



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2018-0069095
(43) 공개일자 2018년06월22일

- | | |
|---|--|
| <p>(51) 국제특허분류(Int. Cl.)
G10L 19/008 (2014.01) G10L 19/005 (2013.01)
G10L 19/02 (2006.01) G10L 19/16 (2013.01)</p> <p>(52) CPC특허분류
G10L 19/008 (2013.01)
G10L 19/005 (2013.01)</p> <p>(21) 출원번호 10-2018-7016512(분할)</p> <p>(22) 출원일자(국제) 2014년07월16일
심사청구일자 없음</p> <p>(62) 원출원 특허 10-2016-7004615
원출원일자(국제) 2014년07월16일
심사청구일자 2016년02월22일</p> <p>(85) 번역문제출일자 2018년06월11일</p> <p>(86) 국제출원번호 PCT/EP2014/065283</p> <p>(87) 국제공개번호 WO 2015/010996
국제공개일자 2015년01월29일</p> <p>(30) 우선권주장
EP13177365 2013년07월22일
유럽특허청(EPO)(EP)
(뒷면에 계속)</p> | <p>(71) 출원인
프라운호퍼 게젤샤프트 쭈르 피르데롱 데어 안겐
반텐 포르슈 에. 베.
독일 80686 뮌헨 한자슈트라쎄 27 체</p> <p>(72) 발명자
보르스, 크리스티안
독일 91058 에를랑겐, 엘리세스파에스 스트라쎄 5
에르텔, 크리스티안
독일 90542 에켄탈, 뉘른베르거 스트라쎄 24
힐퍼트, 요하네스
독일 90411 뉘른베르크, 헤른후트스트라쎄 46</p> <p>(74) 대리인
특허법인 정안</p> |
|---|--|

전체 청구항 수 : 총 15 항

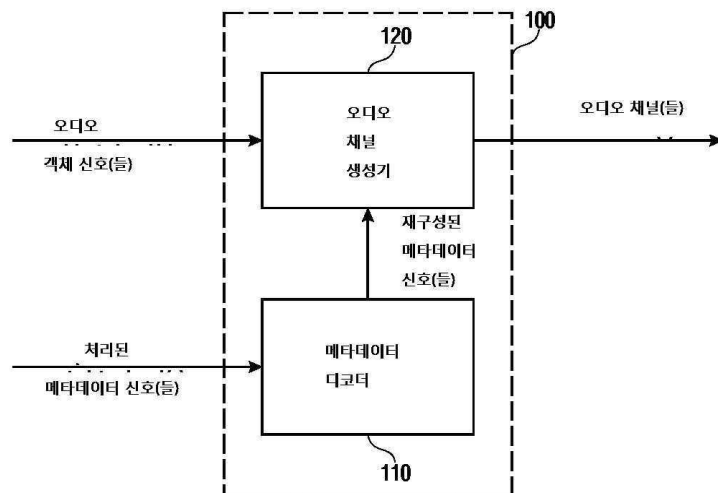
(54) 발명의 명칭 저 지연 객체 메타데이터 코딩을 위한 장치 및 방법

(57) 요약

하나 이상의 오디오 채널들을 생성하기 위한 장치가 제공된다. 장치는 제어 신호(b)에 따라 하나 이상의 처리된 메타데이터 신호들(z_1, \dots, z_N)로부터 하나 이상의 재구성된 메타데이터 신호들(x_1', \dots, x_N')을 생성하기 위한 메타데이터 디코더(110)를 포함하고, 하나 이상의 재구성된 메타데이터 신호(x_1', \dots, x_N') 각각은 하나 이

(뒷면에 계속)

대표도 - 도1



상의 오디오 객체 신호들의 오디오 객체 신호와 연관된 정보를 나타내고, 메타데이터 디코더(110)는 하나 이상의 재구성된 메타데이터 신호들(x_1' , ..., x_N') 각각에 대해 복수의 재구성된 메타데이터 샘플($x_1'(n)$, ..., $x_N'(n)$)을 결정함으로써 상기 하나 이상의 재구성된 메타데이터 신호들(x_1' , ..., x_N')을 생성하도록 구성된다. 더욱이, 장치는 하나 이상의 오디오 객체 신호들에 따라 그리고 하나 이상의 재구성된 메타데이터 신호들(x_1' , ..., x_N')에 따라 하나 이상의 오디오 채널들을 생성하기 위한 오디오 채널 생성기(120)를 포함한다. 메타데이터 디코더(110)는 하나 이상의 처리된 메타데이터 신호들(z_1 , ..., z_N) 각각의 복수의 처리된 메타데이터 샘플($z_1(n)$, ..., $z_N(n)$)을 수신하도록 구성된다. 더욱이, 메타데이터 디코더(110)는 제어 신호(b)를 수신하도록 구성된다. 더욱이, 메타데이터 디코더(110)는 하나 이상의 재구성된 메타데이터 신호(x_1' , ..., x_N')의 각 재구성된 메타데이터 신호(x_i')의 복수의 재구성된 메타데이터 샘플($x_i'(1)$, ..., $x_i'(n-1)$, $x_i'(n)$)의 각 재구성된 메타데이터 샘플($x_i'(n)$)을 결정하도록 구성되어, 제어 신호(b)가 제 1 상태($b(n)=0$)를 나타낼 때, 상기 재구성된 메타데이터 샘플($x_i'(n)$)은 하나 이상의 처리된 메타데이터 신호(z_i) 중 하나의 처리된 메타데이터 신호의 상기 처리된 메타데이터 샘플($z_i(n)$) 중 하나와 상기 재구성된 메타데이터 신호(x_i')의 다른 이미 생성된 재구성된 메타데이터 샘플($x_i'(n-1)$)의 합이다.

(52) CPC특허분류

G10L 19/02 (2013.01)

G10L 19/16 (2013.01)

(30) 우선권주장

EP13177367 2013년07월22일

유럽특허청(EPO)(EP)

EP13177378 2013년07월22일

유럽특허청(EPO)(EP)

EP13189279 2013년10월18일

유럽특허청(EPO)(EP)

명세서

청구범위

청구항 1

하나 이상의 오디오 채널들을 생성하기 위한 장치(100)로서,
 제어 신호(b)에 따라 하나 이상의 처리된 메타데이터 신호들(z_1, \dots, z_N)로부터 하나 이상의 재구성된 메타데이터 신호들(x_1', \dots, x_N')을 생성하기 위한 메타데이터 디코더(110;901)로서, 상기 하나 이상의 재구성된 메타데이터 신호들(x_1', \dots, x_N') 각각은 하나 이상의 오디오 객체 신호들의 오디오 객체 신호와 연관된 정보를 나타내고, 상기 메타데이터 디코더(110;901)는 상기 하나 이상의 재구성된 메타데이터 신호들(x_1', \dots, x_N') 각각에 대해 복수의 재구성된 메타데이터 샘플들($x_1'(n), \dots, x_N'(n)$)을 결정함으로써 상기 하나 이상의 재구성된 메타데이터 신호들(x_1', \dots, x_N')을 생성하도록 구성되는, 메타데이터 디코더(110; 901), 및
 상기 하나 이상의 오디오 객체 신호들에 따라 그리고 상기 하나 이상의 재구성된 메타데이터 신호들(x_1', \dots, x_N')에 따라 상기 하나 이상의 오디오 채널들을 생성하기 위한 오디오 채널 생성기(120)를 포함하고,
 상기 메타데이터 디코더(110;901)는 상기 하나 이상의 처리된 메타데이터 신호들(z_1, \dots, z_N) 각각의 복수의 처리된 메타데이터 샘플들($z_1(n), \dots, z_N(n)$)을 수신하도록 구성되고,
 상기 메타데이터 디코더(110;901)는 제어 신호(b)를 수신하도록 구성되고,
 상기 메타데이터 디코더(110;901)는 상기 제어 신호(b)를 분석하도록 구성되어, 상기 제어 신호(b)가 제 1 상태($b(n)=0$)를 나타내는지 또는 상기 제어 신호(b)가 상기 제 1 상태($b(n)=0$)와 상이한 제 2 상태($b(n)=1$)를 나타내는지 결정하고,
 만약 상기 제어 신호(b)가 제 1 상태($b(n)=0$)를 나타내는 경우, 상기 메타데이터 디코더(110;901)는 상기 하나 이상의 재구성된 메타데이터 신호들(x_1', \dots, x_N')의 각 재구성된 메타데이터 신호(x_i')의 상기 복수의 재구성된 메타데이터 샘플들($x_i'(1), \dots, x_i'(n-1), x_i'(n)$)의 각 재구성된 메타데이터 샘플($x_i'(n)$)을 결정하도록 구성되어, 상기 재구성된 메타데이터 샘플($x_i'(n)$)은 상기 하나 이상의 처리된 메타데이터 신호들(z_i) 중 하나의 처리된 메타데이터 신호의 상기 처리된 메타데이터 샘플들($z_i(n)$) 중 하나와 상기 재구성된 메타데이터 신호(x_i')의 다른 이미 생성된 재구성된 메타데이터 샘플($x_i'(n-1)$)의 합이고,
 만약 상기 제어 신호(b)가 상기 제 2 상태($b(n)=1$)를 나타내는 경우, 상기 메타데이터 디코더(110;901)는 상기 재구성된 메타데이터 샘플($x_i'(n)$)이 상기 하나 이상의 처리된 메타데이터 신호들(z_1, \dots, z_N)의 상기 하나의 처리된 메타데이터 신호(z_i)의 상기 처리된 메타데이터 샘플들($z_i(1), \dots, z_i(n)$)의 상기 하나의 처리된 메타데이터 샘플($z_i(n)$)이라는 것을 결정하도록 구성되는,
 하나 이상의 오디오 채널들을 생성하기 위한 장치.

청구항 2

제 1항에 있어서,
 상기 메타데이터 디코더(110; 901)는 상기 처리된 메타데이터 신호들(z_1, \dots, z_N) 중 2개 이상을 수신하도록 구성되고, 상기 재구성된 메타데이터 신호들(x_1', \dots, x_N') 중 2개 이상을 생성하도록 구성되고,
 상기 메타데이터 디코더(110; 901)는 2개 이상의 메타데이터 디코더 서브 유닛들(911,...,91N)을 포함하고,
 상기 2개 이상의 메타데이터 디코더 서브 유닛들(911,...,91N)의 각각(91i; 91i')은 가산기(910) 및 선택기

(930)를 포함하고,

상기 2개 이상의 메타데이터 디코더 서브 유닛들(911, ..., 91N)의 각각(91i; 91i')은 상기 2개 이상의 처리된 메타데이터 신호들(z_1, \dots, z_N) 중 하나(z_i)의 상기 복수의 처리된 메타데이터 샘플들($z_i(1) \dots, z_i(n-1), z_i(n)$)을 수신하도록 구성되고, 2개 이상의 재구성된 메타데이터 신호들(x_1', \dots, x_N') 중 하나(x_i')를 생성하도록 구성되고,

상기 메타데이터 디코더 서브 유닛(91i; 91i')의 상기 가산기(910)는 합 값($s_i(n)$)을 얻기 위해 상기 2개 이상의 처리된 메타데이터 신호들(z_1, \dots, z_N) 중 상기 하나(z_i)의 상기 처리된 메타데이터 샘플들($z_i(1) \dots, z_i(n)$)의 하나($z_i(n)$)와 상기 2개 이상의 재구성된 메타데이터 신호들(x_1', \dots, x_N') 중 상기 하나(x_i')의 다른 이미 생성된 재구성된 메타데이터 샘플($x_i'(n-1)$)을 가산하도록 구성되고,

상기 메타데이터 디코더 서브 유닛(91i; 91i')의 상기 선택기(930)는 상기 처리된 메타데이터 샘플($z_i(n)$) 중 상기 하나, 상기 합 값($s_i(n)$) 및 상기 제어 신호를 수신하도록 구성되고, 상기 선택기(930)는 상기 재구성된 메타데이터 신호(x_i')의 상기 복수의 메타데이터 샘플들($x_i'(1) \dots, x_i'(n-1), x_i'(n)$) 중 하나를 결정하도록 구성되어, 상기 제어 신호(b)가 상기 제 1 상태($b(n)=0$)를 나타낼 때, 상기 재구성된 메타데이터 샘플($x_i'(n)$)은 상기 합 값($s_i(n)$)이고, 상기 제어 신호가 상기 제 2 상태($b(n)=1$)를 나타낼 때, 상기 재구성된 메타데이터 샘플($x_i'(n)$)은 상기 처리된 메타데이터 신호들($z_1(1) \dots, z_i(n)$)의 상기 하나($z_i(n)$)인, 하나 이상의 오디오 채널들을 생성하기 위한 장치.

청구항 3

제 1항에 있어서,

상기 하나 이상의 재구성된 메타데이터 신호들(x_1', \dots, x_N') 중 적어도 하나는 상기 하나 이상의 오디오 객체 신호들 중 하나에 관한 위치 정보를 나타내고,

상기 오디오 채널 생성기(120)는 상기 하나 이상의 오디오 객체 신호들 중 상기 하나에 따라, 그리고 상기 위치 정보에 따라 상기 하나 이상의 오디오 채널들 중 적어도 하나를 생성하도록 구성되는, 하나 이상의 오디오 채널들을 생성하기 위한 장치.

청구항 4

제 1항에 있어서,

상기 하나 이상의 재구성된 메타데이터 신호들(x_1', \dots, x_N') 중 적어도 하나는 상기 하나 이상의 오디오 객체 신호들 중 하나의 볼륨을 나타내고,

상기 오디오 채널 생성기(120)는 상기 하나 이상의 오디오 객체 신호들 중 상기 하나에 따라, 그리고 상기 볼륨에 따라 상기 하나 이상의 오디오 채널들 중 적어도 하나를 생성하도록 구성되는, 하나 이상의 오디오 채널들을 생성하기 위한 장치.

청구항 5

인코딩된 오디오 데이터를 디코딩하기 위한 장치로서,

상기 인코딩된 오디오 데이터를 수신하기 위한 입력 인터페이스(1100)로서, 상기 인코딩된 오디오 데이터는 복수의 인코딩된 채널들 또는 복수의 인코딩된 객체들 또는 상기 복수의 객체들에 관련된 압축 메타데이터를 포함하는, 입력 인터페이스(1100), 및

제 1항 내지 제 4항 중 어느 한 항에 따른 장치(100)를 포함하고,

제 1항 내지 제 4항 중 어느 한 항에 따른 상기 장치(100)의 상기 메타데이터 디코더(110; 901)는 상기 압축된 메타데이터를 압축 해제하기 위한 메타데이터 압축 해제기(400)이고,

제 1항 내지 제 4항 중 어느 한 항에 따른 상기 장치(100)의 상기 오디오 채널 생성기(120)는 상기 복수의 인코

딩된 채널들 및 상기 복수의 인코딩된 객체들을 디코딩하기 위한 코어 디코더(1300)를 포함하고,

상기 오디오 채널 생성기(120)는 상기 객체들로부터의 오디오 데이터와 상기 디코딩된 채널들을 포함하는 다수의 출력 채널들(1205)을 얻기 위해 상기 압축 해제된 메타데이터를 이용하여 상기 복수의 디코딩된 객체들을 처리하기 위한 객체 프로세서(1200)를 더 포함하고,

상기 오디오 채널 생성기(120)는 상기 다수의 출력 채널들(1205)을 출력 포맷으로 변환하기 위한 후치 프로세서(1700)를 더 포함하는, 하나 이상의 오디오 채널들을 생성하기 위한 장치.

청구항 6

하나 이상의 인코딩된 오디오 신호들 및 하나 이상의 처리된 메타데이터 신호들을 포함하는 인코딩된 오디오 정보를 생성하기 위한 장치(250)로서,

하나 이상의 원본 메타데이터 신호들을 수신하고 상기 하나 이상의 처리된 메타데이터 신호들을 결정하기 위한 메타데이터 인코더(210; 801; 802)로서, 상기 하나 이상의 원본 메타데이터 신호들 각각은 복수의 원본 메타데이터 샘플들을 포함하고, 상기 하나 이상의 원본 메타데이터 신호들 각각의 상기 원본 메타데이터 샘플들은 하나 이상의 오디오 객체 신호들 중 하나의 오디오 객체 신호와 연관된 정보를 나타내는, 메타데이터 인코더(210; 801; 802), 및

상기 하나 이상의 인코딩된 오디오 신호들을 얻기 위해 상기 하나 이상의 오디오 객체 신호들을 인코딩하기 위한 오디오 인코더(220)를 포함하고,

상기 메타데이터 인코더(210; 801; 802)는 상기 하나 이상의 처리된 메타데이터 신호들(z_1, \dots, z_N)의 각 처리된 메타데이터 신호(z_i)의 복수의 처리된 메타데이터 샘플들($z_i(1), \dots, z_i(n-1), z_i(n)$)의 각 처리된 메타데이터 샘플($z_i(n)$)을 결정하도록 구성되어, 제어 신호(b)가 제 1 상태($b(n)=0$)를 나타낼 때, 상기 처리된 메타데이터 샘플($z_i(n)$)은 상기 하나 이상의 원본 메타데이터 신호들 중 하나(x_i)의 복수의 원본 메타데이터 샘플들 중의 하나($x_i(n)$)와 상기 처리된 메타데이터 신호(z_i)의 다른 이미 생성된 처리된 메타데이터 샘플 사이의 차이 또는 양자화된 차이를 나타내고,

상기 메타데이터 인코더(210; 801; 802)는 상기 처리된 메타데이터 샘플($z_i(n)$)을 결정하도록 구성되어, 상기 제어 신호(b)가 상기 제 1 상태와 상이한 제 2 상태($b(n)=1$)을 나타낼 때, 상기 처리된 메타데이터 샘플($z_i(n)$)은 상기 하나 이상의 원본 메타데이터 신호들 중 상기 하나(x_i)의 상기 원본 메타데이터 샘플들($x_i(1), \dots, x_i(n)$) 중의 상기 하나($x_i(n)$)이거나, 상기 원본 메타데이터 샘플들($x_i(1), \dots, x_i(n)$) 중의 상기 하나($x_i(n)$)의 양자화된 표현($q_i(n)$)인,

인코딩된 오디오 정보를 생성하기 위한 장치.

청구항 7

제 6항에 있어서,

상기 메타데이터 인코더(210; 801; 802)는 상기 원본 메타데이터 신호들(x_1, \dots, x_N) 중 2개 이상을 수신하도록 구성되고, 상기 처리된 메타데이터 신호들(z_1, \dots, z_N) 중 2개 이상을 생성하도록 구성되고,

상기 메타데이터 인코더(210; 801; 802)는 2개 이상의 DPCM 인코더들(811, ..., 81N)을 포함하고,

상기 2개 이상의 DPCM 인코더들(811, ..., 81N) 각각은 차이 샘플($y_i(n)$)을 얻기 위해 상기 2개 이상의 원본 메타데이터 신호들(x_1, \dots, x_N)의 하나(x_i)의 상기 원본 메타데이터 샘플들($x_i(1), \dots, x_i(n)$)의 하나($x_i(n)$)와 상기 2개 이상의 재구성된 메타데이터 신호들(z_1, \dots, z_N)의 하나(z_i)의 다른 이미 생성된 처리된 메타데이터 샘플 사이의 차이 또는 양자화된 차이를 결정하도록 구성되고,

상기 메타데이터 인코더(210; 801; 802)는 선택기(830)를 더 포함하고, 상기 선택기(830)는 상기 처리된 메타데이터 신호(z_i)의 상기 복수의 처리된 메타데이터 샘플들($z_i(1), \dots, z_i(n-1), z_i(n)$)의 하나를 결정하도록 구성

되어, 상기 제어 신호(b)가 상기 제 1 상태(b(n)=0)를 나타낼 때, 상기 처리된 메타데이터 샘플($z_i(n)$)은 상기 차이 샘플($y_i(n)$)이고, 상기 제어 신호가 상기 제 2 상태(b(n)=1)를 나타낼 때, 상기 처리된 메타데이터 샘플($z_i(n)$)은 상기 원본 메타데이터 샘플들($x_i(1), \dots, x_i(n)$)의 상기 하나($x_i(n)$)이거나, 상기 원본 메타데이터 샘플들($x_i(1), \dots, x_i(n)$)의 상기 하나($x_i(n)$)의 양자화된 표현($q_i(n)$)인, 인코딩된 오디오 정보를 생성하기 위한 장치.

청구항 8

제 6항에 있어서,

상기 하나 이상의 원본 메타데이터 신호들 중 적어도 하나는 상기 하나 이상의 오디오 객체 신호들 중 하나에 관한 위치 정보를 나타내고,

상기 메타데이터 인코더(210; 801; 802)는 상기 위치 정보를 나타내는 상기 하나 이상의 원본 메타데이터 신호들 중 상기 적어도 하나에 따라 상기 하나 이상의 처리된 메타데이터 신호들의 적어도 하나를 생성하도록 구성되는, 인코딩된 오디오 정보를 생성하기 위한 장치.

청구항 9

제 6항에 있어서,

상기 하나 이상의 원본 메타데이터 신호들 중 적어도 하나는 상기 하나 이상의 오디오 객체 신호들 중 하나의 볼륨을 나타내고,

상기 메타데이터 인코더(210; 801; 802)는 상기 볼륨 정보를 나타내는 상기 하나 이상의 원본 메타데이터 신호들 중 상기 적어도 하나에 따라 상기 하나 이상의 처리된 메타데이터 신호들의 적어도 하나를 생성하도록 구성되는, 인코딩된 오디오 정보를 생성하기 위한 장치.

청구항 10

제 6항에 있어서, 상기 메타데이터 인코더(210; 801; 802)는 상기 제어 신호가 상기 제 1 상태(b(n) = 0)를 나타낼 때 비트들의 제 1 수로 그리고 상기 제어 신호가 상기 제 2 상태(b(n) = 1)를 나타낼 때 비트들의 제 2 수로 상기 하나 이상의 처리된 메타데이터 신호들(z_1, \dots, z_N)의 하나($z_i()$)의 상기 처리된 메타데이터 샘플들($z_i(1), \dots, z_i(n)$) 각각을 인코딩하도록 구성되는, 인코딩된 오디오 정보를 생성하기 위한 장치.

청구항 11

오디오 출력 데이터(501)를 얻기 위해 오디오 입력 데이터(101)를 인코딩하기 위한 장치로서,

복수의 오디오 채널들, 복수의 오디오 객체들 및 상기 복수의 오디오 객체들 중 하나 이상에 관련된 메타데이터를 수신하기 위한 입력 인터페이스(1100),

복수의 사전-믹싱된 채널들을 얻기 위해 상기 복수의 객체들 및 상기 복수의 채널들을 믹싱하기 위한 믹서(200)로서, 각 사전-믹싱된 채널은 채널의 오디오 데이터 및 적어도 하나의 객체의 오디오 데이터를 포함하는, 믹서(200), 및

제 6항 내지 제 10항 중 어느 한 항에 따른 장치(250)를 포함하고,

제 6항 내지 제 10항 중 어느 한 항에 따른 상기 장치(250)의 상기 오디오인코더(220)는 코어 인코더 입력 데이터를 코어 인코딩하기 위한 코어 인코더(300)이고,

제 6항 내지 제 10항 중 어느 한 항에 따른 상기 장치(250)의 상기 메타데이터 인코더(210; 801; 802)는 상기 복수의 오디오 객체들 중 상기 하나 이상에 관련된 상기 메타데이터를 압축하기 위한 메타데이터 압축기(400)인, 오디오 입력 데이터(101)를 인코딩하기 위한 장치.

청구항 12

시스템으로서,

하나 이상의 인코딩된 오디오 신호들 및 하나 이상의 처리된 메타데이터 신호들을 포함하는 인코딩된 오디오 정보를 생성하기 위한 장치(250), 및

상기 하나 이상의 인코딩된 오디오 신호들 및 상기 하나 이상의 처리된 메타데이터 신호들을 수신하고, 상기 하나 이상의 인코딩된 오디오 신호들에 따라, 그리고 상기 하나 이상의 처리된 메타데이터 신호들에 따라 하나 이상의 오디오 채널들을 생성하기 위한 제 1항 내지 제 4항 중 어느 한 항에 따른 장치(100)를 포함하고,

인코딩된 오디오 정보를 생성하기 위한 상기 장치(250)는,

하나 이상의 원본 메타데이터 신호들을 수신하고 상기 하나 이상의 처리된 메타데이터 신호들을 결정하기 위한 메타데이터 인코더(210; 801; 802)로서, 상기 하나 이상의 원본 메타데이터 신호들 각각은 복수의 원본 메타데이터 샘플들을 포함하고, 상기 하나 이상의 원본 메타데이터 신호들 각각의 상기 원본 메타데이터 샘플들은 하나 이상의 오디오 객체 신호들 중 하나의 오디오 객체 신호와 연관된 정보를 나타내는, 메타데이터 인코더(210; 801; 802), 및

상기 하나 이상의 인코딩된 오디오 신호들을 얻기 위해 상기 하나 이상의 오디오 객체 신호들을 인코딩하기 위한 오디오 인코더(220)를 포함하고,

상기 메타데이터 인코더(210; 801; 802)는 상기 하나 이상의 처리된 메타데이터 신호들(z_1, \dots, z_N)의 각 처리된 메타데이터 신호(z_i)의 복수의 처리된 메타데이터 샘플들($z_i(1), \dots, z_i(n-1), z_i(n)$)의 각 처리된 메타데이터 샘플($z_i(n)$)을 결정하도록 구성되어, 제어 신호(b)가 제 1 상태($b(n)=0$)를 나타낼 때, 상기 처리된 메타데이터 샘플($z_i(n)$)은 상기 하나 이상의 원본 메타데이터 신호들(x_i) 중 하나의 원본 메타데이터 신호의 복수의 원본 메타데이터 샘플들($x_i(n)$) 중의 하나와 상기 처리된 메타데이터 신호(z_i)의 다른 이미 생성된 처리된 메타데이터 샘플의 하나 사이의 차이 또는 양자화된 차이를 나타내고, 상기 제어 신호가 상기 제 1 상태와 상이한 제 2 상태($b(n)=1$)을 나타낼 때, 상기 처리된 메타데이터 샘플($z_i(n)$)은 상기 하나 이상의 원본 메타데이터 신호들(x_i) 중 상기 하나의 원본 메타데이터 신호의 상기 원본 메타데이터 샘플들($x_i(1), \dots, x_i(n)$)의 상기 하나($x_i(n)$)이거나, 상기 원본 메타데이터 샘플들($x_i(1), \dots, x_i(n)$)의 상기 하나($x_i(n)$)의 양자화된 표현($q_i(n)$)인,

시스템.

청구항 13

하나 이상의 오디오 채널들을 생성하기 위한 방법으로서,

제어 신호(b)에 따라 하나 이상의 처리된 메타데이터 신호들(z_1, \dots, z_N)로부터 하나 이상의 재구성된 메타데이터 신호들(x_1', \dots, x_N')을 생성하는 단계로서, 상기 하나 이상의 재구성된 메타데이터 신호들(x_1', \dots, x_N') 각각은 하나 이상의 오디오 객체 신호들 중 하나의 오디오 객체 신호와 연관된 정보를 나타내고, 상기 하나 이상의 재구성된 메타데이터 신호들(x_1', \dots, x_N')을 생성하는 단계는 상기 하나 이상의 재구성된 메타데이터 신호들(x_1', \dots, x_N') 각각에 대해 복수의 재구성된 메타데이터 샘플들($x_1'(n), \dots, x_N'(n)$)을 결정함으로써 수행되는, 하나 이상의 재구성된 메타데이터 신호들(x_1', \dots, x_N')을 생성하는 단계, 및

상기 하나 이상의 오디오 객체 신호들에 따라 그리고 상기 하나 이상의 재구성된 메타데이터 신호들(x_1', \dots, x_N')에 따라 상기 하나 이상의 오디오 채널들을 생성하는 단계를 포함하고,

상기 하나 이상의 재구성된 메타데이터 신호들(x_1', \dots, x_N')을 생성하는 단계는:

상기 하나 이상의 처리된 메타데이터 신호들(z_1, \dots, z_N) 각각의 복수의 처리된 메타데이터 샘플들($z_1(n), \dots, z_N(n)$)을 수신함으로써,

상기 제어 신호(b)를 수신함으로써,

상기 제어 신호(b)가 제 1 상태($b(n)=0$)를 나타내는지 또는 상기 제어 신호(b)가 상기 제 1 상태($b(n)=0$)와 상이한 제 2 상태($b(n)=1$)를 나타내는지 결정하기 위해 상기 제어 신호(b)를 분석함으로써 수행되고,

만약 제어 신호(b)가 제 1 상태(b(n)=0)를 나타내는 경우, 상기 하나 이상의 재구성된 메타데이터 신호들(x_1' , ..., x_N')의 각 재구성된 메타데이터 신호(x_i')의 상기 복수의 재구성된 메타데이터 샘플들($x_i'(1)$, ..., $x_i'(n-1)$, $x_i'(n)$)의 각 재구성된 메타데이터 샘플($x_i'(n)$)이 결정되어, 상기 재구성된 메타데이터 샘플($x_i'(n)$)은 상기 하나 이상의 처리된 메타데이터 신호들(z_i) 중 하나의 처리된 메타데이터 신호의 상기 처리된 메타데이터 샘플들($z_i(n)$) 중 하나와 상기 재구성된 메타데이터 신호(x_i')의 다른 이미 생성된 재구성된 메타데이터 샘플($x_i'(n-1)$)의 합이고,

만약 상기 제어 신호(b)가 상기 제 2 상태(b(n)=1)를 나타내는 경우, 상기 재구성된 메타데이터 샘플($x_i'(n)$)은 상기 하나 이상의 처리된 메타데이터 신호들(z_1, \dots, z_N) 중 상기 하나의 처리된 메타데이터 신호(z_i)의 상기 처리된 메타데이터 샘플들($z_i(1) \dots, z_i(n)$) 중 상기 하나의 처리된 메타데이터 샘플($z_i(n)$)이라고 결정되는,

하나 이상의 오디오 채널들을 생성하기 위한 방법.

청구항 14

하나 이상의 인코딩된 오디오 신호들 및 하나 이상의 처리된 메타데이터 신호들을 포함하는 인코딩된 오디오 정보를 생성하기 위한 방법으로서,

하나 이상의 원본 메타데이터 신호들을 수신하는 단계,

상기 하나 이상의 처리된 메타데이터 신호들을 결정하는 단계,

상기 하나 이상의 인코딩된 오디오 신호들을 얻기 위해 상기 하나 이상의 오디오 객체 신호들을 인코딩하는 단계를 포함하고,

상기 하나 이상의 원본 메타데이터 신호들 각각은 복수의 원본 메타데이터 샘플들을 포함하고, 상기 하나 이상의 원본 메타데이터 신호들 각각의 상기 원본 메타데이터 샘플들은 하나 이상의 오디오 객체 신호들 중 하나의 오디오 객체 신호와 연관된 정보를 나타내고,

상기 하나 이상의 처리된 메타데이터 신호들을 결정하는 단계는 상기 하나 이상의 처리된 메타데이터 신호들(z_1, \dots, z_N)의 각 처리된 메타데이터 신호(z_i)의 복수의 처리된 메타데이터 샘플들($z_i(1), \dots, z_i(n-1), z_i(n)$)의 각 처리된 메타데이터 샘플($z_i(n)$)을 결정하는 단계를 포함하여, 제어 신호(b)가 제 1 상태(b(n)=0)를 나타낼 때, 상기 처리된 메타데이터 샘플($z_i(n)$)은 상기 하나 이상의 원본 메타데이터 신호들 중 하나(x_i)의 복수의 원본 메타데이터 샘플들 중의 하나($x_i(n)$)와 상기 처리된 메타데이터 신호(z_i)의 다른 이미 생성된 처리된 메타데이터 샘플 사이의 차이 또는 양자화된 차이를 나타내고,

상기 제어 신호 (b)가 상기 제 1 상태와 상이한 제 2 상태(b(n)=1)를 나타낼 때, 상기 처리된 메타데이터 샘플 ($z_i(n)$)은 상기 하나 이상의 처리된 메타데이터 신호들(x_i) 중 상기 하나의 원본 메타데이터 신호의 상기 원본 메타데이터 샘플들($x_i(1), \dots, x_i(n)$)의 상기 하나($x_i(n)$)이거나, 상기 원본 메타데이터 샘플들($x_i(1), \dots, x_i(n)$) 중 상기 하나($x_i(n)$)의 양자화된 표현($q_i(n)$)이라고 결정되는,

인코딩된 오디오 정보를 생성하기 위한 방법.

청구항 15

컴퓨터 또는 신호 프로세서 상에서 실행될 때 제 13항 또는 제 14항의 방법을 구현하기 위한 컴퓨터 프로그램을 포함하는 컴퓨터 판독가능 매체.

발명의 설명

기술 분야

본 발명은 오디오 인코딩/디코딩에 관한 것으로, 특히 공간 오디오 코딩 및 공간 오디오 객체 코딩에 관한 것으로, 더 구체적으로 효율적인 객체 메타데이터 코딩을 위한 장치 및 방법에 관한 것이다.

[0001]

배경 기술

- [0002] 공간 오디오 코딩 툴은 종래 기술에 잘 알려져 있고, 예를 들어 MPEG 서라운드 표준에서 표준화된다. 공간 오디오 코딩은 재생 설정에서의 그들의 위치에 의해 식별되는 5 또는 7 채널들, 예를 들어 좌측 채널, 중앙 채널, 우측 채널, 좌측 서라운드 채널, 우측 서라운드 채널 및 저주파수 개선 채널과 같은 원본 입력 채널들로부터 시작한다. 공간 오디오 인코더는 통상적으로 원본 채널들로부터 하나 이상의 다운믹스 채널들을 도출하고, 추가적으로 채널 코히어런스(coherence) 값들, 채널간 위상차, 채널간 시간 차 등에서의 채널간 레벨 차이와 같이 공간 큐와 관련된 파라미터 데이터를 도출한다. 하나 이상의 다운 믹스 채널은 원본 입력 채널의 근사화된 버전인 출력 채널을 결국 얻기 위해 다운 믹스 채널 및 연관된 파라미터 데이터를 디코딩하는 공간 오디오 디코더에 공간 큐를 나타내는 파라메트릭 부가 정보와 함께 전송된다. 출력 설정에서 채널의 배치는 통상적으로 고정되고, 예를 들면, 5.1 포맷, 7.1 포맷 등이다.
- [0003] 이러한 채널 기반 오디오 포맷은, 각 채널이 주어진 위치에서 특정 스피커에 관련하는 다중 채널 오디오 콘텐츠를 저장하거나 전송하기 위해 널리 사용된다. 이러한 종류의 포맷들의 충실한 재생은, 스피커가 오디오 신호의 재생시 사용된 스피커와 동일한 위치에 배치된 스피커 설정을 요구한다. 스피커들의 수가 증가하는 것이 진정 실감나는 3D 오디오 장면의 재생을 개선하지만, 이러한 요건을 충족하는 것이 점점 더 어려워진다 - 특히 거실과 같은 거주 환경에서.
- [0004] 특정 스피커 설정을 가질 필요는, 스피커 신호가 특히 재생 설정을 위해 렌더링되는 객체 기반 접근법에 의해 극복될 수 있다.
- [0005] 예를 들어, 공간 오디오 객체 코딩 툴은 종래 기술에 널리 공지되어 있으며, MPEG SAOC 표준(SAOC = 공간 오디오 객체 코딩)에서 표준화되어 있다. 원본 채널들에서 시작하는 공간 오디오 코딩과 대조적으로, 공간 오디오 객체 코딩은 특정 렌더링 재생 설정에 대해 자동으로 지정되지 않는 오디오 객체에서 시작한다. 대신, 재생 장면의 오디오 객체의 배치는 가요성이고, 공간 오디오 객체 코딩 디코더에 특정 렌더링 정보를 입력하여 사용자에게 의해 결정될 수 있다. 대안적으로 또는 추가적으로, 특정 오디오 객체가 시간이 지남에 따라 일반적으로 배치될 재생 설정에서의 위치를 갖는 정보는 추가적인 부가 정보 또는 메타데이터로서 전송될 수 있다. 특정 데이터 압축을 얻기 위해, 오디오 객체의 개수는 특정 다운믹스 정보에 따라 객체를 다운믹싱하여 입력 객체로부터 하나 이상의 전송 채널을 계산하는 SAOC 인코더에 의해 인코딩된다. 더욱이, SAOC 인코더는 객체 레벨 차이(OLD), 객체 코히어런스 값 등과 같은 객체 간 큐를 나타내는 파라메트릭 부가 정보를 계산한다. SAC(SAC = 공간 오디오 코딩)에서와 같이, 객체 간 파라미터 데이터는 개별적인 시간/주파수 타일들(tiles)에 대해 계산되는데, 즉, 1024 또는 2048개의 샘플들, 24, 32, 64 등을 포함하는 오디오 신호의 특정 프레임에 대해, 주파수 대역은, 결국, 파라메트릭 데이터가 각 프레임 및 각 주파수에 대해 존재하도록 고려된다. 예를 들어, 오디오 피스(piece)가 20 프레임을 가질 때, 그리고 각 프레임이 32개의 주파수 대역으로 세분화될 때, 시간/주파수 타일의 수는 640이다.
- [0006] 객체 기반 방법에서, 음장은 이산 오디오 객체에 의해 기술된다. 이것은 특히 3D 공간에서 각 음원의 시변 위치를 나타내는 객체 메타데이터를 필요로 한다.
- [0007] 종래 기술에서의 제 1 메타데이터 코딩 개념은, 여전히 개발 [1] 하에 있는 오디오 장면 설명, 공간 사운드 설명 교환 포맷(SpatDIF)이다. 이것은 객체 기반 사운드 장면에 대한 교환 포맷으로서 설계되고, 객체 궤적에 대한 임의의 압축 방법을 제공하지 않는다. SpatDIF는 객체 메타데이터 [2]를 구성하기 위해 텍스트 기반의 오픈 사운드 제어(OSC) 포맷을 사용한다. 하지만, 단순한 텍스트 기반의 표현은 객체 궤적의 압축 전송을 위한 옵션이 아니다.
- [0008] 종래 기술의 또 다른 메타데이터 개념은 오디오 장면 설명 포맷(ASDF) [3], 동일한 단점을 갖는 텍스트 기반의 솔루션이다. 데이터는 확장 가능한 마크 업 언어(XML) [4,5]의 서브셋인 동기화 멀티미디어 통합 언어(SMIL)의 확장에 의해 구성된다.
- [0009] 종래 기술에서 추가 메타데이터 개념은 장면(AudioBIFS)을 위한 오디오 이진 포맷, MPEG-4 규격 [6,7]의 부분인 이진 포맷이다. 이것은 시청각 3D 장면과 대화형 가상 현실 응용 [8]의 설명을 위해 개발된 xML 기반의 가상 현실 모델링 언어(VRML)와 밀접하게 관련있다. 복합 AudioBIFS 규격은 객체의 움직임 경로를 규정하기 위해 장면 그래프를 사용한다. AudioBIFS의 주요 단점은, 데이터 스트림에 대한 제한 시스템 지연 및 랜덤 액세스가 요구되는 실시간 동작을 위해 설계되지 않는다. 또한, 객체의 위치의 인코딩은 인간 청취자의 제한된 국부화 성능을 이용하지 않는다. 시청각 장면 내의 고정 청취자 위치의 경우, 객체 데이터는 비트의 더 낮은 수로 양자화될 수

있다 [9]. 따라서 AudioBIFS에 인가된 객체 메타데이터의 인코딩은 데이터 압축에 대해서는 효율적이지 못하다.

[0010] 따라서 개선된 경우, 효율적인 객체 메타데이터 코딩 개념이 제공되는 것이 크게 인식된다.

발명의 내용

[0011] 본 발명의 목적은, 제 1항에 따른 장치, 제 6항에 따른 장치, 제 12항에 따른 시스템, 제 13항에 따른 방법, 제 14항에 따른 방법, 및 제 15항에 따른 컴퓨터 프로그램에 의해 해결된다. 하나 이상의 오디오 채널을 생성하기 위한 장치가 제공된다.

[0012] 장치는 제어 신호(b)에 따라 하나 이상의 처리된 메타데이터 신호들(z_1, \dots, z_N)로부터 하나 이상의 재구성된 메타데이터 신호(x_1', \dots, x_N')를 생성하기 위한 메타데이터 디코더를 포함하고, 하나 이상의 재구성된 메타데이터 신호(x_1', \dots, x_N') 각각은 하나 이상의 오디오 객체 신호의 오디오 객체 신호와 연관된 정보를 나타내고, 메타데이터 디코더는 하나 이상의 재구성된 메타데이터 신호(x_1', \dots, x_N') 각각에 대해 복수의 재구성된 메타데이터 샘플($x_1'(n), \dots, x_N'(n)$)을 결정함으로써 하나 이상의 재구성된 메타데이터 신호(x_1', \dots, x_N')를 생성하도록 구성된다. 더욱이, 장치는 하나 이상의 오디오 객체 신호에 따라 그리고 하나 이상의 재구성된 메타데이터 신호(x_1', \dots, x_N')에 따라 하나 이상의 오디오 채널을 생성하기 위한 오디오 채널 생성기를 포함한다. 메타데이터 디코더는 하나 이상의 처리된 메타데이터 신호들(z_1, \dots, z_N) 각각의 복수의 처리된 메타데이터 샘플($z_1(n), \dots, z_N(n)$)을 수신하도록 구성된다. 또한, 메타데이터 디코더는 제어 신호(b)를 수신하도록 구성된다.

[0013] 또한, 메타데이터 디코더는 하나 이상의 재구성된 메타데이터 신호(x_1', \dots, x_N')의 각 재구성된 메타데이터 신호(x_i')의 복수의 재구성된 메타데이터 샘플($x_i'(1), \dots, x_i'(n-1), x_i'(n)$)의 각 재구성된 메타데이터 샘플($x_i'(n)$)을 결정하도록 구성되어, 제어 신호(b)가 제 1 상태($b(n)=0$)를 나타낼 때, 상기 재구성된 메타데이터 샘플($x_i'(n)$)은 하나 이상이 처리된 메타데이터 신호(z_i)의 하나의 처리된 메타데이터 샘플($z_i(n)$) 중 하나와 상기 재구성된 메타데이터 신호(x_i')의 다른 이미 생성된 재구성된 메타데이터 샘플($x_i'(n-1)$)의 합이고, 제어 신호가 제 1 상태와 상이한 제 2 상태($b(n)=1$)를 나타낼 때, 상기 재구성된 메타데이터 샘플($x_i'(n)$)은 하나 이상의 처리된 메타데이터 신호들(z_1, \dots, z_N)의 상기 하나(z_i)의 처리된 메타데이터 샘플($z_i(1), \dots, z_i(n)$)의 상기 하나($z_i(n)$)이다.

[0014] 또한, 하나 이상의 인코딩된 오디오 신호 및 하나 이상의 처리된 메타데이터 신호를 포함하는 인코딩된 오디오 정보를 생성하기 위한 장치가 제공된다. 장치는 하나 이상의 원본 메타데이터 신호를 수신하고, 하나 이상의 원본 메타데이터 신호를 결정하기 위한 메타데이터 인코더를 포함하고, 하나 이상의 메타데이터 신호 각각은 복수의 원본 메타데이터 샘플을 포함하고, 하나 이상의 원본 메타데이터 신호 각각의 원본 메타데이터 샘플은 하나 이상의 오디오 객체 신호의 오디오 객체 신호와 연관된 정보를 나타낸다.

[0015] 또한, 장치는 하나 이상의 인코딩된 오디오 신호를 얻기 위해 하나 이상의 오디오 객체 신호를 인코딩하기 위한 오디오 인코더를 포함한다.

[0016] 메타데이터 인코더는 하나 이상의 처리된 메타데이터 신호(z_i, \dots, z_N)의 각 처리된 메타데이터 신호(z_i)의 복수의 처리된 메타데이터 샘플들($z_i(1), \dots, z_i(n-1), z_i(n)$)의 각 처리된 메타데이터 샘플($z_i(n)$)을 결정하도록 구성되어, 제어 신호(b)가 제 1 상태($b(n)=0$)를 나타낼 때, 상기 재구성된 메타데이터 샘플($z_i(n)$)은 하나 이상의 원본 메타데이터 신호들(x_i) 중 하나의 원본 메타데이터 신호의 복수의 원본 메타데이터 샘플들($x_i(n)$) 및 상기 처리된 메타데이터 신호(z_i)의 다른 이미 생성된 처리된 메타데이터 샘플의 하나 사이의 차이 또는 양자화된 차이를 나타내고, 제어 신호가 제 1 상태와 상이한 제 2 상태($b(n)=1$)를 나타낼 때, 상기 처리된 메타데이터 샘플($z_i(n)$)은 하나 이상의 처리된 메타데이터 신호들(x_i) 중 상기 하나의 원본 메타데이터 신호의 상기 원본 메타데이터 샘플들($x_i(1), \dots, x_i(n)$)의 상기 하나($x_i(n)$)이거나, 상기 원본 메타데이터 샘플들($x_i(1), \dots, x_i(n)$)의 상기 하나($x_i(n)$)의 양자화된 표현($q_i(n)$)이다.

- [0017] 실시예들에 따라, 객체 메타데이터를 위한 데이터 압축 개념들이 제공되고, 이것은 제한된 데이터율로 전송 채널들을 위한 효율적인 압축 메커니즘을 달성한다. 각각 인코더 및 디코더에 의해 추가 지연이 도입되지 않는다. 더욱이, 순수한 방위각 변화들, 예를 들어 카메라 회전들에 대한 양호한 압축률이 달성된다. 더욱이, 제공된 개념들은 불연속적인 궤적들, 예를 들어 위치 도약들을 지원한다. 더욱이, 낮은 디코딩 복잡도가 실현된다. 더욱이, 제한된 재초기화 시간을 갖는 랜덤 액세스가 달성된다.
- [0018] 더욱이, 하나 이상의 오디오 채널들을 생성하기 위한 방법이 제공된다. 방법은:
- [0019] - 제어 신호(b)에 따라 하나 이상의 처리된 메타데이터 신호들(z_1, \dots, z_N)로부터 하나 이상의 재구성된 메타데이터 신호들(x_1', \dots, x_N')을 생성하는 단계로서, 하나 이상의 재구성된 메타데이터 신호(x_1', \dots, x_N') 각각은 하나 이상의 오디오 객체 신호들의 오디오 객체 신호와 연관된 정보를 나타내고, 하나 이상의 재구성된 메타데이터 신호들(x_1', \dots, x_N')을 생성하는 단계는 상기 하나 이상의 재구성된 메타데이터 신호들(x_1', \dots, x_N') 각각에 대해 복수의 재구성된 메타데이터 샘플($x_1'(n), \dots, x_N'(n)$)을 결정함으로써 수행되는, 하나 이상의 재구성된 메타데이터 신호들(x_1', \dots, x_N')을 생성하는 단계, 및
- [0020] - 하나 이상의 오디오 객체 신호들에 따라 그리고 하나 이상의 재구성된 메타데이터 신호들(x_1', \dots, x_N')에 따라 하나 이상의 오디오 채널들을 생성하는 단계를 포함한다.
- [0021] 하나 이상의 재구성된 메타데이터 신호들(x_1', \dots, x_N')을 생성하는 단계는 하나 이상의 처리된 메타데이터 신호들(z_1, \dots, z_N) 각각의 복수의 처리된 메타데이터 샘플($z_1(n), \dots, z_N(n)$)을 수신함으로써, 제어 신호(b)를 수신함으로써, 그리고 하나 이상의 재구성된 메타데이터 신호(x_1', \dots, x_N')의 각 재구성된 메타데이터 신호(x_i')의 복수의 재구성된 메타데이터 샘플($x_i'(1), \dots, x_i'(n-1), x_i'(n)$)의 각 재구성된 메타데이터 샘플($x_i'(n)$)을 결정함으로써 수행되어, 제어 신호(b)가 제 1 상태($b(n)=0$)를 나타낼 때, 재구성된 메타데이터 샘플($x_i'(n)$)은 하나 이상의 처리된 메타데이터 신호(z_i) 중 하나의 처리된 메타데이터 신호의 상기 처리된 메타데이터 샘플($z_i(n)$) 중 하나와 재구성된 메타데이터 신호(x_i')의 다른 이미 생성된 재구성된 메타데이터 샘플($x_i'(n-1)$)의 합이고, 제어 신호가 제 1 상태와 상이한 제 2 상태($b(n)=1$)를 나타낼 때, 상기 재구성된 메타데이터 샘플($x_i'(n)$)은 하나 이상의 처리된 메타데이터 신호들(z_1, \dots, z_N)의 하나의 처리된 메타데이터 신호(z_i)의 처리된 메타데이터 샘플($z_i(1) \dots, z_i(n)$)의 하나의 처리된 메타데이터 샘플($z_i(n)$)이다.
- [0022] 더욱이, 하나 이상의 인코딩된 오디오 신호들 및 하나 이상의 처리된 메타데이터 신호들을 포함하는 인코딩된 오디오 정보를 생성하기 위한 방법이 제공된다. 방법은:
- [0023] 하나 이상의 원본 메타데이터 신호들을 수신하는 단계,
- [0024] 하나 이상의 처리된 메타데이터 신호들을 결정하는 단계,
- [0025] 하나 이상의 인코딩된 오디오 신호들을 얻기 위해 하나 이상의 오디오 객체 신호들을 인코딩하는 단계를 포함한다.
- [0026] 하나 이상의 원본 메타데이터 신호들 각각은 복수의 원본 메타데이터 샘플들을 포함하고, 하나 이상의 원본 메타데이터 신호들 각각의 원본 메타데이터 샘플들은 하나 이상의 오디오 객체 신호들의 오디오 객체 신호와 연관된 정보를 나타낸다. 하나 이상의 처리된 메타데이터 신호들을 결정하는 단계는 하나 이상의 처리된 메타데이터 신호(z_i, \dots, z_N)의 각 처리된 메타데이터 신호(z_i)의 복수의 처리된 메타데이터 샘플들($z_i(1), \dots, z_i(n-1), z_i(n)$)의 각 처리된 메타데이터 샘플($z_i(n)$)을 결정하는 단계를 포함하여, 제어 신호(b)가 제 1 상태($b(n)=0$)를 나타낼 때, 상기 재구성된 메타데이터 샘플($z_i(n)$)은 상기 하나 이상의 원본 메타데이터 신호들(x_i) 중 하나의 원본 메타데이터 신호의 복수의 원본 메타데이터 샘플들($x_i(n)$) 및 처리된 메타데이터 신호(z_i)의 다른 이미 생성된 처리된 메타데이터 샘플의 하나 사이의 차이 또는 양자화된 차이를 나타내고, 제어 신호가 제 1 상태와 상이한 제 2 상태($b(n)=1$)를 나타낼 때, 상기 처리된 메타데이터 샘플($z_i(n)$)은 하나 이상의 처리된 메타데이터 신호들(x_i) 중 상기 하나의 원본 메타데이터 신호의 원본 메타데이터 샘플들($x_i(1), \dots, x_i(n)$)의 상기 하나($x_i(n)$)이거나, 원본 메타데이터 샘플들($x_i(1), \dots, x_i(n)$)의 상기 하나($x_i(n)$)의 양자화된 표현($q_i(n)$)이다.

- [0027] 또, 컴퓨터 또는 신호 프로세서에서 실행될 때 전문한 방법을 구현하기 위한 컴퓨터 프로그램이 제공된다.
- [0028] 다음에서, 본 발명의 실시예들은 도면들을 참조하여 더 구체적으로 기재된다.

도면의 간단한 설명

- [0029] 도 1은 실시예에 따라 하나 이상의 오디오 채널들을 생성하기 위한 장치를 도시한 도면.
- 도 2는 실시예에 따라 인코딩된 오디오 정보를 생성하기 위한 장치를 도시한 도면.
- 도 3은 실시예에 따른 시스템을 도시한 도면.
- 도 4는 방위각, 상승각 및 반경에 의해 표현된 원점으로부터 3차원 공간에서의 오디오 객체의 위치를 도시한 도면.
- 도 5는 오디오 채널 생성기에 의해 간주된 오디오 객체들 및 스피커 설정의 위치를 도시한 도면.
- 도 6은 차동 펄스 코드 변조 인코더를 도시한 도면.
- 도 7은 차동 펄스 코드 변조 디코더를 도시한 도면.
- 도 8a는 실시예에 따른 메타데이터 인코더를 도시한 도면.
- 도 8b는 다른 실시예에 따른 메타데이터 인코더를 도시한 도면.
- 도 9a는 실시예에 따른 메타데이터 디코더를 도시한 도면.
- 도 9b는 실시예에 따른 메타데이터 디코더 서브유닛을 도시한 도면.
- 도 10은 3D 오디오 인코더의 제 1 실시예를 도시한 도면.
- 도 11은 3D 오디오 디코더의 제 1 실시예를 도시한 도면.
- 도 12는 3D 오디오 인코더의 제 2 실시예를 도시한 도면.
- 도 13은 3D 오디오 디코더의 제 2 실시예를 도시한 도면.
- 도 14는 3D 오디오 인코더의 제 3 실시예를 도시한 도면.
- 도 15는 3D 오디오 인코더의 제 3 실시예를 도시한 도면.

발명을 실시하기 위한 구체적인 내용

- [0030] 도 2는 일 실시예에 따라 하나 이상의 인코딩된 오디오 신호와 하나 이상의 메타데이터 신호들을 포함하는 인코딩된 오디오 정보를 생성하기 위한 장치 (250)를 도시한다. 장치(250)는 하나 이상의 원본 메타데이터 신호를 수신하고, 하나 이상의 원본 메타데이터 신호들을 결정하기 위한 메타데이터 인코더(210)를 포함하고, 하나 이상의 원본 메타데이터 신호들 각각은 복수의 원본 메타데이터 샘플들을 포함하고, 하나 이상의 원본 메타데이터 신호들 각각의 원본 메타데이터 샘플들은 하나 이상의 오디오 객체 신호의 오디오 객체 신호와 관련된 정보를 나타낸다.
- [0031] 또한, 장치(250)는 하나 이상의 인코딩된 오디오 신호를 얻기 위해 하나 이상의 오디오 객체 신호를 인코딩하기 위한 오디오 인코더(220)를 포함한다. 메타데이터 인코더(210)는 하나 이상의 메타데이터 신호(z_i, \dots, z_N)의 각 처리된 메타데이터 신호(z_i)의 복수의 처리된 메타데이터 샘플들($z_i(1), \dots, z_i(n-1), z_i(n)$)의 각 처리된 메타데이터 샘플($z_i(n)$)을 결정하도록 구성되어, 제어 신호(b)가 제 1 상태($b(n)=0$)를 나타낼 때, 상기 재구성된 메타데이터 샘플($z_i(n)$)은 하나 이상의 원본 메타데이터 신호들(x_i) 중 하나의 원본 메타데이터 신호의 복수의 원본 메타데이터 샘플들($x_i(n)$) 및 상기 처리된 메타데이터 신호(z_i)의 다른 이미 생성된 처리된 메타데이터 샘플 중 하나 사이의 차이 또는 양자화된 차이를 나타내고, 제어 신호가 제 2 상태($b(n)=1$)를 나타낼 때, 상기 처리된 메타데이터 샘플($z_i(n)$)은 하나 이상의 처리된 메타데이터 신호들(x_i) 중 상기 하나의 원본 메타데이터 신호의 원본 메타데이터 샘플들($x_i(n), \dots, x_i(n)$)의 상기 하나($x_i(n)$)이거나, 원본 메타데이터 샘플들($x_i(n), \dots, x_i(n)$)의 상기 하나($x_i(n)$)의 양자화된 표현($q_i(n)$)이다.

- [0032] 도 1은 실시예에 따라, 하나 이상의 오디오 채널을 생성하기 위한 장치(100)를 도시한다.
- [0033] 장치(100)는 제어 신호(b)에 따라 하나 이상의 처리된 메타데이터 신호(z_1, \dots, z_N)로부터 하나 이상의 재구성된 메타데이터 신호들(x_1', \dots, x_i')을 생성하기 위한 메타데이터 디코더(110)를 포함하고, 하나 이상의 재구성된 메타데이터 신호들(x_1', \dots, x_i') 각각은 하나 이상의 오디오 객체 신호들의 오디오 객체 신호와 관련된 정보를 나타내고, 메타데이터 디코더(110)는 하나 이상의 재구성된 메타데이터 신호들(x_1', \dots, x_i') 각각에 대해 복수의 재구성된 메타데이터 샘플들($x_1'(n), \dots, x_i'(n)$)을 결정함으로써 하나 이상의 재구성된 메타데이터 신호들(x_1', \dots, x_i')을 생성하도록 구성된다. 또한, 장치(100)는 하나 이상의 오디오 객체 신호들에 따라 그리고 하나 이상의 재구성된 메타데이터 신호들(x_1', \dots, x_i')에 따라 하나 이상의 오디오 채널을 생성하기 위한 오디오 채널 생성기(120)를 포함한다.
- [0034] 메타데이터 디코더(110)는 하나 이상의 처리된 메타데이터 신호들(z_1, \dots, z_N)의 복수의 처리된 메타데이터 샘플들($z_1(n), \dots, z_N(n)$)을 수신하도록 구성된다. 더욱이, 메타데이터 디코더(110)는 제어 신호(b)를 수신하도록 구성된다.
- [0035] 또한, 메타데이터 디코더(110)는 하나 이상의 인코딩된 오디오 신호를 얻기 위해 하나 이상의 오디오 객체 신호를 인코딩하기 위한 오디오 인코더(220)를 포함한다.
- [0036] 더욱이, 메타데이터 디코더(110)는 하나 이상의 재구성된 메타데이터 신호(x_1', \dots, x_N')의 각 재구성된 메타데이터 신호(x_i')의 복수의 재구성된 메타데이터 샘플들($x_1'(1), \dots, x_i'(n-1), x_i'(n)$)의 각 재구성된 메타데이터 샘플($x_i'(n)$)을 결정하도록 구성되어, 제어 신호(b)가 제 1 상태($b(n)=0$)를 나타낼 때, 상기 재구성된 메타데이터 샘플($x_i'(n)$)은 하나 이상의 처리된 메타데이터 신호들(z_1, \dots, z_N) 중 하나의 처리된 메타데이터 신호의 하나의 처리된 메타데이터 샘플들($z_i(n)$)의 하나 및 상기 재구성된 메타데이터 신호(x_1')의 다른 이미 생성된 재구성된 메타데이터 샘플($x_1'(n-1)$)의 합이고, 제어 신호가 제 2 상태($b(n)=1$)를 나타낼 때, 상기 재구성된 메타데이터 샘플($x_i'(n)$)은 하나 이상의 처리된 메타데이터 신호들(z_1, \dots, z_N) 중 상기 하나(z_i)의 처리된 메타데이터 샘플들($z_i(1), \dots, z_i(n)$)의 상기 하나($z_i(n)$)이다.
- [0037] 메타데이터 샘플을 참조하면, 메타데이터 샘플이 메타데이터 샘플 값뿐 아니라, 이에 관련되는 시간의 인스턴트(instant)에 의해 특징되는 것을 주목해야 한다. 예를 들어, 그러한 시간 인스턴트는 오디오 시퀀스 또는 유사한 것의 시작에 상대적일 수 있다. 예를 들어, 인덱스(n 또는 k)는 메타데이터 신호에서 메타데이터 샘플의 위치를 식별할 수 있고, 이에 의해, 시간(시작 시간에 상대적임)의 (상대적인) 인스턴트가 표시된다. 두 개의 메타데이터 샘플들이 상이한 시간 인스턴트들에 관련될 때, 상이한 메타데이터 샘플들이고, 심지어 메타데이터 샘플 값들이 동일하더라도, 종종 그러할 수 있다는 것을 주목해야 한다.
- [0038] 상기 실시예는 오디오 객체 신호와 연관된 메타데이터 정보(메타데이터 신호로 구성)가 종종 느리게 변한다는 사실에 기초한다.
- [0039] 예를 들어, 메타데이터 신호는 오디오 객체(예를 들어, 방위각, 양각 또는 오디오 객체의 위치를 정의하는 반경)의 위치 정보를 나타낼 수 있다. 대부분의 시간에, 오디오 객체의 위치가 변하지 않거나 단지 느리게 변하는 것이 가정될 수 있다.
- [0040] 또는 메타데이터 신호는 예를 들어, 오디오 객체의 볼륨(예를 들어, 이득)을 나타낼 수 있고, 또한, 대부분의 시간에 오디오 객체의 볼륨이 느리게 변한다는 것이 가정될 수 있다.
- [0041] 이 때문에, 모든 시간의 인스턴트에 (완전한) 메타데이터 정보를 송신할 필요가 없다.
- [0042] 대신에, (완전한) 메타데이터 정보는, 예를 들어, 일부 실시예에 따라 단지 특정 시간 인스턴트에서, 예를 들어 주기적으로, 예를 들어, 시간의 모든 N 번째 인스턴트에서, 예를 들어, 시간 0, N, 2N, 3N, 등에서의 지점에서 송신될 수 있다.
- [0043] 예를 들어, 실시예에서, 메타데이터 신호는 3D 공간에서의 오디오 객체의 위치를 지정한다. 메타데이터 신호들 중 제 1 신호는, 예를 들어, 오디오 객체의 위치의 방위각을 지정할 수 있다. 메타데이터 신호들 중 제 2 신호

는, 예를 들어, 오디오 객체의 위치의 양각을 지정할 수 있다. 메타데이터신호들 중 제 3 신호는, 예를 들어, 오디오 객체의 거리에 관한 반경을 지정할 수 있다.

- [0044] 방위각, 양각 및 반경은 원점으로부터 3D 공간에서 오디오 객체의 위치를 모호하게 정의한다. 이는 도 4를 참조하여 도시된다.
- [0045] 도 4는 방위각, 양각 및 반경에 의해 표현된 원점(400)으로부터 3차원(3D) 공간에서의 오디오 객체의 위치(410)를 도시한다.
- [0046] 양각은, 예를 들면, 원점으로부터 객체 위치로의 직선과 xy 평면(x 축 및 Y 축에 의해 정의되는 평면) 상으로의 이 직선의 법선 투사 사이의 각도를 지정한다. 방위각은, 예를 들면, x 축과 상기 법선 투사 사이의 각도를 정의한다. 방위각과 양각을 지정함으로써, 기점(400) 및 오디오 객체의 위치(410)를 통한 직선(415)이 정의될 수 있다. 또한 반경을 지정함으로써, 오디오 객체의 정확한 위치(410)가 정의될 수 있다.
- [0047] 실시예에서, 방위각은 범위에 대해 정의된다: $-180^\circ < \text{방위각} \leq 180^\circ$, 양각은 범위에 대해 정의된다: $90^\circ \leq \text{양각} \leq -90^\circ$ 및 반경은 예를 들어 미터[m](0m보다 크거나 같은) 단위로 정의될 수 있다.
- [0048] 다른 실시예에서, 예를 들면, xyz 좌표계에서 오디오 객체 위치의 모든 x 값이 제로보다 크거나 같은 것으로 간주 될 수 있는 경우, 방위각은 범위에 대해 정의될 수 있고: $-90^\circ \leq \text{방위각} \leq 90^\circ$, 양각은 범위에 대해 정의될 수 있고: $-90^\circ \leq \text{양각} \leq 90^\circ$, 반경은, 예를 들면, 미터[m] 단위로 정의될 수 있다.
- [0049] 다른 실시예에서, 메타데이터 신호는, 방위각이 범위에 대해 정의될 수 있고: $-128^\circ \leq \text{방위각} \leq 128^\circ$, 양각이 범위에 대해 정의될 수 있고: $-32^\circ \leq \text{양각} \leq 32^\circ$, 반경이, 예를 들면, 로그 스케일 상에서 정의될 수 있도록 스케일링될 수 있다. 일부 실시예에서, 원본 메타데이터 신호, 처리된 메타데이터 신호 및 재구성된 메타데이터 신호 각각은 하나 이상의 오디오 객체 신호들 중 하나의 볼륨의 스케일링된 정보 및/또는 위치 정보의 스케일링된 표현을 포함 할 수 있다.
- [0050] 오디오 채널 생성기(120)는, 예를 들어, 하나 이상의 오디오 객체 신호에 따라 그리고 재구성된 메타데이터 신호에 따라, 하나 이상의 오디오 채널을 생성하도록 구성될 수 있고, 재구성된 메타데이터 신호는 예를 들어, 오디오 객체의 위치를 나타낼 수 있다.
- [0051] 도 5는 오디오 채널 생성기에 의해 가정된 오디오 객체의 위치 및 스피커 설정을 도시한다. xyz 좌표계의 원점(500)이 도시된다. 또한, 제 1 오디오 객체의 위치(510)와 제 2 오디오 객체의 위치(520)가 도시되어 있다. 또한, 도 5는, 오디오 채널 생성기(120)가 네 개의 스피커를 위한 네 개의 오디오 채널을 생성하는 경우 시나리오를 도시한다. 오디오 채널 생성기(120)는 네 개의 스피커(511, 512, 513 및 514)가 도 5에 도시된 위치에 배치되어 있다고 가정한다.
- [0052] 도 5에서, 제 1 오디오 객체는 스피커(511 및 512)의 가정된 위치에 가까운 위치(510)에 위치되고, 스피커(513 및 514)로부터 멀리 떨어져 위치된다. 그러므로, 오디오 채널 생성기(120)는, 제 1 오디오 객체(510)가 스피커(511 및 512)에 의해 재생되지만, 스피커(513 및 514)에 의해 재생되지 않도록 4개의 오디오 채널들을 생성할 수 있다.
- [0053] 다른 실시예에서, 오디오 채널 생성기(120)는, 제 1 오디오 객체(510)가 스피커(511 및 512)에 의해 높은 볼륨으로 그리고 스피커(513 및 514)에 의해 낮은 볼륨으로 재생되도록 4개의 오디오 채널을 생성할 수 있다.
- [0054] 또한, 제 2 오디오 객체는 스피커(513 및 514)의 가정된 위치에 가까운 위치(520)에 위치되고, 스피커(511 및 512)로부터 멀리 떨어지게 위치된다. 그러므로, 오디오 채널 생성기(120)는 제 2 오디오 객체(520)가 스피커(513 및 514)에 의해 재생되지만, 스피커(511 및 512)에 의해 재생되지 않도록 4개의 오디오 채널들을 생성할 수 있다.
- [0055] 다른 실시예에서, 오디오 채널 생성기(120)는, 제 2 오디오 객체(520)가 스피커(513 및 514)에 의해 높은 볼륨으로 그리고 스피커(511 및 512)에 의해 낮은 볼륨으로 재생되도록 4개의 오디오 채널을 생성할 수 있다.
- [0056] 대안적인 실시예에서, 두 개의 메타데이터 신호만이 오디오 객체의 위치를 지정하는데 사용된다. 예를 들어, 모든 오디오 객체가 단일 평면 내에 위치되는 것으로 가정할 때, 단지 방위각과 반경은, 예를 들어, 지정될 수 있다.
- [0057] 또 다른 실시예에서, 각각의 오디오 객체에 대해, 단일의 메타데이터 신호가 인코딩되고, 위치 정보로서 송신된다. 예를 들어, 단지 방위각은 오디오 객체(예를 들어, 모든 오디오 객체는 중심점으로부터 동일한 거리를 갖는

동일 평면상에 위치하며, 따라서 동일한 반경을 갖는 것으로 가정되는 것이 가정될 수 있다)에 대한 위치 정보로서 지정될 수 있다. 방위 정보는, 예를 들어, 오디오 객체가 좌측 스피커에 가깝고 우측 스피커로부터 멀리 위치하는지를 결정하기에 충분할 수 있다. 이러한 상황에서, 오디오 채널 생성기(120)는, 예를 들어, 오디오 객체가 좌측 스피커 뿐 아니라 우측 스피커에 의해 재생되도록 하나 이상의 오디오 채널을 생성할 수 있다.

[0058] 예를 들어, 벡터 기반 진폭 패닝(VBAP)은, 스피커의 오디오 채널 각각 내에서의 오디오 객체 신호의 가중치를 결정하는데 이용될 수 있다(예를 들어, [11]을 참조). 예를 들어, VBAP에 대하여, 오디오 객체가 가상 소스에 관한 것임이 가정된다.

[0059] 실시예에서, 추가 메타데이터 신호는 볼륨을 지정할 수 있고, 예를 들면, 각 오디오 객체에 대한 이득(예를 들어, 데시벨 [dB]로 표시)을 지정할 수 있다.

[0060] 예를 들어, 도 5에서, 제 1 이득 값은 위치(520)에 위치한 제 2 오디오 객체에 대한 다른 추가의 메타데이터 신호에 의해 지정된 제 2 이득 값보다 높은 위치(510)에 위치한 제 1 오디오 객체에 대한 추가 메타데이터 신호에 의해 지정될 수 있다. 그러한 상황에서, 스피커(511 및 512)는, 스피커(513 및 514)가 제 2 오디오 객체를 재생하는 볼륨보다 더 높은 볼륨을 갖는 제 1 오디오 객체를 재생할 수 있다.

[0061] 실시예는, 또한 오디오 객체의 그러한 이득 값이 종종 느리게 변화한다고 가정한다. 따라서, 시간의 모든 지점에서 이러한 메타데이터 정보를 송신할 필요가 없다. 대신에, 메타데이터 정보는 시간의 특정 지점에서만 전송된다. 시간의 중간 지점에서, 메타데이터 정보는, 예를 들어, 송신된, 이전의 메타데이터 샘플 및 후속 메타데이터 샘플을 사용하여 근사화될 수 있다. 예를 들어, 선형 보간은 중간 값들의 근사에 이용될 수 있다. 예를 들어, 각 오디오 객체들의 이득, 방위각, 양각 및/또는 반경은, 메타 데이터가 송신되지 않은 경우 시간의 지점에 대해 근사화될 수 있다.

[0062] 이러한 접근법에 의해, 메타데이터의 송신률에 상당한 절감이 달성될 수 있다.

[0063] 도 3은 실시예에 따른 시스템을 도시한다. 시스템은 전술한 바와 같이 하나 이상의 인코딩된 오디오 신호와 하나 이상의 처리된 메타데이터 신호들을 포함하는 인코딩된 오디오 정보를 생성하기 위한 장치(250)를 포함한다. 또한, 시스템은 하나 이상의 인코딩된 음성 신호 및 상기 하나 이상의 처리된 메타데이터 신호를 수신하고, 하나 이상의 인코딩된 오디오 신호에 따라 그리고 하나 이상의 처리된 메타데이터 신호에 따라 하나 이상의 오디오 채널들을 생성하기 위한 장치(100)를 포함한다. 예를 들어, 하나 이상의 인코딩된 오디오 신호는, 인코딩을 위한 장치(250)가 하나 이상의 오디오 객체를 인코딩하기 위한 SAOC 인코더를 사용할 때, 하나 이상의 오디오 객체 신호를 획득하기 위해 종래 기술에 따른 SAOC 디코더를 이용함으로써 하나 이상의 오디오 채널을 생성하기 위한 장치(100)에 의해 디코딩될 수 있다. 실시예들은, 차동 펄스 코드 변조 개념이 확장될 수 있고, 이러한 확장된 개념이 오디오 객체에 대한 메타데이터 신호를 인코딩하기에 적합하다는 발견에 기초한다. 차동 펄스 코드 변조(DPCM) 방법은 차동 송신 [10]을 통해 양자화 및 리던던시(redundancy)를 통해 무관함(irrelevance)을 감소하는 시간 신호를 느리게 변화시키기 위한 확립된 방법이다. DPCM 인코더는 도 6에 도시되어 있다. 도 6의 DPCM 인코더에서, 입력 신호(x)의 실제 입력 샘플{x(n)}은 감산 유닛(610)에 공급된다. 감산 유닛의 다른 입력에서, 다른 값이 감산 유닛에 공급된다. 또한 이 다른 값이 이전에 수신된 샘플{x(n-1)}이지만, 양자화 예러 또는 다른 예러가 다른 입력에서의 값이 이전 샘플{x(n-1)}에 정확하게 일치하지 않는 결과를 가질 수 있다고 가정될 수 있다. x(n-1)로부터 그러한 가능한 편차로 인해, 감산기의 다른 입력은 $x^*(n-1)$ 으로 언급될 수 있다. 감산 유닛은 상이한 차동 값{d(n)}을 얻기 위해 x(n)로부터 $x^*(n-1)$ 을 감산한다. d(n)은 출력 신호(y)의 다른 출력 샘플{y(n)}을 획득하기 위해 양자화기(620)에서 양자화된다. 일반적으로, y(n)은 d(n)과 동일하거나, d(n)에 가까운 값이다. 또한, y(n)은 또한 가산기(630)로 공급된다. 더욱이, $x^*(n-1)$ 은 가산기(630)에 공급된다. d(n)이 감산기(n) = x(n) - $x^*(n-1)$ 으로부터 초래될 때, 그리고 y(n)이 d(n)과 동일하거나 적어도 가까운 값일 때, 가산기(630)의 출력{x'(n)}은 x(n)과 동일하거나, x(n)에 적도 가깝다. $x^*(n)$ 은 유닛(640)의 샘플링 기간 동안 유지되고, 그 후, 처리는 다음 샘플{x(N + 1)}로 계속된다. 도 7은 대응하는 DPCM 디코더를 도시한다. 도 7에서, DPCM 인코더로부터의 출력 신호(y)의 샘플{y(n)}은 가산기(710)에 공급된다. y(n)은 재구성될 신호{x(n)}의 차분 값을 나타낸다. 가산기(710)의 다른 입력에, 이전에 재구성된 샘플{x'(n-1)}은 가산기(710)에 공급된다. 가산기의 출력{x'(n)}은 가산기{x'(n) = x'(n-1) + y(n)}으로부터 초래된다. x'(n-1)이 일반적으로 x(n-1)과 동일하거나 이에 적어도 가까울 때, 그리고 y(n)이 일반적으로 x(n)-x(n-1)과 동일하거나 이에 가까울 때, 가산기(710)의 출력{x'(n)}은 일반적으로 x(n)과 동일하거나 이에 가깝다. x'(n)은 유닛(740)의 샘플링 기간 동

안 유지되고, 그 후, 처리는 다음의 샘플 $\{y_{(n+1)}\}$ 로 계속된다. DPCM 압축 방법이 앞서 언급된 필요한 대부분의 특징을 충족하지만, 랜덤 액세스를 허용하지 않는다. 도 8a는 일 실시예에 따른 메타데이터 인코더(801)를 도시한다. 도 8a의 메타데이터 인코더(801)에 의해 이용된 인코딩 방법은 전형적인 DPCM 인코딩 방법의 확장이다.

[0064] 도 8a의 메타데이터 인코더(801)는 하나 이상의 DPCM 인코더(811, ..., 81N)을 포함한다. 예를 들어, 메타데이터 인코더(801)가 N 원본 메타데이터 신호를 수신하도록 구성될 때, 메타데이터 인코더(801)는, 예를 들어, 정확하게 N DPCM 인코더를 포함할 수 있다. 실시예에서, N DPCM 인코더는 도 6에 대해 기재된 바와 같이 구현된다.

[0065] 실시예에서, N DPCM 인코더 각각은 N 원본 메타데이터 신호 $\{x_1, \dots, x_n\}$ 중 하나의 원본 메타데이터 신호의 메타데이터 샘플 $\{x_i(n)\}$ 을 수신하도록 구성되고, 상기 DPCM 인코더에 공급되는 상기 원본 메타데이터 신호 $\{x_i\}$ 의 메타데이터 샘플 $\{x_i(n)\}$ 각각에 대한 메타데이터 차이 신호 $\{y_i\}$ 의 차이 샘플 $\{y_i(n)\}$ 로서 차이 값을 생성한다. 실시예에서, 차이 샘플 $\{y_i(n)\}$ 을 생성하는 것은 예를 들어, 도 6을 참조하여 기재된 바와 같이 수행될 수 있다.

[0066] 도 8a의 메타데이터 인코더(801)는 제어 신호 $\{b(n)\}$ 를 수신하도록 구성되는 선택기(830)("A")를 더 포함한다. 더욱이, 선택기(830)는 차이 신호 $\{y_1, \dots, y_n\}$ 의 N 메타데이터를 수신하도록 구성된다. 또한, 도 8a의 실시예에서, 메타데이터 인코더(801)는 N 양자화된 메타데이터 신호 $\{q_1, \dots, q_n\}$ 를 얻기 위해 N 원본 메타데이터 신호 $\{x_1, \dots, x_n\}$ 를 양자화하는 양자화기(820)를 포함한다. 그러한 실시예에서, 양자화기는 N 양자화된 메타데이터 신호를 선택기(830)에 공급하도록 구성될 수 있다. 선택기(830)는 제어 신호 $\{b(n)\}$ 에 따라 DPCM 인코딩된 차이 메타데이터 신호 $\{q_i\}$ 로부터 처리된 메타데이터 신호 $\{z_i\}$ 를 생성하도록 구성될 수 있다. 예를 들어, 제어 신호(b)가 제 1 상태(예를 들면, $b(n) = 0$)에 있을 때, 선택기(830)는 처리된 메타데이터 신호 $\{z_i\}$ 의 메타데이터 샘플 $\{z_i(n)\}$ 로서 메타데이터 차이 신호 $\{y_i\}$ 의 차이 샘플 $\{y_i(n)\}$ 을 출력하도록 구성될 수 있다. 제어 신호(b)가 제 1 상태와 상이한 제 2 상태(예를 들면, $b(n) = 1$)에 있을 때, 선택기(830)는 처리된 메타데이터 신호 $\{z_i\}$ 의 메타데이터 샘플 $\{z_i(n)\}$ 로서 양자화된 메타데이터 신호 $\{q_i\}$ 의 메타데이터 샘플 $\{q_i(n)\}$ 을 출력하도록 구성될 수 있다.

[0067] 도 8b는 다른 실시예에 따른 메타데이터 인코더(802)를 도시한다. 도 8b의 실시예에서, 메타데이터 인코더(802)는 양자화기(820)를 포함하지 않고, N 양자화 메타데이터 신호 $\{q_1, \dots, q_n\}$ 대신에, N 원본 메타데이터 신호 $\{x_1, \dots, x_n\}$ 은 선택기(830)에 직접 공급된다. 이러한 실시예에서, 예를 들어, 제어 신호(b)가 제 1 상태(예를 들면, $b(n) = 0$)일 때, 선택기(830)는 처리된 메타데이터 신호 $\{z_i\}$ 의 메타데이터 샘플 $\{z_i(n)\}$ 로서 메타데이터 차이 신호 $\{y_i\}$ 의 차이 샘플 $\{y_i(n)\}$ 을 출력하도록 구성될 수 있다. 제어 신호(b)가 제 1 상태와 상이한 제 2 상태(예를 들면, $b(n) = 1$)에 있을 때, 선택기(830)는 처리된 메타데이터 신호 $\{z_i\}$ 의 메타데이터 샘플 $\{z_i(n)\}$ 로서 원본 메타데이터 신호 $\{x_i\}$ 의 메타데이터 샘플 $\{x_i(n)\}$ 을 출력하도록 구성될 수 있다. 도 9a는 실시예에 따른 메타데이터 디코더(901)를 도시한다. 도 9a에 따른 메타데이터 인코더는 도 8a 및 도 8b의 메타데이터 인코더에 대응한다. 도 9a의 메타데이터 디코더(901)는 하나 이상의 메타데이터 디코더 서브 유닛(911, ..., 91N)을 포함한다. 메타데이터 디코더(901)는 하나 이상의 처리된 메타데이터 신호 $\{z_1, \dots, z_n\}$ 를 수신하도록 구성된다. 또한, 메타데이터 디코더(901)는 제어 신호(b)를 수신하도록 구성된다. 메타데이터 디코더는 제어 신호(b)에 따라 하나 이상의 처리된 메타데이터 신호 $\{z_1, \dots, z_n\}$ 로부터 하나 이상의 재구성된 메타데이터 신호 $\{x_1', \dots, x_n'\}$ 를 생성하도록 구성된다. 실시예에서, N 처리된 메타데이터 신호 $\{z_1, \dots, z_n\}$ 각각은 메타데이터 디코더 서브 유닛(911, ..., 91N)의 다른 하나에 공급된다. 또한, 실시예에 따라, 제어 신호(b)는 메타데이터 디코더 서브 유닛(911, ..., 91N) 각각에 공급된다. 실시예에 따라, 메타데이터 디코더 서브 유닛(911, ..., 91N)의 개수는 메타데이터 디코더(901)인 수신되는 처리된 메타데이터 신호들 $\{z_1, \dots, z_n\}$ 의 수와 동일하다. 도 9b는 실시예에 따라 도 9a의 메타데이터 디코더 서브 유닛(911, ..., 91N)의 메타데이터 디코더 서브 유닛(91i)을 도시한다. 메타데이터 디코더 서브 유닛(91i)은 단일 처리된 메타데이터 신호 $\{z_i\}$ 에 대한 디코딩을 수행하도록 구성된다. 메타데이터 디코더 서브 유닛(91i)은 선택기(930)("B") 및 가산기(910)를 포함한다. 메타데이터 디코더 서브 유닛(91i)은 제어 신호 $\{b(n)\}$ 에 따라 수신된 처리된 메타데이터 신호 $\{z_i\}$ 로부터 재구성된 메타데이터 신호 $\{x_i'\}$ 를

생성하도록 구성된다. 이것은 예를 들어, 다음과 같이, 실현될 수 있다 : 재구성된 메타데이터 신호 $\{x_i'(n-1)\}$ 의 마지막으로 재구성된 메타데이터 샘플 $\{x_i'(n-1)\}$ 은 또한 가산기(910)에 공급된다. 더욱이, 처리된 메타데이터 신호 $\{z_i(n)\}$ 의 실제 메타데이터 샘플 $\{z_i(n)\}$ 은 또한 가산기(910)에 공급된다. 가산기는 최종 재구성된 메타데이터 샘플 $\{x_i'(n-1)\}$ 및 실제 메타데이터 샘플 $\{z_i(n)\}$ 을 추가하도록 구성되어, 선택기(930)에 공급되는 합 값 $\{s_i(n)\}$ 을 얻는다. 또한, 실제 메타데이터 샘플 $\{z_i(n)\}$ 도 가산기(930)에 공급된다. 선택기는 제어 신호(b)에 따라 재구성된 메타데이터 신호 $\{x_i'(n)\}$ 의 실제 메타데이터 샘플 $\{x_i'(n)\}$ 로서 가산기(910)로부터의 합 값 $\{s_i(n)\}$ 또는 실제 메타데이터 샘플 $\{z_i(n)\}$ 을 선택하도록 구성된다. 예를 들어, 제어 신호(b)가 제 1 상태(예를 들면, $b(n) = 0$)에 있을 때, 제어 신호(b)는 실제 메타데이터 샘플 $\{z_i(n)\}$ 이 차이 값이므로, 합 값 $\{s_i(n)\}$ 이 재구성된 메타데이터 신호 $\{s_i'\}$ 의 실제 메타데이터 샘플 $\{s_i'(n)\}$ 이라는 것을 나타낸다. 선택기(830)는, 제어 신호(b)가 제 1 상태(예를 들면, $b(n) = 0$)에 있을 때, 재구성된 메타데이터 신호 $\{x_i'\}$ 의 실제 메타데이터 샘플 $\{s_i'(n)\}$ 로서 합 값 $\{s_i'(n)\}$ 을 선택하도록 구성된다. 제어 신호(b)가 제 1 상태와 상이한 제 2 상태(예를 들면, $b(n) = 1$)에 있을 때, 제어 신호(b)는 실제 메타데이터 샘플 $\{z_i(n)\}$ 이 차이 값이 아니므로, 실제 메타데이터 샘플 $\{z_i(n)\}$ 이 재구성된 메타데이터 신호 $\{x_i'\}$ 의 정확히 실제 메타데이터 샘플 $\{x_i'(n)\}$ 이라는 것을 나타낸다. 선택기(830)는, 제어 신호가 제 2 상태($b(n) = 1$)에 있을 때, 재구성된 메타데이터 신호 $\{x_i'\}$ 의 실제 메타데이터 샘플 $\{x_i'(n)\}$ 로서 실제 메타데이터 샘플 $\{z_i(n)\}$ 을 선택하도록 구성된다. 실시예에 따르면, 메타데이터 디코더 서브 유닛(91i')은 유닛(920)을 더 포함한다. 유닛(920)은 샘플링된 주기의 지속기간 동안 재구성된 메타데이터 신호의 실제 메타데이터 샘플 $\{x_i'(n)\}$ 을 유지하도록 구성된다. 실시예에서, 이것은, $x_i'(n)$ 이 생성될 때, 생성된 $x_i'(n)$ 이 너무 일찍 공급되지 않아서, $z_i(n)$ 이 차이 값일 때, $x_i'(n)$ 이 $x_i'(n-1)$ 에 기초하여 실제로 생성되는 것을 보장한다. 도 9b의 실시예에서, 선택기(930)는 제어 신호에 $\{b(n)\}$ 에 따라 수신된 신호 성분 $\{z_i(n)\}$ 및, 지연된 출력 성분(재구성된 메타데이터 신호의 이미 생성된 메타데이터 샘플)과 수신된 신호 성분 $\{z_i(n)\}$ 의 선형 조합으로부터 메타데이터 샘플 $\{x_i'(n)\}$ 을 생성할 수 있다.

[0068] 다음에서, DPCM 인코딩된 신호는 $y_i(n)$ 로서 표시되고, B의 제 2 입력 신호(합 신호)는 $s_i(n)$ 로서 표시된다. 대응하는 입력 성분들에만 의존하는 출력 성분들에 대해, 인코더 및 디코더 출력은 다음과 같이 주어진다:

[0069]
$$z_i(n) = A(x_i(n), v_i(n), b(n))$$

[0070]
$$x_i'(n) = B(z_i(n), s_i(n), b(n))$$

[0071] 위에서 스케치된 일반적인 방법에 대한 일 실시예에 따른 해결책은 DPCM 인코딩된 신호와 양자화된 입력 신호 사이를 스위칭하는 $b(n)$ 을 사용하는 것이다. 간략화 이유로 인해 시간 인덱스(n)를 생략하면, 기능 블록(A 및 B)이 다음과 같이 주어진다:

[0072] 메타데이터 인코더(801, 802)에서, 선택기(830)(A)는 선택한다:

[0073] A: $z_i(z_i, y_i, b) = y_i, b = 0$ 인 경우 (z_i 는 차이 값을 나타냄)

[0074] A: $z_i(z_i, y_i, b) = y_i, b = 1$ 인 경우 (z_i 는 차이 값을 나타내지 않음)

[0075] 메타데이터 디코더 서브 유닛(91i, 91i')에서, 선택기(930)(B)는 선택한다 :

[0076] B: $x_i'(z_i, y_i, b) = s_i, b = 0$ 인 경우 (z_i 는 차이 값을 나타냄)

[0077] B: $x_i'(z_i, y_i, b) = z_i, b = 1$ 인 경우 (z_i 는 차이 값을 나타내지 않음)

[0078] 이것은, $b(n)$ 이 1과 동일할 때마다 양자화된 입력 신호를 송신하고, b 가 $b(n)$ 이 0일 때마다 송신하도록 허용된다. 후자의 경우에, 디코더는 DPCM 디코더가 된다.

[0079] 객체 메타데이터의 송신에 적용되는 경우, 이 메커니즘은 랜덤 액세스를 위해 디코더에 의해 이용될 수 있는 압축되지 않은 객체 위치를 규칙적으로 송신하기 위해 사용된다.

[0080] 바람직한 실시예에서, 보다 적은 비트는 메타데이터 샘플을 인코딩하는데 사용된 비트의 수보다 차분 값을 인코딩하는데 사용된다. 이들 실시예는 대부분 (예를 들어, N) 후속 메타데이터 샘플이 약간 다를 수 있다는 발견에 기초한다. 예를 들어, 메타데이터 샘플의 한 종류가 예를 들어, 8 비트만큼 인코딩되는 경우, 이 메타데이터 샘플은 256 상이한 값으로부터 하나를 취할 수 있다. (예를 들어, N) 후속 메타데이터 값의 일반적으로 약간의 변화로 인해서, 차이 값, 예를 들어 5 비트에 의해, 인코딩하는 데 충분한 것으로 간주될 수 있다. 따라서, 차이 값이 송신되는 경우에도, 송신된 비트의 수가 감소될 수 있다.

[0081] 실시예에서, 메타데이터 인코더(210)는 제 1 제어 신호가 제 1 상태($b(n) = 0$)를 나타내는 비트의 제 1 수 및 제 2 제어 신호가 제 2 상태($b(n) = 1$)를 나타내는 비트의 제 2 수로 하나 이상의 처리된 메타데이터 신호(z_1, \dots, z_N)의 하나의 $z_i()$ 의 각 처리된 메타데이터 샘플들($z_1(n), \dots, z_N(n)$) 각각을 인코딩하도록 구성된다.

[0082] 바람직한 실시예에서, 하나 이상의 차이 값이 송신되고, 하나 이상의 차이 값의 각각은 메타데이터 샘플들 각각보다 더 적은 비트들로 인코딩되고, 차이 값 각각은 정수 값이다. 실시예에 따라, 메타데이터 인코더(110)는 제 1 비트 수로 하나 이상의 처리된 메타데이터 신호들 중 하나의 메타데이터 샘플 중 하나 이상을 인코딩하도록 구성되고, 하나 이상의 처리된 메타데이터 신호 중 상기 하나의 상기 하나 이상의 메타데이터 샘플들 각각은 정수를 나타낸다. 더욱이, 메타데이터 인코더(110)는 비트의 제 2 수로 하나 이상의 차이 값을 인코딩하도록 구성되고, 상기 하나 이상의 차이 값 각각은 정수를 나타내고, 비트의 제 2 수는 비트의 제 1 수보다 작다.

[0083] 예를 들어, 실시예에서, 메타데이터 샘플들이 8 비트로 인코딩되는 방위각을 나타낼 수 있는 것을 고려하자. 예를 들어, 방위각이 $-90 \leq \text{방위각} \leq 90$ 일 수 있다. 따라서, 방위각은 181개의 상이한 값을 취할 수 있다. 그러나, (예를 들어, N) 후속 방위각 샘플만이 예를 들어, 단지 ± 15 만큼 차이가 있다고 가정할 수 있으면, 5 비트 ($2^5 = 32$)는 상이한 값들로 인코딩할 정도로 충분할 수 있다. 차이 값이 정수로서 표현되면, 차이 값은 송신될 추가 값들을 적합한 값 범위로 자동으로 변환한다.

[0084] 예를 들어, 제 1 오디오 객체의 제 1 방위각 값이 60° 이고 후속 값이 45° 로부터 75° 까지 변화하는 경우를 고려해보자. 더욱이, 제 2 오디오 객체의 제 2 방위각 값이 -30° 이고, 그 후속 값이 -45° 로부터 -15° 까지 변화하는 것을 고려해보자. 제 1 오디오 객체의 후속 값과 제 2 오디오 객체의 후속 값 모두에 대한 차이 값을 결정함으로써, 제 1 방위각 값 및 제 2 방위각 값의 차이 값은 -15° 으로부터 $+15^\circ$ 까지의 값 범위에 있어서, 5 비트는 차이 값 각각을 인코딩하기에 충분하고, 차이 값을 인코딩한 비트 시퀀스는, 제 1 방위각의 차이 값과 제 2 방위각 값의 차이 값에 대해 동일한 의미를 갖는다.

[0085] 다음에서, 실시예에 따른 객체 메타데이터와 실시예에 따른 심볼 표현이 기재된다.

[0086] 인코딩된 객체 메타데이터가 프레임 단위로 송신된다. 이들 객체 메타데이터 프레임들은, 동적 객체 데이터가 마지막으로 송신된 프레임 아래의 변화를 포함하는 경우 인트라코딩된 객체 데이터 또는 동적 객체 데이터를 포함할 수 있다. 객체 메타데이터 프레임에 대한 다음 구문의 일부 또는 모든 부분은, 예를 들어, 이용될 수 있다:

	No. of bits	Mnemonic
object_metadata() { has_intracoded_object_metadata; if (has_intracoded_object_metadata) { intracoded_object_metadata (); } else { dynamic_object_metadata(); } }	1	bslbf

[0087] 다음에서, 실시예에 따른 인트라코딩된 객체 데이터가 설명된다.

[0088]

[0089] 인코딩된 객체 메타데이터의 랜덤 액세스는 규칙적인 그리드(예를 들어, 길이 1024의 32 프레임마다) 상에서 샘플링된 양자화된 값을 포함하는 인트라코딩된 객체 데이터 ("I-프레임")를 통해 실현된다. 예를 들어, 이들 I-프레임은 다음의 구문을 가질 수 있고, 여기서 position_azimuth, position_elevation, position_radius, 및 gain_factor는 현재 양자화된 값을 지정한다.

	No. of bits	Mnemonic
intracoded_object_metadata()		
{		
if (num_objects>1) {		
fixed_azimuth;	1	bslbf
if (fixed_azimuth) {		
default_azimuth;	8	tcimsbf
}		
else {		
common_azimuth;	1	bslbf
if (common_azimuth) {		
default_azimuth;	8	tcimsbf
}		
else {		
for (o=1:num_objects) {		
position_azimuth[o];	8	tcimsbf
}		
}		
}		
fixed_elevation;	1	bslbf
if (fixed_azimuth) {		
default_elevation;	6	tcimsbf
}		
else {		
common_elevation;	1	bslbf
if (common_azimuth) {		
default_elevation;	6	tcimsbf
}		
else {		
for (o=1:num_objects) {		
position_azimuth[o];	6	tcimsbf
}		
}		
}		
fixed_radius;	1	bslbf
if (fixed_azimuth) {		
default_radius;	4	tcimsbf
}		
else {		
common_radius;	1	bslbf
if (common_azimuth) {		
default_radius;	4	tcimsbf
}		
else {		
for (o=1:num_objects) {		
position_radius [o];	4	tcimsbf
}		
}		
}		
}		

[0090]

}		
fixed_gain;	1	bslbf
if (fixed_azimuth) {		
default_gain;	7	tcimsbf
}		
else {		
common_gain;	1	bslbf
if (common_azimuth) {		
default_gain;	7	tcimsbf
}		
else {		
for (o=1:num_objects) {		
gain_factor [o];	7	tcimsbf
}		
}		
}		
else {		
position_azimuth;	8	tcimsbf
position_elevation;	6	tcimsbf
position_radius;	4	tcimsbf
gain_factor;	7	tcimsbf
}		
}		

[0091]

[0092] 다음에서, 실시예에 따른 동적 객체 데이터가 설명된다. DPCM 데이터는, 예를 들어, 다음 구문을 가질 수 있는 동적 객체 프레임에서 송신된다:

	No. of bits	Mnemonic
dynamic_object_metadata() { flag_absolute;	1	bslbf
for (o=1:num_objects) { has_object_metadata;	1	bslbf
if (has_object_metadata) { single_dynamic_object_metadata(flag_absolute); } } }		

[0093]

	No. of bits	Mnemonic
single_dynamic_object_metadata(flag_absolute){ if (flag_absolute){ if (!fixed_azimuth*){		

[0094]

position_azimuth;	8	tcimsbf
} if (!fixed_elevation*) { position_elevation;	6	tcimsbf
} if (!fixed_radius*) { position_radius;	4	tcimsbf
} if (!fixed_gain*) { gain_factor;	7	tcimsbf
} else { nbits;	3	uimsbf
if (!fixed_azimuth*) { flag_azimuth;	1	bslbf
if (flag_azimuth) { position_azimuth_difference ;	num_bits	tcimsbf
} } if (!fixed_elevation*) { flag_elevation;	1	bslbf
if (flag_elevation) { position_elevation_difference ;	min(num_bits,7)	tcimsbf
} } if (!fixed_radius*) { flag_radius;	1	bslbf
if (flag_radius) { position_radius_difference ;	min(num_bits,5)	tcimsbf
} } if (!fixed_gain*) { flag_gain;	1	bslbf
if (flag_gain) { gain_factor_difference ;	min(num_bits,8)	tcimsbf
} } } }		
Note: num_bits = nbits + 2; Footnote *: Given by the preceding intracoded_object_data()-frame		

[0095]

[0096] 특히, 실시예에서, 상기 매크로는, 예를 들어 다음의 의미를 가질 수 있다:

[0097] 실시예에 따른 object_data() 페이로드들의 정의:

[0098] **has_intracoded_object_metadata** 프레임이 인트라코딩되거나 상이하게 코딩되어 있는지 여부를 나타냄.

[0099] 실시예에 따른 intracoded_object_metadata() 페이로드들의 정의:

[0100] **fixed_azimuth** 방위각 값이 모든 객체를 위해 고정되고 dynamic_object_metadata()의 경우에 송신되지

않는 지의 여부를 나타냄

- [0101] **default_azimuth** 고정된 값 또는 공통 방위각 값을 정의함
- [0102] **common_azimuth** 사용되는 공통 방위각 각이 모든 개체에 사용되는 지의 여부를 나타냄
- [0103] **position_azimuth** 공통 값위각 값이 없는 경우, 각 객체에 대한 값이 송신됨
- [0104] **fixed_elevation** 양각 값이 모든 객체에 대해 고정되고 `dynamic_object_metadata()`의 경우에 송신되지 않는 지의 여부를 나타냄
- [0105] **default_elevation** 고정되거나 공통 양각의 값을 정의
- [0106] **common_elevation** 공통 양각 값이 각 모든 객체에 대해 사용되는지 여부를 나타냄
- [0107] **position_elevation** 공통 양각 값이 없는 경우, 각 객체에 대한 값이 송신됨
- [0108] **fixed_radius** 반경이 모든 객체에 대해 고정되고 `dynamic_object_metadata()`의 경우에 송신되지 않는 지의 여부를 나타냄
- [0109] **default_radius** 공통 반경 값을 정의
- [0110] **common_radius** 공통 반경 값이 모든 객체에 대해 사용되는 지의 여부를 나타냄
- [0111] **position_radius** 공통 반경 값이 없는 경우, 각 객체에 대한 값이 송신됨
- [0112] **fixed_gain** 이득 인자가 모든 객체에 대해 고정되고 `dynamic_object_metadata()`의 경우에 송신되지 않는 지의 여부를 나타냄
- [0113] **default_gain** 고정되거나 공통 이득 인자 값을 정의
- [0114] **common_gain** 일반적인 이득 값이 모든 객체에 대해 사용되는 지의 여부를 나타냄
- [0115] **gain_factor** 공통 이득 값이 없는 경우, 각 객체에 대한 값이 송신됨
- [0116] **position_azimuth** 단지 하나의 객체가 존재하는 경우, 이것은 방위각임
- [0117] **position_elevation** 단지 하나의 객체가 존재하는 경우, 이것은 양각임
- [0118] **position_radius** 단지 하나의 객체가 존재하는 경우, 이것은 반경임
- [0119] **gain_factor** 단지 하나의 객체가 존재하는 경우 이것은 이득 인자임
- [0120] 실시예에 따른 `dynamic_object_metadata()` 페이로드들의 정의:
- [0121] **flag_absolute** 성분의 값이 상이하게 또는 절대 값으로 송신되는 지의 여부를 나타냄
- [0122] **has_object_metadata** 비트 스트림에 존재하는 객체 데이터가 있는 지의 여부를 나타냄
- [0123] 실시예에 따른 `single_dynamic_object_metadata()` 페이로드들의 정의:
- [0124] **position_azimuth** 값이 고정되지 않은 경우 방위각의 절대 값
- [0125] **position_elevation** 값이 고정되지 않은 경우 양각의 절대 값
- [0126] **position_radius** 값이 고정되지 않은 경우 반경의 절대값
- [0127] **gain_factor** 값이 고정되지 않은 경우 이득 인자의 절대 값
- [0128] **nbits** 얼마나 많은 비트가 차분 값을 표현하는데 요구되는 지
- [0129] **flag_azimuth** 방위각 값이 변하는 지의 여부를 나타내는 객체 당 플래그
- [0130] **position_azimuth_difference** 이전과 활성 값 사이의 차이
- [0131] **flag_elevation** 양각 값이 변하는 지의 여부를 나타내는 객체 당 플래그
- [0132] **position_elevation_difference** 이전과 활성 값 사이의 차이 값
- [0133] **flag_radius** 반경이 변하는 지의 여부를 나타내는 객체 당 플래그

- [0134] **position_radius_difference** 이진과 활성 값 사이의 차이
- [0135] **flag_gain** 이득 반경이 변하는 지의 여부를 나타내는 객체 당 플래그
- [0136] **gain_factor_difference** 이진과 활성 값 사이의 차이
- [0137] 종래 기술에서, 낮은 비트율에서 허용가능한 오디오 품질이 얻어지도록 한 편으로 채널 코딩과 다른 한 편으로 객체 코딩을 조합하는 융통성있는 기술은 존재하지 않는다.
- [0138] 이 제한은 3D 오디오 코덱 시스템에 의해 극복된다. 이제, 3D 오디오 코덱 시스템이 설명된다.
- [0139] 도 10은 본 발명의 실시예에 따른 3D 오디오 인코더를 도시한다. 3D 오디오 인코더는 오디오 출력 데이터(501)를 얻기 위해 오디오 입력 데이터(101)를 인코딩하기 위해 구성된다. 3D 오디오 인코더는 CH에 의해 표시된 복수의 오디오 채널 및 OBJ로 표시된 복수의 오디오 객체를 수신하기 위한 입력 인터페이스를 포함한다. 또한, 도 10에 도시된 바와 같이, 입력 인터페이스(1100)는 하나 이상의 복수의 오디오 객체(OBJ)에 관련된 메타데이터를 추가적으로 수신한다. 또한, 3D 오디오 인코더는 복수의 사전-믹싱된 채널들을 얻기 위해 복수의 객체들과 복수의 채널들을 믹싱하기 위한 믹서(200)를 포함하고, 각 사전-믹싱된 채널은 채널의 오디오 데이터 및 적어도 하나의 객체의 오디오 데이터를 포함한다.
- [0140] 또한, 3D 오디오 인코더는 코어 인코더 입력 데이터를 코어 인코딩하기 위한 코어 인코더(300), 하나 이상의 복수의 오디오 객체에 관련된 메타데이터를 압축하기 위한 메타데이터 압축기(400)를 포함한다.
- [0141] 또한, 3D 오디오 인코더는 여러 동작 모드들 중 하나에서 믹서를 제어하기 위한 모드 제어기(600), 코어 인코더 및/또는 출력 인터페이스(500)를 포함할 수 있고, 제 1 모드에서, 코어 인코더는 믹서에 의한 어떠한 상호 작용 없이, 즉 믹서(200)에 의한 어떠한 믹싱 없이 입력 인터페이스(1100)에 의해 수신된 복수의 오디오 객체들 및 복수의 오디오 채널들을 인코딩하도록 구성된다. 하지만, 믹서(200)가 활성화된 제 2 모드에서, 코어 인코더는 복수의 믹싱된 채널들, 즉 블록(200)에 의해 생성된 출력을 인코딩한다. 이러한 후자의 경우에, 어떠한 객체 데이터도 더 이상 인코딩하지 않는 것이 바람직하다. 그 대신, 오디오 객체들의 위치들을 나타내는 메타데이터는 메타데이터에 의해 표시된 채널들 상에 객체들을 렌더링하기 위해 믹서(200)에 의해 이미 사용된다. 즉, 믹서(200)는 오디오 객체들을 사전 렌더링하기 위해 복수의 오디오 객체들에 관련된 메타데이터를 이용하고, 사전-렌더링된 오디오 객체들은 믹서의 출력에서 믹싱된 채널들을 얻기 위해 채널들과 믹싱된다. 이 실시예에서, 임의의 객체들은 송신될 필요가 없을 수 있고, 이것은 또한 블록(400)에 의한 출력으로서 압축된 메타데이터에 적용된다. 하지만, 인터페이스(1100)에 입력된 모든 객체들이 믹싱되는 것은 아니고, 특정 양의 객체들이 믹싱되면, 그럼에도 불구하고 나머지 비-믹싱된 객체들 및 연관된 메타데이터만이 각각 코어 인코더(300) 또는 메타데이터 압축기(400)에 송신된다.
- [0142] 도 10에서, 메타데이터 압축기(400)는 전술한 실시예 중 하나에 따라 인코딩된 오디오 정보를 생성하기 위한 장치(250)의 메타데이터 인코더(210)이다. 또한, 도 10에서, 믹서(200) 및 코어 인코더(300)는 전술한 실시예 중 하나에 따라 인코딩된 오디오 정보를 생성하기 위한 장치(250)의 오디오 인코더(220)를 함께 형성한다.
- [0143] 도 12는 추가적으로, SAOC 인코더(800)를 포함하는 3D 오디오 인코더의 추가 실시예를 도시한다. SAOC 인코더(800)는 공간 오디오 객체 인코더 입력 데이터로부터 하나 이상의 전송 채널들과 파라메트릭 데이터를 생성하기 위해 구성된다. 도 12에 도시된 바와 같이, 공간 오디오 객체 인코더 입력 데이터는 사전-렌더러/믹서에 의해 처리되지 않은 객체이다. 대안적으로, 사전-렌더러/믹서가, 개별 채널/객체 코딩이 활성화된 모드 1로서 우회된다고 가정하면, 입력 인터페이스(1100)에 입력된 모든 객체는 SAOC 인코더(800)에 의해 인코딩된다.
- [0144] 또한, 도 12에 도시된 바와 같이, 코어 인코더(300)는 USAC 인코더, 즉 MPEG-USAC 표준(USAC = 통합 음성 및 오디오 코딩)에서 정의되고 표준화된 인코더로서 바람직하게 구현된다. 도 12에 도시된 전체 3D 오디오 인코더의 출력은 개별적인 데이터 유형에 대한 컨테이너-형 구조를 갖는 MPEG 4 데이터 스트림이다. 또한, 메타데이터는 "OAM" 데이터로서 표시되고, 도 10에서의 메타데이터 압축기(400)는 QAM 인코더(400)에 대응하여, USAC 인코더(300)에 입력되는 압축된 OAM 데이터를 얻고, USAC 인코더(300)는 도 12에서 알 수 있듯이, 인코딩된 채널/객체 데이터를 가질 뿐 아니라 압축된 OAM 데이터를 갖는 MP4 출력 데이터 스트림을 얻기 위해 출력 인터페이스를 추가적으로 포함한다.
- [0145] 도 12에서, OAM 인코더(400)는 전술한 실시예 중 하나에 따라 인코딩된 오디오 정보를 생성하기 위한 장치(250)의 메타데이터 인코더(210)이다. 또한, 도 12에서, SAOC 인코더(800) 및 USAC 인코더(300)는 전술한 실시예 중 하나에 따라 인코딩된 오디오 정보를 생성하기 위한 장치(250)의 오디오 인코더(220)를 함께 형성한다.

- [0146] 도 14는 3D 오디오 인코더의 추가 실시예를 도시하고, 여기서 도 12와 대조적으로, SAOC 인코더는 SAOC 인코딩 알고리즘을 가지고, 이 모드에서 활성화되지 않는 사전-렌더러/믹서(200)에 제공된 채널들을 인코딩하고, 또는 대안적으로 사전-렌더링된 채널들에 객체들을 더한 것을 SAOC 인코딩하도록 구성될 수 있다. 따라서, 도 14에서, SAOC 인코더(800)는 상이한 3개의 유형의 입력 데이터, 즉 어떠한 사전-렌더링된 객체들을 갖지 않은 채널들, 사전-렌더링된 객체들을 갖는 채널들 또는 객체를 단독으로 동작할 수 있다. 또한, 처리를 위해 SAOC 인코더(800)가 디코더 측에서 동일한 데이터, 즉, 원본 OAM 데이터가 아니라 손실 압축에 의해 얻어진 데이터를 이용하도록 도 14에서 추가 OAM 디코더(420)를 제공하는 것이 바람직하다.
- [0147] 도 14에서, 3D 오디오 인코더는 여러 개별 모드에서 동작할 수 있다.
- [0148] 도 10의 상황에서 논의된 바와 같이 제 1 및 제 2 모드뿐 외에도, 3D 오디오 인코더는, 사전-렌더러/믹서(200)가 활성화되지 않을 때 코어 인코더가 개별 객체들로부터 하나 이상의 전송 채널들을 생성하는 제 3 모드로 추가적으로 동작할 수 있다. 대안적으로 또는 추가적으로, 이러한 제 3 모드에서, SAOC 인코더(800)는 즉, 다시 도 10의 믹서(200)에 대응하는 사전-렌더러/믹서(200)가 활성화되지 않을 때 원본 채널로부터 하나 이상의 대안적인 또는 추가적인 전송 채널을 생성할 수 있다.
- [0149] 마지막으로, SAOC 인코더(800)는, 3D 오디오 인코더가 제 4 모드로 구성될 때, 사전-렌더러/믹서에 의해 생성된 채널에 사전-렌더링된 객체들을 더한 것을 인코딩할 수 있다. 따라서, 제 4 모드에서, 가장 낮은 비트율 응용들은, 채널들 및 객체들이 "SAOC-SI"로서 도 3 및 도 5에 표시된 바와 같이 개별적인 SAOC 전송 채널들 및 연관된 부가 정보로 완전히 변환되었고, 추가적으로 이러한 제 4 모드에서 어떠한 압축된 메타데이터도 송신될 필요가 없다는 점으로 인해 양호한 품질을 제공할 것이다.
- [0150] 도 14에서, OAM 인코더(400)는 전술한 실시예 중 하나에 따라 인코딩된 오디오 정보를 생성하기 위한 장치(250)의 메타데이터 인코더(210)이다. 또한, 도 14에서, SAOC 인코더(800) 및 USAC 인코더(300)는 전술한 실시예 중 하나에 따라 인코딩된 오디오 정보를 생성하기 위한 장치(250)의 오디오 인코더(220)를 함께 형성한다.
- [0151] 실시예에 따라, 오디오 출력 데이터(501)를 얻기 위해 오디오 입력 데이터(101)를 인코딩하기 위한 장치가 제공된다. 오디오 입력 데이터(101)를 인코딩하기 위한 장치는:
- [0152] - 복수의 오디오 채널들, 복수의 오디오 객체들, 및 복수의 오디오 객체들 중 하나 이상에 관련된 메타데이터를 수신하기 위한 입력 인터페이스(1100),
- [0153] - 복수의 사전-믹싱된 채널들을 얻기 위해 복수의 객체들 및 복수의 채널들을 믹싱하기 위한 믹서(200)로서, 각 사전-믹싱된 채널은 채널의 오디오 데이터 및 적어도 하나의 객체의 오디오 데이터를 포함하는, 믹서(200), 및
- [0154] - 전술한 바와 같이 메타데이터 인코더 및 오디오 인코더를 포함하는 인코딩된 오디오 정보를 생성하기 위한 장치(250)를 포함한다.
- [0155] 인코딩된 오디오 정보를 생성하기 위한 장치(250)의 오디오 인코더(220)는 코어 인코더 입력 데이터를 코어 인코딩하기 위한 코어 인코더(300)이다.
- [0156] 인코딩된 오디오 정보를 생성하기 위한 장치(250)의 메타데이터 인코더 (210)는 하나 이상의 복수의 오디오 객체에 관련된 메타데이터를 압축하기 위한 메타데이터 압축기(400)이다.
- [0157] 도 11은 본 발명의 실시예에 따른 3D 오디오 디코더를 도시한다. 3D 오디오 디코더는 입력으로서, 인코딩된 오디오 데이터, 즉, 도 10의 데이터(501)를 수신한다.
- [0158] 3D 오디오 디코더는 메타데이터 압축 해제기(1400), 코어 디코더(1300), 객체 프로세서(1200), 모드 제어기(1600) 및 후치 프로세서(1700)를 포함한다.
- [0159] 특히, 3D 오디오 디코더는 인코딩된 오디오 데이터를 디코딩하기 위해 구성되고, 입력 인터페이스는 인코딩된 오디오 데이터를 수신하기 위해 구성되고, 인코딩된 오디오 데이터는 복수의 인코딩된 채널 및 복수의 인코딩된 객체들 및 복수의 객체들에 관련된 압축된 메타데이터를 특정 모드에서 수신하기 위해 구성된다.
- [0160] 또한, 코어 디코더(1300)는 복수의 인코딩된 채널들 및 복수의 인코딩된 객체들을 디코딩하기 위해 구성되고, 추가적으로, 메타데이터 압축 해제기는 압축된 메타데이터를 압축 해제하기 위해 구성된다.
- [0161] 또한, 객체 프로세서(1200)는 객체 데이터 및 디코딩된 채널들을 포함하는 미리 결정된 수의 출력 채널들을 얻기 위해 압축 해제된 메타데이터를 이용하여 코어 디코더(1300)에 의해 생성된 복수의 디코딩된 객체들을 처리

하기 위해 구성된다. 1205로 표시된 이들 출력 채널들은 후치 프로세서(1700)에 입력된다. 후치 프로세서(1700)는 다수의 출력 채널들(1205)을, 5.1, 7.1 등의 출력 포맷과 같은 스피커 출력 포맷 또는 입체 음향 출력 포맷일 수 있는 특정 출력 포맷으로 변환하기 위해 구성된다. 바람직하게는, 3D 오디오 디코더는 모드 표시를 검출하기 위해 인코딩된 데이터를 분석하기 위해 구성되는 모드 제어기(1600)를 포함한다. 그러므로, 모드 제어기(1600)는 도 11에서의 입력 인터페이스(1100)에 연결된다. 그러나, 대안적으로, 모드 제어기가 반드시 있을 필요는 없다. 그 대신에, 융통성있는 오디오 디코더는 사용자 입력 또는 임의의 다른 제어와 같은 임의의 다른 유형의 제어 데이터에 의해 사전 설정될 수 있다. 도 11에서의 3D 오디오 디코더 및 바람직하게 모드 제어기(1600)에 의해 제어된 3D 오디오 디코더는 객체 프로세서를 우회하고 복수의 디코딩된 채널들을 후치 프로세서(1700)에 공급하도록 구성된다. 이것은, 즉 모드 2가 도 10의 3D 오디오 인코더에 적용될 때, 사전-렌더링된 채널들이 수신되는 모드 2에서의 동작이다. 대안적으로, 모드 1이 3D 오디오 인코더에 적용될 때, 즉 3D 오디오 인코더가 개별적인 채널/객체 코딩을 수행할 때, 객체 프로세서(1200)는 우회하지 않고, 복수의 디코딩된 채널들 및 복수의 디코딩된 객체들은 메타데이터 압축 해제기(1400)에 의해 생성된 압축 해제된 메타데이터와 함께 객체 프로세서(1200)에 공급된다.

[0162] 바람직하게, 모드 1 또는 모드 2가 적용되는 지의 여부의 표시는 인코딩된 오디오 데이터에 포함되고, 모드 제어기(1600)는 모드 표시를 검출하기 위해 인코딩된 데이터를 분석한다. 모드 1은, 모드 표시가 인코딩된 오디오 데이터가 인코딩된 채널들 및 인코딩된 객체들을 포함한다는 것을 표시할 때 사용되고, 모드 2는, 모드 표시가 인코딩된 오디오 데이터가 어떠한 오디오 객체들도 포함하지 않는다는 것, 즉 도 10의 3D 오디오 인코더의 모드 2에 의해 얻어진 사전-렌더링된 채널들만을 포함한다는 것을 나타낼 때 적용된다.

[0163] 도 11에서, 메타데이터 압축 해제기(1400)는 전술한 실시예 중 어느 하나에 따라 하나 이상의 오디오 채널을 생성하기 위한 장치(100)의 메타데이터 디코더(110)이다. 또한, 도 11에서, 코어 디코더(1300), 객체 프로세서(1200) 및 후치 프로세서(1700)는 상술한 실시예 중 하나에 따라 하나 이상의 오디오 채널을 생성하기 위한 장치(100)의 오디오 디코더(120)를 함께 형성한다.

[0164] 도 13은 도 11의 3D 디코더와 비교된 바람직한 실시예를 도시하고, 도 13의 실시예는 도 12의 3D 오디오 인코더에 대응한다. 도 11의 3D 오디오 디코더 구현에 더하여, 도 13에서의 3D 오디오 디코더는 SAOC 디코더(1800)를 포함한다. 더욱이, 도 11의 객체 프로세서(1200)는 별개의 객체 렌더러(1210) 및 믹서(1220)로서 구현되는 한편, 모드에 따라, 객체 렌더러(1210)의 기능은 SAOC 디코더(1800)에 의해 또한 구현될 수 있다.

[0165] 또한, 후치 프로세서(1700)는 입체 음향 렌더러(1710) 또는 포맷 변환기(1720)로서 구현된다. 대안적으로, 도 11의 데이터(1205)의 직접 출력은 또한 1730에 의해 도시된 바와 같이 구현될 수 있다. 따라서, 유연성을 갖고 더 작은 포맷이 요구되는 경우 후치-처리하기 위해 22.2 또는 32와 같은 가장 높은 수의 채널들 상에서 디코더에서의 처리를 수행하는 것이 바람직하다. 하지만, 5.1 포맷과 같은 작은 포맷만이 요구된다는 것이 바로 도입 부로부터 명백하게 될 때, SAOC 디코더 및/또는 USAC 디코더를 통한 특정 제어가 불필요한 업믹싱 동작들 및 후속 다운믹싱 동작들을 피하기 위해 적용될 수 있다는 것이 슛컷(1727)에 의해 도 11 또는 도 6에 의해 표시된 바와 같이 바람직하다.

[0166] 본 발명의 바람직한 실시예에서, 객체 프로세서(1200)는 SAOC 디코더(1800)를 포함하고, SAOC 디코더는 코어 디코더에 의해 출력된 하나 이상의 전송 채널들 및 연관된 파라메트릭 데이터를 디코딩하기 위해, 그리고 복수의 렌더링된 오디오 객체들을 얻기 위해 압축 해제된 메타데이터를 이용하기 위해 구성된다. 이 때문에, OAM 출력은 박스(1800)에 연결된다.

[0167] 또한, 객체 프로세서(1200)는 코어 디코더에 의해 출력된 디코딩된 객체들을 렌더링하도록 구성되고, 이러한 디코딩된 객체들은 SAOC 전송 채널에서 인코딩되지 않고 객체 렌더러(1210)로 표시된 바와 같이 일반적으로 단일 채널링된 요소들에서 개별적으로 인코딩된다. 더욱이, 디코더는 믹서의 출력을 스피커들에 출력하기 위해 출력(1730)에 대응하는 출력 인터페이스를 포함한다.

[0168] 추가 실시예에서, 객체 프로세서(1200)는 인코딩된 오디오 신호 또는 인코딩된 오디오 채널을 표현하는 하나 이상의 전송 채널들 및 연관된 파라메트릭 부가 정보를 디코딩하기 위한 공간 오디오 객체 코딩 디코더(1800)를 포함하고, 공간 오디오 객체 코딩 디코더는 연관된 파라메트릭 정보 및 압축 해제된 메타데이터를, 예를 들어, SAOC의 더 이른 버전에 정의된 것과 같이 출력 포맷을 직접 렌더링하기 위해 사용가능한 트랜스코딩된 파라메트릭 부가 정보로 트랜스코딩하도록 구성된다. 후치 프로세서(1700)는 디코딩된 전송 채널들 및 트랜스코딩된 파라메트릭 부가 정보를 이용하여 출력 포맷의 오디오 채널들을 계산하기 위해 구성된다. 후치 프로세서에 의해 수행되는 처리는 MPEG 서라운드 처리와 유사할 수 있거나 또는 BCC 처리 등과 같은 임의의 다른 처리일 수

있다.

- [0169] 추가 실시예에서, 객체 프로세서(1200)는 디코딩된(코어 디코더에 의해) 전송 채널들 및 파라메트릭 부가 정보를 이용하여 출력 포맷에 대한 채널 신호들을 직접 업믹싱 및 렌더링하도록 구성된 공간 오디오 객체 코딩 디코더(1800)를 포함한다.
- [0170] 더욱이, 그리고 중요하게, 도 11의 객체 프로세서(1200)는. 채널들로 믹싱된 사전-렌더링된 객체들이 존재할 때, 즉 도 10의 믹서(200)가 활성화될 때 USAC 디코더(1300)에 의해 출력된 데이터를 입력으로서 직접 수신하는 믹서(1220)를 추가적으로 포함한다. 추가적으로, 믹서(1220)는 SAOC 디코딩 없이 객체 렌더링을 수행하는 객체 렌더러로부터 데이터를 수신한다. 더욱이, 믹서는 SAOC 디코더 출력 데이터, 즉 SAOC 렌더링된 객체들을 수신한다.
- [0171] 믹서(1220)는 출력 인터페이스(1730), 입체 음향 렌더러(1710), 및 포맷 변환기(1720)에 연결된다. 입체 음향 렌더러(1710)는 헤드 관련 전달 함수들 또는 입체 음향 룬 임펄스 응답(BRIR)을 사용하여 출력 채널들을 두 개의 입체 음향 채널로 렌더링하기 위해 구성된다. 포맷 변환기(1720)는 출력 채널들을 믹서의 출력 채널(1205)보다 낮은 수의 채널들을 갖는 출력 포맷으로 변환하기 위해 구성되고, 포맷 변환기(1720)는 5.1 스피커들 등과 같은 재생 레이아웃에 관한 정보를 요구한다.
- [0172] 도 13에서, OAM 디코더(1400)는 전술한 실시예의 하나에 따라 하나 이상의 오디오 채널을 생성하기 위한 장치(100)의 메타데이터 디코더(110)이다. 또한, 도 13에서, 객체 렌더러(1210), USAC 디코더(1300) 및 믹서(1220)는 전술한 실시예의 하나에 따라 하나 이상의 오디오 채널을 생성하기 위한 장치(100)의 오디오 디코더(120)를 함께 형성한다.
- [0173] 도 15의 3D 오디오 디코더는, SAOC 디코더가 렌더링된 객체들을 생성할 수 없지만 렌더링된 채널들을 생성할 수 있다는 점에서 도 13의 3D 오디오 디코더와 상이하고, 이것은 도 14의 3D 오디오 인코더가 사용되었고 채널들/사전-렌더링된 객체들과 SAOC 인코더(800) 입력 인터페이스 사이의 연결(900)이 활성화될 때 그러하다.
- [0174] 또한, 벡터 기반 진폭 페닝(VBAP) 스테이지(1810)가 구성되고, SAOC 디코더로부터 재생 레이아웃에 관한 정보를 수하고, 렌더링 매트릭스를 SAOC 디코더에 출력하여, SAOC 디코더는 결국, 1205의 높은 채널 포맷, 즉 32 스피커에서 믹서의 어떠한 추가 동작 없이 렌더링된 채널들을 제공할 수 있다.
- [0175] VBAP 블록은 바람직하게 렌더링 매트릭스를 도출하도록 디코딩된 OAM 데이터를 수신한다. 더 일반적으로, 바람직하게는 재생 레이아웃뿐만 아니라, 입력 신호가 재생 레이아웃 상에서 렌더링되어야 하는 위치의 기하학적 정보를 요구한다. 이 기하학적 입력 데이터는 SAOC를 이용하여 송신된 채널들에 대한 객체들 또는 채널 위치 정보에 대한 OAM 데이터일 수 있다.
- [0176] 하지만, 단지 특정 출력 인터페이스가 요구되면, VBAP 상태(1810)는 예컨대, 5.1 출력에 대한 요구된 렌더링 매트릭스를 미리 제공할 수 있다. SAOC 디코더(1800) SAOC 전송 채널, 연관된 파라 메트릭 데이터와 압축 해제된 메타데이터로부터 직접 렌더링을 수행하고, 믹서(1220)의 어떠한 상호 작용 없이 요구된 출력 포맷으로의 직접 렌더링을 수행한다. 하지만, 모드들 사이의 특정 믹스가 적용될 때, 즉 여러 채널들이 SAOC 인코딩되지만 모든 채널들이 SAOC 인코딩되지 않은 경우, 또는 여러 객체들이 SAOC 인코딩되지만 모든 객체들이 SAOC 인코딩되지 않은 경우, 또는 채널들을 갖는 사전-렌더링된 객체들의 특정 양이 SAOC 디코딩되고 나머지 채널들이 SAOC 처리되지 않을 때, 믹서는 개별적인 입력 부분들로부터, 즉 코어 디코더(1300)로부터 직접, 객체 렌더러(1210) 및 SAOC 디코더(1800)로부터 직접 데이터를 수집할 것이다.
- [0177] 도 15에서, OAM 디코더(1400)는 전술한 실시예의 하나에 따라 하나 이상의 오디오 채널을 생성하기 위한 장치(100)의 메타데이터 디코더(110)이다. 또한, 도 15에서, 객체 렌더러(1210), USAC 디코더(1300), 및 믹서(1220)는 전술한 실시예의 하나에 따라 하나 이상의 오디오 채널을 생성하기 위한 장치(100)의 오디오 디코더(120)를 함께 형성한다.
- [0178] 인코딩된 오디오 데이터를 디코딩하기 위한 장치가 제공된다. 인코딩된 오디오 데이터를 디코딩하기 위한 장치는
- [0179] - 인코딩된 오디오 데이터를 수신하기 위한 입력 인터페이스(1100)로서, 인코딩된 오디오 데이터는 복수의 인코딩된 채널들 또는 복수의 인코딩된 객체들 또는 복수의 객체들에 관련된 압축 메타데이터를 포함하는, 입력 인터페이스(1100), 및
- [0180] - 전술한 바와 같이 하나 이상의 오디오 채널들을 생성하기 위한 오디오 채널 생성기(120) 및 메타데이터 디코

더(110)를 포함하는 장치(100)를

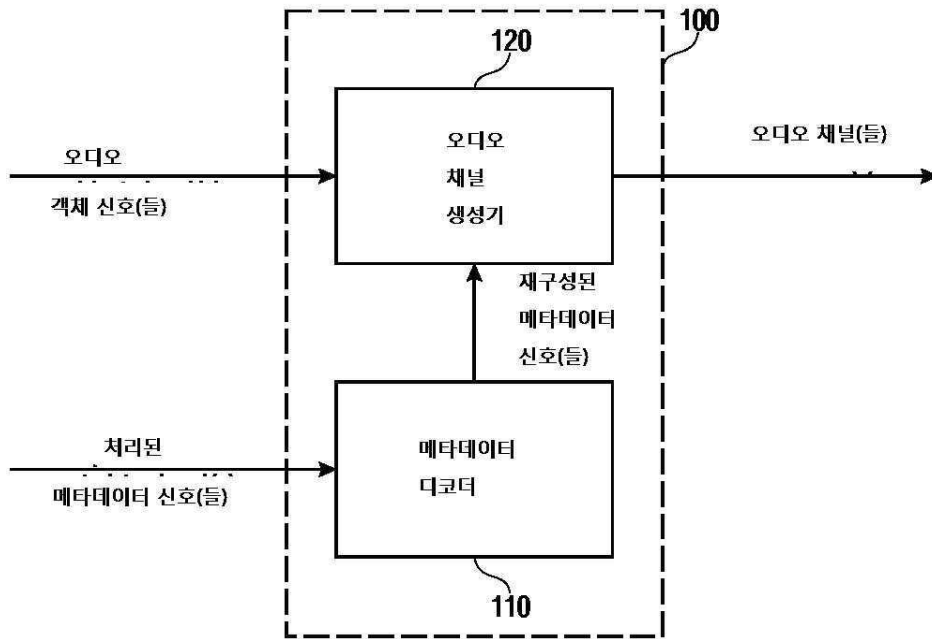
- [0181] 포함한다.
- [0183] 하나 생성 이상의 오디오 채널을 생성하기 위한 장치(100)의 메타데이터 디코더(110)는 압축된 메타데이터를 압축 해제하기 위한 메타데이터 압축 해제기(400)이다.
- [0184] 하나 이상의 오디오 채널을 생성하기 위한 장치(100)의 오디오 채널 생성기(120)는 복수의 인코딩된 채널들 및 복수의 인코딩된 객체들을 디코딩하기 위한 코어 디코더(1300)를 포함한다.
- [0185] 또한, 오디오 채널 생성기(120)는 객체들로부터의 오디오 데이터 및 디코딩된 채널들을 포함하는 다수의 출력 채널들(1205)을 얻기 위해 압축 해제된 메타데이터를 이용하여 복수의 디코딩된 객체들을 처리하기 위한 객체 프로세서(1200)를 더 포함한다.
- [0186] 또한, 오디오 채널 생성기(120)는 다수의 출력 채널들(1205)을 출력 포맷으로 변환하기 위한 후치 프로세서(1700)를 포함한다.
- [0187] 몇몇 양상들이 장치의 정황에서 기재되었지만, 이들 양상들이 또한 대응하는 방법의 설명을 나타내고, 여기서 블록 또는 디바이스가 방법 단계 또는 방법 단계의 특징에 대응한다는 것이 명확하다. 유사하게, 방법 단계의 정황에서 기재된 양상들은 또한 대응하는 블록 또는 항목 또는 대응하는 장치의 특징의 설명을 나타낸다.
- [0188] 본 발명의 압축 해제된 신호는 디지털 저장 매체 상에 저장될 수 있거나 인터넷과 같은 무선 송신 매체 또는 유선 송신과 같은 저장 매체 상에서 송신될 수 있다.
- [0189] 특정 구현 조건들에 따라, 본 발명의 실시예들은 하드웨어 또는 소프트웨어로 구현될 수 있다. 구현은 디지털 저장 매체, 예를 들어, 플로피 디스크, DVD, CD, ROM, PROM, EPROM, EEPROM, 또는 FLASH 메모리를 이용하여 수행될 수 있는데, 이러한 디지털 저장 매체는 그 위에 저장된 전자적으로 판독가능한 제어 신호들을 갖고, 각 방법이 수행되도록 프로그래밍가능 컴퓨터 시스템과 협력한다(또는 협력할 수 있다). 그러므로, 디지털 저장 매체는 컴퓨터 판독가능할 수 있다.
- [0190] 본 발명에 따른 몇몇 실시예들은, 본 명세서에 기재된 방법들 중 하나가 수행되도록, 프로그래밍가능 컴퓨터 시스템과 협력할 수 있는, 전자적으로 판독가능한 제어 신호들을 갖는 비-임시 데이터 캐리어를 포함한다.
- [0191] 일반적으로, 본 발명의 실시예들은 프로그램 코드를 갖는 컴퓨터 프로그램 제품으로서 구현될 수 있고, 프로그램 코드는, 컴퓨터 프로그램이 컴퓨터 상에서 실행될 때 방법들 중 하나를 수행하기 위해 동작가능하다. 프로그램 코드는 예를 들어, 기계 판독가능한 캐리어 상에 저장될 수 있다.
- [0192] 다른 실시예들은 기계 판독가능한 캐리어 상에 저장된, 본 명세서에 기재된 방법들 중 하나를 수행하기 위한 컴퓨터 프로그램을 포함한다.
- [0193] 즉, 그러므로, 본 발명의 방법의 실시예는, 컴퓨터 프로그램이 컴퓨터 상에서 실행될 때, 본 명세서에 기재된 방법들 중 하나를 수행하기 위한 프로그램 코드를 갖는 컴퓨터 프로그램이다.
- [0194] 그러므로, 본 발명의 방법들의 추가 실시예는 본 명세서에 기재된 방법들 중 하나를 수행하기 위한 컴퓨터 프로그램을 그 위에 리코딩되게 포함하는 데이터 캐리어(또는 디지털 저장 매체, 또는 컴퓨터-판독가능 매체)이다. 데이터 캐리어, 디지털 저장 매체 또는 리코딩된 매체는 일반적으로 실체적(tangible)이고 및/또는 비-과도적이다.
- [0195] 그러므로, 본 발명의 방법의 추가 실시예는 본 명세서에 기재된 방법들 중 하나를 수행하기 위한 컴퓨터 프로그램을 나타내는 신호들의 시퀀스 또는 데이터 스트림이다. 예를 들어, 신호들의 시퀀스들 또는 데이터 스트림은 데이터 통신 연결부를 통해, 예를 들어, 인터넷을 통해, 전송되도록 구성될 수 있다.
- [0196] 추가 실시예는 본 명세서에 기재된 방법들 중 하나를 수행하도록 프로그래밍되고, 구성되거나 적용된 처리 수단, 예를 들어, 컴퓨터, 또는 프로그래밍가능 논리 디바이스를 포함한다.
- [0197] 추가 실시예는 본 명세서에 기재된 방법들 중 하나를 수행하기 위한 컴퓨터 프로그램이 그 위에 설치된 컴퓨터를 포함한다.
- [0198] 몇몇 실시예들에서, 프로그래밍가능 논리 디바이스(예를 들어, 전계 프로그래밍가능 게이트 어레이)는 본 명세

서에 기재된 방법들의 기능들 중 몇몇 또는 전부를 수행하는데 사용될 수 있다. 몇몇 실시예들에서, 전체 프로그래밍가능 게이트 어레이는 본 명세서에 기재된 방법들 중 하나를 수행하기 위해 마이크로프로세서와 협력할 수 있다. 일반적으로, 방법들은 임의의 하드웨어 장치에 의해 바람직하게 수행된다.

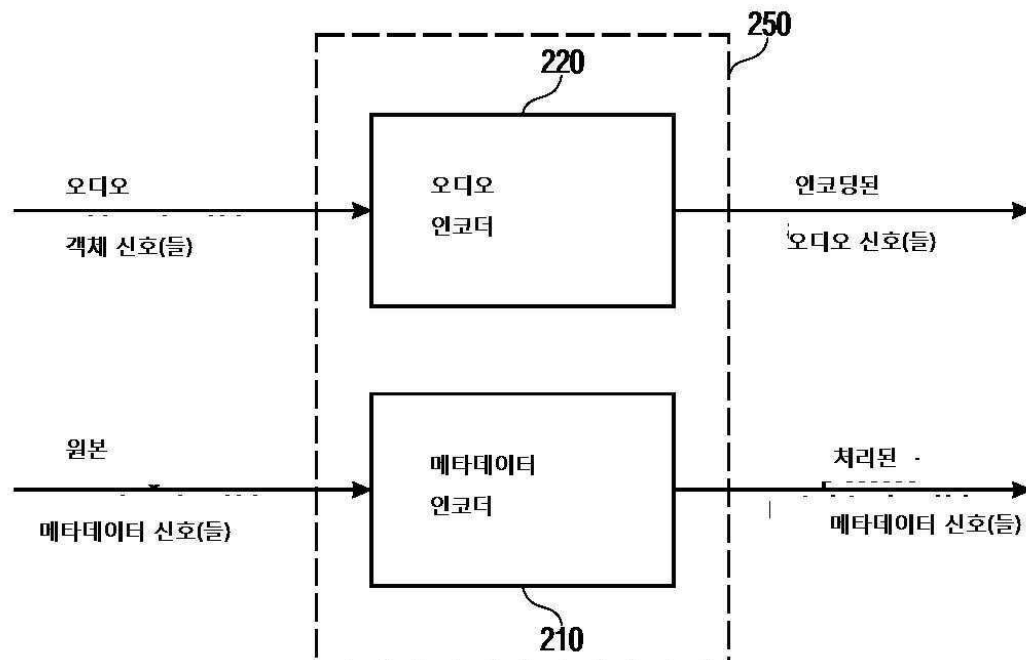
- [0199] 전술한 실시예들은 본 발명의 원리들을 위해 단지 예시적이다. 본 명세서에 기재된 세부사항들 및 배치들의 변형들 및 변경들이 당업자에게 명백하다는 것이 이해된다. 그러므로, 본 명세서에서 실시예들의 기재 및 설명에 의해 제공된 특정 세부사항들에 의해서가 아니라 다음의 특허 청구항들의 범주에 의해서만 제한되도록 의도된다.
- [0200] 인용 문헌들
- [0201] [1] Peters, N., Lossius, T. and Schacher J. C., "SpatDIF: Principles, Specification, and Examples", 9th Sound and Music Computing Conference, Copenhagen, Denmark, Jul. 2012.
- [0202] [2] Wright, M., Freed, A., "Open Sound Control: A New Protocol for Communicating with Sound Synthesizers", International Computer Music Conference, Thessaloniki, Greece, 1997.
- [0203] [3] Matthias Geier, Jens Ahrens, and Sascha Spors. (2010), "Object-based audio reproduction and the audio scene description format", *Org. Sound*, Vol. 15, No. 3, pp. 219-227, December 2010.
- [0204] [4] W3C, "Synchronized Multimedia Integration Language (SMIL 3.0)", Dec. 2008.
- [0205] [5] W3C, "Extensible Markup Language (xML) 1.0 (Fifth Edition)", Nov. 2008.
- [0206] [6] MPEG, "ISO/IEC International Standard 14496-3 - Coding of audio-visual objects, Part 3 Audio", 2009.
- [0207] [7] Schmidt, J.; Schroeder, E. F. (2004), "New and Advanced Features for Audio Presentation in the MPEG-4 Standard", 116th AES Convention, Berlin, Germany, May 2004
- [0208] [8] Web3D, "International Standard ISO/IEC 14772-1:1997 - The Virtual Reality Modeling Language (VRML), Part 1: Functional specification and EUC-KR encoding", 1997.
- [0209] [9] Sporer, T. (2012), "Codierung räumlicher Audiosignale mit leichtgewichtigen Audio-Objekten", Proc. Annual Meeting of the German Audiological Society (DGA), Erlangen, Germany, Mar. 2012.
- [0210] [10] Cutler, C. C. (1950), "Differential Quantization of Communication Signals", US Patent US2605361, Jul. 1952.
- [0211] [11] Ville Pulkki, "Virtual Sound Source Positioning Using Vector Base Amplitude Panning"; *J. Audio Eng. Soc.*, Volume 45, Issue 6, pp. 456-466, June 1997.

도면

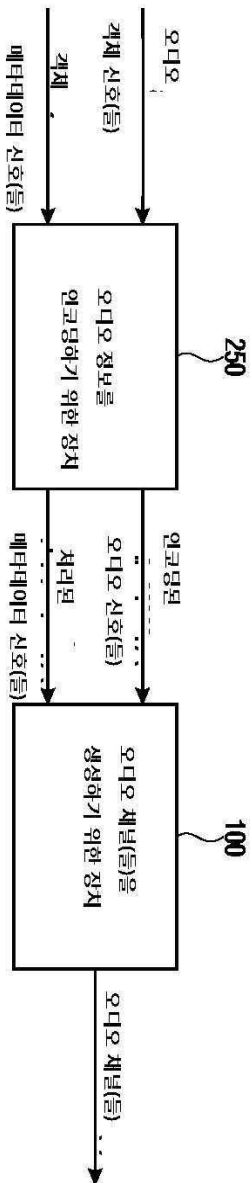
도면1



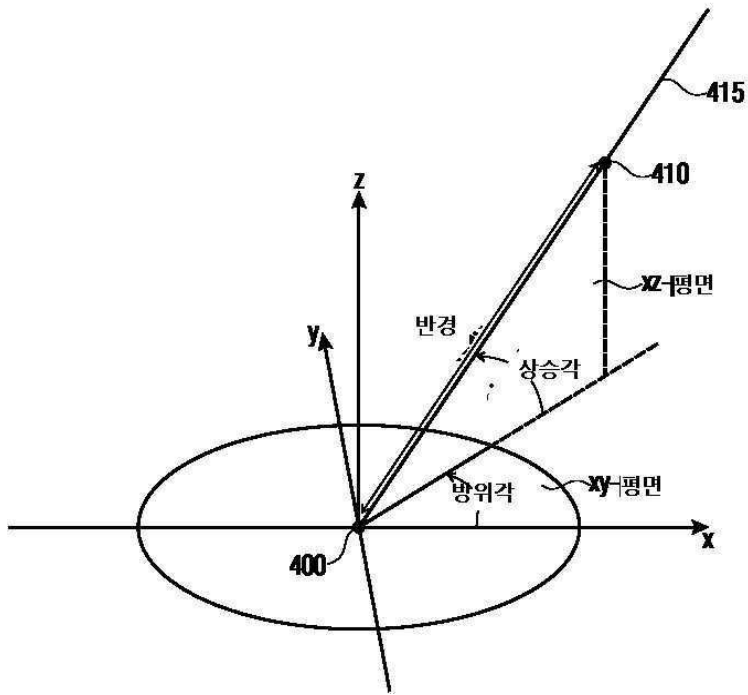
도면2



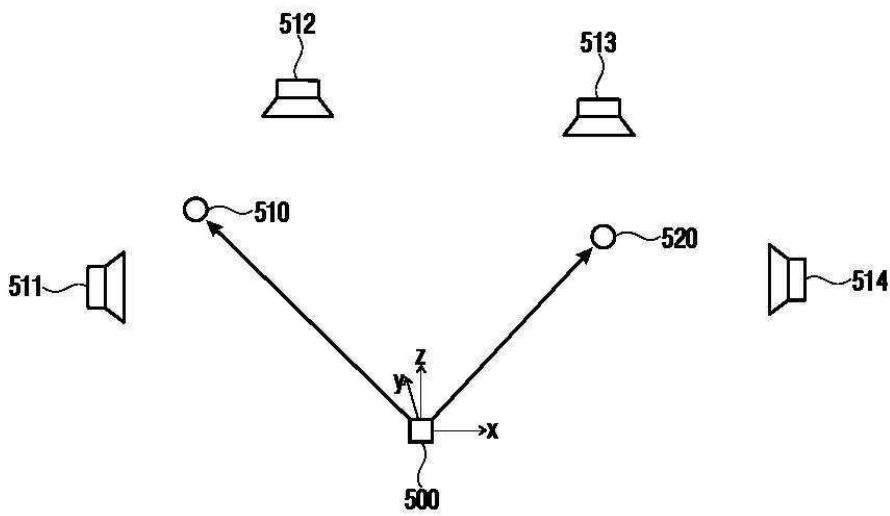
도면3



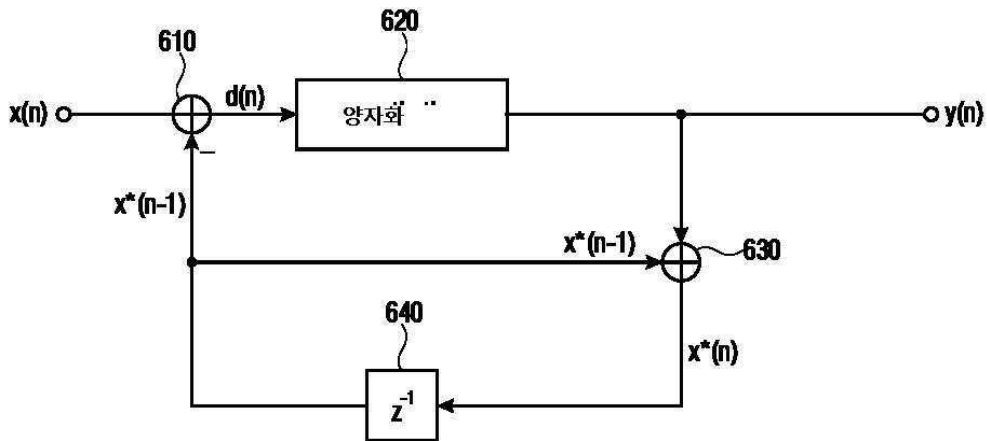
도면4



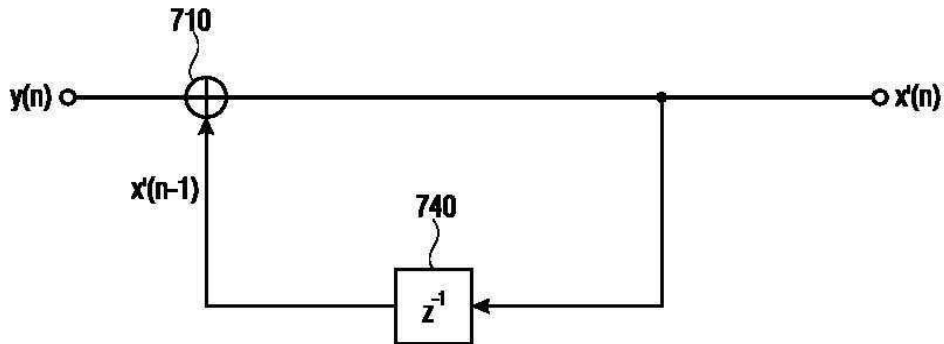
도면5



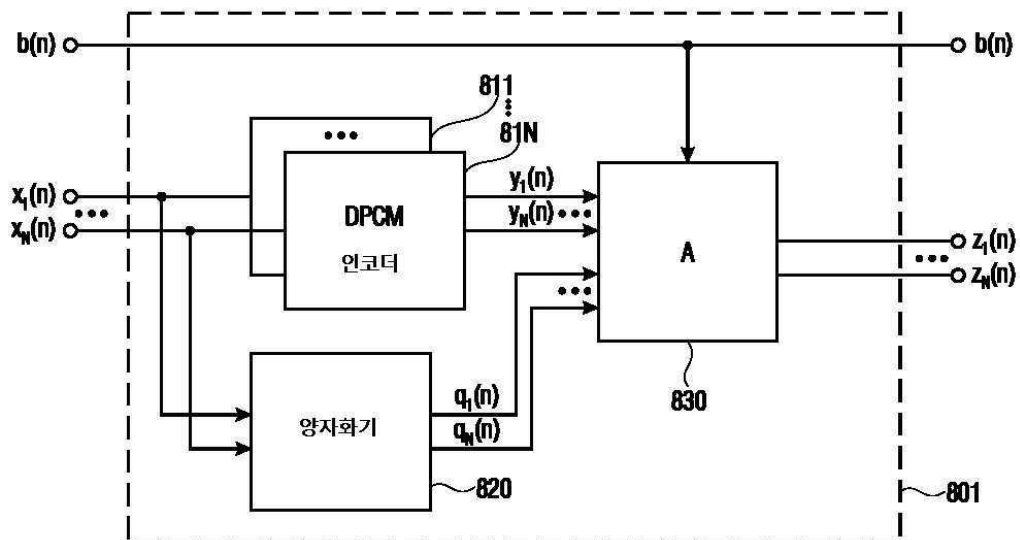
도면6



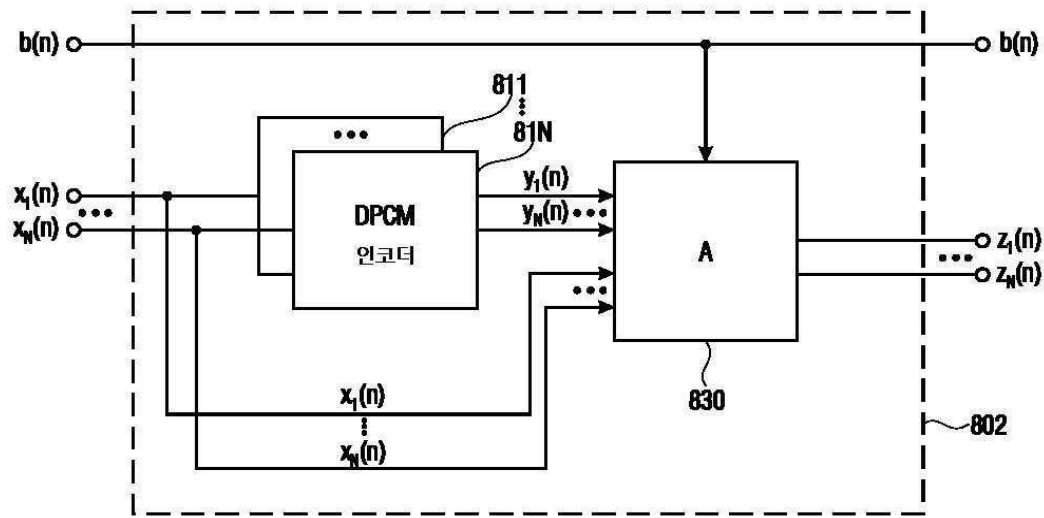
도면7



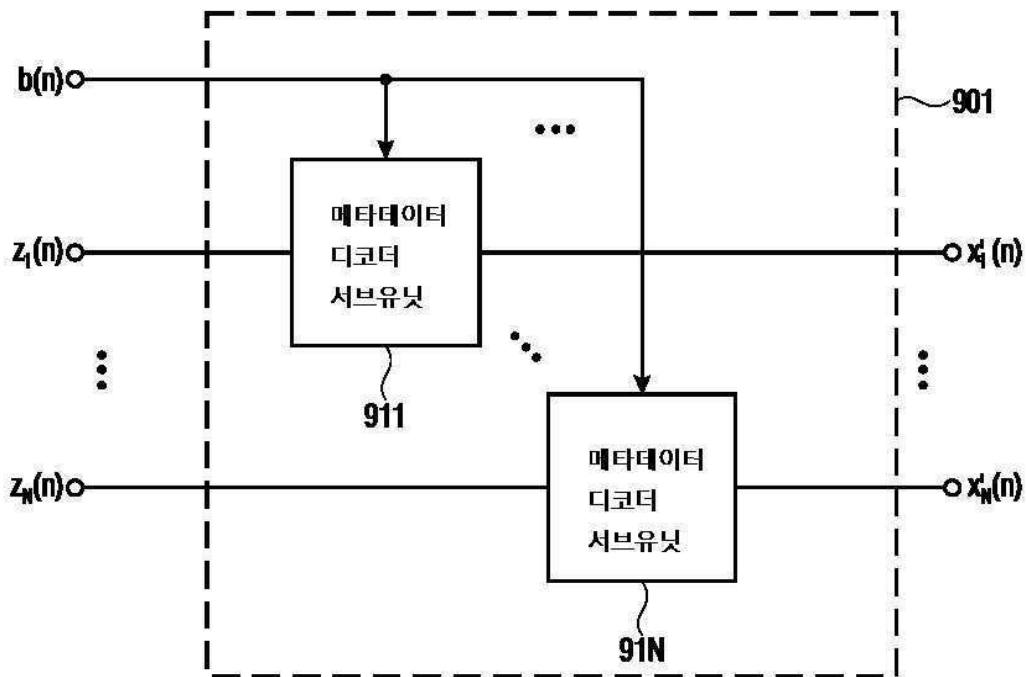
도면8a



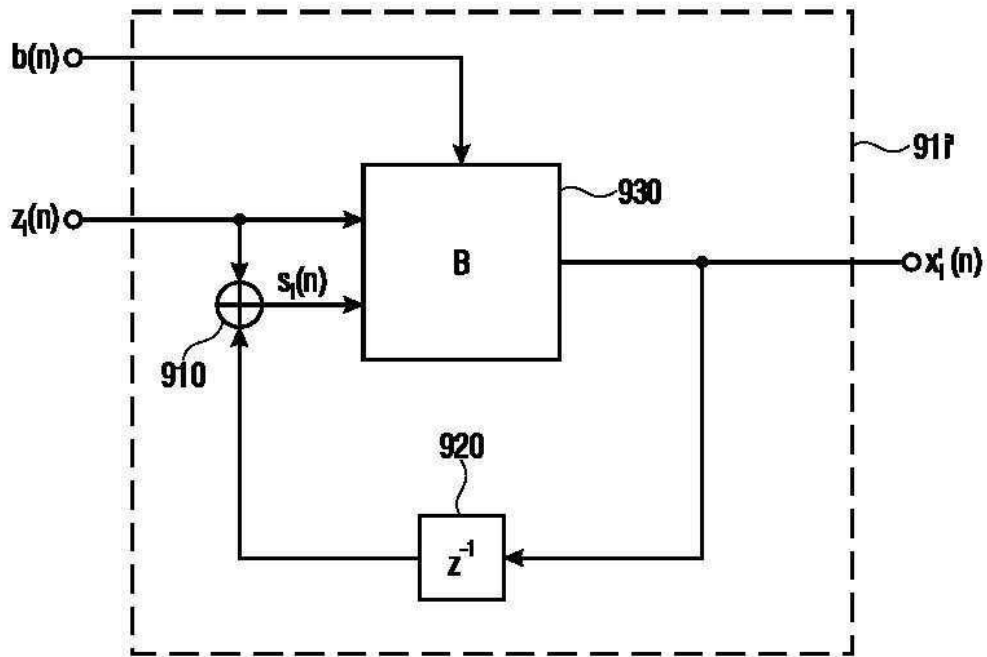
도면8b



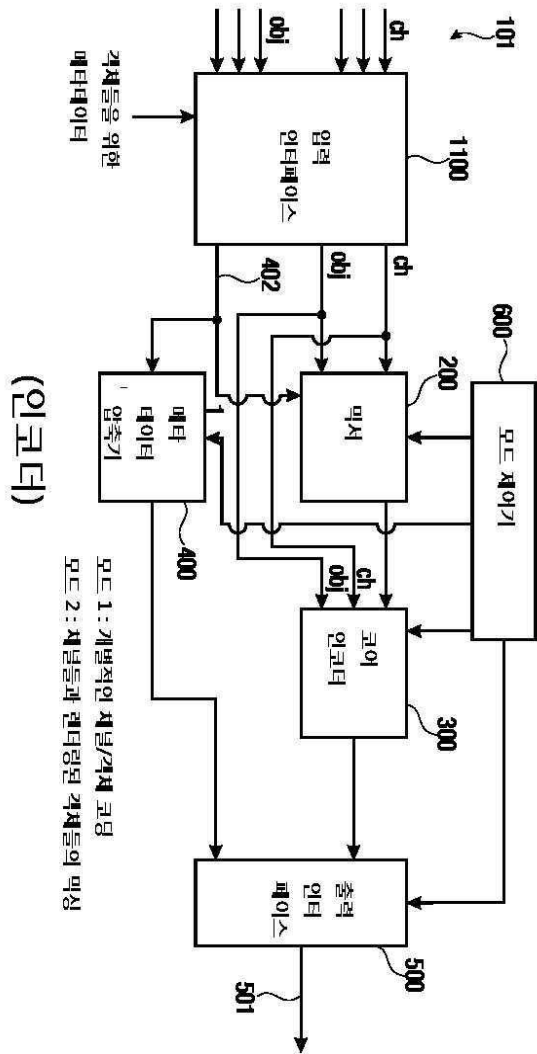
도면9a



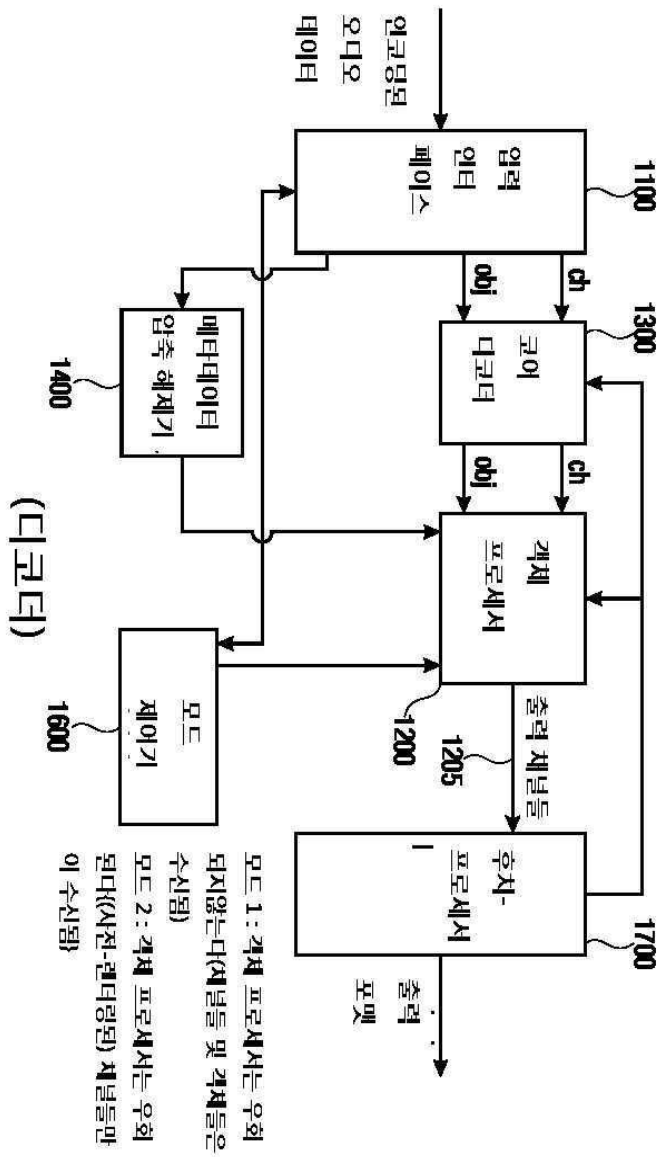
도면9b



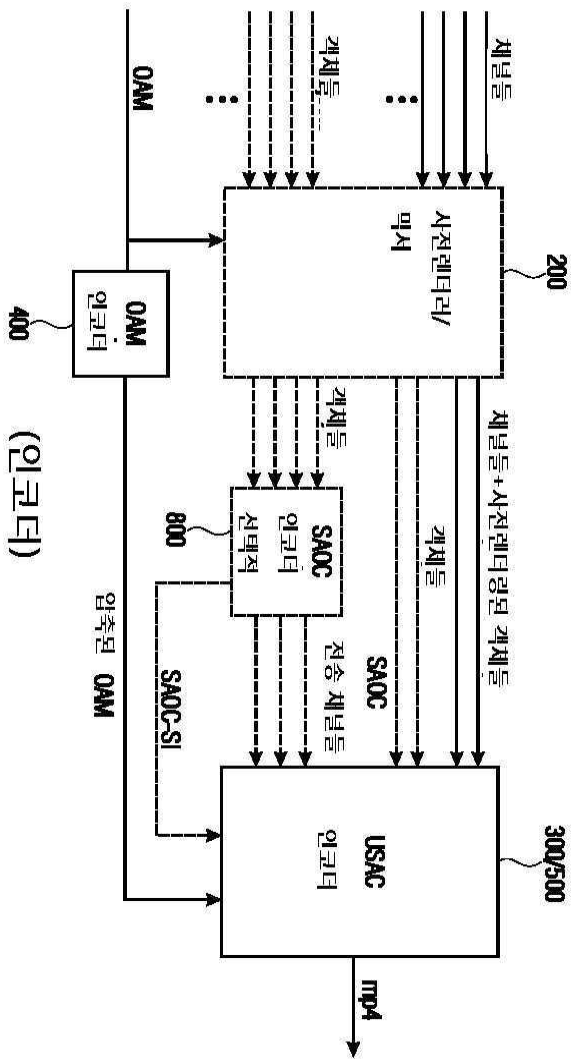
도면10



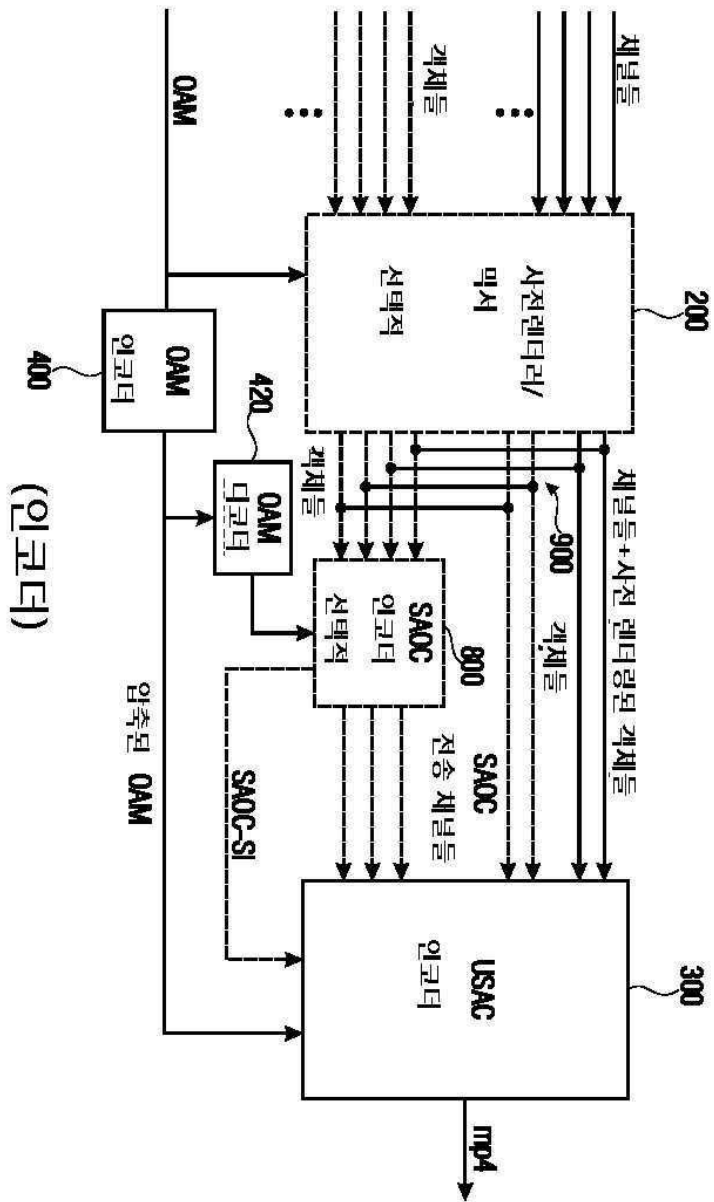
도면11



도면12



도면14



도면15

