(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2025-0087591
(43) 공개일자 2025년06월16일

(51) 국제특허분류(Int. Cl.)
G10L 19/008 (2014.01) *G10L 19/16* (2013.01)
H04S 7/00 (2006.01)
(52) CPC특허분류
G10L 19/008 (2020.08)
G10L 19/167 (2013.01)
(21) 출원번호 10-2025-7014302
(22) 출원일자(국제) 2023년09월15일
심사청구일자 없음
(85) 번역문제출일자 2025년04월29일
(86) 국제출원번호 PCT/US2023/074310
(87) 국제공개번호 WO 2024/076828
국제공개일자 2024년04월11일
(30) 우선권주장
63/378,497 2022년10월05일 미국(US)
63/578,534 2023년08월24일 미국(US)

(71) 출원인
돌비 레버러토리즈 라이센싱 코오포레이션
미합중국, 캘리포니아 94103, 샌프란시스코, 마켓 스트리트 1275
돌비 인터네셔널 에이비
아일랜드 디02 브이케이60 더블린 그랜드 커널 독 랜즈 블록 씨 서 존 로저슨스 키 77
(72) 발명자
크조엘링, 크리스토퍼
미국 94103 캘리포니아 샌프란시스코 마켓 스트리트 1275 돌비 레버러토리즈, 인크. 내
푸른하겐, 헤이코
미국 94103 캘리포니아 샌프란시스코 마켓 스트리트 1275 돌비 레버러토리즈, 인크. 내
(뒷면에 계속)
(74) 대리인
양영준, 김연송, 백만기

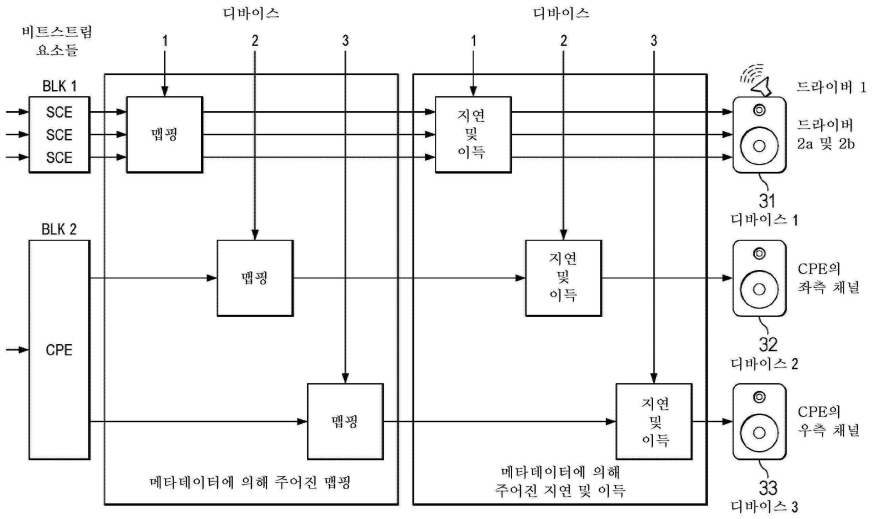
전체 청구항 수 : 총 25 항

(54) 발명의 명칭 **파라메트릭의 유연한 렌더링 구성 데이터를 이용한 오디오 비트스트림들의 인코딩 및 디코딩을 위한 방법, 장치, 및 매체**

(57) 요 약

복수의 오디오 신호들을 포함하는 오디오 프로그램으로부터 인코딩된 비트스트림을 생성하기 위한 방법으로서, 복수의 오디오 신호들 각각에 대해, 각각의 오디오 신호가 연관되는 재생 디바이스를 표시하는 정보를 수신하는 단계, 각각의 재생 디바이스에 대해, 각각의 재생 디바이스와 연관된 지연, 이득, 및 등화 곡선 중 적어도 하나를 표시하는 정보를 수신하는 단계, 복수의 오디오 신호들로부터, 2개 이상의 관련 오디오 신호들의 그룹을 결정하는 단계, 공동으로 코딩된 오디오 신호들을 획득하기 위해 그룹의 2개 이상의 관련 오디오 신호들에 하나 이상의 공동 코딩 툴을 적용하는 단계, 및 공동으로 코딩된 오디오 신호들, 공동으로 코딩된 오디오 신호들이 연관되는 재생 디바이스들의 표시, 및 공동으로 코딩된 오디오 신호들이 연관되는 각각의 재생 디바이스들과 연관된 지연 및 이득의 표시들을 인코딩된 비트스트림의 독립 블록으로 결합하는 단계를 포함하는, 방법.

대 표 도



(52) CPC특허분류
H04S 7/302 (2013.01)
(72) 발명자
세펠드트, 알란 제이.
미국 94103 캘리포니아 샌프란시스코 마켓 스트리트 1275 돌비 레버러토리즈, 인크. 내
란도, 조슈아 비.
미국 94103 캘리포니아 샌프란시스코 마켓 스트리트 1275 돌비 레버러토리즈, 인크. 내

명 세 서

청구범위

청구항 1

복수의 오디오 신호들을 포함하는 오디오 프로그램으로부터 인코딩된 비트스트림을 생성하기 위한 방법으로서,

상기 복수의 오디오 신호들 각각에 대해, 각각의 오디오 신호가 연관되는 재생 디바이스를 표시하는 정보를 수신하는 단계;

각각의 재생 디바이스에 대해, 상기 각각의 재생 디바이스와 연관된 지연, 이득, 및 등화 곡선 중 적어도 하나를 표시하는 정보를 수신하는 단계;

상기 복수의 오디오 신호들로부터, 2개 이상의 관련 오디오 신호들의 그룹을 결정하는 단계;

공동으로 코딩된 오디오 신호들을 획득하기 위해 상기 그룹의 상기 2개 이상의 관련 오디오 신호들에 하나 이상의 공동 코딩 툴을 적용하는 단계; 및

상기 공동으로 코딩된 오디오 신호들, 상기 공동으로 코딩된 오디오 신호들이 연관되는 상기 재생 디바이스들의 표시, 및 상기 공동으로 코딩된 오디오 신호들이 연관되는 상기 각각의 재생 디바이스들과 연관된 지연, 이득, 및 등화 곡선 중 적어도 하나를 표시하는 정보를 인코딩된 비트스트림의 독립 블록으로 결합하는 단계를 포함하는, 방법.

청구항 2

제1항에 있어서,

상기 정보는 상기 지연 및 이득을 표시하는, 방법.

청구항 3

제1항에 있어서,

상기 각각의 재생 디바이스와 연관된 상기 지연, 이득 및/또는 등화 곡선은 청취자의 위치에 대한 상기 각각의 재생 디바이스의 위치에 의존하는, 방법.

청구항 4

제1항 내지 제3항 중 어느 한 항에 있어서,

상기 각각의 재생 디바이스와 연관된 상기 지연, 이득 및/또는 등화 곡선은 다른 재생 디바이스들의 위치들에 대한 상기 각각의 재생 디바이스의 위치에 의존하는, 방법.

청구항 5

제1항 내지 제4항 중 어느 한 항에 있어서,

상기 지연, 이득, 및/또는 등화 곡선은 동적으로 가변적인, 방법.

청구항 6

제5항에 있어서,

상기 지연, 이득 및/또는 등화 곡선은 상기 청취자의 위치의 변화에 응답하여 조정되는, 방법.

청구항 7

제5항 또는 제6항에 있어서,

상기 지연, 이득 및/또는 등화 곡선은 상기 재생 디바이스의 위치의 변화에 응답하여 조정되는, 방법.

청구항 8

제5항 내지 제7항 중 어느 한 항에 있어서,

상기 지연, 이득 및/또는 등화 곡선은 상기 다른 재생 디바이스들 중 하나 이상의 위치의 변화에 응답하여 조정되는, 방법.

청구항 9

제1항 내지 제8항 중 어느 한 항에 있어서,

상기 복수의 오디오 신호들로부터, 상기 2개 이상의 관련 오디오 신호들의 그룹의 일부가 아닌 오디오 신호를 결정하는 단계를 더 포함하는, 방법.

청구항 10

제9항에 있어서,

상기 2개 이상의 관련 오디오 신호들의 그룹의 일부가 아닌 상기 오디오 신호에 대해, 상기 오디오 신호가 연관되는 상기 재생 디바이스와 연관된 상기 지연, 이득, 및/또는 등화 곡선을 적용하는 단계를 더 포함하는, 방법.

청구항 11

제10항에 있어서,

상기 2개 이상의 관련 오디오 신호들의 그룹의 일부가 아닌 상기 오디오 신호를 독립적으로 코딩하는 단계, 및 독립적으로 코딩된 오디오 신호 및 상기 독립적으로 코딩된 오디오 신호가 연관되는 상기 재생 디바이스의 표시를 상기 인코딩된 비트스트림의 별개의 독립적으로 디코딩가능한 서브세트로 결합하는 단계를 더 포함하는, 방법.

청구항 12

제9항에 있어서,

상기 2개 이상의 관련 오디오 신호들의 그룹의 일부가 아닌 상기 오디오 신호를 독립적으로 코딩하는 단계, 및 독립적으로 코딩된 오디오 신호, 상기 독립적으로 코딩된 오디오 신호가 연관되는 상기 재생 디바이스의 표시, 및 상기 독립적으로 코딩된 오디오 신호가 연관되는 상기 재생 디바이스와 연관된 상기 지연, 이득, 및/또는 등화 곡선의 표시를 상기 인코딩된 비트스트림의 별개의 독립적으로 디코딩가능한 서브세트로 결합하는 단계를 더 포함하는, 방법.

청구항 13

인코딩된 비트스트림의 프레임으로부터 재생 디바이스와 연관된 하나 이상의 오디오 신호를 디코딩하기 위한 방법 - 상기 프레임은 인코딩된 데이터의 하나 이상의 독립 블록을 포함함 - 으로서,

상기 인코딩된 비트스트림으로부터, 상기 재생 디바이스와 연관된 상기 하나 이상의 오디오 신호에 대응하는 인코딩된 데이터의 독립 블록을 식별하는 단계;

상기 인코딩된 비트스트림으로부터, 식별된 인코딩된 데이터의 독립 블록을 추출하는 단계;

추출된 인코딩된 데이터의 독립 블록이 2개 이상의 공동으로 코딩된 오디오 신호들을 포함한다고 결정하는 단계;

상기 재생 디바이스와 연관된 상기 하나 이상의 오디오 신호를 획득하기 위해 상기 2개 이상의 공동으로 코딩된 오디오 신호들에 하나 이상의 공동 디코딩 툴을 적용하는 단계;

상기 추출된 인코딩된 데이터의 독립 블록으로부터, 상기 재생 디바이스와 연관된 지연, 이득 및 등화 곡선 중 적어도 하나를 결정하는 단계; 및

상기 재생 디바이스와 연관된 상기 지연, 이득 및/또는 등화 곡선을 상기 재생 디바이스와 연관된 상기 하나 이상의 오디오 신호에 적용하는 단계를 포함하는, 방법.

청구항 14

제13항에 있어서,

상기 재생 디바이스와 연관된 상기 결정된 지연, 이득 및/또는 등화 곡선은 청취자의 위치에 대한 상기 재생 디바이스의 위치에 의존하는, 방법.

청구항 15

제13항 또는 제14항에 있어서,

상기 재생 디바이스와 연관된 상기 결정된 지연, 이득 및/또는 등화 곡선은 다른 재생 디바이스들에 대한 상기 재생 디바이스의 위치에 의존하는, 방법.

청구항 16

제13항 내지 제15항 중 어느 한 항에 있어서,

상기 재생 디바이스의 상기 결정된 지연, 이득, 및/또는 등화 곡선은 동적으로 가변적인, 방법.

청구항 17

제16항에 있어서,

상기 재생 디바이스와 연관된 상기 결정된 지연, 이득, 및/또는 등화 곡선이 상기 재생 디바이스와 연관된 이전에 결정된 지연, 이득, 및/또는 등화 곡선과 상이할 때, 상기 방법은 상기 재생 디바이스와 연관된 상기 이전에 결정된 지연, 이득, 및/또는 등화 곡선과 상기 재생 디바이스와 연관된 상기 결정된 지연, 이득, 및/또는 등화 곡선 사이를 보간하는 단계를 더 포함하는, 방법.

청구항 18

제17항에 있어서,

상기 결정된 지연, 이득 및/또는 등화 곡선은 청취자의 위치의 변화로 인해 상기 이전에 결정된 지연, 이득 및/또는 등화 곡선과 상이한, 방법.

청구항 19

제17항 또는 제18항에 있어서,

상기 결정된 지연, 이득 및/또는 등화 곡선은 상기 재생 디바이스의 위치의 변화로 인해 상기 이전에 결정된 지연, 이득 및/또는 등화 곡선과 상이한, 방법.

청구항 20

제17항 내지 제19항 중 어느 한 항에 있어서,

상기 결정된 지연, 이득 및/또는 등화 곡선은 상기 다른 재생 디바이스들 중 하나 이상의 위치의 변화로 인해 상기 이전에 결정된 지연, 이득 또는 등화 곡선과 상이한, 방법.

청구항 21

제13항 내지 제20항 중 어느 한 항에 있어서,

상기 인코딩된 비트스트림의 프레임은 인코딩된 데이터의 2개 이상의 독립 블록들을 포함하고, 상기 방법은,

상기 독립 블록들 중 하나 이상이 상기 재생 디바이스와 연관되지 않은 오디오 신호들을 포함한다고 결정하는 단계; 및

상기 재생 디바이스와 연관되지 않은 오디오 신호들을 포함하는 상기 하나 이상의 독립 블록을 무시하는 단계를 더 포함하는, 방법.

청구항 22

제13항 내지 제21항 중 어느 한 항에 있어서,

하나 이상의 공동 디코딩 툴을 적용하는 단계는 상기 재생 디바이스와 연관되는 상기 공동으로 코딩된 오디오 신호들의 서브세트를 식별하는 단계, 및 상기 공동으로 코딩된 오디오 신호들의 해당 서브세트만을 재구성하여 상기 재생 디바이스와 연관된 상기 하나 이상의 오디오 신호를 획득하는 단계를 포함하는, 방법.

청구항 23

제13항 내지 제21항 중 어느 한 항에 있어서,

하나 이상의 공동 디코딩 툴을 적용하는 단계는 상기 공동으로 코딩된 오디오 신호들 각각을 재구성하는 단계, 상기 재생 디바이스와 연관된 재구성된 공동으로 코딩된 오디오 신호들의 서브세트를 식별하는 단계, 및 상기 재생 디바이스와 연관된 상기 재구성된 공동으로 코딩된 오디오 신호들의 서브세트로부터 상기 재생 디바이스와 연관된 상기 하나 이상의 오디오 신호를 획득하는 단계를 포함하는, 방법.

청구항 24

장치로서,

제1항 내지 제23항 중 어느 한 항의 방법을 수행하도록 구성된, 장치.

청구항 25

비일시적 컴퓨터 판독가능 저장 매체로서,

실행될 때, 하나 이상의 디바이스로 하여금 제1항 내지 제23항 중 어느 한 항의 방법을 수행하게 하는 명령어들의 시퀀스를 포함하는, 비일시적 컴퓨터 판독가능 저장 매체.

발명의 설명

기 술 분 야

[0001] **관련 출원들에 대한 상호 참조**

[0002] 본 출원은 2022년 10월 5일자로 출원된 미국 가출원 일련 번호 제63/378,497호 및 2023년 8월 24일자로 출원된 미국 가출원 일련 번호 제63/578,534호로부터의 우선권의 이익을 주장하며, 이들 둘 다는 그 전체가 본 명세서에 참조로 포함된다.

[0003] **기술분야**

[0004] 본 개시내용은 일반적으로 오디오 신호 처리에 관한 것으로서, 보다 구체적으로는, 디바이스들 사이의 몰입형 오디오 프로그램들(immersive audio programs)의 오디오 신호들의 낮은 레이턴시 교환(low latency interchange)을 위한 오디오 소스 코딩 및 디코딩에 관한 것이다.

배 경 기 술

[0005] 오디오의 스트리밍은 오늘날의 사회에서 흔한 것이다. 오디오 스트리밍은 품질에 대한 사용자의 기대가 높아지면서 점점 더 까다로워지고 있지만, 또한 사용자의 설정도 스피커들의 수뿐만 아니라 스피커들의 타입들에 대해 더 복잡해지고 있다. 스트리밍은 통상적으로 적어도 무선 링크를 통해 일부 부분들에서 행해지고, 이는 이어서 양호한 품질을 갖기 위해 무선 링크에 대한 소정의 요건을 부여하며, 아마도 많은 경험이 있을 것 같지만, 항상 그러한 것은 아니다.

[0006] 따라서, 특정 포맷이 클라우드/서버로부터 그리고 후속하여 무선(또는, 일부 경우들에서는 유선) 링크들을 통한 분배를 위해 보다 적합한 낮은 레이턴시 포맷으로 트랜스코딩된 디바이스 상에서 스트리밍되는 이용 사례들에 대한 교환 포맷을 정의할 필요가 있다. 예시적인 이용 사례들은 가정 내 접속성(in-home connectivity)은 물론, 전화-자동차 접속성(phone-to-automotive connectivity)이지만, 포맷은 단일 디바이스로부터 하나 이상의 접속된 디바이스로의 오디오 신호들의 낮은 레이턴시 분배가 요망되는 임의의 시나리오에서 유익할 수 있다.

[0007] 무선으로 전송되고 스트리밍되는 오디오 정보에 더하여, 스트림 내에 통합되는 다른 타입의 정보도 존재할 수 있다. 그러한 다른 타입의 정보는 또한 무선 링크의 품질에 의해 영향을 받을 것이고, 오디오에 대한 것과 유

사한 단점들을 가질 수 있다.

[0008] 따라서, 다른 타입의 정보 또는 신호들과 결합된 상이한 타입들의 스트리밍된 오디오에 대한 무선 스트리밍과 연관된 문제점들을 극복하는 것이 유리할 것이다.

발명의 내용

[0009] 본 개시내용의 목적은 오디오 및 다른 타입들의 정보의 무선 스트리밍으로 상기의 문제점을 적어도 부분적으로 극복하는 것이다.

[0010] 본 개시내용의 제1 양태에 따르면, 복수의 오디오 신호들을 포함하는 오디오 프로그램으로부터 인코딩된 비트스트림을 생성하기 위한 방법이 제공되고, 방법은 복수의 오디오 신호들 각각에 대해, 각각의 오디오 신호가 연관되는 재생 디바이스를 표시하는 정보를 수신하는 단계, 각각의 재생 디바이스에 대해, 각각의 재생 디바이스와 연관된 지연, 이득, 및 등화 곡선 중 적어도 하나를 표시하는 정보를 수신하는 단계, 복수의 오디오 신호들로부터, 2개 이상의 관련 오디오 신호들의 그룹을 결정하는 단계, 공동으로 코딩된 오디오 신호들을 획득하기 위해 그룹의 2개 이상의 관련 오디오 신호들에 하나 이상의 공동 코딩 툴을 적용하는 단계, 및 공동으로 코딩된 오디오 신호들, 공동으로 코딩된 오디오 신호들이 연관되는 재생 디바이스들의 표시, 및 공동으로 코딩된 오디오 신호들이 연관되는 각각의 재생 디바이스들과 연관된 지연, 이득, 및 등화 곡선 중 적어도 하나를 표시하는 정보를 인코딩된 비트스트림의 독립 블록으로 결합하는 단계를 포함한다.

[0011] 본 개시내용의 제2 양태에 따르면, 인코딩된 비트스트림의 프레임으로부터 재생 디바이스와 연관된 하나 이상의 오디오 신호를 디코딩하기 위한 방법이 제공되고, 프레임은 인코딩된 데이터의 하나 이상의 독립 블록을 포함하고, 방법은 인코딩된 비트스트림으로부터, 재생 디바이스와 연관된 하나 이상의 오디오 신호에 대응하는 인코딩된 데이터의 독립 블록을 식별하는 단계, 인코딩된 비트스트림으로부터, 식별된 인코딩된 데이터의 독립 블록을 추출하는 단계, 추출된 인코딩된 데이터의 독립 블록이 2개 이상의 공동으로 코딩된 오디오 신호들을 포함한다고 결정하는 단계, 재생 디바이스와 연관된 하나 이상의 오디오 신호를 획득하기 위해 2개 이상의 공동으로 코딩된 오디오 신호들에 하나 이상의 공동 디코딩 툴을 적용하는 단계, 추출된 인코딩된 데이터의 독립 블록으로부터, 재생 디바이스와 연관된 지연, 이득 및 등화 곡선 중 적어도 하나를 결정하는 단계, 재생 디바이스와 연관된 지연, 이득 및/또는 등화 곡선을 재생 디바이스와 연관된 하나 이상의 오디오 신호에 적용하는 단계를 포함한다.

[0012] 본 개시내용의 제3 양태에 따르면, 제1 및/또는 제2 양태에 따른 방법을 수행하도록 구성된 장치가 제공된다.

[0013] 본 개시내용의 제4 양태에 따르면, 실행될 때, 하나 이상의 디바이스로 하여금 제1 및/또는 제2 양태에 따른 방법 중 어느 하나의 방법을 수행하게 하는 명령어들의 시퀀스를 포함하는 비일시적 컴퓨터 판독가능 저장 매체가 제공된다.

[0014] 본 개시내용의 추가의 예들은 종속 청구항들에서 정의된다.

[0015] 일부 예들에서, 각각의 재생 디바이스와 연관된 지연, 이득 및/또는 등화 곡선은 청취자의 위치에 대한 각각의 재생 디바이스의 위치 및/또는 다른 재생 디바이스들의 위치들에 대한 각각의 재생 디바이스에 의존한다. 일부 예들에서, 지연, 이득, 및/또는 등화 곡선은 동적으로 가변적이다.

[0016] 일부 예들에서, 지연, 이득 및/또는 등화 곡선은 청취자의 위치의 변화, 재생 디바이스의 위치의 변화 및/또는 다른 재생 디바이스들 중 하나 이상의 위치의 변화에 응답하여 조정된다.

[0017] 일부 예들에서, 복수의 오디오 신호들로부터, 2개 이상의 관련 오디오 신호들의 그룹의 일부가 아닌 오디오 신호를 결정하는 단계를 포함하는 방법이 제공된다.

[0018] 일부 예들에서, 2개 이상의 관련 오디오 신호들의 그룹의 일부가 아닌 오디오 신호에 대해, 오디오 신호가 연관되는 재생 디바이스와 연관된 지연, 이득, 및/또는 등화 곡선을 적용하는 단계를 포함하는 방법이 제공된다.

[0019] 일부 예들에서, 2개 이상의 관련 오디오 신호들의 그룹의 일부가 아닌 오디오 신호를 독립적으로 코딩하는 단계, 및 독립적으로 코딩된 오디오 신호 및 독립적으로 코딩된 오디오 신호가 연관되는 재생 디바이스의 표시를 인코딩된 비트스트림의 별개의 독립적으로 디코딩가능한 서브세트로 결합하는 단계를 포함하는 방법이 제공된다.

[0020] 본 개시내용에서, 프레임은 모든 신호들 전체의 시간 슬라이스를 나타낸다. 블록 스트림은 세션의 지속기간 동안의 신호들의 집합을 나타낸다. 블록은 블록 스트림의 하나의 프레임을 나타낸다. 주어진 샘플링 주파수를

갖는 디지털 오디오의 경우, 프레임 크기는 임의의 오디오 신호에 대한 프레임에서의 오디오 샘플들의 수와 등가이다. 프레임 크기는 일반적으로 세션의 지속기간 동안 일정하게 유지된다.

[0021] 웨이크 워드(wake word)는 하나의 단어, 또는 2개 이상의 단어를 고정된 순서로 포함하는 어구를 포함할 수 있다.

[0022] 청구항들에서 포함하는 본 개시내용 전반에 걸쳐, 표현 "시스템"은 디바이스, 시스템, 또는 서브시스템을 나타내기 위하여 넓은 의미로 이용된다. 예를 들어, 디코더를 구현하는 서브시스템은 디코더 시스템이라고 지칭될 수 있고, 그러한 서브시스템을 포함하는 시스템(예를 들어, 서브시스템이 M개의 입력들을 생성하고 다른 X-M개의 입력들이 외부 소스로부터 수신되는, 다수의 입력들에 응답하여 X개의 출력 신호들을 생성하는 시스템)이 또한 디코더 시스템이라고 지칭될 수 있다.

도면의 간단한 설명

[0023] 본 개시내용의 예들은 첨부 도면들을 참조하여 더 상세히 설명될 것이다. 이하의 도면들에서, 유사한 참조 번호들은 유사한 요소들을 지칭하는 데 이용된다. 이하의 도면들은 다양한 예들을 도시하지만, 하나 이상의 구현은 도면들에 도시된 예들로 제한되지 않는다.

도 1은 가정 내 접속성 낮은 레이턴시 트랜스코딩의 예를 도시한다.

도 2는 자동차 접속성 오디오 스트리밍의 예를 도시한다.

도 3은 무선 디바이스들의 이용에 의한 증강 TV 오디오 스트리밍의 예를 도시한다.

도 4는 간단한 무선 스피커를 이용한 오디오 스트리밍의 예를 도시한다.

도 5는 비트스트림 요소들로부터 디바이스로의 메타데이터 맵핑의 예를 도시한다.

도 6은 오디오 스트리밍의 간단한 유연한 렌더링을 전개하는 방법의 예를 도시한다.

도 7은 유연한 렌더링 데이터를 비트스트림 요소들에 맵핑하는 예를 도시한다.

도 8은 오디오를 재생하고 다수의 디바이스들로 커맨드들을 청취할 때 에코-참조들(echo-references)의 시그널링의 예를 도시한다.

도 9는 프레임들, 블록들, 및 패킷들이 어떻게 서로 관련되는지의 예를 도시한다.

도 10은 현재의 MPEG-4 구조에서의 추가의 코덱 정보의 예를 도시한다.

도 11은 현재의 MPEG-4 구조에서의 더 추가의 코덱 정보의 예를 도시한다.

도 12는 청취 능력과 음성 인식을 통합하는 예를 도시한다.

도 13은 다수의 블록들을 포함하는 프레임의 예를 도시한다.

도 14는 다수의 블록들을 포함하는 비트스트림의 예를 도시한다.

도 15는 상이한 우선순위를 갖는 블록들의 예를 도시한다.

도 16은 상이한 우선순위를 갖는 다수의 블록들을 포함하는 비트스트림의 예를 도시한다.

도 17은 상이한 우선순위들을 갖는 프레임의 예를 도시한다.

도 18은 상이한 우선순위들을 갖는 프레임들을 포함하는 비트스트림의 예를 도시한다.

발명을 실시하기 위한 구체적인 내용

[0024] 이제, 본 발명의 원리들이 도면들에 도시된 다양한 예들을 참조하여 설명된다. 이러한 예들의 묘사는 단지 본 기술분야의 통상의 기술자가 본 발명을 더 잘 이해하고 추가로 구현할 수 있게 하기 위한 것이며, 본 발명의 범위를 어떤 식으로든 제한하려는 것은 아님을 이해해야 한다.

[0025] 도 1에서, 몰입형 오디오 스트림은 클라우드 또는 서버(10)로부터 스트리밍되고, TV 또는 허브 디바이스(20) 상에서 디코딩된다. 몰입형 오디오 스트림은, 예를 들어, Dolby Digital Plus, AC-4 등을 포함하는 임의의 기존 포맷으로 코딩될 수 있다. 출력은, 바람직하게는 로컬 무선 접속, 예를 들어, WiFi 소프트 액세스 포인트 또는 블루투스 접속을 통해 접속되는 접속된 디바이스들(30)로의 추가 송신을 위해 낮은 레이턴시 교환 포맷으로 후

속적으로 트랜스코딩된다. 낮은 레이턴시는 통상적으로, 예를 들어, 프레임 크기, 샘플링 레이트, 하드웨어 및/또는 소프트웨어 계산 자원들 등과 같은 다양한 인자들에 의존하지만, 낮은 레이턴시는 통상적으로 40ms, 20ms 또는 10ms 미만일 것이다.

[0026] 다른 전형적인 이용 사례에서, 전화는 클라우드 또는 서버로부터 몰입형 오디오 스트림을 페치하고, 교환 포맷으로 트랜스코딩하고, 후속하여 접속된 자동차에 송신한다. 도 2의 예시된 예에서, 모바일 디바이스(예를 들어, 전화 또는 태블릿)(20)는 서버(10)에 접속되어 몰입형 오디오 스트림을 수신하고, 낮은 레이턴시 트랜스코드를 낮은 레이턴시 교환 포맷으로 수행하고, 트랜스코딩된 신호를 몰입형 오디오 재생을 지원하는 자동차(30)에 송신한다. 몰입형 오디오 스트림의 예는 돌비 아트모스 포맷(Dolby Atmos format)의 오디오를 포함하는 스트림이고, 몰입형 오디오 재생을 지원하는 자동차(30)의 예는 돌비 아트모스 몰입형 포맷을 재생하도록 구성된 자동차이다.

[0027] 일반적으로, 교환 포맷은 바람직하게는, 낮은 레이턴시, 낮은 인코드 및 디코드 복잡도, 높은 품질로 스케일링하는 능력, 및 합리적인 코딩 효율을 갖는다. 포맷은 바람직하게는 또한, 구성가능한 레이턴시를 지원하여, 레이턴시가 가변 접속성 조건들 하에서 동작가능하도록 효율 및 또한 에러 복원에 대해 트레이드될 수 있다.

[0028] **청취 능력을 갖거나 갖지 않는 무선 스피커들을 갖는 허브 또는 증강 디스플레이**

[0029] 도 3에 예시된 것은 무선 스피커들(30)의 세트를 구동하는 허브(20), 또는 가능하게는 내장된 스피커들(30)을 갖는 디스플레이(예를 들어, 텔레비전, 또는 TV)(20)가 여러 무선 스피커들(30)로 증강되는 예이다. 증강은, 디스플레이(20)가 스피커들(30)을 포함하는 예들에서, 디스플레이(20)가 또한 오디오 재생의 일부임을 제안한다. 무선 스피커들/디바이스들(30)은 동일한 완전한 신호(도 3의 좌측에 도시된 브로드캐스트 모드) 또는 특정 디바이스에 맞게 조정된 개별 스트림(도 3의 우측에 도시된 유니캐스트-멀티포인트 모드)을 수신하고 있을 수 있다.

[0030] 각각의 스피커는 동일한 채널의 상이한 주파수 범위들을 커버하거나 정규 맵핑(canonical mapping)에서 상이한 채널들에 대응하는 다수의 드라이버들을 포함할 수 있다. 예를 들어, 스피커는 2개의 드라이버들을 가질 수 있고, 그 중 하나는 상승된 스피커를 에뮬레이트하기 위해 높이 채널에 대응하는 신호를 출력하는 상향 발사 드라이버(upwards firing driver)일 수 있다. 무선 디바이스들은 청취 능력을 가질 수 있고(예를 들어, "스마트 스피커"), 그에 따라, 스피커가 하나 이상의 에코-참조를 수신할 것을 요구할 수 있는 에코-관리를 요구할 수 있다. 에코-참조는 (예를 들어, 동일한 스피커/디바이스에 대해) 국지적일 수 있거나, 그 대신에 근처의 다른 스피커/디바이스로부터의 관련 신호를 나타낼 수도 있다.

[0031] 스피커들은 임의의 위치들에 배치될 수 있으며, 이 경우 소위 "유연한 렌더링(flexible rendering)"이 수행될 수 있고, 이에 의해 렌더링은 (예를 들어, 정규 가정(canonical assumption)과는 대조적으로, 이에 의해 라우드스피커들이 고정된 미리 정의된 위치들에 위치되는 것으로 가정되는) 스피커들의 실제 위치를 고려한다. 유연한 렌더링은 허브 또는 TV에서 발생할 수 있고, 후속해서, 렌더링된 신호는, 브로드캐스트 모드에서, 스피커/디바이스에 전송되고, 여기서, 렌더링된 신호들 각각은 각각의 디바이스에 전송되고, 각각의 디바이스는 적절한 신호를 추출하여 출력하거나, 개개의 스트림으로서 개개의 디바이스에 전송된다. 대안적으로, 유연한 렌더링은 각각의 디바이스 상에서 국지적으로 발생할 수 있으며, 이에 의해 각각의 디바이스는 완전한 몰입형 프로그램의 표현, 예를 들어, 7.1.4 채널 기반 몰입형 표현을 수신하고, 그 표현으로부터 출력 신호를 각각의 디바이스에 대해 적절하게 렌더링한다.

[0032] 무선 디바이스들은 디바이스에 의해 제공되는 소프트 액세스 포인트(Soft Access Point)를 통해 또는 가정 내의 로컬 액세스 포인트를 통해 허브 또는 TV에 접속될 수 있다. 이것은 비트레이트 및 레이턴시에 대해 상이한 요건들을 부과할 수 있다.

[0033] **자동차 응용들에서의 모바일 디바이스 투영**

[0034] 다른 이용 사례에서, 모바일 디바이스(예를 들어, 폰 또는 태블릿)(20)는 클라우드 또는 서버(10)로부터 몰입형 오디오 스트림을 페치하고, 몰입형 오디오 스트림을 교환 포맷으로 트랜스코딩하고, 후속하여 트랜스코딩된 스트림을 접속된 자동차(connected car)(30)에 송신한다. 도 2의 예에서, Dolby Atmos-가능형 전화(20)는 낮은 레이턴시 트랜스코딩 및 Dolby Atmos-가능형 자동차(3)로의 송신을 위해, 서버(10)에 접속하고 Dolby Atmos 스트림을 수신한다. 이 이용 사례에서, 거실 이용 사례들에서보다 더 높은 비트레이트들이 이용가능할 수 있고, 또한 무선 채널의 상이한 특성들이 있을 수 있다. 자동차 이용 사례에서 더 높은 비트레이트에 대한 가능한 이유들은, 예컨대, 이웃들, 다른 룸들 등으로부터의 모든 무선 디바이스들이 무선 환경에 잡음을 추가하고 있는

거실에 비해, 자동차가 일종의 패러데이 케이지(faraday cage)로서 작용하고 외부 무선 교란으로부터 자동차 내부 환경을 차폐하기 때문에, 잡음이 적은 무선 환경에 기인한다. 동시에, 통상적으로 가정 내의 모든 무선 디바이스들이 이용가능한 대역폭에 대해 경쟁하고 있는 거실 무선 디바이스들을 포함하는 거실 이용 사례에 비해, 이용가능한 무선 대역폭에 대해 경쟁하고 있는 자동차 내의 매우 적은 무선 디바이스들이 통상적으로 존재한다.

[0035] 이러한 이용 사례의 경우, 모바일 디바이스에 의해 트랜스코딩되고 자동차에 송신될 신호는 채널 기반 몰입형 표현, 객체 기반 표현, 장면 기반 표현(예를 들어, 그리고 앰비소닉스(Ambisonics) 표현), 또는 심지어 상이한 표현들의 조합일 수 있다는 점이 구상된다. 이 예에서, 일반적으로 완전한 프리젠테이션이 모바일 디바이스로부터 단일 엔드-포인트(예를 들어, 자동차)로 이전될 것이므로, 상이한 렌더링 아키텍처들 및 브로드캐스트 대 멀티포인트는 관련되지 않을 수 있다.

[0036] **교환 포맷의 설명**

[0037] 몰입형 교환 포맷은 지각적으로 동기부여된(perceptually motivated) 양자화 및 코딩을 갖는 수정된 이산 코사인 변환(modified discrete cosine transform)(MDCT) 상에 구축된다. 그것은 구성가능한 레이턴시, 예를 들어, 주어진 샘플링 레이트에서 상이한 변환 크기들에 대한 지원을 가진다. 예시적인 프레임 크기들은 48kHz 및 44.1kHz의 샘플링 레이트들에서 128, 256, 512, 1024 및 120, 240, 480, 960 및 192, 384, 768 샘플들이다.

[0038] 포맷은 몰입형 채널 구성들(예를 들어, 5.1.2, 5.1.4, 7.1.2, 7.1.4, 9.1.6, 및 22.2, 또는 ISO/IEC 23091-3:2018, Table-2에 명시된 바와 같은 고유 채널들로 구성되는 임의의 다른 채널 구성을 포함하지만 이에 제한되지 않음)을 포함하는 모노, 스테레오, 5.1 및 다른 채널 구성들을 지원할 수 있다. 포맷은 또한 앰비소닉스(예를 들어, 1차 또는 고차)와 같은 객체 기반 오디오 및 장면 기반 표현들을 지원할 수 있다. 포맷은 또한 (예를 들어, MPEG-4 Audio Standard라고도 지칭될 수 있는 ISO/IEC 14496-3, MPEG-4 System Standard라고도 지칭될 수 있는 ISO 14496-1, ISO Base Media File Format Standard라고도 지칭될 수 있는 ISO 14496-12, 및/또는 MP4 File Format Standard라고도 지칭될 수 있는 ISO 14496-14에 기술된 것들과 같은) 기존 포맷들과의 통합에 적합한 시그널링 스킴을 이용할 수 있다.

[0039] 더욱이, 시스템은 신택스 요소들의 부분들을 스킵하고, 주어진 스피커에 대한 관련 부분들만을 디코딩하는 능력에 대한 지원, 지연 정렬, 레벨 조정 및 등화(equalization)와 같은 메타데이터 제어 플렉스 렌더링 양태들(metadata controlled flex-rendering aspects)에 대한 지원, 청취 능력을 갖는 (스마트 스피커에서의 드라이버/라우드스피커 각각에 독립적으로 공급할 수 있게 하는 신호들의 세트가 전송되는 시나리오를 포함하는) 스마트 스피커들의 이용에 대한 지원, 에코-참조들의 시그널링에 의한 에코-관리에 대한 연관된 지원, 50% 중첩 윈도우들은 물론, 낮은 중첩 윈도우들 모두를 갖는 MDCT 도메인에서의 양자화 및 코딩에 대한 지원, 및/또는 양자화 잡음, 예를 들어, TNS(Temporal Noise Shaping)의 시간 쉐이핑(temporal shaping)을 수행하기 위한 MDCT 도메인에서의 주파수 축을 따른 필터링에 대한 지원을 가질 수 있다. 따라서, 다양한 예들에서, 중첩 윈도우들은 대칭 또는 비대칭이다.

[0040] 개시된 교환 포맷은, 일부 예들에서, 유연한 렌더링 이용 사례에서 신호들 사이의 증가된 상관을 이용하기 위한 개선된 공동 채널 코딩(joint channel coding), 채널 요소들에 걸친 공유 스케일 인자들에 의한 개선된 코딩 효율, MDCT 도메인에서의 고주파수 재구성 및 잡음 추가 기술들의 포함, 당면한 이용 사례들에 대한 더 나은 효율 및 적합성을 허용하기 위한 레거시 코딩 구조들에 대한 다양한 개선들을 제공할 수 있다.

[0041] 브로드캐스트 모드에서 개별 재생 디바이스들을 제어하는 일부 예들에서, 스킵가능한 블록들 및 메타데이터가 이용될 수 있다. 브로드캐스트 모드 설정에서, 각각의 무선 디바이스는 특정 디바이스에 대응하는 특정 드라이버들/채널들에 대한 완전한 오디오의 관련 부분을 재생할 필요가 있을 것이다. 이것은 디바이스가 스트림의 어떤 부분들이 그 디바이스와 관련이 있는지를 알고, 브로드캐스트되는 완전한 오디오 스트림으로부터 그 부분들을 추출할 필요가 있다는 것을 암시한다. 낮은 복잡도의 디코딩 동작들을 가능하게 하기 위해, 특정 디바이스에 대한 디코더가 디코딩에 관련되지 않은 요소들을, 디바이스 및 디바이스 상의 주어진 드라이버들에 관련된 요소들로 효율적으로 스킵할 수 있는 방식으로 스트림이 구성되는 것이 바람직하다.

[0042] 그러나, (압축 관점으로부터) 상이한 디바이스들(예를 들어, 가장 단순한 시나리오에서 스테레오 프리젠테이션의 좌측 및 우측 채널)로 향하는 신호들 사이의 공동 코딩을 여전히 허용하는 이점들이 있을 수 있다는 점에 유의해야 한다.

[0043] 따라서, 포맷은: 상이한 스피커들/디바이스들에 대응하는 신호들 사이에 공동 코딩 기술들을 적용하기 위한 능력을 유지하면서, 특정 디바이스에 대해서만 관련된 부분들의 효율적인 디코딩을 가능하게 하기 위해 비트스트

림에서의 "스킵가능한 블록들"을 가능하게 하고, 특정 디바이스에 대한 하나 이상의 스킵가능한 블록의 유연한 맵핑을 가능하게 하는 메타데이터를 포함할 수 있다.

[0044] 도 4에서, 임의의 셋업의 예가 도시된다. 3개의 접속된 무선 디바이스들(31, 32, 33)이 있고, 그 중 2개는 단일 채널 스피커들(32, 33)(오디오의 하나의 채널을 의미함)이고, 하나의 스피커(31)는 3개의 상이한 드라이버들을 갖는 더 진보된 스피커(31)이며, 이에 의해 3개의 상이한 신호들에 대해 동작한다. 제1 스피커(31)는 이 예에서 3개의 개별 신호에 대해 동작하는 반면, 제2 및 제3 스피커들(32, 33)은 스테레오 표현(예를 들어, 좌측 및 우측)에 대해 동작한다. 이와 같이, 스테레오 표현의 신호들의 공동 코딩을 행하는 것이 유익할 수 있다.

[0045] 상기의 시나리오가 주어지면, 포맷은 제1 디바이스가 그 스피커에 대한 신호들을 디코딩하기 위한 스트림의 관련 부분들만을 추출할 수 있도록 스킵가능한 블록들을 포함하는 비트스트림을 특정하고, 디코더 복잡도 대 효율은 각각의 스피커가 단일 신호를 출력하기 위해 2개의 신호들을 디코딩할 필요가 있는 신호를 구성함으로써 스테레오 쌍에 대해 트레이드오프될 수 있고, 그 장점은 공동 코딩이 수행될 수 있는 것이다.

[0046] 포맷은 복수의 스킵가능한 블록들 중의 특정 블록들을 하나 이상의 디바이스에 맵핑하는 일반적이고 유연한 표현을 가능하게 하는 메타데이터 포맷을 특정한다. 이것은 도 5에 예시되어 있는데, 여기서 맵핑은 각각의 디바이스(31, 32, 33)를 하나 이상의 비트스트림 요소에 연관시키는 행렬로서 표현될 수 있고, 따라서 주어진 디바이스에 대한 디코더는 어떤 비트스트림 요소들을 출력하고 디코딩할지를 알 것이다.

[0047] 예를 들어, 도 5의 예에서, 제1 블록 또는 스킵 블록 Blk1은 3개의 단일 채널 요소(디바이스 1(31)의 각각의 드라이버에 대해 하나씩)를 포함하는 반면, 제2 블록 또는 스킵 블록 Blk2는 디바이스 2(32) 및 디바이스 3(33)에 의해 출력될 신호들의 공동 코딩된 버전들을 포함할 수 있는 채널 쌍 요소를 포함한다. 디바이스 1(31)은 맵핑 메타데이터를 추출하고, 그가 필요로 하는 신호들이 스킵 블록 1 Blk1에 있는 것으로 결정한다. 따라서, 스킵 블록 1 Blk1을 추출하고, 그 안의 3개의 단일 채널 요소를 디코딩하고, 이들을 드라이버들(1, 2a, 2b) 각각에 제공한다.

[0048] 또한, 디바이스 1(31)은 스킵 블록 2 Blk2를 무시한다. 유사하게, 디바이스 2(32)는 맵핑 메타데이터를 추출하고, 그것이 요구하는 신호들이 스킵 블록 2 Blk2에 있다고 결정한다. 따라서, 디바이스 2(32)는 스킵 블록 1 Blk1을 스킵하고, 스킵 블록 2 Blk2를 추출한다. 디바이스 2(32)는 채널 쌍 요소를 디코딩하고, CPE의 좌측 채널 출력을 그것의 드라이버들에 제공한다. 유사하게, 디바이스 3(33)은 맵핑 메타데이터를 추출하고, 그것이 요구하는 신호들이 스킵 블록 2 Blk2에 있다고 결정하고, 따라서 디바이스 3(33)은 스킵 블록 1 Blk1을 스킵하고 또한 스킵 블록 2 Blk2를 추출한다. 디바이스 3(33)은 채널 쌍 요소를 디코딩하고, CPE의 우측 채널 출력을 그것의 드라이버들에 제공한다.

[0049] 일부 예들에서, 디바이스 2(32)는 스킵 블록 2 Blk2로부터의 신호들의 서브세트만을 요구한다고 결정할 수 있다. 그러한 예들에서, 가능할 때, 디바이스 2(32)는 스킵 블록 2 Blk2에서의 신호들을 완전히 디코딩하기 위해 요구되는 동작들의 서브세트만을 수행할 수 있다. 구체적으로, 도 5의 예에서, 디바이스 2(32)는 CPE의 좌측 채널을 추출하기 위해 요구되는 그러한 처리 동작들만을 수행할 수 있으며, 따라서 계산 복잡도의 감소를 가능하게 한다. 유사하게, 디바이스 3(33)은 CPE의 우측 채널을 추출하기 위해 요구되는 그러한 처리 동작들만을 수행할 수 있다.

[0050] 예를 들어, 도 5의 경우, CPE가 공동 채널 코딩을 이용하여 코딩되는 시간들, 및 그것이 독립 채널 코딩을 이용하여 코딩되는 다른 시간들이 존재할 수 있다. CPE가 독립적인 채널 코딩을 이용하여 코딩되는 시간들 동안, 디바이스 2(32)는 CPE의 제1(예를 들어, 좌측) 채널만을 추출할 수 있는 반면, 디바이스 3(33)은 CPE의 제2(예를 들어, 우측) 채널만을 추출할 수 있다.

[0051] 다른 예에서, CPE의 채널들은 공동 채널 코딩을 이용하여 코딩될 수 있으며, 이 경우에 디바이스들 2(32) 및 3(33)은 CPE의 2개의 중간 채널을 추출해야 한다. 그러나, 디바이스 2(32)는 중간 채널들로부터 좌측 채널을 추출하기 위해 요구되는 그러한 동작들만을 수행함으로써 감소된 계산 복잡도로 여전히 동작할 수 있다. 유사하게, 디바이스 3(33)은 중간 디코딩된 채널들로부터 우측 채널을 추출하기 위해 요구되는 그러한 동작들만을 수행함으로써 감소된 계산 복잡도로 동작할 수 있다.

[0052] 공동 채널 코딩의 상세에 따라, 다른 최적화가 가능할 수 있다. 상이한 디바이스들/디코더들 각각의 아이덴티티는 시스템 초기화 또는 셋업 단계 동안 정의될 수 있다. 그러한 셋업은 일반적인 것이며, 일반적으로 방의 음향, 스위트 스폿(sweet spot)까지의 스피커 거리 등을 측정하는 것을 수반한다.

[0053] **유연한 렌더링(프리-인코드 및 포스트 디코드)의 분배 및 지연 등의 적용. 포스트 디코더**

[0054] 특정 디바이스들/스피커들에 관련 신호들을 전송하기 전에 허브/TV에서 유연한 렌더링이 적용되는 이용 사례의 경우, 렌더링은, 코딩 관점에서, 예를 들어, 그러한 신호들의 공동 코딩을 고려할 때, 코딩하기 더 어려운 신호들을 생성할 수 있다. 한 가지 이유는 유연한 렌더링이 (예를 들어, 다른 스피커들 및 청취자에 대한 스피커의 배치에 따라) 상이한 디바이스들에 대해 상이한 지연들, 등화 및/또는 이득 조정들을 적용할 수 있다는 것이다. 또한, 정보, 예를 들어, 이득 및 지연을 초기 셋업으로 미리 설정하고, 유연하게 등화만을 렌더링하는 것도 가능할 것이다. 미리 설정된 정보 및 유연한 렌더링 정보의 다른 변형들이 또한 다른 예들에서 가능할 수 있다. 본 문서에서, "이득"이라는 용어는, 특정 레벨 조정들(예를 들어, 증폭)만으로 제한되는 것이 아니라, 임의의 레벨 조정(예를 들어, 감쇠, 증폭, 또는 통과(pass-through))을 의미하는 것으로 해석되어야 한다는 점에 유의한다.

[0055] 도 6의 예에서, 우측 채널(33) 및 좌측 채널(32) 스피커들에는 상이한 레이턴시들이 주어져서, 청취자에 대해 상이한 배치들을 반영한다(예를 들어, 스피커들(32, 33)은 청취자에게 등거리가 아닐 수 있으므로, 상이한 스피커들(32, 33)로부터의 코히어런트 사운드들이 동시에 청취자에게 도달하도록 스피커들(32, 33)로부터 출력된 신호들에 상이한 레이턴시들이 적용될 수 있다). 상이한 스피커들(31, 32, 33)을 통한 재생을 위해 의도된 코히어런트 신호들에 대한 그러한 상이한 레이턴시들의 도입은 그러한 신호들의 공동 코딩을 어렵게 만든다.

[0056] 이러한 과제를 해결하기 위해, 유연한 렌더링 프로세스의 양태들은 신호의 디코딩 후에 엔드포인트 디바이스에서 파라미터화되고 적용될 수 있다.

[0057] 도 7의 예에서, 각각의 디바이스에 대한 지연 및 이득 값들은 파라미터화되고, 각각의 스피커들(31, 32, 33)에 전송된 인코딩된 신호들에 포함된다. 각각의 신호들은 각각의 디바이스들(31, 32, 33)에 의해 디코딩될 수 있고, 그 후, 각각의 디코딩된 신호들에 파라미터화된 이득 및 지연 값들을 도입할 수 있다.

[0058] 상이한 디바이스들(31, 32, 33)에 대한 코딩된 신호들이 분리가능한 블록들(예를 들어, 스킵가능한 블록들)에서 전송되는 도 7에 도시된 것과 같은 예에서, 파라미터들(예를 들어, 지연 및 이득)은 또한 분리가능한 블록들에서 전송될 수 있고, 따라서, 디바이스(31, 32, 33)는 그 디바이스(31, 32, 33)에 대해 요구되는 파라미터들의 서브세트만을 추출하고, 그 디바이스(31, 32, 33)에 대해 요구되지 않는 파라미터들을 무시(및 스킵)할 수 있다. 그러한 경우에, 어느 파라미터들이 어느 블록들에 포함되는지를 표시하는 맵핑 메타데이터가 각각의 디바이스에 제공될 수 있다.

[0059] 도 7은 지연 및 이득 파라미터들만을 나타내지만, 등화 파라미터들과 같은 다른 파라미터들도 포함될 수 있다는 점에도 유의한다. 등화 파라미터들은, 예를 들어, 상이한 주파수 영역들에 적용될 복수의 이득들, 재생 디바이스(31, 32, 33)에 의해 적용될 미리 결정된 등화 곡선의 표시, 무한 임펄스 응답(IIR) 또는 유한 임펄스 응답(FIR) 필터 계수들의 하나 이상의 세트, 바이쿼드 필터 계수들(biquad filter coefficients)의 세트, 파라메트릭 등화기의 특성들을 특정하는 파라미터들뿐만 아니라, 본 기술분야의 통상의 기술자에게 알려진 등화를 특정하기 위한 다른 파라미터들을 포함할 수 있다.

[0060] 또한, 유연한 렌더링 양태들의 파라미터화는 정적일 필요는 없고, (예를 들어, 오디오 프로그램의 재생 동안 청취자가 이동하는 경우에) 동적일 수 있다. 이와 같이, 파라미터들이 동적으로 변하는 것을 허용하는 것이 바람직할 수 있다. 오디오 프로그램 동안 하나 이상의 파라미터가 변하는 경우에, 디바이스는 이전의 지연 및/또는 이득 파라미터들과 업데이트된 지연 및/또는 이득 파라미터들 사이를 보간하여 매끄러운 전이를 제공할 수 있다. 이것은 시스템이 청취자의 위치를 동적으로 추적하고, 그에 대응하여 동적 렌더링을 위해 스위트 스폿을 업데이트하는 상황들에서 특히 유용할 수 있다.

[0061] 또한 주목되었고, 후속하여 설명될 바와 같이, 유연한 렌더링이 적용될 때, 채널들 사이의 상관의 증가된 레벨이 발생할 수 있고, 이것은 더 유연한 공동 코딩에 의해 이용될 수 있다.

[0062] **에코-참조 코딩 및 시그널링**

[0063] 도 8에 나타낸 바와 같은 이용 사례의 경우, 도 8의 좌측에 나타낸 브로드캐스트 모드 또는 도 8의 우측에 나타낸 유니캐스트 멀티포인트 모드에서 신호를 수신하면서 함께 동작하는 다수의 디바이스/스피커(30)가 있고, 동시에 "청취" 능력을 가능하게 하는 디바이스(30) 상에 마이크로폰들(40)이 있는 경우, 에코-관리에 대한 필요성이 발생한다.

[0064] 다수의 스피커/디바이스로 에코-관리를 수행할 때, 단지 로컬 스피커 디바이스를 넘어선 것으로부터의 에코-참

조를 이용하는 것이 유익할 수 있다. 예로서, 하나의 디바이스는 다른 디바이스에 근접하게 배치될 수 있고, 따라서, 근접한 디바이스로부터의 신호는 활성 마이크로폰과의 디바이스의 에코-관리에 영향을 줄 것이다. 브로드캐스트 모드 이용 사례에서, 각각의 디바이스는 모든 디바이스들에 대한 신호들을 수신한다. 디바이스가 다른 디바이스들에 대한 신호들을 가질 때, 이들 신호를 에코-참조로서 이용하는 것이 유익할 수 있다. 이를 위해, 신호들 중 어느 것이 어느 다른 디바이스들에 대한 에코-참조들로서 이용될 수 있는지를 특정 디바이스에 시그널링하는 것이 필요하다.

[0065] 일 예에서, 이것은 특정 스피커/디바이스에 의해 (전체 세트로부터) 플레이아웃될(played out) 채널/신호를 맵핑할 뿐만 아니라, 특정 스피커/디바이스에 대한 (전체 세트로부터의) 에코-참조로서 이용될 채널/신호를 맵핑하는 메타데이터를 제공함으로써 이루어질 수 있다. 그러한 메타데이터 또는 시그널링은 동적일 수 있어서, 바람직한 에코-참조의 표시가 시간에 따라 변할 수 있게 한다.

[0066] 각각의 디바이스/스피커가 플레이아웃할 특정 신호만을 수신하는 이용 사례의 경우, 적절한 에코-참조를 제공하기 위하여, 각각의 디바이스/스피커에 추가적인 신호(예를 들어, 에코-참조 신호)를 송신할 필요가 있을 수 있다. 다시, 그렇게 하기 위해, 각각의 디바이스가 플레이아웃을 위한 적절한 신호 및 에코-관리를 위한 적절한 신호들을 선택할 수 있도록 디바이스-특정 시그널링을 제공하는 것이 필요하다.

[0067] 에코-관리를 위한 신호는 에코 관리를 수행하기 위한 디바이스에 의해서만 이용되고, 청취자를 위해 디바이스에 의해 플레이아웃되지 않기 때문에, 에코-참조 신호는 디바이스에 의한 재생을 위해 의도되는 신호와는 상이하게 코딩되거나 표현될 수 있다. 구체적으로, 성공적인 에코-관리는 청취자를 위한 플레이아웃을 위해 정상적으로 이용될 것보다 더 낮은 레이트로 코딩된 신호로 달성될 수 있기 때문에, 전혀 청취자에게 플레이아웃하기에 적합하지 않을 수 있지만, 오디오 신호의 필요한 특징들을 캡처하는 신호의 파라메트릭 표현과 같은 추가적인 압축 툴들은 상당히 감소된 송신 비용으로 양호한 에코-관리를 제공할 수 있다.

[0068] **다양한 신택스 요소들**

[0069] 일부 예들에서, 오디오의 전송을 최적화하기 위해 블록들이 이용된다. 설명된 포맷에서, 각각의 프레임은, 스킵가능한 블록들과 관련하여 전술한 바와 같이, 블록들로 분할될 수 있다. 블록은 그것이 속하는 프레임 번호, 동일한 ID를 갖는 상이한 프레임들로부터의 연속 블록들을 블록 스트림에 연관시키는 데 이용될 수 있는 블록 ID, 및 재송신을 위한 우선순위에 의해 식별될 수 있다. 상기의 일 예는 다수의 프레임들 N-2, N-1 및 N을 갖는 스트림이고, 프레임 N은 도 13에 도시된 ID1, ID2, ID3으로서 식별된 다수의 블록을 포함한다. 비트스트림 포맷인 이 스트림의 예가 도 14에 도시되어 있다. 일부 예들에서, 몰입형 오디오 프로그램의 오디오 신호들에 대해, 블록의 블록 ID는 전체 몰입형 오디오 프로그램의 신호들의 어느 세트가 그 블록에 의해 운반되는지를 표시할 수 있다.

[0070] 설명된 포맷에 대한 이용 사례는 낮은 레이턴시에서 Wifi와 같은 무선 네트워크를 통한 오디오의 신뢰성 있는 송신이다. 예를 들어, Wifi는 패킷 기반 네트워크 프로토콜을 이용한다. 패킷 크기는 보통 제한된다. IP 네트워크들에서의 전형적인 최대 패킷 크기는 1500 바이트이다. 스트림의 블록 기반 아키텍처는 송신을 위해 패킷들을 어셈블링할 때 유연성을 허용한다. 예를 들어, 더 작은 프레임들을 갖는 패킷들은 다른 프레임들로부터 재송신된 블록들로 채워질 수 있다. 큰 프레임들은 이들을 패킷화하기 전에 블록 경계들 상에서 분할되어 네트워크 프로토콜 계층 상의 패킷들 사이의 종속성을 감소시킬 수 있다.

[0071] 도 9는 프레임들, 블록들, 및 패킷들 사이의 관계를 도시한다. 프레임은 시작 시간, 종료 시간, 및 종료 시간과 시작 시간의 차이인 지속기간을 갖는 오디오 신호의 연속 세그먼트를 나타내는 오디오 데이터, 바람직하게는 모든 오디오 데이터를 운반한다. 연속 세그먼트는 ISO/IEC 14496-3, 서브파트 4, 섹션 4.5.2.1.1에 따른 시간 기간을 포함할 수 있다. 섹션 4.5.2.1.1은 raw_data_block()의 콘텐츠를 기술한다. 프레임은 또한, 예를 들어, 더 낮은 데이터 레이트로 인코딩된 그 세그먼트의 중복 표현들을 운반할 수 있다. 인코딩 후에, 그 프레임은 블록들로 분할될 수 있다. 블록들은 패킷 기반 네트워크를 통한 송신을 위해 패킷들로 결합될 수 있다. 상이한 프레임들로부터의 블록들은 단일 패킷으로 결합될 수 있고/있거나 비순차적으로 전송될 수 있다.

[0072] 예에서, 블록들은 개별 디바이스들을 어드레싱하기 위해 이용된다. 데이터의 패킷들은 개별 디바이스들 또는 관련 디바이스 그룹들에 의해 수신된다. 스킵가능한 블록들의 개념은 개별 디바이스들 또는 관련 디바이스 그룹들을 어드레싱하는 데 이용될 수 있다. 네트워크가 패킷들을 상이한 디바이스들로 전송할 때 브로드캐스트 모드에서 동작하더라도, 오디오의 처리(예를 들어, 디코딩, 렌더링 등)는 그 디바이스로 어드레싱되는 블록들로 감소될 수 있다. 모든 다른 블록들은, 동일한 패킷 내에서 수신되더라도, 단순히 스킵될 수 있다. 일부 예들

에서, 블록들은 그들의 디코드 또는 프리젠테이션 시간에 기초하여 올바른 순서로 가져올 수 있다. 더 낮은 우선순위를 갖는 재송신된 블록들은 더 높은 우선순위를 갖는 동일한 블록이 또한 수신되었다면 제거될 수 있다. 블록들의 스트림은 이후 디코더에 공급될 수 있다.

[0073] 일부 예들에서, 스트림 및 디바이스들의 구성은 대역외 전송된다. 코덱은 오디오 스트림들이 비교적 높은 레이트로 그러나 낮은 레이턴시로 송신되는 접속을 셋업하는 것을 가능하게 한다. 그러한 접속의 구성은 그러한 접속의 지속기간에 걸쳐 안정적으로 유지될 수 있다. 그 경우, 오디오 스트림의 구성 부분을 만드는 대신에, 그것은 대역외 전송될 수 있다. 그러한 대역외 송신을 위해, 심지어 상이한 네트워크들, 또는 네트워크 프로토콜들이 이용될 수 있다. 예를 들어, 오디오 스트림은 낮은 레이턴시 송신을 위해 사용자 데이터그램 프로파일(UDP)을 이용할 수 있는 반면, 구성은 구성의 신뢰성 있는 송신을 보장하기 위해 송신 제어 프로토콜(TCP)을 이용할 수 있다.

[0074] **MPEG-4 오디오와 관련하여 코덱을 가능하게 하는 것**

[0075] 본 기술의 하나의 특정 응용은 MPEG-4 오디오 내에서의 이용이다. MPEG-4 오디오에서, 상이한 코덱 기술들에 대해 상이한 AOT(Audio Object Type)들이 정의된다. 본 문서에서 설명된 포맷의 경우, 새로운 AOT가 정의될 수 있고, 이는 그 포맷에 대한 특정 시그널링 및 데이터를 허용한다. 또한, MPEG-4에서의 디코더의 구성은 DecoderSpecificInfo() 페이로드에서 행해지고 이는 결과적으로 AudioSpecificConfig() 페이로드를 운반한다. 후자에서, 특정 AOT에 대한 특정 정보뿐만 아니라, 샘플링 레이트 및 채널 구성들과 같은, 특정 포맷의 특정 일반 시그널링 불가지론(certain general signaling agnostic)이 정의된다. 전체 스트림이 단일 디바이스에 의해 디코딩되는 종래의 포맷들에 대해, 이것은 이치에 맞을 수 있다. 그러나, 단일 스트림이 여러 디코더로 송신되는 브로드캐스트 모드에서, 각각의 개별 디코더는 스트림의 부분들만을 디코딩하며, (디바이스의 출력 기능을 셋업하기 위한 수단으로서) 업프론트 채널 구성 시그널링(up-front channel configuration signaling)은 차선일 수 있다.

[0076] 도 10은 종래의 MPEG-4 하이 레벨 구조(흑색)를 회색으로 수정하여 나타낸 것이다. 브로드캐스트 이용 사례를 지원하기 위해, 특정 이용 사례에 대해 시그널링이 재정의되는 codecSepcificConfig()(여기서, "codec"은 일반 플레이스홀더 명칭일 수 있음)가 정의되어, 특정 채널 요소를 특정 디바이스에 맵핑할 뿐만 아니라 다른 관련 정적 파라미터를 포함할 수 있다. "0"의 값을 갖는 MPEG-4 요소 channelConfiguration은 codecSpecificConfig에서 정의된 채널 구성으로서 정의된다. 따라서, 이 값은 코덱 특정 구성 내부의 채널 구성들의 시그널링의 수정을 가능하게 하는 데 이용될 수 있다.

[0077] 또한, MPEG-4의 사상에서는, codecSpecificConfig()가 디코딩가능한 경우에, 당면한 특정 디코더에 대해 원시 페이로드들(raw payloads)이 지정된다. 그러나, 포맷은 동적 메타데이터가 원시 페이로드들의 일부인 것을 보장하고, 모든 원시 페이로드들에 대해 길이 정보가 이용가능한 것을 보장하며, 따라서 디코더는 특정 디바이스에 관련되지 않은 요소들을 쉽게 스킵할 수 있다.

[0078] 도 11에서, MPEG-4에서 정의된 바와 같은 예시적인 raw_data_block의 일부가 우측에 주어진다. 원시 데이터 블록은 주어진 순서로 채널 요소들(단일 채널 요소들(SCE) 또는 채널 쌍 요소들(CPE))을 포함한다. 그러나, 당면한 출력 디바이스에 관련되지 않을 수 있는 이들 채널 요소의 부분들을 스킵하고 싶어하는 디코더는, 종래의 MPEG-4 오디오 신택스에서, 관련 부분들을 추출할 수 있도록 모든 채널 요소를 파싱(및 어느 정도 또한 디코딩)해야만 한다. 도 11의 좌측에 도시된 새로운 raw_data_block에서, 콘텐츠는, 디코더가 관련없는 부분들을 스킵하고, 메타데이터에 의해 지시된 채널 요소들만을 당면한 디바이스에 관련된 것으로 디코딩할 수 있도록, 스킵가능한 블록들로 구성될 것이다. 예에서, 스킵가능한 블록들은 raw_data_block 및 관련 정보를 포함한다.

[0079] 재송신을 위해 블록들을 이용하는 것도 가능할 것이다. 각각의 프레임은 상기의 스킵가능한 블록들 섹션에서 설명된 바와 같이 블록들로 분할될 수 있다. 블록은, 도 15 및 도 16에 나타낸 바와 같이, 그것이 속하는 프레임 번호, 동일한 ID를 갖는 상이한 프레임들로부터의 연속 블록들을 블록 스트림에 연관시키는데 이용될 수 있는 블록 ID, 및 재송신을 위한 우선순위에 의해 식별될 수 있다. 예를 들어, (예를 들어, 도 15 및 도 16에서 우선순위 0으로 표시된) 재전송을 위한 높은 우선순위는, 동일한 블록 ID 및 프레임 카운터를 갖지만 재송신을 위한 더 낮은 우선순위(예를 들어, 도 15 및 도 16에서 우선순위 1로 표시됨)를 갖는 다른 블록보다 수신기에서 블록이 선호된다는 것을 시그널링한다. 도 15 및 도 16의 예에서와 같이, 우선순위 인덱스가 증가함에 따라 우선순위가 감소할 수 있는 반면(예를 들어, 우선순위 1은 우선순위 0보다 낮은 우선순위일 수 있음), 다른 예들에서, 우선순위 인덱스가 증가함에 따라 우선순위가 증가할 수 있다는 점에 유의한다(예를 들어, 우선순위 0은 우선순위 1보다 낮은 우선순위임). 또 다른 예들은 통상의 기술자에게 명백할 것이다.

[0080] 신택스는 오디오 요소들의 재송신을 지원할 수 있다. 재송신을 위해, 다양한 품질 레벨들이 지원될 수 있다. 따라서, 도 17 및 도 18에 예시된 바와 같이, 동일한 프레임 카운터 및 블록 ID로 수신된 블록들이 중복되고 따라서 상호 배타적인 디코더에 대한 것이기 때문에, 재송신된 블록들은 동일한 블록 ID를 갖는 어느 블록들이 디코더에 대해 우선순위를 가져야 하는지를 나타내기 위해 '우선순위' 플래그를 운반할 수 있다.

[0081] 블록들의 재송신은 감소된 데이터 레이트로 행해질 수 있다. 그러한 감소된 데이터 레이트는 오디오 신호의 신호 대 잡음비를 감소시키거나, 오디오 신호의 대역폭을 감소시키거나, 오디오 신호의 채널 카운트를 감소시키거나(예를 들어, 그 전체가 참조로 포함되는 미국 특허 제11,289,103호에 설명된 바와 같이), 이들의 임의 조합에 의해 달성될 수 있다. 디코더가 최상의 가능한 품질을 제공하는 오디오 블록을 선택하기 위해, 최고 품질 신호들을 제공하는 블록은 최고 디코딩 우선순위를 가질 수 있고, 제2 최고 품질 신호들을 갖는 블록은 제2 최고 우선순위를 가질 수 있는 등이다.

[0082] 블록들은 또한 동일한 품질 레벨에서 재송신될 수 있다. 그러한 경우, 우선순위는 재송신된 블록의 레이턴시를 반영할 수 있다.

[0083] **코어 코딩 효율을 개선하기 위한 툴들**

[0084] 추가의 예들은 채널 요소들을 통해 MDCT 양자화 스케일 인자들을 공유하는 것을 가능하게 할 수 있다. 일부 경우들에서, 연관된 사이드 비트 레이트를 감소시키기 위해 특정 채널들에 걸쳐 MDCT 스케일 인자들을 공유하는 것이 가능할 수 있다. 더욱이, 스케일 인자 공유는, 예를 들어, 7.1.4 입력에 대해서와 같이, 상이한 채널 요소들로 확장될 수 있다. 공유된 스케일 인자들의 이용은 3개의 활성 구성을 허용하는 2개의 시그널링 비트에 의해 표시될 수 있다. 하나의 가능한 구성은 좌측 수평 채널들, 좌측 최상부 채널들, 우측 수평 채널들, 및 우측 최상부 채널들에서 스케일 인자들을 공유하는 것일 것이다. 스케일 인자들이 스킵 블록들 내에서만 공유되는 것을 보장하기 위해, 특정 신택스가 스킵가능한 블록들 개념에 정렬될 수 있다.

[0085] 2개보다 많은 채널의 공동 코딩을 갖는 것도 가능할 것이다. 도 4, 도 5, 도 6, 및 도 7에 예시된 예에서, 블록 Blk1은 3개의 SCE를 운반하는데, 하나는 이 시나리오에서 고려되는 스마트 스피커에서의 드라이버들 각각에 대한 것이다. 그러한 상황에서, (스테레오 예측, 예를 들어, MPEG-D USAC라고 지칭되는 ISO/IEC 23003-3에 기술된 바와 같은 MDCT-기반 복합 예측 스테레오 툴을 포함하도록 확장될 수 있는, CPE에 의해 제공되는 바와 같은) 단지 2개의 채널뿐만 아니라, 예를 들어, ETSI TS 103 190에 도입된 SAP(Stereo Audio Processing) 툴에 기술된 바와 같은, 2개 초과의 채널의 공동 코딩을 가능하게 하는 것이 유익할 수 있다.

[0086] 유사하게, 유연한 렌더링의 경우, 디바이스들에 걸친 신호들 사이에 많은 양의 상관이 존재할 수 있다는 점에 유의한다. 그러한 신호에 대해, 다수의 디바이스를 커버하는 스킵 블록을 구성하고, 앞서 개략적으로 기술한 툴을 이용하여 2개 초과의 채널에 걸쳐 공동 채널 코딩이 적용될 수 있게 하는 것이 유익할 수 있다.

[0087] 특정 변환 길이들에 대해 이용될 수 있는 다른 공동 채널 코딩 툴은 채널 결합(channel coupling)이며, 여기서 합성 채널(composite channel) 및 스케일 인자 정보가 중간 및 높은 주파수들에 대해 송신된다. 이것은 양호한 품질 범위에서 재생을 위한 비트레이트 감소를 제공할 수 있다. 예를 들어, 낮은 레이턴시 코딩을 위해 256개 샘플 주변의 프레임 길이에 대응하는 256의 변환 길이를 갖는 프레임들에 대해 채널 결합 툴을 이용하는 것이 유익할 수 있다.

[0088] 본 개시내용은 또한 대역제한된 신호들의 효율적인 코딩을 허용할 것이다. 스마트 스피커의 상이한 드라이버들을 공급하는 별개의 신호들이 전송되는 시나리오에서, 예를 들어, 우퍼(woofer), 중간 범위(mid-range), 및 트위터(tweeter)를 갖는 3-웨이 드라이버 구성의 경우와 같이, 이들 신호들 중 일부는 대역제한될 수 있다. 따라서, 그러한 대역제한된 신호들의 효율적인 인코딩이 바람직하며, 이는 개선된 코딩 효율 및/또는 감소된 계산 복잡도(예를 들어, 우퍼 피드에 대한 대역제한된 IMDCT의 이용을 가능하게 함)을 갖는 그러한 시나리오들을 처리하기 위해 신택스의 잠재적 수정들뿐만 아니라 특별히 튜닝된 심리음향학 모델들(psychoacoustic models) 및 비트 할당 전략들로 변환될 수 있다.

[0089] 일부 예들에서, MDCT 도메인에서의 고주파수 재구성 및 잡음 추가의 조합이 가능할 것이다. 현대의 오디오 코덱들은 전형적으로 파라메트릭 코딩 기술들을 지원하도록 설계되고, 유사하게 여기서 예를 들어, MDCT 도메인에 고주파수 재구성 방법들을 포함하여 낮은 레이턴시를 유지하고, MDCT 도메인에서 잡음 추가 스킴들을 수행하여 재구성된 고대역에서 토널 대 잡음비(tonal to noise ratio)를 관리하는 것을 구상할 수 있다. 그러한 파라메트릭 코딩 기술들은 특히 더 낮은 동작 포인트들에 대해, 그리고 특히 재송신이 FEC(Forward Error Correction) 스킴의 일부로서 발생하는 시나리오들에서 유용할 수 있고, 여기서 전형적으로 메인 신호가 송신되고, 그 후 동

일한 신호가 지연된 방식으로 더 낮은 비트레이트로 재송신된다.

[0090] 이러한 예들은 인코더에서 피크 복잡도 감소를 더 허용할 것이다. 이것은 일정한 비트레이트 전송 채널에 대한 버퍼 모델 제한들이 준비되지 않은 경우에 적용될 수 있다. 버퍼 모드를 갖는 일정한 비트레이트에서, 양자화 및 카운팅 후에, 인코딩될 프레임을 위해 이용될 결과적인 비트들의 수가 허용된 비트들의 한계 위에 있는 것이 발생할 수 있다. 비트 요건을 충족시키기 위해 적어도 하나의 새로운, 더 거친 양자화 및 비트 카운팅 단계가 수행될 필요가 있다. 이 제한이 더 완화된 경우, 인코더는 제1 양자화 결과를 유지할 수 있고, 약간 더 높은 순간 비트레이트(instantaneous bitrate)로 행해질 것이다. 여전히 버퍼 모델 인코더에 의한 일정한 비트레이트의 것과 유사한 인코더 거동은, 버퍼 요건들을 준수하는 인코더에 따라 업데이트되는 가상 버퍼 충만도(virtual buffer fullness)와 함께, 동일한 비트 저장소 제어 메커니즘을 이용함으로써 달성될 수 있다. 이것은 추가적인 양자화 및 비트 카운팅 단계들을 절약할 것이고, 결과적인 오디오 품질은 약간 증가된 전체 비트레이트의 단점을 갖는 일정한 비트레이트 경우에 비해 동일하거나 더 양호할 것이다.

[0091] **리턴 채널 개념**

[0092] 재생 및 청취 이용 사례들의 경우, 스마트 스피커들(30)은 모노 또는 공간 사운드필드를 (예를 들어, 모노, 스테레오, A-포맷, B-포맷, 또는 임의의 다른 등방성 또는 이방성 채널 포맷으로) 캡처하도록 구성되는 하나 이상의 마이크로폰(40)(예를 들어, 마이크로폰 또는 마이크로폰 어레이)을 가질 수 있고, 코덱이 그러한 포맷들의 효율적인 코딩을 낮은 레이턴시 방식으로 수행할 수 있는 것에 대한 필요성이 존재하여, 동일한 코덱이, 예를 들어, 도 12에 예시된 리턴 채널에 대해 이용되는 바와 같이 스마트 스피커 디바이스(30)로의 브로드캐스트/송신을 위해 이용된다.

[0093] 웨이크 워드 검출(wake word detection)이 스마트 스피커 디바이스 상에서 수행되는 동안, 음성 인식은 전형적으로 클라우드에서 행해지고, 기록된 음성의 적합한 세그먼트는 웨이크 워드 검출에 의해 트리거된다는 것에 또한 유의해야 한다.

[0094] 이러한 맥락에서, 예를 들어, 코덱 내의 중간 포맷/표현에 대해 음성 분석 시스템들을 생성함으로써 시스템 복잡도를 전반적으로 절약하는 것이 흥미로울 수 있다. 인간 대 인간 대화가 진행 중이 아닌 가장 간단한 이용 사례의 경우, 음성 인식 작업에 특유한 적합한 표현이 정의될 수 있는데, 그 이유는 그것이 다른 인간에 의해 청취되지 않아야 하기 때문이다. 이것은 대역 에너지들, 멜-주파수 켑스트럴 계수(Mel-frequency Cepstral Coefficient)(MFCC)들 등, 또는 MDCT 코딩된 스펙트럼들의 낮은 비트레이트 버전일 수 있다.

[0095] 인간 대 인간 대화가 진행 중이고, 웨이크 워드 및 요구된 음성 인식이 인터리빙되는 이용 사례에서, 이에 대한 스트림의 관련 부분들을 추출할 수 있기를 원할 것이다. 그러한 추출은 메인 음성 신호와 병렬로 전송되는 단순히 추가적인 데이터의 계층화된 코딩 구조를 수반할 수 있다. 또한, 기존의 디코딩된 MDCT 스펙트럼들로부터 관련 표현으로의 정의된 트랜스코드를 구상하고, 그렇게 함으로써 구조를 단순화한다. 본질적으로, 디코더는 음성 인식 표현을 (병렬로) 또한 디코딩해야 한다는 것을 시그널링할 때까지 인간 청취가능 오디오를 디코딩하고 출력할 것이다. 이 표현은 계층화된 스트림, 간단히 동일한 완전한 데이터의 대안 디코드, 또는 간단히 스트림에서의 추가적인 표현의 디코드 및 출력의 "필드 오프(peeled off)"일 수 있다. 이 이용 사례에서, 시그널링은 수신 측의 디코더가 음성 인식 관련 표현을 출력해야 함을 나타내는 인에이블링 피스(enabling piece)이다.

[0096] **열거된 예들**

[0097] 이하에서, 청구항들이 아닌 7개의 열거된 예시적인 예들(EEE-A들, EEE-B들, EEE-C들, EEE-D들, EEE-E들, EEE-F들, 및 EEE-G들)의 세트들이 본 명세서에 개시된 예들의 양태들을 기술한다.

[0098] EEE-A1. 오디오 신호를 디코딩하기 위한 방법으로서,

[0099] 적어도 하나의 프레임을 포함하는 비트스트림을 수신하는 단계 - 적어도 하나의 프레임의 각각의 프레임은 복수의 블록들을 포함함 -;

[0100] 시그널링 데이터로부터, 출력 디바이스의 디바이스 정보에 기초하여, 디코딩할 때 스킵될 복수의 블록들 중의 하나 이상의 블록의 부분들을 식별하는 정보를 결정하는 단계; 및

[0101] 하나 이상의 블록의 식별된 부분들을 스킵하면서 비트스트림을 디코딩하는 단계를 포함하는, 방법.

[0102] EEE-A2. EEE-A1의 방법으로서, 디코딩할 때 스킵될 복수의 블록들 중의 하나 이상의 블록의 부분들을 식별하는

정보는 복수의 출력 디바이스들 중의 각각의 출력 디바이스를 하나 이상의 비트스트림 요소에 연관시키는 행렬을 포함하는, 방법.

[0103] EEE-A3. EEE-A2의 방법으로서, 하나 이상의 비트스트림 요소는 대응하는 연관된 출력 디바이스에 대한 비트스트림의 디코딩을 위해 요구되는, 방법.

[0104] EEE-A4. EEE-A1 내지 EEE-A3 중 어느 하나의 방법으로서, 출력 디바이스는 무선 디바이스, 모바일 디바이스, 태블릿, 단일-채널 스피커, 및/또는 멀티-채널 스피커 중 적어도 하나를 포함할 수 있는, 방법.

[0105] EEE-A5. EEE-A1 내지 EEE-A4 중 어느 하나의 방법으로서, 식별된 부분은 적어도 하나의 블록을 포함하는, 방법.

[0106] EEE-A6. EEE-A1 내지 EEE-A5 중 어느 하나의 방법으로서, 출력 디바이스는 제1 출력 디바이스이고, 방법은 비트스트림의 하나 이상의 신호 사이의 공동 코딩 기술들을 제2 출력 디바이스 및 제3 출력 디바이스에 적용하는 단계를 더 포함하는, 방법.

[0107] EEE-A7. EEE-A1 내지 EEE-A6 중 어느 하나의 방법으로서, 각각의 출력 디바이스 및/또는 디코더의 아이덴티티는 시스템 초기화 단계 동안 정의되는, 방법.

[0108] EEE-A8. EEE-A1 내지 EEE-A7 중 어느 하나의 방법으로서, 시그널링 데이터는 비트스트림의 메타데이터로부터 결정되는, 방법.

[0109] EEE-A9. EEE-A1 내지 EEE-A8 중 어느 하나의 방법을 수행하도록 구성된 장치.

[0110] EEE-A10. 실행될 때, 하나 이상의 디바이스로 하여금 EEE-A1 내지 EEE-A8 중 어느 하나의 방법을 수행하게 하는 명령어들의 시퀀스를 포함하는 비일시적 컴퓨터 판독가능 저장 매체.

[0111] EEE-B1. 복수의 오디오 신호들을 포함하는 오디오 프로그램으로부터 인코딩된 비트스트림을 생성하기 위한 방법으로서,

[0112] 복수의 오디오 신호들 각각에 대해, 각각의 오디오 신호가 연관되는 재생 디바이스를 표시하는 정보를 수신하는 단계;

[0113] 각각의 재생 디바이스에 대해, 각각의 재생 디바이스와 연관된 지연, 이득 및 등화 곡선 중 적어도 하나를 표시하는 정보를 수신하는 단계;

[0114] 복수의 오디오 신호들로부터, 2개 이상의 관련 오디오 신호들의 그룹을 결정하는 단계;

[0115] 그룹의 2개 이상의 관련 오디오 신호들에 하나 이상의 공동 코딩 툴을 적용하여 공동 코딩된 오디오 신호들을 획득하는 단계;

[0116] 공동 코딩된 오디오 신호들, 공동 코딩된 오디오 신호들이 연관되는 재생 디바이스들의 표시, 및 공동 코딩된 오디오 신호들이 연관되는 각각의 재생 디바이스들과 연관된 지연 및 이득의 표시들을 인코딩된 비트스트림의 독립 블록으로 결합하는 단계를 포함하는, 방법.

[0117] EEE-B2. EEE-B1의 방법으로서, 각각의 재생 디바이스와 연관된 지연, 이득 및/또는 등화 곡선은 청취자의 위치에 대한 각각의 재생 디바이스의 위치에 의존하는, 방법.

[0118] EEE-B3. EEE-B1 또는 EEE-B2의 방법으로서, 각각의 재생 디바이스와 연관된 지연, 이득 및/또는 등화 곡선은 다른 재생 디바이스들의 위치들에 대한 각각의 재생 디바이스의 위치에 의존하는, 방법.

[0119] EEE-B4. EEE-B1 내지 EEE-B3 중 어느 하나의 방법으로서, 지연, 이득, 및/또는 등화 곡선은 동적으로 가변적인, 방법.

[0120] EEE-B5. EEE-B4의 방법으로서, 지연, 이득 및/또는 등화 곡선은 청취자의 위치의 변화에 응답하여 조정되는, 방법.

[0121] EEE-B6. EEE-B4 또는 EEE-B5의 방법으로서, 지연, 이득 및/또는 등화 곡선은 재생 디바이스의 위치의 변화에 응답하여 조정되는, 방법.

[0122] EEE-B7. EEE-B4 내지 EEE-B6 중 어느 하나의 방법으로서, 지연, 이득 및/또는 등화 곡선은 다른 재생 디바이스들 중 하나 이상의 위치의 변화에 응답하여 조정되는, 방법.

[0123] EEE-B8. EEE-B1 내지 EEE-B7 중 어느 하나의 방법으로서, 복수의 오디오 신호들로부터, 2개 이상의 관련 오디오

신호들의 그룹의 일부가 아닌 오디오 신호를 결정하는 단계를 더 포함하는, 방법.

[0124] EEE-B9. EEE-B8의 방법으로서, 2개 이상의 관련 오디오 신호들의 그룹의 일부가 아닌 오디오 신호에 대해, 오디오 신호가 연관되는 재생 디바이스와 연관되는 지연, 이득, 및/또는 등화 곡선을 적용하는 단계를 더 포함하는, 방법.

[0125] EEE-B10. EEE-B9의 방법으로서, 2개 이상의 관련 오디오 신호들의 그룹의 일부가 아닌 오디오 신호를 독립적으로 코딩하는 단계, 및 독립적으로 코딩된 오디오 신호와, 독립적으로 코딩된 오디오 신호가 연관되는 재생 디바이스의 표시를 인코딩된 비트스트림의 별개의 독립적으로 디코딩가능한 서브세트로 결합하는 단계를 더 포함하는, 방법.

[0126] EEE-B11. EEE-B8의 방법으로서, 2개 이상의 관련 오디오 신호들의 그룹의 일부가 아닌 오디오 신호를 독립적으로 코딩하는 단계, 및 독립적으로 코딩된 오디오 신호, 독립적으로 코딩된 오디오 신호가 연관되는 재생 디바이스의 표시, 및 독립적으로 코딩된 오디오 신호가 연관되는 재생 디바이스와 연관된 지연, 이득, 및/또는 등화 곡선의 표시를 인코딩된 비트스트림의 별개의 독립적으로 디코딩가능한 서브세트로 결합하는 단계를 더 포함하는, 방법.

[0127] EEE-B12. 인코딩된 비트스트림의 프레임으로부터 재생 디바이스와 연관된 하나 이상의 오디오 신호를 디코딩하기 위한 방법 - 프레임은 인코딩된 데이터의 하나 이상의 독립 블록을 포함함 - 으로서,

[0128] 인코딩된 비트스트림으로부터, 재생 디바이스와 연관된 하나 이상의 오디오 신호에 대응하는 인코딩된 데이터의 독립 블록을 식별하는 단계;

[0129] 인코딩된 비트스트림으로부터, 식별된 인코딩된 데이터의 독립 블록을 추출하는 단계;

[0130] 추출된 독립적인 인코딩된 데이터 블록이 2개 이상의 공동 코딩된 오디오 신호들을 포함한다고 결정하는 단계;

[0131] 재생 디바이스와 연관된 하나 이상의 오디오 신호를 획득하기 위해 2개 이상의 공동 코딩된 오디오 신호들에 하나 이상의 공동 디코딩 툴을 적용하는 단계;

[0132] 추출된 인코딩된 데이터의 독립 블록으로부터, 재생 디바이스와 연관된 지연, 이득 및 등화 곡선 중 적어도 하나를 결정하는 단계;

[0133] 재생 디바이스와 연관된 지연, 이득, 및/또는 등화 곡선을 재생 디바이스와 연관된 하나 이상의 오디오 신호에 적용하는 단계를 포함하는, 방법.

[0134] EEE-B13. EEE-B12의 방법으로서, 재생 디바이스와 연관된 결정된 지연, 이득 및/또는 등화 곡선은 청취자의 위치에 대한 재생 디바이스의 위치에 의존하는, 방법.

[0135] EEE-B14. EEE-B12 또는 EEE-B13의 방법으로서, 재생 디바이스와 연관된 결정된 지연, 이득 및/또는 등화 곡선은 다른 재생 디바이스들에 대한 재생 디바이스의 위치에 의존하는, 방법.

[0136] EEE-B15. EEE-B12 내지 EEE-B14 중 어느 하나의 방법으로서, 재생 디바이스의 결정된 지연, 이득 및/또는 등화 곡선은 동적으로 가변적인, 방법.

[0137] EEE-B16. EEE-B15의 방법으로서, 재생 디바이스와 연관된 결정된 지연, 이득 및/또는 등화 곡선이 재생 디바이스와 연관된 이전에 결정된 지연, 이득 및/또는 등화 곡선과 상이할 때, 방법은 재생 디바이스와 연관된 이전에 결정된 지연, 이득 및/또는 등화 곡선과 재생 디바이스와 연관된 결정된 지연, 이득 및/또는 등화 곡선 사이를 보간하는 단계를 더 포함하는, 방법.

[0138] EEE-B17. EEE-B16의 방법으로서, 결정된 지연, 이득 및/또는 등화 곡선은 청취자의 위치의 변화로 인해 이전에 결정된 지연, 이득 및/또는 등화 곡선과 상이한, 방법.

[0139] EEE-B18. EEE-B16 또는 EEE-B17의 방법으로서, 결정된 지연, 이득 및/또는 등화 곡선은 재생 디바이스의 위치의 변화로 인해 이전에 결정된 지연, 이득 및/또는 등화 곡선과 상이한, 방법.

[0140] EEE-B19. EEE-B16 내지 EEE-B18 중 어느 하나의 방법으로서, 결정된 지연, 이득 및/또는 등화 곡선은 다른 재생 디바이스들 중 하나 이상의 위치의 변화로 인해 이전에 결정된 지연, 이득 또는 등화 곡선과 상이한, 방법.

[0141] EEE-B20. EEE-B12 내지 EEE-B19 중 어느 하나의 방법으로서, 인코딩된 비트스트림의 프레임은 인코딩된 데이터의 2개 이상의 독립 블록들을 포함하고, 방법은,

[0142] 독립 블록들 중의 하나 이상이 재생 디바이스와 연관되지 않은 오디오 신호들을 포함한다고 결정하는 단계; 및

[0143] 재생 디바이스와 연관되지 않은 오디오 신호들을 포함하는 하나 이상의 독립 블록을 무시하는 단계를 더 포함하는, 방법.

[0144] EEE-B21. EEE-B12 내지 EEE-B20 중 어느 하나의 방법으로서, 하나 이상의 공동 디코딩 툴을 적용하는 단계는 재생 디바이스와 연관되는 공동 코딩된 오디오 신호들의 서브세트를 식별하는 단계, 및 재생 디바이스와 연관된 하나 이상의 오디오 신호를 획득하기 위해 공동 코딩된 오디오 신호들의 서브세트만을 재구성하는 단계를 포함하는, 방법.

[0145] EEE-B22. EEE-B12 내지 EEE-B20 중 어느 하나의 방법으로서, 하나 이상의 공동 디코딩 툴을 적용하는 단계는 공동 코딩된 오디오 신호들 각각을 재구성하는 단계, 재생 디바이스와 연관된 재구성된 공동 코딩된 오디오 신호들의 서브세트를 식별하는 단계, 및 재생 디바이스와 연관된 재구성된 공동 코딩된 오디오 신호들의 서브세트로부터 재생 디바이스와 연관된 하나 이상의 오디오 신호를 획득하는 단계를 포함하는, 방법.

[0146] EEE-B23. EEE-B1 내지 EEE-B22 중 어느 하나의 방법을 수행하도록 구성된 장치.

[0147] EEE-B24. 실행될 때, 하나 이상의 디바이스로 하여금 EEE-B1 내지 EEE-B22 중 어느 하나의 방법을 수행하게 하는 명령어들의 시퀀스를 포함하는 비일시적 컴퓨터 판독가능 저장 매체.

[0148] EEE-C1. 복수의 오디오 신호들을 포함하는 오디오 프로그램의 인코딩된 비트스트림의 프레임을 생성하기 위한 방법 - 프레임은 인코딩된 데이터의 2개 이상의 독립 블록들을 포함함 - 으로서,

[0149] 복수의 오디오 신호들 중 하나 이상에 대해, 하나 이상의 오디오 신호들이 연관되는 재생 디바이스를 표시하는 정보를 수신하는 단계;

[0150] 표시된 재생 디바이스에 대해, 하나 이상의 추가적인 연관된 재생 디바이스를 표시하는 정보를 수신하는 단계;

[0151] 표시된 하나 이상의 추가적인 연관된 재생 디바이스와 연관된 하나 이상의 오디오 신호를 수신하는 단계;

[0152] 재생 디바이스와 연관된 하나 이상의 오디오 신호를 인코딩하는 단계;

[0153] 표시된 하나 이상의 추가적인 연관된 재생 디바이스와 연관된 하나 이상의 오디오 신호를 인코딩하는 단계;

[0154] 재생 디바이스와 연관된 하나 이상의 인코딩된 오디오 신호 및 하나 이상의 추가적인 연관된 재생 디바이스를 표시하는 시그널링 정보를 제1 독립 블록으로 결합하는 단계;

[0155] 하나 이상의 추가적인 연관된 재생 디바이스와 연관된 하나 이상의 인코딩된 오디오 신호를 하나 이상의 추가적인 독립 블록으로 결합하는 단계; 및

[0156] 제1 독립 블록 및 하나 이상의 추가적인 독립 블록을 인코딩된 비트스트림의 프레임으로 결합하는 단계를 포함하는, 방법.

[0157] EEE-C2. EEE-C1의 방법으로서, 복수의 오디오 신호들은 재생 디바이스 또는 하나 이상의 추가적인 연관된 재생 디바이스와 연관되지 않은 오디오 신호들의 하나 이상의 그룹을 포함하고,

[0158] 재생 디바이스 또는 하나 이상의 추가적인 연관된 재생 디바이스와 연관되지 않은 오디오 신호들의 하나 이상의 그룹 각각을 각각의 독립 블록으로 인코딩하는 단계; 및

[0159] 하나 이상의 그룹 각각에 대한 각각의 독립 블록을 인코딩된 비트스트림의 프레임으로 결합하는 단계를 더 포함하는, 방법.

[0160] EEE-C3. EEE-C1 또는 EEE-C2의 방법으로서, 표시된 하나 이상의 추가적인 연관된 재생 디바이스와 연관된 하나 이상의 오디오 신호는 재생 디바이스에 대한 에코-관리를 수행하기 위한 에코-참조들로서 이용되도록 특별히 의도되는, 방법.

[0161] EEE-C4. EEE-C3의 방법으로서, 에코-참조들로서 이용되도록 의도된 하나 이상의 오디오 신호는 재생 디바이스와 연관된 하나 이상의 오디오 신호보다 더 적은 데이터를 이용하여 송신되는, 방법.

[0162] EEE-C5. EEE-C3 또는 EEE-C4의 방법으로서, 하나 이상의 오디오 신호는 파라메트릭 코딩 툴들을 이용하여 인코딩된 에코-참조들로서 이용되도록 의도되는, 방법.

[0163] EEE-C6. EEE-C1 또는 EEE-C2의 방법으로서, 하나 이상의 다른 재생 디바이스와 연관된 하나 이상의 오디오 신호

는 하나 이상의 다른 재생 디바이스로부터의 재생에 적합한, 방법.

[0164] EEE-C7. 인코딩된 비트스트림의 프레임으로부터 재생 디바이스와 연관된 하나 이상의 오디오 신호를 디코딩하기 위한 방법 - 프레임은 인코딩된 데이터의 2개 이상의 독립 블록들을 포함하고, 재생 디바이스는 하나 이상의 마이크로폰을 포함함 - 으로서,

[0165] 인코딩된 비트스트림으로부터, 재생 디바이스와 연관된 하나 이상의 오디오 신호에 대응하는 인코딩된 데이터의 독립 블록을 식별하는 단계;

[0166] 인코딩된 비트스트림으로부터, 식별된 인코딩된 데이터의 독립 블록을 추출하는 단계;

[0167] 식별된 인코딩된 데이터의 독립 블록으로부터 재생 디바이스와 연관된 하나 이상의 오디오 신호를 추출하는 단계;

[0168] 인코딩된 비트스트림으로부터, 하나 이상의 다른 재생 디바이스와 연관된 하나 이상의 오디오 신호에 대응하는 인코딩된 데이터의 하나 이상의 다른 독립 블록을 식별하는 단계;

[0169] 인코딩된 데이터의 하나 이상의 다른 독립 블록으로부터, 하나 이상의 다른 재생 디바이스와 연관된 하나 이상의 오디오 신호를 추출하는 단계;

[0170] 재생 디바이스의 하나 이상의 마이크로폰을 이용하여 하나 이상의 오디오 신호를 캡처하는 단계; 및

[0171] 하나 이상의 캡처된 오디오 신호에 응답하여 재생 디바이스에 대한 에코-관리를 수행하기 위한 에코-참조들로서 하나 이상의 다른 재생 디바이스와 연관된 하나 이상의 추출된 오디오 신호를 이용하는 단계를 포함하는, 방법.

[0172] EEE-C8. EEE-C7의 방법으로서,

[0173] 인코딩된 비트스트림이 인코딩된 데이터의 하나 이상의 추가적인 독립 블록을 포함한다고 결정하는 단계; 및

[0174] 인코딩된 데이터의 하나 이상의 추가적인 독립 블록을 무시하는 단계를 더 포함하는, 방법.

[0175] EEE-C9. EEE-C8의 방법으로서, 인코딩된 데이터의 하나 이상의 추가적인 독립 블록을 무시하는 단계는 인코딩된 데이터의 추가적인 하나 이상의 독립 블록을 추출하지 않고 인코딩된 데이터의 하나 이상의 추가적인 독립 블록을 스킵하는 단계를 포함하는, 방법.

[0176] EEE-C10. EEE-C7 내지 EEE-C9 중 어느 하나의 방법으로서, 하나 이상의 다른 재생 디바이스와 연관된 하나 이상의 오디오 신호는 재생 디바이스에 대한 에코-관리를 수행하기 위한 에코-참조들로서 이용되도록 특별히 의도되는, 방법.

[0177] EEE-C11. EEE-C10의 방법으로서, 에코-참조들로서 이용되도록 특별히 의도된 하나 이상의 오디오 신호는 재생 디바이스와 연관된 하나 이상의 오디오 신호보다 더 적은 데이터를 이용하여 송신되는, 방법.

[0178] EEE-C12. EEE-C10 또는 EEE-C11의 방법으로서, 에코-참조들로서 이용되도록 특별히 의도된 하나 이상의 오디오 신호는 하나 이상의 오디오 신호의 파라메트릭 표현들로부터 재구성되는, 방법.

[0179] EEE-C13. EEE-C7 내지 EEE-C9 중 어느 하나의 방법으로서, 하나 이상의 다른 재생 디바이스와 연관된 하나 이상의 오디오 신호는 하나 이상의 다른 재생 디바이스로부터의 재생에 적합한, 방법.

[0180] EEE-C14. EEE-C7의 방법으로서, 인코딩된 신호는 재생 디바이스에 대한 에코-참조들로서 이용할 하나 이상의 다른 재생 디바이스를 표시하는 시그널링 정보를 포함하는, 방법.

[0181] EEE-C15. EEE-C14의 방법으로서, 현재 프레임에 대한 시그널링 정보에 의해 표시된 하나 이상의 다른 재생 디바이스는 이전 프레임에 대한 에코-참조들로서 이용되는 하나 이상의 다른 재생 디바이스와 상이한, 방법.

[0182] EEE-C16. EEE-C1 내지 EEE-C15 중 어느 하나의 방법을 수행하도록 구성된 장치.

[0183] EEE-C17. 실행될 때, 하나 이상의 디바이스로 하여금 EEE-C1 내지 EEE-C15 중 어느 하나의 방법을 수행하게 하는 명령어들의 시퀀스를 포함하는 비일시적 컴퓨터 판독가능 저장 매체.

[0184] EEE-D1. 오디오 신호를 송신하기 위한 방법으로서,

[0185] 비트스트림의 부분들을 포함하는 데이터의 패킷들을 생성하는 단계 - 비트스트림은 복수의 프레임들을 포함하고, 복수의 프레임들 중의 각각의 프레임은 복수의 블록들을 포함함 - 를 포함하고, 생성하는 단계는,

[0186] 복수의 블록들 중의 하나 이상의 블록으로 데이터의 패킷을 어셈블링하는 단계 - 상이한 프레임들로부터의 블록들은 단일 패킷으로 결합되고/되거나 비순차적으로 송신됨 -; 및

[0187] 패킷 기반 네트워크를 통해 데이터의 패킷들을 송신하는 단계를 포함하는, 방법.

[0188] EEE-D2. EEE-D1의 방법으로서, 복수의 블록들 중의 각각의 블록은 식별 정보를 포함하는, 방법.

[0189] EEE-D3. EEE-D2의 방법으로서, 식별 정보는 블록 ID, 블록과 연관된 대응하는 프레임 번호, 및/또는 재송신을 위한 우선순위 중 적어도 하나를 포함하는, 방법.

[0190] EEE-D4. EEE-D1 내지 EEE-D3 중 어느 하나의 방법으로서, 복수의 프레임들 중의 각각의 프레임은 시작 시간, 종료 시간, 및 지속기간을 갖는 오디오 신호의 연속 세그먼트를 나타내는 모든 오디오 데이터를 운반하는, 방법.

[0191] EEE-D5. 오디오 신호를 디코딩하기 위한 방법으로서,

[0192] 비트스트림의 부분들을 포함하는 데이터의 패킷들을 수신하는 단계 - 비트스트림은 복수의 프레임들을 포함하고, 복수의 프레임들 중의 각각의 프레임은 복수의 블록들을 포함함 -;

[0193] 디바이스에 어드레싱된 복수의 블록들 중의 블록들의 세트를 결정하는 단계; 및

[0194] 디바이스에 어드레싱된 블록들의 세트를 디코딩하고, 디바이스에 어드레싱되지 않은 복수의 블록들 중의 블록들의 디코딩을 스킵하는 단계를 포함하는, 방법.

[0195] EEE-D6. 오디오 스트림을 송신하기 위한 방법으로서,

[0196] 오디오 스트림을 송신하는 단계 - 오디오 스트림은 복수의 프레임들을 포함하고, 복수의 프레임들 중의 각각의 프레임은 복수의 블록들을 포함하고, 송신하는 단계는 오디오 스트림에 대한 구성 정보를 대역외 송신하는 단계를 포함함 - 를 포함하는, 방법.

[0197] EEE-D7. EEE-D6의 방법으로서, 오디오 스트림에 대한 구성 정보를 대역외 송신하는 단계는,

[0198] 제1 네트워크 및/또는 제1 네트워크 프로토콜을 통해 오디오 스트림을 송신하는 단계; 및

[0199] 제2 네트워크 및/또는 제2 네트워크 프로토콜을 통해 구성 정보를 송신하는 단계를 포함하는, 방법.

[0200] EEE-D8. EEE-D7의 방법으로서, 제1 네트워크 프로토콜은 사용자 데이터그램 프로토콜(UDP)이고, 제2 네트워크 프로토콜은 송신 제어 프로토콜(TCP)인, 방법.

[0201] EEE-D9. 오디오 신호를 디코딩하기 위한 방법으로서,

[0202] 비트스트림을 수신하는 단계 - 비트스트림은,

[0203] 정적 구성 양태들의 시그널링에 대응하는 정보;

[0204] 정적 메타데이터를 포함함 -; 및

[0205] 정보 및/또는 정적 메타데이터에 기초하여 하나 이상의 채널 요소를 하나 이상의 디바이스에 맵핑하는 단계를 포함하는, 방법.

[0206] EEE-D10. EEE-D9의 방법으로서, 비트스트림은 비트스트림을 디코딩하도록 구성된 복수의 디코더들에 의해 수신되고, 복수의 디코더들 중의 각각의 디코더는 비트스트림의 일부를 디코딩하도록 구성되는, 방법.

[0207] EEE-D11. EEE-D9 또는 EEE-D10의 방법으로서, 비트스트림은 동적 메타데이터를 더 포함하는, 방법.

[0208] EEE-D12. EEE-D9 내지 EEE-D11 중 어느 하나의 방법으로서, 비트스트림은 복수의 블록들을 포함하고, 복수의 블록들 중의 각각의 블록은,

[0209] 디코딩 동안 블록의 일부가 스킵될 수 있게 하는 정보 - 일부는 디바이스에 대해 필요하지 않음 -; 및

[0210] 동적 메타데이터를 포함하는, 방법.

[0211] EEE-D13. 오디오 신호의 블록들을 재송신하기 위한 방법으로서,

[0212] 비트스트림의 하나 이상의 블록을 송신하는 단계 - 비트스트림은 복수의 블록들을 포함하고, 비트스트림의 하나 이상의 블록들 각각은 이전에 송신되었음 - 를 포함하고,

[0213] 하나 이상의 블록 각각은 디코딩 우선순위 표시자를 포함하는, 방법.

[0214] EEE-D14. EEE-D13의 방법으로서, 디코딩 우선순위 표시자는 비트스트림의 하나 이상의 블록을 디코딩하기 위한 우선순위의 순서를 디코더에 표시하는, 방법.

[0215] EEE-D15. EEE-D13 또는 EEE-D14의 방법으로서, 하나 이상의 블록 중의 각각의 블록은 동일한 블록 ID를 포함하는, 방법.

[0216] EEE-D16. EEE-D13 내지 EEE-D15 중 어느 하나의 방법으로서, 비트스트림의 하나 이상의 블록의 송신은 이전 전송과 비교하여 데이터 레이트를 감소시킴으로써 송신되는, 방법.

[0217] EEE-D17. EEE-D16의 방법으로서, 데이터 레이트를 감소시키는 것은 오디오 신호의 신호 대 잡음비를 감소시키는 것, 오디오 신호의 대역폭을 감소시키는 것, 및/또는 오디오 신호의 채널 카운트를 감소시키는 것 중 적어도 하나를 포함하는, 방법.

[0218] EEE-E1. 복수의 오디오 신호들을 포함하는 오디오 프로그램의 인코딩된 비트스트림의 프레임을 생성하기 위한 방법 - 프레임은 인코딩된 데이터의 하나 이상의 독립 블록을 포함함 - 으로서,

[0219] 복수의 오디오 신호들 각각에 대해, 각각의 오디오 신호가 연관되는 재생 디바이스를 표시하는 정보를 수신하는 단계;

[0220] 각각의 재생 디바이스와 연관된 하나 이상의 오디오 신호를 인코딩하여 하나 이상의 인코딩된 오디오 신호를 획득하는 단계;

[0221] 각각의 재생 디바이스와 연관된 하나 이상의 인코딩된 오디오 신호를 프레임의 제1 독립 블록으로 결합하는 단계;

[0222] 복수의 오디오 신호들 중의 하나 이상의 다른 오디오 신호를 하나 이상의 추가적인 독립 블록으로 인코딩하는 단계; 및

[0223] 제1 독립 블록 및 하나 이상의 추가적인 독립 블록을 인코딩된 비트스트림의 프레임으로 결합하는 단계를 포함하는, 방법.

[0224] EEE-E2. EEE-E1의 방법으로서, 2개 이상의 오디오 신호들이 재생 디바이스와 연관되고, 2개 이상의 오디오 신호들 각각은 재생 디바이스의 각각의 드라이버에 의한 재생을 위해 의도된 대역제한된 신호이고, 대역제한된 신호들 각각에 대해 상이한 인코딩 기술들이 이용되는, 방법.

[0225] EEE-E3. EEE-E2의 방법으로서, 대역제한된 신호들 각각에 대해 상이한 심리음향학 모델 및/또는 상이한 비트 할당 기술이 이용되는, 방법.

[0226] EEE-E4. EEE-E1 내지 EEE-E3 중 어느 하나의 방법으로서, 인코딩된 신호의 순간 프레임 레이트는 가변적이고, 버퍼 충만도 모델(buffer fullness model)에 의해 제약된다.

[0227] EEE-E5. EEE-E1의 방법으로서, 각각의 재생 디바이스와 연관된 하나 이상의 오디오 신호를 인코딩하는 단계는 각각의 재생 디바이스와 연관된 하나 이상의 오디오 신호 및 하나 이상의 추가적인 재생 디바이스와 연관된 하나 이상의 추가적인 오디오 신호를 프레임의 제1 독립 블록으로 공동으로 인코딩하는 단계를 포함하는, 방법.

[0228] EEE-E6. EEE-E5의 방법으로서, 하나 이상의 오디오 신호 및 하나 이상의 추가적인 오디오 신호를 공동으로 인코딩하는 단계는 2개 이상의 오디오 신호들에 걸쳐 하나 이상의 스케일 인자를 공유하는 단계를 포함하는, 방법.

[0229] EEE-E7. EEE-E6의 방법으로서, 2개 이상의 오디오 신호들은 공간적으로 관련되는, 방법.

[0230] EEE-E8. EEE-E7의 방법으로서, 2개 이상의 공간적으로 관련된 오디오 신호들은 좌측 수평 채널, 좌측 최상부 채널, 우측 수평 채널, 또는 우측 최상부 채널을 포함하는, 방법.

[0231] EEE-E9. EEE-E5의 방법으로서, 하나 이상의 오디오 신호 및 하나 이상의 추가적인 오디오 신호를 공동으로 인코딩하는 단계는 결합 툴을 적용하는 단계를 포함하고, 결합 툴은,

[0232] 2개 이상의 오디오 신호들을 지정된 주파수를 초과하는 합성 신호로 결합하는 것; 및

[0233] 2개 이상의 오디오 신호들 각각에 대해, 합성 신호의 에너지와 각각의 각자의 신호의 에너지를 관련시키는 스케일 인자들을 결정하는 것을 포함하는, 방법.

[0234] EEE-E10. EEE-E5의 방법으로서, 하나 이상의 오디오 신호 및 하나 이상의 추가적인 오디오 신호를 공동으로 인코딩하는 단계는 2개보다 많은 신호들에 공동 코딩 툴을 적용하는 단계를 포함하는, 방법.

[0235] EEE-E11. 인코딩된 비트스트림의 프레임으로부터 재생 디바이스와 연관된 하나 이상의 오디오 신호를 디코딩하기 위한 방법 - 프레임은 인코딩된 데이터의 하나 이상의 독립 블록을 포함함 - 으로서,

[0236] 인코딩된 비트스트림으로부터, 재생 디바이스와 연관된 하나 이상의 오디오 신호에 대응하는 인코딩된 데이터의 독립 블록을 식별하는 단계;

[0237] 인코딩된 비트스트림으로부터, 식별된 인코딩된 데이터의 독립 블록을 추출하는 단계;

[0238] 인코딩된 데이터의 독립 블록으로부터 재생 디바이스와 연관된 하나 이상의 오디오 신호를 디코딩하여 하나 이상의 디코딩된 오디오 신호를 획득하는 단계;

[0239] 인코딩된 비트스트림으로부터, 하나 이상의 추가적인 오디오 신호에 대응하는 인코딩된 데이터의 하나 이상의 추가적인 독립 블록을 식별하는 단계; 및

[0240] 인코딩된 데이터의 하나 이상의 추가적인 독립 블록을 디코딩 또는 스킵하는 단계를 포함하는, 방법.

[0241] EEE-E12. EEE-E11의 방법으로서, 2개 이상의 오디오 신호들이 재생 디바이스와 연관되고, 2개 이상의 오디오 신호들 각각은 재생 디바이스의 각각의 드라이버에 의한 재생을 위해 의도된 대역제한된 신호이고, 2개 이상의 오디오 신호들을 디코딩하기 위해 상이한 디코딩 기술들이 이용되는, 방법.

[0242] EEE-E13. EEE-E12의 방법으로서, 대역제한된 신호들 각각을 인코딩하기 위해 상이한 심리음향학 모델 및/또는 상이한 비트 할당 기술이 이용된, 방법.

[0243] EEE-E14. EEE-E11 또는 EEE-E13의 방법으로서, 인코딩된 비트스트림의 순간 프레임 레이트는 가변적이고, 버퍼 충만도 모델에 의해 제약되는, 방법.

[0244] EEE-E15. EEE-E11의 방법으로서, 재생 디바이스와 연관된 하나 이상의 오디오 신호를 디코딩하는 단계는 인코딩된 데이터의 독립 블록으로부터 각각의 재생 디바이스와 연관된 하나 이상의 오디오 신호 및 하나 이상의 추가적인 재생 디바이스와 연관된 하나 이상의 추가적인 오디오 신호를 공동으로 디코딩하는 단계를 포함하는, 방법.

[0245] EEE-E16. EEE-E15의 방법으로서, 하나 이상의 오디오 신호 및 하나 이상의 추가적인 오디오 신호를 공동으로 디코딩하는 단계는 2개 이상의 오디오 신호들에 걸쳐 공유되는 스케일 인자들을 추출하는 단계를 포함하는, 방법.

[0246] EEE-E17. EEE-E16의 방법으로서, 2개 이상의 오디오 신호들은 공간적으로 관련되는, 방법.

[0247] EEE-E18. EEE-E17의 방법으로서, 2개 이상의 공간적으로 관련된 오디오 신호들은 좌측 수평 채널, 좌측 최상부 채널, 우측 수평 채널, 또는 우측 최상부 채널을 포함하는, 방법.

[0248] EEE-E19. EEE-E15의 방법으로서, 하나 이상의 오디오 신호 및 하나 이상의 추가적인 오디오 신호를 공동으로 디코딩하는 단계는 결합해제 툴(decoupling tool)을 적용하는 단계를 포함하는, 방법.

[0249] EEE-E20. EEE-E19의 방법으로서, 결합해제 툴은,

[0250] 지정된 주파수 미만인 독립적으로 디코딩된 신호들을 추출하는 것;

[0251] 지정된 주파수를 초과하는 합성 신호를 추출하는 것;

[0252] 합성 신호로부터 지정된 주파수를 초과하는 각각의 결합해제된 신호들 및 합성 신호의 에너지와 각각의 신호들의 에너지들을 관련시키는 스케일 인자들을 결정하는 것; 및

[0253] 공동으로 디코딩된 신호들을 획득하기 위해 각각의 독립적으로 디코딩된 신호를 각각의 결합해제된 신호와 결합하는 것을 포함하는, 방법.

[0254] EEE-E21. EEE-E15의 방법으로서, 하나 이상의 오디오 신호 및 하나 이상의 추가적인 오디오 신호를 공동 디코딩하는 단계는 2개보다 많은 오디오 신호를 추출하기 위해 공동 디코딩 툴을 적용하는 단계를 포함하는, 방법.

[0255] EEE-E22. EEE-E11의 방법으로서, 재생 디바이스와 연관된 하나 이상의 오디오 신호를 디코딩하는 단계는 오디오 신호들이 코딩된 것과 동일한 도메인에서 오디오 신호들에 대역폭 확장을 적용하는 단계를 포함하는, 방법.

[0256] EEE-E23. EEE-E22의 방법으로서, 도메인은 수정된 이산 코사인 변환(MDCT) 도메인인, 방법.

[0257] EEE-E24. EEE-E22 또는 EEE-E23의 방법으로서, 대역폭 확장은 적응적 잡음 추가를 포함하는, 방법.

[0258] EEE-E25. EEE-E1 내지 EEE-E24 중 어느 하나의 방법을 수행하도록 구성된 장치.

[0259] EEE-E26. 실행될 때, 하나 이상의 디바이스로 하여금 EEE-E1 내지 EEE-E24 중 어느 하나의 방법을 수행하게 하는 명령어들의 시퀀스를 포함하는 비일시적 컴퓨터 판독가능 저장 매체.

[0260] EEE-F1. 하나 이상의 마이크로폰을 갖는 디바이스에 의해 수행되는, 인코딩된 비트스트림을 생성하기 위한 방법으로서,

[0261] 하나 이상의 마이크로폰에 의해, 하나 이상의 오디오 신호를 캡처하는 단계;

[0262] 캡처된 오디오 신호들을 분석하여 웨이크 워드의 존재를 결정하는 단계;

[0263] 웨이크 워드의 존재를 검출 시에,

[0264] 캡처된 오디오 신호들에 대해 음성 인식 작업이 수행되어야 함을 나타내는 플래그를 설정하는 단계;

[0265] 캡처된 오디오 신호들을 인코딩하는 단계;

[0266] 인코딩된 오디오 신호들 및 플래그를 인코딩된 비트스트림으로 어셈블링하는 단계를 포함하는, 방법.

[0267] EEE-F2. EEE-F1의 방법으로서, 하나 이상의 마이크로폰은 모노 또는 공간적 사운드필드를 캡처하도록 구성되는, 방법.

[0268] EEE-F3. EEE-F2의 방법으로서, 공간 사운드필드는 A-포맷 또는 B-포맷인, 방법.

[0269] EEE-F4. EEE-F1 내지 EEE-F3 중 어느 하나의 방법으로서, 캡처된 오디오 신호들은 음성 인식 작업을 수행하는 데에만 이용되도록 의도되는, 방법.

[0270] EEE-F5. EEE-F4의 방법으로서, 캡처된 오디오 신호들은, 캡처된 오디오 신호들이 디코딩될 때, 디코딩된 오디오 신호들의 품질이 음성 인식 작업을 수행하기에는 충분하지만 인간 청취에는 충분하지 않도록 인코딩되는, 방법.

[0271] EEE-F6. EEE-F4 또는 EEE-F5의 방법으로서, 캡처된 오디오 신호들은 캡처된 오디오 신호들을 인코딩하기 전에 대역 에너지들, 멜-주파수 켑스트럴 계수들, 또는 수정된 이산 코사인 변환(MDCT) 스펙트럼 계수들 중 하나 이상을 포함하는 표현으로 변환되는, 방법.

[0272] EEE-F7. EEE-F1 내지 EEE-F3 중 어느 하나의 방법으로서, 캡처된 오디오 신호들은 인간 청취 및 음성 인식 작업을 수행하는 데 이용하는 것 둘 다를 위해 의도되는, 방법.

[0273] EEE-F8. EEE-F7의 방법으로서, 캡처된 오디오 신호들은, 캡처된 오디오 신호들이 디코딩될 때, 디코딩된 오디오 신호들의 품질이 인간 청취를 위해 충분하도록 인코딩되는, 방법.

[0274] EEE-F9. EEE-F7의 방법으로서, 캡처된 오디오 신호들을 인코딩하는 단계는 캡처된 오디오 신호의 제1 인코딩된 표현 및 캡처된 오디오 신호의 제2 인코딩된 표현을 생성하는 단계를 포함하고, 제1 인코딩된 표현은 캡처된 오디오 신호들이 제1 인코딩된 표현으로부터 디코딩될 때, 디코딩된 오디오 신호들의 품질이 인간 청취를 위해 충분하도록 생성되고, 제2 인코딩된 표현은 캡처된 오디오 신호 신호들이 제2 인코딩된 표현으로부터 디코딩될 때, 디코딩된 오디오 신호들의 품질이 음성 인식 작업을 수행하기에는 충분하지만 인간 청취에는 충분하지 않도록 생성되는, 방법.

[0275] EEE-F10. EEE-F9의 방법으로서, 캡처된 오디오 신호들의 제2 인코딩된 표현을 생성하는 단계는 캡처된 오디오 신호들을 인코딩하기 전에, 캡처된 오디오 신호들을 파라메트릭 표현, 거친 파형 표현, 또는 대역 에너지들, 멜-주파수 켑스트럴 계수들, 또는 수정된 이산 코사인 변환(MDCT) 스펙트럼 계수들 중 하나 이상을 포함하는 표현 중 하나 이상으로 변환하는 단계를 포함하는, 방법.

[0276] EEE-F11. EEE-F9 또는 EEE-F10의 방법으로서, 인코딩된 오디오 신호들을 비트스트림으로 어셈블링하는 단계는 제1 인코딩된 표현을 인코딩된 비트스트림의 제1 독립 블록에 삽입하는 단계, 및 제2 인코딩된 표현을 인코딩된 비트스트림의 제2 독립 블록에 삽입하는 단계를 포함하는, 방법.

[0277] EEE-F12. EEE-F9 또는 EEE-F10의 방법으로서, 제1 인코딩된 표현은 인코딩된 비트스트림의 제1 계층에 포함되고, 제2 인코딩된 표현은 인코딩된 비트스트림의 제2 계층에 포함되며, 제1 및 제2 계층들은 인코딩된 비

트스트림의 단일 블록에 포함되는, 방법.

[0278] EEE-F13. EEE-F1 내지 EEE-F12 중 어느 하나의 방법으로서, 웨이크 워드의 존재가 검출되지 않을 때,

[0279] 캡처된 오디오 신호들에 대해 음성 인식 작업이 수행되지 않아야 함을 나타내도록 플래그를 설정하는 단계;

[0280] 캡처된 오디오 신호들을 인코딩하는 단계;

[0281] 인코딩된 오디오 신호들 및 플래그를 인코딩된 비트스트림으로 어셈블링하는 단계를 더 포함하는, 방법.

[0282] EEE-F14. 오디오 신호를 디코딩하기 위한 방법으로서,

[0283] 인코딩된 오디오 신호들 및 음성 인식 작업이 수행되어야 하는지를 나타내는 플래그를 포함하는 인코딩된 비트스트림을 수신하는 단계;

[0284] 인코딩된 오디오 신호들을 디코딩하여 디코딩된 오디오 신호들을 획득하는 단계; 및

[0285] 플래그가 음성 인식 작업이 수행되어야 함을 나타낼 때, 디코딩된 오디오 신호들에 대해 음성 인식 작업을 수행하는 단계를 포함하는, 방법.

[0286] EEE-F15. EEE-F14의 방법으로서, 디코딩된 오디오 신호들은 음성 인식 작업을 수행하는데만 이용되도록 의도되는, 방법.

[0287] EEE-F16. EEE-F15의 방법으로서, 디코딩된 오디오 신호들의 품질은 음성 인식 작업을 수행하기에는 충분하지만 인간 청취에는 충분하지 않은, 방법.

[0288] EEE-F17. EEE-F15 또는 EEE-F16의 방법으로서, 디코딩된 오디오 신호들은 캡처된 오디오 신호들을 인코딩하기 전에 대역 에너지들, 멜-주파수 켑스트럴 계수들, 또는 수정된 이산 코사인 변환(MDCT) 스펙트럼 계수들 중 하나 이상을 포함하는 표현으로 되어 있는, 방법.

[0289] EEE-F18. EEE-F14의 방법으로서, 캡처된 오디오 신호들은, 캡처된 오디오 신호들이 디코딩될 때, 디코딩된 오디오 신호들의 품질이 인간 청취에 충분하도록 인코딩되는, 방법.

[0290] EEE-F19. EEE-F18의 방법으로서, 인코딩된 오디오 신호들은 하나 이상의 오디오 신호의 제1 인코딩된 표현 및 하나 이상의 오디오 신호의 제2 인코딩된 표현을 포함하는, 방법.

[0291] EEE-F20. EEE-F18의 방법으로서, 제1 표현으로부터 디코딩된 오디오 신호들의 품질은 인간 청취에 충분하고, 제2 표현으로부터 디코딩된 오디오 신호들의 품질은 음성 인식 작업을 수행하기에는 충분하지만 인간 청취에는 충분하지 않은, 방법.

[0292] EEE-F21. EEE-F19 또는 EEE-F20의 방법으로서, 제1 표현은 인코딩된 비트스트림의 제1 독립 블록에 있고, 제2 표현은 인코딩된 비트스트림의 제2 독립 블록에 있는, 방법.

[0293] EEE-F22. EEE-F19 또는 EEE-F20의 방법으로서, 제1 표현은 인코딩된 비트스트림의 제1 계층에 있고, 제2 인코딩된 표현은 인코딩된 비트스트림의 제2 계층에 포함되며, 제1 및 제2 계층들은 인코딩된 비트스트림의 단일 블록에 포함되는, 방법.

[0294] EEE-F23. EEE-F18 내지 EEE-F22 중 어느 하나의 방법으로서, 인코딩된 오디오 신호들을 디코딩하는 단계는 제2 표현만을 디코딩하고, 제1 표현을 무시하는 단계를 포함하는, 방법.

[0295] EEE-F24. EEE-F18 내지 EEE-F23 중 어느 하나의 방법으로서, 제2 인코딩된 표현으로부터 디코딩된 오디오 신호들은 파라메트릭 표현, 파형 표현, 또는 대역 에너지들, 멜-주파수 켑스트럴 계수들, 또는 수정된 이산 코사인 변환(MDCT) 스펙트럼 계수들 중 하나 이상을 포함하는 표현으로 되어 있는, 방법.

[0296] EEE-F25. EEE-F1 내지 EEE-F24 중 어느 하나의 방법을 수행하도록 구성된 장치.

[0297] EEE-F26. 실행될 때, 하나 이상의 디바이스로 하여금 EEE-F1 내지 EEE-F24 중 어느 하나의 방법을 수행하게 하는 명령어들의 시퀀스를 포함하는 비일시적 컴퓨터 판독가능 저장 매체.

[0298] EEE-G1. 하나 이상의 재생 디바이스로의 낮은 레이턴시 송신을 위해 몰입형 오디오 프로그램의 오디오 신호들을 인코딩하기 위한 방법으로서,

[0299] 몰입형 오디오 프로그램의 복수의 시간 도메인 오디오 신호들을 수신하는 단계;

[0300] 프레임 크기를 선택하는 단계;

[0301] 프레임 크기에 응답하여 시간 도메인 오디오 신호들의 프레임을 추출하는 단계 - 시간 도메인 오디오 신호들의 프레임은 시간 도메인 오디오 신호들의 이전 프레임과 중첩됨 -;

[0302] 오디오 신호들을 중첩 프레임들로 분할하는 단계;

[0303] 시간 도메인 오디오 신호들의 프레임을 주파수 도메인 신호들로 변환하는 단계;

[0304] 주파수 도메인 신호들을 코딩하는 단계;

[0305] 지각적으로 동기부여된 양자화 툴을 이용하여 코딩된 주파수 도메인 신호들을 양자화하는 단계;

[0306] 양자화되고 코딩된 주파수 도메인 신호들을 프레임 내의 하나 이상의 독립 블록으로 어셈블링하는 단계; 및

[0307] 하나 이상의 독립 블록을 인코딩된 프레임으로 어셈블링하는 단계를 포함하는, 방법.

[0308] EEE-G2. EEE-G1의 방법으로서, 복수의 오디오 신호들은 정의된 채널 구성을 갖는 채널 기반 신호들을 포함하는, 방법.

[0309] EEE-G3. EEE-G2의 방법으로서, 채널 구성은 모노, 스테레오, 5.1, 5.1.2, 5.1.4, 7.1.2, 7.1.4, 9.1.6, 또는 22.2 중 하나인, 방법.

[0310] EEE-G4. EEE-G1 내지 EEE-G3 중 어느 하나의 방법으로서, 복수의 오디오 신호들은 하나 이상의 객체 기반 신호를 포함하는, 방법.

[0311] EEE-G5. EEE-G1 내지 EEE-G4 중 어느 하나의 방법으로서, 복수의 오디오 신호들은 몰입형 오디오 프로그램의 장면 기반 표현을 포함하는, 방법.

[0312] EEE-G6. EEE-G1 내지 EEE-G5 중 어느 하나의 방법으로서, 선택된 프레임 크기는 128, 256, 512, 1024, 120, 240, 480, 또는 960 샘플들 중 하나인, 방법.

[0313] EEE-G7. EEE-G1 내지 EEE-G6 중 어느 하나의 방법으로서, 시간 도메인 오디오 신호들의 프레임과 시간 도메인 오디오 신호의 이전 프레임 사이의 중첩은 50% 이하인, 방법.

[0314] EEE-G8. EEE-G1 내지 EEE-G7 중 어느 하나의 방법으로서, 변환은 수정된 이산 코사인 변환(MDCT)인, 방법.

[0315] EEE-G9. EEE-G1 내지 EEE-G8 중 어느 하나의 방법으로서, 복수의 오디오 신호들 중 2개 이상은 공동으로 코딩되는, 방법.

[0316] EEE-G10. EEE-G1 내지 EEE-G9 중 어느 하나의 방법으로서, 각각의 독립 블록은 하나 이상의 재생 디바이스에 대한 인코딩된 신호들을 포함하는, 방법.

[0317] EEE-G11. EEE-G1 내지 EEE-G10 중 어느 하나의 방법으로서, 적어도 하나의 독립 블록은 2개 이상의 재생 디바이스에 대한 인코딩된 신호들을 포함하고, 인코딩된 신호들은 공동 코딩된 오디오 신호들을 포함하는, 방법.

[0318] EEE-G12. EEE-G1 내지 EEE-G11 중 어느 하나의 방법으로서, 적어도 하나의 독립 블록은 재생 디바이스의 상이한 드라이버들로부터의 재생을 위해 의도된 상이한 대역폭들을 커버하는 복수의 인코딩된 신호들을 포함하는, 방법.

[0319] EEE-G13. EEE-G1 내지 EEE-G12 중 어느 하나의 방법으로서, 적어도 하나의 독립 블록은 재생 디바이스에 의해 수행되는 에코-관리에 이용하기 위한 인코딩된 에코-참조 신호를 포함하는, 방법.

[0320] EEE-G14. EEE-G1 내지 EEE-G13 중 어느 하나의 방법으로서, 양자화된 주파수 도메인 신호들을 코딩하는 단계는 다음과 같은 툴들, 즉, 시간 잡음 성형(TNS), 공동 채널 코딩, 신호들에 걸친 스케일 인자들의 공유, 고주파수 재구성을 위한 제어 파라미터들의 결정, 및 잡음 치환을 위한 제어 파라미터들의 결정 중 하나 이상을 적용하는 단계를 포함하는, 방법.

[0321] EEE-G15. EEE-G1 내지 EEE-G14 중 어느 하나의 방법으로서, 하나 이상의 독립 블록은 재생 디바이스의 지연, 이득 및 등화 중 하나 이상을 제어하기 위한 파라미터들을 포함하는, 방법.

[0322] EEE-G16. 인코딩된 신호로부터 몰입형 오디오 프로그램의 오디오 신호들을 디코딩하기 위한 낮은 레이턴시 방법으로서,

[0323] 하나 이상의 독립 블록을 포함하는 인코딩된 프레임을 수신하는 단계;

[0324] 하나 이상의 독립 블록으로부터, 양자화되고 코딩된 주파수 도메인 신호들을 추출하는 단계;

[0325] 양자화되고 코딩된 주파수 도메인 신호들을 역 양자화하는 단계;

[0326] 역 양자화된 주파수 도메인 신호들을 디코딩하는 단계;

[0327] 디코딩된 주파수 도메인 신호들을 역 변환하여 시간 도메인 신호들을 획득하는 단계; 및

[0328] 몰입형 오디오 프로그램의 복수의 오디오 신호들을 제공하기 위해 시간 도메인 신호들을 이전 프레임으로부터의 시간 도메인 신호들과 중첩하고 가산하는 단계를 포함하는, 방법.

[0329] EEE-G17. EEE-G16의 방법으로서, 복수의 오디오 신호들은 정의된 채널 구성을 갖는 채널 기반 신호들을 포함하는, 방법.

[0330] EEE-G18. EEE-G17의 방법으로서, 채널 구성은 모노, 스테레오, 5.1, 5.1.2, 5.1.4, 7.1.2, 7.1.4, 9.1.6, 또는 22.2 중 하나인, 방법.

[0331] EEE-G19. EEE-G16 내지 EEE-G18 중 어느 하나의 방법으로서, 복수의 오디오 신호들은 하나 이상의 객체 기반 신호를 포함하는, 방법.

[0332] EEE-G20. EEE-G16 내지 EEE-G19 중 어느 하나의 방법으로서, 복수의 오디오 신호들은 몰입형 오디오 프로그램의 장면 기반 표현을 포함하는, 방법.

[0333] EEE-G21. EEE-G16 내지 EEE-G20 중 어느 하나의 방법으로서, 시간 도메인 샘플들의 프레임은 128, 256, 512, 1024, 120, 240, 480, 또는 960 샘플들 중 하나를 포함하는, 방법.

[0334] EEE-G22. EEE-G16 내지 EEE-G21 중 어느 하나의 방법으로서, 이전 프레임과의 중첩은 50% 이하인, 방법.

[0335] EEE-G23. EEE-G16 내지 EEE-G22 중 어느 하나의 방법으로서, 역 변환은 역 수정된 이산 코사인 변환(IMDCT)인, 방법.

[0336] EEE-G24. EEE-G16 내지 EEE-G23 중 어느 하나의 방법으로서, 각각의 독립 블록은 하나 이상의 재생 디바이스에 대한 양자화되고 코딩된 주파수 도메인 신호들을 포함하는, 방법.

[0337] EEE-G25. EEE-G16 내지 EEE-G24 중 어느 하나의 방법으로서, 적어도 하나의 독립 블록은 2개 이상의 재생 디바이스들을 위한 양자화되고 코딩된 주파수 도메인 신호들을 포함하고, 양자화되고 코딩된 주파수 도메인 신호들은 공동 코딩된 오디오 신호들인, 방법.

[0338] EEE-G26. EEE-G16 내지 EEE-G25 중 어느 하나의 방법으로서, 적어도 하나의 독립 블록은 재생 디바이스의 상이한 드라이버들로부터의 재생을 위해 의도된 상이한 대역폭들을 커버하는 복수의 양자화되고 코딩된 주파수 도메인 신호들을 포함하는, 방법.

[0339] EEE-G27. EEE-G16 내지 EEE-G26 중 어느 하나의 방법으로서, 적어도 하나의 독립 블록은 재생 디바이스에 의해 수행되는 에코-관리에 이용하기 위한 인코딩된 에코-참조 신호를 포함하는, 방법.

[0340] EEE-G28. EEE-G16 내지 EEE-G27 중 어느 하나의 방법으로서, 역 양자화된 주파수 도메인 신호들을 디코딩하는 단계는 다음과 같은 디코딩 툴들, 즉, 시간 잡음 성형(TNS), 공동 채널 디코딩, 신호에 걸친 스케일 인자들의 공유, 고주파수 재구성, 및 잡음 치환 중 하나 이상을 적용하는 단계를 포함하는, 방법.

[0341] EEE-G29. EEE-G16 내지 EEE-G28 중 어느 하나의 방법으로서, 하나 이상의 독립 블록은 재생 디바이스의 지연, 이득 및 등화 중 하나 이상을 제어하기 위한 파라미터들을 포함하는, 방법.

[0342] EEE-G30. EEE-G16 내지 EEE-G29 중 어느 하나의 방법으로서, 방법은 재생 디바이스에 의해 수행되고, 하나 이상의 독립 블록으로부터 양자화되고 코딩된 신호들을 추출하는 단계는 재생 디바이스에 의한 재생을 위해 양자화되고 코딩된 주파수 도메인 신호들을 포함하는 블록들만을 선택하는 단계, 및 다른 재생 디바이스들에 의한 재생을 위해 양자화되고 코딩된 주파수 도메인 신호들을 포함하는 독립 블록들을 무시하는 단계를 포함하는, 방법.

[0343] EEE-G31. EEE-G1 내지 EEE-G30 중 어느 하나의 방법을 수행하도록 구성된 장치.

[0344] EEE-G32. 실행될 때, 하나 이상의 디바이스로 하여금 EEE-G1 내지 EEE-G30 중 어느 하나의 방법을 수행하게 하

는 명령어들의 시퀀스를 포함하는 비일시적 컴퓨터 판독가능 저장 매체.

[0345] 아래의 청구항들 및 본 명세서의 설명에서, 용어들 포함하는(comprising), 구성되는(comprised of), 또는 포함하는(which comprises) 중 어느 하나는 적어도 뒤따르는 요소들/특징들을 포함하지만, 다른 것들을 배제하지 않는 것을 의미하는 개방 용어이다. 따라서, "포함하는"이라는 용어는, 청구항들에서 이용될 때, 그 후에 열거되는 수단들 또는 요소들, 또는 단계들로 제한되는 것으로 해석되어서는 안된다. 예를 들어, A 및 B를 포함하는 디바이스의 표현의 범위는 요소들 A 및 B로만 구성되는 디바이스들로 제한되지 않아야 한다. 본 명세서에서 이용되는 바와 같은 용어들 포함하는(including) 또는 포함하는(which includes) 또는 포함하는(that includes) 중 어느 하나는 또한 적어도 그 용어를 따르는 요소들/특징들을 포함하는 것을 또한 의미하지만, 다른 것들을 배제하지 않는 개방 용어이다. 따라서, 포함하는(including)은 포함하는(comprising)과 동의어이고 이를 의미한다.

[0346] 본 발명의 예들의 상기의 설명에서, 개시를 간소화하고 다양한 본 발명의 양태들 중 하나 이상의 이해를 돕기 위해 다양한 특징들이 때때로 단일 예, 도면, 또는 그 설명에서 함께 그룹화된다는 것을 알아야 한다. 그러나, 이러한 개시내용의 방법은 각각의 청구항에서 명시적으로 인용되는 것보다 더 많은 특징들이 요구된다는 의도를 반영하는 것으로서 해석되지 않아야 한다. 오히려, 이하의 청구항들이 반영하는 바와 같이, 본 발명의 양태들은 단일의 전술한 개시된 예의 모든 특징들보다 적은 것에 있다. 따라서, 상세한 설명에 후속하는 청구항들은 이로써 본 상세한 설명에 명시적으로 포함되고, 각각의 청구항은 그 자체로 본 발명의 개별적인 예로서 존재한다.

[0347] 더욱이, 본 명세서에 설명된 일부 예들이 다른 예들에 포함된 다른 특징들이 아니라 일부를 포함하지만, 본 기술분야의 통상의 기술자라면 잘 알 것인 바와 같이, 상이한 예들의 특징들의 조합들이 포괄되고 상이한 예들을 형성하는 것으로 보아야 한다. 예를 들어, 이하의 청구항들에서, 청구된 예들 중 임의의 것이 임의의 조합으로 이용될 수 있다.

[0348] 더욱이, 예들 중 일부는 컴퓨터 시스템의 프로세서에 의해 또는 기능을 수행하는 다른 수단에 의해 구현될 수 있는 방법 또는 방법의 요소들의 조합으로서 본 명세서에 설명된다. 따라서, 그러한 방법 또는 방법의 요소를 수행하기 위한 필요한 명령어들을 갖는 프로세서는 방법 또는 방법의 요소를 수행하기 위한 수단을 형성한다. 또한, 장치의 본 명세서에서 설명된 요소는 요소에 의해 수행된 기능을 수행하기 위한 수단의 예이다.

[0349] 더욱이, 본 명세서에 설명되는 예들 중 일부는 유선 및/또는 무선 시스템과 같은 분배 및/또는 송신 시스템에서 가능하게는 구현되는 것으로 해석되어야만 하는 연결된 해결책들을 개시하고 있다. 예를 들어, 3G, 4G는 물론 5G와 같은, 임의의 전기, 광학, 및/또는 모바일 시스템의 이용에 의해.

[0350] 따라서, 본 발명의 구체적인 예들이 설명되었지만, 본 기술분야의 통상의 기술자들은 다른 그리고 추가의 수정들이 이루어질 수 있다는 것을 인식할 것이고, 모든 그러한 변경들 및 수정들을 청구하도록 의도된다. 예를 들어, 위에 주어진 임의의 공식들은 이용될 수 있는 절차들을 대표할 뿐이다. 기능성은 블록도들로부터 추가되거나 삭제될 수 있고, 동작들은 기능 블록들 사이에서 교환될 수 있다. 단계들은 설명된 방법들에 추가되거나 삭제될 수 있다.

[0351] 위에 개시된 시스템들, 디바이스들, 및 방법들은 소프트웨어, 펌웨어, 하드웨어, 또는 이들의 조합으로서 구현될 수 있다. 예를 들어, 본 출원의 양태들은 디바이스, 하나보다 많은 디바이스를 포함하는 시스템, 방법, 컴퓨터 프로그램 제품 등에서 적어도 부분적으로 구현될 수 있다.

[0352] 하드웨어 구현에서, 상기의 설명에서 언급된 기능 유닛들 사이의 작업들의 분할은 반드시 물리적 유닛들로의 분할에 대응하지는 않고; 반대로, 하나의 물리적 컴포넌트는 다수의 기능들을 가질 수 있고, 하나의 작업은 협력하여 여러 물리적 컴포넌트들에 의해 수행될 수 있다.

[0353] 특정 컴포넌트들 또는 모든 컴포넌트들은 디지털 신호 프로세서 또는 마이크로프로세서에 의해 실행되는 소프트웨어로서 구현되거나, 하드웨어로서 또는 주문형 집적 회로로서 구현될 수 있다. 그러한 소프트웨어는 컴퓨터 저장 매체(또는 비일시적 매체) 및 통신 매체(또는 일시적 매체)를 포함할 수 있는 컴퓨터 판독가능 매체 상에 분산될 수 있다.

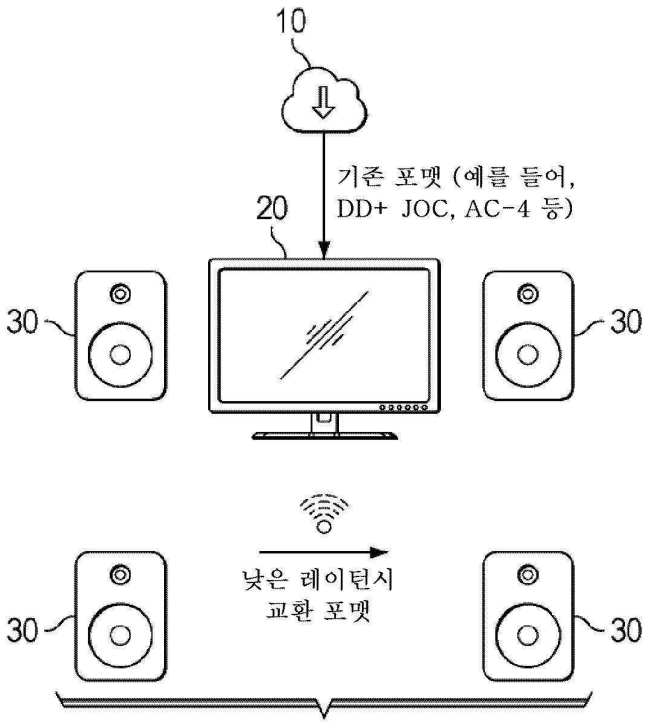
[0354] 본 기술분야의 통상의 기술자에게 잘 알려진 바와 같이, 용어 컴퓨터 저장 매체는 컴퓨터 판독가능 명령어들, 데이터 구조들, 프로그램 모듈들, 또는 다른 데이터와 같은 정보의 저장을 위한 임의의 방법 또는 기술로 구현되는 휘발성 및 비휘발성, 이동식 및 비이동식 매체 둘 다를 포함한다. 컴퓨터 저장 매체는 RAM, ROM, EEPROM, 플래시 메모리 또는 기타 메모리 기술, CD-ROM, DVD(digital versatile disk) 또는 다른 광학 디스크 저장소,

자기 카세트, 자기 테이프, 자기 디스크 저장소 또는 다른 자기 저장 디바이스, 또는 원하는 정보를 저장하는 데 이용될 수 있고 컴퓨터에 의해 액세스될 수 있는 임의의 다른 매체를 포함하지만 이에 제한되는 것은 아니다.

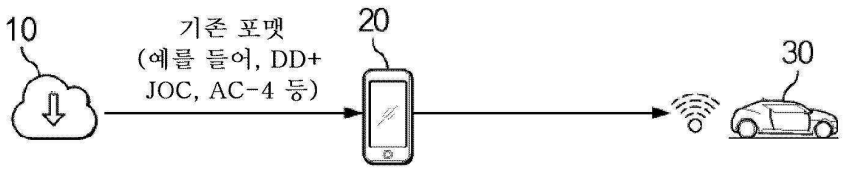
[0355] 또한, 통신 매체는 전형적으로 컴퓨터 판독가능 명령어들, 데이터 구조들, 프로그램 모듈들, 또는 다른 데이터를 반송파 또는 다른 전송 메커니즘과 같은 변조된 데이터 신호로 구현하고 임의의 정보 전달 매체를 포함한다는 것이 통상의 기술자에게 잘 알려져 있다.

도면

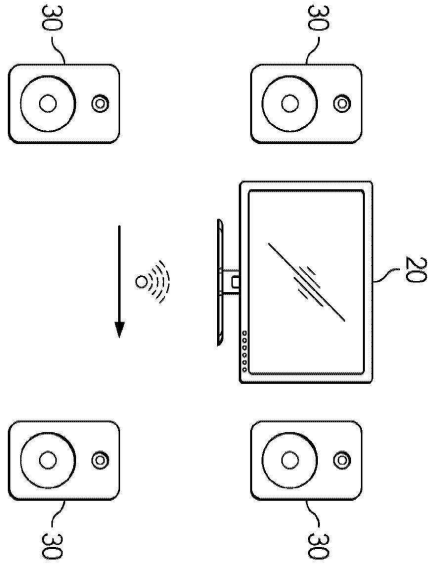
도면1

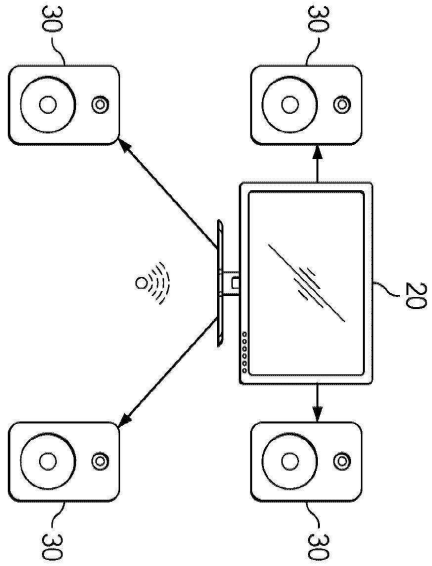


도면2

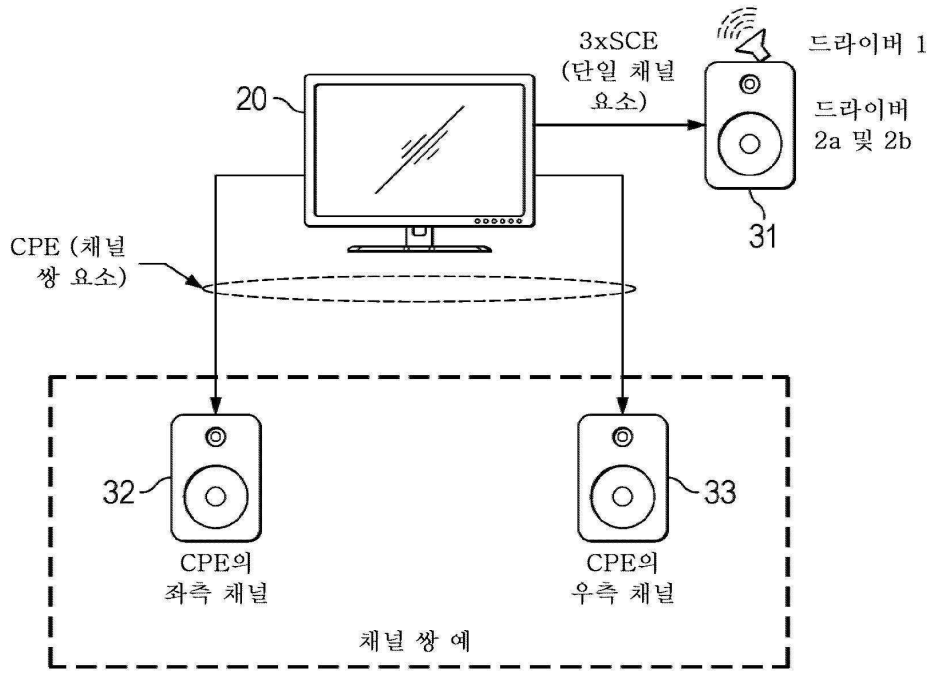


도면3

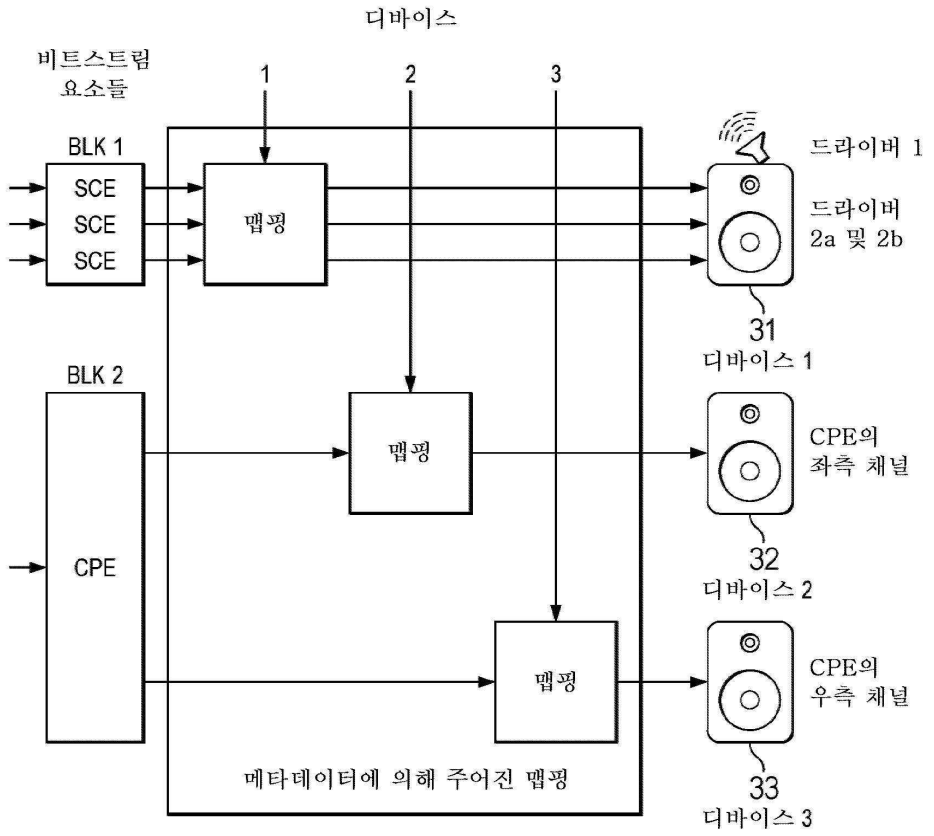




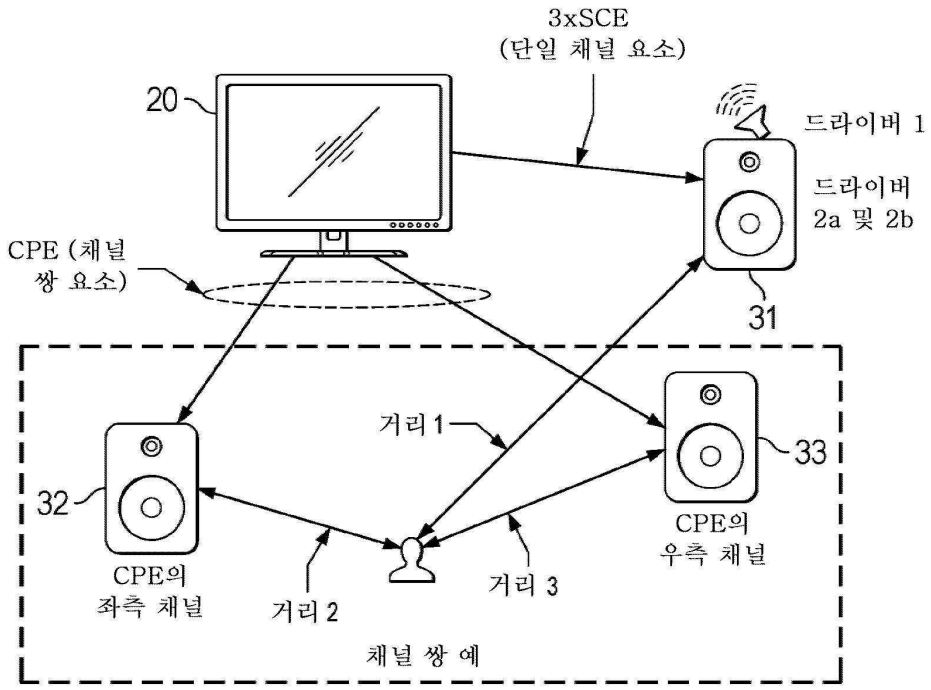
도면4


3xSCE
(단일 채널
요소)
드라이버 1
20
드라이버
2a 및 2b
31
CPE (채널
쌍 요소)
32
33
CPE의
좌측 채널
CPE의
우측 채널
채널 쌍 예

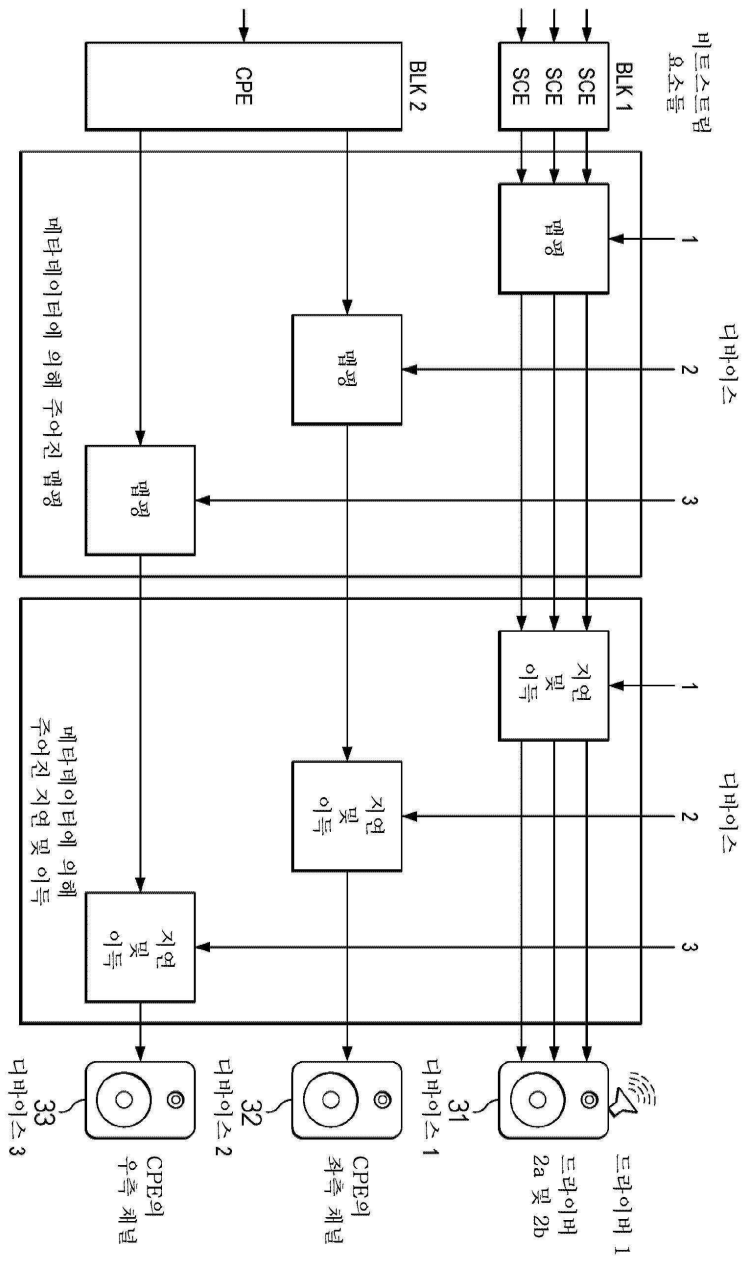
도면5


디바이스
비트스트림
요소들
1
2
3
BLK 1
SCE
SCE
SCE
맵핑
드라이버 1
드라이버
2a 및 2b
31
디바이스 1
BLK 2
맵핑
CPE의
좌측 채널
CPE
32
디바이스 2
맵핑
CPE의
우측 채널
메타데이터에 의해 주어진 맵핑
33
디바이스 3

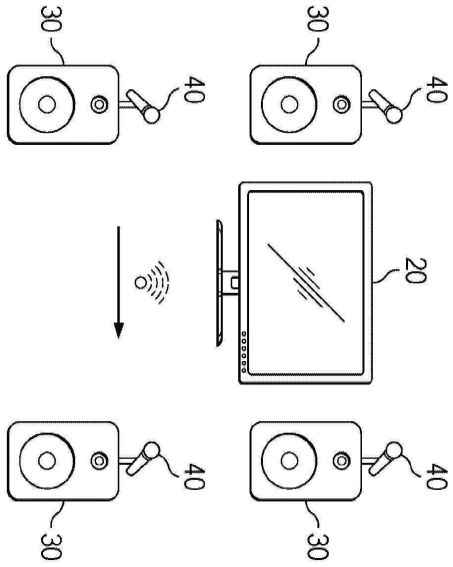
도면6

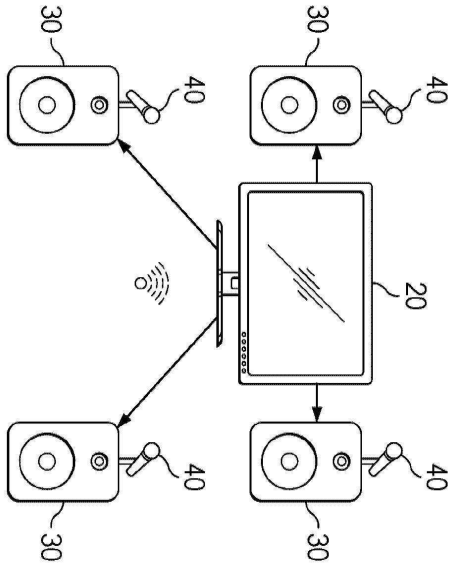

3xSCE
(단일 채널 요소)
20
드라이버 1
드라이버
2a 및 2b
CPE (채널
쌍 요소)
31
거리 1
32
33
CPE의
좌측 채널
CPE의
우측 채널
거리 2
거리 3
채널 쌍 예

도면7

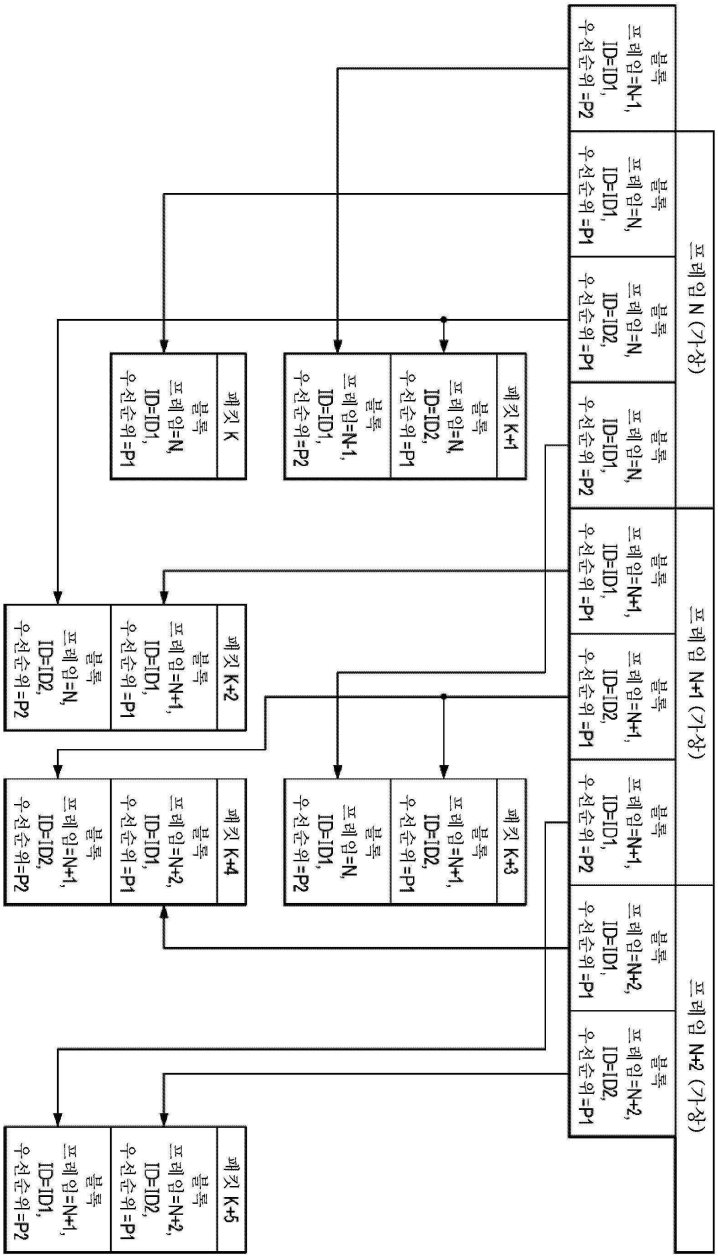

비트스트림 요소들
BLK 1
SCE
SCE
SCE
BLK 2
CPE
디바이스
1
2
3
맵핑
맵핑
맵핑
메타데이터에 의해 주어진 맵핑
디바이스
1
2
3
지연 및 이득
지연 및 이득
지연 및 이득
메타데이터에 의해 주어진 지연 및 이득
트라이버 1
트라이버 2a 및 2b
디바이스 1
31
CPE의 좌측 채널
디바이스 2
32
CPE의 우측 채널
디바이스 3
33

도면8





도면9


프레임 N (가상)
프레임 N+1 (가상)
프레임 N+2 (가상)
블록 프레임=N-1, ID=ID1, 우선순위=P2
블록 프레임=N, ID=ID1, 우선순위=P1
블록 프레임=N, ID=ID2, 우선순위=P1
블록 프레임=N, ID=ID1, 우선순위=P2
블록 프레임=N+1, ID=ID1, 우선순위=P1
블록 프레임=N+1, ID=ID2, 우선순위=P1
블록 프레임=N+1, ID=ID1, 우선순위=P2
블록 프레임=N+2, ID=ID1, 우선순위=P1
블록 프레임=N+2, ID=ID2, 우선순위=P1
패킷 K
블록 프레임=N, ID=ID1, 우선순위=P1
패킷 K+1
블록 프레임=N-1, ID=ID1, 우선순위=P2
블록 프레임=N, ID=ID2, 우선순위=P1
패킷 K+2
블록 프레임=N+1, ID=ID1, 우선순위=P1
블록 프레임=N, ID=ID2, 우선순위=P2
패킷 K+3
블록 프레임=N, ID=ID1, 우선순위=P2
블록 프레임=N+1, ID=ID2, 우선순위=P1
패킷 K+4
블록 프레임=N+2, ID=ID1, 우선순위=P1
블록 프레임=N+1, ID=ID2, 우선순위=P2
패킷 K+5
블록 프레임=N+2, ID=ID2, 우선순위=P1
블록 프레임=N+1, ID=ID1, 우선순위=P2

도면10

신택스	비트 수	연상 기호

AudioSpecificConfig()
{
audioObjectType=GetAudioObjectType();
samplingFrequencyIndex;
if (samplingFrequencyIndex == 0xf) {
samplingFrequency;
}
channelConfiguration;
	4
	24
	4
		bslbf
		uimsbf
		bslbf

도면11

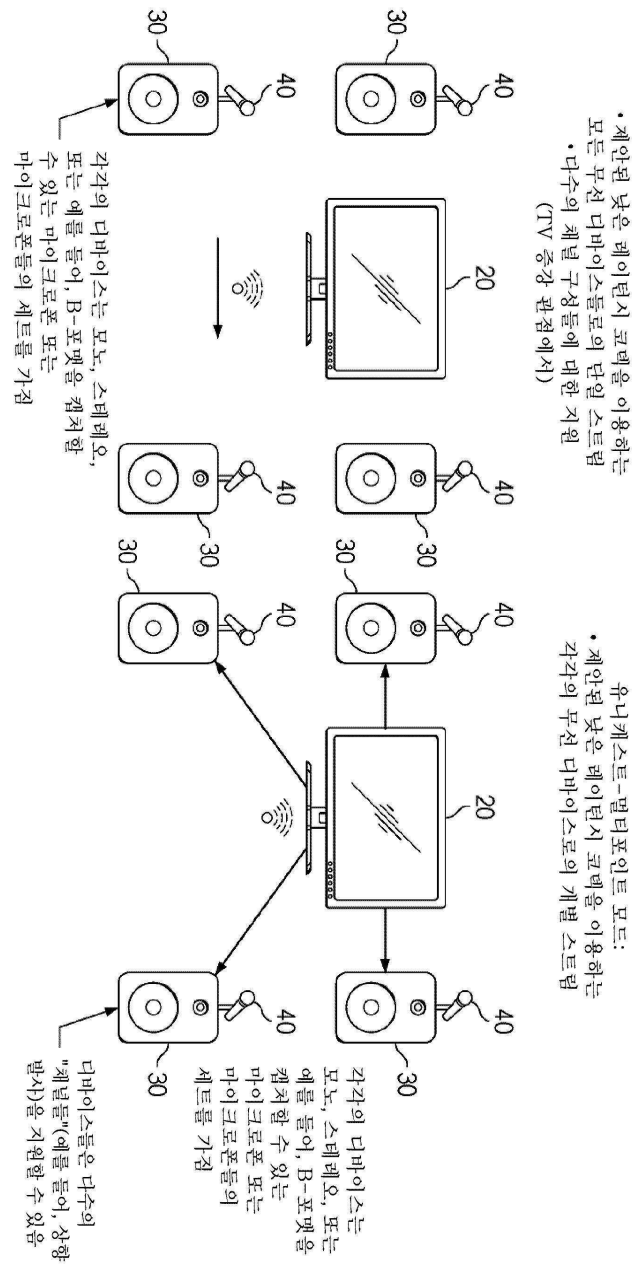
raw_data_block_codec(configData)

- MPEG-4에서의 기존의 원시 데이터 블록들과 유사함
- 주어진 디바이스에 대해 필요하지 않은 데이터 블록의 부분을 스킵가능하게 하는 추가적인 정보를 가짐
- 동적 메타데이터

신택스	비트 수	연상 기호

raw_data_block()
{
while((id= id_syn_ele) != ID_END){
switch (id) {
case ID_SCE: single_channel_element();
break;
case ID_CPE: channel_pair_element();
break;
case ID_CCE: coupling_channel_element();
break;
case ID_LFE: lffe_channel_element();
break;
case ID_DSE: data_stream_element();
break;
case ID_PCE: program_config_element();
break;
case ID_FIL: fill_element();
	3	uimsbf

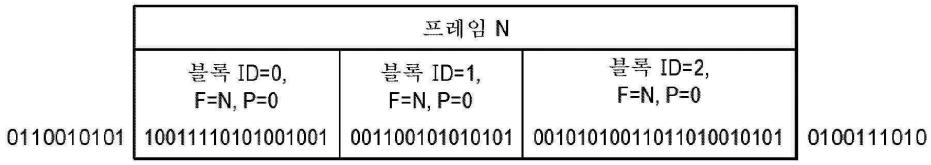
도면12


브로드캐스트 모드:
• 제안된 낮은 레이턴시 코덱을 이용하는 모든 무선 디바이스들로의 단일 스트림
• 다수의 채널 구성들에 대한 지원
(TV 증강 관점에서)
유니캐스트-멀티포인트 모드:
• 제안된 낮은 레이턴시 코덱을 이용하는 각각의 무선 디바이스로의 개별 스트림
각각의 디바이스는 모노, 스테레오, 또는 예를 들어, B-포맷을 캡처할 수 있는 마이크로폰들의 세트를 가짐
디바이스들은 다수의 "채널들"(예를 들어, 상향 방사)을 지원할 수 있음
20
30
40

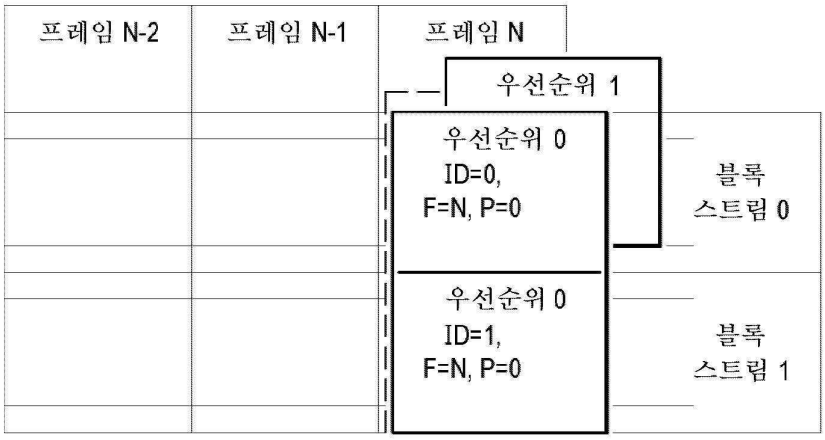
도면13



도면14



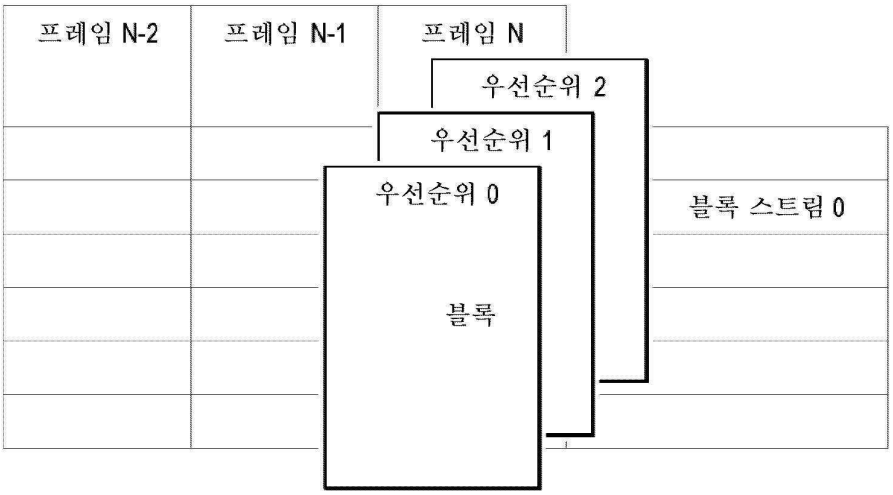
도면15



도면16



도면17


프레임 N-2
프레임 N-1
프레임 N
우선순위 2
우선순위 1
우선순위 0
블록
블록 스트림 0

도면18


프레임 N
블록 ID=0, F=N, P=0
블록 ID=0, F=N, P=1
블록 ID=0, F=N, P=2
0110010101
0010101001101101001010 1
10011110101001001
001100101010101
0100111010