



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2012-0018778
(43) 공개일자 2012년03월05일

(51) 국제특허분류(Int. Cl.)
G10L 19/00 (2006.01) G10L 19/14 (2006.01)
(21) 출원번호 10-2011-7028264
(22) 출원일자(국제) 2010년04월28일
심사청구일자 2011년11월25일
(85) 번역문제출일자 2011년11월25일
(86) 국제출원번호 PCT/EP2010/055717
(87) 국제공개번호 WO 2010/125104
국제공개일자 2010년11월04일
(30) 우선권주장
61/173,456 2009년04월28일 미국(US)

(71) 출원인
프라운호퍼 게젤샤프트 쭈르 피르데룽 데어 안겐반텐 포르슘 에. 베.
독일 80686 뮌헨 한자슈트라쎄 27 체
프리드리히-알렉산더-우니베르지테트 에를랑겐-뉘른베르크
독일, 테-91054 에를랑겐, 쉬로스플라츠 4
돌비 인터네셔널 에이비
네덜란드 1101 씨엔 암스트레담 주이두스트 헤리커베르그백 1-35 3이 아폴로 빌딩
(72) 발명자
헤어레 위르겐
독일 부켄호프 91054 할러스트라쎄 24
호엘처 안드레아스
독일 에를랑겐 91054 오베르 칼스트라쎄 23
(뒷면에 계속)
(74) 대리인
배성렬, 김원식

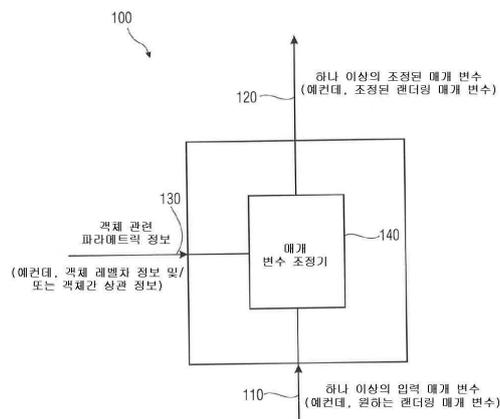
전체 청구항 수 : 총 34 항

(54) 발명의 명칭 **다운믹스 신호 표현에 기초하여 업믹스 신호 표현을 제공하기 위해 하나 이상의 조정된 매개 변수를 제공하는 장치, 오디오 신호 디코더, 오디오 신호 트랜스코더, 오디오 신호 인코더, 오디오 비트스트림, 객체 관련 파라메트릭 정보를 이용하는 방법 및 컴퓨터 프로그램**

(57) 요약

다운믹스 신호 표현 및 객체 관련 파라메트릭 정보에 기초하여 업믹스 신호 표현을 제공하기 위해 하나 이상의 조정된 매개 변수를 제공하는 장치는 매개 변수 조정기를 포함한다. 매개 변수 조정기는 하나 이상의 입력 매개 변수를 수신하여, 이에 기초하여 하나 이상의 조정된 매개 변수를 제공하도록 구성된다. 매개 변수 조정기는 하나 이상의 입력 매개 변수 및 객체 관련 파라메트릭 정보에 따라 하나 이상의 조정된 매개 변수를 제공하도록 구성됨으로써, 최적화되지 않은 매개 변수의 사용으로 유발되는 업믹스 신호 표현의 왜곡은 미리 정해진 편차 이상만큼 적어도 최적의 매개 변수에서 벗어난 입력 매개 변수에 대해 감소된다.

대표도 - 도1



(72) 발명자

데렌티브 레오니드

독일 에를랑겐 91056 아파트 11 압 유로파카날 36

카스트너 토르스텐

독일 스톡카임 라이트쉬 96342 도르프스트라세 17

팔크 코르넬리아

독일 뉘른베르크 90491 카슬러 스트라세 12

푸른하겐 헤이코

스웨덴 슌드비베르그 17265 규테리바켄 17

앵데가르트 조나스

스웨덴 스톡홀름 11543 벤스트림스베겐 6

리더르부쉬 팔코

독일 뉘른베르크 90419 아담 크래프트 스트라세 57

특허청구의 범위

청구항 1

다운믹스 신호 표현(212;312;524) 및 객체 관련 파라메트릭 정보(214;314;520)에 기초하여 업믹스 신호 표현(\hat{y}^1 내지 \hat{y}^N ; 316; 522,524; 522,574)을 제공하기 위해 하나 이상의 조정된 매개 변수(120; 222; 324; r_m' , $r_{lim,m}$)를 제공하는 장치(100; 240; 320; 550)에 있어서,

하나 이상의 입력 매개 변수(110; 242; 322; 552,554; r_i)를 수신하여, 이에 기초하여 하나 이상의 조정된 매개 변수(120;222;324;542)를 제공하도록 구성되는 매개 변수 조정기(140;240)를 포함하는데,

상기 매개 변수 조정기는 상기 하나 이상의 입력 매개 변수 및 상기 객체 관련 파라메트릭 정보(130; 214a,214b,214c;314;520)에 따라 상기 하나 이상의 조정된 매개 변수를 제공하도록 구성됨으로써, 최적화되지 않은 매개 변수의 사용으로 유발되는 상기 업믹스 신호 표현의 왜곡은 미리 정해진 편차 이상만큼 적어도 최적의 매개 변수에서 벗어난 입력 매개 변수에 대해 감소되는 것을 특징으로 하는 하나 이상의 조정된 매개 변수를 제공하는 장치.

청구항 2

청구항 1에 있어서,

상기 장치는, 상기 입력 매개 변수(110; 242; 322; 552,554; r_i)로서, 상기 업믹스 신호 표현(\hat{y}^1 내지 \hat{y}^N ; 316; 522,524; 522,574)에 의해 나타내는 하나 이상의 오디오 채널에서 다수의 오디오 객체 신호(x_1 내지 x_N)의 원하는 강도 스케일링을 나타내는 원하는 렌더링 매개 변수(r_i)를 수신하도록 구성되고;

상기 매개 변수 조정기는 상기 하나 이상의 원하는 렌더링 매개 변수(r_i)에 따라 하나 이상의 실제 렌더링 매개 변수(r_m' , $r_{lim,m}$)를 제공하도록 구성되는 것을 특징으로 하는 하나 이상의 조정된 매개 변수를 제공하는 장치.

청구항 3

청구항 2에 있어서,

상기 매개 변수 조정기는 상기 객체 관련 파라메트릭 정보(130; 214a,214b,214c;314;520) 및, 상기 다운믹스 신호 표현에 대한 상기 오디오 객체 신호(x_1 내지 x_N)의 기여를 나타내는 다운믹스 정보(214b; d_i)에 따라 하나 이

상의 렌더링 매개 변수 제한 값 (\hat{r}_m^2)을 획득하도록 구성됨으로써, 왜곡 매트릭(dm1(m), dm2(m), dm5(m), dm6(m), DM1, DM2, DM3, DM4, DM5, DM6)은 상기 렌더링 매개 변수 제한 값으로 정의되는 제한에 따르는 렌더링 매개 변수 값에 대한 미리 정해진 범위 내에 있고,

상기 매개 변수 조정기는 상기 원하는 렌더링 매개 변수(r_i) 및 상기 하나 이상의 렌더링 매개 변수 제한 값에 따라 상기 실제 렌더링 매개 변수(r_m' , $r_{lim,m}$)를 획득하도록 구성됨으로써, 상기 실제 렌더링 매개 변수는 상기 렌더링 매개 변수 제한 값으로 정의되는 제한에 따르는 것을 특징으로 하는 하나 이상의 조정된 매개 변수를 제공하는 장치.

청구항 4

청구항 2 또는 3에 있어서,

상기 매개 변수 조정기는 상기 하나 이상의 렌더링 매개 변수 제한 값 (\hat{r}_m^2)을 획득하도록 구성됨으로써, 상기 하나 이상의 렌더링 매개 변수 제한 값에 따르는 하나 이상의 렌더링 매개 변수(r_m' , $r_{lim,m}$)를 이용하여 렌더링 되는 다수의 객체 신호의 렌더링된 중첩에서의 객체 신호(x_1 내지 x_N)의 상대 기여는 단지 미리 정해진 차이만큼

다운믹스 신호(212:312:524)에서의 상기 객체 신호(x_1 내지 x_N)의 상대 기여와 상이한 것을 특징으로 하는 하나 이상의 조정된 매개 변수를 제공하는 장치.

청구항 5

청구항 4에 있어서,

상기 매개 변수 조정기는 상기 하나 이상의 렌더링 매개 변수 값(r_m)을 결정하도록 구성됨으로써, 식

$$dm_1(m) = \frac{r_m^2 \cdot \sum_{i=1}^N d_i^2 \cdot X_i}{d_m^2 \cdot \sum_{i=1}^N r_i^2 \cdot X_i} \leq T$$

이 객체 인덱스(m)에 의해 명시되는 하나 이상의 오디오 객체에 대해 충족되고,

r_m 은 상기 업믹스 신호의 주어진 채널 (\hat{y}_1 to \hat{y}_N)에 대해 객체 인덱스(m)를 가진 오디오 객체의 객체 신호의 기여를 나타내는 렌더링 매개 변수를 명시하고,

d_m 은 상기 다운믹스 신호에서 인덱스(m)를 가진 상기 객체의 객체 신호(x_1 내지 x_N)의 기여를 나타내는 다운믹스 매개 변수를 명시하며,

X_i 는 객체 인덱스(m)를 가진 상기 오디오 객체의 에너지 측정을 명시하며, 상기 에너지 측정은 상기 객체 관련 파라메트릭 정보에 의해 결정되는 것을 특징으로 하는 하나 이상의 조정된 매개 변수를 제공하는 장치.

청구항 6

청구항 2 또는 3에 있어서,

상기 매개 변수 조정기는 상기 하나 이상의 렌더링 매개 변수 제한 값 (\hat{r}_m^2)을 획득하도록 구성됨으로써, 상기 다운믹스 신호 표현에 의해 나타내는 다운믹스 신호와, 상기 하나 이상의 렌더링 매개 변수 제한 값 (\hat{r}_m^2)에 따르는 하나 이상의 렌더링 매개 변수(r_m)를 이용하여 렌더링되는 렌더링된 신호 사이의 코히어런스를 나타내는 왜곡 측정(DM3)이 미리 정해진 범위 내에 있는 것을 특징으로 하는 하나 이상의 조정된 매개 변수를 제공하는 장치.

청구항 7

청구항 6에 있어서,

상기 매개 변수 조정기는 상기 하나 이상의 렌더링 매개 변수 제한 값 (\hat{r}_m^2)을 획득하도록 구성됨으로써, 상기

$$DM_3 = 1 - \min\left(\frac{|c_{12}|}{\sqrt{c_{11} \cdot c_{22}}}, 1\right)$$

왜곡 측정 은 미리 정해진 값을 취하고,

C는

$$C = M \cdot E \cdot M^* = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix} \text{로 정의되며;}$$

$$M = \begin{pmatrix} r_1 & r_2 & \cdots & r_N \\ d_1 & d_2 & \cdots & d_N \end{pmatrix}$$

은 렌더링 매개 변수(r_1 내지 r_n)의 제 1 행 및, 상기 다운믹스 신호 표현에 대한 상기 오디오 객체 신호의 기여를 나타내는 다운믹스 매개 변수(d_1 내지 d_n)의 제 2 행을 포함하는 매트릭스이며;

E 는 상기 객체 관련 파라메트릭 정보의 매개 변수(OLD, IOC)를 이용하여 획득되는 객체 공분산 매트릭스이며, "*" 는 복소 공액 연산자인 것을 것을 특징으로 하는 하나 이상의 조정된 매개 변수를 제공하는 장치.

청구항 8

청구항 2에 있어서,

상기 매개 변수 조정기는 원하는 렌더링 매개 변수(r_m)의 제공과 최적의 렌더링 매개 변수($r_{opt,m}$)의 제공 사이의 선형 조합을 계산하여, 실제 렌더링 매개 변수($r_{lim,m}$)를 획득하도록 구성되고,

상기 매개 변수 조정기는 미리 정해진 임계 매개 변수 T 및 왜곡 매트릭(dm1, dm2, dm3, dm4, dm5, dm6)에 따라 선형 조합에 대한 원하는 렌더링 매개 변수(r_m) 및 최적의 렌더링 매개 변수($r_{opt,m}$)의 기여를 결정하도록 구성되며, 상기 왜곡 매트릭은 상기 다운믹스 신호 표현에 기초하여 상기 업믹스 신호 표현을 획득하기 위해 상기 최적의 렌더링 매개 변수($r_{opt,m}$) 보다는 상기 하나 이상의 원하는 렌더링 매개 변수(r_m)를 이용함으로써 유발되는 왜곡을 나타내는 것을 특징으로 하는 하나 이상의 조정된 매개 변수를 제공하는 장치.

청구항 9

청구항 8에 있어서,

상기 매개 변수 조정기는 상기 업믹스 신호의 주어진 채널에 대해 객체 인덱스(m)를 가진 객체의 객체 신호의

$$r_{lim,m}^2 = \frac{T-1}{dm_x(m)} (r_m^2 - r_{opt,m}^2) + r_{opt,m}^2$$

기여를 나타내는 실제 렌더링 매개 변수($r_{lim,m}$)를 획득하기 위해 식

을

평가하도록 구성되고,

T는 미리 정해진 왜곡 임계 매개 변수를 명시하며,

$dm_x(m)$ 은 상기 업믹스 신호의 주어진 채널에 대해 객체 인덱스(m)를 가진 오디오 객체의 객체 신호의 원하는 기여를 나타내는 원하는 렌더링 매개 변수(r_m)와 관련된 왜곡 매트릭을 명시하며;

$r_{opt,m}$ 은 상기 업믹스 신호의 상기 주어진 채널에 대해 객체 인덱스(m)를 가진 오디오 객체의 객체 신호의 최적의 기여를 나타내는 최적의 렌더링 매개 변수를 명시하는 것을 특징으로 하는 하나 이상의 조정된 매개 변수를 제공하는 장치.

청구항 10

청구항 8 또는 9에 있어서,

상기 매개 변수 조정기는 상기 왜곡 매트릭을 획득하도록 구성됨으로써, 상기 왜곡 매트릭은 상기 원하는 렌더링 매개 변수에 따라 렌더링되는 다수의 객체 신호의 렌더링된 중첩에서의 주어진 객체 신호의 상대 기여와, 상기 주어진 객체 신호를 포함하는 다운믹스 신호에서의 상기 주어진 객체 신호의 상대 기여 사이의 관계에 의존하는 것을 특징으로 하는 하나 이상의 조정된 매개 변수를 제공하는 장치.

청구항 11

청구항 8, 9 또는 10 중 어느 한 항에 있어서,

상기 매개 변수 조정기는 상기 왜곡 매트릭(dm_1)을 획득하도록 구성됨으로써, 상기 왜곡 매트릭은 상기 원하는 렌더링 매개 변수(r_m)에 따라 렌더링되는 다수의 객체 신호의 렌더링된 중첩에서의 주어진 객체 신호(x_1 내지 x_N)의 상대 기여와, 상기 주어진 객체 신호(x_1 내지 x_N)를 포함하는 다운믹스 신호에서의 상기 주어진 객체 신호(x_1 내지 x_N)의 상대 기여 사이의 비율에 의존하는 것을 특징으로 하는 하나 이상의 조정된 매개 변수를 제공하

는 장치.

청구항 12

청구항 8 내지 11 중 어느 한 항에 있어서,

$$dm_x(m) = dm_1(m) = \frac{r_m^2 \sum_{i=1}^{N_{ob}} d_i^2 X_i}{d_m^2 \sum_{i=1}^{N_{ob}} r_i^2 X_i}$$

상기 매개 변수 조정기는 에 따라 상기 왜곡 메트릭($dm_x(m)$)을 계산하도록 구성되고,

r_m 및 r_i 은 제각기 객체 인덱스(m) 및 (i)를 가진 오디오 객체와 관련된 원하는 렌더링 매개 변수를 명시하며;

d_m 및 d_i 는 상기 다운믹스 신호 표현의 다운믹스 신호에 대해 제각기 인덱스(m) 및 (i)를 가진 오디오 객체의 객체 신호의 기여를 나타내는 다운믹스 매개 변수를 명시하며;

N_{ob} 는 고려 하에 오디오 객체의 수를 명시하며;

X_i 는 객체 인덱스(i)를 가진 상기 오디오 객체의 상기 객체 신호와 관련된 에너지 측정을 명시하는 것을 특징으로 하는 하나 이상의 조정된 매개 변수를 제공하는 장치.

청구항 13

청구항 8, 9 또는 10 중 어느 한 항에 있어서,

상기 매개 변수 조정기는 상기 왜곡 메트릭(dm_2)을 획득하도록 구성됨으로써, 상기 왜곡 메트릭은 상기 원하는 렌더링 매개 변수(r_m)에 따라 렌더링되는 다수의 객체 신호의 렌더링된 중첩에서의 주어진 객체 신호(x_1 내지 x_N)의 상대 기여와, 상기 주어진 객체 신호(x_1 내지 x_N)를 포함하는 다운믹스 신호에서의 상기 주어진 객체 신호(x_1 내지 x_N)의 상대 기여의 차에 의존하는 것을 특징으로 하는 하나 이상의 조정된 매개 변수를 제공하는 장치.

청구항 14

청구항 8 내지 13 중 어느 한 항에 있어서,

상기 매개 변수 조정기는 상기 왜곡 메트릭(dm_2)을 계산하도록 구성됨으로써, 상기 왜곡 메트릭은 마스크 대 신호비(msr)에 의존하고, 상기 왜곡 메트릭(dm_2)은 상기 마스크 대 신호비가 증가할 경우에 감소하며, 이는 보다 적은 왜곡을 나타내는 것을 특징으로 하는 하나 이상의 조정된 매개 변수를 제공하는 장치.

청구항 15

청구항 8 내지 10 또는 11 또는 12 중 어느 한 항에 있어서,

상기 매개 변수 조정기는 다음 식에 따라 상기 왜곡 메트릭을 계산하도록 구성되며:

$$dm_x(m) = dm_2(m) = \frac{\left(r_m^2 \cdot \sum_{i=1}^N d_i^2 \cdot X_i - d_m^2 \cdot \sum_{i=1}^N r_i^2 \cdot X_i \right) \cdot X_m}{msr \cdot \left(\sum_{i=1}^N r_i^2 \cdot X_i \right) \cdot \left(\sum_{i=1}^N d_i^2 \cdot X_i \right)}$$

또는

$$dm_x = dm_2'(m) = \frac{P_{Noise}}{Mask} = \frac{E\{|y_{m,ideal} - \hat{y}_{m,actual}|^2\}}{msr \cdot P_{total}} =$$

$$\frac{\left| r_m^2 \cdot \sum_{i=1}^N d_i^2 \cdot X_i + d_m^2 \cdot \sum_{i=1}^N r_i^2 \cdot X_i - 2 \cdot d_m r_m \cdot \sqrt{\left(\sum_{i=1}^N r_i^2 \cdot X_i \right) \cdot \left(\sum_{i=1}^N d_i^2 \cdot X_i \right)} \right| \cdot X_m}{msr \cdot \left(\sum_{i=1}^N r_i^2 \cdot X_i \right) \cdot \left(\sum_{i=1}^N d_i^2 \cdot X_i \right)}$$

r_m 및 r_i 는 제각기 객체 인덱스(m) 및 (i)를 가진 오디오 객체와 관련된 원하는 렌더링 매개 변수를 명시하며;

d_m 및 d_i 는 상기 다운믹스 신호 표현의 다운믹스 신호에 대해 제각기 인덱스(m) 및 (i)를 가진 오디오 객체의 객체 신호의 기여를 나타내는 다운믹스 매개 변수를 명시하며;

N은 고려 하에 오디오 객체의 수를 명시하며;

X_i 및 X_m 은 제각기 객체 인덱스(i) 및 (m)를 가진 상기 오디오 객체의 상기 객체 신호와 관련된 에너지 측정을 명시하며;

msr은 마스크 대 신호비를 정의하는 것을 특징으로 하는 하나 이상의 조정된 매개 변수를 제공하는 장치.

청구항 16

청구항 1 내지 15 중 어느 한 항에 있어서,

상기 매개 변수 조정기는 지각 저하의 계산 측정에 따라 상기 하나 이상의 조정된 매개 변수를 제공하도록 구성됨으로써, 최적이지 아닌 매개 변수의 사용으로 유발되고, 상기 지각 저하의 계산 측정에 의해 나타내는 상기 업믹스 신호 표현의 지각적 평가 왜곡이 제한되는 것을 특징으로 하는 하나 이상의 조정된 매개 변수를 제공하는 장치.

청구항 17

청구항 1 내지 16 중 어느 한 항에 있어서,

상기 매개 변수 조정기는 상기 다운믹스 신호 표현에 의해 나타내는 다운믹스 신호에 대한 기초를 형성하는 하나 이상의 원래의 객체 신호의 개개의 속성을 나타내는 개개의 객체 속성 정보를 수신하도록 구성되고;

상기 매개 변수 조정기는 상기 개개의 객체 속성 정보를 고려하여, 상기 조정된 매개 변수를 제공하도록 구성됨으로써, 이상적으로 렌더링된 업믹스 신호 표현에 대한 상기 업믹스 신호 표현의 왜곡이 미리 정해진 편차 이상만큼 적어도 최적의 매개 변수에서 벗어나는 입력 매개 변수에 대해 감소되는 것을 특징으로 하는 하나 이상의 조정된 매개 변수를 제공하는 장치.

청구항 18

청구항 17에 있어서,

상기 매개 변수 조정기는 상기 하나 이상의 조정된 매개 변수를 제공하기 위해, 개개의 객체 속성 정보로서, 객체 신호 음조 정보를 수신하여 고려하도록 구성되는 것을 특징으로 하는 하나 이상의 조정된 매개 변수를 제공하는 장치.

청구항 19

청구항 18에 있어서,

상기 매개 변수 조정기는 상기 수신된 객체 신호 음조 정보 및 상기 수신된 객체 파워 정보(OLD,P)에 따라 이상

적으로 렌더링된 업믹스 신호의 음조(N)를 추정하도록 구성되고;

상기 매개 변수 조정기는 상기 하나 이상의 조정된 매개 변수를 제공하여, 상기 추정된 음조와 상기 하나 이상의 입력 매개 변수를 이용하여 획득된 업믹스 신호의 음조의 차와 비교했을 때 상기 추정된 음조와 상기 하나 이상의 조정된 매개 변수를 이용하여 획득된 업믹스 신호의 음조의 차를 감소시키거나, 상기 추정된 음조와, 미리 정해진 범위 내에 상기 하나 이상의 조정된 매개 변수를 이용하여 획득된 업믹스 신호의 음조의 차를 유지하도록 구성되는 것을 특징으로 하는 하나 이상의 조정된 매개 변수를 제공하는 장치.

청구항 20

청구항 1 내지 19 중 어느 한 항에 있어서,

상기 매개 변수 조정기는 상기 입력 매개 변수의 시간 및 주파수 변형 조정을 수행하도록 구성되는 것을 특징으로 하는 하나 이상의 조정된 매개 변수를 제공하는 장치.

청구항 21

청구항 1 내지 20 중 어느 한 항에 있어서,

상기 매개 변수 조정기는 또한 상기 하나 이상의 조정된 매개 변수를 제공하기 위한 상기 다운믹스 신호 표현을 고려하도록 구성되는 것을 특징으로 하는 하나 이상의 조정된 매개 변수를 제공하는 장치.

청구항 22

청구항 1 내지 21 중 어느 한 항에 있어서,

상기 매개 변수 조정기는 다수의 타입의 아티팩트를 나타내는 왜곡 측정의 가중 조합인 전체 왜곡 측정을 획득하도록 구성되고;

상기 매개 변수 조정기는 상기 전체 왜곡 측정을 획득하도록 구성됨으로써, 상기 전체 왜곡 측정이 상기 다운믹스 신호 표현에 기초하여 상기 업믹스 신호 표현을 획득하기 위해 최적의 렌더링 매개 변수보다는 하나 이상의 상기 입력 렌더링 매개 변수를 이용함으로써 유발되는 왜곡의 측정인 것을 특징으로 하는 하나 이상의 조정된 매개 변수를 제공하는 장치.

청구항 23

청구항 22에 있어서,

상기 매개 변수 조정기는 상기 전체 왜곡 측정을 획득하기 위해 다음의 왜곡 측정:

- 오디오 객체의 음색의 기생 변경(parasitic change)을 나타내는 측정;
- 오디오 객체와 관련된 객체 신호의 기생 변조를 나타내는 측정;
- 기생 음악 톤의 존재를 나타내는 측정;
- 기생 변조 노이즈의 존재를 나타내는 측정 중 적어도 둘을 조합하도록 구성되는 것을 특징으로 하는 하나 이상의 조정된 매개 변수를 제공하는 장치.

청구항 24

업믹스 신호 표현으로서, 다운믹스 신호 표현(212;312), 객체 관련 파라메트릭 정보(214;314) 및 원하는 렌더링 정보(242;322)에 기초하여 다수의 업믹스 오디오 채널(\hat{y}_1 내지 \hat{y}_N ; 316)을 제공하는 오디오 신호 디코더(220,240; 300)에 있어서,

상기 객체 관련 파라메트릭 정보(214;314) 및, 상기 객체 관련 파라메트릭 정보에 의해 나타난 오디오 객체의 다수의 객체 신호를 상기 업믹스 오디오 채널에 할당하는 것을 나타내는 실제 렌더링 정보(222; 324)에 따라 상기 다운믹스 신호 표현(212;312)에 기초하여 상기 업믹스 오디오 채널(\hat{y}_1 내지 \hat{y}_N ; 316)을 획득하도록 구성

되는 업믹서(220; 310); 및

청구항 1 내지 23 중 어느 한 항에 따라 하나 이상의 조정된 매개 변수를 제공하는 장치(100; 240; 320)를 포함하는데, 상기 하나 이상의 조정된 매개 변수를 제공하는 장치는 상기 하나 이상의 입력 매개 변수(110)로서 상기 원하는 렌더링 정보(242;322)를 수신하여, 상기 실제 렌더링 정보로서 상기 하나 이상의 조정된 매개 변수(222; 324)를 제공하도록 구성되며; 및

상기 하나 이상의 조정된 매개 변수를 제공하는 장치는 상기 하나 이상의 조정된 매개 변수를 제공하도록 구성됨으로써, 최적의 렌더링 매개 변수($r_{opt,m}$)에서 벗어나는 상기 실제 렌더링 매개 변수($r_{lim,m}$)의 사용으로 유발되

는 상기 업믹스 오디오 채널(\hat{y}_1 내지 \hat{y}_N ; 316)의 왜곡이 미리 정해진 편차 이상만큼 상기 최적의 렌더링 매개 변수($r_{opt,m}$)에서 벗어나는 적어도 원하는 렌더링 매개 변수(r_i)에 대해 감소되는 것을 특징으로 하는 오디오 신호 디코더.

청구항 25

업믹스 신호 표현(522)으로서, 다운믹스 신호 표현(524), 객체 관련 파라메트릭 정보(520) 및 원하는 렌더링 정보(552,554)에 기초하여 채널 관련 파라메트릭 정보를 제공하는 오디오 신호 트랜스코더(500;560)에 있어서,

상기 객체 관련 파라메트릭 정보(520) 및, 상기 객체 관련 파라메트릭 정보(522)에 의해 나타낸 오디오 객체의 다수의 객체 신호를 상기 채널 관련 파라메트릭 정보에 의해 나타내는 업믹스 오디오 채널에 할당하는 것을 나타내는 실제 렌더링 정보(542)에 따라 상기 다운믹스 신호 표현(524)에 기초하여 상기 채널 관련 파라메트릭 정보(522)를 획득하도록 구성되는 보조 정보 트랜스코더(540); 및

청구항 1 내지 23 중 어느 한 항에 따라 하나 이상의 조정된 매개 변수(542)를 제공하는 장치(100; 550)를 포함하는데, 상기 하나 이상의 조정된 매개 변수를 제공하는 장치는 상기 하나 이상의 입력 매개 변수(110)로서 상기 원하는 렌더링 정보(552;554)를 수신하여, 상기 실제 렌더링 정보(542)로서 상기 하나 이상의 조정된 매개 변수(120)를 제공하도록 구성되며; 및

상기 하나 이상의 조정된 매개 변수를 제공하는 장치는 상기 하나 이상의 조정된 매개 변수(120)를 제공하도록 구성됨으로써, 최적의 렌더링 매개 변수에서 벗어나는 상기 실제 렌더링 매개 변수(542)의 사용으로 유발되는 상기 업믹스 오디오 채널의 왜곡이 미리 정해진 편차 이상만큼 상기 최적의 렌더링 매개 변수에서 벗어나는 적어도 원하는 렌더링 매개 변수(552,554)에 대해 감소되는 것을 특징으로 하는 오디오 신호 트랜스코더.

청구항 26

다운믹스 신호 표현 및 객체 관련 파라메트릭 정보에 기초하여 업믹스 신호 표현을 제공하기 위해 하나 이상의 조정된 매개 변수를 제공하는 방법에 있어서,

하나 이상의 입력 매개 변수를 수신하여, 이에 기초하여 하나 이상의 조정된 매개 변수를 제공하는 단계를 포함하는데,

상기 하나 이상의 조정된 매개 변수 상기 하나 이상의 입력 매개 변수 및 상기 객체 관련 파라메트릭 정보에 따라 제공됨으로써, 최적화되지 않은 매개 변수의 사용으로 유발되는 상기 업믹스 신호 표현의 왜곡은 미리 정해진 편차 이상만큼 적어도 최적의 매개 변수에서 벗어난 입력 매개 변수에 대해 감소되는 것을 특징으로 하는 하나 이상의 조정된 매개 변수를 제공하는 방법.

청구항 27

업믹스 신호 표현으로서, 다운믹스 신호 표현, 객체 관련 파라메트릭 정보 및 원하는 렌더링 정보에 기초하여 다수의 업믹스 오디오 채널을 제공하는 방법에 있어서,

청구항 26에 따라 하나 이상의 조정된 매개 변수를 제공하는 단계로서, 상기 원하는 렌더링 정보는 상기 하나 이상의 입력 매개 변수로서 수신되고, 상기 하나 이상의 조정된 매개 변수는 실제 렌더링 정보로서 제공되며, 상기 하나 이상의 조정된 매개 변수는 최적의 렌더링 매개 변수에서 벗어나는 상기 실제 렌더링 매개 변수의 사용으로 유발되는 상기 업믹스 오디오 채널의 왜곡이 미리 정해진 편차 이상만큼 적어도 상기 최적의 렌더링 매개 변수에서 벗어난 원하는 렌더링 매개 변수에 대해 감소되도록 제공되는 상기 제공하는 단계; 및

상기 객체 관련 파라메트릭 정보 및, 상기 객체 관련 파라메트릭 정보에 의해 나타낸 오디오 객체의 다수의 객체 신호를 상기 업믹스 오디오 채널에 할당하는 것을 나타내는 상기 실제 렌더링 정보에 따라 상기 다운믹스 신호 표현에 기초하여 상기 업믹스 오디오 채널을 획득하는 단계를 포함하는 것을 특징으로 하는 다수의 업믹스 오디오 채널을 제공하는 방법.

청구항 28

업믹스 신호 표현으로서, 다운믹스 신호 표현, 객체 관련 파라메트릭 정보 및 원하는 렌더링 정보에 기초하여 채널 관련 파라메트릭 정보를 제공하는 방법에 있어서,

청구항 26에 따라 하나 이상의 조정된 매개 변수를 제공하는 단계로서, 상기 원하는 렌더링 정보는 상기 하나 이상의 입력 매개 변수로서 수신되고, 상기 하나 이상의 조정된 매개 변수는 실제 렌더링 정보로서 제공되며, 상기 하나 이상의 조정된 매개 변수는 최적의 렌더링 매개 변수에서 벗어나는 상기 실제 렌더링 매개 변수의 사용으로 유발되는 상기 업믹스 오디오 채널의 왜곡이 미리 정해진 편차 이상만큼 적어도 상기 최적의 렌더링 매개 변수에서 벗어난 원하는 렌더링 매개 변수에 대해 감소되도록 제공되는 상기 제공하는 단계; 및

상기 객체 관련 파라메트릭 정보 및, 상기 객체 관련 파라메트릭 정보에 의해 나타낸 오디오 객체의 다수의 객체 신호를 상기 채널 관련 파라메트릭 정보에 의해 나타내는 업믹스 오디오 채널에 할당하는 것을 나타내는 상기 실제 렌더링 정보에 따라 상기 다운믹스 신호 표현에 기초하여 상기 업믹스 오디오 채널을 나타내는 상기 채널 관련 파라메트릭 정보를 획득하는 단계를 포함하는 것을 특징으로 하는 채널 관련 파라메트릭 정보를 제공하는 방법.

청구항 29

다수의 객체 신호(x_1 내지 x_N)에 기초하여 다운믹스 신호 표현(614) 및 객체 관련 파라메트릭 정보(616)를 제공하는 오디오 신호 인코더(600)에 있어서,

상기 객체 신호(x_1 내지 x_N)와 관련된 다운믹스 계수(d_1 내지 d_N)에 따라 하나 이상의 다운믹스 신호를 제공하여, 상기 하나 이상의 다운믹스 신호가 다수의 객체 신호의 중첩을 포함하도록 구성되는 다운믹서(620);

객체 신호(x_1 내지 x_N)의 레벨차 및 상관 특성을 나타내는 객체간 관계 보조 정보(OLD, IOC) 및, 개개의 객체 신호(x_1 내지 x_N)의 하나 이상의 개개의 속성을 나타내는 개개의 객체 보조 정보를 제공하도록 구성되는 보조 정보 제공기(630)를 포함하는 것을 특징으로 하는 오디오 신호 인코더.

청구항 30

청구항 29에 있어서,

상기 보조 정보 제공기(630)는 상기 개개의 객체 보조 정보를 제공하여, 상기 개개의 객체 보조 정보가 상기 개개의 객체 신호(x_1 내지 x_N)의 음조를 나타내도록 구성되는 것을 특징으로 하는 오디오 신호 인코더.

청구항 31

다수의 객체 신호에 기초하여 다운믹스 신호 표현 및 객체 관련 파라메트릭 정보를 제공하는 방법에 있어서,

상기 객체 신호와 관련된 다운믹스 계수에 따라 하나 이상의 다운믹스 신호를 제공하여, 상기 하나 이상의 다운믹스 신호가 다수의 객체 신호의 중첩을 포함하는 단계; 및

객체 신호의 레벨차 및 상관 특성을 나타내는 객체간 관계 보조 정보를 제공하는 단계; 및

상기 개개의 객체 신호의 하나 이상의 개개의 속성을 나타내는 개개의 객체 보조 정보를 제공하는 단계를 포함하는 것을 특징으로 하는 다운믹스 신호 표현 및 객체 관련 파라메트릭 정보를 제공하는 방법.

청구항 32

인코딩된 형식으로 다수의 객체 신호(x_1 내지 x_N)를 표현하는 오디오 비트스트림(700)에 있어서,

하나 이상의 다운믹스 신호를 표현하는 다운믹스 신호(710) 표현으로서, 상기 다운믹스 신호 중 적어도 하나는

다수의 객체 신호의 중첩을 포함하는 다운믹스 신호(710) 표현; 및
 객체 신호의 레벨차 및 상관 특성을 나타내는 객체간 관계 보조 정보(720); 및
 상기 개개의 객체 신호의 하나 이상의 개개의 속성을 나타내는 개개의 객체 보조 정보(730)를 포함하는 것을 특징으로 하는 오디오 비트스트림.

청구항 33

청구항 32에 있어서,
 상기 개개의 객체 보조 정보는 상기 개개의 객체 신호의 음조를 나타내는 것을 특징으로 하는 오디오 비트스트림.

청구항 34

청구항 26, 27, 28 또는 31 항 중 어느 한 항에 따른 방법을 실행하기 위한 컴퓨터 프로그램.

명세서

기술분야

- [0001] 본 발명에 따른 실시예들은 다운믹스 신호 표현(representation) 및 객체 관련 파라메트릭 정보에 기초하여 업믹스 신호 표현을 제공하기 위해 하나 이상의 조정된 매개 변수를 제공하는 장치에 관한 것이다.
- [0002] 본 발명에 따른 다른 실시예는 오디오 신호 디코더에 관한 것이다.
- [0003] 본 발명에 따른 다른 실시예는 오디오 신호 트랜스코더에 관한 것이다.
- [0004] 본 발명에 따른 또 다른 실시예는 하나 이상의 조정된 매개 변수를 제공하는 방법에 관한 것이다.
- [0005] 또 다른 실시예는, 업믹스 신호 표현으로서, 다운믹스 신호 표현, 객체 관련 파라메트릭 정보 및 원하는 렌더링(rendering) 정보에 기초하여 다수의 업믹스 오디오 채널을 제공하는 방법에 관한 것이다.
- [0006] 또 다른 실시예는, 업믹스 신호 표현으로서, 다운믹스 신호 표현, 객체 관련 파라메트릭 정보 및 원하는 렌더링 정보에 기초하여 다운믹스 신호 표현 및 채널 관련 파라메트릭 정보를 제공하는 방법에 관한 것이다.
- [0007] 본 발명에 따른 또 다른 실시예는 오디오 신호 인코더, 인코딩된 오디오 신호 표현을 제공하는 방법 및 오디오 비트스트림에 관한 것이다.
- [0008] 또 다른 실시예는 대응하는 컴퓨터 프로그램에 관한 것이다.
- [0009] 본 발명에 따른 또 다른 실시예는 오디오 신호 처리의 왜곡을 방지하는 방법, 장치 및 컴퓨터 프로그램에 관한 것이다.

배경기술

- [0010] 오디오 처리, 오디오 전송 및 오디오 저장의 기술 분야에서는, 청각 인상(hearing impression)을 개선하기 위해 다중 채널 콘텐츠를 처리하기 위한 소망이 증가하고 있다. 다중 채널 오디오 콘텐츠의 사용은 사용자에 대한 상당한 개선을 제공한다. 예컨대, 엔터테인먼트 애플리케이션에서 향상된 사용자 만족도를 제공하는 3차원 청각 인상이 획득될 수 있다. 그러나, 다중 채널 오디오 콘텐츠는 또한 스피커 명료도(speaker intelligibility)가 다중 채널 오디오 재생을 이용하여 향상될 수 있기 때문에 전문적인 환경에서, 예컨대, 전화 회의 애플리케이션에 유용하다.
- [0011] 그러나, 또한, 다중 채널 애플리케이션에 의해 유발된 과도한 자원 부하를 방지하기 위해 오디오 품질 및 비트레이트 요건 사이에 양호한 트레이드오프(tradeoff)를 갖는 것이 바람직하다.
- [0012] 최근에, 다수의 오디오 객체를 포함하는 오디오 장면의 비트레이트 효율적인 전송 및/또는 저장을 위한 파라메트릭 기술이 제안되었는데, 예컨대, Binaural Cue Coding(Type I)(예컨대, 참고 문헌 [BCC] 참조), Joint Source Coding(예컨대, 참고 문헌 [JSC] 참조), 및 MPEG Spatial Audio Object Coding(SAOC)(예컨대, 참고 문헌 [SAOC1], [SAOC2] 참조).

- [0013] 이들 기술은 과형 일치에 의해서보다는 원하는 출력 오디오 장면을 지각적으로 재구성하는 것을 목표로 한다.
- [0014] 도 8은 이와 같은 시스템(여기서는: MPEG SAOC)의 시스템 개요를 도시한다. 도 8에 도시된 MPEG SAOC 시스템(800)은 SAOC 인코더(810) 및 SAOC 디코더(820)를 포함한다. SAOC 인코더(810)는, 예컨대, 시간-도메인 신호 또는 시간-주파수-도메인 신호(예컨대, 푸리에 타입 변환의 변환 계수의 세트의 형식, 또는 QMF 부대역 신호의 형식)로서 나타낼 수 있는 다수의 객체 신호 x_1 내지 x_N 를 수신한다. SAOC 인코더(810)는 전형적으로 또한 객체 신호 x_1 내지 x_N 와 관련되는 다운믹스 계수 d_1 내지 d_N 를 수신한다. 다운믹스 계수의 별도의 세트는 다운믹스 신호의 각 채널에 이용 가능할 수 있다. SAOC 인코더(810)는 전형적으로 관련된 다운믹스 계수 d_1 내지 d_N 에 따라 객체 신호 x_1 내지 x_N 를 조합하여 다운믹스 신호의 채널을 획득하도록 구성된다. 전형적으로, 객체 신호 x_1 내지 x_N 보다 적은 다운믹스 채널이 존재한다. SAOC 디코더(820) 측에서 객체 신호의 분리(또는 별도의 처리)를 허용하기 위해, SAOC 인코더(810)는 (다운믹스 채널로 지정되는) 하나 이상의 다운믹스 신호(812) 및 보조 정보(814)의 양방을 제공한다. 보조 정보(814)는 디코더측 객체 특정 처리를 허용하기 위해 객체 신호 x_1 내지 x_N 의 속성을 나타낸다.
- [0015] SAOC 디코더(820)는 하나 이상의 다운믹스 신호(812) 및 보조 정보(814)의 양방을 수신하도록 구성된다. 또한, SAOC 디코더(820)는 전형적으로 원하는 렌더링 설정을 나타내는 사용자 상호 작용 정보 및/또는 사용자 제어 정보(822)를 수신하도록 구성된다. 예컨대, 사용자 상호 작용 정보/사용자 제어 정보(822)는 객체 신호 x_1 내지 x_N 를 제공하는 객체의 원하는 공간 배치 및 스피커 설정을 나타낼 수 있다.
- [0016] SAOC 디코더(820)는 예컨대 다수의 디코딩된 업믹스 채널 신호 \hat{y}_1 내지 \hat{y}_M 를 제공하도록 구성된다. 업믹스 채널 신호는 예컨대 멀티 스피커 렌더링 장치의 개별 스피커와 관련될 수 있다. SAOC 디코더(820)는, 예컨대, 하나 이상의 다운믹스 신호(812) 및 보조 정보(814)에 기초하여 객체 신호 x_1 내지 x_N 를 적어도 대략 재구성하여, 재구성된 객체 신호(820b)를 획득하도록 구성되는 객체 분리기(820a)를 포함할 수 있다. 그러나, 재구성된 객체 신호(820b)는, 예컨대, 보조 정보(814)가 비트레이트 제한(bitrate constraints)으로 인해 완전한 재구성에 대해 매우 충분하지 않기 때문에 원래의 객체 신호 x_1 내지 x_N 에서 약간 벗어날 수 있다. SAOC 디코더(820)는, 재구성된 객체 신호(820b) 및 사용자 상호 작용 정보/사용자 제어 정보(822)를 수신하여, 이에 기초하여, 업믹스 채널 신호 \hat{y}_1 내지 \hat{y}_M 를 제공하도록 구성될 수 있는 믹서(820c)를 더 포함할 수 있다. 믹서(820c)는 사용자 상호 작용 정보/사용자 제어 정보(822)를 이용하여 업믹스 채널 신호 \hat{y}_1 내지 \hat{y}_M 에 대한 별도의 재구성된 객체 신호(820b)의 기여를 결정하도록 구성될 수 있다. 사용자 상호 작용 정보/사용자 제어 정보(822)는, 예컨대, 업믹스 채널 신호 \hat{y}_1 내지 \hat{y}_M 에 대한 별도의 재구성된 객체 신호(822)의 기여를 결정하는 (또한 렌더링 계수로서 지정되는) 렌더링 매개 변수를 포함할 수 있다.
- [0017] 그러나, 많은 실시예들에서, 도 8에서 객체 분리기(820a)에 의해 표시되는 객체 분리, 및 도 8에서 믹서(820c)에 의해 표시되는 믹싱은 단일 단계에서 수행된다. 이러한 목적으로, 업믹스 채널 신호 \hat{y}_1 내지 \hat{y}_M 상에서의 하나 이상의 다운믹스 신호(812)의 직접 매핑을 나타내는 전체 매개 변수가 계산될 수 있다. 이들 매개 변수는 보조 정보 및 사용자 상호 작용 정보 및/또는 사용자 제어 정보(820)에 기초하여 계산될 수 있다.
- [0018] 이제 도 9a, 9b 및 9c를 참조하면, 다운믹스 신호 표현 및 객체 관련 보조 정보에 기초하여 업믹스 신호 표현을 획득하기 위한 여러 장치가 설명될 것이다. 도 9a는 SAOC 디코더(920)를 포함하는 MPEG SAOC 시스템(900)의 개략적인 블록도를 도시한 것이다. SAOC 디코더(920)는, 별도의 기능적 블록으로서, 객체 디코더(922) 및 믹서/렌더러(renderer)(926)를 포함한다. 객체 디코더(922)는 다운믹스 신호 표현(예컨대, 시간 도메인 또는 시간-주파수-도메인에 나타낸 하나 이상의 다운믹스 신호의 형식) 및 객체 관련 보조 정보(예컨대, 객체 메타 데이터의 형식)에 따라 다수의 재구성된 객체 신호(924)를 제공한다. 믹서/렌더러(924)는 다수의 N 객체와 관련되는 재구성된 객체 신호(924)를 수신하여, 이에 기초하여, 하나 이상의 업믹스 채널 신호(928)를 제공한다. SAOC 디코더(920)에서, 객체 신호(924)의 추출은, 믹싱/렌더링 기능에서 객체 디코딩 기능의 분리를 허용하지만, 비교적 고 계산 복잡도를 제공하는 믹싱/렌더링과는 별도로 수행된다.
- [0019] 이제 도 9b를 참조하면, SAOC 디코더(950)를 포함하는 다른 MPEG SAOC 시스템(930)이 간략히 논의될 것이다.

SAOC 디코더(950)는 다운믹스 신호 표현(예컨대, 하나 이상의 다운믹스 신호의 형식) 및 객체 관련 보조 정보(예컨대, 객체 메타 데이터의 형식)에 따라 다수의 업믹스 채널 신호(958)를 제공한다. SAOC 디코더(950)는 조합된 객체 디코더 및 믹서/렌더러를 포함하며, 이는 객체 디코딩 및 믹싱/렌더링의 분리 없이 조인트(joint) 믹싱 프로세스에서 업믹스 채널 신호(958)를 획득하도록 구성되며, 상기 조인트 업믹스 프로세스에 대한 매개 변수는 객체 관련 보조 정보 및 렌더링 정보의 양방에 의존한다. 조인트 업믹스 프로세스는 또한 객체 관련 보조 정보의 부분인 것으로 고려되는 다운믹스 정보에 의존한다.

- [0020] 상술한 바를 요약하기 위해, 업믹스 채널 신호(928, 958)의 제공은 1 단계 프로세스 또는 2 단계 프로세스에서 수행될 수 있다.
- [0021] 이제 도 9c를 참조하면, MPEG SAOC 시스템(960)이 설명될 것이다. SAOC 시스템(960)은 SAOC 디코더 보다는 SAOC 대 MPEG 서라운드 트랜스코더(980)를 포함한다.
- [0022] SAOC 대 MPEG 서라운드 트랜스코더는 객체 관련 보조 정보(예컨대, 객체 메타 데이터의 형식) 및, 선택적으로, 하나 이상의 다운믹스 신호 상의 정보 및 렌더링 정보를 수신하도록 구성되는 보조 정보 트랜스코더(982)를 포함한다. 보조 정보 트랜스코더는 또한 수신된 데이터에 기초하여 MPEG 서라운드 보조 정보(예컨대, MPEG 서라운드 비트스트림의 형식)를 제공하도록 구성된다. 따라서, 보조 정보 트랜스코더(982)는, 객체 인코더로부터 수신되는 객체 관련 (파라메트릭) 보조 정보를, 렌더링 정보 및, 선택적으로 하나 이상의 다운믹스 신호의 내용에 관한 정보를 고려하는 채널 관련 (파라메트릭) 보조 정보로 변환하도록 구성된다.
- [0023] 선택적으로, SAOC 대 MPEG 서라운드 트랜스코더(980)는, 예컨대, 다운믹스 신호 표현에 의해 나타내는 하나 이상의 다운믹스 신호를 조작하여, 조작된 다운믹스 신호 표현(988)을 획득하도록 구성될 수 있다. 그러나, SAOC 대 MPEG 서라운드 트랜스코더(980)의 출력 다운믹스 신호 표현(988)이 SAOC 대 MPEG 서라운드 트랜스코더의 입력 다운믹스 신호 표현과 동일하도록 다운믹스 신호 조작기(986)는 생략될 수 있다. 다운믹스 신호 조작기(986)는, 예컨대, 채널 관련 MPEG 서라운드 보조 정보(984)가 SAOC 대 MPEG 서라운드 트랜스코더(980)의 입력 다운믹스 신호 표현에 기초하여 원하는 청각 인상을 제공하도록 허용하지 않을 경우에 이용될 수 있으며, 이는 일부 렌더링 별자리(rendering constellations)의 경우일 수 있다.
- [0024] 따라서, SAOC 대 MPEG 서라운드 트랜스코더(980)로 입력되는 렌더링 정보에 따라 오디오 객체를 나타내는 다수의 업믹스 채널 신호가 MPEG 서라운드 비트스트림(984) 및 다운믹스 신호 표현(988)을 수신하는 MPEG 서라운드 디코더를 이용하여 생성될 수 있도록 SAOC 대 MPEG 서라운드 트랜스코더(980)는 다운믹스 신호 표현(988) 및 MPEG 서라운드 비트스트림(984)을 제공한다.
- [0025] 상술한 바를 요약하기 위해, SAOC 인코딩된 오디오 신호를 디코딩하는 여러 개념이 이용될 수 있다. 어떤 경우에, 다운믹스 신호 표현 및 객체 관련 파라메트릭 보조 정보에 따라 업믹스 채널 신호(예컨대, 업믹스 채널 신호(928, 958)를 제공하는 SAOC 디코더가 이용된다. 이러한 개념에 대한 예들은 도 9a 및 9b에서 알 수 있다. 대안적으로, SAOC 인코딩된 오디오 정보는, 다운믹스 신호 표현(예컨대, 다운믹스 신호 표현(988)) 및, MPEG 서라운드 디코더에 의해 원하는 업믹스 채널 신호를 제공하기 위해 이용될 수 있는 채널 관련 보조 정보(예컨대, 채널 관련 MPEG 서라운드 비트스트림(984))를 획득하도록 트랜스코딩될 수 있다.
- [0026] MPEG SAOC 시스템(800)에서, 이의 시스템 개요는 도 8에 주어지며, 일반적인 처리는 주파수 선택 방식으로 실행되고, 각 주파수 대역 내에서 다음과 같이 설명될 수 있다:
- [0027]
 - N 입력 오디오 객체 신호 x_1 내지 x_N 는 SAOC 인코더 처리의 부분으로서 다운믹스된다. 모노 다운믹스의 경우, 다운믹스 계수는 d_1 내지 d_N 으로 나타낸다. 게다가, SAOC 인코더(810)는 입력 오디오 객체의 속성을 나타내는 보조 정보(814)를 추출한다. MPEG SAOC의 경우, 서로에 대한 객체 파워(power)의 관계는 이와 같은 보조 정보의 가장 기본적 형식이다.
- [0028]
 - 다운믹스 신호(또는 신호)(812) 및 보조 정보(814)는 전송되고, 및/또는 저장된다. 이를 위해, 다운믹스 오디오 신호는 MPEG-1 Layer II 또는 III(또한, ".mp3"로 알려져 있음), MPEG Advanced Audio Coding(AAC), 또는 어떤 다른 오디오 코더와 같은 잘 알려진 지각 오디오 코더를 이용하여 압축될 수 있다.
- [0029]
 - 수신단에서, SAOC 디코더(820)는 개념적으로 전송된 보조 정보(814)(및, 당연히, 하나 이상의 다운믹스 신호(812))를 이용하여 원래의 객체 신호("객체 분리")를 복원하려고 시도한다. 그 후, 이들 근사화된 객체 신호(또

한 재구성된 객체 신호(820b)로 지정됨)는 렌더링 매트릭스를 이용하여 (예컨대, 업믹스 채널 신호 \hat{y}_1 내지 \hat{y}_M 로 나타낼 수 있는) M 오디오 출력 채널로 나타내는 타겟 장면으로 믹싱된다. 모노 출력의 경우, 렌더링 매트릭스 계수는 r_1 내지 r_N 으로 주어진다.

[0030] • 효과적으로, 객체 신호의 분리는 좀처럼 실행되지 않는데(또는 결코 실행되지 않음), 그 이유는 (객체 분리기 (820a)에 의해 표시되는) 분리 단계 및 (믹서(820c)에 의해 표시되는) 믹싱 단계의 양방이 단일 트랜스코딩 단계로 조합되어, 종종 결과적으로 계산 복잡도의 엄청난 감소를 유발시키기 때문이다.

[0031] 이와 같은 기법은 전송 비트레이트(몇몇 다운믹스 채널 플러스 N 이산 객체 오디오 신호 또는 이산 시스템 대신에 일부 보조 정보를 전송하는데만 필요함) 및 계산 복잡도(처리 복잡도는 주로 오디오 객체의 수보다는 출력 채널의 수에 관계함)의 양방의 관점에서 대단히 효율적임이 발견되었다. 수신단에서 사용자에게 대한 추가적 이점은 자신의 선택(모노, 스테레오, 서라운드, 가상 헤드폰 재생 등)의 렌더링 설정을 선택하는 자유 및 사용자 상호 작용의 특징(feature): 렌더링 매트릭스를 포함하여, 출력 장면이 뜻, 개인 선호 또는 다른 기준에 따라 사용자에게 의해 상호 작용하게 설정되고 변경될 수 있다. 예컨대, 다른 잔여 토크와의 구별을 최대화하기 위해 한 공간 영역에 한 그룹으로부터의 토크를 함께 위치시킬 수 있다. 이러한 상호 작용은 디코더 사용자 인터페이스를 제공함으로써 달성된다.

[0032] 각 전송된 사운드 객체에 대해, 그의 상대 레벨 및 (비모노 렌더링에 대해) 렌더링하는 공간 위치가 조정될 수 있다. 이것은 사용자가 관련된 그래픽 사용자 인터페이스(GUI) 슬라이더의 위치(예컨대: 객체 레벨 = 5dB, 객체 위치 = - 30deg)를 변경할 시에 실시간으로 발생할 수 있다.

[0033] 그러나, 업믹스 신호 표현(예컨대, 업믹스 채널 신호 \hat{y}_1 내지 \hat{y}_M)의 제공을 위한 매개 변수의 디코더측 선택은 어떤 경우에는 가청 왜곡을 가져온다는 것이 발견되었다.

발명의 내용

해결하려는 과제

[0034] 이러한 상황을 고려하여, 본 발명의 목적은 업믹스 신호 표현(예컨대, 업믹스 채널 신호 \hat{y}_1 내지 \hat{y}_M 의 형식)을 제공할 때에 가청 왜곡을 감소시키거나 심지어 방지하는 것을 허용하는 개념을 생성하기 위한 것이다.

과제의 해결 수단

[0035] 이런 문제는, 청구항 1에 따라 다운믹스 신호 표현 및 객체 관련 파라메트릭 정보에 기초하여 업믹스 신호 표현을 제공하기 위해 하나 이상의 조정된 매개 변수를 제공하는 장치, 청구항 24에 따른 오디오 신호 디코더, 청구항 25에 따른 오디오 신호 트랜스코더, 청구항 26, 27 및 28에 따른 방법, 청구항 29에 따른 오디오 신호 인코더, 청구항 31에 따른 방법, 청구항 32에 따른 오디오 비트스트림 및 청구항 34에 따른 컴퓨터 프로그램에 의해 해결된다.

[0036] 본 발명에 따른 실시예는 다운믹스 신호 표현 및 객체 관련 파라메트릭 정보에 기초하여 업믹스 신호 표현을 제공하기 위해 하나 이상의 조정된 매개 변수를 제공하는 장치를 생성한다. 이 장치는 하나 이상의 입력 매개 변수(예컨대, 렌더링 계수 또는 원하는 렌더링 매트릭스에 대한 설명)를 수신하여, 이에 기초하여 하나 이상의 조정된 매개 변수를 제공하도록 구성되는 매개 변수 조정기(예컨대, 렌더링 계수 조정기)를 포함한다. 매개 변수 조정기는 하나 이상의 입력 매개 변수 및 객체 관련 파라메트릭 정보에 따라(예컨대, 하나 이상의 다운믹스 계수, 및/또는 하나 이상의 객체 레벨 차이 값, 및/또는 하나 이상의 객체간 상관 값에 따라) 하나 이상의 조정된 매개 변수를 제공하도록 구성됨으로써, 최적화되지 않은 매개 변수의 사용으로 유발되는 업믹스 신호 표현의 왜곡은 미리 정해진 편차 이상만큼 적어도 최적의 매개 변수에서 벗어난 입력 매개 변수에 대해 감소된다.

[0037] 본 발명에 따른 이런 실시예는 부적절하게 선택된 입력 매개 변수에 의해 유발된 오디오 신호 왜곡이 업믹스 신호 표현을 제공하기 위해 조정된 매개 변수를 제공함으로써 감소될 수 있고, 조정된 매개 변수의 제공이 객체 관련 파라메트릭 정보를 고려하여 양호한 정확도로 수행될 수 있다는 사상에 기초한다. 객체 관련 파라메트릭 정보의 이용은 입력 매개 변수의 이용으로 유발된 가청 왜곡의 추정치 측정을 획득할 수 있도록 하여, 결과적으로 입력 매개 변수와 비교했을 때 미리 정해진 범위 내에 가청 왜곡을 유지시키는데 적합하거나 가청 왜곡을 감

소시키는데 적합한 조정된 매개 변수를 제공할 수 있는 것으로 발견되었다. 객체 관련 정보는, 예컨대, 오디오 객체의 속성을 나타내고, 및/또는 객체의 인코더측 처리에 관한 정보를 제공한다.

[0038] 따라서, 부적절한 매개 변수(예컨대, 부적절한 렌더링 계수)의 이용으로 유발된 바람직하지 않고 종종 성가신 오디오 신호 왜곡은 하나 이상의 조정된 매개 변수를 제공하여 감소될 수 있거나, 심지어 방지될 수 있으며, 매개 변수의 조정을 위한 객체 관련 파라메트릭 정보의 고려는 가청 왜곡의 비교적 신뢰성이 높은 추정을 하여 오디오 신호 왜곡의 효과적 감소 및/또는 제한을 확실히 하는데 도움을 준다.

[0039] 바람직한 실시예에서, 장치는, 입력 매개 변수로서, 업믹스 신호 표현에 의해 나타내는 하나 이상의 채널에서 다수의 오디오 객체 신호의 원하는 강도 스케일링을 나타내는 원하는 렌더링 매개 변수를 수신하도록 구성된다. 이 경우에, 매개 변수 조정기는 하나 이상의 원하는 렌더링 매개 변수에 따라 하나 이상의 실제 렌더링 매개 변수를 제공하도록 구성된다. 부적절한 렌더링 매개 변수의 선택은 이와 같이 부적절하게 선택된 렌더링 매개 변수를 이용하여 획득되는 업믹스 신호 표현의 상당한 (및 종종 가청) 저하를 가져오는 것으로 발견되었다. 또한, 렌더링 매개 변수는 객체 관련 파라메트릭 정보에 따라 효율적으로 조정될 수 있는 것으로 발견되었는데, 그 이유는 객체 관련 파라메트릭 정보가 (입력 매개 변수로 규정될 수 있는) 렌더링 매개 변수의 주어진 선택에 의해 도입되는) 왜곡의 추정을 하기 때문이다.

[0040] 바람직한 실시예에서, 매개 변수 조정기는 객체 관련 파라메트릭 정보 및, 다운믹스 신호 표현에 대한 오디오 객체 신호의 기여를 나타내는 다운믹스 정보에 따라 하나 이상의 렌더링 매개 변수 제한 값을 획득하도록 구성됨으로써, 왜곡 메트릭(distortion metric)은 렌더링 매개 변수 제한 값으로 정의되는 제한에 따르는 렌더링 매개 변수 값에 대한 미리 정해진 범위 내에 있도록 한다. 이 경우에, 매개 변수 조정기는 원하는 렌더링 매개 변수 및 하나 이상의 렌더링 매개 변수 제한 값에 따라 실제 렌더링 매개 변수를 획득하도록 구성됨으로써, 실제 렌더링 매개 변수는 렌더링 매개 변수 제한 값으로 정의되는 제한에 따르도록 한다. 컴퓨팅 렌더링 매개 변수 제한 값은 가청 왜곡이 왜곡 메트릭에 따라 허용 가능한 범위 내에 있음을 확실히 하기 위해 계산상 간단하고 신뢰 가능한 메커니즘을 구성한다.

[0041] 바람직한 실시예에서, 매개 변수 조정기는 하나 이상의 렌더링 매개 변수 제한 값을 획득하도록 구성됨으로써, 하나 이상의 렌더링 매개 변수 제한 값에 따르는 렌더링 매개 변수를 이용하여 렌더링되는 다수의 객체 신호의 렌더링된 중첩(superposition)에서의 객체 신호의 상대 기여는 단지 미리 정해진 차이만큼 다운믹스 신호에서의 객체 신호의 상대 기여와 상이하도록 한다. 객체 신호의 렌더링된 중첩에서의 객체 신호의 기여가 다운믹스 신호에서의 객체 신호의 기여와 유사할 경우에는 왜곡이 전형적으로 상당히 작지만, 상기 상대 기여의 강한 차이는 전형적으로 가청 왜곡을 가져오는 것으로 발견되었다. 이것은 다운믹스 신호 표현에서의 객체 신호의 (상대) 레벨과 비교했을 때에 객체 신호의 (상대) 레벨의 강한 변경이 종종 이상적인 방식으로 서로 다른 오디오 객체의 객체 신호를 분리할 수 없기 때문에 종종 아티팩트(artifacts)를 가져온다는 사실에 기인한다. 따라서, 객체 신호의 상대 기여가 렌더링 매개 변수의 선택에 의해서만 알맞게 변경되도록 렌더링 매개 변수를 조정하는 양호한 결과를 가져온다는 것이 발견되었다.

[0042] 다른 실시예에서, 매개 변수 조정기는 하나 이상의 렌더링 매개 변수 제한 값을 획득하도록 구성됨으로써, 다운믹스 신호 표현에 의해 나타내는 다운믹스 신호와, 하나 이상의 렌더링 매개 변수 제한 값에 따르는 하나 이상의 렌더링 매개 변수를 이용하여 렌더링되는 렌더링된 신호 사이의 코히어런스(coherence)를 나타내는 왜곡 측정이 미리 정해진 범위 내에 있도록 한다. 매개 변수 조정기의 입력 매개 변수를 형성하는 원하는 렌더링 매개 변수의 선택은 충분한 "유사성"이 다운믹스 신호 표현에 의해 나타내는 다운믹스 신호와 렌더링된 신호 사이에서 유지되도록 행해져야 하는데, 그 이유는 그렇지 않으면 업믹스 프로세스에서 가청 아티팩트를 획득하는 위험이 매우 높기 때문이다.

[0043] 또 다른 바람직한 실시예에서, 매개 변수 조정기는 (매개 변수 조정기의 입력 매개 변수를 형성할 수 있는) 원하는 렌더링 매개 변수의 제곱(square)과, (예컨대, 왜곡 메트릭을 최소화하는 렌더링 매개 변수로서 정의될 수 있는) 최적의 렌더링 매개 변수의 제곱 사이의 선형 조합을 계산하여, (장치에 의해 조정된 매개 변수로서 출력될 수 있는 실제 렌더링 매개 변수를 획득하도록 구성된다. 이 경우에, 매개 변수 조정기는 미리 정해진 임계 매개 변수 T 및 왜곡 메트릭에 따라 선형 조합에 대한 원하는 렌더링 매개 변수 및 최적의 렌더링 매개 변수의 기여를 결정하도록 구성되며, 왜곡 메트릭은 다운믹스 신호 표현에 기초하여 업믹스 신호 표현을 획득하기 위해 최적의 렌더링 매개 변수 보다는 하나 이상의 원하는 렌더링 매개 변수를 이용함으로써 유발되는 왜곡을 나타낸다. 이러한 개념은 원하는 렌더링 매개 변수의 충분한 영향을 여전히 유지하면서 수락 가능한 측정치까지 왜곡을 감소시킨다. 이러한 개념에 따르면, 최적의 렌더링 매개 변수와 원하는 렌더링 매개 변수 사이에 상당히 양

호한 타협(compromise)이 가청 왜곡을 제한하는 원하는 정도를 고려하여 찾아질 수 있다

- [0044] 바람직한 실시예에서, 매개 변수 조정기는 지각 저하의 계산 측정에 따라 하나 이상의 조정된 매개 변수를 제공하도록 구성됨으로써, 최적이지 아닌 매개 변수의 사용으로 유발되고, 지각 저하의 계산 측정에 의해 나타내는 업믹스 신호 표현의 지각적 평가 왜곡이 제한되도록 한다. 이러한 방식으로, 매개 변수는 청각 인상에 따라 조정되어, 사용자의 욕망에 따라 매개 변수를 조정할 시에 여전히 충분한 유연성을 제공하면서 수락할 수 없는 나쁜 청각 인상을 방지하는 것으로 달성될 수 있다.
- [0045] 바람직한 실시예에서, 매개 변수 조정기는 다운믹스 신호 표현에 의해 나타내는 다운믹스 신호에 대한 기초를 형성하는 하나 이상의 원래의 객체 신호의 속성을 나타내는 객체 속성 정보를 수신하도록 구성된다. 이러한 경우에, 매개 변수 조정기는 조정된 매개 변수를 제공하기 위해 객체 속성 정보를 고려하도록 구성됨으로써, 업믹스 신호 표현 내에 포함되는 객체 신호의 속성에 대한 업믹스 신호 표현의 왜곡이 미리 정해진 편차 이상만큼 적어도 최적의 매개 변수에서 벗어나는 입력 매개 변수에 대해 감소되도록 한다. 본 발명에 따른 이러한 실시예는 하나 이상의 원래의 객체 신호의 속성이 입력 매개 변수가 적절한지 조정되어야 하는지를 평가하는데 이용될 수 있다는 발견에 기초하는데, 그 이유는 업믹스 신호의 특성이 하나 이상의 원래의 객체 신호의 속성에 관계되도록 업믹스 신호를 제공하는 것이 바람직하고, 그렇지 않으면 지각 인상이 많은 경우에 상당히 저하되기 때문이다.
- [0046] 바람직한 실시예에서, 매개 변수 조정기는 하나 이상의 조정된 매개 변수를 제공하기 위해, 객체 속성 정보로서, 객체 신호 음조(tonality) 정보를 수신하고 고려하도록 구성된다. 객체 신호의 음조는 지각 인상에 상당한 영향을 미치는 수량(quantity)이고, 음조 인상을 상당히 변화시키는 매개 변수의 선택은 양호한 청각 인상을 갖기 위해 방지되어야 하는 것으로 발견되었다.
- [0047] 바람직한 실시예에서, 매개 변수 조정기는 수신된 객체 신호 음조 정보 및 수신된 객체 파워 정보에 따라 이상적으로 렌더링된 업믹스 신호의 음조를 추정하도록 구성된다. 이 경우에, 매개 변수 조정기는 하나 이상의 조정된 매개 변수를 제공하여, 추정된 음조와 입력 매개 변수를 이용하여 획득된 업믹스 신호의 음조 사이의 차와 비교했을 때 추정된 음조와 하나 이상의 조정된 매개 변수를 이용하여 획득된 업믹스 신호의 음조 사이의 차를 감소시키거나, 추정된 음조와, 미리 정해진 범위 내에 하나 이상의 조정된 매개 변수를 이용하여 획득된 업믹스 신호의 음조 사이의 차를 유지하도록 구성된다. 이러한 개념을 이용하여, 청각 인상의 저하에 대한 측정이 렌더링 매개 변수의 적절한 조정을 허용하는 높은 계산 효율로 획득될 수 있다.
- [0048] 바람직한 실시예에서, 매개 변수 조정기는 입력 매개 변수의 시간 및 주파수 변형 조정을 수행하도록 구성된다. 따라서, 조정된 매개 변수를 획득하기 위해 입력 매개 변수의 조정은 조정이 실제로 청각 인상의 개선을 가져오거나 청각 인상의 상당한 저하를 방지하는 그런 시간 구간 또는 주파수 영역에 대해서만 수행될 수 있다.
- [0049] 또 다른 바람직한 실시예에서, 매개 변수 조정기는 또한 하나 이상의 조정된 매개 변수를 제공하기 위한 다운믹스 신호 표현을 고려하도록 구성된다. 다운믹스 신호 표현을 고려함으로써, 청각 인상의 가능한 왜곡에 대한 더욱 정확한 추정이 획득될 수 있다.
- [0050] 바람직한 실시예에서, 매개 변수 조정기는 다수의 아티팩트의 타입을 나타내는 왜곡 측정의 조합인 전체 왜곡 측정을 획득하도록 구성된다. 이 경우에, 매개 변수 조정기는 전체 왜곡 측정이 다운믹스 신호 표현에 기초하여 업믹스 신호 표현을 획득하기 위한 최적의 렌더링 매개 변수보다는 하나 이상의 입력 렌더링 매개 변수를 이용함으로써 유발되는 왜곡의 측정이도록 전체 왜곡 측정을 획득하기 위해 구성된다. 다수의 아티팩트의 타입을 나타내는 다수의 왜곡 측정을 조합함으로써, 청각 인상을 조정하기 위한 잘 제어된 메카니즘이 생성된다.
- [0051] 본 발명에 따른 다른 실시예는, 업믹스 신호 표현으로서, 다운믹스 신호 표현, 객체 관련 파라메트릭 정보 및 원하는 렌더링 정보에 기초하여 다수의 업믹스 오디오 채널을 제공하는 오디오 신호 디코더를 생성한다. 오디오 신호 디코더는 객체 관련 파라메트릭 정보 및 실제 렌더링 정보에 따라 다운믹스 신호 표현에 기초하여 업믹스 오디오 채널을 획득하도록 구성되는 업믹서를 포함하는데, 상기 실제 렌더링 정보는 객체 관련 파라메트릭 정보에 의해 나타낸 오디오 객체의 다수의 객체 신호를 업믹스 오디오 채널에 할당하는 것을 나타낸다. 오디오 신호 디코더는 또한 상술한 바와 같이 하나 이상의 조정된 매개 변수를 제공하는 장치를 포함한다. 하나 이상의 조정된 매개 변수를 제공하는 장치는 하나 이상의 입력 매개 변수로서 원하는 렌더링 정보를 수신하여, 실제 렌더링 정보로서 하나 이상의 조정된 매개 변수를 제공하도록 구성된다. 하나 이상의 조정된 매개 변수를 제공하는 장치는 또한 최적의 렌더링 매개 변수에서 벗어나는 실제 렌더링 매개 변수의 사용으로 유발되는 업믹스 오디오 채널의 왜곡이 미리 정해진 편차 이상만큼 최적의 렌더링 매개 변수에서 벗어나는 적어도 원하는 렌더링 매개

변수에 대해 감소되도록 하나 이상의 조정된 매개 변수를 제공하기 위해 구성된다.

[0052] 오디오 신호 디코더에서 하나 이상의 조정된 매개 변수를 제공하는 장치의 사용으로, 부적절하게 선택된 원하는 렌더링 정보로 오디오 디코딩을 수행함으로써 유발되는 강한 가청 왜곡의 생성이 방지된다.

[0053] 본 발명에 따른 다른 실시예는, 업믹스 신호 표현으로서, 다운믹스 신호 표현, 객체 관련 파라메트릭 정보 및 원하는 렌더링 정보에 기초하여 채널 관련 파라메트릭 정보를 제공하는 오디오 신호 트랜스코더를 생성한다. 오디오 신호 트랜스코더는 객체 관련 파라메트릭 정보 및 실제 렌더링 정보에 따라 다운믹스 신호 표현에 기초하여 채널 관련 파라메트릭 정보를 획득하도록 구성되는 보조 정보 트랜스코더를 포함하는데, 상기 실제 렌더링 정보는 객체 관련 파라메트릭 정보에 의해 나타낸 오디오 객체의 다수의 객체 신호를 업믹스 오디오 채널에 할당하는 것을 나타낸다. 오디오 신호 디코더는 또한 상술한 바와 같이 하나 이상의 조정된 매개 변수를 제공하는 장치를 포함한다. 하나 이상의 조정된 매개 변수를 제공하는 장치는 하나 이상의 입력 매개 변수로서 원하는 렌더링 정보를 수신하여, 실제 렌더링 정보로서 하나 이상의 조정된 매개 변수를 제공하도록 구성된다. 또한, 하나 이상의 조정된 매개 변수를 제공하는 장치는 최적의 렌더링 매개 변수에서 벗어나는 실제 렌더링 매개 변수의 사용으로 유발되는 (다운믹스 신호 정보와 조합하여) 채널 관련 파라메트릭 정보에 의해 나타내는 업믹스 오디오 채널의 왜곡이 미리 정해진 편차 이상만큼 최적의 렌더링 매개 변수에서 벗어나는 적어도 원하는 렌더링 매개 변수에 대해 감소되도록 하나 이상의 조정된 매개 변수를 제공하기 위해 구성된다. 조정된 매개 변수를 제공하는 개념은 또한 오디오 신호 트랜스코더와 조합하여 사용하는데 적합한 것으로 발견되었다.

[0054] 본 발명에 따른 다른 실시예는 하나 이상의 조정된 매개 변수를 제공하는 방법, 오디오 신호를 디코딩하는 방법 및 오디오 신호를 트랜스코딩하는 방법을 생성한다. 상기 방법은 상술한 장치와 동일한 핵심 아이디어에 기초한다.

[0055] 본 발명에 따른 다른 실시예는 다수의 객체 신호에 기초하여 다운믹스 신호 표현 및 객체 관련 파라메트릭 정보를 제공하는 오디오 신호 인코더를 생성한다. 오디오 인코더는 상기 객체 신호와 관련된 다운믹스 계수에 따라 하나 이상의 다운믹스 신호를 제공하여, 하나 이상의 다운믹스 신호가 다수의 객체 신호의 중첩을 포함하도록 구성되는 다운믹서를 포함한다. 오디오 인코더는 또한 객체 신호의 레벨차 및 상관 특성을 나타내는 객체간 관계 보조 정보 및, 개개의 객체 신호의 하나 이상의 개개의 속성을 나타내는 개개의 객체 보조 정보를 제공하도록 구성되는 보조 정보 제공기를 포함한다. 오디오 신호 인코더에 의해 객체간 관계 보조 정보 및 개개의 객체 보조 정보의 양방의 제공은 다중 채널 오디오 신호 디코더측에서 가청 왜곡을 효율적으로 감소시키거나 심지어 방지하는 것으로 발견되었다. 객체간 관계 보조 정보가 디코더측에서 객체 신호를 분리하기 위해 이용되지만, 개개의 객체 보조 정보는 객체 신호의 개개의 특성이 디코더측에 유지되는지를 판단하는데 이용될 수 있으며, 이는 왜곡이 허용 가능한 공차 내에 있음을 나타낸다.

[0056] 바람직한 실시예에서, 보조 정보 제공기는 개개의 객체 보조 정보를 제공하여 개개의 객체 보조 정보가 개개의 객체의 음조를 나타내도록 구성된다. 개개의 객체의 음조는 왜곡의 디코더측 제한(decoder-sided limitation)을 허용하는 심리 음향으로 중요한 수량인 것으로 발견되었다.

[0057] 본 발명에 따른 다른 실시예는 오디오 신호를 인코딩하는 방법을 생성한다.

발명의 효과

[0058] 본 발명에 따른 다른 실시예는 인코딩된 형식으로 다수의 (오디오) 객체 신호를 나타내는 오디오 비트스트림을 생성한다. 오디오 비트스트림은 하나 이상의 다운믹스 신호를 나타내는 다운믹스 신호 표현을 포함하며, 다운믹스 신호 중 적어도 하나는 다수의 (오디오) 객체 신호의 중첩을 포함한다. 오디오 비트스트림은 또한 객체 신호의 레벨차 및 상관 특성을 나타내는 객체간 관계 보조 정보 및, 개개의 객체 신호의 하나 이상의 개개의 속성을 나타내는 개개의 객체 보조 정보를 포함한다. 상술한 바와 같이, 이와 같은 오디오 비트스트림은 다중 채널 오디오 신호의 재구성을 허용하며, 렌더링 매개 변수의 부적절한 설정에 의해 유발된 가청 왜곡은 인식되어 감소되거나 심지어 제거될 수 있다.

[0059] 본 발명에 따른 다른 실시예는 상술한 방법을 구현하기 위한 컴퓨터 프로그램을 생성한다.

도면의 간단한 설명

[0060] 후속하여, 본 발명에 따른 실시예들이 첨부한 도면을 참조로 기술될 것이다.

도 1은 다운믹스 신호 표현 및 객체 관련 파라메트릭 정보에 기초하여 업믹스 신호 표현을 제공하기 위해 하나

이상의 조정된 매개 변수를 제공하는 장치의 개략적인 블록도를 도시한 것이다.

도 2는 본 발명의 실시예에 따른 MPEG SAOC 시스템의 개략적인 블록도를 도시한 것이다.

도 3은 본 발명의 다른 실시예에 따른 MPEG SAOC 시스템의 개략적인 블록도를 도시한 것이다.

도 4는 다운믹스 신호 및 믹스 신호에 대한 객체 신호의 기여의 개략적 표현을 도시한 것이다.

도 5a는 본 발명의 실시예에 따른 모노 다운믹스 기반 SAOC 대 MPEG 서라운드 트랜스코더의 개략적인 블록도를 도시한 것이다.

도 5b는 본 발명의 실시예에 따른 스테레오 다운믹스 기반 SAOC 대 MPEG 서라운드 트랜스코더의 개략적인 블록도를 도시한 것이다.

도 6은 본 발명의 실시예에 따른 오디오 신호 인코더의 개략적인 블록도를 도시한 것이다.

도 7은 본 발명의 실시예에 따른 오디오 비트스트림의 개략적인 블록도를 도시한 것이다.

도 8은 기준 MPEG SAOC 시스템의 개략적인 블록도를 도시한 것이다.

도 9a는 별도의 디코더 및 믹서를 이용한 기준 MPEG SAOC 시스템의 개략적인 블록도를 도시한 것이다.

도 9b는 통합된 디코더 및 믹서를 이용한 기준 MPEG SAOC 시스템의 개략적인 블록도를 도시한 것이다.

도 9c는 SAOC 대 MPEG 트랜스코더를 이용한 기준 MPEG SAOC 시스템의 개략적인 블록도를 도시한 것이다.

발명을 실시하기 위한 구체적인 내용

[0061]

1. 도 1에 따라 하나 이상의 조정된 매개 변수를 제공하는 장치

[0062]

다음에는 다운믹스 신호 표현 및 객체 관련 파라메트릭 정보에 기초하여 업믹스 신호 표현을 제공하기 위해 하나 이상의 조정된 매개 변수를 제공하는 장치(100)가 도 1을 참조로 기술될 것이다. 도 1은 하나 이상의 입력 매개 변수(110)를 수신하도록 구성되는 그러한 장치(100)의 개략적인 블록도를 도시한 것이다. 입력 매개 변수(110)는 예컨대 원하는 렌더링 매개 변수일 수 있다. 장치(100)는 또한 이에 기초하여 하나 이상의 조정된 매개 변수(120)를 제공하도록 구성된다. 조정된 매개 변수는 예컨대 조정된 렌더링 매개 변수일 수 있다. 장치(100)는 객체 관련 파라메트릭 정보(130)를 수신하도록 더 구성된다. 객체 관련 파라메트릭 정보(130)는 예컨대 객체 레벨차 정보 및/또는 다수의 객체를 나타내는 객체간 상관 정보일 수 있다. 장치(100)는 하나 이상의 입력 매개 변수(110)를 수신하여, 이에 기초하여, 하나 이상의 조정된 매개 변수(120)를 제공하도록 구성되는 매개 변수 조정기(140)를 포함한다. 매개 변수 조정기(140)는 하나 이상의 입력 매개 변수(110) 및 객체 관련 파라메트릭 정보(130)에 따라 하나 이상의 조정된 매개 변수(120)를 제공하도록 구성됨으로써, 다운믹스 신호 표현 및 객체 관련 파라메트릭 정보(130)에 기초하여 업믹스 신호 표현을 제공하는 장치에서 최적화되지 않은 매개 변수(예컨대, 하나 이상의 입력 매개 변수(110))의 사용으로 유발되는 업믹스 신호 표현의 왜곡은 미리 정해진 편차 이상만큼 적어도 최적의 매개 변수에서 벗어난 입력 매개 변수(110)에 대해 감소된다.

[0063]

따라서, 장치(100)는 하나 이상의 입력 매개 변수(110)를 수신하여, 이에 기초하여, 하나 이상의 조정된 매개 변수(120)를 제공한다. 장치(100)는, 명시적 또는 암시적으로, 하나 이상의 입력 매개 변수(110)가 다운믹스 신호 표현 및 객체 관련 파라메트릭 정보(130)에 기초하여 업믹스 신호 표현의 제공을 제어하는데 이용되었을 경우에 하나 이상의 입력 매개 변수(110)의 변경 사용이 수락할 수 없는 높은 왜곡을 유발시키는지 판단한다. 따라서, 적어도 하나 이상의 입력 매개 변수(110)가 유익하지 않은 방식으로 선택될 경우에, 조정된 매개 변수(120)는 전형적으로 하나 이상의 입력 매개 변수(110)보다 업믹스 신호 표현의 제공을 위한 그러한 장치를 조정하는데 더 적합하다.

[0064]

따라서, 장치(100)는 전형적으로 하나 이상의 조정된 매개 변수(120)에 따라 업믹스 신호 표현 제공기에 의해 제공되는 업믹스 신호 표현의 지각 인상을 개선한다. 하나 이상의 조정된 매개 변수를 도출하도록 하나 이상의 입력 매개 변수의 조정을 위한 객체 관련 파라메트릭 정보의 사용은 양호한 결과를 가져오는 것으로 발견되었는데, 그 이유는 하나 이상의 조정된 매개 변수(120)는 객체 관련 파라메트릭 정보(130)에 대응하지만, 객체 관련 파라메트릭 정보(130)에 대한 원하는 관계를 방해하는(violate) 매개 변수는 전형적으로 가청 왜곡을 생성시키기 때문이다. 객체 관련 파라메트릭 정보는 예컨대 하나 이상의 다운믹스 신호에 대한 (다수의 오디오 객체로부터의) 객체 신호의 기여를 나타내는 다운믹스 매개 변수를 포함할 수 있다. 객체 관련 파라메트릭 정보는 또한, 대안적으로 또는 부가적으로, 객체 신호의 특성을 나타내는 객체 레벨차 매개 변수 및/또는 객체간 상관 매개

변수를 포함할 수 있다. 객체 신호의 인코더측 처리를 나타내는 매개 변수 및 오디오 객체 자신의 특성을 나타내는 매개 변수의 양방은 매개 변수 조정기(120)에 의해 사용을 위한 유용한 정보로 간주될 수 있는 것으로 발견되었다. 그러나, 다른 객체 관련 파라메트릭 정보(130)는 대안적으로 또는 부가적으로 장치(100)에 의해 이용될 수 있다.

[0065] 그러나, 매개 변수 조정기(140)는 하나 이상의 입력 매개 변수(110)에 기초하여 하나 이상의 조정된 매개 변수(120)를 제공하기 위해 부가적인 정보를 이용할 수 있다. 예컨대, 매개 변수 조정기(140)는 선택적으로 다운믹스 계수, 하나 이상의 다운믹스 신호 또는 어떤 부가적인 정보를 평가하여, 하나 이상의 조정된 매개 변수(120)의 제공을 더욱 개선할 수 있다.

[0066] 2. 도 2에 따른 시스템

[0067] 다음에는, 도 2의 MPEG SAOC 시스템(200)이 상세히 설명될 것이다.

[0068] MPEG SAOC 시스템(200)의 양호한 이해를 제공하기 위해, 원하는 시스템 사양 및 설계 고려에 대한 개요가 주어질 것이다. 그 다음, 시스템의 구조적 개요가 주어질 것이다. 더욱이, 다수의 SAOC 왜곡 매트릭이 논의되고, 왜곡의 제한을 위한 이들 SAOC 왜곡 매트릭의 적용이 설명될 것이다. 부가적으로, 시스템(200)의 추가적 확장이 논의될 것이다.

[0069] 2.1 시스템 설계 고려

[0070] 상술한 바와 같이, 다수의 오디오 객체를 포함하는 오디오 장면의 비트레이트 효율적인 전송/저장을 위한 파라메트릭 기술은 전형적으로 전송 비트레이트 및 계산 복잡도의 양방의 관점에서 효율적일 수 있다. 수신단에서 이와 같은 시스템의 사용자에게 대한 추가적 이점은 자신의 선택(모노, 스테레오, 서라운드, 가상 헤드폰 재생 등)의 렌더링 설정을 선택하는 자유 및 사용자 상호 작용의 특징: 렌더링 매트릭스를 포함하여, 출력 장면이 뜻, 개인 선호 또는 다른 기준에 따라 상호 작용하게 설정되고 변경될 수 있다. 예컨대, 다른 잔여 토크와의 구별을 최대화하기 위해 한 공간 영역에 한 그룹으로부터의 토크를 함께 위치시킬 수 있다. 이러한 상호 작용은 디코더 사용자 인터페이스를 제공함으로써 달성된다.

[0071] 각 전송된 사운드 객체에 대해, 그의 상대 레벨 및 (비모노 렌더링에 대해) 렌더링하는 공간 위치가 조정될 수 있다. 이것은 사용자가 관련된 그래픽 사용자 인터페이스(GUI) 슬라이더의 위치(예컨대: 객체 레벨 = +5dB, 객체 위치 = -30deg)를 변경할 시에 실시간으로 발생할 수 있다. 그러나, 다운믹스 분리/믹스 기반 매개 변수 접근법으로 인해, 렌더링된 오디오 출력의 주관적 품질(subjective quality)은 렌더링 매개 변수 설정에 의존하는 것으로 발견되었다. 상대 객체 레벨의 변경은 공간 렌더링 위치의 변경보다 더 최종 오디오 품질에 영향을 미치는("리패닝(re-panning)") 것으로 발견되었다. 또한, 상대 매개 변수에 대한 극단적인 설정(extreme settings)(예컨대, +20dB)은 출력 품질을 수락할 수 없게 할 수 있는 것으로 발견되었다. 이것이 간단히 이러한 기법을 기본화하는 지각 가정의 일부를 방해하는 결과이지만, 그것은 사용자 인터페이스에 대한 설정에 따라 나쁜 사운드 및 아티팩트를 생성하는 상업적 제품에 대해 여전히 수락할 수 없다. 따라서, 예컨대 시스템(200)과 같이 본 발명에 따른 실시예들은 사용자 인터페이스의 설정과 무관하게 수락할 수 없는 저하를 회피하는 문제를 다룬다(사용자 인터페이스의 설정은 "입력 매개 변수"로 간주될 수 있다).

[0072] 다음에는, SAOC 왜곡을 회피하는 접근법에 관한 어떤 상세 사항이 논의될 것이다. 여기에서 제시되는 SAOC 왜곡 제한에 대한 접근법은 다음의 개념에 기초한다.

[0073] • 현저한 SAOC의 왜곡은 (입력 매개 변수로 간주될 수 있는) 렌더링 계수의 부적절한 선택에 대해 나타난다. 이러한 선택은 보통 (예컨대, 상호 작용 애플리케이션을 위한 실시간 그래픽 사용자 인터페이스(GUI)를 통해) 사용자에게 의해 상호 작용 방식으로 행해진다. 그래서, 부가적인 처리 단계가 도입되어, 사용자에게 의해 공급된 렌더링 계수를 수정하여 (예컨대, 이들을 어떤 계산에 기초하여 제한하여), SAOC 렌더링 엔진에 대한 이들 수정된 계수를 이용한다. 예컨대, 사용자에게 의해 공급된 렌더링 계수는 입력 매개 변수로 간주될 수 있고, SAOC 렌더링 엔진에 대한 수정된 계수는 수정된 매개 변수로 간주될 수 있다.

[0074] • 생성된 SAOC 오디오 출력의 과도한 저하를 제어하기 위해, (또한 왜곡 측정 DM으로 명시되는) 지각 저하의 계산 측정을 개발하는 것이 바람직하다. 이러한 왜곡 측정은 어떤 기준을 충족해야 하는 것으로 발견되었다:

[0075] o 왜곡 측정은 SAOC 디코딩 엔진의 내부 매개 변수로부터 쉽게 계산할 수 있어야 한다. 예컨대, 왜곡 측정을 획득하기 위해 여부 필터뱅크 계산은 필요치 않는 것이 바람직하다.

- [0076] o 왜곡 측정 값은 주관적으로 지각된 사운드 품질과 상관 관계(지각 저하)가 있어야 하며, 즉 심리 음향의 기초와 일치해야 한다. 이를 위해, 왜곡 측정의 계산은 바람직하게는 지각적 오디오 코딩 및 처리로부터 일반적으로 알려져 있는 바와 같이 주파수 선택 방식으로 행해질 수 있다.
- [0077] 다수의 SAOC 왜곡 측정이 규정되어 계산될 수 있는 것으로 발견되었다. 그러나, SAOC 왜곡 측정은 바람직하게는 렌더링된 SAOC 품질의 정확한 평가가 되기 위해서는 특정 기본적인 요인을 고려하여, 종종 (그러나 필요치는 않지만) 어떤 공통성을 가져야 하는 것으로 발견되었다:
- [0078] • 이들은 다운믹스 계수를 고려한다. 이들은 하나 이상의 다운믹스 신호 내의 각 오디오 객체의 상대 믹싱 부분(relative mixing fractions)을 결정한다. 배경 정보로서, 발생하는 SAOC 왜곡은 다운믹스와 렌더링 계수 사이의 관계에 의존하는 것으로 발견되었음에 주목되어야한다: 렌더링 계수에 의해 규정되는 상대 객체 기여가 다운믹스 내의 상대 객체 기여와 실질적으로 상이하면, (수정된 매개 변수를 이용하는) SAOC 디코딩 엔진은 다운믹스 신호의 상당한 조정을 수행하여 그것을 렌더링된 출력으로 변환시킬 필요가 있다. 이것은 SAOC 왜곡을 생성시키는 것으로 발견되었다.
- [0079] • 이들은 렌더링 계수를 고려한다. 이들은 하나 이상의 렌더링된 출력 신호의 각각에 대한 각 오디오 객체의 상대 출력 강도를 결정한다. 배경 정보로서, 발생하는 SAOC 왜곡은 또한 서로에 대한 객체 파워의 관계에 의존하는 것으로 발견되었음에 주목되어야한다. 시간적으로 어떤 지점에서의 객체가 다른 객체보다 많이 높은 파워를 가질 경우(및 이러한 객체의 다운믹스 계수가 너무 작지 않을 경우), 이러한 객체는 다운믹스보다 우위를 차지하여, 렌더링된 출력 신호에서 매우 잘 재생된다. 대조적으로, 약한 객체는 다운믹스에서만 매우 약하게 표현되어, 상당한 왜곡 없이 높은 출력 레벨까지 가져올 수 없다.
- [0080] • 이들은 다른 객체에 관한 각 객체의 (상대) 객체 파워/레벨을 고려한다. 이러한 정보는 예컨대 SAOC 객체 레벨차(OLDs)로서 설명된다. 배경 정보로서, 발생하는 SAOC 왜곡은 개개의 객체 신호의 속성에 더 의존하는 것으로 발견되었음에 주목되어야한다. 일례로서, 보다 고 레벨로 렌더링된 출력의 음조 본성(tonal nature)의 객체를 부스팅하는 것(반면에, 다른 객체는 더욱 많은 노이즈형 본성일 수 있음)은 상당한 지각 왜곡을 생성할 것이다.
- [0081] • 이 외에, 원래의 객체 신호의 속성에 관한 다른 정보가 고려될 수 있다. 그 후, 이들은 SAOC 보조 정보의 부분으로서 SAOC 인코더에 의해 전송될 수 있다. 예컨대, 각 객체 항목의 노이즈 또는 음조에 관한 정보는 SAOC 보조 정보의 부분으로서 전송되어, 왜곡 제한을 위해 이용될 수 있다
- [0082] 2.2 시스템 개요
- [0083] 상기 고려에 기초하여, MPEG SAOC 시스템(200)에 대한 개요가 이제 본 발명의 양호한 이해를 위해 주어질 것이다. 도 2에 따른 SAOC 시스템(200)은 도 8에 따른 MPEG SAOC 시스템(800)의 확장 버전이기 때문에, 상술한 것이 또는 적용하는 것에 주목되어야 한다. 더욱이, MPEG SAOC 시스템(200)은 도 9a, 9b 및 9c에 도시된 구현 대안(900, 930, 960)에 따라 수정될 수 있음에 주목되어야 하고, 객체 인코더는 SAOC 인코더에 대응하고, 사용자 상호 작용 정보/사용자 제어 정보(822)는 렌더링 제어 정보/렌더링 계수에 대응한다.
- [0084] 더욱이, MPEG SAOC 시스템(100)의 SAOC 디코더는 별도의 객체 디코더 및 믹서/렌더러 장치(920), 통합 객체 디코더 및 믹서/렌더러 장치(930) 또는 SAOC 대 MPEG 서라운드 트랜스코더(980)로 대체될 수 있다.
- [0085] 이제 도 2를 참조하면, MPEG SAOC 시스템(200)은 SAOC 인코더(210)를 포함하는 것으로 보여질 수 있으며, SAOC 인코더(210)는 1 내지 N개의 수를 가진 다수의 객체와 관련되는 다수의 객체 신호(x_1 내지 x_N)를 수신하도록 구성된다. SAOC 인코더(210)는 또한 다운믹스 계수(d_1 내지 d_N)를 수신하도록 (또는 그렇지 않으면 획득하도록) 구성된다. 예컨대, SAOC 인코더(210)는 SAOC 인코더(210)에 의해 제공되는 다운믹스 신호(212)의 각 채널에 대한 다운믹스 계수(d_1 내지 d_N) 중 한 세트를 획득할 수 있다. SAOC 인코더(210)는 예컨대 객체 신호(x_1 내지 x_N)의 가중된 조합을 획득하여 다운믹스 신호를 획득하도록 구성될 수 있으며, 객체 신호(x_1 내지 x_N)의 각각은 그의 관련된 다운믹스 계수(d_1 내지 d_N)로 가중된다. SAOC 인코더(210)는 또한 서로 다른 객체 신호 사이의 관계를 나타내는 객체간 관계 정보를 획득하도록 구성된다. 예컨대, 객체간 관계 정보는 예컨대 OLD 매개 변수의 형식의 객체 레벨차 정보 및, 예컨대 IOC 매개 변수의 형식의 객체간 상관 관계 정보를 포함한다. 따라서, 그 후, SAOC 인코더(200)는 하나 이상의 다운믹스 신호(212)를 제공하도록 구성되며, 이 신호의 각각은 각각의 다운믹스 신

호(또는 다중 채널 다운믹스 신호(212)의 채널)에 관련되는 다운믹스 매개 변수의 세트에 따라 가중되는 하나 이상의 객체 신호의 가중된 조합을 포함한다. SAOC 인코더(210)는 또한 보조 정보(214)를 제공하도록 구성되며, 보조 정보(214)는 (예컨대, 객체 레벨차 매개 변수 및 객체간 상관 관계 매개 변수의 형식의) 객체간 관계 정보를 포함한다. 보조 정보(214)는 또한, 예컨대, 다운믹스 이득 매개 변수 및 다운믹스 채널 레벨차 매개 변수의 형식의) 다운믹스 매개 변수 정보를 포함한다. 보조 정보(214)는 개개의 객체 속성을 나타낼 수 있는 선택적 객체 속성 보조 정보를 더 포함할 수 있다. 선택적 객체 속성 보조 정보에 관한 상세 사항은 아래에서 논의될 것이다.

[0086] MPEG SAOC 시스템(200)은 또한 SAOC 디코더(820)의 기능을 포함할 수 있는 SAOC 디코더(220)를 포함한다. 따라서, SAOC 디코더(220)는 하나 이상의 다운믹스 신호(212) 및 보조 정보(214) 뿐만 아니라 수정된 (또는 "조정된" 또는 "실제") 렌더링 계수(222)를 수신하여, 이에 기초하여, 하나 이상의 업믹스 채널 신호(\hat{y}_1 내지 \hat{y}_{N_c})를 제공한다.

[0087] MPEG SAOC 시스템(200)은 또한, 하나 이상의 입력 매개 변수, 즉 렌더링 제어 정보 또는 렌더링 계수(242)를 나타내는 입력 매개 변수에 따라 하나 이상의 수정된 (또는 "조정된" 또는 "실제") 매개 변수, 즉 수정된 렌더링 계수(222)를 제공하는 장치(240)를 포함한다. 장치(240)는 또한 보조 정보(214)의 적어도 일부를 수신하도록 구성된다. 예컨대, 장치(240)는 객체 파워(예컨대, 객체 신호(x_1 내지 x_N)의 파워)를 나타내는 매개 변수(214a)를 수신하도록 구성된다. 예컨대, 매개 변수(214a)는 (또한 OLDs로 명시되는) 객체 레벨차 매개 변수를 포함할 수 있다. 장치(240)는 또한 바람직하게는 다운믹스 계수를 나타내는 보조 정보(214)의 매개 변수(214b)를 수신한다. 예컨대, 매개 변수(214b)는 다운믹스 계수(d_1 내지 d_N)를 나타낸다. 선택적으로, 장치(240)는 개개의 객체 속성 보조 정보를 구성하는 부가적인 매개 변수(214c)를 더 수신할 수 있다.

[0088] 장치(240)는 일반적으로 (예컨대, 사용자 인터페이스로부터 수신될 수 있거나, 예컨대, 사용자 입력에 따라 계산되어 사전 설정된 정보로서 제공될 수 있는) 입력 렌더링 계수(242)에 기초하여 수정된 렌더링 계수(222)를 제공하도록 구성됨으로써, SAOC 디코더(220)에 의해 최적화되지 않은 렌더링 매개 변수의 사용으로 유발되는 업믹스 신호 표현의 왜곡은 감소된다. 환언하면, 수정된 렌더링 계수(222)는 입력 렌더링 계수(242)의 수정된 버전이며, 매개 변수(214a 및 214b)에 따라 변경이 행해짐으로써, (업믹스 신호 표현을 형성하는) 업믹스 채널 신호(\hat{y}_1 내지 \hat{y}_{N_c})의 모든 가청 왜곡이 감소되거나 제한되도록 한다.

[0089] 하나 이상의 조정된 매개 변수(242)를 제공하는 장치(240)는 예컨대 입력 렌더링 계수(242)를 수신하여, 이에 기초하여 수정된 렌더링 계수(222)를 제공하는 렌더링 계수 조정기(250)를 포함할 수 있다. 이를 위해, 렌더링 계수 조정기(250)는 입력 렌더링 계수(242)의 사용에 의해 유발되는 왜곡을 나타내는 왜곡 측정(252)을 수신할 수 있다. 왜곡 측정(252)은 예컨대 매개 변수(214a, 214b) 및 입력 렌더링 계수(242)에 따라 왜곡 계산기(260)에 의해 제공될 수 있다.

[0090] 그러나, 렌더링 계수 조정기(250) 및 왜곡 계산기(260)의 기능은 또한 단일 기능 유닛에 통합될 수 있음으로써, 수정된 렌더링 계수(222)가 왜곡 측정(252)의 명시적인 계산 없이 제공되도록 한다. 오히려, 왜곡 측정을 감소시키거나 제한하는 암시적인 메카니즘이 적용될 수 있다.

[0091] MPEG SAOC 시스템(200)의 기능에 관해, 업믹스 채널 신호(\hat{y}_1 내지 \hat{y}_{N_c})의 형식의 출력인 업믹스 신호 표현은 양호한 지각적 품질로 생성되는 것에 주목되어야 하는데, 그 이유는, 기존 시스템(800)에서 사용자 상호 작용 정보/사용자 제어 정보(822)의 부적절한 선택에 의해 유발된 가청 왜곡이 렌더링 계수의 수정 또는 조정에 의해 방지되기 때문이다. 수정 또는 조정이 장치(240)에 의해 수행됨으로써, 지각 인상의 심각한 저하가 방지되거나, 지각 인상의 저하가 입력 렌더링 계수(242)가 SAOC 디코더(220)에 의해 (수정 또는 조정 없이) 직접 이용되는 경우와 비교했을 때 적어도 감소된다.

[0092] 다음에는, 발명의 개념의 기능이 간략히 요약될 것이다. 왜곡 측정(DM)이 주어지면, 주어진 신호에 대한 왜곡 측정 값을 계산하고, SAOC 디코딩 알고리즘을 수정하여 (실제 이용된 렌더링 계수(212)를 제한하여) 오디오 출력의 과도한 왜곡이 방지될 수 있음으로써, 왜곡 측정 값은 어떤 임계치를 초과하지 않는다. 이러한 개념에 따른 시스템(200)은 도 2에 도시되고, 상기의 일부 상세 사항에 설명되었다.

[0093] 시스템(200)에 관해, 다음의 의견이 행해질 수 있다:

- [0094] • 원하는 렌더링 계수(242)는 사용자 또는 다른 인터페이스에 의해 입력된다.
- [0095] • SAOC 디코딩 엔진(220)에 적용되기 전에, 렌더링 계수(242)는 렌더링 계수 조정기(250)에 의해 수정되고, 렌더링 계수 조정기(250)는 왜곡 계산기(260)로부터 공급되는 하나 이상의 계산된 왜곡 측정(252)을 이용한다.
- [0096] • 왜곡 계산기(260)는 보조 정보(214)(예컨대, 상대 객체 파워/OLDS, 다운믹스 계수, 및 선택적으로 객체 신호 속성 정보)로부터 정보(예컨대, 매개 변수(214a, 214b))를 평가한다. 추가적으로, 그것은 원하는 렌더링 계수 입력(242)에 기초한다.
- [0097] 바람직한 실시예에서, 장치(240)는 왜곡 측정에 기초하여 렌더링 계수를 수정하도록 구성된다. 바람직하게는, 렌더링 계수는 예컨대 주파수 선택 가중치를 이용하여 주파수 선택 방식으로 조정된다.
- [0098] 렌더링 계수의 수정은 이러한 프레임(예컨대, 현재 프레임)에 기초할 수 있거나, 렌더링 계수가 프레임 바이 프레임 기준(frame-by-frame basis)으로 시간이 지남에 따라 조정될 수 있을 뿐만 아니라, 시간이 지남에 따라 처리/제어(예컨대, 시간이 지남에 따라 평활)될 수 있으며, 아마도 서로 다른 공격/감쇠(attack/decay) 시간 상수는 동적 범위 압축기/리미터에 대해 동일하게 적용될 수 있다.
- [0099] 일부 실시예에서, 왜곡 측정은 주파수 선택적일 수 있다.
- [0100] 일부 실시예에서, 왜곡 측정은 다음의 특성 중 하나 이상을 고려할 수 있다:
- [0101] • 각 객체의 파워/에너지/레벨;
- [0102] • 다운믹스 계수;
- [0103] • 렌더링 계수; 및/또는
- [0104] • 적용 가능하다면, 추가적인 객체 속성 보조 정보.
- [0105] 일부 실시예에서, 왜곡 측정은 객체마다 계산될 수 있고, 전체 왜곡에 도달하도록 조합될 수 있다.
- [0106] 일부 실시예에서, 추가적인 객체 속성 보조 정보(214c)는 선택적으로 평가될 수 있다. 추가적인 객체 속성 보조 정보(214c)는 예컨대 SAOC 인코더(210) 내에서의 향상된 SAOC 인코더에서 추출될 수 있다. 추가적인 객체 속성 보조 정보는 예컨대 향상된 SAOC 비트스트림에 삽입될 수 있으며, 이는 도 7을 참조로 설명될 것이다. 또한, 추가적인 객체 속성 보조 정보는 향상된 SAOC 디코더에 의해 왜곡 제한을 위해 이용될 수 있다.
- [0107] 특별한 경우에, 노이즈/음조는 추가적인 객체 속성 보조 정보에 의해 나타내는 객체 속성으로 이용될 수 있다. 이러한 경우에, 노이즈/음조는 보조 정보에 저장하기 위해 다른 객체 매개 변수(예컨대, OLDS)보다 훨씬 더 거친 주파수 해상도로 전송될 수 있다. 극단적인 경우에, 노이즈/음조 객체 속성 보조 정보는 객체마다 하나의 정보(예컨대, 광대역 특성)로 전송될 수 있다.
- [0108] 2.3 SAOC 왜곡 메트릭
- [0109] 다음에는, 다수의 서로 다른 왜곡 측정이 기술될 것이며, 이는 예컨대 왜곡 계산기(260)를 이용하여 획득될 수 있다. 렌더링 계수의 제한을 위한 이들 왜곡 측정의 적용에 관한 상세 사항은 아래의 섹션 2.4에서 논의될 것이다.
- [0110] 환언하면, 이 섹션은 수개의 왜곡 측정에 대해 설명한다. 이들은 개별적으로 이용될 수 있거나, 예컨대, 개별 왜곡 메트릭 값의 가중된 가산에 의해 복합물(compound), 더욱 복잡한 왜곡 메트릭을 형성하도록 조합될 수 있다. 여기서, 용어 "왜곡 측정" 및 "왜곡 메트릭"은 유사한 수량을 명시하며, 대부분의 경우에 구별될 필요가 없음에 주목되어야 한다.
- [0111] 다음에는, 다수의 왜곡 메트릭이 설명될 것이며, 이는 왜곡 계산기(260)에 의해 평가될 수 있고, 입력 렌더링 계수(242)에 기초하여 수정된 렌더링 계수(222)를 획득하기 위해 렌더링 계수 조정기(250)에 의해 이용될 수 있다.
- [0112] 2.3.1 왜곡 측정 # 1
- [0113] 다음에는, (또한 왜곡 측정 #.1에 명시된) 제 1 왜곡 측정이 설명될 것이다.

[0114] 개념적 단순성을 위해, N-1-1 SAOC 시스템(예컨대, 모노 다운믹스 신호(212) 및 단일 업믹스 채널(신호))이 고려될 것이다. N 입력 오디오 객체는 모노 신호로 다운믹스되고, 모노 출력으로 렌더링된다. 도 8에 제공된 바와 같이, 다운믹스 계수는 $D_1 \dots D_N$ 으로 나타내고, 렌더링 계수는 $r_1 \dots r_N$ 으로 나타낸다. 다음의 식에서, 시간 인덱스는 단순성을 생략되었다. 마찬가지로, 주파수 인덱스는 생략되며, 식이 서브밴드 신호에 관계함에 주목한다. 아래의 식의 일부에서, 소문자는 계수 또는 신호를 나타내고, 대문자는 대응하는 파워를 나타내며, 이는 식의 문맥에서 알 수 있다. 또한, 신호는 때때로 시간 도메인에서 보다는 대응하는 시간 주파수 도메인 계수로 나타내는 것에 주목되어야 한다.

[0115] 객체 #m(가칭 객체 인덱스 m)가 관심 객체, 예컨대, 상대 레벨에서 증가되어, 전체 사운드 품질을 제한하는 가장 우세한 객체임을 가정한다. 그 후, 이상적인 원하는 출력 신호(업믹스 채널 신호)가 다음에 의해 주어진다.

$$\hat{y}_i = [x_m \cdot r_m] + [\sum_{i=1; i \neq m}^N x_i \cdot r_i] \quad (1)$$

[0116] 여기서, 제 1 항은 출력 신호에 대한 관심 객체의 원하는 기여인 반면에, 제 2 항은 모든 다른 객체("간섭")로부터의 기여를 나타낸다.

[0117]

[0118] 그러나, 사실상, 다운믹스 프로세스로 인해, 출력 신호는 다음에 의해 주어진다.

$$y_i = t \cdot \sum_{i=1}^N x_i \cdot d_i = [x_m \cdot t \cdot d_m] + [\sum_{i=1; i \neq m}^N x_i \cdot t \cdot d_i] \quad (2)$$

[0119] 즉, 다운믹스 신호는 후속하여 MPEG 서라운드 디코더에서 "m2"에 대응하는 트랜스코딩 계수 t만큼 스케일링된다. 다시말하면, 이것은 제 1 항(출력 신호에 대한 객체 신호의 실제 기여) 및 제 2 항(다른 객체 신호에 의한 실제 "간섭")으로 분할될 수 있다. 여기서, SAOC 시스템(예컨대, SAOC 디코더(220) 및 선택적으로 또한 장치(240))은 트랜스코딩 계수 t를 동적으로 결정함으로써, 실제 렌더링된 출력 신호의 파워가 이상적 신호의 파워와 일치되도록 한다.

$$\hat{Y}_1 = Y_1 \Rightarrow t^2 = \frac{\sum_{i=1}^N r_i^2 \cdot X_i}{\sum_{i=1}^N d_i^2 \cdot X_i} \quad (3)$$

[0121] 왜곡 측정(DM)은 객체 #m의 이상적인 전원 기여와 그의 실제 파워 기여 사이의 관계를 계산하여 정의될 수 있다:

$$dm_1(m) = \frac{P_{ideal}}{P_{actual}} = \frac{r_m^2}{d_m^2 \cdot t^2} = \frac{r_m^2 \cdot \sum_{i=1}^N d_i^2 \cdot X_i}{d_m^2 \cdot \sum_{i=1}^N r_i^2 \cdot X_i} \quad (4)$$

[0123] 여기서, $\sum_{i=1}^N r_i^2 \cdot X_i$ 은 최종 렌더링된 신호의 파워를 나타내고, $\sum_{i=1}^N d_i^2 \cdot X_i$ 은 다운믹스 신호의 파워이다. 실제 구현에서, X_i 값은 SAOC 보조 정보(214)의 부분으로 전송되는 대응하는 OLDi(Object Level Difference) 값으로 직접 대체될 수 있다.

[0125] dm_1 의 양호한 해석을 위해, 그 정의는 다음과 같이 다시 공식화될 수 있다:

$$dm_1(m) = \frac{r_m^2 \cdot \sum_{i=1}^N d_i^2 \cdot X_i}{d_m^2 \cdot \sum_{i=1}^N r_i^2 \cdot X_i} = \frac{\frac{r_m^2 \cdot X_m}{\sum_{i=1}^N r_i^2 \cdot X_i}}{\frac{d_m^2 \cdot X_m}{\sum_{i=1}^N d_i^2 \cdot X_i}} \quad (4a)$$

[0126]

[0127] 효과적으로, 이것은 왜곡 메트릭이 이상적으로 렌더링된(출력) 신호 대 다운믹스(입력) 신호의 상대 객체 파워 기여의 비인 것을 의미한다. 이것은 SAOC 기법이 많은 요인에 의해 상대 객체 파워를 변경할 필요가 없을 때에 가장 잘 작동한다는 발견을 가져온다.

[0128] dm_1 값의 증가는 사운드 객체 #m에 대한 사운드 품질의 감소를 나타낸다. 모든 렌더링 계수가 일반적인 요인에 의해 스케일링되거나, 모든 다운믹스 계수가 마찬가지로 스케일링될 경우에 dm_1 의 값이 일정하게 유지하는 것으로 발견되었다. 또한, 객체 #m에 대한 렌더링 계수의 증가(그 상대 레벨의 증가)는 왜곡을 증가시키는 것으로 발견되었다. dm_1 의 값은 다음과 같이 해석될 수 있다:

- [0129] • 1의 값은 객체 #m에 대한 이상적 품질을 나타내고;
- [0130] • 1 이상의 dm_1 값의 증가는 품질의 감소를 나타내며;
- [0131] • 1 이하의 dm_1 의 값은 객체 #m에 대한 품질을 더 개선하지 않는다.

[0132] 결과적으로, 사운드 장면 품질(즉, 모든 객체에 대한 품질)의 전체 측정은 다음과 같이 계산될 수 있다:

$$DM_1 = \frac{\sum_{m=1}^N w(m) \cdot \max[dm_1(m), 1]}{\sum_{m=1}^N w(m)} \quad (5)$$

[0133]

[0134] 이 식에서, $w(m)$ 은 오디오 장면 내의 특정 객체의 중요성 및 민감도에 관한 객체 #m의 가중치를 나타낸다. 일례로서, 그 후, $w(m)$ 은 객체 파워/라우드니스(loudness) $w(m) = (r_m^2 X_m)^\alpha$ 에 따라 선택되며, 여기서, α 는 이러한 객체의 심리 음향 라우드니스 성장을 대충 에뮬레이트(emulate)하는 0.25로 선택될 수 있다. 더욱이, $w(m)$ 은 음조 및 마스킹 현상을 고려한다. 대안적으로, $w(m)$ 은 1로 설정되어, DM_1 의 계산을 용이하게 할 수 있다.

[0135] 2.3.2 왜곡 측정 #2

[0136] 대체 왜곡 측정은 식(4)로부터 개시하여 NMR(Noise-to-Mask-Ratio)의 스타일로 지각 측정을 형성하도록, 즉 노이즈/간섭 및 마스킹 임계값 사이의 관계를 계산하도록 구성될 수 있다:

$$dm_2(m) = \frac{P_{Noise}}{Mask} = \frac{P_{ideal} - P_{actual}}{msr \cdot P_{total}} = \frac{(r_m^2 - d_m^2 \cdot t^2) \cdot X_m}{msr \cdot \sum_{i=1}^N r_i^2 \cdot X_i} = \frac{\left(r_m^2 \cdot \sum_{i=1}^N d_i^2 \cdot X_i - d_m^2 \cdot \sum_{i=1}^N r_i^2 \cdot X_i \right) \cdot X_m}{msr \cdot \left(\sum_{i=1}^N r_i^2 \cdot X_i \right) \cdot \left(\sum_{i=1}^N d_i^2 \cdot X_i \right)} \quad (6)$$

[0137]

[0138] 이 식에서, msr 은 음조에 의존하는 전체 오디오 신호의 Noise-to-Mask-Ratio이다. dm_2 의 값의 증가는 사운드 객체 #m에 대한 더욱 높은 왜곡을 나타낸다. 다시 말하면, 모든 렌더링 계수가 일반적인 요인에 의해 스케일링되거나, 모든 다운믹스 계수가 마찬가지로 스케일링될 경우에 dm_2 의 값은 일정하게 유지한다. dm_2 의 값 범위는 다음과 같이 해석될 수 있다:

- [0139] • 0의 값은 객체 #m에 대한 이상적 품질을 나타내고;

[0140] • 1 이상의 dm_2 값의 증가는 점진적 가청 저하를 나타내며;

[0141] • 1 이하의 dm_2 의 값은 객체 #m에 대한 불분명한 품질을 나타낸다.

[0142] 결과적으로, 사운드 장면 품질(즉, 모든 객체에 대한 품질)의 전체 측정은 다음과 같이 계산될 수 있다:

$$DM_2 = \frac{\sum_{m=1}^N w(m) \cdot \max[dm_2(m), 1]}{\sum_{m=1}^N w(m)} \quad (7)$$

[0143]

[0144] 다시 말하면, $w(m)$ 은 오디오 장면 내의 특정 객체의 중요성/레벨/라우드니스에 관한 객체 #m의 가중치를 나타내며, 전형적으로 $w(m) = (r_m^2 X_m)^\alpha$ 로 선택되며, $\alpha = 0.25$ 이다.

[0145] 식(6) 상의 왜곡 측정은 파워의 차로서 왜곡을 계산한다(이것은 "스펙트럼 차를 가진 NMR" 측정에 대응한다). 대안적으로, 왜곡은 부가적 믹스된 곱의 항(additional mixed product term)을 포함하는 다음의 측정에 이르는 과형 기초로 계산될 수 있다:

$$dm_2'(m) = \frac{P_{Noise}}{Mask} = \frac{E\{|y_{m,ideal} - \hat{y}_{m,actual}|^2\}}{msr \cdot P_{total}} = \frac{r_m^2 \cdot \sum_{i=1}^N d_i^2 \cdot X_i + d_m^2 \cdot \sum_{i=1}^N r_i^2 \cdot X_i - 2 \cdot d_m r_m \cdot \sqrt{(\sum_{i=1}^N r_i^2 \cdot X_i) \cdot (\sum_{i=1}^N d_i^2 \cdot X_i)}}{msr \cdot (\sum_{i=1}^N r_i^2 \cdot X_i) \cdot (\sum_{i=1}^N d_i^2 \cdot X_i)} \quad (8)$$

[0146]

[0147] 2.3.3 왜곡 측정 #3

[0148] 다운믹스 신호와 렌더링된 신호 사이의 코히어런스(coherence)를 나타내는 제 3 왜곡 측정이 제공된다. 더욱 높은 코히어런스는 더 양호한 주관적 사운드 품질을 생성한다. 부가적으로, IOC 데이터가 SAOC 디코더에 제공될 경우에 입력 오디오 객체의 상관 관계가 고려될 수 있다.

[0149] SAOC 매개 변수(예컨대, 객체 레벨차 매개 변수 및 객체간 상관 관계 매개 변수를 포함할 수 있는 매개 변수 (214a))로부터, 객체 공분산(covariance)의 모델이 결정될 수 있다

[0150] $\mathbf{E} = \sqrt{\mathbf{OLD}^T \cdot \mathbf{OLD} \cdot \mathbf{IOC}}$

[0151] 왜곡 측정을 계산하기 위해, 렌더(render) 및 다운믹스 계수를 포함하는 매트릭스 \mathbf{M} 은 어셈블된다(\mathbf{M} 은 N-1-2 SAOC 시스템에 대한 렌더링 매트릭스로서 해석될 수 있다)

[0152] $\mathbf{M} = \begin{pmatrix} r_1 & r_2 & \cdots & r_N \\ d_1 & d_2 & \cdots & d_N \end{pmatrix}$

[0153] 다운믹스와 렌더링된 신호 \mathbf{C} 사이의 공분산은 이때 다음과 같다

[0154] $\mathbf{C} = \mathbf{M} \cdot \mathbf{E} \cdot \mathbf{M}^* = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix}$

[0155] 왜곡 측정 DM_3 는 다음과 같이 정의된다

$$DM_3 = 1 - \min\left(\frac{|c_{12}|}{\sqrt{c_{11} \cdot c_{22}}}, 1\right)$$

[0156]

[0157] DM_3 의 값은 다음과 같이 해석될 수 있다:

- [0158] • 값은 범위 [0 .. 1]내에 있고, 다운믹스와 렌더링된 신호 사이의 공분산을 나타낸다.
- [0159] • 0의 값은 이상적 품질을 나타낸다.
- [0160] • DM_3 의 값의 증가는 품질의 감소를 나타낸다.

[0161] 2.3.4 왜곡 측정 #4

[0162] 2.3.4.1 개요

[0163] 이러한 접근법은 (주어진 다운믹스 DMX 로부터 계산되는) 타겟 렌더링 에너지(UPMIX)와 최적의 다운믹스 에너지 사이의 평균 가중 비율을 왜곡 측정으로 이용하도록 제안한다.

[0164] 상세 사항에 대해, 또한 도 4에 대한 참조가 행해지며, 도 4는 다운믹스(DMX), 최적의 다운믹스 에너지(DMX_{opt}) 및 타겟 렌더링 에너지(UPMIX)의 그래픽 표현을 도시한다.

[0165] 2.3.4.2 명명법(Nomenclature)

- [0166] $ch = \{1, 2, \dots, N_{ch}\}$ 업믹스 채널에 대한 인덱스
- [0167] $dx = \{1, 2\}$ 다운믹스 채널에 대한 인덱스
- [0168] $ob = \{1, 2, \dots, N_{ob}\}$ 오디오 객체에 대한 인덱스
- [0169] $pb = \{1, 2, \dots, N_{pb}\}$ 매개 변수 대역에 대한 인덱스
- [0170] $r_{ch, ob, pb} = r(ch, ob, pb)$ 채널 ch , 오디오 객체 ob 및 매개 변수 대역 pb 에 대한 렌더링 매트릭스
- [0171] $d_{dx, ob, pb} = d(dx, ob, pb)$ 다운믹스 채널 dx , 오디오 객체 ob 및 매개 변수 대역 pb 에 대한 다운믹스 매트릭스
- [0172] $w_{ob, pb} = w(ob, pb)$ 매개 변수 대역 pb 에 대한 오디오 객체 ob 의 중요성/레벨/라우드니스를 나타내는 가중치
- [0173] $NRG_{pb} = NRG(pb)$ 주파수 대역 pb 에 대한 최고 에너지를 가진 오디오 객체의 절대 객체 에너지
- [0174] $OLD_{ob, pb} = OLD(ob, pb)$ 한 오디오 객체 ob 와, 대응하는 주파수 대역 pb 에 대한 최고 에너지를 가진 객체 사이의 강도 차를 나타내는 객체 레벨 차
- [0175] $IOC_{obi, obj, pb} = IOC(ob_i, ob_j, pb)$ 오디오 객체의 두 채널 사이의 상관 관계를 나타내는 객체간 상관 관계.

[0176] 2.3.4.3 알고리즘

[0177] 왜곡 측정 #4을 획득하기 위한 알고리즘의 단계는 다음에서 간략히 설명될 것이다:

[0178] • 업믹스 및 다운믹스 상대 에너지의 계산:

[0179]
$$\hat{r}_{ch,ob,pb}^2 = OLD_{ob,pb} \cdot r_{ch,ob,pb}^2, \quad \hat{d}_{dx,ob,pb}^2 = OLD_{ob,pb} \cdot d_{dx,ob,pb}^2.$$

[0180] •
$$\sum_{ob=1}^{N_{ob}} \tilde{r}_{ch,ob,pb}^2 = 1 \quad \text{및} \quad \sum_{ob=1}^{N_{ob}} \tilde{d}_{dm,ob,pb}^2 = 1$$
 이도록 에너지의 정규화:

[0181]
$$\tilde{r}_{ch,ob,pb}^2 = \frac{\hat{r}_{ch,ob,pb}^2}{\sum_{ob=1}^{N_{ob}} \hat{r}_{ch,ob,pb}^2}, \quad \tilde{d}_{dm,ob,pb}^2 = \frac{\hat{d}_{dm,ob,pb}^2}{\sum_{ob=1}^{N_{ob}} \hat{d}_{dm,ob,pb}^2}.$$

[0182] • 각 업믹스 채널 및 대역에 대한 최적의 다운믹스 $d_{ch,ob,pb}^{2(opt)}$ 의 구성

[0183]
$$d_{ch,ob,pb}^{2(opt)} = \alpha_{ch,ob,pb} \cdot \tilde{d}_{1,ob,pb}^2 + \beta_{ch,ob,pb} \cdot \tilde{d}_{2,ob,pb}^2.$$

[0184] 곱셈 상수(multiplicative constants) $\alpha_{ch,ob,pb}, \beta_{ch,ob,pb}$ 는 선형 식의 오버디파인 시스템

(overdefined system)을 풀이하여 다음 조건:
$$\left\| d_{ch,ob,pb}^{2(opt)} - \tilde{r}_{ch,ob,pb}^2 \right\| \xrightarrow{\alpha, \beta} 0$$
 을 충족하도록 계산된다.

[0185] • 왜곡 측정의 계산:

[0186]
$$DM_4 = \sum_{ob=1}^{N_{ob}} \sum_{ch=1}^{N_{ch}} \left| 1 - \frac{\tilde{r}_{ch,ob,pb}^2}{d_{ch,ob,pb}^{2(opt)}} \right| w_{ob,pb} \hat{r}_{ch,ob,pb}^2$$

[0187] 2.3.4.4 왜곡 제어

[0188] 왜곡 제어는 왜곡 측정 DM4에 따라 하나 이상의 렌더링 계수(들)를 제한함으로써 달성된다.

[0189] (i) 측정은 스테레오 다운믹스 케이스에만 관련이 있고, (2) 그것은 #dx=1 및 #ch=1에 대한 DM1로 감소될 수 있다.

[0190] 2.3.4.5 속성

[0191] 다음에는, 왜곡 측정 수 4를 계산하기 위한 개념의 속성이 간략히 요약될 것이다. 개념은,

- [0192] • 이상적 트랜스코딩을 추정하고
- [0193] • 스테레오 다운믹스를 처리할 수 있으며;
- [0194] • 다중 채널 렌더링에 대한 일반화를 허용한다.

[0195] 2.3.5 왜곡 측정 #5

[0196] 트랜스코딩 계수 t의 대체 계산이 제시된다. 그것은 t의 확장으로서 해석될 수 있고, 객체간 코히어런스(IOC)의 통합을 특징으로 하는 트랜스코딩 매트릭스 T에 이르고, 동시에 현재 매트릭 DM#1 및 DM#2를 스테레오 다운믹스 및 다중 채널 업믹스로 확장한다. 트랜스코딩 계수 t의 현재 구현은 실제 렌더링된 출력 신호의 파워를 이상적 렌더링된 신호의 파워에 일치시키는 것을 고려한다. 즉,

$$t^2 = \frac{\sum_{i=1}^N r_i^2 X_i}{\sum_{i=1}^N d_i^2 X_i}$$

[0197]

[0198] 공분산 매트릭스 E의 통합은 객체간 코히어런스를 고려하는 t, 즉 트랜스코딩 매트릭스 T에 대한 수정된 공식화를 산출한다. E의 요소는 다음과 같은 SAOC 매개 변수(214)로부터 계산된다.

$$e_{ij} = \sqrt{OLD_i OLD_j} IOC_{ij}$$

[0199]

[0200] 트랜스코딩 매트릭스는 $TDx \approx Rx$ 이도록 렌더링된 출력 신호로의 다운믹스의 변환을 나타낸다. 그것은 평균 계급 오차의 최소화를 통해 획득되고, 다음을 산출한다:

$$T = RED^* (DED^*)^{-1}$$

[0201]

$$H = RED^* \quad \text{또는} \quad h_{ij} = \sum_{l=1}^N \sum_{m=1}^N r_{il} d_{jm} e_{lm}$$

[0202]

$$\text{및 } V = DED^* \quad \text{또는} \quad v_{ij} = \sum_{l=1}^N \sum_{m=1}^N d_{il} d_{jm} e_{lm}$$

[0203]

[0204] dm_1 의 스타일로 하지만 지금은 객체 m의 모든 다운믹스/렌더링 조합(n,k)에 대한 왜곡 측정이 다음에 의해 주어진다:

$$dm_5^n(m, n, k) = \frac{r_{m,k}^2 v_{n,n}}{d_{m,n}^2 h_{k,n}}$$

[0205]

[0206] 좌측 및 우측 다운믹스 채널에 대한 $dm_1(m)$ 을 별도로 고려함으로써, 다음에 이른다:

$$dm_L(m, k) = \frac{r_{m,k}^2 v_{1,1}}{d_{m,1}^2 h_{k,1}} \quad \text{및} \quad dm_R(m, k) = \frac{r_{m,k}^2 v_{2,2}}{d_{m,2}^2 h_{k,2}}$$

[0207]

[0208] 두 다운믹스/업믹스 경로 중 더 양호한 것이 렌더링된 출력의 품질에 관련이 있어, 측정은 최소값에 상응하는 것으로 추정될 수 있다, 즉

$$dm_5'(m, k) = \min[dm_L, dm_R]$$

[0209]

[0210] 인덱스 k로 명시되는 모든 출력 채널의 전체 측정은 다음과 같이 계산될 수 있다:

$$dm_5(m) = \frac{\sum_{k=1}^{N_{Ch}} dm_5'(m, k) r_{m,k}^2 X_m}{\sum_{k=1}^{N_{Ch}} r_{m,k}^2 e_{k,k}}$$

[0211]

[0212] 모든 객체에 대한 전체 측정은 다음에 의해 획득될 수 있다:

$$DM_5 = \frac{\sum_{m=1}^N w(m) \max[dm_5(m), 1]}{\sum_{m=1}^N w(m)} \quad \text{전과 같이} \quad w(m) = [r_m^2 X_m]^\alpha$$

[0213]

[0214] t 내지 T의 유사한 확장은 dm_2 및 dm'_2 에 대해 가능하다.

[0215] 2.3.6 왜곡 측정 #6

[0216] 다음에는, 제 6 왜곡 측정이 설명될 것이다.

[0217] $e_i(t)$ 를 객체 신호 #i의 제곱 힐베르트 엔벨로프(squared Hilbert envelope)이고, P_i 를 객체 신호 #i의 파워(양자 모두 전형적으로 서브대역내에 있음)인 것으로 하면, 음조/노이즈형의 측정 N은 다음과 같은 힐베르트 엔벨로프의 정규화된 분산 추정(normalized variance estimate)으로부터 획득될 수 있다:

$$N_i = \frac{\text{var}\{e_i\}}{P_i^2}$$

[0218]

[0219] 대안적으로, 또한 힐베르트 엔벨로프 차 신호의 파워/분산은 힐베르트 엔벨로프 자체의 분산 대신에 이용될 수 있다. 하여튼, 측정은 시간이 지남에 따라 엔벨로프 파동의 강도를 나타낸다.

[0220] 이러한 음조/노이즈형 측정 N은 이상적으로 렌더링된 신호 혼합 및 실제 SAOC 렌더링된 사운드 혼합의 양방에 대해 결정될 수 있고, 왜곡 측정은 양방의 차, 예컨대 다음으로부터 계산될 수 있다:

$$DM_6 = |N_{ideal} - N_{actual}|^\beta$$

[0221]

[0222] 여기서, β 는 매개 변수이다(예컨대, $\beta = 2$).

[0223] 2.3.7. 기준 장면 및 SAOC 렌더링된 장면에 대한 소스 신호 이미지의 에너지를 계산

[0224] 왜곡 측정에 이용되는 기준 및 SAOC 렌더링된 장면 내의 소스 이미지의 객체 에너지를 계산하기 위해서는, 그것이 "왜곡 측정 5"에서 행해질 시에 SAOC 렌더링된 장면에 대한 트랜스코딩 매트릭스 T뿐만 아니라, 기준 장면 및 렌더링된 장면 양방에 대한 소스 신호의 상관 관계를 고려할 필요가 있다.

[0225] 주의: 대문자의 신호의 표기법은 여기서 신호의 매트릭스 표기법을 반영하고, 이전 챕터에서와 같이 신호의 에너지를 반영하지 않는다.

[0226] 임의 소스 x_m 에 대해, 모든 소스 x_i 내의 x_m 의 신호 부분은 다음과 같이 계산될 수 있다:

[0227] 모든 소스 신호 x_i 를 관심 객체 x_m 에 상관된 신호 부분 $x_{i\parallel m}$ 및 x_m 에 상관되지 않는 부분 $x_{i\perp m}$ 으로 분할한다.

이것은 모든 신호 x_i , 즉 $x_i = x_{i\parallel m} + x_{i\perp m}$ 로의 x_m 의 부공간 투영(subspace projection)에 의해 행해질 수 있다. 상관된 부분은 다음에 의해 주어진다:

$$x_{i\parallel m} = \frac{x_m^T x_i}{x_m^T x_m} x_m = \frac{IOC_{i,m}}{\|x_m\|^2} x_m = g_{i,m} x_m$$

[0228]

[0229] 2.3.7.1 기준 장면 y 내의 소스 x_m 의 이미지로부터 P_{ideal, x_m} 를 계산

[0230] $Y = RX$ 및 $X = X_{\perp m} + X_{\parallel m}$, 모든 렌더링된 채널에 대한 소스 x_m 의 이미지 y_{x_m} 는 $Y_{x_m} = RX_{\parallel m}$ 를 통해 계산될 수 있다. 여기서,

[0231]

$$X_{\parallel m} = \begin{pmatrix} x_{1\parallel m}^T \\ x_{2\parallel m}^T \\ \vdots \\ x_{N\parallel m}^T \end{pmatrix} = \begin{pmatrix} g_{1,m} x_m^T \\ g_{2,m} x_m^T \\ \vdots \\ g_{N,m} x_m^T \end{pmatrix}$$

[0232]

Y_{x_m} 는 다음에 의해 계산될 수 있다:

[0233]

$$Y_{x_m} = R X_{\parallel m} = \begin{pmatrix} r_{ch_1, x_1} & r_{ch_1, x_2} & \vdots & r_{ch_1, x_N} \\ r_{ch_2, x_1} & r_{ch_2, x_2} & \vdots & r_{ch_2, x_N} \\ \dots & \dots & \ddots & r_{N_{ch-1}, x_N} \\ r_{N_{ch}, x_1} & r_{N_{ch}, x_2} & r_{N_{ch}, x_{n-1}} & r_{N_{ch}, x_N} \end{pmatrix} \begin{pmatrix} g_{1,m} x_m^T \\ g_{2,m} x_m^T \\ \vdots \\ g_{N,m} x_m^T \end{pmatrix}$$

[0234]

그래서, 기준 장면 내의 소스 이미지 Y_{x_m} 의 에너지 P_{ideal, x_m} 는 다음과 같을 것이다:

[0235]

$$P_{ideal, x_m} = \begin{pmatrix} \|r_{ch_1, x_1} g_{1,m} + r_{ch_1, x_2} g_{2,m} + \dots + r_{ch_1, x_N} g_{N,m}\|^2 \|x_m\|^2 \\ \dots \\ \|r_{N_{ch}, x_1} g_{1,m} + r_{N_{ch}, x_2} g_{2,m} + \dots + r_{N_{ch}, x_N} g_{N,m}\|^2 \|x_m\|^2 \end{pmatrix}$$

[0236]

2.3.7.2 SAOC 렌더링된 장면 \hat{y} 내의 소스 \hat{y}_{x_m} 의 이미지로부터 P_{actual, x_m} 를 계산

[0237]

이것은 P_{actual, x_m} 에 대해서와 동일한 방식으로 행해질 수 있다. 렌더링된 장면 내의 모든 채널에 대해 T에 의한 트랜스코딩 매트릭스 및 D에 의한 다운믹스 매트릭스 \hat{y}_{x_m} 는 다음과 같을 것이다:

[0238]

$$\hat{Y}_{x_m} = T^{0.5} D X_{\parallel m}$$

[0239]

$$D = \begin{pmatrix} d_{11} & \dots & d_{1N} \\ d_{21} & \dots & d_{2N} \end{pmatrix} \quad \text{및} \quad T = \begin{pmatrix} t_{11} & t_{12} \\ \vdots & \vdots \\ t_{N_{ch}1} & t_{N_{ch}2} \end{pmatrix} \quad \text{를 이용하여}$$

[0240]

$$\hat{Y}_{x_m} = \begin{pmatrix} \sqrt{t_{11}} d_{11} + \sqrt{t_{12}} d_{21} & \sqrt{t_{11}} d_{12} + \sqrt{t_{12}} d_{22} & \dots & \sqrt{t_{11}} d_{1N} + \sqrt{t_{12}} d_{2N} \\ \sqrt{t_{21}} d_{11} + \sqrt{t_{22}} d_{21} & \sqrt{t_{21}} d_{12} + \sqrt{t_{22}} d_{22} & \dots & \sqrt{t_{21}} d_{1N} + \sqrt{t_{22}} d_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \sqrt{t_{N_{ch}1}} d_{11} + \sqrt{t_{N_{ch}2}} d_{21} & \sqrt{t_{N_{ch}1}} d_{12} + \sqrt{t_{N_{ch}2}} d_{22} & \dots & \sqrt{t_{N_{ch}1}} d_{1N} + \sqrt{t_{N_{ch}2}} d_{2N} \end{pmatrix} \begin{pmatrix} g_{1,m} x_m^T \\ g_{2,m} x_m^T \\ \vdots \\ g_{N,m} x_m^T \end{pmatrix}$$

[0241] 그래서, 기준 장면 내의 소스 이미지 \hat{Y}_{x_m} 의 에너지 P_{actual, x_m} 는 다음과 같을 것이다:

$$P_{actual, x_m} = \left(\begin{array}{c} \left\| g_{1,m}(\sqrt{t_{11}}d_{11} + \sqrt{t_{12}}d_{21}) + g_{2,m}(\sqrt{t_{11}}d_{12} + \sqrt{t_{12}}d_{22}) + \dots + g_{N,m}(\sqrt{t_{11}}d_{1N} + \sqrt{t_{12}}d_{2N}) \right\|^2 \|x_m\|^2 \\ \dots \\ \left\| g_{1,m}(\sqrt{t_{N_{ch1}}}d_{11} + \sqrt{t_{N_{ch2}}}d_{21}) + g_{2,m}(\sqrt{t_{N_{ch1}}}d_{12} + \sqrt{t_{N_{ch2}}}d_{22}) + \dots + g_{N,m}(\sqrt{t_{N_{ch1}}}d_{1N} + \sqrt{t_{N_{ch2}}}d_{2N}) \right\|^2 \|x_m\|^2 \end{array} \right)$$

[0243] 2.3.7.3. 왜곡 측정을 계산

[0244] dm_1 의 스타일의 왜곡 측정은 모든 객체 m 에 대해 계산되어 렌더링 채널 k 를 다음으로서 출력할 수 있다:

$$dm_7'(m, k) = \frac{P_{ideal}}{P_{actual}} = \frac{\|r_{k1}IOC_{1m} + \dots + r_{kN}IOC_{Nm}\|^2}{\left\| (\sqrt{t_{k1}}d_{11} + \sqrt{t_{k2}}d_{21})IOC_{1m} + \dots + (\sqrt{t_{k1}}d_{1N} + \sqrt{t_{k2}}d_{2N})IOC_{Nm} \right\|^2}$$

$$dm_7(m) = \frac{\sum_{k=1}^{N_{ch}} dm_7'(m, k) r_{m,k}^2 \|x_m\|^2}{\sum_{k=1}^{N_{ch}} r_{m,k}^2 e_{k,k}}$$

$$DM_7 = \frac{\sum_{m=1}^N w(m) \max[dm_7(m), 1]}{\sum_{m=1}^N w(m)} \quad \text{전과 같이} \quad w(m) = [r_m^2 X_m]^\alpha$$

[0247] 2.3.8 객체 신호 속성

[0248] 다음에는, 객체 신호 속성의 예가 설명되며, 이는 예컨대 왜곡 측정을 획득하기 위해 장치(250) 또는 아티팩트 감소부(artifact reduction)(320)에 의해 이용될 수 있다.

[0249] SAOC 처리에서, 여러 개의 오디오 객체 신호는 최종 렌더링된 출력을 생성하는데 이용되는 다운믹스 신호로 다운믹스된다. 음조 객체 신호가 동일한 신호 파워의 더욱 많은 노이즈형의 제 2 객체 신호와 믹싱되면, 그 결과는 노이즈형으로 되는 경향이 있다. 제 2 객체 신호가 더욱 높은 파워를 가질 경우에는 동일하게 유지된다. 단지, 제 2 객체 신호가 실질적으로 제 1 객체 신호보다 낮은 파워를 가질 경우에는 그 결과가 음조되는 경향이 있다. 동일한 방식으로, 렌더링된 SAOC 출력 신호의 음조/노이즈형은 대부분 적용된 렌더링 계수와 무관하게 다운믹스 신호의 음조/노이즈형에 의해 결정된다. 양호한 주관적 출력 품질을 달성하기 위해, 또한 실제 렌더링된 신호의 음조/노이즈형은 이상적으로 렌더링된 신호의 음조/노이즈형에 가까워야 한다. 왜곡 측정에서 이러한 개념을 이용하기 위해, 비트스트림의 부분으로서 각 객체의 음조/노이즈형에 관한 정보를 전송할 필요가 있다. 그 후, 이상적으로 렌더링된 출력의 음조/노이즈형 N 은 SAOC 디코더에서 각 객체 N_i 의 음조/노이즈형 및 그 객체 파워 P_i 의 함수로서 추정될 수 있다. 즉

[0250] $N = f(N_1, P_1, N_2, P_2, N_3, P_3, \dots)$

[0251] 이상적으로 렌더링된 출력의 음조/노이즈형 N 은 왜곡 측정을 계산하기 위해 실제 렌더링된 출력 신호의 음조/노이즈형과 비교된다. 일례로서, 다음의 함수 $f()$ 가 이용될 수 있다:

$$N = \frac{\sum_i N_i \cdot P_i^\alpha}{(\sum_i P_i)^\alpha}$$

[0252] 이는 객체 음조/노이즈형 값 및 객체 파워를 신호의 혼합의 음조/노이즈형 값을 추정하는 단일 출력으로 조합한다. 매개 변수 α 는 주어진 음조/노이즈형 측정(예컨대, $\alpha=2$)에 대한 추정 절차의 정확도를 최적화하도록 선택

될 수 있다. 음조/노이즈형에 기초한 적절한 왜곡 메트릭은 왜곡 측정 #6으로서 섹션 2.3.6에서 설명된다.

[0254] 2.4 왜곡 제한 기법

[0255] 2.4.1 왜곡 제한 기법의 개요

[0256] 다음에는, 다수의 왜곡 제한 기법의 짧은 개요가 주어질 것이다. 상술한 바와 같이, 렌더링 계수 조정기(250)는 입력 렌더링 계수(242)를 수신하여, 이에 기초하여, SAOC 디코더(220)에 의해 이용하기 위한 수정된 렌더링 계수(222)를 제공한다.

[0257] 수정된 렌더링 계수를 제공하기 위한 서로 다른 개념은 구별될 수 있으며, 이 개념은 또한 일부 실시예에서 조합될 수 있다. 제 1 개념에 따르면, 하나 이상의 렌더링 매개 변수 제한값은 보조 정보(214)의 하나 이상의 매개 변수에 따라(즉, 객체 관련 파라메트릭 정보(214)에 따라) 제 1 단계에서 획득된다. 그 다음, 실제 "(수정된 또는 조정된)" 렌더링 계수(222)는 원하는 렌더링 매개 변수(242) 및 하나 이상의 렌더링 매개 변수 제한값에 따라 획득됨으로써, 실제 렌더링 매개 변수가 렌더링 매개 변수 제한값으로 규정된 한계치에 따르도록 한다. 따라서, 렌더링 매개 변수 제한값을 초과하는 그런 렌더링 매개 변수는 렌더링 매개 변수 제한값에 따르도록 조정(수정)된다. 이러한 제 1 개념은 구현하기가 쉽지만, 때때로 약간 저하된 사용자 만족을 가져오는데, 그 이유는 원하는 렌더링 매개 변수(242)의 사용자의 선택이 사용자 규정된 원하는 렌더링 매개 변수(242)가 렌더링 매개 변수 제한값을 초과할 경우에는 고려되지 않기 때문이다.

[0258] 제 2 개념에 따르면, 매개 변수 조정기는 실제 렌더링 매개 변수를 획득하기 위해 원하는 렌더링 매개 변수의 제공과 최적의 렌더링 매개 변수의 제공 사이의 선형 조합을 계산한다. 이 경우에, 매개 변수 조정기는 미리 정해진 임계값 매개 변수 및 (상술한 바와 같은) 왜곡 메트릭에 따라 선형 조합에 대한 원하는 렌더링 매개 변수 및 최적의 렌더링 매개 변수의 기여를 결정하도록 구성된다.

[0259] 게다가, 왜곡 측정(왜곡 메트릭)이 객체간 관계 속성 및/또는 개개의 객체 속성을 이용하여 계산되는지가 구별될 수 있다. 일부 실시예에서는, 객체간 관계 속성만이 평가되지만, (단일 객체에만 관계되는) 개개의 객체 속성은 고려되지 않는다. 일부 다른 실시예에서는, 개개의 객체 속성만이 고려되지만, 객체간 관계 속성은 고려되지 않는다. 그러나, 일부 실시예에서는, 객체간 관계 속성 및 개개의 객체 속성의 양방의 조합이 평가된다.

[0260] 이전의 고려 및 또한 서로 다른 왜곡 측정에 대한 상기 논의에 기초하여, 다음의 부섹션에서 설명되는 바와 같이 왜곡을 제한하기 위한 많은 기법이 정의될 것이다. 왜곡을 제한하기 위한 이들 기법은 입력 렌더링 계수(242)에 따라 수정된 렌더링 계수를 획득하기 위해 렌더링 계수 조정기(250)에 의해 적용될 수 있다.

[0261] 2.4.2 왜곡 제한 기법 #1

[0262] 부섹션 2.3.1에서, 간단한 왜곡 측정이 객체 #m의 이상적 파워 기여와 그 실제 파워 기여 사이의 관계를 계산하여 정의되었다(식 4):

$$dm_1(m) = \frac{P_{ideal}}{P_{actual}} = \frac{r_m^2}{d_m^2 \cdot t^2} = \frac{r_m^2 \cdot \sum_{i=1}^N d_i^2 \cdot X_i}{d_m^2 \cdot \sum_{i=1}^N r_i^2 \cdot X_i} \quad (4)$$

[0263]

[0264] 이 식에서, SAOC 렌더러의 제어 하에 있는 유일한 변수는 트랜스코딩 프로세스에서 이용되는 렌더링 계수이다. 그래서, 생성된 왜곡 메트릭이 어떤 임계값 T를 초과하지 않으면, 이것은 대응하는 렌더링 매트릭스 계수에 조건을 부과한다:

$$dm_1(m) = \frac{r_m^2 \cdot \sum_{i=1}^N d_i^2 \cdot X_i}{d_m^2 \cdot \sum_{i=1}^N r_i^2 \cdot X_i} \leq T \Leftrightarrow r_m^2 \leq \hat{r}_m^2 = T \cdot \frac{d_m^2 \cdot \sum_{i=1, i \neq m}^N r_i^2 \cdot X_i}{\left| \sum_{i=1}^N d_i^2 \cdot X_i - T \cdot d_m^2 \cdot X_m \right|} \quad (6.1.a)$$

[0265]

[0266] 모든 \hat{r}_m^2 에 대한 솔루션을 찾기 위해, 한 세트의 선형 식 $\mathbf{Ax} = \mathbf{b}$ 이 설정될 수 있다.

$$\mathbf{x} = \begin{bmatrix} \hat{r}_1^2 \\ \hat{r}_2^2 \\ \vdots \\ \hat{r}_N^2 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ \sum_{i=1}^N r_i^2 \end{bmatrix} \quad \text{및} \quad \mathbf{A} = \begin{bmatrix} -c_1 & d_1^2 X_2 & \cdots & d_1^2 X_N \\ d_2^2 X_1 & -c_2 & \cdots & d_2^2 X_N \\ \vdots & \vdots & \ddots & \vdots \\ d_N^2 X_1 & d_N^2 X_2 & \cdots & -c_N \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

$$c_m = \frac{1}{T} \left(\sum_{i=1}^N d_i^2 \cdot X_i - T \cdot d_m^2 \cdot X_m \right)$$

A의 제 1 N 행은 식(6.1.a)로부터 직접 유도된다. 부가적으로, 제약 조건(constraint)은 새로운 (제한된) 렌더링 계수의 에너지가 사용자 특정 계수의 에너지와 동일하도록 부가된다. 그 후, (렌더링 매개 변수 제한값으로 간주될 수 있는) \hat{r}_m^2 에 대한 솔루션이 다음과 같이 획득된다:

$$\mathbf{x} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{b}$$

이것으로 시작하면, 제 1 단순한 왜곡 제한 기법이 다음과 같이 보여질 수 있다: 렌더링 매트릭스 계수(242)가 사용자 인터페이스로부터 SAOC 디코더에 제공될 시에 렌더링 매트릭스 계수(242)를 이용하는 대신에, 객체 #m에 대해 효과적으로 이용되는 렌더링 계수 r_m' , 222는 예컨대 SAOC 디코딩 프로세스에 이용되기 전에 프레임 기준으로 렌더링 계수 조정기(240)에 의해 수정/제한된다:

$$r_m'^2 = \min(r_m^2, \hat{r}_m^2)$$

제한 프로세스는 각 특정 프레임 내의 개개의 객체 에너지에 의존함에 주목한다. 이러한 접근법은 단순하고, 다음의 작은 결점을 갖는다:

- 그것은 상대 객체 라우드니스와 지각 마스킹을 고려하지 않는다.
- 그것은 특정 객체를 부스트(boost)하는 효과만을 캡처(capture)하지만, 객체 이득을 감소하여 효과를 캡처하지 않는다. 이것은 또한 dm 값의 하한(lower bound)을 지정하여 처리될 수 있다.

2.4.3 제한 기법 #2

2.4.3.1 제한 기법 개요

이 섹션은 다음의 양태를 고려하는 제한 함수를 설명한다:

- 왜곡 측정은 제한 임계치에 의해 한정되고,
- 제한된 렌더링 매트릭스의 유도는 제한 함수 및 초기 렌더링 매트릭스에 대한 거리에 기초한다.

이러한 제한 함수(또는 제한 기법)은 예컨대 왜곡 계산기(260)와 함께 렌더링 계수 조정기(250)에 의해 수행될 수 있다.

왜곡 측정은 렌더링 매트릭스의 함수이기 때문에,

- (예컨대, 입력 렌더링 계수(242)로 나타내는) 초기 렌더링 매트릭스는 초기 왜곡 측정을 산출하고,
- 최적의 왜곡 측정은 최적의 렌더링 매트릭스를 산출하지만, 초기 렌더링 매트릭스에 대한 이러한 최적의 렌더링 매트릭스의 거리는 최적일 아닐 수 있으며,
- 왜곡 측정은 초기 렌더링 매트릭스에 대한 렌더링 매트릭스의 거리에 비례하는 역선형이며,

[0286] • 어떤 임계치에 대해, (예컨대, 조정되거나 수정된 렌더링 계수(222)로 나타내는) 제한된 렌더링 매트릭스는 초기 및 최적 작업(working) 포인트 사이의 보간법(예컨대, 선형 보간법)을 통해 유도된다.

[0287] 부가적으로, 각 작업 포인트에서 렌더링된 신호의 파워는 다음과 같도록 거의 일정한 것으로 추정될 수 있다.

$$[0288] \sum_{i=1}^{N_{ob}} r_i^2 X_i \approx \sum_{i=1}^{N_{ob}} r_{lim,i}^2 X_i \approx \sum_{i=1}^{N_{ob}} r_{opt,i}^2 X_i$$

[0289] 제한 기법 #2은 다음에 논의되는 바와 같이 서로 다른 왜곡 측정과 함께 이용될 수 있다.

[0290] 2.4.3.2 왜곡 측정의 제한 #1

[0291] 각 매개 변수 대역에 대해, 관심 객체 m에 대한 왜곡 측정 $dm_1(m)$ 은 다음과 같이 정의된다:

$$[0292] dm_1(m) = \frac{r_m^2 \sum_{i=1}^{N_{ob}} d_i^2 X_i}{d_m^2 \sum_{i=1}^{N_{ob}} r_i^2 X_i}$$

[0293] 최적의 렌더링 매트릭스는 $dm_1(m)$ 을 최적의 값으로 설정할 때, 즉 $dm_{1,opt}(m)=1$ 일 때 생성한다.

$$[0294] r_{opt,m}^2 = d_m^2 \frac{\sum_{i=1}^{N_{ob}} r_i^2 X_i}{\sum_{i=1}^{N_{ob}} d_i^2 X_i}$$

[0295] 따라서, 최적의 렌더링 매트릭스 값 $r_{opt,m}^2$ 은 연립 방정식을 이용하여 획득될 수 있으며, 여기서, r_i^2 은 $r_{opt,i}^2$ 으로 대체된다.

[0296] $dm_1(m)$ 에 대한 사전 규정된 임계치 T로, 제한된 렌더링 매트릭스는 다음에 의해 주어진다:

$$[0297] r_{lim,m}^2 = \frac{T-1}{dm_1(m)} (r_m^2 - r_{opt,m}^2) + r_{opt,m}^2$$

[0298] 2.4.3.3 왜곡 측정의 제한 #2a

[0299] 또한 때때로 간단히 " $dm_{2a}(m)$ "으로 명시되는 왜곡 측정 $dm_{2a}(m)$ 은 다음과 같이 정의된다:

$$[0300] dm_{2a}(m) = \frac{\left(r_m^2 \sum_{i=1}^{N_{ob}} d_i^2 X_i - d_m^2 \sum_{i=1}^{N_{ob}} r_i^2 X_i \right) X_m}{msr \sum_{i=1}^{N_{ob}} r_i^2 X_i \sum_{i=1}^{N_{ob}} d_i^2 X_i} = \frac{\frac{r_m^2 X_m}{\sum_{i=1}^{N_{ob}} r_i^2 X_i} - \frac{d_m^2 X_m}{\sum_{i=1}^{N_{ob}} d_i^2 X_i}}{msr}$$

[0301] 객체 m 및 각 매개 변수 대역에 대해, 어떤 매개 변수 대역 pb에 대해, 마스크 대 신호비 $msr(pb)$ 는 렌더링된 신호의 파워의 함수이다.

$$[0302] msr(pb) = \left[\sum_{i=1}^{N_{ob}} r_i^2 X_i M_k \right]_{k=\max(pb)} = \left[\sum_{i=1}^{N_{ob}} r_i^2 X_i \right]_{k=\max(pb)} [M_k]_{k=\max(pb)}$$

[0303] 왜곡 측정에 대한 최적의 값은 0이다. 즉 $dm_{2a,opt}(m)=0$. 이것은 어떤 에러를 도입하지 않은 완전한 트랜스코딩

프로세스에 상응한다. 그래서, 최적의 렌더링 매트릭스는 다음을 산출한다:

$$r_{opt,m}^2 = d_m^2 \frac{\sum_{i=1}^{N_{ob}} r_i^2 X_i}{\sum_{i=1}^{N_{ob}} d_i^2 X_i}$$

[0304]

$dm_{2a}(m) = T$ 이면, 수정된 렌더링 계수(222)로 나타낼 수 있는 제한된 렌더링 매트릭스는 다음과 같이 된다:

$$r_{lim,m}^2 = \frac{T-1}{dm_{2a}(m)} (r_m^2 - r_{opt,m}^2) + r_{opt,m}^2$$

[0306]

2.4.3.4 왜곡 측정의 제한 #2b

[0307]

또한 때때로 간단히 " $dm_{2b}(m)$ "으로 명시되는 왜곡 측정 $dm_{2b}(m)$ 은 또한, 입력 렌더링 계수(242)에 따라 수정된 렌더링 계수(222)로 나타낼 수 있는 제한된 렌더링 매트릭스를 획득하기 위해 장치(240)에 의해 이용될 수 있다.

[0308]

2.4.3.5 왜곡 측정의 제한 #4

[0309]

왜곡 측정 $dm_4(m)$ 은 다음과 같이 정의된다:

[0310]

$$dm_4(m) = \left| 1 - \frac{r_m^2 \sum_{i=1}^{N_{ob}} d_i^2 X_i}{d_m^2 \sum_{i=1}^{N_{ob}} r_i^2 X_i} \right|$$

[0311]

객체 m 및 각 매개 변수 대역에 대해, 그의 최적의 값 $dm_{4,opt}(m)=0$. 따라서, 최적 및 제한된 렌더링 매트릭스는 다음을 생성한다:

[0312]

$$r_{opt,m}^2 = d_m^2 \frac{\sum_{i=1}^{N_{ob}} r_i^2 X_i}{\sum_{i=1}^{N_{ob}} d_i^2 X_i} \quad \text{및}$$

[0313]

$$r_{lim,m}^2 = \frac{T-1}{dm_4(m)} (r_m^2 - r_{opt,m}^2) + r_{opt,m}^2$$

[0314]

따라서, 장치(240)는 입력 렌더링 계수(242)에 따라 및 또한 제 4 왜곡 측정 $dm_4(m)$ 과 동일할 수 있는 왜곡 측정(252)에 따라 수정된 렌더링 계수(222)를 제공할 수 있다.

[0315]

2.4.4 제한 기법 #3

[0316]

식(6.1.a)에 상응하여, 객체 m 에 대한 제한된 렌더링 계수는 다음과 같이 왜곡 측정 #3에 대해 계산될 수 있다. 약어(abbreviations)로,

[0317]

$$c_1 = \sum_{i=1}^N \sum_{j=1}^N d_i d_j e_{ij}, \quad c_2 = \sum_{i=1, i \neq m}^N r_i e_{im}, \quad c_3 = \sum_{i=1, i \neq m}^N \sum_{j=1, j \neq m}^N r_i r_j e_{ij}, \quad c_4 = \sum_{i=1}^N d_i e_{mi}$$

[0318]

$$c_5 = \sum_{i=1, i \neq m}^N \sum_{j=1}^N r_i d_j e_{ij}$$

[0319] 및

[0320] 이차 방정식은 다음과 같이 설정된다:

$$\hat{r}_m^2 \left((1-T)^2 \cdot c_1 e_{mm} - c_4^2 \right) + \hat{r}_m \cdot 2 \cdot \left((1-T)^2 \cdot c_1 c_2 - c_4 c_5 \right) + (1-T)^2 \cdot c_1 c_3 - c_5^2 = a \cdot \hat{r}_m^2 + b \cdot \hat{r}_m + c = 0$$

[0321]

[0322] (포지티브) 솔루션은 다음과 같다:

$$\hat{r}_m = \frac{-b + \sqrt{b^2 - 4ac}}{2a} \quad (6.2.a)$$

[0323]

[0324] 따라서, 장치(240)는 렌더링 매개 변수 제한값 \hat{r}_m 을 포함할 수 있고, 상기 렌더링 매개 변수 제한값에 따라 조정 (또는 수정된) 렌더링 계수(222)를 제한할 수 있다.

[0325] 2.4.5 추가적인 선택적 개선

[0326] 장치(240)에 의해 개별적으로 또는 조합하여 수행되는 렌더링 계수(222)를 제한하기 위한 상술한 개념은 더 개선될 수 있다. 예컨대, M-채널 렌더링의 일반화가 수행될 수 있다. 이를 위해, 렌더링 계수의 제공/파워의 합은 단일 렌더링 계수 대신에 이용될 수 있다.

[0327] 또한, 스테레오 다운믹스에 대한 일반화가 수행될 수 있다. 이를 위해, 다운믹스 계수의 제공/파워의 합은 단일 다운믹스 계수 대신에 이용될 수 있다.

[0328] 일부 실시예에서, 왜곡 메트릭은 저하 제어(degradation control)에 이용되는 단일 주파수를 통해 하나로 조합될 수 있다. 대안적으로, 어떤 경우에는 각 주파수 대역에 대해 개별적으로 왜곡 제어를 하는 것이 더 좋을 수 있다(더 간단할 수 있다).

[0329] 서로 다른 개념들은 실제로 왜곡 제어를 행하기 위해 적용될 수 있다. 예컨대, 하나 이상의 렌더링 계수가 제한될 수 있다. 대안적으로 또는 부가적으로, (예컨대, MPEG 서라운드 디코딩의) m2 매트릭스 계수는 제한될 수 있다. 대안적으로 또는 부가적으로, 상대 객체 이득은 제한될 수 있다.

[0330] 3. 도 3에 따른 실시예

[0331] 다음에는, SAOC 디코더의 다른 실시예가 도 3을 참조로 설명될 것이다. 이해를 용이하게 하기 위해, 기본 고려에 대한 간단한 논의가 먼저 주어질 것이다. (ISO/IEC 23003-2로서의 표준화 하에) "공간 오디오 객체 코딩" (SAOC) 시스템의 출력은 오디오 객체의 속성 및, 렌더링 매트릭스와 다운믹스 매트릭스 사이의 관계에 의존하는 아티팩트를 나타낼 수 있다. 이러한 문제를 논의하기 위해, 다운믹스와 렌더링 매트릭스가 동일한 치수를 갖는 경우가 여기서 일반성의 손실 없이 고려된다. 상응하는 고려는 다운믹스 및 렌더링된 장면의 채널의 수가 서로 다를 경우에 적용한다.

[0332] 일반적으로, 아티팩트의 위험은 렌더링 매트릭스가 다운믹스 매트릭스와 상당히 다르게 될 때에 증가하는 것으로 발견되었다. 서로 다른 타입의 아티팩트는 구별될 수 있다.

[0333] 1. "효과적인" 렌더링 매트릭스가 SAOC 디코더로 입력되는 원하는 렌더링 매트릭스와 다른 (객체의 효과적 달성 감소 또는 이득이 렌더링 매트릭스에 지정된 것과 다른) 렌더링의 결점. 이것은 전형적으로 어떤 매개 변수 대역 내의 객체의 중복의 효과이다.

[0334] 2. 객체의 음색의 바람직하지 않은 및 아마 시간 변형 변경. 이 아티팩트는 특히 "누수(leakage)"가 1에서 언급될 때에 심각하다. 단지 단일 매개 변수 대역에 대해 국부적으로 발생한다.

[0335] 3. SAOC 디코더에서 시간 및 주파수 변형 신호 처리에 의해 유발되는 변조된 객체 신호, 음악적 음색, 또는 변조된 노이즈와 같은 아티팩트.

[0336] 모든 타입의 아티팩트를 최소화하는 것이 바람직한 것으로 발견되었다.

[0337] 이러한 문제를 처리하여, 아티팩트를 최소화하는 일반화된 접근법은 SAOC 디코더로 송신되기 전에 원하는 렌더링 매트릭스의 시간-주파수-변형 후처리를 사용하는 것이다. 이러한 접근법은 도 3에 도시된다.

- [0338] 도 3은 SAOC 디코더 장치(300)의 개략적인 블록도를 도시한다. SAOC 디코더(300)는 또한 간단히 오디오 신호 디코더로 명시될 수 있다. 오디오 신호 디코더(300)는 SAOC 디코더 코어(310)를 포함하며, SAOC 디코더 코어(310)는 다운믹스 신호 표현(312) 및 SAOC 비트스트림(314)을 수신하여, 이에 기초하여, 예컨대 다수의 업믹스 오디오 채널의 표현의 형식으로 렌더링된 장면에 대한 설명(316)을 제공하도록 구성된다.
- [0339] 오디오 신호 디코더(300)는 또한, 예컨대, 하나 이상의 입력 매개 변수에 따라 하나 이상의 조정된 매개 변수를 제공하는 장치의 형태로 제공될 수 있는 아티팩트 감소부(320)를 포함한다. 아티팩트 감소부(320)는 원하는 렌더링 매트릭스에 관한 정보(322)를 수신하도록 구성된다. 정보(322)는 예컨대 아티팩트 감소부의 입력 매개 변수를 형성할 수 있는 다수의 원하는 렌더링 매개 변수의 형식을 취할 수 있다. 아티팩트 감소부(320)는 다운믹스 신호 표현(312) 및 SAOC 비트스트림(314)을 수신하도록 더 구성되며, SAOC 비트스트림(314)은 객체 관련 파라메트릭 정보를 운반할 수 있다. 아티팩트 감소부(320)는 원하는 렌더링 매트릭스에 관한 정보(322)에 따라 (예컨대, 다수의 조정된 렌더링 매개 변수의 형식으로) 수정된 렌더링 매트릭스(324)를 제공하도록 더 구성된다.
- [0340] 결과적으로, SAOC 디코더 코어(310)는 다운믹스 신호 표현(312), SAOC 비트스트림(314) 및 수정된 렌더링 매트릭스(324)에 따라 렌더링된 장면의 표현(316)을 제공하도록 구성될 수 있다.
- [0341] 다음에는, 오디오 신호 디코더의 기능에 관한 일부 상세 사항이 제공될 것이다. 주어진 원하는 렌더링 매트릭스에 대한 SAOC 시스템의 잠재적으로 제한된 분리 능력 때문에 아티팩트의 위험을 평가하기 위해, (다운믹스 신호 표현(312)으로 나타내는) 다운믹스 신호 및 SAOC 비트스트림(314)의 양방을 고려하는 것이 바람직한 것으로 발견되었다. 항상 이용할 수 있는 이러한 정보로, 예컨대, 렌더링 매트릭스의 수정에 의해 이들 아티팩트를 완화하기를 시도할 수 있다. 이것은 아티팩트 감소부(320)에 의해 수행된다. 완화를 위한 고급 전략은 SAOC 시스템의 시간 및 주파수 선택성의 한계(중복) 및 지각 효과의 양방을 고려하는 것이다. 즉, 이들은 원하는 출력 신호와 유사한 렌더링 신호 사운드를 만들려고 하면서, 가청 아티팩트를 가능한 작게 해야 한다.
- [0342] 도 3에 도시된 오디오 신호 디코더(300)에 이용되는 아티팩트 감소를 위한 바람직한 접근법은 상술한 서로 다른 타입의 아티팩트를 평가하는 왜곡 측정의 가중된 조합인 전체 왜곡 측정에 기초한다. 이들 가중치는 상술한 서로 다른 타입의 아티팩트 사이에서 적절한 트레이드오프(tradeoff)를 결정한다. 이들 서로 다른 타입의 아티팩트에 대한 가중치는 SAOC 시스템이 이용되는 애플리케이션에 의존할 수 있음에 주목되어야 한다.
- [0343] 환언하면, 아티팩트 감소부(320)는 다수의 타입의 아티팩트에 대한 왜곡 측정을 획득하도록 구성될 수 있다. 예컨대, 아티팩트 감소부(320)는 상술한 왜곡 측정(dm_1 내지 dm_6)의 일부를 적용할 수 있다. 대안적으로 또는 부가적으로, 아티팩트 감소부(320)는 이 섹션에서 논의된 바와 같이 다른 타입의 아티팩트를 나타내는 추가적 왜곡 측정을 이용할 수 있다. 또한, 아티팩트 감소부는 (예컨대, 섹션 2.4.2, 2.4.3 및 2.4.4 하에) 상술된 왜곡 제한 기법 또는 비교할 만한 아티팩트 제한 기법 중 하나 이상을 이용하여 원하는 렌더링 매트릭스(242)에 기초하여 수정된 렌더링 매트릭스(324)를 획득하도록 구성될 수 있다.
- [0344] 4. 도 5a 및 5b에 따른 오디오 신호 트랜스코더
- [0345] 4.1 도 5a에 따른 오디오 신호 트랜스코더
- [0346] 상술한 개념은 오디오 신호 디코더 및 오디오 신호 트랜스코더의 양방에 적용될 수 있음에 주목되어야 한다. 도 2 및 3을 참조하면, 이 개념은 오디오 신호 디코더와 함께 설명되었다. 다음에는, 발명의 개념의 사용에 대해 오디오 신호 트랜스코더와 함께 간략히 논의될 것이다.
- [0347] 이 문제에 관해, 오디오 신호 디코더 및 오디오 신호 트랜스코더의 유사성이 이미 도 9a, 9b 및 9c와 관련하여 논의되어, 도 9a, 9b 및 9c에 대해 행해진 설명은 발명의 개념에 적용할 수 있음에 주목되어야 한다.
- [0348] 도 5a는 MPEG 서라운드 디코더(510)와 함께 오디오 신호 트랜스코더(500)의 개략적인 블록도를 도시한 것이다. 알 수 있는 바와 같이, SAOC 대 MPEG 서라운드 트랜스코더일 수 있는 오디오 신호 트랜스코더(500)는 SAOC 비트스트림(520)을 수신하여, 이에 기초하여, 다운믹스 신호 표현(524)에 영향을 미치지 않고 (또는 수정하지 않고) MPEG 서라운드 비트스트림(522)을 제공하도록 구성된다. 오디오 신호 트랜스코더(500)는 SAOC 비트스트림(520)을 수신하여, SAOC 비트스트림(530)으로부터 원하는 SAOC 매개 변수를 추출하도록 구성되는 SAOC 파싱(parsing)(530)을 포함한다. 오디오 신호 트랜스코더(500)는 또한 SAOC 파싱(530)에 의해 제공되는 SAOC 매개 변수를 수신하도록 구성되는 장면 렌더링 엔진(540) 및, 실제 렌더링(매트릭스) 정보로 간주될 수 있고, 예컨대, 다수의 조정된 (또는 수정된) 렌더링 매개 변수의 형식으로 표현될 수 있는 렌더링 매트릭스 정보(542)를 포함한다. 장면 렌더링 엔진(540)은 상기 SAOC 매개 변수 및 렌더링 매트릭스(542)에 따라 MPEG 서라운드

비트스트림(522)을 제공하도록 구성된다. 이를 위해, 장면 렌더링 엔진(540)은 (또한 파라메트릭 정보로서 명시되는) 채널 관련 매개 변수인 MPEG 서라운드 비트스트림 매개 변수(522)를 계산하도록 구성된다. 따라서, 장면 렌더링 엔진(540)은, 객체 관련 파라메트릭 정보를 구성하는 SAOC 비트스트림(520)의 매개 변수를, 실제 렌더링 매트릭스(542)에 따라 채널 관련 파라메트릭 정보를 구성하는 MPEG 서라운드 비트스트림의 매개 변수로 변환(또는 "트랜스코더")하도록 구성된다.

[0349] 오디오 신호 트랜스코더(500)는 또한, 예컨대, 재생 구성에 관한 정보(552) 및 객체 위치에 관한 정보(554)의 형식으로 원하는 렌더링 매트릭스에 관한 정보를 수신하도록 구성되는 렌더링 매트릭스 생성부(550)를 포함한다. 대안적으로, 렌더링 매트릭스 생성부(550)는 원하는 렌더링 매개 변수(예컨대, 렌더링 매트릭스 엔트리)에 관한 정보를 수신할 수 있다. 렌더링 매트릭스 생성부는 또한 SAOC 비트스트림(520)(또는, 적어도, SAOC 비트스트림(520)에 의해 나타내는 객체 관련 파라메트릭 정보의 서브세트)을 수신하도록 구성된다. 렌더링 매트릭스 생성부(550)는 또한 수신된 정보에 기초하여 실제 (조정된 또는 수정된) 렌더링 매트릭스(542)를 제공하도록 구성된다. 렌더링 매트릭스 생성부(550)는 장치(100) 또는 장치(240)의 기능을 대신할 수 있다.

[0350] MPEG 서라운드 디코더(510)는 전형적으로 다운믹스 신호 정보(524) 및, 장면 렌더링 엔진(540)에 의해 제공되는 MPEG 서라운드 스트림(522)에 기초하여 다수의 업믹스 채널 신호를 획득하도록 구성된다.

[0351] 요약하면, 오디오 신호 트랜스코더(500)는 MPEG 서라운드 스트림(522)를 제공하여, MPEG 서라운드 스트림(522)이 다운믹스 신호 표현(524)에 기초하여 업믹스 신호 표현의 제공을 허용하도록 구성되며, 업믹스 신호 표현은 실제로 MPEG 서라운드 디코더(510)에 의해 제공된다. 렌더링 매트릭스 생성부(550)는 장면 렌더링 엔진(540)에 의해 이용되는 렌더링 매트릭스(542)를 조정하여, MPEG 서라운드 디코더(510)에 의해 생성되는 업믹스 신호 표현이 수락할 수 없는 가청 왜곡을 포함하지 않도록 한다.

[0352] 4.2 도 5b에 따른 오디오 신호 트랜스코더

[0353] 도 5b는 오디오 신호 트랜스코더(560) 및 MPEG 서라운드 디코더(510)의 다른 장치를 도시한 것이다. 도 5b의 장치는 도 5a의 장치와 매우 유사하여, 동일한 의미 및 신호는 동일한 참조 번호로 명시됨에 주목되어야 한다. 오디오 신호 트랜스코더(560)는 오디오 신호 트랜스코더(560)가 다운믹스 트랜스코더(570)를 포함한다는 점에서 오디오 신호 트랜스코더(500)와 상이하며, 다운믹스 트랜스코더(570)는 입력 다운믹스 표현(524)을 수신하여, MPEG 서라운드 디코더(510)에 공급되는 수정된 다운믹스 표현(574)을 제공하도록 구성된다. 다운믹스 신호 표현의 수정은 원하는 오디오 결과의 정의에 더 많은 유연성을 획득하기 위해 행해진다. 이것은 MPEG 서라운드 비트스트림(522)이 MPEG 서라운드 디코더(510)에 의해 출력되는 업믹스 채널 신호 상으로의 MPEG 서라운드 디코더(510)의 입력 신호의 일부 매핑을 나타낼 수 없다는 사실에 기인한다. 따라서, 다운믹스 트랜스코더(570)를 이용한 다운믹스 신호 표현의 수정은 유연성 증대를 가져올 수 있다.

[0354] 다시말하면, 렌더링 매트릭스 생성부(550)는 장치(100) 또는 장치(240)의 기능을 대신하여, MPEG 서라운드 디코더(510)에 의해 제공되는 업믹스 신호 표현의 가청 왜곡이 확실히 매우 작게 유지되게 할 수 있다.

[0355] 5. 도 6에 따른 오디오 신호 인코더

[0356] 다음에는, 오디오 신호 인코더(600)이 도 6을 참조로 설명되며, 도 6은 이와 같은 오디오 신호 인코더의 개략적인 블록도를 도시한 것이다. 오디오 신호 인코더(600)는 다수의 객체 신호(612a, 612N)(또한 x_1 내지 x_N 으로 명시됨)를 수신하여, 이에 기초하여, 다운믹스 신호 표현(614) 및 객체 관련 파라메트릭 정보(616)를 제공하도록 구성된다. 오디오 신호 인코더(600)는 객체 신호와 관련된 다운믹스 계수(d_1 내지 d_N)에 따라 (다운믹스 신호 표현(614)을 구성하는) 하나 이상의 다운믹스 신호를 제공하여, 하나 이상의 다운믹스 신호가 다수의 객체 신호의 중첩을 포함하도록 구성되는 다운믹서(620)를 포함한다. 오디오 신호 인코더(600)는 또한 2 이상의 객체 신호(612a 내지 612N)의 레벨차 및 상관 특성을 나타내는 객체간 관계 보조 정보를 제공하도록 구성되는 보조 정보 제공기(630)를 포함한다. 보조 정보 제공기(630)는 또한 개개의 객체 신호의 하나 이상의 개개의 속성을 나타내는 개개의 객체 보조 정보를 제공하도록 구성된다.

[0357] 따라서, 오디오 신호 인코더(600)는 객체 관련 파라메트릭 정보(616)를 제공하여, 객체 관련 파라메트릭 정보가 객체간 관계 보조 정보 및 개개의 객체 보조 정보의 양방을 포함하도록 한다.

[0358] 객체 신호와 단일 객체 신호의 개개의 특성 사이의 관계를 나타내는 그런 객체 관련 파라메트릭 정보는 상술한 바와 같이 오디오 신호 디코더에 다중 채널 오디오 신호의 제공을 허용하는 것으로 발견되었다. 객체간 관계 보조 정보는, 다운믹스 신호 표현으로부터 적어도 대략 개개의 객체 신호를 추출하기 위해 객체 관련 파라메트릭

정보(616)를 수신하는 오디오 신호 디코더에 의해 이용될 수 있다. 또한 객체 관련 파라메트릭 정보(614) 내에 포함되는 개개의 객체 보조 정보는 업믹스 프로세스가 너무 강한 신호 왜곡을 가져오기 때문에 업믹스 매개 변수(예컨대, 렌더링 매개 변수)를 조정할 필요가 있는지를 검증하도록 오디오 신호 디코더에 의해 이용될 수 있다.

[0359] 바람직하게는, 보조 정보 제공기(630)는 개개의 객체 보조 정보를 제공하여 개개의 객체 보조 정보가 개개의 객체 신호의 음조를 나타내도록 구성된다. 음조 정보는 업믹스 프로세스가 상당한 왜곡을 가져오는지의 여부를 평가하기 위한 신뢰할 수 있는 기준으로 이용될 수 있는 것으로 발견되었다.

[0360] 또한, 오디오 신호 인코더(600)에는 오디오 신호 인코더에 대해 여기서 논의된 어떤 특징 및 기능이 추가되고, 다운믹스 신호 표현(614) 및 객체 관련 파라메트릭 정보(616)는 오디오 신호 인코더(600)에 의해 제공되어, 발명의 오디오 신호 디코더에 대해 논의된 특성을 포함하는 것에 주목되어야 한다.

[0361] 6. 도 7에 따른 오디오 비트스트림

[0362] 본 발명에 따른 실시예는 오디오 비트스트림(700)을 생성하며, 이의 개략적 표현은 도 7에 도시된다. 오디오 비트스트림은 인코딩된 형식으로 다수의 객체 신호를 나타낸다.

[0363] 오디오 비트스트림(700)은 하나 이상의 다운믹스 신호를 나타내는 다운믹스 신호 표현(710)을 포함하며, 다운믹스 신호의 적어도 하나는 다수의 객체 신호의 중첩을 포함한다. 오디오 비트스트림(700)은 또한 객체 신호의 레벨차 및 상관 특성을 나타내는 객체간 관계 보조 정보(720)를 포함한다. 오디오 비트스트림은 또한 (다운믹스 신호 표현(710)에 대한 기초를 형성하는) 개개의 객체 신호의 하나 이상의 개개의 속성을 나타내는 개개의 객체 보조 정보(730)를 포함한다.

[0364] 객체간 관계 보조 정보 및 개개의 객체 정보는 전적으로 객체 관련 파라메트릭 보조 정보로 간주될 수 있다.

[0365] 바람직한 실시예에서, 개개의 객체 보조 정보는 개개의 객체 신호의 음조를 나타낸다.

[0366] 당연히, 오디오 비트스트림(700)이 전형적으로 여기서 논의된 바와 같이 오디오 신호 인코더에 의해 제공되고, 여기서 논의된 바와 같이 오디오 신호 디코더에 의해 평가된다. 오디오 비트스트림은 오디오 신호 인코더 및 오디오 신호 디코더에 대해 논의된 바와 같은 특성을 포함할 수 있다. 따라서, 오디오 비트스트림(700)은 여기서 논의된 바와 같이 오디오 신호 디코더를 이용하여 다중 채널 오디오 신호의 제공에 적합할 수 있다.

[0367] 7. 결론

[0368] 본 발명에 따른 실시예들은 단일한 원래의 객체 신호가 몇몇 전송된 다운믹스 신호로부터 완전히 재구성될 수 없다는 사실에서 기인하는 상술한 왜곡 문제를 감소시키거나 방지하기 위한 솔루션을 제공한다. 따라서, 적용되는 이러한 문제에 대한 더욱 간단한 솔루션이 있다:

[0369] • 단순한 접근법은 상대 객체 이득의 범위를, 예컨대 +/-12dB로 제한하는 것이다. 큰 객체 이득 설정이 가청 저하(예컨대: 한 객체를 20dB만큼 부스트하지만, 다른 객체 레벨을 0dB에 둠)에 이르게 할 수 있는 것이 사실이지만, 그러나, 이것은 반드시 필요치 않다. 일례로서, 모든 상대 객체 레벨을 동일한 인수만큼의 부스팅은 손상되지 않은 시스템의 출력을 산출한다.

[0370] • 더 정교한 뷰(elaborated view)는 상대 객체 레벨의 차를 볼 수 있다. 두 오디오 객체의 렌더링을 위해, 양방의 상대 객체 레벨의 차는 실제로 렌더링된 출력에서 가능한 저하에 대한 훅(hook)을 제공한다. 그러나, 이러한 아이디어는 2 이상의 렌더링된 오디오 객체로 일반화하는 방법이 명확하지 않다.

[0371] 이러한 상황을 고려하여, 본 발명에 따른 실시예들은 이러한 문제를 처리하여, 불만족스런 사용자 경험을 방지하기 위한 수단을 제공한다. 본 발명에 따른 일부 실시예들은 이전의 섹션에서 논의된 것보다 더 정교한 솔루션을 가지고 있다.

[0372] 따라서, 부적절한 렌더링 매개 변수가 사용자에게 의해 제공될지라도, 본 발명을 이용하여 양호한 청각 인상이 획득될 수 있다.

[0373] 일반적으로, 본 발명에 따른 실시예들은 오디오 신호를 인코딩하거나 인코딩된 오디오 신호를 디코딩하는 장치, 방법 또는 컴퓨터 프로그램, 또는 상술한 바와 같은 (예컨대, 오디오 비트스트림의 형식의) 인코딩된 오디오 신호에 관한 것이다.

[0374] 8. 구현 대안

- [0375] 일부 양태가 장치와 관련하여 설명되었지만, 이들 양태는 또한 대응하는 방법에 대한 설명을 명백히 나타내며, 여기서, 블록 또는 디바이스는 방법 단계 또는 방법 단계의 특징에 대응한다. 유사하게도, 방법 단계와 관련하여 설명되는 양태는 또한 대응하는 장치의 대응하는 블록 또는 항목 또는 특징에 대한 설명을 나타낸다. 방법 단계의 일부 또는 모두는 예컨대, 마이크로프로세서, 프로그램 가능한 컴퓨터 또는 전자 회로와 같은 하드웨어 장치에 의해(또는 이용하여) 실행될 수 있다. 일부 실시예들에서, 가장 중요한 방법 단계 중 일부의 하나 이상은 이와 같은 장치에 의해 실행될 수 있다.
- [0376] 발명의 인코딩된 오디오 신호 또는 오디오 비트스트림은 디지털 저장 매체 상에 저장될 수 있거나, 무선 전송 매체와 같은 전송 매체 또는 인터넷과 같은 유선 전송 매체 상에서 전송될 수 있다.
- [0377] 어떤 구현 요건에 따라, 본 발명의 실시예들은 하드웨어 또는 소프트웨어에서 구현될 수 있다. 이런 구현은 디지털 저장 매체, 예컨대, 플로피 디스크, DVD, 블루레이, CD, ROM, PROM, EPROM, EEPROM 또는 플래시 메모리를 이용하여 실행될 수 있으며, 이들은 전자식 판독 가능한 제어 신호를 저장하여, 각각의 방법이 실행되도록 하는 프로그램 가능한 컴퓨터 시스템과 협력한다 (또는 협력할 수 있다). 그래서, 디지털 저장 매체는 컴퓨터 판독 가능할 수 있다.
- [0378] 본 발명에 따른 일부 실시예들은 여기에 설명된 방법 중 하나가 수행되도록 프로그램 가능한 컴퓨터 시스템과 협력할 수 있는 전자식 판독 가능한 제어 신호를 가진 데이터 캐리어를 포함한다.
- [0379] 일반적으로, 본 발명의 실시예들은 프로그램 코드를 가진 컴퓨터 프로그램 제품으로서 구현될 수 있으며, 이 프로그램 코드는 컴퓨터 프로그램 제품이 컴퓨터 상에서 실행할 시에 방법 중 하나를 수행하기 위해 동작 가능하다. 프로그램 코드는, 예컨대, 기계 판독 가능한 캐리어 상에 저장될 수 있다.
- [0380] 다른 실시예들은, 기계 판독 가능한 캐리어 상에 저장되고, 여기에 설명된 방법 중 하나를 실행하는 컴퓨터 프로그램을 포함한다.
- [0381] 그래서, 환언하면, 발명의 방법의 실시예는, 컴퓨터 프로그램이 컴퓨터 상에서 실행할 시에, 여기에 설명된 방법 중 하나를 실행하기 위한 프로그램 코드를 가진 컴퓨터 프로그램이다.
- [0382] 그래서, 발명의 방법의 추가 실시예는, 여기에 설명된 방법 중 하나를 실행하기 위한 컴퓨터 프로그램을 기록한 데이터 캐리어 (또는 디지털 저장 매체, 또는 컴퓨터 판독 가능한 매체)이다.
- [0383] 그래서, 발명의 방법의 추가 실시예는 여기에 설명된 방법 중 하나를 실행하기 위한 컴퓨터 프로그램을 나타내는 데이터 스트림 또는 신호의 시퀀스이다. 데이터 스트림 또는 신호의 시퀀스는, 예컨대, 데이터 통신 접속을 통해, 예컨대, 인터넷을 통해 전송되도록 구성될 수 있다.
- [0384] 추가 실시예는, 여기에 설명된 방법 중 하나를 실행하기 위해 구성되거나 적용되는 처리 수단, 예컨대, 컴퓨터, 또는 프로그램 가능한 논리 디바이스를 포함한다.
- [0385] 추가 실시예는 여기에 설명된 방법 중 하나를 실행하기 위한 컴퓨터 프로그램을 설치한 컴퓨터를 포함한다.
- [0386] 일부 실시예들에서, 프로그램 가능한 논리 디바이스 (예컨대, 필드 프로그램 가능 게이트 어레이)는 여기에 설명된 방법의 일부 또는 모든 기능을 실행하는데 이용될 수 있다. 일부 실시예들에서, 필드 프로그램 가능 게이트 어레이는 여기에 설명된 방법 중 하나를 실행하기 위해 마이크로프로세서와 협력할 수 있다. 일반적으로, 이들 방법은 바람직하게는 어떤 하드웨어 장치에 의해 실행된다.
- [0387] 상술한 실시예들은 단지 본 발명의 원리를 위해 예시한 것이다. 여기에 설명된 배치 및 상세 사항의 수정 및 변형은 당업자에게는 자명한 것으로 이해된다. 그래서, 여기의 실시예의 설명을 통해 제시된 특정 상세 사항에 의해 제한되지 않고, 첨부한 특허청구범위의 범주에 의해서만 제한되는 것으로 의도된다.

[0388] 참고 문헌

- [0389] [BCC] C. Faller and F. Baumgarte, "Binaural Cue Coding - Part II: Schemes and applications," IEEE Trans. on Speech and Audio Proc., vol. 11, no. 6, Nov. 2003
- [0390] [JSC] C. Faller, "Parametric Joint-Coding of Audio Sources", 120th AES Convention, Paris, 2006,

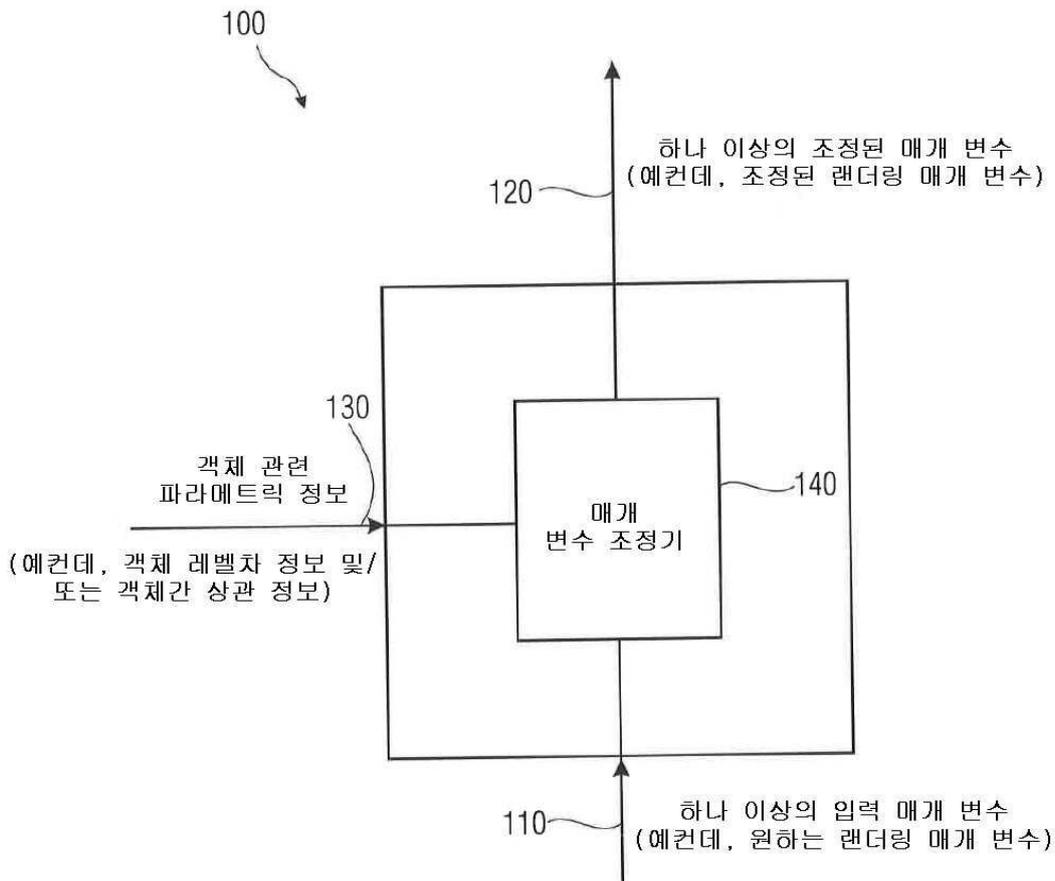
Preprint 6752

[0391] [SAOC1] J. Herre, S. Disch, J. Hilpert, O. Hellmuth: "From SAC To SAOC - Recent Developments in Parametric Coding of Spatial Audio", 22nd Regional UK AES Conference, Cambridge, UK, April 2007

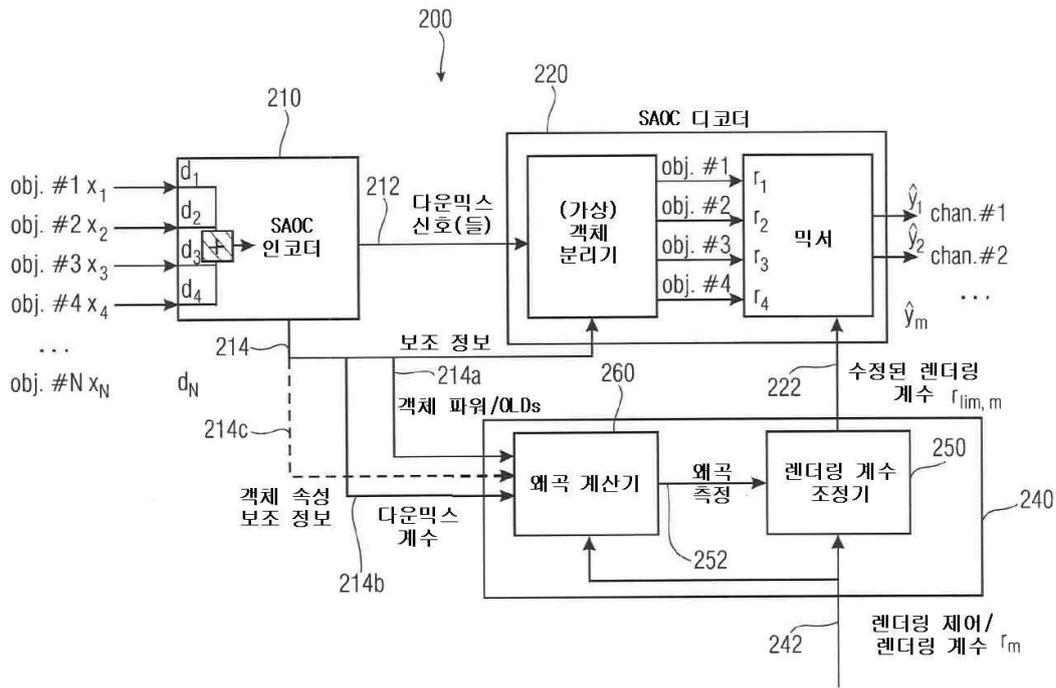
[0392] [SAOC2] J. Engdegard, B. Resch, C. Falch, O. Hellmuth, J. Hilpert, A. Holzer, L. Terentiev, J. Breebaart, J. Koppens, E. Schuijers and W. Oomen: " Spatial Audio Object Coding (SAOC) - The Upcoming MPEG Standard on Parametric Object Based Audio Coding", 124th AES Convention, Amsterdam 2008, Preprint 7377

도면

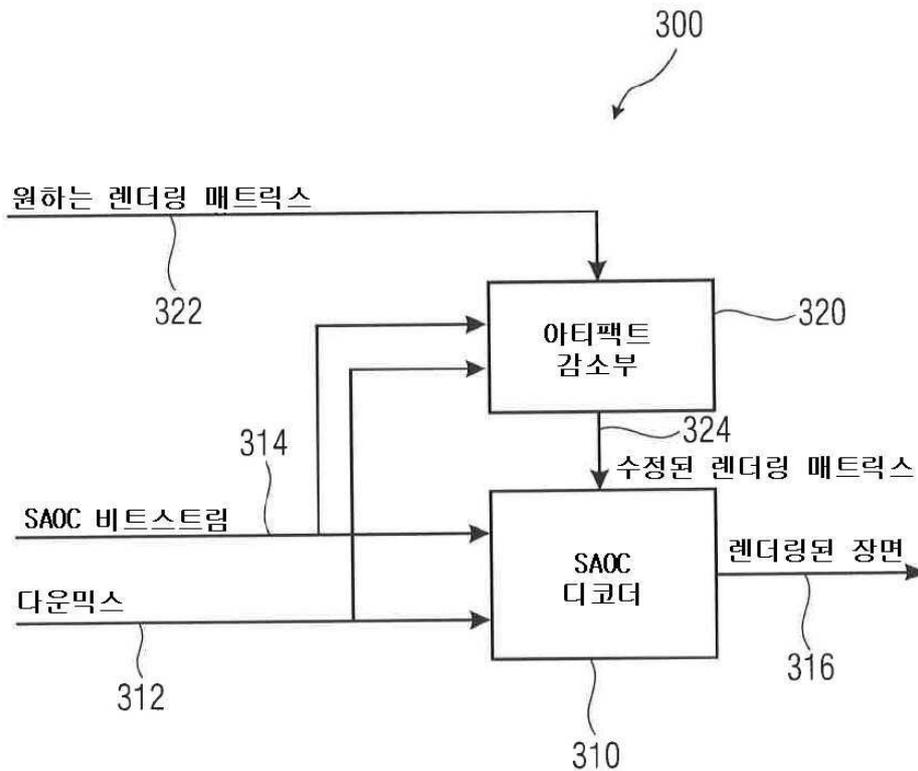
도면1



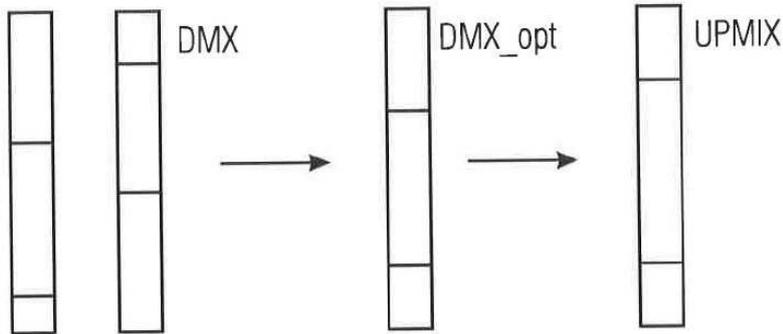
도면2



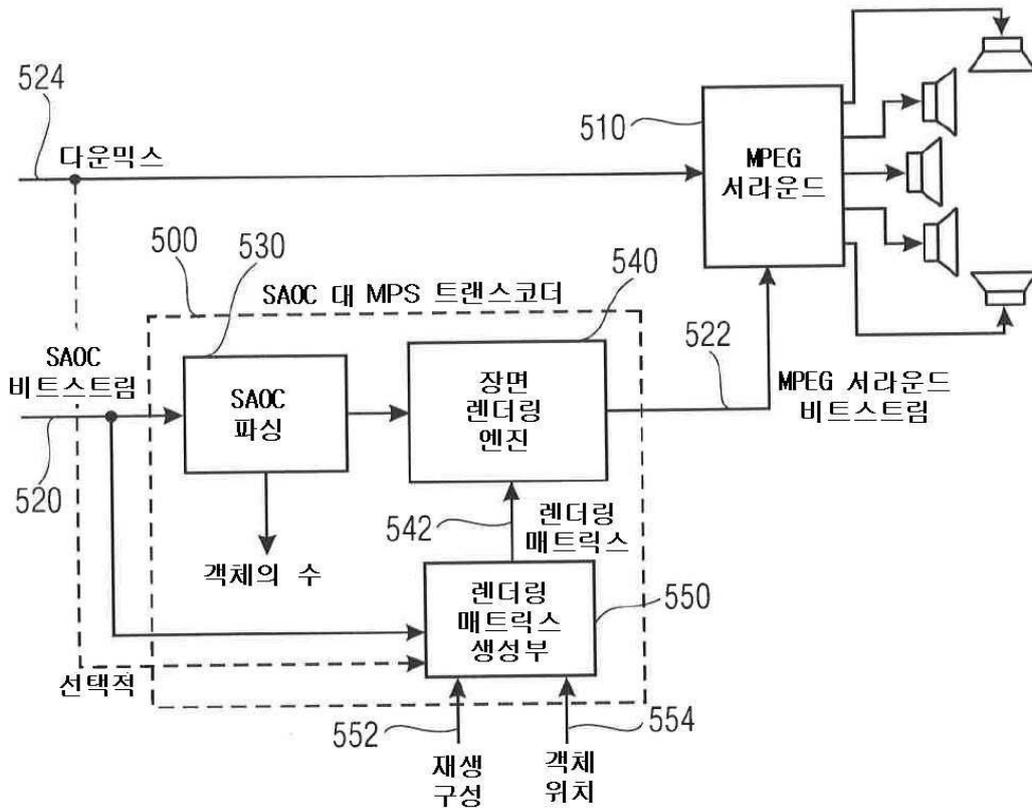
도면3



도면4

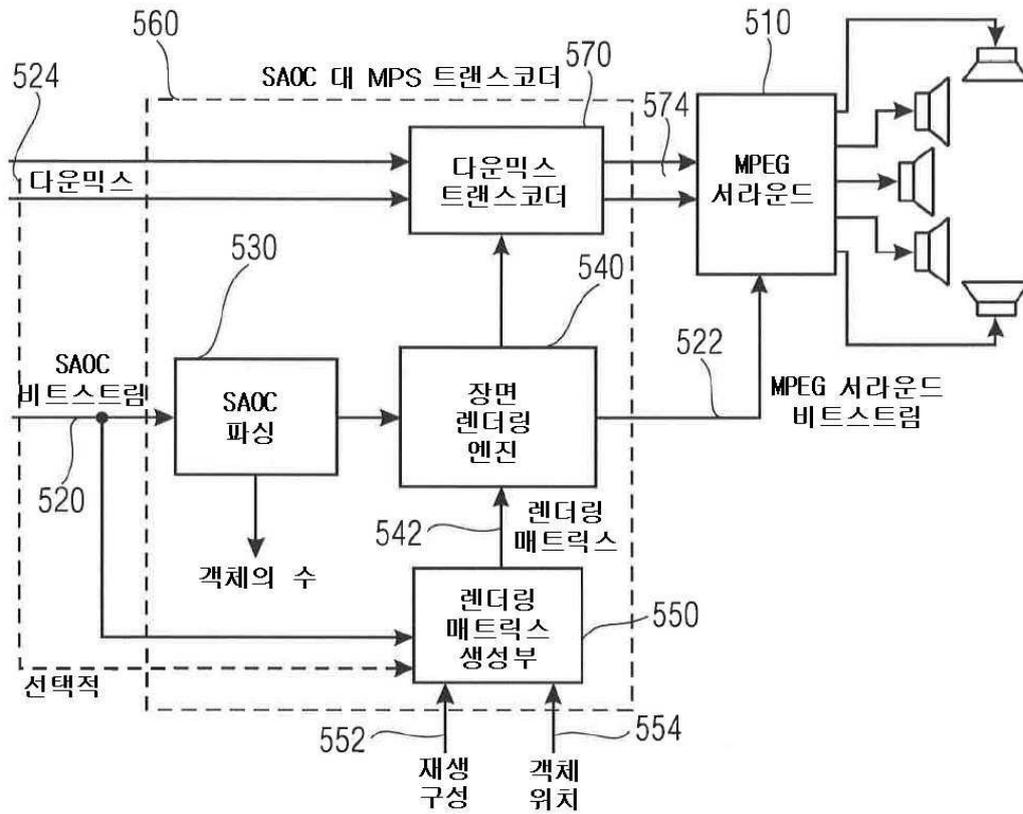


도면5a



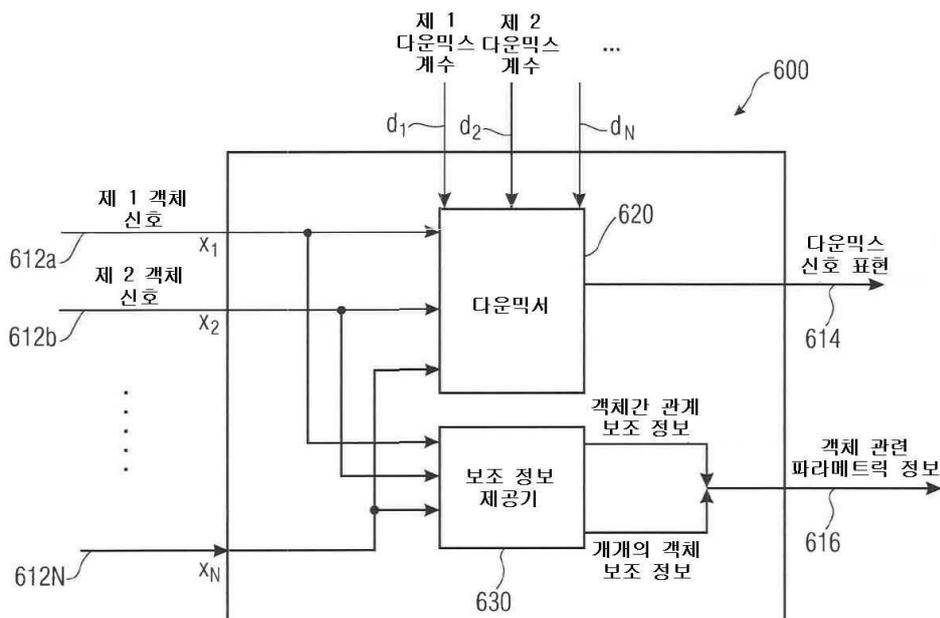
(A) 모노 다운믹스 기반 트랜스코더

도면5b

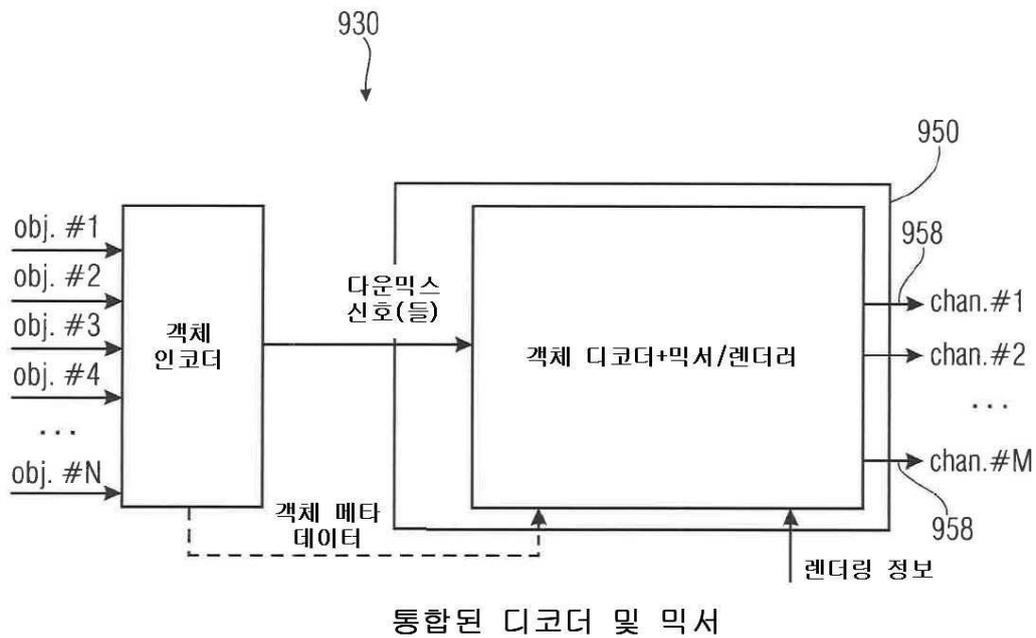


(B) 스테레오 다운믹스 기반 트랜스코더

도면6



도면9b



도면9c

