



(19) 대한민국특허청(KR)  
(12) 등록특허공보(B1)

(45) 공고일자 2018년03월14일  
(11) 등록번호 10-1838056  
(24) 등록일자 2018년03월07일

(51) 국제특허분류(Int. Cl.)  
G10L 19/008 (2014.01) G10L 19/24 (2013.01)  
H04S 3/00 (2006.01)  
(52) CPC특허분류  
G10L 19/008 (2013.01)  
G10L 19/24 (2013.01)  
(21) 출원번호 10-2016-7025844  
(22) 출원일자(국제) 2015년03월20일  
심사청구일자 2016년09월20일  
(85) 번역문제출일자 2016년09월20일  
(65) 공개번호 10-2016-0124422  
(43) 공개일자 2016년10월27일  
(86) 국제출원번호 PCT/EP2015/055914  
(87) 국제공개번호 WO 2015/140291  
국제공개일자 2015년09월24일  
(30) 우선권주장  
14305411.2 2014년03월21일  
유럽특허청(EPO)(EP)  
(56) 선행기술조사문헌  
WD1-HOA Text of MPEG-H 3D Audio. ISO/IEC  
JTC1/SC29/WG11 N14264. 2014.02.21.\*  
Erik Hellerud, et al. Spatial redundancy in  
Higher Order Ambisonics and its use for  
lowdelay lossless compression. IEEE  
International Conference on Acoustics, Speech  
and Signal Processing. 2009. pp.26\*  
\*는 심사관에 의하여 인용된 문헌

(73) 특허권자  
둘비 인터네셔널 에이비  
네덜란드 1101 씨엔 암스트레담 주이두스트 헤리  
커베르그벡 1-35 3이 아폴로 빌딩  
(72) 발명자  
코르돈, 스벤  
독일 31515 분스토르프 뮐렌캠프스트라세 50에이  
크뤼거, 알렉산더  
독일 30655 하노버 포이어바흐스트라세 16  
뷔볼트, 올리버  
독일 30559 하노버 자르브뤼케너 스트라세 4  
(74) 대리인  
양영준, 백만기, 정은진

전체 청구항 수 : 총 2 항

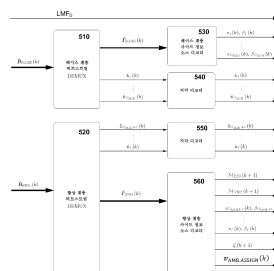
심사관 : 이남숙

(54) 발명의 명칭 고차 앰비소닉스(HOA) 신호를 압축하는 방법, 압축된 HOA 신호를 압축 해제하는 방법, HOA 신호를 압축하기 위한 장치, 및 압축된 HOA 신호를 압축 해제하기 위한 장치

(57) 요약

HOA 계수 시퀀스들의 입력 시간 프레임들( $C(k)$ )을 가진 입력 HOA 표현인 HOA 신호를 압축하는 방법은 입력 시간 프레임들의 공간 HOA 인코딩과, 후속 지각 인코딩 및 소스 인코딩을 포함한다. 각각의 입력 시간 프레임은 우세 사운드 신호들( $X_{ps}(k-1)$ )의 프레임 및 주변 HOA 컴포넌트( $\tilde{C}_{AMB}(k-1)$ )의 프레임으로 분해된다(802). 주변 HOA (뒷면에 계속)

대표도 - 도5



컴포넌트( $\mathbf{c}_{AMB}(k-1)$ )는 계층화 모드에서, 하위 위치들 내의 입력 HOA 표현( $\mathbf{c}_n(k-1)$ )의 제1 HOA 계수 시퀀스들, 및 나머지 상위 위치들 내의 제2 HOA 계수 시퀀스들( $\mathbf{c}_{AMB,n}(k-1)$ )을 포함한다. 제2 HOA 계수 시퀀스들은 입력 HOA 표현과 우세 사운드 신호들의 HOA 표현 사이의 잔차의 HOA 표현의 부분이다.

(52) CPC특허분류

*H04S 3/008* (2013.01)

*H04S 2400/01* (2013.01)

*H04S 2420/11* (2013.01)

## 명세서

### 청구범위

#### 청구항 1

사운드 또는 사운드 필드의 압축된 고차 앰비소닉스(HOA)(Higher Order Ambisonics) 표현을 디코딩하는 방법으로서,

상기 압축된 HOA 표현을 포함하는 비트 스트림을 수신하는 단계;

상기 압축된 HOA 표현에 관한 다수의 계층이 존재하는지를 결정하는 단계;

다수의 계층이 존재한다는 결정에 기초하여, 디코딩된 HOA 표현들의 시퀀스를 획득하기 위해 상기 비트 스트림으로부터 상기 압축된 HOA 표현을 디코딩하는 단계

를 포함하고,

상기 디코딩된 HOA 표현들의 시퀀스의 제1 부분집합은 상기 디코딩된 HOA 표현들의 시퀀스의 인덱스들의 제1 세트에 대응하고, 상기 디코딩된 HOA 표현들의 시퀀스의 제2 부분집합은 상기 디코딩된 HOA 표현들의 시퀀스의 인덱스들의 제2 세트에 대응하고,

인덱스들의 상기 제1 세트는  $O_{\text{MIN}}$  채널들에 기초하고,

인덱스들의 상기 제1 세트 내의 각각의 인덱스에서는, 상기 제1 부분집합 내의 대응하는 디코딩된 HOA 표현은 대응하는 주변 HOA 컴포넌트에만 기초하여 결정되고,

인덱스들의 상기 제2 세트는 상기 다수의 계층 중 적어도 하나의 계층에 기초하여 결정되는 디코딩 방법.

#### 청구항 2

[청구항 2은(는) 설정등록료 납부시 포기되었습니다.]

제1항에 있어서, 인덱스들의 상기 제1 세트는  $1 \leq n \leq O_{\text{MIN}}$  에 기초하여 결정되고, 인덱스들의 상기 제2 세트는  $O_{\text{MIN}} + 1 \leq n \leq O$  에 기초하여 결정되며,  $O$  는 채널들의 총수를 나타내고,  $O_{\text{MIN}}$  은 1과  $O$  사이의 수를 나타내는 디코딩 방법.

#### 청구항 3

[청구항 3은(는) 설정등록료 납부시 포기되었습니다.]

제1항에 있어서, 인덱스  $n$  및 프레임  $k$ 에서는,  $n$ 이 인덱스들의 상기 제1 세트 내에 있을 때, 상기 제1 부분집합은 대응하는 주변 사운드 컴포넌트  $\hat{c}_{\text{AMB},n}(k-1)$  에 기초하여 결정되고,  $n$ 이 인덱스들의 상기 제2 세트 내에 있을 때, 상기 제2 부분집합은 대응하는 주변 사운드 컴포넌트  $\hat{c}_{n,\text{AMB}}(k-1)$  및 대응하는 우세 사운드 컴포넌트  $\hat{c}_{n,\text{PS}}(k-1)$  의 부가에 기초하여 결정되고, 상기 디코딩된 HOA 표현들은 적어도 부분적으로

$$\tilde{c}_n(k-1) = \begin{cases} \hat{c}_{\text{AMB},n}(k-1) & \text{인덱스들의 상기 제1 세트 내의 } n \text{에 대해서} \\ \hat{c}_n(k-1) = \hat{c}_{\text{PS},n}(k-1) + \hat{c}_{\text{AMB},n}(k-1), & \text{인덱스들의 상기 제2 세트 내의 } n \text{에 대해서} \end{cases}$$

에 의해 표현되는 디코딩 방법.

#### 청구항 4

[청구항 4은(는) 설정등록료 납부시 포기되었습니다.]

제1항에 있어서,  $N_{\text{MIN}} \leq N$  일 때  $O_{\text{MIN}} = (N_{\text{MIN}} + 1)^2$  이고,  $N$  은 상기 압축된 HOA 표현의 입력 프레임들의 오더(order)인, 디코딩 방법.

#### 청구항 5

[청구항 5은(는) 설정등록료 납부시 포기되었습니다.]

제1항에 있어서, 상기 비트 스트림에 다수의 계층의 지시가 시그널링되는 디코딩 방법.

#### 청구항 6

[청구항 6은(는) 설정등록료 납부시 포기되었습니다.]

제1항에 있어서, 상기 다수의 계층은 베이스 계층 및 적어도 하나의 향상 계층을 포함하는 디코딩 방법.

#### 청구항 7

[청구항 7은(는) 설정등록료 납부시 포기되었습니다.]

제1항에 있어서, 프레임  $k$ 에서는, 상기 디코딩된 HOA 표현들의 시퀀스가 주변 할당 벡터( $v_{\text{AMBASSIGN}}(k)$ ), 방향 표현 및 각각의 양자화된 방향의 인덱스를 포함하는 제1 튜플 세트  $\mathcal{M}_{\text{DIR}}(k+1)$ , 및 벡터 기반 표현 및 상기 벡터 기반 표현의 방향 분포를 정의하는 벡터의 인덱스를 포함하는 제2 튜플 세트  $\mathcal{M}_{\text{VEC}}(k+1)$ 에 기초하여 결정되는 디코딩 방법.

#### 청구항 8

사운드 또는 사운드 필드의 압축된 고차 앰비소닉스(HOA) 표현을 디코딩하기 위한 장치로서,

상기 압축된 HOA 표현을 포함하는 비트 스트림을 수신하기 위한 수신기;

다수의 계층이 존재한다는 결정에 기초하여, 디코딩된 HOA 표현들의 시퀀스를 획득하기 위해 상기 비트 스트림 으로부터 상기 압축된 HOA 표현을 디코딩하기 위한 오디오 디코더

를 포함하고,

상기 디코딩된 HOA 표현들의 시퀀스의 제1 부분집합은 상기 디코딩된 HOA 표현들의 시퀀스의 인덱스들의 제1 세트에 대응하고, 상기 디코딩된 HOA 표현들의 시퀀스의 제2 부분집합은 상기 디코딩된 HOA 표현들의 시퀀스의 인덱스들의 제2 세트에 대응하며,

인덱스들의 상기 제1 세트는  $O_{\text{MIN}}$  채널들에 기초하고,

인덱스들의 상기 제1 세트 내의 각각의 인덱스에서는, 상기 제1 부분집합 내의 대응하는 디코딩된 HOA 표현은 대응하는 주변 HOA 컴포넌트에만 기초하여 결정되고,

인덱스들의 상기 제2 세트는 상기 다수의 계층 중 적어도 하나의 계층에 기초하여 결정되는 디코딩 장치.

#### 청구항 9

[청구항 9은(는) 설정등록료 납부시 포기되었습니다.]

제8항에 있어서, 인덱스들의 상기 제1 세트는  $1 \leq n \leq O_{\text{MIN}}$  에 기초하여 결정되고, 인덱스들의 상기 제2 세트는  $O_{\text{MIN}} + 1 \leq n \leq O$  에 기초하여 결정되고,  $O$  는 채널들의 총수를 나타내고,  $O_{\text{MIN}}$  은 1과  $O$  사이의 수를 나타내는 디코딩 장치.

#### 청구항 10

[청구항 10은(는) 설정등록료 납부시 포기되었습니다.]

제9항에 있어서, 인덱스  $n$  및 프레임  $k$ 에서는,  $n$ 이 인덱스들의 상기 제1 세트 내에 있을 때, 상기 제1 부분집합

은 대응하는 주변 사운드 컴포넌트  $\hat{c}_{AMB,n}(k-1)$  에 기초하여 결정되고,  $n$ 이 인덱스들의 상기 제2 세트 내에 있을 때, 상기 제2 부분집합은 대응하는 주변 사운드 컴포넌트  $\hat{c}_{n,AMB}(k-1)$  및 대응하는 우세 사운드 컴포넌트  $\hat{c}_{n,PS}(k-1)$  의 부가에 기초하여 결정되고, 상기 디코딩된 HOA 표현들은 적어도 부분적으로

$$\tilde{c}_n(k-1) = \begin{cases} \hat{c}_{AMB,n}(k-1) & \text{인덱스들의 상기 제1 세트 내의 } n \text{에 대해서} \\ \hat{c}_n(k-1) = \hat{c}_{PS,n}(k-1) + \hat{c}_{AMB,n}(k-1), & \text{인덱스들의 상기 제2 세트 내의 } n \text{에 대해서} \end{cases}$$

에 의해 표현되는 디코딩 장치.

#### 청구항 11

[청구항 11은(는) 설정등록료 납부시 포기되었습니다.]

제8항에 있어서,  $N_{MIN} \leq N$  일 때  $O_{MIN} = (N_{MIN} + 1)^2$  이고,  $N$ 은 상기 압축된 HOA 표현의 입력 프레임들의 오더인, 디코딩 장치.

#### 청구항 12

[청구항 12은(는) 설정등록료 납부시 포기되었습니다.]

제8항에 있어서, 상기 비트 스트림에 다수의 계층의 지시가 시그널링되는 디코딩 장치.

#### 청구항 13

[청구항 13은(는) 설정등록료 납부시 포기되었습니다.]

제8항에 있어서, 상기 다수의 계층은 베이스 계층 및 적어도 하나의 향상 계층을 포함하는 디코딩 장치.

#### 청구항 14

[청구항 14은(는) 설정등록료 납부시 포기되었습니다.]

제8항에 있어서, 상기 오디오 디코더는, 프레임  $k$ 에 대해, 주변 할당 벡터( $\mathbf{v}_{AMB,ASSIGN}(k)$ ), 방향 표현 및 각각의 양자화된 방향의 인덱스를 포함하는 제1 튜플 세트  $\mathcal{M}_{DIR}(k+1)$ , 및 벡터 기반 표현 및 상기 벡터 기반 표현의 방향 분포를 정의하는 벡터의 인덱스를 포함하는 제2 튜플 세트  $\mathcal{M}_{VEC}(k+1)$ 에 기초하여 상기 디코딩된 HOA 표현들의 시퀀스를 결정하도록 더 구성되는 디코딩 장치.

#### 청구항 15

[청구항 15은(는) 설정등록료 납부시 포기되었습니다.]

제8항에 있어서,

상기 오디오 디코더는, 채널 재할당 동안, 프레임  $k$ 에서 활성인 계수 시퀀스들에 관한 인덱스들의 제3 세트 ( $\mathcal{J}_{AMB,ACT}(k)$ ), 및 프레임  $(k-1)$ 에서 각각이 인에이블링될 필요가 있고, 디스에이블링될 필요가 있고, 활성으로 유지될 필요가 있는, 계수 시퀀스들에 관한 인덱스들의 제2 세트( $\mathcal{J}_E(k-1), \mathcal{J}_D(k-1), \mathcal{J}_U(k-1)$ )를 생성하도록 더 구성되는 디코딩 장치.

#### 청구항 16

삭제

#### 청구항 17

삭제

#### 청구항 18

삭제

청구항 19

삭제

청구항 20

삭제

청구항 21

삭제

## 발명의 설명

## 기술 분야

[0001] 본 발명은 고차 앰비소닉스(HOA)(Higher Order Ambisonics) 신호를 압축하는 방법, 압축된 HOA 신호를 압축 해제하는 방법, HOA 신호를 압축하기 위한 장치, 및 압축된 HOA 신호를 압축 해제하기 위한 장치에 관한 것이다.

## 배경 기술

[0002] 고차 앰비소닉스(HOA)는 3차원 사운드를 표현하는 가능성을 제공한다. 다른 알려진 기법들은 웨이브 필드 합성(WFS)(wave field synthesis), 또는 22.2와 유사한 채널 기반 접근법들이다. 그러나 채널 기반 방법과 상반되게, HOA 표현은 특정한 라우드스피커 세트-업과 무관한 장점을 제공한다. 그러나 이런 유연성은 특정한 라우드스피커 세트-업에 대한 HOA 표현의 재생에 요구되는 디코딩 프로세스가 희생된다. 필요한 라우드스피커의 수가 통상 매우 큰 WFS 접근법과 비교하여, HOA는 또한, 단지 소수의 라우드스피커로 구성되는 세트-업들로 렌더링될 수 있다. HOA의 추가 장점은 동일 표현이 또한 헤드폰들로의 바이노럴 렌더링(binaural rendering)을 위한 임의의 수정 없이 이용될 수 있다는 점이다.

[0003] HOA는 생략된(truncated) 구면 조화 함수(Spherical Harmonics)(SH) 전개에 의한 복합 조화 평면파(complex harmonic plane wave) 진폭들의 소위 공간 밀도의 표현에 기초한다. 각각의 전개 계수(expansion coefficient)는 시간 도메인 함수에 의해 등가적으로 표현될 수 있는 각 주파수의 함수이다. 따라서, 일반성의 손실 없이, 완전한 HOA 사운드 필드 표현은 실제로  $O$  시간 도메인 함수들로 구성된다고 가정될 수 있으며,  $O$ 는 전개 계수들의 수를 나타낸다. 이러한 시간 도메인 함수들은 이하에서 HOA 계수 시퀀스들 또는 HOA 채널들로서 등가적으로 지칭될 것이다. 보통, 구면 좌표계(Spherical coordinate system)는 x축이 정면 위치를 가리키고, y축이 좌측을 가리키며, z 축이 상측을 가리키는 경우에 이용된다. 공간  $\mathbf{x} = (r, \theta, \phi)^T$  내의 위치는 반경  $r > 0$ (즉, 좌표 원점까지의 거리), 극 축 z로부터 측정된 경사 각도  $\theta \in [0, \pi]$ , 및 x축으로부터 x-y 평면에서 시계 반대 방향으로 측정된 방위 각도  $\phi \in [0, 2\pi]$ 에 의해 표현된다. 또한,  $(\cdot)^T$ 는 이항(transposition)을 표시한다.

[0004] HOA 코딩의 더 상세한 설명은 다음에 제공된다.

[0005]  $\mathcal{F}_t(\cdot)$ 로 표시되는 시간에 대한 음압의 푸리에 변환, 즉  $P(\omega, \mathbf{x}) = \mathcal{F}_t(p(t, \mathbf{x})) = \int_{-\infty}^{\infty} p(t, \mathbf{x}) e^{-i\omega t} dt$  ( $\omega$ 는 각 주파수를 표시하고,  $i$ 는 허수 단위를 나타냄)은  $P(\omega = kc_s, r, \theta, \phi) = \sum_{n=0}^N \sum_{m=-n}^n A_n^m(k) j_n(kr) S_n^m(\theta, \phi)$ 에 따라 구면 조화함수의 급수들로 전개될 수 있다. 여기서,  $c_s$ 는 사운드의 속도를 나타내고,  $k$ 는 각 파수(angular wavenumber)를 나타내며, 이것은  $k = \frac{\omega}{c_s}$ 에 의해 각 주파수  $\omega$ 와 관련된다. 더욱이,  $j_n(\cdot)$ 는 제1종(first kind)의 구면 베셀 함수를 표시하고,  $S_n^m(\theta, \phi)$ 는 오더(order)  $n$  및 차수(degree)  $m$ 의 실가(real valued) 구면 조화 함수를 표시한다. 전개 계수  $A_n^m(k)$ 는 각 파수  $k$ 에만 의존한다. 음압이 공간적으로 대역-제한된다는 것이 암시적으로 가정됨에 유의한다. 따라서, 급수는 HOA 표현의 오더로 불리는 상한  $N$ 에서 오더 인덱스  $n$ 에 대해 생략된다. 사운드 필드가 상이한

각 주파수들  $\omega$ 의 무한 수의 조화 평면 파들의 중첩에 의해 표현되고 각도 튜플(angle tuple)  $(\theta, \phi)$ 에 의해 지정된 모든 가능한 방향으로부터 도달하는 경우, 각각의 평면파 복소 진폭 함수  $C(\omega, \theta, \phi)$ 는 다음의 구면 조화 함수 전개에 의해 표현될 수 있다:

$$C(\omega = kc_s, \theta, \phi) = \sum_{n=0}^N \sum_{m=-n}^n C_n^m(k) S_n^m(\theta, \phi)$$

여기서 오더 계수들  $C_n^m(k)$ 은  $A_n^m(k) = i^n C_n^m(k)$ 에 의해 전개 계수  $A_n^m(k)$ 와 관련된다.

개별 계수들  $C_n^m(\omega = kc_s)$ 이 각 주파수  $\omega$ 의 함수들인 것으로 가정하면, 역 푸리에 변환( $\mathcal{F}^{-1}(\cdot)$ 로 표시됨)의 적용은 각각의 오더  $n$  및 차수  $m$ 에 대한 시간 도메인 함수들  $c_n^m(t) = \mathcal{F}_t^{-1}(C_n^m(\omega/c_s)) = \frac{1}{2\pi} \int_{-\infty}^{\infty} C_n^m\left(\frac{\omega}{c_s}\right) e^{i\omega t} d\omega$  을 제공하고, 이것은  $c(t) = [c_0^0(t) \ c_1^{-1}(t) \ c_1^0(t) \ c_1^1(t) \ c_2^{-2}(t) \ c_2^{-1}(t) \ c_2^0(t) \ \dots \ c_N^{N-1}(t) \ c_N^N(t)]^T$ 에 의해 단일 벡터  $\mathbf{c}(t)$ 에 수집될 수 있다.

벡터  $\mathbf{c}(t)$  내의 시간 도메인 함수  $c_n^m(t)$ 의 위치 인덱스는  $n(n+1) + 1 + m$ 에 의해 주어진다. 벡터  $\mathbf{c}(t)$  내의 요소들의 전체 수는  $O = (N+1)^2$ 에 의해 주어진다. 함수들  $c_n^m(t)$ 의 이산 시간 버전들은 임베디드스 계수 시퀀스들로 지칭된다. 프레임 기반 HOA 표현은 다음과 같이 이들 시퀀스 모두를 길이 B의 프레임들  $\mathbf{C}(k)$  및 프레임 인덱스  $k$ 로 분할함으로써 획득된다:

$$\mathbf{C}(k) := [c((kB+1)T_s) \ c((kB+2)T_s) \ \dots \ c((kB+B)T_s) \ ]$$

$$\mathbf{C}(k) = \begin{bmatrix} \mathbf{c}_1(k) \\ \mathbf{c}_2(k) \\ \vdots \\ \mathbf{c}_O(k) \end{bmatrix}$$

여기서,  $T_s$ 는 샘플링 주기를 표시한다. 프레임  $\mathbf{C}(k)$  자체는 이후 행들의 구성(composition)  $\mathbf{c}_i(k)$  ( $i = 1, \dots, O$ )으로서 표현될 수 있고,

$\mathbf{c}_i(k)$ 은 위치 인덱스  $i$ 를 갖는 임베디드스 시퀀스의 프레임을 표시한다. HOA 표현의 공간 해상도(spatial resolution)는 전개(expansion)의 증가하는 최대 오더  $N$ 에 따라 개선된다. 유감스럽게도, 전개 계수들의 수  $O$ 는 오더  $N$ 에 따라 2차식으로(quadratically), 특히  $O = (N+1)^2$ 으로 증가한다. 예를 들어, 오더  $N=4$ 를 이용하는 전형적인 HOA 표현은  $O=25$  HOA(전개) 계수들을 요구한다. 이러한 고려 사항에 따르면, HOA 표현의 송신을 위한 전체 비트레이트는, 원하는 단일 채널 샘플링 레이트  $f_s$  및 샘플당 비트들의 수  $N_b$ 가 주어지면,  $O \cdot f_s \cdot N_b$ 에 의해 결정된다. 따라서, 오더  $N = 4$ 의 HOA 표현을, 샘플당  $N_b = 16$ 비트를 이용하여  $f_s = 48kHz$  샘플링 레이트로 송신하는 것은,  $19.2Mbits/s$ 의 비트레이트를 초래하며, 이것은 예를 들어, 스트리밍과 같은 많은 실제 애플리케이션에서 매우 크다. 따라서, HOA 표현들의 압축이 매우 바람직하다. 이전에, HOA 사운드 필드 표현의 압축은 유럽 특허 출원들 EP2743922A, EP2665208A 및 EP2800401A에서 제안되었다. 이러한 접근법들은, 사운드 필드 분석을 수행하고 주어진 HOA 표현을 방향성 컴포넌트 및 잔차 주변 컴포넌트(residual ambient component)로 분해하는 것을 통상 갖는다. 한편, 최종 압축된 표현은 복수의 양자화된 신호들을 포함한다고 가정되며, 이는 주변 HOA 컴포넌트의 관련 계수 시퀀스들 및 방향 신호들의 지각 코딩으로부터 발생한다. 한편, 양자화된 신호들과 관련된 추가 사이드 정보가 포함된다고 가정되며, 그 사이드 정보는 그것의 압축된 버전으로부터 HOA 표현의 재구성을 위해 필요하다.

또한, 유사한 방법은 ISO/IEC JTC1/SC29/WG11 N14264(MPEG-H 3D 오디오, 2014년 1월, 산호세의 규격 초안 1-HOA 텍스트)에 기술되며, 여기서 방향 컴포넌트는 소위 우세 사운드 컴포넌트(predominant sound component)로 연장된다. 방향 컴포넌트로서, 우세 사운드 컴포넌트는 방향 신호들로부터 최초 HOA 표현의 일부를 예측하기 위한 일부 예측 파라미터들과 함께, 방향 신호들, 즉, 청취자들에게 나쁜 영향을 미친다고 가정되는 대응하는 방향을 갖는 모노럴 신호(monaural signal)들에 의해 부분적으로 표현된다고 가정된다. 또한, 우세 사운드 컴



포넌트는 소위 벡터 기반 신호들에 의해 표현되기로 되어 있고, 이는 벡터 기반 신호들의 방향 분포를 정의하는 대응하는 벡터를 갖는 모너럴 신호들을 의미한다. 알려진 압축된 HOA 표현은  $I$  양자화된 모노럴 신호들 및 일부 추가 사이드 정보로 구성되며, 여기서  $I$  양자화된 모너럴 신호들 중에서 고정된 수  $0_{\text{MIN}}$ 는 주변 HOA 컴포넌트  $\mathbf{C}_{\text{AMB}}(k-2)$ 의 제1  $0_{\text{MIN}}$  계수 시퀀스들의 공간 변환된 버전을 표현한다. 나머지  $I - 0_{\text{MIN}}$  신호들의 유형은 연속 프레임들 사이에서 변할 수 있고, 방향성 벡터 기반의 엠프티(empty)이거나, 또는 주변 HOA 컴포넌트  $\mathbf{C}_{\text{AMB}}(k-2)$ 의 추가 계수 시퀀스를 표현할 수 있다.

[0014]

HOA 계수 시퀀스의 입력 시간 프레임들( $\mathbf{C}^{(k)}$ )로 HOA 신호 표현을 압축하는 알려진 방법은 입력 시간 프레임들의 공간 HOA 인코딩과, 후속 지각 인코딩 및 소스 인코딩을 포함한다. 도 1a에 도시된 바와 같이, 공간 HOA 인코딩은 방향 및 벡터 추정 블록(101)에서 HOA 신호의 방향 및 벡터 추정 처리를 수행하는 단계를 포함하고, 여기서 방향 신호를 위한 제1 토폴 세트들  $\mathcal{M}_{\text{DIR}}(k)$  및 벡터 기반 신호들을 위한 제2 토폴 세트들  $\mathcal{M}_{\text{VEC}}(k)$ 을 포함하는 데이터가 획득된다. 제1 토폴 세트들 각각은 방향 신호와 각각의 양자화된 방향의 인덱스를 포함하고, 제2 토폴 세트들 각각은 벡터 기반 신호 및 상기 신호들의 방향 분포를 정의하는 벡터의 인덱스를 포함한다. 다음 단계는 HOA 계수 시퀀스들의 각각의 입력 시간 프레임을 복수의 우세 사운드 신호  $\mathbf{x}_{\text{PS}}(k-1)$ 의 프레임 및 주변 HOA 컴포넌트  $\mathbf{C}_{\text{AMB}}(k-1)$ 의 프레임으로 분해하는(103) 단계이고, 우세 사운드 신호들  $\mathbf{x}_{\text{PS}}(k-1)$ 은 상기 방향 사운드 신호들 및 상기 벡터 기반 사운드 신호들을 포함한다. 상기 분해하는 단계는 또한, 예측 파라미터들  $\xi(k-1)$  및 타깃 할당 벡터  $\mathbf{v}_{\text{A,T}}(k-1)$ 를 제공한다. 예측 파라미터들  $\xi(k-1)$ 은 우세 사운드 HOA 컴포넌트들을 강화하기 위해 우세 사운드 신호들  $\mathbf{x}_{\text{PS}}(k-1)$  내의 방향 신호들로부터 HOA 신호 표현의 일부를 어떻게 예측하는지를 기술하고, 타깃 할당 벡터  $\mathbf{v}_{\text{A,T}}(k-1)$ 는 우세 사운드 신호들을 주어진 수  $I$ 의 채널들에 어떻게 할당하는에 대한 정보를 포함한다. 주변 HOA 컴포넌트  $\mathbf{C}_{\text{AMB}}(k-1)$ 은 타깃 할당 벡터  $\mathbf{v}_{\text{A,T}}(k-1)$ 에 의해 제공되는 정보에 따라 수정되고(104), 주변 HOA 컴포넌트의 어느 계수 시퀀스들이 주어진 수  $I$ 의 채널들에서 송신될 것인지는, 우세 사운드 신호가 얼마나 많은 채널을 차지하는지에 따라 결정된다. 수정된 주변 HOA 컴포넌트  $\mathbf{C}_{\text{M,A}}(k-2)$  및 시간적으로 예측된 수정된 주변 HOA 컴포넌트  $\mathbf{C}_{\text{P,M,A}}(k-1)$ 가 획득된다. 또한, 최종 할당 벡터  $\mathbf{v}_{\text{A}}(k-2)$ 는 타깃 할당 벡터  $\mathbf{v}_{\text{A,T}}(k-1)$  내의 정보로부터 획득된다. 상기 분해 단계로부터 획득된 우세 사운드 신호들  $\mathbf{x}_{\text{PS}}(k-1)$ , 및 수정된 주변 HOA 컴포넌트  $\mathbf{C}_{\text{M,A}}(k-2)$  및 시간적으로 예측된 수정된 주변 HOA 컴포넌트  $\mathbf{C}_{\text{P,M,A}}(k-1)$ 의 결정된 계수 시퀀스들은, 최종 할당 벡터  $\mathbf{v}_{\text{A}}(k-2)$ 에 의해 제공된 정보를 이용하여 주어진 수의 채널들에 할당되고, 전송 신호들  $\mathbf{y}_i(k-2)$  ( $i = 1, \dots, I$ ) 및 예측된 전송 신호들  $\mathbf{y}_{\text{P},i}(k-2)$  ( $i = 1, \dots, I$ )이 획득된다. 그 후, 이득 제어(또는 정규화)가 전송 신호들  $\mathbf{y}_i(k-2)$  및 예측된 전송 신호들  $\mathbf{y}_{\text{P},i}(k-2)$ 에 대해 수행되고, 이득 수정된 전송 신호들  $\mathbf{z}_i(k-2)$ , 지수들  $e_i(k-2)$  및 예외 플래그들 ( $\beta_i(k-2)$ )이 획득된다.

[0015]

도 1b에 도시된 바와 같이, 지각 인코딩 및 소스 인코딩은 이득 수정된 전송 신호들  $\mathbf{z}_i(k-2)$ 의 지각 코딩을 포함하고, 지각 인코딩된 전송 신호들  $\tilde{\mathbf{z}}_i(k-2)$ ,  $i = 1, \dots, I$ 이 획득되고, 상기 지수들  $e_i(k-2)$  및 예외 플래그들 ( $\beta_i(k-2)$ ), 제1 및 제2 토폴 세트들  $\mathcal{M}_{\text{DIR}}(k)$  및  $\mathcal{M}_{\text{VEC}}(k)$ , 예측 파라미터들  $\xi(k-1)$  및 최종 할당 벡터  $\mathbf{v}_{\text{A}}(k-2)$ 를 포함하는 사이드 정보를 인코딩하여, 인코딩된 사이드 정보  $\tilde{\mathbf{r}}(k-2)$ 가 획득된다. 결국, 지각 인코딩된 전송 신호들  $\tilde{\mathbf{z}}_i(k-2)$  및 인코딩된 사이드 정보가 비트스트림으로 멀티플렉싱된다.

### 발명의 내용

[0016]

제안된 HOA 압축 방법의 한 가지 단점은 이것이 모놀리식(monolithic)(즉, 논-스케일러블(non-scalable)) 압축된 HOA 표현을 제공한다는 점이다. 그러나, 브로드캐스팅 또는 인터넷 스트리밍과 같은 소정의 애플리케이션에서는, 압축된 표현을 저품질 베이스 계층(BL)(base layer) 및 고품질 향상 계층(EL)(enhancement layer)으로 분할할 수 있는 것이 바람직하다. 베이스 계층은 HOA 표현의 저품질 압축 버전을 제공하기로 되어 있고, 이것



은 항상 계층과 무관하게 디코딩될 수 있다. 그러한 BL은 전형적으로 송신 에러들에 대하여 매우 강건(robust)해야 하고, 열악한 송신 상태하에서도 압축 해제된 HOA 표현의 소정의 최소 품질을 보장하기 위해 낮은 데이터 레이트에서 송신되어야 한다. EL은 압축 해제된 HOA 표현의 품질을 향상시키기 위해 추가 정보를 포함한다.

[0017] 본 발명은 (저품질) 베이스 계층 및 (고품질) 향상 계층을 포함하는 압축된 표현을 제공할 수 있도록 기존 HOA 압축 방법을 수정하기 위한 해결책을 제공한다. 또한, 본 발명은 본 발명에 따라 압축되는 저품질 베이스 계층을 적어도 포함하는 압축된 표현을 디코딩할 수 있도록 기존 HOA 압축 해제 방법을 수정하기 위한 해결책을 제공한다.

[0018] 하나의 개선책은 자납식(self contained)(저품질) 베이스 계층을 획득하는 것과 관련된다. 본 발명에 따르면, 주변 HOA 컴포넌트  $\mathbf{C}_{\text{AMB}}(k-2)$ 의 (일반성 상실 없는) 제1  $O_{\text{MIN}}$  계수 시퀀스들의 공간 변환된 버전을 포함하기로 되어 있는  $O_{\text{MIN}}$  채널들은 베이스 계층으로 이용된다. 베이스 계층을 형성하기 위해 제1  $O_{\text{MIN}}$  채널들을 선택하는 이점은 그들의 시간 불변 유형에 있다. 그러나 종래에, 각각의 신호들에는 사운드 신(sound scene)에 필수적인 임의의 우세 사운드 컴포넌트들이 부족하다. 이것은 주변 HOA 컴포넌트  $\mathbf{C}_{\text{AMB}}(k-1)$ 의 종래의 계산으로부터도 명백하고, 이런 계산은

### 수학적 1

[0019]  $\mathbf{C}_{\text{AMB}}(k-1) = \mathbf{C}(k-1) - \mathbf{C}_{\text{PS}}(k-1)$

[0020] 에 따라, 최초 HOA 표현  $\mathbf{C}(k-1)$ 에서 우세 사운드 HOA 표현  $\mathbf{C}_{\text{PS}}(k-1)$ 의 차감에 의해 수행된다.

[0021] 따라서, 본 발명의 하나의 개선책은 그런 우세 사운드 컴포넌트들의 가산과 관련된다. 본 발명에 따르면, 이 문제에 대한 해결책은 베이스 계층 내에 낮은 공간 해상도에서의 우세 사운드 컴포넌트를 포함시키는 것이다. 이 목적을 위해, 본 발명에 따르는 공간 HOA 인코더에서 HOA 분해 처리에 의해 출력되는 주변 HOA 컴포넌트  $\mathbf{C}_{\text{AMB}}(k-1)$ 는 그것의 수정된 버전에 의해 대체된다. 수정된 주변 HOA 컴포넌트는 공간 변환된 형태로 항상 송신되기로 되어 있는 제1  $O_{\text{MIN}}$  계수 시퀀스들 내에, 최초 HOA 컴포넌트의 계수 시퀀스들을 포함한다. HOA 분해 처리의 이런 개선은 계층화 모드(예를 들어, 듀얼 계층 모드)에서 HOA 압축 작업을 만들기 위한 초기 동작으로 보여질 수 있다. 이 모드는, 예를 들어 베이스 계층 및 향상 계층으로 분할될 수 있는 2개의 비트 스트림 또는 단일 비트 스트림을 제공한다. 이 모드를 이용하거나 이용하지 않는 것은 전체 비트 스트림의 액세스 단위에서 모드 지시 비트(예를 들어, 단일 비트)에 의해 신호화된다.

[0022] 일 실시예에서, 베이스 계층 비트 스트림  $\mathbf{B}_{\text{BASE}}(k-2)$ 은 각각 인코딩된 신호들  $\mathbf{z}_i(k-2)$ ,  $i = 1, \dots, O_{\text{MIN}}$ , 및 대응하는 코딩된 이득 제어 사이드 정보를 포함할 뿐이고, 그 사이드 정보는 지수들  $e_i(k-2)$  및 예외 플래그들  $\beta_i(k-2)$ ,  $i = 1, \dots, O_{\text{MIN}}$ 로 구성된다. 나머지 각각 인코딩된 신호들  $\mathbf{z}_i(k-2)$ ,  $i = O_{\text{MIN}} + 1, \dots, O$ , 및 인코딩된 나머지 사이드 정보는 향상 계층 비트 스트림에 포함된다. 일 실시예에서, 베이스 계층 비트 스트림  $\mathbf{B}_{\text{BASE}}(k-2)$  및 향상 계층 비트 스트림  $\mathbf{B}_{\text{ENH}}(k-2)$ 은 그 후 이전의 전체 비트 스트림  $\mathbf{B}(k-2)$  대신에 함께 송신된다.

[0023] HOA 계수 시퀀스들의 시간 프레임들을 갖는 고차 앰비소닉스(HOA) 신호 표현을 압축하는 방법은 청구항 1에 개시된다. HOA 계수 시퀀스들의 시간 프레임들을 갖는 고차 앰비소닉스(HOA) 신호 표현을 압축하기 위한 장치는 청구항 10에 개시된다.

[0024] HOA 계수 시퀀스들의 시간 프레임들을 갖는 고차 앰비소닉스(HOA) 신호 표현을 압축 해제하는 방법은 청구항 8에 개시된다. HOA 계수 시퀀스들의 시간 프레임들을 갖는 고차 앰비소닉스(HOA) 신호 표현을 압축 해제하기 위한 장치는 청구항 18에 개시된다.

[0025] 컴퓨터로 하여금 HOA 계수 시퀀스들의 시간 프레임들을 갖는 고차 앰비소닉스(HOA) 신호 표현을 압축하는 방법을 수행하게 하기 위한 실행 가능 명령어들을 갖는 비일시적 컴퓨터 판독가능 저장 매체는 청구항 20에 개시된다.

다.

[0026] 컴퓨터로 하여금 HOA 계수 시퀀스들의 시간 프레임들을 갖는 고차 앰비소닉스(HOA) 신호 표현을 압축 해제하는 방법을 수행하게 하기 위한 실행 가능 명령어들을 갖는 비일시적 컴퓨터 판독가능 저장 매체는 청구항 21에 개시된다.

[0027] 본 발명의 유리한 실시예들은 종속 청구항, 하기의 설명 및 도면에 개시된다.

### 도면의 간단한 설명

[0028] 본 발명의 예시적 실시예들이 첨부 도면들을 참조하여 설명된다. 도면에서:

도 1은 HOA 압축기의 종래의 아키텍처의 구조이다;

도 2는 HOA 압축 해제기의 종래의 아키텍처의 구조이다;

도 3은 본 발명의 일 실시예에 따른, HOA 압축기의 공간 HOA 인코딩 및 지각 인코딩 부의 아키텍처의 구조이다;

도 4는 본 발명의 일 실시예에 따른, HOA 압축기의 소스 코더 부의 아키텍처의 구조이다;

도 5는 본 발명의 일 실시예에 따른, HOA 압축 해제기의 지각 디코딩 및 소스 디코딩 부의 아키텍처의 구조이다;

도 6은 본 발명의 일 실시예에 따른, HOA 압축 해제기의 공간 HOA 디코딩 부의 아키텍처의 구조이다;

도 7은 주변 HOA 신호들로부터 수정된 주변 HOA 신호들의 프레임들의 변환이다;

도 8은 HOA 신호를 압축하는 방법의 흐름도이다;

도 9는 압축된 HOA 신호를 압축 해제하는 방법의 흐름도이다; 및

도 10은 본 발명의 일 실시예에 따른, HOA 압축 해제기의 공간 HOA 디코딩 부의 아키텍처의 부분들의 상세도이다.

### 발명을 실시하기 위한 구체적인 내용

[0029] 더 쉬운 이해를 위해, 도 1 및 도 2의 종래 기술 해결책은 이하에서 요약된다.

[0030] 도 1은 HOA 압축기의 종래의 아키텍처의 구조를 도시한다. [4]에 기술된 방법에서, 방향 컴포넌트는 소위 우세 사운드 컴포넌트로 연장된다. 방향 컴포넌트로서, 우세 사운드 컴포넌트는 방향 신호들에 의해 부분적으로 표현된다고 가정되며, 그 방향 신호들은 방향 신호들로부터 최초 HOA 표현의 일부를 예측하기 위한 일부 예측 파라미터들과 함께, 청취자들에게 나쁜 영향을 미친다고 가정되는 대응하는 방향을 갖는 모노럴 신호들을 의미한다. 또한, 우세 사운드 컴포넌트는 소위 벡터 기반 신호들에 의해 표현되기로 되어 있으며, 이는 벡터 기반 신호들의 방향 분포를 정의하는 대응하는 벡터를 갖는 모노럴 신호들을 의미한다. [4]에 제안된 HOA 압축기의 전체 아키텍처는 도 1에 예시된다. 이것은 도 1a에 도시된 공간 HOA 인코딩 부 및 도 1b에 도시된 지각 및 소스 인코딩 부로 세분될 수 있다. 공간 HOA 인코딩은  $I$  신호들과 사이드 정보로 함께 구성되는 제1 압축된 HOA 표현을 제공하며, 사이드 정보는  $I$  신호들의 HOA 표현을 어떻게 생성할지를 기술한다. 지각 및 사이드 정보 소스 코더에서, 언급된  $I$  신호들은 지각 인코딩되고, 사이드 정보는 2개의 코딩된 표현을 멀티플렉싱하기 전에 소스 인코딩을 거친다.

[0031] 통상적으로, 공간 인코딩 작업은 다음과 같다.

[0032] 제1 단계에서, 최초 HOA 표현의  $k$ 번째 프레임  $\mathbf{c}(k)$ 은 튜플 세트들  $\mathcal{M}_{\text{DIR}}(k)$  및  $\mathcal{M}_{\text{VEC}}(k)$ 을 제공하는 방향 및 벡터 추정 처리 블록에 입력된다. 튜플 세트  $\mathcal{M}_{\text{DIR}}(k)$ 는 제1 요소가 방향 신호의 인덱스를 표시하고 제2 요소가 각각의 양자화된 방향을 표시하는 튜플(tuple)들로 구성된다. 튜플 세트  $\mathcal{M}_{\text{VEC}}(k)$ 는 제1 요소가 벡터 기반 신호의 인덱스를 지시하고 제2 요소가 신호들의 방향 분포(즉, 벡터 기반 신호의 HOA 표현이 어떻게 계산되는지)를 정의하는 벡터를 표시하는 튜플들로 구성된다.

[0033] 양쪽 튜플 세트  $\mathcal{M}_{\text{DIR}}(k)$  및  $\mathcal{M}_{\text{VEC}}(k)$ 를 이용하여, 초기 HOA 프레임  $\mathbf{c}(k)$ 은 HOA 분해에서, 모든 우세 사운드(즉,

방향 및 벡터 기반) 신호들의 프레임  $\mathbf{x}_{ps}(k-1)$  및 주변 HOA 컴포넌트의 프레임  $\mathbf{c}_{AMB}(k-1)$ 로 분해된다. 아티팩트 차단을 방지하기 위해 하나의 프레임의 지연이 각각 가산 처리에 오버랩될 예정임에 유의한다. 더욱이, HOA 분해는 우세 사운드 HOA 컴포넌트를 강화하기 위해 방향 신호들로부터 최초 HOA 표현의 부분들을 어떻게 예측하는지를 기술하는 일부 예측 파라미터  $\boldsymbol{\zeta}(k-1)$ 를 출력한다고 가정된다. 또한, HOA 분해 처리 블록에서 결정되었던, 우세 사운드 신호들의  $I$  가용 채널들로의 할당에 대한 정보를 포함하는 타깃 할당 벡터  $\mathbf{v}_{A,T}(k-1)$ 가 제공된다. 영향을 받은 채널들은 점유된 것으로 가정될 수 있으며, 이것은 이 채널들이 각각의 시간 프레임에서 주변 HOA 컴포넌트의 임의의 계수 시퀀스들을 전송하는데 이용 가능하지 않다는 것을 의미한다.

[0034] 주변 컴포넌트 수정 처리 블록에서, 주변 HOA 컴포넌트  $\mathbf{c}_{AMB}(k-1)$ 의 프레임은 타깃 할당 벡터  $\mathbf{v}_{A,T}(k-1)$ 에 의해 제공된 정보에 따라 수정된다. 특히, 주변 HOA 컴포넌트의 어느 계수 시퀀스가 주어진  $I$  채널들에서 송신될 것인지는, 다른 양태들 중에서도 특히 어느 채널들이 이용 가능하고 우세 사운드 신호에 의해 점유되지 않았는지에 대한 정보(타깃 할당 벡터  $\mathbf{v}_{A,T}(k-1)$ 에 포함됨)에 따라 결정된다. 또한, 계수 시퀀스들의 페이드인(fade in) 또는 페이드아웃(fade out)은 선택된 계수 시퀀스들의 인덱스들이 연속 프레임들마다 차이가 있는 경우 수행된다.

[0035] 게다가, 주변 HOA 컴포넌트  $\mathbf{c}_{AMB}(k-2)$ 의 제1  $O_{MIN}$  계수 시퀀스들이 지각 코딩되고 송신되도록 항상 선택된다고 가정하며, 여기서  $O_{MIN} = (N_{MIN} + 1)^2$ 이 되고,  $N_{MIN} \leq N$ 이 되며,  $N$ 은 전형적으로 최초 HOA 표현의 오더보다 작은 오더이다. 이러한 HOA 계수 시퀀스들을 역 상관(de-correlate)시키기 위해, 이들을 일부 사전 정의된 방향들  $\boldsymbol{\Omega}_{MIN,d}, d = 1, \dots, O_{MIN}$ 로부터 침투하는 방향 신호들(즉, 일반 평면파 함수들)로 변환하는 것이 제안된다. 수정된 주변 HOA 컴포넌트  $\mathbf{c}_{M,A}(k-1)$ 과 함께, 시간적으로 예측된 수정된 주변 HOA 컴포넌트  $\mathbf{c}_{P,M,A}(k-1)$ 은 적당한 룩 어헤드(look ahead)를 허용하기 위해 이득 제어 처리 블록에서 차후에 사용되도록 계산된다.

[0036] 주변 HOA 컴포넌트의 수정에 대한 정보는 모든 가능한 유형들의 신호들의 이용 가능한 채널로의 할당에 직접 관련된다. 할당에 대한 최종 정보는 최종 할당 벡터  $\mathbf{v}_A(k-2)$ 에 포함된다. 이 벡터를 계산하기 위해, 타깃 할당 벡터  $\mathbf{v}_{A,T}(k-1)$ 에 포함되는 정보가 이용된다.

[0037] 채널 할당은 할당 벡터  $\mathbf{v}_A(k-2)$ 에 의해 제공되는 정보와 함께,  $\mathbf{x}_{ps}(k-2)$ 에 포함되고  $\mathbf{c}_{M,A}(k-2)$ 에 포함되는 적당한 신호들을  $I$  이용 가능한 채널들에 할당하여, 신호들  $\mathbf{y}_i(k-2), i = 1, \dots, I$ 을 산출한다. 더욱이,  $\mathbf{x}_{ps}(k-1)$  및  $\mathbf{c}_{P,AMB}(k-1)$ 에 포함되는 적당한 신호들은 또한,  $I$  이용 가능한 채널들에 할당되어, 예측된 신호들  $\mathbf{y}_{p,i}(k-2), i = 1, \dots, I$ 을 산출한다. 신호들  $\mathbf{y}_i(k-2), i = 1, \dots, I$ 의 각각은 이득 제어에 의해 최종적으로 처리되고, 여기서 신호 이득은 지각 인코더들에 적합한 값 범위를 달성하도록 원활하게 수정된다. 예측된 신호 프레임들  $\mathbf{y}_{p,i}(k-2), i = 1, \dots, I$ 은 연속 블록들 사이에서 심각한 이득 변동을 방지하기 위해 일종의 룩 어헤드를 허용한다. 이득 수정은 지수들  $e_i(k-2)$  및 예외 플래그들  $\beta_i(k-2), i = 1, \dots, I$ 로 구성되는 이득 제어 사이드 정보를 이용하여 공간 디코더에서 복원된다고 가정된다.

[0038] 도 2는 [4]에서 제안된 바와 같이, HOA 압축 해제기의 종래의 아키텍처의 구조를 도시한다. 통상적으로, HOA 압축 해제기는 HOA 압축기 컴포넌트들의 대응 부분들로 구성되고, 이것은 분명히 반대 순서로 배열된다. 이것은 도 2a에 도시된 지각 및 소스 디코딩 부 및 도 2b에 도시된 공간 HOA 디코딩 부로 세분될 수 있다.

[0039] 지각 및 사이드 정보 소스 디코더에서, 비트 스트림은  $I$  신호들의 지각 코딩된 표현, 및 코딩된 사이드 정보로 먼저 디멀티플렉싱되고, 코딩된 사이드 정보는 그것의 HOA 표현이 어떻게 생성되는지를 기술한다. 연속해서,  $I$  신호들의 지각 디코딩 및 사이드 정보의 디코딩이 수행된다. 그 후, 공간 HOA 디코더는  $I$  신호 및 사이드 정보로부터 재구성된 HOA 표현을 생성한다.

- [0040] 통상적으로, 공간 HOA 디코딩 작업은 다음과 같다.
- [0041] 공간 HOA 디코더에서, 각각 디코딩된 신호들  $\hat{\mathbf{z}}_i(k)$ ,  $i \in \{1, \dots, I\}$  각각은 연관된 이득 정정 지수  $\mathbf{e}_i(k)$  및 이득 정정 예외 플래그  $\beta_i(k)$ 와 함께 역 이득 제어 처리 블록에 먼저 입력된다.  $i$ 번째 역 이득 제어 처리는 이득 정정된 신호 프레임  $\hat{\mathbf{y}}_i(k)$ 을 제공한다.
- [0042] 모든  $I$  이득 정정된 신호 프레임들  $\hat{\mathbf{y}}_i(k)$ ,  $i \in \{1, \dots, I\}$ 은 할당 벡터  $\mathbf{v}_{\text{AMB,ASSIGN}}(k)$  및 토폴 세트들  $\mathcal{M}_{\text{DIR}}(k+1)$  및  $\mathcal{M}_{\text{VEC}}(k+1)$ 와 함께 채널 제한당에 전달된다. 토폴 세트들  $\mathcal{M}_{\text{DIR}}(k+1)$  및  $\mathcal{M}_{\text{VEC}}(k+1)$ 은 (공간 HOA 인코딩을 위해) 위에 정의되었고, 할당 벡터  $\mathbf{v}_{\text{AMB,ASSIGN}}(k)$ 는  $I$  컴포넌트들로 구성되고, 이것은 각각의 송신 채널이 주변 HOA 컴포넌트의 계수 시퀀스들을 포함하는지와 어느 계수 시퀀스들을 포함하는지를 각각의 송신 채널에 대해 나타낸다. 채널 제한당에서, 이득 정정된 신호 프레임들  $\hat{\mathbf{y}}_i(k)$ 은 모든 우세 사운드 신호들(즉, 모든 방향 및 벡터 기반 신호들)의 프레임  $\hat{\mathbf{X}}_{\text{PS}}(k)$  및 주변 HOA 컴포넌트의 중간 표현의 프레임  $\mathbf{C}_{\text{LAMB}}(k)$ 을 재구성하도록 재분배된다. 또한,  $k$ 번째 프레임에서 활성화된, 주변 HOA 컴포넌트의 계수 시퀀스들의 인덱스들의 세트  $\mathcal{I}_{\text{AMB,ACT}}(k)$ , 및  $(k-1)$ 번째 프레임에서 인에이블되고 디스에이블되며 활성화로 유지될 필요가 있는, 주변 HOA 컴포넌트의 계수 인덱스들의 세트들  $\mathcal{I}_{\text{E}}(k-1)$ ,  $\mathcal{I}_{\text{D}}(k-1)$  및  $\mathcal{I}_{\text{U}}(k-1)$ 이 제공된다.
- [0043] 우세 사운드 합성에서, 우세 사운드 컴포넌트  $\hat{\mathbf{C}}_{\text{PS}}(k-1)$ 의 HOA 표현은 예측 파라미터들의 토폴 세트  $\mathcal{M}_{\text{DIR}}(k+1)$  및 세트  $\boldsymbol{\zeta}(k+1)$ , 토폴 세트  $\mathcal{M}_{\text{VEC}}(k+1)$  및 세트들  $\mathcal{I}_{\text{E}}(k-1)$ ,  $\mathcal{I}_{\text{D}}(k-1)$  및  $\mathcal{I}_{\text{U}}(k-1)$ 를 이용하여 모든 우세 사운드 신호들의 프레임들  $\hat{\mathbf{X}}_{\text{PS}}(k)$ 로부터 계산된다.
- [0044] 주변 합성(Ambience Synthesis)에서, 주변 HOA 컴포넌트 프레임  $\hat{\mathbf{C}}_{\text{AMB}}(k-1)$ 은  $k$ 번째 프레임에서 활성화된 주변 HOA 컴포넌트의 계수 시퀀스들의 인덱스들의 세트  $\mathcal{I}_{\text{AMB,ACT}}(k)$ 를 이용하여, 주변 HOA 컴포넌트의 중간 표현의 프레임  $\mathbf{C}_{\text{LAMB}}(k)$ 으로부터 생성된다. 하나의 프레임의 지연이 우세 사운드 HOA 컴포넌트와의 동기화에 기인하여 도입됨에 유의한다. 결국, HOA 구성에서, 우세 사운드 HOA 컴포넌트의 주변 HOA 컴포넌트 프레임  $\hat{\mathbf{C}}_{\text{AMB}}(k-1)$  및 프레임  $\hat{\mathbf{C}}_{\text{PS}}(k-1)$ 은 디코딩된 HOA 프레임  $\hat{\mathbf{C}}(k-1)$ 을 제공하기 위해 중첩된다.
- [0045] 상기 HOA 압축 및 압축 해제 방법의 개략적 설명으로부터 명백하게 된 바와 같이, 압축된 표현은  $I$  양자화된 모노럴 신호들 및 일부 추가 사이드 정보로 구성된다. 이들  $I$  양자화된 모노럴 신호들 중에서 고정된 수  $O_{\text{MIN}}$ 은 주변 HOA 컴포넌트  $\mathbf{C}_{\text{AMB}}(k-2)$ 의 제1  $O_{\text{MIN}}$  계수 시퀀스들의 공간 변환된 버전을 표현한다. 나머지  $I - O_{\text{MIN}}$  신호들의 유형은 연속 프레임들 사이에 변할 수 있으며, 이것은 방향성, 벡터 기반, 엠프티(empty)이거나 또는 주변 HOA 컴포넌트  $\mathbf{C}_{\text{AMB}}(k-2)$ 의 추가 계수 시퀀스들을 표현한다. 있는 그대로 보면, 압축된 HOA 표현은 모노리식인 것을 의미한다. 특히, 설명된 표현을 저품질 베이스 계층 및 향상 계층으로 어떻게 분할할지가 하나의 문제이다.
- [0046] 개시된 본 발명에 따르면, 저품질 베이스 계층을 위한 후보는 주변 HOA 컴포넌트  $\mathbf{C}_{\text{AMB}}(k-2)$ 의 제1  $O_{\text{MIN}}$  계수 시퀀스들의 공간 변환된 버전을 포함하는  $O_{\text{MIN}}$  채널들이다. 이러한(일반성 상실 없이: 먼저)  $O_{\text{MIN}}$  채널들이 저품질 베이스 계층을 형성하기 위해서는 이들의 시분변 유형을 선택하는 것이 좋다. 그러나 각각의 신호들은 사운드 신에 필수적인 임의의 우세 사운드 컴포넌트가 부족하다. 이것은 또한, 주변 HOA 컴포넌트  $\mathbf{C}_{\text{AMB}}(k-1)$ 의 계산에서 알 수 있으며, 이 계산은
- [0047] [수학식 1]
- [0048]  $\mathbf{C}_{\text{AMB}}(k-1) = \mathbf{C}(k-1) - \mathbf{C}_{\text{PS}}(k-1)$

[0049] 에 따라 최초 HOA 표현  $\mathbf{C}(k-1)$  으로부터 우세 사운드 HOA 표현  $\mathbf{C}_{PS}(k-1)$  의 차감에 의해 수행된다.

[0050] 이 문제에 대한 해결책은 낮은 공간 해상도에서의 우세 사운드 컴포넌트들을 베이스 계층에 포함시키는 것이다.

[0051] HOA 압축에 대한 제안된 보정은 이하에 기술된다.

[0052] 도 3은 본 발명의 일 실시예에 따른 HOA 압축기의 공간 HOA 인코딩 및 지각 인코딩 부의 아키텍처의 구조를 도시한다. 낮은 공간 해상도에서의 우세 사운드 컴포넌트들을 또한 베이스 계층 내에 포함시키기 위해, 공간 HOA 인코더(도 1a 참조)에서 HOA 분해 처리에 의해 출력되는 주변 HOA 컴포넌트  $\mathbf{C}_{AMB}(k-1)$  는 수정된 버전

## 수학식 2

$$\tilde{\mathbf{C}}_{AMB}(k-1) = \begin{bmatrix} \tilde{\mathbf{c}}_{AMB,1}(k-1) \\ \tilde{\mathbf{c}}_{AMB,2}(k-1) \\ \vdots \\ \tilde{\mathbf{c}}_{AMB,O}(k-1) \end{bmatrix}$$

[0054] 에 의해 대체되고, 이들의 요소들은

## 수학식 3

$$\tilde{\mathbf{c}}_{AMB,n}(k-1) = \begin{cases} \mathbf{c}_n(k-1) & \text{for } 1 \leq n \leq O_{MIN} \\ \mathbf{c}_{AMB,n}(k-1) & \text{for } O_{MIN} + 1 \leq n \leq O \end{cases}$$

[0056] 에 의해 주어진다.

[0057] 다시 말해서, 공간 변환된 형태로 항상 송신되기로 되어 있는 주변 HOA 컴포넌트의 제1  $O_{MIN}$  계수 시퀀스들은 최초 HOA 컴포넌트의 계수 시퀀스들에 의해 대체된다. 공간 HOA 인코더의 다른 처리 블록들은 변하지 않은 채 유지될 수 있다.

[0058] HOA 분해 처리의 이런 변경이 소위 "듀얼 계층" 또는 "2 계층" 모드에서 HOA 압축 작업을 하는 초기 동작으로 보일 수 있다는 점에 유의하는 것이 중요하다. 이 모드는 저품질 베이스 계층 및 향상 계층으로 분할될 수 있는 비트 스트림을 제공한다. 이 모드를 이용하거나 이용하지 않는 것은 전체 비트 스트림의 액세스 단위에서 단일 비트에 의해 신호화될 수 있다.

[0059] 아래에 더 자세히 설명되는 바와 같이, 베이스 계층 및 향상 계층에 비트 스트림을 제공하도록 멀티플렉싱하는 비트 스트림의 결과적인 가능한 수정은 도 3 및 4에 예시된다.

[0060] 베이스 계층 비트 스트림  $\mathbf{B}_{BASE}(k-2)$  은 지각 인코딩된 신호  $\tilde{\mathbf{z}}_i(k-2)$ ,  $i = 1, \dots, O_{MIN}$ , 및 대응하는 코딩된 이득 제어 사이드 정보를 포함할 뿐이고, 이 사이드 정보는 지수들  $e_i(k-2)$  및 예외 플래그  $\beta_i(k-2)$ ,  $i = 1, \dots, O_{MIN}$  로 구성된다. 나머지 지각 인코딩된 신호들  $\tilde{\mathbf{z}}_i(k-2)$ ,  $i = O_{MIN} + 1, \dots, O$ , 및 인코딩된 나머지 사이드 정보는 향상 계층 비트 스트림에 포함된다. 베이스 계층 및 향상 계층 비트 스트림들  $\mathbf{B}_{BASE}(k-2)$  및  $\mathbf{B}_{ENH}(k-2)$  은 그 후 이전의 전체 비트 스트림  $\mathbf{B}(k-2)$  대신에 함께 송신된다.

[0061] 도 3 및 도 4에는, HOA 계수 시퀀스들의 입력 시간 프레임들 ( $\mathbf{C}^{(k)}$ )을 가진 입력 HOA 표현인 HOA 신호를 압축하기 위한 장치가 도시된다. 상기 장치는 도 3에 도시된, 입력 시간 프레임의 공간 HOA 인코딩 및 후속 지각 인코딩을 위한 공간 HOA 인코딩 및 지각 인코딩 부와, 도 4에 도시된, 소스 인코딩을 위한 소스 코더 부를 포함한다. 공간 HOA 인코딩 및 지각 인코딩 부는 방향 및 벡터 추정 블록(301), HOA 분해 블록(303), 주변 컴포넌트 수정 블록(304), 채널 할당 블록(305), 및 복수의 이득 제어 블록(306)을 포함한다.

[0062] 방향 및 벡터 추정 블록(301)은 HOA 신호의 방향 및 벡터 추정 처리를 수행하도록 적응되고, 여기서 방향 신호



들을 위한 제1 투플 세트들  $\mathcal{M}_{\text{DIR}}(k)$  및 벡터 기반 신호들을 위한 제2 투플 세트들  $\mathcal{M}_{\text{VEC}}(k)$ 을 포함하는 데이터가 획득되고, 제1 투플 세트들  $\mathcal{M}_{\text{DIR}}(k)$  각각은 방향 신호와 각각의 양자화된 방향의 인덱스를 포함하고, 제2 투플 세트들  $\mathcal{M}_{\text{VEC}}(k)$  각각은 벡터 기반 신호 및 상기 신호들의 방향 분포를 정의하는 벡터의 인덱스를 포함한다.

[0063] HOA 분해 블록(303)은 HOA 계수 시퀀스들의 각각의 입력 시간 프레임을 복수의 우세 사운드 신호  $\mathbf{x}_{\text{PS}}(k-1)$ 의 프레임 및 주변 HOA 컴포넌트  $\tilde{\mathbf{c}}_{\text{AMB}}(k-1)$ 의 프레임으로 분해하도록 적응되고, 우세 사운드 신호  $\mathbf{x}_{\text{PS}}(k-1)$ 는 상기 방향 사운드 신호들 및 상기 벡터 기반 사운드 신호들을 포함하고, 주변 HOA 컴포넌트  $\tilde{\mathbf{c}}_{\text{AMB}}(k-1)$ 은 입력 HOA 표현과 우세 사운드 신호들의 HOA 표현 사이의 잔차를 표현하는 HOA 계수 시퀀스들을 포함하고, 분해하는 것은 또한 예측 파라미터들  $\xi(k-1)$  및 타깃 할당 벡터  $\mathbf{v}_{\text{A,T}}(k-1)$ 를 제공한다. 예측 파라미터들  $\xi(k-1)$ 은 우세 사운드 HOA 컴포넌트들을 강화하기 위해 우세 사운드 신호들  $\mathbf{x}_{\text{PS}}(k-1)$  내의 방향 신호들로부터 HOA 신호 표현의 일부를 어떻게 예측하는지를 기술하고, 타깃 할당 벡터  $\mathbf{v}_{\text{A,T}}(k-1)$ 는 우세 사운드 신호들을 주어진 수  $I$ 의 채널들에 어떻게 할당하는지에 대한 정보를 포함한다. 주변 컴포넌트 수정 블록(304)은 타깃 할당 벡터  $\mathbf{v}_{\text{A,T}}(k-1)$ 에 의해 제공되는 정보에 따라 주변 HOA 컴포넌트  $\mathbf{c}_{\text{AMB}}(k-1)$ 를 수정하도록 적응되고, 주변 HOA 컴포넌트  $\mathbf{c}_{\text{AMB}}(k-1)$ 의 어느 계수 시퀀스들이 주어진 수  $I$ 의 채널들에서 송신될 것인지는, 우세 사운드 신호들이 얼마나 많은 채널을 차지하는지에 따라 결정되고, 수정된 주변 HOA 컴포넌트  $\mathbf{c}_{\text{M,A}}(k-2)$  및 시간적으로 예측된 수정된 주변 HOA 컴포넌트  $\mathbf{c}_{\text{P,M,A}}(k-1)$ 가 획득되고, 최종 할당 벡터  $\mathbf{v}_{\text{A}}(k-2)$ 는 타깃 할당 벡터  $\mathbf{v}_{\text{A,T}}(k-1)$ 내의 정보로부터 획득된다.

[0064] 채널 할당 블록(305)은 분해로부터 획득된 우세 사운드 신호들  $\mathbf{x}_{\text{PS}}(k-1)$ 과, 수정된 주변 HOA 컴포넌트  $\mathbf{c}_{\text{M,A}}(k-2)$  및 시간적으로 예측된 수정된 주변 HOA 컴포넌트  $\mathbf{c}_{\text{P,M,A}}(k-1)$ 의 결정된 계수 시퀀스들을, 최종 할당 벡터  $\mathbf{v}_{\text{A}}(k-2)$ 에 의해 제공된 정보를 이용하여 주어진 수  $I$ 의 채널들에 할당하도록 적응되고, 전송 신호들  $\mathbf{y}_i(k-2)$ ,  $i = 1, \dots, I$  및 예측된 전송 신호들  $\mathbf{y}_{\text{P},i}(k-2)$ ,  $i = 1, \dots, I$ 이 획득된다.

[0065] 복수의 이득 제어 블록(306)은 전송 신호들  $\mathbf{y}_i(k-2)$  및 예측된 전송 신호들  $\mathbf{y}_{\text{P},i}(k-2)$ 에 대한 이득 제어 (805)를 수행하도록 적응되고, 이득 수정된 전송 신호들  $\mathbf{z}_i(k-2)$ , 지수들  $e_i(k-2)$  및 예외 플래그들  $\beta_i(k-2)$ 이 획득된다.

[0066] 도 4는 본 발명의 일 실시예에 따른 HOA 압축기의 소스 코더 부의 아키텍처의 구조를 도시한다. 도 4에 도시된 바와 같은 소스 코더 부는 지각 코더(310), 2개의 코더(320, 330)(즉, 베이스 계층 사이드 정보 소스 코더(320) 및 향상 계층 사이드 정보 인코더(330))를 가진 사이드 정보 소스 코더 블록, 및 2개의 멀티플렉서(340, 350)(즉, 베이스 계층 비트스트림 멀티플렉서(340) 및 향상 계층 비트스트림 멀티플렉서(350))를 포함한다. 사이드 정보 소스 코더들은 단일 사이드 정보 소스 코더 블록 내에 있을 수 있다.

[0067] 지각 코더(310)는 상기 이득 수정된 전송 신호들  $\mathbf{z}_i(k-2)$ 를 지각 코딩(806)하도록 적응되며, 지각 인코딩된 전송 신호들  $\tilde{\mathbf{z}}_i(k-2)$ ,  $i = 1, \dots, I$ 이 획득된다.

[0068] 사이드 정보 소스 코더들(320, 330)은, 상기 지수들  $e_i(k-2)$  및 예외 플래그들  $\beta_i(k-2)$ , 상기 제1 투플 세트들  $\mathcal{M}_{\text{DIR}}(k)$  및 제2 투플 세트들  $\mathcal{M}_{\text{VEC}}(k)$ , 상기 예측 파라미터  $\xi(k-1)$  및 상기 최종 할당 벡터  $\mathbf{v}_{\text{A}}(k-2)$ 를 포함하는 사이드 정보를 인코딩하도록 적응되고, 인코딩된 사이드 정보  $\tilde{\mathbf{r}}(k-2)$ 가 획득된다. 멀티플렉서들 (340, 350)은 지각 인코딩된 전송 신호들  $\tilde{\mathbf{z}}_i(k-2)$  및 인코딩된 사이드 정보  $\tilde{\mathbf{r}}(k-2)$ 를 멀티플렉싱된 데이터 스트림  $\tilde{\mathbf{B}}(k-2)$ 으로 멀티플렉싱하도록 적응되고, 분해에서 획득된 주변 HOA 컴포넌트  $\tilde{\mathbf{c}}_{\text{AMB}}(k-1)$ 는

$O_{MIN}$  최하위 위치들(즉, 최하위 인덱스들을 갖는 위치들) 내의 입력 HOA 표현  $\mathbf{c}_n(k-1)$  제1 HOA 계수 시퀀스들, 및 나머지 상위 위치들 내의 제2 HOA 계수 시퀀스들  $\mathbf{c}_{AMB,n}(k-1)$  을 포함한다. 수학식 4-6과 관련하여 아래에 설명되는 바와 같이, 제2 HOA 계수 시퀀스들은 입력 HOA 표현과 우세 사운드 신호들의 HOA 표현 사이의 잔차의 HOA 표현의 부분이다. 더욱이, 제1  $O_{MIN}$  지수들  $e_i(k-2), i=1, \dots, O_{MIN}$  및 지수 플래그들  $\beta_i(k-2), i=1, \dots, O_{MIN}$  은 베이스 계층 사이드 정보 소스 코더(320)에서 인코딩되고, 여기서 인코딩된 베이스 계층 사이드 정보  $\tilde{\mathbf{r}}_{BASE}(k-2)$  가 획득되고,  $O_{MIN} = (N_{MIN} + 1)^2$  및  $O = (N+1)^2$  이 되고,  $N_{MIN} \leq N$  및  $O_{MIN} \leq b$  이 되고,  $N_{MIN}$  이 사전 정의된 정수 값이다. 제1  $O_{MIN}$  지각 인코딩된 전송 신호들  $\tilde{\mathbf{z}}_i(k-2), i=1, \dots, O_{MIN}$  및 인코딩된 베이스 계층 사이드 정보  $\tilde{\mathbf{r}}_{BASE}(k-2)$  는 베이스 계층 비트스트림 멀티플렉서(340)(이것은 상기 멀티플렉서들 중 하나임)에서 멀티플렉싱되고, 베이스 계층 비트스트림  $\tilde{\mathbf{B}}_{BASE}(k-2)$  이 획득된다. 베이스 계층 사이드 정보 소스 코더(320)는 사이드 정보 소스 코더들 중 하나이거나, 또는 사이드 정보 소스 코더 블록 내에 있다. 나머지  $I - O_{MIN}$  지수들  $e_i(k-2), i=O_{MIN}+1, \dots, I$  및 예외 플래그들  $\beta_i(k-2), i=O_{MIN}+1, \dots, I$ , 상기 제1 투플 세트들  $\mathcal{M}_{DIR}(k-1)$  및 제2 투플 세트들  $\mathcal{M}_{VEC}(k-1)$ , 상기 예측 파라미터들  $\xi(k-1)$  및 상기 최종 할당 벡터  $\mathbf{v}_A(k-2)$  는 향상 계층 사이드 정보 인코더(330)에서 인코딩되고, 여기서 인코딩된 향상 계층 사이드 정보  $\tilde{\mathbf{r}}_{ENH}(k-2)$  가 획득된다. 향상 계층 사이드 정보 소스 코더(330)는 사이드 정보 소스 코더들 중 하나이거나, 또는 사이드 정보 소스 코더 블록 내에 있다.

[0069] 나머지  $I - O_{MIN}$  지각 인코딩된 전송 신호들  $\tilde{\mathbf{z}}_i(k-2), i=O_{MIN}+1, \dots, I$  및 인코딩된 향상 계층 사이드 정보  $\tilde{\mathbf{r}}_{ENH}(k-2)$  는 향상 계층 비트스트림 멀티플렉서(350)(이것은 또한 상기 멀티플렉서들 중 하나임)에서 멀티플렉싱되고, 여기서 향상 계층 비트스트림  $\tilde{\mathbf{B}}_{ENH}(k-2)$  이 획득된다. 또한, 모드 지시  $LMF_E$  는 멀티플렉서 또는 지시 삽입 블록에 추가된다. 모드 지시  $LMF_E$  는 계층화 모드의 이용을 신호화하고, 이것은 압축된 신호의 정확한 압축 해제를 위해 사용된다.

[0070] 일 실시예에서, 인코딩을 위한 장치는 모드를 선택하도록 적응된 모드 선택기를 더 포함하며, 모드는 모드 지시  $LMF_E$  에 의해 지시되며, 계층화 모드 및 비계층화 모드 중 하나이다. 비계층화 모드에서, 주변 HOA 컴포넌트  $\tilde{\mathbf{c}}_{AMB}(k-1)$  는 입력 HOA 표현과 우세 사운드 신호들의 HOA 표현 사이의 잔차를 표현하는 HOA 계수 시퀀스들만을 포함한다(즉, 입력 HOA 표현의 계수 시퀀스들이 없음).

[0071] HOA 압축 해제에 제안된 보정은 이하에서 설명된다.

[0072] 계층화 모드에서, HOA 압축시 주변 HOA 컴포넌트  $\tilde{\mathbf{c}}_{AMB}(k-1)$  의 수정은 HOA 구성을 적절히 수정함으로써 HOA 압축 해제에서 고려된다.

[0073] HOA 압축 해제기에서는, 베이스 계층 및 향상 계층 비트 스트림들의 디멀티플렉싱 및 디코딩이 도 5에 따라 수행된다. 베이스 계층 비트 스트림  $\tilde{\mathbf{B}}_{BASE}(k)$  은 베이스 계층 사이드 정보 및 지각 인코딩된 신호들의 코딩된 표현으로 디멀티플렉싱된다. 연속해서, 베이스 계층 사이드 정보 및 지각 인코딩된 신호들의 코딩된 표현은 한편으로는 지수들  $e_i(k)$  및 예외 플래그들을 제공하고, 다른 한편으로는 지각 디코딩된 신호들을 제공하도록 디코딩된다. 유사하게, 향상 계층 비트 스트림은 지각 디코딩된 신호들 및 나머지 사이드 정보를 제공하도록 디멀티플렉싱되고 디코딩된다(도 5 참조). 이런 계층화 모드에서, 공간 HOA 디코딩 부는 또한, 공간 HOA 인코딩시 주변 HOA 컴포넌트  $\tilde{\mathbf{c}}_{AMB}(k-1)$  의 수정을 고려하도록 수정되어야 한다. 이 수정은 HOA 구성에서 달성된다.

[0074] 특히, 재구성된 HOA 표현



#### 수학식 4

$$\hat{\mathbf{c}}(k-1) = \hat{\mathbf{c}}_{\text{PS}}(k-1) + \hat{\mathbf{c}}_{\text{AMB}}(k-1)$$

은 그 수정된 버전

#### 수학식 5

$$\tilde{\mathbf{c}}(k-1) = \begin{bmatrix} \tilde{\mathbf{c}}_1(k-1) \\ \tilde{\mathbf{c}}_2(k-1) \\ \vdots \\ \tilde{\mathbf{c}}_O(k-1) \end{bmatrix}$$

에 의해 대체되고,

그 요소들은

#### 수학식 6

$$\tilde{\mathbf{c}}_n(k-1) = \begin{cases} \hat{\mathbf{c}}_{\text{AMB},n}(k-1) & \text{for } 1 \leq n \leq O_{\text{MIN}} \\ \hat{\mathbf{c}}_n(k-1) & \text{for } O_{\text{MIN}} + 1 \leq n \leq O \end{cases}$$

에 의해 주어진다.

이것은 우세 사운드 HOA 컴포넌트가 제1  $O_{\text{MIN}}$  계수 시퀀스들을 위해 주변 HOA 컴포넌트에 추가되지 않는다는 것을 의미하며, 그 이유는 이것이 이미 포함되었기 때문이다. HOA 공간 디코더의 다른 모든 처리 블록들은 변화없이 유지된다.

이하에서는, 저품질 베이스 계층 비트 스트림  $\tilde{\mathbf{B}}_{\text{BASE}}(k)$ 의 순수한 존재시 HOA 압축 해제가 간단히 고려된다.

비트 스트림은 재구성된 신호들  $\hat{\mathbf{z}}_i(k)$ , 및 지수들  $e_i(k)$  및 예외 플래그들  $\beta_i(k)$ ,  $i = 1, \dots, O_{\text{MIN}}$ 로 구성되는 대응하는 이득 제어 사이드 정보를 제공하도록 먼저 디멀티플렉싱되고 디코딩된다. 항상 계층의 부재시, 지각 코딩된 신호들  $\hat{\mathbf{z}}_i(k-2)$ ,  $i = O_{\text{MIN}} + 1, \dots, O$ 은 이용 가능하지 않다. 이런 상황에 대처하는 가능한 방식은  $\hat{\mathbf{z}}_i(k)$ ,  $i = O_{\text{MIN}} + 1, \dots, O$ 을 0으로 설정하는 것이고, 이것은 재구성된 우세 사운드 컴포넌트  $\mathbf{c}_{\text{PS}}(k-1)$ 과 자동으로 0이 되게 한다.

다음 단계에서, 공간 HOA 디코더에서, 제1  $O_{\text{MIN}}$  역 이득 제어 처리 블록들은 이득 정정된 신호 프레임들  $\hat{\mathbf{y}}_i(k)$ ,  $i = 1, \dots, O_{\text{MIN}}$ 을 제공하고, 이것은 채널 재할당에 의해 주변 HOA 컴포넌트의 중간 표현의 프레임  $\mathbf{c}_{\text{LAMB}}(k)$ 을 구성하는데 사용된다.  $k$ 번째 프레임에서 활성인, 주변 HOA 컴포넌트의 계수 시퀀스들의 인덱스들의 세트  $\mathcal{I}_{\text{AMB,ACT}}(k)$ 가 인덱스들  $1, 2, \dots, O_{\text{MIN}}$ 만을 포함함에 유의한다. 주변 합성에서, 제1  $O_{\text{MIN}}$  계수 시퀀스들의 공간 변환은 주변 HOA 컴포넌트 프레임  $\mathbf{c}_{\text{AMB}}(k-1)$ 을 제공하기 위해 반전된다. 결국, 재구성된 HOA 표현은 수학식 6에 따라 계산된다.

도 5 및 도 6은 본 발명의 일 실시예에 따른 HOA 압축 해제기의 아키텍처의 구조를 도시한다. 장치는 도 5에 도시된 지각 디코딩 및 소스 디코딩 부, 도 6에 도시된 공간 HOA 디코딩 부, 및 압축된 HOA 신호가 압축된 베이스 계층 비트스트림  $\tilde{\mathbf{B}}_{\text{BASE}}(k)$  및 압축된 항상 계층 비트스트림을 포함하는 것을 지시하는 계층화 모드 지시  $\text{LMF}_D$ 를 검출하도록 적응되는 모드 검출기를 포함한다.

- [0087] 도 5는 본 발명의 일 실시예에 따른 HOA 압축 해제기의 지각 디코딩 및 소스 디코딩 부의 아키텍처의 구조를 도시한다. 지각 디코딩 및 소스 디코딩 부는 제1 디멀티플렉서(510), 제2 디멀티플렉서(520), 베이스 계층 지각 디코더(540) 및 향상 계층 지각 디코더(550), 베이스 계층 사이드 정보 소스 디코더(530) 및 향상 계층 사이드 정보 소스 디코더(560)를 포함한다.
- [0088] 제1 디멀티플렉서(510)는 압축된 베이스 계층 비트스트림  $\tilde{\mathbf{B}}_{BASE}(k)$ 을 디멀티플렉싱하도록 적응되고, 제1 지각 인코딩된 전송 신호들  $\tilde{\mathbf{z}}_i(k), i = 1, \dots, O_{MIN}$  및 제1 인코딩된 사이드 정보  $\tilde{\mathbf{r}}_{BASE}(k)$ 가 획득된다. 제2 디멀티플렉서(520)는 압축된 향상 계층 비트스트림  $\tilde{\mathbf{B}}_{ENH}(k)$ 을 디멀티플렉싱하도록 적응되며, 제2 지각 인코딩된 전송 신호들  $\tilde{\mathbf{z}}_i(k), i = O_{MIN} + 1, \dots, I$  및 제2 인코딩된 사이드 정보  $\tilde{\mathbf{r}}_{ENH}(k)$ 가 획득된다.
- [0089] 베이스 계층 지각 디코더(540) 및 향상 계층 지각 디코더(550)는 지각 인코딩된 전송 신호들  $\tilde{\mathbf{z}}_i(k), i = 1, \dots, I$ 을 지각 디코딩하도록(904) 적응되며, 지각 디코딩된 전송 신호들  $\hat{\mathbf{z}}_i(k)$ 이 획득되고, 베이스 계층 지각 디코더(540)에서, 베이스 계층의 상기 제1 지각 인코딩된 전송 신호들  $\tilde{\mathbf{z}}_i(k), i = 1, \dots, O_{MIN}$ 이 디코딩되고, 제1 지각 디코딩된 전송 신호들  $\hat{\mathbf{z}}_i(k), i = 1, \dots, O_{MIN}$ 이 획득된다. 향상 계층 지각 디코더(550)에서, 향상 계층의 상기 제2 지각 인코딩된 전송 신호들  $\tilde{\mathbf{z}}_i(k), i = O_{MIN} + 1, \dots, I$ 이 디코딩되고, 상기 제2 지각 디코딩된 전송 신호들  $\hat{\mathbf{z}}_i(k), i = O_{MIN} + 1, \dots, I$ 이 획득된다.
- [0090] 베이스 계층 사이드 정보 소스 디코더(530)는 제1 인코딩된 사이드 정보  $\tilde{\mathbf{r}}_{BASE}(k)$ 를 디코딩하도록(905) 적응되고, 제1 지수들  $\mathbf{e}_i(k), i = 1, \dots, O_{MIN}$  및 제1 예외 플래그들  $\beta_i(k), i = 1, \dots, O_{MIN}$ 이 획득된다.
- [0091] 향상 계층 사이드 정보 소스 디코더(560)는 제2 인코딩된 사이드 정보  $\tilde{\mathbf{r}}_{ENH}(k)$ 를 디코딩하도록(906) 적응되고, 제2 지수들  $\mathbf{e}_i(k), i = O_{MIN} + 1, \dots, I$  및 제2 예외 플래그들  $\beta_i(k), i = O_{MIN} + 1, \dots, I$ 이 획득되고, 추가 데이터가 획득된다. 추가 데이터는 방향 신호들을 위한 제1 토폴 세트  $\mathcal{M}_{DIR}(k+1)$  및 벡터 기반 신호들을 위한 제2 토폴 세트  $\mathcal{M}_{VEC}(k+1)$ 를 포함한다. 제1 토폴 세트  $\mathcal{M}_{DIR}(k+1)$ 의 각각의 토폴은 방향 신호 및 각각의 양자화된 방향의 인덱스를 포함하고, 제2 토폴 세트  $\mathcal{M}_{VEC}(k+1)$ 의 각각의 토폴은 벡터 기반 신호 및 벡터 기반 신호의 방향 분포를 정의하는 벡터의 인덱스를 포함한다. 더욱이, 예측 파라미터들  $\xi(k+1)$  및 주변 할당 벡터  $\mathbf{v}_{AMB,ASSIGN}(k)$ 가 획득되고, 주변 할당 벡터  $\mathbf{v}_{AMB,ASSIGN}(k)$ 는 각각의 송신 채널이 주변 HOA 컴포넌트의 계수 시퀀스들을 포함하는지와 어느 계수 시퀀스들을 포함하는지를 각각의 송신 채널에 대해 나타내는 컴포넌트들을 포함한다.
- [0092] 도 6은 본 발명의 일 실시예에 따른 HOA 압축 해제기의 공간 HOA 디코딩 부의 아키텍처의 구조를 도시한다. 공간 HOA 디코딩 부는 복수의 역 이득 제어 유닛(604), 채널 재할당 블록(605), 우세 사운드 합성 블록(606), 주변 합성 블록(607), 및 HOA 구성 블록(608)을 포함한다.
- [0093] 복수의 역 이득 제어 유닛(604)은 역 이득 제어를 수행하도록 적응되고, 여기서 상기 제1 지각 디코딩된 전송 신호들  $\hat{\mathbf{z}}_i(k), i = 1, \dots, O_{MIN}$ 은 제1 지수들  $\mathbf{e}_i(k), i = 1, \dots, O_{MIN}$  및 제1 예외 플래그들  $\beta_i(k), i = 1, \dots, O_{MIN}$ 에 따라 제1 이득 정정된 신호 프레임들  $\hat{\mathbf{y}}_i(k), i = 1, \dots, O_{MIN}$ 로 변환되고, 제2 지각 디코딩된 전송 신호들  $\hat{\mathbf{z}}_i(k), i = O_{MIN} + 1, \dots, I$ 은 제2 지수들  $\mathbf{e}_i(k), i = O_{MIN} + 1, \dots, I$  및 제2 예외 플래그들  $\beta_i(k), i = O_{MIN} + 1, \dots, I$ 에 따라 제2 이득 정정된 신호 프레임들  $\hat{\mathbf{y}}_i(k), i = O_{MIN} + 1, \dots, I$ 로 변환된다.
- [0094] 채널 재할당 블록(605)은 제1 및 제2 이득 정정된 신호 프레임들  $\hat{\mathbf{y}}_i(k), i = 1, \dots, I$ 을  $I$  채널들에 재분배하도록(911) 적응되고, 여기서 우세 사운드 신호들  $\hat{\mathbf{x}}_{ps}(k)$ 의 프레임들이 재구성되고, 우세 사운드 신호들은 방향 신호들 및 벡터 기반 신호들을 포함하고, 수정된 주변 HOA 컴포넌트  $\tilde{\mathbf{c}}_{I,AMB}(k)$ 가 획득되고, 상기 주변 할당 벡터  $\mathbf{v}_{AMB,ASSIGN}(k)$ 와, 상기 제1 및 제2 토폴 세트들  $\mathcal{M}_{DIR}(k+1)$  및  $\mathcal{M}_{VEC}(k+1)$  내의 정보에 따라 할당이 이루어진다.

- [0095] 더욱이, 채널 재할당 블록(605)은 k번째 프레임에서 활성화된 수정된 주변 HOA 컴포넌트의 계수 시퀀스들의 인덱스들  $\mathcal{I}_{AMB,ACT}(k)$  의 제1 세트, 및 인에이블되고, 디스에이블되고, (k-1)번째 프레임에서 활성화로 유지될 필요가 있는 수정된 주변 HOA 컴포넌트의 계수 시퀀스들의 인덱스들  $\mathcal{I}_E(k-1), \mathcal{I}_D(k-1), \mathcal{I}_U(k-1)$  의 제2 세트를 생성하도록 적응된다.
- [0096] 우세 사운드 합성 블록(606)은 상기 우세 사운드 신호  $\hat{x}_{ps}(k)$ 로부터 우세 HOA 사운드 컴포넌트들  $\hat{\mathbf{c}}_{ps}(k-1)$ 의 HOA 표현을 합성하도록(912) 적응되고, 상기 제1 및 제2 투플 세트들  $\mathcal{M}_{DIR}(k+1)$  및  $\mathcal{M}_{VEC}(k+1)$  예측 파라미터들  $\xi(k+1)$  및 인덱스들  $\mathcal{I}_E(k-1), \mathcal{I}_D(k-1), \mathcal{I}_U(k-1)$  의 제2 세트가 사용된다.
- [0097] 주변 합성 블록(607)은 수정된 주변 HOA 컴포넌트  $\hat{\mathbf{c}}_{l,AMB}(k)$ 로부터 주변 HOA 컴포넌트  $\hat{\mathbf{c}}_{AMB}(k-1)$ 를 합성하도록(913) 적응되고, 제1  $O_{MIN}$  채널들을 위한 역 공간 변환이 이루어지고, 인덱스들  $\mathcal{I}_{AMB,ACT}(k)$ 의 제1 세트가 사용되고, 인덱스들의 제1 세트는 k번째 프레임에서 활성화된 주변 HOA 컴포넌트의 계수 시퀀스들의 인덱스들이다.
- [0098] 계층화 모드 지시  $LMF_D$ 가 적어도 2개의 계층을 가진 계층화 모드를 지시하는 경우, 주변 HOA 컴포넌트는 그  $O_{MIN}$  최하위 위치들(즉, 최하위 인덱스들을 가진 위치들) 내의 압축 해제된 HOA 신호  $\hat{\mathbf{c}}(k-1)$ 의 HOA 계수 시퀀스들, 및 나머지 상위 위치들 내의 잔차의 HOA 표현의 부분인 계수 시퀀스들을 포함한다. 이런 잔차는 압축 해제된 HOA 신호  $\hat{\mathbf{c}}(k-1)$ 와 우세 HOA 사운드 컴포넌트들  $\hat{\mathbf{c}}_{ps}(k-1)$ 의 HOA 표현 사이의 잔차이다.
- [0099] 한편, 계층화 모드 지시  $LMF_D$ 가 단일-계층 모드를 지시하는 경우, 압축 해제된 HOA 신호  $\hat{\mathbf{c}}(k-1)$ 의 HOA 계수 시퀀스들은 구성되지 않고, 주변 HOA 컴포넌트가 압축 해제된 HOA 신호  $\hat{\mathbf{c}}(k-1)$ 와 우세 사운드 컴포넌트  $\hat{\mathbf{c}}_{ps}(k-1)$ 의 HOA 표현 사이의 잔차가 된다.
- [0100] HOA 구성 블록(608)은 우세 사운드 컴포넌트들의 HOA 표현을 주변 HOA 컴포넌트  $\hat{\mathbf{c}}_{ps}(k-1)\hat{\mathbf{c}}_{AMB}(k-1)$ 에 가산하도록 적응되고, 여기서 우세 사운드 신호들의 HOA 표현의 계수들 및 주변 HOA 컴포넌트의 대응하는 계수들이 가산되고, 압축 해제된 HOA 신호  $\hat{\mathbf{c}}'(k-1)$ 가 획득되며,
- [0101] 계층화 모드 지시  $LMF_D$ 가 적어도 2개의 계층을 가진 계층화 모드를 지시하는 경우, 가장 높은  $1-O_{MIN}$  계수 채널들만이 우세 HOA 사운드 컴포넌트들  $\hat{\mathbf{c}}_{ps}(k-1)$  및 주변 HOA 컴포넌트  $\hat{\mathbf{c}}_{AMB}(k-1)$ 의 가산에 의해 획득되고, 압축 해제된 HOA 신호  $\hat{\mathbf{c}}'(k-1)$ 의 가장 낮은  $O_{MIN}$  계수 채널들이 주변 HOA 컴포넌트  $\hat{\mathbf{c}}_{AMB}(k-1)$ 로부터 복사된다. 한편, 계층화 모드 지시  $LMF_D$ 가 단일-계층 모드를 지시하는 경우, 압축 해제된 HOA 신호  $\hat{\mathbf{c}}'(k-1)$ 의 모든 계수 채널들은 우세 HOA 사운드 컴포넌트  $\hat{\mathbf{c}}_{ps}(k-1)$  및 주변 HOA 컴포넌트  $\hat{\mathbf{c}}_{AMB}(k-1)$ 의 가산에 의해 획득된다.
- [0102] 도 7은 주변 HOA 신호들로부터 수정된 주변 HOA 신호들로의 프레임들의 변환을 도시한다.
- [0103] 도 8은 HOA 신호를 압축하는 방법의 흐름도를 보여준다.
- [0104] 고차 앰비소닉스(HOA)(Higher Order Ambisonics) 신호(이는 HOA 계수 시퀀스들의 입력 시간 프레임  $\mathbf{C}(k)$ 을 가진 오더 N의 입력 HOA 표현임)를 압축하는 방법(800)은 입력 시간 프레임의 공간 HOA 인코딩과, 후속 지각 인코딩 및 소스 인코딩을 포함한다.
- [0105] 공간 HOA 인코딩은,
- [0106] 방향 및 벡터 추정 블록(301)에서 HOA 신호의 방향 및 벡터 추정 처리(801)를 수행하는 단계로서, 방향 신호들을 위한 제1 투플 세트들  $\mathcal{M}_{DIR}(k)$  및 벡터 기반 신호들을 위한 제2 투플 세트들  $\mathcal{M}_{VEC}(k)$ 을 포함하는 데이터가 획득되고, 제1 투플 세트들  $\mathcal{M}_{DIR}(k)$  각각은 방향 신호와 각각의 양자화된 방향의 인덱스를 포함하고, 제2 투플

세트들  $\mathcal{M}_{\text{VEC}}(k)$  각각은 벡터 기반 신호 및 상기 신호들의 방향 분포를 정의하는 벡터의 인덱스를 포함하는 단계,

[0107] HOA 분해 블록(303)에서, HOA 계수 시퀀스들의 각각의 입력 시간 프레임을 복수의 우세 사운드 신호  $\mathbf{x}_{\text{ps}}(k-1)$ 의 프레임 및 주변 HOA 컴포넌트  $\tilde{\mathbf{c}}_{\text{AMB}}(k-1)$ 의 프레임으로 분해하는 단계(802)로서, 우세 사운드 신호들  $\mathbf{x}_{\text{ps}}(k-1)$ 이 상기 방향 사운드 신호들 및 상기 벡터 기반 사운드 신호들을 포함하고, 주변 HOA 컴포넌트  $\tilde{\mathbf{c}}_{\text{AMB}}(k-1)$ 가 입력 HOA 표현과 우세 사운드 신호들의 HOA 표현 사이의 잔차를 나타내는 HOA 계수 시퀀스들을 포함하는 단계 - 분해하는 단계(702)는 또한 예측 파라미터들  $\xi(k-1)$  및 타깃 할당 벡터  $\mathbf{v}_{\text{A,T}}(k-1)$ 를 제공하고, 예측 파라미터들  $\xi(k-1)$ 은 우세 사운드 HOA 컴포넌트들을 강화하기 위해 우세 사운드 신호들  $\mathbf{x}_{\text{ps}}(k-1)$ 내의 방향 신호들로부터 HOA 신호 표현의 일부를 어떻게 예측하는지를 기술하고, 타깃 할당 벡터  $\mathbf{v}_{\text{A,T}}(k-1)$ 은 우세 사운드 신호들을 주어진 수  $I$ 의 채널들에 어떻게 할당하는지에 대한 정보를 포함함 -,

[0108] 주변 컴포넌트 수정 블록(304)에서, 타깃 할당 벡터  $\mathbf{v}_{\text{A,T}}(k-1)$ 에 의해 제공되는 정보에 따라 주변 HOA 컴포넌트  $\mathbf{c}_{\text{AMB}}(k-1)$ 를 수정하는 단계(803)로서, 주변 HOA 컴포넌트  $\mathbf{c}_{\text{AMB}}(k-1)$ 의 어느 계수 시퀀스들이 주어진 수  $I$ 의 채널들에서 송신될 것인지는, 우세 사운드 신호들이 얼마나 많은 채널을 차지하는지에 따라 결정되고, 수정된 주변 HOA 컴포넌트  $\mathbf{c}_{\text{M,A}}(k-2)$  및 시간적으로 예측된 수정된 주변 HOA 컴포넌트  $\mathbf{c}_{\text{P,M,A}}(k-1)$ 가 획득되고, 최종 할당 벡터  $\mathbf{v}_{\text{A}}(k-2)$ 가 타깃 할당 벡터  $\mathbf{v}_{\text{A,T}}(k-1)$  내의 정보로부터 획득되는 단계,

[0109] 채널 할당 블록(105)에서, 분해 단계로부터 획득된 우세 사운드 신호들  $\mathbf{x}_{\text{ps}}(k-1)$ , 및 수정된 주변 HOA 컴포넌트  $\mathbf{c}_{\text{M,A}}(k-2)$  및 시간적으로 예측된 수정된 주변 HOA 컴포넌트  $\mathbf{c}_{\text{P,M,A}}(k-1)$ 의 결정된 계수 시퀀스들을, 최종 할당 벡터  $\mathbf{v}_{\text{A}}(k-2)$ 에 의해 제공된 정보를 이용하여 주어진 수  $I$ 의 채널들에 할당하는 단계(804)로서, 전송 신호들  $\mathbf{y}_i(k-2)$ ,  $i=1, \dots, I$  및 예측된 전송 신호들  $\mathbf{y}_{\text{P},i}(k-2)$ ,  $i=1, \dots, I$ 가 획득되는 단계, 및

[0110] 복수의 이득 제어 블록(306)에서 전송 신호들  $\mathbf{y}_i(k-2)$  및 예측된 전송 신호들  $\mathbf{y}_{\text{P},i}(k-2)$ 에 대한 이득 제어(805)를 수행하는 단계로서, 이득 수정된 전송 신호들  $\mathbf{z}_i(k-2)$ , 지수들  $\mathbf{e}_i(k-2)$  및 예외 플래그들  $\beta_i(k-2)$ 가 획득되는 단계

[0111] 를 포함한다.

[0112] 각각 인코딩 및 소스 인코딩은,

[0113] 각각 코더(310)에서, 상기 이득 수정된 전송 신호들  $\mathbf{z}_i(k-2)$ 을 각각 코딩하는 단계(806)로서, 각각 인코딩된 전송 신호들  $\tilde{\mathbf{z}}_i(k-2)$ ,  $i=1, \dots, I$ 이 획득되는 단계,

[0114] 하나 이상의 사이드 정보 소스 코더들(320, 330)에서, 상기 지수들  $\mathbf{e}_i(k-2)$  및 예외 플래그들  $\beta_i(k-2)$ , 상기 제1 토폴 세트들  $\mathcal{M}_{\text{DIR}}(k)$  및 제2 토폴 세트들  $\mathcal{M}_{\text{VEC}}(k)$ , 상기 예측 파라미터들  $\xi(k-1)$  및 상기 최종 할당 벡터  $\mathbf{v}_{\text{A}}(k-2)$ 를 포함하는 사이드 정보를 인코딩하는 단계(807)로서, 인코딩된 사이드 정보  $\tilde{\mathbf{r}}(k-2)$ 가 획득되는 단계, 및

[0115] 각각 인코딩된 전송 신호들  $\tilde{\mathbf{z}}_i(k-2)$  및 인코딩된 사이드 정보  $\tilde{\mathbf{r}}(k-2)$ 를 멀티플렉싱하는 단계(808)로서, 멀티플렉싱된 데이터 스트림  $\tilde{\mathbf{B}}(k-2)$ 이 획득되는 단계

[0116] 를 포함한다.

[0117] 분해 단계(802)에서 획득된 주변 HOA 컴포넌트  $\tilde{\mathbf{c}}_{\text{AMB}}(k-1)$ 는  $O_{\text{MIN}}$  최하위 위치들(즉, 최하위 인덱스들을 갖는 위치들) 내의 입력 HOA 표현  $\mathbf{c}_n(k-1)$ 의 제1 HOA 계수 시퀀스들, 및 나머지 상위 위치들 내의 제2 HOA 계

수 시퀀스들  $\mathbf{c}_{AMB,n}(k-1)$  을 포함한다. 제2 계수 시퀀스들은 입력 HOA 표현과 우세 사운드 신호들의 HOA 표현 사이의 잔차의 HOA 표현의 부분이다.

[0118] 제1  $O_{MIN}$  지수들  $e_i(k-2), i=1, \dots, O_{MIN}$  및 지수 플래그들  $\beta_i(k-2), i=1, \dots, O_{MIN}$  은 베이스 계층 사이드 정보 소스 코더(320)에서 인코딩되고, 여기서 인코딩된 베이스 계층 사이드 정보  $\tilde{\mathbf{r}}_{BASE}(k-2)$  가 획득되고,  $O_{MIN} = (N_{MIN} + 1)^2$  및  $O = (N+1)^2$  이 되고,  $N_{MIN} \leq N$  및  $O_{MIN} \leq I$  이 되고,  $N_{MIN}$  이 사전 정의된 정수 값이다.

[0119] 제1  $O_{MIN}$  지각 인코딩된 전송 신호들  $\tilde{\mathbf{z}}_i(k-2), i=1, \dots, O_{MIN}$  및 인코딩된 베이스 계층 사이드 정보  $\tilde{\mathbf{r}}_{BASE}(k-2)$  는 베이스 계층 비트스트림 멀티플렉서(340)에서 멀티플렉싱되고(809), 여기서 베이스 계층 비트스트림  $\tilde{\mathbf{b}}_{BASE}(k-2)$  이 획득된다. 나머지  $I - O_{MIN}$  지수들  $e_i(k-2), i=O_{MIN}+1, \dots, I$  및 예외 플래그들  $\beta_i(k-2), i=O_{MIN}+1, \dots, I$ , 상기 제1 투플 세트들  $\mathcal{M}_{DIR}(k-1)$  및 제2 투플 세트들  $\mathcal{M}_{VEC}(k-1)$ , 상기 예측 파라미터들  $\xi(k-1)$  및 상기 최종 할당 벡터  $\mathbf{v}_A(k-2)$  (또한, 도면에서는  $\mathbf{v}_{AMB,ASSIGN}(k)$  로 도시됨)는 향상 계층 사이드 정보 인코더(330)에서 인코딩되고, 여기서 인코딩된 향상 계층 사이드 정보  $\tilde{\mathbf{r}}_{ENH}(k-2)$  가 획득된다.

[0120] 나머지  $I - O_{MIN}$  지각 인코딩된 전송 신호들  $\tilde{\mathbf{z}}_i(k-2), i=O_{MIN}+1, \dots, I$  및 인코딩된 향상 계층 사이드 정보  $\tilde{\mathbf{r}}_{ENH}(k-2)$  는 향상 계층 비트스트림 멀티플렉서(350)에서 멀티플렉싱되고(810), 여기서 향상 계층 비트스트림  $\tilde{\mathbf{b}}_{ENH}(k-2)$  이 획득된다.

[0121] 전술한 바와 같이, 계층화 모드의 이용을 신호화하는 모드 지시가 가산된다(811). 모드 지시는 지시 삽입 블록 또는 멀티플렉서에 가산된다.

[0122] 일 실시예에서, 방법은 베이스 계층 비트스트림  $\tilde{\mathbf{b}}_{BASE}(k-2)$ , 향상 계층 비트스트림  $\tilde{\mathbf{b}}_{ENH}(k-2)$  및 모드 지시를 단일 비트스트림으로 멀티플렉싱하는 최종 단계를 더 포함한다.

[0123] 일 실시예에서, 상기 우세 방향 추정은 에너지적으로 우세한 HOA 컴포넌트들의 방향 전력 분포에 의존한다.

[0124] 일 실시예에서, 주변 HOA 컴포넌트를 수정하는 단계에서, 계수 시퀀스들의 페이드 인(fade in) 및 페이드 아웃(fade out)은 선택된 HOA 계수 시퀀스들의 HOA 시퀀스 인덱스들이 연속 프레임들마다 차이가 있는 경우 수행된다.

[0125] 일 실시예에서, 주변 HOA 컴포넌트를 수정하는 단계에서, 주변 HOA 컴포넌트  $\mathbf{c}_{AMB}(k-1)$  의 부분 역상관(decorrelation)이 수행된다.

[0126] 일 실시예에서, 제1 투플 세트들  $\mathcal{M}_{DIR}(k)$  에 포함되는 양자화된 방향은 우세 방향이다.

[0127] 도 9는 압축된 HOA 신호를 압축 해제하는 방법의 흐름도를 보여준다.

[0128] 본 발명의 이 실시예에서, 압축된 HOA 신호를 압축 해제하는 방법(900)은 HOA 계수 시퀀스들의 출력 시간 프레임들  $\tilde{\mathbf{c}}(k-1)$  을 획득하기 위해 지각 디코딩 및 소스 디코딩과, 후속 공간 HOA 디코딩을 포함하고, 상기 방법은 압축된 고차 앰비소닉스(HOA)(Higher Order Ambisonics) 신호가 압축된 베이스 계층 비트스트림  $\tilde{\mathbf{b}}_{BASE}(k)$  및 압축된 향상 계층 비트스트림  $\tilde{\mathbf{b}}_{ENH}(k)$  을 포함하는 것을 지시하는 계층화 모드 지시 LMF를 검출하는 단계(901)를 포함한다.

[0129] 지각 디코딩 및 소스 디코딩은,

[0130] 압축된 베이스 계층 비트스트림을  $\tilde{\mathbf{b}}_{BASE}(k)$  을 디멀티플렉싱하는 단계(902)로서, 제1 지각 인코딩된 전송 신호들  $\tilde{\mathbf{z}}_i(k), i=1, \dots, O_{MIN}$  및 제1 인코딩된 사이드 정보  $\tilde{\mathbf{r}}_{BASE}(k)$  가 획득되는 단계,

[0131] 압축된 향상 계층 비트스트림  $\tilde{\mathbf{b}}_{ENH}(k)$  을 디멀티플렉싱하는 단계(903)로서, 제2 지각 인코딩된 전송 신호들  $\tilde{\mathbf{z}}_i(k), i=O_{MIN}+1, \dots, I$  및 제2 인코딩된 사이드 정보  $\tilde{\mathbf{r}}_{ENH}(k)$  가 획득되는 단계,



- [0132] 지각 인코딩된 전송 신호들  $\check{\mathbf{z}}_i(k), i = 1, \dots, I$  을 지각 디코딩하는 단계(904)로서, 지각 디코딩된 전송 신호들  $\hat{\mathbf{z}}_i(k)$  이 획득되고, 베이스 계층 지각 디코더(540)에서, 베이스 계층의 상기 제1 지각 인코딩된 전송 신호들  $\check{\mathbf{z}}_i(k), i = 1, \dots, O_{\text{MIN}}$  이 디코딩되고, 제1 지각 디코딩된 전송 신호들  $\hat{\mathbf{z}}_i(k), i = 1, \dots, O_{\text{MIN}}$  이 획득되고, 향상 계층 지각 디코더(550)에서, 향상 계층의 상기 제2 지각 인코딩된 전송 신호들  $\check{\mathbf{z}}_i(k), i = O_{\text{MIN}} + 1, \dots, I$  이 디코딩되고, 제2 지각 디코딩된 전송 신호들  $\hat{\mathbf{z}}_i(k), i = O_{\text{MIN}} + 1, \dots, I$  이 획득되는 단계,
- [0133] 베이스 계층 사이드 정보 소스 디코더(530)에서 제1 인코딩된 사이드 정보  $\check{\mathbf{f}}_{\text{BASE}}(k)$  를 디코딩하는 단계(905)로서, 제1 지수들  $e_i(k), i = 1, \dots, O_{\text{MIN}}$  및 제1 예외 플래그들  $\beta_i(k), i = 1, \dots, O_{\text{MIN}}$  이 획득되는 단계, 및
- [0134] 향상 계층 사이드 정보 소스 디코더(560)에서 제2 인코딩된 사이드 정보  $\check{\mathbf{f}}_{\text{ENH}}(k)$  를 디코딩하는 단계(906)로서, 제2 지수들  $e_i(k), i = O_{\text{MIN}} + 1, \dots, I$  및 제2 예외 플래그들  $\beta_i(k), i = O_{\text{MIN}} + 1, \dots, I$  이 획득되고, 추가 데이터가 획득되는 단계 - 추가 데이터는 방향 신호들을 위한 제1 튜플 세트  $\mathcal{M}_{\text{DIR}}(k+1)$  및 벡터 기반 신호들을 위한 제2 튜플 세트  $\mathcal{M}_{\text{VEC}}(k+1)$  를 포함하고, 제1 튜플 세트  $\mathcal{M}_{\text{DIR}}(k+1)$  의 각각의 튜플은 방향 신호 및 각각의 양자화된 방향의 인덱스를 포함하고, 제2 튜플 세트  $\mathcal{M}_{\text{VEC}}(k+1)$  의 각각의 튜플은 벡터 기반 신호 및 벡터 기반 신호의 방향 분포를 정의하는 벡터의 인덱스를 포함하고, 예측 파라미터들  $\xi(k+1)$  및 주변 할당 벡터  $\mathbf{v}_{\text{AMB,ASSIGN}}(k)$  가 획득됨 - 를 포함한다. 주변 할당 벡터  $\mathbf{v}_{\text{AMB,ASSIGN}}(k)$  는 각각의 송신 채널이 주변 HOA 컴포넌트의 계수 시퀀스를 포함하는지와 어느 계수 시퀀스를 포함하는지를 각각의 송신 채널에 대해 나타내는 컴포넌트들을 포함한다.
- [0135] 공간 HOA 디코딩은,
- [0136] 역 이득 제어를 수행하는 단계(910)로서, 상기 제1 지각 디코딩된 전송 신호들  $\hat{\mathbf{z}}_i(k), i = 1, \dots, O_{\text{MIN}}$  이 제1 지수들  $e_i(k), i = 1, \dots, O_{\text{MIN}}$  및 제1 예외 플래그들  $\beta_i(k), i = 1, \dots, O_{\text{MIN}}$  에 따라 제1 이득 정정된 신호 프레임들  $\hat{\mathbf{y}}_i(k), i = 1, \dots, O_{\text{MIN}}$  로 변환되고, 제2 지각 디코딩된 전송 신호들  $\hat{\mathbf{z}}_i(k), i = O_{\text{MIN}} + 1, \dots, I$  이 제2 지수들  $e_i(k), i = O_{\text{MIN}} + 1, \dots, I$  및 제2 예외 플래그들  $\beta_i(k), i = O_{\text{MIN}} + 1, \dots, I$  에 따라 제2 이득 정정된 신호 프레임들  $\hat{\mathbf{y}}_i(k), i = O_{\text{MIN}} + 1, \dots, I$  로 변환되는 단계,
- [0137] 채널 재할당 블록(605)에서, 제1 및 제2 이득 정정된 신호 프레임들  $\hat{\mathbf{y}}_i(k), i = 1, \dots, I$  을  $I$  채널들에 재분배하는 단계(911)로서, 우세 사운드 신호들  $\hat{\mathbf{x}}_{\text{PS}}(k)$  의 프레임들이 재구성되고, 우세 사운드 신호들이 방향 신호들 및 벡터 기반 신호들을 포함하고, 수정된 주변 HOA 컴포넌트  $\check{\mathbf{c}}_{\text{LAMB}}(k)$  가 획득되고, 상기 주변 할당 벡터  $\mathbf{v}_{\text{AMB,ASSIGN}}(k)$  및 상기 제1 및 제2 튜플 세트들  $\mathcal{M}_{\text{DIR}}(k+1)$  및  $\mathcal{M}_{\text{VEC}}(k+1)$  내의 정보에 따라 할당이 이루어지는 단계,
- [0138] 채널 재할당 블록(605)에서,  $k$  번째 프레임에서 활성인 수정된 주변 HOA 컴포넌트의 계수 시퀀스들의 인덱스들  $\mathcal{I}_{\text{AMB,ACT}}(k)$  의 제1 세트, 및 인에이블되고 디스에이블되고  $(k-1)$  번째 프레임에서 활성으로 유지될 필요가 있는 수정된 주변 HOA 컴포넌트의 계수 시퀀스들의 인덱스들  $\mathcal{I}_{\text{E}}(k-1), \mathcal{I}_{\text{D}}(k-1), \mathcal{I}_{\text{U}}(k-1)$  의 제2 세트를 생성하는 단계(911b),
- [0139] 우세 사운드 합성 블록(606)에서, 상기 우세 사운드 신호들  $\hat{\mathbf{x}}_{\text{PS}}(k)$  로부터 우세 HOA 사운드 컴포넌트들  $\hat{\mathbf{c}}_{\text{PS}}(k-1)$  의 HOA 표현을 합성하는 단계(912)로서, 상기 제1 및 제2 튜플 세트들  $\mathcal{M}_{\text{DIR}}(k+1)$  및  $\mathcal{M}_{\text{VEC}}(k+1)$ , 예측 파라미터들  $\xi(k+1)$  및 인덱스들  $\mathcal{I}_{\text{E}}(k-1), \mathcal{I}_{\text{D}}(k-1), \mathcal{I}_{\text{U}}(k-1)$  의 제2 세트가 사용되는 단계,
- [0140] 주변 합성 블록(607)에서, 수정된 주변 HOA 컴포넌트  $\check{\mathbf{c}}_{\text{LAMB}}(k)$  로부터 주변 HOA 컴포넌트  $\hat{\mathbf{c}}_{\text{AMB}}(k-1)$  를 합성하는 단계(913)로서, 제1  $O_{\text{MIN}}$  채널들을 위한 역 공간 변환이 이루어지고, 인덱스들  $\mathcal{I}_{\text{AMB,ACT}}(k)$  의 제1 세트가

사용되고, 인덱스들의 제1 세트가  $k$  번째 프레임에서 활성화된 주변 HOA 컴포넌트의 계수 시퀀스들의 인덱스들이며, 주변 HOA 컴포넌트가 계층화 모드 지시  $LMF_0$ 에 따라서 적어도 2개의 상이한 구성들 중 하나를 갖는 단계, 및

[0141] HOA 구성 블록(608)에서 우세 HOA 사운드 컴포넌트들  $\hat{\mathbf{c}}_{PS}(k-1)$  및 주변 HOA 컴포넌트  $\hat{\mathbf{c}}_{AMB}(k-1)$ 의 HOA 표현을 가산하는 단계(914)로서, 우세 사운드 신호들의 HOA 표현의 계수들 및 주변 HOA 컴포넌트의 대응하는 계수들이 가산되고, 압축 해제된 HOA 신호  $\hat{\mathbf{c}}'(k-1)$ 가 획득되는 단계

[0142] 를 포함하고, 다음의 조건들이 적용된다:

[0143] 계층화 모드 지시  $LMF_0$ 가 적어도 2개의 계층을 가진 계층화 모드를 지시하는 경우, 가장 높은  $I-O_{MIN}$  계수 채널들만이 우세 HOA 사운드 컴포넌트들  $\hat{\mathbf{c}}_{PS}(k-1)$  및 주변 HOA 컴포넌트  $\hat{\mathbf{c}}_{AMB}(k-1)$ 의 가산에 의해 획득되고, 압축 해제된 HOA 신호  $\hat{\mathbf{c}}'(k-1)$ 의 가장 낮은  $O_{MIN}$  계수 채널들이 주변 HOA 컴포넌트  $\hat{\mathbf{c}}_{AMB}(k-1)$ 로부터 복사된다. 그렇지 않고, 계층화 모드 지시  $LMF_0$ 가 단일-계층 모드를 지시하는 경우, 압축 해제된 HOA 신호  $\hat{\mathbf{c}}'(k-1)$ 의 모든 계수 채널들은 우세 사운드 컴포넌트들  $\hat{\mathbf{c}}_{PS}(k-1)$  및 주변 HOA 컴포넌트  $\hat{\mathbf{c}}_{AMB}(k-1)$ 의 가산에 의해 획득된다.

[0144] 계층화 모드 지시  $LMF_0$ 에 따르는 주변 HOA 컴포넌트의 구성은 다음과 같다:

[0145] 계층화 모드 지시  $LMF_0$ 가 적어도 2개의 계층을 가진 계층화 모드를 지시하는 경우, 주변 HOA 컴포넌트는, 그  $O_{MIN}$  최하위 위치들 내의, 압축 해제된 HOA 신호  $\hat{\mathbf{c}}(k-1)$ 의 HOA 계수 시퀀스들, 및 나머지 상위 위치 내의, 압축 해제된 HOA 신호  $\hat{\mathbf{c}}(k-1)$ 와 우세 HOA 사운드 컴포넌트들  $\hat{\mathbf{c}}_{PS}(k-1)$ 의 HOA 표현 사이의 잔차의 HOA 표현의 부분인 계수 시퀀스들을 포함한다.

[0146] 한편, 계층화 모드 지시  $LMF_0$ 가 단일-계층 모드를 지시하는 경우, 주변 HOA 컴포넌트는 압축 해제된 HOA 신호  $\hat{\mathbf{c}}(k-1)$ 와 우세 사운드 컴포넌트  $\hat{\mathbf{c}}_{PS}(k-1)$ 의 HOA 표현 사이의 잔차이다.

[0147] 일 실시예에서, 압축된 HOA 신호 표현은 멀티플렉싱된 비트스트림이고, 압축된 HOA 신호를 압축 해제하는 방법은 압축된 HOA 신호 표현을 디멀티플렉싱하는 초기 단계를 더 포함하고, 여기서 상기 압축된 베이스 계층 비트스트림  $\hat{\mathbf{b}}_{BASE}(k)$ , 상기 압축된 향상 계층 비트스트림  $\hat{\mathbf{b}}_{ENH}(k)$  및 상기 계층화 모드 지시  $LMF_0$ 가 획득된다.

[0148] 도 10은 본 발명의 일 실시예에 따른 HOA 압축 해제기의 공간 HOA 디코딩 부의 아키텍처의 부분들의 상세를 도시한다.

[0149] 유리하게는, 예를 들어 EL이 수신되지 않거나 또는 BL 품질이 충분한 경우, BL만을 디코딩하는 것이 가능하다. 이 경우, EL의 신호들은 디코더에서 0으로 설정될 수 있다. 그 후, 채널 재할당 블록(605)에서 제1 및 제2 이득 정정된 신호 프레임들  $\hat{\mathbf{y}}_i(k), i=1, \dots, I$ 을  $I$  채널들에 재분배하는 단계(911)는, 우세 사운드 신호들  $\hat{\mathbf{x}}_{PS}(k)$ 의 프레임들이 비어있기 때문에, 매우 단순하다. 인에이블되고 디스에이블되고  $(k-1)$ 번째 프레임에서 활성화로 유지될 필요가 있는 수정된 주변 HOA 컴포넌트의 계수 시퀀스들의 인덱스들  $\mathcal{I}_E(k-1), \mathcal{I}_D(k-1), \mathcal{I}_U(k-1)$ 의 제2 세트는 0으로 설정된다. 따라서, 우세 사운드 합성 블록(606)에서, 우세 사운드 신호들  $\hat{\mathbf{x}}_{PS}(k)$ 로부터 우세 HOA 사운드 컴포넌트들  $\hat{\mathbf{c}}_{PS}(k-1)$ 의 HOA 표현을 합성하는 단계(912)는 스킵될 수 있고, 주변 합성 블록(607)에서, 수정된 주변 HOA 컴포넌트  $\hat{\mathbf{c}}_{I,AMB}(k)$ 로부터 주변 HOA 컴포넌트  $\hat{\mathbf{c}}_{AMB}(k-1)$ 를 합성하는 단계(913)는 종래의 HOA 구성에 대응한다.

[0150] HOA 압축을 위한 최초(즉, 모노리식, 논-스케일러블, 비계층화) 모드는, 저품질 베이스 계층 비트 스트림이, 예를 들어 파일 기반 압축에 요구되지 않는 애플리케이션에 여전히 유용할 수 있다. 최초 HOA 컴포넌트  $\mathbf{c}$ 의 공간적으로 변환된 계수 시퀀스들 대신에, 최초 HOA 표현과 방향 HOA 표현 간의 차이인, 주변 HOA 컴포넌트  $\mathbf{c}_{AMB}$ 의

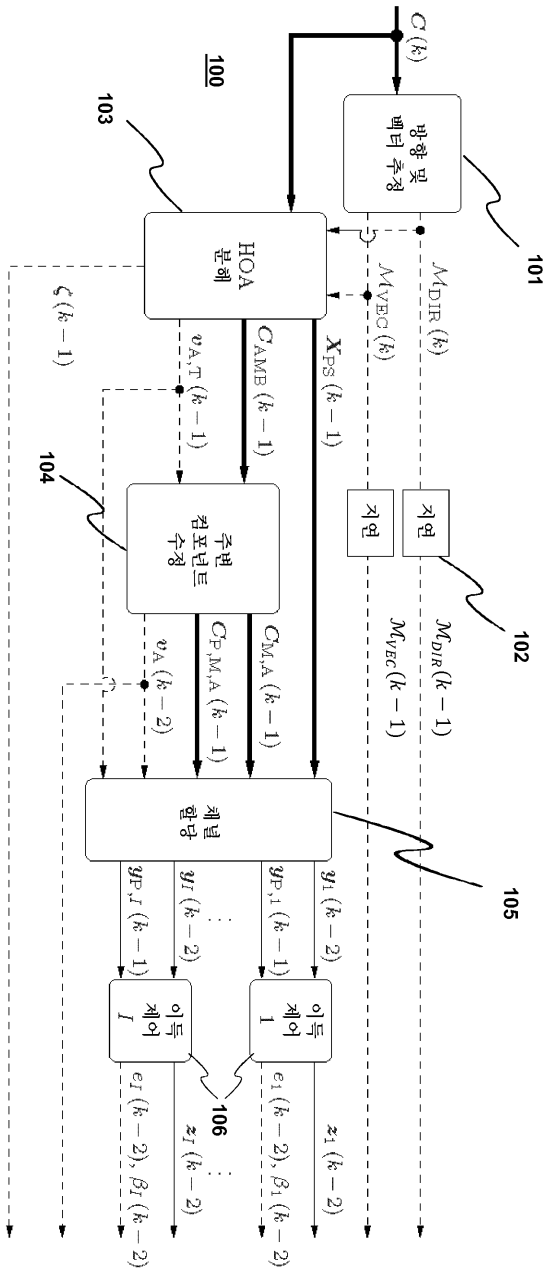


공간적으로 변환된 제1  $O_{\text{MIN}}$  계수 시퀀스들을 지각 코딩하는 주요 장점은, 전자의 경우 지각 코딩될 모든 신호들 사이의 교차 상관관계가 감소된다는 점이다. 신호들  $\mathbf{z}_i, i = 1, \dots, I$  사이의 임의의 교차 상관관계는 공간 디코딩 프로세스 동안 지각 코딩 노이즈의 구조적 중첩을 야기할 수 있고, 동시에 노이즈가 없는 HOA 계수 시퀀스들이 중첩에서 삭제된다. 이 현상은 지각 노이즈 언마스킹(perceptual noise unmasking)으로 알려져 있다. 계층화 모드에서, 신호들  $\mathbf{z}_i, i = 1, \dots, O_{\text{MIN}}$  사이, 또한 신호들  $\mathbf{z}_i, i = 1, \dots, O_{\text{MIN}}$  과  $i = O_{\text{MIN}} + 1, \dots, I$  사이에는 높은 교차 상관관계가 존재하며, 그 이유는 주변 HOA 컴포넌트  $\tilde{\mathbf{c}}_{\text{AMB},n}, n = 1, \dots, O_{\text{MIN}}$ 의 수정된 계수 시퀀스들이 방향 HOA 컴포넌트의 신호들을 포함하기 때문이다(수학식 3 참고). 반대로, 이것은 최초 비계층화 모드에서는 그렇지 않다. 따라서, 계층화 모드에 의해 도입된 송신 강건성은 압축 품질의 희생으로 따라올 수 있다는 것으로 결론지었다. 그러나 압축 품질의 감소는 송신 강건성의 증가와 비교하여 낮다. 위에 보인 바와 같이, 제안된 계층화 모드는 적어도 위에 설명된 상황에서는 유리하다.

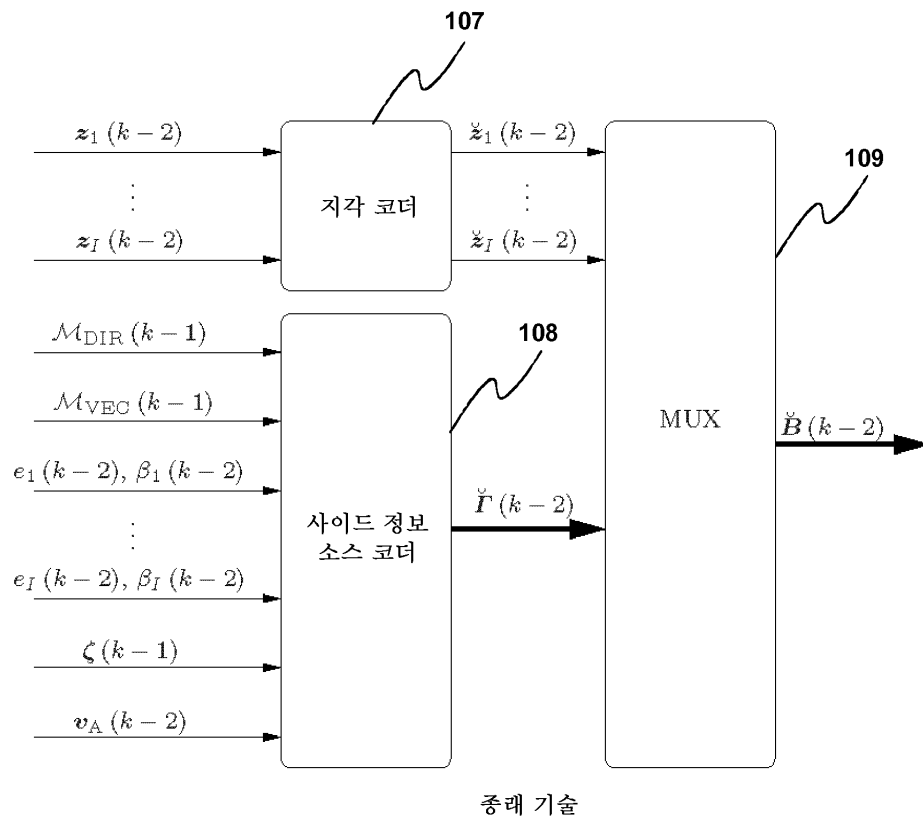
- [0151] 본 발명의 바람직한 실시예들에 적용되는 것처럼 본 발명의 근본적인 신규한 특징들을 도시하고 설명하고 지적하였으나, 본 발명의 정신으로부터 벗어나지 않고도 본 기술분야의 통상의 기술자들에 의해서, 설명한 장치 및 방법에서의 다양한 생략, 치환, 및 변경이 개시된 디바이스의 형태 및 상세 사항에서, 그 디바이스의 동작에서 이루어질 수 있음을 이해할 것이다. 동일한 결과들을 얻기 위해 실질적으로 동일한 방식으로 실질적으로 동일한 기능을 수행하는 요소들의 모든 조합들은 본 발명의 범위 내에 속하는 것이 명확히 의도된다. 하나의 설명된 실시예로부터의 요소들의 다른 실시예로의 치환도 충분히 의도되고 고려된다.
- [0152] 본 발명은 단지 예시적으로 설명되었을 뿐이며, 본 발명의 범위로부터 벗어나지 않고 상세의 변경들이 이루어질 수 있다는 것을 이해할 것이다.
- [0153] 설명 및(적절한 경우에) 청구항들 및 도면들에 개시된 각각의 특징은 독립적으로 또는 임의의 적절한 조합으로 제공될 수 있다. 특징들은 적절한 경우에 하드웨어, 소프트웨어, 또는 이 둘의 조합으로 구현될 수 있다. 적용 가능한 경우, 접속은 무선 접속 또는 유선 접속으로서 실시될 수 있으며 반드시 직접 또는 전용일 필요는 없다.
- [0154] 청구항들에 보이는 참조 번호는 예시만을 목적으로 한 것으로, 청구항들에 한정적인 효과를 미치지 않는다.
- [0155] 참조 문헌들
- [0156] [1] EP12306569.0
- [0157] [2] EP12305537.8(EP2665208A로 공개됨)
- [0158] [3] EP133005558.2
- [0159] [4] ISO/IEC JTC1/SC29/WG11 N14264. Working draft 1-HOA text of MPEG-H 3D audio, January 2014

도면

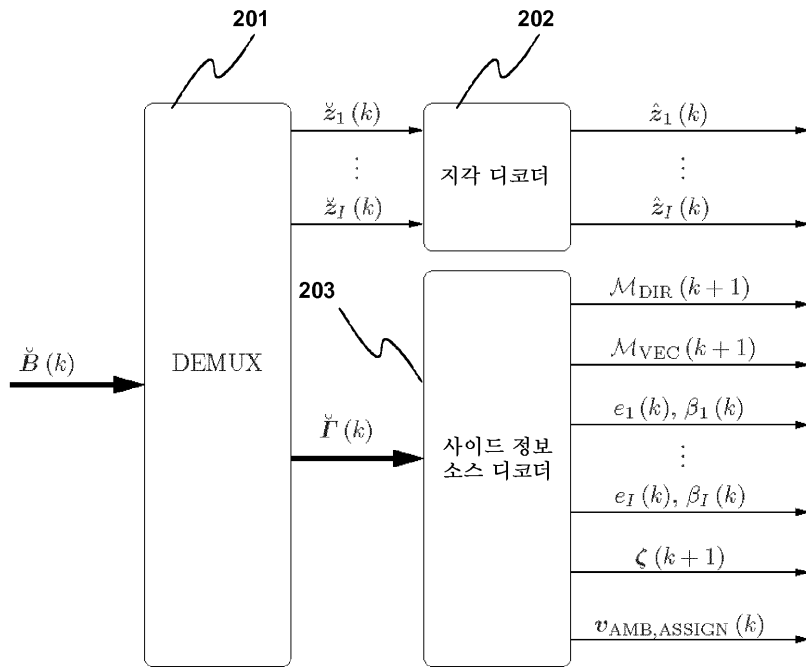
도면1a



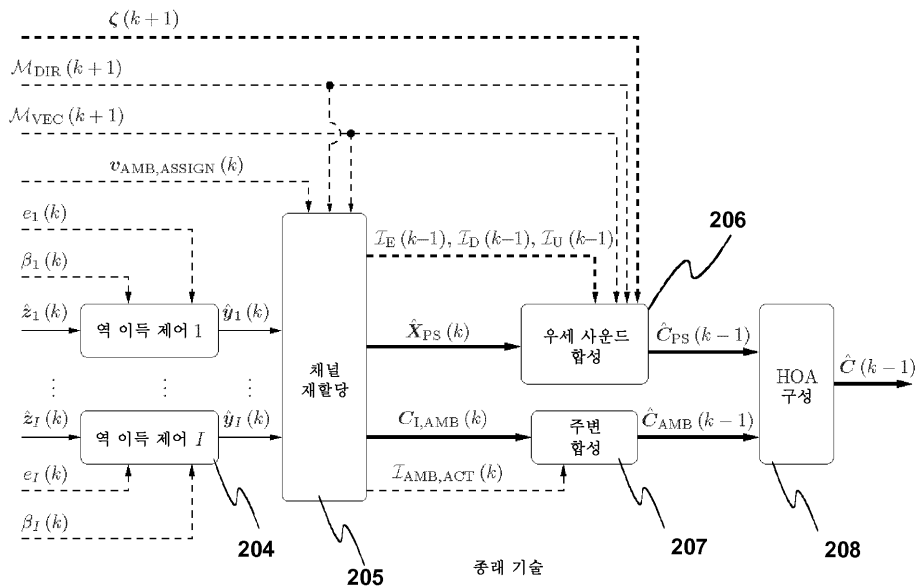
도면1b



도면2

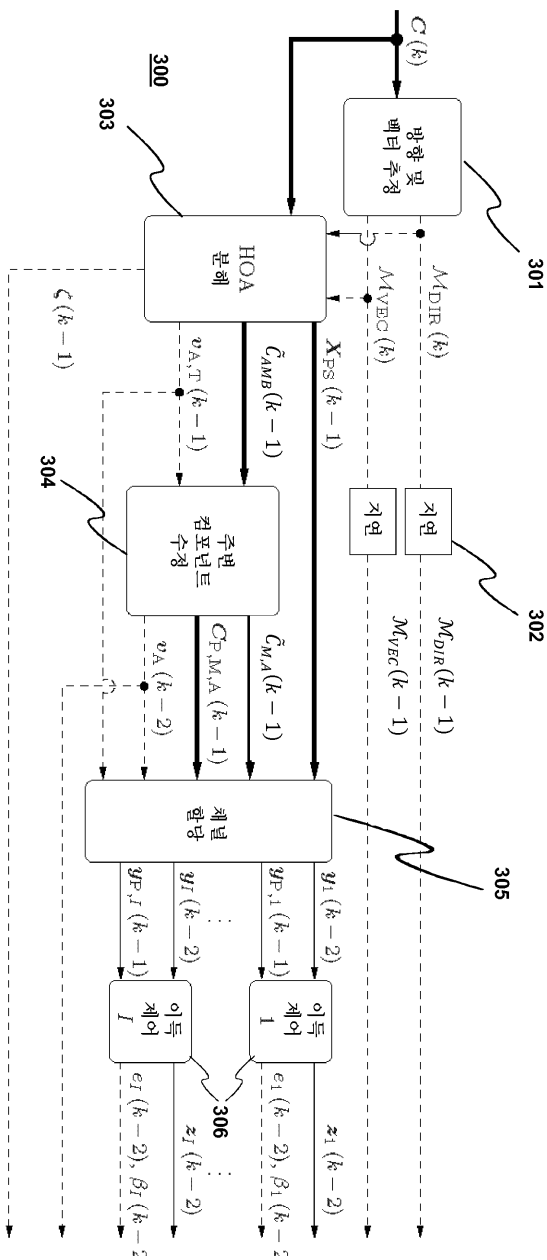


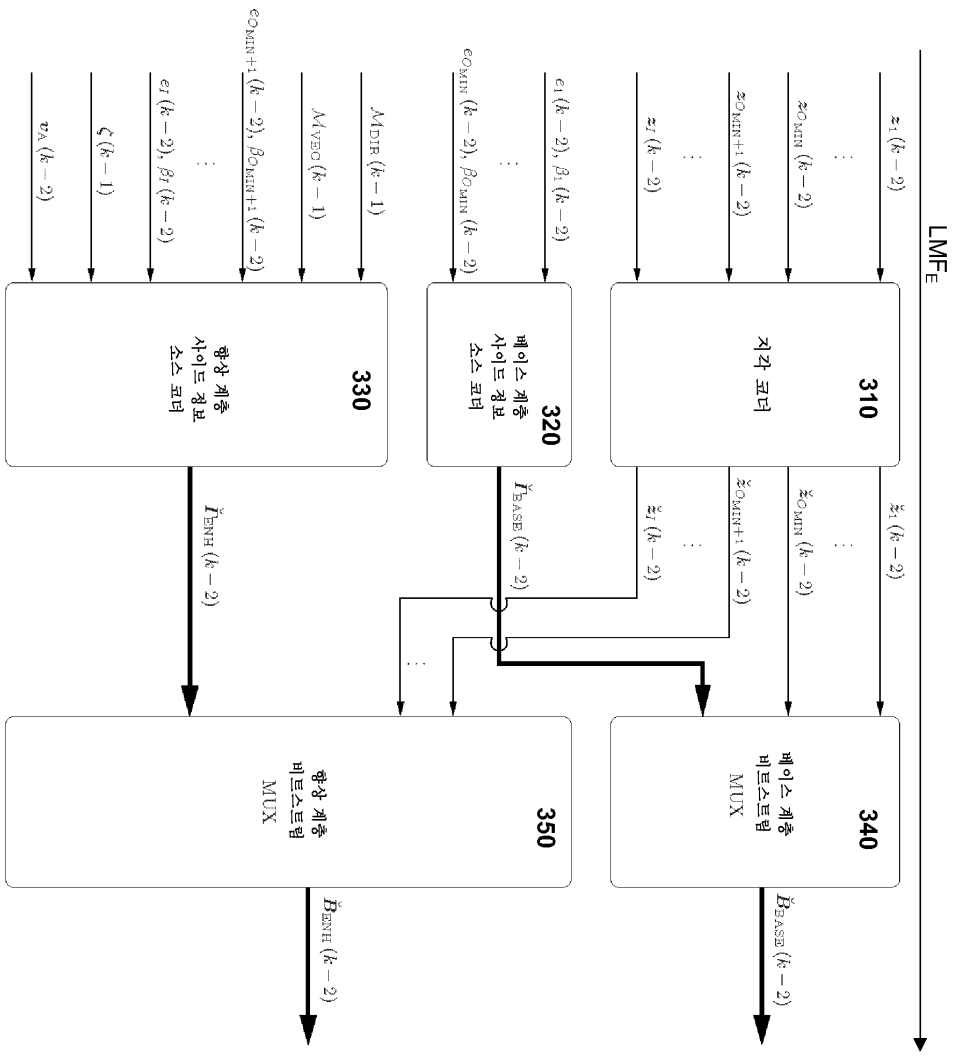
a)



b)

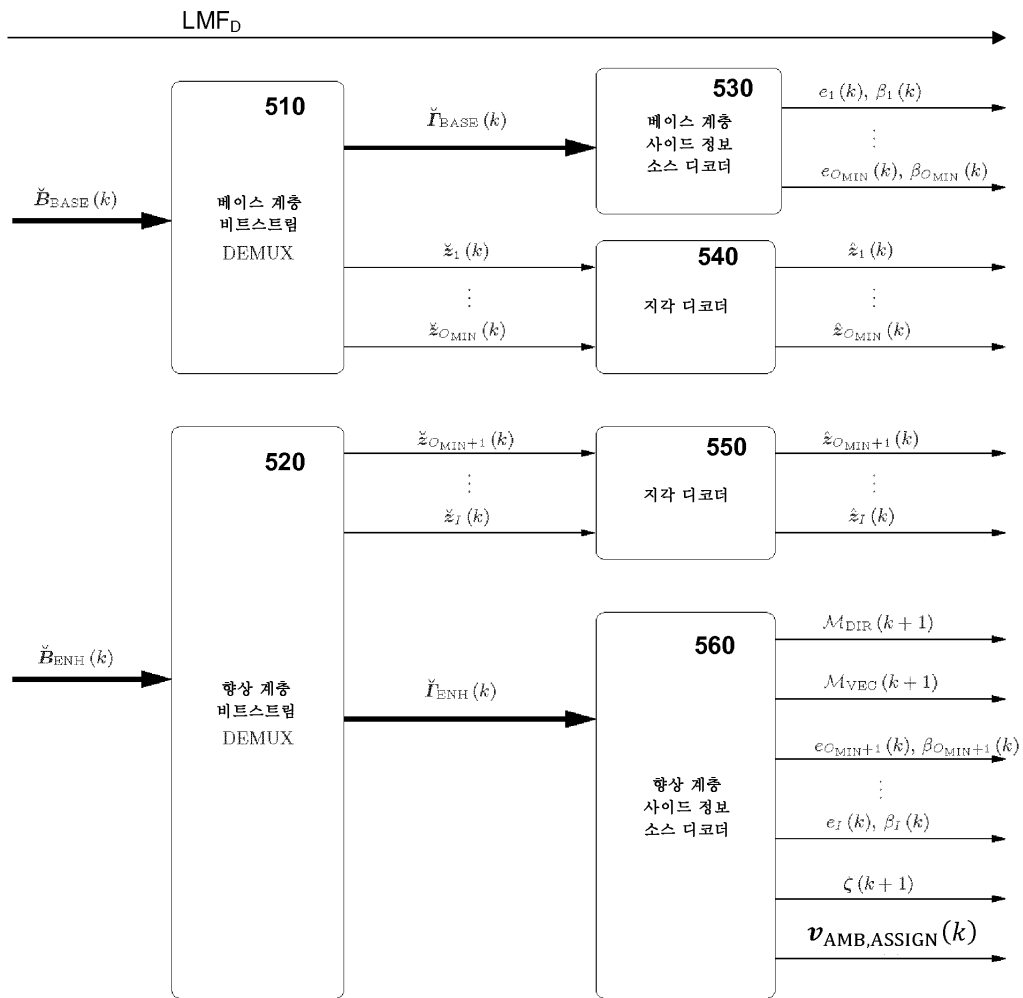
도면3





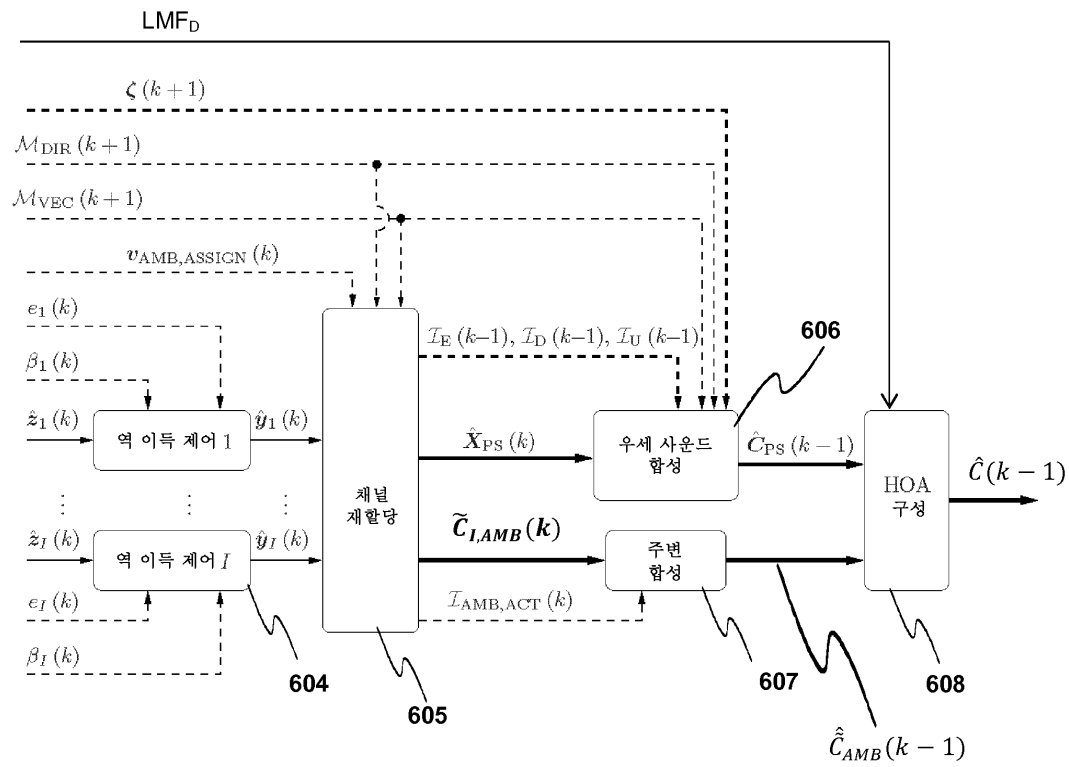
도면4

도면5

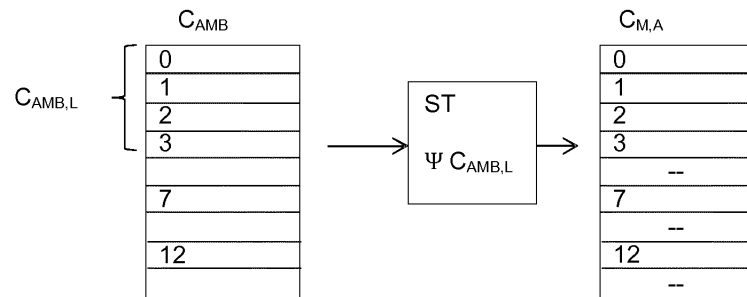




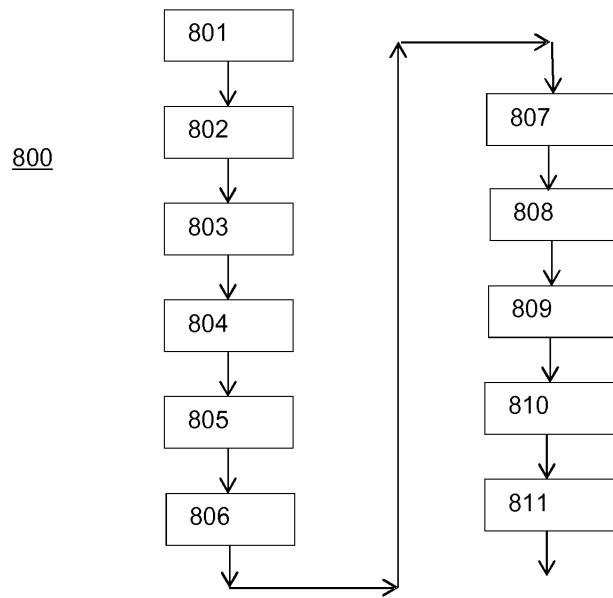
도면6



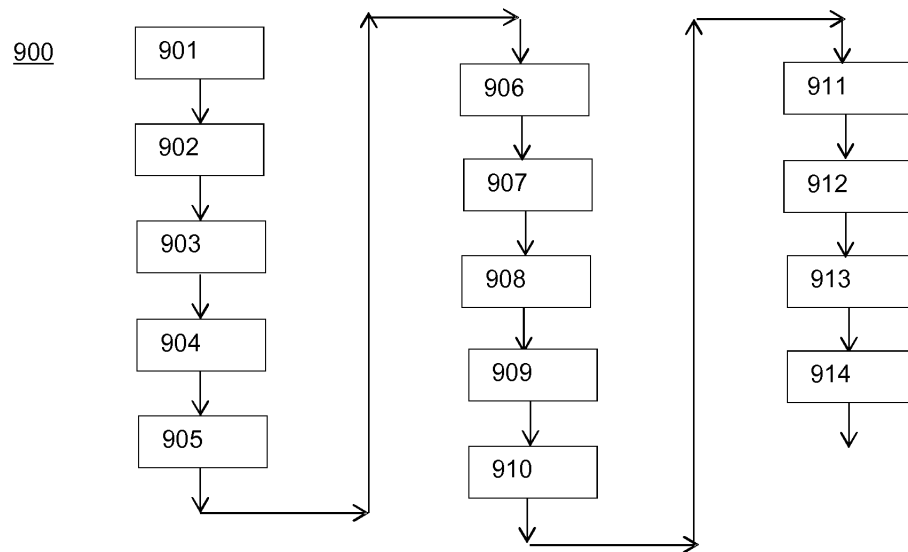
도면7



도면8



도면9



도면10

