



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2021년09월01일
(11) 등록번호 10-2296067
(24) 등록일자 2021년08월25일

- (51) 국제특허분류(Int. Cl.)
HOAS 3/02 (2006.01) GIOL 19/008 (2014.01)
HOAS 3/00 (2006.01)
- (52) CPC특허분류
HOAS 3/02 (2013.01)
GIOL 19/008 (2020.08)
- (21) 출원번호 10-2016-7035529
- (22) 출원일자(국제) 2015년07월02일
심사청구일자 2020년07월01일
- (85) 번역문제출일자 2016년12월19일
- (65) 공개번호 10-2017-0024581
- (43) 공개일자 2017년03월07일
- (86) 국제출원번호 PCT/EP2015/065086
- (87) 국제공개번호 WO 2016/001356
국제공개일자 2016년01월07일
- (30) 우선권주장
14306080.4 2014년07월02일
유럽특허청(EPO)(EP)
14194186.4 2014년11월20일
유럽특허청(EPO)(EP)
- (56) 선행기술조사문헌
EP02665208 A1
EP02743922 A1

- (73) 특허권자
돌비 인터네셔널 에이비
네덜란드 1101 씨엔 암스트레담 주이두스트 헤리
커베르그백 1-35 3이 아폴로 빌딩
- (72) 발명자
크뢰거, 알렉산더
독일 30655 하노버 포이어바흐스트라세 16
고돈, 스텐
독일 31515 분스토르프 뮌헨캄프스트라세 50에이
- (74) 대리인
양영준, 정은진, 백만기

전체 청구항 수 : 총 24 항

심사관 : 권영학

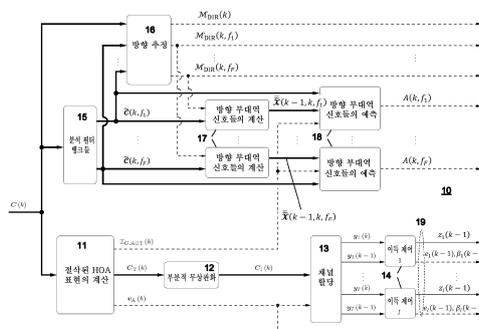
(54) 발명의 명칭 압축된 HOA 표현을 디코딩하기 위한 방법 및 장치와 압축된 HOA 표현을 인코딩하기 위한 방법 및 장치

(57) 요약

HOA(Higher Order Ambisonics) 신호들의 인코딩은 통상적으로 높은 데이터 레이트로 이어진다. 계수 시퀀스들을 갖는 입력 HOA 신호의 프레임들을 낮은 비트 레이트 인코딩하기 위한 방법은, 절삭된 HOA 표현($C_T(k)$)을 계산하는 단계(s110), 활성 계수 시퀀스들($I_{C,ACT}(k)$)을 결정하는 단계(s111), 후보 방향들($M_{DIR}(k)$)을 추정하는 단계

(뒷면에 계속)

대표도



(s16), 입력 HOA 신호를 복수의 주파수 부대역들(f_1, \dots, f_F)로 분할하는 단계(s15), 주파수 부대역들 각각에 대해 후보 방향들($M_{DIR}(k)$)의 서브셋을 활성 방향들($M_{DIR}(k, f_1), \dots, M_{DIR}(k, f_F)$)로서 그리고 각각의 활성 방향에 대해 궤적을 추정하는 단계(s161), 각각의 주파수 부대역에 대해, 활성 방향들에 따라 주파수 부대역의 계수 시퀀스들로부터 방향 부대역 신호들을 계산하는 단계(s17), 각각의 주파수 부대역에 대해, 각각의 활성 계수 시퀀스들($I_{C,ACT}(k)$)을 이용하여 주파수 부대역의 계수 시퀀스들로부터 방향 부대역 신호들을 예측하는데 이용될 수 있는 예측 행렬($A(k, f_1), \dots, A(k, f_F)$)을 계산하는 단계(s18), 및 후보 방향들, 활성 방향들, 예측 행렬들 및 절삭된 HOA 표현을 인코딩하는 단계(s19)를 포함한다.

(52) CPC특허분류

H04S 3/008 (2020.05)

H04S 2420/07 (2013.01)

H04S 2420/11 (2013.01)

명세서

청구범위

청구항 1

압축된 HOA 표현을 디코딩하기 위한 방법으로서,

상기 압축된 HOA 표현으로부터 복수의 절삭된 HOA 계수 시퀀스들($\hat{z}_1(k), \dots, \hat{z}_I(k)$), 상기 절삭된 HOA 계수 시퀀스들의 시퀀스 인덱스들을 나타내거나 포함하는 할당 벡터($\mathbf{v}_{\text{AMBASSIGN}}(k)$), 부대역 관련 방향 정보($M_{\text{DIR}}(k+1, f_1), \dots, M_{\text{DIR}}(k+1, f_F)$), 복수의 예측 행렬들($A(k+1, f_1), \dots, A(k+1, f_F)$), 및 이득 제어 부가 정보($e_1(k), \beta_1(k), \dots, e_I(k), \beta_I(k)$)를 추출하는 단계;

상기 복수의 절삭된 HOA 계수 시퀀스들($\hat{z}_1(k), \dots, \hat{z}_I(k)$), 상기 이득 제어 부가 정보($e_1(k), \beta_1(k), \dots, e_I(k), \beta_I(k)$) 및 상기 할당 벡터($\mathbf{v}_{\text{AMBASSIGN}}(k)$)로부터, 절삭된 HOA 표현($\hat{c}_T(k)$)을 재구성하는 단계;

분석 필터 बैं크들에서, 상기 재구성된 절삭된 HOA 표현($\hat{c}_T(k)$)을 복수의 F개의 주파수 부대역들에 대한 주파수 부대역 표현들($\tilde{c}_T(k, f_1), \dots, \tilde{c}_T(k, f_F)$)로 분해하는 단계;

방향 부대역 합성 블록들에서, 상기 주파수 부대역 표현들 각각에 대해, 상기 재구성된 절삭된 HOA 표현의 각각의 주파수 부대역 표현($\tilde{c}_T(k, f_1), \dots, \tilde{c}_T(k, f_F)$), 상기 부대역 관련 방향 정보($M_{\text{DIR}}(k+1, f_1), \dots, M_{\text{DIR}}(k+1, f_F)$), 및 상기 예측 행렬들($A(k+1, f_1), \dots, A(k+1, f_F)$)로부터, 예측된 방향 HOA 표현($\tilde{c}_D(k, f_1), \dots, \tilde{c}_D(k, f_F)$)을 합성하는 단계;

부대역 조성 블록들에서, 상기 F개의 주파수 부대역들 각각에 대해, 계수 시퀀스가 상기 할당 벡터($\mathbf{v}_{\text{AMBASSIGN}}(k)$)에 포함되는 인덱스 n을 갖는 경우 절삭된 HOA 표현($\tilde{c}_T(k, f_j)$)의 계수 시퀀스들로부터 획득되거나, 그렇지 않으면 방향 부대역 합성 블록들 중 하나에 의해 제공된 예측된 방향 HOA 성분($\tilde{c}_D(k, f_j)$)의 계수 시퀀스들로부터 획득되는 계수 시퀀스들($\tilde{z}_n(k, f_j)$, $n = 1, \dots, O$)을 갖는 디코딩된 부대역 HOA 표현($\tilde{c}(k, f_1), \dots, \tilde{c}(k, f_F)$)을 조성하는 단계; 및

합성 필터 बैं크들에서, 디코딩된 HOA 표현($\tilde{c}(k)$)을 획득하기 위해 상기 디코딩된 부대역 HOA 표현들($\tilde{c}(k, f_1), \dots, \tilde{c}(k, f_F)$)을 합성하는 단계

를 포함하는 방법.

청구항 2

제1항에 있어서, 상기 추출하는 단계는 인코딩된 절삭된 HOA 계수 시퀀스들($\check{z}_1(k), \dots, \check{z}_I(k)$)을 포함하는 인지 코딩된 부분을 획득하는 단계를 포함하고, 상기 절삭된 HOA 계수 시퀀스들($\hat{z}_1(k), \dots, \hat{z}_I(k)$)을 획득하기 위해 상기 인코딩된 절삭된 HOA 계수 시퀀스들($\check{z}_1(k), \dots, \check{z}_I(k)$)을 인지 디코더에서 인지 디코딩하는 단계를 더 포함하는, 방법.

청구항 3

제1항에 있어서, 상기 추출하는 단계는 인코딩된 부가 정보 부분을 획득하는 단계를 포함하고, 부대역 관련 방향 정보($M_{DIR}(k+1, f_1), \dots, M_{DIR}(k+1, f_F)$), 예측 행렬들($A(k+1, f_1), \dots, A(k+1, f_F)$), 이득 제어 부가 정보($e_1(k), \beta_1(k), \dots, e_I(k), \beta_I(k)$), 및 할당 벡터($v_{AMB, ASSIGN}(k)$)를 획득하기 위해 상기 인코딩된 부가 정보 부분을 부가 정보 소스 디코더에서 디코딩하는 단계를 더 포함하는, 방법.

청구항 4

제1항에 있어서, 상기 부대역 관련 방향 정보는, 활성 방향들의 세트($M_{DIR}(k)$)와 제1 및 제2 인덱스를 갖는 인덱스들의 튜플(tuple)들을 포함하는 튜플 세트($M_{DIR}(k+1, f_1), \dots, M_{DIR}(k+1, f_F)$)를 포함하고, 상기 제2 인덱스는 현재의 주파수 부대역에 대한 상기 활성 방향들의 세트($M_{DIR}(k)$) 내의 활성 방향의 인덱스이며, 상기 제1 인덱스는 상기 활성 방향의 궤적 인덱스이고, 궤적은 특정한 사운드 소스의 방향들의 시간적 시퀀스인, 방법.

청구항 5

제1항에 있어서, 적어도 하나의 주파수 부대역 표현은 2개 이상의 주파수 부대역으로 이루어진 부대역 그룹을 포함하는, 방법.

청구항 6

제5항에 있어서, 상기 압축된 HOA 표현으로부터 부대역 그룹 구성 정보가 수신되거나 추출되고, 상기 부대역 그룹 구성 정보는 상기 합성 필터 뱅크들을 셋업하는데 이용되는, 방법.

청구항 7

주어진 개수의 계수 시퀀스들을 갖는 입력 HOA 신호의 프레임들을 인코딩하기 위한 방법으로서 - 각각의 계수 시퀀스는 인덱스를 가짐 - ,

절삭된 HOA 표현에 포함될 활성 계수 시퀀스들($I_{C, ACT}(k)$)의 인덱스들의 세트를 결정하는 단계;

감소된 개수의 비제로 계수 시퀀스들을 갖는 절삭된 HOA 표현($C_T(k)$)을 계산하는 단계;

상기 입력 HOA 신호로부터 제1 세트의 후보 방향들($M_{DIR}(k)$)을 추정하는 단계;

상기 입력 HOA 신호를 복수의 주파수 부대역들(f_1, \dots, f_F)로 분할하는 단계 - 상기 주파수 부대역들의 계수 시퀀스들($\bar{C}(k-1, k, f_1), \dots, \bar{C}(k-1, k, f_F)$)이 획득됨 - ;

상기 주파수 부대역들 각각에 대해 제2 세트의 방향들($M_{DIR}(k, f_1), \dots, M_{DIR}(k, f_F)$)을 추정하는 단계 - 상기 제2 세트의 방향들의 각각의 요소는 제1 및 제2 인덱스를 갖는 인덱스들의 튜플이고, 상기 제2 인덱스는 현재의 주파수 부대역에 대한 활성 방향의 인덱스이며, 상기 제1 인덱스는 상기 활성 방향의 궤적 인덱스이고, 각각의 활성 방향은 또한 상기 입력 HOA 신호의 상기 제1 세트의 후보 방향들($M_{DIR}(k)$)에 포함됨 - ;

상기 주파수 부대역들 각각에 대해, 각각의 주파수 부대역의 상기 제2 세트의 방향들($M_{DIR}(k, f_1), \dots, M_{DIR}(k, f_F)$)에 따라 상기 주파수 부대역의 계수 시퀀스들($\bar{C}(k-1, k, f_1), \dots, \bar{C}(k-1, k, f_F)$)로부터 방향 부대역 신호들($\bar{X}(k-1, k, f_1), \dots, \bar{X}(k-1, k, f_F)$)을 계산하는 단계;

상기 주파수 부대역들 각각에 대해, 각각의 주파수 부대역의 활성 계수 시퀀스들($I_{C, ACT}(k)$)의 인덱스들의 세트를 이용하여 상기 주파수 부대역의 계수 시퀀스들($\bar{C}(k-1, k, f_1), \dots, \bar{C}(k-1, k, f_F)$)로부터 방향 부대역 신호들($\bar{X}(k-1, k, f_1), \dots, \bar{X}(k-1, k, f_F)$)을 예측하기 위해 구성된 예측 행렬($A(k, f_1), \dots, A(k, f_F)$)을 계산하는

단계; 및

상기 제1 세트의 후보 방향들($M_{DIR}(k)$), 상기 제2 세트의 방향들($(M_{DIR}(k, f_1), \dots, M_{DIR}(k, f_F))$), 상기 예측 행렬들($A(k, f_1), \dots, A(k, f_F)$) 및 상기 절삭된 HOA 표현 ($C_T(k)$)을 인코딩하는 단계를 포함하는 방법.

청구항 8

제7항에 있어서, 2개 이상의 부대역으로 이루어진 적어도 하나의 그룹이 생성되고, 상기 적어도 하나의 그룹은 단일 부대역 대신에 이용되고 단일 부대역과 동일한 방식으로 취급되는, 방법.

청구항 9

제7항에 있어서, 상기 절삭된 HOA 표현($C_T(k)$)을 인코딩하는 단계는,

상기 절삭된 HOA 채널 시퀀스들을 부분적으로 무상관화하는 단계;

상기 절삭된 HOA 채널 시퀀스들($y_1(k), \dots, y_I(k)$)을 전송 채널들에 할당하기 위한 채널 할당 단계;

전송 채널들 각각에 관한 이득 제어를 수행하는 단계 - 각각의 전송 채널에 대한 이득 제어 부가 정보 ($e_i(k-1)$, $\beta_i(k-1)$)가 생성됨 - ;

이득 제어된 절삭된 HOA 채널 시퀀스들($z_1(k), \dots, z_I(k)$)을 인지 인코더에서 인코딩하는 단계;

상기 이득 제어 부가 정보($e_i(k-1)$, $\beta_i(k-1)$), 상기 제1 세트의 후보 방향들($M_{DIR}(k)$), 상기 제2 세트의 방향들($M_{DIR}(k, f_1), \dots, M_{DIR}(k, f_F)$), 및 상기 예측 행렬들($A(k, f_1), \dots, A(k, f_F)$)을 부가 정보 소스 코더에서 인코딩하는 단계; 및

인코딩된 HOA 신호 프레임($\bar{B}(k-1)$)을 획득하기 위해 상기 인지 인코더와 상기 부가 정보 소스 코더의 출력들을 멀티플렉싱하는 단계를 포함하는 방법.

청구항 10

제7항에 있어서, 상기 주파수 부대역들 각각에 대해 제2 세트의 방향들($M_{DIR}(k, f_1), \dots, M_{DIR}(k, f_F)$)을 추정하는 단계에서, 주파수 부대역의 방향들은 전체 대역 HOA 신호의 방향들($M_{DIR}(k)$) 중에서만 탐색되는, 방법.

청구항 11

제7항에 있어서, 활성 방향의 궤적을 결정하는 단계를 더 포함하고, 상기 활성 방향은 사운드 소스의 방향이며, 궤적은 특정한 사운드 소스의 방향들의 시간적 시퀀스인, 방법.

청구항 12

제7항에 있어서, 절삭된 HOA 표현은 하나 이상의 계수 시퀀스들이 제로로 설정되는 HOA 신호인, 방법.

청구항 13

HOA 신호를 디코딩하기 위한 장치로서,

압축된 HOA 표현으로부터 복수의 절삭된 HOA 계수 시퀀스들($\hat{z}_1(k), \dots, \hat{z}_I(k)$), 상기 절삭된 HOA 계수 시퀀스들의 시퀀스 인덱스들을 나타내거나 포함하는 할당 벡터($v_{AMBASSIGN}(k)$), 부대역 관련 방향 정보($M_{DIR}(k+1, f_1), \dots, M_{DIR}(k+1, f_F)$), 복수의 예측 행렬들($A(k+1, f_1), \dots, A(k+1, f_F)$), 및 이득 제어 부가 정보

$(e_1(k), \beta_1(k), \dots, e_I(k), \beta_I(k))$ 를 추출하도록 구성된 추출 모듈;

상기 복수의 절삭된 HOA 계수 시퀀스들($\hat{z}_1(k), \dots, \hat{z}_I(k)$), 상기 이득 제어 부가 정보($e_1(k), \beta_1(k), \dots, e_I(k), \beta_I(k)$), 및 상기 할당 벡터($v_{\text{AMB,ASSIGN}}(k)$)로부터, 절삭된 HOA 표현($\hat{c}_T(k)$)을 재구성하도록 구성된 재구성 모듈;

상기 재구성된 절삭된 HOA 표현($\hat{c}_T(k)$)을 복수의 F개의 주파수 부대역에 대한 주파수 부대역 표현들($\tilde{c}_T(k, f_1), \dots, \tilde{c}_T(k, f_F)$)로 분해하도록 구성된 분석 필터 बैं크 모듈;

상기 주파수 부대역 표현들 각각에 대해, 상기 재구성된 절삭된 HOA 표현의 각각의 주파수 부대역 표현($\tilde{c}_T(k, f_1), \dots, \tilde{c}_T(k, f_F)$), 상기 부대역 관련 방향 정보($M_{\text{DIR}}(k+1, f_1), \dots, M_{\text{DIR}}(k+1, f_F)$), 및 상기 예측 행렬들($A(k+1, f_1), \dots, A(k+1, f_F)$)로부터, 예측된 방향 HOA 표현($\tilde{c}_D(k, f_1), \dots, \tilde{c}_D(k, f_F)$)을 합성하도록 구성된 적어도 하나의 방향 부대역 합성 모듈;

상기 F개의 주파수 부대역들 각각에 대해, 계수 시퀀스가 상기 할당 벡터($v_{\text{AMB,ASSIGN}}(k)$)에 포함되는 인덱스 n 을 갖는 경우 절삭된 HOA 표현($\tilde{c}_T(k, f_j)$)의 계수 시퀀스들로부터 획득되거나, 그렇지 않으면 방향 부대역 합성 모듈 중 하나에 의해 제공된 예측된 방향 HOA 성분($\tilde{c}_D(k, f_j)$)의 계수 시퀀스들로부터 획득되는 계수 시퀀스들($\tilde{c}_n(k, f_j)$, $n = 1, \dots, O$)을 갖는 디코딩된 부대역 HOA 표현($\tilde{c}(k, f_1), \dots, \tilde{c}(k, f_F)$)을 조성하도록 구성된 적어도 하나의 부대역 조성 모듈; 및

디코딩된 HOA 표현($\tilde{c}(k)$)을 획득하기 위해 디코딩된 부대역 HOA 표현들($\tilde{c}(k, f_1), \dots, \tilde{c}(k, f_F)$)을 합성하도록 구성된 합성 필터 बैं크 모듈

을 포함하는 장치.

청구항 14

제13항에 있어서, 상기 추출 모듈은, 적어도,

인코딩된 부가 정보 부분 및 인코딩된 절삭된 HOA 계수 시퀀스들($\check{z}_1(k), \dots, \check{z}_I(k)$)을 포함하는 인지 코딩된 부분을 획득하기 위한 디멀티플렉서;

상기 절삭된 HOA 계수 시퀀스들($\hat{z}_1(k), \dots, \hat{z}_I(k)$)을 획득하기 위해 상기 인코딩된 절삭된 HOA 계수 시퀀스들($\check{z}_1(k), \dots, \check{z}_I(k)$)을 인지 디코딩하도록 구성된 인지 디코더; 및

부대역 관련 방향 정보($M_{\text{DIR}}(k+1, f_1), \dots, M_{\text{DIR}}(k+1, f_F)$), 예측 행렬들($A(k+1, f_1), \dots, A(k+1, f_F)$), 이득 제어 부가 정보($e_1(k), \beta_1(k), \dots, e_I(k), \beta_I(k)$), 및 할당 벡터($v_{\text{AMB,ASSIGN}}(k)$)를 획득하기 위해 상기 인코딩된 부가 정보 부분을 디코딩하도록 구성된 부가 정보 소스 디코더

를 포함하는, 장치.

청구항 15

제13항에 있어서, 상기 추출 모듈은 인코딩된 부가 정보 부분을 획득하고, 부대역 관련 방향 정보($M_{\text{DIR}}(k+1, f_1), \dots, M_{\text{DIR}}(k+1, f_F)$), 예측 행렬들($A(k+1, f_1), \dots, A(k+1, f_F)$), 이득 제어 부가 정보($e_1(k), \beta_1(k), \dots, e_I(k), \beta_I(k)$), 및 할당 벡터($v_{\text{AMB,ASSIGN}}(k)$)를 획득하기 위해 상기 인코딩된 부가 정보 부분

을 디코딩하도록 구성된 부가 정보 소스 디코더를 더 포함하는, 장치.

청구항 16

제13항에 있어서, 상기 부대역 관련 방향 정보는, 활성 방향들의 세트($M_{DIR}(k)$)와 제1 및 제2 인덱스를 갖는 인덱스들의 튜플(tuple)들을 포함하는 튜플 세트($M_{DIR}(k+1, f_1), \dots, M_{DIR}(k+1, f_F)$)를 포함하고, 상기 제2 인덱스는 현재의 주파수 부대역에 대한 상기 활성 방향들의 세트($M_{DIR}(k)$) 내의 활성 방향의 인덱스이며, 상기 제1 인덱스는 상기 활성 방향의 궤적 인덱스이고, 궤적은 특정한 사운드 소스의 방향들의 시간적 시퀀스인, 장치.

청구항 17

제13항에 있어서, 적어도 하나의 주파수 부대역 표현은 2개 이상의 주파수 부대역으로 이루어진 부대역 그룹을 포함하는, 장치.

청구항 18

제17항에 있어서, 상기 압축된 HOA 표현으로부터 부대역 그룹 구성 정보가 수신되거나 추출되고, 상기 부대역 그룹 구성 정보는 상기 합성 필터 बैं크들을 셋업하는데 이용되는, 장치.

청구항 19

주어진 개수의 계수 시퀀스들을 갖는 입력 HOA 신호의 프레임들을 인코딩하기 위한 장치로서 -각각의 계수 시퀀스는 인덱스를 가짐-

감소된 개수의 비제로 계수 시퀀스들을 갖는 절삭된 HOA 표현($C_T(k)$)을 계산하도록 구성되고, 또한 상기 절삭된 HOA 표현에 포함된 활성 계수 시퀀스들($I_{C,ACT}(k)$)의 인덱스들의 세트를 결정하도록 구성된 계산 및 결정 모듈;

상기 입력 HOA 신호를 복수의 주파수 부대역들(f_1, \dots, f_F)로 분할하도록 구성된 분석 필터 बैं크 모듈 -상기 주파수 부대역들의 계수 시퀀스들($\bar{C}(k-1, k, f_1), \dots, \bar{C}(k-1, k, f_F)$)이 획득됨-

상기 입력 HOA 신호로부터 제1 세트의 후보 방향들($M_{DIR}(k)$)을 추정하도록 구성되고, 또한 주파수 부대역들 각각에 대해 제2 세트의 방향들($M_{DIR}(k, f_1), \dots, M_{DIR}(k, f_F)$)을 추정하도록 구성된 방향 추정 모듈 -상기 제2 세트의 방향들의 각각의 요소는 제1 및 제2 인덱스를 갖는 인덱스들의 튜플이고, 상기 제2 인덱스는 현재 주파수 부대역에 대한 활성 방향의 인덱스이며, 상기 제1 인덱스는 상기 활성 방향의 궤적 인덱스이고, 각각의 활성 방향은 또한 상기 입력 HOA 신호의 상기 제1 세트의 후보 방향들($M_{DIR}(k)$)에 포함됨-

상기 주파수 부대역들 각각에 대해, 각각의 주파수 부대역의 상기 제2 세트의 방향들($M_{DIR}(k, f_1), \dots, M_{DIR}(k, f_F)$)에 따라 상기 주파수 부대역의 계수 시퀀스들($\bar{C}(k-1, k, f_1), \dots, \bar{C}(k-1, k, f_F)$)로부터 방향 부대역 신호들($\bar{X}(k-1, k, f_1), \dots, \bar{X}(k-1, k, f_F)$)을 계산하도록 구성된 적어도 하나의 방향 부대역 계산 모듈;

상기 주파수 부대역들 각각에 대해, 각각의 주파수 부대역의 활성 계수 시퀀스들($I_{C,ACT}(k)$)의 인덱스들의 세트를 이용하여 상기 주파수 부대역의 계수 시퀀스들($\bar{C}(k-1, k, f_1), \dots, \bar{C}(k-1, k, f_F)$)로부터, 방향 부대역 신호들($\bar{X}(k-1, k, f_1), \dots, \bar{X}(k-1, k, f_F)$)을 예측하기 위해 구성된 예측 행렬($A(k, f_1), \dots, A(k, f_F)$)을 계산하도록 구성된 적어도 하나의 방향 부대역 예측 모듈; 및

상기 제1 세트의 후보 방향들($M_{DIR}(k)$), 상기 제2 세트의 방향들($M_{DIR}(k, f_1), \dots, M_{DIR}(k, f_F)$), 상기 예측 행렬들($A(k, f_1), \dots, A(k, f_F)$) 및 상기 절삭된 HOA 표현($C_T(k)$)을 인코딩하도록 구성된 인코딩 모듈

을 포함하는 장치.

청구항 20

제19항에 있어서, 2개 이상의 부대역으로 이루어진 적어도 하나의 그룹이 생성되고, 상기 적어도 하나의 그룹은 단일 부대역 대신에 이용되고 단일 부대역과 동일한 방식으로 취급되는, 장치.

청구항 21

제19항에 있어서,

상기 절삭된 HOA 채널 시퀀스들을 부분적으로 무상관화하도록 구성된 부분적 무상관화기;

상기 절삭된 HOA 채널 시퀀스들($y_1(k), \dots, y_I(k)$)을 전송 채널들에 할당하도록 구성된 채널 할당 모듈; 및

상기 전송 채널들에 관한 이득 제어를 수행하도록 구성된 적어도 하나의 이득 제어 유닛 - 각각의 전송 채널에 대한 이득 제어 부가 정보($e_i(k-1), \beta_i(k-1)$)가 생성됨 -

을 더 포함하고,

상기 인코딩 모듈은,

상기 이득 제어된 절삭된 HOA 채널 시퀀스들($z_1(k), \dots, z_I(k)$)을 인코딩하도록 구성된 인지 인코더;

상기 이득 제어 부가 정보($e_i(k-1), \beta_i(k-1)$), 상기 제1 세트의 후보 방향들($M_{DIR}(k)$), 상기 제2 세트의 방향들($M_{DIR}(k, f_1), \dots, M_{DIR}(k, f_F)$), 및 상기 예측 행렬들($A(k, f_1), \dots, A(k, f_F)$)을 인코딩하도록 구성된 부가 정보 소스 코더; 및

인코딩된 HOA 신호 프레임($\bar{B}(k-1)$)을 획득하기 위해 상기 인지 인코더 및 상기 부가 정보 소스 코더의 출력들을 멀티플렉싱하도록 구성된 멀티플렉서를 포함하는, 장치.

청구항 22

제19항에 있어서, 상기 방향 추정 모듈은, 상기 주파수 부대역들 각각에 대해 상기 제2 세트의 방향들($M_{DIR}(k, f_1), \dots, M_{DIR}(k, f_F)$)을 추정할 때, 전체 대역 HOA 신호의 방향들($M_{DIR}(k)$) 중에서만 주파수 부대역의 방향들을 탐색하는, 장치.

청구항 23

제19항에 있어서, 활성 방향의 궤적을 결정하도록 구성된 궤적 결정 모듈을 더 포함하고, 상기 활성 방향은 사운드 소스의 방향이며, 궤적은 특정한 사운드 소스의 방향들의 시간적 시퀀스인, 장치.

청구항 24

제19항에 있어서, 절삭된 HOA 표현은 하나 이상의 계수 시퀀스들이 제로로 설정되는 HOA 신호인, 장치.

청구항 25

삭제

발명의 설명

기술 분야

[0001] 본 발명은 주어진 개수의 계수 시퀀스들을 갖는 입력 HOA 신호의 프레임들을 인코딩하기 위한 방법, HOA 신호를 디코딩하기 위한 방법, 주어진 개수의 계수 시퀀스들을 갖는 입력 HOA 신호의 프레임들을 인코딩하기 위한 장치, 및 HOA 신호를 디코딩하기 위한 장치에 관한 것이다.

배경 기술

[0002] HOA(Higher Order Ambisonics)는 음장 합성(WFS; wave field synthesis)이나 "22.2"로서 알려진 것과 유사한 채널 기반 접근법과 같은 기타의 기술들 중에서 3차원 사운드를 표현할 수 있는 한 가능성을 제공한다. 채널 기반 방법과는 대조적으로, HOA 표현은 특정한 확성기 셋업과는 독립적이라는 이점을 제공한다. 이러한 융통성은

특정한 확성기 셋업에 관한 HOA 표현의 재생에 요구되는 디코딩 프로세스를 댓가로 나오는 것이다. 요구되는 확성기의 수가 대개는 매우 큰 WFS 접근 방식에 비해, HOA는 단지 수 개의 확성기만으로 구성된 셋업으로 렌더링 될 수 있다. HOA의 추가 이점은, 헤드폰으로의 쌍이성 렌더링(binaural rendering)을 위한 임의의 수정없이 동일한 표현이 역시 사용될 수 있다는 것이다.

[0003] HOA는, 소위, 절삭된 구면 고조파(SH; Spherical Harmonics) 확장에 의한 복합 고조 평면파 진폭의 공간 밀도의 표현에 기초한다. 각각의 확장 계수는 시간 영역 함수로 등가적으로 표현될 수 있는 각주파수(angular frequency)의 함수이다. 따라서, 일반성을 잃지 않고, 완전한 HOA 음장 표현은 실제로 O 개의 시간 영역 함수들로 구성되는 것으로 이해될 수 있으며, 여기서, O 는 확장 계수의 개수를 나타낸다. 이들 시간 영역 함수는 이하에서는 등가적으로 HOA 계수 시퀀스 또는 HOA 채널이라고 지칭될 것이다.

[0004] 확장의 최대 차수 N 이 증가함에 따라 HOA 표현의 공간 해상도가 향상된다. 불행히도, 확장 계수의 개수 O 는 차수 N 에 따라 2차식으로(quadratically) 증가한다, 특히, $O = (N + 1)^2$ 이다. 차수 $N = 4$ 를 이용한 전형적인 HOA 표현은 $O = 25$ 개의 HOA (확장) 계수들을 요구한다. 상기 고려사항들에 따라, 원하는 단일-채널 샘플링 레이트 f_s 및 샘플 당 비트수 N_b 가 주어지면, HOA 표현의 전송을 위한 총 비트 레이트는 $O \cdot f_s \cdot N_b$ 에 의해 결정된다. 결과적으로, 예를 들어, 샘플링 레이트 $f_s = 48kHz$ 와 샘플당 $N_b = 16$ 비트를 이용하여 차수 $N = 4$ 의 HOA 표현을 전송하는 것은, $19.2Mbits/s$ 의 비트 레이트를 야기하고, 이것은, 예를 들어, 스트리밍 등의 많은 실제 응용에 대해 매우 높다. 따라서, HOA 표현의 압축이 매우 바람직하다.

[0005] HOA 음장 표현의 압축에 대한 다양한 접근법이 [4, 5, 6]에서 제안되었다. 이들 접근법들은, 음장 분석을 수행하고 주어진 HOA 표현을 방향 성분 및 잔여 환경 성분으로 분해한다는 공통점을 갖고 있다. 한편, 최종 압축된 표현은, 환경 HOA 성분의 관련 계수 시퀀스 뿐만 아니라 소위 방향 및 벡터-기반 신호의 인지 코딩(perceptual coding)으로부터 생기는 다수의 양자화된 신호를 포함한다. 또 한편으로는, 이것은, 압축된 버전으로부터의 HOA 표현의 재구성에 필요한 양자화된 신호와 관련된 추가적인 부가 정보를 포함한다.

[0006] 접근법들 [4, 5, 6]에 대한 양자화된 신호의 합리적인 최소 개수는 8이다. 따라서, 각각의 개개의 인지 코더에 대해 32 kbit/s의 데이터 레이트를 가정할 때, 이들 방법들 중 하나에서의 데이터 레이트는 통상적으로 256 kbit/s보다 낮지 않다. 예를 들어, 모바일 디바이스로의 오디오 스트리밍과 유사한 소정의 응용의 경우, 이러한 총 데이터 레이트가 너무 높을 수 있다. 따라서, 명백하게 더 낮은 데이터 레이트, 예를 들어, 128 kbit/s를 해결하기 위한 HOA 압축 방법에 대한 요구가 존재한다.

발명의 내용

해결하려는 과제

과제의 해결 수단

[0007] 음장의 HOA(Higher Order Ambisonics) 표현의 낮은 비트-레이트 압축을 위한 새로운 방법 및 장치가 개시된다.

[0008] 음장의 HOA 표현을 위한 낮은 비트 레이트 압축 방법의 한 주요 양태는, HOA 표현을 복수의 주파수 부대역들로 분해하고, 각각의 주파수 부대역(즉, 부-대역) 내의 계수들을, 절삭된 HOA 표현과 다수의 예측된 방향성 부대역 신호들에 기초하는 표현의 조합으로 근사화하는 것이다.

[0009] 절삭된 HOA 표현은 작은 개수의 선택된 계수 시퀀스들을 포함하고, 여기서 선택은 시간에 따라 변하는 것이 허용된다. 예를 들어, 모든 프레임마다 새로운 선택이 이루어진다. 절삭된 HOA 표현을 나타내는 선택된 계수 시퀀스들은 인지 코딩되고 최종 압축된 HOA 표현의 일부이다. 한 실시 예에서, 코딩 효율을 증가시키고 렌더링시 노이즈 언마스킹(unmasking) 효과를 감소시키기 위해, 선택된 계수 시퀀스들은 인지 코딩 이전에 무상관화(de-correlate)된다. 부분적 무상관화는 미리정의된 개수의 선택된 HOA 계수 시퀀스들에 공간 변환을 적용함으로써 달성된다. 압축해제의 경우, 무상관화는 재상관화에 의해 역전된다. 이러한 부분적인 무상관화의 큰 이점은 압

축해제시 무상관화를 되돌리기 위해 추가적인 부가 정보가 필요하지 않다는 것이다.

- [0010] 근사화된 HOA 표현의 다른 성분은 대응하는 방향들을 갖는 다수의 방향 부대역 신호들에 의해 표현된다. 이들은 절삭된 HOA 표현의 계수 시퀀스들로부터의 예측을 포함하는 파라메타 표현에 의해 코딩된다. 실시예에서, 각각의 방향 부대역 신호는 절삭된 HOA 표현의 계수 시퀀스들의 스케일링된 합에 의해 예측(또는 표현)되고, 여기서, 스케일링은 일반적으로 복소값이다. 압축해제를 위한 방향 부대역 신호들의 HOA 표현을 재합성할 수 있기 위하여, 압축된 표현은 복소값 예측 스케일링 계수들의 양자화된 버전들 뿐만 아니라 방향들의 양자화된 버전들을 포함한다.
- [0011] 한 실시예에서, 주어진 개수의 계수 시퀀스 - 각각의 계수 시퀀스는 인덱스를 가짐 - 를 갖는 입력 HOA 신호의 프레임들을 인코딩(및 그에 의해 압축)하기 위한 방법은,
- [0012] 절삭된 HOA 표현에 포함될 활성 계수 시퀀스들 $I_{C,ACT}(k)$ 의 인덱스들의 세트를 결정하는 단계,
- [0013] 감소된 개수의 비제로 계수 시퀀스들(즉, 입력 HOA 신호보다 적은 수의 비제로 계수 시퀀스들 및 그에 따라 더 많은 수의 제로 계수 시퀀스들)을 갖는 절삭된 HOA 표현 $C_T(k)$ 을 계산하는 단계,
- [0014] 입력 HOA 신호로부터 제1 세트의 후보 방향들 $M_{DIR}(k)$ 을 추정하는 단계,
- [0015] 입력 HOA 신호를 복수의 주파수 부대역으로 분할 - 여기서, 주파수 부대역들의 계수 시퀀스들 $\bar{C}(k-1, k, f_{1,...F})$ 이 획득됨 - 하는 단계,
- [0016] 주파수 부대역들 각각에 대해 제2 세트의 방향들 $M_{DIR}(k, f_1), \dots, M_{DIR}(k, f_F)$ 을 추정 - 제2 세트의 방향들의 각각의 요소는 제1 및 제2 인덱스를 갖는 인덱스들의 튜플이고, 제2 인덱스는 현재의 주파수 부대역에 대한 활성 방향의 인덱스이며 제1 인덱스는 활성 방향의 궤적 인덱스이고, 각각의 활성 방향은 또한 입력 HOA 신호의 제1 세트의 후보 방향들 $M_{DIR}(k)$ 에 포함됨(즉, 제2 세트의 방향들 중의 활성 부대역 방향들은 제1 세트의 전체 대역 방향들의 서브셋임) - 하는 단계,
- [0017] 주파수 부대역들 각각에 대해, 각각의 주파수 부대역의 제2 세트의 방향들 $M_{DIR}(k, f_1), \dots, M_{DIR}(k, f_F)$ 에 따라 주파수 부대역의 계수 시퀀스들 $\bar{C}(k-1, k, f_{1,...F})$ 로부터 방향 부대역 신호들 $\bar{X}(k-1, k, f_1), \dots, \bar{X}(k-1, k, f_F)$ 을 계산하는 단계,
- [0018] 주파수 부대역들 각각에 대해, 각각의 주파수 부대역의 활성 계수 시퀀스들 $I_{C,ACT}(k)$ 의 인덱스들의 세트를 이용하여 주파수 부대역의 계수 시퀀스들 $\bar{C}(k-1, k, f_{1,...F})$ 로부터 방향 부대역 신호 $\bar{X}(k-1, k, f_{1,...F})$ 를 예측하기 위해 구성된 예측 행렬 $A(k, f_1), \dots, A(k, f_F)$ 을 계산하는 단계, 및
- [0019] 제1 세트의 후보 방향들 $M_{DIR}(k)$, 제2 세트의 방향들 $M_{DIR}(k, f_1), \dots, M_{DIR}(k, f_F)$, 예측 행렬들 $A(k, f_1), \dots, A(k, f_F)$ 및 절삭된 HOA 표현 $C_T(k)$ 을 인코딩하는 단계를 포함한다.
- [0020] 제2 세트의 방향들은 주파수 부대역들에 관련된다. 제1 세트의 후보 방향들은 전체 주파수 대역과 관련된다. 유익하게는, 주파수 부대역들 각각에 대해 제2 세트의 방향들을 추정하는 단계에서, 주파수 부대역의 방향들 $M_{DIR}(k, f_1), \dots, M_{DIR}(k, f_F)$ 은, 제2 세트의 부대역 방향들은 제1 세트의 전체 대역 방향들의 서브셋이기 때문에, 전체 대역 HOA 신호의 방향들 $M_{DIR}(k)$ 중에서만 탐색될 필요가 있다. 한 실시예에서, 각각의 튜플 내의 제1 및 제2 인덱스의 순차적 순서는 스왑된다, 즉, 제1 인덱스는 현재 주파수 부대역에 대한 활성 방향의 인덱스이고, 제2 인덱스는 활성 방향의 궤적 인덱스이다.
- [0021] 완전한 HOA 신호는 복수의 계수 시퀀스 또는 계수 채널을 포함한다. 이들 계수 시퀀스들 중 하나 이상이 제로로 설정되는 HOA 신호는 여기서는 절삭된 HOA 표현이라고 불린다. 절삭된 HOA 표현을 계산하거나 생성하는 것은 일반적으로 제로로 설정되거나 설정되지 않는 계수 시퀀스들의 선택을 포함한다. 이 선택은, 다양한 기준에 따라, 예를 들어, 제로로 설정되지 않을 계수 시퀀스로서 최대 에너지를 포함하는 것들을, 또는 인지적으로 가장 관련 있는 것들을 선택하거나, 계수 시퀀스들을 임의로 선택함으로써, 이루어질 수 있다. HOA 신호를 주파수 부대역

들로 분할하는 것은, 예를 들어, 직교 거울 필터(QMF)를 포함하는 분석 필터 बैं크에 의해 수행될 수 있다.

- [0022] 한 실시예에서, 절삭된 HOA 표현 $C_T(k)$ 을 인코딩하는 단계는, 절삭된 HOA 채널 시퀀스들의 부분적 무상관화, (상관된 또는 무상관화된) 절삭된 HOA 채널 시퀀스들 $y_1(k), \dots, y_I(k)$ 을 전송 채널들에 할당하기 위한 채널 할당, 전송 채널들 각각에 관한 이득 제어를 수행 -여기서, 각각의 전송 채널에 대한 이득 제어 부가 정보가 생성됨- 하는 단계, 이득 제어된 절삭된 HOA 채널 시퀀스 $z_1(k), \dots, z_I(k)$ 를 인지 인코더에서 인코딩하는 단계, 이득 제어 부가 정보 $e_i(k-1), \beta_i(k-1)$, 제1 세트의 후보 방향들 $M_{DIR}(k)$, 제2 세트의 방향들 $M_{DIR}(k, f_1), \dots, M_{DIR}(k, f_F)$ 및 예측 행렬들 $A(k, f_1), \dots, A(k, f_F)$ 을 부가 정보 소스 인코더에서 인코딩하는 단계, 및 인지 인코더와 부가 정보 소스 코더의 출력들을 멀티플렉싱하여 인코딩된 HOA 신호 프레임 $\bar{B}(k-1)$ 을 획득하는 단계를 포함한다.
- [0023] 한 실시예에서, 컴퓨터 판독가능한 매체에는, 컴퓨터로 하여금, 입력 HOA 신호의 프레임들을 인코딩 또는 압축하기 위한 상기 방법을 수행하게 하는 실행가능한 명령어들이 저장되어 있다.
- [0024] 한 실시예에서, 주어진 개수의 계수 시퀀스들 -각각의 계수 시퀀스는 인덱스를 가짐- 을 갖는 입력 HOA 신호의 프레임들을 프레임별로 인코딩(및 그에 의해 압축)하기 위한 장치는, 프로세서와, 프로세서에서 실행될 때 입력 HOA 신호의 프레임들을 인코딩 또는 압축하기 위한 전송된 방법의 단계들을 수행하게 하는 소프트웨어 프로그램을 위한 메모리를 포함한다.
- [0025] 또한, 한 실시예에서, 압축된 HOA 표현을 디코딩(및 그에 의해 압축해제)하기 위한 방법은,
- [0026] 압축된 HOA 표현으로부터 복수의 절삭된 HOA 계수 시퀀스들 $\hat{z}_1(k), \dots, \hat{z}_I(k)$, 상기 절삭된 HOA 계수 시퀀스들의 시퀀스 인덱스들을 나타내는(또는 포함하는) 할당 벡터 $v_{AMBASSIGN}(k)$, 부대역 관련 방향 정보 $M_{DIR}(k+1, f_1), \dots, M_{DIR}(k+1, f_F)$, 복수의 예측 행렬 $A(k+1, f_1), \dots, A(k+1, f_F)$, 및 이득 제어 부가 정보 $e_1(k), \beta_1(k), \dots, e_I(k), \beta_I(k)$ 를 추출하는 단계,
- [0027] 복수의 절삭된 HOA 계수 시퀀스들 $\hat{z}_1(k), \dots, \hat{z}_I(k)$, 이득 제어 부가 정보 $e_1(k), \beta_1(k), \dots, e_I(k), \beta_I(k)$ 및 할당 벡터 $v_{AMBASSIGN}(k)$ 로부터, 절삭된 HOA 표현 $\hat{C}_T(k)$ 을 재구성하는 단계,
- [0028] 분석 필터 बैं크에서, 재구성된 절삭된 HOA 표현 $\hat{C}_T(k)$ 을 복수의 F개의 주파수 부대역에 대한 주파수 부대역 표현들 $\tilde{\hat{C}}_T(k, f_1), \dots, \tilde{\hat{C}}_T(k, f_F)$ 로 분해하는 단계,
- [0029] 방향 부대역 합성 블록에서, 주파수 부대역 표현들 각각에 대해, 재구성된 절삭된 HOA 표현의 각각의 주파수 부대역 표현 $\tilde{\hat{C}}_T(k, f_1), \dots, \tilde{\hat{C}}_T(k, f_F)$, 부대역 관련 방향 정보 $M_{DIR}(k+1, f_1), \dots, M_{DIR}(k+1, f_F)$, 및 예측 행렬들 $A(k+1, f_1), \dots, A(k+1, f_F)$ 로부터, 예측된 방향 HOA 표현 $\tilde{\hat{C}}_D(k, f_1), \dots, \tilde{\hat{C}}_D(k, f_F)$ 을 합성하는 단계,
- [0030] F개의 주파수 부대역들 각각에 대해, 계수 시퀀스가 할당 벡터 $v_{AMBASSIGN}(k)$ 에 포함되는(즉, 할당 벡터의 요소인) 인덱스 n을 갖는 경우 절삭된 HOA 표현 $\tilde{\hat{C}}_T(k, f_j)$ 의 계수 시퀀스들로부터 획득되거나, 그렇지 않으면 방향 부대역 합성 블록들 중 하나에 의해 제공된 예측된 방향 HOA 성분 $\tilde{\hat{C}}_D(k, f_j)$ 의 계수 시퀀스들로부터 획득되는 계수 시퀀스들 $\tilde{\hat{c}}_n(k, f_j)$, $n = 1, \dots, O$ 을 갖는 디코딩된 부대역 HOA 표현 $\tilde{\hat{c}}(k, f_1), \dots, \tilde{\hat{c}}(k, f_F)$ 을 조성하는 단계,
- [0031] 합성 필터 बैं크에서, 디코딩된 HOA 표현 $\tilde{\hat{c}}(k)$ 을 획득하기 위해 디코딩된 부대역 HOA 표현들 $\tilde{\hat{c}}(k, f_1), \dots, \tilde{\hat{c}}(k, f_F)$ 을 합성하는 단계를 포함한다.
- [0032] 한 실시예에서, 추출하는 단계는 인지 코딩된 부분 및 인코딩된 부가 정보 부분을 획득하기 위해 압축된 HOA 표

현을 디멀티플렉싱하는 단계를 포함한다. 한 실시예에서, 인지 코딩된 부분은 인지 인코딩된 절삭된 HOA 계수 시퀀스들 $\tilde{z}_1(k), \dots, \tilde{z}_I(k)$ 을 포함하고, 추출하는 단계는 인지 인코딩된 절삭된 HOA 계수 시퀀스들 $\tilde{z}_1(k), \dots, \tilde{z}_I(k)$ 을 인지 디코더에서 디코딩하여 절삭된 HOA 계수 시퀀스들 $\hat{z}_1(k), \dots, \hat{z}_I(k)$ 을 획득하는 단계를 포함한다. 한 실시예에서, 추출하는 단계는, 인코딩된 부가 정보 부분을 부가 정보 소스 디코더에서 디코딩하여, 부대역 관련 방향들의 세트 $M_{DIR}(k+1, f_1), \dots, M_{DIR}(k+1, f_F)$, 예측 행렬들 $A(k+1, f_1), \dots, A(k+1, f_F)$, 이득 제어 부가 정보 $e_1(k), \beta_1(k), \dots, e_I(k), \beta_I(k)$, 및 할당 벡터 $v_{AMBASSIGN}(k)$ 를 획득하는 단계를 포함한다.

[0033] 한 실시예에서, 컴퓨터 판독가능한 매체에는, 컴퓨터로 하여금 우세 방향 신호들의 방향들을 디코딩하기 위한 상기 방법을 수행하게 하는 실행가능한 명령어들이 저장되어 있다.

[0034] 한 실시예에서, 압축된 HOA 표현을 프레임별로 디코딩(및 그에 의해 압축해제)하기 위한 장치는, 프로세서와, 프로세서에서 실행될 때 입력 HOA 신호의 프레임들을 디코딩 또는 압축해제하기 위한 전술된 방법의 단계들을 수행하게 하는 소프트웨어 프로그램을 위한 메모리를 포함한다.

[0035] 한 실시예에서, HOA 신호를 디코딩하기 위한 장치는,

[0036] 디코딩될 HOA 신호 표현에 대한 최대 개수의 방향들 D의 인덱스들을 수신하도록 구성된 제1 모듈, 디코딩될 HOA 신호 표현의 최대 개수의 방향들 D의 방향들을 재구성하도록 구성된 제2 모듈, 부대역당 활성 방향 신호들의 인덱스들을 수신하도록 구성된 제3 모듈, 디코딩될 HOA 신호 표현의 재구성된 방향들 D로부터 부대역당 활성 방향 신호를 재구성하도록 구성된 제4 모듈, 및 부대역들의 방향 신호들을 예측하도록 구성된 제5 모듈을 포함하며, 여기서, 부대역의 현재 프레임에서 방향 신호를 예측하는 것은, 부대역의 선행 프레임의 방향 신호를 결정하는 것을 포함하고, 여기서, 방향 신호의 인덱스가 선행 프레임에서 제로이고 현재 프레임에서 비제로이면 새로운 방향 신호가 생성되고, 방향 신호의 인덱스가 선행 프레임에서 비제로이고 현재 프레임에서 제로이면 이전 방향 신호는 취소되고, 방향 신호의 인덱스가 제1 방향으로부터 제2 방향으로 변하면 방향 신호의 방향은 제1 방향으로부터 제2 방향으로 이동된다.

[0037] 부대역들은 일반적으로 복소값 필터 뱅크로부터 획득된다. 할당 벡터의 한 목적은, 전송된/수신된, 그에 따라 절삭된 HOA 표현에 포함되는 계수 시퀀스들의 시퀀스 인덱스들을 표시하여, 이들 계수 시퀀스들의 최종 HOA 신호로의 할당을 가능케하는 것이다. 바꾸어 말하면, 할당 벡터는, 절삭된 HOA 표현의 계수 시퀀스들 각각에 대해, 최종 HOA 신호의 어느 계수 시퀀스에 대응하는지를 나타낸다. 예를 들어, 절삭된 HOA 표현이 4개의 계수 시퀀스를 포함하고 최종 HOA 신호가 9개의 계수 시퀀스를 갖는다면, 할당 벡터는 (원칙적으로) [1,2,5,7]이 될 수 있으며, 따라서, 절삭된 HOA 표현의 제1, 제2, 제3, 및 제4 계수 시퀀스는, 실제로 최종 HOA 신호의 제1, 제2, 제5 및 제7 계수 시퀀스이다.

[0038] 본 발명의 추가적인 목적, 피쳐 및 이점들은, 첨부된 도면들과 연계한 이하의 설명과 첨부된 청구항들의 고려로부터 명백해질 것이다.

도면의 간단한 설명

[0039] 본 발명의 예시적 실시형태들이 첨부된 도면을 참조하여 설명되며, 여기서,

도 1은 공간 HOA 인코더의 아키텍처이다,

도 2는 방향 추정 블록의 아키텍처이다,

도 3은 인지 부가 정보 소스 인코더이다,

도 4는 인지 부가 정보 소스 디코더이다,

도 5는 공간 HOA 디코더의 아키텍처이다,

도 6은 구면 좌표계이다,

도 7은 방향 추정 처리 블록이다,

도 8은, 절삭된 HOA 표현의 방향, 궤적 인덱스 세트 및 계수들이다,

도 9는 MPEG에서 이용되는 종래의 오디오 인코더이다,

- 도 10은 MPEG에서 이용가능한 개선된 오디오 인코더이다,
- 도 11은 MPEG에서 이용되는 종래의 오디오 디코더이다,
- 도 12는 MPEG에서 이용가능한 개선된 오디오 디코더이다,
- 도 13은 인코딩 방법의 플로차트이다,
- 도 14는 디코딩 방법의 플로차트이다.

발명을 실시하기 위한 구체적인 내용

- [0040] 음장의 HOA 표현들을 위한 제안된 낮은-비트 레이트 압축 방법의 한 주요한 사상은, 2개의 부분들: 절삭된 HOA 표현과 다수의 예측된 방향 부대역 신호들에 기초한 표현의 조합에 의해, 원래의 HOA 표현을 프레임별로 및 부대역별로, 즉, 각각의 HOA 프레임의 개개의 주파수 부대역 내에서 근사화하는 것이다. HOA 기본에 대한 요약이 이하에서 더 제공된다.
- [0041] 근사화된 HOA 표현의 제1 부분은, 작은 개수의 선택된 계수 시퀀스들로 구성된 절삭된 HOA 버전이며, 여기서, 선택은 시간에 따라(예를 들어, 프레임별로) 달라지는 것이 허용된다. 그 다음, 절삭된 HOA 버전을 나타내는 선택된 계수 시퀀스들은 인지 코딩되고 최종 압축된 HOA 표현의 일부이다. 코딩 효율을 증가시키고 렌더링시 노이즈 언마스킹(unmasking) 효과를 감소시키기 위하여, 인지 코딩 이전에 선택된 계수 시퀀스들을 무상관화하는 것이 유의하다. 부분적인 무상관화는, 미리정의된 개수의 선택된 HOA 계수 시퀀스들에 공간 변환을 적용함으로써 달성되며, 이것은 주어진 개수의 가상 확장기 신호들로의 렌더링을 의미한다. 이러한 부분적인 무상관화의 큰 이점은 압축해제시 무상관화를 되돌리기 위해 추가적인 부가 정보가 필요하지 않다는 것이다.
- [0042] 근사화된 HOA 표현의 제2 부분은 대응하는 방향들을 갖는 다수의 방향 부대역 신호들에 의해 표현된다. 그러나, 이들은 통상적으로 코딩되지 않는다. 대신에, 이들은, 제1 부분, 즉, 절삭된 HOA 표현의 계수 시퀀스들로부터의 예측에 의한 파라미터 표현으로서 코딩된다. 특히, 실시예에서, 각각의 방향 부대역 신호는 절삭된 HOA 표현의 계수 시퀀스들의 스케일링된 합에 의해 예측되고, 여기서, 스케일링은 선형이고 일반적으로 복소값이다. 양쪽 부분들은 함께 HOA 신호의 압축된 표현을 형성하고, 그에 따라, 낮은 비트 레이트를 달성한다. 압축해제를 위한 방향 부대역 신호들의 HOA 표현을 재합성할 수 있기 위하여, 압축된 표현은 복소값 예측 스케일링 계수들의 양자화된 버전들 뿐만 아니라 방향들의 양자화된 버전들을 포함한다.
- [0043] 이 맥락에서 특히 중요한 양태들은, 방향 및 복소값 예측 스케일링 계수의 계산, 및 이들을 효율적으로 코딩하는 방법이다.
- [0044] 낮은 비트 레이트 HOA 압축
- [0045] 제안된 낮은 비트 레이트 HOA 압축의 경우, 낮은 비트 레이트 HOA 압축기는, 공간 HOA 인코딩부와, 인지 및 소스 인코딩부로 세분될 수 있다. 공간적 HOA 인코딩부의 예시적인 아키텍처가 도 1에 도시되어 있고, 인지 및 소스 인코딩부분 예시적인 아키텍처가 도 3에 도시되어 있다. 공간적 HOA 인코더(10)는 I 개의 신호들을 포함하는 제1 압축된 HOA 표현을 그 HOA 표현을 생성하는 방법을 기술하는 부가 정보와 함께 제공한다. 인지 및 부가 정보 소스 코더(30)에서, 이들 I 개 신호들은 인지 코더(31)에서 인지 인코딩되고, 부가 정보는 부가 정보 소스 코더(32)에서 소스 인코딩된다. 부가 정보 소스 코더(32)는 코딩된 부가 정보 \bar{F} 를 제공한다. 그 다음, 인지 코더(31) 및 부가 정보 소스 코더(32)에 의해 제공된 2개의 코딩된 표현은 멀티플렉서(33)에서 멀티플렉싱되어 낮은 비트 레이트의 압축된 HOA 데이터 스트림 \bar{B} 를 획득한다.
- [0046] 공간 HOA 인코딩
- [0047] 도 1에 도시된 공간 HOA 인코더는 프레임별 처리를 수행한다. 프레임들은, O 개의 시간-연속적인 HOA 계수 시퀀스들의 부분들로서 정의된다. 예를 들어, 인코딩될 입력 HOA 표현의 k 번째 프레임 $C(k)$ 는 시간-연속적인 HOA 계수 시퀀스들의 벡터 $c(t)$ (수학식 46 참조)에 관해 다음과 같이 정의된다:

수학식 1

$$C(k) := [c((kL+1)T_s) \quad c((kL+2)T_s) \quad \dots \quad c((k+1)LT_s)] \in \mathbb{R}^{O \times L}$$

[0048]

여기서, k 는 프레임 인덱스를 나타내고, L 은 프레임 길이(샘플 단위)를 나타내며, $O = (N+1)^2$ 은 HOA 계수 시퀀스들의 개수를 나타내고, T_s 는 샘플링 기간을 나타낸다.

[0049]

절삭된 HOA 표현의 계산

[0050]

도 1에 도시된 바와 같이, 절삭된 HOA 표현을 계산하는 제1 단계는 원래의 HOA 프레임 $C(k)$ 로부터 절삭된 버전 $C_T(k)$ 을 계산하는 단계(11)를 포함한다. 이 맥락에서의 절삭이란, 입력 HOA 표현의 O 개의 계수 시퀀스들 중 I 개의 특정한 계수 시퀀스들의 선택 및 모든 다른 계수 시퀀스들을 제로로 설정하는 것을 의미한다. 계수 시퀀스들의 선택을 위한 다양한 솔루션이 [4,5,6]에서 알려져 있고, 예를 들어, 인간의 인지와 관련하여 최대의 파워(power) 또는 가장 높은 관련성을 갖는 것들이 해당된다. 선택된 계수 시퀀스들은 절삭된 HOA 버전을 나타낸다. 선택된 계수 시퀀스들의 인덱스들을 포함하는 데이터 세트 $J_{C,ACT}(k)$ 가 생성된다. 그 다음, 이하에서 더

[0051]

설명되는 바와 같이, 절삭된 HOA 버전 $C_T(k)$ 은 부분적으로 무상관화되고(12), 부분적으로 무상관화된 절삭된 HOA 버전 $C_1(k)$ 은, 선택된 계수 시퀀스가 이용가능한 I 개의 전송 채널에 할당되는 채널 할당 (13)을 거칠 것이다. 이하에서 더 설명되는 바와 같이, 이들 계수 시퀀스들은 인지 인코딩되고(30) 최종적으로 압축된 표현의 일부가 된다. 채널 할당 후에 인지 인코딩을 위한 평활 신호들을 얻기 위해, k 번째 프레임에서 선택되지만 $(k+1)$ 번째 프레임에서 선택되지 않는 계수 시퀀스들이 결정된다. 한 프레임에서 선택되고 다음 프레임에서 선택되지 않는 이들 계수 시퀀스들은 페이드 아웃(fade out)된다. 그들의 인덱스들은, $J_{C,ACT}(k)$ 의 서브세트인, 데이터 세트 $J_{C,ACT,OUT}(k)$ 에 포함된다. 마찬가지로 k 번째 프레임에서 선택되지만 $(k-1)$ 번째 프레임에서 선택되지 않은 계수 시퀀스들은 페이드 인(fade in)된다. 그들의 인덱스들은, 역시 $J_{C,ACT}(k)$ 의 서브세트인, 세트 $J_{C,ACT,IN}(k)$ 에 포함된다. 페이딩을 위해, (이하의 수학식 39에서 도입되는 것과 같은) 윈도우 함수 $w_{OA}(l)$, $l = 1, \dots, 2L$ 이 이용될 수 있다.

[0052]

결론적으로, 다음과 같이, 절삭된 버전 $C_T(k)$ 의 HOA 프레임 k 가 각각의 O 개의 개개의 계수 시퀀스 프레임들의 L 개의 샘플로 구성된다면,

수학식 2

$$C_T(k) = \begin{bmatrix} c_{T,1}(k,1) & \dots & c_{T,1}(k,L) \\ c_{T,2}(k,1) & \dots & c_{T,2}(k,L) \\ \vdots & \ddots & \vdots \\ c_{T,O}(k,1) & \dots & c_{T,O}(k,L) \end{bmatrix}$$

[0053]

이 절삭은 계수 시퀀스 인덱스들 $n = 1, \dots, O$ 과 샘플 인덱스들 $l = 1, \dots, L$ 에 대해 다음과 같이 표현할 수 있다:

[0054]

수학식 3

$$c_{T,n}(k) = \begin{cases} c_n(k,l) \cdot w_{OA}(l) & \text{if } n \in J_{CACTIN}(k) \\ c_n(k,l) \cdot w_{OA}(L+1) & \text{if } n \in J_{CACTOUT}(k) \\ c_n(k,l) & \text{if } n \in J_{CACT}(k) \setminus (J_{CACTIN}(k) \cup J_{CACTOUT}(k)) \\ 0 & \text{else} \end{cases}$$

[0055]

[0056]

계수 시퀀스의 선택 기준에는 수 개의 가능성이 있다. 예를 들어, 하나의 유익한 해결책은 대부분의 신호 전력을 나타내는 계수 시퀀스들을 선택하는 것이다. 또 다른 유익한 해결책은 인간의 인지와 관련하여 가장 관련성이 큰 계수 시퀀스들을 선택하는 것이다. 후자의 경우에, 관련성은, 예를 들면, 상이하게 절삭된 표현을 가상 확장기 신호로 렌더링하고, 원래의 HOA 표현에 대응하는 이들 신호와 확장기 신호 간의 에러를 결정하고, 사운드 마스킹 효과를 고려하여 최종적으로 에러의 관련성을 해석함으로써, 결정될 수 있다.

[0057]

한 실시예에서, 세트 $J_{CACT}(k)$ 내의 인덱스를 선택하기 위한 합리적인 전략은, 항상 제1 O_{MIN} 인덱스들, $1, \dots, O_{MIN}$ 을 선택하는 것으로, 여기서 $O_{MIN} = (N_{MIN} + 1)^2 \leq I$ 이고 N_{MIN} 은 절삭된 HOA 표현의 주어진 최소 전체 차수를 나타낸다. 그 다음, 위에서 언급된 기준 중 하나에 따라 세트 $\{O_{MIN}+1, \dots, O_{MAX}\}$ 에서 나머지 $I - O_{MIN}$ 개의 인덱스들을 선택한다, 여기서, $O_{MAX} = (N_{MAX} + 1)^2 \leq O$ 이고 N_{MAX} 는 선택을 위해 고려되는 HOA 계수 시퀀스들의 최대 차수를 나타낸다. O_{MAX} 는 샘플당 전송가능한 계수들의 최대 개수로서 계수들의 총 개수 O 보다 작거나 같다는 점에 유의한다. 이 전략에 따르면, 절삭 처리 블록(11)은 또한 소위 할당 벡터 $\mathbf{v}_A(k) \in \mathbb{N}^{I-O_{MIN}}$ 를 제공하고, 그 요소들 $v_{A,i}(k)$, $i = 1, \dots, I - O_{MIN}$ 은 하기에 따라 설정된다

[0058]

$$v_{A,i}(k) = n$$

[0059]

여기서, n ($n \geq O_{MIN} + 1$)은, 나중에 i 번째 전송 신호 $\mathbf{y}_i(k)$ 에 할당되는, 추가로 선택된 HOA 계수 시퀀스 $\mathbf{c}(k)$ 의 HOA 계수 시퀀스 인덱스를 나타낸다. $\mathbf{y}_i(k)$ 의 정의는 이하의 수학식 10에서 주어진다. $\mathbf{c}_T(k)$ 의 처음 O_{MIN} 개의 행들은 디폴트로 HOA 계수 시퀀스들 $1, \dots, O_{MIN}$ 을 포함하고, $\mathbf{c}_T(k)$ 의 그 다음 $O - O_{MIN}$ (또는, $O = O_{MAX}$ 이면, $O_{MAX} - O_{MIN}$)개의 행들 중에는, 그 인덱스들이 할당 벡터 $\mathbf{v}_A(k)$ 에 저장되어 있는 프레임별로 변하는 HOA 계수 시퀀스들을 포함하는 $I - O_{MIN}$ 개의 행들이 있다. 마지막으로, $\mathbf{c}_T(k)$ 의 나머지 행들은 제로를 포함한다. 결과적으로, 이하에서 설명되는 바와 같이, 이용가능한 I 개의 전송 신호들의 처음(또는, 수학식 10에서와 같이, 마지막) O_{MIN} 개는 디폴트로 HOA 계수 시퀀스들 $1, \dots, O_{MIN}$ 에 할당되고, 나머지 $I - O_{MIN}$ 개의 전송 신호들은, 그 인덱스들이 할당 벡터 $\mathbf{v}_A(k)$ 에 저장되어 있는 프레임별로 변화하는 HOA 계수 시퀀스들에 할당된다.

[0060]

부분적 무상관화

[0061]

제2 단계에서, 후속 인지 인코딩의 효율성을 증가시키고, 렌더링시에 선택된 HOA 계수 시퀀스를 행렬화한 후에 발생하는 코딩 노이즈 언마스킹을 피하기 위하여, 선택된 HOA 계수 시퀀스들의 부분적 무상관화(12)가

실행된다. 예시적인 부분 무상관화(12)는 처음 O_{MIN} 개의 선택된 HOA 계수 시퀀스들에 공간 변환을 적용함으로써 달성되며, 이것은 O_{MIN} 개의 가상 확산기 신호들의 렌더링을 의미한다. 각각의 가상 확산기 위치는 도 6에 도시된 구면 좌표계에 의해 표현되며, 여기서 각각의 위치는 단위 구면 상에 놓여 있다고 가정된다. 즉, 반경 1을 갖는다. 따라서, 위치는 방향 $\Omega_j = (\theta_j, \phi_j)$ 로 등가적으로 표현될 수 있고, 여기서, $1 \leq j \leq O_{\text{MIN}}$ 이며, θ_j 및 ϕ_j 는, 각각 경사각과 방위각을 나타낸다(구 좌표계의 정의에 대해서는 아래를 더 참조). 이들 방향은 가능한한 균일하게 단위 구면 상에 분산되어야 한다(예를 들어, 특정한 방향의 계산에 관한 [2] 참조). HOA는 일반적으로 N_{MIN} 에 의존하여 방향을 정의하기 때문에, 실제로 $\Omega_j^{(N_{\text{MIN}})}$ 을 의미하며, 여기서, Ω_j 가 기재된다는 점에 유의한다.

[0062] 이하에서는, 모든 가상 확산기 신호들의 프레임이 다음과 같이 표기된다

수학식 5

$$W(k) = \begin{bmatrix} w_1(k) \\ w_2(k) \\ \vdots \\ w_{O_{\text{MIN}}}(k) \end{bmatrix}$$

[0063]

여기서, $w_j(k)$ 는 j번째 가상 확산기의 k번째 프레임을 나타낸다. 또한, Ψ_{MIN} 은 가상 방향들 $\Omega_j, 1 \leq j \leq O_{\text{MIN}}$ 에 관한 모드 행렬(mode matrix)을 나타낸다. 모드 행렬은 다음과 같이 정의된다,

[0064]

수학식 6

$$\Psi_{\text{MIN}} := [S_{\text{MIN},1} \ \dots \ S_{\text{MIN},O_{\text{MIN}}}] \in \mathbb{R}^{O_{\text{MIN}} \times O_{\text{MIN}}}$$

[0065]

여기서,

[0066]

수학식 7

$$S_{\text{MIN},i} := [S_0^0(\Omega_i) \ S_1^{-1}(\Omega_i) \ S_1^0(\Omega_i) \ S_1^1(\Omega_i) \ \dots \ S_N^{N-1}(\Omega_i) \ S_N^N(\Omega_i)] \in \mathbb{R}^{O_{\text{MIN}}}$$

[0067]

상기 식은 가상 방향 Ω_i 에 관한 모드 벡터를 나타낸다. 그 요소들 $S_n^m(\cdot)$ 각각은 이하에서 정의된 실수값 구면 고조파 함수를 나타낸다(수학식 48 참조).

[0068]

이 표기를 이용하여, 렌더링 프로세스는 다음과 같은 행렬 곱셈에 의해 공식화될 수 있다,

[0069]

수학식 8

$$W(k) = (\Psi_{\text{MIN}})^{-1} \cdot \begin{bmatrix} c_1(k) \\ \vdots \\ c_{O_{\text{MIN}}}(k) \end{bmatrix}$$

[0070]

[0071] 따라서, 부분적 무상관화(12)의 출력인 중간 표현 $\mathbf{c}_1(k)$ 의 신호들은 다음과 같이 주어진다,

수학식 9

$$\mathbf{c}_{1,n}(k) = \begin{cases} \mathbf{w}_n(k) & \text{if } 1 \leq n \leq O_{\text{MIN}} \\ \mathbf{c}_{T,n}(k) & O_{\text{MIN}} + 1 \leq n \leq O \end{cases}$$

[0072]

[0073] 채널 할당

[0074] 중간 표현 $\mathbf{c}_1(k)$ 의 프레임을 계산한 후, 그 개개의 신호 $\mathbf{c}_{1,n}(k)$, $n \in J_{\text{C,ACT}}(k)$ 는, 인지 인코딩을 위한 전송 신호들 $\mathbf{y}_i(k)$, $i = 1, \dots, I$ 를 제공하기 위해 이용가능한 I 개의 채널들에 할당된다(13). 할당(13)의 한 목적은, 연속된 프레임들 사이에서 선택이 변경되는 경우에 발생할 수 있는, 인지 인코딩될 신호들의 불연속성을 회피하는 것이다. 할당은 다음과 같이 표현될 수 있다,

수학식 10

$$\mathbf{y}_i(k) = \begin{cases} \mathbf{c}_{1,WA,i}(k) & \text{if } 1 \leq i \leq I - O_{\text{MIN}} \\ \mathbf{c}_{1,i-(I-O_{\text{MIN}})}(k) & \text{if } I - O_{\text{MIN}} < i \leq I \end{cases}$$

[0075]

[0076] 이득 제어

[0077] 전송 신호들 $\mathbf{y}_i(k)$ 각각은 최종적으로 이득 제어 유닛(14)에 의해 처리되며, 여기서, 신호 이득은 인지 인코더에 적합한 값 범위를 달성하도록 매끄럽게 수정된다. 이득 수정은 연속적인 블록들 사이에서의 심각한 이득 변화를 피하기 위해 일종의 룩-어헤드(look-ahead)를 요구하며, 그에 따라, 한 프레임의 지연을 도입한다. 각각의 전송 신호 프레임 $\mathbf{y}_i(k)$ 에 대해, 이득 제어 유닛(14)은 지연된 프레임 $\mathbf{y}_i(k-1)$, $i = 1, \dots, I$ 을 수신하거나 생성한다. 이득 제어 후의 수정된 신호 프레임은, $\mathbf{z}_i(k-1)$, $i = 1, \dots, I$ 로 표기된다. 또한, 공간 디코더에서 이루어진 임의의 변경을 되돌릴 수 있기 위하여, 이득 제어 부가 정보가 제공된다. 이득 제어 부가 정보는 지수 $\mathbf{e}_i(k-1)$ 및 예외 플래그 $\beta_i(k-1)$, $i = 1, \dots, I$ 를 포함한다. 이득 제어의 더 상세한 설명은, 예를 들어, [9], Sect.C.5.2.5 또는 [3]에서 이용가능하다. 따라서, 절삭된 HOA 버전(19)은 이득 제어된 신호 프레임들 $\mathbf{z}_i(k-1)$ 및 이득 제어 부가 정보 $\mathbf{e}_i(k-1)$, $\beta_i(k-1)$, $i = 1, \dots, I$ 를 포함한다.

[0078] 분석 필터 बैं크

[0079] 전술된 바와 같이, 근사화된 HOA 표현은, 2개의 부분, 즉, 절삭된 HOA 버전(19), 및 절삭된 HOA 표현의 계수 시퀀스들로부터 예측되는 대응하는 방향들을 갖는 방향 부대역 신호들에 의해 표현되는 성분으로 구성된다. 따라서, 제2 부분의 파라메타 표현을 계산하기 위해, 원래의 HOA 표현 $\mathbf{c}_n(k)$, $n = 1, \dots, O$ 의 개개의 계수 시퀀스의 각각의 프레임은, 먼저, 개개의 부대역 신호들 $\tilde{\mathbf{c}}_n(k, f_1), \dots, \tilde{\mathbf{c}}_n(k, f_F)$ 의 프레임들로 분해된다. 이것은 하나 이상의 분석 필터 बैं크(15)에서 이루어진다. 각각의 부대역 f_j , $j = 1, \dots, F$ 에 대해, 개개의 HOA 계수 시퀀스들의 부대역 신호들의 프레임들은 하기의 부대역 HOA 표현으로 집합될 수 있다,

수학식 11

$$\tilde{c}(k, f_j) = \begin{bmatrix} \tilde{c}_1(k, f_j) \\ \tilde{c}_2(k, f_j) \\ \vdots \\ \tilde{c}_O(k, f_j) \end{bmatrix} \quad \text{for } j = 1, \dots, F$$

[0080]

[0081]

분석 필터 बैं크(15)는 방향 부대역 신호 계산을 위해 방향 추정 처리 블록(16) 및 하나 이상의 계산 블록(17)에 부대역 HOA 표현을 제공한다.

[0082]

원칙적으로, 임의 유형의 필터(즉, 임의의 복소값 필터 बैं크, 예를 들어, QMF, FFT)가 분석 필터 बैं크(15)에서 이용될 수 있다. 분석 및 대응하는 합성 필터 बैं크의 연속적인 적용이 완벽한 재구성 속성으로 알려진 지연 신원(delay identity)을 제공할 것이 요구되지는 않는다. HOA 계수 시퀀스들 $c_n(k)$ 와는 대조적으로, 그들의 부대역 표현 $\tilde{c}_n(k, f_j)$ 은 일반적으로 복소값이라는 점에 유의한다. 또한, 부대역 신호들 $\tilde{c}_n(k, f_j)$ 은 원래의 시간 영역 신호들과 비교해 볼 때 일반적으로 시간적으로 데시메이트(decimate)된다. 결과적으로, 프레임들 $\tilde{c}_n(k, f_j)$ 내의 샘플수는 일반적으로, L인, 시간 영역 신호 프레임들 $c_n(k)$ 내의 샘플수보다 분명히 작다.

[0083]

한 실시예에서, 인간 청각 시스템의 속성에 맞게 처리를 더 양호하게 적용시키기 위해, 2개 이상의 부대역 신호가 부대역 신호 그룹들로 결합된다. 각각의 그룹의 대역폭들은, 예를 들어, 그 부대역 신호들의 개수에 의해 널리 공지된 바크 스케일(Bark scale)로 적합화될 수 있다. 즉, 더 높은 주파수에서 특히, 2개 이상의 그룹들이 하나로 결합될 수 있다. 이 경우, 각각의 부대역 그룹은, 추출된 파라미터의 수가 단일의 부대역의 경우와 동일한 HOA 계수 시퀀스들의 세트 $\tilde{c}_n(k, f_j)$ 로 구성된다라는 점에 유의한다. 한 실시예에서, 그룹화는, 분석 필터 बैं크 블록(15)에 통합될 수 있는 하나 이상의 부대역 신호 그룹화 유닛(명시 적으로 도시되지 않음)에서 수행된다.

[0084]

방향 추정

[0085]

방향 추정 처리 블록(16)은, 입력 HOA 표현을 분석하고 각각의 주파수 부대역 $f_j, j = 1, \dots, F$ 에 대해, 즉, 음장에 주요한 기여를 하는 부대역 일반 평면파의 방향들의 함수 세트 $\mathcal{M}_{\text{DIR}}(k, f_j)$ 를 계산한다. 이 맥락에서, "주요한 기여"라는 용어는 예를 들어 다른 방향들로부터 입사하는 부대역 일반 평면파의 신호 전력보다 높은 신호 전력을 지칭한다. 이것은 또한, 인간의 인지의 측면에서의 높은 관련성을 지칭할 수도 있다. 부대역 그룹화가 이용되는 경우, $\mathcal{M}_{\text{DIR}}(k, f_j)$ 의 계산을 위해, 단일의 부대역 대신에 부대역 그룹이 이용될 수 있다는 점에 유의한다.

[0086]

압축해제 동안에, 예측된 방향 부대역 신호들 내의 아티팩트는, 연속하는 프레임들 사이의 추정된 방향들과 예측 계수들의 변화로 인해 발생할 수 있다. 이러한 아티팩트를 피하기 위하여, 인코딩 동안 방향 부대역 신호들의 방향 추정 및 예측은 연결된 긴 프레임에 관해 수행된다. 연결된 긴 프레임은 현재 프레임과 그 선행물로 구성된다. 압축해제를 위해, 이들 긴 프레임들에 관해 추정된 양은 예측된 방향 부대역 신호들과의 중첩 계산 처리를 수행하는데 이용된다.

[0087]

방향 추정을 위한 직접적인 접근법은 각 부대역을 별개로 처리하는 것이다. 한 실시예에서, 방향 탐색을 위해, 예를 들어, [7]에서 제안된 기술이 적용될 수 있다. 이 접근법은, 각각의 개개의 부대역에 대해, 방향 추정의 부드러운 시간적 궤적을 제공하고 급격한 방향 변화나 시작을 포착할 수 있다. 그러나, 이 공지된 접근법에는 2가지 단점이 있다.

[0088]

첫째, 각각의 부대역에서의 독립적인 방향 추정은, 전체-대역 일반 평면파의 존재시에(예를 들어, 소정 방향의 일시적인 드림 박동), 개개의 하위 방향들에서의 추정 오류가 하나의 단일 방향으로부터의 원하는 전체-대역 버전으로 합해지지 않는 상이한 방향들로부터의 부대역 일반 평면파들로 이어질 수 있다는 바람직하지 않는 효과

를 초래할 수 있다. 특히 소정 방향들로부터의 일시적인 신호들이 흐려진다.

[0089] 둘 째, 낮은 비트 레이트 압축을 얻으려는 의도를 고려할 때, 부가 정보로부터 생기는 총 비트 레이트를 반드시 염두에 두어야 한다. 이하에서, 한 예는, 이러한 단순 접근 방식에 대한 비트 레이트가 다소 높다는 것을 보여 줄 것이다. 예시적으로, 부대역의 개수 F 는 10인 것으로 가정되고 (각각의 세트 $\mathcal{M}_{\text{DIR}}(k, f_j)$ 내의 요소들의 개수에 대응하는) 각각의 부대역의 방향들의 개수는 4인 것으로 가정된다. 또한, [9]에서 제안된 바와 같이, 각각의 부대역에 대해 $Q = 900$ 개의 잠재적인 방향 후보들의 그리드 상에서 탐색을 수행하는 것으로 가정된다. 이것은 단일 방향의 단순 코딩을 위해 $\lceil \log_2(Q) \rceil = 10$ 비트를 요구한다. 초당 약 50 프레임의 프레임 레이트를 가정하면, 결과적인 전체 데이터 레이트는, 방향들의 코딩된 표현에 대해서만 다음과 같다.

[0090]
$$10 \frac{\text{비트}}{\text{방향}} \cdot 4 \frac{\text{방향}}{\text{대역}} \cdot 10 \frac{\text{대역}}{\text{프레임}} \cdot 50 \frac{\text{프레임}}{\text{s}} = 20 \text{ kbit/s}$$

[0091] 초당 25 프레임의 프레임 속도를 가정하더라도, 10 kbit/s의 결과적인 데이터 레이트는 여전히 다소 높다.

[0092] 한 개선으로서, 한 실시예에서, 방향 추정 블록(20)에서 다음과 같은 방향 추정을 위한 방법이 이용된다. 개괄적인 사상은 도 2에 도시되어 있다.

[0093] 제1 단계에서, 전체-대역 방향 추정 블록(21)은, 연결된 긴 프레임을 이용하여, Q 개의 테스트 방향들 $\Omega_{\text{TEST},q}$, $q = 1, \dots, Q$ 로 구성된 방향 그리드상에서, 예비 전체-대역 방향 추정 또는 탐색을 수행한다.

수학식 12

[0094]
$$\bar{\mathbf{C}}(k-1; k) = [\mathbf{C}(k-1) \quad \mathbf{C}(k)]$$

[0095] 여기서, $\mathbf{C}(k)$ 와 $\mathbf{C}(k-1)$ 은 전체-대역의 원래의 HOA 표현의 현재 및 이전 입력 프레임들이다. 이 방향 탐색은, 세트 $\mathcal{M}_{\text{DIR}}(k)$ 에 포함되는, 다수의 $D(k) \leq D$ 개의 방향 후보들 $\Omega_{\text{CAND},d}(k)$, $d = 1, \dots, D(k)$ 을 제공한다, 즉,

수학식 13

[0096]
$$\mathcal{M}_{\text{DIR}}(k) = \{\Omega_{\text{CAND},1}(k), \dots, \Omega_{\text{CAND},D(k)}(k)\}.$$

[0097] 프레임당 방향 후보들의 최대 개수에 대한 전형적인 값은 $D = 16$ 이다. 방향 추정은, 예를 들어, 방향들의 베이시안 추론을 위한 간단한 소스 이동 모델로 입력 HOA 표현의 방향 파워 분포로부터 얻은 정보를 결합하는 사상의 [7]에서 제안된 방법에 의해 달성될 수 있다.

[0098] 제2 단계에서, 방향 탐색은, 부대역(또는 부대역 그룹)마다 부대역 방향 추정 블록(22)에 의해 각각의 개개의 부대역에 대해 수행된다. 그러나, 부대역에 대한 이러한 방향 탐색은 Q 개의 테스트 방향들로 구성되는 초기의 전체 방향 그리드를 고려할 필요가 없고, 단지 각 부대역에 대한 $D(k)$ 개의 방향들만을 포함하는 후보 세트 $\mathcal{M}_{\text{DIR}}(k)$ 만을 고려할 필요가 있다. $D_{\text{SB}}(k, f_j)$ 로 표기되는, f_j 번째 부대역, $j = 1, \dots, F$ 에 대한 방향들의 수는, 통상적으로 D 보다 작은, D_{SB} , 예를 들어, $D_{\text{SB}} = 4$ 보다 크지 않다. 전체-대역 방향 탐색과 마찬가지로, 부대역 관련 방향 탐색은 또한, 현재 및 이전 프레임들로 구성된, 부대역 신호들의 긴 연결된 프레임들에 관해 수행된다,

수학식 14

$$\bar{\mathcal{C}}(k-1; k; f_j) = [\bar{\mathcal{C}}(k-1, f_j) \quad \bar{\mathcal{C}}(k, f_j)] \quad j = 1, \dots, F$$

[0099]

[0100] 원칙적으로, 전체-대역 관련 방향 탐색의 경우와 동일한 베이지안 추론 방법이 부대역 관련 방향 탐색에 적용될 수 있다.

[0101] 특정한 사운드 소스의 방향은 시간에 따라 변할 수도 있다(그러나, 변할 필요는 없다). 특정한 사운드 소스의 시간적 시퀀스는 여기서는 "궤적(trajjectory)"이라고 부른다. 각각의 부대역 관련 방향, 또는 궤적은, 각각, 명확한 인덱스를 가지며, 이것은 상이한 궤적들을 혼합하는 것을 방지하고 연속적인 방향 부대역 신호를 제공한다. 이것은 후술되는 방향 부대역 신호들의 예측에 중요하다. 특히, 이것은 이하에서 더 정의되는 연속적

인 예측 계수 행렬들 $A(k, f_j)$ 간의 시간 의존성을 활용하는 것을 허용한다. 따라서, f_j 번째 부대역에 대한 방향 추정은 튜플 세트 $\mathcal{M}_{\text{DIR}}(k, f_j)$ 를 제공한다. 각각의 튜플은, 한편으로는, 개개의 (활성) 방향 궤적을 식별하는 인덱스 $d \in \mathcal{J}_{\text{DIR}}(k, f_j) \subseteq \{1, \dots, D_{\text{SB}}\}$ 로 구성되고, 다른 한편으로는, 각각의 추정된 방향 $\Omega_{\text{SB},d}(k, f_j)$ 으로 구성된다, 즉,

수학식 15

$$\mathcal{M}_{\text{DIR}}(k, f_j) = \{(d, \Omega_{\text{SB},d}(k, f_j)) \mid d \in \mathcal{J}_{\text{DIR}}(k, f_j)\}.$$

[0102]

[0103] 정의에 따르면, 부대역 방향 탐색은, 전술된 바와 같이, 현재 프레임의 방향 후보들 $\Omega_{\text{CAND},d}(k), d = 1, \dots, D(k)$ 중에서만 수행되기 때문에, 세트 $\{\Omega_{\text{SB},d}(k, f_j) \mid d \in \mathcal{J}_{\text{DIR}}(k, f_j)\}$ 는 각각의 $j = 1, \dots, F$ 에 대한 $\mathcal{M}_{\text{DIR}}(k)$ 의 서브세트이다. 이것은, 각각의 인덱스가 Q 개의 후보 방향들 대신에 $D(k)$, $D(k) \leq Q$ 중에서 하나의 방향을 정의하기 때문에, 방향들에 관한 부가 정보의 더 효율적인 코딩을 허용한다. 인덱스 d는 궤적을 생성하기 위한 후속 프레임에서의 방향들을 추적하는데 이용된다.

[0104] 도 2에 도시되고 전술된 바와 같이, 한 실시예에서 방향 추정 처리 블록(16)은 전체-대역 방향 추정 블록(21)을 갖는 방향 추정 블록(20)을 포함하고, 각각의 부대역 또는 부대역 그룹에 대해, 부대역 방향 추정 블록(22)을 포함한다. 이것은, 도 7에 도시된 바와 같이, 전술된 긴 프레임들을 방향 추정 블록(20)에 제공하는 긴 프레임 생성 블록(23)을 더 포함 할 수 있다. 긴 프레임 생성 블록(23)은, 예를 들어, 하나 이상의 메모리를 이용하여, 각각이 L개의 샘플들의 길이를 갖는 2개의 연속적인 입력 프레임들로부터 긴 프레임을 생성한다. 긴 프레임은, 여기서는 "-"와 2개의 인덱스, k-1 및 k를 가짐으로써 표시된다. 다른 실시예에서, 긴 프레임 생성 블록(23)은 도 1에 도시된 인코더에서 별개의 블록이거나, 다른 블록들에 포함될 수도 있다.

[0105] 방향 부대역 신호들의 계산

[0106] 도 1로 돌아와서, 분석 필터 뱅크(15)에 의해 제공된 부대역 HOA 표현 프레임들 $\bar{\mathcal{C}}(k, f_j) \quad j = 1, \dots, F,$ 은 또한 D_{SB} 하나 이상의 방향 부대역 신호 계산 블록(17)에 입력된다. 방향 부대역 신호 계산 블록(17)에서, 모든 D_{SB} 개의 잠재적인 방향 부대역 신호들 $\bar{\mathbf{x}}_d(k-1; k; f_j) \quad d = 1, \dots, D_{\text{SB}}$ 의 긴 프레임들은 행렬 $\bar{\mathbf{X}}(k-1; k; f_j)$ 에서 다음과 같이 배열된다,

수학식 16

$$\bar{\mathbf{x}}(k-1; k; f_j) = \begin{bmatrix} \bar{\mathbf{x}}_1(k-1; k; f_j) \\ \bar{\mathbf{x}}_2(k-1; k; f_j) \\ \vdots \\ \bar{\mathbf{x}}_{D_{\text{DSB}}}(k-1; k; f_j) \end{bmatrix} \in \mathbb{C}^{D_{\text{DSB}} \times 2L}.$$

[0107]

[0108] 또한, 비활성 방향 부대역 신호들의 프레임들, 즉 그 인덱스 d 가 세트 $\mathcal{J}_{\text{DIR}}(k, f_j)$ 내에 포함되지 않은 긴 신호 프레임들 $\bar{\mathbf{x}}_d(k-1; k; f_j)$ 은 0으로 설정된다.

[0109] 나머지 긴 신호 프레임들 $\bar{\mathbf{x}}_d(k-1; k; f_j)$, 즉 인덱스 $d \in \mathcal{J}_{\text{DIR}}(k, f_j)$ 를 갖는 것들은 행렬 $\bar{\mathbf{X}}_{\text{ACT}}(k-1; k; f_j) \in \mathbb{C}^{D_{\text{DSB}}(k, f_j) \times 2L}$ 내에 수집된다. 그 내부에 포함된 활성 방향 부대역 신호들을 계산하는 한 가성은 그들의 HOA 표현과 원래의 입력 부대역 HOA 표현 간의 오차를 최소화하는 것이다. 그 해는 다음과 같이 주어진다

수학식 17

$$\bar{\mathbf{X}}_{\text{ACT}}(k-1; k; f_j) = (\Psi_{\text{SB}}(k, f_j))^+ \bar{\mathbf{c}}(k-1; k; f_j)$$

[0110]

[0111] 여기서, $(\cdot)^+$ 는 Moore-Penrose 의사 역행렬을 나타내고, $\Psi_{\text{SB}}(k, f_j) \in \mathbb{R}^{O \times D_{\text{DSB}}(k, f_j)}$ 는 세트 $\{\Omega_{\text{SB}, d}(k, f_j) | d \in \mathcal{J}_{\text{DIR}}(k, f_j)\}$ 내의 방향 추정치들에 대한 모드 행렬을 나타낸다. 부대역 그룹들의 경우에, 방향 부대역 신호들의 세트 $\bar{\mathbf{X}}_{\text{ACT}}(k-1; k; f_j)$ 는 그룹의 모든 HOA 표현들 $\bar{\mathbf{c}}(k-1; k; f_j)$ 에 의한 한 행렬 $(\Psi_{\text{SB}}(k, f_j))^+$

[0112]

[0113] 전술된 바와 같이, 근사적인 HOA 표현은 부분적으로 활성 방향 부대역 신호들에 의해 표현되지만, 통상적으로 코딩되지 않는다. 대신에, 현재 설명되는 실시예들에서, 코딩된 표현의 전송을 위한 총 데이터 레이트를 낮게 유지하기 위해 파라메타 표현이 사용된다. 파라메타 표현에서, 각각의 활성 방향 부대역 신호,

즉, 인덱스 $d \in \mathcal{J}_{\text{DIR}}(k, f_j)$ 를 갖는 것들은, 절삭된 부대역 HOA 표현 $\check{\mathbf{c}}_n(k-1, f_j)$ 및 $\check{\mathbf{c}}_n(k, f_j)$ 의 계수 시퀀스들의 가장 합에 의해 예측되고, 여기서, $n \in \mathcal{J}_{\text{C,ACT}}(k-1)$ 이고, 가중치는 일반적으로 복소값이다.

[0114] 따라서, $\overline{\tilde{\mathbf{X}}_p(k-1; k; f_j)}$ 가 $\overline{\tilde{\mathbf{X}}(k-1; k; f_j)}$ 의 예측된 버전을 표현한다고 가정하면, 예측은 다음과 같은 행렬 곱셈에 의해 표현된다

수학식 18

[0115]
$$\overline{\tilde{\mathbf{X}}_p(k-1; k; f_j)} = \mathbf{A}(k, f_j) \overline{\tilde{\mathbf{C}}_T(k-1; k; f_j)},$$

[0116] 여기서, $\mathbf{A}(k, f_j) \in \mathbb{C}^{O \times D_{SB}}$ 는 부대역 f_j 에 대한 모든 가중 인자들(또는, 등가적으로, 예측 계수들)을 갖는 행렬이다. 예측 행렬 $\mathbf{A}(k, f_j)$ 의 계산은 하나 이상의 방향 부대역 예측 블록(18)에서 수행된다. 한 실시예에서, 도 1에 도시된 바와 같이, 부대역당 하나의 방향 부대역 예측 블록(18)이 이용된다. 다른 실시예에서, 단일의 방향 부대역 예측 블록(18)이 복수의 또는 모든 부대역에 대해 이용된다. 부대역 그룹의 경우, 각각의 그룹에 대해 하나의 행렬 $\mathbf{A}(k, f_j)$ 이 계산된다; 그러나, 이것은 그룹의 각각의 HOA 표현 $\overline{\tilde{\mathbf{C}}_T(k-1; k; f_j)}$ 으로 개별적으로 곱해져, 그룹마다 행렬 세트 $\overline{\tilde{\mathbf{X}}_p(k-1; k; f_j)}$ 를 생성한다. 구성당 인덱스 $d \in \mathcal{J}_{DIR}(k, f_j)$ 를 가진 것들을 제외한 $\mathbf{A}(k, f_j)$ 의 모든 행들이 0라는 점에 유의한다. 이것은 활성 방향 부대역 신호만이 예측된다는 것을 의미한다. 또한, 인덱스 $n \in \mathcal{J}_{CACT}(k-1)$ 를 가진 것들을 제외한 $\mathbf{A}(k, f_j)$ 의 모든 열도 0이다. 이것은, 예측을 위해, HOA 압축해제 동안 예측을 위해 전송되고 이용가능한 HOA 계수 시퀀스들만이 고려된다는 것을 의미한다.

[0117] 다음 양태들은 예측 행렬 $\mathbf{A}(k, f_j)$ 의 계산을 위해 반드시 고려되어야 한다.

[0118] 첫째, 원래의 절삭된 부대역 HOA 표현 $\tilde{\mathbf{C}}_T(k, f_j)$ 은 일반적으로 HOA 압축해제에서 이용가능하지 않을 것이다. 대신에, 그 인지 디코딩된 버전 $\tilde{\tilde{\mathbf{C}}}_T(k, f_j)$ 이 이용가능할 것이고 방향 부대역 신호의 예측에 사용될 것이다.

[0119] 낮은 비트 레이트에서, 스펙트럼의 더 낮은 주파수와 중간 주파수가 통상적으로 코딩되는, 전형적인 오디오 코덱(AAC 또는 USAC 등)은 스펙트럼 대역 복제(SBR; spectral band replication)를 이용하는 반면, 높은 주파수 성분(예를 들어, 5kHz에서 시작)은 고주파 엔벨로프에 대한 여분의 부가 정보를 이용하여 더 낮은 주파수 및 중간 주파수로부터 복제된다.

[0120] 그 때문에, 인가 디코딩 이후에 절삭된 HOA 성분 $\tilde{\tilde{\mathbf{C}}}_T(k, f_j)$ 의 재구성된 부대역 계수 시퀀스의 크기는 원래의 것 $\tilde{\mathbf{C}}_T(k, f_j)$ 과 유사하다. 그러나 이것은 위상의 경우에는 해당되지 않는다. 따라서, 고주파수 부대역들에 대해, 복소값 예측 계수들을 이용함으로써 예측에 대한 임의의 위상 관계를 활용하는 것은 의미가 없다. 대신에, 실수값 예측 계수만을 이용하는 것이 더 합리적이다. 특히, f_j 번째 부대역이 SBR에 대한 시작 주파수를 포함하도록 인덱스 j_{SBR} 을 정의하면, 예측 계수의 타입을 다음과 같이 설정하는 것이 유리하다 :

수학식 19

$$A(k, f_j) \in \begin{cases} \mathbb{C}^{O \times D_{SB}} & \text{for } 1 \leq j < j_{SBR} \\ \mathbb{R}^{O \times D_{SB}} & \text{for } j_{SBR} \leq j \leq F \end{cases}$$

[0121]

[0122] 다시 말해, 한 실시예에서, 더 낮은 부대역들에 대한 예측 계수들은 복소값인 반면, 더 높은 부대역들에 대한 예측 계수들은 실수값이다.

[0123] 둘째, 한 실시예에서, 행렬 $A(k, f_j)$ 의 계산의 전략은 그 유형에 적합화된다. 특히, SBR에 의해 영향을 받지 않는 저주파 부대역들 f_j , $1 \leq j < j_{SBR}$ 에 대해, $\bar{\mathcal{X}}(k-1; k; f_j)$ 와 그 예측된 버전 $\bar{\mathcal{X}}_p(k-1; k; f_j)$ 사이의 에러의 유클리드 놈(Euclidean norm)을 최소화함으로써 $A(k, f_j)$ 의 비제로 요소들을 결정하는 것이 가능하다. 인지 코더(31)는 (도시되지 않은) j_{SBR} 을 정의하고 제공한다. 이러한 방식으로, 관련 신호들의 위상 관계가 예측을 위해 명시적으로 활용된다. 부대역 그룹들에 대해, 그룹의 모든 방향 신호들에 대한 예측 에러의 유클리드 놈은 최소화되어야 한다(즉, 최소 제곱 예측 에러).

[0124] SBR에 의해 영향받는 고주파수 부대역들 f_j , $j_{SBR} \leq j \leq F$ 의 경우, 절삭된 HOA 성분 $\tilde{\mathcal{C}}_T(k, f_j)$ 의 재구성된 부대역 계수 시퀀스들의 위상들은 원래의 부대역 계수 시퀀스들의 것과 가장 기초적인 것조차 유사하다고 가정될 수 없기 때문에, 앞서 언급된 기준은 합리적이지 않다.

[0125] 이 경우, 한 가지 해결책은, 예측에 대해, 위상을 무시하고, 대신에 신호 전력에만 집중하는 것이다. 예측 계수의 결정에 대한 합리적인 기준은 다음과 같은 에러를 최소화하는 것이다

수학식 20

$$\left| \bar{\mathcal{X}}(k-1; k; f_j) \right|^2 - |A(k, f_j)|^2 \left| \tilde{\mathcal{C}}_T(k-1; k; f_j) \right|^2$$

[0126]

[0127] 여기서 연산 $|\cdot|^2$ 은 행렬에 요소별로 적용되는 것으로 가정된다. 다시 말하면, 예측 계수는, 절삭된 HOA 성분의 모든 가중 부대역 또는 부대역 그룹 계수 시퀀스들의 전력들의 합이 방향 부대역 신호들의 전력에 가장 가깝도록 선택된다. 이 경우, 이 최적화 문제를 해결하고 예측 행렬 $A(k, f_j), j = 1, \dots, F$ 의 예측 계수를 얻기 위해 비음수 행렬 인수분해 (Nonnegative Matrix Factorization; NMF) 기법(예를 들어, [8]을 참조)이 이용될 수 있다. 그 다음, 이들 행렬들은 인지 및 소스 인코딩 스테이지(30)에 제공된다.

[0128] 인지 및 소스 인코딩

[0129] 전술된 공간적 HOA 코딩 후에, (k-1)번째 프레임에 대한 결과적인 이득 조정된 전송 신호들 $\mathbf{z}_i(k-1)$, $i = 1, \dots, I$ 이 코딩되어 코딩된 표현 $\tilde{\mathbf{z}}_i(k-1)$ 을 획득한다. 이것은 도 3에 도시된 인지 및 소스 인코딩 스테이지(30)에서 인지 코더(31)에 의해 수행된다. 또한, 세트들 $\mathcal{M}_{DIR}(k)$, $\mathcal{M}_{DIR}(k, f_j)$, $j = 1, \dots, F$ 에 포함된 정보, 예측 계수 행렬들 $A(k, f_j) \in \mathbb{C}^{O \times D_{SB}}$, $j = 1, \dots, F$, 이득 제어 파라미터들 $e_i(k-1)$ 및 $\beta_i(k-1)$, $i = 1, \dots, I$, 및

할당 벡터 $\mathbf{v}_A(k-1)$ 는 효율적인 저장 또는 전송을 위한 리던던시를 제거하기 위해 소스 인코딩된다. 이것은 부가 정보 소스 코더(32)에서 수행된다. 결과적인 코딩된 표현 $\tilde{\mathbf{r}}(k-1)$ 은 코딩된 전송 신호 표현 $\tilde{\mathbf{z}}_i(k-1)$, $i = 1, \dots, I$ 과 함께 멀티플렉서(33)에서 멀티플렉싱되어 최종 코딩된 프레임 $\tilde{\mathbf{B}}(k-1)$ 을 제공한다.

[0130] 원칙적으로, 이득 제어 파라미터들의 소스 코딩 및 할당은 [9]와 유사하게 실행될 수 있으므로, 본 설명은, 이하에서 상세히 설명되는 방향 및 예측 파라미터의 코딩에만 집중한다.

[0131] 방향의 코딩

[0132] 개개의 부대역 방향의 코딩을 위해, 상기 설명에 따른 부적합 감소는 선택될 개개의 부대역 방향을 제약하는데 활용될 수 있다. 이미 언급한 바와 같이, 이러한 개개의 부대역 방향은 모든 가능한 테스트 방향들 $\Omega_{\text{TEST},q}$, $q = 1, \dots, Q$ 중에서가 아니라 전체-대역 HOA 표현의 각각의 프레임에 관해 결정된 소수의 후보 중에서 선택된다. 예시적으로, 부대역 방향들의 소스 코딩을 위한 한 가능한 방법이 이하의 알고리즘 1에 요약되어 있다.

알고리즘 1 부대역 방향들의 코딩

```

NoOfGlobalDirs(k) ( coded with  $\lceil \log_2(D) \rceil$  bits )

{Fill GlobalDirGridIndices(k) ( array with NoOfGlobalDirs(k) elements, each coded with  $\lceil \log_2(Q) \rceil$  bits) }
for d = 1 to NoOfGlobalDirs(k) do
    GlobalDirGridIndices(k)[d] = q such that  $\Omega_{\text{FB},d}(k) = \Omega_{\text{TEST},q}$  // global directions
end for

for j = 1 to F do

    {Fill bSubBandDirIsActive(k, f_j) ( bit array with  $D_{\text{SB}}$  elements) }

    for d = 1 to  $D_{\text{SB}}$  do
        if  $d \in \mathcal{I}_{\text{DIR}}(k, f_j)$  then // active directions
            bSubBandDirIsActive(k, f_j)[d] = 1 // per subband
        else
            bSubBandDirIsActive(k, f_j)[d] = 0
        end if
    end for

    {Fill RelDirIndices(k, f_j)
    (array with  $D_{\text{SB}}$ (k, f_j) elements, each coded with  $\lceil \log_2(\text{NoOfGlobalDirs}(k)) \rceil$  bits) }

    for d = 1 to  $D_{\text{SB}}$  do // direction index of
         $d_1 = 1$  // full band
        if bSubBandDirIsActive(k, f_j)[d] = 1 then
            RelDirIndices(k, f_j)[d_1] = i such that  $\Omega_{\text{SB},d}(k, f_j) = \Omega_{\text{FB},i}(k)$ 
             $d_1 = d_1 + 1$ 
        end if
    end for
end for
    
```

[0133] 알고리즘 1의 제1 단계에서, 실제로 부대역 방향들로서 발생하는 모든 전체-대역 방향 후보들의 세트 $\mathcal{M}_{\text{FB}}(k)$ 가 결정된다, 즉,

수학식 21

[0135]
$$\mathcal{M}_{\text{FB}}(k) = \left\{ \begin{array}{l} \Omega_{\text{CAND},d}(k) | \exists j \in \{1, \dots, F\} \text{ and } d \in \mathcal{I}_{\text{DIR}}(k, f_j) \\ \text{따라서 } \Omega_{\text{CAND},d}(k) = \Omega_{\text{SB},d}(k, f_j) \end{array} \right\}$$

[0136] $\text{NoOfGlobalDirs}(k)$ 로 표시된 이 세트의 요소들의 개수는 방향의 코딩된 표현의 첫 번째 부분이다. $\mathcal{M}_{\text{FB}}(k)$ 는 정의에 의해 $\mathcal{M}_{\text{DIR}}(k)$ 의 서브셋이기 때문에, $\text{NoOfGlobalDirs}(k)$ 는 $\lceil \log_2(D) \rceil$ 비트로 코딩될 수 있다. 추가 설

명을 명료화하기 위해, 세트 $\mathcal{M}_{\text{FB}}(k)$ 의 방향들은 $\Omega_{\text{FB},d}(k) \quad d = 1, \dots, \text{NoOfGlobalDirs}(k)$ 로 표기된다, 즉,

수학식 22

$$\mathcal{M}_{\text{FB}}(k) := \{\Omega_{\text{FB},d}(k) | d = 1, \dots, \text{NoOfGlobalDirs}(k)\}$$

[0137]

제2 단계에서, 세트 $\mathcal{M}_{\text{FB}}(k)$ 의 방향들은, 여기서는 그리드라고 하는, 가능한 테스트 방향들 $\Omega_{\text{TEST},q}$ 의 인덱스들 $q = 1, \dots, Q$ 을 이용하여 코딩된다. 각각의 방향 $\Omega_{\text{FB},d}(k) \quad d = 1, \dots, \text{NoOfGlobalDirs}(k)$ 에 대해, 각각의 그리드 인덱스는 $\lceil \log_2(Q) \rceil$ 비트의 크기를 갖는 배열 요소 $\text{GlobalDirGridIndices}(k)[d]$ 로 코딩된다. 모든 코딩된 전체-대역 방향들을 나타내는 전체 배열 $\text{GlobalDirGridIndices}(k)$ 은 $\text{NoOfGlobalDirs}(k)$ 개의 요소들로 구성된다.

[0138]

제3 단계에서, 각각의 부대역 또는 부대역 그룹 $f_j \quad j = 1, \dots, F$ 에 대해, d 번째 방향 부대역 신호($d = 1, \dots, D_{\text{SB}}$)가 활성화인지의 여부, 즉, $d \in \mathcal{J}_{\text{DIR}}(k, f_j)$ 인지 에 대한 정보는 배열 요소 $\text{bSubBandDirIsActive}(k, f_j)[d]$ 로 코딩된다. 총 배열 $\text{bSubBandDirIsActive}(k, f_j)$ 은 D_{SB} 개의 요소들로 구성된다. $d \in \mathcal{J}_{\text{DIR}}(k, f_j)$ 이면, 각각의 부대역 방향 $\Omega_{\text{SB},d}(k, f_j)$ 은, $D_{\text{SB}}(k, f_j)$ 개의 요소들로 구성된 배열 $\text{RelDirIndices}(k, f_j)$ 로의 각각의 전체-대역 방향 $\Omega_{\text{FB},i}(k)$ 의 인덱스 i 에 의해 코딩된다.

[0139]

이 방향 인코딩 방법의 효율성을 보이기 위해, 상기 예에 따른 방향들의 코딩된 표현에 대한 최대 데이터 레이트가 계산된다: $F = 10$ 개의 부대역, 부대역당 $D_{\text{SB}}(k, f_j) = D_{\text{SB}} = 4$ 개의 방향들, $Q = 900$ 개의 잠재적 테스트 방향들, 및 초당 25프레임의 프레임 레이트가 가정된다. 종래의 코딩 방법에서, 요구되는 데이터 레이트는 10 kbit/s였다. 한 실시예에 따른 개선된 코딩 방법에서는, 전체-대역 방향의 개수가 $\text{NoOfGlobalDirs}(k) = D = 8$ 인 것으로 가정하면, $\text{GlobalDirGridIndices}(k)$ 를 코딩하기 위해 프레임당 $D \cdot \lceil \log_2(Q) \rceil = 80$ 비트가, $\text{bSubBandDirIsActive}(k, f_j)$ 를 코딩하기 위해 $D_{\text{SB}} \cdot F = 40$ 비트가, 및 $\text{RelDirIndices}(k, f_j)$ 를 코딩하기 위해 $D_{\text{SB}} \cdot F \cdot \lceil \log_2(\text{NoOfGlobalDirs}(k)) \rceil = 120$ 비트가 필요하다. 그 결과, 데이터 레이트는 240 비트/프레임*25 프레임/s = 6 kbit/s이고, 이것은 10 kbit/s보다 분명히 작다. 전체-대역 방향들의 더 큰 수의 $\text{NoOfGlobalDirs}(k) = D = 16$ 에 대해서도, 7 kbit/s의 데이터 레이트만으로 충분하다.

[0140]

예측 계수 행렬의 코딩

[0141]

예측 계수 행렬의 코딩에 대해, 방향 궤적들 및 결과적으로 방향 부대역 신호의 평활성으로 인해 연속적인 프레임들의 예측 계수들 사이에는 높은 상관관계가 있다는 사실이 활용될 수 있다. 또한, 각각의 예측 계수 행렬 $\mathbf{A}(k, f_j)$ 에 대한 프레임당 $(D_{\text{SB}}(k, f_j) \cdot M_{\text{CACT}}(k-1))$ 개의 잠재적 비제로-요소들의 비교적 많은 수가 존재하며, 여기서, $M_{\text{CACT}}(k-1)$ 는 세트 $\mathcal{J}_{\text{CACT}}(k-1)$ 내의 요소들의 개수를 나타낸다. 전체적으로, 어떠한 부대역 그룹도 이용되지 않는다면 프레임당 코딩될 F 개의 행렬이 존재한다. 부대역 그룹들이 이용된다면, 대응적으로, 프레임

[0142]

당 코딩될 F 보다 적은 개수의 행렬이 존재한다.

[0143] 한 실시예에서, 각각의 예측 계수에 대한 비트수를 낮게 유지하기 위하여, 각각의 복소값 예측 계수는 그 크기와 각도에 의해 표현되고, 그 다음, 각도와 크기는 연속적인 프레임들간에 차분적으로 및 행렬 $A(k, f_j)$ 의 각각의 특정한 요소에 대해 독립적으로 코딩된다. 크기가 구간 $[0,1]$ 내에 있다고 가정하면, 크기 차이는 구간 $[-1,1]$ 내에 있다. 복소수들의 각도들의 차이는 구간 $[-\pi, \pi]$ 내에 있다고 가정될 수 있다. 크기와 각도 차이 양쪽 모두의 양자화를 위해, 각각의 구간은, 예를 들어, 동일한 크기의 2^{N_Q} 개의 부구간들로 세분될 수 있다. 직접적인 코딩은 각각의 크기 및 각도 차이에 대해 N_Q 개의 비트를 필요로 한다.

[0144] 또한, 연속적인 프레임들의 예측 계수들 사이의 전술된 상관관계로 인해, 개별 차이들의 발생 확률은 고도로 불균일하게 분포된다는 것이 실험적으로 밝혀졌다. 특히, 각도뿐만 아니라 크기에서의 작은 차이가 더 큰 쪽보다 훨씬 더 자주 발생한다. 따라서, 예를 들어, 허프만 코딩과 같은, 코딩될 개개의 값의 선형적 확률에 기초한 코딩 방법은, 예측 계수당 평균 비트 수를 상당히 감소시키는데 활용될 수 있다. 다시 말해서, 실수부 및 허수부 대신에, 예측 행렬 $A(k, f_j)$ 에서 값의 크기 및 위상을 차분적으로 인코딩하는 것이 대개 유리하다는 것이 발견되었다. 그러나 실수부 및 허수부의 이용이 허용되는 상황이 나타날 수 있습니다.

[0145] 한 실시예에서, 특별한 액세스 프레임들은 비차분적으로 코딩된 행렬 계수들을 포함하는 소정의 구간에서(애플리케이션에 특유함, 예를 들어, 초당 1회) 전송된다. 이것은, 디코더가 이들 특별한 액세스 프레임들로부터의 차분 디코딩을 재시작하는 것을 허용하므로, 디코딩을위한 무작위 진입을 가능하게 한다.

[0146] 이하에서는, 위에서 구성된 바와 같은 낮은 비트 레이트 압축된 HOA 표현의 압축해제가 설명된다. 또한 압축해제는 프레임별로 동작한다.

[0147] 원칙적으로, 실시예에 따르면, 낮은 비트 레이트 HOA 디코더는, 역순으로 배열된, 전술된 낮은 비트 레이트 HOA 인코더 컴포넌트들의 대응부들을 포함한다. 특히, 낮은 비트 레이트 HOA 디코더는 도 4에 도시된 바와 같은 인지 및 소스 디코딩 부분, 및 도 6에 도시된 바와 같은 공간 HOA 디코딩 부분으로 세분될 수 있다.

[0148] 인지 및 소스 디코딩

[0149] 도 4는, 한 실시예에서의, 인지 및 부가 정보 소스 디코더(40)를 도시한다. 인지 및 부가 정보 소스 디코더(40)에서, 낮은 비트 레이트 압축된 HOA 비트 스트림 \bar{B} 는 먼저 디멀티플렉싱되어(41), I 개의 신호들 \tilde{z}_i , $i = 1, \dots, I$ 의 인지 코딩된 표현과, 그 HOA 표현을 생성하는 방법을 기술하는 부가 정보 \bar{F} 가 된다. 후속해서, I 개의 신호들의 인지 디코딩 및 부가 정보의 디코딩이 수행된다.

[0150] 인지 디코더(42)는 I 개의 신호들 \tilde{z}_i , $i = 1, \dots, I$ 을 인지 디코딩된 신호들 $\hat{z}_i(k)$, $i = 1, \dots, I$ 로 디코딩한다.

[0151] 부가 정보 소스 디코더(43)는, 코딩된 부가 정보 \bar{F} 를, 튜플 세트 $\mathcal{M}_{\text{DIR}}(k+1, f_j)$, $j = 1, \dots, F$, 각각의 부대역 또는 부대역 그룹 f_j ($j = 1, \dots, F$)에 대한 예측 계수 행렬들 $A(k+1, f_j)$, 이득 보정 지수 $e_i(k)$ 및 이득 보정 예외 플래그 $\beta_i(k)$, 및 할당 벡터 $v_{\text{AMBASSIGN}}(k)$ 로 디코딩한다.

[0152] 알고리즘 2는 코딩된 부가 정보 \bar{F} 로부터 튜플 세트 $\mathcal{M}_{\text{DIR}}(k, f_j)$, $j = 1, \dots, F$ 를 생성하는 방법을 예시적으로 요약한다. 부대역 방향들의 디코딩이 이하에서 상세하게 설명된다.

알고리즘 2 부대역 방향들의 디코딩

```

Read NoOfGlobalDirs(k) ( coded with  $\lceil \log_2(D) \rceil$  bits )

{Read GlobalDirGridIndices(k) ( array with NoOfGlobalDirs(k) elements, each coded by  $\lceil \log_2(Q) \rceil$  bits )
(Compute  $\mathcal{M}_{FB}(k)$  )
  for d = 1 to NoOfGlobalDirs(k) do
     $\Omega_{FB,d}(k) = \Omega_{TEST,GlobalDirGridIndices(k)[d]}$ 
  end for

for j = 1 to F do

  {Read bSubBandDirIsActive(k, fj) ( bit array with DSB elements )

  (Compute DSB(k, fj) )

  DSB(k, fj) = 0
  for d = 1 to DSB(k, fj) do
    if bSubBandDirIsActive(k, fj)[d] = 1 then
      DSB(k, fj) = DSB(k, fj) + 1
    end if
  end for

  {Read RelDirIndices(k, fj) (array with DSB(k, fj) elements, each coded with  $\lceil \log_2(\text{NoOfGlobalDirs}(k)) \rceil$  bits ) }
  (Compute  $\mathcal{M}_{DIR}(k, fj)$  )

  for d = 1 to DSB(k, fj) do
    d1 = 1
    if bSubBandDirIsActive(k, fj)[d] = 1 then
       $\Omega_{SB,d}(k, fj) = \Omega_{FB,RelDirIndices(k,f_j)[d_1]}(k)$ 
       $\mathcal{M}_{DIR}(k, fj) = \mathcal{M}_{DIR}(k, fj) \cup \{d, \Omega_{SB,d}(k, fj)\}$ 
      d1 = d1 + 1
    end if
  end for
end for
end for

```

[0153]

[0154] 먼저, 코딩된 부가 정보 $\tilde{\Gamma}$ 로부터 전체-대역 방향들의 개수 $\text{NoOfGlobalDirs}(k)$ 가 추출된다. 전송된 바와 같이, 이들은 또한 부대역 방향들로서 이용된다. 이것은 $\lceil \log_2(D) \rceil$ 비트로 코딩된다.

[0155] 제2 단계에서, $\text{NoOfGlobalDirs}(k)$ 개의 요소들로 구성된 배열 $\text{GlobalDirGridIndices}(k)$ 이 추출되고 각각의 요소는 $\lceil \log_2(Q) \rceil$ 비트들로 코딩된다. 이 배열은 전체-대역 방향들 $\Omega_{FB,d}(k)$, $d = 1, \dots, \text{NoOfGlobalDirs}(k)$ 을 나타내는 그리드 인덱스를 포함하되,

수학식 23

[0156]
$$\Omega_{FB,d}(k) = \Omega_{TEST,GlobalDirGridIndices(k)[d]}$$

[0157] 그 다음, 각각의 부대역 또는 부대역 그룹 f_j , $j = 1, \dots, F$ 에 대해, D_{SB} 개의 요소들로 구성된 배열 $\text{bSubBandDirIsActive}(k, f_j)$ 이 추출되며, 여기서, d 번째 요소 $\text{bSubBandDirIsActive}(k, f_j)[d]$ 는 d 번째 부대역 방향이 활성화인지의 여부를 나타낸다. 또한, 활성화 부대역 방향들의 총 개수 $D_{SB}(k, f_j)$ 가 계산된다.

[0158] 마지막으로, 튜플 세트 $\mathcal{M}_{DIR}(k, f_j)$ 가 각각의 부대역 또는 부대역 그룹 f_j , $j = 1, \dots, F$ 에 대해 계산된다. 이것은 $d \in \mathcal{J}_{DIR}(k, f_j) \subseteq \{1, \dots, D_{SB}\}$ 개개의 (활성) 부대역 방향 궤적을 식별하는 인덱스들, 및 각각의 추정된 방향들 $\Omega_{SB,d}(k, f_j)$ 로 구성된다.

[0159] 그 다음, 각각의 부대역 또는 부대역 그룹 f_j , $j = 1, \dots, F$ 에 대한 예측 계수 행렬들 $A(k+1, f_j)$ 이 코딩된 프레임 $\tilde{\mathbf{B}}(k)$ 으로부터 재구성된다. 한 실시예에서, 재구성은 부대역 또는 부대역 그룹 f_j 마다 다음과 같은 단계들을 포함한다 :

[0160] 먼저, 각각의 행렬 계수의 각도 및 크기 차이가 엔트로피 디코딩에 의해 얻어진다. 그 다음 엔트로피 디코딩된 각도 및 크기 차이가 코딩에 사용된 비트수 N_Q 에 따라 실제 값 범위로 재조정된다. 마지막으로, 현재 예측 계수 행렬 $A(k+1, f_j)$ 은, 재구성된 각도 및 크기 차이를 최신 계수 행렬 $A(k, f_j)$ 의 계수들, 즉, 이전 프레임의 계수 행렬에 더함으로써 생성된다.

[0161] 따라서, 이전 행렬 $A(k, f_j)$ 은 현재 행렬 $A(k+1, f_j)$ 의 디코딩을 위해 알려져야만 한다. 한 실시예에서, 랜덤 액세스를 가능하게 하기 위해, 이들 프레임으로부터의 차분 디코딩을 재시작하기 위해 비차분적으로 코딩된 행렬 계수들을 포함하는 특별한 액세스 프레임이 소정 구간들에서 수신된다.

[0162] 인지 및 부가 정보 소스 디코더(40)는, 인지 디코딩된 신호들 $\hat{z}_i(k) \quad i = 1, \dots, I$, 튜플 세트들 $\mathcal{M}_{\text{DIR}}(k+1, f_j)$, $j = 1, \dots, F$, 예측 계수 행렬들 $A(k+1, f_j)$, 이득 보정 지수들 $e_i(k)$, 이득 보정 예외 플래그들 $\beta_i(k)$, 및 할당 벡터 $v_{\text{AMBASSIGN}}(k)$ 를 후속 공간 HOA 디코더(50)에 출력한다.

[0163] 공간 HOA 디코딩

[0164] 도 5는 한 실시예에서의 예시적인 공간 HOA 디코더(50)를 도시한다. 공간 HOA 디코더(50)는, I 개의 신호들 $\hat{z}_i(k) \quad i = 1, \dots, I$, 및 부가 정보 디코더 (43)에 의해 제공된 전송된 부가 정보로부터 재구성된 HOA 표현을 생성한다. 공간 HOA 디코더(50) 내의 개개의 처리 유닛들이 이하에서 상세하게 설명된다.

[0165] 역 이득 제어

[0166] 공간적 HOA 디코더(50)에서, 인지 디코딩된 신호들 $\hat{z}_i(k) \quad i = 1, \dots, I$, 은, 연관된 이득 보정 지수 $e_i(k)$ 및 이득 보정 예외 플래그 $\beta_i(k)$ 와 함께, 하나 이상의 역 이득 제어 처리 블록(51)에 먼저 입력된다. 역 이득 제어 처리 블록들은 이득 보정된 신호 프레임들 $\hat{y}_i(k), i = 1, \dots, I$ 를 제공한다. 한 실시예에서, I 개의 신호들 $\hat{z}_i(k)$ 각각은, 도 5에서와 같이, 별개의 역 이득 제어 처리 블록(51)에 공급되어, i 번째 역 이득 제어 처리 블록이 이득 보정된 신호 프레임 $\hat{y}_i(k)$ 을 제공하게 한다. 역 이득 제어에 대한 더 상세한 설명은, 예를 들어, [9], 11.4.2.1 절로부터 찾을 수 있다.

[0167] 절삭된 HOA 재구성

[0168] 절삭된 HOA 재구성 블록(52)에서, I 개의 이득 보정된 신호 프레임들 $\hat{y}_i(k), i = 1, \dots, I$ 은 할당 벡터 $v_{\text{AMBASSIGN}}(k)$ 에 의해 제공된 정보에 따라 HOA 계수 시퀀스 행렬에 재분배(즉, 재할당)되어, 절삭된 HOA 표현 $\hat{C}_T(k)$ 이 재구성되게 한다. 할당 벡터 $v_{\text{AMBASSIGN}}(k)$ 는 I 개의 성분들을 포함하고, 이들 성분들은, 각각의 전송 채널에 대해, 자신이 원래의 HOA 성분의 어느 계수 시퀀스를 포함하는지를 나타낸다. 또한, 할당 벡터의 요소들 k 은, k 번째 프레임에 대한 모든 수신된 계수 시퀀스들의 원래의 HOA 성분을 참조하는 인덱스 세트 $J_{\text{CACT}}(k)$ 를 형성한다

수학식 24

$$\mathcal{J}_{\text{C,ACT}}(k) = \{v_{\text{AMB,ASSIGN},i}(k) | i = 1, \dots, I\}.$$

[0169]

절삭된 HOA 표현 $\tilde{\mathbf{C}}_T(k)$ 의 재구성은 다음과 같은 단계들을 포함한다 :

[0170]

먼저, 하기와 같은 디코딩된 중간 표현의 개개의 성분들 $\hat{\mathbf{c}}_{1,n}(k)$, $n = 1, \dots, O$ 은,

[0171]

수학식 25

$$\tilde{\mathbf{C}}_1(k) = \begin{bmatrix} \hat{\mathbf{c}}_{1,1}(k) \\ \vdots \\ \hat{\mathbf{c}}_{1,O}(k) \end{bmatrix}$$

[0172]

할당 벡터 내의 정보에 따라, 0으로 설정되거나 이득 보정된 신호 프레임들 $\hat{\mathbf{y}}_i(k)$ 의 대응하는 성분에 의해 대체 된다, 즉,

[0173]

수학식 26

$$\hat{\mathbf{c}}_{1,n}(k) = \begin{cases} \hat{\mathbf{y}}_i(k) & \text{if } \exists i \in \{1, \dots, I\} \text{ 따라서 } v_{\text{AMB,ASSIGN},i}(k) = n \\ \mathbf{0} & \text{else} \end{cases}$$

[0174]

이것은, 앞서 언급된 바와 같이, 수학식 26에서 n 인, 할당 벡터의 i 번째 요소는, 디코딩된 중간 표현 행렬 $\tilde{\mathbf{C}}_1(k)$ 의 n 번째 라인의 $\hat{\mathbf{c}}_{1,n}(k)$ 가 i 번째 계수 $\hat{\mathbf{y}}_i(k)$ 로 대체됨을 나타낸다는 것을 의미한다.

[0175]

두 번째, $\tilde{\mathbf{C}}_1(k)$ 내의 처음 O_{MIN} 개의 신호들의 재상관은 이들에게 역 공간 변환을 적용함으로써 실행되며, 다음과 같은 프레임을 제공한다

[0176]

수학식 27

$$\tilde{\mathbf{C}}_{T,\text{MIN}}(k) = \Psi_{\text{MIN}} \begin{bmatrix} \hat{\mathbf{c}}_{1,1}(k) \\ \hat{\mathbf{c}}_{1,2}(k) \\ \vdots \\ \hat{\mathbf{c}}_{1,O_{\text{MIN}}}(k) \end{bmatrix}$$

[0177]

여기서 모드 행렬 Ψ_{MIN} 은 수학식 6에서 정의된 바와 같다. 모드 행렬은 각각의 O_{MIN} 또는 N_{MIN} 에 대해 미리 정의된 주어진 방향에 의존하며, 따라서 인코더 및 디코더 양쪽 모두에서 독립적으로 구성될 수 있다. 또한 O_{MIN} (또는 N_{MIN})은 규약에 의해 미리정의된다.

[0178]

[0179] 마지막으로, 재구성된 절삭된 HOA 표현 $\tilde{\mathbf{c}}_T(k)$ 은, 하기 수학식에 따라 재상관된 신호들 $\tilde{\mathbf{c}}_{T,\text{MIN}}(k)$ 과 중간 표현의 신호들 $\hat{\mathbf{c}}_{1,n}(k)$, $n = O_{\text{MIN}} + 1, \dots, O$ 로부터 생성된다

수학식 28

[0180]
$$\tilde{\mathbf{c}}_T(k) = \begin{bmatrix} \tilde{\mathbf{c}}_{T,\text{MIN}}(k) \\ \hat{\mathbf{c}}_{1,O_{\text{MIN}}+1}(k) \\ \vdots \\ \hat{\mathbf{c}}_{1,O}(k) \end{bmatrix} \in \mathbb{R}^{O \times L}$$

[0181] 분석 필터 बैंक

[0182] 예측된 방향 부대역 신호들에 의해 표현되는 제2 HOA 성분을 더 계산하기 위해, 압축해제된 절삭된 HOA 표현 $\tilde{\mathbf{c}}_T(k)$ 의 개개의 계수 시퀀스 n 의 각각의 프레임 $\hat{\mathbf{c}}_{T,n}(k)$, $n = 1, \dots, O$ 은 먼저 하나 이상의 분석 필터 बैंक(53)에서 개개의 부대역 신호들 $\hat{\mathbf{c}}_{T,n}(k, f_j)$, $j = 1, \dots, F$ 의 프레임들로 분해된다. 각각의 부대역 f_j , $j = 1, \dots, F$ 에 대해, 개개의 HOA 계수 시퀀스들의 부대역 신호들의 프레임들은 다음과 같은 부대역 HOA 표현 $\tilde{\mathbf{c}}_T(k, f_j)$ 으로 집합될 수 있다.

수학식 29

[0183]
$$\tilde{\mathbf{c}}_T(k, f_j) = \begin{bmatrix} \hat{\mathbf{c}}_{T,1}(k, f_j) \\ \hat{\mathbf{c}}_{T,2}(k, f_j) \\ \vdots \\ \hat{\mathbf{c}}_{T,O}(k, f_j) \end{bmatrix} \text{ for } j = 1, \dots, F$$

[0184] HOA 공간 디코딩 스테이지에서 적용되는 하나 이상의 분석 필터 बैंक(53)는 HOA 공간 인코딩 스테이지에서의 하나 이상의 분석 필터 बैंक(15)와 동일하고, 부대역 그룹에 대해서는 HOA 공간 인코딩 스테이지로부터의 그룹화가 적용된다. 따라서, 한 실시예에서, 그룹화 정보가 인코딩된 신호에 포함된다. 그룹화 정보에 대한 더 많은 상세사항은 이하에서 제공된다.

[0185] 한 실시예에서, HOA 압축 스테이지(상기 수학식 4 부근을 참조)에서의 절삭된 HOA 표현의 계산에 대해 최대 차수 N_{MAX} 가 고려되며, HOA 압축기 및 압축해제기의 분석 필터 बैंक(15, 53)의 적용은 인덱스들 $n = 1, \dots, O_{\text{MAX}}$ 을 갖는 HOA 계수 시퀀스들 $\hat{\mathbf{c}}_{T,n}(k)$ 만으로 제한된다. 그러면, 인덱스들 $n = O_{\text{MAX}} + 1, \dots, O$ 을 갖는 부대역 신호 프레임들 $\hat{\mathbf{c}}_{T,n}(k, f_j)$ 은 0으로 설정될 수 있다.

[0186] 방향 부대역 HOA 표현의 합성

[0187] 각각의 부대역 또는 부대역 그룹에 대해, 방향 부대역 또는 부대역 그룹 HOA 표현 $\hat{\mathbf{c}}_D(k, f_j)$, $j = 1, \dots, F$ 은 하나 이상의 방향 부대역 합성 블록(54)에서 합성된다. 한 실시예에서, 연속적인 프레임들 간의 방향 및 예측 계수의 변화로 인한 아티팩트를 피하기 위해, 방향 부대역 HOA 표현의 계산은 중첩 가산(overlap add)의 개념에 기초한다. 따라서, 한 실시예에서, f_j 번째 부대역, $j = 1, \dots, F$ 에 관련된 활성 지향성 부대역 신호의 HOA 표현

$\hat{\tilde{\mathbf{c}}}_D(k, f_j)$ 은 페이드 아웃 된 성분 및 페이드 인 된 성분의 합으로서 계산된다:

수학식 30

$$\hat{\tilde{\mathbf{c}}}_D(k, f_j) = \hat{\tilde{\mathbf{c}}}_{D,OUT}(k, f_j) + \hat{\tilde{\mathbf{c}}}_{D,IN}(k, f_j).$$

[0188]

제1 단계에서, 2개의 개개의 성분들을 계산하기 위해, 프레임들 $k_1 \in \{k, k+1\}$ 에 대한 예측 계수 행렬들 $\mathbf{A}(k_1, f_j)$ 및 k 번째 프레임에 대한 절삭된 부대역 HOA 표현 $\hat{\tilde{\mathbf{c}}}_T(k, f_j)$ 과 관련된 모든 방향 부대역 신호들 $\hat{\tilde{\mathbf{x}}}_1(k_1; k; f_j)$ 의 순간 프레임은 다음과 같이 계산된다

수학식 31

$$\hat{\tilde{\mathbf{x}}}_1(k_1; k; f_j) = \mathbf{A}(k_1, f_j) \hat{\tilde{\mathbf{c}}}_T(k, f_j) \quad \text{for } k_1 \in \{k, k+1\}.$$

[0190]

부대역 그룹들에 대해, 각각의 그룹 $\hat{\tilde{\mathbf{c}}}_T(k, f_j)$ 의 HOA 표현들은 고정된 행렬 $\mathbf{A}(k_1, f_j)$ 에 의해 곱해져 그 그룹의 부대역 신호들 $\hat{\tilde{\mathbf{x}}}_1(k_1; k; f_j)$ 을 생성한다.

[0191]

제2 단계에서, 방향 $\Omega_{SB,d}(k, f_j)$ 에 대한 방향 부대역 신호 $\hat{\tilde{\mathbf{x}}}_{1,d}(k_1; k; f_j)$ 의 순시적 부대역 HOA 표현 $\hat{\tilde{\mathbf{c}}}_{D,I}^{(d)}(k_1; k; f_j)$ $d \in \mathcal{M}_{DIR}(k, f_j)$ $j = 1, \dots, F$ 이 다음과 같이 얻어진다

[0192]

수학식 32

$$\hat{\tilde{\mathbf{c}}}_{D,I}^{(d)}(k_1; k; f_j) = \boldsymbol{\psi}(\Omega_{SB,d}(k, f_j)) \hat{\tilde{\mathbf{x}}}_{1,d}(k_1; k; f_j)$$

[0193]

여기서, $\boldsymbol{\psi}(\Omega_{SB,d}(k, f_j)) \in \mathbb{R}^O$ 는 방향 $\Omega_{SB,d}(k, f_j)$ 에 관한 (수학식 7의 모드 벡터로서의) 모드 벡터를 나타낸다.

[0194]

각각의 부대역 그룹에 대해, 수학식 32는 그룹의 모든 신호에 대해 수행되고, 여기서, 행렬 $\boldsymbol{\psi}(\Omega_{SB,d}(k, f_j))$ 는 각각의 그룹에 대해 고정된다.

[0195]

행렬들 $\hat{\tilde{\mathbf{c}}}_{D,OUT}(k, f_j)$, $\hat{\tilde{\mathbf{c}}}_{D,IN}(k, f_j)$, 및 $\hat{\tilde{\mathbf{c}}}_{D,I}^{(d)}(k_1; k; f_j)$ 은 그들의 샘플들로 다음과 같이 구성된다고 가정하자

수학식 33

$$\hat{\tilde{\mathbf{c}}}_{D,OUT}(k, f_j) = \begin{bmatrix} \hat{\tilde{c}}_{D,OUT,1}(k, f_j; 1) & \dots & \hat{\tilde{c}}_{D,OUT,1}(k, f_j; L) \\ \vdots & \ddots & \vdots \\ \hat{\tilde{c}}_{D,OUT,O}(k, f_j; 1) & \dots & \hat{\tilde{c}}_{D,OUT,O}(k, f_j; L) \end{bmatrix} \in \mathbb{R}^{O \times L}$$

[0196]

수학식 34

[0197]

$$\tilde{\mathbf{c}}_{D,IN}(k, f_j) = \begin{bmatrix} \tilde{\mathbf{c}}_{D,IN,1}(k, f_j; 1) & \dots & \tilde{\mathbf{c}}_{D,IN,1}(k, f_j; L) \\ \vdots & \ddots & \vdots \\ \tilde{\mathbf{c}}_{D,IN,O}(k, f_j; 1) & \dots & \tilde{\mathbf{c}}_{D,IN,O}(k, f_j; L) \end{bmatrix} \in \mathbb{R}^{O \times L}$$

수학식 35

[0198]

$$\tilde{\mathbf{c}}_{D,I}^{(d)}(k_1; k; f_j) = \begin{bmatrix} \hat{\mathbf{c}}_{D,I,1}^{(d)}(k-1; k; f_j; 1) & \dots & \hat{\mathbf{c}}_{D,I,1}^{(d)}(k-1; k; f_j; L) \\ \vdots & \ddots & \vdots \\ \hat{\mathbf{c}}_{D,I,O}^{(d)}(k-1; k; f_j; 1) & \dots & \hat{\mathbf{c}}_{D,I,O}^{(d)}(k-1; k; f_j; L) \end{bmatrix} \in \mathbb{R}^{O \times L}$$

[0199]

활성 방향 부대역 신호들의 HOA 표현의 성분들에서 페이딩 아웃되고 페이딩 인된 샘플 값들은 결국 다음과 같이 결정된다

수학식 36

[0200]

$$\tilde{\mathbf{c}}_{D,OUT,n}(k, f_j; l) = \sum_{d \in \mathcal{D}_{DIR}(k, f_j)} \hat{\mathbf{c}}_{D,I,n}^{(d)}(k; k; f_j; l) \cdot w_{OA}(L+l)$$

수학식 37

[0201]

$$\tilde{\mathbf{c}}_{D,IN,n}(k, f_j; l) = \sum_{d \in \mathcal{D}_{DIR}(k+1, f_j)} \hat{\mathbf{c}}_{D,I,n}^{(f_j)}(k+1; k; d; l) \cdot w_{OA}(l)$$

[0202]

여기서, 하기 벡터는

수학식 38

[0203]

$$\mathbf{w}_{OA} = [w_{OA}(1) \ w_{OA}(2) \ \dots \ w_{OA}(2L)]^T \in \mathbb{R}^{2L}$$

[0204]

중첩 가산 윈도우 함수를 나타낸다. 윈도우 함수의 예는 주기적 Hann 윈도우에 의해 주어지고, 그 요소들은 다음과 같이 정의된다

수학식 39

[0205]

$$w_{OA}(l) = \frac{1}{2} \left[1 - \cos \left(2\pi \frac{l-1}{2L} \right) \right]$$

[0206]

부대역 HOA 조성(composition)

[0207]

각각의 부대역 또는 부대역 그룹 $f_j \quad j = 1, \dots, F$ 에 대해, 디코딩된 부대역 HOA 표현 $\tilde{\mathbf{c}}(k, f_j)$ 의 계수 시퀀스들 $\hat{\mathbf{c}}_n(k, f_j)$, $n = 1, \dots, O$ 은, 이전에 전송된 경우에는 절삭된 HOA 표현 $\hat{\mathbf{c}}_T(k, f_j)$ 의 계수 시퀀스의 것으로

설정되고, 그 외의 경우에는 방향 부대역 합성 블록(54)들 중 하나에 의해 제공되는 방향 HOA 성분 $\hat{\tilde{c}}_D(k, f_j)$ 의 것으로 설정된다, 즉,

수학식 40

$$\hat{\tilde{c}}_n(k, f_j) = \begin{cases} \tilde{\tilde{c}}_{T,n}(k, f_j) & \text{if } n \in \mathcal{J}_{C,ACT}(k) \\ \tilde{\tilde{c}}_{D,n}(k, f_j) & \text{else} \end{cases}$$

[0208]

[0209] 이 부대역 조성은 하나 이상의 부대역 조성 블록(55)에 의해 수행된다. 실시예에서, 각각의 부대역 또는 부대역 그룹에 대해 별개의 부대역 조성 블록(55)이 이용되므로, 하나 이상의 방향 부대역 합성 블록(54)들 각각에 대해 이용된다. 한 실시예에서, 방향 부대역 합성 블록(54) 및 그 대응하는 부대역 조성 블록(55)은 단일의 블록으로 통합된다.

[0210] 합성 필터 बैं크

[0211] 최종 단계에서, 디코딩된 HOA 표현은 모든 디코딩된 부대역 HOA 표현들 $\hat{\tilde{c}}(k, f_j), j = 1, \dots, F$ 로부터 합성된다. 압축해제된 HOA 표현 $\tilde{\tilde{c}}(k)$ 의 개개의 시간 영역 계수 시퀀스들 $\tilde{\tilde{c}}_n(k), n = 1, \dots, O$ 은, 압축해제된 HOA 표현 $\tilde{\tilde{c}}(k)$ 을 최종적으로 출력하는 하나 이상의 합성 필터 बैं크(56)에 의해 대응하는 부대역 계수 시퀀스들 $\hat{\tilde{c}}_n(k, f_j), j = 1, \dots, F$ 로부터 합성된다.

[0212] 합성된 시간 영역 계수 시퀀스는 일반적으로 분석 및 합성 필터 बैं크(53, 56)의 연속적인 적용으로 인해 지연을 갖는다는 점에 유의한다.

[0213] 도 8은, 예시적으로, 단일 주파수 부대역 f_1 에 대해, 활성 방향 후보들의 세트, 이들의 선택된 궤적들 및 대응하는 튜플 세트를 도시한다. 프레임 k 에서, 4개의 방향이 주파수 부대역 f_1 에서 활성이다. 방향들은 각각의 궤적 T_1, T_2, T_3 및 T_5 에 속한다. 이전의 프레임 $k-2$ 및 $k-1$ 에서는, 상이한 방향들이 활성이었다, 즉, 각각 T_1, T_2, T_6 및 T_1-T_4 . 프레임 k 에서의 활성 방향 세트 $M_{DIR}(k)$ 는 전체 대역과 관련되며, 수 개의 활성 방향 후보들을 포함한다, 예를 들어, $M_{DIR}(k) = \{\Omega_3, \Omega_8, \Omega_{52}, \Omega_{101}, \Omega_{229}, \Omega_{446}, \Omega_{581}\}$ 이다. 각각의 방향은 임의의 방식으로, 예를 들어, 2개의 각도로 또는 미리정의된 표의 인덱스로서, 표현될 수 있다. 활성 전체-대역 방향들의 세트로부터, 부대역 내에서 실제로 활성인 방향들 및 그들의 대응하는 궤적들은, 튜플 세트들 $M_{DIR}(k, f_j), j = 1, \dots, F$ 에서 각각의 주파수 부대역에 대해 개별적으로 수집된다. 예를 들어, 프레임 k 의 제1 주파수 부대역에서, 활성 방향들은 $\Omega_3, \Omega_{52}, \Omega_{229}$ 및 Ω_{581} 이고, 이들의 연관된 궤적은 각각 T_3, T_1, T_2 및 T_5 이다. 제2 주파수 부대역 f_2 에서, 활성 방향은 전형적으로 Ω_{52} 및 Ω_{229} 뿐이고, 그들의 연관된 궤적은 각각 T_1 및 T_2 이다.

[0214] 다음은, 예시적인 세트 $I_{C,ACT}(k) = \{1, 2, 4, 6\}$ 의 계수 시퀀스들에 대응하는, 예시적인 절삭된 HOA 표현 $C_T(k)$ 의 계수 행렬의 일부이다:

$$C_T(k) = \begin{bmatrix} c_{T,1}(k, 1) & c_{T,1}(k, 2) & c_{T,1}(k, 3) & \dots \\ c_{T,2}(k, 1) & c_{T,2}(k, 2) & c_{T,2}(k, 3) & \dots \\ 0 & 0 & 0 & \dots \\ c_{T,4}(k, 1) & c_{T,4}(k, 2) & c_{T,4}(k, 3) & \dots \\ 0 & 0 & 0 & \dots \\ c_{T,6}(k, 1) & c_{T,6}(k, 2) & c_{T,6}(k, 3) & \dots \\ \dots & \dots & \dots & \dots \end{bmatrix}$$

[0215]

[0216] $I_{C,ACT}(k)$ 에 따라, 행 1, 2, 4 및 6의 계수들만 0으로 설정되지 않는다(그럼에도 불구하고, 이들은 신호에 따라, 0 일 수 있다). 행렬 $C_T(k)$ 의 각각의 열은 샘플을 참조하고, 행렬의 각각의 행은 계수 시퀀스이다. 압축은, 모

는 계수 시퀀스가 인코딩되고 전송되는 것이 아니라, 일부 선택된 계수 시퀀스, 즉, 인덱스가 각각 $I_{c,ACT}(k)$ 및

할당 벡터 $v_A(k)$ 에 포함된 계수 시퀀스만 인코딩되고 전송된다는 것을 포함한다. 디코더에서, 계수들은 압축해제되고 재구성된 절삭된 HOA 표현의 정확한 행렬 행들에 위치한다. 행들에 관한 정보는 할당 벡터 $v_{AMBASSIGN}(k)$

로부터 얻어지며, 이 할당 벡터는 각각의 전송된 계수 시퀀스에 이용되는 전송 채널을 추가로 제공한다. 나머지 계수 시퀀스들은 0으로 채워지고, 수신된 부가 정보에 따라 수신된 (일반적으로 0이 아닌) 계수들로부터 나중에 예측된다, 예를 들어, 부대역 또는 부대역 그룹 관련 예측 행렬들 및 방향들.

[0217] 부대역 그룹화

[0218] 한 실시예에서, 이용된 부대역들은 사람의 청각의 음향심리학적 특성에 적합화된 상이한 대역폭들을 가진다. 대안으로서, 분석 필터 बैं크(53)로부터의 다수의 부대역들이 결합되어 상이한 대역폭들을 갖는 부대역들을 갖춘 적합화된 필터 बैं크를 형성한다. 분석 필터 बैं크(53)로부터의 인접한 부대역들의 그룹은 동일한 파라미터들을 이용하여 처리된다. 결합된 부대역들의 그룹들이 이용된다면, 인코더 측에서 적용되는 대응하는 부대역 구성은 디코더측에 알려져야 한다. 실시예에서, 구성 정보가 전송되고 디코더에 의해 그 합성 필터 बैं크를 셋업하기 위해 이용된다. 실시예에서, 구성 정보는 복수의 미리정의된 알려진 구성들 (예를 들어, 리스트) 중 하나에 대한 식별자를 포함한다.

[0219] 또 다른 실시예에서, 부대역 구성을 정의하기 위해 필요한 비트수를 감소시키는 다음과 같은 융통성있는 해결책이 이용된다. 부대역 구성의 효율적인 인코딩을 위해, 처음의, 끝에서 두 번째의, 및 마지막 부대역 그룹의 데이터는 다른 부대역 그룹들과는 상이하게 취급된다. 또한, 부대역 그룹 대역폭 차이 값들은 인코딩에서 이용된다. 원칙적으로, 부대역 그룹화 정보 코딩 방법은 오디오 신호의 하나 이상의 프레임들에 대해 유효한 부대역 그룹들에 대한 부대역 구성 데이터를 코딩하는데 적합하며, 여기서, 각각의 부대역 그룹은 하나 이상의 인접한 원래의 부대역들의 조합이고 원래의 부대역들의 수는 미리 정의되어 있다. 일 실시예에서, 후속 부대역 그룹의 대역폭은 현재 부대역 그룹의 대역폭보다 크거나 같다. 이 방법은 $N_{SB}-1$ 을 나타내는 고정된 비트수를 갖는 다수의 N_{SB} 개의 부대역 그룹을 코딩하는 단계를 포함하고, $N_{SB} > 1$ 이면, 제1 부대역 그룹 g_1 에 대해 $B_{SB}[1]-1$ 을 나타내는 단항 코드를 갖는 대역폭 값 $B_{SB}[1]$ 을 코딩한다. $N_{SB}=3$ 이면, 고정된 비트수를 갖는 대역폭 차이 값 $\Delta B_{SB}[2] = B_{SB}[2] - B_{SB}[1]$

이 제2 부대역 그룹 g_2 에 대해 코딩된다. $N_{SB} > 3$ 이면, 대응하는 개수의 대역폭 차이

값들 $\Delta B_{SB}[g] = B_{SB}[g] - B_{SB}[g-1]$ 이 단항 코드를 갖는 부대역 그룹들 $g_2, \dots, g_{N_{SB}-2}$ 에 대해 코딩되고, 고정된

수의 비트를 갖는 대역폭 차이 값 $\Delta B_{SB}[N_{SB}-1] = B_{SB}[N_{SB}-1] - B_{SB}[N_{SB}-2]$ 이 마지막 부대역 그룹 $g_{N_{SB}-1}$ 에

대해 코딩된다. 부대역 그룹에 대한 대역폭 값은 인접한 원래의 부대역들의 수로서 표현된다. 마지막 부대역 그룹 $g_{N_{SB}}$ 에 대해, 어떠한 대응하는 값도 코딩된 부대역 구성 데이터에 포함될 필요가 없다.

[0220] 도 9는 종래의 MPEG-H 3D 오디오 인코더의 HOA 인코딩 경로의 일반화된 블록도를 도시한다. 2가지 유형의 우세 사운드 신호가 추출된다: 방향 사운드 추출 블록 DSE에서의 방향 신호들과 VVec 사운드 추출 블록 VSE에서의 벡터-기반 신호 VVec. 벡터-기반 신호 VVec(V-벡터)에 속하는 벡터는 대응하는 벡터-기반 신호에 대한 음장의 공간 분포를 나타낸다. 또한, 환경 성분은 잔류/환경용 계산기(CRA; Calculator for Residuum/Ambience)에 인코딩됨으로써, 방향 사운드 추출 블록 DSE 및 VVec 사운드 추출 블록 VSE로부터의 출력 데이터 중 하나 또는 양쪽 모두가 이용되거나 어느 것도 이용될 수 없다. 환경 신호는 공간 해상도 감소 블록 SRR, 부분적 무상관화 PD 및 이득 제어 GC_A 를 거친다. 박스 내의 블록들은 사운드 장면 분석(SSA; Sound Scene Analysis)에 의해 제어된다. 범용 음성 & 오디오 인코더(USAC3D; Universal Speech & Audio encoder)에 공급되기 전에, 우세 사운드 신호는 각각의 이득 제어 블록들 GC_D , GC_V 에 의해 처리된다. 마지막으로, USAC3D 인코더 ENC_c & HEP_c 는 HOA 공간측 정보를 HOA 확장 페이로드로 패키징한다.

[0221] 도 10은 한 실시예에 따른 MPEG에서 이용가능한 개선된 오디오 인코더를 도시한다. 개시된 기술은 저대역폭을 위한 비트스트림이 공지된 MPEG-H 3D 오디오 포맷의 실제 수퍼셋이도록 하는 방식으로 현재의 MPEG-H 3D 오디오 시스템을 수정한다. 도 9와 비교하여, 사운드 장면 분석 SSA에서, 2개의 새로운 블록을 포함하는 경로가 추가된다. 이들은, 환경 신호에 적용되는 QMF 분석 필터 बैं크 QA_c , 및 방향 부대역 신호의 파라미터의 계산을 위

한 방향 부대역 계산 블록 DSC_c 이다. 이들 파라미터들은 전송된 환경 신호에 기초한 방향 신호들의 합성을 허용한다. 추가로, 누락된 환경 신호의 재생을 허용하는 파라미터들이 계산된다. 합성 프로세스에 대한 부가 정보 파라미터는 USAC3D 인코더 ENC&HEP로 전달되고, 여기에서, 이들은 압축된 출력 신호 $HOA_{c,0}$ 의 HOA 확장 페이로드로 패키징된다. 유익하게도, 압축은 도 9의 구성에 의해 달성되는 종래의 압축보다 효율적이다.

[0222] 도 11은 종래의 MPEG-H 3D 오디오 디코더의 일반화된 블록도를 도시한다. 먼저, HOA 부가 정보가 압축된 입력 비트 스트림 $HOA_{c,1}$ 및 USAC3D로부터 추출되고 HOA 확장 페이로드 디코더 $DEC_c \& HEP_c$ 는 전송 채널 파형 신호를 재생한다. 이들은 대응하는 역 이득 제어 블록들 IGC_b , IGC_v , IGC_A 에 공급된다. 여기서, 인코더에서 적용된 정규화는 역전된다. 대응하는 전송 채널들이 부가 정보와 함께 이용되어 각각 HOA 방향 사운드 합성 블록 DSS 및/또는 VVec 사운드 합성 블록 VSS에서 우세 사운드 신호들(방향 및/또는 벡터-기반)을 합성한다. 세번째 경로에서, 역부분적 무상관화 IPD 및 HOA 환경 합성 HAS 블록들에 의해 환경 성분이 재생된다. 후속하는 HOA 조성 블록 HC_c 는 우세 사운드 성분 및 환경 성분을 결합하여 디코딩된 HOA 신호를 구축한다. 이것은 HOA 렌더러 HR에 공급되어 출력 신호 $HOA'_{d,0}$, 즉, 최종 확성기 피드(feed)를 생성한다.

[0223] 도 12는 한 실시예에 따른 MPEG에서 이용가능한 개선된 오디오 디코더를 도시한다. 인코더에서와 같이, 경로가 추가된다. 이것은, 부대역 신호들의 계산을 위한 디코더측 QMF 분석 블록 QA_b , 및 파라미터적으로 인코딩된 방향 부대역 신호들의 합성을 위한 방향 부대역 신호 합성 블록 DSC_b 를 포함한다. 계산된 부대역 신호들은 대응하는 전송측 정보와 함께 이용되어 방향 신호들의 HOA 표현을 합성한다. 그 후, 합성된 신호 성분은 QMF 합성 필터 बैं크 QS를 이용하여 시간 영역으로 이전된다. 그 출력 신호는 강화된 HOA 조성 블록 HC에 추가로 공급된다. 디코딩된 HOA 출력 신호 $HOA_{d,0}$ 를 제공하기 위한 후속하는 HOA 렌더링 블록 HR은 변하지 않고 남아 있다.

[0224] 이하에서는, HOA(Higher Order Ambisonics)의 일부 기본적인 피쳐들이 설명된다.

[0225] HOA(Higher Order Ambisonics)는 사운드 소스가 없는 것으로 가정되는 관심대상의 콤팩트 영역 내의 음장에 대한 설명에 기초한다. 이 경우, 관심대상 영역 내의 시간 t 와 위치 \mathbf{x} 에서의 음압 $p(t, \mathbf{x})$ 의 시공간적 거동은 균질 파동 방정식에 의해 물리적으로 완전히 결정된다. 이하에서는, 도 6에 도시된 구면 좌표계를 가정한다. 이 좌표계에서, x 축은 정면 위치를 가리키고 y 축은 좌측을 가리키며 z 축은 상부를 가리킨다. 공간에서의 위치 $\mathbf{x} = (r, \theta, \phi)^T$ 는, 반경 $r > 0$ (즉, 좌표 원점까지의 거리), 극축 z (!)으로부터 측정된 경사각 $\theta \in [0, \pi]$, 및 x 축으로부터의 x - y 평면에서 반시계 방향으로 측정된 방위각 $\phi \in [0, 2\pi[$ 으로 나타낸다. 또한, $(\cdot)^T$ 는 전치(transposition)를 나타낸다.

[0226] 그 다음, $\mathcal{F}_t(\cdot)$ 로 표기된 시간에 관한 음압의 푸리에 변환, 즉,

수학식 41

[0227]
$$P(\omega, \mathbf{x}) = \mathcal{F}_t(p(t, \mathbf{x})) = \int_{-\infty}^{\infty} p(t, \mathbf{x}) e^{-i\omega t} dt$$

[0228] 상기 수학식(여기서, ω 는 각주파수를 나타내고 i 는 허수 단위를 나타냄)은 하기 수학식에 따라 구면 고조파 급수로 확장될 수 있다는 것이 나타내어질 수 있다[11].

수학식 42

[0229]
$$P(\omega = kc_g, r, \theta, \phi) = \sum_{n=0}^N \sum_{m=-n}^n A_n^m(k) j_n(kr) S_n^m(\theta, \phi)$$

[0230] 수학식 42에서, c_s 는 사운드의 속도를 나타내고 $k = \frac{\omega}{c_s}$ 에 의해 각주파수 ω 와 관련된 각파수(angular wave number)를 나타낸다. 또한, $j_n(\cdot)$ 는 제1 종 구면 베셀 함수를 나타내고, $S_n^m(\theta, \phi)$ 는, 위에서 정의된 차수 n 및 도수 m 의 실수값 구면 고조파를 나타낸다. 확장 계수 $A_n^m(k)$ 는 각파수 k 에만 의존한다. 음압은 공간적으로 대역-제한된다는 것이 묵시적으로 가정된다는 점에 유의한다. 따라서, 급수는, HOA 표현의 차수라고 불리는 상한 N 에서 차수 인덱스 n 에 관하여 절삭된다.

[0231] 음장이 상이한 각주파수들 ω 의 무한 개수의 고조파 평면파들의 중첩으로 표현되고 각도 튜플 (θ, ϕ) 에 의해 명시된 모든 가능한 방향들에서 도달하면, 각각의 평면파 복소 진폭 함수 $C(\omega, \theta, \phi)$ 는 다음과 같은 구면 고조파 확장에 의해 나타낼 수 있다[10]

수학식 43

[0232]
$$C(\omega = kc_s, \theta, \phi) = \sum_{n=0}^N \sum_{m=-n}^n C_n^m(k) S_n^m(\theta, \phi)$$

[0233] 여기서, 확장 계수들 $C_n^m(k)$ 은 확장 계수들 $A_n^m(k)$ 과 하기 수학식에 의해 관련된다

수학식 44

[0234]
$$A_n^m(k) = i^n C_n^m(k).$$

[0235] 개개의 계수들 $C_n^m(k = \omega/c_s)$ 이 각주파수 ω 의 함수인 것으로 가정하면, ($\mathcal{F}^{-1}(\cdot)$ 으로 표기되는) 역 푸리에 변환의 적용은 각각의 차수 n 및 도수 m 에 대해 하기의 시간 영역 함수들을 제공한다

수학식 45

[0236]
$$c_n^m(t) = \mathcal{F}_t^{-1}(C_n^m(\omega/c_s)) = \frac{1}{2\pi} \int_{-\infty}^{\infty} C_n^m\left(\frac{\omega}{c_s}\right) e^{i\omega t} d\omega$$

[0237] 이들 시간 영역 함수들은 여기서는 연속-시간 HOA 계수 시퀀스라 부르며, 하기의 단일의 벡터 $\mathbf{c}(t)$ 로 집합될 수 있다

수학식 46

[0238]
$$\mathbf{c}(t) = \left[c_0^0(t) \quad c_1^{-1}(t) \quad c_1^0(t) \quad c_1^1(t) \quad c_2^{-2}(t) \quad c_2^{-1}(t) \quad c_2^0(t) \quad c_2^1(t) \quad c_2^2(t) \quad \dots \quad c_N^{N-1}(t) \quad c_N^N(t) \right]^T$$

[0239] 벡터 $\mathbf{c}(t)$ 내에서의 HOA 계수 시퀀스 $c_n^m(t)$ 의 위치 인덱스는 $n(n+1)+1+m$ 으로 주어진다.

[0240] 벡터 $\mathbf{c}(t)$ 내의 요소들의 전체 개수는 $O = (N+1)^2$ 으로 주어진다.

[0241] 최종 Ambisonics 포맷은 하기의 같은 샘플링 주파수 f_s 를 이용한 $\mathbf{c}(t)$ 의 샘플링된 버전을 제공한다

수학식 47

[0242] $\{\mathbf{c}(lT_s)\}_{l \in \mathbb{N}} = \{\mathbf{c}(T_s), \mathbf{c}(2T_s), \mathbf{c}(3T_s), \mathbf{c}(4T_s), \dots\}$

[0243] 여기서, $T_s = 1/f_s$ 는 샘플링 기간을 나타낸다. $\mathbf{c}(lT_s)$ 의 요소들은 여기서는 이산-시간 HOA 계수 시퀀스라 부르며, 항상 실수값인 것으로 보일 수 있다. 이 속성은 또한 명백히 연속-시간 버전 $c_n^m(t)$ 에 대해서도 유효하다.

[0244] 실수값 구면 고조파의 정의

[0245] 실수값 구면 고조파 $S_n^m(\theta, \phi)$ (SN3D 정규화 [1, Ch.3.1]을 가정)는 다음과 같이 주어진다

수학식 48

[0246]
$$S_n^m(\theta, \phi) = \sqrt{(2n+1) \frac{(n-|m|)!}{(n+|m|)!}} P_{n,|m|}(\cos\theta) \text{trg}_m(\phi)$$

[0247] 여기서,

수학식 49

[0248]
$$\text{trg}_m(\phi) = \begin{cases} \sqrt{2}\cos(m\phi) & m > 0 \\ 1 & m = 0 \\ -\sqrt{2}\sin(m\phi) & m < 0 \end{cases}$$

[0249] 연관된 르장드르 함수 $P_{n,m}(x)$ 는 르장드르 다항식 $P_n(x)$ 과 함께 다음과 같이 정의되며,

수학식 50

[0250]
$$P_{n,m}(x) = (1-x^2)^{m/2} \frac{d^m}{dx^m} P_n(x), m \geq 0$$

[0251] [11]에서와 달리, Condon-Shortley 위상 항 $(-1)^m$ 이 없다.

[0252] 한 실시예에서, (복소값 필터 뱅크로부터 획득되는) HOA 신호 표현의 부대역 또는 부대역 그룹 내의 우세 방향 신호들의 방향들의 프레임별 결정 및 효율적인 인코딩을 위한 방법은,

[0253] 각각의 현재 프레임 k에 대해: HOA 신호에서 전체 대역 방향 후보들의 세트 $M_{\text{DIR}}(k)$, 상기 세트 $M_{\text{DIR}}(k)$ 내의 요

소들의 수 NoOfGlobalDirs 및 상기 요소들의 수를 인코딩하는데 요구되는 수 $D(k)=\log_2(\text{NoOfGlobalDirs})$ 를 결정 - 각각의 전체 대역 방향 후보는, 미리정의된 전체 세트의 Q개의 가능한 방향들에 관련된 전역 지수 $q \in [1, \dots, Q]$ 를 가짐 - 하는 단계, 현재 프레임 k의 각각의 부대역 또는 부대역 그룹 j에 대해, 세트 $M_{\text{DIR}}(k)$ 내의 전체 대역 방향 후보들 중 어느 방향들이 활성 부대역 방향으로서 발생하는지를 결정하는 단계, 임의의 부대역 또는 부대역 그룹들에서 활성 부대역 방향으로서 발생하는 이용된 전체 대역 방향 후보들의 세트 $M_{\text{FB}}(k)$ (모두는 HOA 신호에서 전체 대역 방향 후보들의 세트 $M_{\text{DIR}}(k)$ 내에 포함됨) 및 이용된 전체 대역 방향 후보들의 세트 $M_{\text{FB}}(k)$ 내의 요소들의 수 NoOfGlobalDirs(k)를 결정하는 단계, 및 현재 프레임 k의 각각의 부대역 또는 부대역 그룹 j에 대해: 세트 $M_{\text{DIR}}(k)$ 내의 전체 대역 방향 후보들 중에서 $d \in [1, \dots, D]$ 개까지의 방향들 중 어느 방향들이 활성 부대역 방향인지를 결정하고, 활성 부대역 방향들 각각에 대해 궤적 및 궤적 인덱스를 결정하고, 궤적 인덱스를 각각의 활성 부대역 방향에 할당하며, 현재의 부대역 또는 부대역 그룹 j 내의 활성 부대역 방향들 각각을 D(k) 비트를 갖는 상대 인덱스에 의해 인코딩하는 단계를 포함한다.

[0254] 한 실시예에서, 컴퓨터 판독가능한 매체에는, 컴퓨터로 하여금 우세 방향 신호들의 방향들의 프레임별 결정과 효율적인 인코딩을 위한 이 방법을 수행하게 하는 실행가능한 명령어들이 저장되어 있다.

[0255] 또한, 한 실시예에서, HOA 신호 표현의 부대역들 내의 우세 방향 신호들의 방향들의 디코딩을 위한 방법은, 디코딩될 HOA 신호 표현에 대한 최대 개수의 방향들 D의 인덱스들을 수신하는 단계, 디코딩될 HOA 신호 표현의 최대 개수의 방향들 D의 방향들을 재구성하는 단계, 부대역당 활성 방향 신호들의 인덱스들을 수신하는 단계, 디코딩될 HOA 신호 표현의 재구성된 방향들 D로부터 부대역당 활성 방향들을 재구성하는 단계, 및 부대역들의 방향 신호들을 예측하는 단계를 포함하며, 여기서, 부대역의 현재 프레임에서 방향 신호를 예측하는 것은, 부대역의 선행 프레임의 방향 신호를 결정하는 것을 포함하고, 여기서, 방향 신호의 인덱스가 선행 프레임에서 제로이고 현재 프레임에서 비제로이면 새로운 방향 신호가 생성되고, 방향 신호의 인덱스가 선행 프레임에서 비제로이고 현재 프레임에서 제로이면 이전 방향 신호는 취소되고, 방향 신호의 인덱스가 제1 방향으로부터 제2 방향으로 변하면 방향 신호의 방향은 제1 방향으로부터 제2 방향으로 이동된다.

[0256] 한 실시예에서, 도 1 및 도 3에 도시되고 상기에서 논의된 바와 같이, 주어진 개수의 계수 시퀀스 - 각각의 계수 시퀀스는 인덱스를 가짐 - 를 갖는 입력 HOA 신호의 프레임들을 인코딩하기 위한 장치는, 적어도 하나의 하드웨어 프로세서 및 적어도 하나의 소프트웨어 컴포넌트를 유형적으로 구현하는 비일시적, 유형의 컴퓨터 판독 가능한 저장 매체를 포함하고, 적어도 하나의 소프트웨어 컴포넌트는 적어도 하나의 하드웨어 프로세서에서 실행될 때 하드웨어 프로세서로 하여금,

[0257] 감소된 개수의 비제로 계수 시퀀스들을 갖는 절삭된 HOA 표현 $C_T(k)$ 을 계산하고(11),

[0258] 절삭된 HOA 표현에서 포함되는 활성 계수 시퀀스들 $I_{C,ACT}(k)$ 의 인덱스들의 세트를 결정하며(11),

[0259] 입력 HOA 신호로부터 제1 세트의 후보 방향들 $M_{\text{DIR}}(k)$ 을 추정하고(16),

[0260] 입력 HOA 신호를 복수의 주파수 부대역 f_1, \dots, f_F 으로 분할 - 여기서, 주파수 부대역들의 계수 시퀀스들 $\bar{C}(k-1, k, f_1), \dots, \bar{C}(k-1, k, f_F)$ 이 획득됨 - 하며(15),

[0261] 주파수 부대역들 각각에 대해 제2 세트의 방향들 $M_{\text{DIR}}(k, f_1), \dots, M_{\text{DIR}}(k, f_F)$ 을 추정 - 제2 세트의 방향들의 각각의 요소는 제1 및 제2 인덱스를 갖는 인덱스들의 튜플이고, 제2 인덱스는 현재의 주파수 부대역에 대한 활성 방향의 인덱스이며 제1 인덱스는 활성 방향의 궤적 인덱스이고, 각각의 활성 방향은 또한 입력 HOA 신호의 제1 세트의 후보 방향들 $M_{\text{DIR}}(k)$ 에 포함됨 - 하고(16),

[0262] 주파수 부대역들 각각에 대해, 각각의 주파수 부대역의 제2 세트의 방향들 $M_{\text{DIR}}(k, f_1), \dots, M_{\text{DIR}}(k, f_F)$ 에 따라 주파수 부대역의 계수 시퀀스들 $\bar{C}(k-1, k, f_1), \dots, \bar{C}(k-1, k, f_F)$ 로부터 방향 부대역 신호들 $\bar{X}(k-1, k, f_1), \dots, \bar{X}(k-1, k, f_F)$ 을 계산하며(17),

[0263] 주파수 부대역들 각각에 대해, 각각의 주파수 부대역의 활성 계수 시퀀스들 $I_{C,ACT}(k)$ 의 인덱스들의 세트를 이용하

여 주파수 부대역의 계수 시퀀스들 $\bar{\mathbf{c}}(k-1, k, f_1), \dots, \bar{\mathbf{c}}(k-1, k, f_F)$ 로부터 방향 부대역 신호들 $\bar{\mathbf{X}}(k-1, k, f_1), \dots, \bar{\mathbf{X}}(k-1, k, f_F)$ 을 예측하기 위해 구성된 예측 행렬 $A(k, f_1), \dots, A(k, f_F)$ 을 계산하고(18),

[0264] 제1 세트의 후보 방향들 $M_{DIR}(k)$, 제2 세트의 방향들 $M_{DIR}(k, f_1), \dots, M_{DIR}(k, f_F)$, 예측 행렬들 $A(k, f_1), \dots, A(k, f_F)$ 및 절삭된 HOA 표현 $\mathbf{C}_T(k)$ 을 인코딩하게 한다.

[0265] 한 실시예에서, 도 4 및 도 5에 도시되고 상기에서 논의된 바와 같이, 압축된 HOA 표현을 디코딩하기 위한 장치는, 적어도 하나의 하드웨어 프로세서, 및 적어도 하나의 소프트웨어 컴포넌트를 구현하는 비일시적, 유형의, 컴퓨터 판독가능한 저장 매체를 포함하고, 적어도 하나의 소프트웨어 컴포넌트는, 적어도 하나의 하드웨어 프로세서에서 실행될 때, 하드웨어 프로세서로 하여금, 압축된 HOA 표현으로부터, 복수의 절삭된 HOA 계수 시퀀스들 $\hat{\mathbf{z}}_1(k), \dots, \hat{\mathbf{z}}_I(k)$, 상기 절삭된 HOA 계수 시퀀스들의 시퀀스 인덱스를 나타내거나 포함하는 할당 벡터 $\mathbf{v}_{AMB,ASSIGN}(k)$, 부대역 관련 방향 정보 $M_{DIR}(k+1, f_1), \dots, M_{DIR}(k+1, f_F)$, 복수의 예측 행렬 $A(k+1, f_1), \dots, A(k+1, f_F)$, 및 이득 제어 부가 정보 $\mathbf{e}_1(k), \beta_1(k), \dots, \mathbf{e}_I(k), \beta_I(k)$ 를 추출하고(41, 42, 43);

[0266] 복수의 절삭된 HOA 계수 시퀀스들 $\hat{\mathbf{z}}_1(k), \dots, \hat{\mathbf{z}}_I(k)$, 이득 제어 부가 정보 $\mathbf{e}_1(k), \beta_1(k), \dots, \mathbf{e}_I(k), \beta_I(k)$ 및 할당 벡터 $\mathbf{v}_{AMB,ASSIGN}(k)$ 로부터, 절삭된 HOA 표현 $\hat{\mathbf{C}}_T(k)$ 을 재구성하며(51, 52),

[0267] 하나 이상의 분석 필터 बैं크(53)에서, 재구성된 절삭된 HOA 표현 $\hat{\mathbf{C}}_T(k)$ 을 복수의 F개의 주파수 부대역에 대한 주파수 부대역 표현들 $\tilde{\mathbf{C}}_T(k, f_1), \dots, \tilde{\mathbf{C}}_T(k, f_F)$ 로 분해하고,

[0268] 방향 부대역 합성 블록(54)에서, 주파수 부대역 표현들 각각에 대해, 재구성된 절삭된 HOA 표현의 각각의 주파수 부대역 표현 $\tilde{\mathbf{C}}_T(k, f_1), \dots, \tilde{\mathbf{C}}_T(k, f_F)$, 부대역 관련 방향 정보 $M_{DIR}(k+1, f_1), \dots, M_{DIR}(k+1, f_F)$, 및 예측 행렬들 $A(k+1, f_1), \dots, A(k+1, f_F)$ 로부터, 예측된 방향 HOA 표현 $\tilde{\mathbf{C}}_D(k, f_1), \dots, \tilde{\mathbf{C}}_D(k, f_F)$ 을 합성하고(54),

[0269] 부대역 조성 블록(55)에서, F개의 주파수 부대역들 각각에 대해, 계수 시퀀스가 할당 벡터 $\mathbf{v}_{AMB,ASSIGN}(k)$ 에 포함되는 인덱스 n을 갖는 경우 절삭된 HOA 표현 $\tilde{\mathbf{C}}_T(k, f_j)$ 의 계수 시퀀스들로부터 획득되거나, 그렇지 않으면 방향 부대역 합성 블록(54)들 중 하나에 의해 제공된 예측된 방향 HOA 성분 $\tilde{\mathbf{C}}_D(k, f_j)$ 의 계수 시퀀스들로부터 획득되는 계수 시퀀스들 $\tilde{\mathbf{c}}_n(k, f_j)$, $n = 1, \dots, O$ 을 갖는 디코딩된 부대역 HOA 표현 $\tilde{\mathbf{c}}(k, f_1), \dots, \tilde{\mathbf{c}}(k, f_F)$ 을 조성하고(55), 하나 이상의 합성 필터 बैं크(56)에서, 디코딩된 HOA 표현 $\tilde{\mathbf{C}}(k)$ 을 획득하기 위해 디코딩된 부대역 HOA 표현들 $\tilde{\mathbf{c}}(k, f_1), \dots, \tilde{\mathbf{c}}(k, f_F)$ 을 합성하게 한다.

[0270] 한 실시예에서, 주어진 개수의 계수 시퀀스들 - 각각의 계수 시퀀스는 인덱스를 가짐 - 을 갖는 입력 HOA 신호의 프레임들을 인코딩하기 위한 장치(10)는, 감소된 개수의 비제로 계수 시퀀스들을 갖는 절삭된 HOA 표현 $\mathbf{C}_T(k)$ 을 계산하도록 구성되고, 또한 절삭된 HOA 표현에 포함된 활성 계수 시퀀스들($I_{c,ACT}(k)$)의 인덱스들의 세트를 결정하도록 구성된 계산 및 결정 모듈(11);

[0271] 입력 HOA 신호를 복수의 주파수 부대역 f_1, \dots, f_F 으로 분할 - 여기서, 주파수 부대역들의 계수 시퀀스들 $\bar{\mathbf{c}}(k-1, k, f_1), \dots, \bar{\mathbf{c}}(k-1, k, f_F)$ 이 획득됨 - 하도록 구성된 분석 필터 बैं크 모듈(15);

[0272] 입력 방향 HOA 신호로부터 제1 세트의 후보 방향들 $M_{DIR}(k)$ 을 추정하도록 구성되고, 또한 주파수 부대역들 각각에 대해 제2 세트의 방향들 $M_{DIR}(k, f_1), \dots, M_{DIR}(k, f_F)$ 을 추정하도록 - 제2 세트의 방향들의 각각의 요소는 제1

및 제2 인덱스를 갖는 인덱스들의 튜플이고, 제2 인덱스는 현재 주파수 부대역에 대한 활성 방향의 인덱스이며, 제1 인덱스는 활성 방향의 궤적 인덱스이고, 각각의 활성 방향은 또한 입력 HOA 신호의 제1 세트의 후보 방향들 $M_{DIR}(k)$ 에 포함됨 - 구성된 방향 추정 모듈(16); 주파수 부대역들 각각에 대해, 각각의 주파수 부대역의 제2 세트의 방향들 $M_{DIR}(k, f_1), \dots, M_{DIR}(k, f_F)$ 에 따라 주파수 부대역의 계수 시퀀스들

$\bar{C}(k-1, k, f_1), \dots, \bar{C}(k-1, k, f_F)$ 로부터 방향 부대역 신호들 $\bar{X}(k-1, k, f_1), \dots, \bar{X}(k-1, k, f_F)$ 을 계산하도록 구성된 적어도 하나의 방향 부대역 계산 모듈(17); 주파수 부대역들 각각에 대해, 각각의 주파수 부대역의 활성 계수 시퀀스들 $I_{C,ACT}(k)$ 의 인덱스들의 세트를 이용하여 주파수 부대역의 계수 시퀀스들

$\bar{C}(k-1, k, f_1), \dots, \bar{C}(k-1, k, f_F)$ 로부터, 방향 부대역 신호들 $\bar{X}(k-1, k, f_1), \dots, \bar{X}(k-1, k, f_F)$ 을 예측하기 위해 구성된 예측 행렬 $A(k, f_1), \dots, A(k, f_F)$ 을 계산하도록 구성된 적어도 하나의 방향 부대역 예측 모듈(18); 및 제1 세트의 후보 방향들 $M_{DIR}(k)$, 제2 세트의 방향들 $M_{DIR}(k, f_1), \dots, M_{DIR}(k, f_F)$, 예측 행렬들 $A(k, f_1), \dots, A(k, f_F)$, 및 절삭된 HOA 표현 $C_T(k)$ 을 인코딩하도록 구성된 인코딩 모듈(30)을 포함한다.

[0273] 한 실시예에서, 이 장치는, 절삭된 HOA 채널 시퀀스들을 부분적으로 무상관화하도록 구성된 부분적 무상관화기(12); 절삭된 HOA 채널 시퀀스들 $y_1(k), \dots, y_I(k)$ 을 전송 채널들에 할당하도록 구성된 채널 할당 모듈(13); 및 전송 채널들에 관한 이득 제어를 수행 - 여기서, 각각의 전송 채널에 대한 이득 제어 부가 정보 $e_i(k-1), \beta_i(k-1)$ 가 생성됨 - 하도록 구성된 적어도 하나의 이득 제어 유닛(14)을 더 포함한다.

[0274] 한 실시예에서, 인코딩 모듈(30)은, 이득 제어된 절삭된 HOA 채널 시퀀스 $z_1(k), \dots, z_I(k)$ 를 인코딩하도록 구성된 인지 인코더(31);

[0275] 이득 제어 부가 정보 $e_i(k-1), \beta_i(k-1)$, 제1 세트의 후보 방향들 $M_{DIR}(k)$, 제2 세트의 방향들 $M_{DIR}(k, f_1), \dots, M_{DIR}(k, f_F)$, 및 예측 행렬들 $A(k, f_1), \dots, A(k, f_F)$ 을 인코딩하도록 구성된 부가 정보 소스 코더(32); 및

[0276] 인지 인코더(31) 및 부가 정보 소스 코더(32)의 출력들을 멀티플렉싱하여 인코딩된 HOA 신호 프레임 $\bar{B}(k-1)$ 을 획득하도록 구성된 멀티플렉서(33)를 포함한다.

[0277] 한 실시예에서, HOA 신호를 디코딩하기 위한 장치(50)는,

[0278] 압축된 HOA 표현으로부터 방향 정보를 디코딩하기 위한 장치는, 압축된 HOA 표현으로부터, 복수의 절삭된 HOA 계수 시퀀스들 $\hat{z}_1(k), \dots, \hat{z}_I(k)$, 상기 절삭된 HOA 시퀀스들의 시퀀스 인덱스들을 나타내거나 포함하는 할당 벡터 $v_{AMB,ASSIGN}(k)$, 부대역 관련 방향 정보 $M_{DIR}(k+1, f_1), \dots, M_{DIR}(k+1, f_F)$, 복수의 예측 행렬 $A(k+1, f_1), \dots, A(k+1, f_F)$, 및 이득 제어 부가 정보 $e_1(k), \beta_1(k), \dots, e_I(k), \beta_I(k)$ 를 추출하도록 구성된 추출 모듈(40); 복수의 절삭된 HOA 계수 시퀀스 $\hat{z}_1(k), \dots, \hat{z}_I(k)$, 이득 제어 부가 정보 $e_1(k), \beta_1(k), \dots, e_I(k), \beta_I(k)$ 및 할당 벡터 $v_{AMB,ASSIGN}(k)$ 로부터 절삭된 HOA 표현 $\hat{C}_T(k)$ 을 재구성하도록 구성된 재구성 모듈(51, 52); 재구성된 절삭된 HOA 표현 $\hat{C}_T(k)$ 을 복수의 F개의 주파수 부대역에 대한 주파수 부대역 표현들 $\tilde{C}_T(k, f_1), \dots, \tilde{C}_T(k, f_F)$ 로 분해하도록 구성된 분석 필터 뱅크 모듈(53);

[0279] 주파수 부대역 표현들 각각에 대해, 재구성된 절삭된 HOA 표현의 각각의 주파수 부대역 표현 $\tilde{C}_T(k, f_1), \dots, \tilde{C}_T(k, f_F)$, 부대역 관련 방향 정보 $M_{DIR}(k+1, f_1), \dots, M_{DIR}(k+1, f_F)$, 및 예측 행렬들 $A(k+1, f_1), \dots, A(k+1, f_F)$ 로부터, 예측된 방향 HOA 표현 $\tilde{C}_D(k, f_1), \dots, \tilde{C}_D(k, f_F)$ 을 합성하도록 구성된 적어도 하나의 방향 부대역 합성 모듈(54);

- [0280] F개의 주파수 부대역들 각각에 대해, 계수 시퀀스가 할당 벡터 $\mathbf{v}_{\text{AMBASSIGN}}(k)$ 에 포함되는 인덱스 n을 갖는 경우 절삭된 HOA 표현 $\tilde{\mathbf{c}}_T(k, f_j)$ 의 계수 시퀀스들로부터 획득되거나, 그렇지 않으면 방향 부대역 합성 모듈(54)들 중 하나에 의해 제공된 예측된 방향 HOA 성분 $\tilde{\mathbf{c}}_D(k, f_j)$ 의 계수 시퀀스들로부터 획득되는 계수 시퀀스들 $\tilde{\mathbf{c}}_n(k, f_j)$, $n = 1, \dots, O$ 을 갖는 디코딩된 부대역 HOA 표현 $\tilde{\mathbf{c}}(k, f_1), \dots, \tilde{\mathbf{c}}(k, f_F)$ 을 조성하도록 구성된 적어도 하나의 부대역 조성 모듈(55); 및
- [0281] 디코딩된 HOA 표현 $\tilde{\mathbf{c}}(k)$ 을 획득하기 위해 디코딩된 부대역 HOA 표현들 $\tilde{\mathbf{c}}(k, f_1), \dots, \tilde{\mathbf{c}}(k, f_F)$ 을 합성하도록 구성된 합성 필터 뱅크 모듈(56)을 포함한다.
- [0282] 한 실시예에서, 추출 모듈(40)은, 적어도, 인코딩된 부가 정보 부분 및 인코딩된 절삭된 HOA 계수 시퀀스들 $\check{\mathbf{z}}_1(k), \dots, \check{\mathbf{z}}_I(k)$ 을 포함하는 인지 코딩된 부분을 획득하기 위한 디멀티플렉서(41); 인코딩된 절삭된 HOA 계수 시퀀스들 $\check{\mathbf{z}}_1(k), \dots, \check{\mathbf{z}}_I(k)$ 을 인지 디코딩(s42)하여 절삭된 HOA 계수 시퀀스들 $\hat{\mathbf{z}}_1(k), \dots, \hat{\mathbf{z}}_I(k)$ 을 획득하도록 구성된 인지 디코더(42); 및 인코딩된 부가 정보 부분을 디코딩(s43)하여, 부대역 관련 방향 정보 $M_{\text{DIR}}(k+1, f_1), \dots, M_{\text{DIR}}(k+1, f_F)$, 및 예측 행렬들 $A(k+1, f_1), \dots, A(k+1, f_F)$, 이득 제어 부가 정보 $\mathbf{e}_1(k), \beta_1(k), \dots, \mathbf{e}_I(k), \beta_I(k)$ 및 할당 벡터 $\mathbf{v}_{\text{AMBASSIGN}}(k)$ 를 획득하도록 구성된 부가 정보 소스 디코더(43)를 포함한다.
- [0283] 도 13은 한 실시예에서의 낮은 비트 레이트 인코딩 방법의 플로차트를 도시한다. 주어진 개수의 계수 시퀀스들 - 각각의 계수 시퀀스는 인덱스를 가짐 - 을 갖는 입력 HOA 신호의 프레임들을 낮은 비트 레이트 인코딩하기 위한 방법은,
- [0284] 감소된 개수의 비제로 계수 시퀀스들을 갖는 절삭된 HOA 표현 $\mathbf{C}_T(k)$ 을 계산하는 단계 (s110), 절삭된 HOA 표현에 포함된 활성 계수 시퀀스들 $I_{\text{C,ACT}}(k)$ 의 인덱스들의 세트를 결정하는 단계(s111), 입력 HOA 신호로부터 제1 세트의 후보 방향들 $M_{\text{DIR}}(k)$ 을 추정하는 단계(s16), 입력 HOA 신호를 복수의 주파수 부대역들 f_1, \dots, f_F 로 분할 - 여기서, 주파수 부대역들의 계수 시퀀스들 $\bar{\mathbf{c}}(k-1, k, f_1), \dots, \bar{\mathbf{c}}(k-1, k, f_F)$ 이 획득됨 - 하는 단계(s15), 주파수 부대역들 각각에 대해 제2 세트의 방향들 $M_{\text{DIR}}(k, f_1), \dots, M_{\text{DIR}}(k, f_F)$ - 제2 세트의 방향들 각각의 요소는 제1 및 제2 인덱스를 갖는 인덱스들의 튜플이고, 제2 인덱스는 현재의 주파수 부대역에 대한 활성 방향의 인덱스이며, 제1 인덱스는 활성 방향의 궤적 인덱스이고, 각각의 활성 방향은 또한 입력 HOA 신호의 제1 세트의 후보 방향들 $M_{\text{DIR}}(k)$ 에 포함됨 - 을 추정하는 단계(s161),
- [0285] 주파수 부대역들 각각에 대해, 각각의 주파수 부대역의 제2 세트의 방향들 $M_{\text{DIR}}(k, f_1), \dots, M_{\text{DIR}}(k, f_F)$ 에 따라 주파수 부대역의 계수 시퀀스들 $\bar{\mathbf{c}}(k-1, k, f_1), \dots, \bar{\mathbf{c}}(k-1, k, f_F)$ 로부터 방향 부대역 신호들 $\bar{\mathbf{X}}(k-1, k, f_1), \dots, \bar{\mathbf{X}}(k-1, k, f_F)$ 을 계산하는 단계(17),
- [0286] 주파수 부대역들 각각에 대해, 각각의 주파수 부대역의 활성 계수 시퀀스들 $I_{\text{C,ACT}}(k)$ 의 인덱스들의 세트를 이용하여 주파수 부대역의 계수 시퀀스들 $\bar{\mathbf{c}}(k-1, k, f_1), \dots, \bar{\mathbf{c}}(k-1, k, f_F)$ 로부터 방향 부대역 신호들 $\bar{\mathbf{X}}(k-1, k, f_1), \dots, \bar{\mathbf{X}}(k-1, k, f_F)$ 을 예측하기 위해 구성된 예측 행렬 $A(k, f_1), \dots, A(k, f_F)$ 을 계산하는 단계 (s18), 및 제1 세트의 후보 방향 $M_{\text{DIR}}(k)$, 제2 세트의 방향들 $M_{\text{DIR}}(k, f_1), \dots, M_{\text{DIR}}(k, f_F)$, 예측 행렬들 $A(k, f_1), \dots, A(k, f_F)$ 및 절삭된 HOA 표현 $\mathbf{C}_T(k)$ 을 인코딩하는 단계(s19)를 포함한다.
- [0287] 한 실시예에서, 상기 절삭된 HOA 표현 $\mathbf{C}_T(k)$ 을 인코딩하는 단계는,

- [0288] 절삭된 HOA 채널 시퀀스들의 부분적 무상관화(s12), 절삭된 HOA 채널 시퀀스들 $y_1(k), \dots, y_I(k)$ 을 전송 채널들에 할당하기 위한 채널 할당(s13), 전송 채널들 각각에 관한 이득 제어의 수행 -여기서, 각각의 전송 채널에 대한 이득 제어 부가 정보 $e_i(k-1), \beta_i(k-1)$ 가 생성됨-하는 단계(s14),
- [0289] 이득 제어된 절삭된 HOA 채널 시퀀스들 $z_1(k), \dots, z_I(k)$ 을 인지 인코더(31)에서 인코딩하는 단계(s31), 이득 제어 부가 정보 $e_i(k-1), \beta_i(k-1)$, 제1 세트의 후보 방향들 $M_{DIR}(k)$, 제2 세트의 방향들 $M_{DIR}(k, f_1), \dots, M_{DIR}(k, f_F)$, 및 예측 행렬들 $A(k, f_1), \dots, A(k, f_F)$ 을 부가 정보 소스 코더(32)에서 인코딩하는 단계(s32), 및 인지 인코더(31)와 부가 정보 소스 코더(32)의 출력들을 멀티플렉싱(s33)하여 인코딩된 HOA 신호 프레임 $\tilde{B}(k-1)$ 을 획득하는 단계를 포함한다.
- [0290] 한 실시예에서, 주어진 개수의 계수 시퀀스들 -각각의 계수 시퀀스는 인덱스를 가짐-을 갖는 입력 HOA 신호의 프레임들을 인코딩하기 위한 장치는, 프로세서와, 프로세서에 의해 실행될 때 프로세서로 하여금 청구항 8의 단계들을 수행하게 하는 명령어들을 저장한 메모리를 포함한다.
- [0291] 도 14는 한 실시예에서의 디코딩 방법의 플로차트를 도시한다. 낮은 비트 레이트 압축된 HOA 표현을 디코딩하기 위한 이 방법은, 압축된 HOA 표현으로부터 복수의 절삭된 HOA 계수 시퀀스들 $\hat{z}_1(k), \dots, \hat{z}_I(k)$, 상기 절삭된 HOA 계수 시퀀스들의 시퀀스 인덱스들을 나타내는(또는 포함하는) 할당 벡터 $v_{AMB.ASSIGN}(k)$, 부대역 관련 방향 정보 $M_{DIR}(k+1, f_1), \dots, M_{DIR}(k+1, f_F)$, 복수의 예측 행렬 $A(k+1, f_1), \dots, A(k+1, f_F)$, 및 이득 제어 부가 정보 $e_1(k), \beta_1(k), \dots, e_I(k), \beta_I(k)$ 를 추출하는 단계(s41, s42, s43),
- [0292] 복수의 절삭된 HOA 계수 시퀀스들 $\hat{z}_1(k), \dots, \hat{z}_I(k)$, 이득 제어 부가 정보 $e_1(k), \beta_1(k), \dots, e_I(k), \beta_I(k)$, 및 할당 벡터 $v_{AMB.ASSIGN}(k)$ 로부터, 절삭된 HOA 표현 $\hat{C}_T(k)$ 을 재구성하는 단계(s51, s52), 분석 필터뱅크(53)에서, 재구성된 절삭된 HOA 표현 $\hat{C}_T(k)$ 을 복수의 F개의 주파수 부대역에 대한 주파수 부대역 표현들 $\tilde{\hat{C}}_T(k, f_1), \dots, \tilde{\hat{C}}_T(k, f_F)$ 로 분해하는 단계(s53), 방향 부대역 합성 블록(54)에서, 주파수 부대역 표현들 각각에 대해, 재구성된 절삭된 HOA 표현의 각각의 주파수 부대역 표현 $\tilde{\hat{C}}_T(k, f_1), \dots, \tilde{\hat{C}}_T(k, f_F)$, 부대역 관련 방향 정보 $M_{DIR}(k+1, f_1), \dots, M_{DIR}(k+1, f_F)$, 및 예측 행렬들 $A(k+1, f_1), \dots, A(k+1, f_F)$ 로부터, 예측된 방향 HOA 표현 $\tilde{\hat{C}}_D(k, f_1), \dots, \tilde{\hat{C}}_D(k, f_F)$ 을 합성하는 단계(s54),
- [0293] 부대역 조성 블록(55)에서, F개의 주파수 부대역을 각각에 대해, 계수 시퀀스가 할당 벡터 $v_{AMB.ASSIGN}(k)$ 에 포함되는 인덱스 n을 갖는 경우 절삭된 HOA 표현 $\tilde{\hat{C}}_T(k, f_j)$ 의 계수 시퀀스들로부터 획득되거나, 그렇지 않으면 방향 부대역 합성 블록(54)들 중 하나에 의해 제공된 예측된 방향 HOA 성분 $\tilde{\hat{C}}_D(k, f_j)$ 의 계수 시퀀스들로부터 획득되는 계수 시퀀스들 $\tilde{\hat{c}}_n(k, f_j)$, $n = 1, \dots, O$ 을 갖는 디코딩된 부대역 HOA 표현 $\tilde{\hat{c}}(k, f_1), \dots, \tilde{\hat{c}}(k, f_F)$ 을 조성하는 단계(s55), 합성 필터뱅크(56)에서, 디코딩된 HOA 표현 $\tilde{\hat{c}}(k)$ 을 획득하기 위해 디코딩된 부대역 HOA 표현들 $\tilde{\hat{c}}(k, f_1), \dots, \tilde{\hat{c}}(k, f_F)$ 을 합성하는 단계(s56)를 포함한다.
- [0294] 실시예에서, 추출하는 단계는, 인지 코딩된 부분과 인코딩된 부가 정보 부분을 획득하기 위해 압축된 HOA 표현을 디멀티플렉싱하는 단계(s41), 인코딩된 절삭된 HOA 계수 시퀀스들을 인지 디코딩하는 단계(s42), 및 인코딩된 부가 정보를 부가 정보 소스 디코더(43)에서 디코딩하는 단계(s43) 중 하나 이상을 포함한다. 실시예에서, 복수의 절삭된 HOA 계수 시퀀스들로부터 절삭된 HOA 표현 $\hat{C}_T(k)$ 을 재구성하는 단계는, 역 이득 제어를 수행하는 단계(s51)와 절삭된 HOA 표현 $\hat{C}_T(k)$ 을 재구성하는 단계(s52) 중 하나 이상을 포함한다.

[0295] 한 실시예에서, 컴퓨터 판독가능한 매체에는, 컴퓨터로 하여금 우세 방향 신호들의 방향들을 디코딩하기 위한 상기 방법을 수행하게 하는 실행가능한 명령어들이 저장되어 있다.

[0296] 한 실시예에서, 압축된 HOA 신호를 디코딩하기 위한 장치는, 프로세서와, 프로세서에 의해 실행될 때 프로세서로 하여금 청구항 1의 단계들을 수행하게 하는 명령어들을 저장한 메모리를 포함한다.

[0297] 동일한 결과를 달성하기 위해 실질적으로 동일한 방식으로 실질적으로 동일한 기능을 수행하는 요소들의 모든 조합은 본 발명의 범위 내에 있고, 설명 및 (적절한 경우) 청구항들에 개시된 각각의 피쳐 및 도면들은 독립적으로 또는 임의의 적절한 조합으로 제공될 수 있다는 것을 명시적으로 의도한다. 특징들은, 적절하다면, 하드웨어, 소프트웨어, 또는 이들의 조합으로 구현될 수도 있다. 접속은, 적용가능하다면, 무선 접속 또는 유선으로서 구현될 수 있고, 반드시 직접적 또는 전용 접속일 필요는 없다. 한 실시예에서, 추출 모듈, 이득 제어 유닛, 부대역 신호 그룹화 유닛, 처리 유닛 및 기타의 것들과 같은, 앞서 언급된 모듈들 또는 유닛들 각각은, 적어도 부분적으로, 적어도 하나의 실리콘 컴포넌트를 이용함으로써 하드웨어로 구현된다.

[0298] [참조문헌]

[1] Jérôme Daniel. *Représentation de champs acoustiques, application à la transmission et à la reproduction de scènes sonores complexes dans un contexte multimédia*. PhD thesis, Université Paris 6, 2001.

[2] Jörg Fliege and Ulrike Maier. A two-stage approach for computing cubature formulae for the sphere. Technical report, Fachbereich Mathematik, Universität Dortmund, 1999. Node numbers are found at <http://www.mathematik.uni-dortmund.de/lxx/research/projects/fliege/nodes/nodes.html>.

[3] Sven Kordon and Alexander Krueger. Adaptive value range control for HOA signals. Patent application (Technicolor Internal Reference: PD130016), July 2013.

[4] Alexander Krueger and Sven Kordon. Intelligent signal extraction and packing for compression of HOA sound field representations. Patent application EP 13305558.2 (Technicolor Internal Reference: PD130015), filed 29. April 2013.

[5] A. Krueger, S. Kordon, and J. Boehm. HOA compression by decomposition into directional and ambient components. Published patent application EP2743922 (Technicolor Internal Reference: PD120055), December 2012.

[6] Alexander Krüger, Sven Kordon, Johannes Boehm, and Jan-Mark Batke. Method and apparatus for compressing and decompressing a higher order ambisonics signal representation. Published patent application EP2665208 (Technicolor Internal Reference: PD120015), May 2012.

[7] Alexander Krüger. Method and apparatus for robust sound source direction tracking based on Higher Order Ambisonics. Published patent application EP2738962 (Technicolor Internal Reference: PD120049), November 2012.

[8] Daniel D. Lee and H. Sebastian Seung. Learning the parts of objects by nonnegative matrix factorization. *Nature*, 401:788–791, 1999.

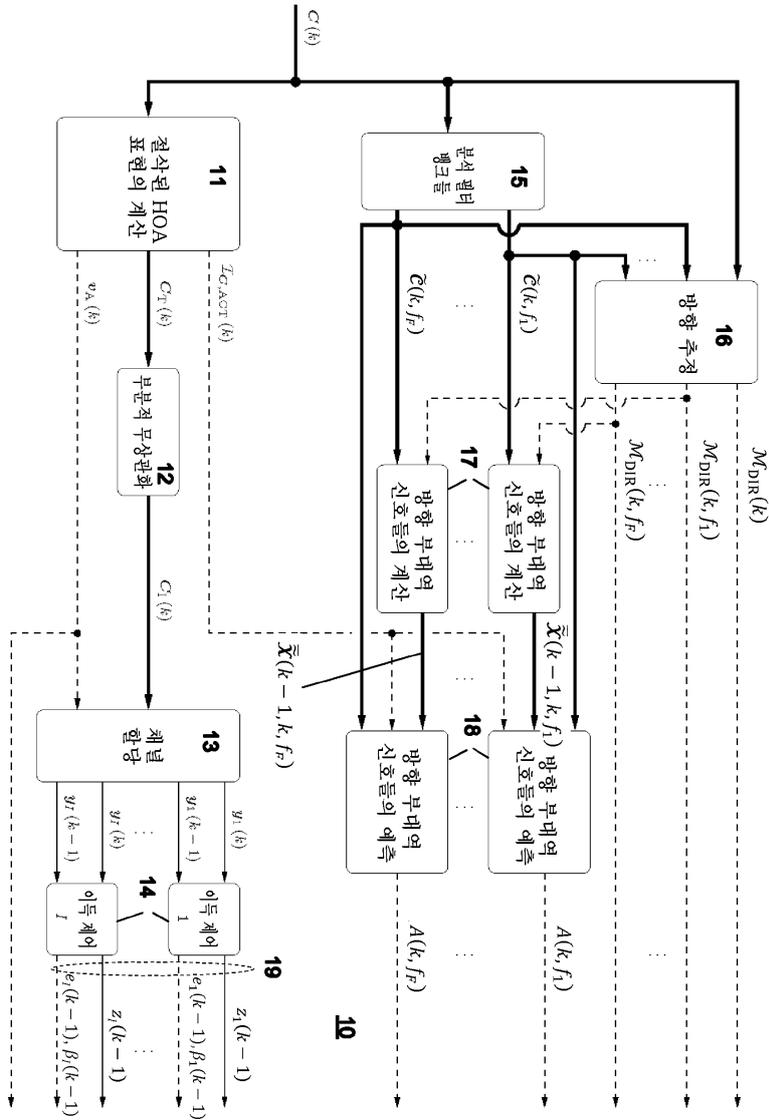
[9] ISO/IEC JTC 1/SC 29 N. Text of ISO/IEC 23008-3/CD, MPEG-H 3d audio, April 2014.

[10] Boaz Rafaely. Plane-wave decomposition of the sound field on a sphere by spherical convolution. *J. Acoust. Soc. Am.*, 4(116):2149–2157, October 2004.

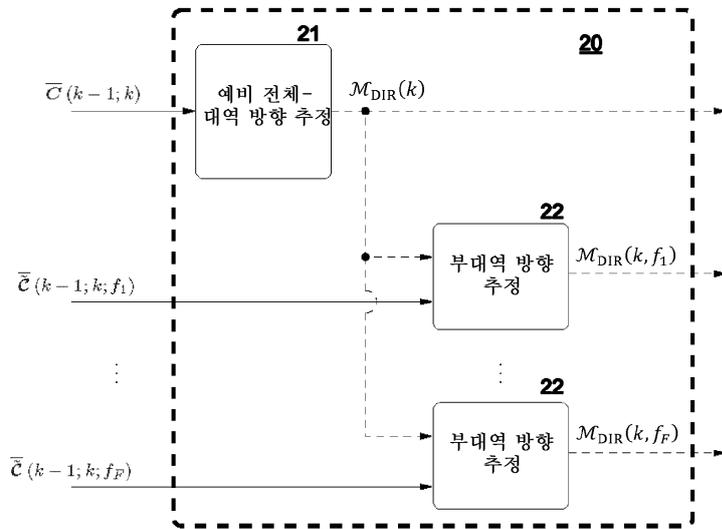
[11] Earl G. Williams. *Fourier Acoustics*, volume 93 of *Applied Mathematical Sciences*. Academic Press, 1999.

[0299]

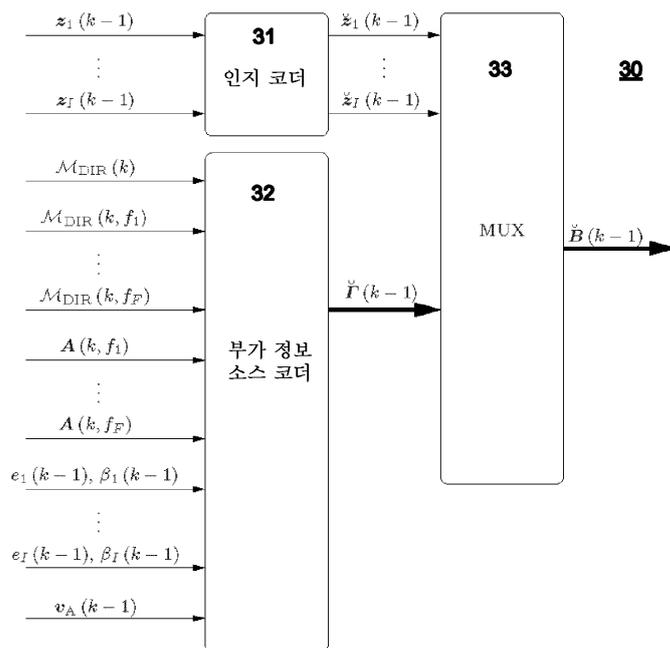
도면
도면1



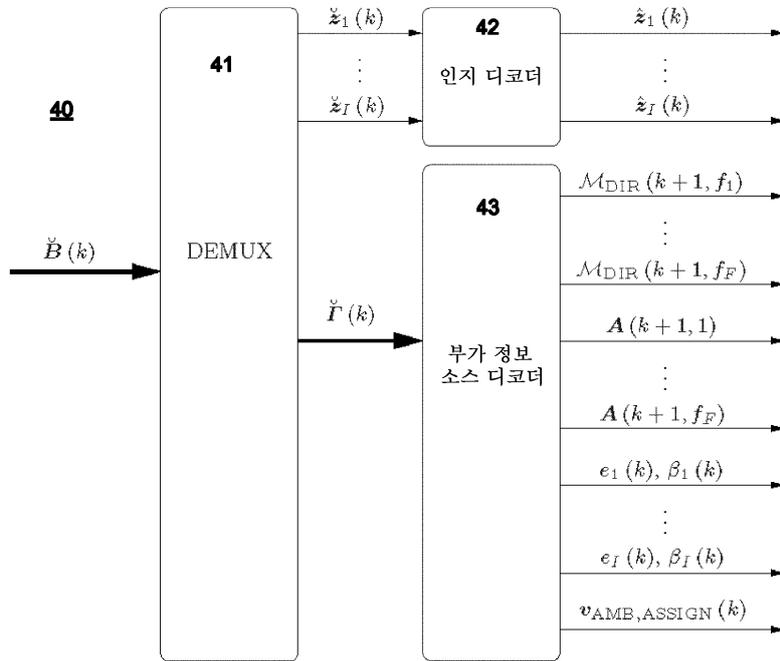
도면2



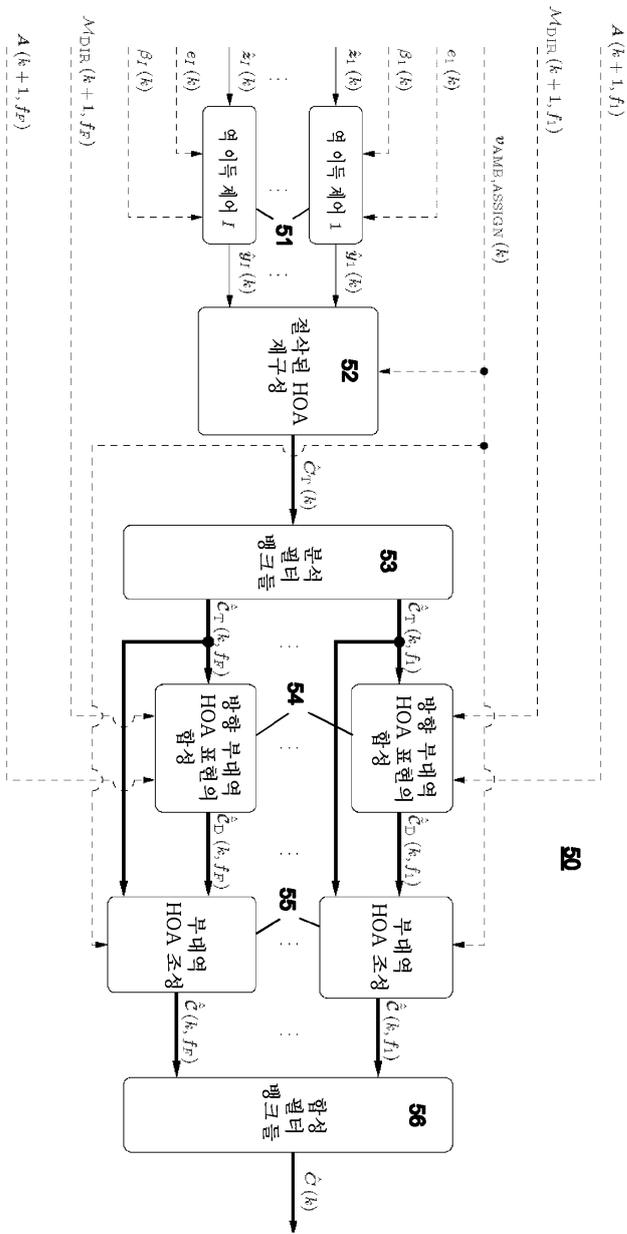
도면3



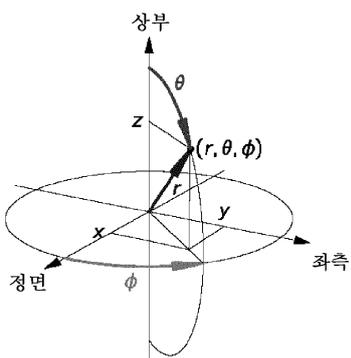
도면4



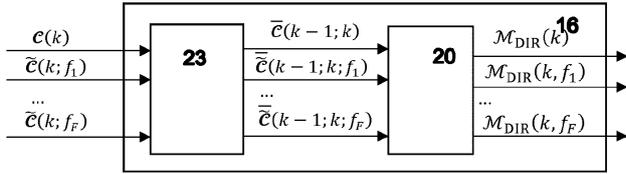
도면5



도면6

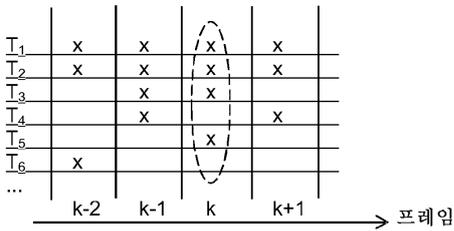


도면7



도면8

$$M_{DIR}(k) = \{\Omega_3, \Omega_8, \Omega_{52}, \Omega_{101}, \Omega_{229}, \Omega_{446}, \Omega_{581}\}$$

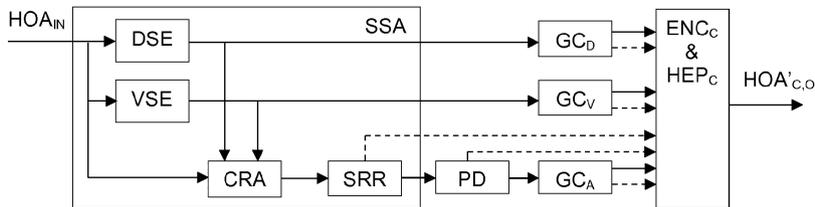


$$M_{DIR}(k, f_1) = \{(1, \Omega_{52}), (2, \Omega_{229}), (3, \Omega_3), (5, \Omega_{581})\}$$

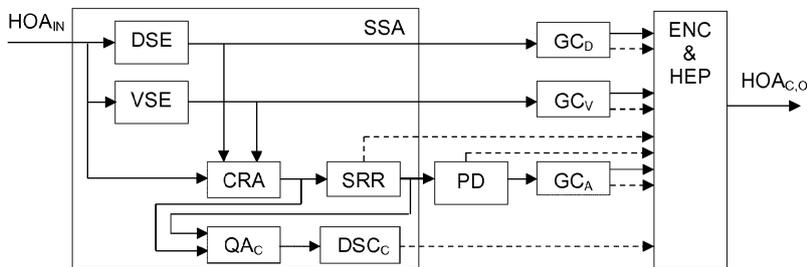
$$M_{DIR}(k, f_2) = \{(1, \Omega_{52}), (2, \Omega_{229})\}$$

...

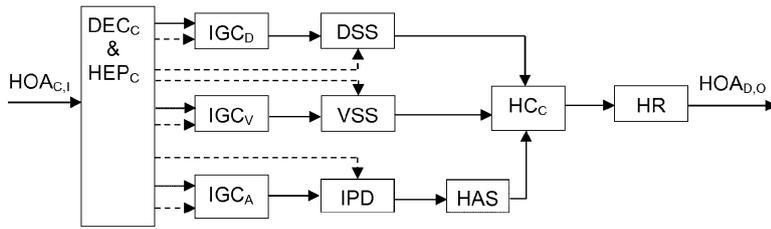
도면9



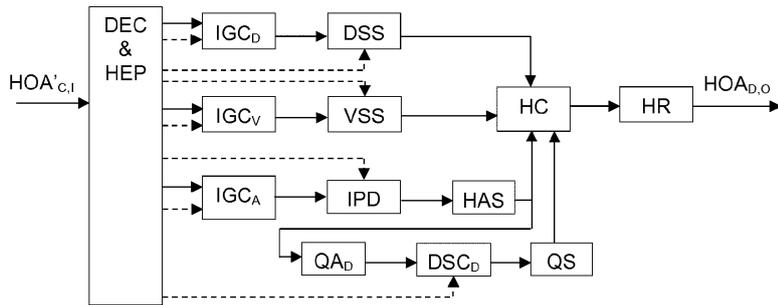
도면10



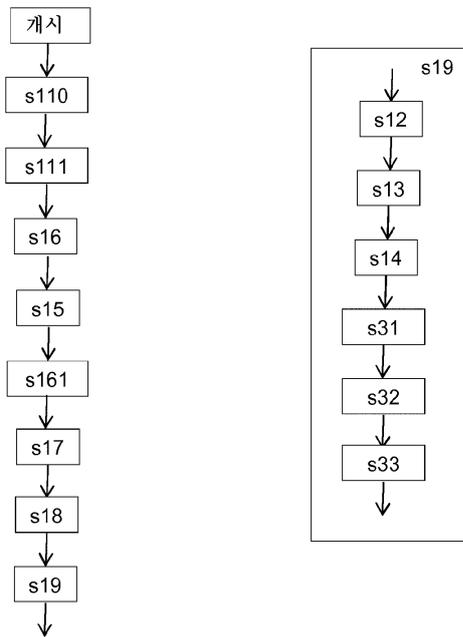
도면11



도면12



도면13



도면14

