(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2022년04월01일
(11) 등록번호 10-2381202
(24) 등록일자 2022년03월28일

(51) 국제특허분류(Int. Cl.)
H04S 3/02 (2006.01) *G10L 19/008* (2014.01)
(52) CPC특허분류
H04S 3/02 (2013.01)
G10L 19/008 (2020.08)
(21) 출원번호 10-2016-7036547
(22) 출원일자(국제) 2015년06월22일
심사청구일자 2020년06월19일
(85) 번역문제출일자 2016년12월27일
(65) 공개번호 10-2017-0023867
(43) 공개일자 2017년03월06일
(86) 국제출원번호 PCT/EP2015/063914
(87) 국제공개번호 WO 2015/197514
국제공개일자 2015년12월30일
(30) 우선권주장
14306024.2 2014년06월27일
유럽특허청(EPO)(EP)
(56) 선행기술조사문헌
KR1020120070521 A
WO2013171083 A1

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

전체 청구항 수 : 총 18 항 심사관 : 권영학

(54) 발명의 명칭 **HOA 데이터 프레임 표현의 압축을 위해 비차분 이득 값들을 표현하는 데 필요하게 되는 비트들의 최저 정수 개수를 결정하는 장치**

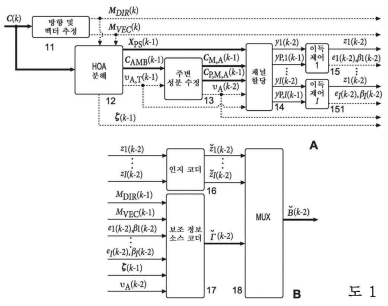
(57) 요 약

HOA 데이터 프레임 표현을 압축할 때, 각각의 채널 신호가 인지 인코딩(16)되기 전에 각각의 채널 신호에 대해 이득 제어(15, 151)가 적용된다. 이득 값들이 차분 방식으로 보조 정보로서 전송된다. 그렇지만, 이러한 스트리밍되는 압축된 HOA 데이터 프레임 표현의 디코딩을 시작하기 위해, 절대 이득 값들이 필요하게 되고, 절대 이득 값들은 최소 수의 비트들로 코딩되어야만 한다. 이러한 비트들의 최저 정수 개수(β_e)를 결정하기 위해, HOA 데이터 프레임 표현($\boldsymbol{C}(k)$)이 공간 영역에서 단위 구면 상에 있는 가상 스피커 신호들로 렌더링되고, 이어서 HOA 데이터 프레임 표현($\boldsymbol{C}(k)$)의 정규화가 뒤따른다. 이어서 비트들의 최저 정수 개수가 (AA)로 설정된다.

대 표 도

(AA)

$$\beta_e = \left\lceil \log_2\left(\left\lfloor \log_A\left(\sqrt{N_{MAX}}\cdot O\right)\right\rfloor + 1\right)\right\rceil$$



도 1

(52) CPC특허분류
H04S 2420/11 (2013.01)

명 세 서

청구범위

청구항 1

HOA 데이터 프레임 표현$(\boldsymbol{C}(k))$의 압축을 위해 HOA 데이터 프레임들의 채널 신호들에 대해 진폭 변화들에 대응하는 비차분 이득 값들의 표현들을 2의 지수(2^e)로서 나타내기 위한 비트들의 최저 정수 개수 β_e를 결정하는 방법으로서,

각각의 프레임 내의 각각의 채널 신호는 일군의 샘플 값들을 포함하고, 상기 HOA 데이터 프레임들의 각각의 HOA 데이터 프레임의 각각의 채널 신호$(\boldsymbol{y}_1(k-2), \ldots, \boldsymbol{y}_I(k-2))$에 차분 이득 값이 할당되며, 상기 차분 이득 값은 현재 HOA 데이터 프레임$((k-2))$에서의 채널 신호의 제1 샘플 값들의, 이전 HOA 데이터 프레임$((k-3))$에서의 채널 신호의 제2 샘플 값들에 대한, 진폭들의 변화를 야기하고, 그 결과 얻어진 이득 적응된 채널 신호들이 인코더에서 인코딩되며,

상기 HOA 데이터 프레임 표현은 공간 영역에서 O개의 가상 스피커 신호들 $w_j(t)$로 렌더링되었고, 상기 가상 스피커들의 위치들은 단위 구면 상에 있고 그 단위 구면 상에 균일하게 분포되는 것을 목표로 하며, 상기 렌더링은 행렬 곱셈 $\boldsymbol{w}(t) = (\boldsymbol{\Psi})^{-1} \cdot \boldsymbol{c}(t)$에 의해 표현되고, $\boldsymbol{w}(t)$는 가상 스피커 신호들 모두를 포함하는 벡터이며, $\boldsymbol{\Psi}$는 가상 스피커 위치들 모드 행렬이고, $\boldsymbol{c}(t)$는 상기 HOA 데이터 프레임 표현의 대응하는 HOA 계수 시퀀스들의 벡터이며,

상기 HOA 데이터 프레임 표현 $(\boldsymbol{C}(k))$은 $\|\boldsymbol{w}(t)\|_\infty = \max_{1 \le j \le O} |w_j(t)| \le 1 \quad \forall t$ 이도록 정규화되었고,

상기 방법은

a) 상기 채널 신호들에서의 우세 음향 신호들$(\boldsymbol{x}(t))$을 표현하기 위해, HOA 계수 시퀀스들의 벡터 $\boldsymbol{c}(t)$를 믹싱 행렬(mixing matrix) $\boldsymbol{A}$와 곱하는 것 - 이 믹싱 행렬 $\boldsymbol{A}$의 유클리드 노름(norm)은 '1' 이하이고, 믹싱 행렬 $\boldsymbol{A}$는 정규화된 HOA 데이터 프레임 표현의 계수 시퀀스들의 선형 결합을 나타냄 -;

b) 상기 채널 신호들에서의 주변 성분 $\boldsymbol{c}_{\mathrm{AMB}}(t)$를 표현하기 위해, 상기 우세 음향 신호들을 상기 정규화된 HOA 데이터 프레임 표현으로부터 차감하고, 상기 주변 성분 $\boldsymbol{c}_{\mathrm{AMB}}(t)$의 계수 시퀀스들의 적어도 일부를 선택하며 - $\|\boldsymbol{c}_{\mathrm{AMB}}(t)\|_2{}^2 \le \|\boldsymbol{c}(t)\|_2{}^2$임 -, $\boldsymbol{w}_{\mathrm{MIN}}(t) = \boldsymbol{\Psi}_{\mathrm{MIN}}^{-1} \cdot \boldsymbol{c}_{\mathrm{AMB,MIN}}(t)$를 계산하는 것에 의해 얻어진 최소 주변 성분 $\boldsymbol{c}_{\mathrm{AMB,MIN}}(t)$를 변환하는 것 - $\|\boldsymbol{\Psi}_{\mathrm{MIN}}^{-1}\|_2 < 1$이고 $\boldsymbol{\Psi}_{\mathrm{MIN}}$은 상기 최소 주변 성분 $\boldsymbol{c}_{\mathrm{AMB,MIN}}(t)$에 대한 모드 행렬임 -;

c) 공간 변환이 적용되는 상기 주변 성분 $\boldsymbol{c}_{\mathrm{AMB}}(t)$의 계수 시퀀스들에 관련되는 상기 HOA 계수 시퀀스들 $\boldsymbol{c}(t)$의 일부를 선택하는 것 - 상기 선택된 계수 시퀀스들의 수를 나타내는 최소 차수 N_{MIN}은

$N_{\mathrm{MIN}} \leq 9$임 -

에 의해 채널 신호들을 형성하는 단계; 및

$\beta_{\mathrm{e}} = \lceil \log_2(\lceil \log_2(\sqrt{K_{\mathrm{MAX}}} \cdot O) \rceil + 1) \rceil$에 기초하여 비트들의 상기 정수 개수 β_{e}를 결정하는 단계 - $K_{\mathrm{MAX}} = \max_{1 \leq N \leq N_{\mathrm{MAX}}} K\left(N, \boldsymbol{\Omega}_1^{(N)}, \ldots, \boldsymbol{\Omega}_O^{(N)}\right)$이고, N은 차수이며, N_{MAX}는 최대 관심 차수이고, $\boldsymbol{\Omega}_1^{(N)}, \ldots, \boldsymbol{\Omega}_O^{(N)}$은 상기 가상 스피커들의 방향들이며, $O = (N+1)^2$은 HOA 계수 시퀀스들의 수이고, K는 상기 모드 행렬의 유클리드 노름의 제곱 $\|\boldsymbol{\Psi}\|_2^{\,2}$과 O 사이의 비임 -

를 포함하는, 방법.

청구항 2

제1항에 있어서, 상기 변환된 최소 주변 성분에 부가하여, 상기 주변 성분 $\boldsymbol{c}_{\mathrm{AMB}}(t)$의 비변환된 주변 계수 시퀀스들이 상기 채널 신호 $(\boldsymbol{y}_1(k-2), \ldots, \boldsymbol{y}_I(k-2))$에 포함되는, 방법.

청구항 3

제1항에 있어서, 상기 HOA 데이터 프레임들의 특정 HOA 데이터 프레임들의 상기 채널 신호들과 연관된 상기 비차분 이득 값들의 표현들(2^e)은 보조 정보로서 전송되고, 상기 표현들의 각각의 표현은 β_{e}개의 비트들로 나타내어지는, 방법.

청구항 4

제1항에 있어서, 비트들의 상기 정수 개수 β_{e}는 $\beta_{\mathrm{e}} = \lceil \log_2(\lceil \log_2(\sqrt{K_{\mathrm{MAX}}} \cdot O) \rceil + e_{\mathrm{MAX}} + 1) \rceil$로 설정되고, $e_{\mathrm{MAX}} > 0$는 이득 제어 이전의 채널 신호의 샘플 값들의 진폭들이 문턱값보다 낮다는 결정에 기초하여 비트들의 상기 개수 β_{e}를 증가시키는 역할을 하는, 방법.

청구항 5

제1항에 있어서, $\sqrt{K_{\mathrm{MAX}}} = 1.5$인, 방법.

청구항 6

제1항에 있어서, 상기 믹싱 행렬 $\boldsymbol{A}$는, 모노럴 우세 음향 신호들의 방향 분포를 표현하는 모든 벡터들로 형성되는 모드 행렬의 Moore-Penrose 의사 역행렬을 취하는 것에 의해, 원래의 HOA 표현과 상기 우세 음향 신호들의 HOA 표현 사이의 잔차의 유클리드 노름을 최소화하도록 결정되는, 방법.

청구항 7

제1항에 있어서, 상기 O개의 가상 스피커 신호들의 위치들이 β_{e}의 계산을 위해 가정된 위치들과 일치하지 않는다는 결정에 기초하여,

- 상기 일치하지 않는 가상 스피커 위치들에 기초한 모드 행렬 $\boldsymbol{\Psi}$를 계산하는 것;

- 상기 모드 행렬의 유클리드 노름 $\|\boldsymbol{\Psi}\|_2$ 를 계산하는 것;

- 상기 정규화에서 최대 허용 진폭을 대체하는 최대 허용 진폭 값 $\gamma = \min\left(1, \frac{\sqrt{O}\cdot\sqrt{K_{\mathrm{MAX,DES}}}}{\|\boldsymbol{\Psi}\|_2}\right)$ 을 계산하는 것을 포함하며,

$K_{MAX,DES} = \max_{1\le N\le N_{\mathrm{MAX,DES}}} K(N, \boldsymbol{\Omega}^{(N)}_{\mathrm{DES},1}, \ldots, \boldsymbol{\Omega}^{(N)}_{\mathrm{DES},O})$ 이고, N 은 차수이며, $O = (N+1)^2$ 은 HOA 계수 시퀀스들의 수이고, K 는 상기 모드 행렬의 유클리드 노름의 제곱과 O 사이의 비이며, $N_{\mathrm{MAX,DES}}$ 는 관심 차수이고 $\boldsymbol{\Omega}^{(N)}_{\mathrm{DES},1}, \ldots, \boldsymbol{\Omega}^{(N)}_{\mathrm{DES},1}$ 는 각각의 차수에 대해, 상기 비차분 이득 값들의 밑이 '2'인 지수들 (e) 을 코딩하기 위해 β_e 가 $\beta_e = \lceil \log_2(\lceil \log_2(\sqrt{K_{MAX,DES}}\cdot O)\rceil + 1)\rceil$ 에 의해 선택되었도록, 상기 HOA 데이터 프레임 표현 $(\boldsymbol{C}(k))$ 의 상기 압축의 구현을 위해 가정되었던 상기 가상 스피커들의 방향들인, 방법.

청구항 8

HOA 데이터 프레임 표현 $(\boldsymbol{C}(k))$ 의 압축을 위해 HOA 데이터 프레임들의 채널 신호들에 대해 진폭 변화들에 대응하는 비차분 이득 값들의 표현들을 2의 지수 (2^e) 로서 나타내기 위한 비트들의 최저 정수 개수 $\beta_{\mathbf{e}}$ 를 결정하는 장치로서,

각각의 프레임 내의 각각의 채널 신호는 일군의 샘플 값들을 포함하고, 상기 HOA 데이터 프레임들의 각각의 HOA 데이터 프레임의 각각의 채널 신호 $(\boldsymbol{y}_1(k-2), \ldots, \boldsymbol{y}_I(k-2))$ 에 차분 이득 값이 할당되며, 상기 차분 이득 값은 현재 HOA 데이터 프레임 $((k-2))$ 에서의 채널 신호의 제1 샘플 값들의, 이전 HOA 데이터 프레임 $((k-3))$ 에서의 채널 신호의 제2 샘플 값들에 대한, 진폭들의 변화를 야기하고, 그 결과 얻어진 이득 적응된 채널 신호들이 인코더에서 인코딩되며,

상기 HOA 데이터 프레임 표현 $(\boldsymbol{C}(k))$ 은 공간 영역에서 O 개의 가상 스피커 신호들 $w_j(t)$ 로 렌더링되었고, 상기 가상 스피커들의 위치들은 단위 구면 상에 있고 그 단위 구면 상에 균일하게 분포되는 것을 목표로 하며, 상기 렌더링은 행렬 곱셈 $\boldsymbol{w}(t) = (\boldsymbol{\Psi})^{-1}\cdot\boldsymbol{c}(t)$ 에 의해 표현되고, $\boldsymbol{w}(t)$ 는 가상 스피커 신호들 모두를 포함하는 벡터이며, $\boldsymbol{\Psi}$ 는 가상 스피커 위치들 모드 행렬이고, $\boldsymbol{c}(t)$ 는 상기 HOA 데이터 프레임 표현의 대응하는 HOA 계수 시퀀스들의 벡터이며,

상기 HOA 데이터 프레임 표현 $(\boldsymbol{C}(k))$ 은 $\|\boldsymbol{w}(t)\|_\infty = \max_{1\le j\le O}|w_j(t)| \le 1 \quad \forall t$ 이도록 정규화되었고,

상기 장치는

a) 상기 채널 신호들에서의 우세 음향 신호들 $(\boldsymbol{x}(t))$ 을 표현하기 위해, 상기 HOA 계수 시퀀스들의 벡터 $\boldsymbol{c}(t)$ 를 믹싱 행렬 $\boldsymbol{A}$ 와 곱하는 것 - 이 믹싱 행렬 $\boldsymbol{A}$ 의 유클리드 노름은 '1' 이하이고, 믹싱 행렬 $\boldsymbol{A}$ 는 정규화된 HOA 데이터 프레임 표현의 계수 시퀀스들의 선형 결합을 나타냄 -;

b) 상기 채널 신호들에서의 주변 성분 $\boldsymbol{c}_{\mathrm{AMB}}(t)$ 를 표현하기 위해, 상기 우세 음향 신호들을 상기 정규화된

HOA 데이터 프레임 표현으로부터 차감하고, 상기 주변 성분 $\boldsymbol{c}_{\mathrm{AMB}}(t)$의 계수 시퀀스들의 적어도 일부를 선택하는 것 - $\|\boldsymbol{c}_{\mathrm{AMB}}(t)\|_2{}^2 \leq \|\boldsymbol{c}(t)\|_2{}^2$임 -, $\boldsymbol{w}_{\mathrm{MIN}}(t) = \boldsymbol{\Psi}_{\mathrm{MIN}}^{-1} \cdot \boldsymbol{c}_{\mathrm{AMB,MIN}}(t)$를 계산하는 것에 의해 얻어진 최소 주변 성분 $\boldsymbol{c}_{\mathrm{AMB,MIN}}(t)$를 변환하고 - $\|\boldsymbol{\Psi}_{\mathrm{MIN}}^{-1}\|_2 < 1$이고 $\boldsymbol{\Psi}_{\mathrm{MIN}}$은 상기 최소 주변 성분 $\boldsymbol{c}_{\mathrm{AMB,MIN}}(t)$에 대한 모드 행렬임 -;

c) 공간 변환이 적용되는 상기 주변 성분 $\boldsymbol{c}_{\mathrm{AMB}}(t)$의 계수 시퀀스들에 관련되는 상기 HOA 계수 시퀀스들 $\boldsymbol{c}(t)$의 일부를 선택하는 것 - 상기 선택된 계수 시퀀스들의 수를 나타내는 최소 차수 N_{MIN}은 $N_{\mathrm{MIN}} \leq 9$임 -

에 의해 상기 채널 신호들 $(\boldsymbol{y}_1(k-2), \ldots, \boldsymbol{y}_I(k-2))$을 형성하도록 구성된 프로세서; 및

$\beta_{\mathrm{e}} = \lceil \log_2(\lceil \log_2(\sqrt{K_{\mathrm{MAX}}} \cdot O) \rceil + 1) \rceil$에 기초하여 비트들의 상기 정수 개수 β_{e}를 결정하도록 구성된 프로세서 - $K_{\mathrm{MAX}} = \max_{1 \leq N \leq N_{\mathrm{MAX}}} K\left(N, \boldsymbol{\Omega}_1^{(N)}, \ldots, \boldsymbol{\Omega}_O^{(N)}\right)$이고, N은 차수이며, N_{MAX}는 최대 관심 차수이고, $\boldsymbol{\Omega}_1^{(N)}, \ldots, \boldsymbol{\Omega}_O^{(N)}$은 상기 가상 스피커들의 방향들이며, $O = (N+1)^2$은 HOA 계수 시퀀스들의 수이고, K는 상기 모드 행렬의 유클리드 노름의 제곱 $\|\boldsymbol{\Psi}\|_2{}^2$과 O 사이의 비임 -

를 포함하는, 장치.

청구항 9

제8항에 있어서, 상기 변환된 최소 주변 성분에 부가하여, 상기 주변 성분 $\boldsymbol{c}_{\mathrm{AMB}}(t)$의 비변환된 주변 계수 시퀀스들이 상기 채널 신호 $(\boldsymbol{y}_1(k-2), \ldots, \boldsymbol{y}_I(k-2))$에 포함되는, 장치.

청구항 10

제8항에 있어서, 상기 HOA 데이터 프레임들의 특정 HOA 데이터 프레임들의 상기 채널 신호들과 연관된 상기 비차분 이득 값들의 표현들(2^e)은 보조 정보로서 전송되고, 상기 표현들의 각각의 표현은 β_{e}개의 비트들로 나타내어지는, 장치.

청구항 11

제8항에 있어서, 비트들의 상기 정수 개수 β_{e}는 $\beta_{\mathrm{e}} = \lceil \log_2(\lceil \log_2(\sqrt{K_{\mathrm{MAX}}} \cdot O) \rceil + e_{\mathrm{MAX}} + 1) \rceil$로 설정되고, $e_{\mathrm{MAX}} > 0$는 이득 제어 이전의 채널 신호의 샘플 값들의 진폭들이 문턱값보다 낮다는 결정에 기초하여 비트들의 상기 개수 β_{e}를 증가시키는 역할을 하는, 장치.

청구항 12

제8항에 있어서, $\sqrt{K_{\mathrm{MAX}}} = 1.5$인, 장치.

청구항 13

제8항에 있어서, 상기 믹싱 행렬 $\boldsymbol{A}$는, 모노럴 우세 음향 신호들의 방향 분포를 표현하는 모든 벡터들로 형성되는 모드 행렬의 Moore-Penrose 의사 역행렬을 취하는 것에 의해, 원래의 HOA 표현과 상기 우세 음향 신호들의 HOA 표현 사이의 잔차의 유클리드 노름을 최소화하도록 결정되는, 장치.

청구항 14

제8항에 있어서, 상기 O개의 가상 스피커 신호들의 위치들이 β_{e}의 계산을 위해 가정된 위치들과 일치하지 않는다는 결정에 기초하여, 상기 프로세서가:

- 상기 일치하지 않는 가상 스피커 위치들에 기초한 모드 행렬 $\boldsymbol{\Psi}$를 계산하고;

- 상기 모드 행렬의 유클리드 노름 $\|\boldsymbol{\Psi}\|_2$를 계산하고;

- 상기 정규화에서 최대 허용 진폭을 대체하는 최대 허용 진폭 값 $\gamma = \min\left(1, \frac{\sqrt{O}\cdot\sqrt{K_{\mathrm{MAX,DES}}}}{\|\boldsymbol{\Psi}\|_2}\right)$을 계산하도록 더 구성되며,

$K_{MAX,DES} = \max_{1\le N\le N_{\mathrm{MAX,DES}}} K(N, \boldsymbol{\Omega}^{(N)}_{\mathrm{DES},1}, \ldots, \boldsymbol{\Omega}^{(N)}_{\mathrm{DES},O})$이고, N은 차수이며, $O = (N+1)^2$은 HOA 계수 시퀀스들의 수이고, K는 상기 모드 행렬의 유클리 노름의 제곱과 O 사이의 비이며, $N_{\mathrm{MAX,DES}}$는 관심 차수이고 $\boldsymbol{\Omega}^{(N)}_{\mathrm{DES},1}, \ldots, \boldsymbol{\Omega}^{(N)}_{\mathrm{DES},1}$는 각각의 차수에 대해, 상기 비차분 이득 값들의 밑이 '2'인 지수들 (e)을 코딩하기 위해 β_e가 $\beta_e = \lceil \log_2(\lceil \log_2(\sqrt{K_{MAX,DES}} \cdot O) \rceil + 1) \rceil$에 의해 선택되었도록, 상기 HOA 데이터 프레임 표현$(\boldsymbol{C}(k))$의 상기 압축의 구현을 위해 가정되었던 상기 가상 스피커들의 방향들인, 장치.

청구항 15

음향 또는 음장의 압축된 HOA(Higher Order Ambisonics) 음향 표현을 디코딩하는 방법으로서,

상기 압축된 HOA 표현을 포함하는 비트 스트림을 수신하고, 상기 압축된 HOA 표현을 디코딩하여 인지 디코딩된 신호들 $\hat{\boldsymbol{z}}_i(k),\ \ i = 1, \ldots, I$, 연관된 이득 보정 지수 $e_i(k)$ 및 이득 보정 예외 플래그 $\beta_i(k)$를 결정하는 단계; 및

상기 인지 디코딩된 신호들 $\hat{\boldsymbol{z}}_i(k),\ \ i = 1, \ldots, I$, 상기 연관된 이득 보정 지수 $e_i(k)$ 및 상기 이득 보정 예외 플래그 $\beta_i(k)$에 대해 역 이득 제어 처리를 수행하는 것에 의해 이득 보정된 신호 프레임들 $\hat{\boldsymbol{y}}_i(k),\ \ i = 1, \ldots, I$를 제공하는 단계;

우세 음향 신호들의 프레임 $\hat{\boldsymbol{X}}_{\mathrm{PS}}(k)$ 및 주변 HOA 성분의 중간 표현의 프레임 $\boldsymbol{C}_{\mathrm{I,AMB}}(k)$를 재구성하기 위해, 채널 재할당 동안 상기 이득 보정된 신호 프레임들 $\hat{\boldsymbol{y}}_i(k),\ \ i = 1, \ldots, I$를 재분포시키는 단계 - 이전 프레임에서 전송 채널의 신호에 적용되는 비트들의 최저 정수 개수 β_{e}는 $\beta_{\mathrm{e}} = \left\lceil \log_2\left(\left\lceil \log_2\left(\sqrt{K_{\mathrm{MAX}}} \cdot O\right)\right\rceil + 1\right)\right\rceil$에 기초하고,

$K_{\mathrm{MAX}} = \max_{1 \le N \le N_{\mathrm{MAX}}} K\left(N, \boldsymbol{\Omega}_1^{(N)}, \ldots, \boldsymbol{\Omega}_O^{(N)}\right)$ 이고, N 은 차수이며, N_{MAX} 는 최대 관심 차수이고, $\boldsymbol{\Omega}_1^{(N)}, \ldots, \boldsymbol{\Omega}_O^{(N)}$ 은 가상 스피커들의 방향들이며, $O = (N+1)^2$ 은 HOA 계수 시퀀스들의 수이고, K 는 모드 행렬의 유클리드 노름의 제곱 $\|\boldsymbol{\Psi}\|_2^2$ 과 O 사이의 비임 -

를 포함하는, 방법.

청구항 16

제15항에 있어서, $K_{\mathrm{MAX}} = 1.5$ 인, 방법.

청구항 17

음향 또는 음장의 압축된 HOA(Higher Order Ambisonics) 음향 표현을 디코딩하는 장치로서,

상기 압축된 HOA 표현을 포함하는 비트 스트림을 수신하고, 상기 압축된 HOA 표현을 디코딩하여 인지 디코딩된 신호들 $\hat{\boldsymbol{z}}_i(k), \quad i = 1, \ldots, I$, 연관된 이득 보정 지수 $e_i(k)$ 및 이득 보정 예외 플래그 $\beta_i(k)$ 를 결정하도록 구성된 프로세서를 포함하고,

상기 프로세서는 상기 인지 디코딩된 신호들 $\hat{\boldsymbol{z}}_i(k), \quad i = 1, \ldots, I$, 상기 연관된 이득 보정 지수 $e_i(k)$ 및 상기 이득 보정 예외 플래그 $\beta_i(k)$ 에 대해 역 이득 제어 처리를 수행하는 것에 의해 이득 보정된 신호 프레임들 $\hat{\boldsymbol{y}}_i(k), \quad i = 1, \ldots, I$ 를 제공하도록 더 구성되고,

상기 프로세서는 우세 음향 신호들의 프레임 $\hat{\boldsymbol{X}}_{\mathrm{PS}}(k)$ 및 주변 HOA 성분의 중간 표현의 프레임 $\boldsymbol{C}_{\mathrm{I,AMB}}(k)$ 를 재구성하기 위해, 채널 재할당 동안 상기 이득 보정된 신호 프레임들 $\hat{\boldsymbol{y}}_i(k), \quad i = 1, \ldots, I$ 를 재분포시키도록 더 구성되고 - 이전 프레임에서 전송 채널의 신호에 적용되는 비트들의 최저 정수 개수 β_{e} 는 $\beta_{\mathrm{e}} = \left\lceil \log_2\left(\left\lceil \log_2\left(\sqrt{K_{\mathrm{MAX}}} \cdot O\right)\right\rceil + 1\right)\right\rceil$ 에 기초하고,

$K_{\mathrm{MAX}} = \max_{1 \le N \le N_{\mathrm{MAX}}} K\left(N, \boldsymbol{\Omega}_1^{(N)}, \ldots, \boldsymbol{\Omega}_O^{(N)}\right)$ 이고, N 은 차수이며, N_{MAX} 는 최대 관심 차수이고, $\boldsymbol{\Omega}_1^{(N)}, \ldots, \boldsymbol{\Omega}_O^{(N)}$ 은 가상 스피커들의 방향들이며, $O = (N+1)^2$ 은 HOA 계수 시퀀스들의 수이고, K 는 모드 행렬의 유클리드 노름의 제곱 $\|\boldsymbol{\Psi}\|_2^2$ 과 O 사이의 비임 -

을 포함하는, 장치.

청구항 18

제17항에 있어서, $K_{\mathrm{MAX}} = 1.5$ 인, 장치.

청구항 19

삭제

청구항 20

삭제

청구항 21

삭제

발명의 설명

기 술 분 야

[0001] 본 발명은 HOA 데이터 프레임 표현의 압축을 위해 상기 HOA 데이터 프레임들 중 특정 HOA 데이터 프레임들의 채널 신호들과 연관된 비차분 이득 값들을 표현하는 데 필요하게 되는 비트들의 최저 정수 개수를 결정하는 장치에 관한 것이다.

배 경 기 술

[0002] HOA라고 표기되는 고차 앰비소닉스(Higher Order Ambisonics)는 3차원 음향을 표현할 하나의 가능성을 제공한다. 다른 기법들은 WFS(wave field synthesis) 또는 22.2와 같은 채널 기반 접근법들이다. 채널 기반 방법들과 달리, HOA 표현은 특정 스피커 설정(loudspeaker set-up)에 독립적이라는 장점을 제공한다. 그렇지만, 이 유연성은 특정의 스피커 설정에서 HOA 표현의 재생을 위해 필요하게 되는 디코딩 프로세스의 대가이다. 요구된 스피커들의 수가 보통 아주 많은 WFS 접근법과 비교하여, HOA는 또한 단지 몇 개의 스피커들만으로 이루어진 설정에 맞춰 렌더링될 수 있다. HOA의 추가적인 장점은 헤드폰에 대한 바이노럴 렌더링(binaural rendering)을 위해 어떤 수정도 없이 동일한 표현이 또한 이용될 수 있다는 것이다.

[0003] HOA는 절단된 SH(Spherical Harmonics) 전개에 의한 복소 조화 평면파 진폭(complex harmonic plane wave amplitude)들의 공간 밀도의 표현에 기초한다. 각각의 전개 계수(expansion coefficient)는, 시간 영역 함수에 의해 동등하게 표현될 수 있는, 각주파수(angular frequency)의 함수이다. 그러므로, 일반성을 잃지 않고, 완전한 HOA 음장 표현이 실제로는 O 개의 시간 영역 함수들로 이루어진 것으로 가정될 수 있으며, 여기서 O는 전개 계수들의 수를 나타낸다. 이 시간 영역 함수들은 동등하게 이하에서 HOA 계수 시퀀스들이라고 또는 HOA 채널들이라고 지칭될 것이다.

[0004] HOA 표현의 공간 분해능은 전개의 최대 차수 N이 증가함에 따라 향상된다. 안타깝게도, 전개 계수들의 수 O는 차수 N에 따라 이차식으로 증가하고, 상세하게는, $O = (N+1)^2$이다. 예를 들어, 차수 $N = 4$를 사용하는 전형적인 HOA 표현은 $O = 25$개의 HOA (전개) 계수들을 필요로 한다. HOA 표현의 전송을 위한 총 비트 레이트는, 원하는 단일 채널 샘플링 레이트 f_s 및 샘플당 비트 수 N_b가 주어지면, $O \cdot f_s \cdot N_b$에 의해 결정된다. 샘플당 $N_b = 16$ 비트를 이용하는 $f_s = 48\text{kHz}$의 샘플링 레이트로 차수 $N = 4$의 HOA 표현을 전송하면 그 결과 $19.2\ \text{MBits/s}$의 비트 레이트가 되고, 이는 많은 실제 응용분야들, 예컨대, 스트리밍에 대해 아주 높다. 이와 같이, HOA 표현들의 압축이 아주 바람직하다.

[0005] 이전에, HOA 음장 표현들의 압축이 EP 2665208 A1, EP 2743922 A1, EP 2800401 A1 - ISO/IEC JTC1/SC29/WG11, N14264, WD1-HOA Text of MPEG-H 3D Audio, January 2014를 참조 - 에서 제안되었다. 이 접근법들의 공통점은 음장 분석을 수행하고 주어진 HOA 표현을 방향 성분(directional component)과 잔차 주변 성분(residual ambient component)으로 분해한다는 것이다. 최종적인 압축된 표현은, 한편으로는, 방향 및 벡터 기반 신호들은 물론 주변 HOA 성분의 관련성있는 계수 시퀀스들의 인지 코딩(perceptual coding)으로부터 얻어진, 다수의 양자화된 신호들로 이루어져 있는 것으로 가정된다. 다른 한편으로는, 최종적인 압축된 표현은 양자화된 신호들에 관련된 부가의 보조 정보를 포함하고, 이 보조 정보는 HOA 표현을 그의 압축된 버전으로부터 재구성하는 데 필요하게 된다.

[0006] 인지 인코더(perceptual encoder)로 전달되기 전에, 이 중간의 시간 영역 신호들은 값 범위 $[-1,1[$ 내에서 최대 진폭을 가질 필요가 있고, 이는 현재 이용가능한 인지 인코더들의 구현으로부터 생기는 요구사항이다. HOA 표현들을 압축할 때 이 요구사항을 충족시키기 위해, 입력 신호들을 매끄럽게 감쇠시키거나 증폭시키는 이득 제어

처리 유닛(EP 2824661 A1 및 앞서 언급된 ISO/IEC JTC1/SC29/WG11 N14264 문서를 참조)이 인지 인코더들보다 앞서 사용된다. 그에 따른 신호 수정이 가역적이고 프레임 단위로 적용되는 것으로 가정되고, 여기서 상세하게는 연속적인 프레임들 사이에서 신호 진폭들의 변화가 '2'의 멱수인 것으로 가정된다. HOA 압축해제기에서 이 신호 수정을 반대로 하는 것을 용이하게 하기 위해, 대응하는 정규화 보조 정보가 총 보조 정보에 포함된다. 이 정규화 보조 정보는 밑이 '2'인 지수들로 이루어져 있을 수 있고, 이 지수들은 2개의 연속적인 프레임들 사이의 상대 진폭 변화를 나타낸다. 이 지수들은 앞서 언급된 ISO/IEC JTC1/SC29/WG11 N14264 문서에 따라 런 길이 코드(run length code)를 사용하여 코딩되는데, 그 이유는 연속적인 프레임들 사이의 사소한 진폭 변화들이 보다 큰 진폭 변화들보다 가능성이 더 많기 때문이다.

발명의 내용

[0007] 예컨대, 단일의 파일이 어떤 시간적 점프도 없이 처음부터 끝까지 압축해제되는 경우에, HOA 압축해제에서 원래의 신호 진폭들을 재구성하기 위해 차분 코딩된 진폭 변화들을 사용하는 것이 실현가능하다. 그렇지만, 랜덤 액세스를 용이하게 하기 위해, 이전 프레임들로부터의 정보에 관계없이, 원하는 위치로부터(또는 적어도 그 근방에서) 압축해제를 시작하는 것을 가능하게 하기 위해, 독립적인 액세스 단위들이 (전형적으로 비트 스트림인) 코딩된 표현에 존재해야만 한다. 이러한 독립적인 액세스 단위는 첫 번째 프레임부터 현재 프레임까지 이득 제어 처리 유닛에 의해 야기된 총 절대 진폭 변화(즉, 비차분 이득 값)를 포함해야만 한다. 2개의 연속적인 프레임들 사이의 진폭 변화들이 '2'의 멱수인 것으로 가정하면, 총 절대 진폭 변화를 밑이 '2'인 지수에 의해 나타내는 것으로도 충분하다. 이 지수의 효율적인 코딩을 위해, 이득 제어 처리 유닛의 적용 이전에 신호들의 잠재적인 최대 이득들을 아는 것이 필수적이다. 그렇지만, 이것을 아는 것은 압축될 HOA 표현들의 값 범위에 관한 제약조건들의 명시에 많이 의존하고 있다. 안타깝게도, MPEG-H 3D 오디오 문서 ISO/IEC JTC1/SC29/WG11 N14264는, 값 범위들에 대한 어떤 제약조건들도 설정함이 없이, 입력 HOA 표현에 대한 포맷의 설명만을 제공한다.

[0008] 본 발명에 의해 해결될 문제점은 비차분 이득 값들을 표현하는 데 필요하게 되는 비트들의 최저 정수 개수를 제공하는 것이다. 이 문제점은 청구항 1에 개시된 장치에 의해 해결된다.

[0009] 본 발명의 유리한 부가의 실시예들은 각자의 종속 청구항들에 개시되어 있다.

[0010] 본 발명은 HOA 압축기 내에서의 이득 제어 처리 유닛의 적용 이전에 입력 HOA 표현의 값 범위와 신호들의 잠재적인 최대 이득들 사이의 상호관계를 구축한다. 그 상호관계에 기초하여, 첫 번째 프레임부터 현재 프레임까지 이득 제어 처리 유닛에 의해 야기된 수정된 신호들의 총 절대 진폭 변화들(즉, 비차분 이득 값)을 액세스 단위 내에 나타내기 위한 밑이 '2'인 지수의 효율적인 코딩을 위해 - 입력 HOA 표현의 값 범위에 대한 주어진 명시에 대해 - 필요하게 되는 비트들의 양이 결정된다.

[0011] 게다가, 지수의 코딩을 위해 필요하게 되는 비트들의 양의 계산을 위한 규칙이 고정되어 있으면, 본 발명은 주어진 HOA 표현이 요구된 값 범위 제약조건들을 충족시킴으로써 올바르게 압축될 수 있는지를 검증하기 위한 처리를 사용한다.

[0012] 원칙적으로, 본 발명의 장치는 HOA 데이터 프레임 표현의 압축을 위해 상기 HOA 데이터 프레임들 중 특정 HOA 데이터 프레임들의 채널 신호들에 대한 비차분 이득 값들을 표현하는 데 필요하게 되는 비트들의 최저 정수 개수 β_e를 결정하는 데 적합하고, 여기서 각각의 프레임 내의 각각의 채널 신호는 일군의 샘플 값들을 포함하고, 여기서 상기 HOA 데이터 프레임들의 각각의 HOA 데이터 프레임의 각각의 채널 신호에 차분 이득 값이 할당되며, 이러한 차분 이득 값은 현재 HOA 데이터 프레임에서의 채널 신호의 샘플 값들의, 이전 HOA 데이터 프레임에서의 그 채널 신호의 샘플 값들에 대한, 진폭들의 변화를 야기하고, 여기서 이러한 이득 적응된 채널 신호들이 인코더에서 인코딩되며,

[0013] 그리고 여기서 상기 HOA 데이터 프레임 표현은 공간 영역에서 O개의 가상 스피커 신호들 $w_j(t)$로 렌더링되었고, 여기서 가상 스피커들의 위치들은 단위 구면 상에 있고 그 단위 구면 상에 균일하게 분포되는 것을 목표로 하며, 상기 렌더링은 행렬 곱셈 $\boldsymbol{w}(t) = (\boldsymbol{\Psi})^{-1} \cdot \boldsymbol{c}(t)$에 의해 표현되고, 여기서 $\boldsymbol{w}(t)$는 가상 스피커 신호들 모두를 포함하는 벡터이며, $\boldsymbol{\Psi}$는 가상 스피커 위치들 모드 행렬이고, $\boldsymbol{c}(t)$는 상기 HOA 데이터 프레임 표

현의 대응하는 HOA 계수 시퀀스들의 벡터이며,

[0014] 그리고 여기서 상기 HOA 데이터 프레임 표현은 $\|\boldsymbol{w}(t)\|_\infty = \max_{1\le j\le O}|w_j(t)| \le 1 \quad \forall t$ 이도록 정규화되었고,

[0015] 상기 장치는

[0016] - 상기 정규화된 HOA 데이터 프레임 표현으로부터 동작들 a), b), c) 중 하나 이상에 의해 상기 채널 신호들을 형성하는 수단:

[0017] a) 상기 채널 신호들에서의 우세 음향 신호들을 표현하기 위해, 상기 HOA 계수 시퀀스들의 벡터 $\boldsymbol{c}(t)$를 믹싱 행렬 $\boldsymbol{A}$와 곱하는 동작 - 이 믹싱 행렬 $\boldsymbol{A}$의 유클리드 노름은 '1' 이하이고, 믹싱 행렬 $\boldsymbol{A}$는 상기 정규화된 HOA 데이터 프레임 표현의 계수 시퀀스들의 선형 결합을 나타냄 -;

[0018] b) 상기 채널 신호들에서의 주변 성분 $\boldsymbol{c}_{\mathrm{AMB}}(t)$를 표현하기 위해, 상기 우세 음향 신호들을 상기 정규화된 HOA 데이터 프레임 표현으로부터 차감하고, 상기 주변 성분 $\boldsymbol{c}_{\mathrm{AMB}}(t)$의 계수 시퀀스들의 적어도 일부를 선택하며 - $\|\boldsymbol{c}_{\mathrm{AMB}}(t)\|_2^{\,2} \le \|\boldsymbol{c}(t)\|_2^{\,2}$임 -, $\boldsymbol{c}_{\mathrm{AMB,MIN}}(t)$를 계산하는 것에 의해 얻어진 최소 주변 성분 $\boldsymbol{w}_{\mathrm{MIN}}(t) = \boldsymbol{\Psi}_{\mathrm{MIN}}^{-1} \cdot \boldsymbol{c}_{\mathrm{AMB,MIN}}(t)$를 변환하는 동작 - $\|\boldsymbol{\Psi}_{\mathrm{MIN}}^{-1}\|_2 < 1$이고 $\boldsymbol{\Psi}_{\mathrm{MIN}}$은 상기 최소 주변 성분 $\boldsymbol{c}_{\mathrm{AMB,MIN}}(t)$에 대한 모드 행렬임 -;

[0019] c) 상기 HOA 계수 시퀀스들 $\boldsymbol{c}(t)$의 일부를 선택하는 동작 - 상기 선택된 계수 시퀀스들은 공간 변환이 적용되는 주변 HOA 성분의 계수 시퀀스들에 관련되고, 상기 선택된 계수 시퀀스들의 수를 나타내는 최소 차수 N_{MIN}은 $N_{\mathrm{MIN}} \le 9$임 -;

[0020] - 상기 채널 신호들에 대한 상기 비차분 이득 값들을 표현하기 위해 필요하게 되는 비트들의 상기 최저 정수 개수 β_e를 $\beta_e = \lceil \log_2(\lceil \log_2(\sqrt{K_{\mathrm{MAX}}} \cdot O)\rceil + 1)\rceil$로 설정하는 수단 -

[0021] $K_{\mathrm{MAX}} = \max_{1\le N\le N_{\mathrm{MAX}}} K\left(N, \boldsymbol{\Omega}_1^{(N)}, \ldots, \boldsymbol{\Omega}_O^{(N)}\right)$이고, N은 차수이며, N_{MAX}는 최대 관심 차수이고, $\boldsymbol{\Omega}_1^{(N)}, \ldots, \boldsymbol{\Omega}_O^{(N)}$은 상기 가상 스피커들의 방향들이며, $O = (N+1)^2$은 HOA 계수 시퀀스들의 수이고, K는 상기 모드 행렬의 유클리드 노름의 제곱 $\|\boldsymbol{\Psi}\|_2^{\,2}$과 O 사이의 비임 - 을 포함한다.

도면의 간단한 설명

[0022] 본 발명의 예시적인 실시예들이 첨부 도면들을 참조하여 기술된다.

도 1은 HOA 압축기를 나타낸 도면.

도 2는 HOA 압축해제기를 나타낸 도면.

도 3은 HOA 차수들 $N = 1, \ldots, 29$에 대한 가상 방향들 $\boldsymbol{\Omega}_j^{(N)},\ 1 \le j \le O,$의 스케일링 값들 K를 나타낸 도면.

도 4는 HOA 차수들 $N_{\mathrm{MIN}} = 1, \ldots, 9$에 대한 가상 방향들 $\boldsymbol{\Omega}_{\mathrm{MIN},d}$, $d = 1, \ldots, O_{\mathrm{MIN}}$의 역 모드 행렬들 $\boldsymbol{\Psi}^{-1}$의 유클리드 노름들을 나타낸 도면.

도 5는 위치들 $\boldsymbol{\Omega}_j^{(N)},\ 1 \le j \le O$ - 단, $O = (N+1)^2$임 - 에 있는 가상 스피커들의 신호들의 최대 허용 크기

γ_{dB}의 결정을 나타낸 도면.

도 6은 구면 좌표계를 나타낸 도면.

발명을 실시하기 위한 구체적인 내용

[0023] 비록 명시적으로 기술되어 있지 않지만, 이하의 실시예들은 임의의 컴비네이션 또는 서브컴비네이션으로 이용될 수 있다.

[0024] 이하에서, 앞서 언급된 문제점이 발생하는 보다 상세한 정황을 제공하기 위해 HOA 압축 및 압축해제의 원리가 제시된다. 이 제시에 대한 기초는 MPEG-H 3D 오디오 문서 ISO/IEC JTC1/SC29/WG11 N14264 - EP 2665208 A1, EP 2800401 A1 및 EP 2743922 A1도 참조 - 에 기술된 처리이다. N14264에서, '방향 성분'이 '우세 음향 성분'으로 확장된다. 방향 성분으로서, 우세 음향 성분은, 방향 신호들 - 청취자에 충돌하는 것으로 가정되는 대응하는 방향을 갖는 모노럴 신호(monaural signal)들을 의미함 - 로부터 원래의 HOA 표현의 일부분들을 예측하는 어떤 예측 파라미터들과 함께, 방향 신호들에 의해 부분적으로 표현되는 것으로 가정된다. 그에 부가하여, 우세 음향 성분은 '벡터 기반 신호들' - 벡터 기반 신호들의 방향 분포를 정의하는 대응하는 벡터를 갖는 모노럴 신호들을 의미함 - 에 의해 표현되도록 되어 있다.

[0025] *HOA 압축*

[0026] EP 2800401 A1에 기술된 HOA 압축기의 전체적인 아키텍처가 도 1에 예시되어 있다. 이는 도 1의 A에 도시된 공간 HOA 인코딩 부분과 도 1의 B에 도시된 인지 및 소스 인코딩 부분을 갖는다. 공간 HOA 인코더는 I개의 신호들로 이루어져 있는 제1 압축된 HOA 표현을, 그의 HOA 표현을 어떻게 생성해야 하는지를 기술하는 보조 정보와 함께, 제공한다. 인지 및 보조 정보 소스 코더들에서, I개의 신호들이 인지 인코딩되고 보조 정보는 소스 인코딩을 거치며, 그 후에 2개의 코딩된 표현들을 멀티플렉싱한다.

[0027] *공간 HOA 인코딩*

[0028] 제1 단계에서, 원래의 HOA 표현의 현재 제 k 프레임 $\boldsymbol{C}(k)$가, 튜플 세트들 $\mathcal{M}_{\mathrm{DIR}}(k)$ 및 $\mathcal{M}_{\mathrm{VEC}}(k)$를 제공하는 것으로 가정되는, 방향 및 벡터 추정 처리 단계 또는 스테이지(11)에 입력된다. 튜플 세트 $\mathcal{M}_{\mathrm{DIR}}(k)$는 제1 요소가 방향 신호의 인덱스를 나타내고 제2 요소가 각자의 양자화된 방향을 나타내는 튜플들로 이루어져 있다. 튜플 세트 $\mathcal{M}_{\mathrm{VEC}}(k)$는 제1 요소가 벡터 기반 신호의 인덱스를 나타내고 제2 요소가 신호들의 방향 분포 - 즉, 벡터 기반 신호의 HOA 표현이 어떻게 계산되는지 - 를 정의하는 벡터를 나타내는 튜플들로 이루어져 있다.

[0029] 튜플 세트들 $\mathcal{M}_{\mathrm{DIR}}(k)$ 및 $\mathcal{M}_{\mathrm{VEC}}(k)$ 둘 다를 사용하여, 초기 HOA 프레임 $\boldsymbol{C}(k)$가 HOA 분해 단계 또는 스테이지(12)에서 모든 우세 음향(즉, 방향 및 벡터 기반) 신호들의 프레임 $\boldsymbol{X}_{\mathrm{PS}}(k-1)$과 주변 HOA 성분의 프레임 $\boldsymbol{C}_{\mathrm{AMB}}(k-1)$으로 분해된다. 블록킹 아티팩츠(blocking artefact)를 피하기 위해 중첩-가산(overlap-add) 처리로 인한 하나의 프레임의 지연에 유의한다. 게다가, HOA 분해 단계/스테이지(12)는, 우세 음향 HOA 성분을 강화하기 위해, 방향 신호들로부터 원래의 HOA 표현의 일부분들을 어떻게 예측해야 하는지를 나타내는 어떤 예측 파라미터들 $\zeta(k-1)$을 출력하는 것으로 가정된다. 그에 부가하여, HOA 분해 처리 단계 또는 스테이지(12)에서 결정된, 우세 음향 신호들의 I개의 이용가능 채널들에의 할당에 관한 정보를 포함하는 대상 할당 벡터 $\boldsymbol{v}_{\mathrm{A,T}}(k-1)$이 제공되는 것으로 가정된다. 영향을 받는 채널들은 점유되는 - 각자의 시간 프레임에서 주변 HOA 성분의 임의의 계수 시퀀스들을 전송하는 데 이용가능하지 않다는 것을 의미함 - 것으로 가정될 수 있다.

[0030] 주변 성분 수정 처리 단계 또는 스테이지(13)에서, 주변 HOA 성분의 프레임 $\mathbf{C}_{\mathrm{AMB}}(k-1)$이 대상 할당 벡터 $\boldsymbol{v}_{\mathrm{A,T}}(k-1)$에 의해 제공되는 정보에 따라 수정된다. 상세하게는, (양태들 중에서도 특히) 어느 채널들이 이용

가능하고 우세 음향 신호들에 의해 이미 점유되어 있는지에 관한 (대상 할당 벡터 $\boldsymbol{v}_{\mathrm{A,T}}(k-1)$에 포함된) 정보에 따라, 주변 HOA 성분의 어느 계수 시퀀스들이 주어진 I 개의 채널들에서 전송되어야 하는지가 결정된다. 그에 부가하여, 선택된 계수 시퀀스들의 인덱스들이 연속적인 프레임들 사이에서 변하면, 계수 시퀀스들의 페이드인(fade-in) 및 페이드아웃(fade-out)이 수행된다.

[0031] 게다가, 주변 HOA 성분 $\boldsymbol{C}_{\mathrm{AMB}}(k-2)$의 처음 O_{MIN} 개의 계수 시퀀스들이 인지 코딩되어 전송되도록 항상 선택되고, 여기서 $O_{\mathrm{MIN}}=(N_{\mathrm{MIN}}+1)^2$ 이고, $N_{\mathrm{MIN}} \leq N$ 은 전형적으로 원래의 HOA 표현의 차수보다 더 작은 차수이다. 이 HOA 계수 시퀀스들을 무상관화(de-correlate)시키기 위해, 그들이 단계/스테이지(13)에서 어떤 미리 정의된 방향들 $\boldsymbol{\Omega}_{\mathrm{MIN},d}$, $d=1,\ldots,O_{\mathrm{MIN}}$ 으로부터 충돌하는 방향 신호들(즉, 일반적인 평면파 함수들)로 변환될 수 있다.

[0032] 수정된 주변 HOA 성분 $\boldsymbol{C}_{\mathrm{M,A}}(k-1)$과 함께, 시간 예측된 수정된 주변 HOA 성분 $\boldsymbol{C}_{\mathrm{P,M,A}}(k-1)$이 단계/스테이지(13)에서 계산되고, 타당한 예견을 가능하게 하기 위해 이득 제어 처리 단계들 또는 스테이지들(15, 151)에서 사용되며, 주변 HOA 성분의 수정에 관한 정보는 채널 할당 단계 또는 스테이지(14)에서 모든 가능한 유형들의 신호들을 이용가능한 채널들에 할당하는 것에 직접 관련되어 있다. 그 할당에 관한 최종 정보는 최종 할당 벡터 $\boldsymbol{v}_{\mathrm{A}}(k-2)$에 포함되는 것으로 가정된다. 단계/스테이지(13)에서 이 벡터를 계산하기 위해, 대상 할당 벡터 $\boldsymbol{v}_{\mathrm{A,T}}(k-1)$에 포함된 정보가 이용된다.

[0033] 단계/스테이지(14)에서의 채널 할당은, 할당 벡터 $\boldsymbol{v}_{\mathrm{A}}(k-2)$에 의해 제공되는 정보를 사용해, 프레임 $\boldsymbol{X}_{\mathrm{PS}}(k-2)$에 포함된 그리고 프레임 $\boldsymbol{C}_{\mathrm{M,A}}(k-2)$에 포함된 적절한 신호들을 I 개의 이용가능한 채널들에 할당하여, 신호 프레임들 $\boldsymbol{y}_i(k-2)$, $i=1,\ldots,I$ 를 산출한다. 게다가, 프레임 $\boldsymbol{X}_{\mathrm{PS}}(k-1)$에 그리고 프레임 $\boldsymbol{C}_{\mathrm{P,AMB}}(k-1)$에 포함된 적절한 신호들이 또한 I 개의 이용가능한 채널들에 할당되어, 예측된 신호 프레임들 $\boldsymbol{y}_{\mathrm{P},i}(k-1)$, $i=1,\ldots,I$ 를 산출한다.

[0034] 신호 프레임들 $\boldsymbol{y}_i(k-2)$, $i=1,\ldots,I$ 각각은 이득 제어(15, 151)에 의해 최종적으로 처리되어, 지수들 $e_i(k-2)$ 및 예외 플래그들 $\beta_i(k-2)$, $i=1,\ldots,I$ 가 얻어지고, 신호 이득이 인지 인코더 단계들 또는 스테이지들(16)에 적당한 값 범위를 달성하는 것 등을 위해 매끄럽게 수정되는 신호들 $\boldsymbol{z}_i(k-2)$, $i=1,\ldots,I$ 가 얻어진다. 단계들/스테이지들(16)은 대응하는 인코딩된 신호 프레임들 $\breve{\boldsymbol{z}}_i(k-2)$, $i=1,\ldots,I$ 를 출력한다. 예측된 신호 프레임들 $\boldsymbol{y}_{\mathrm{P},i}(k-1)$, $i=1,\ldots,I$ 는 연속적인 블록들 사이의 심각한 이득 변화들을 피하기 위해 일종의 예견을 가능하게 한다. 보조 정보 데이터 $\mathcal{M}_{\mathrm{DIR}}(k-1)$, $\mathcal{M}_{\mathrm{VEC}}(k-1)$, $e_i(k-2)$, $\beta_i(k-2)$, $\zeta(k-1)$ 및 $\boldsymbol{v}_{\mathrm{A}}(k-2)$는 보조 정보 소스 코더 단계 또는 스테이지(17)에서 소스 코딩되어, 인코딩된 보조 정보 프레임 $\breve{\boldsymbol{\Gamma}}(k-2)$가 얻어진다. 멀티플렉서(18)에서, 프레임 $(k-2)$의 인코딩된 신호들 $\breve{\boldsymbol{z}}_i(k-2)$와 이 프레임에 대한 인코딩된 보조 정보 데이터 $\breve{\boldsymbol{\Gamma}}(k-2)$가 결합되어, 출력 프레임 $\breve{\boldsymbol{B}}(k-2)$가 얻어진다.

[0035] 공간 HOA 디코더에서, 단계들/스테이지들(15, 151)에서의 이득 수정들은 지수들 $e_i(k-2)$ 및 예외 플래그들

$\beta_i(k-2)$ $i=1,\ldots,I$, 로 이루어져 있는 이득 제어 보조 정보를 사용하여 복귀되는 것으로 가정된다.

[0036] *HOA 압축해제*

[0037] EP 2800401 A1에 기술된 HOA 압축해제기의 전체적인 아키텍처가 도 2에 예시되어 있다. HOA 압축해제기는, 역순으로 배열되어 있고 도 2의 A에 도시된 인지 및 소스 디코딩 부분과 도 2의 B에 도시된 공간 HOA 디코딩 부분을 포함하는, HOA 압축기 컴포넌트들의 대응물들로 이루어져 있다.

[0038] (인지 및 보조 정보 소스 디코더를 나타내는) 인지 및 소스 디코딩 부분에서, 디멀티플렉싱 단계 또는 스테이지(21)는 비트 스트림으로부터 입력 프레임 $\breve{\boldsymbol{B}}(k)$를 수신하고 I 개의 신호들의 인지 코딩된 표현 $\breve{\boldsymbol{z}}_i(k)$, $i=1,\ldots,I$ 및 그의 HOA 표현을 어떻게 생성해야 하는지를 기술하는 코딩된 보조 정보 데이터 $\breve{\boldsymbol{\Gamma}}(k)$를 제공한다. $\breve{\boldsymbol{z}}_i(k)$ 신호들은 인지 디코더 단계 또는 스테이지(22)에서 인지 디코딩되어, 디코딩된 신호들 $\hat{\boldsymbol{z}}_i(k)$ $i=1,\ldots,I$가 얻어진다. 코딩된 보조 정보 데이터 $\breve{\boldsymbol{\Gamma}}(k)$는 보조 정보 소스 디코더 단계 또는 스테이지(23)에서 디코딩되어, 데이터 세트들 $\mathcal{M}_{\mathrm{DIR}}(k+1)$, $\mathcal{M}_{\mathrm{VEC}}(k+1)$, 지수들 $e_i(k)$, 예외 플래그들 $\beta_i(k)$, 예측 파라미터들 $\zeta(k+1)$ 및 할당 벡터 $\boldsymbol{v}_{\mathrm{AMB,ASSIGN}}(k)$가 얻어진다. $\boldsymbol{v}_{\mathrm{A}}$와 $\boldsymbol{v}_{\mathrm{AMB,ASSIGN}}$ 사이의 차이에 관해서는, 앞서 언급된 MPEG 문서 N14264를 참조하기 바란다.

[0039] *공간 HOA 디코딩*

[0040] 공간 HOA 디코딩 부분에서, 인지 디코딩된 신호들 $\hat{\boldsymbol{z}}_i(k)$ $i=1,\ldots,I$ 각각은 역 이득 제어 처리 단계 또는 스테이지(24, 241)에, 그의 연관된 이득 보정 지수 $e_i(k)$ 및 이득 보정 예외 플래그 $\beta_i(k)$와 함께, 입력된다. 제 i 역 이득 제어 처리 단계/스테이지는 이득 보정 신호 프레임 $\hat{\boldsymbol{y}}_i(k)$를 제공한다.

[0041] I 개의 이득 보정된 신호 프레임들 $\hat{\boldsymbol{y}}_i(k)$ $i=1,\ldots,I$ 모두가, 할당 벡터 $\boldsymbol{v}_{\mathrm{AMB,ASSIGN}}(k)$ 및 튜플 세트들 $\mathcal{M}_{\mathrm{DIR}}(k+1)$ 및 $\mathcal{M}_{\mathrm{VEC}}(k+1)$과 함께, 채널 재할당 단계 또는 스테이지(25)에 피드된다 - 튜플 세트들 $\mathcal{M}_{\mathrm{DIR}}(k+1)$ 및 $\mathcal{M}_{\mathrm{VEC}}(k+1)$의 앞서 기술된 정의를 참조 -. 할당 벡터 $\boldsymbol{v}_{\mathrm{AMB,ASSIGN}}(k)$는, 각각의 전송 채널에 대해, 전송 채널이 주변 HOA 성분의 계수 시퀀스를 포함하는지 그리고 전송 채널이 어느 것을 포함하는지를 나타내는 I 개의 성분들로 이루어져 있다. 채널 재할당 단계/스테이지(25)에서, 이득 보정된 신호 프레임들 $\hat{\boldsymbol{y}}_i(k)$는 모든 우세 음향 신호들(즉, 모든 방향 및 벡터 기반 신호들)의 프레임 $\hat{\boldsymbol{X}}_{\mathrm{PS}}(k)$ 및 주변 HOA 성분의 중간 표현의 프레임 $\boldsymbol{C}_{\mathrm{I,AMB}}(k)$를 재구성하기 위해 재분포된다. 그에 부가하여, 제 k 프레임에서 활성인 주변 HOA 성분의 계수 시퀀스들의 인덱스들의 세트 $\mathcal{I}_{\mathrm{AMB,ACT}}(k)$, 및 제 $(k-1)$ 프레임에서 인에이블되고, 디스에이블되며, 활성인 채로 있어야만 하는 주변 HOA 성분의 계수 인덱스들의 데이터 세트들 $\mathcal{I}_{\mathrm{E}}(k-1)$, $\mathcal{I}_{\mathrm{D}}(k-1)$ 및 $\mathcal{I}_{\mathrm{U}}(k-1)$이 제공된다.

[0042] 우세 음향 합성(predominant sound synthesis) 단계 또는 스테이지(26)에서, 우세 음향 성분 $\hat{\boldsymbol{C}}_{\mathrm{PS}}(k-1)$의 HOA 표현이 모든 우세 음향 신호들의 프레임 $\hat{\boldsymbol{X}}_{\mathrm{PS}}(k)$로부터, 튜플 세트 $\mathcal{M}_{\mathrm{DIR}}(k+1)$, 예측 파라미터들의 세트

$\zeta(k+1)$, 튜플 세트 $\mathcal{M}_{\mathrm{VEC}}(k+1)$ 그리고 데이터 세트들 $\mathcal{I}_{\mathrm{E}}(k-1)$, $\mathcal{I}_{\mathrm{D}}(k-1)$ 및 $\mathcal{I}_{\mathrm{U}}(k-1)$을 사용하여, 계산된다.

[0043] 앰비언스 합성(ambience synthesis) 단계 또는 스테이지(27)에서, 주변 HOA 성분 프레임 $\widetilde{\boldsymbol{C}}_{\mathrm{AMB}}(k-1)$이 주변 HOA 성분의 중간 표현의 프레임 $\boldsymbol{C}_{\mathrm{I,AMB}}(k)$로부터, 제 k 프레임에서 활성인 주변 HOA 성분의 계수 시퀀스들의 인덱스들의 세트 $\mathcal{I}_{\mathrm{AMB,ACT}}(k)$를 사용하여, 생성된다. 우세 음향 HOA 성분과의 동기화로 인해 하나의 프레임의 지연이 유입된다.

[0044] 마지막으로, HOA 조성(HOA composition) 단계 또는 스테이지(28)에서, 주변 HOA 성분 프레임 $\widetilde{\boldsymbol{C}}_{\mathrm{AMB}}(k-1)$및 우세 음향 HOA 성분의 프레임 $\widetilde{\boldsymbol{C}}_{\mathrm{PS}}(k-1)$이 디코딩된 HOA 프레임 $\widetilde{\boldsymbol{C}}(k-1)$을 제공하기 위해 중첩된다.

[0045] 그 후에, 공간 HOA 디코더는, I 개의 신호들 및 보조 정보로부터, 재구성된 HOA 표현을 생성한다.

[0046] 인코딩측에서, 주변 HOA 성분이 방향 신호들로 변환된 경우에, 그 변환이 디코더측에서 단계/스테이지(27)에서 반대로 행해진다.

[0047] HOA 압축기 내에서 이득 제어 처리 단계들/스테이지들(15, 151) 이전의 신호들의 잠재적인 최대 이득들은 입력 HOA 표현의 값 범위에 많이 의존한다. 그러므로, 처음에 입력 HOA 표현에 대한 의미있는 값 범위가 정의되고, 이어서 이득 제어 처리 단계들/스테이지들에 들어가기 전에 신호들의 잠재적인 최대 이득들에 대해 결론을 내린다.

[0048] *입력 HOA 표현의 정규화*

[0049] 본 발명의 방법을 사용하기 위해, (총) 입력 HOA 표현 신호의 정규화가 이전에 수행되어야만 한다. HOA 압축을 위해, 프레임 단위가 수행되고, 여기서 원래의 입력 HOA 표현의 제 k 프레임 $\boldsymbol{C}(k)$는 고차 앰비소닉스의 기초 섹션에서 수학식 54에 명시된 시간 연속적 HOA 계수 시퀀스들의 벡터 $\boldsymbol{c}(t)$와 관련하여 수학식 1로서 정의되고,

[0050] [수학식 1]

[0051] $$\boldsymbol{C}(k) := \left[\boldsymbol{c}\left((kL+1)T_{\mathrm{S}}\right) \quad \boldsymbol{c}\left((kL+2)T_{\mathrm{S}}\right) \quad \ldots \quad \boldsymbol{c}\left((k+1)LT_{\mathrm{S}}\right)\right] \in \mathbb{R}^{O\times L}$$

[0052] 여기서 k는 프레임 인덱스를 나타내고, L은 프레임 길이(단위: 샘플)를 나타내며, $O=(N+1)^2$은 HOA 계수 시퀀스들의 수를 나타내고, T_{S}는 샘플링 주기를 나타낸다.

[0053] EP 2824661 A1에서 언급된 바와 같이, 실제적 관점에서 본 HOA 표현의 의미있는 정규화는 개개의 HOA 계수 시퀀스들 $c_n^m(t)$의 값 범위에 제약조건들을 부과하는 것에 의해 달성되지 않는데, 그 이유는 이 시간 영역 함수들이 렌더링 후에 스피커들에 의해 실제로 재생되는 신호들이 아니기 때문이다. 그 대신에, HOA 표현을 O개의 가상 스피커 신호들 $w_j(t)$, $1\le j\le O$로 렌더링하는 것에 의해 획득되는, '등가의 공간 영역 표현'을 고려하는 것이 보다 편리하다. 각자의 가상 스피커 위치들이 구면 좌표계에 의해 표현되는 것으로 가정되고, 여기서 각각의 위치는 단위 구면 상에 있는 것으로 그리고 '1'의 반경을 갖는 것으로 가정된다. 그러므로, 위치들이 차수 의존적 방향들 $\boldsymbol{\Omega}_j^{(N)}=(\theta_j^{(N)},\phi_j^{(N)})$, $1\le j\le O$에 의해 등가적으로 표현될 수 있고, 여기서 $\theta_j^{(N)}$ 및 $\phi_j^{(N)}$은, 각각, 경사들 및 방위각들을 나타낸다(또한 도 6 및 구면 좌표계의 정의에 대한 그의 설명을 참조). 이 방향들은 단위 구면 상에 가능한 균일하게 분포되어야만 하고, 예컨대, J. Fliege, U. Maier, "A two-stage approach for computing cubature formulae for the sphere", Technical report, Fachbereich Mathematik, University of Dortmund, 1999를 참조하기 바란다. 노드 번호들은 특정 방향들의 계산에 대한

http://www.mathematik.uni-dortmund.de/lsx/research/projects/fliege/nodes/nodes.html에 있다. 이 위치들은 일반적으로 '구면 상에서의 균일한 분포'의 정의의 종류에 의존하고, 따라서 명확하지 않다.

[0054] HOA 계수 시퀀스들에 대한 값 범위들을 정의하는 것보다 가상 스피커 신호들에 대한 값 범위들을 정의하는 것의 장점은 전자에 대한 값 범위가 PCM 표현을 가정하는 종래의 스피커 신호들에 대해서와 같이 똑같이 직관적으로 구간 $[-1,1[$로 설정될 수 있다는 것이다. 이것은 공간적으로 균일하게 분포된 양자화 오차를 가져오며, 따라서 유리하게도 실제의 청취에 대해 관련성 있는 영역(domain)에서 양자화가 적용된다. 이와 관련하여 중요한 측면은 샘플당 비트 수가 종래의 스피커 신호들에 대해 전형적인 것만큼 낮게 - 즉, 16으로 - 선택될 수 있고, 이는, 보통 보다 높은 샘플당 비트 수(예컨대, 24 또는 심지어 32)가 요구되는, HOA 계수 시퀀스들의 직접 양자화와 비교하여 효율을 증가시킨다는 것이다.

[0055] 공간 영역에서의 정규화 프로세스를 상세히 기술하기 위해, 모든 가상 스피커 신호들이 벡터로 수학식 2로서 요약되고

[0056] [수학식 2]

[0057] $$\boldsymbol{w}(t) := [w_1(t) \quad \dots \quad w_O(t)]^T$$

[0058] 여기서 $(\cdot)^T$는 전치(transposition)를 나타낸다. 가상 방향들 $\boldsymbol{\Omega}_j^{(N)},\ 1 \le j \le O$에 대한 모드 행렬을 $\boldsymbol{\Psi}$ - 이는 수학식 3에 의해 정의되고

[0059] [수학식 3]

[0060] $$\boldsymbol{\Psi} := [S_1 \quad \dots \quad S_O] \in \mathbb{R}^{O \times O}$$

[0061] 여기서

[0062] [수학식 4]

[0063] $$\boldsymbol{S}_j :=$$

[0064] $$[S_0^0(\boldsymbol{\Omega}_j^{(N)}) \quad S_1^{-1}(\boldsymbol{\Omega}_j^{(N)}) \quad S_1^0(\boldsymbol{\Omega}_j^{(N)}) \quad S_1^1(\boldsymbol{\Omega}_j^{(N)}) \quad \dots \quad S_N^{N-1}(\boldsymbol{\Omega}_j^{(N)}) \quad S_N^N(\boldsymbol{\Omega}_j^{(N)})]^T$$ 임 -로 나타내면,

[0065] 렌더링 프로세스는 행렬 곱셈

[0066] [수학식 5]

[0067] $$\boldsymbol{w}(t) = (\boldsymbol{\Psi})^{-1} \cdot \boldsymbol{c}(t)$$

[0068] 로서 수식화될 수 있다.

[0069] 이 정의들을 사용하여, 가상 스피커 신호들에 대한 타당한 요구사항은 수학식 6이고:

[0070] [수학식 6]

[0071] $$\|\boldsymbol{w}(lT_S)\|_\infty = \max_{1 \le j \le O} |w_j(lT_S)| \le 1 \qquad \forall l$$

[0072] 수학식 6은 각각의 가상 스피커 신호의 크기가 범위 $[-1,1[$ 내에 있어야만 한다는 것을 의미한다. 시간 t의 시각(time instant)은 샘플 인덱스 l 및 상기 HOA 데이터 프레임들의 샘플 값들의 샘플 주기 T_S에 의해 표현된다.

[0073] 스피커 신호들의 총 전력은 결과적으로 조건 수학식 7을 충족시킨다.

[0074] [수학식 7]

[0075] $$\|\boldsymbol{w}(lT_S)\|_2^2 = \sum_{j=1}^{O} |w_j(lT_S)|^2 \leq O \qquad \forall l$$

[0076] HOA 데이터 프레임 표현의 렌더링 및 정규화는 도 1의 A의 입력 $\boldsymbol{C}(k)$ 의 전단(upstream)에서 수행된다.

[0077] *이득 제어 이전의 신호 값 범위에 대한 결과*

[0078] 입력 HOA 표현의 정규화가 *입력 HOA 표현의 정규화* 섹션에서의 설명에 따라 수행되는 것으로 가정하면, HOA 압축기에서의 이득 제어 처리 유닛(15, 151)에 입력되는, 신호들 $\boldsymbol{y}_i$, $i = 1, \ldots, I$ 의 값 범위가 이하에서 고려된다. 이 신호들은 HOA 계수 시퀀스들, 또는 우세 음향 신호들 $\boldsymbol{x}_{\mathrm{PS},d}$, $d = 1, \ldots, D$, 및/또는 주변 HOA 성분 $\boldsymbol{c}_{\mathrm{AMB},n}$, $n = 1, \ldots, O$ 의 특정의 계수 시퀀스들 - 이들의 일부에 공간 변환이 적용됨 - 중 하나 이상을 이용가능한 I 개의 채널들에 할당하는 것에 의해 생성된다. 그러므로, 수학식 6에서의 정규화 가정 하에서 이 언급된 상이한 신호 유형들의 가능한 값 범위를 분석하는 것이 필요하다. 모든 종류의 신호들이 원래의 HOA 계수 시퀀스들로부터 즉각 계산되기 때문에, 그들의 가능한 값 범위들이 검토된다.

[0079] 하나 이상의 HOA 계수 시퀀스들만이 I 개의 채널들에 포함되어 있는 경우가 도 1의 A 및 도 2의 B에 도시되어 있지 않으며, 즉 이러한 경우에, HOA 분해, 주변 성분 수정 및 대응하는 합성 블록들이 필요하지 않게 된다.

[0080] *HOA 표현의 값 범위에 대한 결과*

[0081] 시간 연속적 HOA 표현이 수학식 8에 의해 가상 스피커 신호들로부터 획득되고,

[0082] [수학식 8]

[0083] $$\boldsymbol{c}(t) = \boldsymbol{\Psi}\boldsymbol{w}(t)$$

[0084] 수학식 8은 수학식 5에서의 연산의 역연산(inverse operation)이다. 그러므로, 모든 HOA 계수 시퀀스들의 총 전력은 수학식 8 및 수학식 7을 사용하여 다음과 같이 제한된다:

[0085] [수학식 9]

[0086] $$\|\boldsymbol{c}(lT_S)\|_2^2 \leq \|\boldsymbol{\Psi}\|_2^2 \cdot \|\boldsymbol{w}(lT_S)\|_2^2 \leq \|\boldsymbol{\Psi}\|_2^2 \cdot O$$

[0087] 구면 조화함수 함수들의 N3D 정규화의 가정 하에서, 모드 행렬의 유클리드 노름의 제곱은 수학식 10a와 같이 써질 수 있고,

[0088] [수학식 10a]

[0089] $$\|\boldsymbol{\Psi}\|_2^2 = K \cdot O$$

[0090] 여기서

[0091] [수학식 10b]

[0092] $$K = \frac{\|\boldsymbol{\Psi}\|_2^2}{O}$$

[0093] 은 모드 행렬의 유클리드 노름의 제곱과 HOA 계수 시퀀스들의 수 O 사이의 비이다. 이 비는 특정 HOA 차수 N 및 특정 가상 스피커 방향들 $\boldsymbol{\Omega}_j^{(N)}$, $1 \leq j \leq O$ 에 의존하며, 수학식 10c와 같이 각자의 파라미터 목록을 비에 부가하는 것에 의해 표현될 수 있다.

[0094] [수학식 10c]

[0095] $$K = K(N, \boldsymbol{\Omega}_1^{(N)}, \ldots, \boldsymbol{\Omega}_O^{(N)})$$

[0096] 도 3은 HOA 차수들 $N = 1, \ldots, 29$ 에 대한 앞서 언급된 Fliege 등의 논문에 따른 가상 방향들 $\boldsymbol{\Omega}_j^{(N)},\ 1 \le j \le O,$ 에 대한 K 의 값들을 나타내고 있다.

[0097] 모든 이전의 인수(argument)들과 고려사항들을 결합하는 것은 다음과 같이 HOA 계수 시퀀스들의 크기에 대한 상한을 제공하고:

[0098] [수학식 11]

[0099] $$\boxed{\|\boldsymbol{c}(lT_S)\|_\infty \le \|\boldsymbol{c}(lT_S)\|_2 \le \sqrt{K} \cdot O}$$

[0100] 여기서 첫 번째 부등식은 노름 정의들로부터 직접 나온 것이다.

[0101] 유의할 중요한 점은 수학식 6에서의 조건이 수학식 11에서의 조건을 암시하지만, 그 반대가 성립하지 않는다 - 즉, 수학식 11이 수학식 6을 암시하지 않는다 - 는 것이다.

[0102] 추가적인 중요한 측면은, 거의 균일하게 분포된 가상 스피커 위치들의 가정 하에서, 가상 스피커 위치들과 관련하여 모드 벡터들을 표현하는 모드 행렬 $\boldsymbol{\Psi}$ 의 열 벡터들이 서로 거의 직교이고 각각이 $N+1$ 의 유클리드 노름을 갖는다는 것이다. 이 특성은 공간 변환이 곱셈 상수(multiplicative constant)를 제외하고는 유클리드 노름을 거의 보존한다는 것을 의미하며, 즉,

[0103] [수학식 12]

[0104] $$\|\boldsymbol{c}(lT_S)\|_2 \approx (N+1)\|\boldsymbol{w}(lT_S)\|_2$$

[0105] 이다.

[0106] 실제 노름 $\|\boldsymbol{c}(lT_S)\|_2$ 가 수학식 12에서의 근사화와 많이 상이할수록, 모드 벡터들에 대한 직교성 가정이 많이 위반된다.

[0107] *우세 음향 신호들의 값 범위에 대한 결과*

[0108] 양 유형의 우세 음향 신호들(방향 및 벡터 기반)의 공통점은 HOA 표현에 대한 그들의 기여가 $N+1$ 의 유클리드 노름을 갖는 단일의 벡터 $\boldsymbol{v}_1 \in \mathbb{R}^O$ 에 의해 기술된다는 것이며, 즉

[0109] [수학식 13]

[0110] $$\|\boldsymbol{v}_1\|_2 = N+1$$

[0111] 이다. 방향 신호의 경우에, 이 벡터는 특정의 신호 소스 방향 $\boldsymbol{\Omega}_{S,1}$ 에 대한 모드 벡터에 대응하며, 즉

[0112] [수학식 14]

[0113] $$\boldsymbol{v}_1 = \boldsymbol{S}(\boldsymbol{\Omega}_{S,1})$$

[0114] [수학식 15]

[0115] $$:= [S_0^0(\boldsymbol{\Omega}_{S,1}) \quad S_1^{-1}(\boldsymbol{\Omega}_{S,1}) \quad S_1^0(\boldsymbol{\Omega}_{S,1}) \quad S_1^1(\boldsymbol{\Omega}_{S,1}) \quad \ldots \quad S_N^{N-1}(\boldsymbol{\Omega}_{S,1}) \quad S_N^N(\boldsymbol{\Omega}_{S,1})]^T$$

[0116] 이다. 이 벡터는 HOA 표현에 의해 신호 소스 방향 $\boldsymbol{\Omega_{S,1}}$ 으로의 방향 빔을 기술한다. 벡터 기반 신호의 경우에, 벡터 $\boldsymbol{v_1}$ 은 임의의 방향에 대해 모드 벡터로 제약되지 않고, 그러므로 모노럴 벡터 기반 신호의 보다 일반적인 방향 분포를 기술할 수 있다.

[0117] 이하에서 D 개의 우세 음향 신호들 $\boldsymbol{x_d(t)}$, $\boldsymbol{d = 1, \dots, D}$ - 이들은 수학식 16에 따라 벡터 $\boldsymbol{x(t)}$ 에 모여 있을 수 있음 - 의 일반적인 경우가 고려된다.

[0118] [수학식 16]

[0119] $$\boldsymbol{x(t) = [x_1(t) \quad x_2(t) \quad \dots \quad x_D(t)]^T}$$

[0120] 이 신호들은 모노럴 우세 음향 신호들 $\boldsymbol{x_d(t)}$, $\boldsymbol{d = 1, \dots, D}$ 의 방향 분포를 나타내는 모든 벡터들 $\boldsymbol{v_d}$, $\boldsymbol{d = 1, \dots, D}$ 로 형성되는 수학식 17의 행렬에 기초하여 결정되어야만 한다.

[0121] [수학식 17]

[0122] $$\boldsymbol{V := [v_1 \quad v_2 \quad \dots \quad v_D]}$$

[0123] 우세 음향 신호들 $\boldsymbol{x(t)}$ 의 의미있는 추출을 위해, 이하의 제약조건들이 수식화된다:

[0124] a) 각각의 우세 음향 신호가, 수학식 18과 같이, 원래의 HOA 표현의 계수 시퀀스들의 선형 결합으로서 획득되며,

[0125] [수학식 18]

[0126] $$\boldsymbol{x(t) = A \cdot c(t)}$$

[0127] 여기서 $\boldsymbol{A \in \mathbb{R}^{D \times O}}$ 는 믹싱 행렬을 나타낸다.

[0128] b) 믹싱 행렬 $\boldsymbol{A}$ 는 그의 유클리드 노름이 '1'의 값을 초과하지 않도록 - 즉,

[0129] [수학식 19]

[0130] $$\boldsymbol{\|A\|_2 \overset{!}{\leq} 1}$$

[0131] 이도록 - 그리고 원래의 HOA 표현과 우세 음향 신호들의 HOA 표현 사이의 잔차의 유클리드 노름의 제곱(또는 동등하게 전력)이 원래의 HOA 표현의 유클리드 노름의 제곱(또는 동등하게 전력) 이하이도록 - 즉,

[0132] [수학식 20]

[0133] $$\boldsymbol{\|c(t) - V \cdot x(t)\|_2^2 \overset{!}{\leq} \|c(t)\|_2^2}$$

[0134] - 이도록 선택되어야만 한다.

[0135] 수학식 18을 수학식 20에 삽입함으로써, 수학식 20이 수학식 21의 제약조건과 동등하다는 것을 알 수 있고,

[0136] [수학식 21]

[0137] $$\boldsymbol{\|I - V \cdot A\|_2 \overset{!}{\leq} 1}$$

[0138] 여기서 $\boldsymbol{I}$ 는 항등 행렬을 나타낸다.

[0139] 수학식 18 및 수학식 19에서의 제약조건들로부터 그리고 유클리드 행렬 및 벡터 노름들의 호환성으로부터, 우세 음향 신호들의 크기들에 대한 상한이 수학식 18, 수학식 19 및 수학식 11을 사용하여 다음과 같이 구해진다.

[0140] [수학식 22]

[0141] $$\|\boldsymbol{x}(lT_S)\|_\infty \le \|\boldsymbol{x}(lT_S)\|_2$$

[0142] [수학식 23]

[0143] $$\le \|\boldsymbol{A}\|_2 \|\boldsymbol{c}(lT_S)\|_2$$

[0144] [수학식 24]

[0145] $$\le \sqrt{K} \cdot O \quad ,$$

[0146] 그러므로, 우세 음향 신호들이, 수학식 25와 같이, 원래의 HOA 계수 시퀀스들과 동일한 범위에 있도록 보장된다 (수학식 11과 비교).

[0147] [수학식 25]

[0148] $$\boxed{\|\boldsymbol{x}(lT_S)\|_\infty \le \sqrt{K} \cdot O}$$

[0149] *믹싱 행렬의 선택에 대한 예*

[0150] 제약조건(수학식 20)을 충족시키는 믹싱 행렬을 어떻게 결정해야 하는지의 일 예는, 수학식 26과 같이, 추출 이후의 잔차의 유클리드 노름이 최소화되도록 우세 음향 신호들을 계산하는 것에 의해 얻어진다.

[0151] [수학식 26]

[0152] $$\boldsymbol{x}(t) = \mathrm{argmin}_{\boldsymbol{x}(t)} \|\boldsymbol{V} \cdot \boldsymbol{x}(t) - \boldsymbol{c}(t)\|_2$$

[0153] 수학식 26에서의 최소화 문제에 대한 해결책은 수학식 27에 의해 주어지고,

[0154] [수학식 27]

[0155] $$\boldsymbol{x}(t) = \boldsymbol{V}^+ \boldsymbol{c}(t)$$

[0156] 여기서 $(\cdot)^+$ 는 Moore-Penrose 의사 역행렬(pseudo-inverse)을 나타낸다. 수학식 27과 수학식 18을 비교해 보면, 이 경우에, 믹싱 행렬이 행렬 $\boldsymbol{V}$ 의 Moore-Penrose 의사 역행렬과 같은 것으로 된다 - 즉, $\boldsymbol{A} = \boldsymbol{V}^+$ 임 -.

[0157] 그럼에도 불구하고, 행렬 $\boldsymbol{V}$ 는 여전히 제약조건(수학식 19)을 충족시키도록, 즉 수학식 28이도록, 선택되어야만 한다.

[0158] [수학식 28]

[0159] $$\boxed{\|\boldsymbol{V}^+\|_2 \overset{!}{\le} 1}$$

[0160] 행렬 $\boldsymbol{V}$ 가 어떤 소스 신호 방향들 $\boldsymbol{\Omega}_{S,d}$, $d = 1, \dots, D$ 에 대한 모드 행렬 - 즉, 수학식 29 - 인, 방향 신호들만의 경우에,

[0161] [수학식 29]

[0162] $$\boldsymbol{V} = [\boldsymbol{S}(\boldsymbol{\Omega}_{S,1}) \quad \boldsymbol{S}(\boldsymbol{\Omega}_{S,2}) \quad \dots \quad \boldsymbol{S}(\boldsymbol{\Omega}_{S,D})]$$

[0163] 제약조건(수학식 28)은, 임의의 2개의 이웃하는 방향들의 거리가 너무 작지 않도록, 소스 신호 방향들 $\boldsymbol{\Omega}_{S,d}$, $d = 1, \ldots, D$ 를 선택하는 것에 의해 충족될 수 있다.

[0164] *주변 HOA 성분의 계수 시퀀스들의 값 범위에 대한 결과*

[0165] 주변 HOA 성분은 원래의 HOA 표현을 우세 음향 신호들의 HOA 표현으로부터 차감하는 것에 의해 - 즉, 수학식 30에 의해 - 계산된다.

[0166] [수학식 30]

[0167] $$\boldsymbol{c}_{\mathrm{AMB}}(t) = \boldsymbol{c}(t) - \boldsymbol{V} \cdot \boldsymbol{x}(t)$$

[0168] 우세 음향 신호들의 벡터 $\boldsymbol{x}(t)$ 가 기준(수학식 20)에 따라 결정되면, 다음과 같이 결론지어질 수 있다:

[0169] [수학식 31]

[0170] $$\|\boldsymbol{c}_{\mathrm{AMB}}(lT_{\mathrm{S}})\|_{\infty} \leq \|\boldsymbol{c}_{\mathrm{AMB}}(lT_{\mathrm{S}})\|_{2}$$

[0171] [수학식 32]

[0172] $$\overset{(30)}{=} \|\boldsymbol{c}(lT_{\mathrm{S}}) - \boldsymbol{V} \cdot \boldsymbol{x}(lT_{\mathrm{S}})\|_{2}$$

[0173] [수학식 33]

[0174] $$\overset{(20)}{\leq} \|\boldsymbol{c}(lT_{\mathrm{S}})\|_{2}$$

[0175] [수학식 34]

[0176] $$\overset{(11)}{=} \sqrt{K} \cdot O$$

[0177] *주변 HOA 성분의 공간 변환된 계수 시퀀스들의 값 범위*

[0178] EP 2743922 A1에서 그리고 앞서 언급된 MPEG 문서 N14264에서 제안된 HOA 압축 처리에서의 추가적인 측면은 주변 HOA 성분의 처음 O_{MIN} 개의 계수 시퀀스들이 항상 전송 채널들에 할당되도록 선택되고, 여기서 $O_{\mathrm{MIN}} = (N_{\mathrm{MIN}} + 1)^2$ 이고 $N_{\mathrm{MIN}} \leq N$ 은 전형적으로 원래의 HOA 표현의 차수보다 더 작은 차수이다. 이 HOA 계수 시퀀스들을 무상관화시키기 위해, 이들이 (*입력 HOA 표현의 정규화* 섹션에 기술된 개념과 유사하게) 어떤 미리 정의된 방향들 $\boldsymbol{\Omega}_{\mathrm{MIN},d}$, $d = 1, \ldots, O_{\mathrm{MIN}}$ 로부터 충돌하는 가상 스피커 신호들로 변환될 수 있다.

[0179] 차수 인덱스 $n \leq N_{\mathrm{MIN}}$ 을 갖는 주변 HOA 성분의 모든 계수 시퀀스들의 벡터를 $\boldsymbol{c}_{\mathrm{AMB,MIN}}(t)$ 에 의해 그리고 가상 방향들 $\boldsymbol{\Omega}_{\mathrm{MIN},d}$, $d = 1, \ldots, O_{\mathrm{MIN}}$ 에 대한 모드 행렬을 $\boldsymbol{\Psi}_{\mathrm{MIN}}$ 에 의해 정의하면, 모든 가상 스피커 신호들의 벡터 $\boldsymbol{w}_{\mathrm{MIN}}(t)$ (모든 가상 스피커 신호들은 그에 의해 정의됨)는 수학식 35에 의해 획득된다:

[0180] [수학식 35]

[0181] $$\boldsymbol{w}_{\mathrm{MIN}}(t) = \boldsymbol{\Psi}_{\mathrm{MIN}}^{-1} \cdot \boldsymbol{c}_{\mathrm{AMB,MIN}}(t)$$

[0182] 그러므로, 유클리드 행렬과 벡터 노름들의 호환성을 사용하여,

[0183] [수학식 36]

[0184] $$\|\boldsymbol{w}_{\mathrm{MIN}}(lT_S)\|_\infty \le \|\boldsymbol{w}_{\mathrm{MIN}}(lT_S)\|_2$$

[0185] [수학식 37]

[0186] $$\overset{(35)}{\le} \|\boldsymbol{\Psi}_{\mathrm{MIN}}^{-1}\|_2 \cdot \left\|\boldsymbol{c}_{\mathrm{AMB,MIN}}(lT_S)\right\|_2$$

[0187] [수학식 38]

[0188] $$\overset{(34)}{\le} \|\boldsymbol{\Psi}_{\mathrm{MIN}}^{-1}\|_2 \cdot \sqrt{K} \cdot O$$.

[0189] 앞서 언급된 MPEG 문서 N14264에서, 가상 방향들 $\boldsymbol{\Omega}_{\mathrm{MIN},d}$, $d = 1,\dots,O_{\mathrm{MIN}}$ 은 앞서 언급된 Fliege 등의 논문에 따라 선택된다. 모드 행렬들 $\boldsymbol{\Psi}_{\mathrm{MIN}}$ 의 역행렬의 각자의 유클리드 노름들이 차수들 $N_{\mathrm{MIN}} = 1,\dots,9$ 에 대해 도 4에 예시되어 있다. 수학식 39라는 것을 알 수 있다.

[0190] [수학식 39]

[0191] $$\|\boldsymbol{\Psi}_{\mathrm{MIN}}^{-1}\|_2 < 1 \quad \text{for } N_{\mathrm{MIN}} = 1,\dots,9$$

[0192] 그렇지만, 이것은 $N_{\mathrm{MIN}} > 9$ 에 대해 일반적으로 성립하지 않으며, 여기서 $\|\boldsymbol{\Psi}_{\mathrm{MIN}}^{-1}\|_2$ 의 값들은 전형적으로 '1'보다 훨씬 더 크다. 그럼에도 불구하고, 적어도 $1 \le N_{\mathrm{MIN}} \le 9$ 에 대해, 가상 스피커 신호들의 진폭들이 수학식 40에 의해 제한된다.

[0193] [수학식 40]

[0194] $$\boxed{\|\boldsymbol{w}_{\mathrm{MIN}}(lT_S)\|_\infty \overset{(38),\text{도 4}}{\le} \sqrt{K} \cdot O \quad \text{for } 1 \le N_{\mathrm{MIN}} \le 9}$$

[0195] 이 HOA 표현으로부터 생성된 가상 스피커 신호들의 진폭들이 '1'의 값을 초과하지 않을 것을 요구하는, 조건(수학식 6)을 충족시키도록 입력 HOA 표현을 제약하는 것에 의해, 이득 제어 이전의 신호들의 진폭들이 다음과 같은 조건들 하에서 값 $\sqrt{K} \cdot O$ (수학식 25, 수학식 34 및 수학식 40을 참조)을 초과하지 않을 것임이 보장될 수 있다:

[0196] a) 모든 우세 음향 신호들의 벡터 $x(t)$ 가 수학식/제약조건 18, 수학식/제약조건 19 및 수학식/제약조건 20에 따라 계산된다;

[0197] b) 앞서 언급된 Fliege 등의 논문에서 정의된 것들이 가상 스피커 위치들로서 사용되는 경우, 공간 변환이 적용되는 주변 HOA 성분의 제1 계수 시퀀스들의 수 O_{MIN} 을 결정하는 최소 차수 N_{MIN}은 '9'보다 낮아야만 한다.

[0198] 이득 제어 이전의 신호들의 진폭들이 최대 관심 차수 N_{MAX}까지의 임의의 차수 N 에 대해 값 $\sqrt{K_{\mathrm{MAX}}} \cdot O$ 를 초과하지 않을 것으로 - 즉, $1 \le N \le N_{\mathrm{MAX}}$일 것으로 - 추가로 결론내려질 수 있으며, 여기서 수학식 41a이다.

[0199] [수학식 41a]

[0200] $$K_{\mathrm{MAX}} = \max_{1 \le N \le N_{\mathrm{MAX}}} K\left(N, \boldsymbol{\Omega}_1^{(N)}, \dots, \boldsymbol{\Omega}_O^{(N)}\right)$$

[0201] 상세하게는, 도 3으로부터 결론내려질 수 있다. 초기 공간 변환을 위한 가상 스피커 방향들 $\boldsymbol{\Omega}_j^{(N)}, 1 \le j \le O,$ 가

Fliege 등의 논문에서의 분포에 따라 선택되는 것으로 가정되는 경우 그리고 그에 부가하여, 최대 관심 차수가 (예컨대, MPEG 문서 N14264에서와 같이) $N_{MAX} = 29$ 인 것으로 가정되는 경우, 이득 제어 이전의 신호들의 진폭들이 값 $1.5\ O$ 를 초과하지 않을 것인데, 그 이유는 이 특수한 경우에 $\sqrt{K_{MAX}} < 1.5$ 이기 때문이다. 즉, $\sqrt{K_{MAX}} = 1.5$ 가 선택될 수 있다.

[0202] K_{MAX} 는 최대 관심 차수 N_{MAX} 및 가상 스피커 방향들 $\boldsymbol{\Omega}_j^{(N)}, 1 \le j \le O$ 에 의존하고, 수학식 41b에 의해 표현될 수 있다.

[0203] [수학식 41b]

[0204] $$K_{MAX} = K_{MAX}\left(\left\{\boldsymbol{\Omega}_1^{(N)}, \ldots, \boldsymbol{\Omega}_O^{(N)} \mid 1 \le N \le N_{MAX}\right\}\right)$$

[0205] 그러므로, 인지 코딩 이전의 신호들이 구간 $[-1,1]$ 내에 있도록 하기 위해 이득 제어에 의해 적용되는 최소 이득은 $2^{e_{MIN}}$ 에 의해 주어지고, 여기서 수학식 41c이다.

[0206] [수학식 41c]

[0207] $$e_{MIN} = -\lceil \log_2(\sqrt{K_{MAX}} \cdot O) \rceil < 0$$

[0208] 이득 제어 이전의 신호들의 진폭들이 너무 작은 경우, 그들을 $2^{e_{MAX}}$ 까지의 인수를 사용해 매끄럽게 증폭시키는 것이 가능하고, 여기서 $e_{MAX} \ge 0$ 이 코딩된 HOA 표현 내에서 보조 정보로서 전송되는 것이 MPEG 문서 N14264에 제안되어 있다.

[0209] 이와 같이, 액세스 단위 내에서 첫 번째 프레임부터 현재 프레임까지 이득 제어 처리 유닛에 의해 야기된 수정된 신호의 총 절대 진폭 변화를 기술하는, 밑이 '2'인 각각의 지수가 구간 $[e_{MIN}, e_{MAX}]$ 내의 임의의 정수값을 취할 수 있다. 결과적으로, 그것을 코딩하는 데 필요하게 되는 비트들의 (최저 정수) 개수 β_e 가 수학식 42에 의해 주어진다.

[0210] [수학식 42]

[0211] $$\beta_e = \lceil \log_2(|e_{MIN}| + e_{MAX} + 1) \rceil = \lceil \log_2(\lceil \log_2(\sqrt{K_{MAX}} \cdot O) \rceil + e_{MAX} + 1) \rceil$$

[0212] 이득 제어 이전 신호들의 진폭들이 너무 작지는 않은 경우, 수학식 42가 수학식 42a로 단순화될 수 있다:

[0213] [수학식 42a]

[0214] $$\beta_e = \lceil \log_2(|e_{MIN}| + 1) \rceil = \lceil \log_2(\lceil \log_2(\sqrt{K_{MAX}} \cdot O) \rceil + 1) \rceil$$

[0215] 이 비트 수 β_e 는 이득 제어 단계들/스테이지들(15,...,151)의 입력에서 계산될 수 있다.

[0216] 지수에 대해 이 비트 수 β_e 를 사용하는 것은 HOA 압축기 이득 제어 처리 유닛들(15, ..., 151)에 의해 야기된 모든 가능한 절대 진폭 변화들이 포착될 수 있도록 보장하여, 압축된 표현 내의 어떤 미리 정의된 진입점들에서 압축해제를 시작하는 것을 가능하게 한다.

[0217] 압축된 HOA 표현의 압축해제를 HOA 압축해제기에서 시작할 때, 어떤 데이터 프레임들에 대한 보조 정보에 할당

되고 수신된 데이터 스트림 $\check{B}$ 중에서 디멀티플렉서(21)로부터 수신되는 총 절대 진폭 변화들을 나타내는 비차분 이득 값들이, 이득 제어 단계들/스테이지들(15,...,151)에서 수행된 처리와 정반대의 방식으로, 올바른 이득 제어를 적용하기 위해 역 이득 제어 단계들 또는 스테이지들(24,..., 241)에서 사용된다.

[0218] *추가 실시예*

[0219] *HOA 압축* 섹션, *공간 HOA 인코딩* 섹션, *HOA 압축해제* 섹션 및 *공간 HOA 디코딩* 섹션에 기술된 바와 같은 특정의 HOA 압축/압축해제 시스템들을 구현할 때, 지수의 코딩을 위한 비트들의 양 β_e는 스케일링 인자 $K_{\mathrm{MAX,DES}}$에 의존하여 수학식 42에 따라 설정되어야만 하며, 스케일링 인자 자체는 압축될 HOA 표현들의 원하는 최대 차수 $N_{\mathrm{MAX,DES}}$ 및 특정 가상 스피커 방향들 $\boldsymbol{\Omega}_{\mathrm{DES},1}^{(N)}, \ldots, \boldsymbol{\Omega}_{\mathrm{DES},O}^{(N)}$, $1 \leq N \leq N_{\mathrm{MAX}}$에 의존한다.

[0220] 예를 들어, $N_{\mathrm{MAX,DES}} = 29$인 것으로 가정하고 Fliege 등의 논문에 따라 가상 스피커 방향들을 선택할 때, 타당한 선택은 $\sqrt{K_{\mathrm{MAX,DES}}} = 1.5$일 것이다. 그 상황에서, 동일한 가상 스피커 방향들 $\boldsymbol{\Omega}_{\mathrm{DES},1}^{(N)}, \ldots, \boldsymbol{\Omega}_{\mathrm{DES},O}^{(N)}$를 사용하여 *입력 HOA 표현의 정규화* 섹션에 따라 정규화되는, $1 \leq N \leq N_{\mathrm{MAX}}$를 갖는 차수 N의 HOA 표현들에 대해 올바른 압축이 보장된다. 그렇지만, (효율로 인해) 가상 스피커 신호들에 의해 PCM 포맷으로 동등하게 표현되는 HOA 표현의 경우에, 그리고 가상 스피커들의 방향들 $\boldsymbol{\Omega}_j^{(N)}, 1 \leq j \leq O$이 시스템 설계 스테이지에서 가정된 가상 스피커 방향들 $\boldsymbol{\Omega}_{\mathrm{DES},1}^{(N)}, \ldots, \boldsymbol{\Omega}_{\mathrm{DES},O}^{(N)}$와 상이하게 선택되는 경우, 이 보장이 주어지지 않을 수 있다.

[0221] 가상 스피커 위치들의 이 상이한 선택으로 인해, 이 가상 스피커 신호들의 진폭들이 구간 $[1,1[$ 내에 있더라도, 이득 제어 이전의 신호들의 진폭들이 값 $\sqrt{K_{\mathrm{MAX,DES}}} \cdot O$를 초과하지 않을 것이라고 더 이상 보장될 수 없다. 그리고 따라서 이 HOA 표현이 MPEG 문서 N14264에 기술된 처리에 따른 압축에 대해 적절한 정규화를 갖는다고 보장될 수 없다.

[0222] 이 상황에서, 각자의 HOA 표현이 MPEG 문서 N14264에 기술된 처리에 따른 압축에 적당하도록 하기 위해, 가상 스피커 위치들을 알고 있는 것에 기초하여, 가상 스피커 신호들의 최대 허용 진폭을 제공하는 시스템을 가지는 것이 유리하다. 도 5에, 이러한 시스템이 예시되어 있다. 이 시스템은 가상 스피커 위치들 $\boldsymbol{\Omega}_j^{(N)}, 1 \leq j \leq O$ - 단, $O = (N+1)^2$이고 $N \in \mathbb{N}_0$임 - 를 입력으로서 취하고, 가상 스피커 신호들의 최대 허용 진폭 γ_{dB}(데시벨로 측정됨)를 출력으로서 제공한다. 단계 또는 스테이지(51)에서, 가상 스피커 위치들에 대한 모드 행렬 $\boldsymbol{\Psi}$가 수학식 3에 따라 계산된다. 다음 단계 또는 스테이지(52)에서, 모드 행렬의 유클리드 노름 $\|\boldsymbol{\Psi}\|_2$가 계산된다. 제3 단계 또는 스테이지(53)에서, 수학식 43과 같이, 진폭 γ가 가상 스피커 위치들의 수의 제곱근과 $K_{\mathrm{MAX,DES}}$의 제곱근의 곱과 모드 행렬의 유클리드 노름 사이의 비율(quotient)과 '1' 중의 최소값으로서 계산된다.

[0223] [수학식 43]

[0224] $$\gamma = \min\left(1, \frac{\sqrt{O} \cdot \sqrt{K_{\mathrm{MAX,DES}}}}{\|\boldsymbol{\Psi}\|_2}\right)$$

[0225] 수학식 44에 의해 데시벨 단위의 값이 획득된다.

[0226] [수학식 44]

[0227] $$\gamma_{dB} = 20\log_{10}(\gamma)$$

[0228] 설명을 위해: 이상의 유도들로부터, HOA 계수 시퀀스들의 크기가 값 $\sqrt{K_{MAX,DES}} \cdot O$ 를 초과하지 않는 경우, 즉 수학식 45인 경우,

[0229] [수학식 45]

[0230] $$\|\boldsymbol{c}(lT_S)\|_\infty \leq \sqrt{K_{MAX,DES}} \cdot O$$

[0231] 이득 제어 처리 유닛들(15, 151) 이전의 모든 신호들이 그에 따라, 적절한 HOA 압축에 대한 요구사항인, 이 값을 초과하지 않는다는 것을 알 수 있다.

[0232] 수학식 9로부터, HOA 계수 시퀀스들의 크기가 수학식 46에 의해 제한된다는 것을 알 수 있다.

[0233] [수학식 46]

[0234] $$\|\boldsymbol{c}(lT_S)\|_\infty \leq \|\boldsymbol{c}(lT_S)\|_2 \leq \|\boldsymbol{\Psi}\|_2 \cdot \|\boldsymbol{w}(lT_S)\|_2$$

[0235] 결과적으로, γ 가 수학식 43에 따라 설정되고 PCM 포맷의 가상 스피커 신호들이 수학식 47을 충족시키는 경우,

[0236] [수학식 47]

[0237] $$\|\boldsymbol{w}(lT_S)\|_\infty \leq \gamma$$

[0238] 수학식 7로부터 당연히 수학식 48로 되고,

[0239] [수학식 48]

[0240] $$\|\boldsymbol{w}(lT_S)\|_2 \leq \gamma \cdot \sqrt{O}$$

[0241] 요구사항(수학식 45)이 충족된다.

[0242] 즉, 수학식 6에서의 '1'의 최대 크기 값이 수학식 47에서의 최대 크기 값 γ 에 의해 대체된다.

[0243] *고차 앰비소닉스의 기초*

[0244] 고차 앰비소닉스(HOA)는 음원들이 없는 것으로 가정되는 콤팩트한 관심 영역 내에서의 음장의 설명에 기초한다. 그 경우에, 시간 t 및 위치 x 에서의 음압 $p(t,x)$ 의 시공간 거동이 동차 파동 방정식(homogeneous wave equation)에 의해 물리적으로 완전히 결정된다. 이하에서, 도 6에 도시된 바와 같은 구면 좌표계가 가정된다. 사용되는 좌표계에서, x 축은 전방 위치를 가리키고, y 축은 좌측을 가리키며, z 축은 상방을 가리킨다. 공간에서의 위치 $x = (r,\theta,\phi)^T$ 는 반경 $r > 0$ (즉, 좌표 원점까지의 거리)에 의해 표현되고, 경사각 $\theta \in [0,\pi]$ 는 극축(polar axis) z 로부터 측정되며, 방위각 $\phi \in [0,2\pi[$ 는 $x-y$ 평면에서 x 축으로부터 시계반대방향으로 측정된다. 게다가, $(\cdot)^T$ 는 전치를 나타낸다.

[0245] 이어서, "Fourier Acoustics" 교재로부터 $\mathcal{F}_t(\cdot)$ 에 의해 표기된 시간에 대한 음압의 푸리에 변환, 즉, 수학식 49가

[0246] [수학식 49]

[0247] $$P(\omega,x) = \mathcal{F}_t\big(p(t,x)\big) = \int_{-\infty}^{\infty} p(t,x)e^{-i\omega t}dt$$

[0248] - 여기서 ω는 각주파수를 나타내고 i는 허수 단위를 나타냄 - 수학식 50에 따라 구면 조화함수의 급수로 전개될 수 있다는 것을 알 수 있고,

[0249] [수학식 50]

[0250] $$P(\omega = kc_s, r, \theta, \phi) = \sum_{n=0}^{N} \sum_{m=-n}^{n} A_n^m(k) j_n(kr) S_n^m(\theta, \phi)$$

[0251] 여기서 c_s는 음속을 나타내고 k 는 $k = \frac{\omega}{c_s}$ 에 의해 각주파수 ω에 관련되는 각파수(angular wave number)를 나타낸다. 게다가, $j_n(\cdot)$는 제1종의 구면 베셀 함수(spherical Bessel function)들을 나타내고, $S_n^m(\theta, \phi)$는, *실수값 구면 조화함수의 정의* 섹션에서 정의되는, 계수(order) n 및 차수(degree) m 의 실수값 구면 조화함수를 나타낸다. 전개 계수들 $A_n^m(k)$는 각파수 k에만 의존한다. 유의할 점은, 음압이 공간적으로 대역 제한되어 있는 것으로 암시적으로 가정되었다는 것이다. 이와 같이, 급수가 차수 인덱스 n과 관련하여, HOA 표현의 차수라고 불리우는, 상한 N에서 절단된다.

[0252] 음장이 각도 튜플(angle tuple) (θ, ϕ)에 의해 명시되는 모든 가능한 방향들부터 도착하는 상이한 각주파수들 ω의 무한 수의 조화 평면파들의 중첩에 의해 표현되는 경우, 각자의 평면파 분해 복소 진폭 함수 $C(\omega, \theta, \phi)$가 수학식 51의 구면 조화함수 전개에 의해 표현될 수 있고

[0253] [수학식 51]

[0254] $$C(\omega = kc_s, \theta, \phi) = \sum_{n=0}^{N} \sum_{m=-n}^{n} C_n^m(k) S_n^m(\theta, \phi)$$

[0255] 여기서 전개 계수들 $C_n^m(k)$는 수학식 52에 의해 전개 계수들 $A_n^m(k)$에 관련되어 있다는 것을 알 수 있다(B. Rafaely, "Plane-wave decomposition of the sound field on a sphere by spherical convolution", J. Acoust. Soc. Am., vol.4(116), pages 2149-2157, October 2004를 참조)

[0256] [수학식 52]

[0257] $$A_n^m(k) = i^n C_n^m(k)$$

[0258] 개개의 계수들 $C_n^m(k = \omega/c_s)$가 각주파수 ω의 함수들인 것으로 가정하면, 역푸리에 변환($\mathcal{F}^{-1}(\cdot)$에 의해 표기됨)의 적용은 각각의 계수 n 및 차수 m에 대한 수학식 53의 시간 영역 함수들을 제공한다:

[0259] [수학식 53]

[0260] $$c_n^m(t) = \mathcal{F}_t^{-1}\left(C_n^m(\omega/c_s)\right) = \frac{1}{2\pi} \int_{-\infty}^{\infty} C_n^m\left(\frac{\omega}{c_s}\right) e^{i\omega t} d\omega$$

[0261] 이 시간 영역 함수들은 여기서, 수학식 54에 의해 단일의 벡터 $\boldsymbol{c}(t)$에 모여 있을 수 있는, 연속 시간 HOA 계수 시퀀스들이라고 지칭된다.

[0262] [수학식 54]

[0263] $$\boldsymbol{c}(t) =$$

[0264] $$\left[c_0^0(t) \;\; c_1^{-1}(t) \;\; c_1^0(t) \;\; c_1^1(t) \;\; c_2^{-2}(t) \;\; c_2^{-1}(t) \;\; c_2^0(t) \;\; c_2^1(t) \;\; c_2^2(t) \;\; \ldots \;\; c_N^{N-1}(t) \;\; c_N^N(t)\right]^T$$

[0265] 벡터 $\boldsymbol{c}(t)$ 내에서의 HOA 계수 시퀀스 $c_n^m(t)$의 위치 인덱스는 $n(n+1)+1+m$에 의해 주어진다. 벡터 $\boldsymbol{c}(t)$ 내의 요소들의 총수는 $O=(N+1)^2$에 의해 주어진다.

[0266] 최종적인 앰비소닉스 포맷은 샘플링 주파수 f_s를 사용한 $\boldsymbol{c}(t)$의 샘플링된 버전을 수학식 55로서 제공하고,

[0267] [수학식 55]

[0268] $$\{\boldsymbol{c}(lT_s)\}_{l\in\mathbb{N}}=\{\boldsymbol{c}(T_s),\boldsymbol{c}(2T_s),\boldsymbol{c}(3T_s),\boldsymbol{c}(4T_s),\dots\}$$

[0269] 여기서 $T_s=1/f_s$는 샘플링 주기를 나타낸다. $\boldsymbol{c}(lT_s)$의 요소들은, 항상 실수값이라는 것을 알 수 있는, 이산 시간 HOA 계수 시퀀스들이라고 지칭된다. 이 특성은 연속 시간 버전들 $c_n^m(t)$에 대해서도 성립한다.

[0270] *실수값 구면 조화함수의 정의*

[0271] 실수값 구면 조화함수 $S_n^m(\theta,\phi)$(J. Daniel, "Representation de champs acoustiques, application a la transmission et a la reproduction de scenes sonores complexes dans un contexte multimedia", PhD thesis, Universite Paris, 6, 2001, chapter 3.1에 따라 SN3D 정규화를 가정함)는 수학식 56에 의해 주어지고

[0272] [수학식 56]

[0273] $$S_n^m(\theta,\phi)=\sqrt{(2n+1)\frac{(n-|m|)!}{(n+|m|)!}}\;P_{n,|m|}(cos\theta)\;\mathrm{trg}_m(\phi)$$

[0274] 여기서

[0275] [수학식 57]

[0276] $$\mathrm{trg}_m(\phi)=\begin{cases}\sqrt{2}cos(m\phi) & m>0\\ 1 & m=0\\ -\sqrt{2}sin(m\phi) & m<0\end{cases}$$

[0277] 연관된 르장드르 함수(Legendre function)들 $P_{n,m}(x)$는 수학식 58로서 정의되고,

[0278] [수학식 58]

[0279] $$P_{n,m}(x)=(1-x^2)^{m/2}\frac{\mathrm{d}^m}{\mathrm{d}x^m}P_n(x),m\geq 0$$

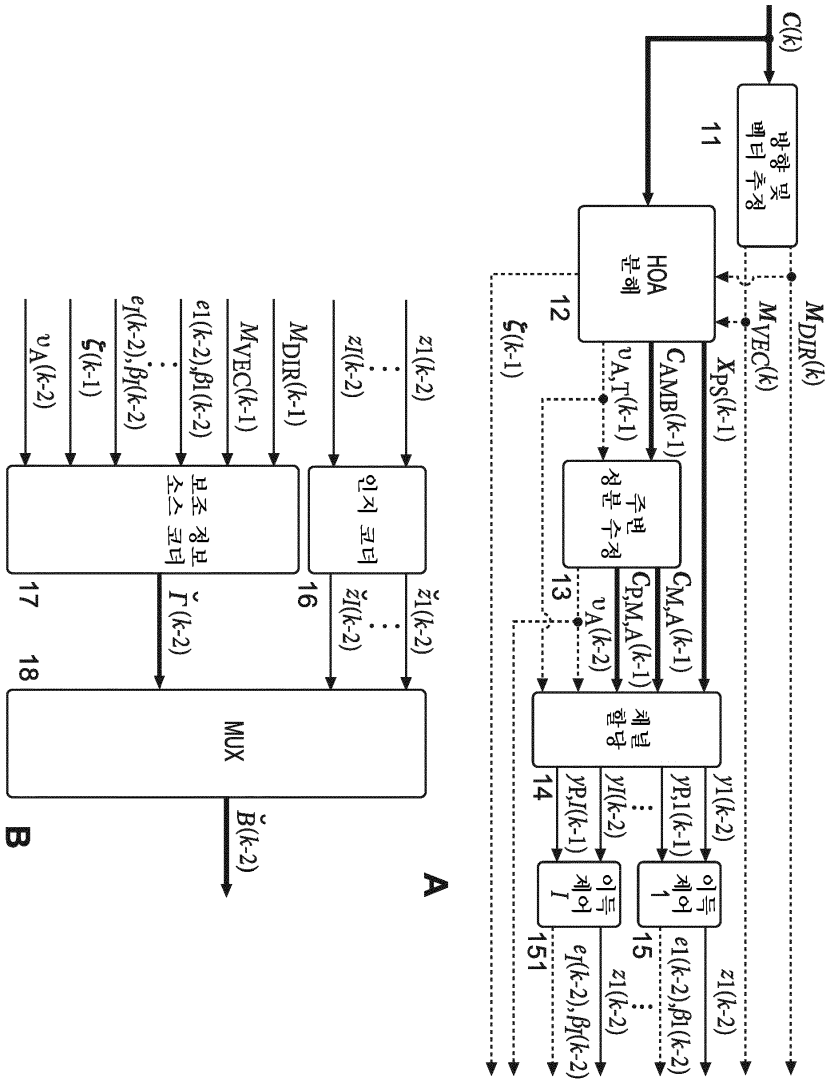
[0280] 수학식 58은 르장드르 다항식(Legendre polynomial) $P_n(x)$를 갖지만, E.G. Williams, "Fourier Acoustics", vol.93 of Applied Mathematical Sciences, Academic Press, 1999에서와 달리, 콘던-쇼틀리 위상(Condon-Shortley phase) 항 $(-1)^m$을 갖지 않는다.

[0281] 본 발명의 처리는 단일의 프로세서 또는 전자 회로에 의해, 또는 병렬로 동작하는 그리고/또는 본 발명의 처리의 상이한 부분들에서 동작하는 몇 개의 프로세서들 또는 전자 회로들에 의해 수행될 수 있다.

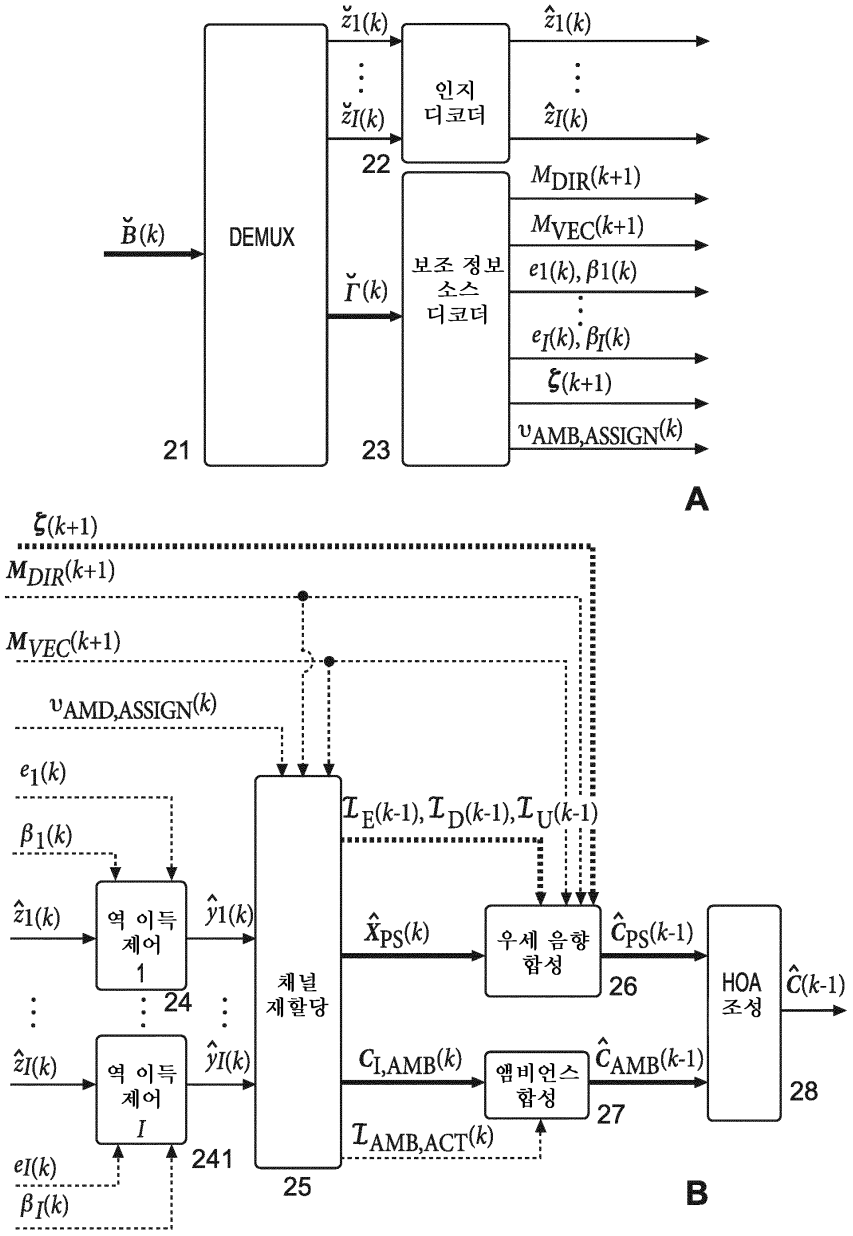
[0282] 프로세서 또는 프로세서들을 작동시키는 명령어들은 하나 이상의 메모리들에 저장될 수 있다.

도면

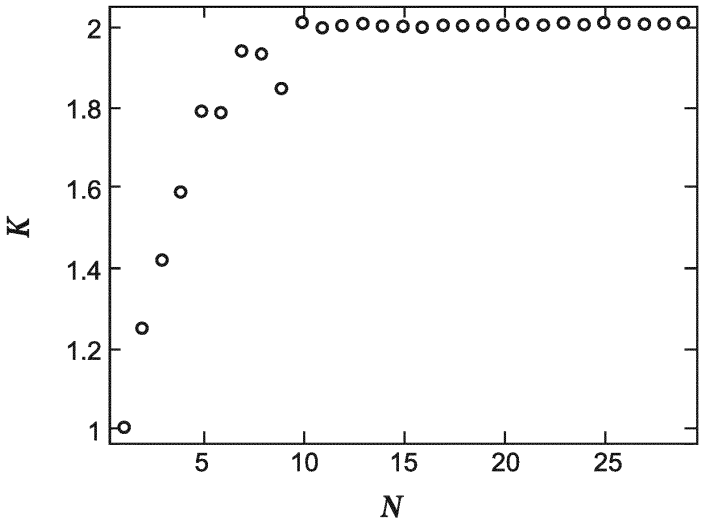
도면1


C(k)
방향 및 벡터 추정
11
HOA 분해
12
ξ(k-1)
M_DIR(k)
M_VEC(k)
X_PS(k-1)
C_AMB(k-1)
υ_A,T(k-1)
주변 성분 추정
13
C_P,M,A(k-1)
C_M,A(k-1)
υ_A(k-2)
채널 할당
14
y_P,1(k-1)
y_P,I(k-1)
y_1(k-2)
y_I(k-2)
이득 제어 1
이득 제어 I
151
15
e_1(k-2),β_1(k-2)
e_I(k-2),β_I(k-2)
z_1(k-2)
z_I(k-2)
인지 코더
16
보조 정보 코더
17
18
MUX
B
A
z_1(k-2)
z_I(k-2)
M_DIR(k-1)
M_VEC(k-1)
e_1(k-2),β_1(k-2)
e_I(k-2),β_I(k-2)
ξ(k-1)
υ_A(k-2)
Γ̌(k-2)
ž_1(k-2)
ž_I(k-2)
B̌(k-2)

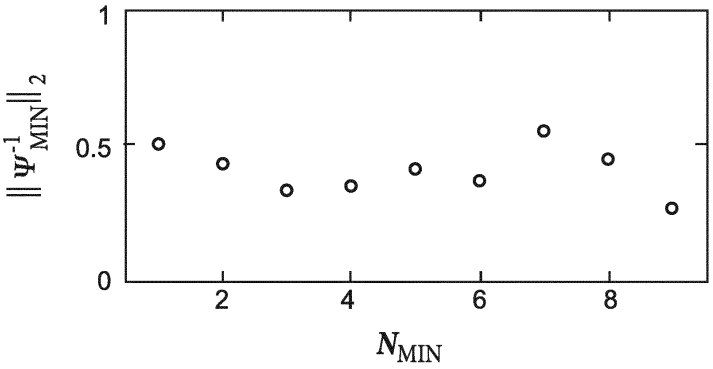
도면2


DEMUX
21
인지 디코더
22
보조 정보 소스 디코더
23
A
역 이득 제어 1
24
역 이득 제어 I
241
채널 재할당
25
우세 음향 합성
26
앰비언스 합성
27
HOA 조성
28
B

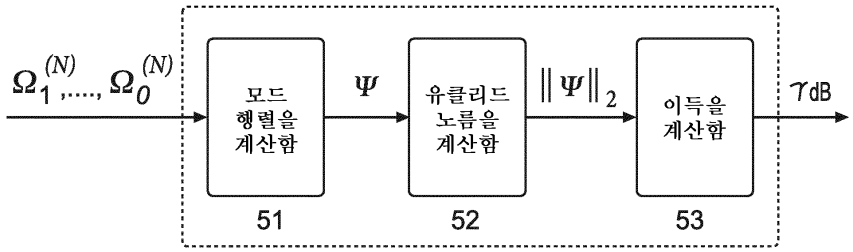
도면3


2
1.8
1.6
1.4
1.2
1
K
5
10
15
20
25
N

도면4


1
0.5
0
$\|\Psi^{-1}_{MIN}\|_2$
2
4
6
8
N_{MIN}

도면5


$\Omega_1^{(N)}, \ldots, \Omega_0^{(N)}$
모드 행렬을 계산함
51
Ψ
유클리드 노름을 계산함
52
$\|\Psi\|_2$
이득을 계산함
53
τ_{dB}

도면6

