



등록특허 10-2715677



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

- (51) 국제특허분류(Int. Cl.)
G10L 19/008 (2014.01) *G10L 19/16* (2013.01)
- (52) CPC특허분류
G10L 19/008 (2020.08)
G10L 19/167 (2013.01)
- (21) 출원번호 10-2018-7012716
- (22) 출원일자(국제) 2016년10월07일
심사청구일자 2021년10월06일
- (85) 번역문제출일자 2018년05월03일
- (65) 공개번호 10-2018-0066136
- (43) 공개일자 2018년06월18일
- (86) 국제출원번호 PCT/EP2016/073969
- (87) 국제공개번호 WO 2017/060410
국제공개일자 2017년04월13일
- (30) 우선권주장
15306589.1 2015년10월08일
유럽특허청(EPO)(EP)
(뒷면에 계속)
- (56) 선행기술조사문현
JP2015525897 A
(뒷면에 계속)
- 전체 청구항 수 : 총 12 항

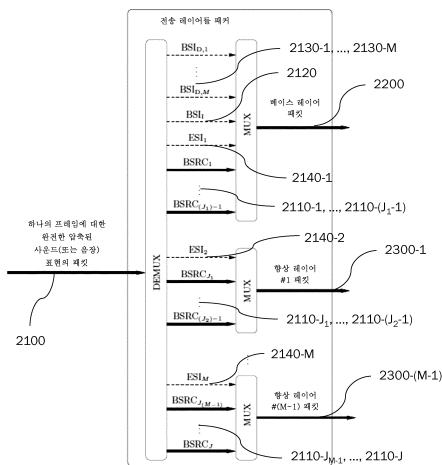
- (45) 공고일자 2024년10월11일
(11) 등록번호 10-2715677
(24) 등록일자 2024년10월04일
- (73) 특허권자
돌비 인터네셔널 에이비
아일랜드 디02 브이케이60 더블린 그랜드 커널 독
랜즈 블록 씨 서 존 로저슨스 키 77
- (72) 발명자
코르돈, 스벤
독일 31515 분스토프 디이 뮐렌캄프스트라세 50
에이
크뤼거, 알렉산더
독일 30655 하노버 포이어바흐스트라세 16
- (74) 대리인
양영준, 이민호, 백만기

심사관 : 이남숙

(54) 발명의 명칭 **압축된 사운드 또는 음장 표현들에 대한 계층화된 코딩****(57) 요 약**

본 문서는 사운드 또는 음장의 압축된 사운드 표현의 계층화된 인코딩 방법에 관한 것이다. 압축된 사운드 표현은 복수의 컴포넌트들을 포함하는 기본 압축된 사운드 표현, 기본 압축된 사운드 표현을 사운드 또는 음장의 기본 재구성된 사운드 표현으로 디코딩하기 위한 기본 보조 정보, 및 기본 재구성된 사운드 표현을 개선시키기 위

(뒷면에 계속)

대 표 도 - 도2

한 파라미터들을 포함하는 향상 보조 정보를 포함한다. 본 방법은 복수의 컴포넌트들을 복수의 컴포넌트 그룹들로 세분하고 복수의 그룹들 각각을 복수의 계층적 레이어들의 각자의 계층적 레이어에 배정하는 단계 - 그룹들의 수는 레이어들의 수에 대응하고, 복수의 레이어들은 베이스 레이어 및 하나 이상의 계층적 향상 레이어를 포함함 -, 기본 보조 정보를 베이스 레이어에 추가하는 단계, 및 향상 보조 정보로부터 복수의 향상 보조 정보 부분들을 결정하고 복수의 향상 보조 정보 부분들 각각을 복수의 레이어들의 각자의 레이어에 배정하는 단계 - 각각의 향상 보조 정보 부분은 각자의 레이어 및 각자의 레이어보다 하위인 임의의 레이어들에 포함된 데이터로부터 획득 가능한 재구성된 사운드 표현을 개선시키기 위한 파라미터들을 포함함 - 를 포함한다. 본 문서는 추가로 사운드 또는 음장의 압축된 사운드 표현을 디코딩하는 방법 - 압축된 사운드 표현은 베이스 레이어 및 하나 이상의 계층적 향상 레이어를 포함하는 복수의 계층적 레이어들에 인코딩됨 - 은 물론, 압축된 사운드 표현의 계층화된 코딩을 위한 인코더 및 디코더에 관한 것이다.

(56) 선행기술조사문헌

KR1020150038156 A

WO2015140293 A1*

Johannes Boehm, et al., Detailed Technical Description of 3D Audio Phase 2 Reference Model 0 for HOA technologies, ISO/IEC JTC1/SC29/WG11 MPEG2014/M35057, 2014.10.19.*

WD1-HOA Text of MPEG-H 3D Audio. ISO/IEC JTC1/SC29/WG11 N14264. 2014.02.21.

*는 심사관에 의하여 인용된 문헌

(30) 우선권주장

15306653.5 2015년10월15일

유럽특허청(EPO)(EP)

62/361,461 2016년07월12일 미국(US)

62/361,416 2016년07월12일 미국(US)

명세서

청구범위

청구항 1

계층화된 인코딩을 사용하여 복수의 계층적 레이어들에 인코딩되는 사운드 또는 음장의 압축된 고차 앰비소닉스 (Higher Order Ambisonics)(HOA) 사운드 표현을 디코딩하는 방법으로서,

베이스 레이어(base layer) 및 적어도 두 개의 계층적 향상 레이어(hierarchical enhancement layers)를 포함하는 상기 복수의 계층적 레이어들에 대응하는 상기 압축된 HOA 표현을 포함하는 비트스트림을 수신하는 단계 - 상기 복수의 계층적 레이어들 중 적어도 하나는 상기 사운드 또는 음장의 기본 압축된 사운드 표현(basic compressed sound representation)의 컴포넌트들을 배정받았고, 상기 컴포넌트들은 복수의 모노럴 신호들에 대응함 -, 및

상기 베이스 레이어와 연관되어 있는 기본 보조 정보(basic side information)에 기초하여 그리고 상기 적어도 두 개의 계층적 향상 레이어와 연관되어 있는 향상 보조 정보(enhancement side information)에 기초하여 상기 압축된 HOA 표현을 디코딩하는 단계

를 포함하고,

상기 기본 보조 정보는 상기 복수의 모노럴 신호들 중 다른 모노럴 신호들과 독립적으로 디코딩될 상기 복수의 모노럴 신호들 중 제1 개별 모노럴 신호들에 관련된 기본 독립적 보조 정보(basic independent side information)를 포함하는, 방법.

청구항 2

삭제

청구항 3

제1항에 있어서, 상기 기본 보조 정보는 상기 복수의 모노럴 신호들 중 다른 모노럴 신호들에 의존하여 디코딩될 상기 복수의 모노럴 신호들 중 제2 개별 모노럴 신호들에 관련된 기본 종속적 보조 정보(basic dependent side information)를 추가로 포함하는, 방법.

청구항 4

제3항에 있어서, 상기 기본 종속적 보조 정보는 상기 음장 내에서 방향성으로 분포된(directionally distributed) 벡터 기반 신호들을 포함하고, 상기 방향성 분포는 벡터에 의해 명시되는, 방법.

청구항 5

제4항에 있어서, 상기 벡터의 컴포넌트들은 0으로 설정되고 상기 압축된 벡터 표현의 일부가 아닌, 방법.

청구항 6

삭제

청구항 7

삭제

청구항 8

제1항에 있어서, 상기 향상 보조 정보는 공간 예측, 서브대역 방향성 신호 합성, 및 파라메트릭 앰비언스 복제 (parametric ambience replication) 중 적어도 하나에 관련된 파라미터들을 포함하는, 방법.

청구항 9

제1항에 있어서, 상기 향상 보조 정보는 방향성 신호들로부터 상기 사운드 또는 음장의 누락 부분들의 예측을 가능하게 하는 정보를 포함하는, 방법.

청구항 10

삭제

청구항 11

계층화된 인코딩을 사용하여 복수의 계층적 레이어들에 인코딩되는 사운드 또는 음장의 압축된 고차 앰비소닉스 (HOA) 사운드 표현을 디코딩하기 위한 장치로서,

베이스 레이어 및 적어도 두 개의 계층적 향상 레이어를 포함하는 상기 복수의 계층적 레이어들에 대응하는 상기 압축된 HOA 표현을 포함하는 비트스트림을 수신하기 위한 수신기 - 상기 복수의 계층적 레이어들 중 적어도 하나는 상기 사운드 또는 음장의 기본 압축된 사운드 표현의 컴포넌트들을 배정받았고, 상기 컴포넌트들은 복수의 모노럴 신호들에 대응함 -, 및

상기 베이스 레이어와 연관되어 있는 기본 보조 정보에 기초하여 그리고 상기 적어도 두 개의 계층적 향상 레이어와 연관되어 있는 향상 보조 정보에 기초하여 상기 압축된 HOA 표현을 디코딩하기 위한 디코더

를 포함하고,

상기 기본 보조 정보는 상기 복수의 모노럴 신호들 중 다른 모노럴 신호들과 독립적으로 디코딩될 상기 복수의 모노럴 신호들 중 제1 개별 모노럴 신호들에 관련된 기본 독립적 보조 정보를 포함하는, 장치.

청구항 12

삭제

청구항 13

제11항에 있어서, 상기 기본 보조 정보는 상기 복수의 모노럴 신호들 중 다른 모노럴 신호들에 의존하여 디코딩될 상기 복수의 모노럴 신호들 중 제2 개별 모노럴 신호들에 관련된 기본 종속적 보조 정보를 추가로 포함하는, 장치.

청구항 14

제13항에 있어서, 상기 기본 종속적 보조 정보는 상기 음장 내에서 방향성으로 분포된 벡터 기반 신호들을 포함하고, 상기 방향성 분포는 벡터에 의해 명시되는, 장치.

청구항 15

제14항에 있어서, 상기 벡터의 컴포넌트들은 0으로 설정되고 상기 압축된 벡터 표현의 일부가 아닌, 장치.

청구항 16

삭제

청구항 17

삭제

청구항 18

제11항에 있어서, 상기 향상 보조 정보는 공간 예측, 서브대역 방향성 신호 합성, 및 파라메트릭 앰비언스 복제 중 적어도 하나에 관련된 파라미터들을 포함하는, 장치.

청구항 19

제11항에 있어서, 상기 향상 보조 정보는 방향성 신호들로부터 상기 사운드 또는 음장의 누락 부분들의 예측을 가능하게 하는 정보를 포함하는, 장치.

청구항 20

삭제

발명의 설명

기술 분야

관련 출원의 상호 참조

[0001] 본 출원은 2015년 10월 8일자로 출원된 유럽 특허 출원 제15306589.1호 및 2015년 10월 15일자로 출원된 유럽 특허 출원 제15306653.5호, 그리고 미국 특허 출원 제62/361,461호 및 제62/361,416호 - 이들은 참조에 의해 그 전체가 본원에 원용됨 - 에 대한 우선권을 주장한다.

[0003] 본 문서는 계층화된 오디오 코딩을 위한 방법들 및 장치들에 관한 것이다. 상세하게는, 본 문서는 압축된 사운드(또는 음장) 표현들, 예를 들어, 고차 앰비소닉스(Higher-Order Ambisonics)(HOA) 사운드(또는 음장) 표현들의 계층화된 오디오 코딩을 위한 방법들 및 장치들에 관한 것이다.

배경 기술

[0004] 시변 조건들을 갖는 전송 채널을 통해 사운드(또는 음장) 표현을 스트리밍하기 위해, 계층화된 코딩은 수신된 사운드 표현의 품질을 전송 조건들에 적응시키기 위한, 그리고 상세하게는 원하지 않는 신호 드롭아웃(signal dropout)들을 회피하기 위한 수단이다.

[0005] 계층화된 코딩의 경우, 사운드(또는 음장) 표현은 보통 비교적 작은 크기의 고 우선순위 베이스 레이어(base layer)와 점감하는 우선순위(decremental priority)들 및 임의적 크기들을 갖는 부가의 향상 레이어(enhancement layer)들로 세분된다. 각각의 향상 레이어는 전형적으로 사운드(또는 음장) 표현의 품질을 개선시키기 위해 하위 레이어들 전부의 정보를 보완하는 증분적 정보(incremental information)를 포함하는 것으로 가정된다. 개별 레이어들의 전송에 대한 에러 방지(error protection)의 양은 그들의 우선순위에 기초하여 제어된다. 상세하게는, 베이스 레이어는 높은 에러 방지를 제공받으며, 이는 베이스 레이어의 작은 크기로 인해 타당하고 무난한 것이다.

[0006] 그렇지만, 예를 들어, 압축된 HOA 사운드 또는 음장 표현들과 같은, 특수 유형들의 압축된 사운드 또는 음장 표현들(그의 확장된 버전들)에 대한 계층화된 코딩 스키마(layered coding scheme)들이 필요하다.

[0007] 본 문서는 이상의 문제들을 다룬다. 상세하게는, 압축된 사운드 또는 음장 표현의 계층화된 코딩을 위한 방법들 및 인코더들/디코더들이 기술된다.

발명의 내용

[0008] 일 양태에 따르면, 사운드 또는 음장의 압축된 사운드 표현의 계층화된 인코딩 방법이 기술된다. 압축된 사운드 표현은 복수의 컴포넌트들을 포함하는 기본 압축된 사운드 표현(basic compressed sound representation)을 포함할 수 있다. 복수의 컴포넌트들은 상보적 컴포넌트들일 수 있다. 압축된 사운드 표현은 기본 압축된 사운드 표현을 사운드 또는 음장의 기본 재구성된 사운드 표현(basic reconstructed sound representation)으로 디코딩하기 위한 기본 보조 정보(basic side information)를 추가로 포함할 수 있다. 압축된 사운드 표현은 또한 기본 재구성된 사운드 표현을 개선(예컨대, 향상)시키기 위한 파라미터들을 포함하는 향상 보조 정보를 추가로 포함할 수 있다. 본 방법은 복수의 컴포넌트들을 복수의 컴포넌트 그룹들로 세분(예컨대, 그룹화)하는 단계를 포함할 수 있다. 본 방법은 복수의 그룹들 각각을 복수의 계층적 레이어들의 각자의 레이어에 배정(assign)(예컨대, 추가)하는 단계를 추가로 포함할 수 있다. 배정은 각자의 그룹들 및 레이어들 간의 대응관계를 표시할 수 있다. 각자의 레이어에 배정된 컴포넌트들은 그 레이어에 포함된다고 말해질 수 있다. 그룹들의 수는 레이어들의 수에 대응할(예컨대, 그와 동일할) 수 있다. 복수의 레이어들은 베이스 레이어 및 하나 이상의 계층적 향상 레이어를 포함할 수 있다. 복수의 계층적 레이어들은 베이스 레이어로부터 제1 향상 레이어, 제2 향상 레이어 등을 거쳐 전체 최상위(overall highest) 향상 레이어(전체 최상위 레이어)까지 순서화될 수 있다. 본 방법은 기본 보조 정보를 베이스 레이어에 추가하는 단계(예를 들어, 전송 또는 저장을 위해, 예컨대, 기본 보조 정보를 베이스 레이어에 포함시키는 단계, 또는 기본 보조 정보를 베이스 레이어에 할당(allocate)하는 단계)를 추가로 포함할 수 있다. 본 방법은 향상 보조 정보로부터 복수의 향상 보조 정보 부분들을 결정하는 단계를 추

가로 포함할 수 있다. 본 방법은 또한 복수의 향상 보조 정보 부분들 각각을 복수의 레이어들의 각자의 레이어에 배정(예컨대, 추가)하는 단계를 추가로 포함할 수 있다. 각각의 향상 보조 정보 부분은 각자의 레이어 및 각자의 레이어보다 하위인 임의의 레이어들에 포함된(예컨대, 배정된 또는 추가된) 데이터로부터 획득가능한 재구성된(예컨대, 압축해제된) 사운드 표현을 개선시키기 위한 파라미터들을 포함할 수 있다. 전송 채널을 통해 전송하기 위해 또는, 예를 들어, CD, DVD, 또는 Blu-ray Disc™과 같은, 적당한 저장 매체에 저장하기 위해, 계층화된 인코딩이 수행될 수 있다.

[0009] 이상과 같이 구성될 때, 제안된 방법은, 복수의 컴포넌트들은 물론, 앞서 기재된 특성들을 갖는 기본 및 향상 보조 정보(예컨대, 독립적 기본 보조 정보 및 향상 보조 정보)를 포함하는 압축된 사운드 표현들에 계층화된 코딩을 효율적으로 적용하는 것을 가능하게 한다. 상세하게는, 제안된 방법은 각각의 레이어가 그 해당 레이어(layer in question)까지의 임의의 레이어들에 포함된 컴포넌트들로부터 재구성된 사운드 표현을 재구성하기 위한 적당한 보조 정보를 포함하도록 보장한다. 거기에서, 그 해당 레이어까지의 레이어들이란 그 해당 레이어까지의, 예를 들어, 베이스 레이어, 제1 향상 레이어, 제2 향상 레이어 등을 포함하는 것으로 이해된다. 따라서, 실제의 최상위 사용가능 레이어(예컨대, 최상위 사용가능 레이어 아래의 레이어들 전부 및 최상위 사용가능 레이어 자체가 유효하게 수신되도록, 유효하게 수신되지 않은 최하위 레이어 아래의 레이어)에 상관없이, 디코더는, 재구성된 사운드 표현이 완전한(complete)(예컨대, 전체(full)) 사운드 표현과 상이할 수 있을지라도, 재구성된 사운드 표현을 개선 또는 향상시킬 수 있을 것이다. 상세하게는, 실제의 최상위 사용가능 레이어에 상관없이, 디코더가 실제의 최상위 사용가능 레이어까지의 레이어들에 포함된 컴포넌트들 전부에 기초하여 획득가능한 재구성된 사운드 표현을 개선 또는 향상시키기 위해 단일 레이어만에 대한(즉, 최상위 사용가능 레이어에 대한) 향상 보조 정보의 페이로드를 디코딩하는 것으로 충분하다. 즉, 각각의 시간 구간(예컨대, 프레임)에 대해, 향상 보조 정보의 단일 페이로드만 디코딩되면 된다. 다른 한편으로, 제안된 방법은 계층화된 코딩을 적용할 때 달성될 수 있는 요구 대역폭의 감소를 충분히 이용하는 것을 가능하게 한다.

[0010] 실시예들에서, 기본 압축된 사운드 표현의 컴포넌트들은 모노럴 신호들(예컨대, 전송 신호들 또는 모노럴 전송 신호들)에 대응할 수 있다. 모노럴 신호들은 HOA 표현의 우세 사운드 신호들 또는 계수 시퀀스들 중 어느 하나를 나타낼 수 있다. 모노럴 신호들은 양자화될 수 있다.

[0011] 실시예들에서, 기본 보조 정보는 복수의 컴포넌트들 중 하나 이상을, 다른 컴포넌트들과 독립적으로, 개별적으로 디코딩(예컨대, 압축해제)하는 것을 명시하는 정보를 포함할 수 있다. 예를 들어, 기본 보조 정보는, 다른 모노럴 신호들과 독립적으로, 개별 모노럴 신호들에 관련된 보조 정보를 표현할 수 있다. 따라서, 기본 보조 정보는 독립적 기본 보조 정보(independent basic side information)라고 지칭될 수 있다.

[0012] 실시예들에서, 향상 보조 정보는 향상 보조 정보를 나타낼 수 있다. 향상 보조 정보는 기본 압축된 사운드 표현 및 기본 보조 정보로부터 획득가능한 기본 재구성된 사운드 표현을 개선(예컨대, 향상)시키기 위한 기본 압축된 사운드 표현에 대한 예측 파라미터들을 포함할 수 있다.

[0013] 실시예들에서, 본 방법은 복수의 레이어들의 데이터(예컨대, 각자의 레이어들에 배정되거나 추가된, 또는 각자의 레이어들에 다른 방식으로 포함된 데이터)의 전송을 위한 전송 스트림(transport stream)을 생성하는 단계를 추가로 포함할 수 있다. 베이스 레이어는 최상위 전송 우선순위를 가질 수 있고, 계층적 향상 레이어들은 점감하는 전송 우선순위들을 가질 수 있다. 즉, 베이스 레이어로부터 제1 향상 레이어로, 제1 향상 레이어로부터 제2 향상 레이어로, 이하 마찬가지로 전송 우선순위가 저하될 수 있다. 복수의 레이어들의 데이터의 전송에 대한 여러 방지의 양은 각자의 전송 우선순위들에 따라 제어될 수 있다. 그에 의해, 적어도 다수의 하위 레이어들이 신뢰성있게 전송되면서, 다른 한편으로 상위 레이어들에 과도한 여러 방지를 적용하지 않는 것에 의해 전체 요구 대역폭을 감소시키는 것이 보장될 수 있다.

[0014] 실시예들에서, 본 방법은, 복수의 레이어들 각각에 대해, 각자의 레이어의 데이터를 포함하는 전송 레이어 패킷을 생성하는 단계를 추가로 포함할 수 있다. 예를 들어, 각각의 시간 구간(예컨대, 프레임)에 대해, 복수의 레이어들 각각에 대한 각자의 전송 레이어 패킷이 생성될 수 있다.

[0015] 실시예들에서, 압축된 사운드 표현은 기본 압축된 사운드 표현을 기본 재구성된 사운드 표현으로 디코딩하기 위한 부가 기본 보조 정보를 추가로 포함할 수 있다. 부가 기본 보조 정보는 복수의 컴포넌트들 중 하나 이상을 각자의 다른 컴포넌트들에 의존하여 디코딩하는 것을 명시하는 정보를 포함할 수 있다. 본 방법은 부가 기본 보조 정보를 복수의 부가 기본 보조 정보 부분들로 분해하는 단계를 추가로 포함할 수 있다. 본 방법은 또한 부가 기본 보조 정보 부분들을 베이스 레이어에 추가하는 단계(예를 들어, 전송 또는 저장을 위해, 예컨대, 부가 기본 보조 정보 부분들을 베이스 레이어에 포함시키는 단계, 또는 부가 기본 보조 정보 부분들을 베이스 레이어에 포함하는 단계)를 포함하는 단계를 추가로 포함할 수 있다.

이어에 할당하는 단계)를 추가로 포함할 수 있다. 각각의 부가 기본 보조 정보 부분은 각자의 레이어에 대응할 수 있고, 각자의 레이어에 배정된 하나 이상의 컴포넌트를 각자의 레이어 및 각자의 레이어보다 하위인 임의의 레이어들에 배정된 각자의 다른 컴포넌트들에(만) 의존하여 디코딩하는 것을 명시하는 정보를 포함할 수 있다. 즉, 각각의 부가 기본 보조 정보 부분은 각자의 레이어보다 상위인 레이어들에 배정된 임의의 다른 컴포넌트들을 참조하지 않고 그 부가 기본 보조 정보 부분이 대응하는 각자의 레이어 내의 컴포넌트들을 명시한다.

[0016] 이와 같이 구성될 때, 제안된 방법은 부분들 전부를 베이스 레이어에 추가하는 것에 의해 부가 기본 보조 정보의 단편화(fragmentation)를 회피한다. 환언하면, 부가 기본 보조 정보 부분들 전부가 베이스 레이어에 포함된다. 부가 기본 보조 정보의 분해는 각각의 레이어에 대해 상위 레이어들 내의 컴포넌트들에 대한 지식을 요구하지 않는 부가 기본 보조 정보 부분이 이용가능하도록 보장한다. 따라서, 실제의 최상위 사용가능 레이어에 상관없이, 디코더가 최상위 사용가능 레이어까지의 레이어들에 포함된 부가 기본 보조 정보를 디코딩하는 것으로 충분하다.

[0017] 실시예들에서, 부가 기본 보조 정보는 복수의 컴포넌트들 중 하나 이상을 다른 컴포넌트들에 의존하여 디코딩(예컨대, 압축해제)하는 것을 명시하는 정보를 포함할 수 있다. 예를 들어, 부가 기본 보조 정보는 다른 모노럴 신호들에 의존하여 개별 모노럴 신호들에 관련된 보조 정보를 표현할 수 있다. 따라서, 부가 기본 보조 정보는 종속적 기본 보조 정보(dependent basic side information)라고 지칭될 수 있다.

[0018] 실시예들에서, 연속적 시간 구간들, 예를 들어, 동일한 크기의 시간 구간들에 대해 압축된 사운드 표현이 처리될 수 있다. 연속적 시간 구간들은 프레임들일 수 있다. 따라서, 본 방법은 프레임 기반으로(on a frame basis) 동작할 수 있으며, 즉, 압축된 사운드 표현은 프레임 단위로(in a frame-wise manner) 인코딩될 수 있다. 각각의 연속적 시간 구간에 대해(예컨대, 각각의 프레임에 대해) 압축된 사운드 표현이 이용가능할 수 있다. 즉, 압축된 사운드 표현이 획득되는 압축 동작이 프레임 기반으로 동작할 수 있다.

[0019] 실시예들에서, 본 방법은, 각각의 레이어에 대해, 그 레이어에 배정된 기본 압축된 사운드 표현의 컴포넌트들을 표시하는 구성 정보를 생성하는 단계를 추가로 포함할 수 있다. 따라서, 디코더는 수신된 데이터 페이로드들의 불필요한 파싱 없이 디코딩에 필요한 정보에 쉽게 액세스할 수 있다.

[0020] 다른 양태에 따르면, 사운드 또는 음장의 압축된 사운드 표현의 계층화된 인코딩 방법이 기술된다. 압축된 사운드 표현은 복수의 컴포넌트들을 포함하는 기본 압축된 사운드 표현을 포함할 수 있다. 복수의 컴포넌트들은 상보적 컴포넌트들일 수 있다. 압축된 사운드 표현은 기본 압축된 사운드 표현을 사운드 또는 음장의 기본 재구성된 사운드 표현으로 디코딩하기 위한 기본 보조 정보(예컨대, 독립적 기본 보조 정보) 및 제3 정보(예컨대, 종속적 기본 보조 정보)를 추가로 포함할 수 있다. 기본 보조 정보는 복수의 컴포넌트들 중 하나 이상을, 다른 컴포넌트들과 독립적으로, 개별적으로 디코딩하는 것을 명시하는 정보를 포함할 수 있다. 부가 기본 보조 정보는 복수의 컴포넌트들 중 하나 이상을 각자의 다른 컴포넌트들에 의존하여 디코딩하는 것을 명시하는 정보를 포함할 수 있다. 본 방법은 복수의 컴포넌트들을 복수의 컴포넌트 그룹들로 세분(예컨대, 그룹화)하는 단계를 포함할 수 있다. 본 방법은 복수의 그룹들 각각을 복수의 계층적 레이어들의 각자의 레이어에 배정(예컨대, 추가)하는 단계를 추가로 포함할 수 있다. 배정은 각자의 그룹들 및 레이어들 간의 대응관계를 표시할 수 있다. 각자의 레이어에 배정된 컴포넌트들은 그 레이어에 포함된다고 말해질 수 있다. 그룹들의 수는 레이어들의 수에 대응할(예컨대, 그와 동일할) 수 있다. 복수의 레이어들은 베이스 레이어 및 하나 이상의 계층적 향상 레이어를 포함할 수 있다. 본 방법은 기본 보조 정보를 베이스 레이어에 추가하는 단계(예를 들어, 전송 또는 저장을 위해, 예컨대, 기본 보조 정보를 베이스 레이어에 포함시키는 단계, 또는 기본 보조 정보를 베이스 레이어에 할당하는 단계)를 추가로 포함할 수 있다. 본 방법은 부가 기본 보조 정보를 복수의 부가 기본 보조 정보 부분들로 분해하는 단계 및 부가 기본 보조 정보 부분들을 베이스 레이어에 추가하는 단계(예를 들어, 전송 또는 저장을 위해, 예컨대, 부가 기본 보조 정보 부분들을 베이스 레이어에 포함시키는 단계, 또는 부가 기본 보조 정보 부분들을 베이스 레이어에 할당하는 단계)를 추가로 포함할 수 있다. 각각의 부가 기본 보조 정보 부분은 각자의 레이어에 대응할 수 있고, 각자의 레이어에 배정된 하나 이상의 컴포넌트를 각자의 레이어 및 각자의 레이어보다 하위인 임의의 레이어들에 배정된 각자의 다른 컴포넌트들에 의존하여 디코딩하는 것을 명시하는 정보를 포함할 수 있다.

[0021] 이와 같이 구성될 때, 제안된 방법은, 각각의 레이어에 대해, 임의의 상위 레이어들의 유효한 수신 또는 디코딩(또는 일반적으로 그에 대한 지식)을 요구하지 않고, 각자의 레이어까지의 임의의 레이어에 포함된 컴포넌트들을 디코딩하기 위해 적절한 부가 기본 보조 정보가 이용가능하도록 보장한다. 압축된 HOA 표현의 경우에, 제안된 방법은 벡터 코딩 모드에서 최상위 사용가능 레이어까지의 레이어들에 속하는 컴포넌트들 전부에 대해 적당

한 V-벡터가 이용가능하도록 보장한다. 상세하게는, 제안된 방법은 상위 레이어들에서의 컴포넌트들에 대응하는 V-벡터의 요소들이 명시적으로 시그널링되지 않는 경우를 제외한다. 그에 따라, 최상위 사용가능 레이어까지의 레이어들에 속하는 임의의 컴포넌트들을 디코딩(예컨대, 압축해제)하는 데 최상위 사용가능 레이어까지의 레이어들에 포함된 정보로 충분하다. 그에 의해, 상위 레이어들이 디코더에 의해 유효하게 수신되지 않았을 수 있더라도 하위 레이어들에 대한 각자의 재구성된 HOA 표현들의 적절한 압축해제가 보장된다. 다른 한편으로, 제안된 방법은 계층화된 코딩을 적용할 때 달성될 수 있는 요구 대역폭의 감소를 충분히 이용하는 것을 가능하게 한다.

[0022] 이 양태의 실시예들은 전술한 양태의 실시예들에 관련될 수 있다.

[0023] 다른 양태에 따르면, 사운드 또는 음장의 압축된 사운드 표현을 디코딩하는 방법이 기술된다. 압축된 사운드 표현은 복수의 계층적 레이어들에 인코딩되었을 수 있다. 복수의 계층적 레이어들은 베이스 레이어 및 하나 이상의 계층적 향상 레이어를 포함할 수 있다. 복수의 레이어들은 사운드 또는 음장의 기본 압축된 사운드 표현의 컴포넌트들을 배정받았을 수 있다. 환연하면, 복수의 레이어들은 기본 압축된 보조 정보의 컴포넌트들을 포함할 수 있다. 컴포넌트들은 각자의 컴포넌트 그룹들 내의 각자의 레이어들에 배정될 수 있다. 복수의 컴포넌트들은 상보적 컴포넌트들일 수 있다. 베이스 레이어는 기본 압축된 사운드 표현을 디코딩하기 위한 기본 보조 정보를 포함할 수 있다. 각각의 레이어는 각자의 레이어 및 각자의 레이어보다 하위인 임의의 레이어들에 포함된 데이터로부터 획득가능한 기본 재구성된 사운드 표현을 개선시키기 위한 파라미터들을 포함하는 향상 보조 정보 부분을 포함할 수 있다. 본 방법은 복수의 계층적 레이어들에 각각 대응하는 데이터 페이로드들을 수신하는 단계를 포함할 수 있다. 본 방법은 기본 압축된 사운드 표현을 사운드 또는 음장의 기본 재구성된 사운드 표현으로 디코딩하기 위해 사용될 복수의 레이어들 중 최상위 사용가능 레이어를 표시하는 제1 레이어 인덱스를 결정하는 단계를 추가로 포함할 수 있다. 본 방법은, 기본 보조 정보를 사용하여, 최상위 사용가능 레이어 및 최상위 사용가능 레이어보다 하위인 임의의 레이어들에 배정된 컴포넌트들로부터 기본 재구성된 사운드 표현을 획득하는 단계를 추가로 포함할 수 있다. 본 방법은 기본 재구성된 사운드 표현을 개선(예컨대, 향상)시키기 위해 어느 향상 보조 정보 부분이 사용되어야만 하는지를 표시하는 제2 레이어 인덱스를 결정하는 단계를 추가로 포함할 수 있다. 본 방법은 또한, 제2 레이어 인덱스를 참조하여, 기본 재구성된 사운드 표현으로부터 사운드 또는 음장의 재구성된 사운드 표현을 획득하는 단계를 추가로 포함할 수 있다.

[0024] 이와 같이 구성될 때, 제안된 방법은, 이용가능한(예컨대, 유효하게 수신된) 정보를 가능한 한 최대한으로 사용하여, 재구성된 사운드 표현이 최적의 품질을 갖도록 보장한다.

[0025] 실시예들에서, 기본 압축된 사운드 표현의 컴포넌트들은 모노럴 신호들(예컨대, 모노럴 전송 신호들)에 대응할 수 있다. 모노럴 신호들은 HOA 표현의 우세 사운드 신호들 또는 계수 시퀀스들 중 어느 하나를 나타낼 수 있다. 모노럴 신호들은 양자화될 수 있다.

[0026] 실시예들에서, 기본 보조 정보는 복수의 컴포넌트들 중 하나 이상을, 다른 컴포넌트들과 독립적으로, 개별적으로 디코딩(예컨대, 압축해제)하는 것을 명시하는 정보를 포함할 수 있다. 예를 들어, 기본 보조 정보는, 다른 모노럴 신호들과 독립적으로, 개별 모노럴 신호들에 관련된 보조 정보를 표현할 수 있다. 따라서, 기본 보조 정보는 독립적 기본 보조 정보라고 지칭될 수 있다.

[0027] 실시예들에서, 향상 보조 정보는 향상 보조 정보를 나타낼 수 있다. 향상 보조 정보는 기본 압축된 사운드 표현 및 기본 보조 정보로부터 획득가능한 기본 재구성된 사운드 표현을 개선(예컨대, 향상)시키기 위한 기본 압축된 사운드 표현에 대한 예측 파라미터들을 포함할 수 있다.

[0028] 실시예들에서, 본 방법은, 각각의 레이어에 대해, 각자의 레이어가 유효하게 수신되었는지 여부를 결정하는 단계를 추가로 포함할 수 있다. 본 방법은 제1 레이어 인덱스를 유효하게 수신되지 않은 최하위 레이어 바로 아래의 레이어의 레이어 인덱스로서 결정하는 단계를 추가로 포함할 수 있다.

[0029] 실시예들에서, 제2 레이어 인덱스를 결정하는 단계는 제2 레이어 인덱스를 제1 레이어 인덱스와 동일하도록 결정하는 단계, 또는 재구성된 사운드 표현을 획득할 때 어떠한 향상 보조 정보도 사용하지 말 것을 표시하는 인덱스 값을 제2 레이어 인덱스로서 결정하는 단계 중 어느 하나를 포함할 수 있다. 후자의 경우에, 재구성된 사운드 표현은 기본 재구성된 사운드 표현과 동일할 수 있다.

[0030] 실시예들에서, 연속적 시간 구간들, 예를 들어, 동일한 크기의 시간 구간들에 대한 데이터 페이로드들이 수신되고 처리될 수 있다. 연속적 시간 구간들은 프레임들일 수 있다. 따라서, 본 방법은 프레임 기반으로 동작할 수 있다. 본 방법은, 연속적 시간 구간들에 대한 압축된 사운드 표현들이 서로 독립적으로 디코딩될 수 있는

경우, 제2 레이어 인덱스를 제1 레이어 인덱스와 동일하도록 결정하는 단계를 추가로 포함할 수 있다.

[0031] 실시예들에서, 연속적 시간 구간들, 예를 들어, 동일한 크기의 시간 구간들에 대한 데이터 페이로드들이 수신되고 처리될 수 있다. 연속적 시간 구간들은 프레임들일 수 있다. 따라서, 본 방법은 프레임 기반으로 동작할 수 있다. 본 방법은, 연속적 시간 구간들 중 주어진 시간 구간에 대해, 연속적 시간 구간들에 대한 압축된 사운드 표현들이 서로 독립적으로 디코딩될 수 없는 경우, 각각의 레이어에 대해, 각자의 레이어가 유효하게 수신되었는지 여부를 결정하는 단계를 추가로 포함할 수 있다. 본 방법은 주어진 시간 구간에 대한 제1 레이어 인덱스를 주어진 시간 구간에 선행하는 시간 구간의 제1 레이어 인덱스 및 유효하게 수신되지 않은 최하위 레이어 바로 아래의 레이어의 레이어 인덱스 중 작은 것으로 결정하는 단계를 추가로 포함할 수 있다.

[0032] 실시예들에서, 본 방법은, 주어진 시간 구간에 대해, 연속적 시간 구간들에 대한 압축된 사운드 표현들이 서로 독립적으로 디코딩될 수 없는 경우, 주어진 시간 구간에 대한 제1 레이어 인덱스가 선행하는 시간 구간에 대한 제1 레이어 인덱스와 동일한지 여부를 결정하는 단계를 추가로 포함할 수 있다. 본 방법은, 주어진 시간 구간에 대한 제1 레이어 인덱스가 선행하는 시간 구간에 대한 제1 레이어 인덱스와 동일한 경우, 주어진 시간 구간에 대한 제2 레이어 인덱스를 주어진 시간 구간에 대한 제1 레이어 인덱스와 동일하도록 결정하는 단계를 추가로 포함할 수 있다. 본 방법은, 주어진 시간 구간에 대한 제1 레이어 인덱스가 선행하는 시간 구간에 대한 제1 레이어 인덱스와 동일하지 않은 경우, 재구성된 사운드 표현을 획득할 때 어떠한 향상 보조 정보도 사용하지 말 것을 표시하는 인덱스 값을 제2 레이어 인덱스로서 결정하는 단계를 추가로 포함할 수 있다.

[0033] 실시예들에서, 베이스 레이어는, 각자의 레이어에 대응하고 각자의 레이어에 배정된 컴포넌트들 중 하나 이상의 컴포넌트를 각자의 레이어 및 각자의 레이어보다 하위인 임의의 레이어들에 배정된 다른 컴포넌트들에 의존하여 디코딩하는 것을 명시하는 정보를 포함하는, 적어도 하나의 부가 기본 보조 정보 부분을 포함할 수 있다. 본 방법은, 각각의 부가 기본 보조 정보 부분에 대해, 부가 기본 보조 정보 부분을, 그 각자의 레이어 및 각자의 레이어보다 하위인 임의의 레이어들에 배정된 컴포넌트들을 참조하여 디코딩하는 단계를 추가로 포함할 수 있다. 본 방법은 부가 기본 보조 정보 부분을 최상위 사용가능 레이어 및 최상위 사용가능 레이어와 각자의 레이어 사이의 임의의 레이어들에 배정된 컴포넌트들을 참조하여 정정하는 단계를 추가로 포함할 수 있다. 기본 재구성된 사운드 표현은, 기본 보조 정보 및 최상위 사용가능 레이어까지의 레이어들에 대응하는 부가 기본 보조 정보 부분들로부터 획득된 정정된 부가 기본 보조 정보 부분들을 사용하여, 최상위 사용가능 레이어 및 최상위 사용가능 레이어보다 하위인 임의의 레이어들에 배정된 컴포넌트들로부터 획득될 수 있다.

[0034] 실시예들에서, 부가 기본 보조 정보는 복수의 컴포넌트들 중 하나 이상을 다른 컴포넌트들에 의존하여 디코딩(예컨대, 압축해제)하는 것을 명시하는 정보를 포함할 수 있다. 예를 들어, 부가 기본 보조 정보는 다른 모노럴 신호들에 의존하여 개별 모노럴 신호들에 관련된 보조 정보를 표현할 수 있다. 따라서, 부가 기본 보조 정보는 종속적 기본 보조 정보라고 지칭될 수 있다.

[0035] 다른 양태에 따르면, 사운드 또는 음장의 압축된 사운드 표현을 디코딩하는 방법이 기술된다. 압축된 사운드 표현은 복수의 계층적 레이어들에 인코딩되었을 수 있다. 복수의 계층적 레이어들은 베이스 레이어 및 하나 이상의 계층적 향상 레이어를 포함할 수 있다. 복수의 레이어들은 사운드 또는 음장의 기본 압축된 사운드 표현의 컴포넌트들을 배정받았을 수 있다. 환언하면, 복수의 레이어들은 기본 압축된 보조 정보의 컴포넌트들을 포함할 수 있다. 컴포넌트들은 각자의 컴포넌트 그룹들 내의 각자의 레이어들에 배정될 수 있다. 복수의 컴포넌트들은 상보적 컴포넌트들일 수 있다. 베이스 레이어는 기본 압축된 사운드 표현을 디코딩하기 위한 기본 보조 정보를 포함할 수 있다. 베이스 레이어는, 각자의 레이어에 대응하고 각자의 레이어에 배정된 컴포넌트들 중 하나 이상의 컴포넌트를 각자의 레이어 및 각자의 레이어보다 하위인 임의의 레이어들에 배정된 다른 컴포넌트들에 의존하여 디코딩하는 것을 명시하는 정보를 포함하는, 적어도 하나의 부가 기본 보조 정보 부분을 추가로 포함할 수 있다. 본 방법은 복수의 계층적 레이어들에 각각 대응하는 데이터 페이로드들을 수신하는 단계를 포함할 수 있다. 본 방법은 기본 압축된 사운드 표현을 사운드 또는 음장의 기본 재구성된 사운드 표현으로 디코딩하기 위해 사용될 복수의 레이어들 중 최상위 사용가능 레이어를 표시하는 제1 레이어 인덱스를 결정하는 단계를 추가로 포함할 수 있다. 본 방법은, 각각의 부가 기본 보조 정보 부분에 대해, 부가 기본 보조 정보 부분을 그 각자의 레이어 및 각자의 레이어보다 하위인 임의의 레이어들에 배정된 컴포넌트들을 참조하여 디코딩하는 단계를 추가로 포함할 수 있다. 본 방법은, 각각의 부가 기본 보조 정보 부분에 대해, 부가 기본 보조 정보 부분을 최상위 사용가능 레이어 및 최상위 사용가능 레이어와 각자의 레이어 사이의 임의의 레이어들에 배정된 컴포넌트들을 참조하여 정정하는 단계를 추가로 포함할 수 있다. 기본 재구성된 사운드 표현은, 기본 보조 정보 및 최상위 사용가능 레이어까지의 레이어들에 대응하는 부가 기본 보조 정보 부분들로부터 획득된 정정된 부가 기본 보조 정보 부분들을 사용하여, 최상위 사용가능 레이어 및 최상위 사용가능 레이어보다 하위인 임의의

레이어들에 배정된 컴포넌트들로부터 획득될 수 있다. 본 방법은 디코딩 동안 제1 레이어 인덱스와 동일한 또는 항상 보조 정보의 생략을 표시하는 제2 레이어 인덱스를 결정하는 단계를 추가로 포함할 수 있다.

[0036] 이와 같이 구성될 때, 제안된 방법은 기본 압축된 사운드 표현을 디코딩하는 데 궁극적으로 사용되는 부가 기본 보조 정보가 중복적 요소들을 포함하지 않도록 보장하고, 그에 의해 기본 압축된 사운드 표현의 실제의 디코딩을 보다 효율적으로 만든다.

[0037] 이 양태의 실시예들은 전술한 양태의 실시예들에 관련될 수 있다.

[0038] 다른 양태에 따르면, 사운드 또는 음장의 압축된 사운드 표현의 계층화된 인코딩을 위한 인코더가 기술된다. 압축된 사운드 표현은 복수의 컴포넌트들을 포함하는 기본 압축된 사운드 표현을 포함할 수 있다. 복수의 컴포넌트들은 상보적 컴포넌트들일 수 있다. 압축된 사운드 표현은 기본 압축된 사운드 표현을 사운드 또는 음장의 기본 재구성된 사운드 표현으로 디코딩하기 위한 기본 보조 정보를 추가로 포함할 수 있다. 압축된 사운드 표현은 또한 기본 재구성된 사운드 표현을 개선(예컨대, 항상)시키기 위한 파라미터들을 포함하는 항상 보조 정보를 추가로 포함할 수 있다. 인코더는 앞서 언급된 제1 양태 및 앞서 언급된 제2 양태에 따른 방법들의 방법 단계들의 일부 또는 전부를 수행하도록 구성된 프로세서를 포함할 수 있다.

[0039] 다른 양태에 따르면, 사운드 또는 음장의 압축된 사운드 표현을 디코딩하기 위한 디코더가 기술된다. 압축된 사운드 표현은 복수의 계층적 레이어들에 인코딩되었을 수 있다. 복수의 계층적 레이어들은 베이스 레이어 및 하나 이상의 계층적 항상 레이어를 포함할 수 있다. 복수의 레이어들은 사운드 또는 음장의 기본 압축된 사운드 표현의 컴포넌트들을 배정받았을 수 있다. 환언하면, 복수의 레이어들은 기본 압축된 보조 정보의 컴포넌트들을 포함할 수 있다. 컴포넌트들은 각자의 컴포넌트 그룹들 내의 각자의 레이어들에 배정될 수 있다. 복수의 컴포넌트들은 상보적 컴포넌트들일 수 있다. 베이스 레이어는 기본 압축된 사운드 표현을 디코딩하기 위한 기본 보조 정보를 포함할 수 있다. 각각의 레이어는 각자의 레이어 및 각자의 레이어보다 하위인 임의의 레이어들에 포함된 데이터로부터 획득가능한 기본 재구성된 사운드 표현을 개선(예컨대, 항상)시키기 위한 파라미터들을 포함하는 항상 보조 정보 부분을 포함할 수 있다. 디코더는 앞서 언급된 제3 양태 및 앞서 언급된 제4 양태에 따른 방법들의 방법 단계들의 일부 또는 전부를 수행하도록 구성된 프로세서를 포함할 수 있다.

[0040] 다른 양태들에 따르면, 방법들, 장치들 및 시스템들은 사운드 또는 음장의 압축된 고차 앰비소닉스(HOA) 사운드 표현을 디코딩하는 것에 관한 것이다. 본 장치는 베이스 레이어 및 하나 이상의 계층적 항상 레이어를 포함하는 복수의 계층적 레이어들에 대응하는 압축된 HOA 표현을 포함하는 비트스트림을 수신하도록 구성된 수신기를 가질 수 있거나 본 방법은 그 비트스트림을 수신할 수 있다. 복수의 레이어들은 사운드 또는 음장의 기본 압축된 사운드 표현의 컴포넌트들을 배정받았고, 컴포넌트들은 각자의 컴포넌트 그룹들 내의 각자의 레이어들에 배정된다. 본 장치는 베이스 레이어와 연관되어 있는 기본 보조 정보에 기초하여 그리고 하나 이상의 계층적 항상 레이어와 연관되어 있는 항상 보조 정보에 기초하여 압축된 HOA 표현을 디코딩하도록 구성된 디코더를 가질 수 있고, 본 방법은 그 압축된 HOA 표현을 디코딩할 수 있다. 기본 보조 정보는 다른 모노럴 신호들과 독립적으로 디코딩될 제1 개별 모노럴 신호들에 관련된 기본 독립적 보조 정보를 포함할 수 있다. 하나 이상의 계층적 항상 레이어 각각은 각자의 레이어들 및 각자의 레이어보다 하위인 임의의 레이어들에 포함된 데이터로부터 획득가능한 기본 재구성된 사운드 표현을 개선시키기 위한 파라미터들을 포함하는 항상 보조 정보의 일부분을 포함할 수 있다.

[0041] 기본 독립적 보조 정보는 제1 개별 모노럴 신호들이 입사 방향을 갖는 방향성 신호(directional signal)를 나타낸다는 것을 표시할 수 있다. 기본 보조 정보는 다른 모노럴 신호들에 의존하여 디코딩될 제2 개별 모노럴 신호들에 관련된 기본 종속적 보조 정보를 추가로 포함할 수 있다. 기본 종속적 보조 정보는 음장 내에서 방향성으로 분포된(directionally distributed) 벡터 기반 신호들을 포함할 수 있으며, 여기서 방향성 분포(directional distribution)는 벡터에 의해 명시된다. 벡터의 컴포넌트들은 0으로 설정되고 압축된 벡터 표현의 일부가 아니다.

[0042] 기본 압축된 사운드 표현의 컴포넌트들은 HOA 표현의 우세 사운드 신호들 또는 계수 시퀀스들 중 어느 하나를 나타내는 모노럴 신호들에 대응할 수 있다. 비트스트림은 복수의 계층적 레이어들에 각각 대응하는 데이터 패이로드들을 포함한다. 항상 보조 정보는 공간 예측, 서브대역 방향성 신호 합성, 및 파라메트릭 앰비언스 복제(parametric ambience replication) 중 적어도 하나에 관련된 파라미터들을 포함할 수 있다. 항상 보조 정보는 방향성 신호들로부터 사운드 또는 음장의 누락 부분들의 예측을 가능하게 하는 정보를 포함할 수 있다. 각각의 레이어에 대해, 각자의 레이어가 유효하게 수신되었는지 여부 및 유효하게 수신되지 않은 최하위 레이어 바로 아래의 레이어의 레이어 인덱스가 추가로 결정될 수 있다.

- [0043] 다른 양태에 따르면, 소프트웨어 프로그램이 기술된다. 소프트웨어 프로그램은 프로세서 상에서 실행되도록 그리고 컴퓨팅 디바이스 상에서 실행될 때 본 문서에 개요가 기술된 방법 단계들 중 일부 또는 전부를 수행하도록 적합화될 수 있다.
- [0044] 또 다른 양태에 따르면, 저장 매체가 기술된다. 저장 매체는 프로세서 상에서 실행되도록 그리고 컴퓨팅 디바이스 상에서 실행될 때 본 문서에 개요가 기술된 방법 단계들 중 일부 또는 전부를 수행하도록 적합화된 소프트웨어 프로그램을 포함할 수 있다.
- [0045] 이상의 양태들 또는 그의 실시예들 중 임의의 것에 관해 이루어진 진술들은, 통상의 기술자가 알게 될 것인 바와 같이, 각자의 다른 양태들 또는 그들의 실시예들에도 적용된다. 이 진술들을 모든 양태 또는 실시예에 대해 반복하는 것이 간결함을 위해 생략되었다.
- [0046] 본 방법들 및 장치들은, 본 문서에 개요가 기술된 그들의 바람직한 실시예들을 비롯하여, 단독으로 또는 본 문서에 개시된 다른 방법들 및 시스템들과 결합하여 사용될 수 있다. 게다가, 본 문서에 개요가 기술된 방법들 및 장치들의 모든 양태들이 임의적으로 조합될 수 있다. 상세하게는, 청구항들의 피처(feature)들이 임의적인 방식으로 서로 조합될 수 있다.
- [0047] 방법 단계들 및 장치 피처들은 많은 방식들로 상호교환될 수 있다. 상세하게는, 통상의 기술자가 알 것인 바와 같이, 개시된 방법의 상세들이 방법의 단계들의 일부 또는 전부를 실행하도록 적합화된 장치로서 구현될 수 있고 그 반대일 수도 있다.

도면의 간단한 설명

- [0048] 본 발명이 이하에서 예시적으로 첨부 도면들을 참조하여 설명된다.
- 도 1은 본 개시내용의 실시예들에 따른 계층화된 인코딩 방법의 일 예를 예시한 플로차트;
- 도 2는 본 개시내용의 실시예들에 따른 인코더 스테이지의 일 예를 개략적으로 예시한 블록 다이어그램;
- 도 3은 본 개시내용의 실시예들에 따른, 복수의 계층적 레이어들로 인코딩된 사운드 또는 음장의 압축된 사운드 표현을 디코딩하는 방법의 일 예를 예시한 플로차트;
- 도 4a 및 도 4b는 본 개시내용의 실시예들에 따른 디코더 스테이지의 예들을 개략적으로 예시한 블록 다이어그램;
- 도 5는 본 개시내용의 실시예들에 따른 인코더의 하드웨어 구현의 일 예를 개략적으로 예시한 블록 다이어그램;
- 도 6은 본 개시내용의 실시예들에 따른 디코더의 하드웨어 구현의 일 예를 개략적으로 예시한 블록 다이어그램.

발명을 실시하기 위한 구체적인 내용

- [0049] 먼저, 본 개시내용에 따른 방법들 및 인코더들/디코더들이 적용가능한 압축된 사운드(또는 음장) 표현(이후부터, 간결함을 위해 압축된 사운드 표현이라고 지칭됨)이 기술될 것이다. 일반적으로, 완전한 압축된 사운드(또는 음장) 표현(이후부터, 간결함을 위해 완전한 압축된 사운드 표현이라고 지칭됨)은 다음과 같은 3개의 컴포넌트: 기본 압축된 사운드(또는 음장) 표현(이후부터, 간결함을 위해 기본 압축된 사운드 표현이라고 지칭됨), 기본 보조 정보, 및 항상 보조 정보를 포함할 수 있다(예컨대, 이들로 이루어져 있을 수 있다).
- [0050] 기본 압축된 사운드 표현 자체는 다수의 컴포넌트들(예컨대, 상보적 컴포넌트들)을 포함한다(예컨대, 이들로 이루어져 있다). 기본 압축된 사운드 표현은 완전한 압축된 사운드 표현의 두드러지게 가장 큰 비율(percentag e)을 차지할 수 있다. 기본 압축된 사운드 표현은 원래의 HOA 표현의 우세 사운드 신호들 또는 계수 시퀀스들 중 어느 하나를 나타내는 모노럴 전송 신호들로 이루어져 있을 수 있다.
- [0051] 기본 보조 정보는 기본 압축된 사운드 표현을 디코딩하는 데 필요하며 기본 압축된 사운드 표현과 비교하여 훨씬 더 작은 크기인 것으로 가정될 수 있다. 이는 대부분 비중복 부분(disjoint portion)들로 이루어져 있을 수 있으며, 비중복 부분들 각각은 기본 압축된 사운드 표현의 단지 하나의 특정 컴포넌트의 압축해제를 명시한다. 기본 보조 정보는 독립적 기본 보조 정보라고 알려져 있을 수 있는 제1 파트 및 부가 기본 보조 정보라고 알려져 있을 수 있는 제2 파트로 이루어져 있을 수 있다.
- [0052] 독립적 기본 보조 정보 및 부가 기본 보조 정보인 제1 및 제2 파트들 둘 다는 기본 압축된 사운드 표현의 특정의 컴포넌트들의 압축해제를 명시할 수 있다. 제2 파트는 임의적이며 생략될 수 있다. 이 경우에, 압축된 사

운드 표현은 제1 파트(예컨대, 기본 보조 정보)를 포함한다고 말해질 수 있다.

[0053] 제1 파트(예컨대, 기본 보조 정보)는 다른 (상보적) 컴포넌트들과 독립적으로 기본 압축된 사운드 표현의 개별 (상보적) 컴포넌트들을 기술하는 보조 정보를 포함할 수 있다. 상세하게는, 제1 파트(예컨대, 기본 보조 정보)는 복수의 컴포넌트들 중 하나 이상을, 다른 컴포넌트들과 독립적으로, 개별적으로 디코딩하는 것을 명시할 수 있다. 따라서, 제1 파트는 독립적 기본 보조 정보라고 지칭될 수 있다.

[0054] 제2 (임의적) 파트는, 부가 기본 보조 정보라고도 알려져 있는, 보조 정보를 포함할 수 있고, 기본 압축된 사운드 표현의 개별 (상보적) 컴포넌트들을 다른 (상보적) 컴포넌트들에 의존하여 기술할 수 있다. 이 제2 파트는 종속적 기본 보조 정보라고도 지칭될 수 있다. 상세하게는, 종속성(dependence)은 다음과 같은 특성들을 가질 수 있다:

- 기본 압축된 사운드 표현의 각각의 개별 (상보적) 컴포넌트에 대한 종속적 기본 보조 정보는, 기본 압축된 사운드 표현에 다른 특정 (상보적) 컴포넌트들이 포함되어 있지 않을 때, 그의 가장 큰 크기(extent)를 달성할 수 있다.

[0055] - 부가의 특정 (상보적) 컴포넌트들이 기본 압축된 사운드 표현에 추가되는 경우에, 고려된 개별 (상보적) 컴포넌트에 대한 종속적 기본 보조 정보는 원래의 종속적 기본 보조 정보의 서브세트로 될 수 있고, 그에 의해 그의 크기를 감소시킬 수 있다.

[0056] 향상 보조 정보가 또한 임의적이다. 이는 기본 압축된 사운드 표현을 개선 또는 향상(예컨대, 파라미터적으로 개선 또는 향상)시키는 데 사용될 수 있다. 그의 크기가 또한 기본 압축된 사운드 표현의 크기보다 훨씬 더 작은 것으로 가정될 수 있다.

[0057] 따라서, 실시예들에서, 압축된 사운드 표현은 복수의 컴포넌트들을 포함하는 기본 압축된 사운드 표현, 기본 압축된 사운드 표현을 사운드 또는 음장의 기본 재구성된 사운드 표현으로 디코딩(예컨대, 압축해제)하기 위한 기본 보조 정보, 및 기본 재구성된 사운드 표현을 개선 또는 향상(예컨대, 파라미터적으로 개선 또는 향상)시키기 위한 파라미터들을 포함하는 향상 보조 정보를 포함할 수 있다. 압축된 사운드 표현은 기본 압축된 사운드 표현을 기본 재구성된 사운드 표현으로 디코딩(예컨대, 압축해제)하기 위한 부가 기본 보조 정보를 추가로 포함할 수 있으며, 부가 기본 보조 정보는 복수의 컴포넌트들 중 하나 이상을 각자의 다른 컴포넌트들에 의존하여 디코딩하는 것을 명시하는 정보를 포함할 수 있다.

[0058] 이러한 유형의 완전한 압축된 사운드 표현의 일 예는 MPEG-H 3D 오디오 표준의 예비 버전(참고문헌 1), 제12장 및 부록 C.5에 의해 명시된 바와 같은 압축된 고차 앰비소닉스(HOA) 음장 표현에 의해 주어진다. 즉, 압축된 사운드 표현은 사운드 또는 음장의 압축된 HOA 사운드(또는 음장) 표현에 대응할 수 있다.

[0059] 이 예에서, 기본 압축된 음장 표현(기본 압축된 사운드 표현)은 다수의 컴포넌트들을 포함할 수 있다(예컨대, 다수의 컴포넌트들로 식별될 수 있다). 컴포넌트들은 모노럴 신호들일 수 있다(예컨대, 모노럴 신호들에 대응할 수 있다). 모노럴 신호들은 양자화된 모노럴 신호들일 수 있다. 모노럴 신호들은 주변 HOA 음장 컴포넌트 (ambient HOA sound field component)의 우세 사운드 신호들 또는 계수 시퀀스들 중 어느 하나를 나타낼 수 있다.

[0060] 기본 보조 정보는, 그 중에서도 특히, 이 모노럴 신호들 각각에 대해 모노럴 신호가 음장에 어떻게 공간적으로 기여하는지를 기술할 수 있다. 예를 들어, 기본 보조 정보는 우세 사운드 신호를, 특정 입사 방향을 갖는 일반 평면파(general plane wave)를 의미하는, 순수 방향성 신호(purely directional signal)로서 명시할 수 있다. 대안적으로, 기본 보조 정보는 모노럴 신호를 특정 인덱스를 갖는 원래의 HOA 표현의 계수 시퀀스로서 명시할 수 있다. 기본 보조 정보는, 앞서 살펴본 바와 같이, 제1 파트와 제2 파트로 추가로 분리될 수 있다.

[0061] 제1 파트는 특정 개별 모노럴 신호들에 관련된 보조 정보(예컨대, 독립적 기본 보조 정보)이다. 이 독립적 기본 보조 정보는 다른 모노럴 신호들의 존재와 독립적이다. 이러한 보조 정보는, 예를 들어, 특정 입사 방향을 갖는 방향성 신호(예컨대, 일반 평면파를 의미함)를 표현하는 모노럴 신호를 명시할 수 있다. 대안적으로, 모노럴 신호는 특정 인덱스를 갖는 원래의 HOA 표현의 계수 시퀀스로서 명시될 수 있다. 제1 파트는 독립적 기본 보조 정보라고 지칭될 수 있다. 일반적으로, 제1 파트(예컨대, 기본 보조 정보)는 복수의 모노럴 신호들 중 하나 이상을, 다른 모노럴 신호들과 독립적으로, 개별적으로 디코딩하는 것을 명시할 수 있다.

[0062] 제2 파트는 특정 개별 모노럴 신호들에 관련된 보조 정보(예컨대, 부가 기본 보조 정보)이다. 이 보조 정보는 다른 모노럴 신호들의 존재에 의존한다. 예를 들어, 모노럴 신호들이 벡터 기반 신호들인 것으로 명시되는 경

우, 이러한 보조 정보가 이용될 수 있다(예컨대, 참고문헌 1, 섹션 12.4.2.4.4를 참조). 이 신호들은 음장 내에서 방향성으로 분포되며, 여기서 방향성 분포는 벡터에 의해 명시될 수 있다. 특정 모드(예컨대, CodedVVecLength = 1을 참조)에서, 이 벡터의 특정의 컴포넌트들은 암시적으로 0으로 설정되고 압축된 벡터 표현의 일부가 아니다. 이 컴포넌트들은 원래의 HOA 표현의 계수 시퀀스들의 인덱스들과 동일한 인덱스들을 갖는 컴포넌트들이고 기본 압축된 사운드 표현의 일부이다. 이는, 벡터의 개별 컴포넌트들이 코딩되는 경우, 그들의 총수가 기본 압축된 사운드 표현에 의존할 수 있다는 것을 의미한다. 상세하게는, 총수는 원래의 HOA 표현이 어느 계수 시퀀스들을 포함하는지에 의존할 수 있다.

[0064] 원래의 HOA 표현의 계수 시퀀스들이 기본 압축된 사운드 표현에 포함되지 않은 경우, 각각의 벡터 기반 신호에 대한 종속적 기본 보조 정보는 벡터 컴포넌트들 전부로 이루어져 있고 그의 가장 큰 크기를 갖는다. 특정 인덱스들을 갖는 원래의 HOA 표현의 계수 시퀀스들이 기본 압축된 사운드 표현에 추가되는 경우에, 그 인덱스들을 갖는 벡터 컴포넌트들이 각각의 벡터 기반 신호에 대한 보조 정보로부터 제거되고, 그에 의해 벡터 기반 신호들에 대한 종속적 기본 보조 정보의 크기를 감소시킨다.

[0065] 항상 보조 정보(예컨대, 향상 보조 정보)는 (광대역) 공간 예측에 관련된 파라미터들(참고문헌 1, 섹션 12.4.2.4.3을 참조) 및/또는 서브대역 방향성 신호 합성 및 파라메트릭 앰비언스 복제에 관련된 파라미터들을 포함할 수 있다.

[0066] (광대역) 공간 예측에 관련된 파라미터들은 방향성 신호들로부터 음장의 누락 부분들을 (선형적으로) 예측하는데 사용될 수 있다.

[0067] 서브대역 방향성 신호 합성 및 파라메트릭 앰비언스 복제는 수정안[참고문헌 2, 섹션 1을 참조]을 갖는 MPEG-H 3D 오디오 표준에 최근에 도입된 압축 도구들이다. 이 2개의 도구는 공간적으로 불완전한 또는 결함있는 압축된 HOA 표현을 보완하기 위해 부가의 모노럴 신호들의 주파수 종속적 파라메트릭 예측(frequency-dependent parametric-prediction)이 공간적으로 분산될 수 있게 한다. 예측은 기본 압축된 사운드 표현의 계수 시퀀스들에 기초할 수 있다.

[0068] 유의할 중요한 점은 음장에 대한 앞서 언급된 상보적 기여가, 부가의 양자화된 신호들에 의해서가 아니라 오히려 비교할 수 있을 정도로 훨씬 더 작은 크기의 추가의 보조 정보에 의해, 압축된 HOA 표현 내에 표현된다는 것이다. 따라서, 2개의 언급된 코딩 도구는 낮은 데이터 레이트들에서 HOA 표현들의 압축에 특히 적합하다.

[0069] 앞서 언급된 구조를 갖는 하나 이상의 모노럴 신호의 압축된 표현의 제2 예는, 기본 압축된 표현으로 간주될 수 있는, 특정 상부 주파수까지의 비중복 주파수 대역(disjoint frequency band)들에 대한 코딩된 스펙트럼 정보; (예컨대, 코딩된 주파수 대역들의 수 및 폭에 의해) 코딩된 스펙트럼 정보를 명시하는 기본 보조 정보; 및 기본 압축된 표현에서 고려되지 않은 상위 주파수 대역들에 대한 스펙트럼 정보를 기본 압축된 표현으로부터 어떻게 파라미터적으로 재구성할지를 기술하는, 스펙트럼 대역 복제(Spectral Band Replication)(SBR)의 파라미터들을 포함하는(예컨대, 이들로 이루어진) 항상 보조 정보를 포함한다.

[0070] 본 개시내용은 앞서 언급된 구조를 갖는 완전한 압축된 사운드(또는 음장) 표현의 계층화된 코딩을 위한 방법을 제안한다.

[0071] 압축이 연속적 시간 구간들에 대한 (데이터 패킷들 또는 등가적으로 프레임 페이로드들의 형태의) 압축된 표현들을 제공한다는 의미에서, 압축은 프레임 기반일 수 있다. 시간 구간들은 동일하거나 상이한 크기들을 가질 수 있다. 이 데이터 패킷들은 유효성 플래그, 그들의 크기를 표시하는 값은 물론 실제의 압축된 표현 데이터를 포함하는 것으로 가정될 수 있다. 이하에서, 제한하려는 의도 없이, 압축이 프레임 기반이라고 가정될 것이다. 게다가, 달리 언급되지 않는 한 그리고 제한하려는 의도 없이, 단일 프레임의 처리에 중점을 둘 것이고, 따라서 프레임 인덱스가 생략될 것이다.

[0072] 고려 중인 완전한 압축된 사운드(또는 음장) 표현의 각각의 프레임 페이로드는 $BSRC_j$, $j=1, \dots, J$ 에 의해 표기되는 J 개의 데이터 패킷 - 각각은 기본 압축된 사운드 표현의 하나의 컴포넌트에 대한 것임 - 을 포함하는 것으로 가정된다. 게다가, 프레임 페이로드는 기본 압축된 사운드 표현의 특정의 컴포넌트들($BSRC_j$)을 다른 컴포넌트들과 독립적으로 명시하는 BSI_1 에 의해 표기된 독립적 기본 보조 정보(기본 보조 정보)를 갖는 패킷을 포함하는 것으로 가정된다. 임의로, 프레임 페이로드는, 그에 부가하여, 기본 압축된 사운드 표현의 특정의 컴포넌트들($BSRC_j$)을 다른 컴포넌트들에 의존하여 명시하는 BSI_D 에 의해 표기된 종속적 기본 보조 정보(부가 기본 보조 정보)를 갖는 패킷을 포함하는 것으로 가정된다.

- [0073] 2개의 데이터 패킷(BSI_I 및 BSI_D) 내에 포함된 정보는 기본 보조 정보의 단일 데이터 패킷(BSI)으로 임의로 그룹화될 수 있다. 단일 데이터 패킷(BSI)은, 그 중에서도 특히, J개의 부분들을 포함하는 것으로 말해질 수 있고, 그 각각은 기본 압축된 사운드 표현의 하나의 특정의 컴포넌트(BSRC_j)를 명시한다. 이 부분들 각각은 차례로 독립적 보조 정보의 일부분 그리고, 임의로, 종속적 보조 정보의 일부분을 포함하는 것으로 말해질 수 있다.
- [0074] 궁극적으로, 프레임 페이로드는 완전한 기본 압축된 사운드 표현으로부터 재구성된 사운드(또는 음장)를 어떻게 개선 또는 향상시킬지에 대한 설명을 갖는 ESI에 의해 표기된 향상 보조 정보 페이로드(향상 보조 정보)를 포함할 수 있다.
- [0075] 계층화된 코딩을 위한 제안된 해결책은 전송을 위한 데이터 패킷들의 패킹(packing)을 포함하는 압축 파트는 물론 수신기 및 압축해제 파트 둘 다를 가능하게 하는 데 요구된 단계들을 다루고 있다. 각각의 파트는 이하에서 상세히 기술될 것이다.
- [0076] 먼저, (예컨대, 전송을 위한) 압축 및 패킹이 기술될 것이다. 상세하게는, 계층화된 코딩의 경우의 완전한 압축된 사운드(또는 음장) 표현의 컴포넌트들 및 요소들이 기술될 것이다.
- [0077] 도 1은 압축 및 패킹 방법(예컨대, 인코딩 방법, 또는 사운드 또는 음장의 압축된 사운드 표현의 계층화된 인코딩 방법)의 일 예의 플로차트를 개략적으로 예시하고 있다. 개별 페이로드들을 베이스 레이어 및 (M-1)개의 향상 레이어에 배정(예컨대, 할당)하는 것은 전송 레이어들 패커(transport layers packer)에 의해 달성될 수 있다. 도 2는 개별 페이로드들의 배정/할당의 일 예의 블록 다이어그램을 개략적으로 예시하고 있다.
- [0078] 앞서 살펴본 바와 같이, 완전한 압축된 사운드 표현(2100)은, 예를 들어, 기본 압축된 사운드 표현을 포함하는 압축된 HOA 표현에 관련될 수 있다. 완전한 압축된 사운드 표현(2100)은 복수의 컴포넌트들(예컨대, 모노럴 신호들)(2110-1, ..., 2110-J), 독립적 기본 보조 정보(기본 보조 정보)(2120), 임의적인 향상 보조 정보(향상 보조 정보)(2140), 및 임의적인 종속적 기본 보조 정보(부가 기본 보조 정보)(2130)를 포함할 수 있다. 기본 보조 정보(2120)는 기본 압축된 사운드 표현을 사운드 또는 음장의 기본 재구성된 사운드 표현으로 디코딩하기 위한 정보일 수 있다. 기본 보조 정보(2120)는 하나 이상의 컴포넌트(예컨대, 모노럴 신호)를, 다른 컴포넌트들과 독립적으로, 개별적으로 디코딩하는 것을 명시하는 정보를 포함할 수 있다. 향상 보조 정보(2140)는 기본 재구성된 사운드 표현을 개선(예컨대, 향상)시키기 위한 파라미터들을 포함할 수 있다. 부가 기본 보조 정보(2130)는 기본 압축된 사운드 표현을 기본 재구성된 사운드 표현으로 디코딩하기 위한 (추가의) 정보일 수 있으며, 복수의 컴포넌트들 중 하나 이상을 각자의 다른 컴포넌트들에 의존하여 디코딩하는 것을 명시하는 정보를 포함할 수 있다.
- [0079] 도 2는 하나의 베이스 레이어(기본 레이어) 및 하나 이상의 (계층적) 향상 레이어를 포함하는, 복수의 계층적 레이어들이 존재하는, 기반을 이루는 가정을 예시하고 있다. 예를 들어, 총 M개의 레이어, 즉 하나의 베이스 레이어 및 M-1개의 향상 레이어가 있을 수 있다. 복수의 계층적 레이어들은 순차적으로 증가하는 레이어 인덱스를 갖는다. 레이어 인덱스의 최저값(예컨대, 레이어 인덱스 1)은 베이스 레이어에 대응한다. 레이어들이 베이스 레이어로부터 향상 레이어들을 거쳐 전체 최상위 향상 레이어(즉, 전체 최상위 레이어)까지 순서화된다는 것이 추가로 이해된다.
- [0080] 제안된 방법은 프레임 기반으로(즉, 프레임 단위로) 수행될 수 있다. 상세하게는, 연속적 시간 구간들, 예를 들어, 동일한 크기의 시간 구간들에 대한 압축된 사운드 표현(2100)이 압축될 수 있다. 각각의 시간 구간은 프레임에 대응할 수 있다. 각각의 연속적 시간 구간(예컨대, 프레임)에 대해 이하에 기술되는 단계들이 수행될 수 있다.
- [0081] 도 1의 S1010에서, 복수의 컴포넌트들(2110)이 복수의 컴포넌트 그룹들로 세분된다. 복수의 그룹들 각각은 이어서 복수의 계층적 레이어들의 각자의 레이어에 배정(예컨대, 추가, 또는 할당)된다. 거기에서, 그룹들의 수는 레이어들의 수에 대응한다. 예를 들어, 그룹들의 수는 레이어들의 수와 동일할 수 있고, 따라서 각각의 레이어에 대해 하나의 컴포넌트 그룹이 있다. 앞서 살펴본 바와 같이, 복수의 레이어들은 베이스 레이어 및 하나 이상의(예컨대, M-1개의) 계층적 향상 레이어를 포함할 수 있다.
- [0082] 환연하면, 기본 압축된 사운드 표현은 개별 레이어들에 배정될 파트들로 세분된다. 일반성을 잊지 않고, $J_{m-1} \leq j < J_m$ 에 대해 컴포넌트들(BSRC_j)이 제m 레이어에 배정되도록, 그룹화는 M+1개의 숫자들($J_m, m=0, \dots, M$) 고, $J_0 = 1$ 이며 $J_M = J+1$ 임)에 의해 기술될 수 있다.

- [0083] S1020에서, 컴포넌트 그룹들이 그 각자의 레이어들에 배정된다. S1030에서, 기본 보조 정보(2120)가 베이스 레이어(즉, 복수의 계층적 레이어들 중 최하위 레이어)에 추가(예컨대, 할당)된다.
- [0084] 즉, 그의 작은 크기로 인해, 그의 불필요한 단편화를 회피하기 위해, 완전한 기본 보조 정보(기본 보조 정보 및 임의적인 부가 기본 보조 정보)를 베이스 레이어에 포함시키는 것이 제안된다.
- [0085] 고려 중인 압축된 사운드 표현이 종속적 기본 보조 정보(부가 기본 보조 정보)를 포함하는 경우, 본 방법은 부가 기본 보조 정보를 복수의 부가 기본 보조 정보 부분들(2130-1, ..., 2130-M)로 분해하는 단계(도 1에 도시되지 않음)를 추가로 포함할 수 있다. 부가 기본 보조 정보 부분들이 이어서 베이스 레이어에 추가(예컨대, 할당)될 수 있다. 환언하면, 부가 기본 보조 정보 부분들이 베이스 레이어에 포함될 수 있다. 각각의 부가 기본 보조 정보 부분은 각자의 레이어에 대응할 수 있고, 각자의 레이어에 배정된 하나 이상의 컴포넌트를 각자의 레이어 및 각자의 레이어보다 하위인 임의의 레이어들에 배정된 다른 컴포넌트들에 의존하여 디코딩하는 것을 명시하는 정보를 포함할 수 있다.
- [0086] 따라서, 독립적 기본 보조 정보(BSI_I)(기본 보조 정보)(2120)는 배정을 위해 변경되지 않은 채로 있지만, 한편으로는 수신기측에서의 올바른 디코딩을 가능하게 하기 위해 그리고 다른 한편으로는 전송될 종속적 기본 보조 정보의 크기를 감소시키기 위해, 종속적 기본 보조 정보가 특히 계층화된 코딩을 위해 핸들링되어야만 한다. 종속적 기본 보조 정보를 BSI_{D,m}, m = 1, ..., M에 의해 표기된 M개의 파트(부분)로 분해하는 것이 제안되고, 여기서, 고려 중인 압축된 사운드 표현에 대해 임의적인 종속적 기본 보조 정보가 존재한다고 가정할 때, 제m 파트는 제m 레이어에 배정된 기본 압축된 사운드 표현의 컴포넌트들(BSRC_j, J_{m-1} ≤ j < J_m) 각각에 대한 종속적 기본 보조 정보를 포함한다. 각자의 종속적 보조 정보가 존재하지 않는 경우에, 파트들(BSI_{D,m})의 압축된 사운드 표현이 비어있는 것으로 가정될 수 있다. 종속적 기본 보조 정보의 각각의 파트(BSI_{D,m})는, 제m 레이어까지의 레이어들 전부에 포함된(즉, 레이어들(j = 1, ..., m) 전부에 포함된), 컴포넌트들(BSRC_j, 1 ≤ j < J_m) 전부에 의존할 수 있다.
- [0087] 독립적 기본 보조 정보 패킷(BSI_I)이 무시할 수 있을 정도로 작은 크기를 갖는 경우, 이를 전체적으로 유지하고 이를 베이스 레이어에 추가(배정)하는 것이 타당하다. 임의로, 종속적 기본 보조 정보에 대해서와 유사한 분해가 또한 독립적 기본 보조 정보에 대해 행해져, 패킷들(BSI_{I,m}, m = 1, ..., M)을 제공할 수 있다. 이것은 독립적 기본 보조 정보의 파트들을 기본 압축된 사운드 표현의 대응하는 컴포넌트들을 갖는 레이어들에 추가(할당)하는 것에 의해 베이스 레이어의 크기를 감소시키는 데 유용하다.
- [0088] S1040에서, 복수의 향상 보조 정보 부분들(2140-1, ..., 2140-M)이 결정될 수 있다. 각각의 향상 보조 정보 부분은 각자의 레이어 및 각자의 레이어보다 하위인 임의의 레이어들에 포함된 데이터로부터 획득가능한 재구성된 사운드 표현을 개선(예컨대, 향상)시키기 위한 파라미터들을 포함할 수 있다.
- [0089] 이 단계를 수행하는 이유는, 계층화된 코딩의 경우에, 각각의 레이어에 대해 추가로 향상 보조 정보가 계산되어야만 하는 것을 실현하는 것이 중요하기 때문인데, 왜냐하면 예비적인 압축해제된 사운드(또는 음장)를 향상시키는 것 - 그렇지만 이는 압축해제를 위해 이용가능한 레이어들에 의존함 - 이 의도되어 있기 때문이다. 상세하게는, 주어진 최상위 디코딩가능 레이어(최상위 사용가능 레이어)에 대한 예비적인 압축해제된 사운드(또는 음장)는 최상위 디코딩가능 레이어 및 최상위 디코딩가능 레이어 아래의 임의의 레이어들에 포함된 컴포넌트들에 의존한다. 따라서, 압축은, ESI_m, m = 1, ..., M에 의해 표기된, M개의 개별 향상 보조 정보 데이터 패킷(향상 보조 정보 부분)을 제공해야만 하고, 여기서 제m 데이터 패킷(ESI_m) 내의 향상 보조 정보는, 예컨대, 베이스 레이어 및 m보다 더 낮은 인덱스들을 갖는 향상 레이어들에 포함된 데이터 전부(예컨대, 제m 레이어 및 제m 레이어 아래의 임의의 레이어들에 포함된 데이터 전부)로부터 획득된 사운드(또는 음장) 표현을 향상시키도록 계산된다.
- [0090] S1050에서, 복수의 향상 보조 정보 부분들(2140-1, ..., 2140-M)이 복수의 레이어들에 배정(예컨대, 추가, 또는 할당)된다. 복수의 향상 보조 정보 부분들 각각이 복수의 레이어들의 각자의 레이어에 배정된다. 예를 들어, 복수의 레이어들 각각은 각자의 향상 보조 정보 부분을 포함한다.
- [0091] 기본 및/또는 향상 보조 정보를 각자의 레이어들에 배정하는 것은 인코딩 방법에 의해 생성되는 구성 정보에 표시될 수 있다. 환언하면, 기본 및/또는 향상 보조 정보와 각자의 레이어들 사이의 대응관계는 구성 정보에 표시될 수 있다. 게다가, 구성 정보는, 각각의 레이어에 대해, 그 레이어에 배정된(예컨대, 그 레이어에 포함된)

기본 압축된 사운드 표현의 컴포넌트들을 표시할 수 있다. 부가 기본 보조 정보 부분들은 베이스 레이어에 포함되지만, 베이스 레이어와 상이한 레이어들에 대응할 수 있다.

[0092] 요약하면, 압축 스테이지에서, 하기의 조성(composition)을 갖는, FRAME에 의해 표기된, 프레임 데이터 패킷이 제공된다:

수학식 1

$$\text{FRAME} = [\text{BSRC}_1 \dots \text{BSRC}_J \text{ BSI}_1 \text{ BSI}_{D,1} \dots \text{BSI}_{D,M} \text{ ESI}_1 \dots \text{ESI}_M]$$

[0094] 게다가, 패킷들(BSI_1 및 $m = 1, \dots, M$ 에 대한 $\text{BSI}_{D,m}$)이 단일 패킷(BSI)으로 결합될 수 있으며, 이 경우에, FRAME에 의해 표기된, 프레임 데이터 패킷은 하기의 조성을 가질 것이다:

수학식 2

$$\text{FRAME} = [\text{BSRC}_1 \text{ BSRC}_2 \dots \text{BSRC}_J \text{ BSI} \text{ ESI}_1 \text{ ESI}_2 \dots \text{ESI}_M]$$

[0096] 프레임 데이터 패킷을 갖는 개별 페이로드들의 순서화는 일반적으로 임의적일 수 있다.

[0097] 개별 데이터 패킷들은 이어서 페이로드들 내에 그룹화될 수 있고, 이 페이로드들은 유효성 플래그, 그들의 크기를 표시하는 값은 물론 실제의 압축된 표현 데이터를 포함하는 특수 데이터 패킷들로서 정의된다. 페이로드들의 사용은 수신기측에서의 간단한 디멀티플렉스(de-multiplex)를 가능하게 하고, 오래된 페이로드(obsolete payload)들을, 그들을 파싱할 필요 없이, 폐기할 수 있다는 장점을 제공한다. 하나의 가능한 그룹화는 하기에 의해 주어진다:

[0098] - 각각의 BSRC_j 패킷($j = 1, \dots, J$)을 \overline{BP}_j 로 표기된 개별 페이로드에 배정(예컨대, 할당)하는 것.

[0099] - 제 m 향상 보조 정보 데이터 패킷(ESI_m) 및 제 m 종속적 보조 정보 데이터 패킷($\text{BSI}_{D,m}$)을 \overline{EP}_m ($m = 1, \dots, M$)에 의해 표기된 하나의 향상 페이로드에 배정(예컨대, 할당)하는 것.

[0100] - 독립적 기본 보조 정보(BSI_1) 패킷을 \overline{BSIP} 에 의해 표기된 별개의 보조 정보 페이로드에 배정하는 것.

[0101] 임의로, 독립적 기본 보조 정보의 크기가 큰 경우, 그의 컴포넌트들의 각각의 제 m 컴포넌트($\text{BSI}_{1,m}$, $m = 1, \dots, M$)는 향상 페이로드(\overline{EP}_m)에 배정(예컨대, 할당)될 수 있다. 이 경우에, 보조 정보 페이로드(\overline{BSIP})는 비어 있고 무시될 수 있다.

[0102] 다른 옵션은 종속적 기본 보조 정보 데이터 패킷들($\text{BSI}_{D,m}$) 전부를 보조 정보 페이로드(\overline{BSIP})에 배정하는 것이다, 이는 종속적 기본 보조 정보의 크기가 작은 경우에 타당하다.

[0103] 궁극적으로, 하기의 조성을 갖는, FRAME에 의해 표기된, 프레임 데이터 패킷이 제공될 수 있다:

수학식 3

$$\text{FRAME} = [\overline{BP}_1 \dots \overline{BP}_J \text{ } \overline{BSIP} \text{ } \overline{EP}_1 \dots \overline{EP}_M]$$

[0105] 프레임 데이터 패킷을 갖는 개별 페이로드들의 순서화는 일반적으로 임의적일 수 있다.

[0106] 본 방법은 복수의 레이어들 각각에 대해, 각자의 레이어의 데이터(예컨대, 베이스 레이어에 대한 컴포넌트들, 기본 보조 정보 및 향상 보조 정보, 또는 하나 이상의 향상 레이어에 대한 컴포넌트들 및 향상 보조 정보)를 포

함하는 전송 레이어 패킷(예컨대, 베이스 레이어 패킷(2200) 및 M-1개의 향상 레이어 패킷(2300-1, ..., 2300-(M-1))을 생성하는 단계(도 1에 도시되지 않음)를 추가로 포함할 수 있다.

[0107] 상이한 레이어들에 대한 전송 레이어 패킷들은 상이한 전송 우선순위들을 가질 수 있다. 따라서, 본 방법은 복수의 레이어들의 데이터의 전송을 위한 전송 스트림을 생성하는 단계 - 베이스 레이어는 최상위 전송 우선순위를 갖고 계층적 향상 레이어들은 점감하는 전송 우선순위들을 가짐 - (도 1에 도시되지 않음)를 추가로 포함할 수 있다. 거기에서, 보다 높은 전송 우선순위들은 보다 큰 정도의 에러 방지에 대응할 수 있고, 그 반대일 수도 있다.

[0108] 단계들이 특정 다른 단계들을 전제조건들로서 요구하지 않는 한, 앞서 언급된 단계들이 임의의 순서로 수행될 수 있고 도 1에 예시된 예시적인 순서는 비제한적인 것으로 이해된다.

[0109] 도 3은 디코딩 또는 압축해제(언패킹)를 위해 사운드 또는 음장의 압축된 사운드 표현을 디코딩하는 방법을 예시하고 있다. 대응하는 수신기 및 압축해제 스테이지의 예들이 도 4a 및 도 4b의 블록 다이어그램들에 개략적으로 예시되어 있다.

[0110] 이상으로부터 알 수 있는 바와 같이, 압축된 사운드 표현이 복수의 계층적 레이어들에 인코딩될 수 있다. 복수의 레이어들은 기본 압축된 사운드 표현의 컴포넌트들을 배정받았을 수 있고(예컨대, 포함할 수 있고), 컴포넌트들은 각자의 컴포넌트 그룹들 내의 각자의 레이어들에 배정된다. 베이스 레이어는 기본 압축된 사운드 표현을 디코딩하기 위한 기본 보조 정보를 포함할 수 있다. 각각의 레이어는 각자의 레이어 및 각자의 레이어보다 하위인 임의의 레이어들에 포함된 데이터로부터 획득가능한 기본 재구성된 사운드 표현을 개선시키기 위한 파라미터들을 포함하는 앞서 언급된 향상 보조 정보 부분들 중 하나를 포함할 수 있다.

[0111] 제안된 방법은 프레임 기반으로(즉, 프레임 단위로) 수행될 수 있다. 상세하게는, 연속적 시간 구간들, 예를 들어, 동일한 크기의 시간 구간들에 대한 사운드 또는 음장의 복원된 표현이 생성될 수 있다. 시간 구간들은, 예를 들어, 프레임들일 수 있다. 각각의 연속적인 시간 구간들(예컨대, 프레임들)에 대해 이하에 기술되는 단계들이 수행될 수 있다.

[0112] S3010에서, 복수의 레이어들에 대응하는 데이터 페이로드들(예컨대, 전송 레이어 패킷들)이 수신된다. 데이터 페이로드들은 사운드 또는 음장의 압축된 HOA 표현을 포함하는 비트스트림의 일부로서 수신될 수 있으며, 이 표현은 복수의 계층적 레이어들에 대응한다. 계층적 레이어들은 베이스 레이어 및 하나 이상의 계층적 향상 레이어를 포함한다. 복수의 레이어들은 사운드 또는 음장의 기본 압축된 사운드 표현의 컴포넌트들을 배정받았다. 컴포넌트들은 각자의 컴포넌트 그룹들 내의 각자의 레이어들에 배정된다.

[0113] 완전한 압축된 사운드 표현의 수신된 프레임 패킷을 제공하기 위해 개별 레이어 패킷들이 멀티플렉싱될 수 있다. 수신된 프레임 패킷은

수학식 4

$$[BSI_I \ BSI_{D,1} \ \dots \ BSI_{D,M} \ ESI_1 \ BSRC_1 \ \dots \ BSRC_{(J_1)-1} \ \dots \ ESI_M \ BSRC_{J_{(M-1)}} \ \dots \ BSRC_J]$$

[0114] 예 의해 표시될 수 있다.

[0115] 패킷들(BSI_I 및 m = 1, ..., M에 대한 BSI_{D,m})이 단일 패킷(BSI)으로 결합되는 대안의 경우에,

수학식 5

$$[BSI \ ESI_1 \ BSRC_1 \ \dots \ BSRC_{(J_1)-1} \ \dots \ ESI_M \ BSRC_{J_{(M-1)}} \ \dots \ BSRC_J]$$

[0116] 예 의해 표시되는 완전한 압축된 사운드 표현의 수신된 프레임 패킷을 제공하기 위해 개별 레이어 패킷들이 멀티플렉싱될 수 있다.

[0117] 페이로드들의 관점에서, 수신된 프레임 패킷은

수학식 6

$$[0120] \quad \text{FRAME} = [\overline{BP}_1 \quad \dots \quad \overline{BP}_J \quad \overline{BSIP} \quad \overline{EP}_1 \quad \dots \quad \overline{EP}_M]$$

[0121] 예 의해 주어질 수 있다.

[0122] 수신된 프레임 패킷은 이어서 압축해제기 또는 디코더(4100)에 전달될 수 있다. 개별 레이어의 전송이 에러가 없는 경우, (예컨대, 항상 보조 정보 부분에 대응하는) 적어도 포함된 항상 보조 정보 페이로드(\overline{EP}_m) 부분의 유효성 플래그가 "참(true)"으로 설정된다. 개별 레이어의 전송으로 인한 에러의 경우에, 이 레이어에서의 적어도 항상 보조 정보 페이로드 내의 유효성 플래그는 "거짓(false)"으로 설정된다. 따라서, 레이어 패킷의 유효성은 포함된 항상 보조 정보 페이로드의 유효성으로부터(예컨대, 그의 유효성 플래그로부터) 결정될 수 있다.

[0123] 압축해제기(4100)에서, 수신된 프레임 패킷이 디멀티플렉싱될 수 있다. 이를 위해, 개별 페이로드들의 데이터의 불필요한 파싱을 회피하기 위해 각각의 페이로드의 크기에 관한 정보가 이용될 수 있다.

[0124] S3020에서, 기본 압축된 사운드 표현을 사운드 또는 음장의 기본 재구성된 사운드 표현으로 디코딩하기 위해 사용될 복수의 레이어들 중에서 최상위 레이어(예컨대, 최상위 사용가능 레이어, 또는 최상위 디코딩가능 레이어)를 표시하는 제1 레이어 인덱스가 결정된다.

[0125] 더욱이, S3020에서, 기본 사운드 표현의 압축해제를 위해 사용될 최상위 레이어(최상위 사용가능 레이어)의 값(예컨대, 레이어 인덱스)(N_B)이 선택될 수 있다. 기본 사운드 표현의 압축해제를 위해 실제로 사용될 최상위 항상 레이어는 $N_B - 1$ 에 의해 주어진다. 각각의 레이어가 정확히 하나의 항상 보조 정보 페이로드(항상 보조 정보 부분)를 포함하기 때문에, 포함된 레이어(containing layer)가 유효한지(예컨대, 유효하게 수신되었는지) 여부는 항상 보조 정보 페이로드에 기초하여 결정될 수 있다. 따라서, 선택은 항상 보조 정보 페이로드들(ESI_m , $m = 1, \dots, M$ (또는 그에 대응하여, \overline{EP}_m , $m = 1, \dots, M$)) 전부를 사용하여 달성을 수 있다.

[0126] S3030에서, 기본 재구성된 사운드 표현이 획득된다. 기본 재구성된 사운드 표현이, 기본 보조 정보를 사용하여(또는 일반적으로, 기본 보조 정보를 사용하여), 제1 레이어 인덱스에 의해 표시된 최상위 사용가능 레이어 및 이 최상위 사용가능 레이어보다 하위인 임의의 레이어들에 배정된 컴포넌트들로부터 획득될 수 있다.

[0127] 기본 압축된 사운드 표현 컴포넌트들(BSRC₁, ..., BSR_j)의 페이로드들이, 기본 보조 정보 페이로드들(예컨대, BSI 또는 BSI₁ 및 BSI_{D,m}, $m = 1, \dots, M$ (그 전부) 및 값(N_B)과 함께, 기본 표현 압축해제 처리 유닛(4200)에 제공될 수 있다. 기본 표현 압축해제 처리 유닛(4200)(도 4a 및 도 4b에 예시됨)은 최하위 N_B 개의 레이어, 즉 베이스 레이어 및 $N_B - 1$ 개의 항상 레이어(즉, 제1 레이어 인덱스에 의해 표시된 레이어까지의 레이어들) 내에 포함된 그 기본 압축된 사운드 표현 컴포넌트들만을 사용하여 기본 사운드(또는 음장) 표현을 재구성한다. 대안적으로, 최하위 N_B 개의 레이어에 포함된 기본 압축된 사운드 표현 컴포넌트들의 페이로드들만이 각자의 기본 보조 정보 페이로드들과 함께 기본 표현 압축해제 처리 유닛(4200)에 제공될 수 있다.

[0128] 기본 압축된 사운드(또는 음장) 표현의 어느 컴포넌트들이 개별 레이어들에 포함되는지에 관한 요구된 정보는 구성 정보를 갖는 데이터 패킷으로부터 압축해제기(4100)에 알려지는 것으로 가정되며, 이 구성 정보는 프레임 데이터 패킷들 이전에 송신 및 수신되는 것으로 가정된다.

[0129] 종속적 보조 정보 데이터 패킷들(BSI_{D,m}, $m = 1, \dots, N_B$) 및 항상 보조 정보 데이터 패킷(ESI_{N_E})을 제공하기 위해, 항상 페이로드들 전부가 값(N_E) 및 값(N_B)과 함께 압축해제기(4100)의 부분 파서(4400)(도 4b를 참조)에 입력될 수 있다. 파서는 실제 압축해제에 사용되지 않을 페이로드들 및 데이터 패킷들 전부를 폐기할 수 있다. N_E 의 값이 0인 경우, 항상 보조 정보 데이터 패킷들 전부가 비어있는 것으로 가정될 수 있다.

[0130] 베이스 레이어가 각자의 레이어에 대응하는 적어도 하나의 종속적 기본 보조 정보 페이로드(부가 기본 보조 정보 부분)를 포함하는 경우, 각각의 개별 종속적 기본 보조 정보 페이로드(예컨대, BSI_{D,m}, $m = 1, \dots, N_B$ (부가

기본 보조 정보 부분))의 디코딩은 (i) 부가 기본 보조 정보 부분을 그 각자의 레이어 및 각자의 레이어보다 하위인 임의의 레이어들에 배정된 컴포넌트들을 참조하는 것에 의해 디코딩하는 것(예비적 디코딩), 및 (ii) 부가 기본 보조 정보 부분을 최상위 사용가능 레이어 및 최상위 사용가능 레이어와 각자의 레이어 사이의 임의의 레이어들에 배정된 컴포넌트들을 참조하여 정정하는 것(정정)을 포함할 수 있다. 거기에서, 각자의 레이어에 대응하는 부가 기본 보조 정보는 각자의 레이어에 배정된 컴포넌트들 중 하나 이상의 컴포넌트를 각자의 레이어 및 각자의 레이어보다 하위인 임의의 레이어들에 배정된 다른 컴포넌트들에 의존하여 디코딩하는 것을 명시하는 정보를 포함한다.

[0131] 이어서, 기본 재구성된 사운드 표현은, 기본 보조 정보 및 최상위 사용가능 레이어까지의 레이어들에 대응하는 부가 기본 보조 정보 부분들로부터 획득된 정정된 부가 기본 보조 정보 부분들을 사용하여, 최상위 사용가능 레이어 및 최상위 사용가능 레이어보다 하위인 임의의 레이어들에 배정된 컴포넌트들로부터 획득(예컨대, 생성)될 수 있다.

[0132] 상세하게는, 각각의 페이로드($BSI_{D,m}$, $m = 1, \dots, N_B$)의 예비적 디코딩은, 인코딩 스테이지에서 가정된, 처음 m 개의 레이어에 포함된 처음 $J_m - 1$ 개의 기본 압축된 사운드 표현 컴포넌트($BSRC_1, \dots, BSRC_{(J_m)-1}$)에 대한 그의 종속성을 이용하는 것을 포함할 수 있다.

[0133] 각각의 페이로드($BSI_{D,m}$, $m = 1, \dots, N_B$)의 연속적 정정은 기본 사운드 컴포넌트가, 예비적 디코딩에 대해 가정된 것보다 더 많은 컴포넌트들인, 처음 $N_B > m$ 개의 레이어에 포함된 처음 $J_{N_B} - 1$ 개의 기본 압축된 사운드 표현 컴포넌트($BSRC_1, \dots, BSRC_{(J_{N_B})-1}$)로부터 최종적으로 재구성된다는 것을 고려하는 것을 포함할 수 있다. 따라서, 정정은 오래된 정보를 폐기하는 것에 의해 달성될 수 있으며, 이는 특정 상보적 컴포넌트들이 기본 압축된 사운드 표현에 추가되는 경우, 각각의 개별 (상보적) 컴포넌트에 대한 종속적 기본 보조 정보가 원래의 종속적 기본 보조 정보의 서브세트로 된다는 종속적 기본 보조 정보의 초기에 가정된 특성으로 인해 가능하다.

[0134] S3040에서, 제2 레이어 인덱스가 결정될 수 있다. 제2 레이어 인덱스는 기본 재구성된 사운드 표현을 개선(예컨대, 향상)시키기 위해 사용되어야만 하는 향상 보조 정보 부분(들)을 표시할 수 있다.

[0135] 제1 레이어 인덱스에 부가하여, 압축해제를 위해 사용될 향상 보조 정보 페이로드(제2 향상 정보 부분)의 인덱스(제2 레이어 인덱스)(N_E)가 결정될 수 있다. 제2 레이어 인덱스(N_E)는 항상 제1 레이어 인덱스(N_B)와 동일하거나 0일 수 있다. 향상이 항상 최상위 사용가능 레이어로부터 획득된 기본 사운드 표현에 따라 달성될 수 있거나 전혀 달성되지 않을 수 있다.

[0136] S3050에서, 제2 레이어 인덱스를 참조하여, 기본 재구성된 사운드 표현으로부터 사운드 또는 음장의 재구성된 사운드 표현이 획득(예컨대, 생성)된다.

[0137] 즉, 재구성된 사운드 표현은 기본 재구성된 사운드 표현을 (파라미터적으로) 개선 또는 향상시키는 것에 의해, 예컨대, 제2 레이어 인덱스에 의해 표시된 향상 보조 정보(향상 보조 정보 부분)를 사용하는 것에 의해, 획득된다. 이하에서 추가로 살펴보는 바와 같이, 제2 레이어 인덱스는 이 스테이지에서 어떠한 향상 보조 정보도 전혀 사용하지 말라는 것을 표시할 수 있다. 그러면, 재구성된 사운드 표현은 기본 재구성된 사운드 표현에 대응할 것이다.

[0138] 이를 위해, 재구성된 기본 사운드 표현은, 향상 보조 정보 페이로드들(ESI_1, \dots, ESI_M) 전부, 기본 보조 정보 페이로드들(예컨대, BSI 또는 BSI_1 및 $BSI_{D,m}$, $m = 1, \dots, M$), 및 값(N_E)과 함께, 향상된 표현 압축해제 처리 유닛(4300)(도 4a 및 도 4b에 예시됨)에 제공되고, 향상된 표현 압축해제 처리 유닛(4300)은, 향상 보조 정보 페이로드(ESI_{N_E})만을 사용하고 다른 향상 보조 정보 페이로드들 전부를 폐기하는 것에 의해, 최종적인 향상된 사운드(또는 음장) 표현(2100')을 계산한다. 대안적으로, 향상 보조 정보 페이로드들 전부 대신에, 향상 보조 정보 페이로드(ESI_{N_E})만이 향상된 표현 압축해제 처리 유닛(4300)에 제공될 수 있다. N_E 의 값이 0인 경우, 향상 보조 정보 페이로드들 전부가 폐기되고(또는 대안적으로, 어떠한 향상 보조 정보 페이로드도 제공되지 않고) 재구성된 최종적인 향상된 사운드 표현(2100')은 재구성된 기본 사운드 표현과 동일하다. 향상 보조 정보 페이

로드(ESI_{N_E})는 부분 파서(4400)에 의해 획득되었을 수 있다.

[0139] 도 3은 또한 일반적으로 베이스 레이어와 연관되어 있는 기본 보조 정보에 기초하여 그리고 하나 이상의 계층적 향상 레이어와 연관되어 있는 향상 보조 정보에 기초하여 압축된 HOA 표현을 디코딩하는 것을 예시하고 있다.

[0140] 단계들이 특정 다른 단계들을 전제조건들로서 요구하지 않는 한, 앞서 언급된 단계들이 임의의 순서로 수행될 수 있고 도 3에 예시된 예시적인 순서는 비제한적인 것으로 이해된다.

[0141] 다음에, 단계들(S3020 및 S3040)에서의 압축해제를 위한 레이어 선택(제1 및 제2 레이어 인덱스들의 선택)의 상세들이 기술될 것이다.

[0142] 제1 레이어 인덱스를 결정하는 것은, 각각의 레이어에 대해, 각자의 레이어가 유효하게 수신되었는지 여부를 결정하는 것을 포함할 수 있다. 제1 레이어 인덱스를 결정하는 것은 제1 레이어 인덱스를 유효하게 수신되지 않은 최하위 레이어 바로 아래의 레이어의 레이어 인덱스로서 결정하는 것을 추가로 포함할 수 있다. 레이어가 유효하게 수신되었는지 여부는 그 레이어의 향상 보조 정보 페이로드가 유효하게 수신되었는지 여부를 평가하는 것에 의해 결정될 수 있다. 이것은 차례로 향상 보조 정보 페이로드들 내의 유효성 플래그들을 평가하는 것에 의해 행해질 수 있다.

[0143] 제2 레이어 인덱스를 결정하는 것은 일반적으로 제2 레이어 인덱스를 제1 레이어 인덱스와 동일하도록 결정하는 것, 또는 재구성된 사운드 표현을 획득할 때 어떠한 향상 보조 정보도 사용하지 말라는 것을 표시하는 인덱스 값을 제2 레이어 인덱스(예컨대, 인덱스 값 0)로서 결정하는 것 중 어느 하나를 포함할 수 있다.

[0144] 프레임 데이터 패킷들 전부가 서로 독립적으로 압축해제될 수 있는 경우에, 기본 사운드 표현의 압축해제를 위해 실제로 사용될 최상위 레이어(최상위 사용가능 레이어)의 번호(N_B) 및 압축해제를 위해 사용될 향상 보조 정보 페이로드의 인덱스(N_E) 둘 다가 유효한 향상 보조 정보 페이로드의 가장 높은 번호(L)로 설정될 수 있으며, 유효한 향상 보조 정보 페이로드 자체는 향상 보조 정보 페이로드들 내의 유효성 플래그들을 평가하는 것에 의해 결정될 수 있다. 각각의 향상 보조 정보 페이로드의 크기에 대한 지식을 이용하는 것에 의해, 페이로드들의 유효성을 결정하기 위한 페이로드들의 실제 데이터의 복잡한 파싱이 회피될 수 있다.

[0145] 즉, 연속적 시간 구간들에 대한 압축된 사운드 표현들이 독립적으로 디코딩될 수 있는 경우, 제2 레이어 인덱스는 제1 레이어 인덱스와 동일하도록 결정될 수 있다. 이 경우에, 재구성된 기본 사운드 표현은 최상위 사용가능 레이어의 향상 보조 정보 페이로드에 기초하여 향상될 수 있다.

[0146] 프레임간 종속성(inter-frame dependency)들을 갖는 차분 압축해제(differential decompression)가 이용되는 경우에, 이전 프레임으로부터의 결정이 또한 고려되어야만 한다. 차분 압축해제에서, 보통 독립적인 프레임 데이터 패킷들이 규칙적인 시간 구간들로, 이 시각(time instant)들로부터 압축해제를 시작하는 것을 가능하게 하도록, 전송되고, 여기서 값들(N_B 및 N_E)의 결정은 프레임 독립적으로 되고 앞서 기술된 바와 같이 수행된다는 것에 유의한다.

[0147] 제안된 프레임 종속적 결정을 상세히 설명하기 위해, 제k 프레임에 대해 유효한 향상 보조 정보 페이로드의 가장 높은 번호(예컨대, 레이어 인덱스)는 $L(k)$ 로 표기되고, 기본 사운드 표현의 압축해제를 위해 선택 및 사용될 가장 높은 레이어 번호(예컨대, 레이어 인덱스)는 $N_B(k)$ 로 표기되며, 압축해제를 위해 사용될 향상 보조 정보 페이로드의 번호(예컨대, 레이어 인덱스)는 $N_E(k)$ 로 표기된다.

[0148] 이 표기법을 사용하여, $N_B(k)$ 로 되어 있는 기본 사운드 표현의 압축해제를 위해 사용될 가장 높은 레이어 번호는

수학식 7

$$N_B(k) = \min(N_B(k-1), L(k))$$

[0150]에 따라 계산될 수 있다.

[0151] $L(k)$ 및 $N_B(k-1)$ 보다 크지 않은 $N_B(k)$ 를 선택하는 것에 의해, 기본 사운드 표현의 차분 압축해제를 위해 요구된

정보 전부가 이용가능하도록 보장된다.

[0152] 즉, 연속적 시간 구간들(예컨대, 프레임들)에 대한 압축된 사운드 표현들이 서로 독립적으로 디코딩될 수 없는 경우, 제1 레이어 인덱스를 결정하는 것은, 각각의 레이어에 대해, 각자의 레이어가 유효하게 수신되었는지 여부를 결정하는 것, 및 주어진 시간 구간에 대한 제1 레이어 인덱스를 주어진 시간 구간에 선행하는 시간 구간의 제1 레이어 인덱스 및 유효하게 수신되지 않은 최하위 레이어 바로 아래의 레이어의 레이어 인덱스 중 작은 것으로서 결정하는 것을 포함할 수 있다.

[0153] 압축해제를 위해 사용될 향상 보조 정보 페이로드의 번호($N_E(k)$)는

수학식 8

$$N_E(k) = \begin{cases} N_B(k) & \text{if } N_B(k) = N_B(k-1) \\ 0 & \text{else} \end{cases}$$

[0154] [0155]에 따라 결정될 수 있다.

[0156] 거기에서, $N_E(k)$ 에 대해 0을 선택하는 것은 재구성된 기본 사운드 표현이 향상 보조 정보를 사용하여 개선 또는 향상되지 않아야 한다는 것을 표시한다.

[0157] 이것은 상세하게는, 기본 사운드 표현의 압축해제를 위해 사용될 가장 높은 레이어 번호($N_B(k)$)가 변하지 않는 한, 동일한 대응하는 향상 레이어 번호가 선택된다는 것을 의미한다. 그렇지만, $N_B(k)$ 의 변화의 경우에, $N_E(k)$ 를 0으로 설정하는 것에 의해 향상이 디스에이블된다. 향상 보조 정보의 가정된 차분 압축해제로 인해, $N_B(k)$ 에 따른 그의 변화가 가능하지 않은데, 그 이유는 그것이 수행되지 않은 것으로 가정되는 이전 프레임에서의 대응하는 향상 보조 정보 레이어의 압축해제를 요구할 것이기 때문이다.

[0158] 즉, 연속적 시간 구간들(예컨대, 프레임들)에 대한 압축된 사운드 표현들이 서로 독립적으로 디코딩될 수 없는 경우, 제2 레이어 인덱스를 결정하는 것은 주어진 시간 구간에 대한 제1 레이어 인덱스가 선행하는 시간 구간에 대한 제1 레이어 인덱스와 동일한지 여부를 결정하는 것을 포함할 수 있다. 주어진 시간 구간에 대한 제1 레이어 인덱스가 선행하는 시간 구간에 대한 제1 레이어 인덱스와 동일한 경우, 주어진 시간 구간에 대한 제2 레이어 인덱스는 주어진 시간 구간에 대한 제1 레이어 인덱스와 동일하게 결정(예컨대, 선택)될 수 있다. 다른 한편으로, 주어진 시간 구간에 대한 제1 레이어 인덱스가 선행하는 시간 구간에 대한 제1 레이어 인덱스와 동일하지 않은 경우, 재구성된 사운드 표현을 획득할 때 어떠한 향상 보조 정보도 사용하지 말 것을 표시하는 인덱스 값이 제2 레이어 인덱스로서 결정(예컨대, 선택)될 수 있다.

[0159] 대안적으로, 압축해제에서 $N_E(k)$ 까지의 번호들을 갖는 향상 보조 정보 페이로드들 전부가 별별로 압축해제되는 경우, 수학식 4에서의 선택 규칙이

수학식 9

$$N_E(k) = N_B(k)$$

[0160] [0161]로 대체될 수 있다.

[0162] 마지막으로, 차분 압축해제의 경우 최상위 사용된 레이어의 번호(N_B)가 독립적 프레임 데이터 패킷들에서만 증가할 수 있는 반면, 모든 프레임에서 감소가 가능하다는 것에 유의한다.

[0163] 압축된 사운드 표현의 계층화된 인코딩의 제안된 방법이 압축된 사운드 표현의 계층화된 인코딩을 위한 인코더에 의해 구현될 수 있다는 것이 이해된다. 이러한 인코더는 앞서 기술된 각자의 단계들을 수행하도록 적합화된 각자의 유닛들을 포함할 수 있다. 이러한 인코더(5000)의 일 예가 도 5에 개략적으로 예시되어 있다. 예를 들어, 이러한 인코더(5000)는 앞서 언급된 S1010을 수행하도록 적합화된 컴포넌트 세분 유닛(5010), 앞서 언급된 S1020을 수행하도록 적합화된 컴포넌트 배정 유닛(5020), 앞서 언급된 S1030을 수행하도록 적합화된 기본 보조

정보 배정 유닛(5030), 앞서 언급된 S1040을 수행하도록 적합화된 향상 보조 정보 파티셔닝 유닛(5040), 및 앞서 언급된 S1050을 수행하도록 적합화된 향상 보조 정보 배정 유닛(5050)을 포함할 수 있다. 이러한 인코더의 각자의 유닛들이 상기 각자의 유닛들 각각에 의해 수행되는 처리를 수행하도록 적합화된, 즉 앞서 언급된 단계들 중 일부 또는 전부는 물론 제안된 인코딩 방법의 임의의 추가 단계들을 수행하도록 적합화된 컴퓨팅 디바이스의 프로세서(5100)에 의해 구현될 수 있다는 것이 추가로 이해된다. 인코더 또는 컴퓨팅 디바이스는 프로세서(5100)에 의해 액세스 가능한 메모리(5200)를 추가로 포함할 수 있다.

[0164] 복수의 계층적 레이어들에 인코딩되는 압축된 사운드 표현을 디코딩하는 제안된 방법이 복수의 계층적 레이어들에 인코딩되는 압축된 사운드 표현을 디코딩하기 위한 디코더에 의해 구현될 수 있다는 것이 추가로 이해된다. 이러한 디코더는 앞서 기술된 각자의 단계들을 수행하도록 적합화된 각자의 유닛들을 포함할 수 있다. 이러한 디코더(6000)의 일 예가 도 6에 개략적으로 예시되어 있다. 예를 들어, 이러한 디코더(6000)는 앞서 언급된 S3010을 수행하도록 적합화된 수신 유닛(6010), 앞서 언급된 S3020을 수행하도록 적합화된 제1 레이어 인덱스 결정 유닛(6020), 앞서 언급된 S3030을 수행하도록 적합화된 기본 재구성 유닛(6030), 앞서 언급된 S3040을 수행하도록 적합화된 제2 레이어 인덱스 결정 유닛(6040), 및 앞서 언급된 S3050을 수행하도록 적합화된 향상된 재구성 유닛(6050)을 포함할 수 있다. 이러한 디코더의 각자의 유닛들이 상기 각자의 유닛들 각각에 의해 수행되는 처리를 수행하도록 적합화된, 즉 앞서 언급된 단계들 중 일부 또는 전부는 물론 제안된 디코딩 방법의 임의의 추가 단계들을 수행하도록 적합화된 컴퓨팅 디바이스의 프로세서(6100)에 의해 구현될 수 있다는 것이 추가로 이해된다. 디코더 또는 컴퓨팅 디바이스는 프로세서(6100)에 의해 액세스 가능한 메모리(6200)를 추가로 포함할 수 있다.

[0165] 설명 및 도면들이 제안된 방법들 및 장치들의 원리들을 예시하는 것에 불과하다는 것에 유의해야 한다. 따라서 본 기술분야의 통상의 기술자가, 비록 본원에 명시적으로 기술 또는 도시되지는 않았지만, 본 발명의 원리들을 구현하고 그의 사상 및 범주 내에 포함되는 다양한 구성들을 고안할 수 있을 것임을 알게 될 것이다. 게다가, 본원에 열거된 모든 예들은 주로 읽는 사람이 제안된 방법들 및 장치들의 원리들 및 발명자들에 의해 기술을 발전시키는 데 기여된 개념들을 이해하는 데 도움을 주기 위해 명확히 교육적 목적으로만 의도된 것이며, 이러한 특별히 열거된 예들 및 조건들로 제한되지 않는 것으로 해석되어야 한다. 더욱이, 본 발명의 원리들, 양태들, 및 실시예들은 물론 그의 특정 예들을 열거하는 본원에서의 진술들 전부가 그의 등가물들을 포함하도록 의도되어 있다.

[0166] 본 문서에 기술된 방법들 및 장치들은 소프트웨어, 펌웨어 및/또는 하드웨어로서 구현될 수 있다. 특정 컴포넌트들은, 예컨대, 디지털 신호 프로세서 또는 마이크로프로세서 상에서 실행되는 소프트웨어로서 구현될 수 있다. 다른 컴포넌트들은, 예컨대, 하드웨어로서 그리고/또는 ASIC(application specific integrated circuit)들로서 구현될 수 있다. 기술된 방법들 및 장치들에서 나오는 신호들은 랜덤 액세스 메모리 또는 광학 저장 매체와 같은 매체 상에 저장될 수 있다. 이들은 라디오 네트워크(radio network)들, 위성 네트워크들, 무선 네트워크(wireless network)들 또는 유선 네트워크들, 예컨대, 인터넷과 같은, 네트워크들을 통해 전송될 수 있다.

[0167] 참고문헌 1:

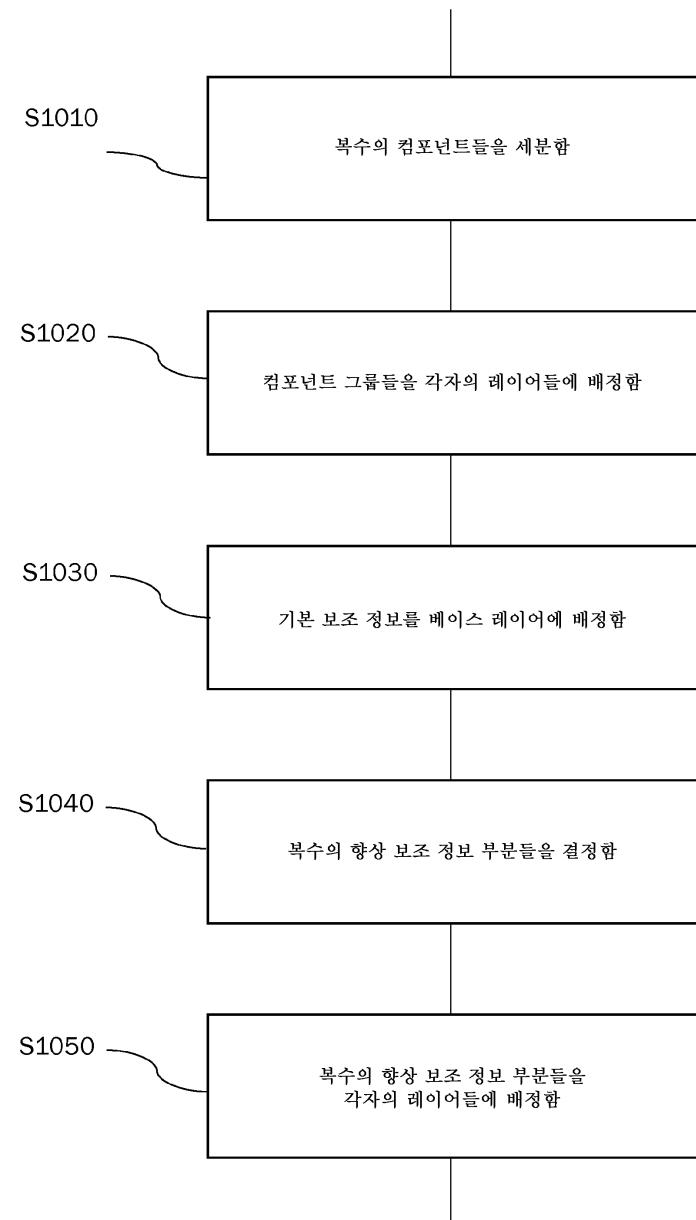
[0168] ISO/IEC JTC1/SC29/WG11 23008-3:2015(E). Information technology - High efficiency coding and media delivery in heterogeneous environments - Part 3: 3D audio, February 2015.

[0169] 참고문헌 2:

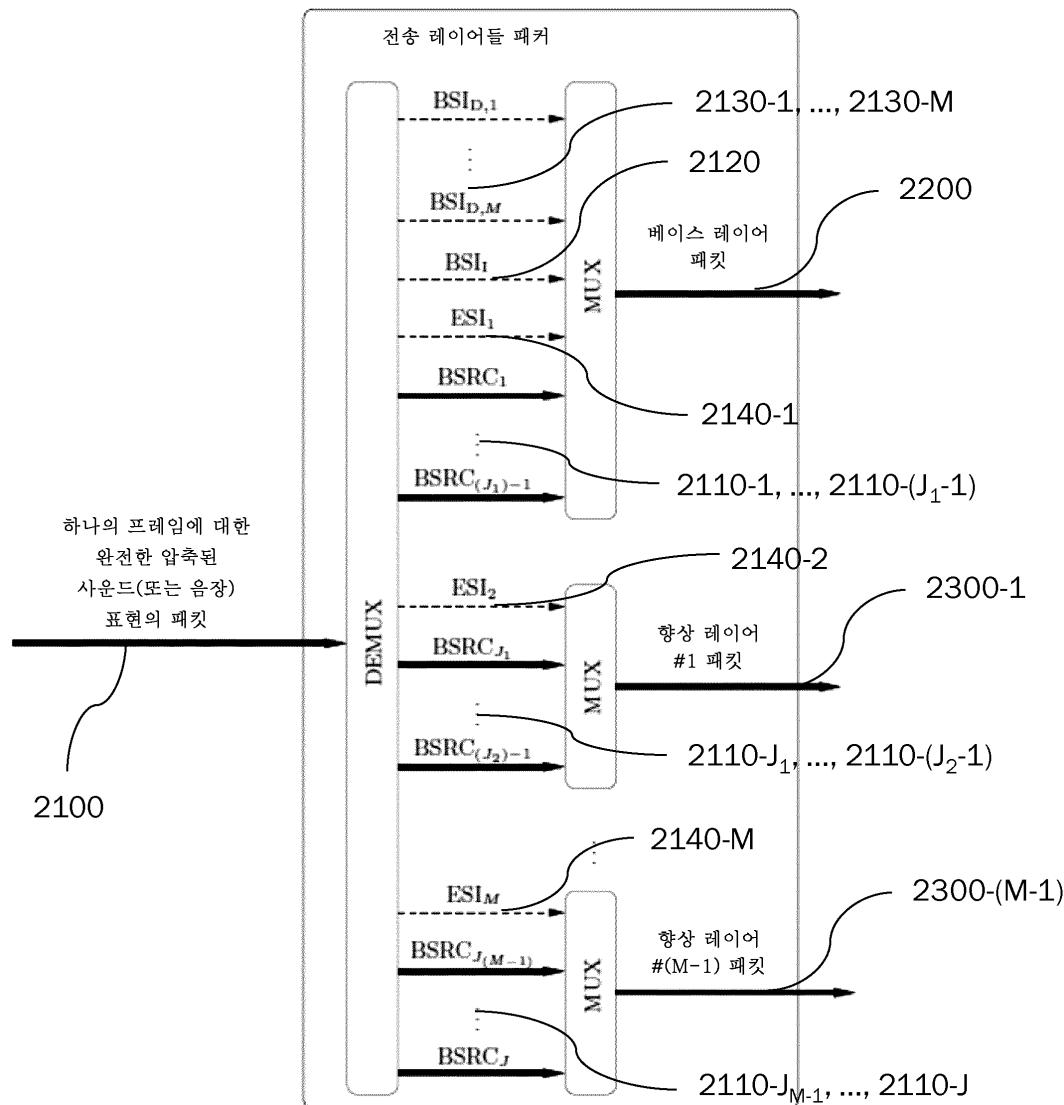
[0170] ISO/IEC JTC1/SC29/WG11 23008-3:2015/PDAM3. Information technology - High efficiency coding and media delivery in heterogeneous environments - Part 3: 3D audio, AMENDMENT 3: MPEG-H 3D Audio Phase 2, July 2015.

도면

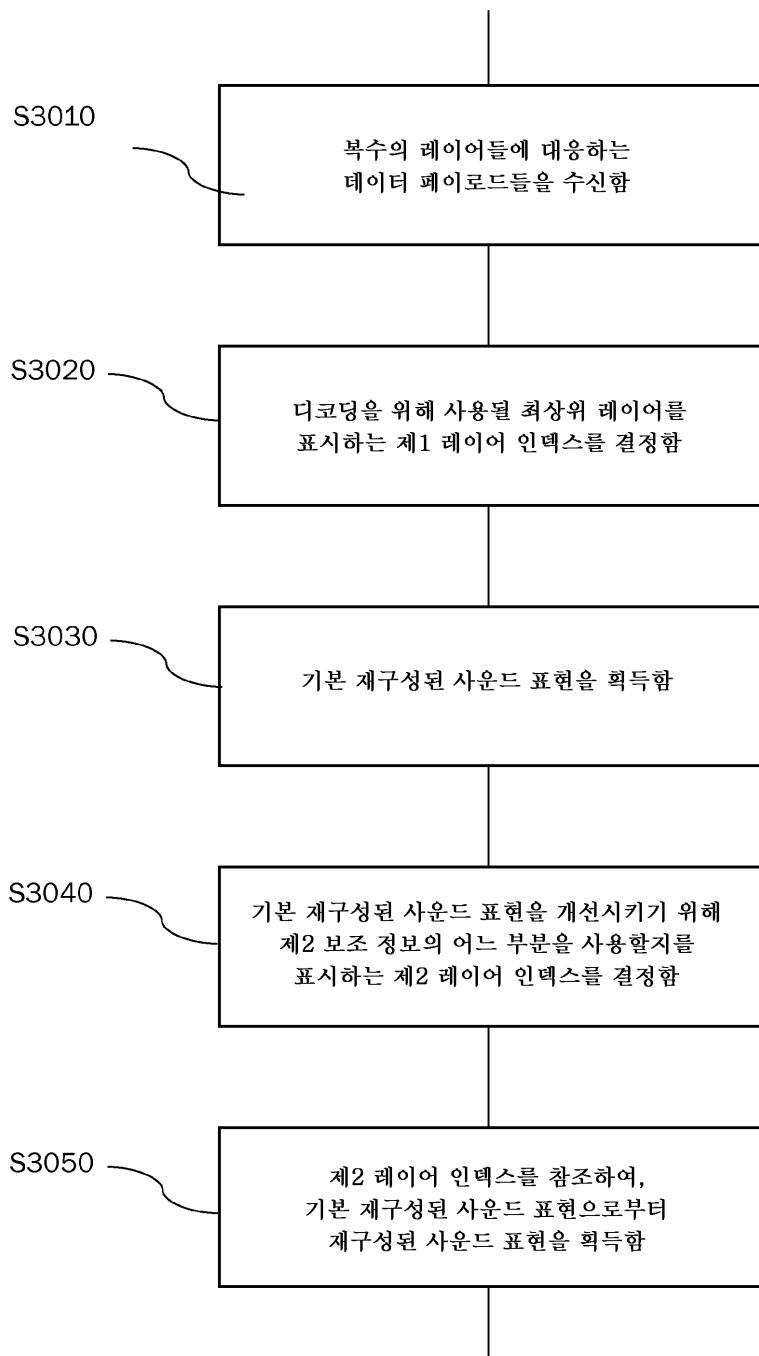
도면1



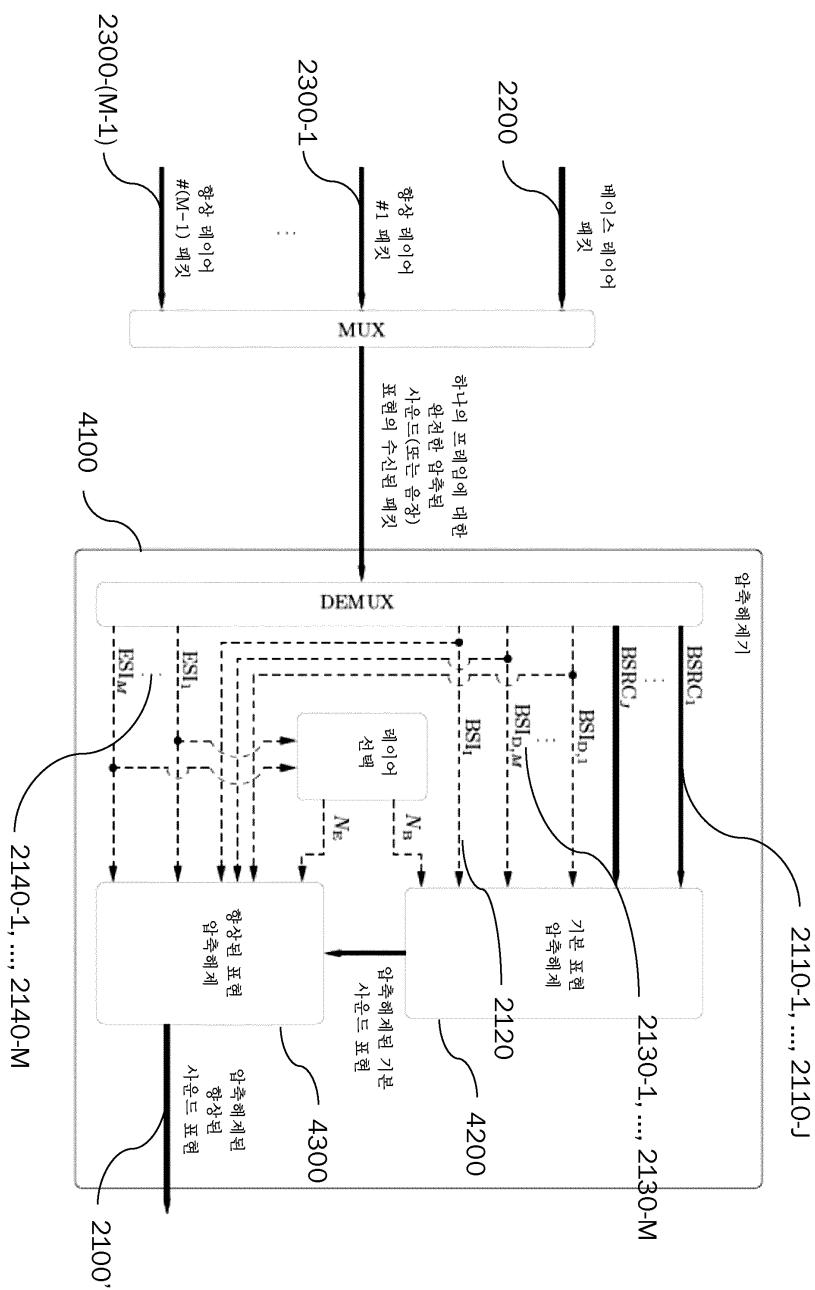
도면2



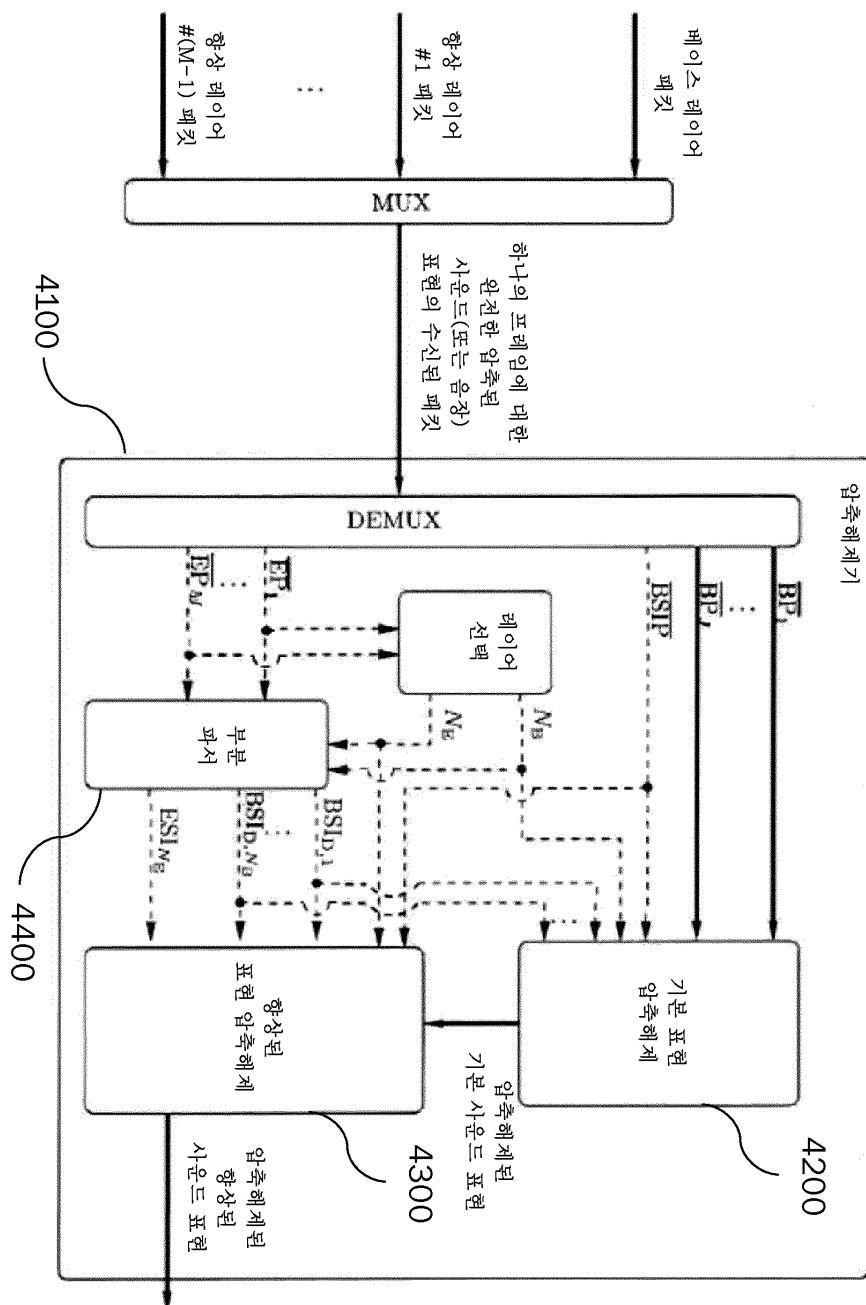
도면3



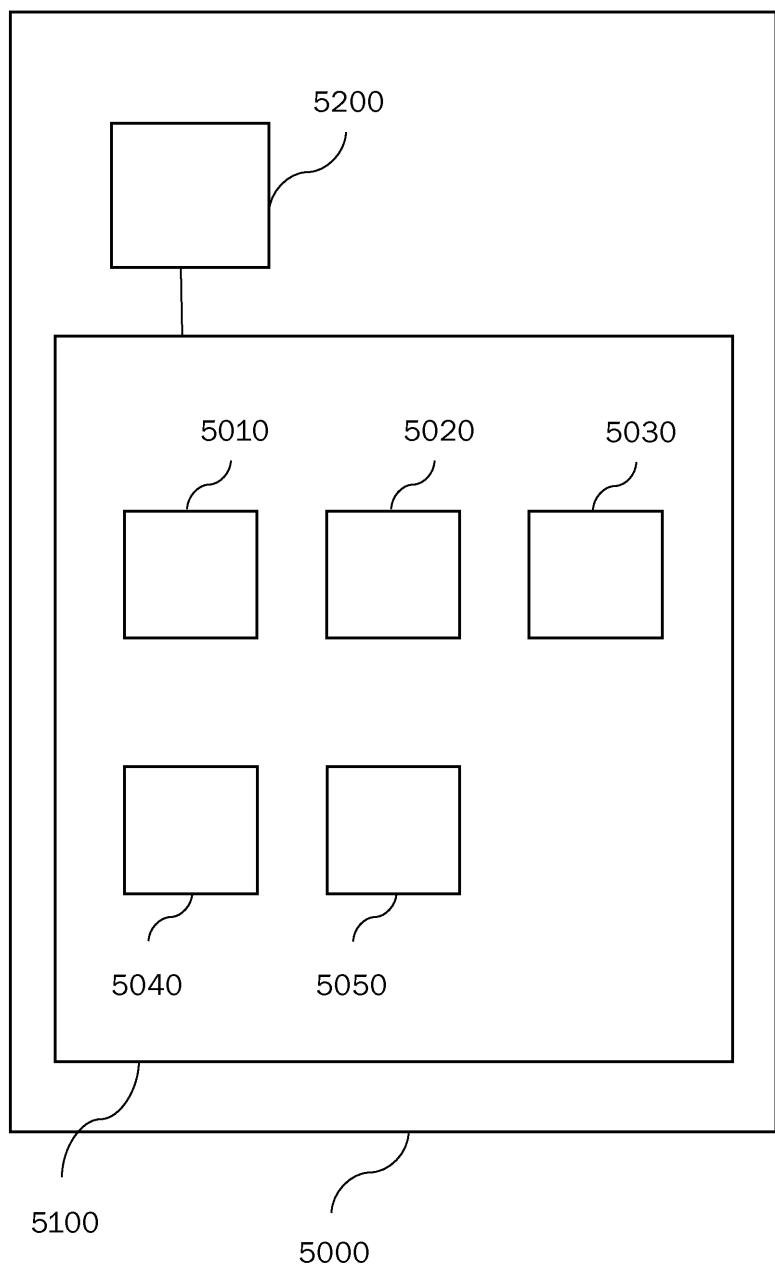
도면4a



도면4b



도면5



도면6

