 측에서, 오디오 디코딩 유닛 (569) 은 앰비소닉스 오더에 대해 배경 오디오-채널들을 재구성할 수도 있다.

[1316] 도 49d(i) 및 도 49d(ii) 는 오디오 인코딩 디바이스 (567D) 및 오디오 디코딩 디바이스 (569D) 를 각각 나타낸다. 오디오 인코딩 디바이스 (567D) 및 오디오 디코딩 디바이스 (569D) 는, 이 개시물의 양태들에 따라, 하나 이상의 방향-기반 구별성 결정들을 수행하도록 구성될 수도 있다. 고차 앰비소닉스 (HOA) 는 공간 퓨리에 변환에 기초하여 음장의 모든 방향적 정보를 기술하기 위한 방법이다. 앰비소닉스 오더 N 이 더 높을 수록, 공간 분해능이 더 높아지고, 구면 조화 (SH) 계수들 $(N+1)^2$ 의 수가 더 커지고, 데이터 송신 및 저장을 위해 요구되는 대역폭이 더 커지게 된다. HOA 의 대역폭 요건들은 다소 높기 때문에, 많은 애플리케이션들에 대해, 대역폭 감소가 소망된다.

[1317] 이전 설명들은, SVD (특이값 분해) 또는 관련 프로세스들이 어떻게 공간 오디오 압축에 대해 이용될 수도 있는지를 설명하였다. 본 명세서에서 설명된 기술들은 전경 엘리먼트들로서 또한 알려진 돌출하는 엘리먼트들을 선택하기 위한 향상된 알고리즘을 제시한다. HOA 오디오 프레임의 그것의 U, S, 및 V 행렬로의 SVD-기반 분해 후에, 본 기술들은 U 행렬의 처음 K 개의 채널들 [U(:,1:K)*S(1:K,1:K)] 에 배타적으로 K 개의 돌출하는 엘리먼트들의 선택의 기초를 두고 있다. 이는 최고의 에너지를 갖는 오디오 엘리먼트들을 선택하는 것을 초래한다. 하지만, 그들 엘리먼트들이 또한 방향성이라는 것은 보장되지 않는다. 따라서, 본 기술들은 높은 에너지를 가지고 또한 방향성인 사운드 엘리먼트들을 발견하는 것을 지향한다. 이것은 S 행렬로 V 행렬을 가중함으로써 잠재적으로 달성된다. 그 다음, 이 결과적인 행렬의 각 행에 대해, (상위 오더의 HOA 계수들과 연관되는) 더 높게 인덱싱된 엘리먼트들이 제곱되고 합산되어, 행 당 하나의 값이 발생한다 (도 40h 와 관련하여 설명된 의사-코드에서의 합 VS]. 의사-코드에서 표현된 워크플로우에 따르면, 제 5 인덱스에서 시작하는 상위 오더의 앰비소닉스 계수들이 고려된다. 이들 값들은 그들의 사이즈에 따라 분류되고, 분류 인덱스는 원래의 U, S, 및 V 행렬을 이에 따라 재배열하기 위해 사용된다. 이 개시물에서 이전에 기술된 SVD-기반 압축 알고리즘은 그러면 추가적인 수정 없이 적용될 수 있다.

[1318] 도 49e(i) 및 도 49e(ii) 는 오디오 인코딩 디바이스 (571E) 및 오디오 디코딩 디바이스 (573E) 를 각각 나타내는 블록 다이어그램이다. 오디오 인코딩 디바이스 (571E) 및 오디오 디코딩 디바이스 (573E) 는, 오디오 인코딩 디바이스 (571E) 가 S^2 행렬 및 V 행렬을 생성하기 위해 HOA 계수들의 전력 스펙트럼 밀도 행렬 (PDS) 에 대한 특이값 분해를 수행하는 점을 제외하고는, 도 49 내지 도 49d(ii) 의 예들에 대해 상기 설명된 기술들의 다양한 양태들을 수행할 수도 있다. S^2 행렬은 제곱된 S 행렬을 나타낼 수도 있고, 그래서, S^2 행렬은 S 행렬을 획득하기 위해 제곱근 연산을 거칠 수도 있다. 오디오 인코딩 디바이스 (571E) 는, 일부 경우들에서, (V' 행렬로서 표시될 수도 있는) 양자화된 V 행렬을 획득하기 위해 V 행렬에 대해 양자화를 수행할 수도 있다.

[1319] 오디오 인코딩 디바이스 (571E) 는 SV' 행렬을 생성하기 위해 양자화된 V' 행렬에 의해 S 행렬을 먼저 곱합으로

써 U 행렬을 획득할 수도 있다. 오디오 인코딩 디바이스 (571E) 는 다음으로 SV' 행렬의 의사-역을 획득하고, 그 다음, U 행렬을 획득하기 위해 SV' 행렬의 의사-역에 의해 HOA 계수들을 곱할 수도 있다. 계수들 그 자신들보다는 HOA 계수들의 전력 스펙트럼 밀도에 대해 SVD 를 수행함으로써, 오디오 인코딩 디바이스 (571E) 는, 프로세서 사이클들 및 저장 공간 중 하나 이상의 면에서 SVD 를 수행하는 계산적 복잡성을 잠재적으로 감소시키는 한편, SVD 가 직접 HOA 계수들에 적용된 것처럼 동일한 소스 오디오 인코딩 효율을 달성한다.

[1320] 오디오 디코딩 디바이스 (573E) 는, 오디오 디코딩 디바이스 (573E) 가 HOA 계수들에 직접적으로 보다는 HOA 계수들의 전력 스펙트럼 밀도에 대한 SVD 의 적용을 통해 달성된 HOA 계수들의 분해들로부터 HOA 계수들을 재구성할 수도 있다는 점을 제외하고는, 상기 기술된 그들 오디오 디코딩 디바이스들과 유사할 수도 있다.

[1321] 도 50a 및 도 50b 는 이 개시물에서 설명된 기술들에 따라 배경 컨텐츠의 오더를 잠재적으로 감소시키기 위한 2 가지 상이한 접근법들 중의 하나를 각각 나타내는 블록도들이다. 도 50 의 예에서 나타난 바와 같이, 제 1 접근법은 $\tilde{\eta}$ 이 N 보다 적을 (<) 때, 오더를 N 으로부터 $\tilde{\eta}$ 으로 감소시키기 위해 $U_{BG} * S_{BG} * V^T$ 벡터들에 대해 오더-감소를 채용할 수도 있다. 즉, 도 40b 내지 도 40j 의 예에서 나타난 오더 감소 유닛 (528A) 은 $\tilde{\eta}$ 이 N 보다 적을 (<) 때, $U_{BG} * S_{BG} * V^T$ 벡터들의 오더 N 을 절단하거나 그 외에 감소시키기 위해 오더-감소를 수행할 수도 있다.

[1322] 대안적인 접근법으로서, 오더 감소 유닛 (528A) 은, 도 50b 의 예에서 나타난 바와 같이, $(\tilde{\eta}+1)^2$ 이 되도록 로우들을 제거하는 V^T 에 대한 절단 (truncation) 을 수행할 수도 있고, 이는 예시의 용이함을 위해 도 40b 의 예에서는 나타내지 않았다. 다르게 말하면, 오더 감소 유닛 (528A) 은 V_{BG} 행렬을 유효하게 생성하기 위해 V^T 행렬의 하나 이상의 오더들을 제거할 수도 있다. V_{BG} 행렬의 사이즈는 $(\tilde{\eta}+1)^2 \times (N+1)^2$ -D 이고, 여기서, V_{BG} 행렬은 그 다음, 사이즈 $M \times (\tilde{\eta}+1)^2$ 의 $U_{BG} * S_{BG} * V^T$ 벡터들을 생성하기 위해 절단을 유효하게 수행함에 있어서, $U_{BG} * S_{BG} * V^T$ 벡터들을 생성할 때 V^T 행렬 대신에 이용된다.

[1323] 도 51 은 구면 조화 계수들 (701) 을 압축하기 위해 이 개시물에서 설명된 기술들의 다양한 양태들을 구현할 수도 있는 오디오 인코딩 디바이스 (700A) 의 구별되는 성분 압축 경로의 예들을 나타내는 블록 다이어그램이다.
도 51 의 예에서, 구별되는 성분 압축 경로는, SHC (701) 에 의해 표현되는 음장의 구별되는 성분들을 압축하는 오디오 인코딩 디바이스 (700A) 의 프로세싱 경로를 지칭할 수도 있다. 배경 성분 압축 경로로서 지칭될 수도 있는 또 다른 경로는 SHC (701) 의 배경 성분들을 압축하는 오디오 인코딩 디바이스 (700A) 의 프로세싱 경로를 나타낼 수도 있다.

[1324] 예시의 용이함을 위해 나타내지 않았지만, 배경 성분 압축 경로는 SHC (701) 의 분해들에 대해서 보다는 SHC (701) 에 대해 직접 작용할 수도 있다. 이것은, U_{BG}, S_{BG} 및 V_{BG} 행렬들로부터의 배경 성분들을 개조한 다음 이들 개조된 배경 성분들의 (예컨대, AAC 인코더를 이용한) 심리 음향적 인코딩의 몇몇 형태를 수행하기보다는, 배경 성분 프로세싱 경로는 (도 4 의 예에서 나타난 오디오 인코딩 디바이스 (20) 에 대해 상기 설명된 바와 같이) SHC (701) 에 대해 직접 심리 음향적 인코더를 이용하여 이들 배경 성분들을 압축하는 것을 동작할 수도 있다는 점을 제외하고는, 도 49 내지 도 49c 에 대해 상기 설명된 것과 유사하다. SHC (701) 에 대해 직접 심리 음향적 인코딩을 수행함으로써, 개조된 배경 성분들에 대해 심리 음향적 인코딩을 수행하는 것에 비해, 불연속성들이 감소될 수도 있는 한편, (배경 성분들을 압축하기 위해 필요한 동작들의 면에서) 컴퓨테이션 복잡성을 또한 감소시킨다. 비록 구별되는 및 배경의 면에서 지칭되었지만, 이 개시물에서 "돌출하는" 이란 용어는 "구별되는" 대신에 사용될 수도 있고, "주변의" 라는 용어는 "배경" 대신에 사용될 수도 있다.

[1325] 어느 경우에도, 구면 조화 계수들 (701) ("SHC (701)") 은 $M \times (N+1)^2$ 의 사이즈를 갖는 계수들의 행렬을 포함할 수도 있고, 여기서, M 은 오디오 프레임에서의 샘플들의 수 (일부 경우들에서, 1024) 를 나타내고, N 은 계수들이 대응하는 기초 함수의 최고 오더를 나타낸다. 상기 언급된 바와 같이, N 은 통상적으로 총 1024 × 25 계수들에 대해 4 로 설정된다. 특정 오더, 서브-오더 조합에 대응하는 SHC (701) 의 각각은 채널로 지칭될 수도 있다. 예를 들어, 제 1 오더, 제로 서브-오더 기초 함수에 대응하는 M 개의 샘플 계수들의 전부가 채널을 나타낼 수도 있는 한편, 제로 오더, 제로 서브-오더 기초 함수에 대응하는 계수들은 다른 채널을 나타낼

수도 있는 등이다. SHC (701) 는 또한 이 개시물에서 고차 앰비소닉스 (HOA) 컨텐츠 (701) 로서 또는 SH 신호 (701) 로서 지칭될 수도 있다.

[1326] 도 51 의 예에서 나타난 바와 같이, 오디오 인코딩 디바이스 (700A) 는 분석 유닛 (702), 벡터 기반 합성 유닛 (704), 벡터 감소 유닛 (706), 심리 음향 인코딩 유닛 (708), 계수 감소 유닛 (710) 및 압축 유닛 (712) ("압축 유닛 (712)") 을 포함한다. 분석 유닛 (702) 은, 음장의 구별되는 성분들 (D) (703) 및 배경 성분들의 총 수 (BG_{TOT}) (705) 를 식별하도록 SHC (701) 에 대해 분석을 수행하도록 구성된 유닛을 나타낼 수도 있다. 상기 기술된 오디오 인코딩 디바이스들에 비해, 오디오 인코딩 디바이스 (700A) 는 SHC (701) 의 분해들에 대해 이 결정을 수행하지 않고, SHC (701) 에 대해 직접 이 결정을 수행한다.

[1327] 벡터 기반 합성 유닛 (704) 은, SVD 의 경우들에서, M × $(N+1)^2$ 의 사이즈를 갖는 [US] 행렬 (707) 및 $(N+1)^2$ × $(N+1)^2$ 의 사이즈를 갖는 [V] 행렬 (709) 을 생성하기 위해, SVD, KLT, PCA 또는 임의의 다른 벡터 기반 합성과 같은, SHC 701) 에 대한, 몇몇 형태의 벡터 기반 합성을 수행하도록 구성된 유닛을 나타낸다. [US] 행렬 (707) 은 SHC (701) 에 대한 SVD 의 적용을 통해 생성된 [U] 행렬 및 [S] 행렬의 행렬 곱으로부터 발생하는 행렬을 나타낼 수도 있다.

[1328] 벡터 감소 유닛 (706) 은, [US] 행렬 (707) 및 [V] 행렬 (709) 의 나머지 벡터들의 각각이 음장의 구별되는 또는 돌출하는 성분을 식별하도록, [US] 행렬 (707) 및 [V] 행렬 (709) 의 벡터들의 수를 감소시키도록 구성된 유닛을 나타낼 수도 있다. 벡터 감소 유닛 (706) 은 구별되는 성분들의 수 (D) (703) 에 기초하여 이 감소를 수행할 수도 있다. 구별되는 성분들의 수 (D) (703) 는, 사실상, 수들의 어레이를 나타낼 수도 있고, 여기서, 각 수는 행렬들 (707 및 709) 의 상이한 구분 벡터들을 식별한다. 벡터 감소 유닛 (706) 은 사이즈 M × D 의 감소된 [US] 행렬 (711) 및 사이즈 $(N+1)^2$ × D 의 감소된 [V] 행렬 (713) 을 출력할 수도 있다.

[1329] 비록 예시의 용이함을 위해 나타내지는 않았지만, [V] 행렬 (709) 의 보간은 상기 더 상세하게 설명된 것과 유사한 방식으로 [V] 행렬 (709) 의 감소 이전에 발생할 수도 있다. 더욱이, 비록 예시의 용이함을 위해 나타내지는 않았지만, 감소된 [US] 행렬 (711) 및/또는 감소된 [V] 행렬 (712) 의 리오더링은 상기 더 상세하게 설명된 방식의 것이다. 따라서, 본 기술들은 (도 51 의 예에서 나타내지는 않았지만 상기 설명된 전술한 기술들의 에러 투영 또는 임의의 다른 양태와 같은) 이들 및 다른 점들에 제한되어서는 안된다.

[1330] 심리 음향 인코딩 유닛 (708) 은 비트스트림 (715) 을 생성하기 위해 [US] 행렬 (711) 에 대해 심리 음향적 인코딩을 수행하도록 구성된 유닛을 나타낸다. 계수 감소 유닛 (710) 은 감소된 [V] 행렬 (713) 의 채널들의 수를 감소시키도록 구성된 유닛을 나타낼 수도 있다. 다르게 말하면, 계수 감소 유닛 (710) 은, 방향성 정보를 적게 가지거나 전혀 가지지 않는 (감소된 [V] 행렬 (713) 을 형성하는) 별개의 V 벡터들의 그들 계수들을 제거하도록 구성된 유닛을 나타낼 수도 있다. 상기 기술된 바와 같이, 일부 예들에서, (위에서 N_{BG} 로서 표시된) 제 1 및 제로 오더 기초 함수들에 대응하는 별개의 V 벡터들의 그들 계수들은 적은 방향성 정보를 제공하고, 따라서, (상기 "오더 감소" 로서 지칭되는 것을 통해) 별개의 V 벡터들로부터 제거될 수 있다. 이 예에서, N_{BG} 에 대응하는 이들 계수들을 식별하기 위해서 뿐만 아니라 [$(N_{BG}+1)^2$+1, $(N+1)^2$] 의 세트로부터 (가변 TotalOfAddAmbHOAChan 에 의해 표시될 수도 있는) 추가적인 HOA 채널들을 식별하기 위해서 더 큰 유연성이 제공될 수도 있다. 분석 유닛 (702) 은 $(N_{BG}+1)^2$ 뿐만 아니라 TotalOfAddAmbHOAChan 을 식별할 수도 있는 BG_{TOT} 를 결정하기 위해 SHC (701) 를 분석할 수도 있다. 계수 감소 유닛 (710) 은 그 다음, 사이즈 ($(N+1)^2$ - (BG_{TOT}) × D 의 작은 [V] 행렬 (717) 을 생성하기 위해 감소된 [V] 행렬 (713) 로부터 $(N_{BG}+1)^2$ 및 TotalOfAddAmbHOAChan 에 대응하는 그들 계수들을 제거할 수도 있다.

[1331] 압축 유닛 (712) 은 그 다음, 작은 [V] 행렬 (717) 을 압축하기 위해 상기 언급된 스칼라 양자화 및/또는 허프만 인코딩을 수행할 수도 있고, 사이드 채널 정보 (719) ("사이드 채널 정보 (719)") 로서 그 압축된 작은 [V] 행렬 (717) 을 출력한다. 압축 유닛 (712) 은 도 10 내지 도 10o(ii) 의 예에서 나타난 것과 유사한 방식으로 사이드 채널 정보 (719) 를 출력할 수도 있다. 일부 예들에서, 상기 기술된 것들과 유사한 비트스트림 생성 유닛은 사이드 채널 정보 (719) 를 비트스트림 (715) 내로 통합할 수도 있다. 또한, 비트스트림 (715) 으로서 지칭되었지만, 오디오 인코딩 디바이스 (700A) 는, 상기 언급된 바와 같이, 또 다른 비트스트림을 발생시키는 배경 성분 프로세싱 경로를 포함할 수도 있고, 여기서, 상기 설명된 것들과 유사한 비트스트림 생성 유

닛은 비트스트림 (715) 및 배경 성분 프로세싱 경로에 의해 출력되는 비트스트림을 포함하는 상기 설명된 비트스트림 (17) 과 유사한 비트스트림을 생성할 수도 있다.

[1332] 이 개시물에서 설명된 기술들에 따르면, 분석 유닛 (702) 은 음장의 구별되는 성분을 나타내기 위해 사용될 벡터, 즉 감소된 [V] 행렬 (713) 의 벡터의 계수들의 제 1 비-제로 세트를 결정하도록 구성될 수도 있다. 일부 예들에서, 분석 유닛 (702) 은 감소된 [V] 행렬 (713) 을 형성하는 매 벡터의 계수들의 전부가 사이드 채널 정보 (719) 에 포함되어야 하는 것을 결정할 수도 있다. 분석 유닛 (702) 은 따라서 제로와 동일한 BG_{TOT} 를 설정할 수도 있다.

[1333] 오디오 인코딩 디바이스 (700A) 는 따라서, "디코딩된 벡터들" 로서 표시된 테이블에 대해 상기 설명된 것과 상호적으로 역의 방식으로 유효하게 작용할 수도 있다. 또한, 오디오 인코딩 디바이스 (700A) 는, 그 중에서 복수의 구성 모드들이 선택되었던 (하나 이상의 프레임들을 포함할 수도 있는) 액세스 유닛의 헤더에서 신택스 엘리먼트를 특정할 수도 있다. 비록 액세스 유닛 당 특정되는 것으로서 설명되었지만, 분석 유닛 (702) 은 프레임 당 또는 임의의 다른 주기적 기초 또는 비-주기적 기초 (예컨대, 전체 비트스트림에 대해 한번) 로 이 신택스 엘리먼트를 특정할 수도 있다. 어느 경우에도, 이 신택스 엘리먼트는, 이 구별되는 성분의 방향성 양태들을 나타내기 위해 감소된 [V] 행렬 (713) 의 계수들의 비-제로 세트를 명시하기 위해 4 개의 구성 모드들이 선택되었던 것을 나타내는 2 비트들을 포함할 수도 있다. 신택스 엘리먼트는 "codedVVecLength" 로서 표시될 수도 있다. 이러한 방식으로, 오디오 인코딩 디바이스 (700A) 는 비트스트림에서 작은 [V] 행렬 (717) 을 명시하기 위해 4 개의 구성 모드들이 사용되었던 비트스트림에서 시그널링 또는 그 외에 명시할 수도 있다.
비록 4 개의 구성 모드들에 대해 설명되었지만, 본 기술들은 4 개의 구성 모드들에 제한되어서는 아니되고, 단일 구성 모드 또는 복수의 구성 모드들을 포함하는 임의의 수의 구성 모드들이 가능하다.

[1334] 본 기술들의 다양한 양태들은 따라서, 오디오 인코딩 디바이스 (700A) 로 하여금 다음의 조항들에 따라 동작하도록 구성되는 것을 가능하게 할 수도 있다.

[1335] 조항 133149-1F. 디바이스는 벡터의 계수들의 비-제로 세트를 명시하기 위해 복수의 구성 모드들 중 하나의 구성 모드를 선택하고 (벡터는 음장을 서술하고 음장의 구별되는 성분을 나타내는 복수의 구면 조화 계수들로부터 분해된다), 복수의 구성 모드들 중 선택된 하나의 구성 모드에 기초하여 벡터의 계수들의 비-제로 세트를 명시하도록 구성된 하나 이상의 프로세서들을 포함한다.

[1336] 조항 133149-2F. 조항 133149-1F 의 디바이스, 여기서 복수의 구성 모드들 중 하나의 구성 모드는 계수들의 비-제로 세트가 계수들 모두를 포함한다고 나타낸다.

[1337] 조항 133149-3F. 조항 133149-1F 의 디바이스, 여기서 복수의 구성 모드들 중 하나의 구성 모드는 계수들의 비-제로 세트가 복수의 구면 조화 계수들 중 하나 이상의 구면 조화 계수들이 대응하는 기저 함수의 오더보다 큰 오더에 대응하는 계수들의 비-제로 세트를 포함한다고 표시한다.

[1338] 조항 133149-4F. 조항 133149-1F 의 디바이스, 여기서 복수의 구성 모드들 중 하나의 구성 모드는 계수들의 비-제로 세트가 복수의 구면 조화 계수들 중 하나 이상의 구면 조화 계수들이 대응하는 기저 함수의 오더보다 큰 오더에 대응하는 계수들의 비-제로 세트를 포함하고 복수의 구면 조화 계수들 중 하나 이상의 구면 조화 계수들이 대응하는 기저 함수의 오더보다 큰 오더에 대응하는 계수들 중 적어도 하나의 계수는 제외한다고 나타낸다.

[1339] 조항 133149-5F. 조항 133149-1F 의 디바이스, 여기서 복수의 구성 모드들 중 하나의 구성 모드는 계수들 중 적어도 하나의 계수를 제외하고 계수들의 비-제로 세트가 계수들 모두를 포함한다고 나타낸다.

[1340] 조항 133149-6F. 조항 133149-1F 의 디바이스, 여기서 하나 이상의 프로세서들은 비트스트림에 복수의 구성 모드들 중 선택된 하나의 구성 모드를 명시하도록 더 구성된다.

[1341] 조항 133149-1G. 디바이스는 복수의 구성 모드들 중 하나의 구성 모드에 따라 벡터의 계수들의 비-제로 세트를 추출하기 위해 복수의 구성 모드들 중 하나의 구성 모드를 결정하고 (벡터는 음장을 서술하고 음장의 구별되는 성분을 나타내는 복수의 구면 조화 계수들로부터 분해된다), 복수의 구성 모드들 중 획득된 하나의 구성 모드에 기초하여 벡터의 계수들의 비-제로 세트를 추출하도록 구성된 하나 이상의 프로세서들을 포함한다.

[1342] 조항 133149-2G. 조항 133149-1G 의 디바이스, 여기서 복수의 구성 모드들 중 하나의 구성 모드는 계수들의 비-제로 세트가 계수들 모두를 포함한다고 나타낸다.

[1343] 조항 133149-3G. 조항 133149-1G 의 디바이스, 여기서 복수의 구성 모드들 중 하나의 구성 모드는 계수들의 비-제로 세트가 복수의 구면 조화 계수들 중 하나 이상의 구면 조화 계수들이 대응하는 기저 함수의 오더보다 큰 오더에 대응하는 계수들의 비-제로 세트를 포함한다고 나타낸다.

[1344] 조항 133149-4G. 조항 133149-1G 의 디바이스, 여기서 복수의 구성 모드들 중 하나의 구성 모드는 계수들의 비-제로 세트가 복수의 구면 조화 계수들 중 하나 이상의 구면 조화 계수들이 대응하는 기저 함수의 오더보다 큰 오더에 대응하는 계수들의 비-제로 세트를 포함하고 복수의 구면 조화 계수들 중 하나 이상의 구면 조화 계수들이 대응하는 기저 함수의 오더보다 큰 오더에 대응하는 계수들 중 적어도 하나의 계수는 제외한다고 나타낸다.

[1345] 조항 133149-5G. 조항 133149-1G 의 디바이스, 여기서 복수의 구성 모드들 중 하나의 구성 모드는 계수들 중 적어도 하나의 계수를 제외하고 계수들의 비-제로 세트가 계수들 모두를 포함한다고 나타낸다.

[1346] 조항 133149-6G. 조항 133149-1G 의 디바이스, 하나 이상의 프로세서들은, 복수의 구성 모드들 중 하나의 구성 모드를 결정하는 경우, 비트스트림에서 시그널링된 값에 기초하여 복수의 구성 모드들 중 하나의 구성 모드를 결정하도록 더 구성된다.

[1347] 도 52 는 SHC (701) 를 재구성하거나 거의 재구성하는 본 개시물에 설명된 기술들의 다양한 양상들을 구현할 수도 있는 오디오 디코딩 디바이스 (750A) 의 다른 예를 예시하는 블록 다이어그램이다. 도 52 의 예에서, 오디오 디코딩 디바이스 (750A) 는, 추출 유닛 (542) 이 비트스트림 (715') (이는, 비트스트림 (715') 이 또한 SHC_{BG} (752) 의 오디오 인코딩된 버전을 포함한다는 점을 제외하고, 도 51 의 예에 대해 전술된 비트스트림 (715) 과 유사하다) 및 사이드 채널 정보 (719) 를 수신한다는 것을 제외하고, 도 41d 의 예에 도시된 오디오 디코딩 디바이스 (540D) 와 유사하다. 이러한 이유로, 추출 유닛은 "추출 유닛 (542')" 으로 표시된다.

[1348] 또한, 추출 유닛 (542') 은 추출 유닛 (542') 이 V 압축해제 유닛 (555) (이는 도 52 의 예에서 "V 압축해재 유닛 (555')" 으로 도시된다) 의 수정된 형태를 포함한다는 점에서 추출 유닛 (542) 과 상이하다. V 압축 유닛 (555') 은 사이드 채널 정보 (719) 및 codedWecLength (754) 라고 표시된 신택스 엘리먼트를 수신한다. 추출 유닛 (542') 은 비트스트림 (715') 으로부터 (그리고, 일 예에서, 비트스트림 (715') 내에 포함된 액세스 유닛 헤더로부터) codedWecLength (754) 를 파싱한다. V 압축해제 유닛 (555') 은 앞서 설명된 구성 모드들 (760) 중 어느 하나에 따라 동작하도록 구성가능한 모드 구성 유닛 (756) ("mode config unit (756)") 및 파싱 유닛 (758) 을 포함한다.

[1349] 모드 구성 유닛 (756) 은 신택스 엘리먼트 (754) 를 수신하고 구성 모드들 (760) 중 하나의 구성 모드를 선택한다. 모드 구성 유닛 (756) 은 그 다음에 구성 모드들 (760) 중 선택된 하나의 구성 모드로 파싱 유닛 (758) 을 구성한다. 파싱 유닛 (758) 은 사이드 채널 정보 (719) 로부터 작은 [V] 벡터들 (717) 의 압축된 형태를 파싱하기 위해 구성 모드들 (760) 중 임의의 하나의 구성 모드에 따라 동작하도록 구성된 유닛을 표현한다. 파싱 유닛 (758) 은 다음의 테이블에서 제시되는 스위치 스테이트먼트에 따라 동작할 수도 있다.

[1350] **테이블 - 디코딩된 벡터들**

신택스	비트들의 수	연상 기호

decodeVVec(i)
{
switch codedVVecLength {
case 0: //벡터 완성
VVecLength = NumOfHoaCoeffs;
for (m=0; m< VVecLength; ++m){ VecCoeff[m] = m+1; }
break;
case 1: //하위 오더들은 제거됨
VVecLength = NumOfHoaCoeffs – MinNumOfCoeffsForAmbHOA;
 	
for (m=0; m< VVecLength; ++m){
VecCoeff[m] = m + MinNumOfCoeffsForAmbHOA + 1;
}
break;
case 2:
VVecLength = NumOfHoaCoeffs – MinNumOfCoeffsForAmbHOA - NumOfAddAmbHoaChan;
n = 0;
for(m=0;m<NumOfHoaCoeffs– MinNumOfCoeffsForAmbHOA; ++m){
c = m + MinNumOfCoeffsForAmbHOA + 1;		

[1351]

```
        if ( ismember(c, AmbCoeffIdx) == 0){
          VecCoeff[n] = c;
           n++;
        }
      }
       break;
    case 3:
       VVecLength = NumOfHoaCoeffs -
NumOfAddAmbHoaChan;
      n = 0;
      for(m=0; m<NumOfHoaCoeffs; ++m){
        c = m + 1;
        if ( ismember(c, AmbCoeffIdx) == 0){
          VecCoeff[n] = c;
           n++;
        }
      }
  }

  if (NbitsQ[i] == 5) { /* 균일한 양자화기 */
    for (m=0; m< VVecLength; ++m){
      VVec(k)[i][m] = (VecValue / 128.0) – 1.0;          8          uimsbf
    }
  }
  else { /* 허프만 디코딩 */
    for (m=0; m< VVecLength; ++m){
      Idx = 5;
      If (CbFlag[i] == 1) {
        idx = (min(3, max(1, ceil(sqrt(VecCoeff[m]) - 1)));
      }
      else if (PFlag[i] == 1) {idx = 4;}
        cid =                                            동적       huffDe
huffDecode(huffmannTable[NbitsQ].codebook[idx]; huffVal);          code
        if ( cid > 0 ) {
          aVal = sgn = (sgnVal * 2) - 1;                 1          bslbf
          if (cid > 1) {
            aVal =  sgn * (2.0^(cid -1 ) + intAddVal);   cid-1      uimsbf
          }
        } else {aVal = 0.0;}
    }
  }
}
```

[1352]

```
주의:
   -  균일한 양자화기를 위한 인코더 함수는 min(255, round((x + 1.0) * 128.0) )

   -  연상 기호 huffDecode 에 대한 비트들의 수는 동적이다
```

[1353]

[1354] 앞서의 신택스 테이블에서, 4 가지 케이스들 (케이스 0-3) 을 갖는 제 1 스위치 스테이트먼트는 계수들의 수의 면에서 작은 [V] 행렬 (717) 의 각각의 벡터의 길이들을 결정하는 방식을 제공한다. 제 1 케이스 (케이스 0) 는 V^{T}_{DIST} 벡터들의 계수들의 모두가 명시된다고 나타낸다. 제 2 케이스 (케이스 1) 는 MinNumOfCoeffsForAmbHOA 보다 큰 오더에 대응하는 V^{T}_{DIST} 벡터의 계수들만이 명시된다고 나타내며, 이는 위에서 $(N_{DIST}+1) - (N_{BG} + 1)$ 로 지칭되는 것을 표시할 수도 있다. 제 3 케이스 (케이스 2) 는 제 2 케이스와 유사하나, NumOfAddAmbHoaChan 으로 식별되는 계수들을 더 감산하며, NumOfAddAmbHoaChan 은 N_{BG} 의 오더를 초과하

는 오더에 대응하는 추가적인 채널들을 명시하기 위한 변수이다 (여기서, "채널들" 은 소정의 오더, 하위-오더 조합에 대응하는 특정 오더를 지칭한다). 제 4 케이스 (케이스 3) 는 NumOfAddAmbHoaChan 으로 식별되는 계수들을 제거한 후에 남겨진 V^{T}_{DIST} 벡터의 계수들만이 명시된다고 나타낸다.

[1355] 이 점에 있어서, 오디오 디코딩 디바이스 (750A) 는 음장의 구별되는 성분을 나타내는 벡터의 계수들의 제 1 비-제로 세트를 결정하기 위해 본 개시물에서 설명된 기술들에 따라 동작할 수도 있으며, 벡터는 음장을 서술하는 복수의 구면 조화 계수들로부터 분해되었다.

[1356] 또한, 오디오 디코딩 디바이스 (750A) 는 복수의 구성 모드들 중 하나의 구성 모드에 따라 벡터의 계수들의 비-제로 세트를 추출하는 복수의 구성 모드들 중 하나의 구성 모드를 결정하기 위해 본 개시물에 설명된 기술들에 따라 동작하고 (벡터는 음장을 서술하고 음장의 구별되는 성분을 나타내는 복수의 구면 조화 계수들로부터 분해되었다), 복수의 구성 모드들 중 획득된 하나의 구성 모드에 기초하여 벡터의 계수들의 비-제로 세트를 추출하도록 구성될 수도 있다.

[1357] 도 53 은 본 개시물에 설명된 기술들의 다양한 양상들을 수행할 수도 있는 오디오 인코딩 디바이스 (570) 의 다른 예를 예시하는 블록 다이어그램이다. 도 53 의 예에서, 오디오 인코딩 디바이스 (570) 는 오디오 인코딩 디바이스들 (510A-510J) 중 하나 이상의 오디오 인코딩 디바이스와 유사할 수도 있다 (여기서 오더 감소 유닛 (528A) 은 사운드 성분 추출 유닛 (20) 에 포함되는 것으로 가정되나 설명의 용이함을 위해 도시되지는 않는다). 그러나, 오디오 인코딩 디바이스 (570) 는 일부 예들에서 분해 유닛 (518) 을 포함할 수도 있는 보다 일반적인 변환 유닛 (572) 을 포함할 수도 있다.

[1358] 도 54 는, 더 상세하게, 도 53 의 예에 도시된 오디오 인코딩 디바이스 (570) 의 일 예시적인 구현을 예시하는 블록 다이어그램이다. 도 54 의 예에 도시된 바와 같이, 오디오 인코딩 디바이스 (570) 의 변환 유닛 (572) 은 회전 유닛 (654) 을 포함한다. 오디오 인코딩 디바이스 (570) 의 음장 성분 추출 유닛 (520) 은 공간 분석 유닛 (650), 컨텐츠-특성들 분석 유닛 (652), 코히어런트 성분들 추출 유닛 (656), 및 확산 성분들 추출 유닛 (658) 을 포함한다. 오디오 인코딩 디바이스 (570) 의 오디오 인코딩 유닛 (514) 은 AAC 코딩 엔진 (660), 및 AAC 코딩 엔진 (162) 을 포함한다. 오디오 인코딩 디바이스 (570) 의 비트스트림 생성 유닛 (516) 은 멀티플렉서 (MUX) (164) 를 포함한다.

[1359] SHC 의 형태로 3D 오디오 데이터를 표현하기 위해 요구되는 - bits/second 의 면에서 - 대역폭은 소비자 이용의 면에서는 엄두도 못 낼 정도로 높은 비용이 들게 할 수도 있다. 예를 들어, 48 kHz 의 샘플링 레이트를 이용하고, 32 비트/동일한 해상도를 갖는 경우, 제 4 차 SHC 표현은 36 Mbits/second (25x48000x32 bps) 의 대역폭을 표현한다. 통상적으로 약 100 kbits/second 인, 스테레오 신호들에 대한 최신 오디오 코딩과 비교하는 경우, 이는 큰 수치이다. 도 54 의 예에서 구현된 기술들은 3D 오디오 표현들의 대역폭을 감소시킬 수도 있다.

[1360] 공간 분석 유닛 (650), 컨텐츠-특성들 분석 유닛 (652), 및 회전 유닛 (654) 은 SHC (511) 를 수신할 수도 있다. 본 개시물의 다른 곳에서 설명된 바와 같이, SHC (511) 는 음장을 표현할 수도 있다. 도 54 의 예에서, 공간 분석 유닛 (650), 컨텐츠-특성들 분석 유닛 (652), 및 회전 유닛 (654) 은 음장의 4 차원 (n=4) 표현을 위해 25 개의 SHC 를 수신할 수도 있다.

[1361] 공간 분석 유닛 (650) 은 음장의 구별되는 성분들은 식별하고 음장의 성분들을 확산시키기 위해 SHC (511) 의 의해 표현되는 음장을 분석할 수도 있다. 음장의 구별되는 성분들은 식별가능한 방향으로부터 나오는 것으로 지각되거나 그렇지 않으면 배경으로부터 구별되거나 음장의 확산 성분들인 사운드들이다. 예를 들어, 개개의 음악 기기에 의해 생성된 사운드는 식별가능한 방향으로부터 나오는 것으로 지각될 수도 있다. 그에 반해서, 사운드의 확산 또는 배경 성분들은 식별가능한 방향으로부터 나오는 것으로 지각되지 않는다. 예를 들어, 숲을 통과하는 바람의 사운드는 음장의 확산 성분일 수도 있다.

[1362] 공간 분석 유닛 (650) 은 (이러한 음장을 레코딩한 것으로 여겨지는 마이크로폰에 대해) 수직 및/또는 수평 축으로 가장 높은 에너지를 갖는 구별되는 성분들의 음장을 정렬하도록 음장을 회전시키는 최적의 각을 식별하려는 시도로 하나 이상의 구별되는 성분들을 식별할 수도 있다. 공간 분석 유닛 (650) 은 이러한 구별되는 성분들이 도 1 및 도 2 의 예들에서 도시된 근본적인 구면 기저 함수들과 보다 잘 맞춰지도록 음장이 회전될 수도 있도록 이러한 최적의 각을 식별할 수도 있다.

[1363] 일부 예들에서, 공간 분석 유닛 (650) 은 확산 사운드들 (이는 낮은 레벨의 방향 또는 하위 오더의 SHC 를 갖는

사운드들을 지칭할 수도 있으며, 1 이하의 오더를 갖는 SHC (511) 를 의미한다) 을 포함하는 SHC (511) 에 의해 표현되는 음장의 비율을 식별하기 위해 일 형태의 확산 분석을 수행하도록 구성된 유닛을 표현할 수도 있다.

일 예로서, 공간 분석 유닛 (650) 은 2007 년 6 월에 J. Audio EnG. Soc, Vol. 55, No. 6 에서 공개된 "Spatial Sound Reproduction with Directional Audio Coding" 이라는 제목의 Ville Pulkki 의 논문에서 설명된 것과 유사한 방식으로 확산 분석을 수행할 수도 있다. 일부 경우들에서, 공간 분석 유닛 (650) 은, 확산 비율을 결정하기 위해 확산 분석을 수행하는 경우, 오직 HOA 계수들의 비-제로 서브세트, 예컨대 SHC (511) 의 제로 및 제 1 차의 것들만을 분석할 수도 있다.

[1364] 컨텐츠-특성들 분석 유닛 (652) 은 SHC (511) 에 적어도 부분적으로 기초하여, SHC (511) 가 음장의 자연적 레코딩에 의해 생성되었는지, 일 예로서, PCM 오브젝트와 같은 오디오 오브젝트로부터 인공적으로 (즉, 합성적으로) 생산되었는지 여부를 결정할 수도 있다. 또한, 컨텐츠-특성들 분석 유닛 (652) 은 그 다음에, SHC (511) 가 음장의 실제 레코딩에 의해 생성되었는지 인공적 오디오 오브젝트로부터 생성되었는지 여부에 적어도 부분적으로 기초하여, 비트스트림 (517) 에 포함시킬 전체 채널들의 수를 결정할 수도 있다. 예를 들어, 컨텐츠-특성들 분석 유닛 (652) 은, SHC (511) 가 실제 음장의 레코딩에 의해 생성되었는지 인공적 오디오 오브젝트로부터 생성되었는지 여부에 적어도 부분적으로 기초하여, 비트스트림 (517) 이 16 개의 채널들을 포함할지를 결정할 수도 있다. 채널들의 각각은 모노 채널일 수도 있다. 컨텐츠-특성들 분석 유닛 (652) 은 비트스트림 (517) 의 출력 비트레이트, 예를 들어, 1.2 Mbps 에 기초하여 비트스트림 (517) 에 포함시킬 전체 채널들의 수의 결정을 더 수행할 수도 있다.

[1365] 또한, 컨텐츠-특성들 분석 유닛 (652) 은, SHC (511) 가 실제 음장의 레코딩에 의해 생성되었는지 인공적 오디오 오브젝트로부터 생성되었는지 여부에 적어도 부분적으로 기초하여, 얼마나 많은 채널들을 음장의 코히어런트 또는 다시 말해 구별되는 성분들에 할당할지, 그리고 얼마나 많은 채널들을 음장의 확산 또는 다시 말해 배경 성분들에 할당할지를 결정할 수도 있다. 예를 들어, SHC (511) 가 일 예로서 Eigenmic 을 이용하여 실제 음장의 레코딩으로부터 생성된 경우, 컨텐츠-특성들 분석 유닛 (652) 은 음장의 코히어런트 성분들에 채널들 중 3 개의 채널들을 할당할 수도 있고, 음장의 확산 성분들에 남은 채널들을 할당할 수도 있다. 이러한 예에서, SHC (511) 가 인공적 오디오 오브젝트로부터 생성된 경우, 컨텐츠-특성들 분석 유닛 (652) 은 음장의 코히어런트 성분들에 채널들 중 5 개의 채널들을 할당할 수도 있고, 음장의 확산 성분들에 남은 채널들을 할당할 수도 있다. 이러한 방식으로, 컨텐츠 분석 블록 (즉, 컨텐츠-특성들 분석 유닛 (652)) 은 음장의 유형 (예를 들어, 확산/방향성 등) 을 결정하고, 차례로 추출할 코히어런트/확산 성분들의 수를 결정할 수도 있다.

[1366] 타겟 비트 레이트는 개개의 AAC 코딩 엔진들 (예를 들어, AAC 코딩 엔진들 (660, 662)) 의 성분들의 수 및 비트 레이트에 영향을 줄 수도 있다. 다시 말해, 컨텐츠-특성들 분석 유닛 (652) 은 비트스트림 (517) 의 출력 비트레이트, 예를 들어, 1.2 Mbps 에 기초하여, 얼마나 많은 채널들을 코히어런트 성분들에 할당할지 및 얼마나 많은 채널들을 확산 성분들에 할당할지의 결정을 더 수행할 수도 있다.

[1367] 일부 예들에서, 음장의 코히어런트 성분들에 할당된 채널들은 음장의 확산 성분들에 할당된 채널들보다 큰 비트 레이트들을 가질 수도 있다. 예를 들어, 비트스트림 (517) 의 최대 비트레이트는 1.2 Mb/sec 일 수도 있다. 이러한 예에서, 코히어런트 성분들에 할당된 4 개의 채널들 및 확산 성분들에 할당된 16 개의 채널들이 있을 수도 있다. 또한, 이러한 예에서, 코히어런트 성분들에 할당된 채널들의 각각은 64 kb/sec 의 최대 비트레이트를 가질 수도 있다. 이러한 예에서, 확산 성분들에 할당된 채널들의 각각은 48 kb/sec 의 최대 비트레이트를 가질 수도 있다.

[1368] 위에서 나타내어진 바와 같이, 컨텐츠-특성들 분석 유닛 (652) 은 SHC (511) 가 실제 음장의 레코딩으로부터 생성되었는지 인공적 오디오 오브젝트로부터 생성되었는지 여부를 결정할 수도 있다. 컨텐츠-특성들 분석 유닛 (652) 은 다양한 방식들로 이러한 결정을 할 수도 있다. 예를 들어, 오디오 인코딩 디바이스 (570) 는 4 오더 SHC 를 이용할 수도 있다. 이러한 예에서, 컨텐츠-특성들 분석 유닛 (652) 은 24 개의 채널들을 코딩하고 25 번째 채널 (이는 벡터로 표현될 수도 있다) 을 예측할 수도 있다. 컨텐츠-특성들 분석 유닛 (652) 은 25 번째 벡터를 결정하기 위해 24 개의 채널들 중 적어도 일부의 채널들에 스칼라들을 적용하고 결과적인 값들을 가산할 수도 있다. 또한, 이러한 예에서, 컨텐츠-특성들 분석 유닛 (652) 은 예측된 25 번째 채널의 정확도를 결정할 수도 있다. 이러한 예에서, 예측된 25 번째 채널의 정확도가 상대적으로 높은 경우 (예를 들어, 정확도가 특정 임계치를 초과한다), SHC (511) 는 합성 오디오 오브젝트로부터 생성될 가능성이 있다. 그에 반해서, 예측된 25 번째 채널들의 정확도가 상대적으로 낮은 경우 (예를 들어, 정확도가 특정 임계치 아래인 경우), SHC (511) 는 레코딩된 음장을 표현할 가능성이 더 높다. 예를 들어, 이러한 예에서, 25 번째 채널의 신호-대-잡음 비 (SNR) 가 100 데시벨 (dbs) 을 넘는 경우, SHC (511) 는 합성 오디오 오브젝트로부터

생성된 음장을 표현할 가능성이 더 높다. 그에 반해서, 아이겐 마이크로폰을 이용하여 레코딩된 음장의 SNR 은 5 dbs 내지 20 dbs 일 수도 있다. 따라서, 실제 직접적인 레코딩으로부터 생성된 SHC (511) 에 의해 표현된 음장과 합성 오디오 오브젝트로부터 생성된 SHC (511) 에 의해 표현된 음장 사이에 구별되는 SNR 비율들에서의 구분이 있을 수도 있다.

[1369] 또한, 컨텐츠-특성들 분석 유닛 (652) 은, SHC (511) 가 실제 음장의 레코딩으로부터 생성되었는지 또는 인공적 오디오 오브젝트로부터 생성되었는지 여부에 적어도 부분적으로 기초하여, V 벡터를 양자화하기 위한 코드북들을 선택할 수도 있다. 다시 말해, 컨텐츠-특성들 분석 유닛 (652) 은, HOA 계수들에 의해 표현되는 음장이 레코딩되었는지 합성인지 여부에 따라, V 벡터를 양자화할 시에 이용하기 위해 상이한 코드북들을 선택할 수도 있다.

[1370] 일부 예들에서, 컨텐츠-특성들 분석 유닛 (652) 은, 정기적으로, SHC (511) 가 실제 음장의 레코딩으로부터 생성되었는지 또는 인공적 오디오 오브젝트로부터 생성되었는지 여부를 결정할 수도 있다. 일부 그러한 예들에서, 정기적 기준은 매 프레임일 수도 있다. 다른 예들에서, 컨텐츠-특성들 분석 유닛 (652) 은 이러한 결정을 한번만 수행할 수도 있다. 또한, 컨텐츠-특성들 분석 유닛 (652) 은, 정기적으로, 전체 채널들의 수 및 코히어런트 성분 채널들과 확산 성분 채널들의 할당을 결정할 수도 있다. 일부 그러한 예들에서, 정기적 기준은 매 프레임일 수도 있다. 다른 예들에서, 컨텐츠-특성들 분석 유닛 (652) 은 이러한 결정을 한번만 수행할 수도 있다. 일부 예들에서, 컨텐츠-특성들 분석 유닛 (652) 은, 정기적으로, V 벡터를 양자화할 시에 이용하기 위한 코드북들을 선택할 수도 있다. 일부 그러한 예들에서, 정기적인 기준은 매 프레임일 수도 있다. 다른 예들에서, 컨텐츠-특성들 분석 유닛 (652) 은 이러한 결정을 한번만 수행할 수도 있다.

[1371] 회전 유닛 (654) 은 HOA 계수들의 회전 동작을 수행할 수도 있다. (예를 들어, 도 55 및 도 55b 에 대하여) 본 개시물의 다른 곳에서 논의된 바와 같이, 회전 동작을 수행하는 것은 SHC (511) 를 나타내는데 요구되는 비트들의 수를 감소시킬 수도 있다. 일부 예들에서, 회전 유닛 (652) 에 의해 수행되는 회전 분석은 특이값 분해 (SVD) 분석의 일 예이다. 주요 성분 분석 ("PCA"), 독립적인 성분 분석 ("ICA"), 및 카루넨-루베 변환 ("KLT") 이 적용가능할 수도 있는 관련된 기술들이다.

[1372] 도 54 의 예에서, 추출 코히어런트 성분들 유닛 (656) 은 회전 유닛 (654) 으로부터 회전된 SHC (511) 를 수신한다. 또한, 추출 코히어런트 성분들 유닛 (656) 은, 회전된 SHC (511) 로부터, 음장의 코히어런트 성분들과 연관된 SHC (511) 를 추출한다.

[1373] 또한, 추출 코히어런트 성분들 유닛 (656) 은 하나 이상의 코히어런트 성분 채널들을 생성한다. 코히어런트 성분 채널들의 각각은 음장의 코히어런트 계수들과 연관된 회전된 SHC (511) 의 상이한 서브세트를 포함할 수도 있다. 도 54 의 예에서, 추출 코히어런트 성분들 유닛 (656) 은 1 개에서 16 개의 코히어런트 성분 채널들을 생성할 수도 있다. 추출 코히어런트 성분들 유닛 (656) 에 의해 생성된 코히어런트 성분 채널들의 개수는 컨텐츠-특성들 분석 유닛 (652) 에 의해 음장의 코히어런트 성분들에 대해 할당된 채널들의 개수에 의해 결정될 수도 있다. 추출 코히어런트 성분들 유닛 (656) 에 의해 생성된 코히어런트 성분 채널들의 비트레이트들은 컨텐츠-특성들 분석 유닛 (652) 에 의해 결정될 수도 있다.

[1374] 유사하게, 도 54 의 예에서, 추출 확산 성분들 유닛 (658) 은 회전 유닛 (654) 으로부터 회전된 SHC (511) 를 수신한다. 또한, 추출 확산 성분들 유닛 (658) 은, 회전된 SHC (511) 로부터, 음장의 확산 성분들과 연관된 회전된 SHC (511) 를 추출한다.

[1375] 또한, 추출 확산 성분들 유닛 (658) 은 하나 이상의 확산 성분 채널들을 생성한다. 확산 성분 채널들의 각각은 음장의 확산 계수들과 연관된 회전된 SHC (511) 의 상이한 서브세트를 포함할 수도 있다. 도 54 의 예에서, 추출 확산 성분들 유닛 (658) 은 1 개에서 9 개의 확산 성분 채널들을 생성할 수도 있다. 추출 확산 성분들 유닛 (658) 에 의해 생성된 확산 성분 채널들의 개수는 컨텐츠-특성들 분석 유닛 (652) 에 의해 음장의 확산 성분들에 대해 할당된 채널들의 개수에 의해 결정될 수도 있다. 추출 확산 성분들 유닛 (658) 에 의해 생성된 확산 성분 채널들의 비트레이트들은 컨텐츠-특성들 분석 유닛 (652) 에 의해 결정될 수도 있다.

[1376] 도 54 의 예에서, AAC 코딩 유닛 (660) 은 추출 코히어런트 성분들 유닛 (656) 에 의해 생성된 코히어런트 성분 채널들을 인코딩하기 위해 AAC 코덱을 이용할 수도 있다. 유사하게, AAC 코딩 유닛 (662) 은 추출 확산 성분들 유닛 (658) 에 의해 생성된 확산 성분 채널들을 인코딩하기 위해 AAC 코덱을 이용할 수도 있다. 멀티플렉서 (664) ("MUX 664") 는 인코딩된 코히어런트 성분 채널들 및 인코딩된 확산 성분 채널들을, 사이드 데이터 (예를 들어, 공간 분석 유닛 (650) 에 의해 결정된 최적 각) 와 함께 멀티플렉싱하여, 비트스트림 (517) 을

생성할 수도 있다.

[1377] 이러한 방식으로, 기술들은 오디오 인코딩 디바이스 (570) 가 음장을 나타내는 구면 조화 계수들이 합성 오디오 오브젝트로부터 생성되는지 여부를 결정하는 것을 가능하게 할 수도 있다.

[1378] 일부 예들에서, 오디오 인코딩 디바이스 (570) 는, 구면 조화 계수들이 합성 오디오 오브젝트로부터 생성되는지 여부에 기초하여, 음장의 구별되는 성분들을 나타내는 구면 조화 계수들의 서브세트를 결정할 수도 있다. 이러한 예들 및 다른 예들에서, 오디오 인코딩 디바이스 (570) 는 구면 조화 계수들의 서브세트를 포함하도록 비트스트림을 생성할 수도 있다. 오디오 인코딩 디바이스 (570) 는, 일부 경우들에서, 구면 조화 계수들의 서브세트를 오디오 인코딩하고, 구면 조화 계수들의 오디오 인코딩된 서브세트를 포함하도록 비트스트림을 생성할 수도 있다.

[1379] 일부 예들에서, 오디오 인코딩 디바이스 (570) 는, 구면 조화 계수들이 합성 오디오 오브젝트로부터 생성되는지 여부에 기초하여, 음장의 배경 성분들을 나타내는 구면 조화 계수들의 서브세트를 결정할 수도 있다. 이러한 예들 및 다른 예들에서, 오디오 인코딩 디바이스 (570) 는 구면 조화 계수들의 서브세트를 포함하도록 비트스트림을 생성할 수도 있다. 이러한 예들 및 다른 예들에서, 오디오 인코딩 디바이스 (570) 는 구면 조화 계수들의 서브세트를 오디오 인코딩하고, 구면 조화 계수들의 오디오 인코딩된 서브세트를 포함하도록 비트스트림을 생성할 수도 있다.

[1380] 일부 예들에서, 오디오 인코딩 디바이스 (570) 는 구면 조화 계수들에 의해 표현되는 음장을 회전시킬 각을 식별하기 위해 구면 조화 계수들에 대해 공간 분석을 수행하고, 회전된 구면 조화 계수들을 생성하기 위해 식별된 각으로 음장을 회전시키는 회전 동작을 수행할 수도 있다.

[1381] 일부 예들에서, 오디오 인코딩 디바이스 (570) 는, 구면 조화 계수들이 합성 오디오 오브젝트로부터 생성되었는지 여부에 기초하여, 음장의 구별되는 성분들을 나타내는 구면 조화 계수들의 제 1 서브세트를 결정하고, 구면 조화 계수들이 합성 오디오 오브젝트로부터 생성되는지 여부에 기초하여, 음장의 배경 성분들을 나타내는 구면 조화 계수들의 제 2 서브세트를 결정할 수도 있다. 이러한 예들 및 다른 예들에서, 오디오 인코딩 디바이스 (570) 는 구면 조화 계수들의 제 2 서브세트를 오디오 인코딩하는데 이용된 것더 높은 타겟 비트레이트를 갖는 구면 조화 계수들의 제 1 서브세트를 오디오 인코딩할 수도 있다.

[1382] 이러한 방식으로, 기술들의 다양한 양상들은 다음의 조항들에 따라 오디오 인코딩 디바이스 (570) 가 SHC (511) 가 합성 오디오 오브젝트로부터 생성되었는지 여부를 결정하는 것을 가능하게 할 수도 있다.

[1383] 조항 132512-1. 오디오 인코딩 디바이스 (570) 와 같은 디바이스, 여기서 하나 이상의 프로세서들은 음장을 나타내는 구면 조화 계수들이 합성 오디오 오브젝트로부터 생성되었는지 여부를 결정하도록 더 구성된다.

[1384] 조항 132512-2. 조항 132512-1 의 디바이스, 여기서 하나 이상의 프로세서들은, 음장을 나타내는 구면 조화 계수들이 합성 오디오 오브젝트로부터 생성되었는지 여부를 결정할 경우, 음장을 나타내는 구면 조화 계수들의 적어도 일부분을 저장하는 프레임화된 구면 조화 계수 행렬로부터 제 1 벡터를 제외하여 감소된 프레임화된 구면 조화 계수 행렬을 획득하도록 더 구성된다.

[1385] 조항 132512-3. 조항 132512-1 의 디바이스, 하나 이상의 프로세서들은, 음장을 나타내는 구면 조화 계수들이 합성 오디오 오브젝트로부터 생성되었는지 여부를 결정할 경우, 음장을 나타내는 구면 조화 계수들의 적어도 일부분을 저장하는 프레임화된 구면 조화 계수 행렬로부터 제 1 벡터를 제외하여 감소된 프레임화된 구면 조화 계수 행렬을 획득하고, 감소된 프레임화된 구면 조화 계수 행렬의 남은 벡터들에 기초하여 감소된 프레임화된 구면 조화 계수 행렬의 벡터를 예측하도록 더 구성된다.

[1386] 조항 132512-4. 조항 132512-1 의 디바이스, 여기서 하나 이상의 프로세서들은, 음장을 나타내는 구면 조화 계수들이 합성 오디오 오브젝트로부터 생성되었는지 여부를 결정할 경우, 음장을 나타내는 구면 조화 계수들의 적어도 일부분을 저장하는 프레임화된 구면 조화 계수 행렬로부터 제 1 벡터를 제외하여 감소된 프레임화된 구면 조화 계수 행렬을 획득하고, 적어도 부분적으로, 감소된 프레임화된 구면 조화 계수 행렬의 남은 벡터들의 합에 기초하여, 감소된 프레임화된 구면 조화 계수 행렬의 벡터를 예측하도록 더 구성된다.

[1387] 조항 132512-5. 조항 132512-1 의 디바이스, 여기서 하나 이상의 프로세서들은, 음장을 나타내는 구면 조화 계수들이 합성 오디오 오브젝트로부터 생성되었는지 여부를 결정할 경우, 적어도 부분적으로, 프레임화된 구면 조화 계수 행렬의 남은 벡터들의 합에 기초하여, 구면 조화 계수들의 적어도 일부분을 저장하는 프레임화된 구면 조화 계수 행렬의 벡터를 예측하도록 더 구성된다.

[1388] 조항 132512-6. 조항 132512-1 의 디바이스, 여기서 하나 이상의 프로세서들은, 음장을 나타내는 구면 조화 계수들이 합성 오디오 오브젝트로부터 생성되었는지 여부를 결정할 경우, 적어도 부분적으로, 프레임화된 구면 조화 계수 행렬의 남은 벡터들의 합에 기초하여, 구면 조화 계수들의 적어도 일부분을 저장하는 프레임화된 구면 조화 계수 행렬의 벡터를 예측하고, 예측된 벡터에 기초하여 에러를 계산하도록 더 구성된다.

[1389] 조항 132512-7. 조항 132512-1 의 디바이스, 여기서 하나 이상의 프로세서들은, 음장을 나타내는 구면 조화 계수들이 합성 오디오 오브젝트로부터 생성되었는지 여부를 결정할 경우, 적어도 부분적으로, 프레임화된 구면 조화 계수 행렬의 남은 벡터들의 합에 기초하여, 구면 조화 계수들의 적어도 일부분을 저장하는 프레임화된 구면 조화 계수 행렬의 벡터를 예측하고, 예측된 벡터 및 프레임화된 구면 조화 계수 행렬의 대응하는 벡터에 기초하여 에러를 계산하도록 더 구성된다.

[1390] 조항 132512-8. 조항 132512-1 의 디바이스, 여기서 하나 이상의 프로세서들은, 음장을 나타내는 구면 조화 계수들이 합성 오디오 오브젝트로부터 생성되었는지 여부를 결정할 경우, 적어도 부분적으로, 프레임화된 구면 조화 계수 행렬의 남은 벡터들의 합에 기초하여, 구면 조화 계수들의 적어도 일부분을 저장하는 프레임화된 구면 조화 계수 행렬의 벡터를 예측하고, 프레임화된 구면 조화 계수 행렬의 예측된 벡터와 대응하는 벡터의 차이의 절대 값의 합으로서 에러를 계산하도록 더 구성된다.

[1391] 조항 132512-9. 조항 132512-1 의 디바이스, 여기서 하나 이상의 프로세서들은, 음장을 나타내는 구면 조화 계수들이 합성 오디오 오브젝트로부터 생성되었는지 여부를 결정하는 경우, 적어도 부분적으로, 프레임화된 구면 조화 계수 행렬의 남은 벡터들의 합에 기초하여, 구면 조화 계수들의 적어도 일부분을 저장하는 프레임화된 구면 조화 계수 행렬의 벡터를 예측하고, 프레임화된 구면 조화 계수 행렬의 예측된 벡터 및 대응하는 벡터에 기초하여 에러를 계산하고, 프레임화된 구면 조화 계수 행렬의 대응하는 벡터의 에너지 및 에러에 기초하여 비율을 계산하고, 비율을 임계치와 비교하여 음장을 나타내는 구면 조화 계수들이 합성 오디오 오브젝트로부터 생성되었는지 여부를 결정하도록 더 구성된다.

[1392] 조항 132512-10. 청구항 4 내지 청구항 9 중 임의의 청구항의 디바이스, 여기서 하나 이상의 프로세서들은, 벡터를 예측할 경우, 구면 조화 계수들의 적어도 일부분을 저장하는 프레임화된 구면 조화 계수 행렬의 제 1 비-제로 벡터를 예측하도록 더 구성된다.

[1393] 조항 132512-11. 청구항 1 내지 청구항 10 중 임의의 청구항의 디바이스, 여기서 하나 이상의 프로세서들은 구면 조화 계수들의 압축된 버전을 저장하는 비트스트림에 구면 조화 계수들이 합성 오디오 오브젝트로부터 생성되었는지 여부의 표시를 명시하도록 더 구성된다.

[1394] 조항 132512-12. 조항 132512-11 의 디바이스, 여기서 표시는 단일 비트이다.

[1395] 조항 132512-13. 조항 132512-1 의 디바이스, 여기서 하나 이상의 프로세서들은, 구면 조화 계수들이 합성 오디오 오브젝트로부터 생성되었는지 여부에 기초하여, 음장의 구별되는 성분들을 나타내는 구면 조화 계수들의 서브세트를 결정하도록 더 구성된다.

[1396] 조항 132512-14. 조항 132512-13 의 디바이스, 여기서 하나 이상의 프로세서들은 구면 조화 계수들의 서브세트를 포함하도록 비트스트림을 생성하도록 더 구성된다.

[1397] 조항 132512-15. 조항 132512-13 의 디바이스, 여기서 하나 이상의 프로세서들은 구면 조화 계수들의 서브세트를 오디오 인코딩하고, 구면 조화 계수들의 오디오 인코딩된 서브세트를 포함하도록 비트스트림을 생성하도록 더 구성된다.

[1398] 조항 132512-16. 조항 132512-1 의 디바이스, 여기서 하나 이상의 프로세서들은, 구면 조화 계수들이 합성 오디오 오브젝트로부터 생성되었는지 여부에 기초하여, 음장의 배경 성분들을 나타내는 구면 조화 계수들의 서브세트를 결정하도록 더 구성된다.

[1399] 조항 132512-17. 조항 132512-16 의 디바이스, 여기서 하나 이상의 프로세서들은 구면 조화 계수들의 서브세트를 포함하도록 비트스트림을 생성하도록 더 구성된다.

[1400] 조항 132512-18. 조항 132512-15 의 디바이스, 여기서 하나 이상의 프로세서들은 구면 조화 계수들의 서브세트를 오디오 인코딩하고, 구면 조화 계수들의 오디오 인코딩된 서브세트를 포함하도록 비트스트림을 생성하도록 더 구성된다.

[1401] 조항 132512-19. 조항 132512-1 의 디바이스, 여기서 하나 이상의 프로세서들은 구면 조화 계수들에 의해

표현되는 음장을 회전시키기 위한 각을 식별하기 위해 구면 조화 계수들에 대해 공간 분석을 수행하고, 식별된 각으로 음장을 회전시키는 회전 동작을 수행하여 회전된 구면 조화 계수들을 생성하도록 더 구성된다.

[1402] 조항 132512-20. 조항 132512-1 의 디바이스, 여기서 하나 이상의 프로세서들은, 구면 조화 계수들이 합성 오디오 오브젝트로부터 생성되었는지 여부에 기초하여, 음장의 구별되는 성분들을 나타내는 구면 조화 계수들의 제 1 서브세트를 결정하고, 구면 조화 계수들이 합성 오디오 오브젝트로부터 생성되었는지 여부에 기초하여, 음장의 배경 성분들을 나타내는 구면 조화 계수들의 제 2 서브세트를 결정하도록 더 구성된다.

[1403] 조항 132512-21. 조항 132512-20 의 디바이스, 여기서 하나 이상의 프로세서들은 구면 조화 계수들의 제 2 서브세트를 오디오 인코딩하는데 이용되는 것더 높은 타겟 비트레이트를 갖는 구면 조화 계수들의 제 1 서브세트를 오디오 인코딩하도록 더 구성된다.

[1404] 조항 132512-22. 조항 132512-1 의 디바이스, 여기서 하나 이상의 프로세서들은 구면 조화 계수들에 대해 특이값 분해를 수행하여, 복수의 구면 조화 계수들의 좌-특이 벡터들을 나타내는 U 행렬, 복수의 구면 조화 계수들의 특이값들을 나타내는 S 행렬, 및 복수의 구면 조화 계수들의 우-특이 벡터들을 나타내는 V 행렬을 생성하도록 더 구성된다.

[1405] 조항 132512-23. 조항 132512-22 의 디바이스, 여기서 하나 이상의 프로세서들은, 구면 조화 계수들이 합성 오디오 오브젝트로부터 생성되는지 여부에 기초하여, 음장의 구별되는 성분들을 나타내는 U 행렬, S 행렬, 및 V 행렬 중 하나 이상의 부분들을 결정하도록 더 구성된다.

[1406] 조항 132512-24. 조항 132512-22 의 디바이스, 여기서 하나 이상의 프로세서들은, 구면 조화 계수들이 합성 오디오 오브젝트로부터 생성되는지 여부에 기초하여, 음장의 배경 성분들을 나타내는 U 행렬, S 행렬, 및 V 행렬 중 하나 이상의 부분들을 결정하도록 더 구성된다.

[1407] 조항 132512-1C. 하나 이상의 프로세서들을 포함하는 오디오 인코딩 디바이스 (570) 와 같은 디바이스, 하나 이상의 프로세서들은, 적어도, 구면 조화 계수들의 벡터의 에너지 및 구면 조화 계수들의 벡터의 예측된 버전 및 구면 조화 계수들의 벡터에 기초하여 도출된 에러의 함수로서 계산된 비율에 기초하여 음장을 나타내는 구면 조화 계수들이 합성 오디오 오브젝트로부터 생성되었는지 여부를 결정하도록 구성된다.

[1408] 전술된 다양한 경우들의 각각에서, 오디오 인코딩 디바이스 (570) 는 방법을 수행하거나 그렇지 않으면, 오디오 인코딩 디바이스 (570) 가 수행하도록 구성되는 방법의 각각을 단계를 수행하는 수단을 포함할 수도 있는 것으로 이해되어야 한다. 일부 경우들에서, 이러한 수단은 하나 이상의 프로세서들을 포함할 수도 있다. 일부 경우들에서, 하나 이상의 프로세서들은 비-일시적 컴퓨터 판독가능 저장 매체에 저장된 명령들에 의해 구성된 특수 용도 프로세서를 표현할 수도 있다. 다시 말해, 인코딩 예들의 세트의 각각에서의 기술들의 다양한 양상들은, 실행되는 경우, 하나 이상의 프로세서들로 하여금 오디오 인코딩 디바이스 (570) 가 수행하도록 구성된 방법을 수행하게 하는 명령들을 저장한 비-일시적 컴퓨터 판독가능 저장 매체를 제공할 수도 있다.

[1409] 도 55 및 도 55b 는 음장 (640) 을 회전시키기 위해 본 개시물에 설명된 기술들의 다양한 양상들을 수행하는 일 예를 예시하는 다이어그램들이다. 도 55 는 본 개시물에 설명된 기술들의 다양한 양상들에 따라 회전하기 전의 음장 (640) 을 예시하는 다이어그램이다. 도 55 의 예에서, 음장 (640) 은 위치 (642A 및 642B) 로 표기되는, 높은 압력의 2 개의 위치들을 포함한다. 이러한 위치 (642A 및 642B) ("위치들 (642)") 는 비-제로 슬로프 (이는 수평 라인들인 제로의 슬로프를 갖는 것과는 다른, 수평이 아닌 라인을 지칭하는 방식이다) 를 갖는 라인 (644) 을 따라 있다. 위치들 (642) 이 x 및 y 좌표들과 함께 z 좌표를 갖는다고 하면, 이러한 음장 (640) 을 정확하게 표현하기 위해 더 높은 차원의 구면 기저 함수들이 요구될 수도 있다 (이러한 더 높은 차원의 구면 기저 함수들이 음장의 상부 및 하부 또는 비-수평 부분들을 서술하기 때문이다). SHC 들 (511) 에 대해 직접적으로 음장 (640) 을 감소시키기 보다는, 오디오 인코딩 디바이스 (570) 는 위치들 (642) 을 연결하는 라인 (644) 이 수평일 때까지 음장 (640) 을 회전시킬 수도 있다.

[1410] 도 55b 는 위치들 (642) 을 연결하는 라인 (644) 이 수평일 때까지 회전된 후의 음장 (640) 을 예시하는 다이어그램이다. 이러한 방식으로 음장 (640) 을 회전시킨 결과, 회전된 음장 (640) 이 z 좌표들을 갖는 압력 (또는 에너지) 의 임의의 위치들을 더 이상 갖지 않는다는 것을 고려하여 SHC (511) 중 상위 오더의 SHC 가 제로들로 명시되도록 SHC (511) 가 도출될 수도 있다. 이러한 방식으로, 오디오 인코딩 디바이스 (570) 는 비-제로 값들을 갖는 SHC (511) 의 개수를 감소시키도록 음장 (640) 을 회전시키거나, 변환시키거나, 좀더 일반적으로 조정할 수도 있다. 본 기술들의 다양한 다른 양상들과 연계하여, 오디오 인코딩 디바이스 (570) 는, 그러면, SHC (511) 중 이러한 상위 오더 SHC 들이 제로 값들을 갖는다는 것을 식별하는 32-비트 부호 수를 시그널

링하기 보다는, SHC (511) 중 이러한 상위 오더 SHC 들이 시그널링되지 않다는 것을 비트스트림 (517) 의 필드로 시그널링할 수도 있다. 오디오 인코딩 디바이스 (570) 는 또한, 종종 전술된 방식으로 방위각 및 높이를 표현함으로써, 음장 (640) 이 어떻게 회전되는지를 나타내는 회전 정보를 비트스트림 (517) 에 명시할 수도 있다. 오디오 인코딩 디바이스와 같은 추출 디바이스가, 그 다음에, SHC (511) 에 기초하여 음장 (640) 을 재생성하는 경우, SHC (511) 중 시그널링되지 않은 SHC 가 제로 값을 갖는다고 암시하고, 음장 (640) 이 도 55 의 예에서 도시된 음장 (640) 을 닮도록 음장 (640) 을 회전시키는 회전을 수행할 수도 있다. 이러한 방식으로, 오디오 인코딩 디바이스 (570) 는 본 개시물에 설명된 기술들에 따라 비트스트림 (517) 에 명시될 요구되는 SHC (511) 의 개수를 감소시킬 수도 있다.

[1411] 음장의 최적의 회전을 결정하기 위해 '공간 압축' 알고리즘이 이용될 수도 있다. 일 실시형태에서, 오디오 인코딩 디바이스 (570) 는 가능한 모든 방위각 및 높이의 조합들 (즉, 위의 예에서 1024x512 조합들) 을 통해 반복하는 알고리즘을 수행하여, 각각의 조합에 대한 음장을 회전시키고 임계 값더 높은 SHC (511) 의 수를 산출할 수도 있다. 임계 값보다 위의 최소 개의 SHC (511) 를 생성하는 방위각/높이 후보 조합은 "최적 회전" 으로 지칭될 수도 있는 것으로 고려될 수도 있다. 이러한 회전된 형태에서, 음장은 음장을 표현하기 위해 최소 개수의 SHC (511) 를 요구할 수도 있고 그러면 컴팩트화된 (compacted) 것으로 여겨질 수도 있다. 일부 경우들에서, 조정은 이러한 최적 회전을 포함할 수도 있고, 전술된 조정 정보는 (방위각 및 고도각들의 면에서) ("최적 회전" 이라고 칭해질 수도 있는) 이러한 회전 정보를 포함할 수도 있다.

[1412] 일부 경우들에서, 오직 방위각 및 고도각만을 명시하는 대신에, 오디오 인코딩 디바이스 (570) 는 일 예로서 오일러 (Euler) 각들의 형태로 추가적인 각들을 명시할 수도 있다. 오일러 각들은 z-축, 이전 x-축, 및 이전 z-축에 대한 회전의 각을 명시한다. 방위각 및 고도각의 조합들에 대해 본 개시물에서 설명되었으나, 본 개시물의 기술들은 오직 방위각 및 고도각만을 명시하는 것으로 제한되어서는 안되고, 위에서 언급된 3 개의 오일러 각들을 포함하여 임의의 개수의 각들을 명시하는 것을 포함할 수도 있다. 이러한 면에서, 오디오 인코딩 디바이스 (570) 는 음장을 서술할 시에 관련있는 정보를 제공하고 비트스트림에 회전 정보로서 오일러 각들을 명시하는 복수의 계층적 엘리먼트들의 개수를 감소시키도록 음장을 회전시킬 수도 있다. 오일러 각들은, 위에서 언급된 바와 같이, 음장이 어떻게 회전되었는지를 서술할 수도 있다. 오일러 각들을 이용하는 경우, 비트스트림 추출 디바이스는, 오일러 각들을 포함하는 회전 정보를 결정하기 위해 비트스트림을 파싱할 수도 있고, 음장을 서술하는 것에 관련있는 정보를 제공하는 복수의 계층적 엘리먼트들의 기초하여 음장을 재생성하는 경우, 오일러 각들에 기초하여 음장을 회전시킬 수도 있다.

[1413] 또한, 일부 경우들에서, 비트스트림 (517) 에 이러한 각들을 명시적으로 명시하기 보다는, 오디오 인코딩 디바이스 (570) 는 회전을 명시하는 각들의 미리-정의된 조합들과 연관된 ("회전 인덱스" 라고 지칭될 수도 있는) 인덱스를 명시할 수도 있다. 다시 말해, 회전 정보는, 일부 경우들에서, 회전 인덱스를 포함할 수도 있다. 이러한 경우들에서, 회전 인덱스의 주어진 값, 예컨대 제로의 값은 회전이 수행되지 않았음을 나타낼 수도 있다. 이러한 회전 인덱스는 회전 테이블과 관련하여 이용될 수 있다. 즉, 오디오 인코딩 디바이스 (570) 는 방위각과 고도각의 조합들의 각각에 대한 엔트리를 포함하는 회전 테이블을 포함할 수도 있다.

[1414] 대안으로, 회전 테이블은 방위각과 고도각의 각각의 조합을 나타내는 각각의 행렬 변환들에 대한 엔트리를 포함할 수도 있다. 즉, 오디오 인코딩 디바이스 (570) 는 방위각 및 고도각의 조합들의 각각에 의해 음장을 회전시키기 위한 각각의 행렬 변환에 대한 엔트리를 갖는 회전 테이블을 저장할 수도 있다. 통상적으로, 오디오 인코딩 디바이스 (570) 는, 다음의 등식에 따라, 회전이 수행되는 경우, SHC (511) 를 수신하고 SHC (511') 를 도출한다.

[1415] $$\begin{bmatrix} SHC \\ 27' \end{bmatrix} = \begin{bmatrix} EncMat_2 \\ (25x32) \end{bmatrix} \begin{bmatrix} InvMat_1 \\ (32x25) \end{bmatrix} \begin{bmatrix} SHC \\ 27 \end{bmatrix}$$

[1416] 위의 등식에서, SHC (511') 는 참조의 제 2 프레임 ($EncMat_2$) 의 면에서 음장을 인코딩하기 위한 인코딩 행렬, 참조의 제 2 프레임 ($InvMat_1$) 의 면에서 SHC (511) 를 다시 음장으로 되돌리기 위한 역 행렬, 및 SHC (511) 의 함수로 계산된다. $EncMat_2$ 은 사이즈가 25x32 인 반면, $InvMat_2$ 은 사이즈가 32x25 이다. SHC (511') 및 SHC (511) 양자 모두는 사이즈가 25 인데, 여기서 SHC (511') 는 핵심적인 오디오 정보를 명시하지 않는 것들의 제거로 인해 더 감소될 수도 있다. $EncMat_2$ 는 각각의 방위각 및 고도각 조합에 대해 달라지는데 반해, $InvMat_1$ 는 각각의 방위각 및 고도각에 대해 고정적으로 있을 수도 있다. 회전 테이블은 각각의 상이

한 $EncMat_2$ 를 $InvMat_1$ 에 곱한 결과를 저장하는 엔트리를 포함할 수도 있다.

[1417] 도 56 은 참조의 제 1 프레임에 따라 캡쳐된 일 예시적인 음장을 예시하는 도면으로, 참조의 제 1 프레임은 참조의 제 2 프레임의 면에서 음장을 표현하기 위해 본 개시물에 설명된 기술들에 따라 회전된다. 도 56 의 예에서, 아이겐 마이크로폰 (646) 을 둘러싸는 음장은 참조의 제 1 프레임을 가정하여 캡쳐되며, 참조의 제 1 프레임은 도 56 의 예에서, X_1, Y_1, 및 Z_1 축들에 의해 표기된다. SHC (511) 는 이러한 참조의 제 1 프레임의 면에서 음장을 서술한다. $InvMat_1$ 는 SHC (511) 를 다시 음장으로 변환하여, 도 56 의 예에서 X_2, Y_2, 및 Z_2 축들에 의해 표기되는 참조의 제 2 프레임에 대해 음장이 회전되는 것을 가능하게 한다. 전술된 $EncMat_2$ 는 음장을 회전시켜 참조의 제 2 프레임의 면에서 이러한 회전된 음장을 서술하는 SHC (511') 를 생성할 수도 있다.

[1418] 임의의 경우에, 위의 등식은 다음과 같이 도출될 수도 있다. 전면이 x-축의 방향인 것으로 여겨지도록, 음장이 소정의 좌표 시스템으로 레코딩된다고 하면, 아이겐 마이크로폰 (또는 다른 마이크로폰 구성들) 의 32 개의 마이크로폰 포지션들은 이러한 참조 좌표 시스템으로부터 정의된다. 음장의 회전은 그러면 이러한 참조 프레임의 회전으로 여겨질 수도 있다. 가정된 참조 프레임에 있어서, SHC (511) 는 다음과 같이 계산될 수도 있다:

[1419]
$$\begin{bmatrix} SHC \\ 27 \end{bmatrix} = \begin{bmatrix} Y_0^0(Pos_1) & Y_0^0(Pos_2) & \cdots & Y_0^0(Pos_{32}) \\ Y_1^{-1}(Pos_1) & \cdots & & Y_1^{-1}(Pos_{32}) \\ \vdots & \ddots & & \\ Y_4^4(Pos_1) & & & Y_4^4(Pos_{32}) \end{bmatrix} \begin{bmatrix} mic_1(t) \\ mic_2(t) \\ \vdots \\ mic_{32}(t) \end{bmatrix}$$

[1420] 위의 등식에서, Y_n^m 은 i 번째 마이크로폰 (여기서, i 는 이 예에서 1-32 이다) 의 포지션 (PoS_i) 에서의 구면 기저 함수들을 표현한다. mic_i 벡터는 시간 (t) 동안에 i 번째 마이크로폰에 대한 마이크로폰 신호를 표기한다. 포지션들 (PoS_i) 은 참조의 제 1 프레임 (즉, 이 예에서 회전 이전의 참조의 프레임) 에서의 마이크로폰의 포지션을 지칭한다.

[1421] 위의 등식은 다르게는 위에서

[1422]
$$[SHC_27] = [E_s(\theta, \varphi)][m_i(t)]$$

[1423] 으로 표기된 수학식들의 면에서 표현될 수도 있다.

[1424] 음장을 회전시키기 위해 (또는 참조의 제 2 프레임에서), 포지션 (Pos_i) 은 참조의 제 2 프레임에서 계산될 것이다. 원래의 마이크로폰 신호들이 존재하는 한, 음장은 임의대로 회전될 수도 있다. 그러나, 원래의 마이크로폰 신호들 ($mic_i(t)$) 은 종종 이용가능하지 않다. 문제는, 그러면, 어떻게 SHC (511) 로부터 마이크로폰 신호들 ($mic_i(t)$) 을 취출하는가일 수도 있다. (32 개의 마이크로폰 아이겐 마이크로폰에서와 같이) T-설계가 이용되는 경우, 이러한 문제에 대한 해법은 다음의 등식을 해결함으로써 달성될 수도 있다:

[1425]
$$\begin{bmatrix} mic_1(t) \\ mic_2(t) \\ \vdots \\ mic_{32}(t) \end{bmatrix} = [InvMat_1] \begin{bmatrix} SHC \\ 27 \end{bmatrix}$$

[1426] 이러한 $InvMat_1$ 는 참조의 제 1 프레임에 대해 명시된 바와 같은 마이크로폰들의 포지션에 따라 계산된 구면 조화 기저 함수들을 명시할 수도 있다. 이러한 등식은 또한, 위에서 언급된 바와 같이, $[m_i(t)] = [E_s(\theta, \varphi)]^{-1}[SHC]$ 으로 표현될 수도 있다.

[1427] 위의 등식에 따라 마이크로폰 신호들 ($mic_i(t)$) 이 취출되면, 음장을 서술하는 마이크로폰 신호들 ($mic_i(t)$) 은 참조의 제 2 프레임에 대응하는 SHC (511') 를 계산하도록 회전되어, 다음의 등식을 초래할 수도 있다:

[1428] $$\begin{bmatrix} SHC \\ 27' \end{bmatrix} = \begin{bmatrix} EncMat_2 \\ (25x32) \end{bmatrix} \begin{bmatrix} InvMat_1 \\ (32x25) \end{bmatrix} \begin{bmatrix} SHC \\ 27 \end{bmatrix}$$

[1429] $EncMat_2$ 는 회전된 위치 (Pos_i') 로부터의 구면 조화 기저 함수들을 명시한다. 이러한 방식으로, $EncMat_2$ 는 방위각과 고도각의 조합을 효과적으로 명시할 수도 있다. 따라서, 회전 테이블이 방위각과 고도각의 각각의 조합에 대한 $\begin{bmatrix} EncMat_2 \\ (25x32) \end{bmatrix} \begin{bmatrix} InvMat_1 \\ (32x25) \end{bmatrix}$ 의 결과를 저장하는 경우, 회전 테이블은 방위각과 고도각의 각각의 조합을 효과적으로 명시한다.

[1430] 위의 등식은 또한:

[1431] $$[SHC\ 27'] = [E_s(\theta_2, \varphi_2)][E_s(\theta_1, \varphi_1)]^{-1}[SHC\ 27]$$

[1432] 과 같이 표현될 수도 있으며, 여기서 θ_2, φ_2 은 θ_1, φ_1 으로 표현되는 제 1 방위각 및 고도각과 상이한 제 2 방위각 및 제 2 고도각을 표현한다. θ_1, φ_1 은 참조의 제 1 프레임에 대응하며, 한편 θ_2, φ_2 는 참조의 제 2 프레임에 대응한다. $InvMat_1$ 는 따라서 $[E_s(\theta_1, \varphi_1)]^{-1}$ 에 대응할 수도 있으며, 한편 $EncMat_2$ 는 $[E_s(\theta_2, \varphi_2)]$ 에 대응할 수도 있다.

[1433] 위의 것은 (n 차의 구면 베셀 함수를 지칭하는) $j_n(\cdot)$ 함수 에 의해 주파수 도메인에서 SHC (511) 의 도출을 표기하는 다양한 등식들로 위에서 표현되는, 필터링 동작을 고려하지 않은 보다 간소화된 버전의 계산을 표현할 수도 있다. 시간 도메인에서, 이러한 $j_n(\cdot)$ 함수는 특정 오더 n 에 특정한 필터링 동작들을 표현한다. 필터링으로, 회전은 오더마다 수행될 수도 있다. 예시를 위해, 다음의 등식들을 고려한다:

[1434] $$a_n^k(t) \square b_n(t) * \left\langle \left[Y_n^m \right] \square \left[m_i(t) \right] \right\rangle$$
$$a_n^k(t) \square \left\langle \left[Y_n^m \right] \square b_n(t) * \left[m_i(t) \right] \right\rangle$$

[1435] 이러한 등식들로부터, 오더들에 대한 회전된 SHC (511') 가 별도로 행해지는데, bn(t) 가 각각의 오더에 대해 상이하기 때문이다. 그 결과, 위의 등식은 회전된 SHC (511') 의 제 1 오더를 계산하기 위해 다음과 같이 변경될 수도 있다:

[1436] $$\begin{bmatrix} 1^{st} \\ Order \\ SHC \\ 27' \end{bmatrix} = \begin{bmatrix} EncMat_2 \\ (3x32) \end{bmatrix} \begin{bmatrix} InvMat_1 \\ (32x3) \end{bmatrix} \begin{bmatrix} 1^{st} \\ Order \\ SHC \\ 27 \end{bmatrix}$$

[1437] SHC (511) 중 3 개의 1 차 SHC (511) 가 있다고 하면, SHC (511' 및 511) 벡터들의 각각은 위의 등식에서 사이즈가 3 이다. 마찬가지로, 2 차에 대해, 다음의 등식이 적용될 수도 있다:

[1438] $$\begin{bmatrix} 2^{nd} \\ Order \\ SHC \\ 27' \end{bmatrix} = \begin{bmatrix} EncMat_2 \\ (5x32) \end{bmatrix} \begin{bmatrix} InvMat_1 \\ (32x5) \end{bmatrix} \begin{bmatrix} 2^{nd} \\ Order \\ SHC \\ 27 \end{bmatrix}$$

[1439] 다시, SHC (511) 중 5 개의 1 차 SHC (511) 가 있다고 하면, SHC (511' 및 511) 벡터들의 각각은 위의 등식에서 사이즈가 5 이다. 다른 오더들에 대한 남은 등식들, 즉 제 3 오더 및 제 4 오더는 전술된 것과 유사할 수도 있으며, $EncMat_2$ 의 로우들의 수, $InvMat_1$ 의 컬럼들의 수, 및 제 3 오더 및 제 4 오더 SHC (511) 및 SHC (511') 벡터들의 사이즈들이 제 3 오더 및 제 4 오더 구면 조화 기저 함수들의 각각의 하위-오더들 (m 곱하기 2

더하기 1) 의 수와 동일하다는 점에서) 행렬들의 사이즈들에 대한 동일한 패턴을 따른다.

[1440] 오디오 인코딩 디바이스 (570) 는 따라서 이른바 최적 회전을 식별하기 위한 시도록 방위각 및 고도각의 모든 조합에 대해 이러한 회전 동작을 수행할 수도 있다. 오디오 인코딩 디바이스 (570) 는, 이러한 회전 동작을 수행한 후에, 임계 값을 넘는 SHC (511') 의 수를 계산할 수도 있다. 일부 경우들에서, 오디오 인코딩 디바이스 (570) 는 오디오 프레임과 같은 시간의 지속기간 동안의 음장을 나타내는 일련의 SHC (511') 를 도출하기 위해 이러한 회전을 수행할 수도 있다. 이러한 시간의 지속기간 동안의 음장을 나타내는 일련의 SHC (511') 를 도출하기 위해 이러한 회전을 수행함으로써, 오디오 인코딩 디바이스 (570) 는 프레임 또는 다른 길이보다 짧은 시간 지속기간들 동안의 음장을 서술하는 SHC (511) 의 각각의 세트에 대해 이를 수행하는 것과 비교하여 수행해야 할 회전 동작들의 수를 감소시킬 수도 있다. 임의의 경우에, 오디오 인코딩 디바이스 (570) 는, 이러한 프로세스에 걸쳐, 임계 값보다 큰 최소 수의 SHC (511') 를 갖는 SHC (511') 를 저장할 수도 있다.

[1441] 그러나, 방위각 및 고도각의 모든 조합에 대해 이러한 회전 동작을 수행하는 것은 프로세서 집약적이거나 시간 소모적일 수도 있다. 그 결과, 오디오 인코딩 디바이스 (570) 는 회전 알고리즘의 "무차별 (brute force)" 구현이라고 특징지어질 수도 있는 것을 수행하지 않을 수도 있다. 대신에, 오디오 인코딩 디바이스 (570) 는 일반적으로 압축을 제공하는 방위각과 고도각의 가능하게는 (통계적 측면에서) 알려진 조합의 서브세트에 대해 회전들을 수행하여, 이러한 서브세트에 맞춘 조합들에 대해 회전들을 더 수행해 서브세트에서의 다른 조합들과 비교하여 보다 좋은 압축을 제공할 수도 있다.

[1442] 다른 대안으로서, 오디오 인코딩 디바이스 (570) 는 오직 알려진 조합들의 서브세트에 대해서만 이러한 회전을 수행할 수도 있다. 다른 대안으로서, 오디오 인코딩 디바이스 (570) 는 조합들의 (공간적) 궤적을 따라, 이러한 조합들의 궤적에 대해 회전들을 수행할 수도 있다. 다른 대안으로서, 오디오 인코딩 디바이스 (570) 는 임계 값을 넘는 비-제로 값들을 갖는 SHC (511') 의 최대 수를 정의하는 압축 임계치를 명시할 수도 있다. 이러한 압축 임계는 검색에 대한 중지 지점을 효과적으로 설정할 수도 있어, 오디오 인코딩 디바이스 (570) 가 회전을 수행하고, 임계치를 넘는 값을 갖는 SHC (511') 의 수가 압축 임계치 이하 (또는 일부 경우들에서 미만) 라고 결정하는 경우, 오디오 인코딩 디바이스 (570) 는 남은 조합들에 대해 임의의 추가적인 회전 동작들을 수행하는 것을 중지한다. 또 다른 대안으로서, 오디오 인코딩 디바이스 (570) 는 계층적으로 배열된 조합들의 트리 (또는 다른 데이터 구조) 를 트래버싱하여, 임계 값보다 큰 비-제로 값을 갖는 SHC (511') 의 수에 따라 현재 조합에 대해 회전 동작들을 수행하고 (예를 들어, 이진 트리들에 있어서) 오른쪽 또는 왼쪽으로 트리를 트래버싱할 수도 있다.

[1443] 이러한 경우에, 이러한 대안들의 각각은 제 1 및 제 2 회전 동작을 수행하고 제 1 및 제 2 회전 동작을 수행한 결과를 비교하여 임계 값보다 큰 비-제로 값을 갖는 SHC (511') 의 최소 수를 초래하는 제 1 및 제 2 회전 동작들 중 하나의 동작을 식별하는 것을 수반한다. 이에 따라, 오디오 인코딩 디바이스 (570) 는 제 1 방위각 및 제 1 고도각에 따라 음장을 회전시키도록 음장에 대해 제 1 회전 동작을 수행하고, 음장을 서술하는 것과 관련되는 정보를 제공하는 제 1 방위각 및 제 1 고도각에 따라 회전된 음장을 나타내는 복수의 계층적 엘리먼트들의 제 1 수를 결정할 수도 있다. 오디오 인코딩 디바이스 (570) 는 또한 제 2 방위각 및 제 2 고도각에 따라 음장을 회전시키도록 음장에 대해 제 2 회전 동작을 수행하고, 음장을 서술하는 것과 관련되는 정보를 제공하는 제 2 방위각 및 제 2 고도각에 따라 회전된 음장을 나타내는 복수의 계층적 엘리먼트들의 제 2 수를 결정할 수도 있다. 또한, 오디오 인코딩 디바이스 (570) 는 복수의 계층적 엘리먼트들의 제 1 수 및 복수의 계층적 엘리먼트들의 제 2 수의 비교에 기초하여 제 1 회전 동작 또는 제 2 회전 동작을 선택할 수도 있다.

[1444] 일부 경우들에서, 회전 알고리즘은 시간의 지속기간에 대해 수행될 수도 있으며, 여기서 후속하는 회전 알고리즘의 실시들은 회전 알고리즘의 과거 실시들에 기초하여 회전 동작들을 수행할 수도 있다. 다시 말해, 회전 알고리즘은 이전의 시간의 지속기간에 대해 음장을 회전시키는 경우 결정된 과거 회전 정보에 기초하여 적응될 수도 있다. 예를 들어, 오디오 인코딩 디바이스 (570) 는 시간의 제 1 지속기간, 예를 들어, 오디오 프레임에 대해 음장을 회전시켜, 이러한 시간의 제 1 지속기간에 대한 SHC (511') 를 식별할 수도 있다. 오디오 인코딩 디바이스 (570) 는 전술된 방식들 중 임의의 방식으로 비트스트림 (517) 에 회전 정보 및 SHC (511') 를 명시할 수도 있다. 이러한 회전 정보는 시간의 제 1 지속기간 동안의 음장의 회전을 서술한다는 점에서 제 1 회전 정보라고 지칭될 수도 있다. 오디오 인코딩 디바이스 (570) 는, 그 다음에, 제 1 회전 정보에 기초하여, 시간의 제 2 지속기간, 예를 들어, 제 2 오디오 프레임 동안에 음장을 회전시켜, 이러한 시간의 제 1 지속기간에 대한 SHC (511') 를 식별할 수도 있다. 오디오 인코딩 디바이스 (570) 는, 일 예로서, 방위각 및 고도각의 "최적" 조합에 대한 검색을 초기화하기 위해 제 2 시간의 지속기간 동안에 제 2 회전 동작을 수행하는 경우 이러한 제 1 회전 정보를 사용할 수도 있다. 오디오 인코딩 디바이스 (570) 는 그러면 비트스트림

(517) 에 시간의 제 2 지속기간에 대한 SHC (511') 및 대응하는 회전 정보 (이는 "제 2 회전 정보" 라고 지칭될 수도 있다) 를 명시할 수도 있다.

[1445] 프로세싱 시간 및/또는 소비를 감소시키기 위해 회전 알고리즘을 구현하는 다수의 상이한 방식들에 대해 전술되었으나, 기술들은 "최적 회전" 이라고 지칭될 수도 있는 식별을 감소시키거나 그렇지 않으면 속도를 높이는 임의의 알고리즘에 대해 수행될 수도 있다. 또한, 기술들은 비-최적 회전들을 식별하나, 종종 속도, 또는 프로세서, 또는 다른 자원 활용의 면에서 측정되는 다른 양상들에서의 성능을 향상시킬 수도 있는 임의의 알고리즘에 대해 수행될 수도 있다.

[1446] 도 57 내지 도 57e 는 각각 본 개시물에 설명된 기술들에 따라 형성된 비트스트림들 (517A-517E) 을 예시하는 다이어그램이다. 도 57a 의 예에서, 비트스트림 (517A) 은 위의 도 53 에 도시된 비트스트림 (517) 의 일 예를 표현할 수도 있다. 비트스트림 (517A) 은 SHC 존재 필드 (670) 및 SHC (511') 를 저장하는 필드 (여기서 필드는 "SHC (511')" 로 표기된다) 를 포함한다. SHC 존재 필드 (670) 는 SHC (511) 의 각각에 대응하는 비트를 포함할 수도 있다. SHC (511') 는, SHC (511) 의 수보다 적을 수도 있는, 비트스트림에 명시된 SHC (511) 를 표현할 수도 있다. 통상적으로, SHC (511') 의 각각은 비-제로 값들을 갖는 SHC (511) 이다.

위에서 언급된 바와 같이, 임의의 주어진 음장의 4 차 표현을 위해서는, $(1+4)^2$ 또는 25 개의 SHC 가 요구된다. 이러한 SHC 중 하나 이상의 SHC 를 제거하고 이를 단일 비트를 갖는 제로 값의 SHC 로 대체하는 것은 31 비트를 절약할 수도 있으며, 31 비트는 음장의 다른 부분들을 더 상세하게 나타내는데 할당될 수도 있거나, 그렇지 않으면 효율적인 대역폭 활용을 가능하게 하기 위해 제거될 수도 있다.

[1447] 도 57b 의 예에서, 비트스트림 (517B) 은 위의 도 53 에 도시된 비트스트림 (517) 의 일 예를 표현할 수도 있다. 비트스트림 (517B) 은 변환 정보 필드 (672) ("변환 정보 (672)") 및 SHC (511') 를 저장하는 필드 (여기서 필드는 "SHC (511')" 로 표기된다) 를 포함한다. 변환 정보 (672) 는, 위에서 언급된 바와 같이, 변환 정보, 회전 정보, 및/또는 음장에 대한 조정을 표기하는 임의의 다른 형태의 정보를 포함할 수도 있다. 일부 경우들에서, 변환 정보 (672) 는 또한 SHC (511') 로 비트스트림 (517B) 에 명시되는 가장 상위 오더의 SHC (511) 를 명시할 수도 있다. 즉, 변환 정보 (672) 는 3 의 오더를 나타낼 수도 있으며, 추출 디바이스는 SHC (511') 가 최대 SHC (511) 를 포함하는 것으로 나타내고 3 의 오더를 갖는 SHC (511) 를 포함하는 것으로 이해할 수도 있다. 추출 디바이스는 그 다음에 4 이상의 오더를 갖는 SHC (511) 를 제로로 설정하도록 구성될 수도 있으며, 그렇게 함으로써 잠재적으로 비트스트림에서 4 이상의 오더의 SHC (511) 의 명시적 시그널링을 제거한다.

[1448] 도 57c 의 예에서, 비트스트림 (517C) 은 위의 도 53 에 도시된 비트스트림 (517) 의 일 예를 표현할 수도 있다. 비트스트림 (517C) 은 변환 정보 필드 (672) ("변환 정보 (672)"), SHC 존재 필드 (670), 및 SHC (511') 를 저장하는 필드 (여기서 필드는 "SHC (511')" 로 표기된다) 를 포함한다. 도 57b 에 대해 전술된 바와 같이 어떤 오더의 SHC (511) 가 시그널링되지 않을 것인지를 이해하도록 구성되기 보다는, SHC 존재 필드 (670) 는 SHC (511) 중 어느 것이 비트스트림 (517C) 에 SHC (511') 로서 명시되는지를 명시적으로 시그널링할 수도 있다.

[1449] 도 57d 의 예에서, 비트스트림 (517D) 은 위의 도 53 에 도시된 비트스트림 (517) 의 일 예를 표현할 수도 있다. 비트스트림 (517D) 은 오더 필드 (674) ("오더 (60)"), SHC 존재 필드 (670), 방위각 플래그 (676) ("AZF (676)"), 고도각 플래그 (678) ("ELF (678)"), 방위각 필드 (680) ("방위각 (680)"), 고도각 필드 (682) ("고도각 (682)"), 및 SHC (511') 를 저장하는 필드 (여기서, 다시, 필드는 "SHC (511')" 이라고 표기된다) 를 포함한다. 오더 필드 (674) 는 SHC (511') 의 오더, 즉 음장을 나타내는데 이용되는 구면 기저 함수의 가장 상위 오더에 대해 위에서 n 으로 표기된 오더를 명시한다. 오더 필드 (674) 는 8 비트 필드인 것으로 도시되나, 3 과 같은 다른 다양한 비트 사이즈들일 수도 있다 (이는 4 개의 오더를 명시하는데 요구되는 비트들의 수이다). SHC 존재 필드 (670) 는 25 비트 필드로 도시된다. 다시, 그러나, SHC 존재 필드 (670) 는 다른 다양한 비트 사이즈들일 수도 있다. SHC 존재 필드 (670) 는 SHC 존재 필드 (670) 가 음장의 4 개의 오더 표현에 대응하는 구면 조화 계수들의 각각에 대해 1 비트를 포함할 수도 있다는 것을 나타내기 위해 25 비트로 도시된다.

[1450] 방위각 플래그 (676) 는 방위각 필드 (680) 가 비트스트림 (517D) 에 존재하는지 여부를 명시하는 1-비트 플래그를 표현한다. 방위각 플래그 (676) 가 1 로 설정된 경우, SHC (511') 에 대한 방위각 필드 (680) 가 비트스트림 (517D) 에 존재한다. 방위각 플래그 (676) 가 제로로 설정된 경우, SHC (511') 의 방위각 필드 (680) 가 비트스트림 (517D) 에 존재하지 않거나 그렇지 않으면 명시되지 않는다. 마찬가지로, 고도각 플래

그 (678) 는 고도각 필드 (682) 가 비트스트림 (517D) 에 존재하는지 여부를 명시하는 1-비트 플래그를 표현한다. 고도각 플래그 (678) 가 1 로 설정된 경우, SHC (511') 에 대한 고도각 필드 (682) 가 비트스트림 (517D) 에 존재한다. 고도각 플래그 (678) 가 제로로 설정된 경우, SHC (511') 에 대한 고도각 필드 (682) 가 비트스트림 (517D) 에 존재하지 않거나 그렇지 않으면 명시되지 않는다. 대응하는 필드가 존재하는 경우 1 을 시그널링하고 대응하는 필드가 존재하지 않는 경우 제로를 시그널링하는 것으로 설명되었으나, 제로가 대응하는 필드가 비트스트림 (517D) 에 명시되는 것을 명시하고 1 이 대응하는 필드가 비트스트림 (517D) 에 명시되지 않음을 명시하도록 종래기술이 반전될 수도 있다. 본 개시물에 설명된 기술들은 따라서 이러한 면으로 제한되지 않아야 한다.

[1451] 방위각 필드 (680) 는, 비트스트림 (517D) 에 존재하는 경우, 방위각을 명시하는 10-비트 필드를 표현한다. 10-비트 필드로 도시되었으나, 방위각 필드 (680) 는 다른 비트 사이즈들일 수도 있다. 고도각 필드 (682) 는, 비트스트림 (517D) 에 존재하는 경우, 고도각을 명시하는 9-비트 필드를 표현한다. 필드들 (680 및 682) 에 명시된 방위각 및 고도각은, 각각, 전술된 회전 정보를 나타내는 플래그들 (676 및 678) 과 연계될 수도 있다. 이러한 회전 정보는 SHC (511) 를 원래의 참조의 프레임으로 복원하도록 음장을 회전시키는데 이용될 수도 있다.

[1452] SHC (511') 필드는 사이즈 (X) 가변 필드로 도시된다. SHC (511') 필드는 SHC 존재 필드 (670) 에 의해 표기되는 바와 같이 비트스트림에 명시되는 SHC (511') 의 수로 인해 달라질 수도 있다. 사이즈 (X) 는 SHC 존재 필드 (670) 곱하기 32-비트에서의 수의 함수로서 도출될 수도 있다 (이는 각각의 SHC (511') 의 사이즈이다).

[1453] 도 57e 의 예에서, 비트스트림 (517E) 은 위의 도 53 에 도시된 비트스트림 (517) 의 다른 예를 표현할 수도 있다. 비트스트림 (517E) 은 오더 필드 (674) ("오더 (60)"), SHC 존재 필드 (670), 및 회전 인덱스 필드 (684), 및 SHC (511') 를 저장하는 필드 (여기서, 다시, 필드는 "SHC (511')" 로 표기된다) 를 포함한다. 오더 필드 (674), SHC 존재 필드 (670), 및 SHC (511') 필드는 전술된 것들과 실질적으로 유사할 수도 있다. 회전 인덱스 필드 (684) 는 고도각 및 방위각의 1024x512 (또는, 다시 말해, 524288) 개의 조합들 중 하나를 명시하는데 이용되는 20-비트 필드를 표현할 수도 있다. 일부 경우들에서, 오직 19-비트만이 이러한 회전 인덱스 필드 (684) 를 명시하는데 이용될 수도 있고, 오디오 인코딩 디바이스 (570) 는 회전 동작이 수행되었는지 여부 (및, 따라서, 회전 인덱스 필드 (684) 가 비트스트림에 존재하는지 여부) 를 나타내기 위해 비트스트림에 추가적인 플래그를 명시할 수도 있다. 이러한 회전 인덱스 필드 (684) 는 위에서 언급된 회전 인덱스를 명시하며, 이는 오디오 인코딩 디바이스 (570) 및 비트스트림 추출 디바이스 양자 모두에 공통인 회전 테이블에서의 엔트리를 지칭할 수도 있다. 이러한 회전 테이블은, 일부 경우들에서, 방위각 및 고도각의 상이한 조합들을 저장할 수도 있다. 대안으로, 회전 테이블은 전술된 행렬을 저장할 수도 있으며, 이는 행렬 형태로 방위각 및 고도각의 상이한 조합들을 효과적으로 저장한다.

[1454] 도 58 은 본 개시물에 설명된 기술들의 회전 양상들을 구현할 시에 도 53 의 예에 도시된 오디오 인코딩 디바이스 (570) 의 예시적인 동작을 예시하는 플로우차트이다. 최초에, 오디오 인코딩 디바이스 (570) 는 전술된 다양한 회전 알고리즘들 하나 이상의 회전 알고리즘에 따라 방위각 및 고도각 조합을 선택할 수도 있다 (800). 오디오 인코딩 디바이스 (570) 는 그 다음에 선택된 방위각 및 고도각에 따라 음장을 회전시킬 수도 있다 (802). 전술된 바와 같이, 오디오 인코딩 디바이스 (570) 는 우선 위에서 언급된 $InvMat_1$ 를 이용하여 SHC (511) 로부터 음장을 도출할 수도 있다. 오디오 인코딩 디바이스 (570) 는 또한 회전된 음장을 나타내는 SHC (511') 를 결정할 수도 있다 (804). 별도의 단계들 또는 동작들로 설명되긴 했으나, 오디오 인코딩 디바이스 (570) 는 방위각 및 고도각 조합의 선택을 나타내는 변환 (이는 $[EncMat_2][InvMat_1]$ 의 결과를 표현할 수도 있다) 을 적용하여, SHC (511) 로부터 음장을 도출하고, 음장을 회전시키고, 회전된 음장을 나타내는 SHC (511') 를 결정할 수도 있다.

[1455] 임의의 경우에, 오디오 인코딩 디바이스 (570) 는 그 다음에 임계 값보다 큰 결정된 SHC (511') 의 수를 컴퓨팅하여, 이러한 수를 이전 방위각 및 고도각에 대한 이전의 반복에 대해 컴퓨팅된 수와 비교할 수도 있다 (806, 808). 제 1 방위각 및 고도각 조합에 대한 제 1 반복에서, 이러한 비교는 미리 정의된 이전의 수 (이는 제로로 설정될 수도 있다) 일 수도 있다. 임의의 경우에, SHC (511') 의 결정된 수가 이전의 수보다 작은 경우 ("예" (808)), 오디오 인코딩 디바이스 (570) 는 SHC (511'), 방위각, 및 고도각을 저장하고, 종종 회전 알고리즘의 이전의 반복으로부터 저장된 이전의 SHC (511'), 방위각, 및 고도각을 대체한다 (810).

[1456] SHC (511') 의 결정된 수가 이전 수보다 적지 않은 경우 ("아니오" (808)), 또는 저장된 SHC (511'), 방위각,

및 고도각 대신에 SHC (511'), 방위각, 및 고도각을 저장한 후에, 오디오 인코딩 디바이스 (570) 는 회전 알고리즘이 완료되었는지 여부를 결정할 수도 있다 (812). 즉, 오디오 인코딩 디바이스 (570) 는, 일 예로서, 방위각 및 고도각의 모든 이용가능한 조합이 평가되었는지 여부를 결정할 수도 있다. 다른 예들에서, 오디오 인코딩 디바이스 (570) 는 오디오 인코딩 디바이스 (570) 가 회전 알고리즘을 수행하는 것을 완료하도록 다른 기준이 충족되었는지 여부 (예컨대, 조합의 정의된 서브세트 모두가 수행되었는지, 주어진 궤적이 트래버싱되었는지 여부, 계층 트리가 리프 노드까지 트래버싱되었는지 여부 등) 를 결정할 수도 있다. 완료되지 않은 경우 ("아니오" (812)), 오디오 인코딩 디바이스 (570) 는 다른 선택된 조합에 대해 위의 프로세스를 수행할 수도 있다 (800-812). 완료된 경우 ("예" (812)), 오디오 인코딩 디바이스 (570) 는 전술된 다양한 방식들 중 하나의 방식으로 저장된 SHC (511'), 방위각, 및 고도각을 비트스트림 (517) 에 명시할 수도 있다 (814).

[1457] 도 59 은 본 개시물에 설명된 기술들의 변환 양상들을 수행할 시에 도 53 의 예에 도시된 오디오 인코딩 디바이스 (570) 의 예시적인 동작을 예시하는 플로우차트이다. 최초에, 오디오 인코딩 디바이스 (570) 는 선형 가역 변환을 나타내는 행렬을 선택할 수도 있다 (820). 선형 가역 변환을 나타내는 행렬의 일 예는 $[EncMat_1][IncMat_1]$ 의 결과인, 위에서 보여진 행렬일 수도 있다. 오디오 인코딩 디바이스 (570) 는 그 다음에 행렬을 음장에 적용하여 음장을 변환시킬 수도 있다 (822). 오디오 인코딩 디바이스 (570) 는 또한 회전된 음장을 나타내는 SHC (511') 를 결정할 수도 있다 (824). 별도의 단계들 및 동작들인 것으로 설명되었으나, 오디오 인코딩 디바이스 (570) 는 변환 (이는 $[EncMat_2][InvMat_1]$ 의 결과를 표현할 수도 있다) 을 적용하여, SHC (511) 로부터 음장을 도출하고, 음장을 변환시키고, 변환 음장을 나타내는 SHC (511') 를 결정할 수도 있다.

[1458] 임의의 경우에, 오디오 인코딩 디바이스 (570) 는 그 다음에 임계 값보다 큰 결정된 SHC (511') 의 수를 컴퓨팅하여, 이러한 수를 변환 행렬의 이전 적용에 대한 이전의 반복에 대해 컴퓨팅된 수와 비교할 수도 있다 (826, 828). SHC (511') 의 결정된 수가 이전의 수보다 작은 경우 ("예" (828)), 오디오 인코딩 디바이스 (570) 는 SHC (511') 및 행렬 (또는 행렬와 연관된 인덱스와 같은 일부 도출물) 을 저장하여, 종종 회전 알고리즘의 이전의 반복으로부터 저장된 이전 SHC (511') 및 행렬 (또는 그것의 도출물) 을 대체한다 (830).

[1459] SHC (511') 의 결정된 수가 이전 수보다 작지 않은 경우 ("아니오" (828)), 또는 이전에 저장된 SHC (511') 및 행렬 대신에 SHC (511') 및 행렬을 저장한 후에, 오디오 인코딩 디바이스 (570) 는 변환 알고리즘이 완료되었는지 여부를 결정할 수도 있다 (832). 즉, 오디오 인코딩 디바이스 (570) 는, 일 예로서, 모든 이용가능한 변환 행렬들이 평가되었는지 여부를 결정할 수도 있다. 다른 예들에서, 오디오 인코딩 디바이스 (570) 는 오디오 인코딩 디바이스 (570) 가 변환 알고리즘을 수행하는 것을 완료하도록 다른 기준이 충족되었는지 여부 (예컨대, 이용가능한 변환 행렬들의 정의된 서브세트 모두가 수행되었는지, 주어진 궤적이 트래버싱되었는지 여부, 계층 트리가 리프 노드까지 트래버싱되었는지 여부 등) 를 결정할 수도 있다. 완료되지 않은 경우 ("아니오" (832)), 오디오 인코딩 디바이스 (570) 는 다른 선택된 변환 행렬에 대해 위의 프로세스를 수행할 수도 있다 (820-832). 완료된 경우 ("예" (832), 오디오 인코딩 디바이스 (570) 는 전술된 다양한 방식들 중 하나의 방식으로 저장된 SHC (511') 및 행렬을 비트스트림 (517) 에 명시할 수도 있다 (834).

[1460] 일부 예들에서, 변환 알고리즘은 단일 반복을 수행하여, 단일 변환 행렬을 평가할 수도 있다. 즉, 변환 행렬은 선형 가역 변환을 나타내는 임의의 행렬을 포함할 수도 있다. 일부 경우들에서, 선형 가역 변환은 음장을 공간 도메인으로부터 주파수 도메인으로 변환시킬 수도 있다. 그러한 선형 가역 변환들의 예들은 이산 푸리에 변환 (discrete Fourier transform; DFT) 을 포함할 수도 있다. DFT 의 적용은 오직 단일 반복만을 수반할 수도 있고, 따라서 변환 알고리즘이 완료되었는지 여부를 결정하기 위한 단계들을 반드시 포함하지는 않는다. 이에 따라, 기술들은 도 59 의 예로 제한되어서는 안된다.

[1461] 다시 말해, 선형 가역 변환의 일 예는 이산 푸리에 변환 (DFT) 이다. 25 개의 SHC (511') 이 DFT 에 의해 동작되어 25 개의 복소 계수들의 세트를 형성할 수 있다. 오디오 인코딩 디바이스 (570) 는 2 의 정수 배수일 25 개의 SHC (511') 들이, 예를 들어, 고속 푸리에 변환 (fast Fourier transform; FFT) 을 적용하여, 잠재적으로 DFT 의 빈 사이즈의 분해능을 증가시키고, 잠재적으로 DFT 의 보다 효율적을 구현을 가질 수 있도록, 제로-패딩할 수도 있다. 일부 경우들에서, DFT 의 분해능을 25 지점들보다 넘게 증가시키는 것이 반드시 요구되는 것은 아니다. 변환 도메인에서, 오디오 인코딩 디바이스 (570) 는 특정 빈에 임의의 스펙트럼 에너지가 있는지 여부를 결정하기 위해 임계치를 적용할 수도 있다. 오디오 인코딩 디바이스 (570) 는, 이러한 맥락에서, 이러한 임계치 아래인 스펙트럼 계수 에너지를 폐기하거나 제로로 할 수도 있고, 오디오 인코딩 디바이스 (570) 는 폐기되거나 제로로 된 SHC (511') 중 하나 이상을 갖는 SHC (511') 를 복원하기 위해 역변환을 적

용할 수도 있다. 즉, 역변환이 적용된 후에, 임계치 아래의 계수들은 존재하지 않고, 그 결과, 음장을 인코딩하는데 보다 적은 비트들이 이용될 수도 있다.

[1462] 하나 이상의 예들에서, 설명된 기능들은 하드웨어, 소프트웨어, 펌웨어, 또는 이들의 임의의 조합으로 구현될 수도 있다. 소프트웨어로 구현되는 경우, 기능들은 하나 이상의 명령들 또는 코드로서 컴퓨터 판독가능 매체 상에 저장되거나 송신되거나 하드웨어-기반 프로세싱 유닛에 의해 실행될 수도 있다. 컴퓨터 판독가능 매체들은, 데이터 저장 매체들과 같은 유형의 매체에 대응하는 컴퓨터 판독가능 저장 매체들, 또는 예를 들어, 통신 프로토콜에 따라, 한 곳에서 다른 곳으로 컴퓨터 프로그램의 전송을 가능하게 하는 임의의 매체를 포함하는 통신 매체들을 포함할 수도 있다. 이러한 방식으로, 컴퓨터 판독가능 매체들은 일반적으로 (1) 비-일시적인 유형의 컴퓨터 판독가능 저장 매체들, 또는 (2) 신호 또는 반송파와 같은 통신 매체에 대응할 수도 있다.

데이터 저장 매체들은 본 개시물에 설명된 기술들의 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 판독가능 매체를 포함할 수도 있다.

[1463] 비제한적인 예로서, 그러한 컴퓨터 판독가능 저장 매체들은 RAM, ROM, EEPROM, CD-ROM 또는 다른 광학 디스크 스토리지, 자기 디스크 스토리지 또는 다른 자기 저장 디바이스들, 플래시 메모리, 또는 원하는 프로그램 코드를 명령들 또는 데이터 구조들의 형태로 이송 또는 저장하기 위해 이용될 수 있고 컴퓨터에 의해 액세스될 수 있는 임의의 다른 매체를 포함할 수 있다. 또한, 임의의 접속은 컴퓨터 판독가능 매체라고 적절히 칭해진다. 예를 들어, 동축 케이블, 광섬유 케이블, 연선, 디지털 가입자 회선 (digital subscriber line; DSL), 또는 적외선, 무선, 및 마이크로파와 같은 무선 기술들을 이용하여 웹사이트, 서버, 또는 다른 원격 소스로부터 명령들이 송신되는 경우, 동축 케이블, 광섬유 케이블, 연선, DSL, 또는 적외선, 무선, 및 마이크로파와 같은 무선 기술들은 매체의 정의 내에 포함된다. 그러나, 컴퓨터 판독가능 저장 매체들 및 데이터 저장 매체들은 접속부들, 반송파들, 신호들, 또는 다른 일시적 매체들을 포함하지 않고, 대신에 비-일시적, 유형의 저장 매체들에 대한 것임이 이해되어야 한다. 본원에서 이용된 디스크 (disk) 와 디스크 (disc) 는, 컴팩트 디스크(compact disc; CD), 레이저 디스크, 광학 디스크, 디지털 다기능 디스크 (digital versatile disc; DVD), 플로피 디스크, 및 블루레이 디스크를 포함하며, 여기서 디스크 (disk) 들은 통상 자기적으로 데이터를 재생하는 반면, 디스크 (disc) 들은 레이저들을 이용하여 광학적으로 데이터를 재생한다. 위의 조합들도 컴퓨터 판독가능 매체들의 범위 내에 포함되어야 한다.

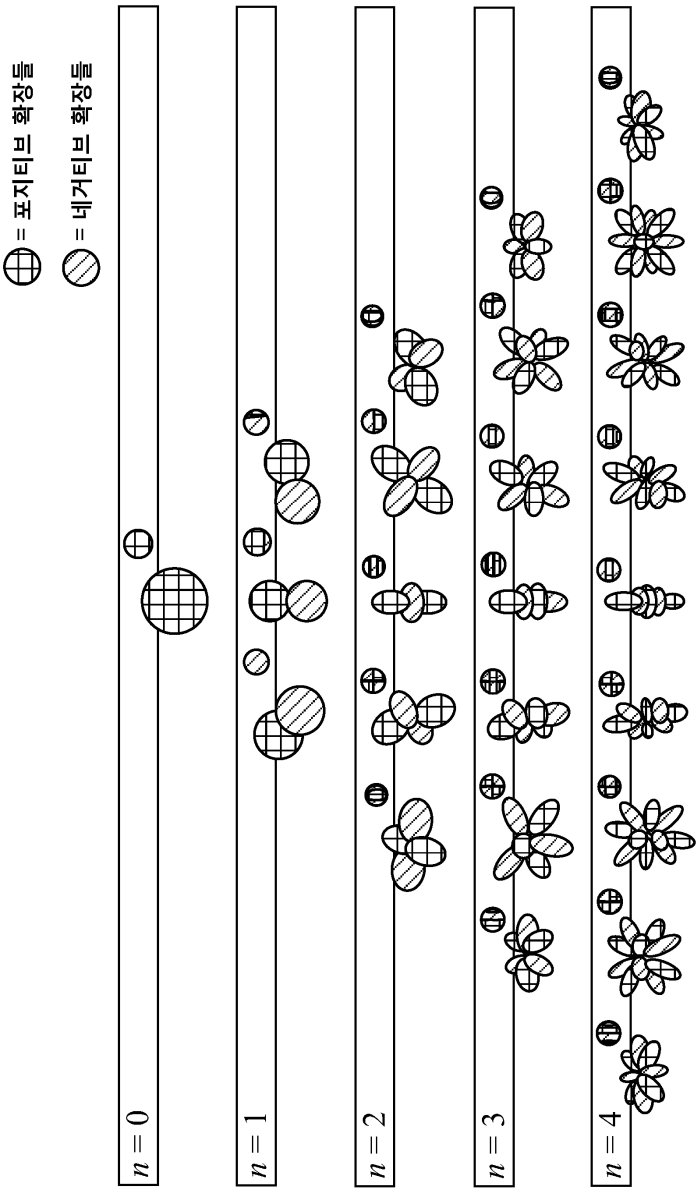
[1464] 명령들은, 하나 이상의 디지털 신호 프로세서 (digital signal processor; DSP) 들, 범용 마이크로프로세서들, 주문형 반도체 (application specific integrated circuit; ASIC) 들, 필드 프로그래머블 로직 어레이 (field programmable logic array; FPGA) 들, 또는 다른 등가의 집적 또는 이산 로직 회로와 같은, 하나 이상의 프로세서들에 의해 실행될 수도 있다. 이에 따라, 본원에서 이용되는 바와 같은 용어 "프로세서" 는 앞서 언급한 구조, 또는 본원에서 설명된 기술들을 구현하기에 적합한 임의의 다른 구조 중 임의의 것을 지칭할 수도 있다. 또한, 일부 양상들에서, 본원에서 설명된 기능성은 인코딩 및 디코딩을 위해 구성된 전용 하드웨어 및/또는 소프트웨어 모듈 내에 제공되거나, 통합 코덱에 통합될 수도 있다. 또한, 기술들은 하나 이상의 회로들 또는 로직 엘리먼트들에서 완전히 구현될 수 있다.

[1465] 본 개시물의 기술들은 무선 핸드셋, 집적 회로 (integrated circuit; IC), 또는 IC 들의 세트 (예를 들어, 칩셋) 를 포함하여, 매우 다양한 디바이스들 또는 장치들로 구현될 수도 있다. 개시된 기술들을 수행하도록 구성된 디바이스들의 기능적 양상들을 강조하기 위해 다양한 성분들, 모듈들, 또는 유닛들이 본 개시물에서 설명되었지만, 반드시 상이한 하드웨어 유닛들에 의한 실현을 요구하지는 않는다. 오히려, 전술된 바와 같이, 다양한 유닛들은, 적합한 소프트웨어 및/또는 펌웨어와 연계하여, 코덱 하드웨어 유닛에 통합되거나 전술한 하나 이상의 프로세서들을 포함하여 상호동작적인 하드웨어 유닛들의 집합에 의해 제공될 수도 있다.

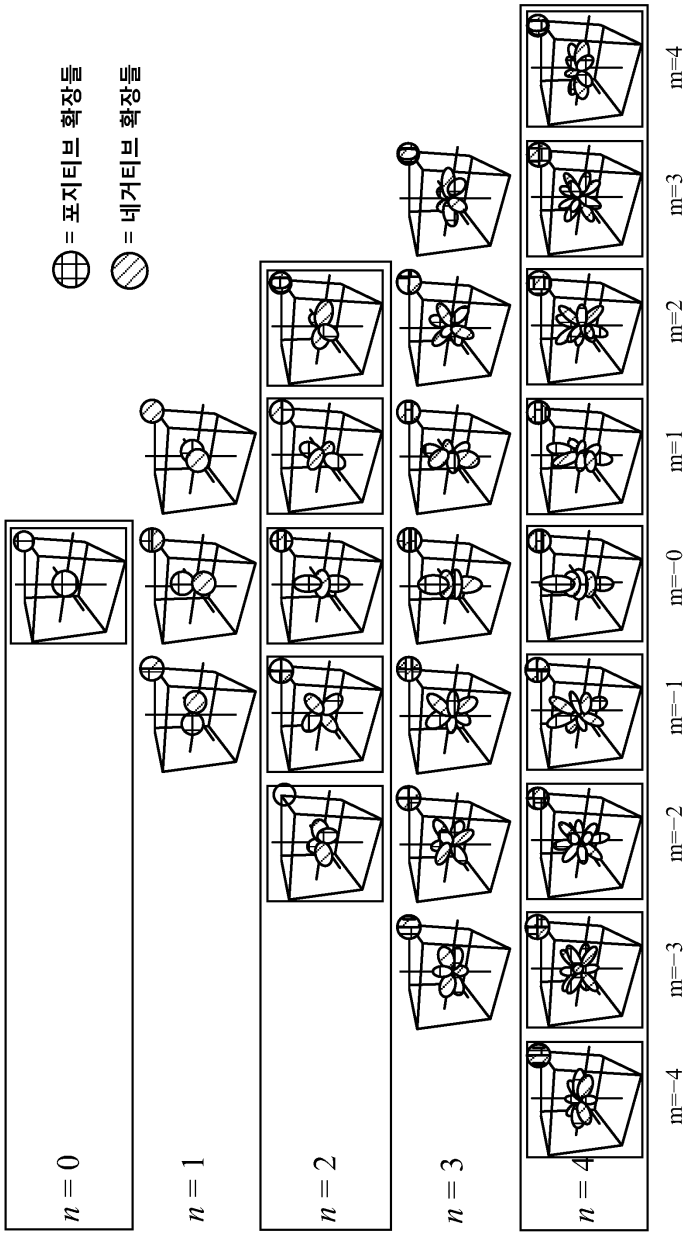
[1466] 기술들이 다양한 실시형태들이 설명되었다. 기술들의 이러한 양상들 및 다른 양상들은 다음의 청구항들의 범위 내에 있다.

도면

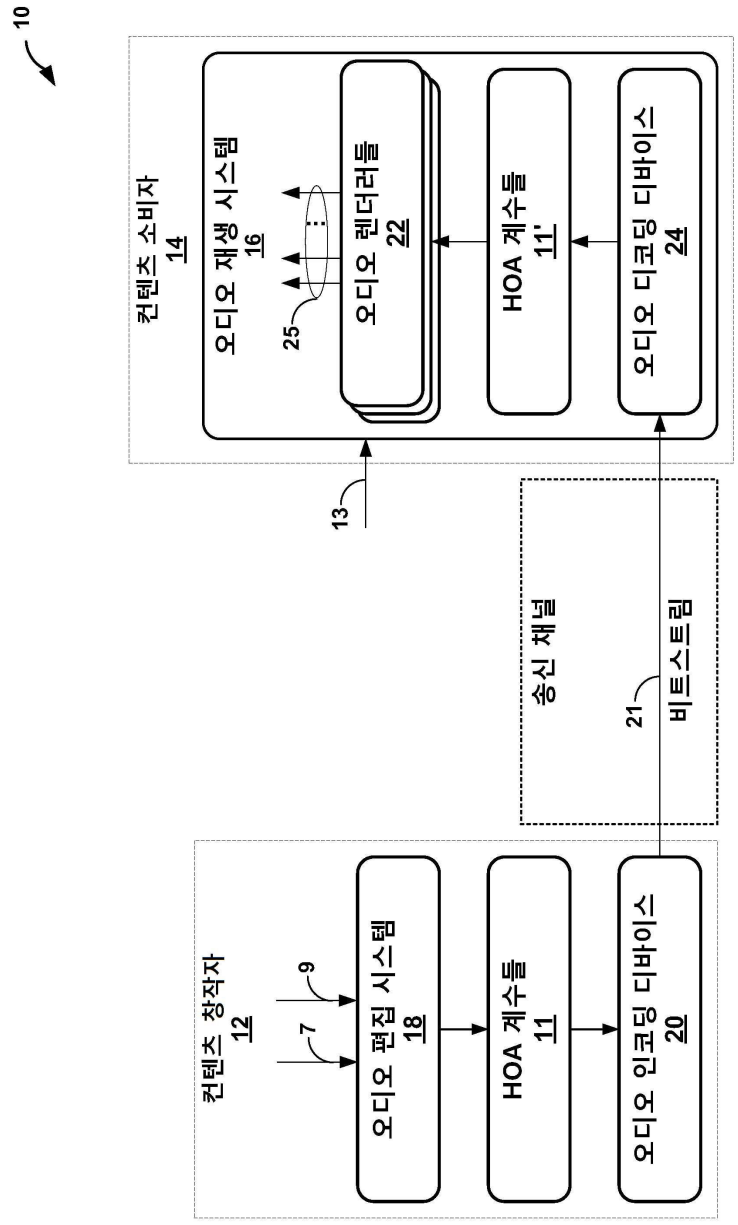
도면1



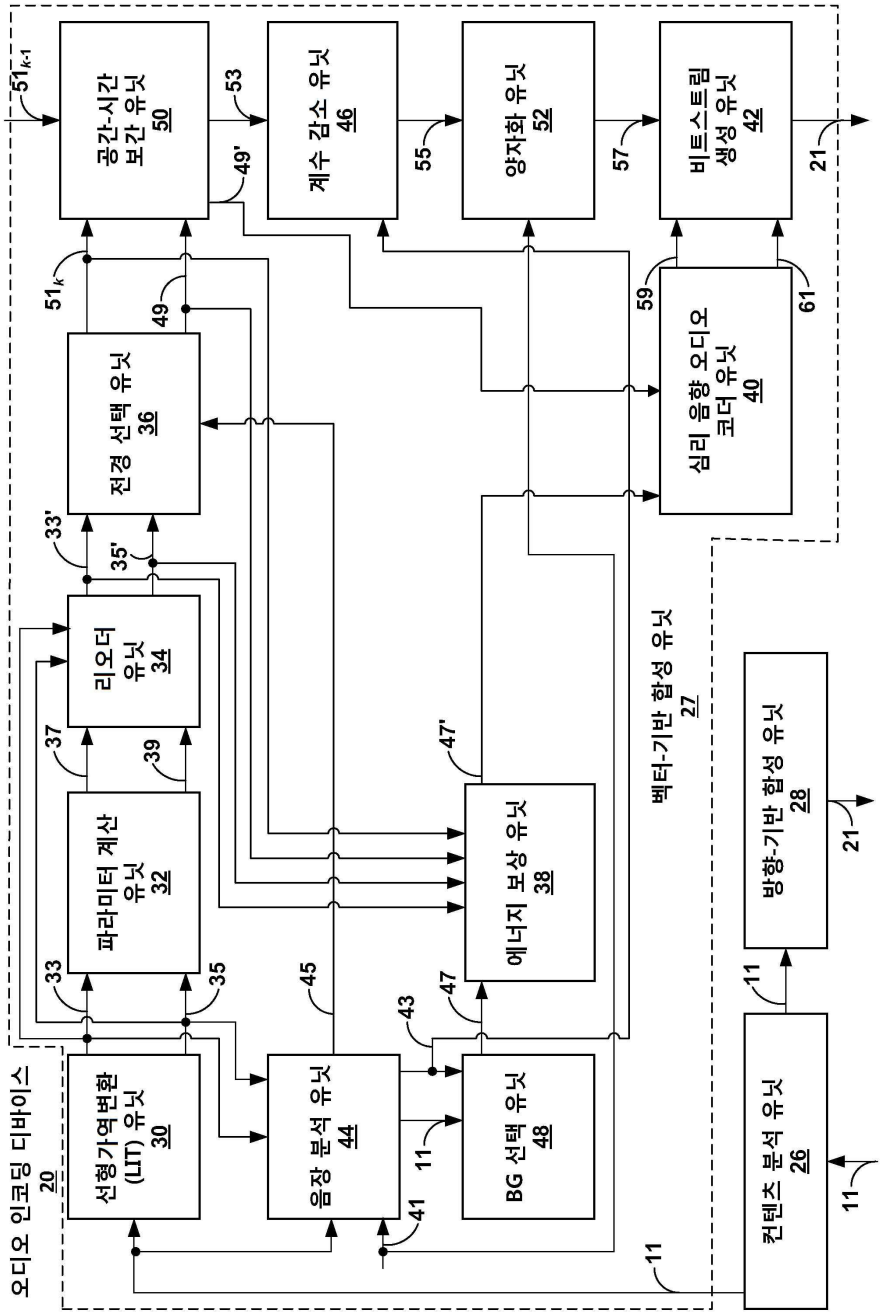
도면2



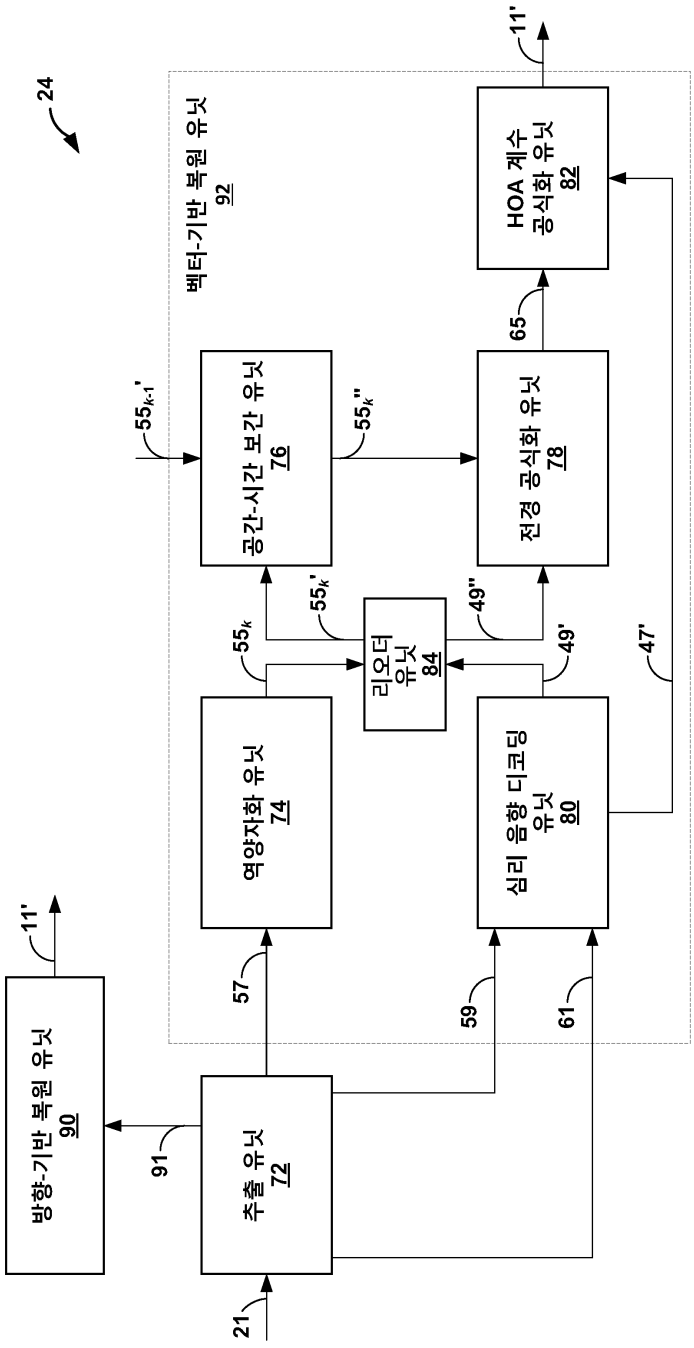
도면3


10
컨텐츠 소비자
14
오디오 재생 시스템
16
25
오디오 렌더러들
22
HOA 계수들
11'
오디오 디코딩 디바이스
24
13
송신 채널
21
비트스트림
컨텐츠 창작자
12
9
7
오디오 편집 시스템
18
HOA 계수들
11
오디오 인코딩 디바이스
20

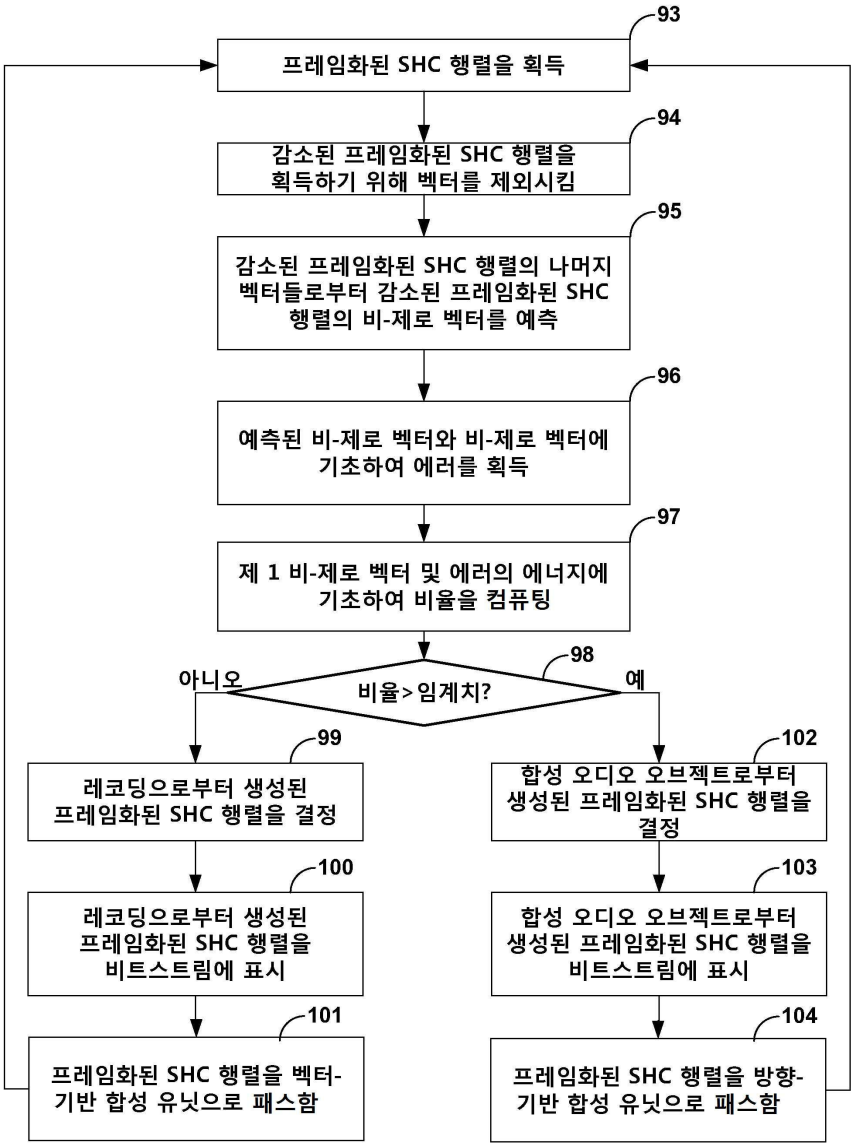
도면4


오디오 인코딩 디바이스
20
선형가역변환 (LIT) 유닛
30
파라미터 계산 유닛
32
리오더 유닛
34
전경 선택 유닛
36
공간-시간 보간 유닛
50
음장 분석 유닛
44
BG 선택 유닛
48
에너지 보상 유닛
38
계수 감소 유닛
46
양자화 유닛
52
비트스트림 생성 유닛
42
심리 음향 오디오 코더 유닛
40
벡터-기반 합성 유닛
27
방향-기반 합성 유닛
28
컨텐츠 분석 유닛
26

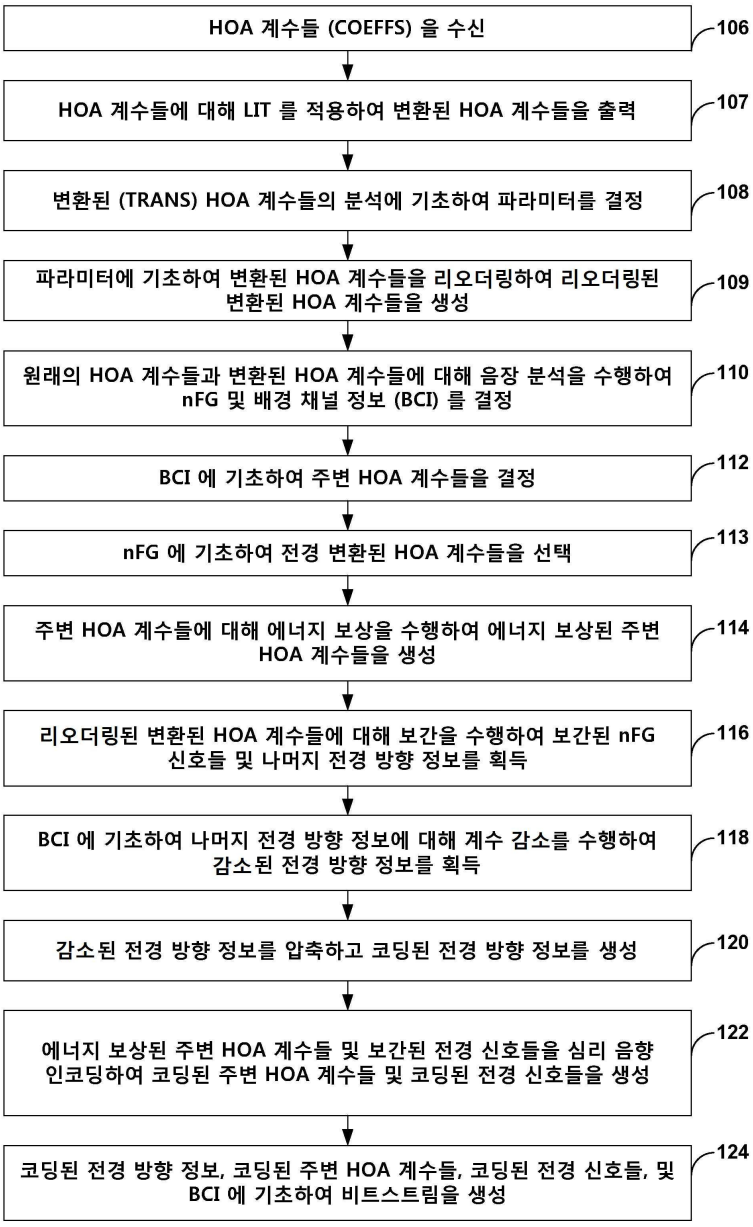
도면5


24
벡터-기반 복원 유닛
92
11'
HOA 계수
공식화 유닛
82
65
55_{k-1}'
공간-시간 보간 유닛
76
55_k"
전경 공식화 유닛
78
55_k
55_k'
리오더
유닛
84
49"
49'
47'
역양자화 유닛
74
심리 음향 디코딩
유닛
80
11'
방향-기반 복원 유닛
90
57
59
61
91
추출 유닛
72
21

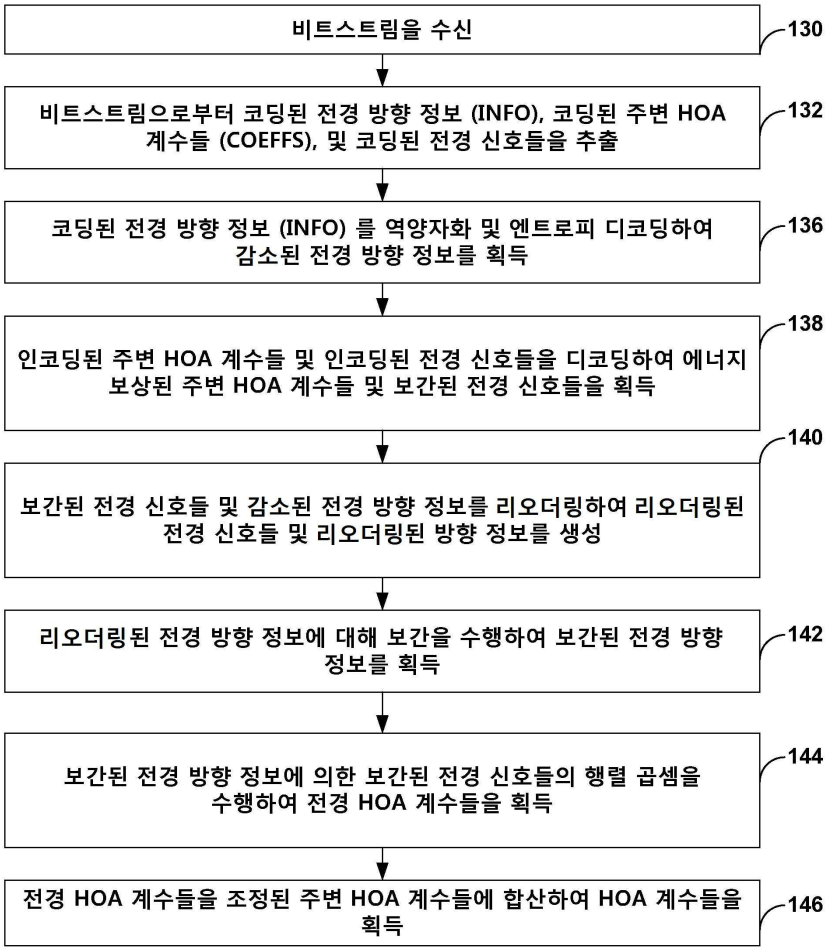
도면6


93
프레임화된 SHC 행렬을 획득
94
감소된 프레임화된 SHC 행렬을 획득하기 위해 벡터를 제외시킴
95
감소된 프레임화된 SHC 행렬의 나머지 벡터들로부터 감소된 프레임화된 SHC 행렬의 비-제로 벡터를 예측
96
예측된 비-제로 벡터와 비-제로 벡터에 기초하여 에러를 획득
97
제 1 비-제로 벡터 및 에러의 에너지에 기초하여 비율을 컴퓨팅
98
비율>임계치?
아니오
예
99
레코딩으로부터 생성된 프레임화된 SHC 행렬을 결정
100
레코딩으로부터 생성된 프레임화된 SHC 행렬을 비트스트림에 표시
101
프레임화된 SHC 행렬을 벡터-기반 합성 유닛으로 패스함
102
합성 오디오 오브젝트로부터 생성된 프레임화된 SHC 행렬을 결정
103
합성 오디오 오브젝트로부터 생성된 프레임화된 SHC 행렬을 비트스트림에 표시
104
프레임화된 SHC 행렬을 방향-기반 합성 유닛으로 패스함

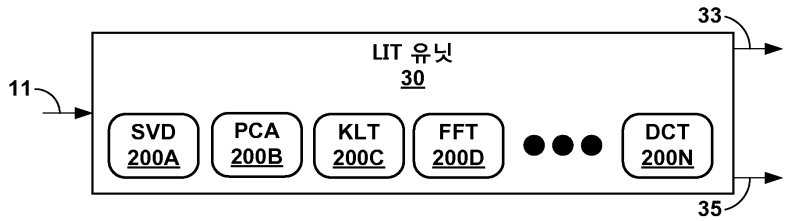
도면7


HOA 계수들 (COEFFS) 을 수신
106
HOA 계수들에 대해 LIT 를 적용하여 변환된 HOA 계수들을 출력
107
변환된 (TRANS) HOA 계수들의 분석에 기초하여 파라미터를 결정
108
파라미터에 기초하여 변환된 HOA 계수들을 리오더링하여 리오더링된 변환된 HOA 계수들을 생성
109
원래의 HOA 계수들과 변환된 HOA 계수들에 대해 음장 분석을 수행하여 nFG 및 배경 채널 정보 (BCI) 를 결정
110
BCI 에 기초하여 주변 HOA 계수들을 결정
112
nFG 에 기초하여 전경 변환된 HOA 계수들을 선택
113
주변 HOA 계수들에 대해 에너지 보상을 수행하여 에너지 보상된 주변 HOA 계수들을 생성
114
리오더링된 변환된 HOA 계수들에 대해 보간을 수행하여 보간된 nFG 신호들 및 나머지 전경 방향 정보를 획득
116
BCI 에 기초하여 나머지 전경 방향 정보에 대해 계수 감소를 수행하여 감소된 전경 방향 정보를 획득
118
감소된 전경 방향 정보를 압축하고 코딩된 전경 방향 정보를 생성
120
에너지 보상된 주변 HOA 계수들 및 보간된 전경 신호들을 심리 음향 인코딩하여 코딩된 주변 HOA 계수들 및 코딩된 전경 신호들을 생성
122
코딩된 전경 방향 정보, 코딩된 주변 HOA 계수들, 코딩된 전경 신호들, 및 BCI 에 기초하여 비트스트림을 생성
124

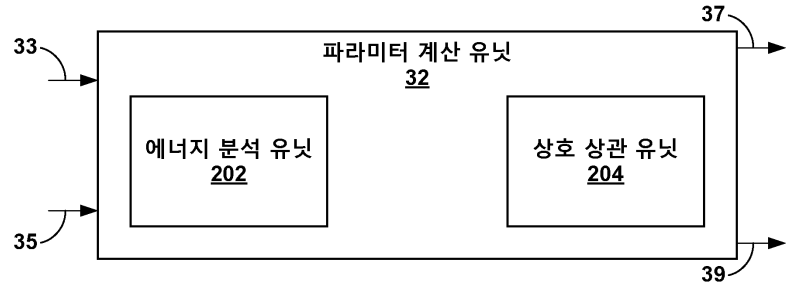
도면8



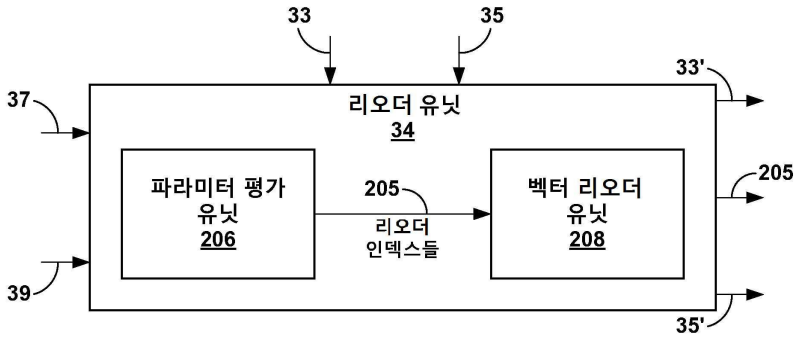
도면9a



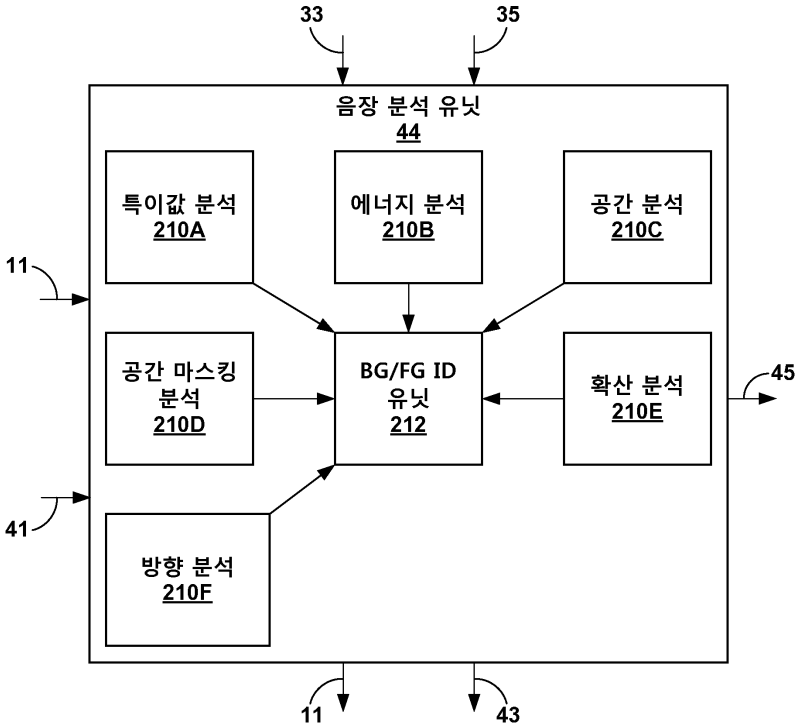
도면9b



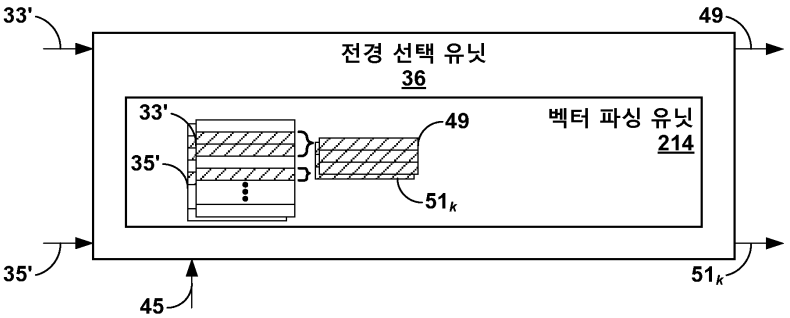
도면9c



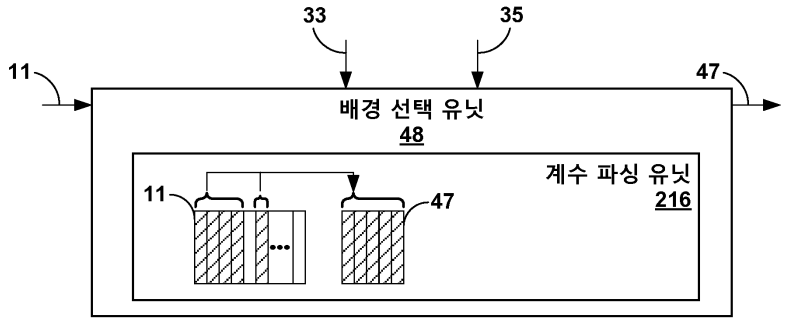
도면9d



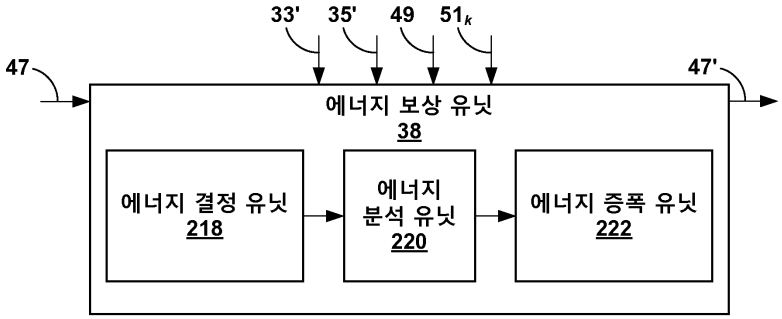
도면9e



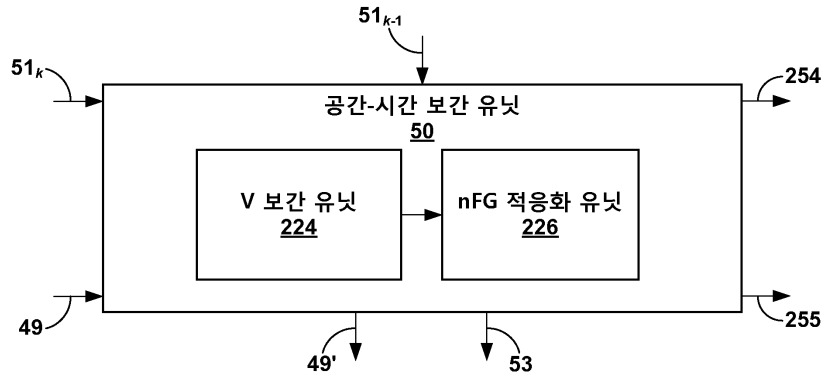
도면9f



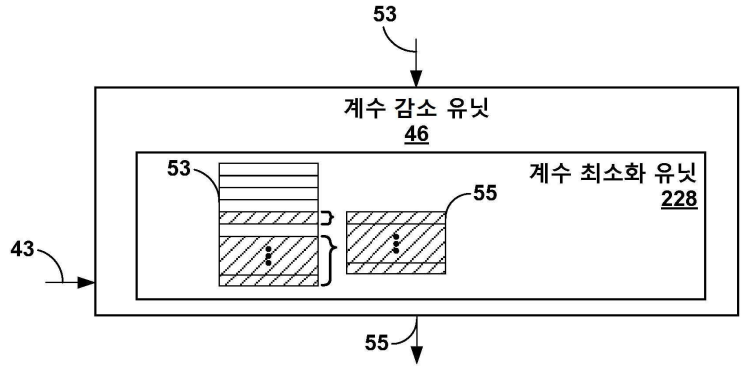
도면9g



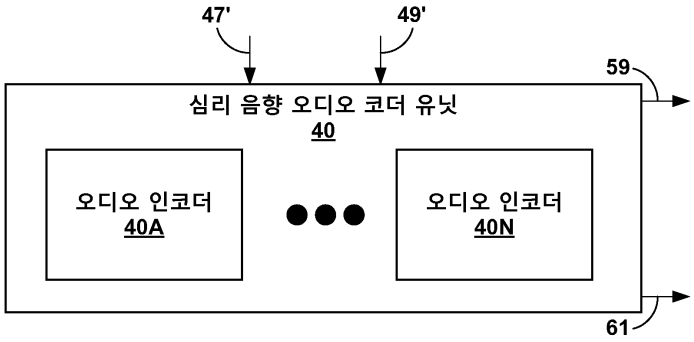
도면9h



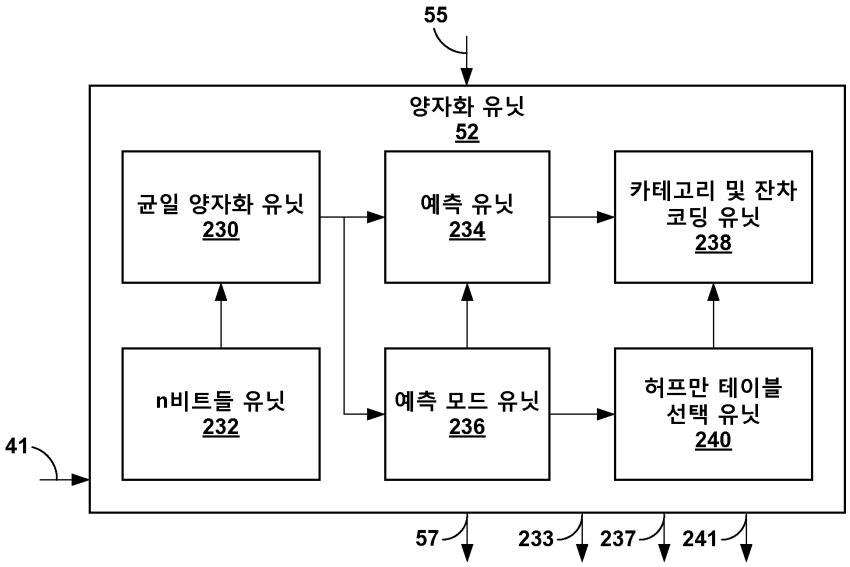
도면9i



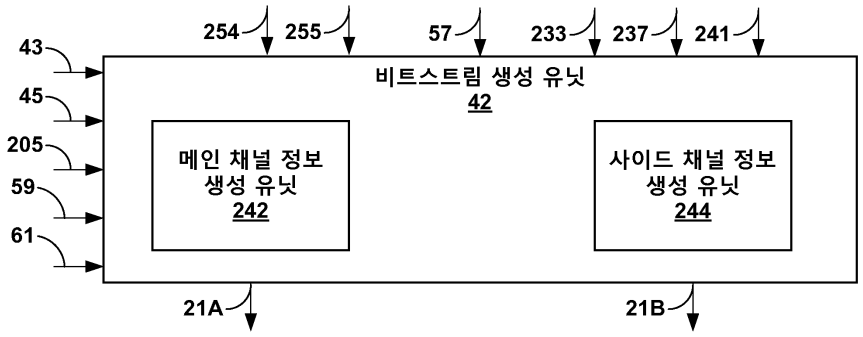
도면9j



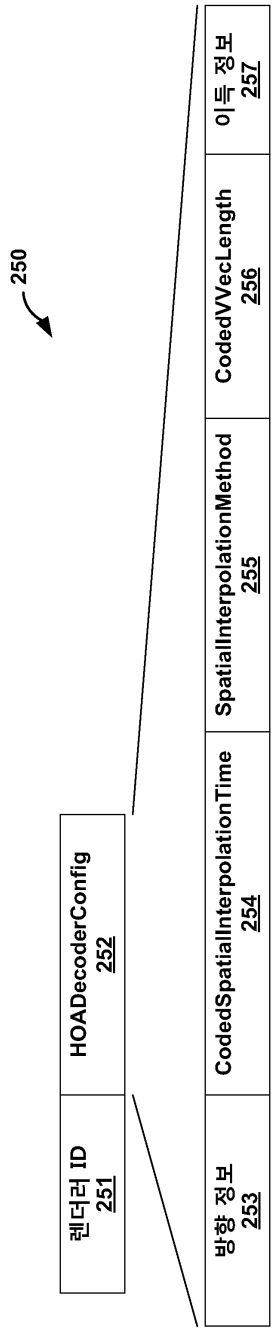
도면9k



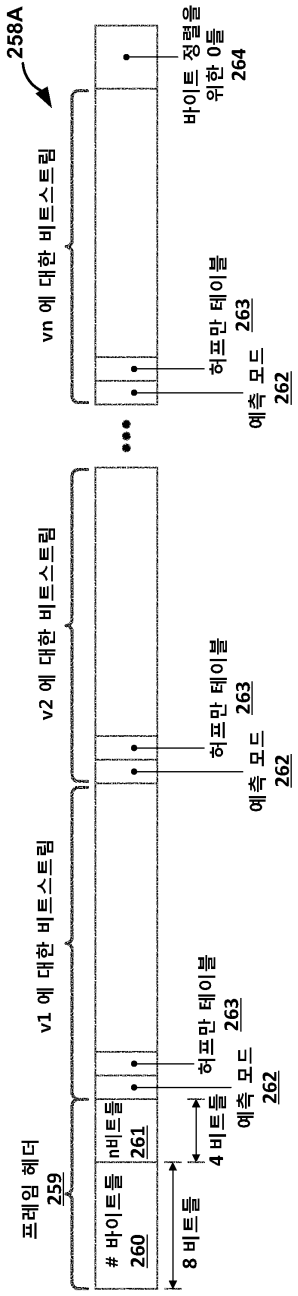
도면9l



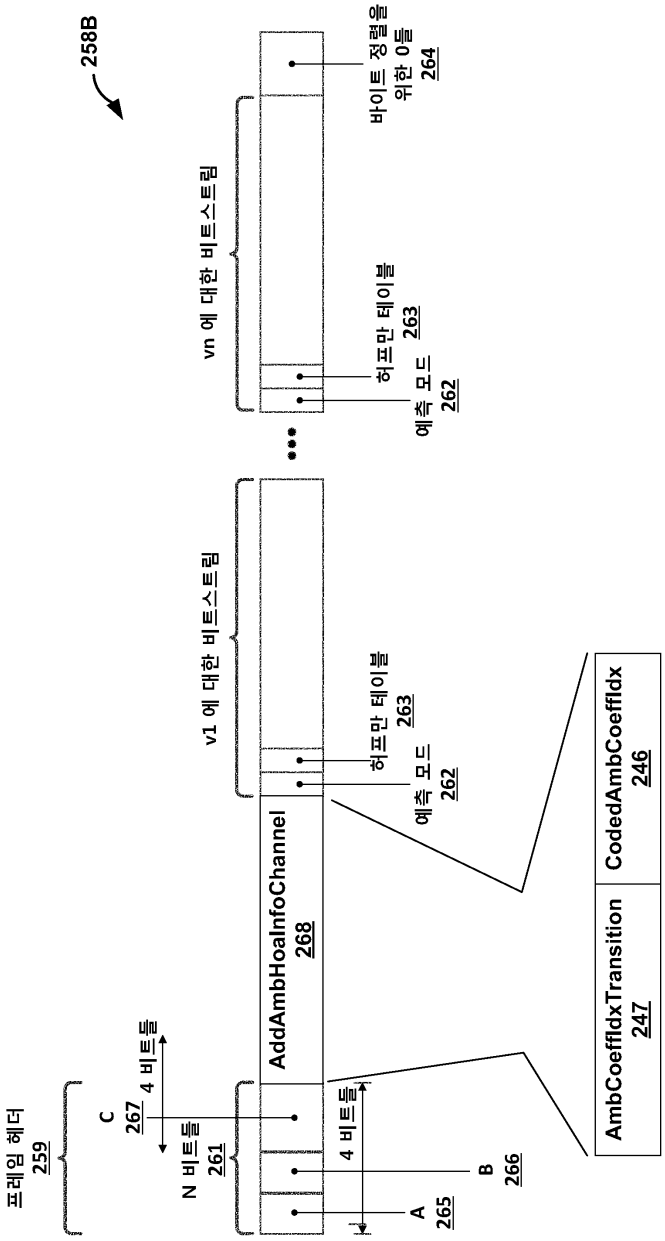
도면10a


250
렌더러 ID
251
HOADecoderConfig
252
방향 정보
253
CodedSpatialInterpolationTime
254
SpatialInterpolationMethod
255
CodedVVecLength
256
이득 정보
257

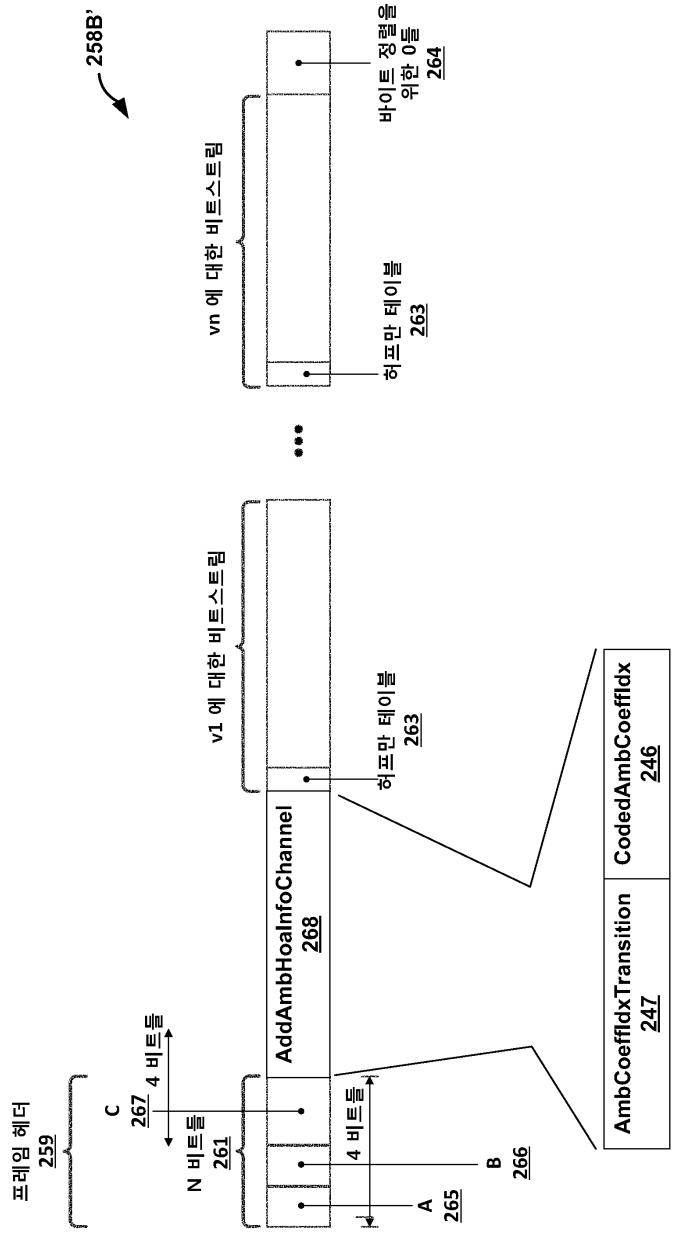
도면10b


258A
프레임 헤더
259
v1 에 대한 비트스트림
v2 에 대한 비트스트림
vn 에 대한 비트스트림
바이트들
260
n비트들
261
8 비트들
4 비트들
허프만 테이블
263
예측 모드
262
바이트 정렬을
위한 0들
264

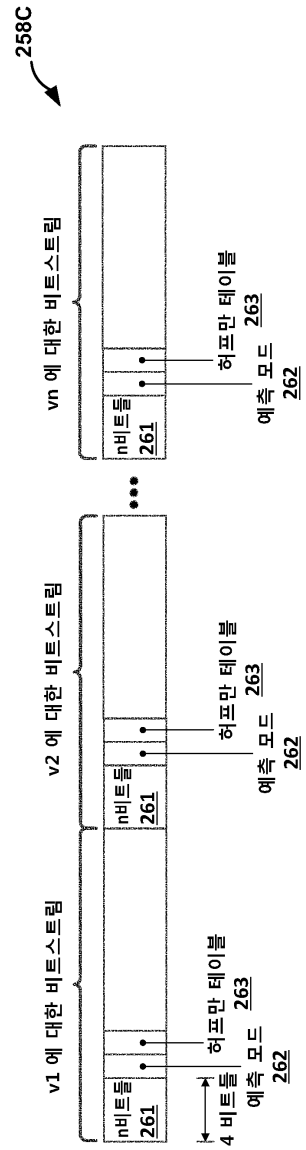
도면10c


258B
프레임 헤더
259
C
267
4 비트들
N 비트들
261
AddAmbHoaInfoChannel
268
4 비트들
A
265
B
266
v1 에 대한 비트스트림
허프만 테이블
263
예측 모드
262
vn 에 대한 비트스트림
허프만 테이블
263
예측 모드
262
바이트 정렬을
위한 0들
264
AmbCoeffIdxTransition
247
CodedAmbCoeffIdx
246

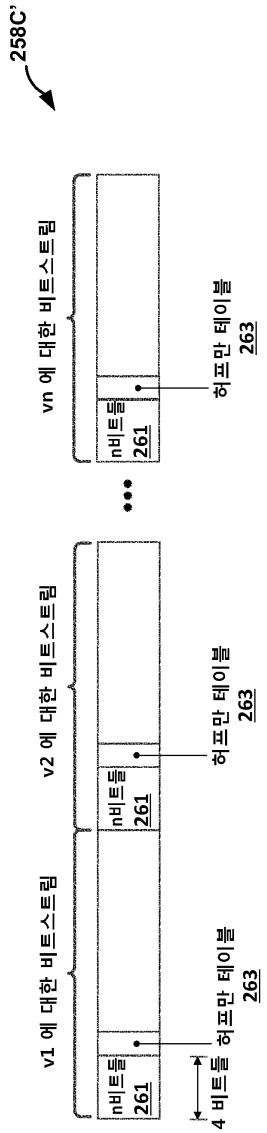
도면10ci


258B'
프레임 헤더
259
C
267
4 비트들
N 비트들
261
AddAmbHoaInfoChannel
268
4 비트들
A
265
B
266
v1 에 대한 비트스트림
허프만 테이블
263
vn 에 대한 비트스트림
허프만 테이블
263
바이트 정렬을 위한 0들
264
AmbCoeffIdxTransition
247
CodedAmbCoeffIdx
246

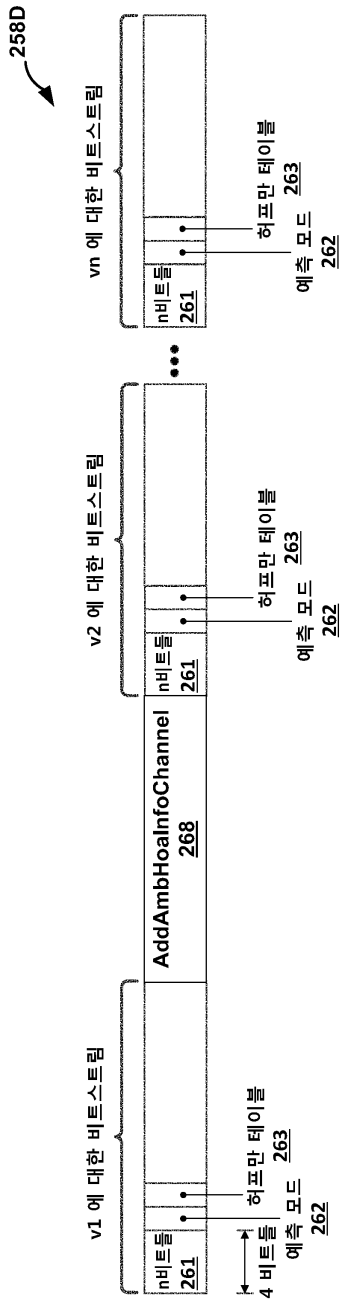
도면10d


258C
v1 에 대한 비트스트림
v2 에 대한 비트스트림
vn 에 대한 비트스트림
nbit들
261
허프만 테이블
263
예측 모드
262
4 비트들

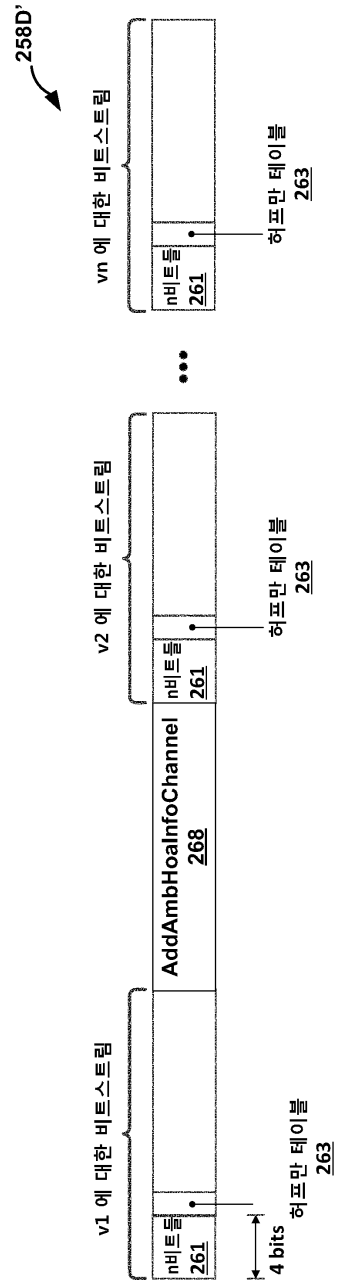
도면10di


258C'
v1 에 대한 비트스트림
v2 에 대한 비트스트림
vn 에 대한 비트스트림
nビ트들
261
허프만 테이블
263
4 비트들 허프만 테이블
263

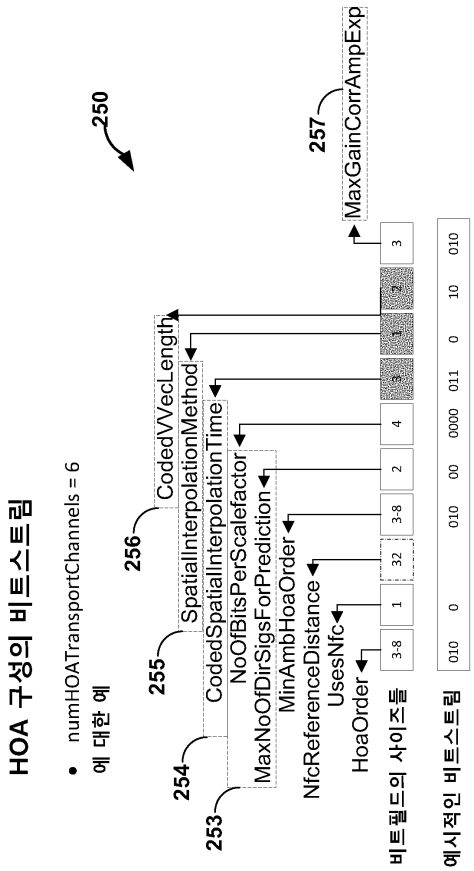
도면10e


258D
v1 에 대한 비트스트림
v2 에 대한 비트스트림
vn 에 대한 비트스트림
nBits
261
AddAmbHoaInfoChannel
268
허프만 테이블
263
예측 모드
262
4 비트들

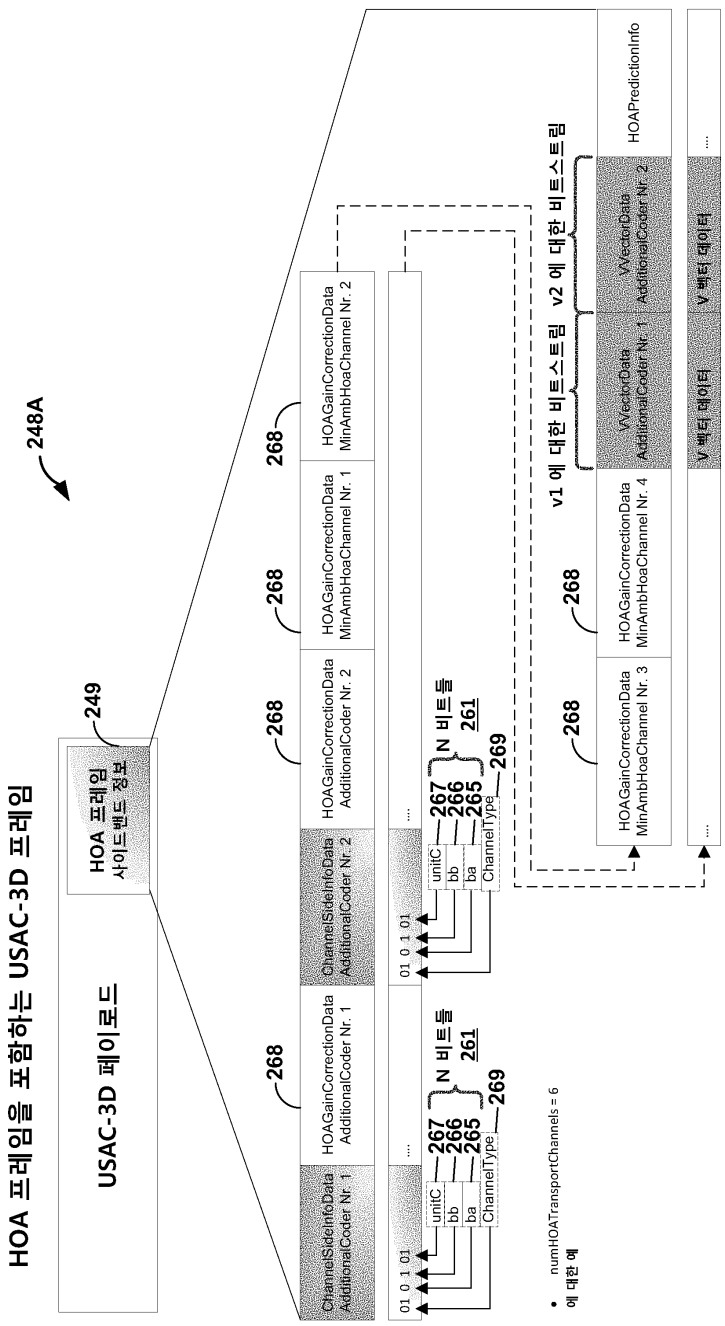
도면10ei


258D'
v1 에 대한 비트스트림
v2 에 대한 비트스트림
vn 에 대한 비트스트림
nbit들
261
AddAmbHoaInfoChannel
268
허프만 테이블
263
4 bits

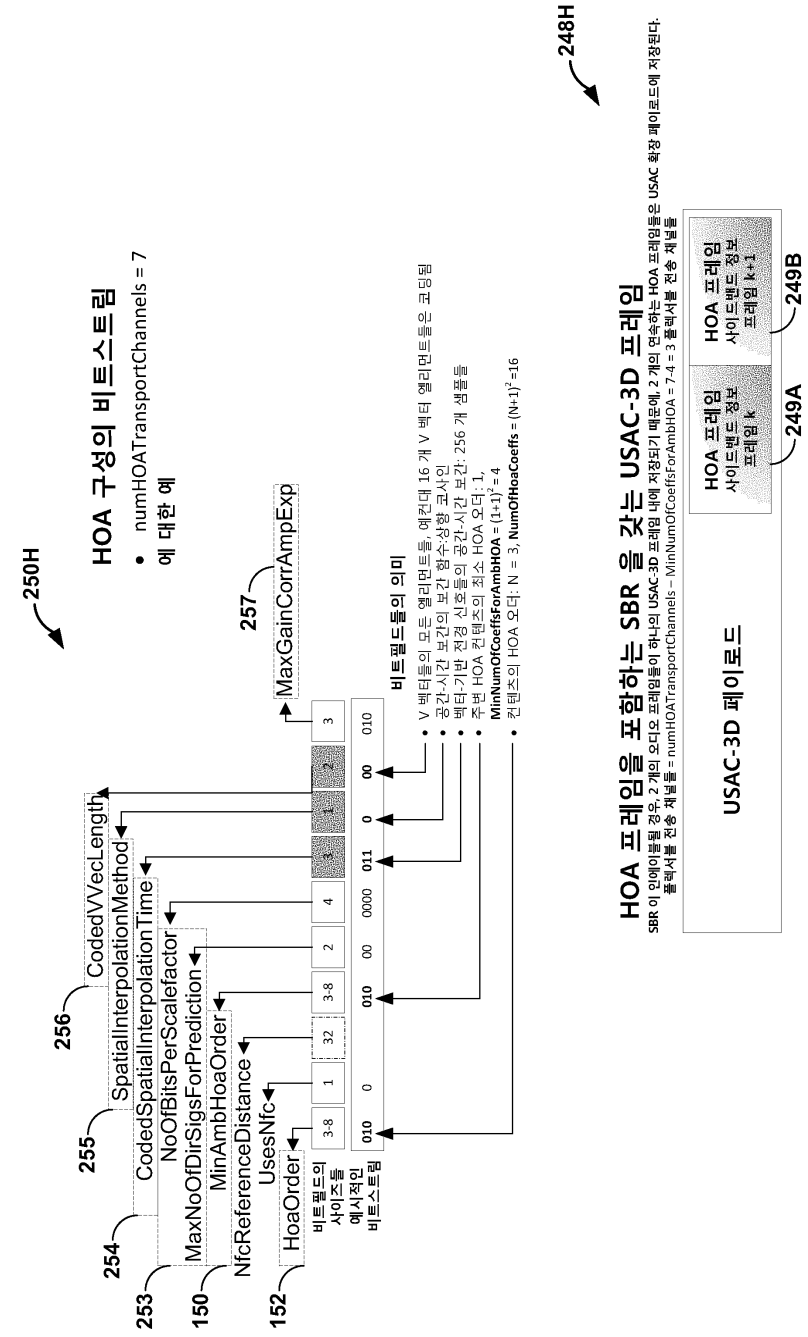
도면10f


HOA 구성의 비트스트림
numHOATransportChannels = 6
에 대한 예
250
256
255
254
253
257
CodedVecLength
SpatialInterpolationMethod
CodedSpatialInterpolationTime
NoOfBitsPerScalefactor
MaxNoOfDirSigsForPrediction
MinAmbHoaOrder
NfcReferenceDistance
UsesNfc
HoaOrder
MaxGainCorrAmpExp
비트필드의 사이즈들
3-8
1
32
3-8
2
4
3
1
2
3
예시적인 비트스트림
010
0
010
00
0000
011
0
10
010

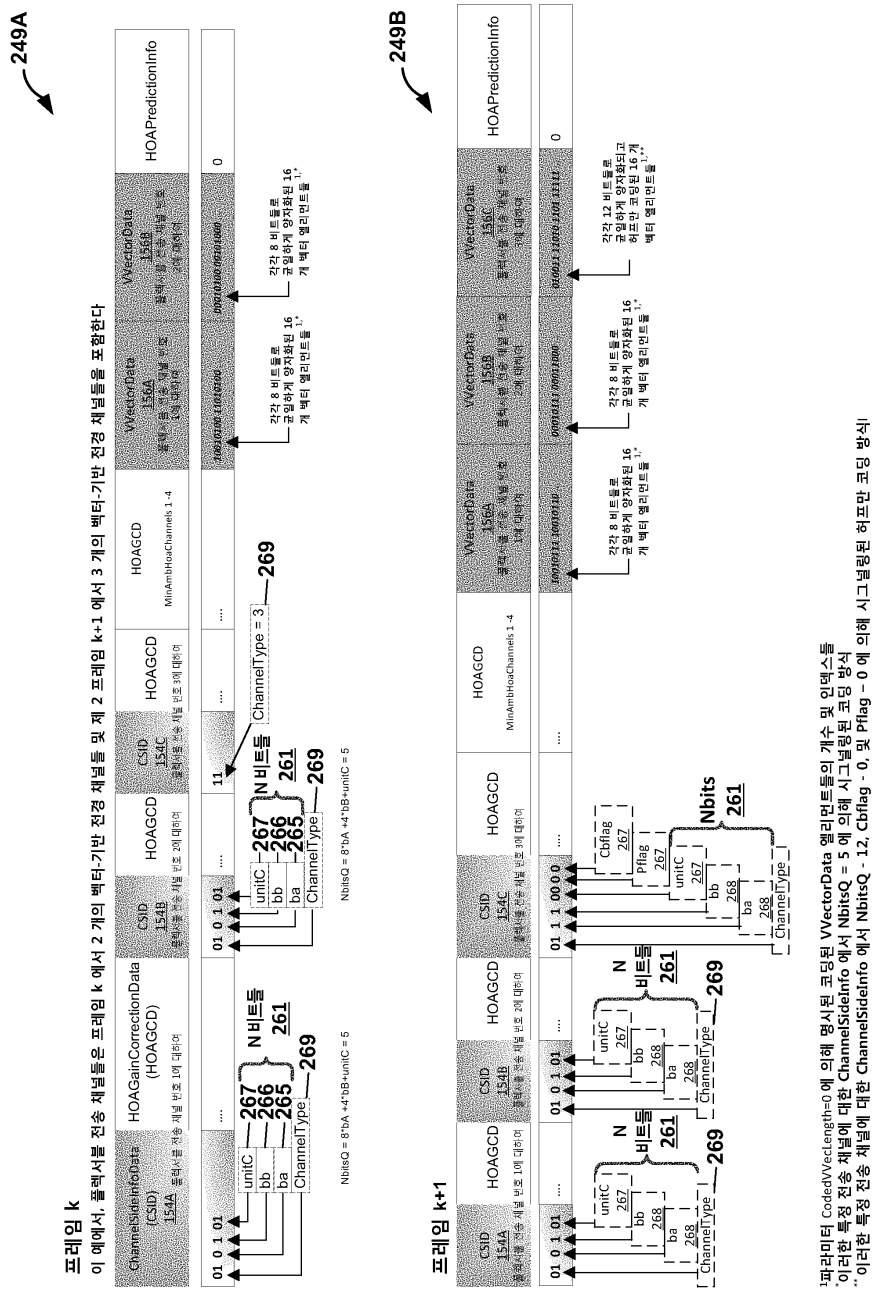
도면10g


HOA 프레임을 포함하는 USAC-3D 프레임
248A
USAC-3D 페이로드
249
HOA 프레임
사이드밴드 정보
268
HOAGainCorrectionData
AdditionalCoder Nr. 1
ChannelSideInfoData
AdditionalCoder Nr. 1
HOAGainCorrectionData
AdditionalCoder Nr. 2
ChannelSideInfoData
AdditionalCoder Nr. 2
HOAGainCorrectionData
MinAmbHoaChannel Nr. 1
HOAGainCorrectionData
MinAmbHoaChannel Nr. 2
HOAGainCorrectionData
MinAmbHoaChannel Nr. 3
HOAGainCorrectionData
MinAmbHoaChannel Nr. 4
VVectorData
AdditionalCoder Nr. 1
VVectorData
AdditionalCoder Nr. 2
HOAPredictionInfo
V 벡터 데이터
V 벡터 데이터
v1 에 대한 비트스트림
v2 에 대한 비트스트림
01 0 1 01
....
unitC
bb
ba
ChannelType
267
266
265
269
N 비트들
261
numHOATransportChannels = 6
에 대한 예

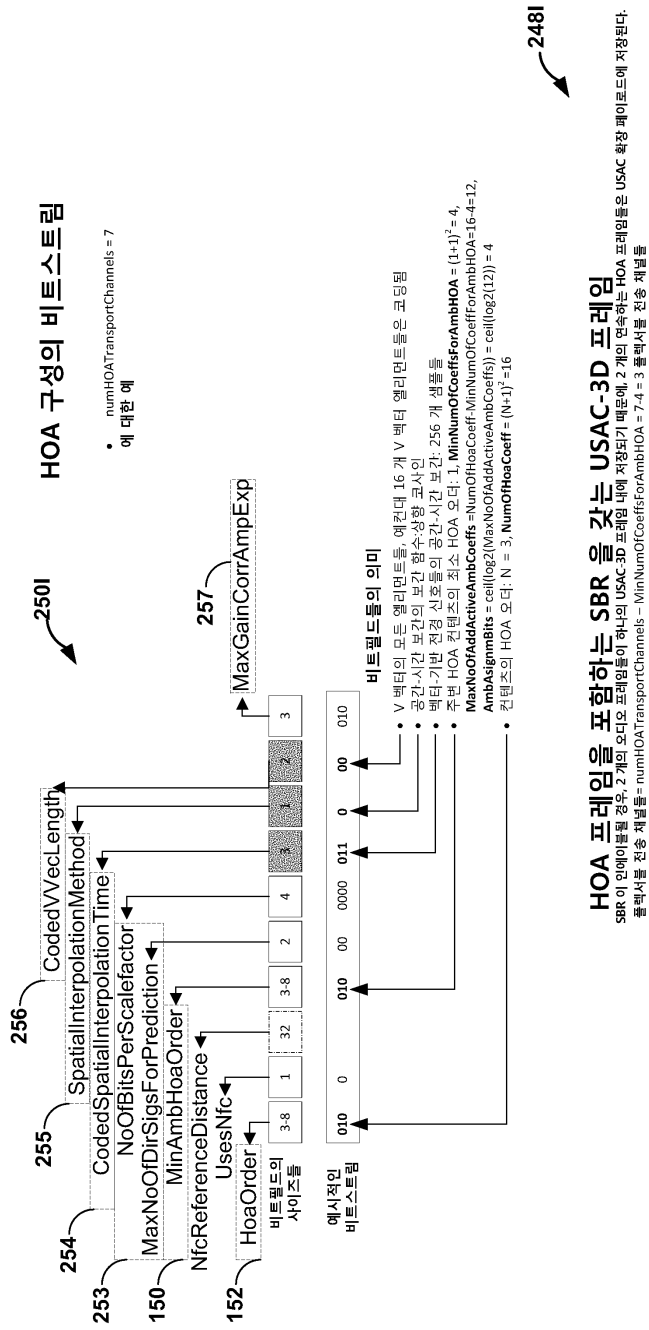
도면10hi


250H
HOA 구성의 비트스트림
• numHOATransportChannels = 7
에 대한 예
256
CodedVVecLength
255
SpatialInterpolationMethod
254
CodedSpatialInterpolationTime
253
NoOfBitsPerScalefactor
150
MaxNoOfDirSigsForPrediction
MinAmbHoaOrder
NfcReferenceDistance
UsesNfc
152
HoaOrder
257
MaxGainCorrAmpExp
비트필드의 사이즈들
예시적인 비트스트림
3-8
1
32
3-8
2
4
3
1
2
3
010
0
010
00
0000
011
0
00
010
비트필드들의 의미
• V 벡터들의 모드 엘리먼트들, 예컨대 16 개 V 벡터 엘리먼트들은 코딩됨
• 공간-시간 보간의 보간 함수-상향 코사인
• 벡터-기반 전경 신호들의 공간-시간 보간: 256 개 샘플들
• 주변 HOA 컨텐츠의 최소 HOA 오더: 1,
MinNumOfCoeffsForAmbHOA = (1+1)² = 4
• 컨텐츠의 HOA 오더: N = 3, NumOfHoaCoeffs = (N+1)²=16
248H
HOA 프레임을 포함하는 SBR 을 갖는 USAC-3D 프레임
SBR 이 있에이블될 경우, 2 개의 오디오 프레임들이 하나의 USAC-3D 프레임 내에 저장되기 때문에, 2 개의 연속하는 HOA 프레임들은 USAC 확장 페이로드에 저장된다.
플렉서블 전송 채널들 = numHOATransportChannels – MinNumOfCoeffsForAmbHOA = 7-4 = 3 플렉서블 전송 채널들
USAC-3D 페이로드
HOA 프레임
사이드밴드 정보
프레임 k
249A
HOA 프레임
사이드밴드 정보
프레임 k+1
249B

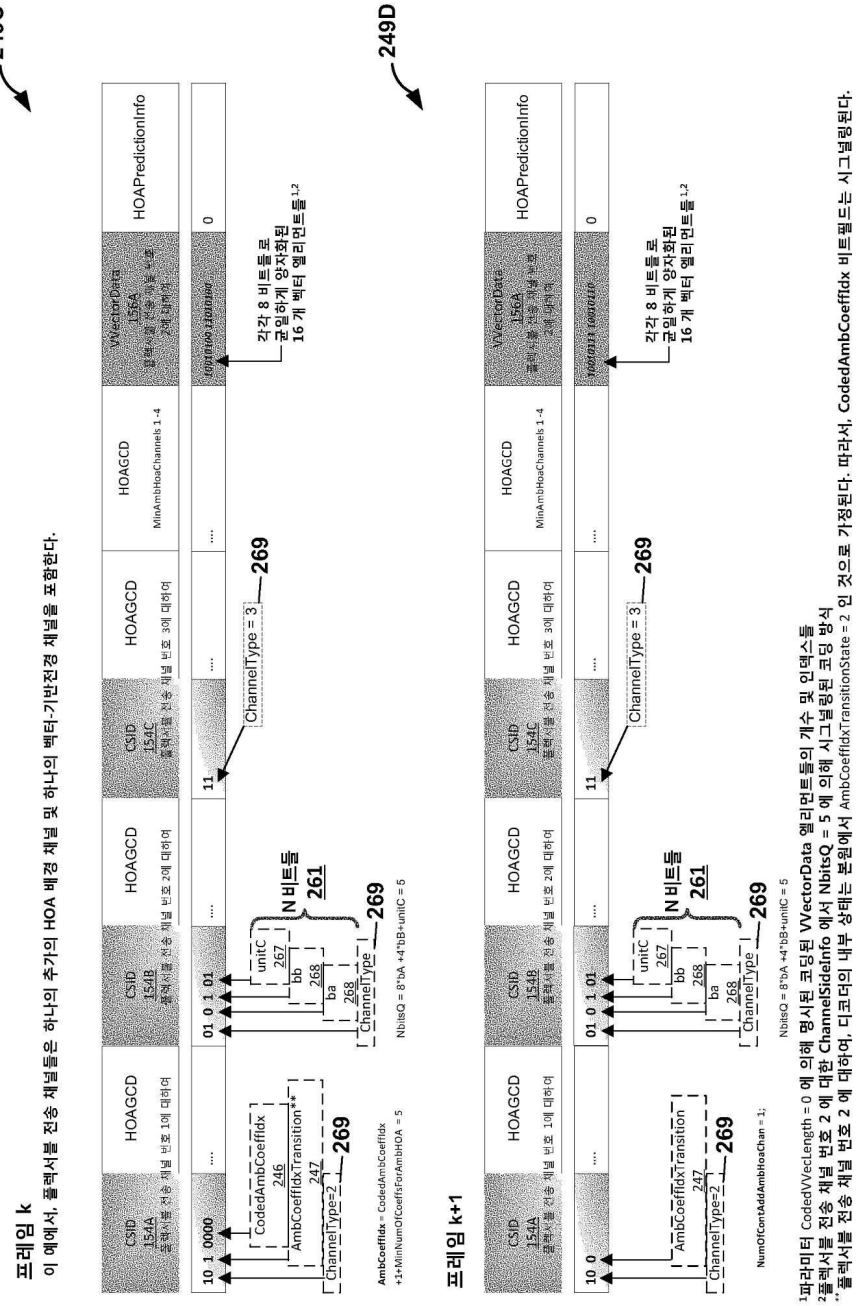
도면10hii


249A
프레임 k
이 예에서, 플렉서블 전송 채널들은 프레임 k 에서 2 개의 벡터-기반 전경 채널들 및 제 2 프레임 k+1 에서 3 개의 벡터-기반 전경 채널들을 포함한다
ChannelSideInfoData (CSID) 154A
HOAGainCorrectionData (HOAGCD) 154B
CSID 154C
HOAGCD MinAmbHoaChannels 1-4
VVectorData 156A
VVectorData 156B
VVectorData 156C
HOAPredictionInfo
unitC
bb
ba
ChannelType
N 비트들
261
269
267
266
265
ChannelType = 3
NbitsQ = 8*bA +4*bB+unitC = 5
각각 8 비트들로 균일하게 양자화된 16 개 벡터 엘리먼트들 1,*
249B
프레임 k+1
CbFlag
PFlag
Nbits
각각 12 비트들로 균일하게 양자화되고 허프만 코딩된 16 개 벡터 엘리먼트들 1,**
1파라미터 CodedVVecLength=0 에 의해 명시된 코딩된 VVectorData 엘리먼트들의 개수 및 인덱스들
*이러한 특정 전송 채널에 대한 ChannelSideInfo 에서 NbitsQ = 5 에 의해 시그널링된 코딩 방식
**이러한 특정 전송 채널에 대한 ChannelSideInfo 에서 NbitsQ - 12, CbFlag - 0, 및 PFlag - 0 에 의해 시그널링된 허프만 코딩 방식

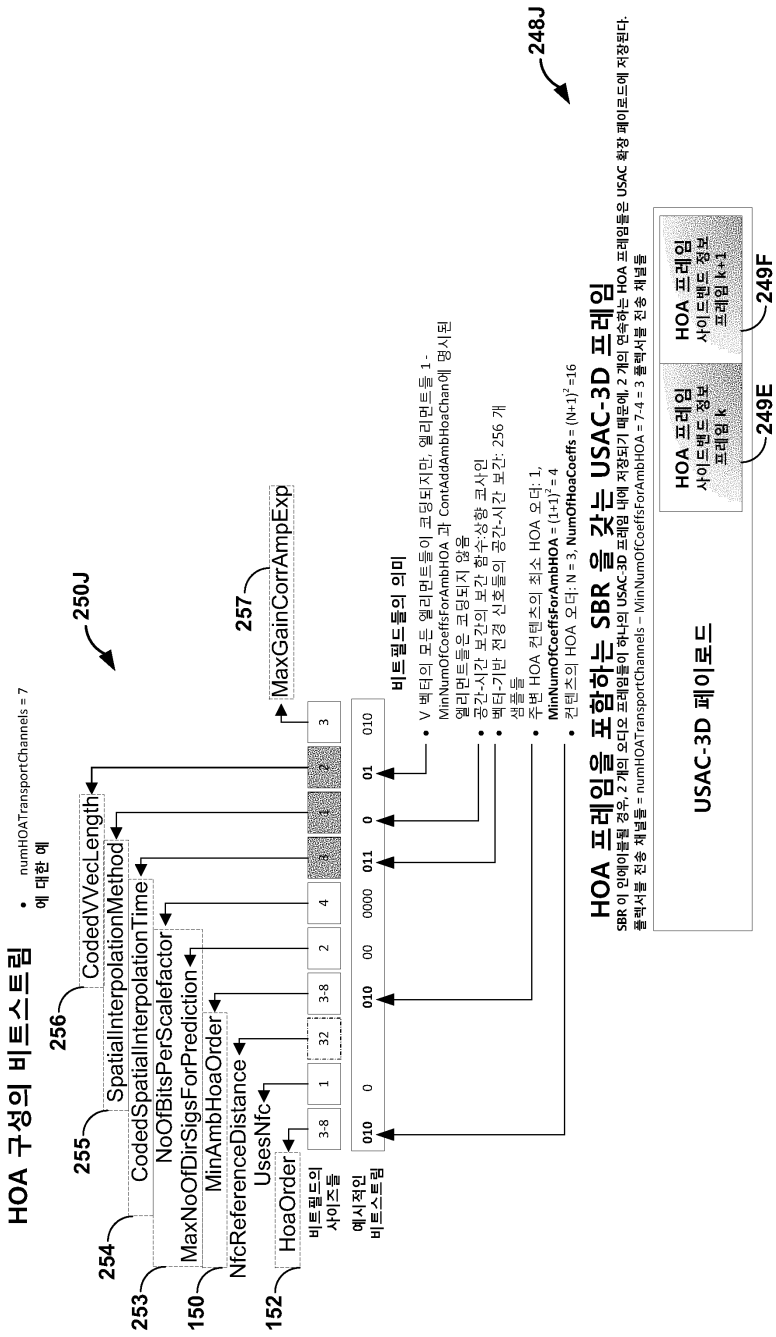
도면10ii


HOA 구성의 비트스트림
250I
numHOATransportChannels = 7
에 대한 예
256
CodedVVecLength
255
SpatialInterpolationMethod
CodedSpatialInterpolationTime
254
NoOfBitsPerScalefactor
253
MaxNoOfDirSigsForPrediction
150
MinAmbHoaOrder
NfcReferenceDistance
UsesNfc
152
HoaOrder
257
MaxGainCorrAmpExp
비트필드의 사이즈들
3-8
1
32
3-8
2
4
3
1
2
3
예시적인 비트스트림
010
0
010
00
0000
011
0
00
010
비트필드들의 의미
V 벡터의 모드 엘리먼트들, 예컨대 16 개 V 벡터 엘리먼트들은 코딩됨
공간-시간 보간의 보간 함수상항 교사인
벡터-기반 전경 신호들의 공간-시간 보간: 256 개 샘플들
주변 HOA 컨텐츠의 최소 HOA 오더: 1, MinNumOfCoeffsForAmbHOA = (1+1)2 = 4,
MaxNoOfAddActiveAmbCoeffs =NumOfHoaCoeff-MinNumOfCoeffForAmbHOA=16-4=12,
AmbAsignmBits = ceil((log2(MaxNoOfAddActiveAmbCoeffs)) = ceil(log2(12)) = 4
컨텐츠의 HOA 오더: N = 3, NumOfHoaCoeff = (N+1)2 =16
248I
HOA 프레임을 포함하는 SBR 을 갖는 USAC-3D 프레임
SBR 이 인에이블될 경우, 2 개의 오디오 프레임들이 하나의 USAC-3D 프레임 내에 저장되기 때문에, 2 개의 연속하는 HOA 프레임들은 USAC 확장 페이로드에 저장된다.
플렉서블 전송 채널들 = numHOATransportChannels – MinNumOfCoeffsForAmbHOA = 7-4 = 3 플렉서블 전송 채널들
USAC-3D 페이로드
HOA 프레임
사이드밴드 정보
프레임 k
249C
HOA 프레임
사이드밴드 정보
프레임 k+1
249D

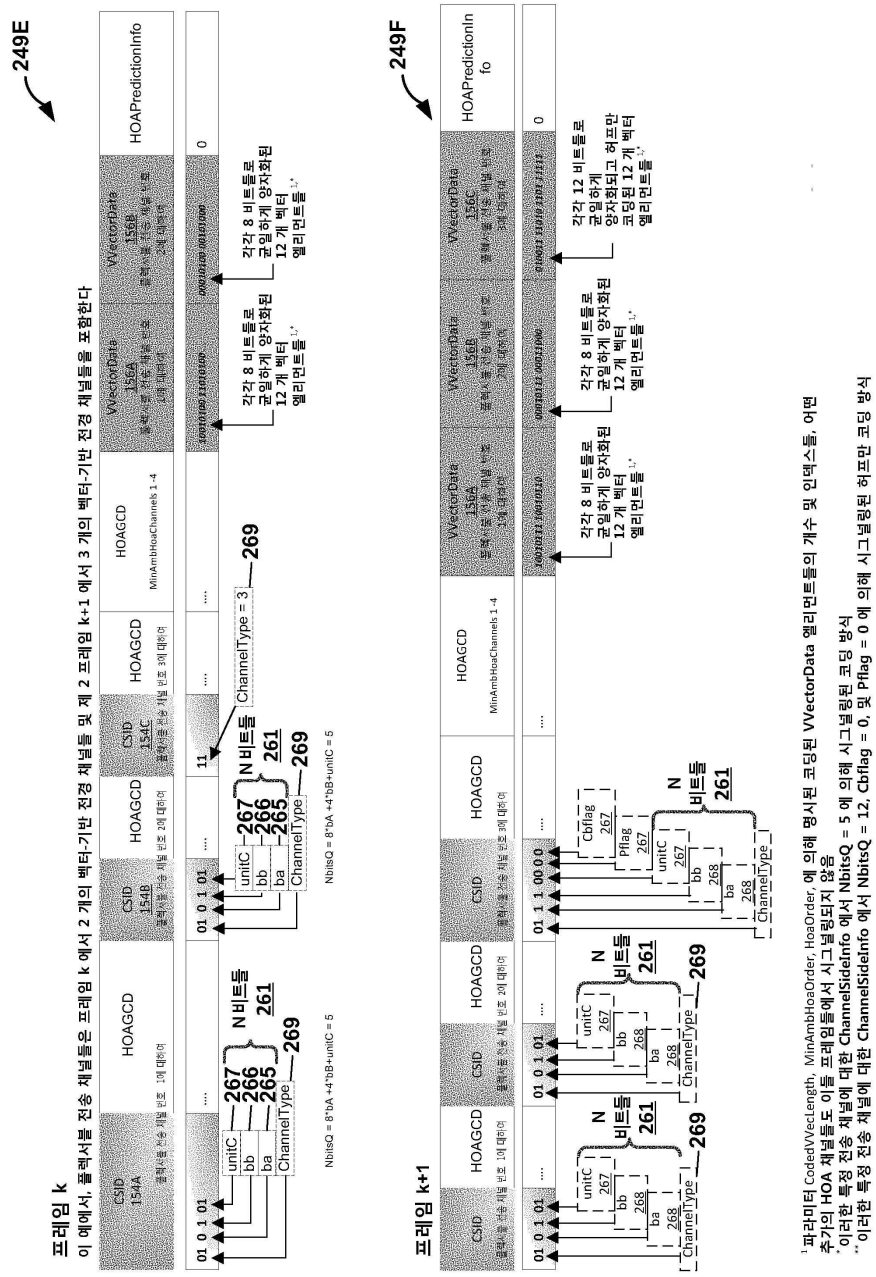
도면10iii


249C
프레임 k
이 예에서, 플렉서블 전송 채널들은 하나의 추가의 HOA 배경 채널 및 하나의 벡터-기반전경 채널을 포함한다.
CSID 154A 플렉서블 전송 채널 번호 1에 대하여
HOAGCD
CSID 154B 플렉서블 전송 채널 번호 2에 대하여
HOAGCD
CSID 154C 플렉서블 전송 채널 번호 3에 대하여
HOAGCD
MinAmbHoaChannels 1-4
VVectorData 156A 플렉서블 전송 채널 번호 3에 대하여
HOAPredictionInfo
10 1 0000
01 0 1 01
11
0
CodedAmbCoeffIdx 246
AmbCoeffIdxTransition* 247
ChannelType=2
269
unitC 267
bb 268
ba 268
N 비트들 261
ChannelType
269
ChannelType = 3
269
각각 8 비트들로 균일하게 양자화된 16 개 벡터 엘리먼트들1,2
AmbCoeffIdx = CodedAmbCoeffIdx +1+MinNumOfCoeffsForAmbHOA - 5
NbitsQ = 8*bA +4*bB+unitC - 5
프레임 k+1
10 0
01 0 1 01
11
0
AmbCoeffIdxTransition* 247
ChannelType=2
269
unitC 267
bb 268
ba 268
N 비트들 261
ChannelType
269
ChannelType = 3
269
각각 8 비트들로 균일하게 양자화된 16 개 벡터 엘리먼트들1,2
NumOfContAddAmbHoaChan = 1;
NbitsQ = 8*bA +4*bB+unitC - 5
249D
1파라미터 CodedVVecLength = 0 에 의해 명시된 코딩된 VVectorData 엘리먼트들의 개수 및 인덱스들
2플렉서블 전송 채널 번호 2 에 대한 ChannelSideInfo 에서 NbitsQ = 5 에 의해 시그널링된 코딩 방식
*플렉서블 전송 채널 번호 2 에 대하여, 디코더의 내부 상태는 본원에서 AmbCoeffIdxTransitionState = 2 인 것으로 가정된다. 따라서, CodedAmbCoeffIdx 비트필드는 시그널링된다.

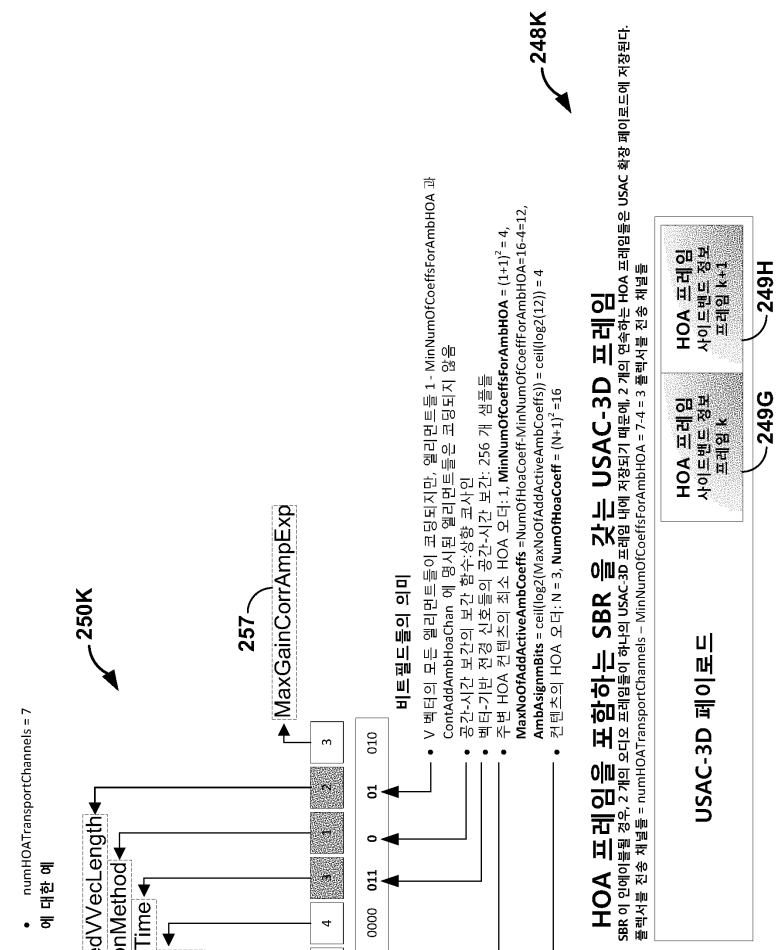
도면10ji


HOA 구성의 비트스트림
numHOATransportChannels = 7
에 대한 예
250J
256
255
254
253
150
152
257
CodedVecLength
SpatialInterpolationMethod
CodedSpatialInterpolationTime
NoOfBitsPerScalefactor
MaxNoOfDirSigsForPrediction
MinAmbHoaOrder
NfcReferenceDistance
UsesNfc
HoaOrder
MaxGainCorrAmpExp
비트필드의 사이즈들
예시적인 비트스트림
3-8
1
32
3-8
2
4
3
1
2
3
010
0
010
00
0000
011
0
01
010
비트필드들의 의미
• V 벡터의 모든 엘리먼트들이 코딩되지만, 엘리먼트들 1-MinNumOfCoeffsForAmbHOA 과 ContAddAmbHoaChan에 명시된 엘리먼트들은 코딩되지 않음
• 공간-시간 보간의 보간 함수상황 코사인
• 벡터-기반 전경 신호들의 공간-시간 보간: 256 개 샘플들
• 주변 HOA 컨텐츠의 최소 HOA 오더: 1, MinNumOfCoeffsForAmbHOA = (1+1)2 = 4
• 컨텐츠의 HOA 오더: N = 3, NumOfHoaCoeffs = (N+1)2 =16
248J
HOA 프레임을 포함하는 SBR 을 갖는 USAC-3D 프레임
SBR 이 안에이블될 경우, 2 개의 오디오 프레임들이 하나의 USAC-3D 프레임 내에 저장되기 때문에, 2 개의 연속하는 HOA 프레임들은 USAC 확장 페이로드에 저장된다.
플렉서블 전송 채널들 = numHOATransportChannels – MinNumOfCoeffsForAmbHOA = 7-4 = 3 플렉서블 전송 채널들
USAC-3D 페이로드
HOA 프레임
사이드밴드 정보
프레임 k
249E
HOA 프레임
사이드밴드 정보
프레임 k+1
249F

도면10jii


249E
프레임 k
이 예에서, 플렉서블 전송 채널들은 프레임 k 에서 2 개의 벡터-기반 전경 채널들 및 제 2 프레임 k+1 에서 3 개의 벡터-기반 전경 채널들을 포함한다
CSID 154A
HOAGCD
CSID 154B
HOAGCD
CSID 154C
HOAGCD
HOAPredictionInfo
VVectorData 156A
VVectorData 156B
VVectorData 156C
unitC
bb
ba
ChannelType
N비트들
261
267
266
265
269
ChannelType = 3
NbitsQ = 8*bA +4*bB+unitC = 5
각각 8 비트들로 균일하게 양자화된 12 개 벡터 엘리먼트들 1,*
249F
프레임 k+1
Cbflag
Pflag
각각 12 비트들로 균일하게 양자화되고 허프만 코딩된 12 개 벡터 엘리먼트들 1,*
1 파라미터 CodedVVecLength, MinAmbHoaOrder, HoaOrder 에 의해 명시된 코딩된 VVectorData 엘리먼트들의 개수 및 인덱스들, 여편
* 추가의 HOA 채널들도 이들 프레임들에서 시그널링되지 않음
* 이러한 특정 전송 채널에 대한 ChannelSideInfo 에서 NbitsQ = 5 에 의해 시그널링된 코딩 방식
** 이러한 특정 전송 채널에 대한 ChannelSideInfo 에서 NbitsQ = 12, Cbflag = 0, 및 Pflag = 0 에 의해 시그널링된 허프만 코딩 방식

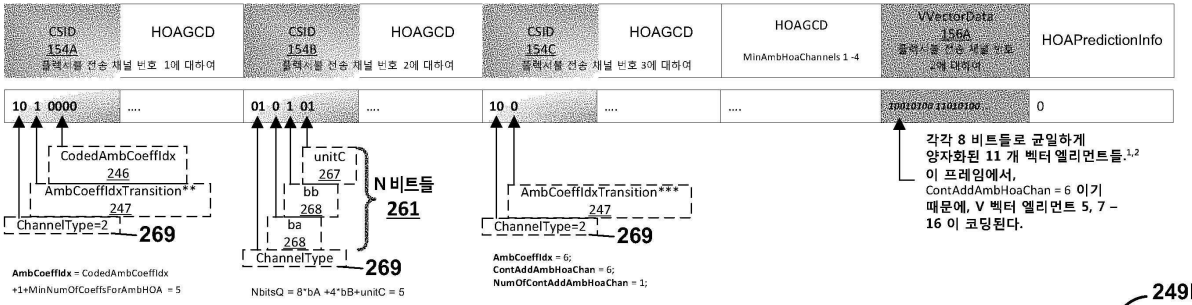
도면10ki


HOA 구성의 비트스트림
numHOATransportChannels = 7 에 대한 예
250K
256
255
254
253
150
152
CodedVVecLength
SpatialInterpolationMethod
CodedSpatialInterpolationTime
NoOfBitsPerScalefactor
MaxNoOfDirSigsForPrediction
MinAmbHoaOrder
NfcReferenceDistance
UsesNfc
HoaOrder
257
MaxGainCorrAmpExp
비트필드의 사이즈들
예시적인 비트스트림
비트필드들의 의미
V 벡터의 모든 엘리먼트들이 코딩되지만, 엘리먼트들 1 - MinNumOfCoeffsForAmbHOA 과
ContAddAmbHoaChan 에 명시된 엘리먼트들은 코딩되지 않음
공간-시간 보간의 보간 함수:상향 코사인
벡터-기반 전경 신호들의 공간-시간 보간: 256 개 샘플들
주변 HOA 컨텐츠의 최소 HOA 오더: 1, MinNumOfCoeffsForAmbHOA = (1+1)² = 4,
MaxNoOfAddActiveAmbCoeffs =NumOfHoaCoeff-MinNumOfCoeffForAmbHOA=16-4=12,
AmbAsignmBits = ceil(log2(MaxNoOfAddActiveAmbCoeffs)) = ceil(log2(12)) = 4
컨텐츠의 HOA 오더: N = 3, NumOfHoaCoeff = (N+1)² =16
248K
HOA 프레임을 포함하는 SBR 을 갖는 USAC-3D 프레임
SBR 이 인에이블될 경우, 2 개의 오디오 프레임들이 하나의 USAC-3D 프레임 내에 저장되기 때문에, 2 개의 연속하는 HOA 프레임들은 USAC 확장 페이로드에 저장된다.
룸렉서블 전송 채널들 = numHOATransportChannels – MinNumOfCoeffsForAmbHOA = 7-4 = 3 룸렉서블 전송 채널들
USAC-3D 페이로드
HOA 프레임
사이드밴드 정보
프레임 k
HOA 프레임
사이드밴드 정보
프레임 k+1
249G
249H

249G

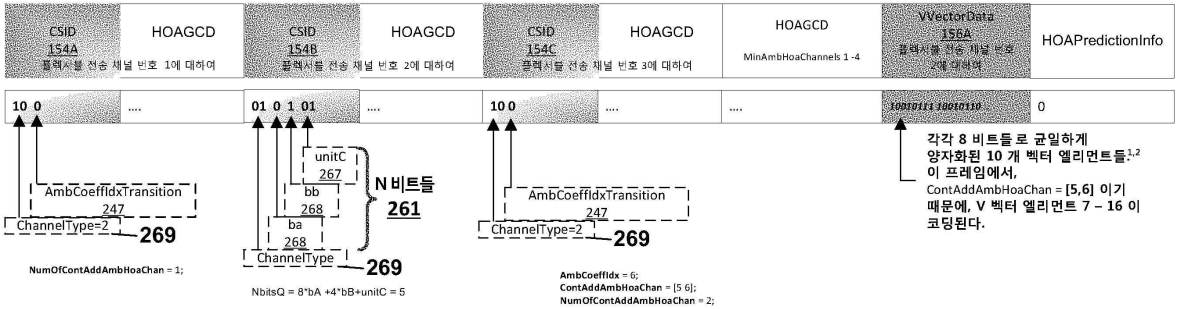
프레임 k

이 예에서, 플렉서블 전송 채널들은 하나의 추가의 HOA 배경 채널 및 하나의 벡터-기반전경 채널을 포함한다.



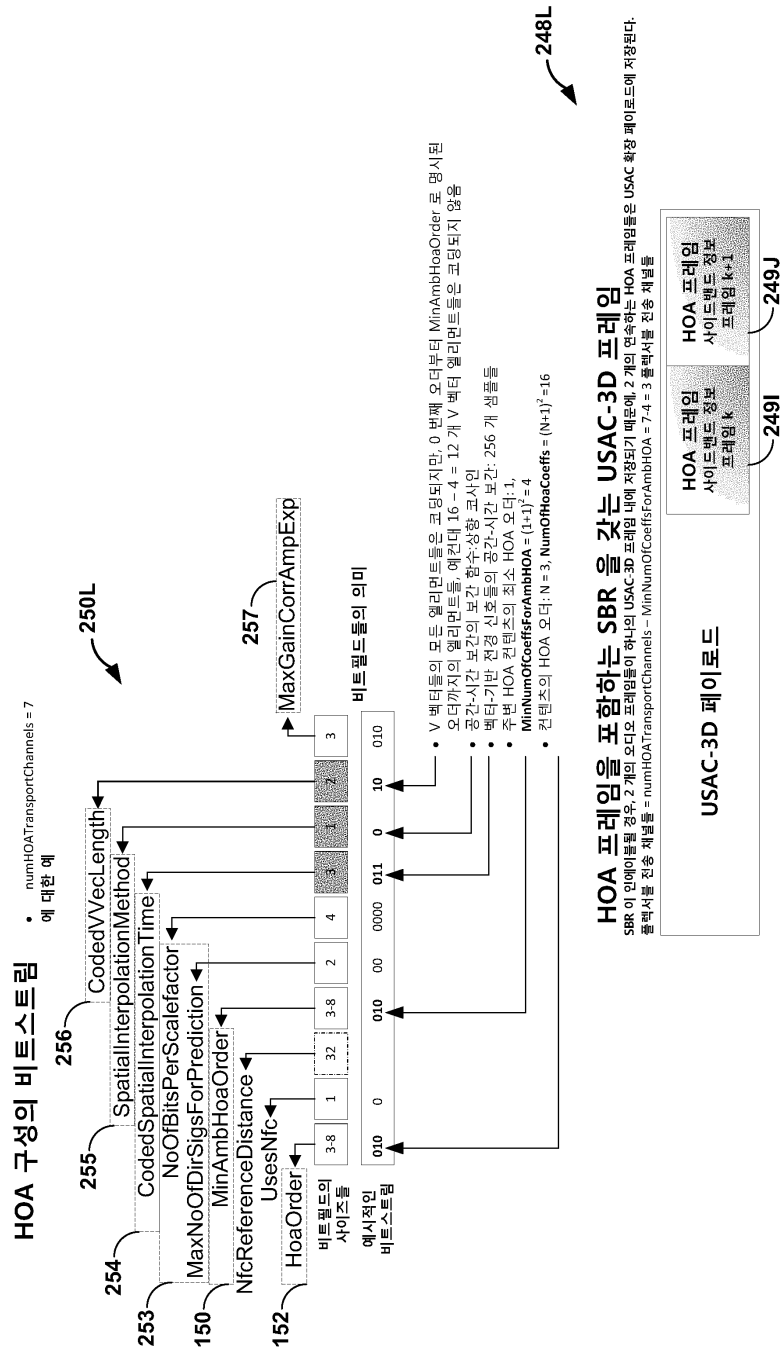
249H

프레임 k+1

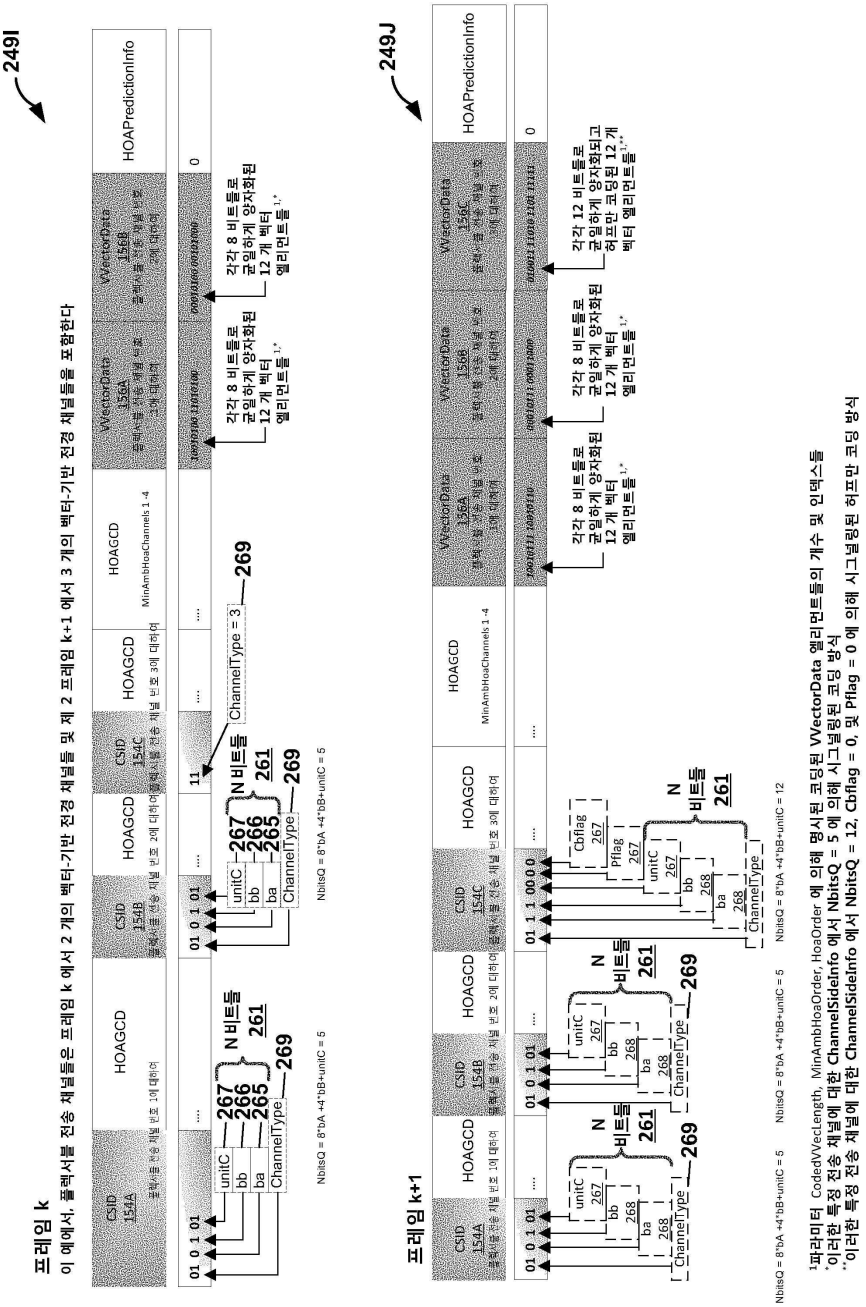


[1]파라미터 CodedVVecLength, MinNumOfCoeffsForAmbHOA , NumOfHoaCoeff 에 의해 명시된 코딩된 VVectorData 엘리먼트들의 개수 및 인덱스들
[2]플렉서블 전송 채널 번호 2 에 대한 ChannelSideInfo 에서 NbitsQ = 5 에 의해 시그널링된 코딩 방식
**플렉서블 전송 채널 번호 1 에 대하여, 디코더의 내부 상태는 본원에서 AmbCoeffIdxTransitionState = 2인 것으로 가정된다. 따라서, CodedAmbCoeffIdx 비트필드는 시그널링된다.
***플렉서블 전송 채널 번호 3 은 추가의 HOA 계수 6 을 운반하는 것으로 이전 프레임에 정의되었다. 따라서, CodedAmbCoeffIdx 비트필드는 시그널링되지 않는다.

도면101i

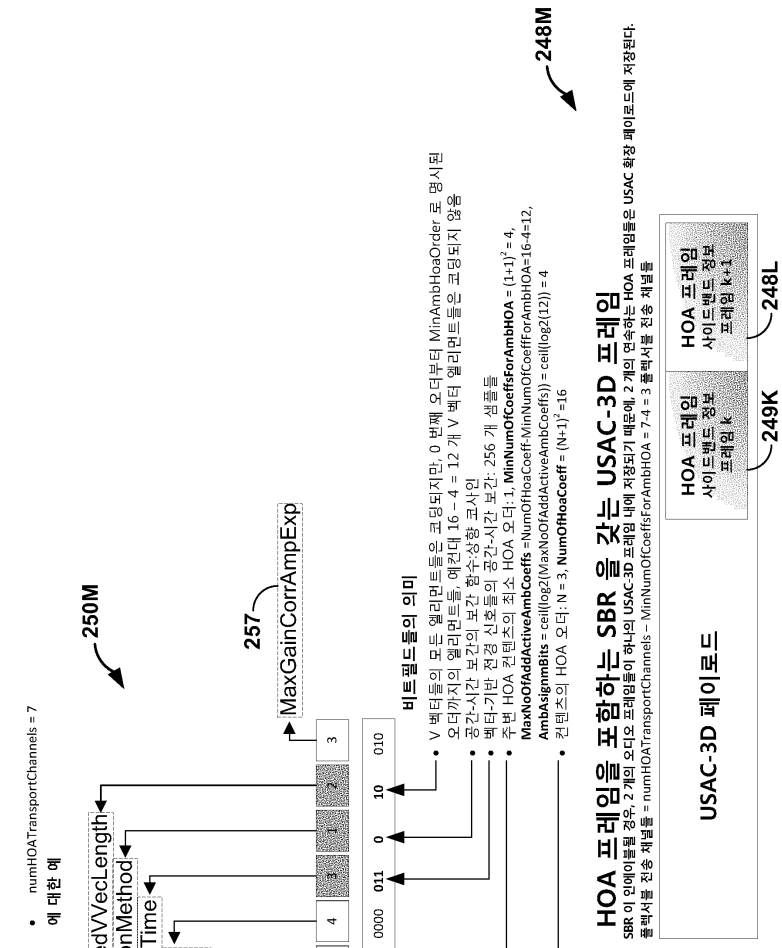

HOA 구성의 비트스트림
• numHOATransportChannels = 7 에 대한 예
250L
256
CodedVVecLength
255
SpatialInterpolationMethod
254
CodedSpatialInterpolationTime
253
NoOfBitsPerScalefactor
150
MaxNoOfDirSigsForPrediction
MinAmbHoaOrder
NfcReferenceDistance
UsesNfc
152
HoaOrder
257
MaxGainCorrAmpExp
비트필드의 사이즈들
예시적인 비트스트림
비트필드들의 의미
• V 벡터들의 모든 엘리먼트들은 코딩되지만, 0 번째 오더부터 MinAmbHoaOrder 로 명시된 오더까지의 엘리먼트들 예컨대 16 – 4 = 12 개 V 벡터 엘리먼트들은 코딩되지 않음
• 공간-시간 보간의 보간 함수상향 코사인
• 벡터-기반 전경 신호들의 공간-시간 보간: 256 개 샘플들
• 주변 HOA 컨텐츠의 최소 HOA 오더: 1,
MinNumOfCoeffsForAmbHOA = (1+1)² = 4
• 컨텐츠의 HOA 오더: N = 3, NumOfHoaCoeffs = (N+1)² =16
248L
HOA 프레임을 포함하는 SBR 을 갖는 USAC-3D 프레임
SBR 이 인에이블될 경우, 2 개의 오디오 프레임들이 하나의 USAC-3D 프레임 내에 저장되기 때문에, 2 개의 연속하는 HOA 프레임들은 USAC 확장 페이로드에 저장된다.
플렉서블 전송 채널들 = numHOATransportChannels – MinNumOfCoeffsForAmbHOA = 7-4 = 3 플렉서블 전송 채널들
USAC-3D 페이로드
HOA 프레임
사이드밴드 정보
프레임 k
249I
HOA 프레임
사이드밴드 정보
프레임 k+1
249J

도면10iii



[1]파라미터 CodedVecLength, MinAmbHoaOrder, HoaOrder 에 의해 형식된 코딩된 VVectorData 엘리먼트들의 개수 및 인덱스들
*이러한 특정 전송 채널에 대한 ChannelSideInfo 에서 NbitsQ = 5 에 의해 시그널링된 코딩 방식
**이러한 특정 전송 채널에 대한 ChannelSideInfo 에서 NbitsQ = 12, CbFlag = 0, 및 Pflag = 0 에 의해 시그널링된 허프만 코딩 방식

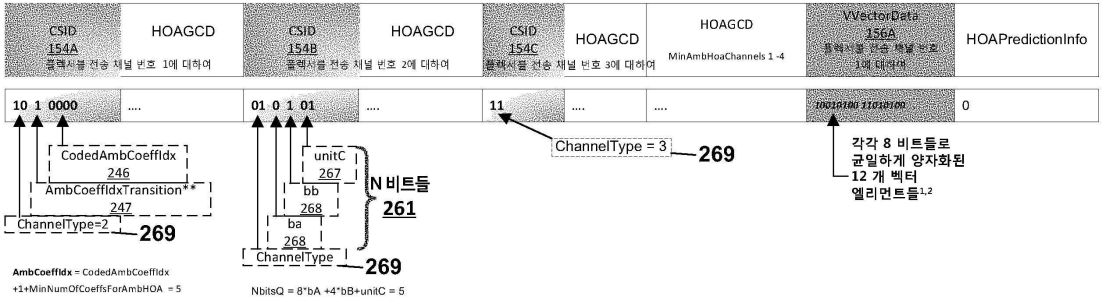
도면10mi


HOA 구성의 비트스트림
• numHOATransportChannels = 7 에 대한 예
250M
256
255
254
253
150
152
CodedVecLength
SpatialInterpolationMethod
CodedSpatialInterpolationTime
NoOfBitsPerScalefactor
MaxNoOfDirSigsForPrediction
MinAmbHoaOrder
NfcReferenceDistance
UsesNfc
HoaOrder
257
MaxGainCorrAmpExp
비트필드의 사이즈들
예시적인 비트스트림
3-8
1
32
3-8
2
4
3
1
2
3
010
0
010
00
0000
011
0
10
010
비트필드들의 의미
• V 벡터들의 모든 엘리먼트들은 코딩되지만, 0 번째 오더부터 MinAmbHoaOrder 로 명시된 오더까지의 엘리먼트들, 예컨대 16 – 4 = 12 개 V 벡터 엘리먼트들은 코딩되지 않음
• 공간-시간 보간의 보간 함수:상향 코사인
• 벡터-기반 전경 신호들의 공간-시간 보간: 256 개 샘플들
• 추변 HOA 컨텐츠의 최소 HOA 오더: 1, MinNumOfCoeffsForAmbHOA = $(1+1)^2$ = 4, MaxNoOfAddActiveAmbCoeffs =NumOfHoaCoeff-MinNumOfCoeffForAmbHOA=16-4=12, AmbAsignmBits = ceil(log2(MaxNoOfAddActiveAmbCoeffs)) = ceil(log2(12)) = 4
• 컨텐츠의 HOA 오더: N = 3, NumOfHoaCoeff = $(N+1)^2$ =16
248M
HOA 프레임을 포함하는 SBR 을 갖는 USAC-3D 프레임
SBR 이 인에이블될 경우, 2 개의 오디오 프레임들이 하나의 USAC-3D 프레임 내에 저장되기 때문에, 2 개의 연속하는 HOA 프레임들은 USAC 확장 페이로드에 저장된다.
플렉서블 전송 채널들 = numHOATransportChannels – MinNumOfCoeffsForAmbHOA = 7-4 = 3 플렉서블 전송 채널들
USAC-3D 페이로드
HOA 프레임 사이드밴드 정보 프레임 k
HOA 프레임 사이드밴드 정보 프레임 k-1
249K
248L

249K

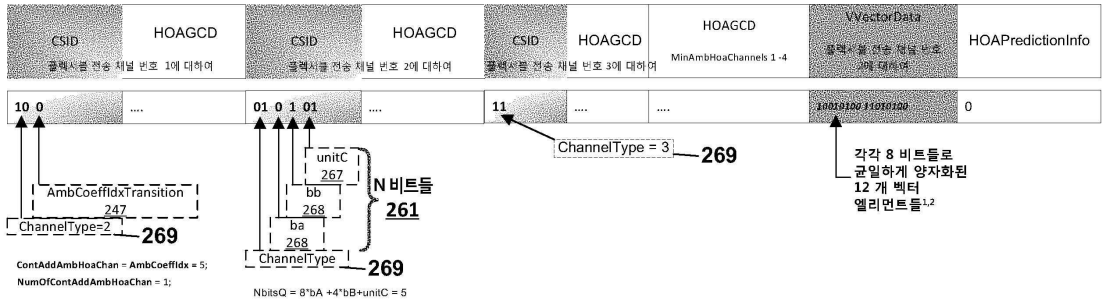
프레임 k

이 예에서, 플렉서블 전송 채널들은 1 개의 추가의 HOA 배경 채널 및 1 개의 벡터-기반 전경 채널을 포함한다



249L

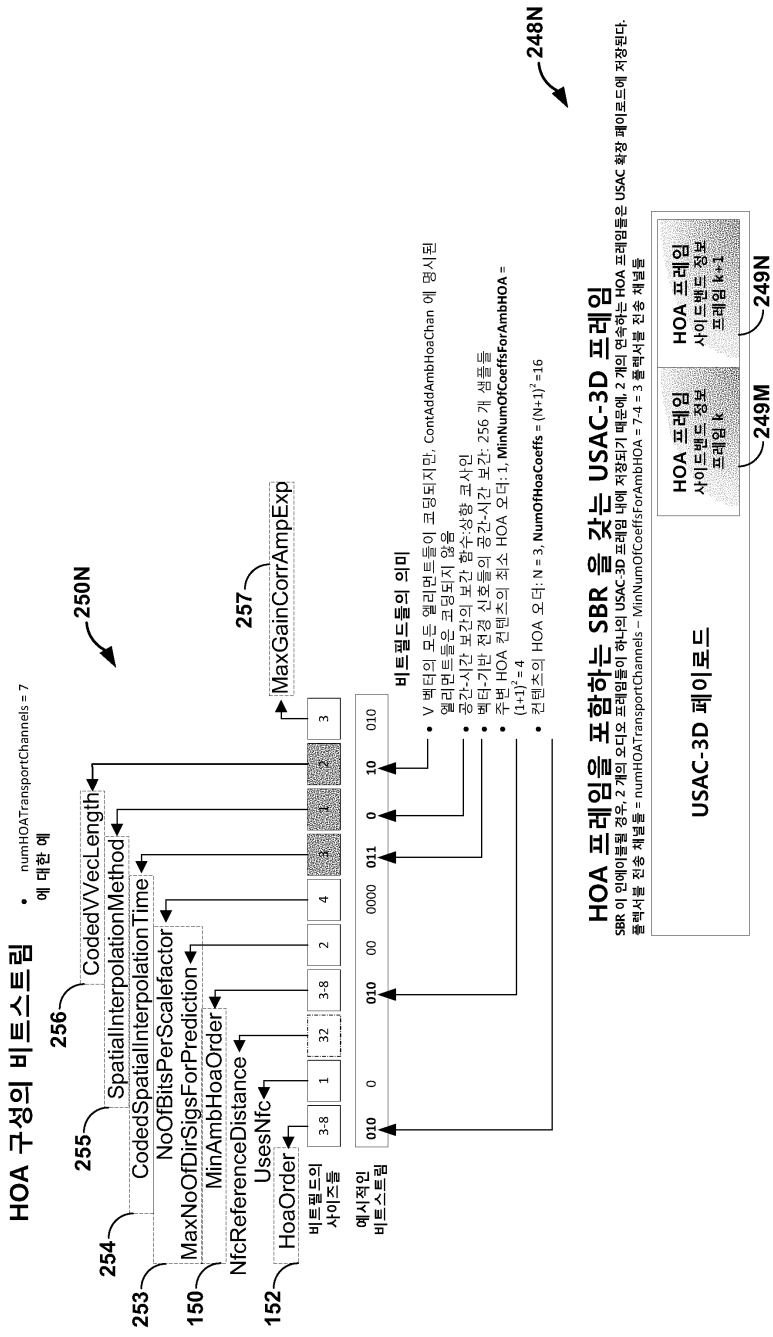
프레임 k+1



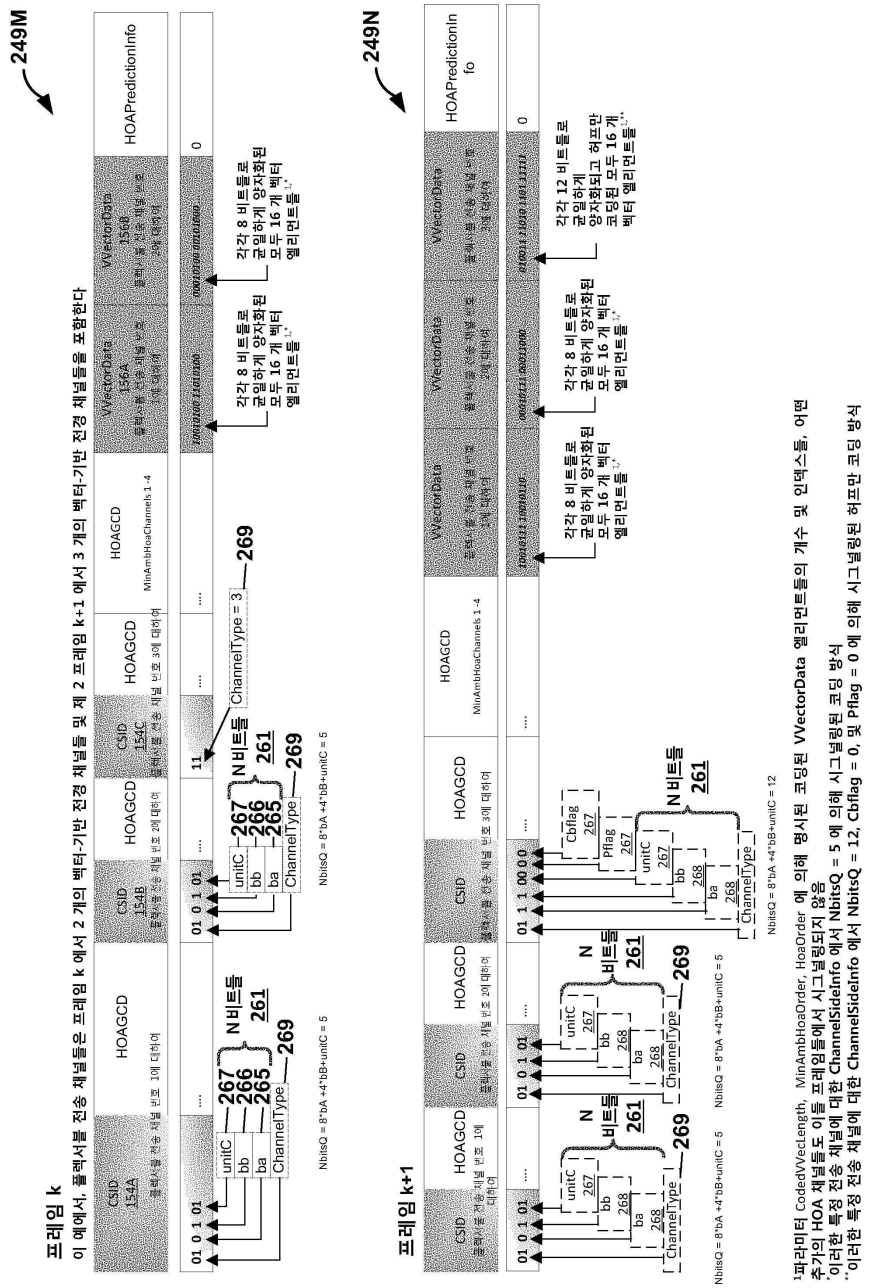
[1]**파라미터** CodedVVecLength, MinAmbHoaOrder, HoaOrder **에 의해 명시된 코딩된 VVectorData 엘리먼트들의 개수 및 인덱스들**
[2]**플렉서블 전송 채널 번호 2 에 대한 ChannelSideInfo 에서 NbitsQ = 5 에 의해 시그널링된 코딩 방식**
[**]**플렉서블 전송 채널 번호 1 에 대하여, 디코더의 내부 상태는 본원에서** AmbCoeffIdxTransitionState = 2**인 것으로 가정된다. 따라서, CodedAmbCoeffIdx 비트필드는 시그널링된다.**

등록특허 10-1929092

도면10ni


HOA 구성의 비트스트림
numHOATransportChannels = 7
에 대한 예
250N
256
255
254
253
150
152
CodedVVecLength
SpatialInterpolationMethod
CodedSpatialInterpolationTime
NoOfBitsPerScalefactor
MaxNoOfDirSigsForPrediction
MinAmbHoaOrder
NfcReferenceDistance
UsesNfc
HoaOrder
257
MaxGainCorrAmpExp
비트필드의 사이즈들
3-8
1
32
3-8
2
4
3
1
2
3
예시적인 비트스트림
010
0
010
00
0000
011
0
10
010
비트필드들의 의미
V 벡터의 모든 엘리먼트들이 코딩되지만, ContAddAmbHoaChan 에 명시된 엘리먼트들을 코딩되지 않음
공간-시간 보간의 보간 함수상향 코사인
벡터-기반 전경 신호들의 공간-시간 보간: 256 개 샘플들
주변 HOA 컨텐츠의 최소 HOA 오더: 1, MinNumOfCoeffsForAmbHOA = (1+1)2 = 4
컨텐츠의 HOA 오더: N = 3, NumOfHoaCoeffs = (N+1)2 =16
248N
HOA 프레임을 포함하는 SBR 을 갖는 USAC-3D 프레임
SBR 이 인에이블될 경우, 2 개의 오디오 프레임들이 하나의 USAC-3D 프레임 내에 저장되기 때문에, 2 개의 연속하는 HOA 프레임들은 USAC 확장 페이로드에 저장된다.
플렉서블 전송 채널들 = numHOATransportChannels – MinNumOfCoeffsForAmbHOA = 7-4 = 3 플렉서블 전송 채널들
USAC-3D 페이로드
HOA 프레임
사이드밴드 정보
프레임 k
HOA 프레임
사이드밴드 정보
프레임 k+1
249M
249N

도면10nii



[1]파라미터 CodedVVecLength, MinAmbHoaOrder, HoaOrder 에 의해 명시된 코딩된 VVectorData 엘리먼트들의 개수 및 인덱스들, 어떤 추가의 HOA 채널들도 이들 프레임들에서 시그널링되지 않음

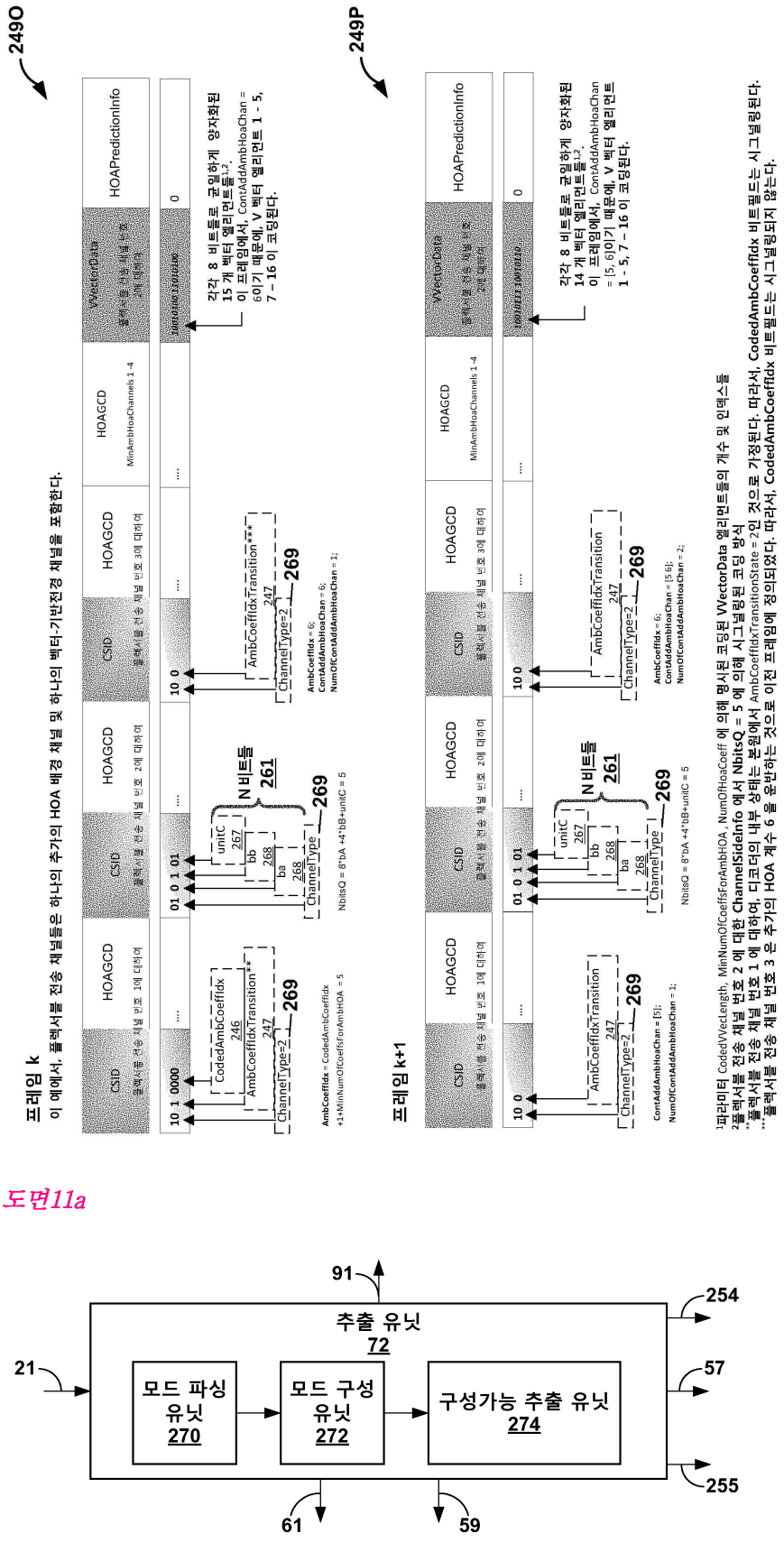
*이러한 특정 전송 채널에 대한 ChannelSideInfo 에서 NbitsQ = 5 에 의해 시그널링된 코딩 방식

**이러한 특정 전송 채널에 대한 ChannelSideInfo 에서 NbitsQ = 12, Cbflag = 0, 및 Pflag = 0 에 의해 시그널링된 허프만 코딩 방식

도면10oi

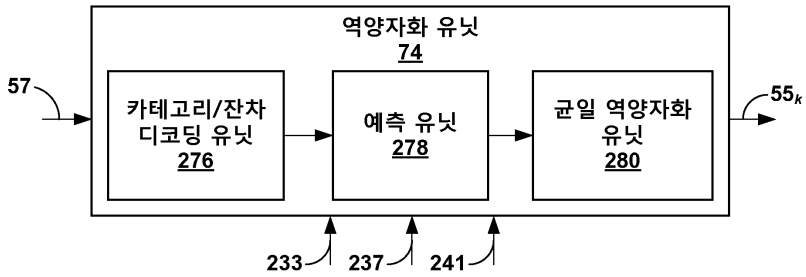
HOA 구성의 비트스트림
• numHOATransportChannels = 7 에 대한 예
2500
256
CodedVecLength
255
SpatialInterpolationMethod
254
CodedSpatialInterpolationTime
253
NoOfBitsPerScalefactor
150
MaxNoOfDirSigsForPrediction
MinAmbHoaOrder
NfcReferenceDistance
UsesNfc
152
HoaOrder
257
MaxGainCorrAmpExp
비트필드의 사이즈들
3-8
1
32
3-8
2
4
3
1
2
3
예시적인 비트스트림
010
0
010
00
0000
011
0
10
010
비트필드들의 의미
• V 벡터의 모드 엘리먼트들이 코딩되지만, ContAddAmbHoaChan 에 명시된 엘리먼트들은 코딩되지 않음
• 공간-시간 보간의 보간 함수-상향 코사인
• 벡터-기반 전경 신호들의 공간-시간 보간: 256 개 샘플들
• 주변 HOA 컨텐츠의 최소 HOA 오더: 1, MinNumOfCoeffsForAmbHOA = $(1+1)^2 = 4$,
MaxNoOfAddActiveAmbCoeffs =NumOfHoaCoeff-MinNumOfCoeffForAmbHOA=16-4=12,
AmbAsigmBits = ceil(log2(MaxNoOfAddActiveAmbCoeffs)) = ceil(log2(12)) = 4
• 컨텐츠의 HOA 오더: N = 3, NumOfHoaCoeff = $(N+1)^2$ =16
2480
USAC-3D frame with SBR including HOAFrame
SBR 이 인에이블될 경우, 2 개의 오디오 프레임들마 하나의 USAC-3D 프레임 내에 저장되기 때문에, 2 개의 연속하는 HOA 프레임들은 USAC 확장 페이로드에 저장된다.
폴헥사블 전송 채널들 = numHOATransportChannels – MinNumOfCoeffsForAmbHOA = 7-4 = 3 폴헥사블 전송 채널들
USAC-3D 페이로드
HOA 프레임
사이드밴드 정보
프레임 k
2490
HOA 프레임
사이드밴드 정보
프레임 k+1
249P

도면10oii

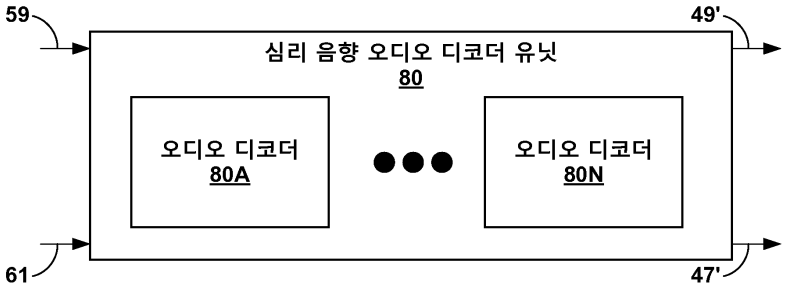


도면11a

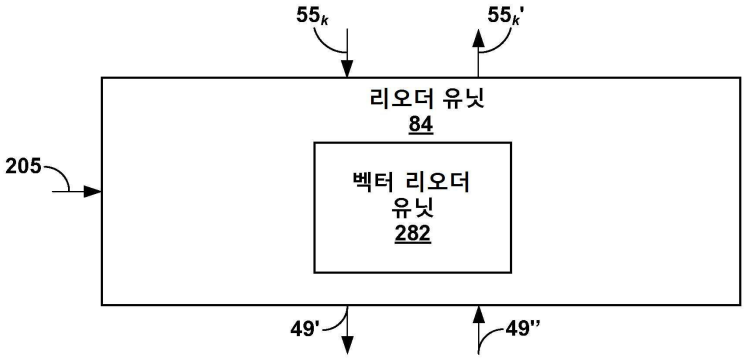
도면11b



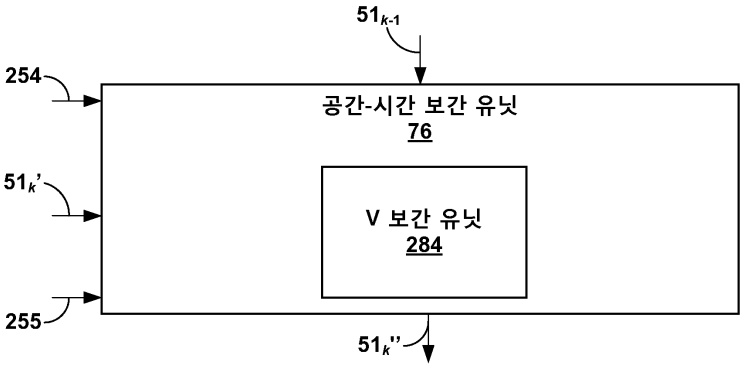
도면11c



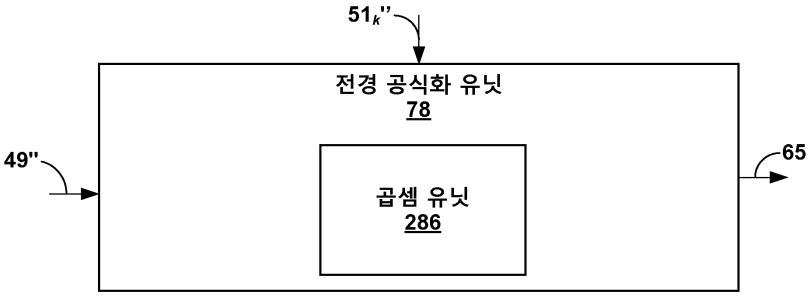
도면11d



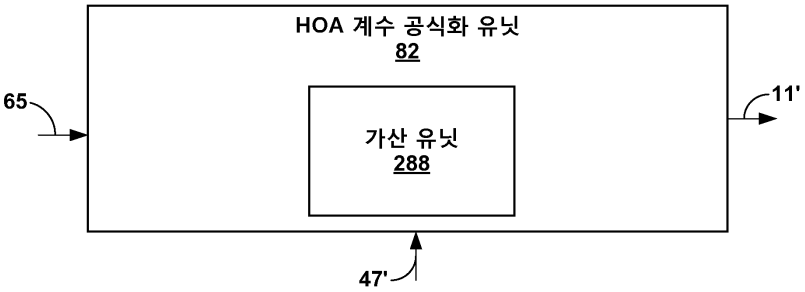
도면11e



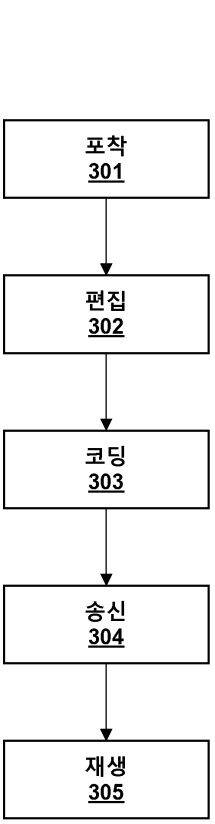
도면11f



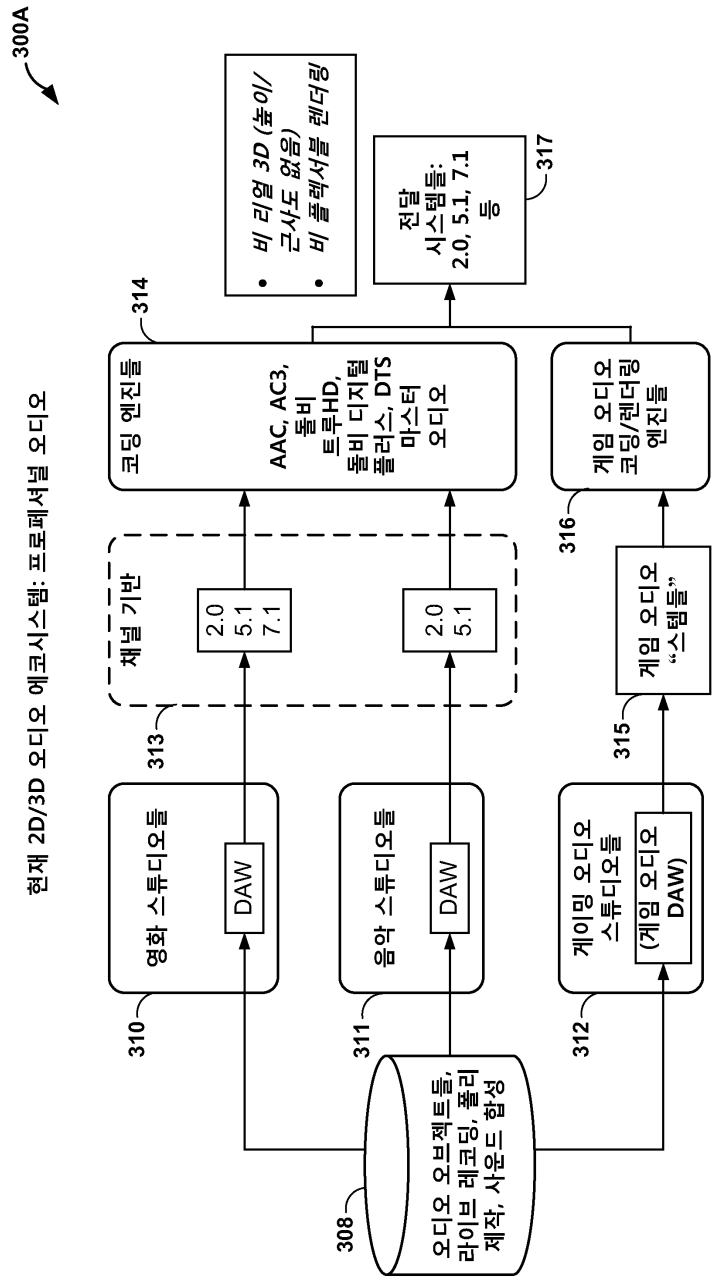
도면11g



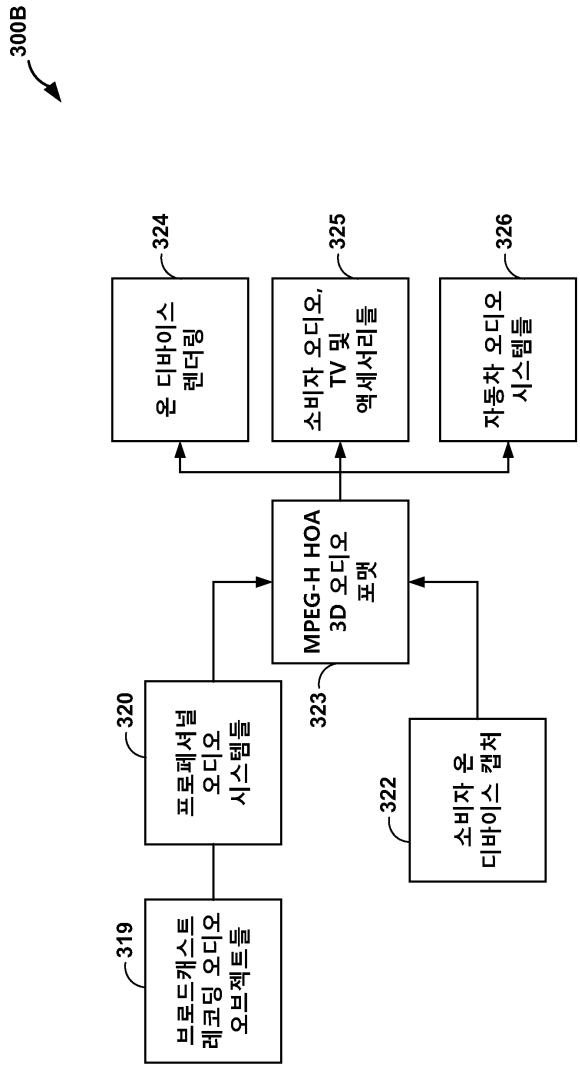
도면12



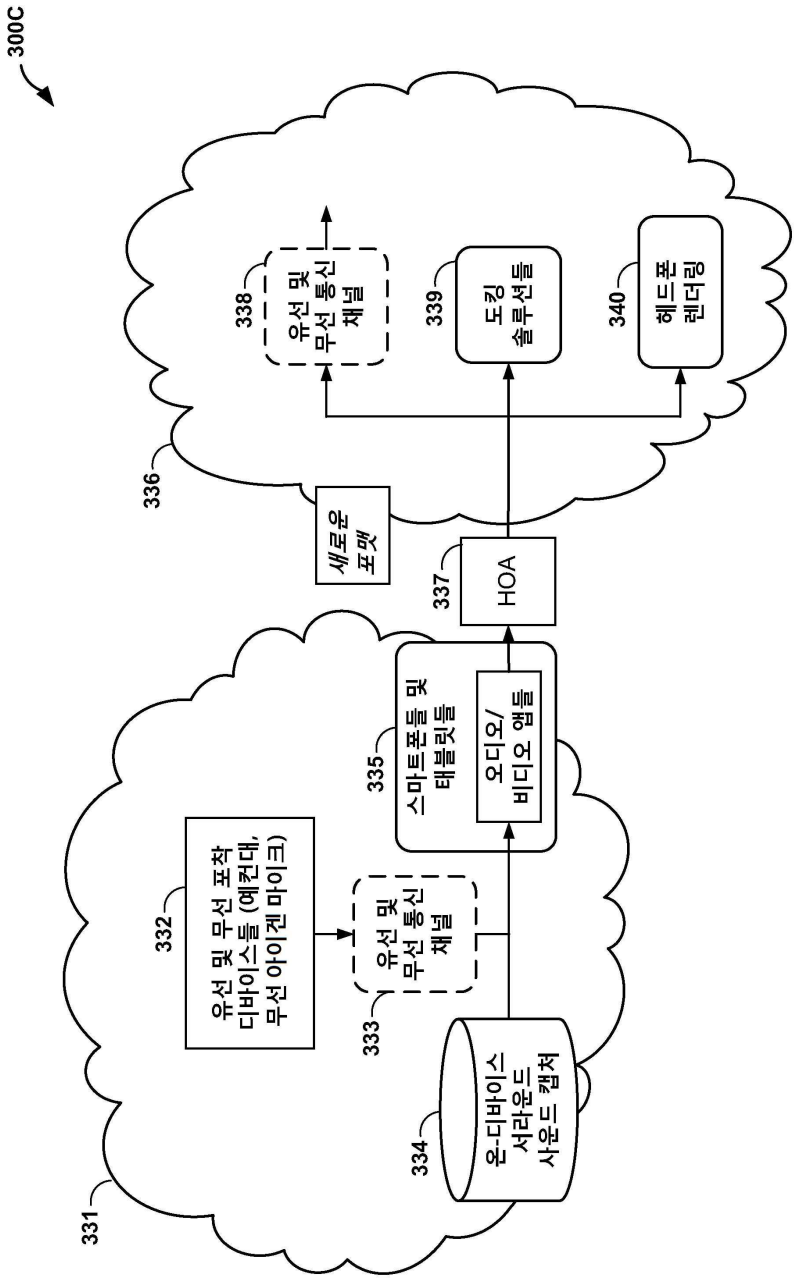
도면13


300A
현재 2D/3D 오디오 에코시스템: 프로페셔널 오디오
영화 스튜디오들
DAW
310
음악 스튜디오들
DAW
311
게이밍 오디오 스튜디오들 (게임 오디오 DAW)
312
오디오 오브젝트들, 라이브 레코딩, 폴리 제작, 사운드 합성
308
채널 기반
313
2.0 5.1 7.1
2.0 5.1
게임 오디오 "스템들"
315
코딩 엔진들
AAC, AC3, 돌비 트루HD, 돌비 디지털 플러스, DTS 마스터 오디오
314
게임 오디오 코딩/렌더링 엔진들
316
• 비 리얼 3D (높이/근사도 없음)
• 비 플랙서블 렌더링
전달 시스템들: 2.0, 5.1, 7.1 등
317

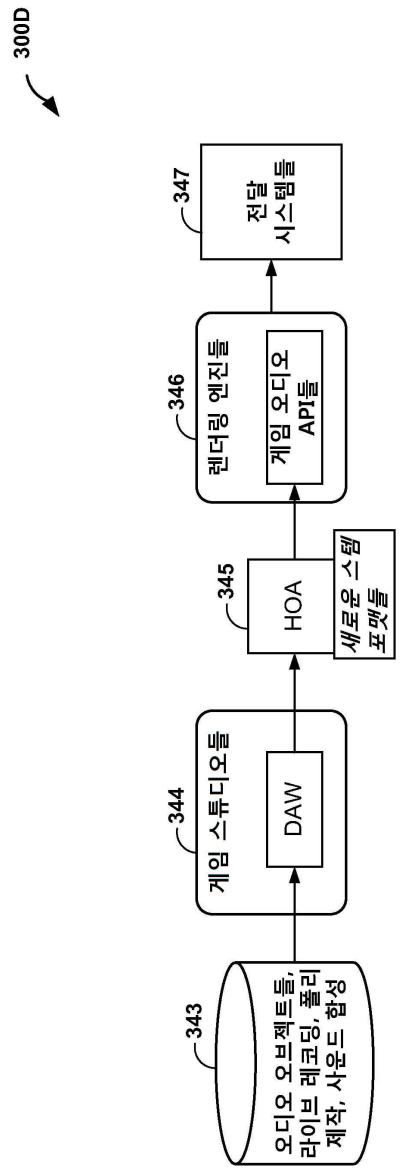
도면14


300B
324
온 디바이스 렌더링
325
소비자 오디오, TV 및 액세서리들
326
자동차 오디오 시스템들
MPEG-H HOA 3D 오디오 포맷
323
320
프로페셔널 오디오 시스템들
322
소비자 온 디바이스 캡처
319
브로드캐스트 레코딩 오디오 오브젝트들

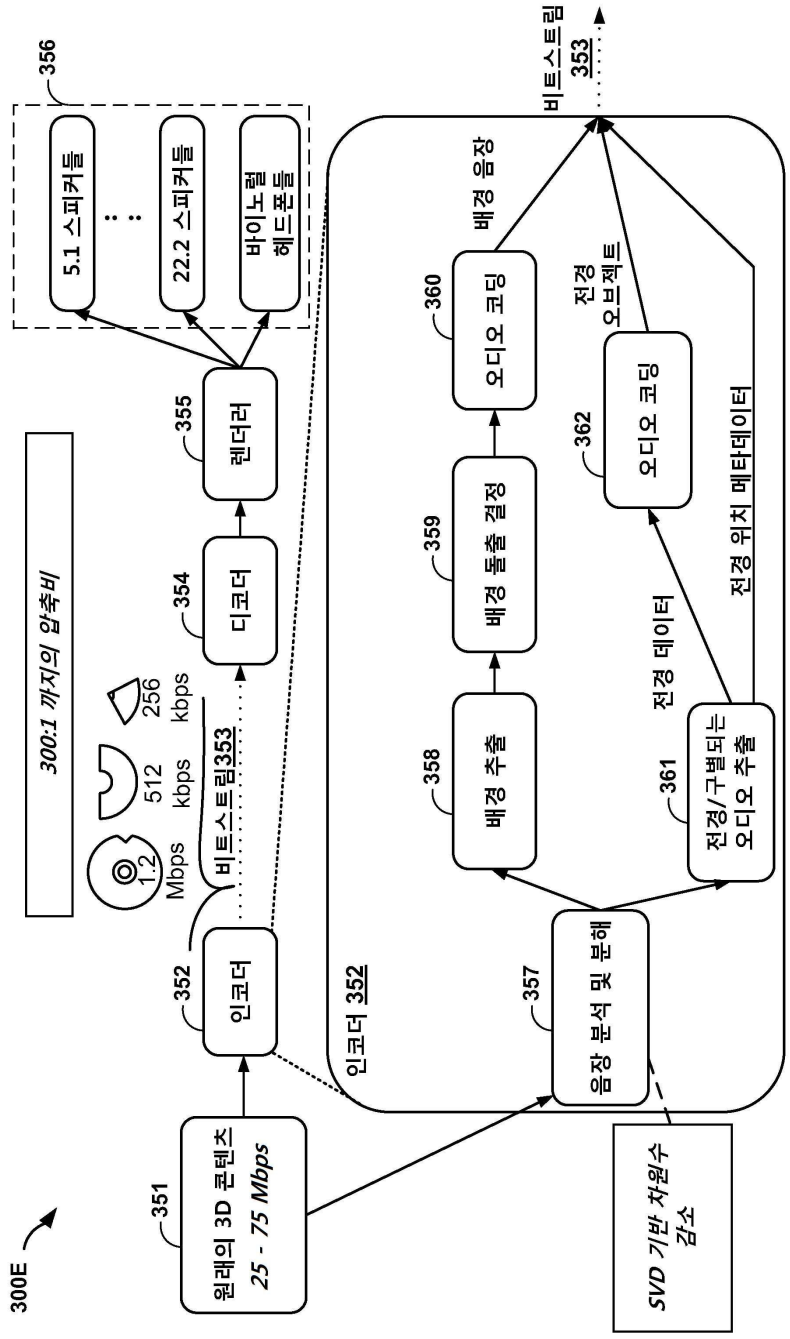
도면15a


300C
331
332
유선 및 무선 포착 디바이스들 (예컨대, 무선 아이겐 마이크)
333
유선 및 무선 통신 채널
334
온-디바이스 서라운드 사운드 캡처
335
스마트폰들 및 태블릿들
오디오/비디오 앱들
337
HOA
새로운 포맷
336
338
유선 및 무선 통신 채널
339
도킹 솔루션들
340
헤드폰 렌더링

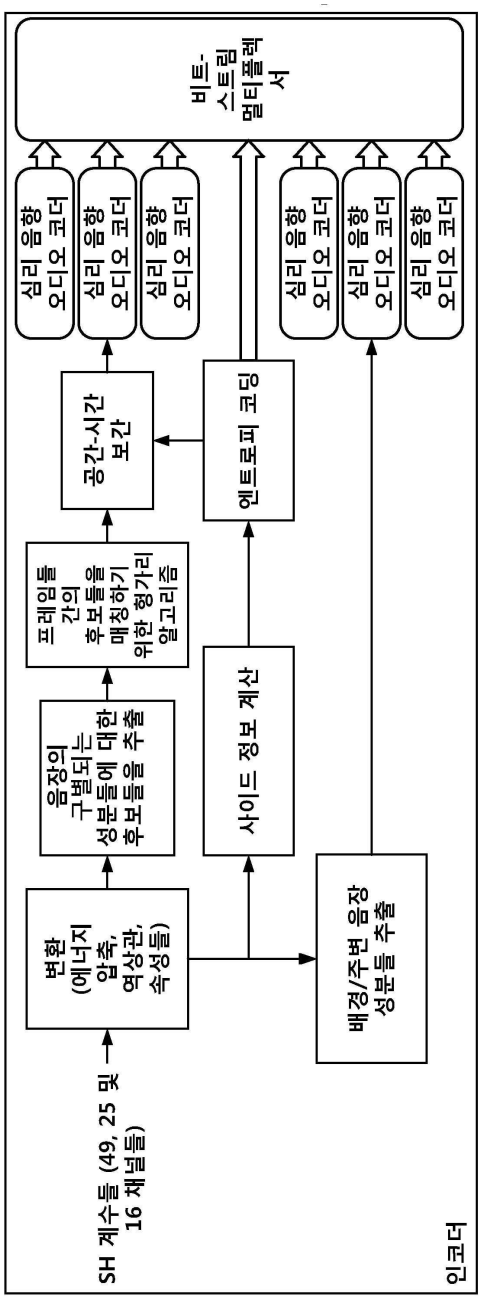
도면15b


300D
343
오디오 오브젝트들, 라이브 레코딩, 폴리 제작, 사운드 합성
344
게임 스튜디오들
DAW
345
HOA
새로운 스템 포맷들
346
렌더링 엔진들
게임 오디오 API들
347
전달 시스템들

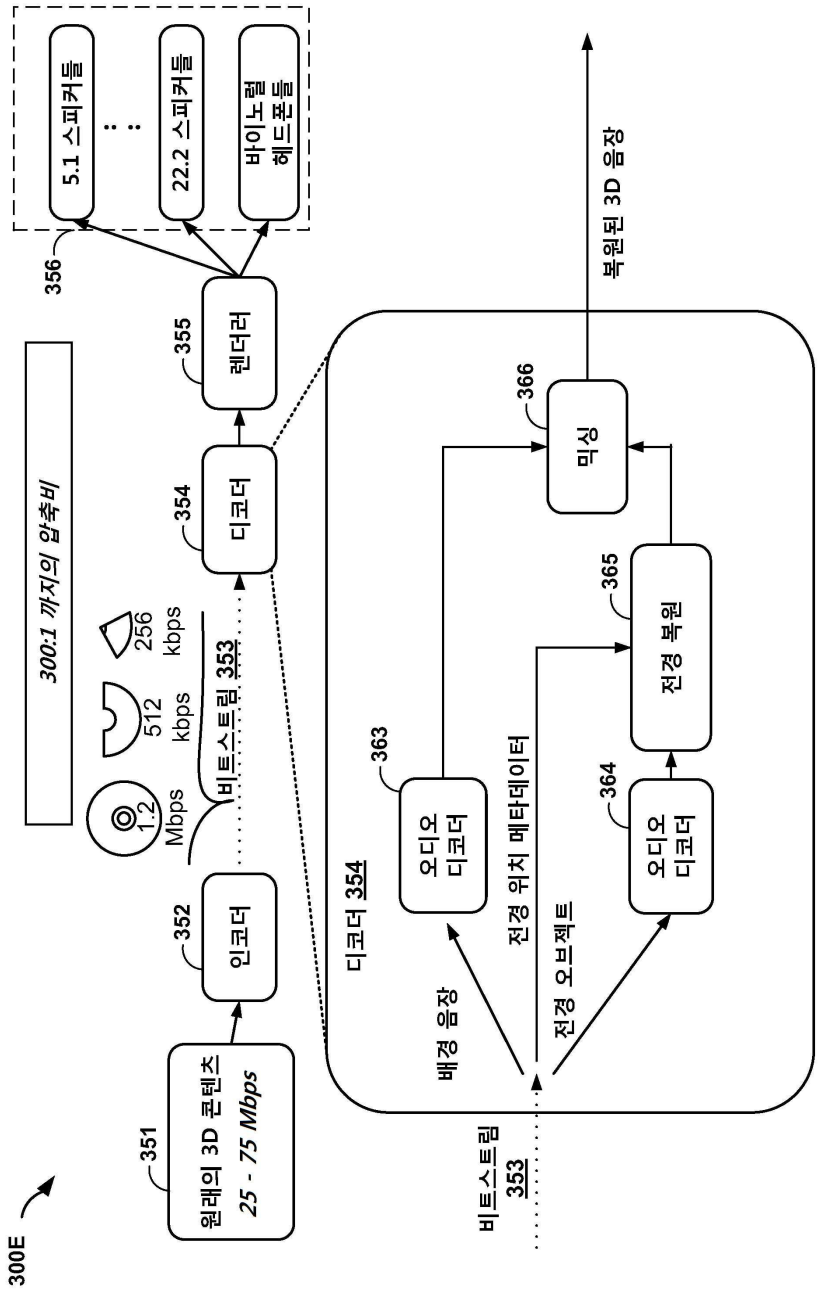
도면16


300E
351
원래의 3D 콘텐츠
25 - 75 Mbps
352
인코더
300:1 까지의 압축비
1.2 Mbps
512 kbps
256 kbps
비트스트림353
354
디코더
355
렌더러
356
5.1 스피커들
22.2 스피커들
바이노럴 헤드폰들
인코더 352
357
음장 분석 및 분해
358
배경 추출
359
배경 돌출 결정
360
오디오 코딩
배경 음장
361
전경/구별되는 오디오 추출
362
오디오 코딩
전경 오브젝트
전경 데이터
전경 위치 메타데이터
비트스트림 353
SVD 기반 차원수 감소

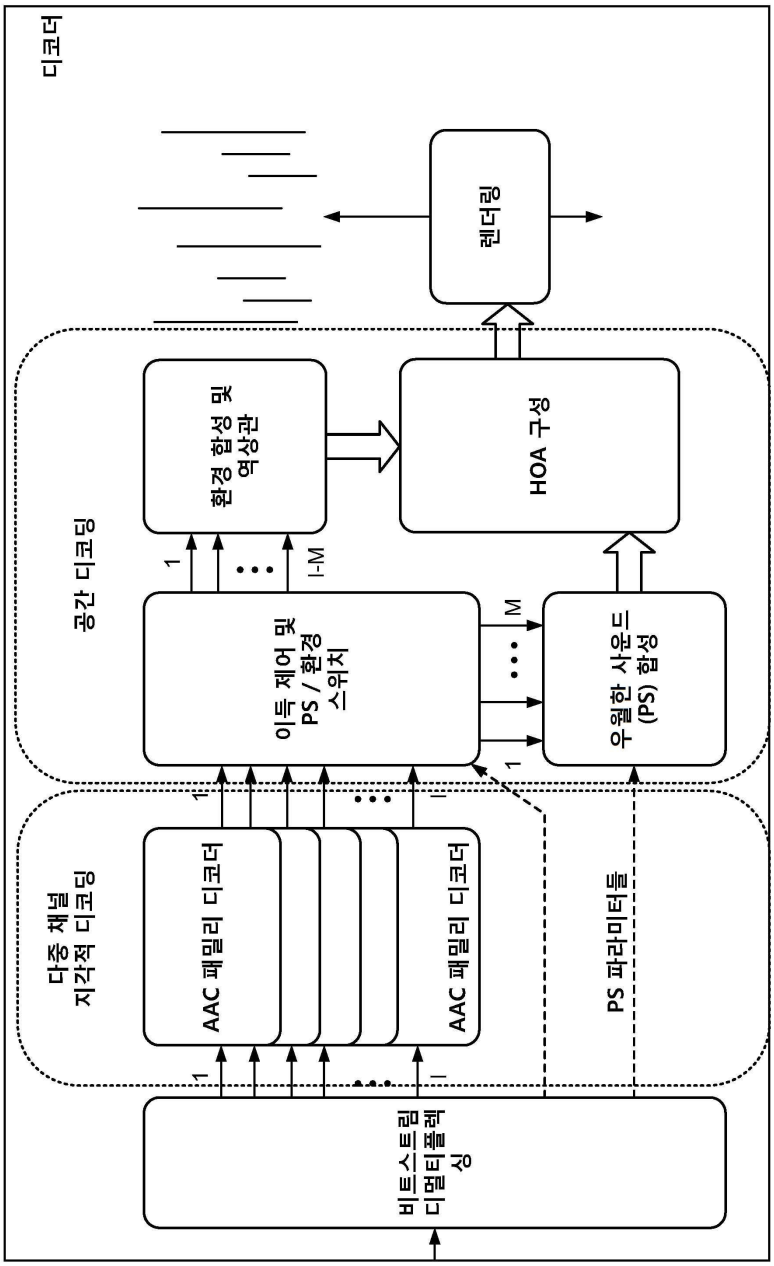
도면17


인코더
SH 계수들 (49, 25 및 16 채널들)
변환 (에너지 압축, 역상관, 속성들)
음장의 구별되는 성분들에 대한 후보들을 추출
프레임들 간의 후보들을 매칭하기 위한 형가리 알고리즘
공간-시간 보간
배경/주변 음장 성분들 추출
사이드 정보 계산
엔트로피 코딩
심리 음향 오디오 코더
심리 음향 오디오 코더
심리 음향 오디오 코더
심리 음향 오디오 코더
심리 음향 오디오 코더
심리 음향 오디오 코더
비트-스트림 멀티플렉서

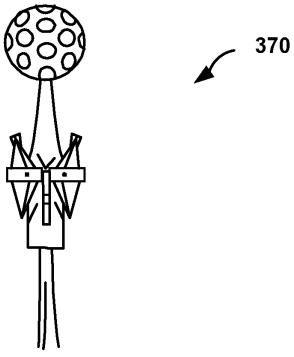
도면18


300E
원래의 3D 콘텐츠
25 - 75 Mbps
351
인코더
352
1.2 Mbps
512 kbps
256 kbps
비트스트림 353
300:1 까지의 압축비
디코더
354
렌더러
355
356
5.1 스피커들
22.2 스피커들
바이노럴 헤드폰들
디코더 354
오디오 디코더
363
배경 음장
전경 위치 메타데이터
전경 오브젝트
오디오 디코더
364
전경 복원
365
믹싱
366
복원된 3D 음장
비트스트림 353

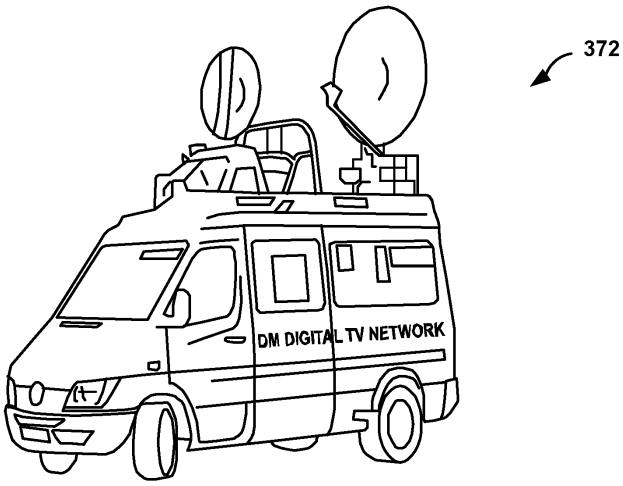
도면19


디코더
렌더링
환경 합성 및 역상관
HOA 구성
공간 디코딩
이득 제어 및 PS / 환경 스위치
우월한 사운드 (PS) 합성
1
I-M
M
1
1
I
다중 채널 지각적 디코딩
AAC 패밀리 디코더
AAC 패밀리 디코더
PS 파라미터들
1
I
비트스트림 디멀티플렉싱

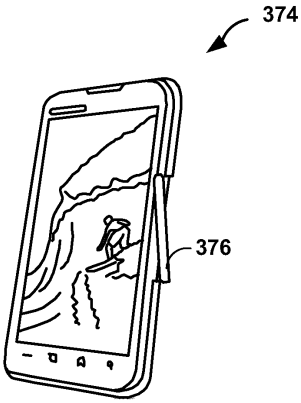
도면20a


370

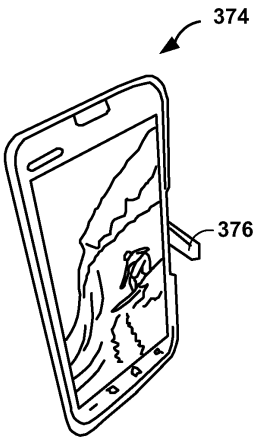
도면20b



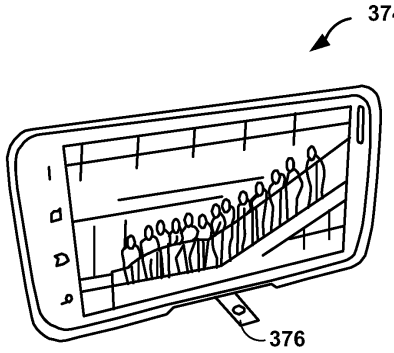
도면20c



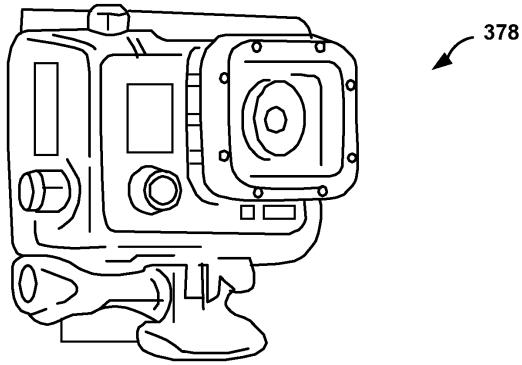
도면20d



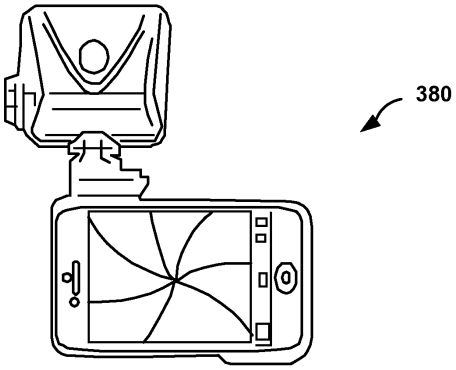
도면20e



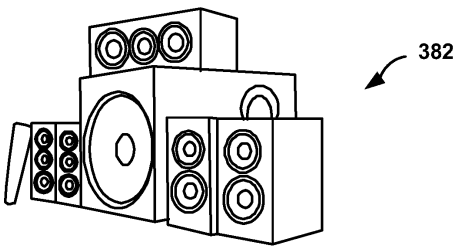
도면20f



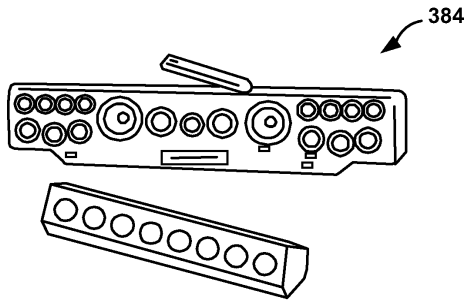
도면20g



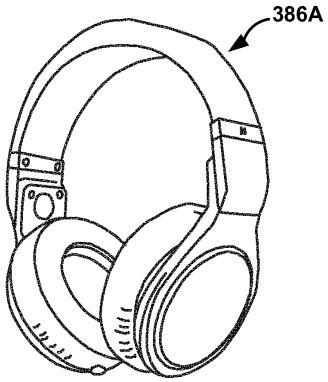
도면21a



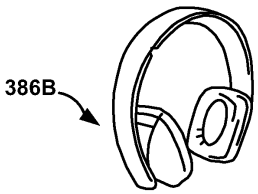
도면21b



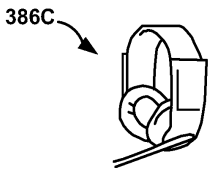
도면21c



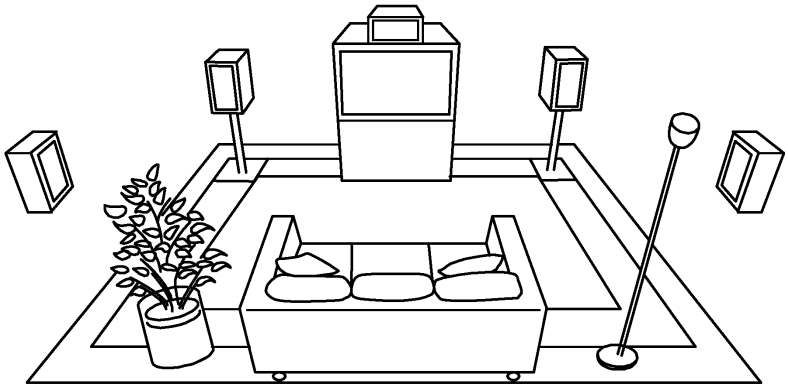
도면21d



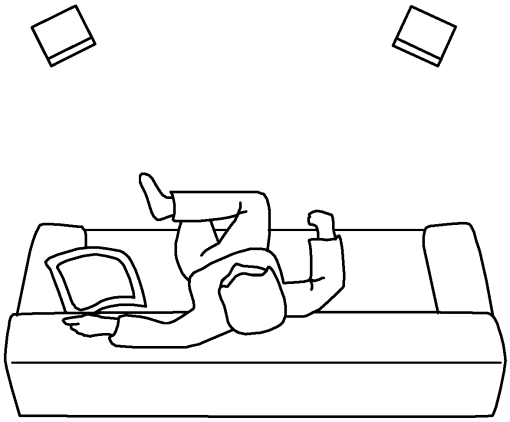
도면21e



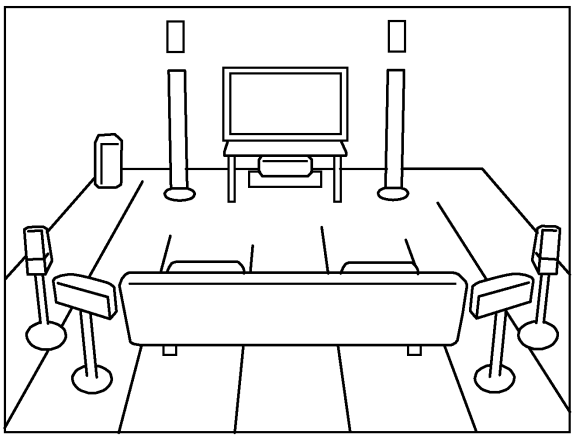
도면22a



도면22b

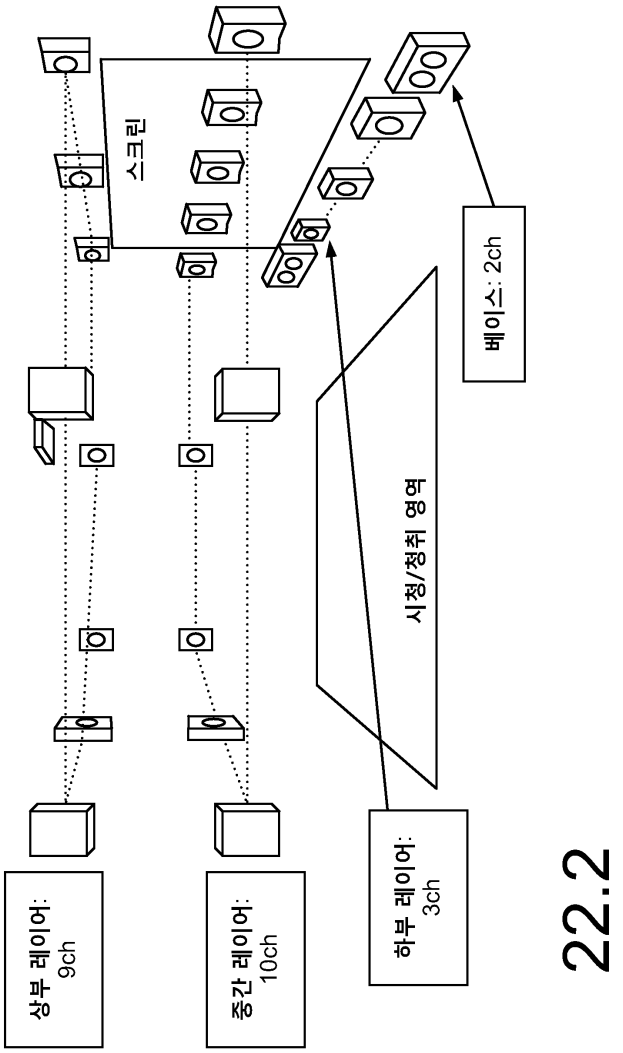


도면22c

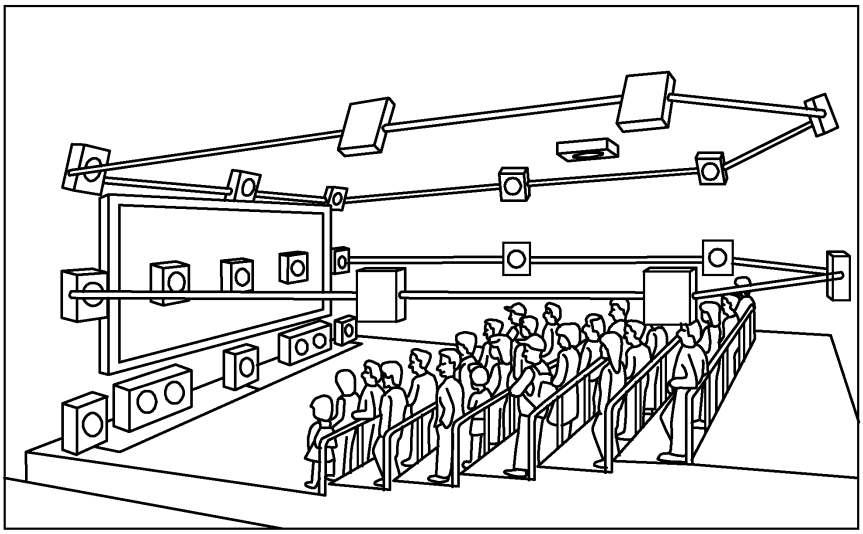


풀 하이트를 갖는 9.1

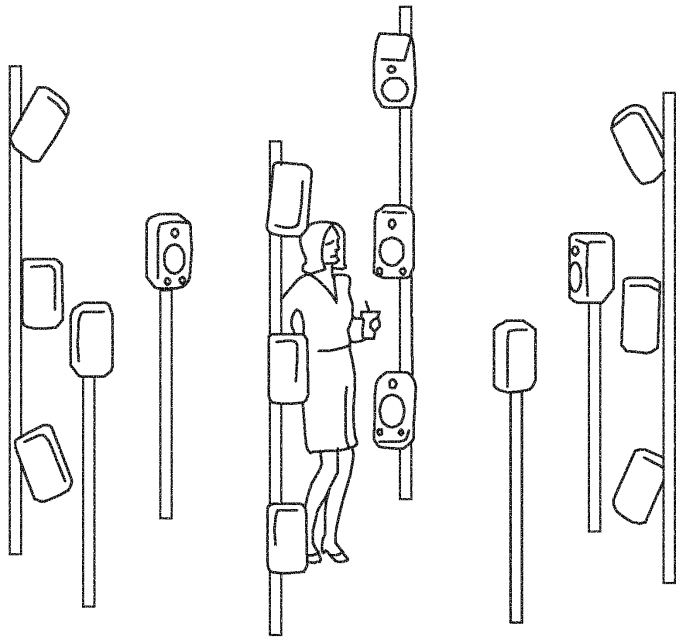
도면22d



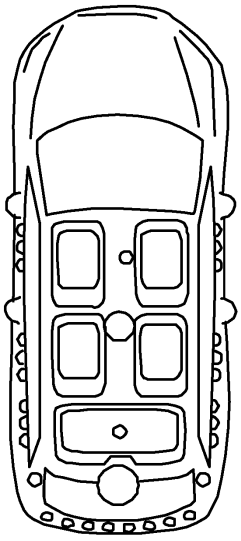
도면22e



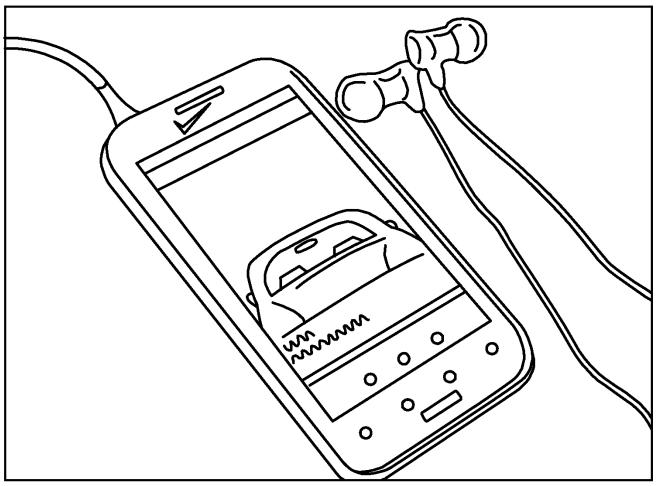
도면22f



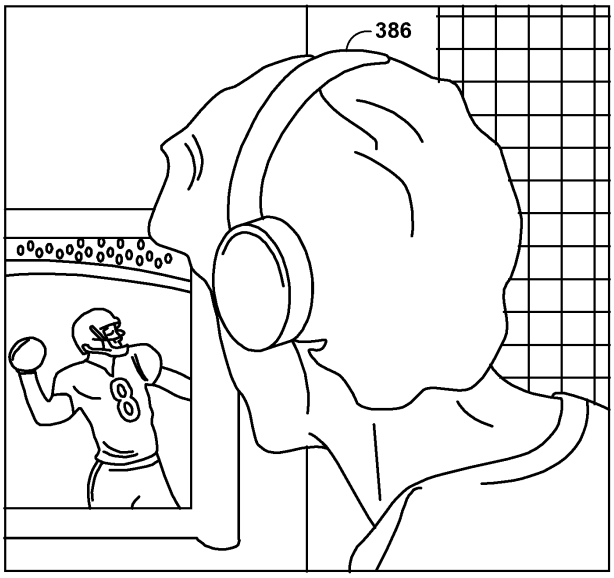
도면22g



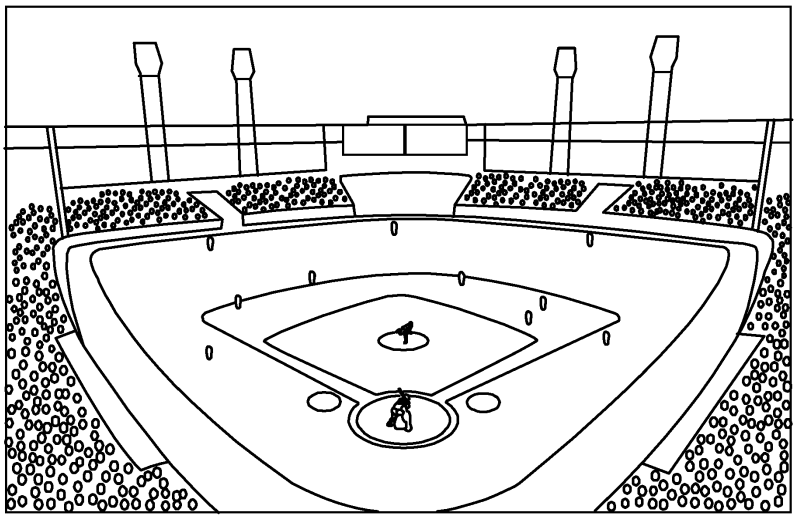
도면22h



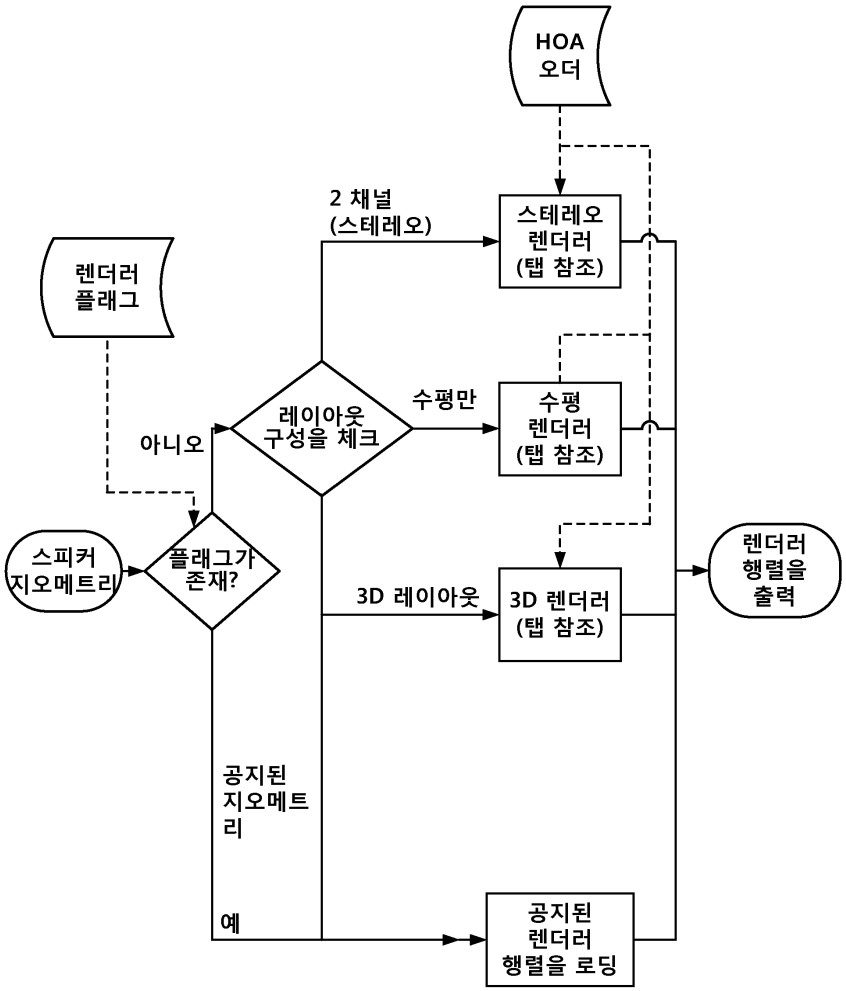
도면23



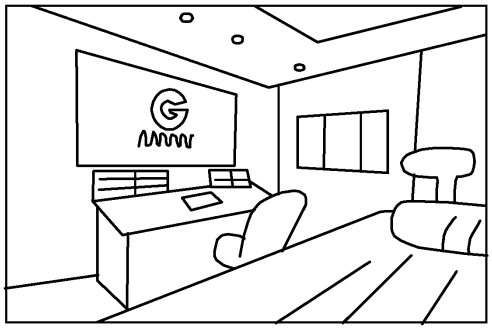
도면24



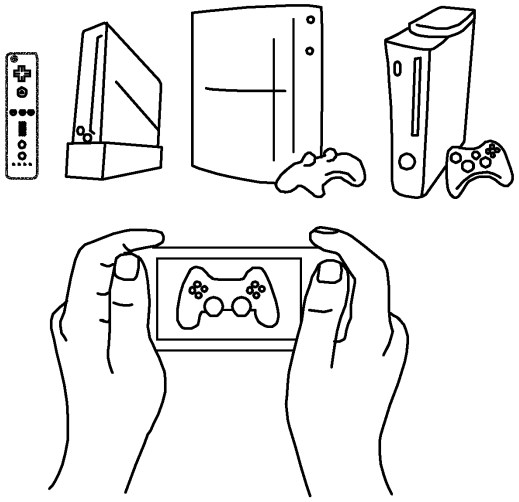
도면25


HOA
오더
2 채널
(스테레오)
스테레오
렌더러
(탭 참조)
렌더러
플래그
레이아웃
구성을 체크
수평만
수평
렌더러
(탭 참조)
아니오
스피커
지오메트리
플래그가
존재?
3D 레이아웃
3D 렌더러
(탭 참조)
렌더러
행렬을
출력
공지된
지오메트
리
예
공지된
렌더러
행렬을 로딩

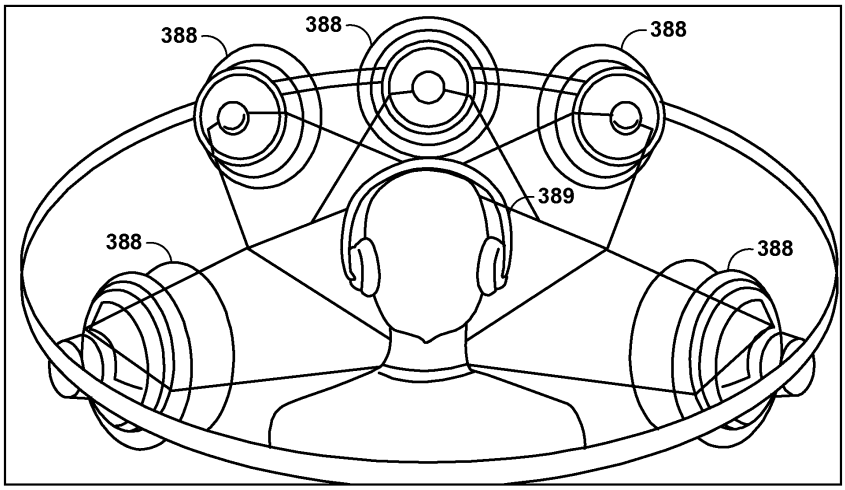
도면26



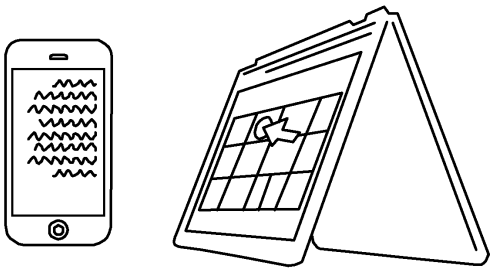
도면27

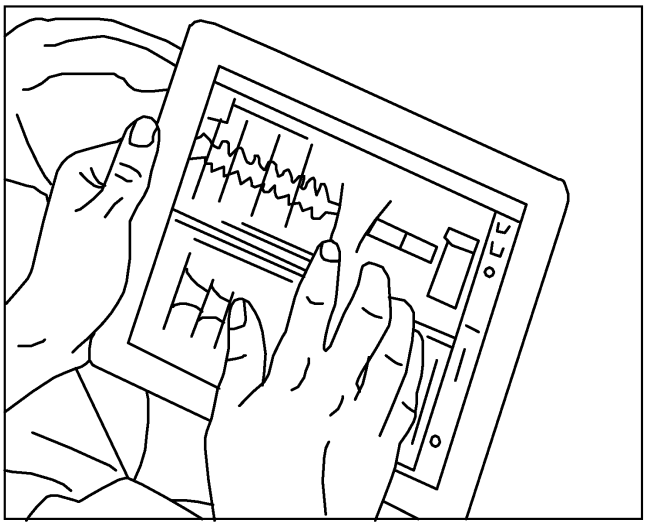


도면28

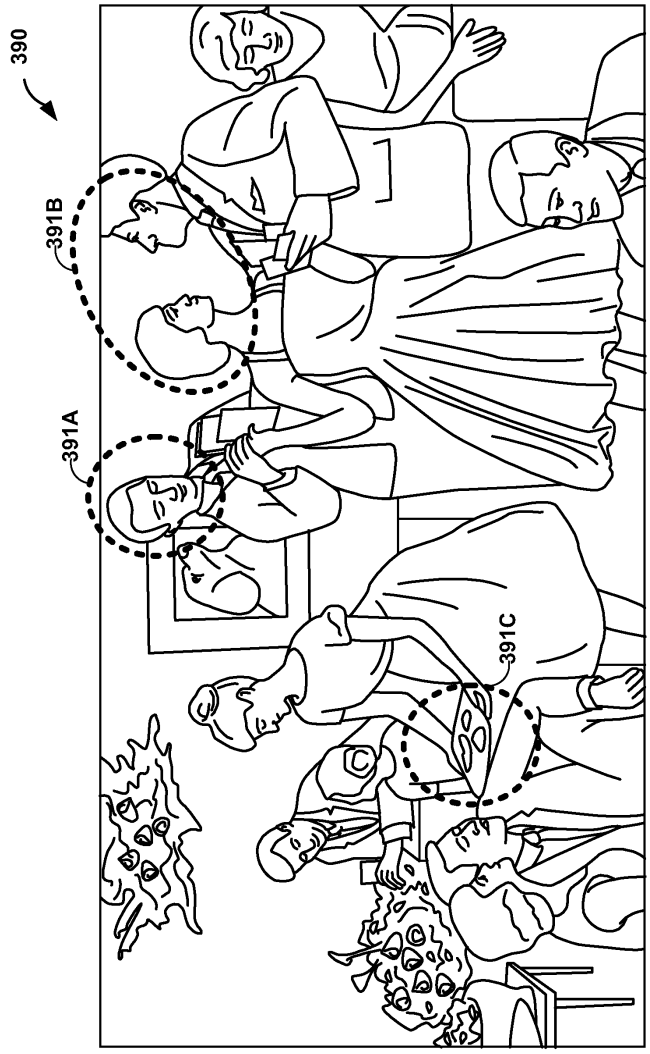


도면29

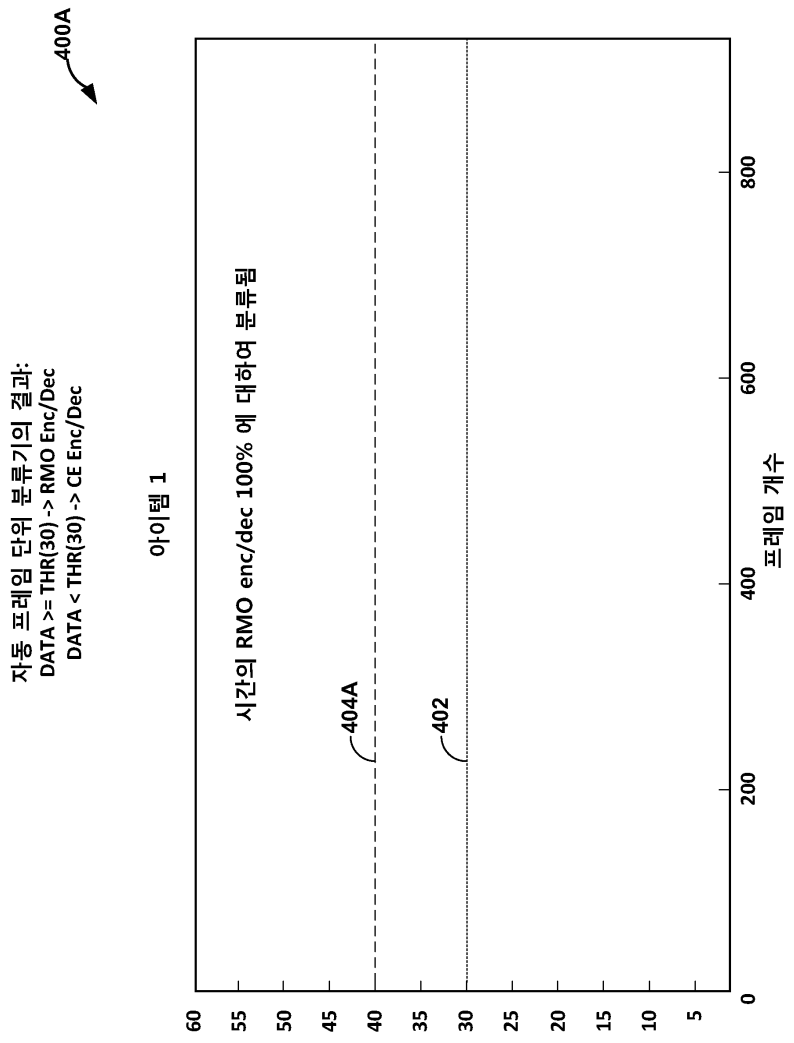




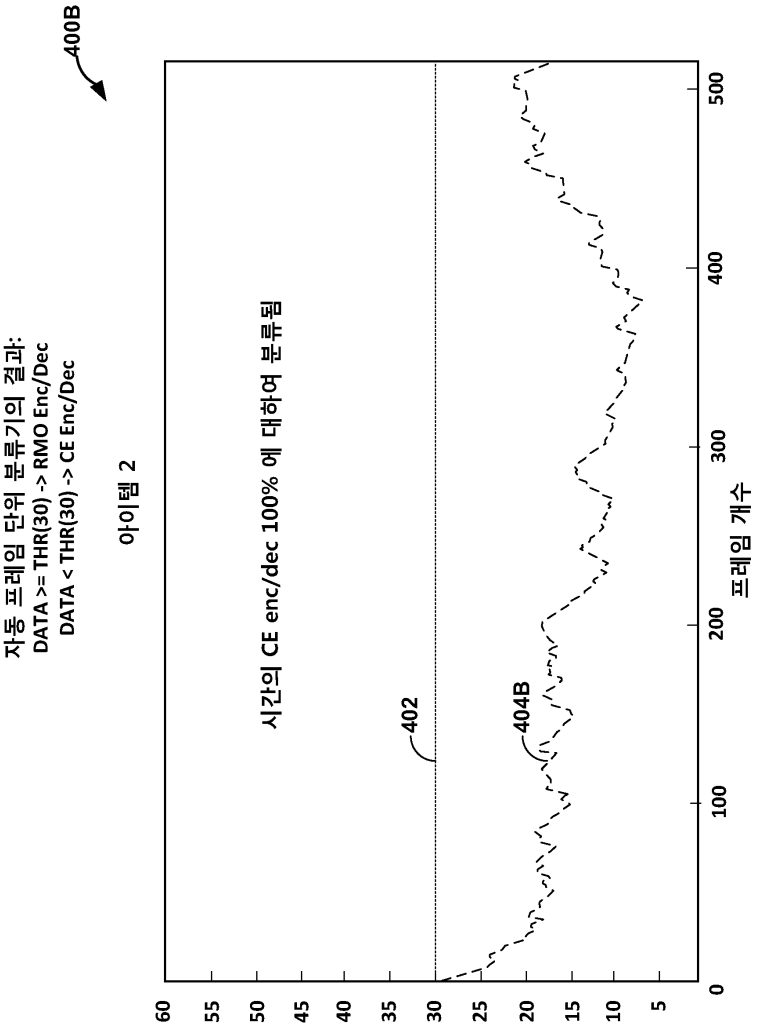
도면30



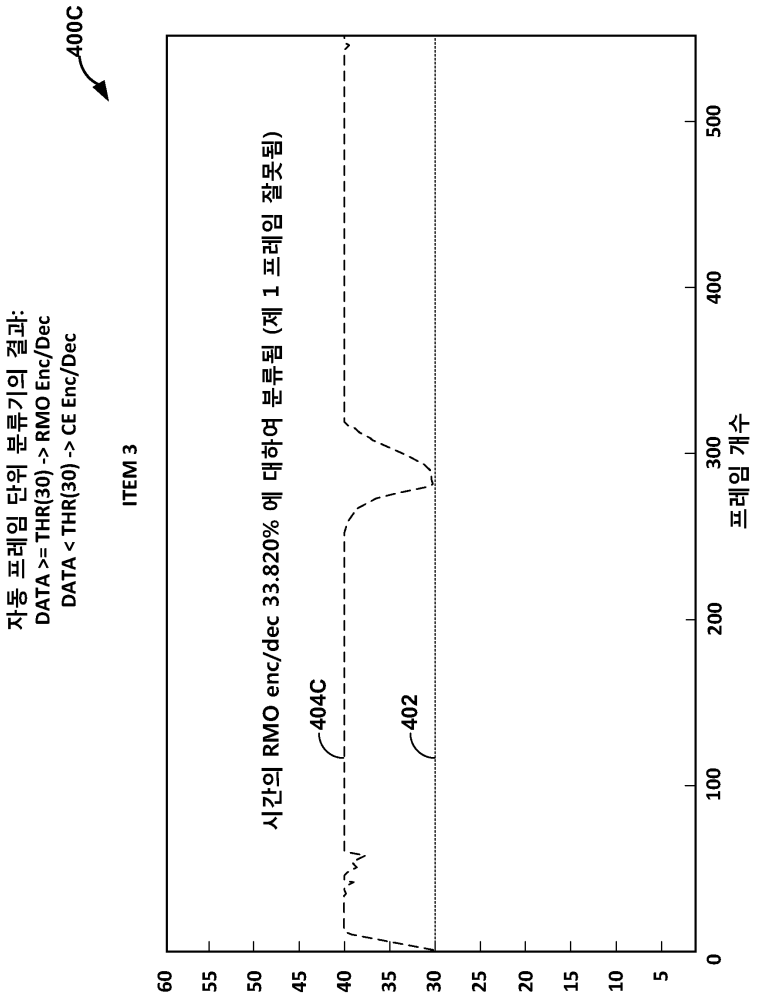
도면31a


400A
자동 프레임 단위 분류기의 결과:
DATA >= THR(30) -> RMO Enc/Dec
DATA < THR(30) -> CE Enc/Dec
아이템 1
시간의 RMO enc/dec 100% 에 대하여 분류됨
404A
402
60
55
50
45
40
35
30
25
20
15
10
5
0
200
400
600
800
프레임 개수

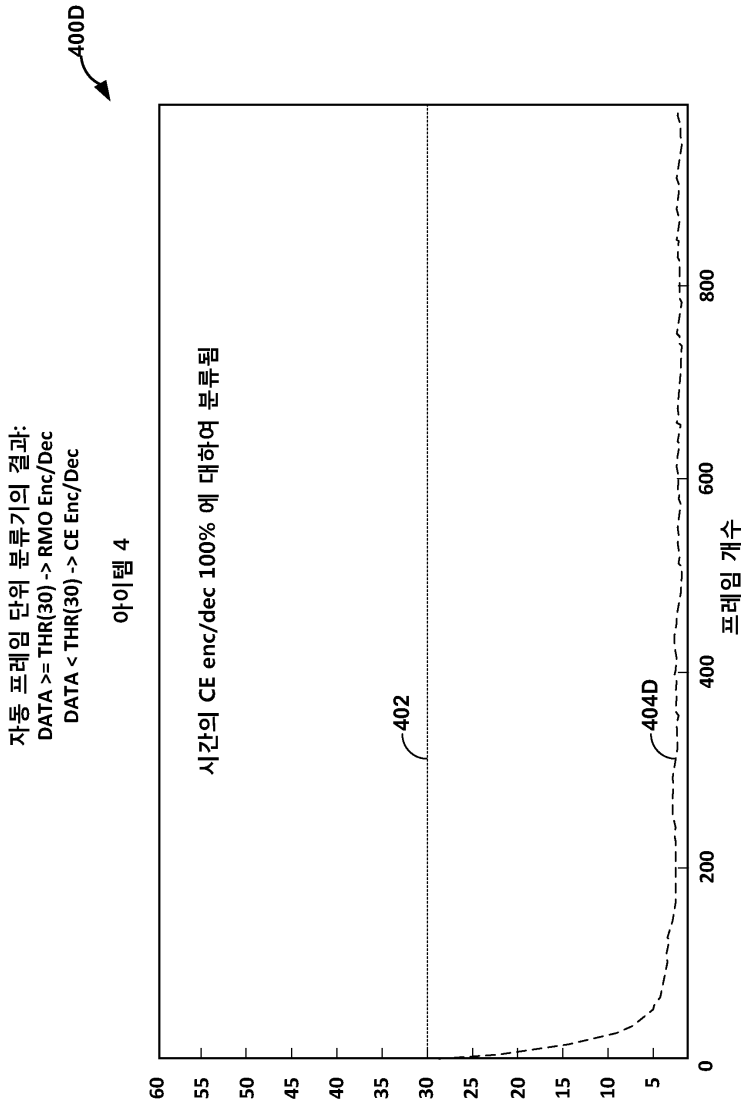
도면31b


400B
자동 프레임 단위 분류기의 결과:
DATA >= THR(30) -> RMO Enc/Dec
DATA < THR(30) -> CE Enc/Dec
아이템 2
시간의 CE enc/dec 100% 에 대하여 분류됨
402
404B
60
55
50
45
40
35
30
25
20
15
10
5
0
100
200
300
400
500
프레임 개수

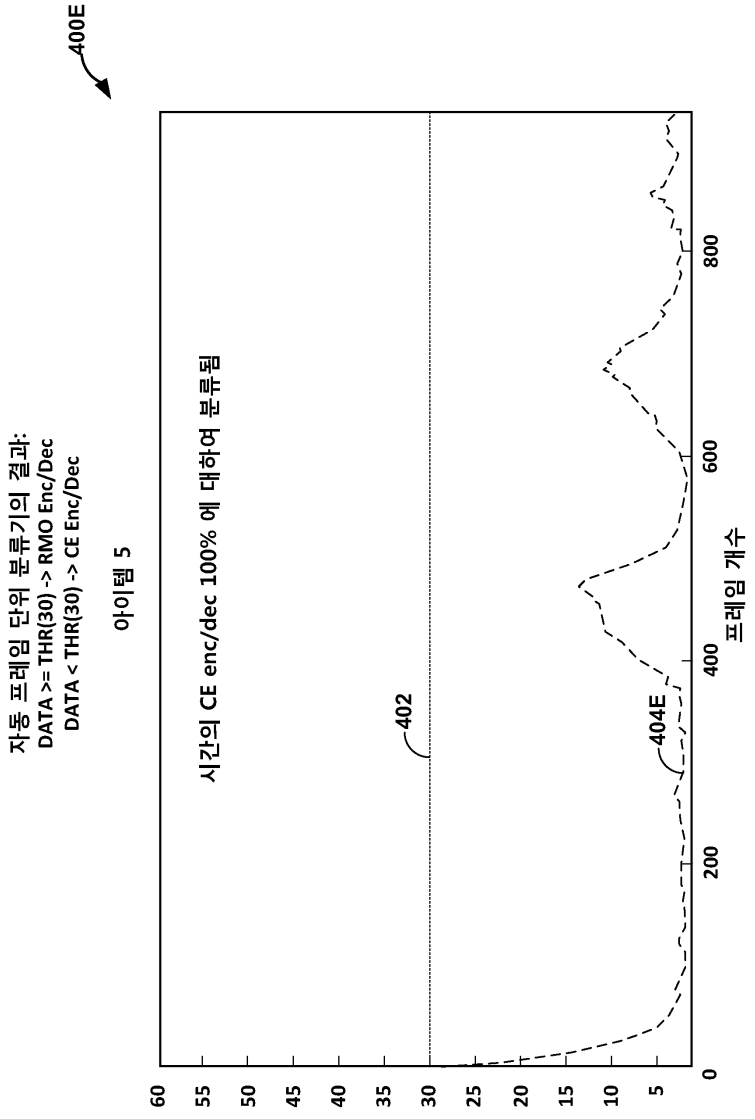
도면31c


400C
자동 프레임 단위 분류기의 결과:
DATA >= THR(30) -> RMO Enc/Dec
DATA < THR(30) -> CE Enc/Dec
ITEM 3
시간의 RMO enc/dec 33.820% 에 대하여 분류됨 (제 1 프레임 잘못됨)
404C
402
60
55
50
45
40
35
30
25
20
15
10
5
0
100
200
300
400
500
프레임 개수

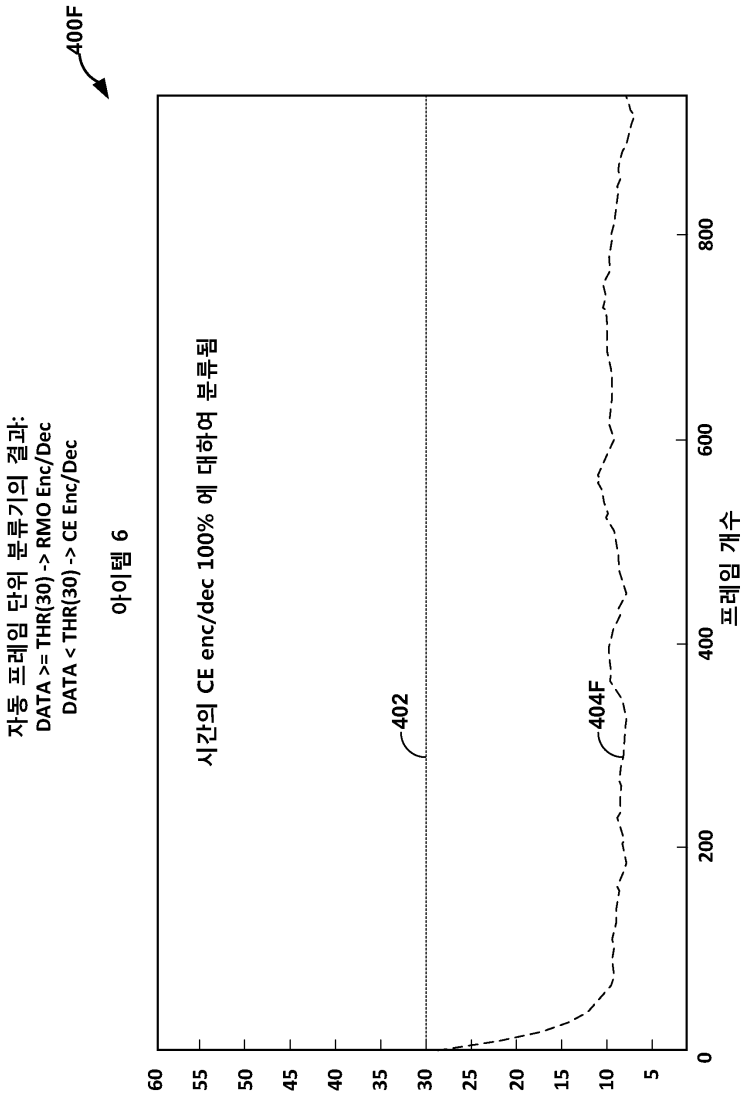
도면31d


400D
자동 프레임 단위 분류기의 결과:
DATA >= THR(30) -> RMO Enc/Dec
DATA < THR(30) -> CE Enc/Dec
아이템 4
시간의 CE enc/dec 100% 에 대하여 분류됨
402
404D
60
55
50
45
40
35
30
25
20
15
10
5
0
200
400
600
800
프레임 개수

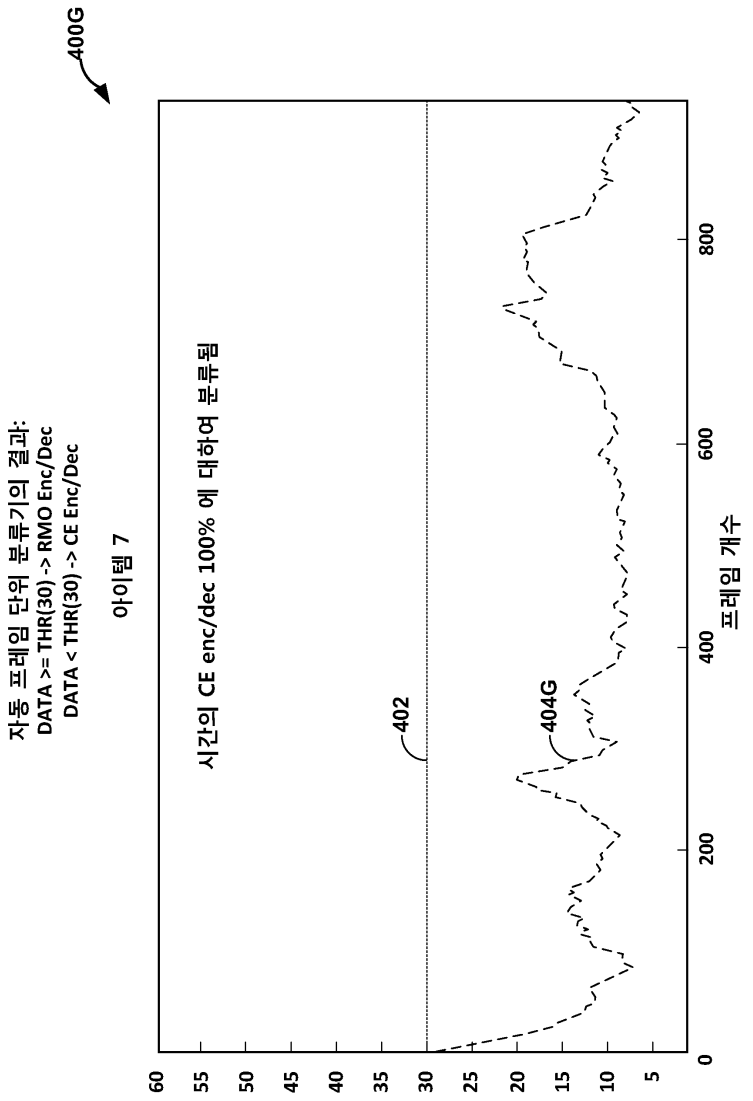
도면31e


400E
자동 프레임 단위 분류기의 결과:
DATA >= THR(30) -> RMO Enc/Dec
DATA < THR(30) -> CE Enc/Dec
아이템 5
시간의 CE enc/dec 100% 에 대하여 분류됨
402
404E
60
55
50
45
40
35
30
25
20
15
10
5
0
200
400
600
800
프레임 개수

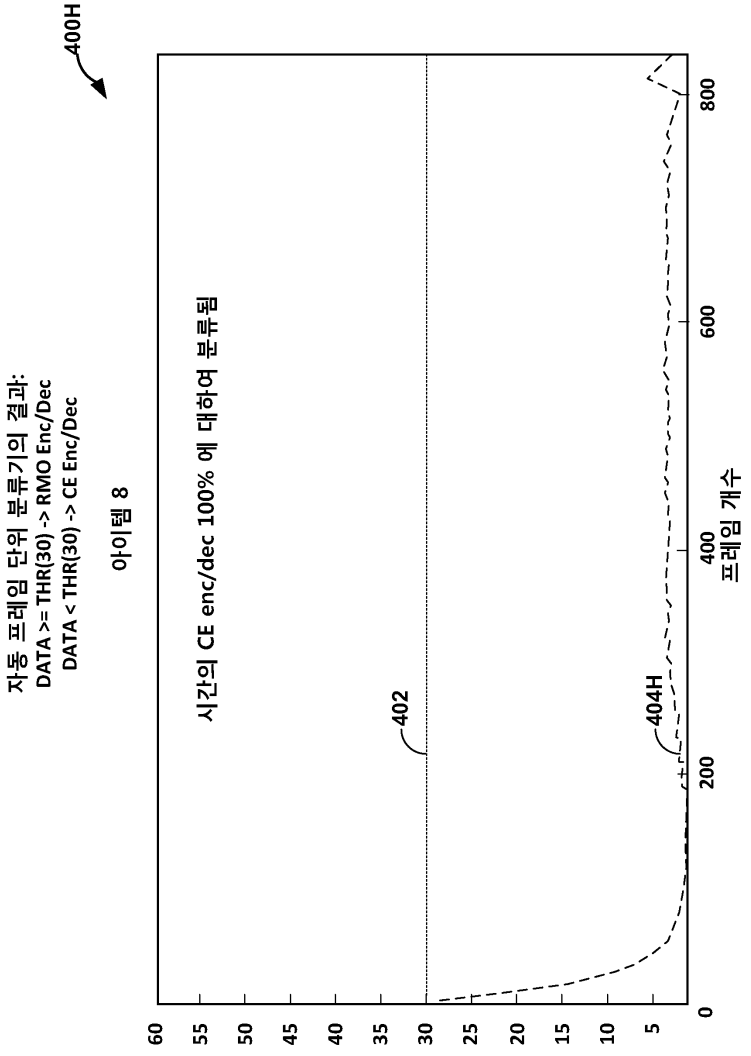
도면31f


400F
자동 프레임 단위 분류기의 결과:
DATA >= THR(30) -> RMO Enc/Dec
DATA < THR(30) -> CE Enc/Dec
아이템 6
시간의 CE enc/dec 100% 에 대하여 분류됨
402
404F
60
55
50
45
40
35
30
25
20
15
10
5
0
200
400
600
800
프레임 개수

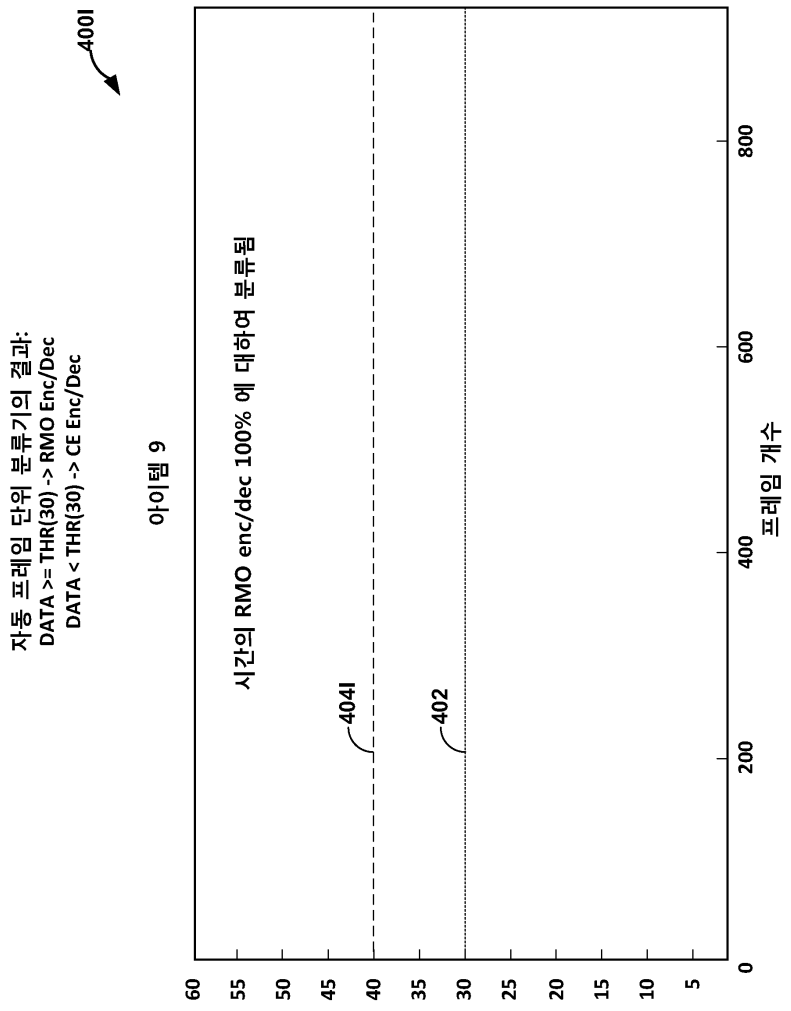
도면31g


400G
자동 프레임 단위 분류기의 결과:
DATA >= THR(30) -> RMO Enc/Dec
DATA < THR(30) -> CE Enc/Dec
아이템 7
시간의 CE enc/dec 100% 에 대하여 분류됨
402
404G
60
55
50
45
40
35
30
25
20
15
10
5
0
200
400
600
800
프레임 개수

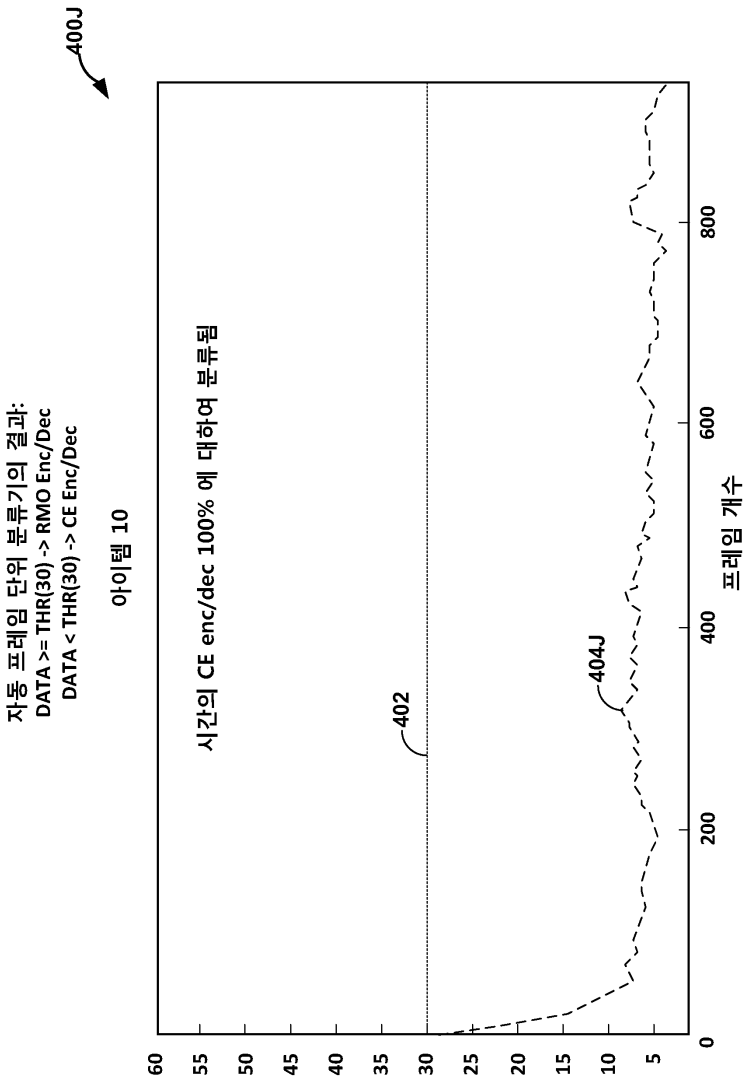
도면31h


400H
자동 프레임 단위 분류기의 결과:
DATA >= THR(30) -> RMO Enc/Dec
DATA < THR(30) -> CE Enc/Dec
아이템 8
시간의 CE enc/dec 100% 에 대하여 분류됨
402
404H
60
55
50
45
40
35
30
25
20
15
10
5
0
200
400
600
800
프레임 개수

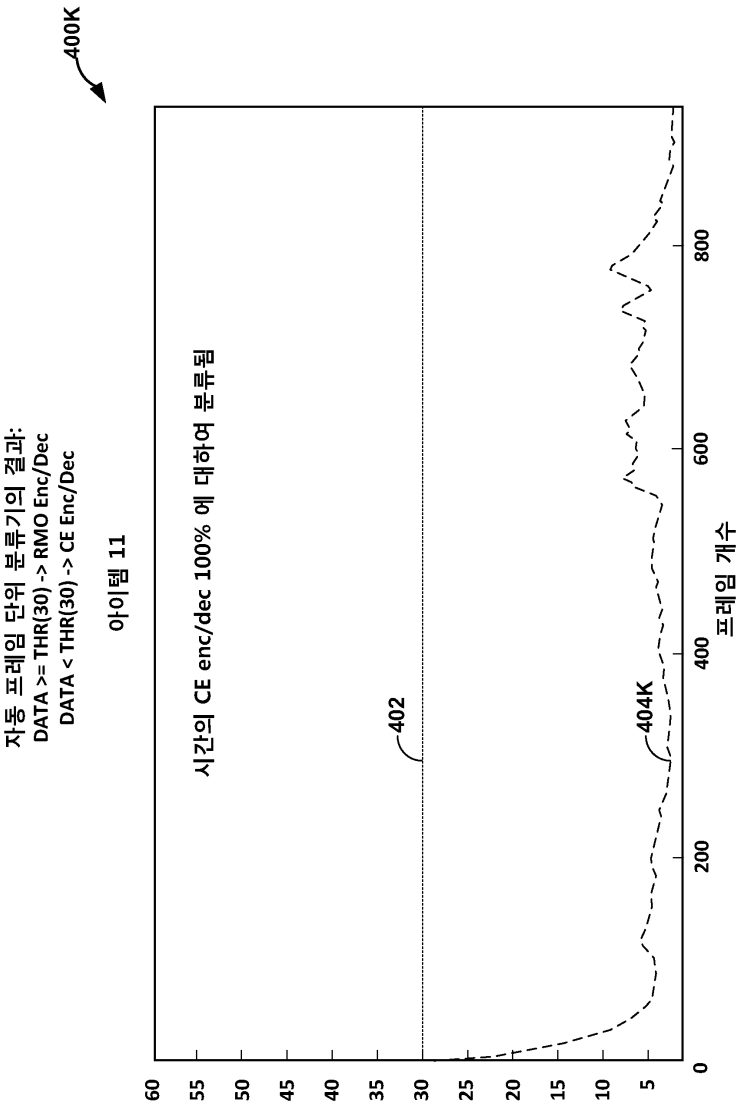
도면31i


400I
자동 프레임 단위 분류기의 결과:
DATA >= THR(30) -> RMO Enc/Dec
DATA < THR(30) -> CE Enc/Dec
아이템 9
시간의 RMO enc/dec 100% 에 대하여 분류됨
404I
402
60
55
50
45
40
35
30
25
20
15
10
5
0
200
400
600
800
프레임 개수

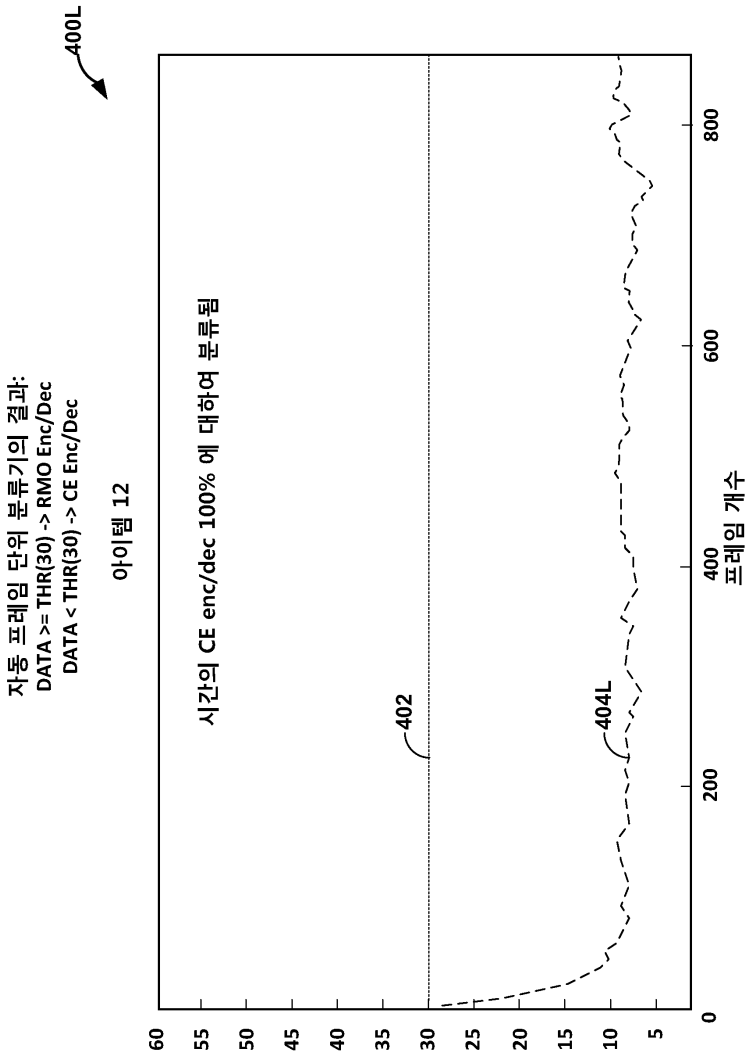
도면31j


400J
자동 프레임 단위 분류기의 결과:
DATA >= THR(30) -> RMO Enc/Dec
DATA < THR(30) -> CE Enc/Dec
아이템 10
시간의 CE enc/dec 100% 에 대하여 분류됨
402
404J
60
55
50
45
40
35
30
25
20
15
10
5
0
200
400
600
800
프레임 개수

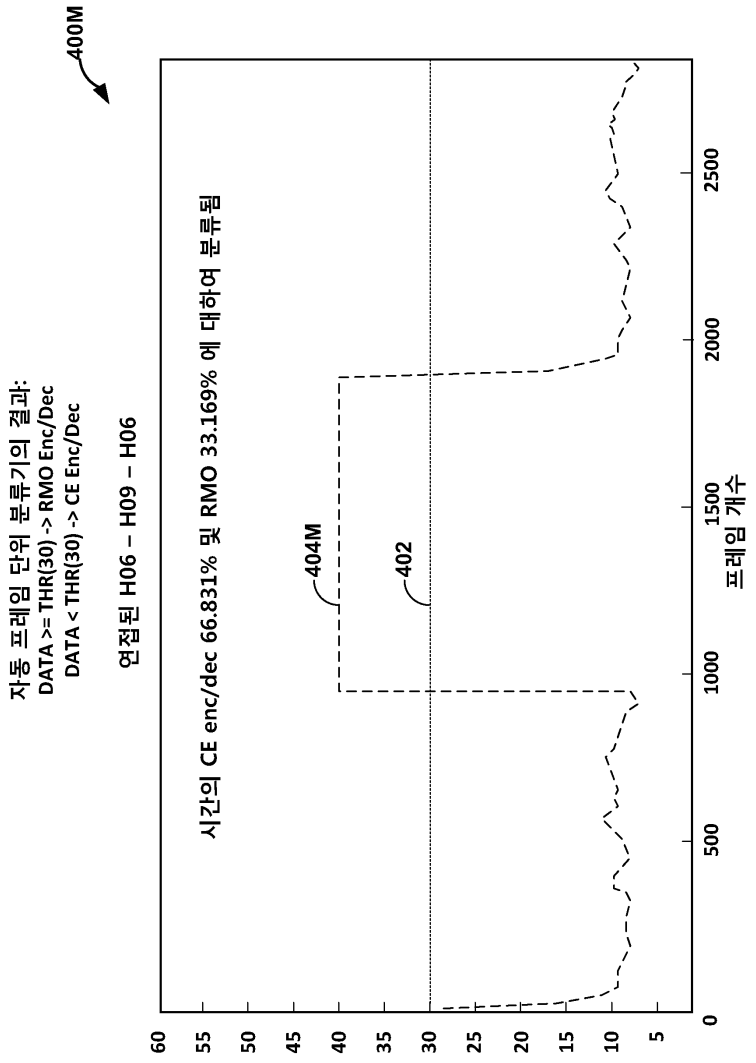
도면31k


400K
자동 프레임 단위 분류기의 결과:
DATA >= THR(30) -> RMO Enc/Dec
DATA < THR(30) -> CE Enc/Dec
아이템 11
시간의 CE enc/dec 100% 에 대하여 분류됨
402
404K
60
55
50
45
40
35
30
25
20
15
10
5
0
200
400
600
800
프레임 개수

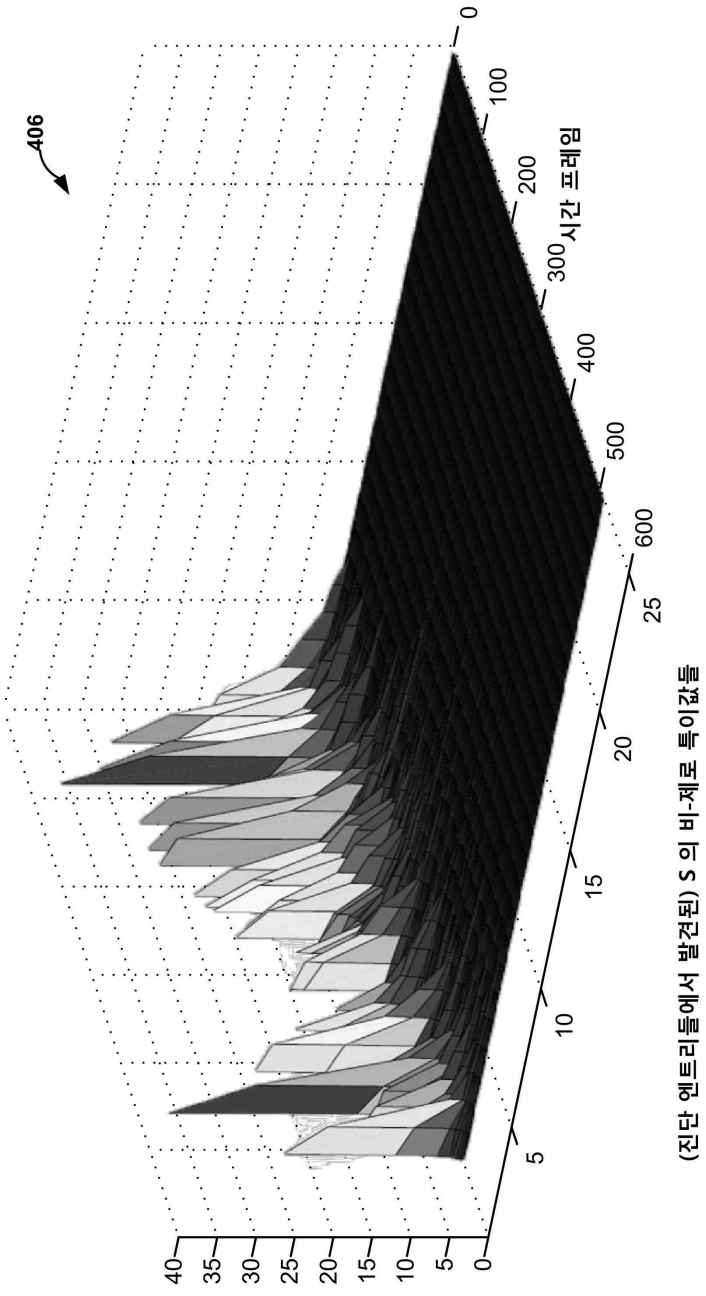
도면311


400L
자동 프레임 단위 분류기의 결과:
DATA >= THR(30) -> RMO Enc/Dec
DATA < THR(30) -> CE Enc/Dec
아이템 12
시간의 CE enc/dec 100% 에 대하여 분류됨
402
404L
60
55
50
45
40
35
30
25
20
15
10
5
0
200
400
600
800
프레임 개수

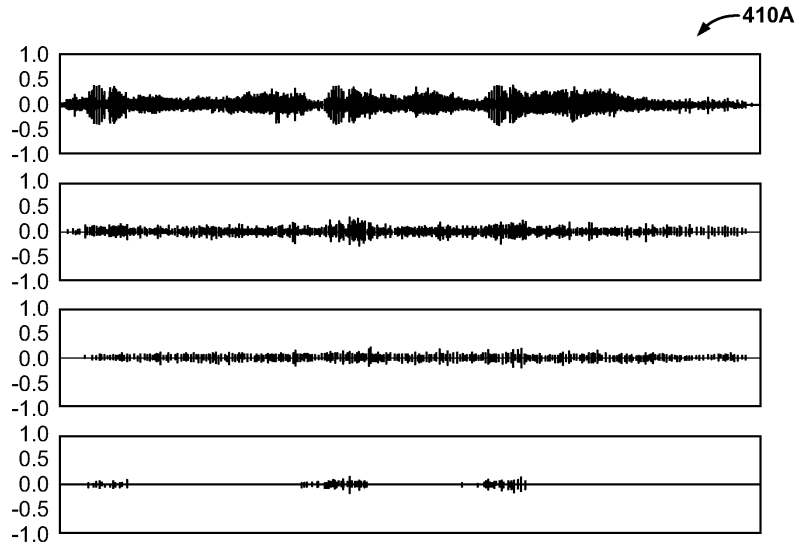
도면31m


400M
자동 프레임 단위 분류기의 결과:
DATA >= THR(30) -> RMO Enc/Dec
DATA < THR(30) -> CE Enc/Dec
연접된 H06 – H09 – H06
시간의 CE enc/dec 66.831% 및 RMO 33.169% 에 대하여 분류됨
404M
402
60
55
50
45
40
35
30
25
20
15
10
5
0
500
1000
1500
2000
2500
프레임 개수

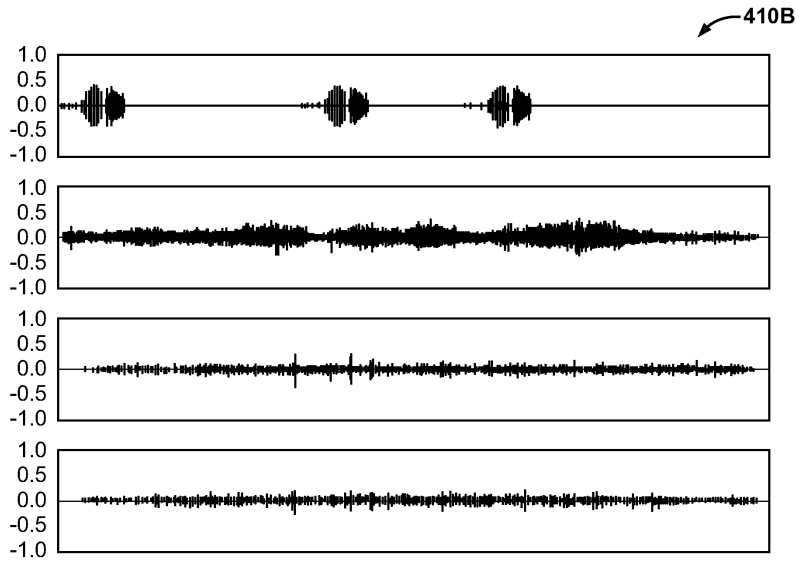
도면32


406
0
100
200
300
400
500
600
시간 프레임
0
5
10
15
20
25
(진단 엔트리들에서 발견된) S 의 비-제로 특이값들
0
5
10
15
20
25
30
35
40

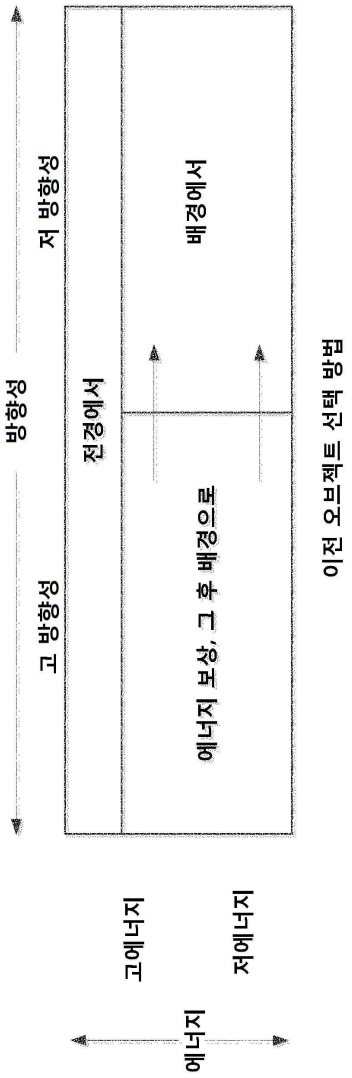
도면33a



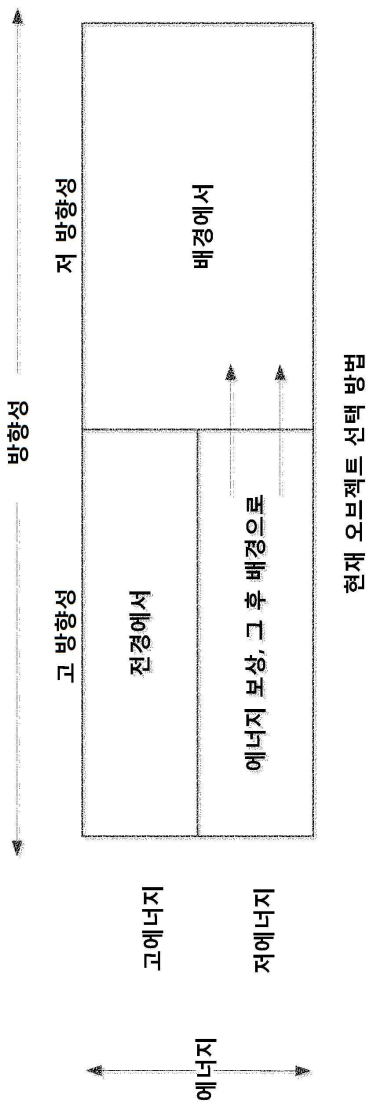
도면33b



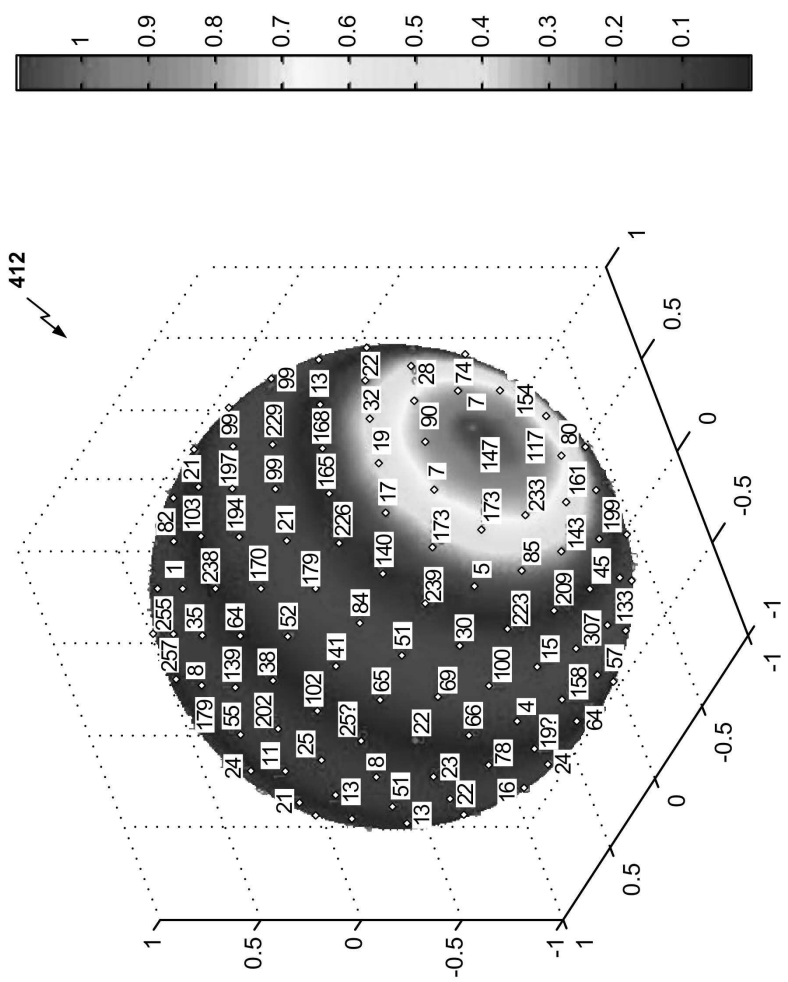
도면34


방향성
저 방향성
고 방향성
전경에서
배경에서
에너지 보상, 그 후 배경으로
이전 오브젝트 선택 방법
고에너지
저에너지
에너지

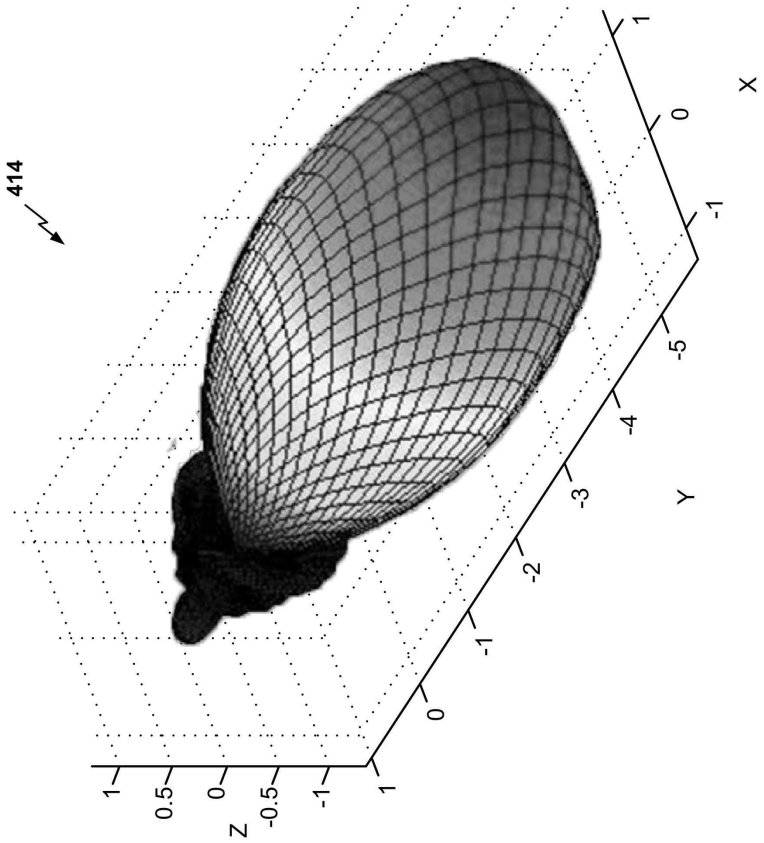
도면35


방향성
고 방향성
저 방향성
전경에서
배경에서
에너지 보상, 그 후 배경으로
현재 오브젝트 선택 방법
고에너지
저에너지
에너지

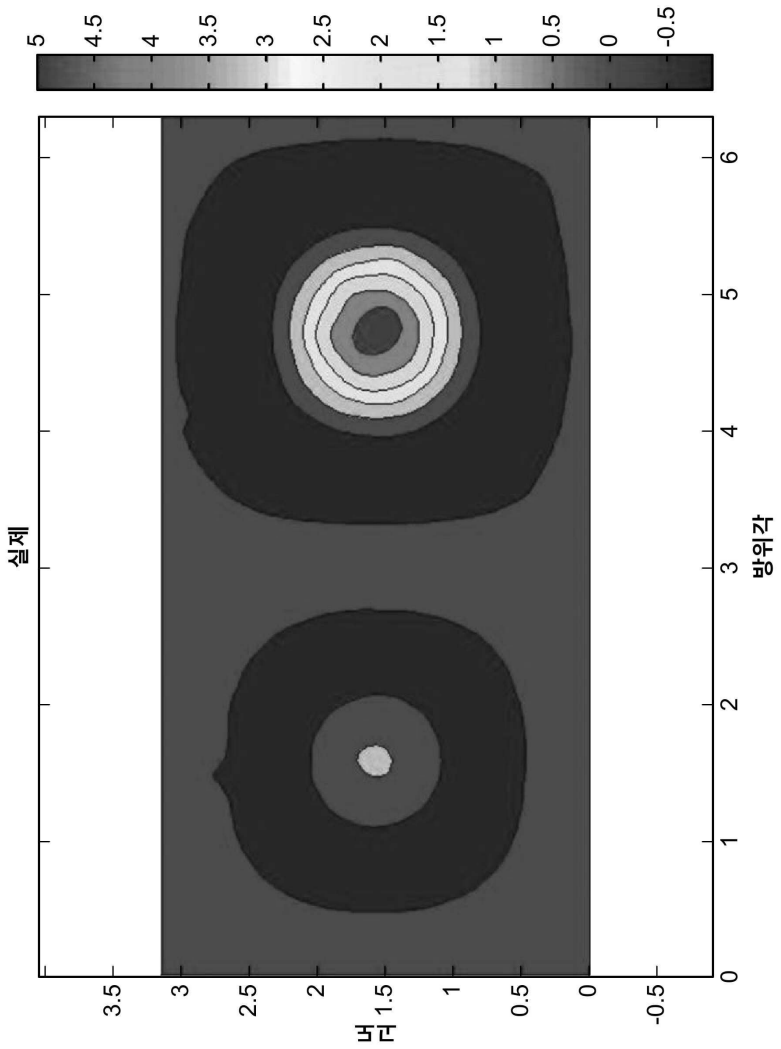
도면36a



도면36b

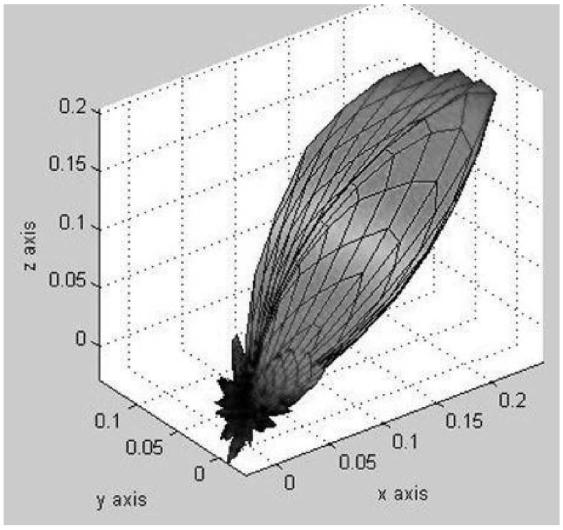


도면36c



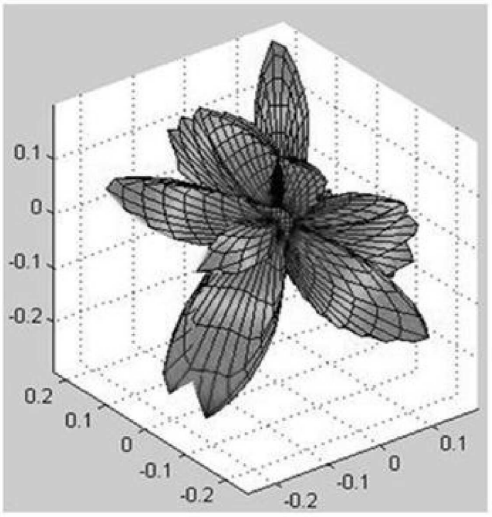
도면36d

벌의 사운드는 방위각 = 0° 및 고도 = 45° 에서 생성된다.



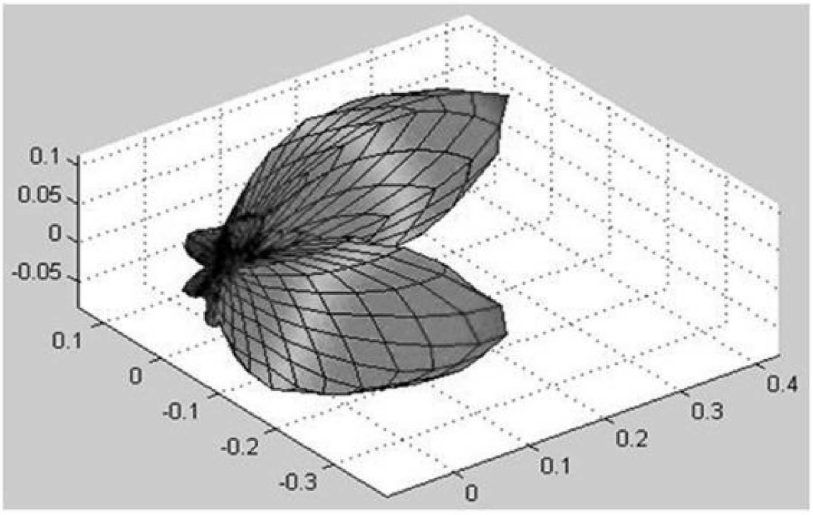
도면36e

헬리콥터 사운드는 하늘에서 생성된다.



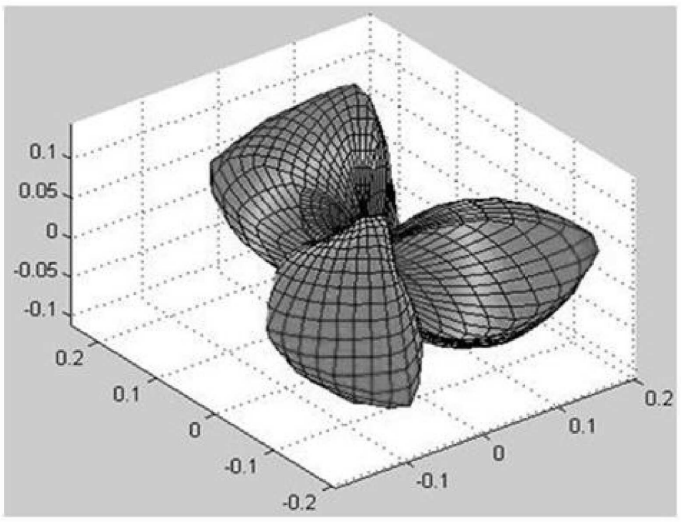
도면36f

현대 전자 음악은 2 개의 상이한 방향들에서 생성된다.

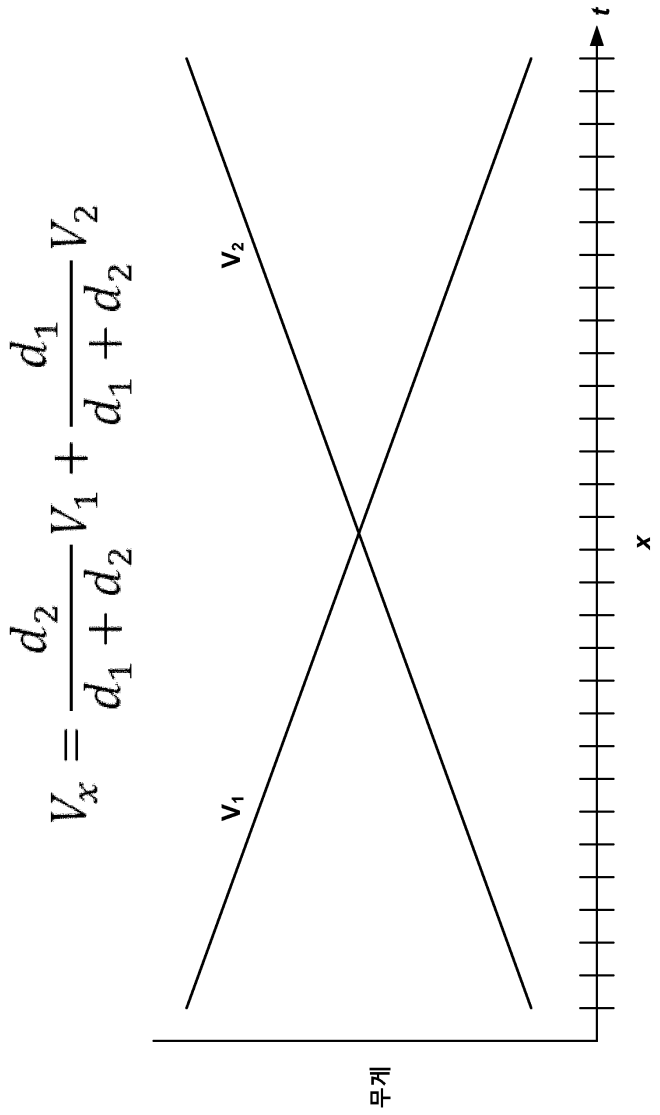


도면36g

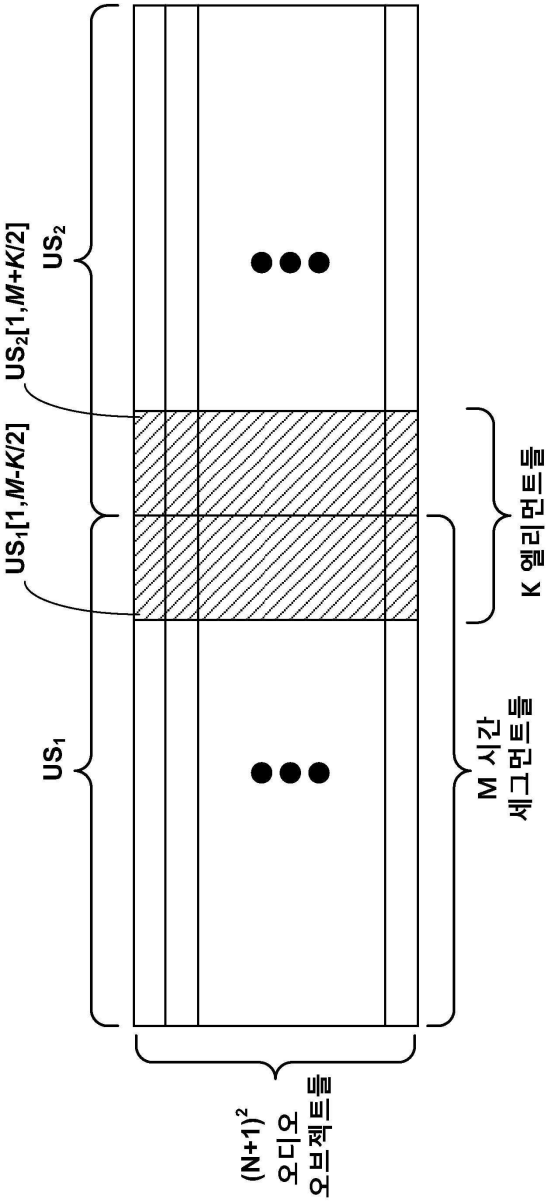
사람들이 체육관에서 소리지르고 있다.



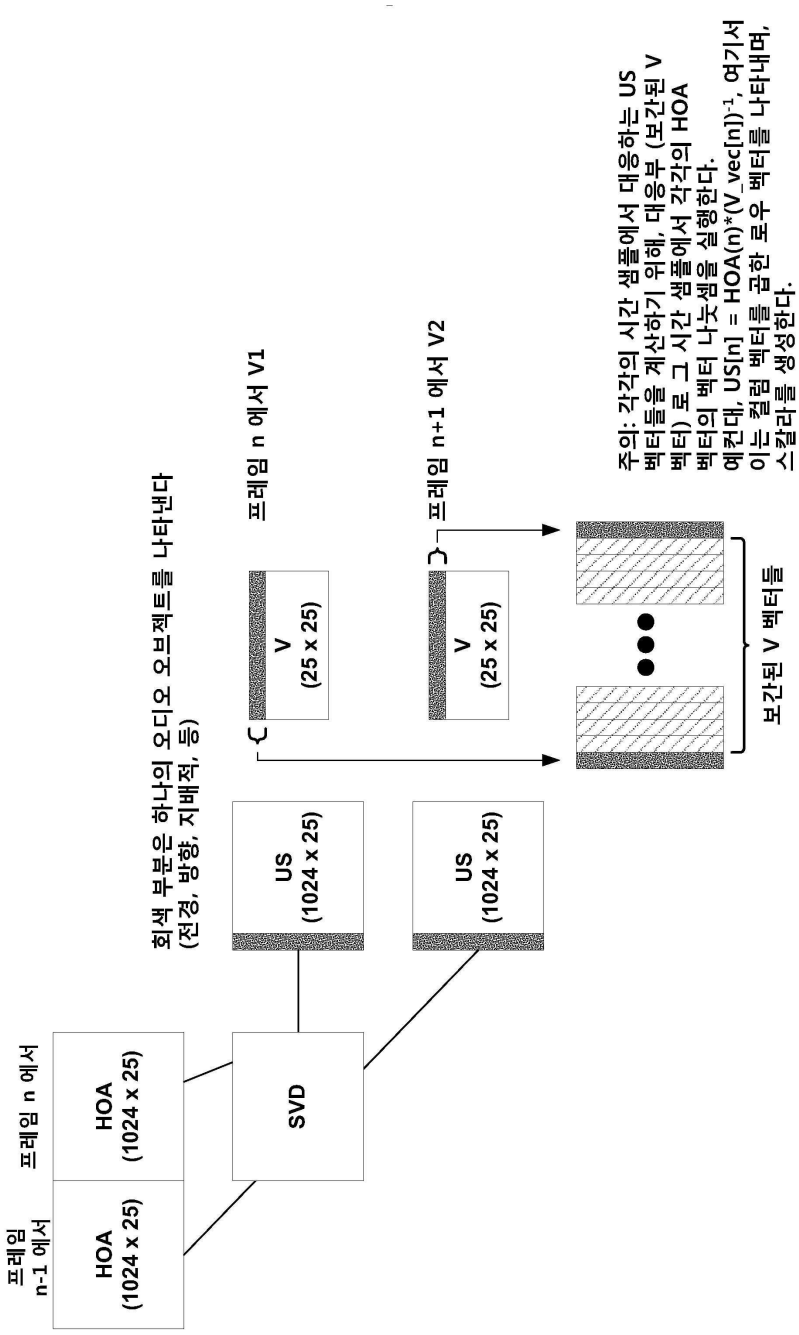
도면37



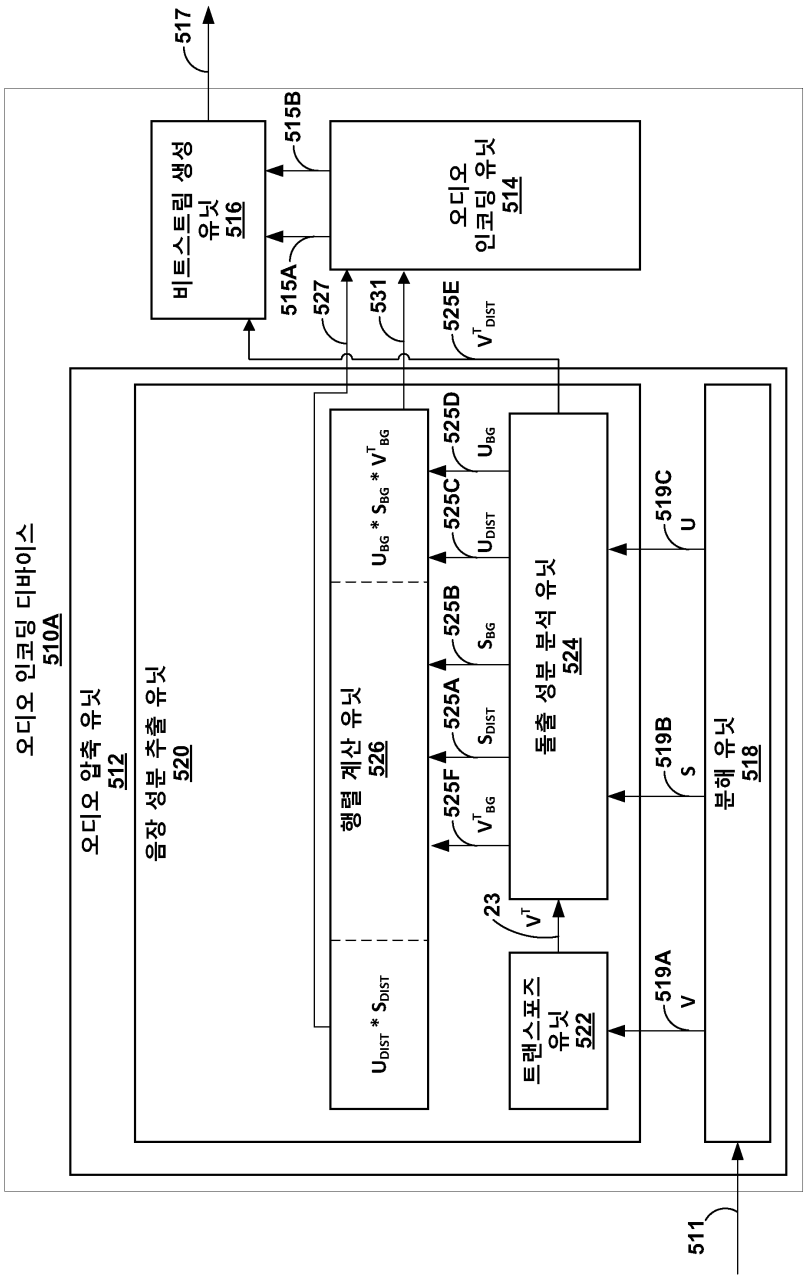
도면38



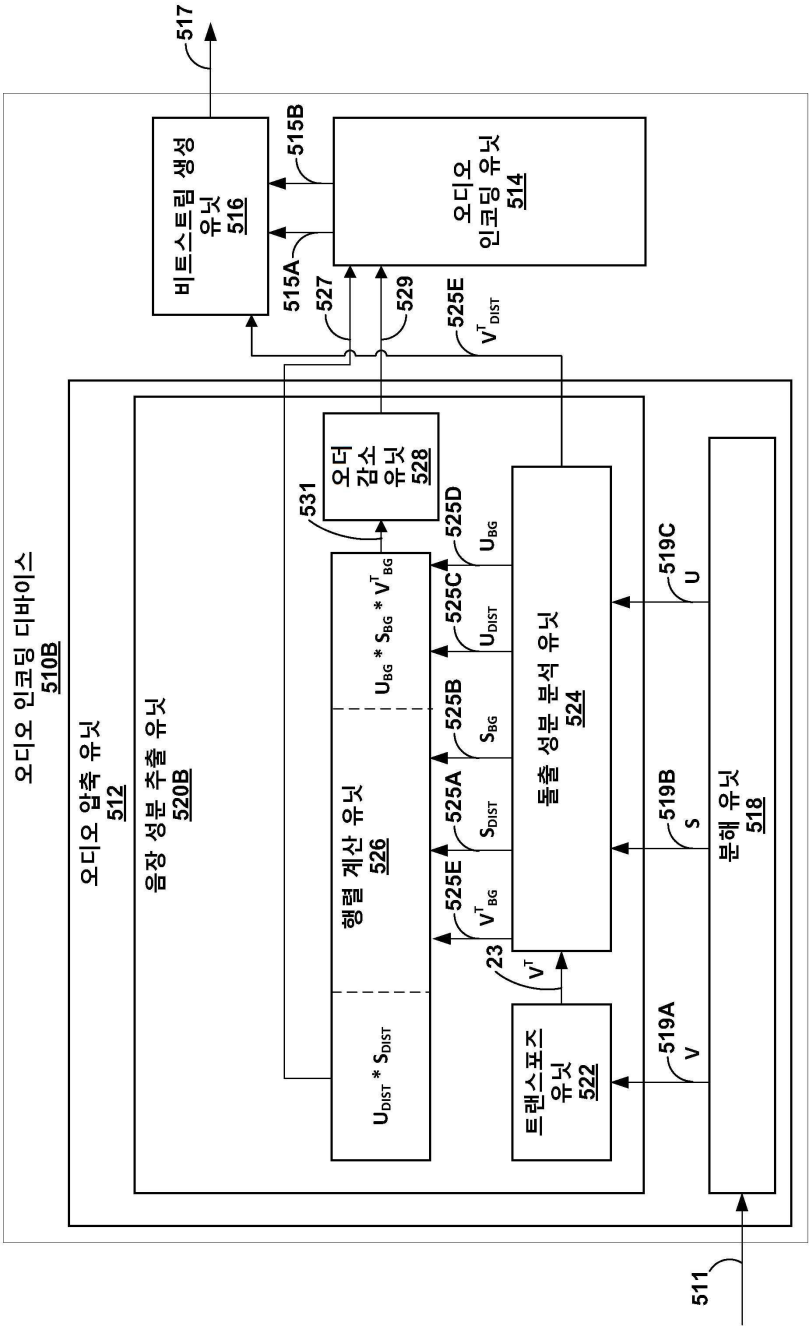
도면39


프레임 n-1 에서
HOA (1024 x 25)
프레임 n 에서
HOA (1024 x 25)
SVD
US (1024 x 25)
US (1024 x 25)
회색 부분은 하나의 오디오 오브젝트를 나타낸다
(전경, 방향, 지배적, 등)
프레임 n 에서 V1
V (25 x 25)
프레임 n+1 에서 V2
V (25 x 25)
보간된 V 벡터들
주의: 각각의 시간 샘플에서 대응하는 US 벡터들을 계산하기 위해, 대응부 (보간된 V 벡터) 로 그 시간 샘플에서 각각의 HOA 벡터의 벡터 나눗셈을 실행한다.
예컨대, US[n] = HOA(n)*(V_vec[n])-1, 여기서 이는 컬럼 벡터를 곱한 로우 벡터를 나타내며, 스칼라를 생성한다.

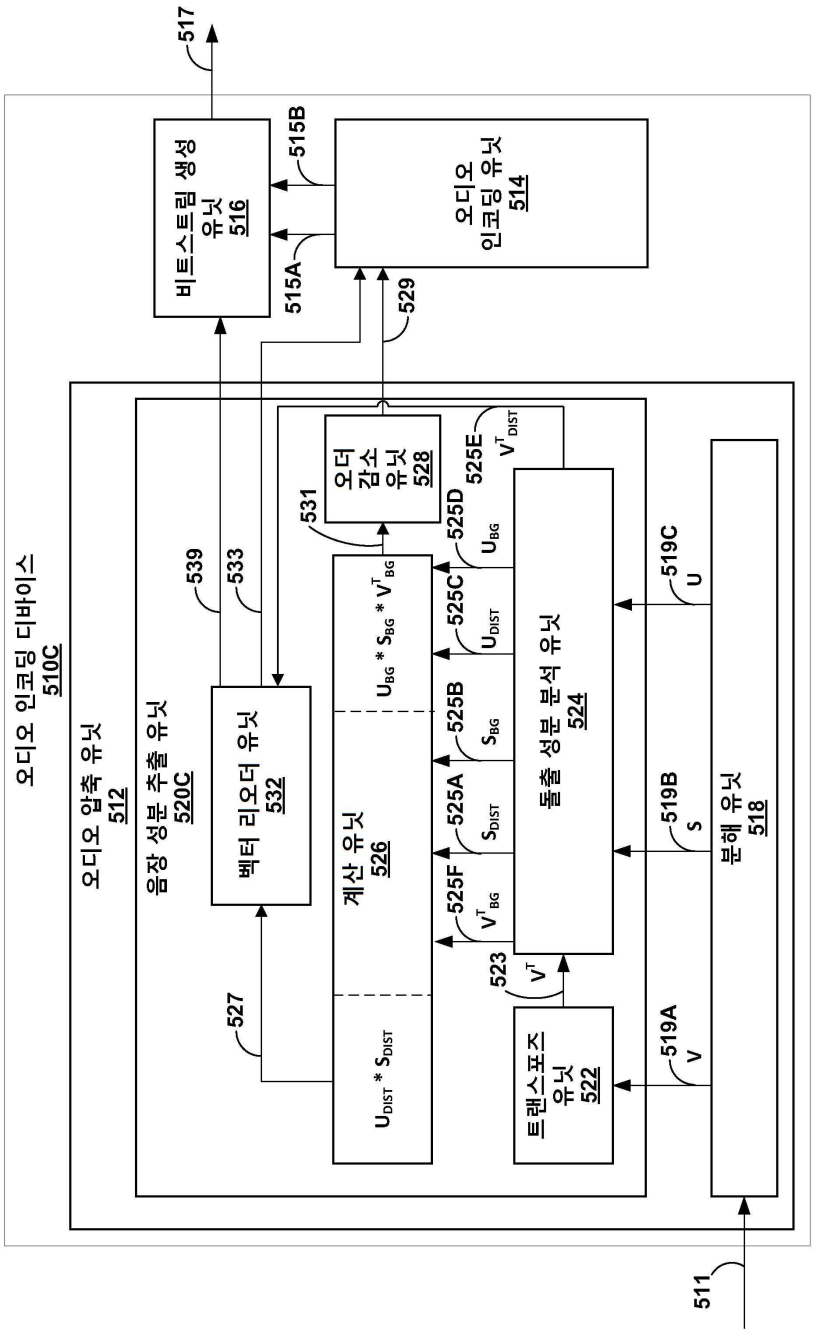
도면40a


517
515B
515A
527
531
525E
V^T_{DIST}
비트스트림 생성 유닛 516
오디오 인코딩 유닛 514
오디오 인코딩 디바이스 510A
오디오 압축 유닛 512
음장 성분 추출 유닛 520
행렬 계산 유닛 526
$U_{BG} * S_{BG} * V^T_{BG}$
$U_{DIST} * S_{DIST}$
525D
525C
525B
525A
525F
U_{BG}
U_{DIST}
S_{BG}
S_{DIST}
V^T_{BG}
돌출 성분 분석 유닛 524
23
V^T
트랜스포즈 유닛 522
519C
U
519B
S
519A
V
분해 유닛 518
511

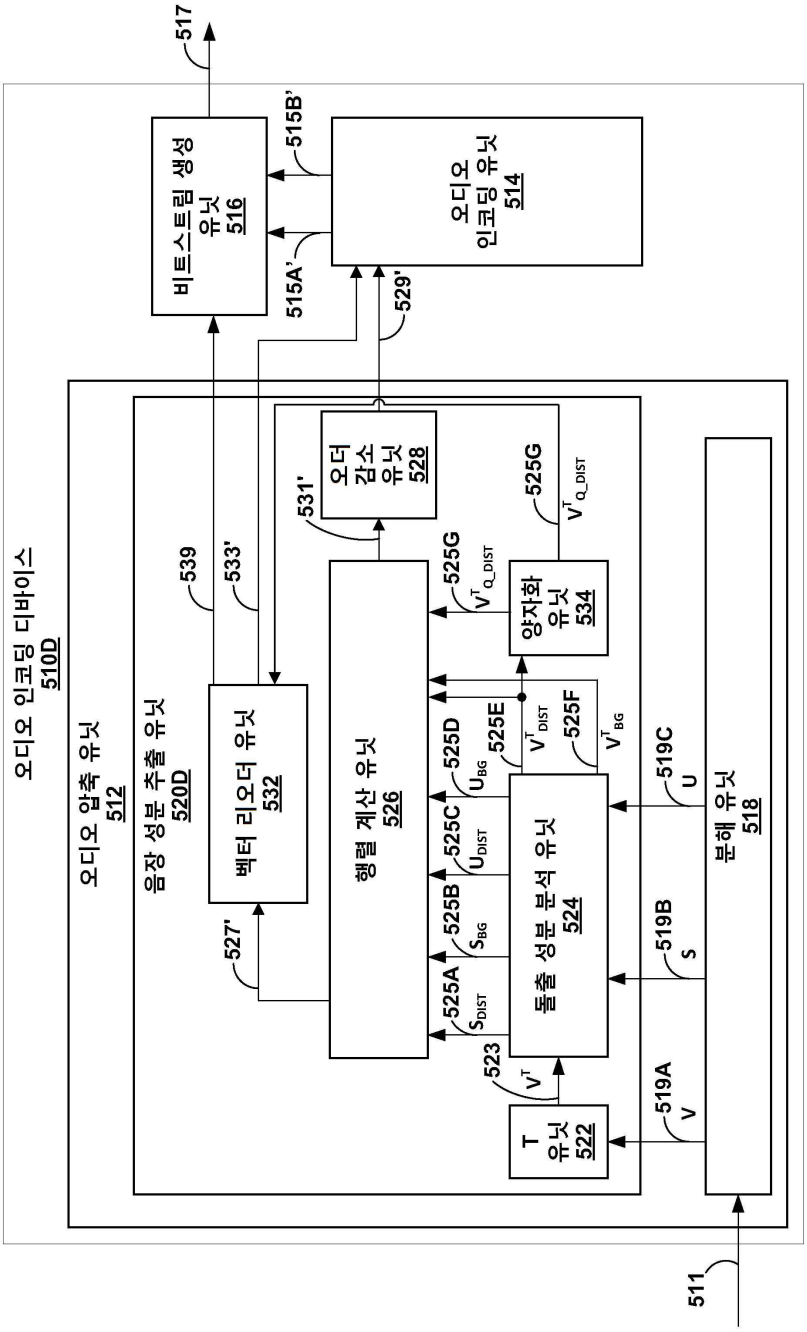
도면40b


517
비트스트림 생성 유닛 516
515B
515A
오디오 인코딩 유닛 514
527
529
525E V^T_{DIST}
오디오 인코딩 디바이스 510B
오디오 압축 유닛 512
음장 성분 추출 유닛 520B
531
오더 감소 유닛 528
525D U_{BG}
525C U_{DIST}
525B S_{BG}
525A S_{DIST}
525E V^T_{BG}
$U_{BG} * S_{BG} * V^T_{BG}$
행렬 계산 유닛 526
$U_{DIST} * S_{DIST}$
23 V^T
트랜스포즈 유닛 522
519A V
519B S
519C U
돌출 성분 분석 유닛 524
분해 유닛 518
511

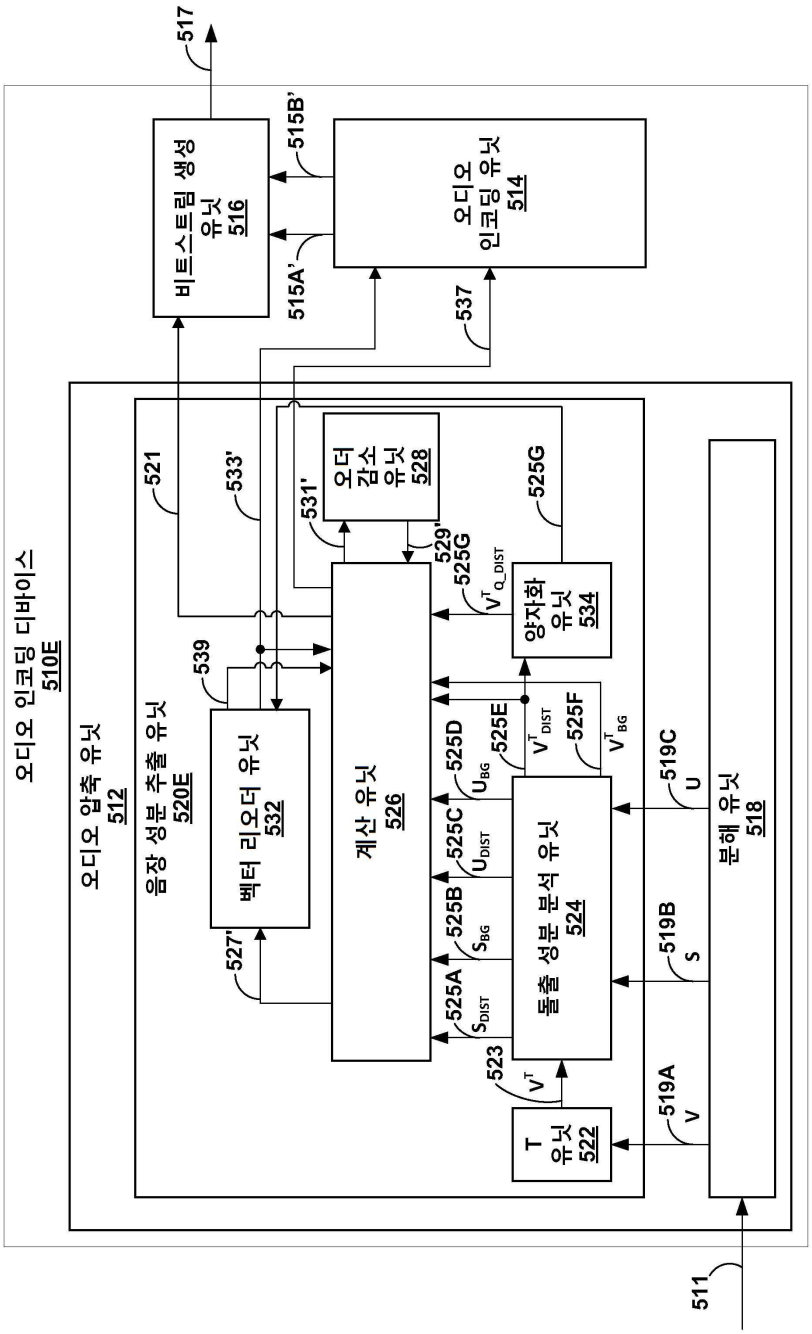
도면40c


517
515B
515A
비트스트림 생성 유닛 516
오디오 인코딩 유닛 514
529
오디오 인코딩 디바이스 510C
오디오 압축 유닛 512
음장 성분 추출 유닛 520C
539
533
531
오더 감소 유닛 528
525E
V^T_{DIST}
525D
U_{BG}
525C
U_{DIST}
525B
S_{BG}
525A
S_{DIST}
525F
V^T_{BG}
519C
U
519B
S
519A
V
벡터 리오더 유닛 532
계산 유닛 526
$U_{BG} * S_{BG} * V^T_{BG}$
$U_{DIST} * S_{DIST}$
돌출 성분 분석 유닛 524
분해 유닛 518
527
523
V^T
트랜스포즈 유닛 522
511

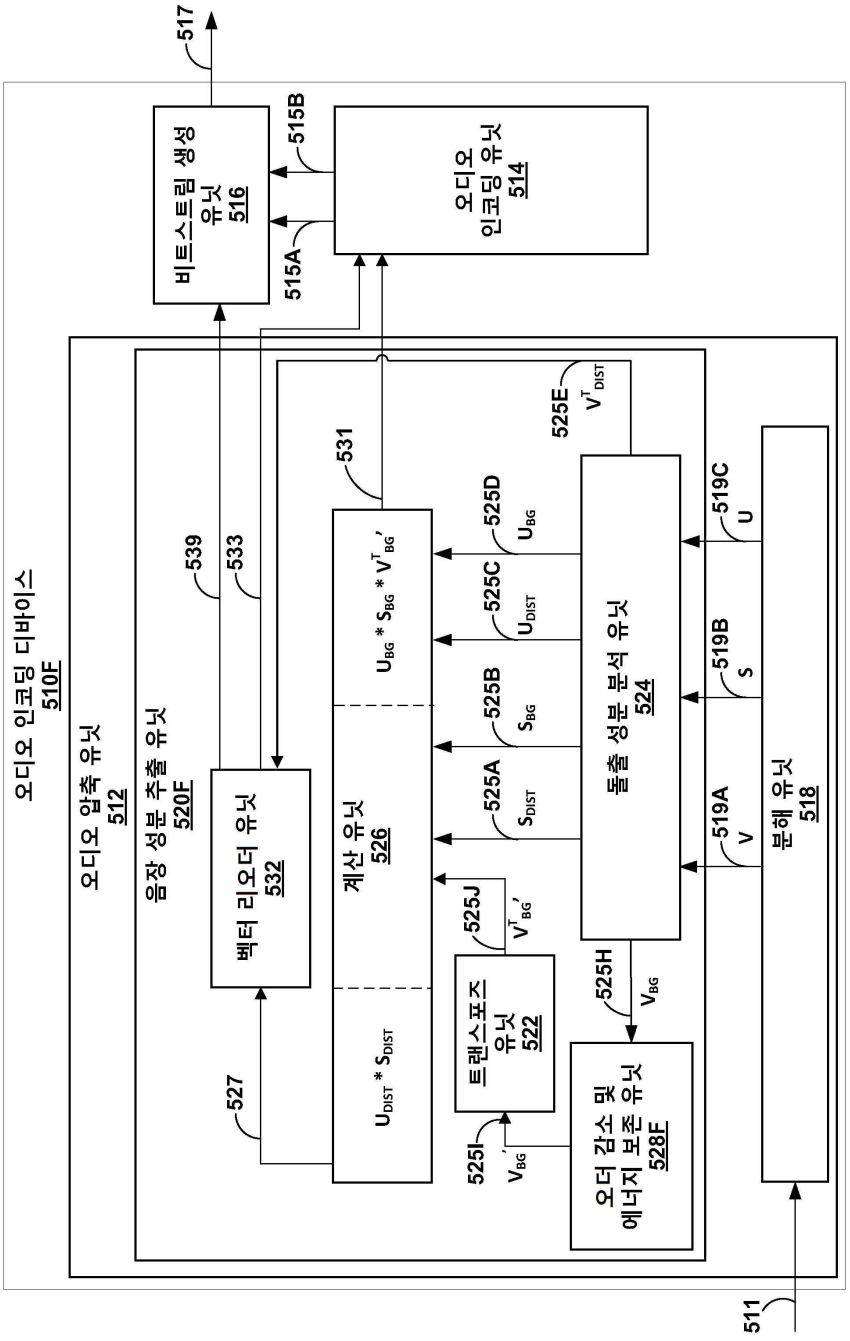
도면40d


517
515B'
515A'
529'
비트스트림 생성 유닛 516
오디오 인코딩 유닛 514
오디오 인코딩 디바이스 510D
오디오 압축 유닛 512
음장 성분 추출 유닛 520D
539
533'
531'
527'
벡터 리오더 유닛 532
행렬 계산 유닛 526
오더 감소 유닛 528
양자화 유닛 534
525A
525B
525C
525D
525E
525F
525G
돌출 성분 분석 유닛 524
T 유닛 522
523
519A
519B
519C
분해 유닛 518
511

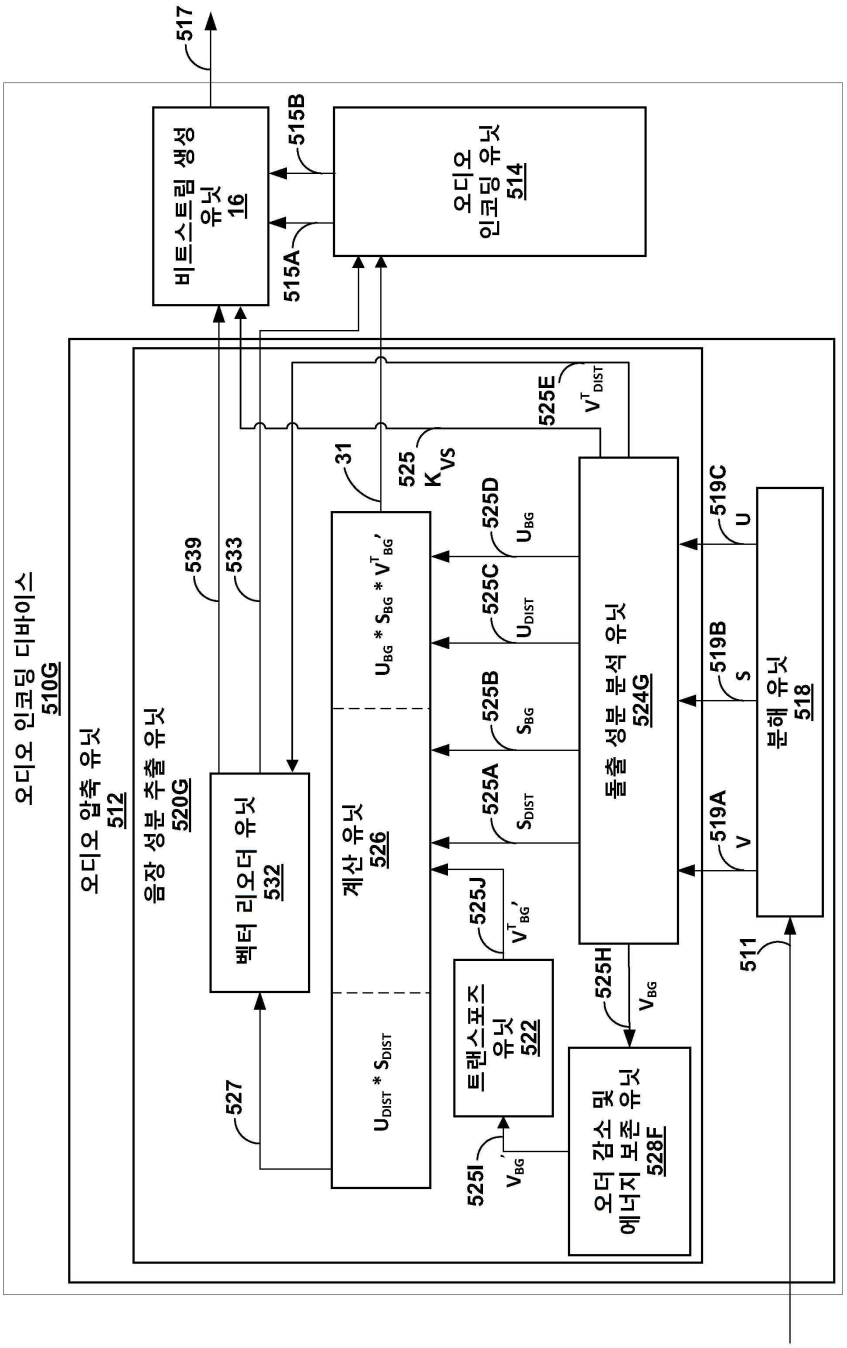
도면40e


517
515B'
515A'
비트스트림 생성 유닛 516
오디오 인코딩 유닛 514
537
오디오 인코딩 디바이스 510E
오디오 압축 유닛 512
521
533'
531'
오더 감소 유닛 528
529'
525G
음장 성분 추출 유닛 520E
539
벡터 리오더 유닛 532
527'
계산 유닛 526
양자화 유닛 534
525D
525E
525F
525C
525B
525A
523
돌출 성분 분석 유닛 524
T 유닛 522
519C
519B
519A
분해 유닛 518
511

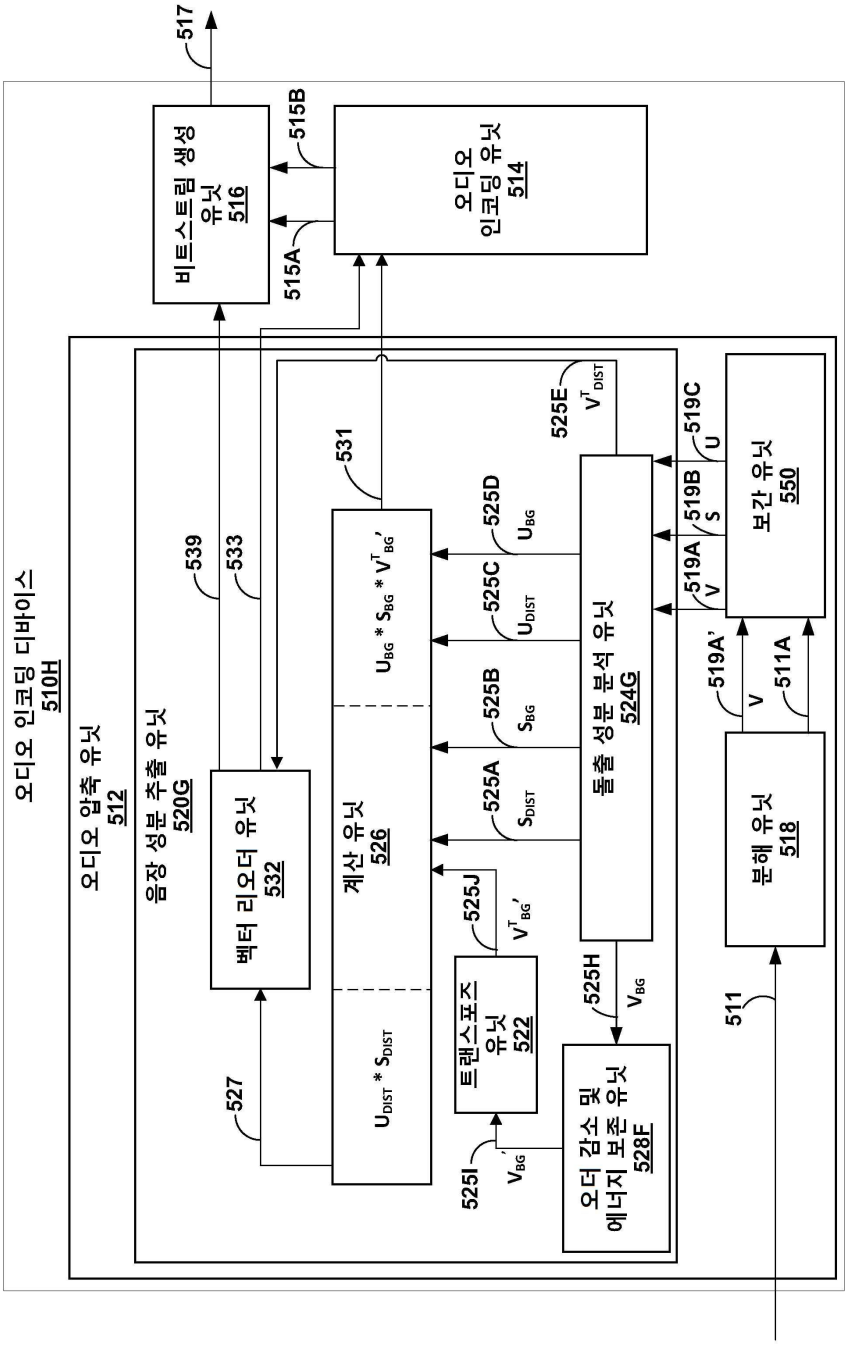
도면40f


517
515B
515A
비트스트림 생성 유닛 516
오디오 인코딩 유닛 514
오디오 인코딩 디바이스 510F
오디오 압축 유닛 512
음장 성분 추출 유닛 520F
539
533
531
525E
V^T_{DIST}
525D
U_{BG}
525C
U_{DIST}
525B
S_{BG}
525A
S_{DIST}
525J
$V^T_{BG}{}'$
525H
V_{BG}
519C
U
519B
S
519A
V
벡터 리오더 유닛 532
계산 유닛 526
$U_{BG} * S_{BG} * V^T_{BG}{}'$
$U_{DIST} * S_{DIST}$
돌출 성분 분석 유닛 524
분해 유닛 518
527
트랜스포즈 유닛 522
525I
$V_{BG}{}'$
오더 감소 및 에너지 보존 유닛 528F
511

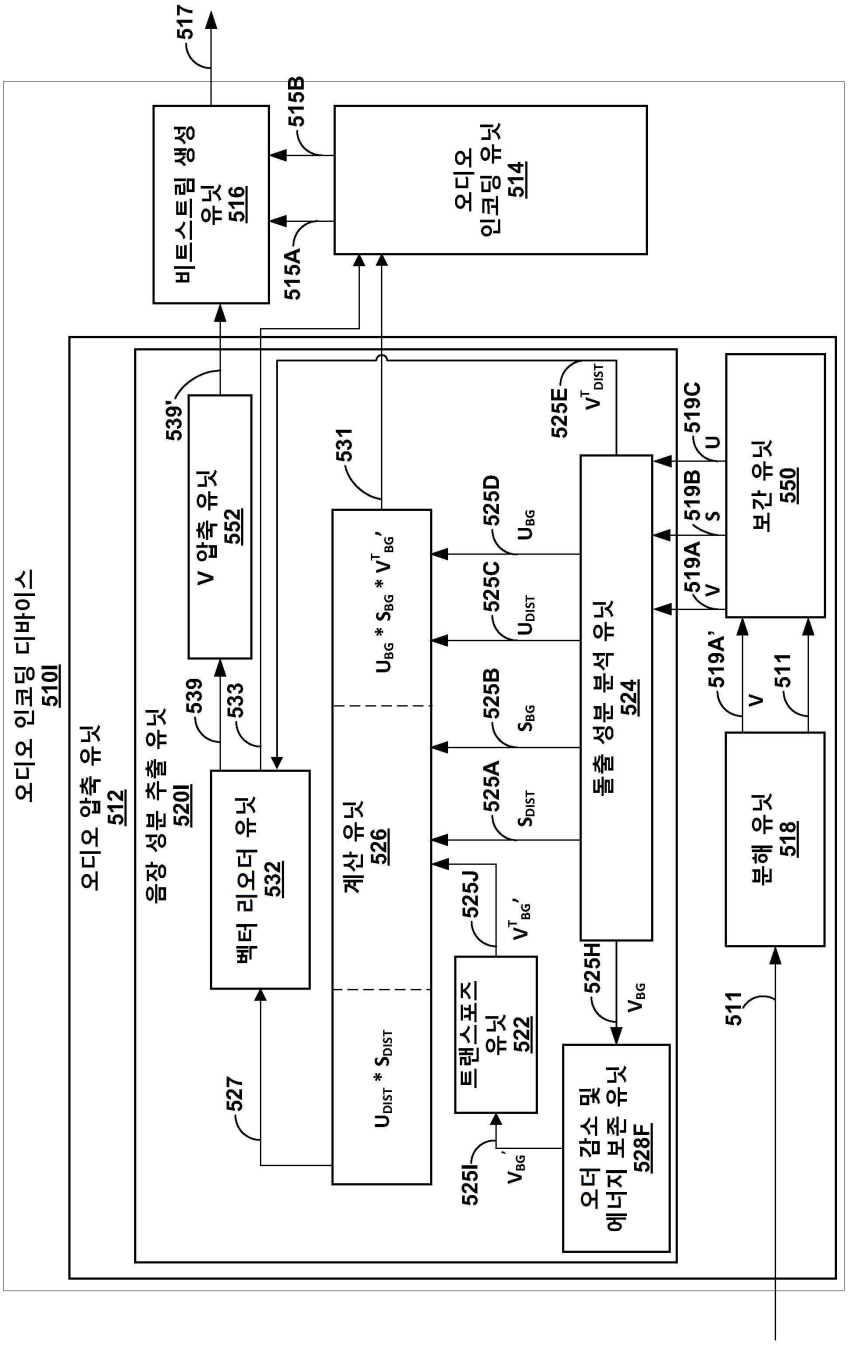
도면40g


517
515B
515A
비트스트림 생성 유닛 16
오디오 인코딩 유닛 514
오디오 인코딩 디바이스 510G
오디오 압축 유닛 512
음장 성분 추출 유닛 520G
539
533
527
벡터 리오더 유닛 532
계산 유닛 526
31
525
K_{VS}
525E
V^T_{DIST}
525D
U_{BG}
525C
U_{DIST}
525B
S_{BG}
525A
S_{DIST}
525J
V^T_{BG}'
525I
V_{BG}'
$U_{BG} * S_{BG} * V^T_{BG}'$
$U_{DIST} * S_{DIST}$
트랜스포즈 유닛 522
오더 감소 및 에너지 보존 유닛 528F
돌출 성분 분석 유닛 524G
525H
V_{BG}
519C
U
519B
S
519A
V
분해 유닛 518
511

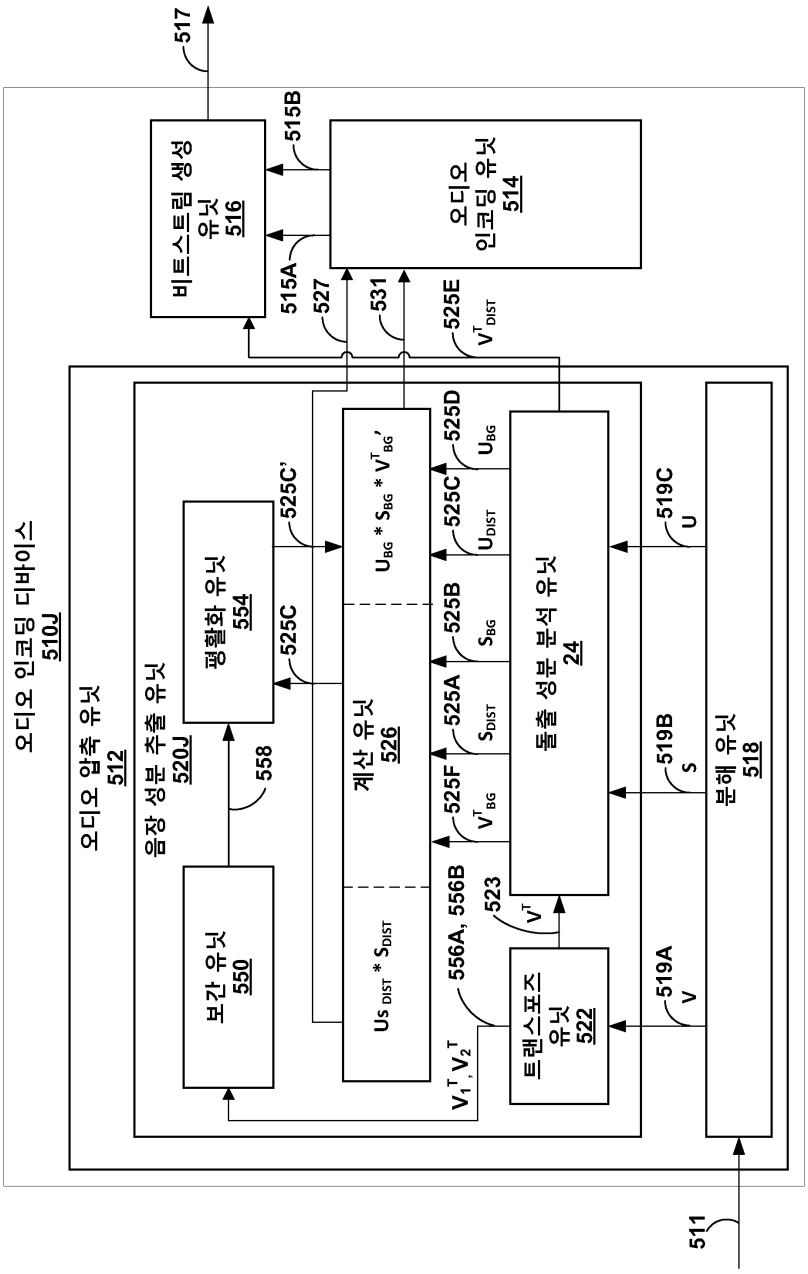
도면40h


517
비트스트림 생성 유닛 516
515B
515A
오디오 인코딩 유닛 514
오디오 인코딩 디바이스 510H
오디오 압축 유닛 512
음장 성분 추출 유닛 520G
539
533
531
525E
519C
519B
519A
519A'
511A
511
527
525I
525J
525H
525D
525C
525B
525A
벡터 리오더 유닛 532
계산 유닛 526
트랜스포즈 유닛 522
오더 감소 및 에너지 보존 유닛 528F
돌출 성분 분석 유닛 524G
분해 유닛 518
보간 유닛 550

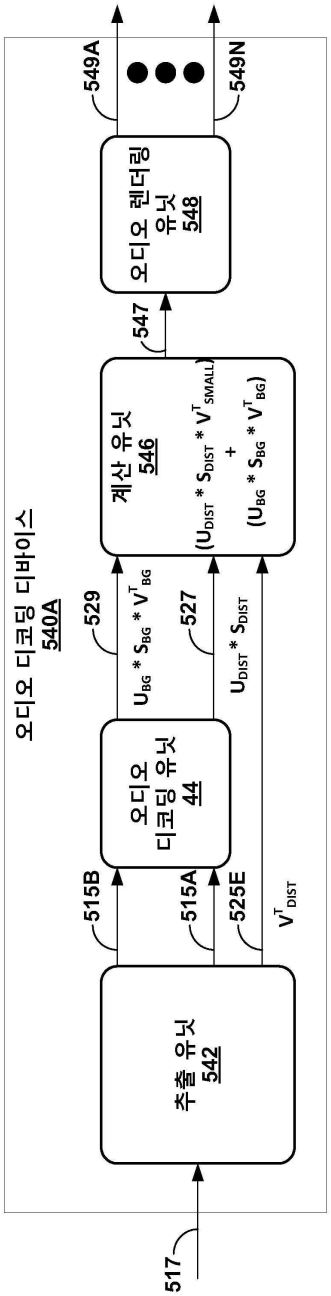
도면40i


517
515B
515A
비트스트림 생성 유닛 516
오디오 인코딩 유닛 514
오디오 인코딩 디바이스 510I
오디오 압축 유닛 512
음장 성분 추출 유닛 520I
539'
539
533
527
V 압축 유닛 552
벡터 리오더 유닛 532
531
계산 유닛 526
$U_{DIST} * S_{DIST}$
$U_{BG} * S_{BG} * V^{T}_{BG}{}'$
525I
$V_{BG}{}'$
트랜스포즈 유닛 522
525J
$V^{T}_{BG}{}'$
525A
S_{DIST}
525B
S_{BG}
525C
U_{DIST}
525D
U_{BG}
525E
V^{T}_{DIST}
돌출 성분 분석 유닛 524
525H
V_{BG}
오디 감소 및 에너지 보존 유닛 528F
519C
U
519B
S
519A
V
보간 유닛 550
519A'
V
511
분해 유닛 518
511

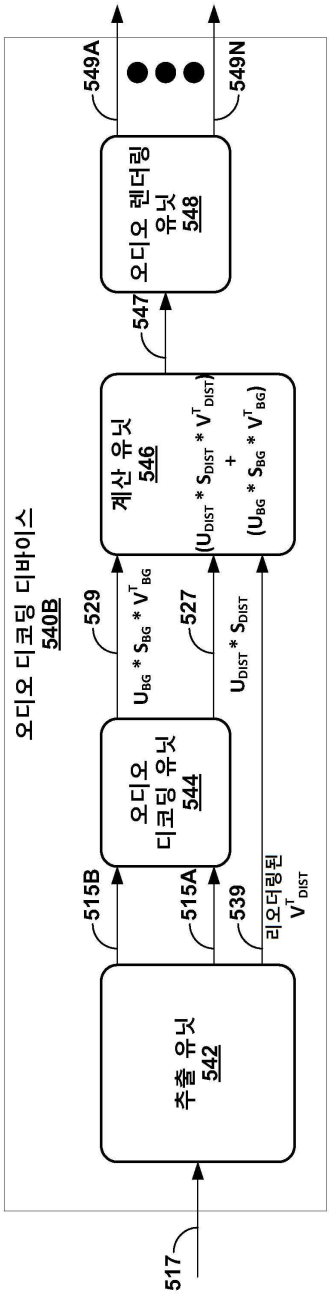
도면40j


517
515B
비트스트림 생성 유닛 516
오디오 인코딩 유닛 514
515A
527
531
525E
V^T_{DIST}
오디오 인코딩 디바이스 510J
오디오 압축 유닛 512
음장 성분 추출 유닛 520J
평활화 유닛 554
525C'
525C
558
보간 유닛 550
계산 유닛 526
$U_{BG} * S_{BG} * V^T_{BG}'$
$U_{S\ DIST} * S_{DIST}$
525D
U_{BG}
525C
U_{DIST}
525B
S_{BG}
525A
S_{DIST}
525F
V^T_{BG}
556A, 556B
V_1^T, V_2^T
523
V^T
돌출 성분 분석 유닛 24
519C
U
519B
S
519A
V
트랜스포즈 유닛 522
분해 유닛 518
511

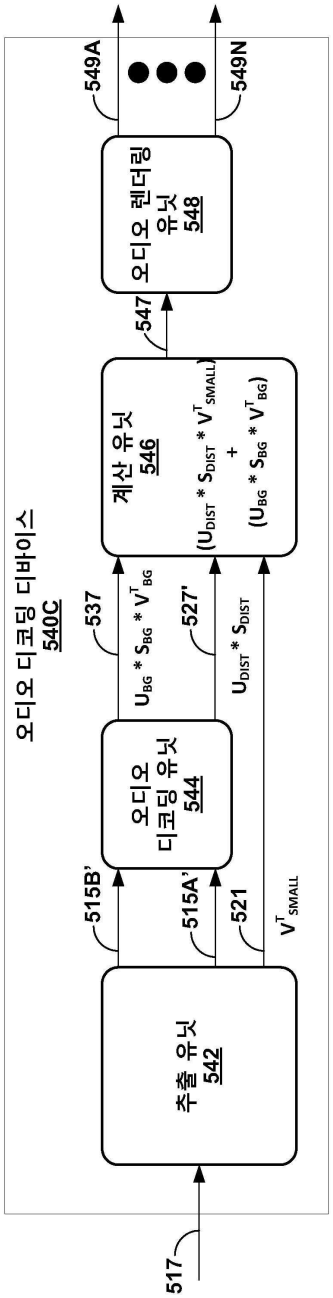
도면41a


오디오 디코딩 디바이스
540A
517
추출 유닛
542
515B
515A
525E
V^T_{DIST}
오디오
디코딩 유닛
44
529
$U_{BG} * S_{BG} * V^T_{BG}$
527
$U_{DIST} * S_{DIST}$
계산 유닛
546
$(U_{DIST} * S_{DIST} * V^T_{SMALL}) + (U_{BG} * S_{BG} * V^T_{BG})$
547
오디오 렌더링
유닛
548
549A
549N

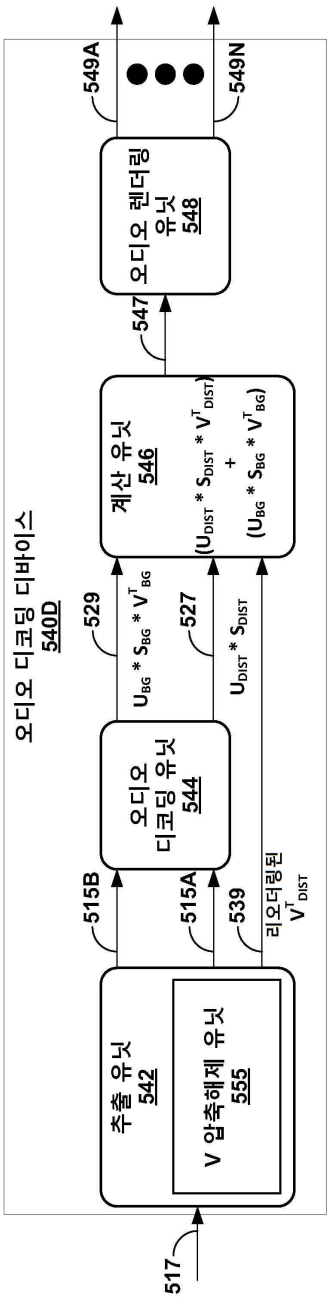
도면41b


오디오 디코딩 디바이스
540B
517
추출 유닛
542
515B
515A
539
리오더링된
V^T_{DIST}
오디오
디코딩 유닛
544
529
$U_{BG} * S_{BG} * V^T_{BG}$
527
$U_{DIST} * S_{DIST}$
계산 유닛
546
$(U_{DIST} * S_{DIST} * V^T_{DIST})$
+
$(U_{BG} * S_{BG} * V^T_{BG})$
547
오디오 렌더링
유닛
548
549A
549N

도면41c


오디오 디코딩 디바이스
540C
517
추출 유닛
542
515B'
515A'
521
V^T_{SMALL}
오디오
디코딩 유닛
544
537
$U_{BG} * S_{BG} * V^T_{BG}$
527'
$U_{DIST} * S_{DIST}$
계산 유닛
546
$(U_{DIST} * S_{DIST} * V^T_{SMALL})$
+
$(U_{BG} * S_{BG} * V^T_{BG})$
547
오디오 렌더링
유닛
548
549A
549N

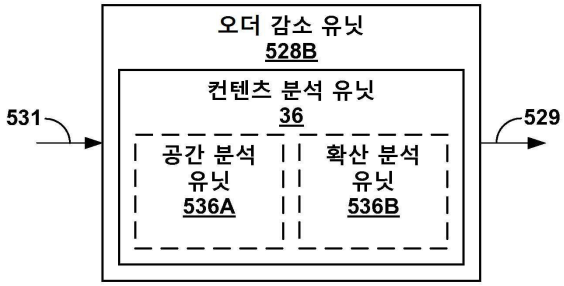
도면41d


549A
549N
오디오 렌더링 유닛 548
547
계산 유닛 546
$(U_{DIST} * S_{DIST} * V^T_{DIST}) + (U_{BG} * S_{BG} * V^T_{BG})$
529
$U_{BG} * S_{BG} * V^T_{BG}$
527
$U_{DIST} * S_{DIST}$
오디오 디코딩 유닛 544
515B
515A
539
리오더링된 V^T_{DIST}
추출 유닛 542
V 압축해제 유닛 555
517
오디오 디코딩 디바이스 540D

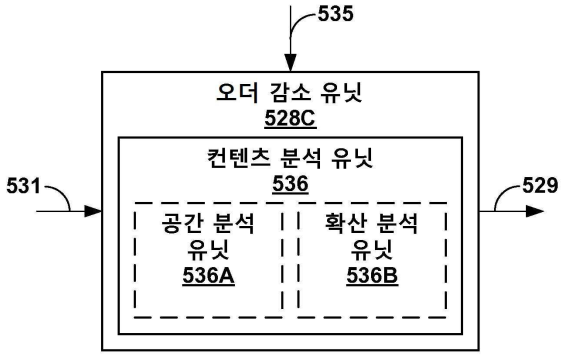
도면42a


535
531
오더 감소 유닛 528A
529

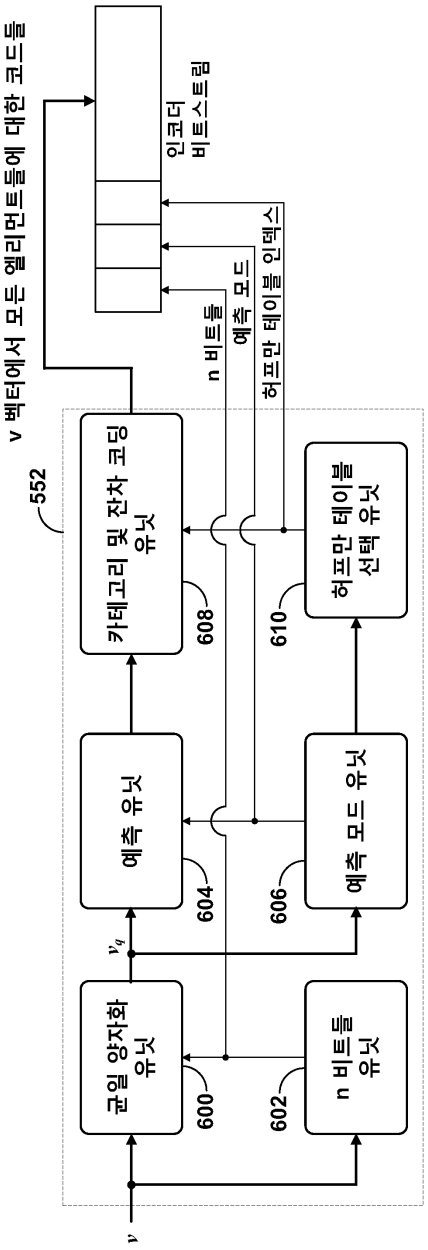
도면42b


오더 감소 유닛
528B
컨텐츠 분석 유닛
36
531
529
공간 분석
유닛
536A
확산 분석
유닛
536B

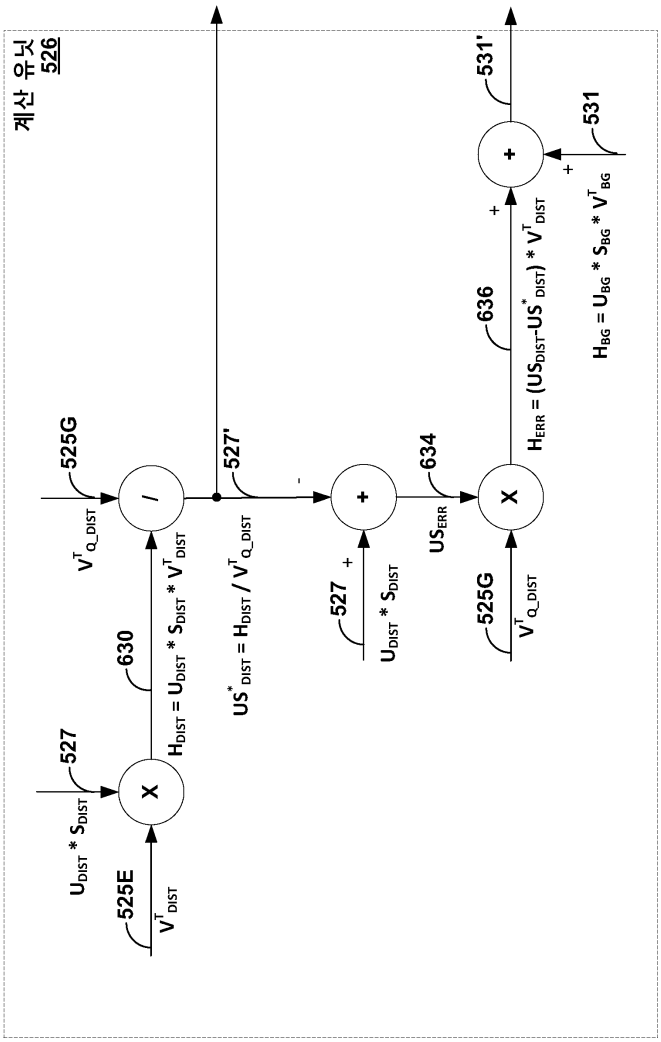
도면42c


535
오더 감소 유닛
528C
컨텐츠 분석 유닛
536
531
529
공간 분석
유닛
536A
확산 분석
유닛
536B

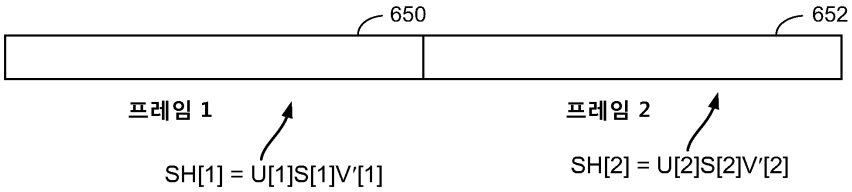
도면43


v 벡터에서 모든 엘리먼트들에 대한 코드들
552
인코더 비트스트림
n 비트들
예측 모드
허프만 테이블 인덱스
카테고리 및 잔차 코딩 유닛
608
허프만 테이블 선택 유닛
610
예측 유닛
604
예측 모드 유닛
606
v_q
균일 양자화 유닛
600
n 비트들 유닛
602
v

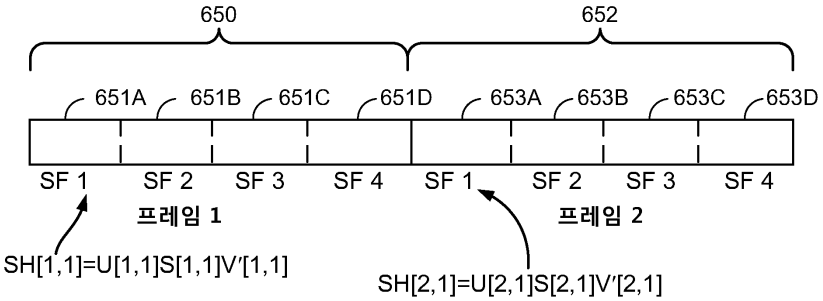
도면44


계산 유닛 526
525G
$V^T_{Q_DIST}$
/
527'
$US^*_{DIST} = H_{DIST} / V^T_{Q_DIST}$
+
634
US_{ERR}
X
636
$H_{ERR} = (US_{DIST} - US^*_{DIST}) * V^T_{DIST}$
+
531'
531
$H_{BG} = U_{BG} * S_{BG} * V^T_{BG}$
527
$U_{DIST} * S_{DIST}$
+
525G
$V^T_{Q_DIST}$
630
$H_{DIST} = U_{DIST} * S_{DIST} * V^T_{DIST}$
527
$U_{DIST} * S_{DIST}$
X
525E
V^T_{DIST}

도면45a


650
652
프레임 1
프레임 2
SH[1] = U[1]S[1]V'[1]
SH[2] = U[2]S[2]V'[2]

도면45b

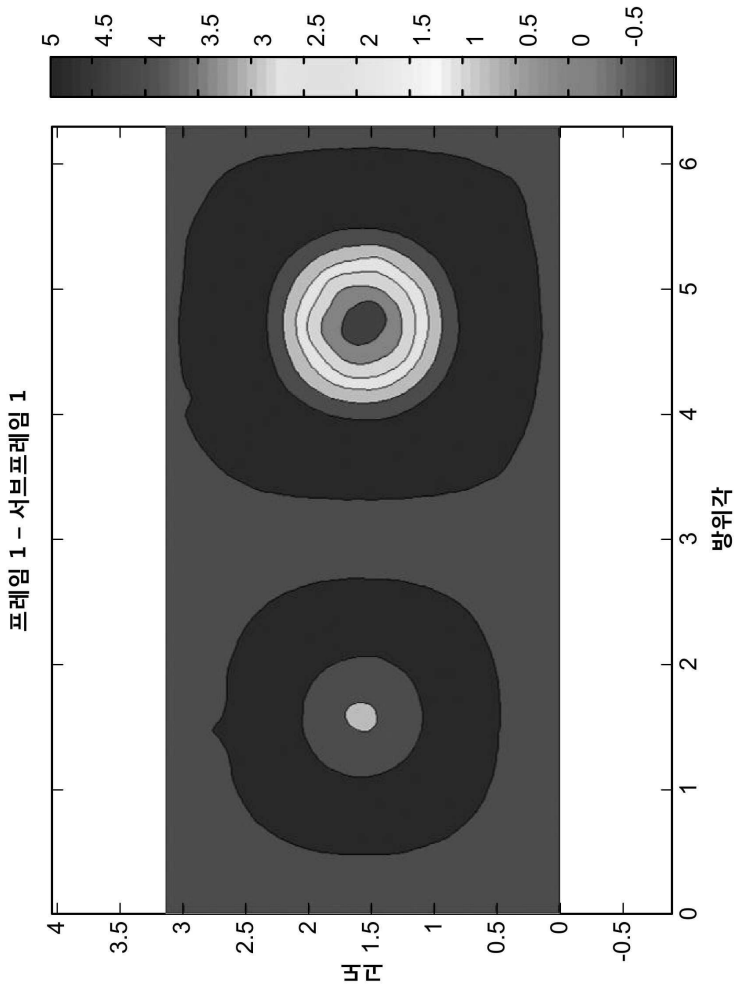


V'[1,2] = interpolation(V'[1,1], V'[2,1]) → U[1,2]S[1,2] = SH[1,2](V'[1,2])-1

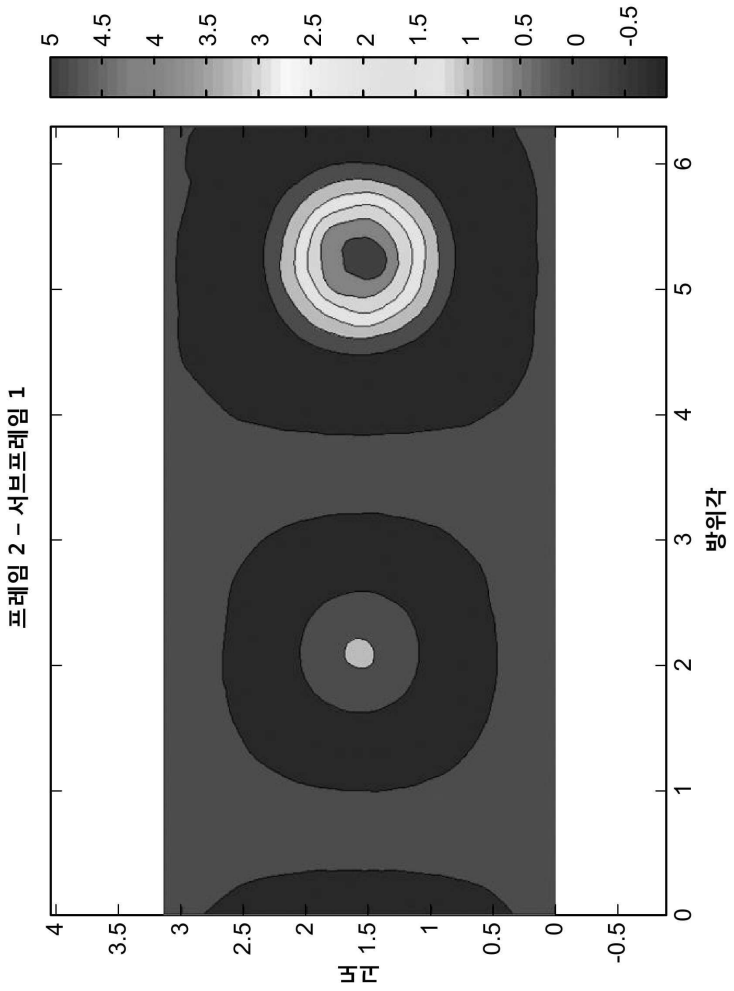
V'[1,3] = interpolation(V'[1,1], V'[2,1]) → U[1,3]S[1,3] = SH[1,3](V'[1,3])-1

V'[1,4] = interpolation(V'[1,1], V'[2,1]) → U[1,3]S[1,3] = SH[1,3](V'[1,3])-1

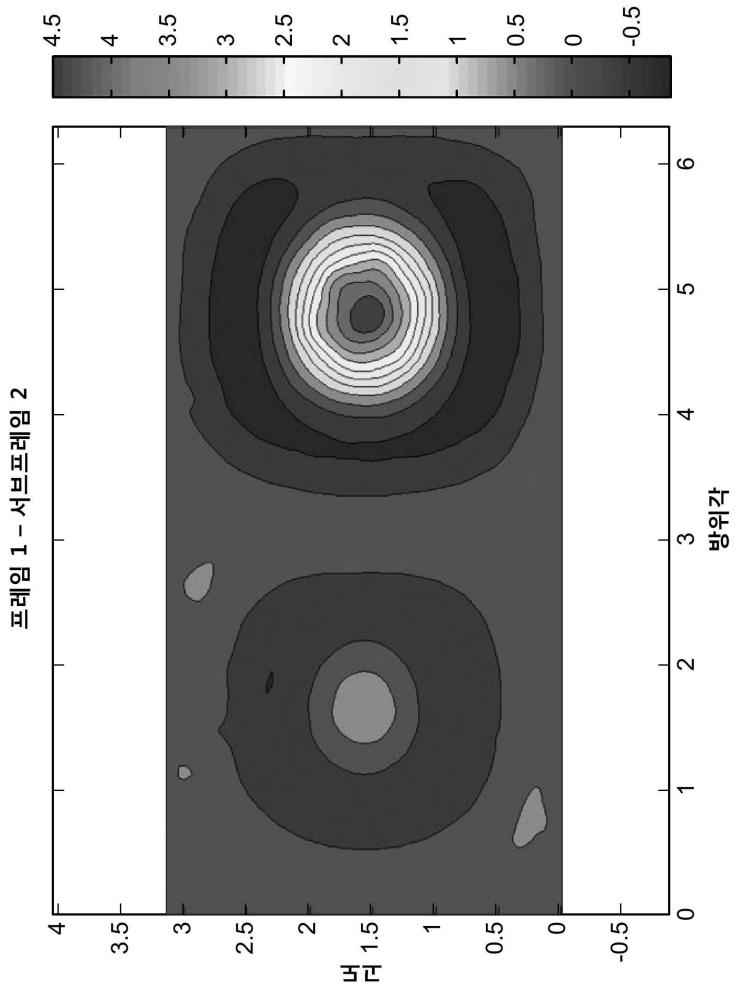
도면46a



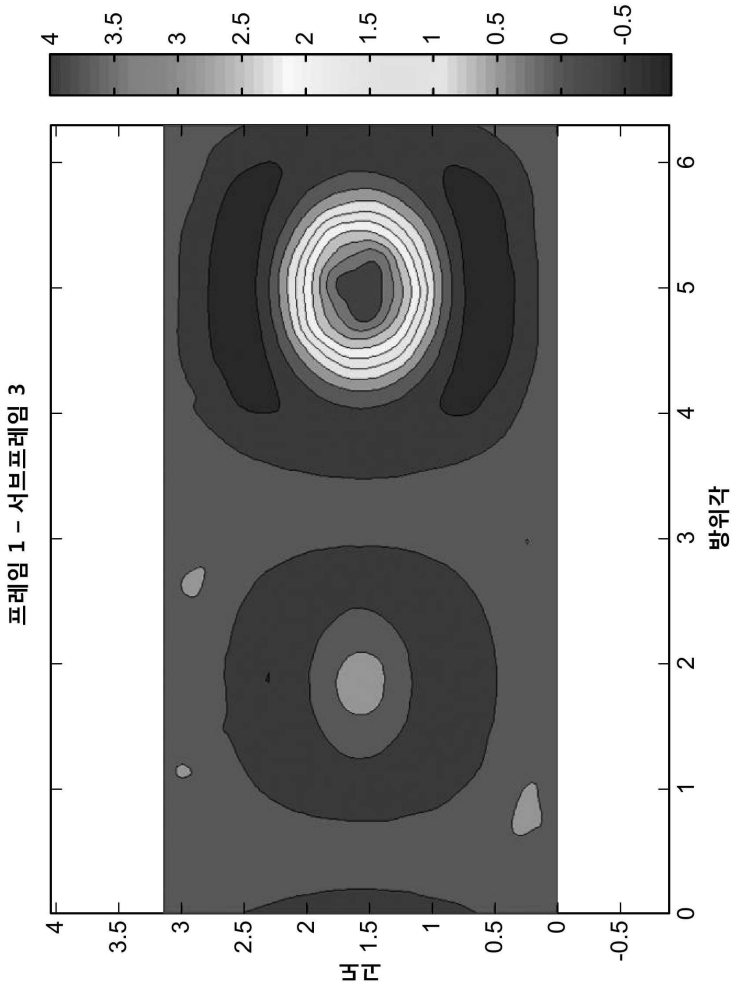
도면46b


5
4.5
4
3.5
3
2.5
2
1.5
1
0.5
0
-0.5
프레임 2 - 서브프레임 1
4
3.5
3
2.5
2
1.5
1
0.5
0
-0.5
범위
0
1
2
3
4
5
6
방위각

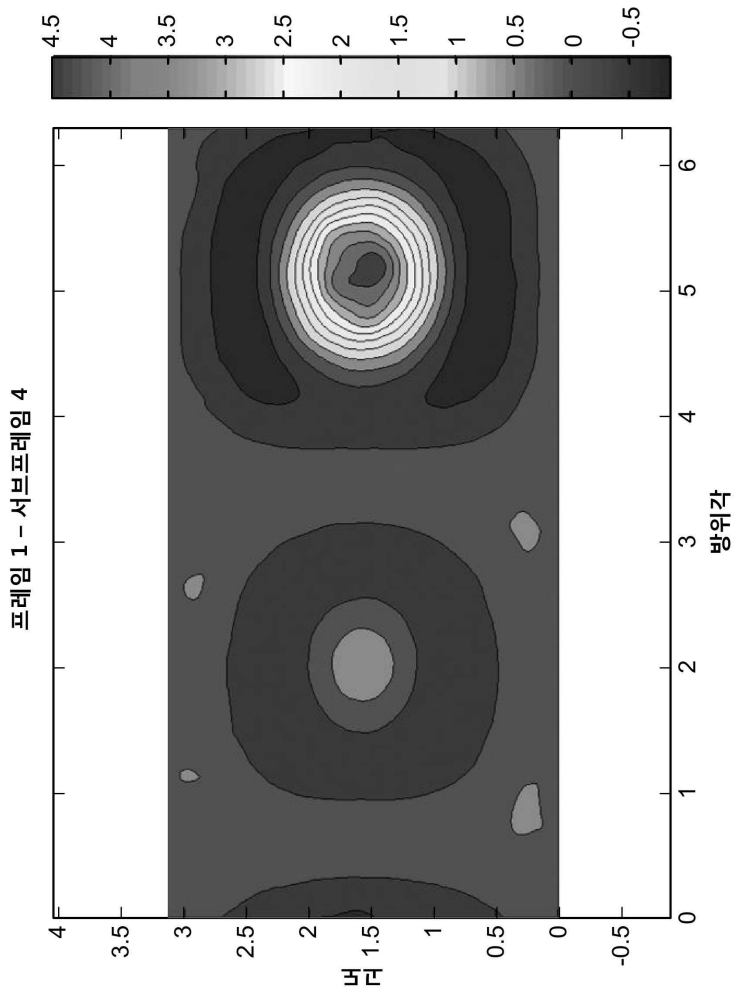
도면46c



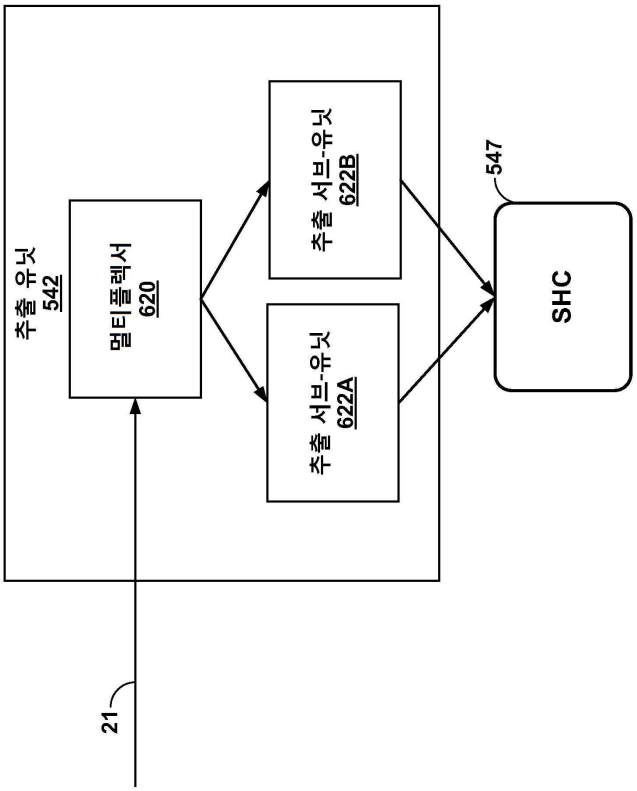
도면46d



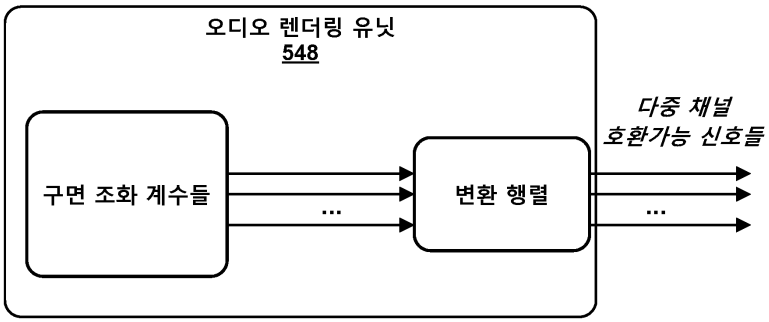
도면46e



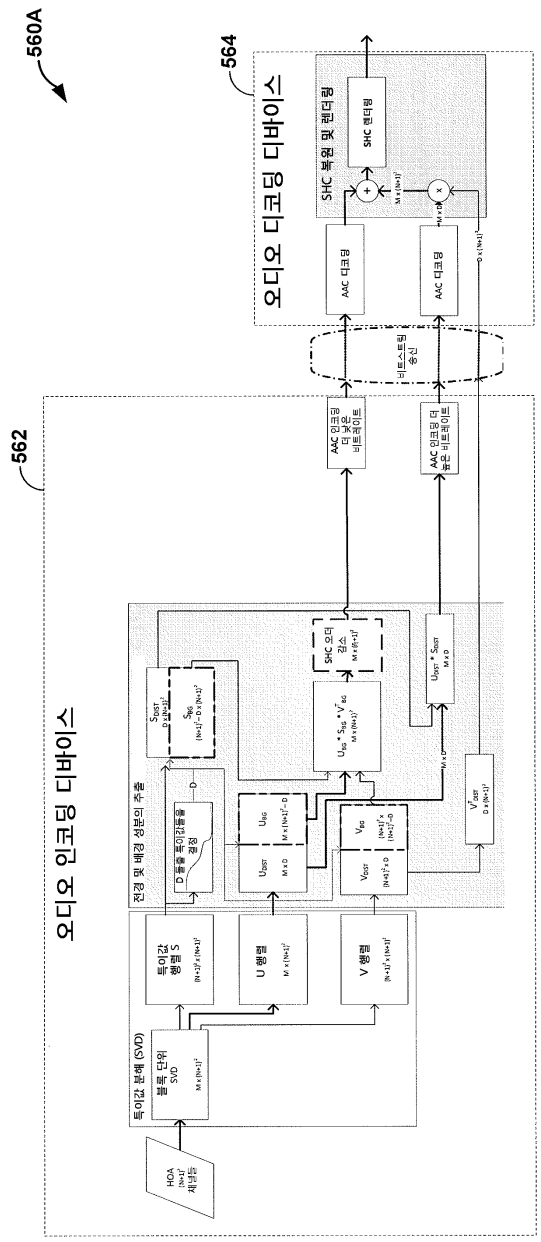
도면47


추출 유닛
542
멀티플렉서
620
추출 서브-유닛
622B
추출 서브-유닛
622A
547
SHC
21

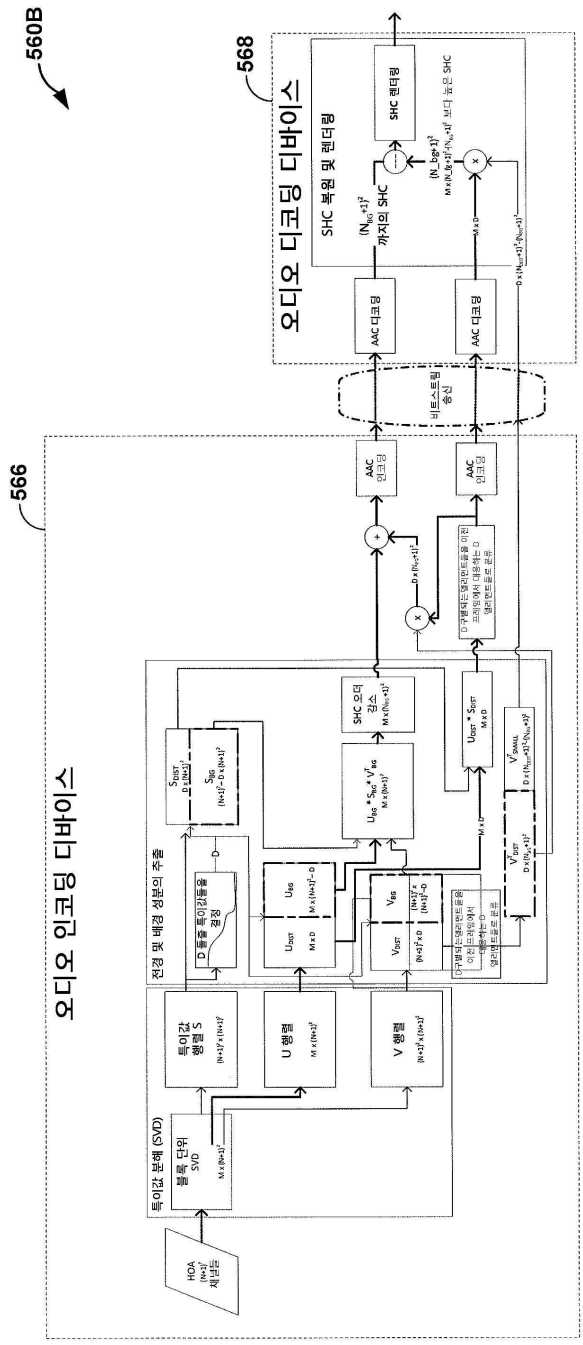
도면48


오디오 렌더링 유닛
548
구면 조화 계수들
...
변환 행렬
다중 채널
호환가능 신호들
...

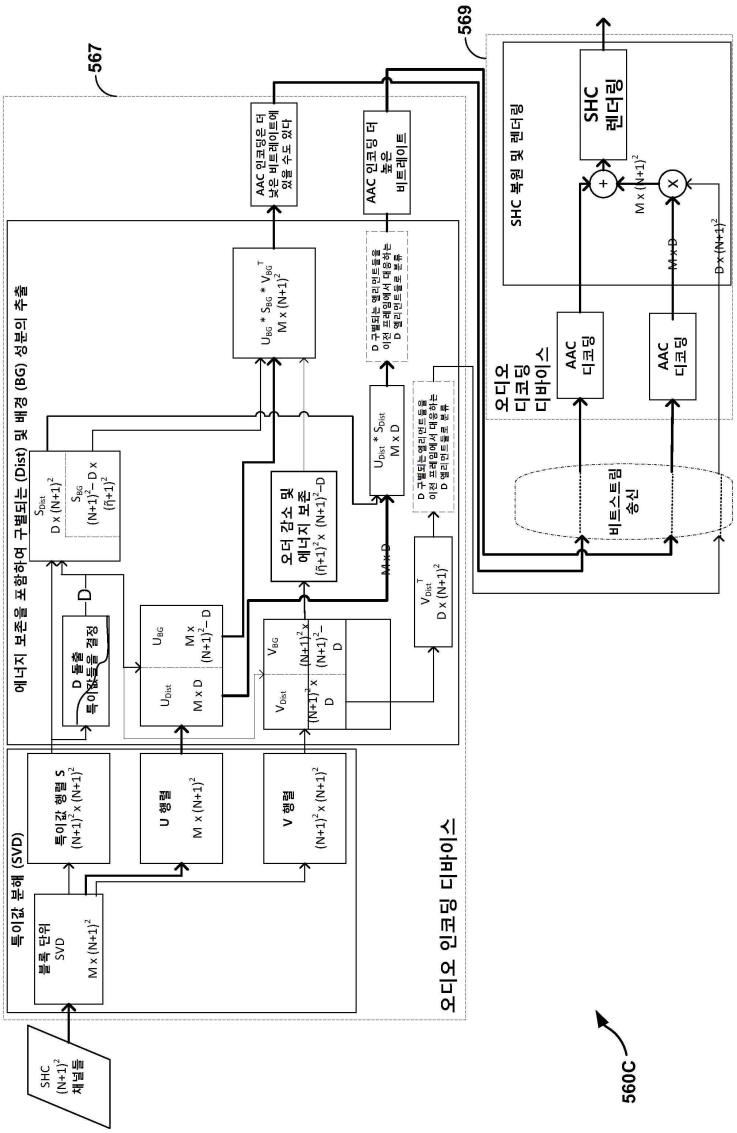
도면49a


560A
562
오디오 인코딩 디바이스
564
오디오 디코딩 디바이스
특이값 분해 (SVD)
블록 단위 SVD
특이값 행렬 S
U 행렬
V 행렬
전경 및 배경 성분의 추출
D 동적 추이값들을 결정
SHC 오더 감소
AAC 인코딩 및 낮은 비트레이트
AAC 인코딩 및 높은 비트레이트
비트스트림 송신
AAC 디코딩
AAC 디코딩
SHC 복원 및 렌더링
SHC 렌더링

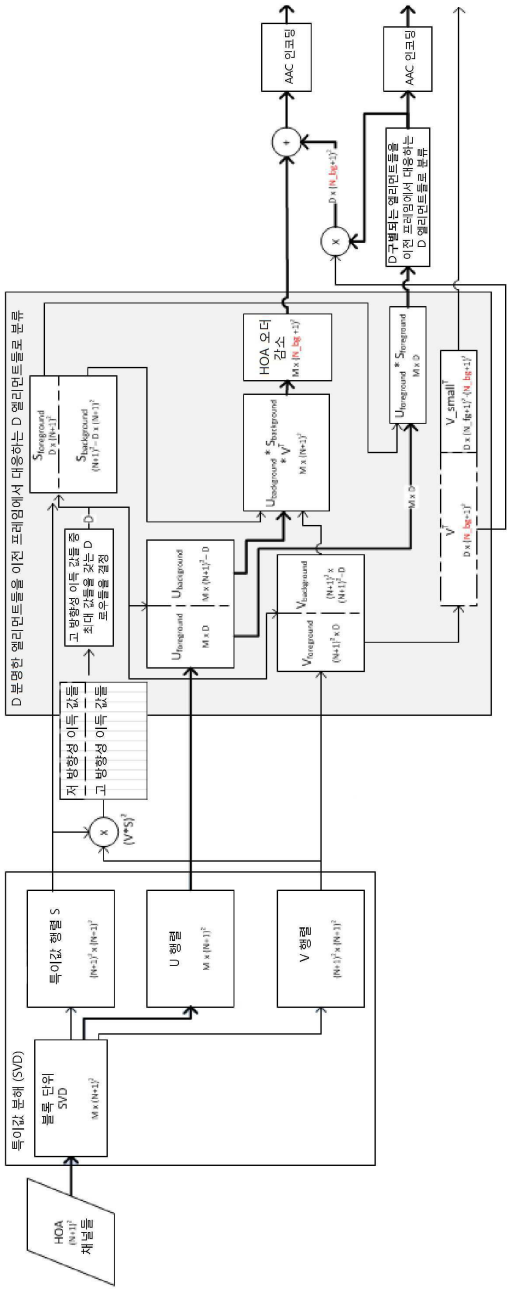
도면49b


560B
568
오디오 디코딩 디바이스
SHC 복원 및 렌더링
AAC 디코딩
SHC 렌더링
비트스트림 수신
566
오디오 인코딩 디바이스
특이값 분해 (SVD)
블록 단위 SVD
특이값 행렬 S
U 행렬
V 행렬
전경 및 배경 정보의 추출
D 추출을 위한 임계값 결정
SHC 오더 감소
AAC 인코딩
HOA 채널들

도면49c

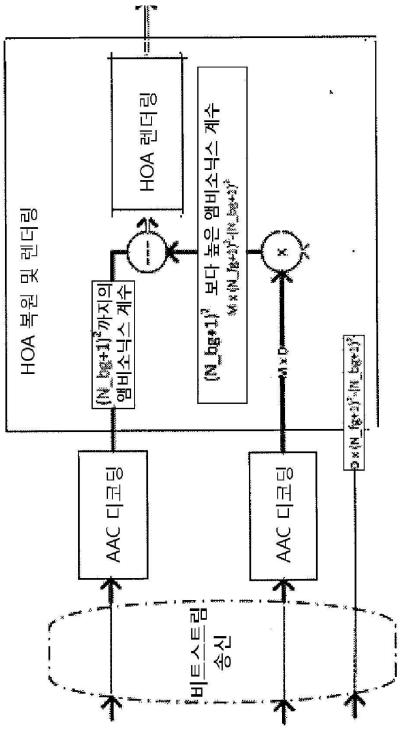

567
569
560C
오디오 인코딩 디바이스
특이값 분해 (SVD)
블록 단위 SVD
특이값 행렬 S
U 행렬
V 행렬
에너지 보존을 포함하여 구별되는 (Dist) 및 배경 (BG) 성분의 추출
D 돌출 특이벡터를 결정
오더 감소 및 에너지 보존
AAC 인코딩 더 낮은 비트레이트에 있을 수도 있다
AAC 인코딩 더 높은 비트레이트
D 구별되는 앰리언트들을 이전 프레임에서 대응하는 D 앰리언트들로 분류
비트스트림 송신
오디오 디코딩 디바이스
AAC 디코딩
SHC 복원 및 렌더링
SHC 렌더링
SHC (N+1)² 채널들

도면49di


HOA 채널들
특이값 분해 (SVD)
블록 단위 SVD
특이값 정렬 S
U 정렬
V 정렬
저 방향성 이득 값들
고 방향성 이득 값들
D 분명한 엘리먼트들을 이전 프레임에서 대응하는 D 엘리먼트들로 분류
고 방향성 이득 값들을 갖는 D 로우들을 결정
HOA 오더 감소
AAC 인코딩
AAC 인코딩
D구별되는 엘리먼트들을 이전 프레임에서 대응하는 D 엘리먼트들로 분류

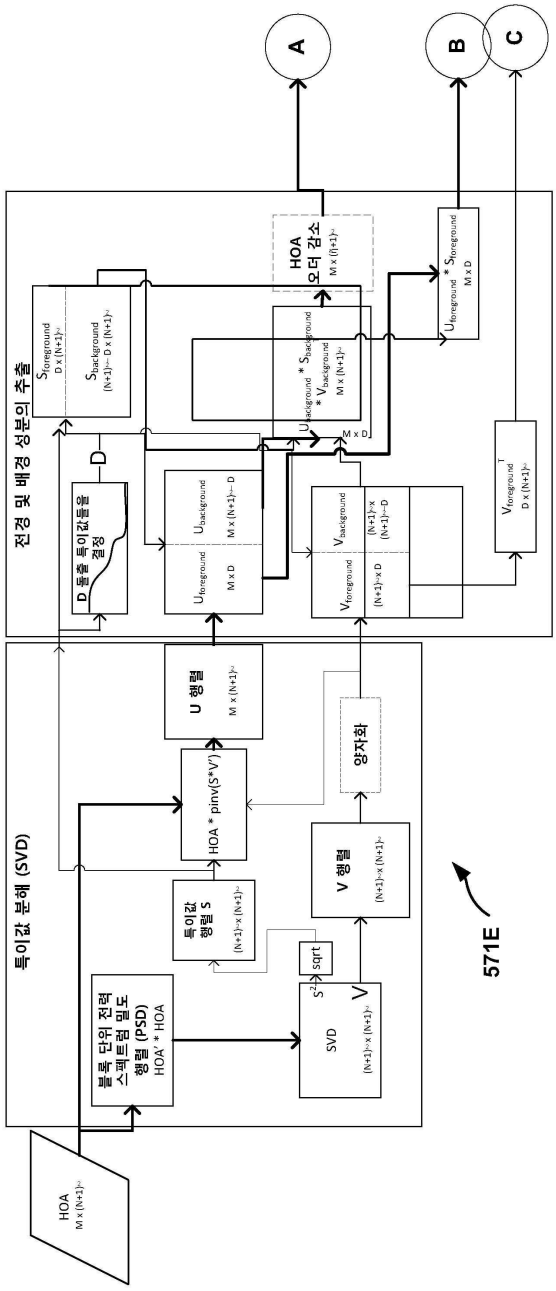
567D

도면49dii

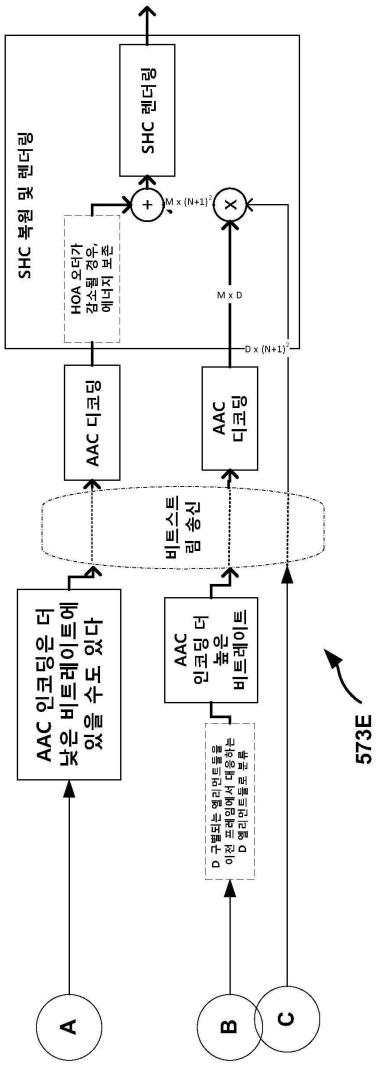


569D

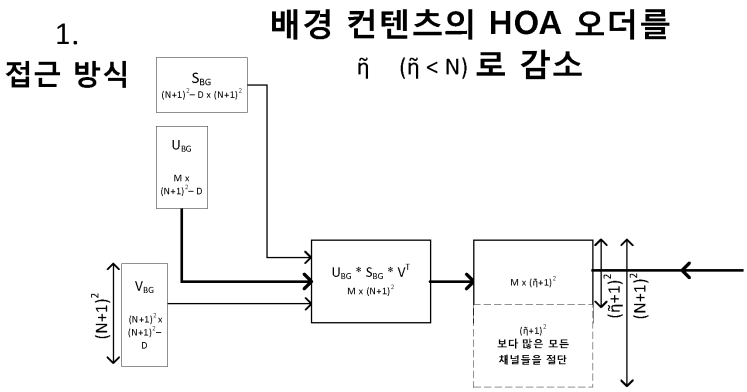
도면49ei


A
B
C
HOA
M x (N+1)²
특이값 분해 (SVD)
블록 단위 전력 스펙트럼 밀도 행렬 (PSD)
HOA' * HOA
SVD
S²
V
(N+1)² x (N+1)²
sqrt
특이값 행렬 S
(N+1)² x (N+1)²
HOA * pinv(S*V')
U 행렬
M x (N+1)²
V 행렬
(N+1)² x (N+1)²
양자화
571E
전경 및 배경 성분의 추출
D 출출 특이값들을 결정
D
U_foreground
M x D
U_background
M x (N+1)² - D
V_foreground
(N+1)² x D
V_background
(N+1)² x (N+1)² - D
S_foreground
D x D
S_background
(N+1)² - D x (N+1)² - D
U_background * S_background * V_background
M x (N+1)²
HOA 오더 감소
M x (l+1)²
U_foreground * S_foreground
M x D
V_foreground^T
D x (N+1)²

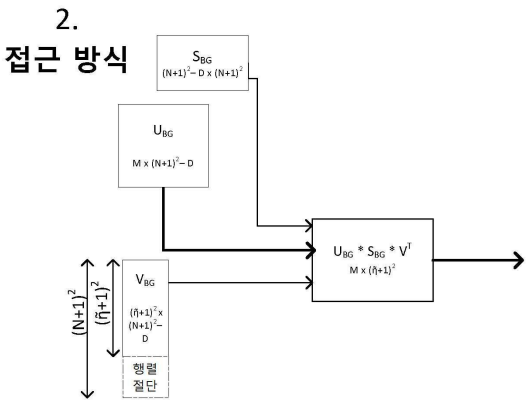
도면49eii



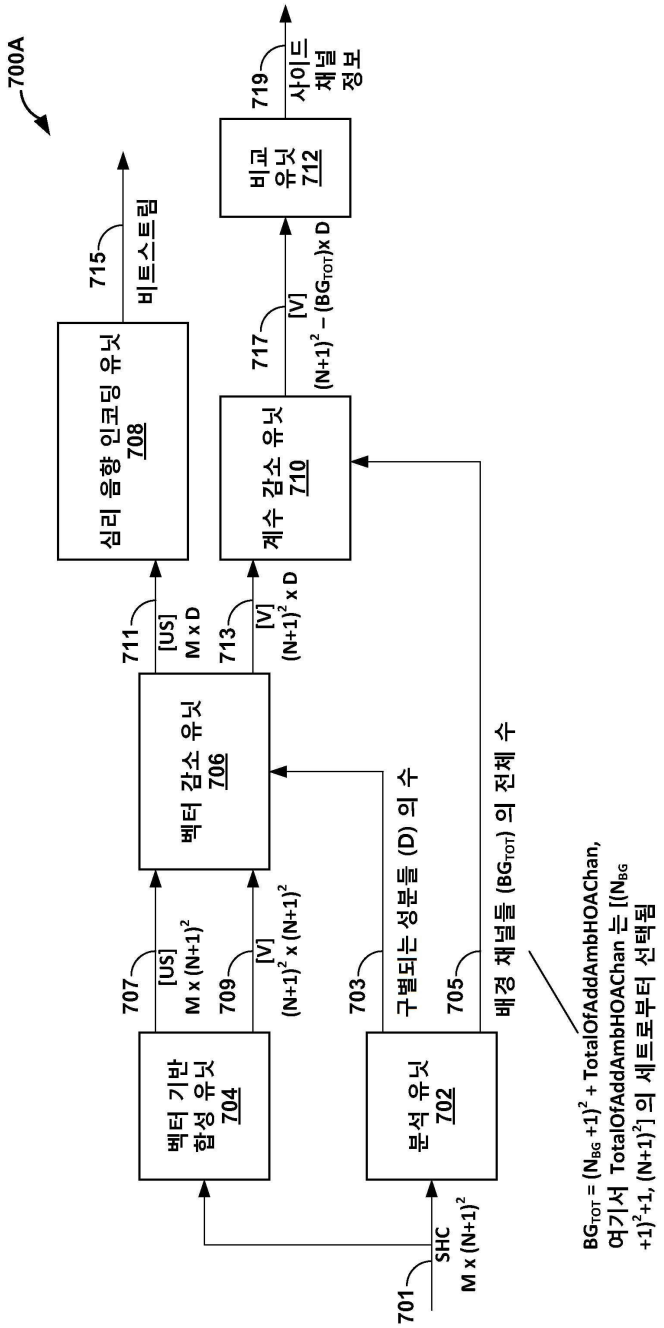
도면50a



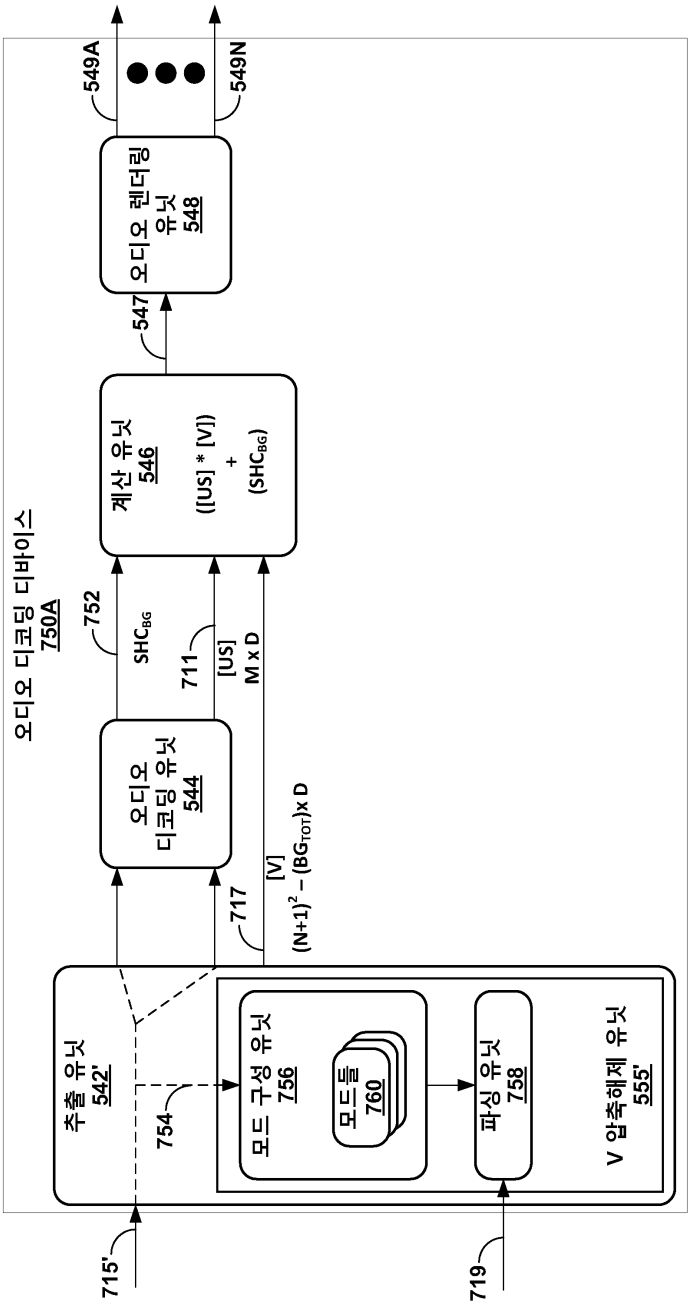
도면50b


2.
접근 방식
S_{BG}
$(N+1)^2 - D \times (N+1)^2$
U_{BG}
$M \times (N+1)^2 - D$
$U_{BG} * S_{BG} * V^T$
$M \times (\hat{n}+1)^2$
V_{BG}
$(\hat{n}+1)^2 \times (N+1)^2 - D$
행렬 절단
$(N+1)^2$
$(\hat{n}+1)^2$

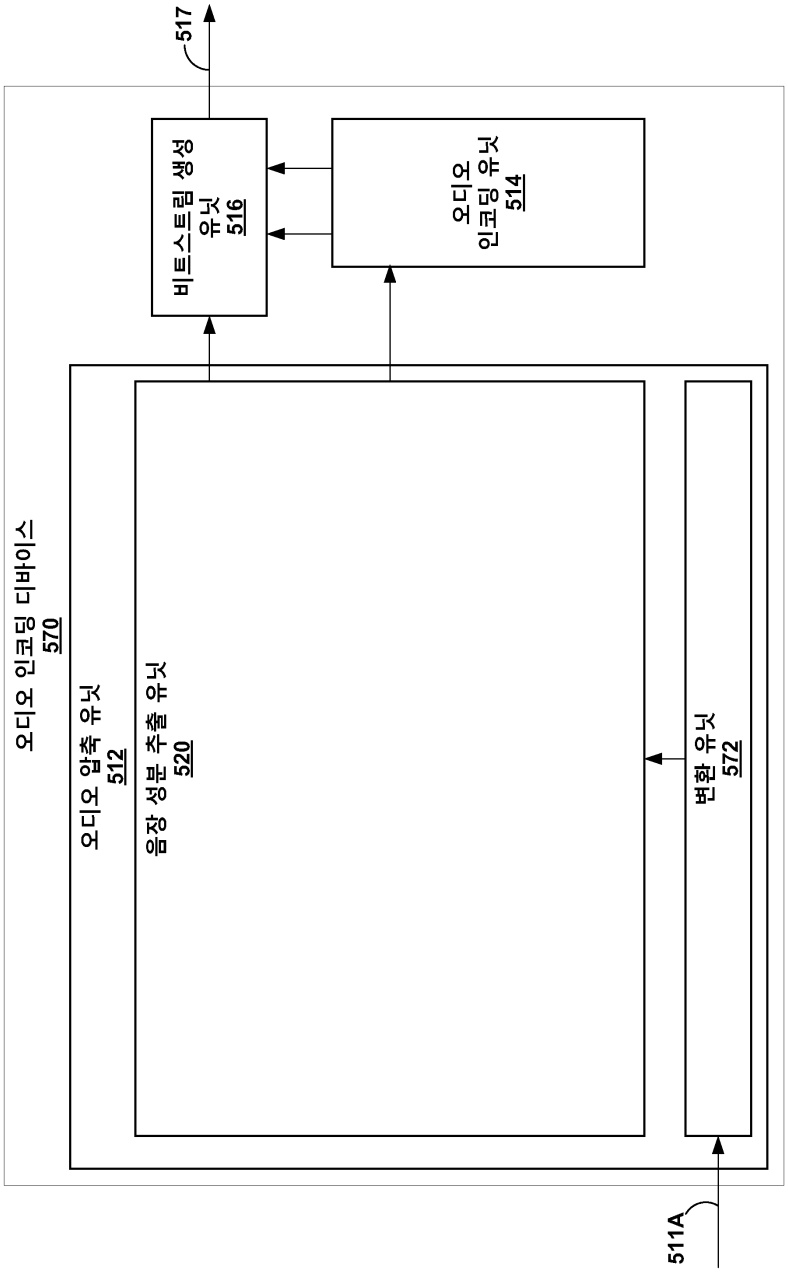
도면51


700A
701
SHC
M x (N+1)²
분석 유닛
702
벡터 기반
합성 유닛
704
707
[US]
M x (N+1)²
709
[V]
(N+1)² x (N+1)²
703
구별되는 성분들 (D) 의 수
705
배경 채널들 (BG_{TOT}) 의 전체 수
BG_{TOT} = $(N_{BG}+1)^2$ + TotalOfAddAmbHOAChan,
여기서 TotalOfAddAmbHOAChan 는 [(N_{BG}
+1)²+1, (N+1)²] 의 세트로부터 선택됨
벡터 감소 유닛
706
711
[US]
M x D
713
[V]
(N+1)² x D
계수 감소 유닛
710
717
[V]
(N+1)² – (BG_{TOT})x D
비교
유닛
712
719
사이드
채널
정보
심리 음향 인코딩 유닛
708
715
비트스트림

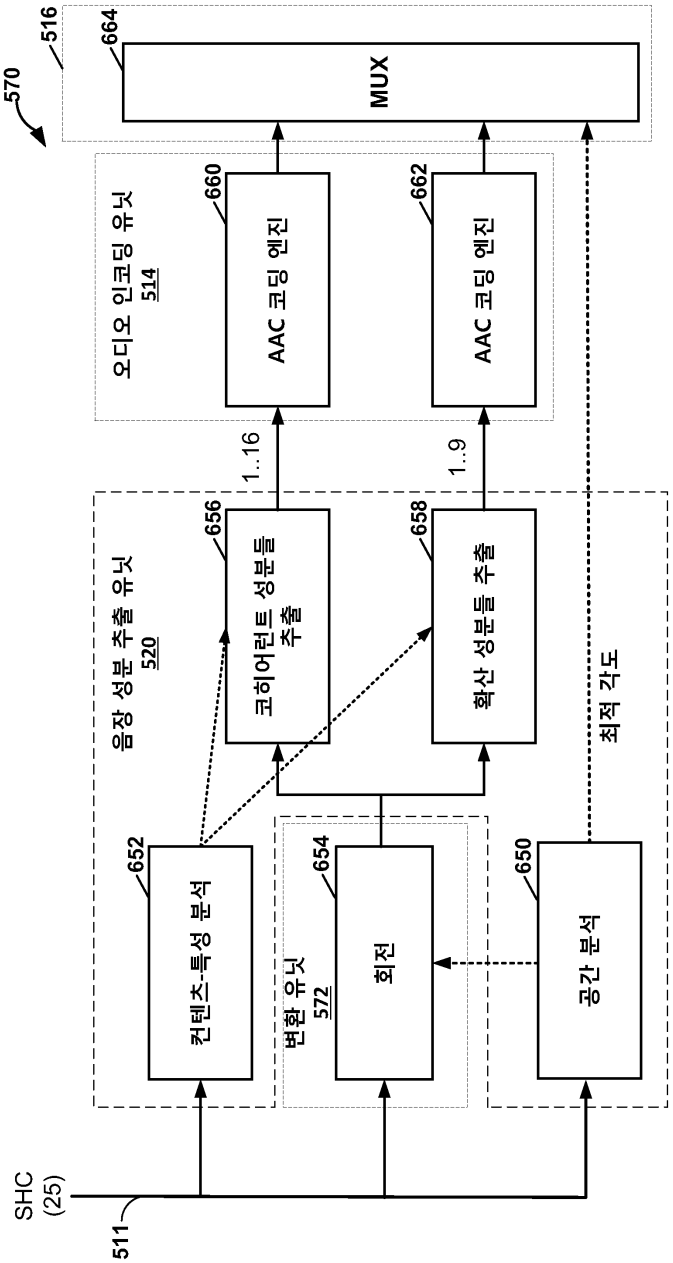
도면52


549A
549N
오디오 렌더링 유닛 548
547
계산 유닛 546
([US] * [V]) + (SHC_{BG})
752
SHC_{BG}
711
[US] M x D
오디오 디코딩 유닛 544
717
[V] $(N+1)^2 - (BG_{TOT})$x D
오디오 디코딩 디바이스 750A
추출 유닛 542'
754
모드 구성 유닛 756
모드들 760
파싱 유닛 758
V 압축해제 유닛 555'
715'
719

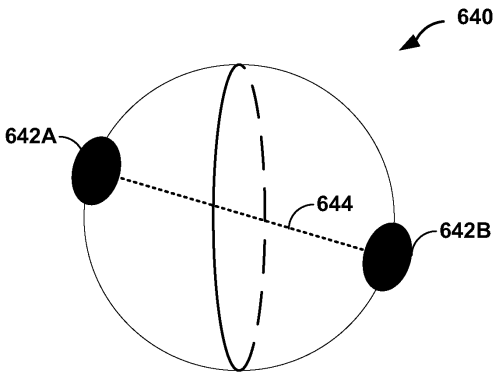
도면53


517
비트스트림 생성 유닛 516
오디오 인코딩 유닛 514
오디오 인코딩 디바이스 570
오디오 압축 유닛 512
음장 성분 추출 유닛 520
변환 유닛 572
511A

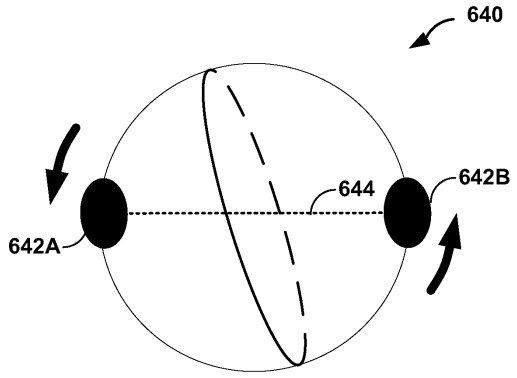
도면54


570
516
664
MUX
오디오 인코딩 유닛
514
660
AAC 코딩 엔진
662
AAC 코딩 엔진
1..16
1..9
음장 성분 추출 유닛
520
656
코히어런트 성분들 추출
658
확산 성분들 추출
최적 각도
652
컨텐츠-특성 분석
654
변환 유닛
572
회전
650
공간 분석
SHC
(25)
511

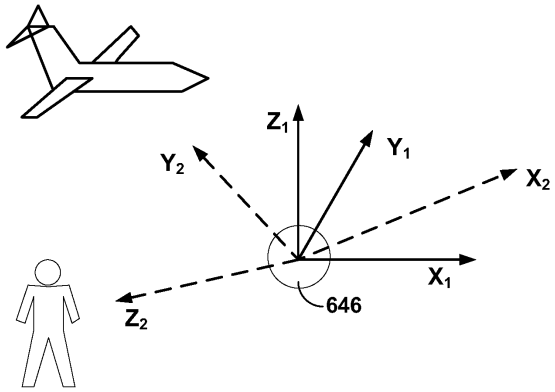
도면55a


640
642A
644
642B

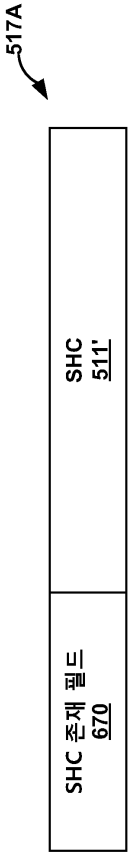
도면55b



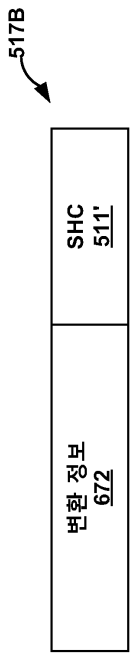
도면56



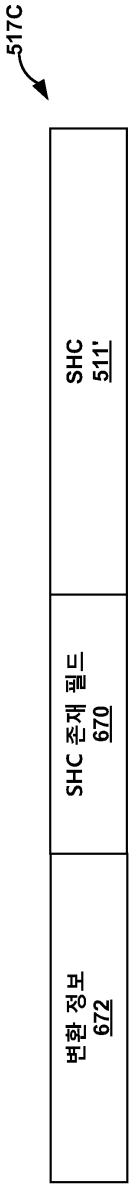
도면57a



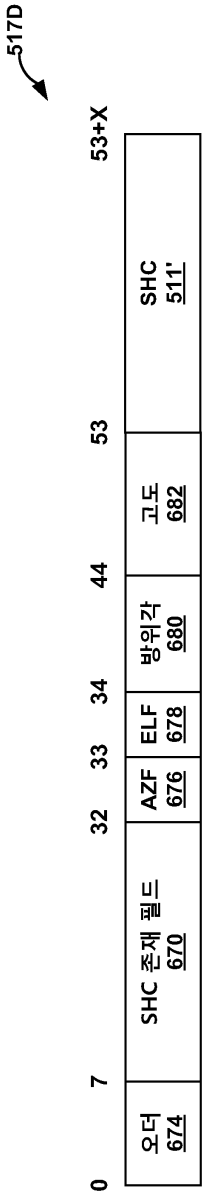
도면57b



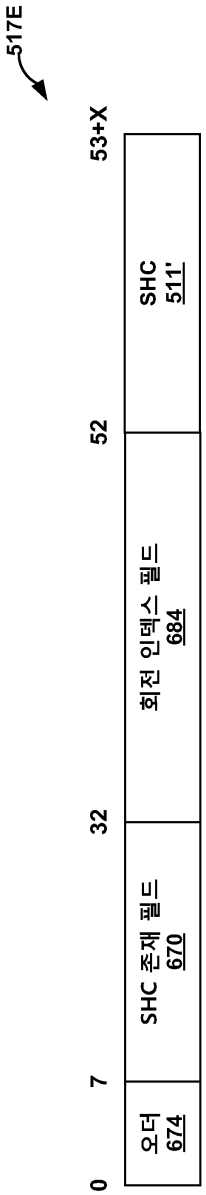
도면57c



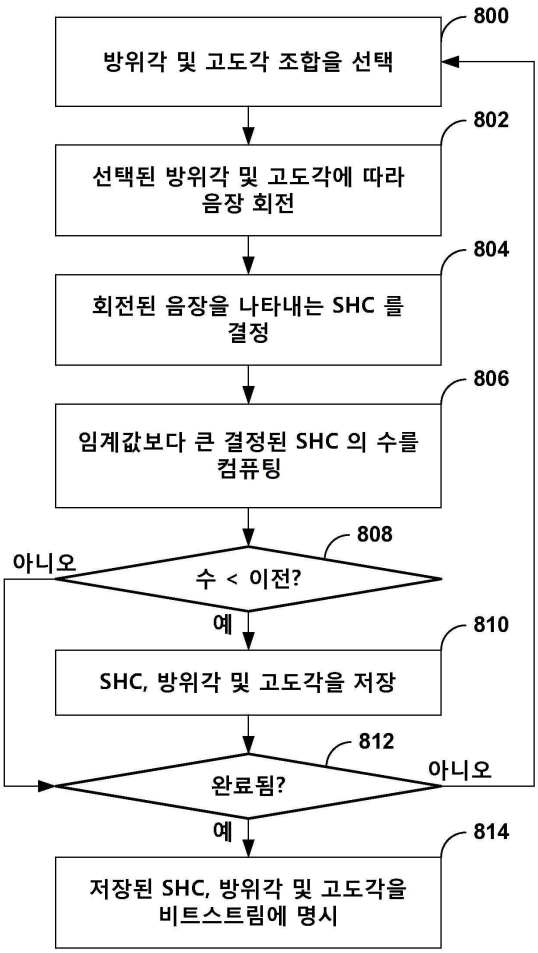
도면57d


517D
0
7
32
33
34
44
53
53+X
오더 674
SHC 존재 필드 670
AZF 676
ELF 678
방위각 680
고도 682
SHC 511'

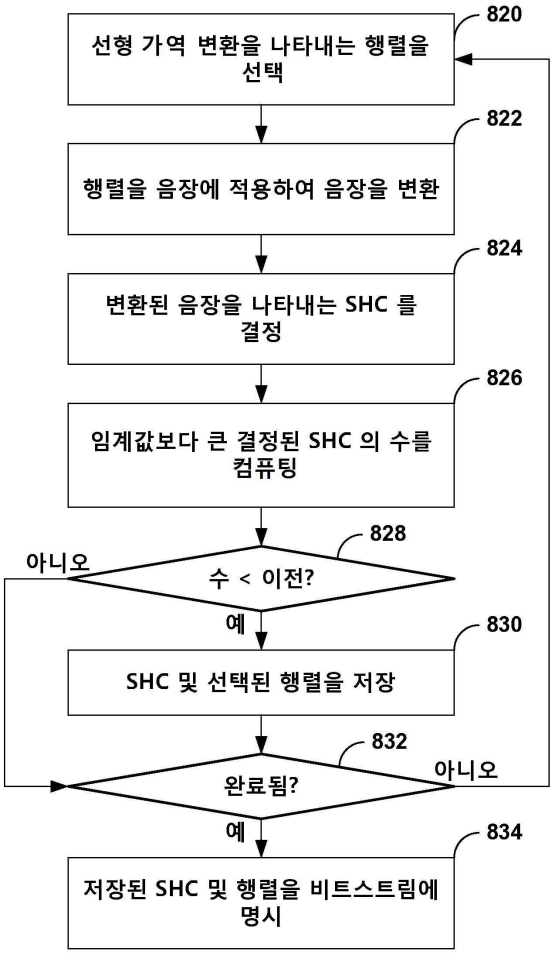
도면57e



도면58


800
방위각 및 고도각 조합을 선택
802
선택된 방위각 및 고도각에 따라 음장 회전
804
회전된 음장을 나타내는 SHC 를 결정
806
임계값보다 큰 결정된 SHC 의 수를 컴퓨팅
808
수 < 이전?
아니오
예
810
SHC, 방위각 및 고도각을 저장
812
완료됨?
아니오
예
814
저장된 SHC, 방위각 및 고도각을 비트스트림에 명시

도면59


820
선형 가역 변환을 나타내는 행렬을 선택
822
행렬을 음장에 적용하여 음장을 변환
824
변환된 음장을 나타내는 SHC 를 결정
826
임계값보다 큰 결정된 SHC 의 수를 컴퓨팅
828
수 < 이전?
아니오
예
830
SHC 및 선택된 행렬을 저장
832
완료됨?
아니오
예
834
저장된 SHC 및 행렬을 비트스트림에 명시