(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2022년08월18일
(11) 등록번호 10-2434299
(24) 등록일자 2022년08월16일

(51) 국제특허분류(Int. Cl.)
H04N 21/845 (2011.01) *H04L 9/40* (2022.01)
H04N 19/30 (2014.01) *H04N 19/70* (2014.01)
H04N 19/85 (2014.01) *H04N 21/2343* (2011.01)
H04N 21/432 (2011.01) *H04N 21/854* (2011.01)
(52) CPC특허분류
H04N 21/8451 (2013.01)
H04L 65/75 (2022.05)
(21) 출원번호 10-2018-7033927
(22) 출원일자(국제) 2017년05월24일
심사청구일자 2020년05월07일
(85) 번역문제출일자 2018년11월22일
(65) 공개번호 10-2019-0010568
(43) 공개일자 2019년01월30일
(86) 국제출원번호 PCT/US2017/034296
(87) 국제공개번호 WO 2017/205522
국제공개일자 2017년11월30일
(30) 우선권주장
62/340,886 2016년05월24일 미국(US)
15/602,962 2017년05월23일 미국(US)
(56) 선행기술조사문헌
US20150110192 A1
KR1020150088899 A
KR1020140117659 A
KR1020160075554 A

(73) 특허권자
퀄컴 인코포레이티드
미국 92121-1714 캘리포니아주 샌 디에고 모어하우스 드라이브 5775
(72) 발명자
왕 예-쿠이
미국 92121-1714 캘리포니아주 샌디에고 모어하우스 드라이브 5775
(74) 대리인
특허법인코리아나

전체 청구항 수 : 총 18 항 심사관 : 홍기완

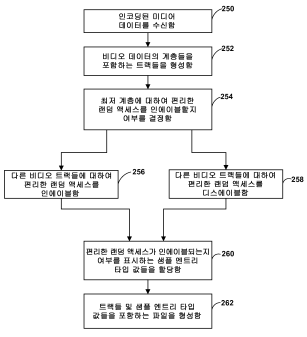
(54) 발명의 명칭 **샘플 엔트리들 및 랜덤 액세스**

(57) 요 약

일 예에서, 비디오 데이터를 취출하기 위한 디바이스는, 비디오 데이터의 최저 서브 계층을 운반하는 비디오 데이터의 기본 계층을 포함하는, 복수의 트랙들의 최저 트랙이 편리한 랜덤 액세스가 인에이블되는 것을 나타내는 샘플들에 대한 샘플 엔트리 타입 값들을 포함할 것으로 결정하는 것에 응답하여, 비디오 데이터를 포함하는 상기

(뒷면에 계속)

대 표 도 - 도7



복수의 트랙들의 다른 트랙들의 각각의 샘플들에 대한 샘플 엔트리 타입 값들을 편리한 랜덤 액세스가 인에이블되는 것을 나타내도록 세팅하고, 그리고 상기 복수의 트랙들의 트랙들에 대한 상기 샘플 엔트리 타입 값들이 편리한 랜덤 액세스가 인에이블되는 것을 나타내도록, 상기 복수의 트랙들을 포함하는 파일을 생성하도록 구성된 하나 이상의 프로세서들을 포함한다.

(52) CPC특허분류
H04N 19/30 (2015.01)
H04N 19/70 (2015.01)
H04N 19/85 (2015.01)
H04N 21/23439 (2013.01)
H04N 21/432 (2013.01)
H04N 21/8456 (2013.01)
H04N 21/85406 (2013.01)

명 세 서

청구범위

청구항 1

비디오 데이터를 포함하는 파일을 생성하는 방법으로서,

단일 계층 또는 다중 계층 비트스트림을 운반하는 복수의 트랙들의 최저 트랙이 편리한 랜덤 액세스가 인에이블되는 것을 나타내는 샘플들에 대한 샘플 엔트리 타입 값들을 포함할 것으로 결정하는 것에 응답하여, 비디오 데이터를 포함하는 상기 복수의 트랙들의 다른 트랙들의 각각의 샘플들에 대한 샘플 엔트리 타입 값들을 편리한 랜덤 액세스가 인에이블되는 것을 나타내도록 세팅하는 단계로서, 상기 최저 트랙은 상기 비디오 데이터의 최저 서브 계층을 운반하는 상기 비디오 데이터의 기본 계층을 포함하고, 상기 편리한 랜덤 액세스는 이전 샘플들로부터의 파라미터 세트들을 페치하지 않고 샘플에서 시작하는 랜덤 액세스를 포함하는, 상기 샘플 엔트리 타입 값들을 편리한 랜덤 액세스가 인에이블되는 것을 나타내도록 세팅하는 단계; 및

상기 복수의 트랙들의 트랙들에 대한 상기 샘플 엔트리 타입 값들이 편리한 랜덤 액세스가 인에이블되는 것을 나타내도록, 상기 복수의 트랙들을 포함하는 파일을 생성하는 단계를 포함하는, 비디오 데이터를 포함하는 파일을 생성하는 방법.

청구항 2

제 1 항에 있어서,

상기 최저 트랙 및 상기 다른 트랙들에 대한 상기 샘플 엔트리 타입 값들은 'hev1' 또는 'hev2' 중 하나를 포함하는, 비디오 데이터를 포함하는 파일을 생성하는 방법.

청구항 3

제 1 항에 있어서,

상기 비디오 데이터는 고효율 비디오 코딩 (HEVC) 또는 계층화된 HEVC (L-HEVC) 중 하나에 따라 인코딩되는, 비디오 데이터를 포함하는 파일을 생성하는 방법.

청구항 4

제 1 항에 있어서,

상기 샘플 엔트리 타입 값들 중 적어도 하나를 클라이언트 디바이스에 전송하는 단계;

상기 샘플 엔트리 타입 값들 중 적어도 하나를 전송한 후에, 상기 클라이언트 디바이스로부터 랜덤 액세스 샘플에 대한 요청을 수신하는 단계; 및

상기 요청에 응답하여, 상기 샘플을 기술하고 상기 샘플에 대응하는 샘플 엔트리 또는 상기 샘플 중 적어도 하나에 포함되는 파라미터 세트들을 전송하는 것을 포함하여, 상기 랜덤 액세스 샘플을 상기 클라이언트 디바이스에 전송하는 단계를 더 포함하는, 비디오 데이터를 포함하는 파일을 생성하는 방법.

청구항 5

비디오 데이터를 포함하는 파일을 생성하기 위한 디바이스로서,

비디오 데이터를 저장하도록 구성된 메모리; 및

회로부 내에 구현되는 하나 이상의 프로세서들을 포함하며,

상기 하나 이상의 프로세서들은,

단일 계층 또는 다중 계층 비트스트림을 운반하는 복수의 트랙들의 최저 트랙이 편리한 랜덤 액세스가 인에이블되는 것을 나타내는 샘플들에 대한 샘플 엔트리 타입 값들을 포함할 것으로 결정하는 것에 응답하여, 비디오 데이터를 포함하는 상기 복수의 트랙들의 다른 트랙들의 각각의 샘플들에 대한 샘플 엔트리 타입 값들을

편리한 랜덤 액세스가 인에이블되는 것을 나타내도록 세팅하는 것으로서, 상기 최저 트랙은 상기 비디오 데이터의 최저 서브 계층을 운반하는 상기 비디오 데이터의 기본 계층을 포함하고, 상기 편리한 랜덤 액세스는 이전 샘플들로부터의 파라미터 세트들을 페치하지 않고 샘플에서 시작하는 랜덤 액세스를 포함하는, 상기 샘플 엔트리 타입 값들을 편리한 랜덤 액세스가 인에이블되는 것을 나타내도록 세팅하며; 그리고

상기 복수의 트랙들의 트랙들에 대한 상기 샘플 엔트리 타입 값들이 편리한 랜덤 액세스가 인에이블되는 것을 나타내도록, 상기 복수의 트랙들을 포함하는 파일을 생성하도록

구성되는, 비디오 데이터를 포함하는 파일을 생성하기 위한 디바이스.

청구항 6

제 5 항에 있어서,

상기 최저 트랙 및 상기 다른 트랙들에 대한 상기 샘플 엔트리 타입 값들은 'hev1' 또는 'hev2' 중 하나를 포함하는, 비디오 데이터를 포함하는 파일을 생성하기 위한 디바이스.

청구항 7

제 5 항에 있어서,

상기 비디오 데이터는 고효율 비디오 코딩 (HEVC) 또는 계층화된 HEVC (L-HEVC) 중 하나에 따라 인코딩되는, 비디오 데이터를 포함하는 파일을 생성하기 위한 디바이스.

청구항 8

제 5 항에 있어서,

상기 하나 이상의 프로세서들은 추가로,

상기 샘플 엔트리 타입 값들 중 적어도 하나를 클라이언트 디바이스에 전송하고;

상기 샘플 엔트리 타입 값들 중 적어도 하나를 전송한 후에, 상기 클라이언트 디바이스로부터 랜덤 액세스 샘플에 대한 요청을 수신하고; 그리고

상기 요청에 응답하여, 상기 샘플을 기술하고 상기 샘플에 대응하는 샘플 엔트리 또는 상기 샘플 중 적어도 하나에 포함되는 파라미터 세트들을 전송하는 것을 포함하여, 상기 랜덤 액세스 샘플을 상기 클라이언트 디바이스에 전송하도록

구성되는, 비디오 데이터를 포함하는 파일을 생성하기 위한 디바이스.

청구항 9

비디오 데이터를 포함하는 파일을 생성하기 위한 디바이스로서,

단일 계층 또는 다중 계층 비트스트림을 운반하는 복수의 트랙들의 최저 트랙이 편리한 랜덤 액세스가 인에이블되는 것을 나타내는 샘플들에 대한 샘플 엔트리 타입 값들을 포함할 것으로 결정하는 것에 응답하여, 비디오 데이터를 포함하는 상기 복수의 트랙들의 다른 트랙들의 각각의 샘플들에 대한 샘플 엔트리 타입 값들을 편리한 랜덤 액세스가 인에이블되는 것을 나타내도록 세팅하는 수단으로서, 상기 최저 트랙은 상기 비디오 데이터의 최저 서브 계층을 운반하는 상기 비디오 데이터의 기본 계층을 포함하고, 상기 편리한 랜덤 액세스는 이전 샘플들로부터의 파라미터 세트들을 페치하지 않고 샘플에서 시작하는 랜덤 액세스를 포함하는, 상기 샘플 엔트리 타입 값들을 편리한 랜덤 액세스가 인에이블되는 것을 나타내도록 세팅하는 수단; 및

상기 복수의 트랙들의 트랙들에 대한 상기 샘플 엔트리 타입 값들이 편리한 랜덤 액세스가 인에이블되는 것을 나타내도록, 상기 복수의 트랙들을 포함하는 파일을 생성하는 수단을 포함하는, 비디오 데이터를 포함하는 파일을 생성하기 위한 디바이스.

청구항 10

제 9 항에 있어서,

상기 최저 트랙 및 상기 다른 트랙들에 대한 상기 샘플 엔트리 타입 값들은 'hev1' 또는 'hev2' 중 하나를 포함

하는, 비디오 데이터를 포함하는 파일을 생성하기 위한 디바이스.

청구항 11

제 9 항에 있어서,

상기 비디오 데이터는 고효율 비디오 코딩 (HEVC) 또는 계층화된 HEVC (L-HEVC) 중 하나에 따라 인코딩되는, 비디오 데이터를 포함하는 파일을 생성하기 위한 디바이스.

청구항 12

제 9 항에 있어서,

상기 샘플 엔트리 타입 값들 중 적어도 하나를 클라이언트 디바이스에 전송하는 수단;

상기 샘플 엔트리 타입 값들 중 적어도 하나를 전송한 후에, 상기 클라이언트 디바이스로부터 랜덤 액세스 샘플에 대한 요청을 수신하는 수단; 및

상기 요청에 응답하여, 상기 샘플을 기술하고 상기 샘플에 대응하는 샘플 엔트리 또는 상기 샘플 중 적어도 하나에 포함되는 파라미터 세트들을 전송하는 것을 포함하여, 상기 랜덤 액세스 샘플을 상기 클라이언트 디바이스에 전송하는 수단을 더 포함하는, 비디오 데이터를 포함하는 파일을 생성하기 위한 디바이스.

청구항 13

명령들이 저장된 컴퓨터 판독가능 저장 매체로서,

상기 명령들은 실행될 경우, 프로세서로 하여금,

단일 계층 또는 다중 계층 비트스트림을 운반하는 복수의 트랙들의 최저 트랙이 편리한 랜덤 액세스가 인에이블되는 것을 나타내는 샘플들에 대한 샘플 엔트리 타입 값들을 포함할 것으로 결정하는 것에 응답하여, 비디오 데이터를 포함하는 상기 복수의 트랙들의 다른 트랙들의 각각의 샘플들에 대한 샘플 엔트리 타입 값들을 편리한 랜덤 액세스가 인에이블되는 것을 나타내도록 세팅하게 하는 것으로서, 상기 최저 트랙은 상기 비디오 데이터의 최저 서브 계층을 운반하는 상기 비디오 데이터의 기본 계층을 포함하고, 상기 편리한 랜덤 액세스는 이전 샘플들로부터의 파라미터 세트들을 페치하지 않고 샘플에서 시작하는 랜덤 액세스를 포함하는, 상기 샘플 엔트리 타입 값들을 편리한 랜덤 액세스가 인에이블되는 것을 나타내도록 세팅하게 하고; 그리고

상기 복수의 트랙들의 트랙들에 대한 상기 샘플 엔트리 타입 값들이 편리한 랜덤 액세스가 인에이블되는 것을 나타내도록, 상기 복수의 트랙들을 포함하는 파일을 생성하게 하는, 컴퓨터 판독가능 저장 매체.

청구항 14

제 13 항에 있어서,

상기 최저 트랙 및 상기 다른 트랙들에 대한 상기 샘플 엔트리 타입 값들은 'hev1' 또는 'hev2' 중 하나를 포함하는, 컴퓨터 판독가능 저장 매체.

청구항 15

제 13 항에 있어서,

상기 비디오 데이터는 고효율 비디오 코딩 (HEVC) 또는 계층화된 HEVC (L-HEVC) 중 하나에 따라 인코딩되는, 컴퓨터 판독가능 저장 매체.

청구항 16

제 13 항에 있어서,

상기 프로세서로 하여금,

상기 샘플 엔트리 타입 값들 중 적어도 하나를 클라이언트 디바이스에 전송하게 하고;

상기 샘플 엔트리 타입 값들 중 적어도 하나를 전송한 후에, 상기 클라이언트 디바이스로부터 랜덤 액세스 샘플에 대한 요청을 수신하게 하며; 그리고

상기 요청에 응답하여, 상기 샘플을 기술하고 상기 샘플에 대응하는 샘플 엔트리 또는 상기 샘플 중 적어도 하나에 포함되는 파라미터 세트들을 전송하는 것을 포함하여, 상기 랜덤 액세스 샘플을 상기 클라이언트 디바이스에 전송하게 하는

명령들을 더 포함하는, 컴퓨터 판독가능 저장 매체.

청구항 17

제 5 항에 있어서,

상기 디바이스는 무선 통신 디바이스이고, 상기 파일을 송신하도록 구성된 송신기를 더 포함하는, 비디오 데이터를 포함하는 파일을 생성하기 위한 디바이스.

청구항 18

제 17 항에 있어서,

상기 무선 통신 디바이스는 셀룰러 전화이고, 상기 파일은 상기 송신기에 의해 송신되고 셀룰러 통신 표준에 따라 변조되는, 비디오 데이터를 포함하는 파일을 생성하기 위한 디바이스.

발명의 설명

기 술 분 야

[0001] 본 출원은 2016 년 5 월 24 일자로 출원된 미국 가출원 제 62/340,886 호의 이점을 청구하며, 그 전체 내용은 참조에 의해 본원에 통합된다.

[0002] 본 개시는 비디오 비트스트림들과 같은 미디어 비트스트림들의 저장 및 전송을 위한 파일 포맷들에 관한 것이다.

배 경 기 술

[0003] 디지털 비디오 능력들은 디지털 텔레비전들, 디지털 직접 브로드캐스트 시스템들, 무선 브로드캐스트 시스템들, 개인 디지털 보조장치들 (PDA들), 랩탑 또는 데스크탑 컴퓨터들, 디지털 카메라들, 디지털 레코딩 디바이스들, 디지털 미디어 플레이어들, 비디오 게이밍 디바이스들, 비디오 게임 콘솔들, 셀룰러 또는 위성 무선 전화기들, 비디오 텔레컨퍼런싱 디바이스들 등을 포함한, 광범위한 디바이스들에 통합될 수 있다. 디지털 비디오 디바이스들은 MPEG-2, MPEG-4, ITU-T H.263 or ITU-T H.264/MPEG-4, 파트 10, 어드밴스드 비디오 코딩 (AVC), 및 그러한 표준들의 확장물들에 의해 정의된 표준들에서 설명된 기술들과 같은 비디오 압축 기술들을 구현하여, 디지털 비디오 정보를 더 효율적으로 송신 및 수신한다.

[0004] 비디오 압축 기술들은 비디오 시퀀스들에 고유한 리던던시를 감소 또는 제거하기 위해 공간 예측 및/또는 시간 예측을 수행한다. 블록 기반 비디오 코딩에 대해, 비디오 프레임 또는 슬라이스는 매크로블록들로 파티셔닝될 수도 있다. 각 매크로블록은 추가로 파티셔닝될 수 있다. 인트라-코딩된 (I) 프레임 또는 슬라이스에서의 매크로블록들은 이웃 매크로블록들에 대하여 공간 예측을 이용하여 인코딩된다. 인터-코딩된 (P 또는 B) 프레임 또는 슬라이스에서의 매크로블록들은 동일한 프레임 또는 슬라이스에서의 이웃 매크로블록들에 대하여 공간 예측, 또는 다른 참조 프레임들에 대하여 시간 예측을 이용할 수도 있다.

[0005] 비디오 데이터가 인코딩된 후에, 비디오 데이터는 송신 또는 저장을 위해 패킷화될 수도 있다. 비디오 데이터는 ISO (International Organization for Standardization) 베이스 미디어 파일 포맷 및 그 확장물들, 예컨대 AVC 파일 포맷과 같은 다양한 표준들 중 임의의 것에 따르는 비디오 파일로 어셈블링될 수도 있다.

발명의 내용

과제의 해결 수단

[0006] 일반적으로, 본 개시는 하나 이상의 트랙들에서 단일 계층 또는 다중 계층 비트스트림을 저장하는 파일들에 대한 편리한 랜덤 액세스 동작들을 인에이블하는 기술들을 설명한다. 본 개시는 하나 이상의 트랙들에서 단일 계층 또는 다중 계층 비트스트림을 저장하는 파일들에 대한, 편리한 랜덤 액세스 동작들을 인에이블하는 샘플

엔트리 설계들에 관한 방법들, 디바이스들, 및 컴퓨터 프로그램 제품들을 제공한다.

[0007] 일 예에서, 비디오 데이터를 취출하는 방법은, 비디오 비트스트림의 샘플에 대한 샘플 엔트리 타입을 기술하는 데이터를 수신하는 단계로서, 상기 샘플 엔트리 타입은 'hev1' 또는 'hev2' 중 하나이고, 상기 샘플은 고효율 비디오 코딩 (HEVC) 또는 계층화된 HEVC (L-HEVC) 중 하나에 따라 인코딩된 비디오 데이터를 포함하며, 하나 이상의 다른 샘플들은 디코딩 순서에 있어서 상기 비디오 비트스트림에서의 상기 샘플에 선행하는, 상기 데이터를 수신하는 단계, 및 상기 샘플 엔트리 타입이 'hev1' 또는 'hev2' 이고 상기 샘플이 HEVC 또는 L-HEVC 중 하나에 따라 인코딩된 비디오 데이터를 포함하는 것에 응답하여, 디코딩 순서에 있어서 상기 비디오 비트스트림의 임의의 이전 샘플들의 파라미터 세트들을 취출하지 않고 상기 샘플에 선행하는 상기 하나 이상의 다른 샘플들 중 임의의 샘플의 비디오 데이터를 취출하지 않고, 상기 샘플을 사용하여 랜덤 액세스를 수행하도록 상기 샘플을 취출하는 단계를 포함한다.

[0008] 다른 예에서, 비디오 데이터를 취출하기 위한 디바이스는, 비디오 비트스트림의 샘플에 대한 샘플 엔트리 타입을 기술하는 데이터를 수신하는 것으로서, 상기 샘플 엔트리 타입은 'hev1' 또는 'hev2' 중 하나이고, 상기 샘플은 고효율 비디오 코딩 (HEVC) 또는 계층화된 HEVC (L-HEVC) 중 하나에 따라 인코딩된 비디오 데이터를 포함하며, 하나 이상의 다른 샘플들은 디코딩 순서에 있어서 상기 비디오 비트스트림에서의 상기 샘플에 선행하는, 상기 데이터를 수신하고, 그리고 상기 샘플 엔트리 타입이 'hev1' 또는 'hev2' 이고 상기 샘플이 HEVC 또는 L-HEVC 중 하나에 따라 인코딩된 비디오 데이터를 포함하는 것에 응답하여, 디코딩 순서에 있어서 상기 비디오 비트스트림의 임의의 이전 샘플들의 파라미터 세트들을 취출하지 않고 상기 샘플에 선행하는 상기 하나 이상의 다른 샘플들 중 임의의 샘플의 비디오 데이터를 취출하지 않고, 상기 샘플을 사용하여 랜덤 액세스를 수행하도록 상기 샘플을 취출하도록 구성된 하나 이상의 프로세서들을 포함한다.

[0009] 또 다른 예에서, 비디오 데이터를 취출하기 위한 디바이스는, 비디오 비트스트림의 샘플에 대한 샘플 엔트리 타입을 기술하는 데이터를 수신하는 수단으로서, 상기 샘플 엔트리 타입은 'hev1' 또는 'hev2' 중 하나이고, 상기 샘플은 고효율 비디오 코딩 (HEVC) 또는 계층화된 HEVC (L-HEVC) 중 하나에 따라 인코딩된 비디오 데이터를 포함하며, 하나 이상의 다른 샘플들은 디코딩 순서에 있어서 상기 비디오 비트스트림에서의 상기 샘플에 선행하는, 상기 데이터를 수신하는 수단, 및 상기 샘플 엔트리 타입이 'hev1' 또는 'hev2' 이고 상기 샘플이 HEVC 또는 L-HEVC 중 하나에 따라 인코딩된 비디오 데이터를 포함하는 것에 응답하여, 디코딩 순서에 있어서 상기 비디오 비트스트림의 임의의 이전 샘플들의 파라미터 세트들을 취출하지 않고 상기 샘플에 선행하는 상기 하나 이상의 다른 샘플들 중 임의의 샘플의 비디오 데이터를 취출하지 않고, 상기 샘플을 사용하여 랜덤 액세스를 수행하도록 상기 샘플을 취출하는 수단을 포함한다.

[0010] 또 다른 예에서, 명령들이 저장된 컴퓨터 판독가능 저장 매체로서, 상기 명령들은 실행될 경우, 프로세서로 하여금 비디오 비트스트림의 샘플에 대한 샘플 엔트리 타입을 기술하는 데이터를 수신하게 하는 것으로서, 상기 샘플 엔트리 타입은 'hev1' 또는 'hev2' 중 하나이고, 상기 샘플은 고효율 비디오 코딩 (HEVC) 또는 계층화된 HEVC (L-HEVC) 중 하나에 따라 인코딩된 비디오 데이터를 포함하며, 하나 이상의 다른 샘플들은 디코딩 순서에 있어서 상기 비디오 비트스트림에서의 상기 샘플에 선행하는, 상기 데이터를 수신하게 하고, 그리고 상기 샘플 엔트리 타입이 'hev1' 또는 'hev2' 이고 상기 샘플이 HEVC 또는 L-HEVC 중 하나에 따라 인코딩된 비디오 데이터를 포함하는 것에 응답하여, 디코딩 순서에 있어서 상기 비디오 비트스트림의 임의의 이전 샘플들의 파라미터 세트들을 취출하지 않고 상기 샘플에 선행하는 상기 하나 이상의 다른 샘플들 중 임의의 샘플의 비디오 데이터를 취출하지 않고, 상기 샘플을 사용하여 랜덤 액세스를 수행하도록 상기 샘플을 취출하게 한다.

[0011] 또 다른 예에서, 비디오 데이터를 포함하는 파일을 생성하는 방법은, 비디오 데이터의 최저 서브 계층을 운반하는 비디오 데이터의 기본 계층을 포함하는, 복수의 트랙들의 최저 트랙이 편리한 랜덤 액세스가 인에이블되는 것을 나타내는 샘플들에 대한 샘플 엔트리 타입 값들을 포함할 것으로 결정하는 것에 응답하여, 비디오 데이터를 포함하는 상기 복수의 트랙들의 다른 트랙들의 각각의 샘플들에 대한 샘플 엔트리 타입 값들을 편리한 랜덤 액세스가 인에이블되는 것을 나타내도록 세팅하는 단계, 및 상기 복수의 트랙들의 트랙들에 대한 상기 샘플 엔트리 타입 값들이 편리한 랜덤 액세스가 인에이블되는 것을 나타내도록, 상기 복수의 트랙들을 포함하는 파일을 생성하는 단계를 포함한다.

[0012] 또 다른 예에서, 비디오 데이터를 포함하는 파일을 생성하기 위한 디바이스는, 비디오 데이터를 저장하도록 구성된 메모리; 및 회로부에 구현된 하나 이상의 프로세서들을 포함하며, 상기 하나 이상의 프로세서들은, 비디오 데이터의 최저 서브 계층을 운반하는 비디오 데이터의 기본 계층을 포함하는, 복수의 트랙들의 최저 트랙이 편리한 랜덤 액세스가 인에이블되는 것을 나타내는 샘플들에 대한 샘플 엔트리 타입 값들을 포함할 것으로 결정하

는 것에 응답하여, 비디오 데이터를 포함하는 상기 복수의 트랙들의 다른 트랙들의 각각의 샘플들에 대한 샘플 엔트리 타입 값들을 편리한 랜덤 액세스가 인에이블되는 것을 나타내도록 세팅하고, 그리고 상기 복수의 트랙들의 트랙들에 대한 상기 샘플 엔트리 타입 값들이 편리한 랜덤 액세스가 인에이블되는 것을 나타내도록, 상기 복수의 트랙들을 포함하는 파일을 생성하도록 구성된다.

[0013] 또 다른 예에서, 비디오 데이터를 포함하는 파일을 생성하기 위한 디바이스는, 비디오 데이터의 최저 서브 계층을 운반하는 비디오 데이터의 기본 계층을 포함하는, 복수의 트랙들의 최저 트랙이 편리한 랜덤 액세스가 인에이블되는 것을 나타내는 샘플들에 대한 샘플 엔트리 타입 값들을 포함할 것으로 결정하는 것에 응답하여, 비디오 데이터를 포함하는 상기 복수의 트랙들의 다른 트랙들의 각각의 샘플들에 대한 샘플 엔트리 타입 값들을 편리한 랜덤 액세스가 인에이블되는 것을 나타내도록 세팅하는 수단, 및 상기 복수의 트랙들의 트랙들에 대한 상기 샘플 엔트리 타입 값들이 편리한 랜덤 액세스가 인에이블되는 것을 나타내도록, 상기 복수의 트랙들을 포함하는 파일을 생성하는 수단을 포함한다.

[0014] 또 다른 예에서, 명령들이 저장된 컴퓨터 판독가능 저장 매체로서, 상기 명령들은 실행될 경우, 프로세서로 하여금, 비디오 데이터의 최저 서브 계층을 운반하는 비디오 데이터의 기본 계층을 포함하는, 복수의 트랙들의 최저 트랙이 편리한 랜덤 액세스가 인에이블되는 것을 나타내는 샘플들에 대한 샘플 엔트리 타입 값들을 포함할 것으로 결정하는 것에 응답하여, 비디오 데이터를 포함하는 상기 복수의 트랙들의 다른 트랙들의 각각의 샘플들에 대한 샘플 엔트리 타입 값들을 편리한 랜덤 액세스가 인에이블되는 것을 나타내도록 세팅하게 하고, 그리고 상기 복수의 트랙들의 트랙들에 대한 상기 샘플 엔트리 타입 값들이 편리한 랜덤 액세스가 인에이블되는 것을 나타내도록, 상기 복수의 트랙들을 포함하는 파일을 생성하게 한다.

[0015] 하나 이상의 예들의 상세들이 첨부 도면들 및 하기의 설명에 개시된다. 다른 특징들, 목적들, 및 이점들은 그 설명 및 도면들로부터, 그리고 청구항들로부터 명백할 것이다.

도면의 간단한 설명

[0016] 도 1 은 네트워크를 통해 미디어 데이터를 스트리밍하기 위한 기술들을 구현하는 예시적인 시스템을 도시하는 블록도이다.

도 2 는 취출 유닛의 컴포넌트들의 예시적인 세트를 도시하는 블록도이다.

도 3 은 예시적인 멀티미디어 컨텐츠의 엘리먼트들을 도시하는 개념도이다.

도 4 는 예시적인 비디오 파일의 엘리먼트들을 도시하는 블록도이다.

도 5 는 데이터를 프로세싱하는 예시적인 기술을 도시하는 흐름도이다.

도 6 은 데이터를 프로세싱하는 예시적인 기술을 도시하는 흐름도이다.

도 7 은 비디오 데이터를 포함하는 파일을 생성하는 예시적인 기술을 도시하는 흐름도이다.

발명을 실시하기 위한 구체적인 내용

[0017] 일반적으로, 본 개시는 하나 이상의 트랙들에서 단일 계층 또는 다중 계층 비트스트림을 저장하는 파일들에 대한 편리한 랜덤 액세스 동작들을 인에이블 하는 기술들을 설명한다. 편리한 랜덤 액세스는 이전 샘플들로부터의 파라미터 세트들을 검색 및/또는 페치할 필요 없는 랜덤 액세스로서 정의될 수도 있다. 즉, 클라이언트 디바이스가 이전 샘플들의 파라미터 세트들을 검색 및/또는 페치하지 않고 샘플 및 후속 샘플들을 요청할 수 있을 때, 비디오 비트스트림의 샘플에 대해 편리한 랜덤 액세스가 인에이블된다. 따라서, 필요한 모든 파라미터 세트들은 샘플 자체 또는 그 샘플에 대응하는 샘플 엔트리에 포함될 수도 있다.

[0018] 본 개시는 하나 이상의 트랙들에서 단일 계층 또는 다중 계층 비트스트림을 저장하는 파일들에 대한, 편리한 랜덤 액세스 동작들을 인에이블하는 샘플 엔트리 설계들에 관한 방법들을 설명한다.

[0019] 비디오 코딩 표준들은 ITU-T H.261, ISO/IEC MPEG-1 비주얼, ITU-T H.262 또는 ISO/IEC MPEG-2 비주얼, ITU-T H.263, ISO/IEC MPEG-4 비주얼, ITU-T H.264 또는 ISO/IEC MPEG-4 AVC (그 스케일러블 비디오 코딩 (SVC) 및 멀티-뷰 비디오 코딩 (MVC) 확장물들을 포함함), 및 또한 ITU-T H.265 및 ISO/IEC 23008-2 로 알려진 고효율 비디오 코딩 (HEVC) (그 스케일러블 코딩 확장물 (즉, 스케일러블 고효율 비디오 코딩, SHVC), 멀티뷰 확장물 (즉, 멀티뷰 고효율 비디오 코딩, MV-HEVC), 및 3D 확장물 (즉, 3D 고효율 비디오 코딩, 3D-HEVC) 를 포함함)

을 포함한다.

[0020] 파일 포맷 표준들은 ISO 기본 미디어 파일 포맷 (ISOBMFF, ISO/IEC 14496-12), 및 ISOBMFF 로부터 도출된 다른 포맷을 포함하고, ISOBMFF 로부터 도출된 다른 포맷은 MPEG-4 파일 포맷 (ISO/IEC 14496-15), 3GPP 파일 포맷 (3GPP TS 26.244), 및 HEVC 및 그 확장물들을 위한 파일 포맷들뿐만 아니라 AVC 및 그 확장물들을 위한 파일 포맷들을 포함하는 ISO/IEC 14496-15 을 포함한다. ISO/IEC 14496-12 및 14496-15 에 대한 최근 개정판의 초안들은 각각, http://phenix.int-evry.fr/mpeg/doc_end_user/documents/111_Geneva/wg11/w15177-v6-w15177.zip 및 http://wg11.sc29.org/doc_end_user/documents/114_San%20Diego/wg11/w15928-v2-w15928.zip 에서 입수가능하다.

[0021] 본 개시의 기술들은 ISO 기본 미디어 파일 포맷, SVC (Scalable Video Coding) 파일 포맷, AVC (Advanced Video Coding) 파일 포맷, 3GPP (Third Generation Partnership Project) 파일 포맷, 및/또는 MVC (Multiview Video Coding) 파일 포맷, 또는 다른 유사한 비디오 파일 포맷들 중 임의의 파일 포맷에 따라 인캡슐레이션된 비디오 데이터를 따르는 비디오 파일들에 적용될 수도 있다.

[0022] HTTP 스트리밍에서, 자주 사용되는 동작들은 HEAD, GET 및 부분 GET 을 포함한다. HEAD 동작은 주어진 URL (uniform resource locator) 또는 URN (uniform resource name) 에 연관된 파일의 헤더를, 그 URL 또는 URN 과 연관된 페이로드를 취출하지 않고, 취출한다. GET 동작은 주어진 URL 또는 URN 과 연관된 전체 파일을 취출한다. 부분 GET 동작은 입력 파라미터로서 바이트 범위를 수신하고 파일의 연속하는 수의 바이트들을 취출하는데, 그 바이트들의 수는 수신된 바이트 범위에 대응한다. 따라서, 무비 프래그먼트 박스들은 HTTP 스트리밍을 위해 제공될 수도 있는데, 부분 GET 동작이 하나 이상의 개개의 무비 프래그먼트 박스들을 얻을 수 있기 때문이다. 무비 프래그먼트에서는, 상이한 트랙들의 여러 트랙 프래그먼트들이 있을 수 있다. HTTP 스트리밍에서, 미디어 프리젠테이션은 클라이언트에 액세스가능한 데이터의 구조화된 컬렉션일 수도 있다. 클라이언트는 스트리밍 서비스를 사용자에게 제시하기 위해 미디어 데이터 정보를 요청하고 다운로드할 수도 있다.

[0023] HTTP 스트리밍을 사용하여 3GPP 데이터를 스트리밍하는 예에서, 멀티미디어 컨텐츠의 비디오 및/또는 오디오 데이터에 대한 다수의 리프리젠테이션들이 있을 수도 있다. 아래에서 설명되는 바와 같이, 상이한 리프리젠테이션들이 상이한 코딩 특성들 (예컨대, 비디오 코딩 표준의 상이한 프로파일들 또는 레벨들), 상이한 코딩 표준들 또는 코딩 표준들의 확장물들 (이를테면 멀티뷰 및/또는 스케일러블 확장물들), 또는 상이한 비트레이트들에 대응할 수도 있다. 그런 리프리젠테이션들의 매니페스트는 미디어 프리젠테이션 디스크립션 (MPD) 데이터 구조에서 정의될 수도 있다. 미디어 프리젠테이션이 HTTP 스트리밍 클라이언트 디바이스에 액세스가능한 데이터의 구조화된 컬렉션에 대응할 수도 있다. HTTP 스트리밍 클라이언트 디바이스는 스트리밍 서비스를 클라이언트 디바이스의 사용자에게 제시하기 위해 미디어 데이터 정보를 요청하고 다운로드할 수도 있다. 미디어 프리젠테이션이 MPD 의 업데이트들을 포함할 수도 있는 MPD 데이터 구조에서 기술될 수도 있다.

[0024] 미디어 프리젠테이션은 하나 이상의 주기들의 시퀀스를 포함할 수도 있다. 각각의 주기는 다음의 주기의 시작까지, 또는 마지막 주기의 경우에는 미디어 프리젠테이션의 말단까지 연장할 수도 있다. 각각의 주기는 동일한 미디어 컨텐츠에 대해 하나 이상의 리프리젠테이션들을 포함할 수도 있다. 리프리젠테이션은 오디오, 비디오, 타임드 (timed) 텍스트, 또는 다른 이러한 데이터의 다수의 대안적 인코딩된 버전들 중 하나일 수도 있다. 리프리젠테이션들은 인코딩 타입들에 의해, 예컨대, 비디오 데이터에 대한 비트레이트, 해상도, 및/또는 코덱과, 오디오 데이터에 대한 비트레이트, 언어, 및/또는 코덱에 의해 상이할 수도 있다. 리프리젠테이션이란 용어는, 멀티미디어 컨텐츠의 특정 주기에 대응하고 특정 방식으로 인코딩되는 인코딩된 오디오 또는 비디오 데이터의 섹션을 지칭하는데 사용될 수도 있다.

[0025] 특정 주기의 리프리젠테이션들은 리프리젠테이션들이 속하는 적응 세트를 나타내는 MPD 에서의 속성에 의해 표시된 그룹에 할당될 수도 있다. 동일한 적응 세트에서의 리프리젠테이션들은, 클라이언트 디바이스가, 예컨대 대역폭 적응을 수행하기 위해, 이들 리프리젠테이션들 사이를 동적으로 그리고 끊김 없이 스위칭할 수 있다는 점에서, 서로에 대한 대안들로서 일반적으로 간주된다. 예를 들어, 특정 주기 동안의 비디오 데이터의 각각의 리프리젠테이션은 동일한 적응 세트에 할당될 수도 있어서, 그 리프리젠테이션들 중 임의의 리프리젠테이션이 대응하는 주기 동안의 멀티미디어 컨텐츠의 비디오 데이터 또는 오디오 데이터와 같은 미디어 데이터를 제시하기 위한 디코딩을 위해 선택될 수도 있다. 하나의 주기 내의 미디어 컨텐츠는, 일부 예들에서, 그룹 0 이 존재한다면, 그 그룹 0 으로부터의 하나의 리프리젠테이션, 또는 각각의 영이 아닌 그룹으로부터의 많아야 하나의 리프리젠테이션의 조합 중 어느 하나에 의해 표현될 수도 있다. 주기의 각각의 리프리젠테이션에 대

한 타이밍 데이터는 그 주기의 시작 시간을 기준으로 나타낼 수도 있다.

[0026] 리프리젠테이션은 하나 이상의 세그먼트들을 포함할 수도 있다. 각각의 리프리젠테이션은 초기화 세그먼트를 포함할 수도 있거나, 또는 리프리젠테이션의 각각의 세그먼트는 자체 초기화될 수도 있다. 존재하는 경우, 초기화 세그먼트는 리프리젠테이션에 액세스하기 위한 초기화 정보를 포함할 수도 있다. 일반적으로, 초기화 세그먼트는 미디어 데이터를 포함하지 않는다. 세그먼트가 식별자, 이를테면 URL (uniform resource locator), URN (uniform resource name), 또는 URI (uniform resource identifier) 에 의해 고유하게 참조될 수도 있다. MPD는 각각의 세그먼트에 대해 식별자들을 제공할 수도 있다. 일부 예들에서, MPD 는 URL, URN, 또는 URI 에 의해 액세스가능한 파일 내의 세그먼트를 위한 데이터에 대응할 수도 있는 범위 속성의 형태로 바이트 범위들을 제공할 수도 있다.

[0027] 상이한 리프리젠테이션들이 상이한 타입들의 미디어 데이터에 대한 실질적으로 동시 취출을 위해 선택될 수도 있다. 예를 들어, 클라이언트 디바이스가 세그먼트들을 취출하는 타임드 텍스트 리프리젠테이션, 비디오 리프리젠테이션, 및 오디오 리프리젠테이션을 선택할 수도 있다. 일부 예들에서, 클라이언트 디바이스는 대역폭 적응을 수행하기 위한 특정 적응 세트들을 선택할 수도 있다. 즉, 클라이언트 디바이스는 비디오 리프리젠테이션들을 포함하는 적응 세트, 오디오 리프리젠테이션들을 포함하는 적응 세트, 및/또는 타임드 텍스트를 포함하는 적응 세트를 선택할 수도 있다. 대안적으로, 클라이언트 디바이스는 특정한 타입들의 미디어 (예컨대, 비디오) 를 위한 적응 세트들을 선택하고, 다른 타입들의 미디어 (예컨대, 오디오 및/또는 타임드 텍스트) 를 위한 리프리젠테이션들을 직접적으로 선택할 수도 있다.

[0028] 도 1 은 네트워크를 통해 미디어 데이터를 스트리밍하기 위한 기술들을 구현하는 예시적인 시스템 (10) 을 도시하는 블록도이다. 이 예에서, 시스템 (10) 은 컨텐츠 준비 디바이스 (20), 서버 디바이스 (60), 및 클라이언트 디바이스 (40) 를 포함한다. 클라이언트 디바이스 (40) 와 서버 디바이스 (60) 는 인터넷을 포함할 수도 있는 네트워크 (74) 에 의해 통신가능하게 커플링된다. 일부 예들에서, 컨텐츠 준비 디바이스 (20) 와 서버 디바이스 (60) 는 또한, 네트워크 (74) 또는 다른 네트워크에 의해 커플링될 수도 있거나, 또는 직접 통신가능하게 커플링될 수도 있다. 일부 예들에서, 컨텐츠 준비 디바이스 (20) 와 서버 디바이스 (60) 는 동일한 디바이스를 포함할 수도 있다.

[0029] 컨텐츠 준비 디바이스 (20) 는, 도 1 의 예에서, 오디오 소스 (22) 및 비디오 소스 (24) 를 포함한다. 오디오 소스 (22) 는, 예를 들어, 오디오 인코더 (26) 에 의해 인코딩될 캡처된 오디오 데이터를 나타내는 전기적 신호들을 생성하는 마이크로폰을 포함할 수도 있다. 대안적으로, 오디오 소스 (22) 는 이전에 레코딩된 오디오 데이터를 저장하는 저장 매체, 컴퓨터화된 합성기와 같은 오디오 데이터 생성기, 또는 오디오 데이터의 임의의 다른 소스를 포함할 수도 있다. 비디오 소스 (24) 는 비디오 인코더 (28) 에 의해 인코딩될 비디오 데이터를 생성하는 비디오 카메라, 이전에 레코딩된 비디오 데이터로 인코딩된 저장 매체, 컴퓨터 그래픽 소스와 같은 비디오 데이터 생성 유닛, 또는 비디오 데이터의 임의의 다른 소스를 포함할 수도 있다. 컨텐츠 준비 디바이스 (20) 는 모든 예들에서 서버 디바이스 (60) 에 통신적으로 커플링될 필요는 없지만, 서버 디바이스 (60) 에 의해 판독되는 별개의 매체에 멀티미디어 컨텐츠를 저장할 수도 있다.

[0030] 원시 (raw) 오디오 및 비디오 데이터는 아날로그 또는 디지털 데이터를 포함할 수도 있다. 아날로그 데이터는 오디오 인코더 (26) 및/또는 비디오 인코더 (28) 에 의해 인코딩되기 전에 디지털화될 수도 있다. 오디오 소스 (22) 는 발화 (speaking) 참가자가 말을 하는 동안 발화 참가자로부터 오디오 데이터를 획득할 수도 있고, 비디오 소스 (24) 는 동시에 발화 참가자의 비디오 데이터를 획득할 수도 있다. 다른 예들에서, 오디오 소스 (22) 는 저장된 오디오 데이터를 포함하는 컴퓨터 판독가능 저장 매체를 포함할 수도 있고, 비디오 소스 (24) 는 저장된 비디오 데이터를 포함하는 컴퓨터 판독가능 저장 매체를 포함할 수도 있다. 이런 방식으로, 본 개시에서 설명된 기법들은 라이브, 스트리밍, 실시간 오디오 및 비디오 데이터에 또는 아카이브된 (archived), 미리 레코딩된 오디오 및 비디오 데이터에 적용될 수도 있다.

[0031] 비디오 프레임들에 대응하는 오디오 프레임들은 일반적으로 비디오 프레임들 내에 포함되는 비디오 소스 (24) 에 의해 캡처된 (또는 생성된) 비디오 데이터와 동시에 오디오 소스 (22) 에 의해 캡처되었던 (또는 생성된) 오디오 데이터를 포함하는 오디오 프레임들이다. 예를 들어, 발화 참가자가 일반적으로 발화에 의해 오디오 데이터를 생성하는 동안, 오디오 소스 (22) 는 오디오 데이터를 캡처하고, 비디오 소스 (24) 는 동시에, 다시 말하면, 오디오 소스 (22) 가 오디오 데이터를 캡처하고 있는 동안, 발화 참가자의 비디오 데이터를 캡처한다.

그래서, 오디오 프레임이 하나 이상의 특정 비디오 프레임들에 시간적으로 대응할 수도 있다. 따라서, 오디오 프레임이 비디오 프레임에 대응한다는 것은 일반적으로 오디오 데이터 및 비디오 데이터가 동일한 시간

에 캡처되었던 상황에 해당하고 그 상황 동안 오디오 프레임 및 비디오 프레임은 동일한 시간에 캡처되었던 오디오 데이터 및 비디오 데이터를 각각 포함한다.

[0032] 몇몇 예들에서, 오디오 인코더 (26) 는, 인코딩된 오디오 프레임에 대한 오디오 데이터가 레코딩되었던 시간을 나타내는 각각의 인코딩된 오디오 프레임에서의 타임스탬프 (timestamp) 를 인코딩할 수도 있고, 그리고 유사하게, 비디오 인코더 (28) 는 인코딩된 비디오 프레임에 대한 비디오 데이터가 레코딩되었던 시간을 나타내는 각각의 인코딩된 비디오 프레임에서의 타임스탬프를 인코딩할 수도 있다. 이러한 예들에서, 비디오 프레임에 대응하는 오디오 프레임은 타임스탬프를 포함하는 오디오 프레임 및 동일한 타임스탬프를 포함하는 비디오 프레임을 포함할 수도 있다. 컨텐츠 준비 디바이스 (20) 는 내부 클록을 포함할 수도 있고, 그 내부 클록으로부터 오디오 인코더 (26) 및/또는 비디오 인코더 (28) 는 타임스탬프들을 생성할 수도 있고, 또는, 오디오 소스 (22) 및 비디오 소스 (24) 는 오디오 및 비디오 데이터를 각각 타임스탬프와 연관시키기 위해 이용할 수도 있다.

[0033] 일부 예들에서, 오디오 소스 (22) 는 오디오 데이터가 레코딩되었던 시간에 대응하는 데이터를 오디오 인코더 (26) 에 전송할 수도 있고, 비디오 소스 (24) 는 비디오 데이터가 레코딩되었던 시간에 대응하는 데이터를 비디오 인코더 (28) 에 전송할 수도 있다. 일부 예들에서, 오디오 인코더 (26) 는, 인코딩된 오디오 데이터의 상대적인 시간적 순서를 나타내기 위해서, 하지만 그 오디오 데이터가 레코딩되었던 절대적 시간을 반드시 나타낼 필요는 없이, 인코딩된 오디오 데이터에서 시퀀스 식별자를 인코딩할 수도 있고, 그리고 유사하게, 비디오 인코더 (28) 는 또한, 인코딩된 비디오 데이터의 상대적인 시간적 순서를 나타내기 위해 시퀀스 식별자들을 이용할 수도 있다. 유사하게, 일부 예들에서, 시퀀스 식별자는 타임스탬프와 맵핑되거나 그렇지 않으면 상관될 수도 있다.

[0034] 오디오 인코더 (26) 는 일반적으로 인코딩된 오디오 데이터의 스트림을 생성하는 한편, 비디오 인코더 (28) 는 인코딩된 비디오 데이터의 스트림을 생성한다. 데이터 (오디오든 비디오든) 의 각각의 개별 스트림은 기본 스트림 (elementary stream) 으로서 지칭될 수도 있다. 기본 스트림은 리프리젠테이션의 단일의, 디지털로 코딩된 (가능하게는 압축된) 컴포넌트이다. 예를 들어, 그 리프리젠테이션의 코딩된 비디오 또는 오디오 부분은 기본 스트림일 수 있다. 기본 스트림은 비디오 파일 내에 인캡슐레이션되기 전에 패킷화된 기본 스트림 (PES) 으로 컨버팅될 수도 있다. 동일한 리프리젠테이션 내에서, 스트림 ID 는 하나의 기본 스트림에 속하는 PES-패킷들을 다른 것들로부터 구별하는데 이용될 수도 있다. 기본 스트림의 데이터의 기본 단위는 패킷화된 기본 스트림 (PES) 패킷이다. 따라서, 코딩된 비디오 데이터는 일반적으로 기본 비디오 스트림들에 대응한다. 유사하게, 오디오 데이터는 하나 이상의 개별 기본 스트림들에 대응한다.

[0035] ITU-T H.264/AVC 및 도래하는 고효율 비디오 코딩 (HEVC) 표준과 같은 많은 비디오 코딩 표준들은, 에러 없는 비트스트림들을 위한 신택스, 시맨틱스, 및 디코딩 프로세스를 정의하고, 그것들 중의 어떤 것은 특정 프로파일 또는 레벨을 따른다. 비디오 코딩 표준들은 통상적으로 인코더를 특정하지 않지만, 인코더는 생성된 비트스트림들이 디코더에 부합하는 표준이라는 것을 보장하는 임무가 주어진다. 비디오 코딩 표준의 측면에서, "프로파일"은 알고리즘들, 특징들 (features), 또는 툴들 및 그것들에 적용되는 제약들의 서브세트에 대응한다. H.264 표준에 의해 정의된 바와 같이, 예를 들어, "프로파일" 은 H.264 표준에 의해 특정되는 전체 비트스트림 신택스의 서브세트이다. "레벨" 은 예를 들어, 디코더 메모리 및 컴퓨테이션과 같은 디코더 리소스 소비의 한계들에 대응하며, 이 한계들은 픽처들의 해상도, 비트 레이트, 및 블록 프로세싱 레이트에 관련된다. 프로파일은 profile_idc (프로파일 표시자) 값으로 시그널링될 수도 있는 한편, 레벨은 level_idc (레벨 표시자) 값으로 시그널링될 수도 있다.

[0036] H.264 표준은, 예를 들어, 주어진 프로파일의 신택스에 의해 부과되는 경계들 내에서, 디코딩된 픽처들의 특정된 사이즈와 같이 비트스트림 내의 신택스 엘리먼트들에 의해 취해진 값들에 의존하여 인코더들 및 디코더들의 성능에서의 큰 변화를 요구하는 것이 여전히 가능하다는 것을 인식한다. H.264 표준은 많은 애플리케이션들에서, 특정 프로파일 내에서 신택스의 모든 가설적 사용들을 처리하는 것이 가능한 디코더를 구현하는 것이 실용적이지도 않고 경제적이지도 않다는 것을 추가로 인식한다. 따라서, H.264 표준은 비트스트림에서 신택스 엘리먼트들의 값들에 부과되는 특정된 제약들의 세트로서 "레벨" 을 정의한다. 이들 제약들은 값들에 관한 간단한 제한들일 수도 있다. 대안적으로, 이들 제약들은 값들의 산술적 조합들 (예컨대, 픽처 폭 곱하기 픽처 높이 곱하기 초당 디코딩되는 픽처들의 수) 에 대한 제약들의 형태를 취할 수도 있다. H.264 표준은 개개의 구현예들이 각각의 지원된 프로파일들에 대해 상이한 레벨을 지원할 수도 있다는 것을 추가로 제공한다.

[0037] 프로파일에 부합하는 디코더는 보통은 프로파일에서 정의된 모든 특징들을 지원한다. 예를 들어, 코딩 특징으로서, B-픽처 코딩은 H.264/AVC 의 베이스라인 프로파일에서 지원되지 않지만 H.264/AVC 의 다른 프로파일들에서 지원된다. 레벨에 부합하는 디코더는 그 레벨에서 정의된 한계들을 넘어 자원들을 요구하지 않는 임의의 비트스트림을 디코딩하는 것이 가능해야 한다. 프로파일들 및 레벨들의 정의들은 해석능력 (interpretability) 에 도움이 될 수도 있다. 예를 들어, 비디오 송신 동안, 한 쌍의 프로파일 및 레벨 정의들은 전체 송신 세션 동안에 협상되고 합의될 수도 있다. 보다 구체적으로는, H.264/AVC 에서, 레벨은 예를 들어, 처리될 필요가 있는 매크로블록들의 수, 디코딩된 픽처 버퍼 (decoded picture buffer; DPB) 사이즈, 코딩된 픽처 버퍼 (coded picture buffer; CPB) 사이즈, 수직 모션 벡터 범위, 2 개의 연속적인 MB 들 당 모션 벡터들의 최대 수, 및 B-블록이 8x8 픽셀들 미만의 서브-매크로블록 파티션들을 가질 수 있는지의 여부에 관한 제한들을 정의할 수도 있다. 이러한 방식으로, 디코더는, 디코더가 비트스트림을 적절히 디코딩하는 것이 가능한지의 여부를 결정할 수도 있다.

[0038] 도 1 의 예에서, 컨텐츠 준비 디바이스 (20) 의 인캡슐레이션 유닛 (30) 은 코딩된 비디오 데이터를 포함하는 기본 스트림들을 비디오 인코더 (28) 로부터 그리고 코딩된 오디오 데이터를 포함하는 기본 스트림들을 오디오 인코더 (26) 로부터 수신한다. 일부 예들에서, 비디오 인코더 (28) 및 오디오 인코더 (26) 는 각각, 인코딩된 데이터로부터 PES 패킷들을 형성하기 위한 패킷화기들을 포함할 수도 있다. 다른 예들에서, 비디오 인코더 (28) 및 오디오 인코더 (26) 는 각각, 인코딩된 데이터로부터 PES 패킷들을 형성하기 위한 각기의 패킷화기들과 인터페이싱할 수도 있다. 또 다른 예들에서, 인캡슐레이션 유닛 (30) 은 인코딩된 오디오 및 비디오 데이터로부터 PES 패킷들을 형성하기 위한 패킷화기들을 포함할 수도 있다.

[0039] 비디오 인코더 (28) 는 멀티미디어 컨텐츠의 비디오 데이터를 다양한 방식들로 인코딩하여, 다양한 비트레이트들에서 그리고 다양한 특성들, 이를테면 픽셀 해상도들, 프레임 레이트들, 다양한 코딩 표준들에 대한 준수 (conformance), 다양한 코딩 표준들에 대한 다양한 프로파일들 및/또는 프로파일들의 레벨들에 대한 준수, (예컨대, 2 차원 또는 3 차원 재생을 위한) 하나 또는 다수의 뷰들을 갖는 리프리젠테이션들, 또는 다른 그런 특성들을 갖는 멀티미디어 컨텐츠의 상이한 리프리젠테이션들을 생성할 수도 있다. 리프리젠테이션은, 본 개시에서 사용된 바와 같이, 오디오 데이터, 비디오 데이터, 텍스트 데이터 (예컨대, 폐쇄된 캡션들에 대해), 또는 다른 이러한 데이터 중 하나를 포함할 수도 있다. 리프리젠테이션은 오디오 기본 스트림 또는 비디오 기본 스트림과 같은 기본 스트림을 포함할 수도 있다. 각각의 PES 패킷은 PES 패킷이 속한 기본 스트림을 식별하는 stream_id 를 포함할 수도 있다. 인캡슐레이션 유닛 (30) 은 기본 스트림들을 다양한 리프리젠테이션들의 비디오 파일들 (예컨대, 세그먼트들) 로 어셈블링하는 것을 담당한다.

[0040] 인캡슐레이션 유닛 (30) 은 오디오 인코더 (26) 및 비디오 인코더 (28) 로부터 리프리젠테이션의 기본 스트림들을 위한 PES 패킷들을 수신하고 그 PES 패킷들로부터 대응하는 네트워크 추상화 계층 (network abstraction layer; NAL) 단위들을 형성한다. H.264/AVC (Advanced Video Coding) 의 예에서, 코딩된 비디오 세그먼트들은 NAL 유닛들로 조직화되며, 그것들은 픽처 전화, 스토리지 (storage), 브로드캐스트, 또는 스트리밍과 같은 "네트워크 친화적 (network-friendly)" 비디오 리프리젠테이션 어드레싱 애플리케이션들을 제공한다. NAL 유닛들은 비디오 코딩 계층 (Video Coding Layer; VCL) NAL 유닛들 및 비-VCL NAL 유닛들로 분류될 수 있다. VCL 유닛들은 코어 압축 엔진을 포함할 수도 있고 블록, 매크로블록, 및/또는 슬라이스 레벨 데이터를 포함할 수도 있다. 다른 NAL 유닛들은 비-VCL NAL 유닛들일 수도 있다. 일부 예들에서, 보통 프라이머리 (primary) 코딩된 픽처로서 제시되는 하나의 시간 인스턴스에서의 코딩된 픽처는, 하나 이상의 NAL 유닛들을 포함할 수도 있는, 액세스 유닛 내에 포함될 수도 있다.

[0041] 비-VCL NAL 유닛들은 무엇보다 파라미터 세트 NAL 유닛들 및 SEI NAL 유닛들을 포함할 수도 있다. 파라미터 세트들은 (시퀀스 파라미터 세트들 (SPS) 에서) 시퀀스-레벨 헤더 정보 및 (픽처 파라미터 세트들 (PPS) 에서) 드물게 변화하는 픽처-레벨 헤더 정보를 포함할 수도 있다. 파라미터 세트들 (예컨대, PPS 및 SPS) 에 있어서, 드물게 변화하는 정보는 각 시퀀스 또는 픽처에 대해 반복될 필요가 없고, 따라서, 코딩 효율이 향상될 수도 있다. 또한, 파라미터 세트들의 사용은 중요한 헤더 정보의 대역외 송신을 가능하게 할 수도 있어, 에러 복구를 위한 중복성 송신에 대한 필요성을 회피할 수도 있다. 대역외 송신 예들에서, 파라미터 세트 NAL 유닛들은 SEI NAL 유닛들과 같은, 다른 NAL 유닛들과는 상이한 채널을 통해 송신될 수도 있다.

[0042] SEI (Supplemental Enhancement Information) 는 VCL NAL 유닛들로부터 코딩된 픽처들 샘플들을 디코딩할 필요는 없지만 디코딩, 디스플레이, 에러 복구 및 기타 목적과 관련된 프로세스를 지원할 수도 있는 정보를 포함할 수도 있다. SEI 메시지는 비-VCL NAL 유닛에 포함될 수도 있다. SEI 메시지들은 일부 표준 사양들의 정규 부분이므로, 표준 준수 디코더 구현에 항상 필수적인 것은 아니다. SEI 메시지들은 시퀀스 레벨 SEI 메

시지들 또는 픽처 레벨 SEI 메시지들일 수도 있다. 일부 시퀀스 레벨 정보는 SVC 예에서의 스케일가능 정보 SEI 메시지들 및 MVC 에서의 뷰 스케일가능 정보 SEI 메시지들과 같은 SEI 메시지들에 포함될 수도 있다. 이러한 예시적인 SEI 메시지들은 동작 포인트의 추출 및 동작 포인트들의 특성들에 관한 정보를 전달할 수도 있다. 또한, 인캡슐레이션 유닛 (30) 은 리프리젠테이션들의 특성들을 기술하는 미디어 프리젠테이션 디스크립터 (MPD) 와 같은, 매니페스트 파일을 형성할 수도 있다. 인캡슐레이션 유닛 (30) 은 확장성 마크업 언어 (extensible markup language; XML) 에 따라 MPD 를 포맷할 수도 있다.

[0043] 인캡슐레이션 유닛 (30) 은 멀티미디어 컨텐츠의 하나 이상의 리프리젠테이션들에 대한 데이터를 매니페스트 파일 (예를 들어, MPD) 과 함께 출력 인터페이스 (32) 에 제공할 수도 있다. 출력 인터페이스 (32) 는 네트워크 인터페이스 또는 저장 매체에 기록하기 위한 인터페이스, 예컨대 범용 직렬 버스 (USB) 인터페이스, CD 또는 DVD 라이터 또는 버너 (burner), 자기적 또는 플래시 저장 매체들로에 대한 인터페이스, 또는 미디어 데이터를 저장하거나 송신하기 위한 다른 인터페이스들을 포함할 수도 있다. 인캡슐레이션 유닛 (30) 은 멀티미디어 컨텐츠의 리프리젠테이션들의 각각의 데이터를 출력 인터페이스 (32) 에 제공할 수도 있는데, 이것은 그 데이터를 네트워크 송신 또는 저장 매체들을 통하여 서버 디바이스 (60) 로 전송할 수도 있다. 도 1 의 예에서, 서버 디바이스 (60) 는 각기의 매니페스트 파일 (66) 및 하나 이상의 리프리젠테이션들 (68A-68N) (리프리젠테이션들 (68)) 을 각각 포함하는 다양한 멀티미디어 컨텐츠 (64) 를 저장하는 저장 매체 (62) 를 포함한다. 일부 예들에서, 출력 인터페이스 (32) 는 또한 데이터를 직접 네트워크 (74) 에 전송할 수도 있다.

[0044] 일부 예들에서, 리프리젠테이션들 (68) 은 적응 세트들로 분리될 수도 있다. 즉, 리프리젠테이션들 (68) 의 다양한 서브세트들은 각기의 특성들의 공통 세트들, 이를테면 코덱, 프로파일 및 레벨, 해상도, 뷰들의 수, 세그먼트들에 대한 파일 포맷, 디코딩되고 예컨대 스피커들에 의해 제시될 오디오 데이터 및/또는 리프리젠테이션으로 디스플레이될 텍스트의 언어 또는 다른 특성들을 식별할 수도 있는 텍스트 타입 정보, 적응 세트에서 리프리젠테이션들에 대한 장면의 관점에서 카메라 각도 또는 실세계 카메라를 기술할 수도 있는 카메라 각도 정보, 특정 청중들에 대한 컨텐츠 적합성 (suitability) 을 기술하는 등급 정보 등을 포함할 수도 있다.

[0045] 매니페스트 파일 (66) 은 특정한 적응 세트들, 및 적응 세트들에 대한 공통 특성들에 대응하는 리프리젠테이션들 (68) 의 서브세트들을 나타내는 데이터를 포함할 수도 있다. 매니페스트 파일 (66) 은 또한 적응 세트들의 개개의 리프리젠테이션들에 대한 개개의 특성들, 이를테면 비트레이트들을 나타내는 데이터를 포함할 수도 있다. 이런 방식으로, 적응 세트는 단순화된 네트워크 대역폭 적응을 제공할 수도 있다. 적응 세트에서의 리프리젠테이션들은 매니페스트 파일 (66) 의 적응 세트 엘리먼트의 자식 (child) 엘리먼트들을 이용하여 나타내어질 수도 있다.

[0046] 서버 디바이스 (60) 는 요청 프로세싱 유닛 (70) 과 네트워크 인터페이스 (72) 를 포함한다. 일부 예들에서, 서버 디바이스 (60) 는 복수의 네트워크 인터페이스들을 포함할 수도 있다. 더욱이, 서버 디바이스 (60) 의 특징들의 임의의 것 또는 모두는 컨텐츠 전달 네트워크의 다른 디바이스들, 이를테면 라우터들, 브릿지들, 프록시 디바이스들, 스위치들, 또는 다른 디바이스들 상에서 구현될 수도 있다. 일부 예들에서, 컨텐츠 전달 네트워크의 중간 디바이스들은 멀티미디어 컨텐츠 (64) 의 데이터를 캐싱하고, 서버 디바이스 (60) 의 그것들에 실질적으로 부합하는 컴포넌트들을 포함할 수도 있다. 일반적으로, 네트워크 인터페이스 (72) 는 네트워크 (74) 를 통해 데이터를 전송하고 수신하도록 구성된다.

[0047] 요청 프로세싱 유닛 (70) 은, 저장 매체 (72) 의 데이터에 대해, 클라이언트 디바이스들, 이를테면 클라이언트 디바이스 (40) 로부터 네트워크 요청들을 수신하도록 구성된다. 예를 들어, 요청 프로세싱 유닛 (70) 은, R. Fielding 등의 RFC 2616, "Hypertext Transfer Protocol - HTTP/1.1", Network Working Group, IETF, 1999 년 6 월에서 기술된 바와 같이, 하이퍼텍스트 전송 프로토콜 (HTTP) 버전 1.1 을 구현할 수도 있다. 다시 말하면, 요청 프로세싱 유닛 (70) 은 HTTP GET 또는 부분 GET 요청들을 수신하고 그 요청들에 응답하여 멀티미디어 컨텐츠 (64) 의 데이터를 제공하도록 구성될 수도 있다. 그 요청들은 리프리젠테이션들 (68) 중 하나의 리프리젠테이션의 세그먼트를, 예컨대, 그 세그먼트의 URL 을 이용하여 특정할 수도 있다. 일부 예들에서, 요청들은 또한 세그먼트의 하나 이상의 바이트 범위들을 특정할 수도 있고, 따라서, 부분 GET 요청들을 포함할 수도 있다. 요청 프로세싱 유닛 (70) 은 리프리젠테이션들 (68) 중 하나의 리프리젠테이션의 세그먼트의 헤더 데이터를 제공하기 위해 HTTP HEAD 요청들을 서비스하도록 추가로 구성될 수도 있다. 어느 경우에도, 요청 프로세싱 유닛 (70) 은 요청된 데이터를 요청 디바이스, 이를테면 클라이언트 디바이스 (40) 에 제공하기 위해 요청들을 처리하도록 구성될 수도 있다.

[0048] 추가적으로 또는 대안적으로, 요청 프로세싱 유닛 (70) 은 eMBMS 와 같은 브로드캐스트 또는 멀티캐스트 프로토

콜을 통해 미디어 데이터를 전달하도록 구성될 수도 있다. 컨텐츠 준비 디바이스 (20) 는 기술된 것과 실질적으로 동일한 방식으로 DASH 세그먼트들 및/또는 서브-세그먼트들을 형성할 수도 있지만, 서버 디바이스 (60) 는 eMBMS 또는 다른 브로드캐스트 또는 멀티캐스트 네트워크 전송 프로토콜을 이용하여 이들 세그먼트들 또는 서브-세그먼트들을 전달할 수도 있다. 예를 들어, 요청 프로세싱 유닛 (70) 은 클라이언트 디바이스 (40) 로부터 멀티캐스트 그룹 참여 요청을 수신하도록 구성될 수도 있다. 즉, 서버 디바이스 (60) 는, 특정 미디어 컨텐츠 (예컨대, 라이브 이벤트의 브로드캐스트) 와 연관된, 클라이언트 디바이스 (40) 를 포함하는, 클라이언트 디바이스들에 멀티캐스트 그룹과 연관된 인터넷 프로토콜 (IP) 어드레스를 광고할 수도 있다. 클라이언트 디바이스 (40) 는, 다시, 멀티캐스트 그룹에 참여하기 위한 요청을 제출할 수도 있다. 이 요청은 네트워크 (74), 예컨대, 네트워크 (74) 를 구성하는 라우터들을 통해 전파될 수도 있어, 그 라우터들은 클라이언트 디바이스 (40) 와 같은 가입 클라이언트 디바이스들에 멀티캐스트 그룹과 연관된 IP 어드레스를 목적지로 하는 트래픽을 다이렉팅하도록 야기될 수도 있다.

[0049] 도 1 의 예에서 도시된 바와 같이, 멀티미디어 컨텐츠 (64) 는, 미디어 프리젠테이션 디스크립션 (MPD) 에 대응할 수도 있는 매니페스트 파일 (66) 을 포함한다. 매니페스트 파일 (66) 은 상이한 대안적 리프리젠테이션들 (68) (예컨대, 상이한 품질들을 갖는 비디오 서비스들) 의 서술들을 포함할 수도 있고, 그 디스크립션은, 예컨대, 코덱 정보, 프로파일 값, 레벨 값, 비트레이트, 및 리프리젠테이션들 (68) 의 다른 서술적 특성들을 포함할 수도 있다. 클라이언트 디바이스 (40) 는 미디어 프리젠테이션의 MPD 를 취출하여 리프리젠테이션들 (68) 의 세그먼트들에 어떻게 액세스할 지를 결정할 수도 있다.

[0050] 특히, 취출 유닛 (52) 은 클라이언트 디바이스 (40) 의 구성 데이터 (configuration data) (미도시) 를 취출하여 비디오 디코더 (48) 의 디코딩 능력들 및 비디오 출력 (44) 의 렌더링 능력들 (rendering capabilities) 을 결정할 수도 있다. 구성 데이터는 클라이언트 디바이스 (40) 의 사용자에 의하여 선택되는 언어 선호도, 클라이언트 디바이스 (40) 의 사용자에 의하여 설정된 심도 선호도들에 대응하는 하나 이상의 카메라 퍼스펙티브들, 및/또는 클라이언트 디바이스 (40) 의 사용자에 의하여 선택된 등급 선호도 중 임의의 것 또는 전부를 또한 포함할 수도 있다. 취출 유닛 (52) 은, 예를 들어, HTTP GET 및 부분 GET 요청들을 제출하도록 구성되는 미디어 클라이언트 또는 웹 브라우저를 포함할 수도 있다. 취출 유닛 (52) 은 클라이언트 디바이스 (40) 의 하나 이상의 프로세서들 또는 프로세싱 유닛들 (미도시) 에 의하여 실행되는 소프트웨어 명령들에 대응할 수도 있다. 몇몇 예들에서, 취출 유닛 (52) 에 대하여 기술된 기능성의 전부 또는 부분들은 하드웨어로, 또는 하드웨어, 소프트웨어, 및/또는 펌웨어의 조합으로 구현될 수도 있는데, 여기서 필요한 하드웨어가 소프트웨어 또는 펌웨어에 대한 명령들을 실행하기 위하여 제공될 수도 있다.

[0051] 취출 유닛 (52) 은 클라이언트 디바이스 (40) 의 디코딩 및 렌더링 능력들을 매니페스트 파일 (66) 의 정보에 의하여 표시되는 리프리젠테이션들 (68) 의 특성들과 비교할 수도 있다. 취출 유닛 (52) 은 처음에 적어도 매니페스트 파일 (66) 의 일부를 취출하여 리프리젠테이션들 (68) 의 특성들을 결정할 수도 있다. 예를 들어, 취출 유닛 (52) 은 본 개시의 기법들에 따라, 하나 이상의 적응 세트들의 특성들을 기술하는 매니페스트 파일 (66) 일부를 요청할 수도 있다. 취출 유닛 (52) 은 클라이언트 디바이스 (40) 의 코딩 및 렌더링 능력들에 의해 만족될 수 있는 특성들을 갖는 리프리젠테이션들 (68) (예컨대, 적응 세트) 의 서브세트를 선택할 수도 있다. 취출 유닛 (52) 은 그 다음, 적응 세트 내의 리프리젠테이션들에 대한 비트레이트들을 결정하고, 네트워크 대역폭의 현재 이용가능한 양을 결정하며, 그리고 네트워크 대역폭에 의하여 만족될 수 있는 비트레이트를 가지는 리프리젠테이션들 중 하나로부터 세그먼트들을 취출할 수도 있다.

[0052] 일반적으로, 더 높은 비트레이트 리프리젠테이션들은 더 높은 품질의 비디오 재생을 도출할 수도 있는 한편, 더 낮은 비트레이트 리프리젠테이션들은 가용 네트워크 대역폭이 감소할 때 충분한 품질의 비디오 재생을 제공할 수도 있다. 이에 따라, 가용 네트워크 대역폭이 상대적으로 높은 경우에는, 취출 유닛 (52) 은 데이터를 상대적으로 높은 비트레이트 리프리젠테이션들로부터 취출할 수도 있는 반면에, 가용 네트워크 대역폭이 낮은 경우에는, 취출 유닛 (52) 은 데이터를 상대적으로 낮은 비트레이트 리프리젠테이션들로부터 취출할 수도 있다.

이러한 방식으로, 클라이언트 디바이스 (40) 는 네트워크 (74) 의 변동하는 네트워크 대역폭 이용가능성에 적응하면서도 멀티미디어 데이터를 네트워크 (74) 를 통해 스트리밍할 수도 있다.

[0053] 추가적으로 또는 대안적으로, 취출 유닛 (52) 은 eMBMS 또는 IP 멀티캐스트와 같은, 브로드캐스트 또는 멀티캐스트 네트워크 프로토콜에 따라 데이터를 수신하도록 구성될 수도 있다. 이러한 예들에서, 취출 유닛 (52) 은 특정 미디어 컨텐츠와 연관된 멀티캐스트 네트워크 그룹에 참여하기 위한 요청을 제출할 수도 있다. 멀티캐스트 그룹에 참여한 후에, 취출 유닛 (52) 은 서버 디바이스 (60) 또는 컨텐츠 준비 디바이스 (20) 에 발행된 추가적인 요청들 없이 멀티캐스트 그룹의 데이터를 수신할 수도 있다. 취출 유닛 (52) 은, 멀티캐스트

그룹의 데이터가 더 이상 필요 없을 때 멀티캐스트 그룹을 떠나기 위한, 예컨대 다른 멀티캐스트 그룹으로 채널들을 변경하거나 재생을 중지하기 위한 요청을 제출할 수도 있다.

[0054] 네트워크 인터페이스 (54) 는, 선택된 리프리젠테이션의 세그먼트들의 데이터를 수신하고 취출 유닛 (52) 에 제공할 수도 있고, 이 취출 유닛 (52) 은 다시 그 세그먼트들을 파일 포맷 프로세싱 유닛 (50) 에 제공할 수도 있다. 파일 포맷 프로세싱 유닛 (50) 은 비디오 파일의 엘리먼트들을 구성성분인 PES 스트림들로 디캡슐레이션하며, PES 스트림들을 패킷해제하여 인코딩된 데이터를 취출하고, 예컨대, 스트림의 PES 패킷 헤더들에 의해 나타내어진 바와 같이, 인코딩된 데이터가 오디오 스트림의 부분인지 또는 비디오 스트림의 부분인지에 따라, 오디오 디코더 (46) 또는 비디오 디코더 (48) 중 어느 일방에 인코딩된 데이터를 전송할 수도 있다. 오디오 디코더 (46) 는 인코딩된 오디오 데이터를 디코딩하고 디코딩된 오디오 데이터를 오디오 출력부 (42) 에 전송하는 한편, 비디오 디코더 (48) 는 인코딩된 비디오 데이터를 디코딩하고 스트림의 복수의 뷰들을 포함할 수도 있는 디코딩된 비디오 데이터를 비디오 출력부 (44) 에 전송한다.

[0055] 비디오 인코더 (28), 비디오 디코더 (48), 오디오 인코더 (26), 오디오 디코더 (46), 인캡슐레이션 유닛 (30), 취출 유닛 (52), 및 파일 포맷 프로세싱 유닛 (50) 각각은 적용가능한 경우에 다양한 적합한 처리 회로들, 예컨대 하나 이상의 마이크로프로세서들, 디지털 신호 프로세서 (DSP) 들, 주문형 집적회로 (ASIC) 들, 필드 프로그래밍가능 게이트 어레이 (FPGA) 들, 이산 로직 회로, 소프트웨어, 하드웨어, 펌웨어 또는 임의의 이들의 조합들 중 임의의 것으로서 구현될 수도 있다. 비디오 인코더 (28) 및 비디오 디코더 (48) 각각은 하나 이상의 인코더들 또는 디코더들 내에 포함될 수도 있는데, 이들 중 어떤 것은 결합된 비디오 인코더/디코더 (CODEC) 의 일부로서 통합될 수도 있다. 마찬가지로, 오디오 인코더 (26) 및 오디오 디코더 (46) 각각은 하나 이상의 인코더들 또는 디코더들 내에 포함될 수도 있는데, 이들 중 임의의 것은 결합된 코덱의 일부로서 통합될 수도 있다. 비디오 인코더 (28), 비디오 디코더 (48), 오디오 인코더 (26), 오디오 디코더 (46), 인캡슐레이션 유닛 (30), 취출 유닛 (52), 및/또는 파일 포맷 프로세싱 유닛 (50) 을 포함하는 장치는 집적 회로, 마이크로프로세서, 및/또는 무선 통신 디바이스, 예컨대 셀룰러 전화기를 포함할 수도 있다.

[0056] 클라이언트 디바이스 (40), 서버 디바이스 (60), 및/또는 컨텐츠 준비 디바이스 (20) 는 본 개시의 기법들에 따라 동작하도록 구성될 수도 있다. 예시를 위해, 본 개시는 클라이언트 디바이스 (40) 와 서버 디바이스 (60) 에 대해 이들 기법들을 기술한다. 하지만, 컨텐츠 준비 디바이스 (20) 가 서버 디바이스 (60) 대신에 (또는 추가적으로) 이들 기법들을 수행하도록 구성될 수도 있다.

[0057] 인캡슐레이션 유닛 (30) 은, NAL 유닛이 속하는 프로그램을 식별하는 헤더와, 예컨대, 오디오 데이터, 비디오 데이터, 또는 NAL 유닛이 대응하는 반송 또는 프로그램 스트림을 기술하는 데이터와 같은 페이로드를 포함하는 NAL 유닛들을 형성할 수도 있다. 예를 들어, H.264/AVC 에서, NAL 유닛은 1-바이트 헤더 및 가변 사이즈의 페이로드를 포함한다. 페이로드에 비디오 데이터를 포함하는 NAL 유닛은 다양한 입도 레벨들 (granularity levels) 의 비디오 데이터를 포함할 수도 있다. 예를 들어, NAL 유닛은 비디오 데이터의 블록, 복수의 블록들, 비디오 데이터의 슬라이스, 또는 비디오 데이터의 전체 픽처를 포함할 수도 있다. 인캡슐레이션 유닛 (30) 은 기본 스트림들의 PES 패킷들의 형태로 비디오 인코더 (28) 로부터 인코딩된 비디오 데이터를 수신할 수도 있다. 인캡슐레이션 유닛 (30) 은 각각의 기본 스트림을 대응하는 프로그램과 연관시킬 수도 있다.

[0058] 인캡슐레이션 유닛 (30) 은 또한 복수의 NAL 유닛들로부터 액세스 유닛들을 어셈블링할 수도 있다. 일반적으로, 액세스 유닛은 비디오 데이터의 프레임, 및 오디오 데이터가 이용가능할 때 그 프레임에 대응하는 오디오 데이터를 나타내기 위한 하나 이상의 NAL 유닛들을 포함할 수도 있다. 액세스 유닛은 일반적으로 하나의 출력 시간 인스턴스에 대한 모든 NAL 유닛들, 예컨대, 하나의 타임 인스턴스에 대한 모든 오디오 및 비디오 데이터를 포함한다. 예를 들어, 각각의 뷰가 20 의 초당 프레임 (fps) 의 프레임 레이트를 가지는 경우에, 각각의 시간 인스턴스는 0.05 초의 시간 간격에 대응할 수도 있다. 이 시간 간격 동안, 동일한 액세스 유닛 (동일한 시간 인스턴스) 의 모든 뷰들에 대한 특정 프레임들은 동시에 렌더링될 수도 있다. 일 예에서, 액세스 유닛은 하나의 시간 인스턴스에서 코딩된 픽처를 포함할 수도 있고, 이는 프라이머리 코딩된 픽처로서 제시될 수도 있다.

[0059] 이에 따라, 액세스 유닛은 공통 시간 인스턴스의 모든 오디오 및 비디오 프레임들, 예컨대, 시간 X 에 대응하는 모든 뷰들을 포함할 수도 있다. 본 개시는 또한, 특정 뷰의 인코딩된 픽처를 "뷰 컴포넌트" 로서 지칭한다. 즉, 뷰 컴포넌트는 특정 시간에서 특정 뷰에 대한 인코딩된 픽처 (또는 프레임) 를 포함할 수도 있다. 따라서, 액세스 유닛은 공통 시간 인스턴스의 모든 뷰 컴포넌트들을 포함하는 것으로서 정의될 수도 있다. 액세스 유닛들의 디코딩 순서는 반드시 출력 또는 디스플레이 순서와 동일할 필요는 없다.

[0060] 미디어 프리젠테이션은 미디어 프리젠테이션 디스크립션 (MPD) 을 포함할 수도 있고, 이 미디어 프리젠테이션 디스크립션은 상이한 대안적인 리프리젠테이션들 (예를 들어 상이한 품질들을 갖는 비디오 서비스들) 의 디스크립션을 포함할 수도 있고, 디스크립션은 예를 들어 코덱 정보, 프로파일 값, 및 레벨 값을 포함할 수도 있다. MPD 는 매니페스트 파일 (66) 과 같은 매니페스트 파일의 일 예이다. 클라이언트 디바이스 (40) 는 어떻게 다양한 프리젠테이션들의 무비 프래그먼트 박스들을 액세스할지를 결정하기 위해 미디어 프리젠테이션의 MPD 를 취출할 수도 있다. 무비 프래그먼트들은 비디오 파일들의 무비 프래그먼트 박스들 (moof boxes) 에 로케이팅될 수도 있다.

[0061] 매니페스트 파일 (66) (이는, 예를 들어, MPD를 포함할 수도 있음) 이 리프리젠테이션들 (68) 의 세그먼트들의 가용성을 광고할 수도 있다. 다시 말하면, MPD 는 리프리젠테이션들 (68) 중 하나의 리프리젠테이션의 첫 번째 세그먼트가 이용가능하게 되는 벽시계 시간 (wall-clock time) 을 표시하는 정보, 뿐만 아니라 리프리젠테이션들 (68) 내의 세그먼트들의 지속기간들을 표시하는 정보를 포함할 수도 있다. 이런 방식으로, 클라이언트 디바이스 (40) 의 취출 유닛 (52) 은 특정 세그먼트에 선행하는 세그먼트들의 시작 시간 뿐만 아니라 지속기간들에 기초하여, 각각의 세그먼트가 이용가능한 경우를 결정할 수도 있다.

[0062] 인캡슐레이션 유닛 (30) 이 수신된 데이터에 기초하여 NAL 유닛들 및/또는 액세스 유닛들을 비디오 파일로 어셈블링한 후에, 인캡슐레이션 유닛 (30) 은 비디오 파일을 출력을 위해 출력 인터페이스 (32) 로 패스 (pass) 할 수도 있다. 일부 예들에서, 인캡슐레이션 유닛 (30) 은 비디오 파일을 로컬로 (locally) 저장할 수도 있고, 또는, 비디오 파일을 직접 클라이언트 디바이스 (40) 에 전송하기보다는 비디오 파일을 출력 인터페이스 (32) 를 통해 원격 서버에 전송할 수도 있다. 출력 인터페이스 (32) 는, 예를 들어, 송신기, 트랜시버, 예컨대 광학 드라이브, 자기 매체 드라이브 (예컨대, 플로피 드라이브) 와 같은 컴퓨터 판독가능 매체에 데이터를 기록하기 위한 디바이스, 범용 직렬 버스 (USB) 포트, 네트워크 인터페이스, 또는 다른 출력 인터페이스를 포함할 수도 있다. 출력 인터페이스 (32) 는 비디오 파일을, 예를 들어 송신 신호, 자기 매체, 광학 매체, 메모리, 플래시 드라이브, 또는 다른 컴퓨터 판독가능 매체와 같은 컴퓨터 판독가능 매체에 출력한다.

[0063] 네트워크 인터페이스 (54) 는 네트워크 (74) 를 통해 NAL 유닛 또는 액세스 유닛을 수신하고 그 NAL 유닛 또는 액세스 유닛을 취출 유닛 (52) 을 경유하여, 파일 포맷 프로세싱 유닛 (50) 에 제공한다. 파일 포맷 프로세싱 유닛 (50) 은 비디오 파일의 엘리먼트들을 구성성분인 PES 스트림들로 디캡슐레이션하며, PES 스트림들을 패킷해제하여 인코딩된 데이터를 취출하고, 예컨대, 스트림의 PES 패킷 헤더들에 의해 나타내어진 바와 같이, 인코딩된 데이터가 오디오 스트림의 부분인지 또는 비디오 스트림의 부분인지에 따라, 오디오 디코더 (46) 또는 비디오 디코더 (48) 중 어느 일방에 인코딩된 데이터를 전송할 수도 있다. 오디오 디코더 (46) 는 인코딩된 오디오 데이터를 디코딩하고 디코딩된 오디오 데이터를 오디오 출력부 (42) 에 전송하는 한편, 비디오 디코더 (48) 는 인코딩된 비디오 데이터를 디코딩하고 스트림의 복수의 뷰들을 포함할 수도 있는 디코딩된 비디오 데이터를 비디오 출력부 (44) 에 전송한다.

[0064] (예컨대, 인캡슐레이션 유닛 (30) 을 통한) 컨텐츠 준비 디바이스 (20) 및 (예컨대, 파일 포맷 프로세싱 유닛 (50) 을 통한) 클라이언트 디바이스 (40) 는 비디오 컨텐츠를 인캡슐레이션 및/또는 디캡슐레이션하기 위해 다수의 파일 포맷들 중 하나 또는 다수를 활용할 수도 있다. ISOBMFF 는 AVC 파일 포맷과 같은 다수의 코덱 인캡슐레이션 포맷들뿐만 아니라, MPEG-4 파일 포맷, 3GPP 파일 포맷 (3GP) 및 DVB 파일 포맷과 같은 다수의 멀티미디어 컨테이너 포맷들에 대한 기준으로 사용된다.

[0065] 오디오 및 비디오와 같은 연속 미디어 외에도 이미지와 같은 정적 미디어 및 메타 데이터는 ISOBMFF를 준수하는 파일에 저장될 수 있다. ISOBMFF 에 따라 구조화된 파일들은, 로컬 미디어 파일 재생, 원격 파일의 점진적 다운로딩, HTTP 를 통한 동적 적응형 스트리밍 (DASH) 를 위한 세그먼트들, 스트리밍될 컨텐츠에 대한 컨테이너들 및 그것의 패킷화 명령들, 및 수신된 실시간 미디어 스트림들의 레코딩을 포함하는, 많은 목적들을 위해 사용될 수도 있다.

[0066] 박스는 4-캐릭터 코딩된 박스 타입, 박스의 바이트 수 및 페이로드를 포함하여 ISOBMFF의 기본 신택스 구조입니다. ISOBMFF 파일은 박스들의 시퀀스로 이루어지고, 박스들은 다른 박스들을 포함할 수도 있다. 무비 박스 ("moov") 는 파일에 존재하는 연속적 미디어 스트림들에 대한 메타데이터를 포함하고, 이들 각각은 파일에서 트랙으로 표현된다. 트랙에 대한 메타데이터는 트랙 박스 ("trak") 에서 인클로징되는 한편, 트랙의 미디어 컨텐츠는 미디어 데이터 박스 ("mdat") 에서 인클로징되거나 아니면 별개의 파일에서 직접적으로 인클로징된다. 트랙들에 대한 미디어 컨텐츠는 오디오 또는 비디오 액세스 유닛들과 같은 샘플들의 시퀀스를 포함하거나 그러한 샘플들의 시퀀스로 이루어진다.

[0067] ISOBMFF 는 다음과 같은 타입들의 트랙들을 명시한다: 기본적 미디어 스트림을 포함하는 미디어 트랙, 미디어 송신 명령들을 포함하거나 아니면 수신된 패킷 스트림을 표현하는 힌트 트랙, 및 시간-동기화된 메타데이터를 포함하는 타임드 메타데이터 트랙.

[0068] 원래 저장용으로 설계되었지만, ISOBMFF 는 스트리밍을 위해, 예컨대 점진적 다운로드 또는 DASH 를 위해 매우 유용함이 입증되었다. 스트리밍 목적으로, ISOBMFF 에서 정의된 무비 프래그먼트 박스들이 사용될 수 있다.

[0069] 각 트랙에 대한 메타데이터는 트랙에서 사용되는 코딩 또는 인캡슐레이션 포맷 및 그 포맷을 프로세싱하기 위해 사용되는 초기화 데이터를 각각 제공하는, 샘플 디스크립션 엔트리들의 리스트를 포함한다. 각 샘플은 트랙의 샘플 디스크립션 엔트리들 중 하나와 연관된다.

[0070] ISOBMFF 는 다양한 메커니즘들로 샘플-특정적 메타데이터를 명시하는 것을 가능하게 한다. 샘플 테이블 박스 (Sample Table box) ("stbl") 내의 특정 박스들은 공통의 필요성들에 응답하도록 표준화되었다. 예를 들어, 싱크 샘플 박스 ("stss") 는 트랙의 랜덤 액세스 샘플들을 리스팅하기 위해 사용된다. 샘플 그룹핑 메커니즘은 파일에서 샘플 그룹 디스크립션 엔트리로서 명시된 동일한 특성을 공유하는 샘플들의 그룹들 내로의 4-문자 그룹핑 타입에 따른 샘플들의 맵핑을 가능하게 한다. 수개의 그룹핑 타입들이 ISOBMFF 에서 명시되었다.

[0071] ISOBMFF 명세는 DASH 에 있어서의 사용을 위해 6 가지 타입들의 스트림 액세스 포인트 (Stream Access Point; SAP) 들을 명시한다. 처음 2 개의 SAP 타입들 (타입들 1 및 2) 은 H.264/AVC 및 HEVC 에서의 IDR 픽처들에 대응한다. 제 3 SAP 타입 (타입 3) 은 오픈-GOP 랜덤 액세스 포인트들 따라서 HEVC 에서의 BLA 또는 CRA 픽처들에 대응한다. 제 4 SAP 타입 (타입 4) 은 GDR 랜덤 액세스 포인트들에 대응한다.

[0072] ISO/IEC 14496-15 에서, 다양한 샘플 엔트리 타입들 (샘플 엔트리 명칭들로도 지칭됨) 이 명시된다.

[0073] HEVC 파일 포맷 (ISO/IEC 14496-15 의 8 절) 에서, 샘플 엔트리 타입들 'hvc1' 및 'hev1' 이 명시된다. 샘플 엔트리 타입 'hev1' 에 대한 제약은 다음과 같이 명시된다:

[0074] 샘플 엔트리 명칭이 'hev1' 인 경우, 다음이 적용된다:

[0075] 샘플이 랜덤 액세스 포인트인 경우, 그 샘플을 디코딩하는데 필요한 모든 파라미터 세트들은 샘플 엔트리 또는 샘플 자체에 포함되어야 한다.

[0076] 그렇지 않으면 (샘플이 랜덤 액세스 포인트가 아닌 경우), 그 샘플을 디코딩하는데 필요한 모든 파라미터 세트들은 이전의 랜덤 액세스 포인트부터 샘플 자체까지 샘플 엔트리 또는 임의의 샘플들에 포함되어야 한다.

[0077] 이 제약의 목적은 또한, 이전 샘플들의 파라미터 세트들을 검색하고 페치할 필요 없는 랜덤 액세스 포인트인 샘플로부터, 편리한 임의 액세스를 인에이블하는 것이다.

[0078] 계층화된 HEVC (L-HEVC) 파일 포맷 (ISO/IEC 14496-15 의 9 절) 에서, 샘플 엔트리 타입들 'hvc2', 'hev2', 'lhv1' 및 'lhe1' 이 명시된다. 샘플 엔트리 타입 'lhe1' 에 대한 제약은 다음과 같이 명시된다:

[0079] 샘플 엔트리 명칭이 'hev1' 인 경우, 다음이 적용된다:

[0080] 아래 제약들은 적어도 일부 계층들에서 IRAP 픽처들을 포함하는 액세스 유닛들에서 편리한 랜덤 액세스를 인에이블하기 위해, 대역 외 파라미터 세트들 (샘플 엔트리들에서) 및 대역 내 파라미터 세트들 (샘플들에서) 의 배치에 대한 제한들을 부과한다. 이들 제약들로, 모든 픽처들이 IRAP 픽처들인, 샘플 엔트리들로 초기화되고 액세스 유닛으로부터 전진하는 파일 판독기는 필요로 하는 모든 파라미터 세트들을 가질 것이다.

[0081] 특정 트랙의 임의의 특정 샘플에 대해, 다른 트랙에서 시간적으로 병치된 (collocated) 샘플은 이 특정 샘플의 것과 동일한 디코딩 시간을 갖는 것으로 정의된다.

[0082] 주어진 샘플, 트랙 및 계층의 IRAP 픽처의 경우, IRAP 픽처를 디코딩하는데 필요한 각 파라미터 세트는 다음 중 하나에 포함되어야 한다:

[0083] a. 소정 트랙 내의 소정 샘플에 적용되는 샘플 엔트리

[0084] b. 소정 계층의 참조 계층을 운반하는 트랙의 초기 샘플의 샘플 엔트리, 여기서, 초기 샘플은 시간적으로 병치된 샘플이 참조 계층의 IRAP 픽처를 포함하는 경우, 소정의 샘플의 시간적으로 병치된 샘플, 또는 참조 계층의 IRAP 픽처를 포함하는 이전 샘플임

[0085] c. 가능하면 추출기들을 사용함으로써, 소정 샘플 자체

[0086] d. 존재하는 경우, 가능하면 추출기들을 사용함으로써, 소정 계층의 참조 계층들을 운반하는 트랙들의 임의의 시간적으로 병치된 샘플

[0087] 주어진 샘플, 트랙 및 계층의 비-IRAP 픽처의 경우, 그 픽처를 디코딩하는데 필요한 각 파라미터 세트는 다음 중 하나에 포함되어야 한다:

[0088] a. 소정 트랙 내의 소정 샘플에 적용되는 샘플 엔트리

[0089] b. 소정 계층의 참조 계층을 운반하는 트랙의 초기 샘플의 샘플 엔트리, 여기서, 초기 샘플은 시간적으로 병치된 샘플이 참조 계층의 IRAP 픽처를 포함하는 경우, 소정의 샘플의 시간적으로 병치된 샘플, 또는 참조 계층의 IRAP 픽처를 포함하는 이전 샘플임

[0090] c. 가능하면 추출기들을 사용함으로써, 소정 계층에서 IRAP 픽처를 포함하는 이전 샘플부터 소정 샘플 자체까지, 소정 트랙에서 임의의 샘플들

[0091] d. 존재하는 경우, 가능하면 추출기들을 사용함으로써, 소정 계층에서 IRAP 픽처를 포함하는 이전 샘플의 시간적으로 병치된 샘플부터 소정 샘플의 시간적으로 병치된 샘플까지, 소정 계층의 참조 계층을 운반하는 트랙에서 임의의 샘플들

[0092] 이 제약의 목적은 또한, 편리한 임의 액세스를 인에이블하는 것이지만, 제약의 설명의 부분으로서 앞서 상세히 설명된 것과 같이, 다중 계층들을 포함하는 비트스트림에 대한 것이다.

[0093] (이하 카피된) ISO/IEC 14496-15 의 표 10 은 HEVC 및 L-HEVC 트랙들에 대한 샘플 엔트리들, 구성들 및 L-HEVC 툴들의 모든 가능한 사용들을 보여준다:

샘플 엔트리 명칭	구성 기록들	의미
'hvc1' 또는 'hev1'	오직 HEVC 구성만	0 보다 큰 nuh_layer_id 를 갖는 NAL 유닛들이 없는 플레인 HEVC 트랙; 추출기들 및 수집기들은 존재하지 않아야 한다.
'hvc1' 또는 'hev1'	HEVC 및 L-HEVC 구성들	0 과 동일한 nuh_layer_id 를 갖는 NAL 유닛들과 0 보다 큰 nuh_layer_id 를 갖는 NAL 유닛들 양자를 갖는 L-HEVC 트랙; 추출기들 및 수집기들은 존재하지 않아야 한다.
'hvc2' 또는 'hev2'	오직 HEVC 구성만	0 보다 큰 nuh_layer_id 를 갖는 NAL 유닛들이 없는 플레인 HEVC 트랙; 추출기들이 존재하고 NAL 유닛들을 참조하는데 사용될 수도 있다; 수집기들은 NAL 유닛들을 포함하고 참조하기 위해 존재할 수도 있다.
'hvc2' 또는 'hev2'	HEVC 및 L-HEVC 구성들	0 과 동일한 nuh_layer_id 를 갖는 NAL 유닛들과 0 보다 큰 nuh_layer_id 를 갖는 NAL 유닛들 양자를 갖는 L-HEVC 트랙; 추출기들 및 수집기들이 존재할 수도 있다; 추출기들은 임의의 NAL 유닛들을 참조할 수도 있다; 수집기들은 임의의 NAL 유닛들을 포함하고 참조하는 양자를 수행할 수도 있다.
'lhv1', 'lhe1'	오직 L-HEVC 구성만	0 보다 큰 nuh_layer_id 를 갖는 NAL 유닛들을 가지고 0 과 동일한 nuh_layer_id 를 갖는 NAL 유닛들이 없는 L-HEVC 트랙; 추출기들이 존재하고 NAL 유닛들을 참조하는데 사용될 수도 있다; 수집기들은 NAL 유닛들을 포함하고 참조하기 위해 존재할 수도 있다.

[0094]

[0095] 본 개시는 (예컨대, ISO/IEC 14496-15 의 8 절 및 9 절에서) HEVC 및 L-HEVC 에 대한 파일 포맷들에서 샘플 입력 타입들의 현재 설계가 다양한 문제들을 제시할 수도 있음을 인식한다. 예를 들어:

[0096] 제 1 잠재적 문제를 기술하기 위해, ISO/IEC 14496-15 상태들의 표 10 의 로우 2 는 다음을 언급하는 것에 유의한다:

'hvc1' 또는 'hev1'	HEVC 및 L-HEVC 구성들	0 과 동일한 nuh_layer_id 를 갖는 NAL 유닛들과 0 보다 큰 nuh_layer_id 를 갖는 NAL 유닛들 양자를 갖는 L-HEVC 트랙; 추출기들 및 수집기들은 존재하지 않아야 한다.

[0097]

[0098] 한편, ISO/IEC 14496-15 의 표 10 의 로우 4 는 다음을 언급한다:

'hvc2' 또는 'hev2'	HEVC 및 L-HEVC 구성들	0 과 동일한 nuh_layer_id 를 갖는 NAL 유닛들과 0 보다 큰 nuh_layer_id 를 갖는 NAL 유닛들 양자를 갖는 L-HEVC 트랙; 추출기들 및 수집기들이 존재할 수도 있다; 추출기들은 임의의 NAL 유닛들을 참조할 수도 있다; 수집기들은 임의의 NAL 유닛들을 포함하고 참조하는 양자를 수행할 수도 있다.

[0099]

[0100] 샘플 엔트리 타입이 'hvc1', 'hev1', 'hvc2' 또는 'hev2' 이고 동일한 엔트리가 양자의 HEVC 및 L-HEVC 구성들을 포함하는 경우, 트랙은 기본 계층과 하나 이상의 향상 계층들 양자의 데이터를 운반한다.

[0101] 그러나, 편리한 랜덤 액세스를 인에이블하지만 다중 계층들을 포함하는 비트스트림들에 대한 제약은 오직 샘플 엔트리 타입 'lhe1' 에 대해서만 명시된다. 이는 다중 계층 L-HEVC 비트스트림이 임의의 샘플 엔트리 타입들 'hvc1', 'hev1', 'hvc2' 및 'hev2' 을 사용하여 저장되는 경우, 편리한 랜덤 액세스가 (이전 샘플들로부터의 파라미터 세트들을 검색하고 페치하는 필요 없이) 인에이블되는 것을 보장하고 표시하는 방법은 없음을 의미한다.

[0102] 본 개시의 기술들은 앞서 논의된 제 1 의 잠재적인 문제를 어드레싱하는데 사용될 수도 있다. 특히, 일 예에서, 다중 계층들의 데이터를 포함하는 트랙에 대한 이전 샘플들로부터의 파라미터 세트들을 검색하고 페치할 필요 없이, 편리한 랜덤 액세스를 인에이블하는 제약이 샘플 엔트리 타입이 양자의 HEVC 및 L-HEVC 구성들을 포함할 때 샘플 엔트리 타입들 'hev1' 및 'hev2' 에 대해 특정된다. 따라서, 컨텐츠 준비 디바이스 (20) 는 필수 파라미터 세트들에 'hev1' 및 'hev2' 의 샘플 엔트리 타입들을 갖는 샘플들 및/또는 샘플 엔트리들이 제공되는 것을 보장하여, 이전 샘플들로부터의 파라미터 세트들의 검색 및 페치가 요구되지 않도록 할 수도 있다.

마찬가지로, 클라이언트 디바이스 (40) 는 'hev1' 또는 'hev2' 의 샘플 엔트리 타입을 갖는 샘플 엔트리 및 샘플을 취출하고 비디오 디코딩 순서로 임의의 이전 샘플들의 파라미터 세트들을 취출하지 않고 랜덤 액세스를 수행할 수도 있다.

[0103] 제 2 잠재적 문제를 기술하기 위해, ISO/IEC 14496-15 상태들의 표 10 의 로우 3 는 다음을 언급하는 것에 유의한다:

'hvc2' 또는 'hev2'	오직 HEVC 구성만	0 보다 큰 nuh_layer_id 를 갖는 NAL 유닛들이 없는 플레인 HEVC 트랙; 추출기들이 존재하고 NAL 유닛들을 참조하는데 사용될 수도 있다; 수집기들은 NAL 유닛들을 포함하고 참조하기 위해 존재할 수도 있다.

[0104]

[0105] 샘플 엔트리 타입이 'hvc2' 또는 'hev2' 이고, 동일한 엔트리가 오직 HEVC 구성만을 포함하는 경우, 다음 시나리오들 중 하나가 발생할 수도 있다: (1) 트랙은 전체 단일 계층 HEVC 비트스트림을 운반하며 - 여기서 VCL NAL 유닛들 모두는 0 과 nuh_layer_id 을 갖는다 - , 일부 추출기들 및/또는 수집기들이 존재한다; 또는 (2) 트랙은 그러한 단일 계층 HEVC 비트스트림의 서브 세트를 운반하고, 서브세트는 (추출기들 및/또는 수집기들이 존재하는지 여부에 관계없이) 오직 0 보다 큰 TemporalId 를 갖는 VCL NAL 유닛들을 포함한다.

[0106] 그러나, 샘플 엔트리 타입들 'hvc2' 및 'hev2' 중 어느 것도 샘플 엔트리 타입 'hev1' 관해서도 유사하게 편리한 랜덤 액세스를 인에이블하도록 제한되지 않았다. 이는 위의 2 개의 시나리오들에 대하여, (이전 샘플들에서의 파라미터 세트들을 검색하고 페치할 필요 없이) 편리한 랜덤 액세스가 인에이블되는 것을 보장하고 표시하는 방법이 없는 것을 의미한다.

[0107] 본 개시는 또한, 제 2 잠재적 문제를 해결하는데 사용될 수도 있는 기술들을 기술한다. 특히, 오직 전체 단일 계층 HEVC 비트스트림 또는 그 서브세트만을 포함하는 트랙에 대한 이전 샘플들로부터의 파라미터 세트들을 검색하고 페치할 필요 없이, 편리한 랜덤 액세스를 인에이블하는 제약이 명시되어, 샘플 엔트리 타입 'hev2' 은

샘플 엔트리가 오직 HEVC 구성만을 포함할 때 사용될 것이다. 따라서, 샘플 엔트리가 오직 HEVC 구성만을 포함 할 때, 컨텐츠 준비 디바이스 (20) 는 'hev2' 의 샘플 엔트리 타입을 특정할 수도 있다. 마찬가지로, 클라이언트 디바이스 (40) 는 샘플 엔트리 타입 'hev2' 를 갖는 샘플에 대해 편리한 랜덤 액세스가 인에이블되고, 또한 샘플이 L-HEVC 구성보다는 오직 HEVC 구성에 따라 인코딩된 비디오 데이터를 포함하는 것으로 결정할 수도 있다.

[0108] 제 3 의 잠재적 문제점을 설명하기 위해, 다중 계층 L-HEVC 비트스트림의 각각의 전체 계층이 현재의 사양에 따라 분리된 트랙에 저장될 때, 베이스 트랙이 오직 HEVC 구성만을 갖는 샘플 엔트리 타입을 사용하고, 다른 트랙들은 오직 L-HEVC 구성만을 갖는 샘플 엔트리 타입 'lhv1' 을 사용하는 것이 가능하고, 기본 트랙은 오직 HEVC 구성만을 갖는 샘플 엔트리 타입 'hvc1' 을 사용하고 다른 트랙들은 오직 L-HEVC 구성만을 갖는 샘플 엔트리 타입 'lhe1' 을 사용하는 것이 또한 가능한 것이 언급된다.

[0109] 제 1 시나리오에서, 편리한 랜덤 액세스는 다른 트랙들에 대해서가 아니라 기본 트랙에 대하여 표시된다. 제 2 시나리오에서, 편리한 랜덤 액세스는 기본 트랙들에 대해서 표시되는 것이 아니라 다른 트랙들에 대하여 표시된다. 제 1 시나리오에 대해 그럴듯한 구실이 존재할 수도 있고, 예컨대 기본 트랙은 편리한 랜덤 액세스를 인에이블하는 오버헤드를 정당화하기에 충분히 중요하고, 강화 계층들을 운반하는 트랙들에 대하여 적은 오버헤드로 편리한 랜덤 액세스를 포기하는 절충안이 양호한 트레이드오프로 고려된다. 그러나, 제 2 시나리오는 강화 계층들을 운반하는 트랙들에 대한 편리한 랜덤 액세스를 인에이블하는 것이 기본 트랙에 대한 편리한 랜덤 액세스 인에이블하는 것을 효과적으로 요구할 것이기 때문에 타당하지 않으며, 기본 트랙에 대해 인에이블된 경우에, 올바른 샘플 엔트리 타입, 즉 'hev1' 을 사용함으로써 이를 표시하지 않을 이유가 없다.

[0110] 마찬가지로, 제 3 의 잠재적 문제는 또한, 다중의 시간적 서브 계층들을 포함하는 단일 계층 HEVC 비트스트림이 다중 트랙들에 의해 운반될 때 적용될 수도 있는데, 여기서 최저 서브 계층을 운반하는 트랙 (그 VCL NAL 유닛들은 0 과 동일한 TemporalId 를 갖는다) 은 편리한 랜덤 액세스가 인에이블되는 것을 나타내는 샘플 엔트리 타입을 사용하는 반면, 다른 트랙은 편리한 랜덤 액세스가 인에이블되는 것을 표시하지 않는 샘플 엔트리 유형을 사용하며, 그 반대도 가능하다.

[0111] 본 개시의 기술들은 또한 제 3 의 잠재적 문제점을 해결할 수도 있다. 특히, 제약은 (HEVC, L-HEVC, 또는 임의의 다른 코덱들로 코딩된) 단일 계층 또는 다중 계층 비트스트림을 운반하는 모든 트랙들에 대해, 모든 트랙들이 (HEVC 및/또는 L-HEVC 구성들의 존재와 함께) 편리한 랜덤 액세스가 인에이블되는 것을 나타내는 샘플 엔트리 타입들을 사용하거나 또는 모든 트랙들이 (HEVC 및/또는 L-HEVC 구성들의 존재와 함께) 편리한 랜덤 액세스가 인에이블되는 것을 나타내지 않는 샘플 엔트리 타입들을 사용하는 것을 요구하는 것으로 명시될 수도 있다. 따라서, 컨텐츠 준비 디바이스 (20) 는 단일 계층 또는 다중 계층 비트스트림을 운반하는 트랙들에 대해, 모든 트랙들이 편리한 랜덤 액세스가 인에이블되는 것을 나타내는 샘플 엔트리 타입들을 사용하거나, 또는 모든 트랙들이 편리한 랜덤 액세스가 인에이블되지 않는 것을 나타내는 샘플 엔트리 타입들을 사용하는 것을 보장할 수도 있다. 이러한 방식으로, 편리한 랜덤 액세스가 트랙들 중 하나에 대해 인에이블된다면, 클라이언트 디바이스 (40) 는 트랙들의 각각에 대해 편리한 랜덤 액세스가 인에이블되는 것으로 결정할 수도 있다.

[0112] 대안적으로, 제약은 (HEVC, L-HEVC, 또는 임의의 다른 코덱들로 코딩된) 단일 계층 또는 다중 계층 비트스트림을 운반하는 모든 트랙들에 대해, (그 VCL NAL 유닛들이 0 과 동일한 TemporalId 을 갖는) 최저 서브 계층을 운반하는 트랙이 편리한 랜덤 액세스가 인에이블되는 것을 나타내는 샘플 엔트리 타입을 (HEVC 및/또는 L-HEVC 구성들의 존재와 함께) 사용할 때, 모든 다른 트랙들이 (HEVC 및/또는 L-HEVC 구성들의 존재와 함께) 편리한 랜덤 액세스가 인에이블되는 것을 나타내는 샘플 엔트리 타입들을 사용해야만 하는 것을 요구하는 것으로 명시될 수도 있다. 따라서, 컨텐츠 준비 디바이스 (20) 는 단일 계층 또는 다중 계층 비트스트림을 운반하는 트랙들에 대해, 기본 계층의 최저 서간 서브-계층을 운반하는 트랙이 편리한 랜덤 액세스가 인에이블되는 것을 나타낼 때, 모든 트랙들이 편리한 랜덤 액세스가 인에이블되는 것을 나타내는 샘플 엔트리 타입들을 사용하는 것, 또는 대안적으로 편리한 랜덤 액세스가 기본 계층의 최저 시간 서브-계층에 대하여 인에이블되지 않을 경우, 트랙들 중 어느 것도 편리한 랜덤 액세스가 인에이블되는 것을 나타내는 샘플 엔트리 타입들을 사용하지 않는 것을 보장할 수도 있다. 이러한 방식으로, 편리한 랜덤 액세스가 기본 계층의 최저 시간 서브 계층을 포함하는 트랙에 대해 인에이블되면, 클라이언트 디바이스 (40) 는 편리한 랜덤 액세스가 트랙들의 각각에 대해 인에이블되는 것을 결정할 수도 있다.

[0113] 클라이언트 디바이스 (40) 는 다양한 방식으로 편리한 랜덤 액세스를 수행하기 위해 본 개시의 기술들을 사용할 수도 있다. 예를 들어, 클라이언트 디바이스 (40) 는 DASH 와 같은 네트워크 스트리밍 프로토콜을 사용하여

서버 디바이스 (60) 로부터 데이터를 요청함으로써 편리한 랜덤 액세스를 수행할 수도 있다. 다른 예들에서, 클라이언트 디바이스 (40) 는 디지털 범용 디스크, 블루레이 디스크, 하드 드라이브, 플래시 메모리 등과 같은 고정된 컴퓨터 판독가능 저장 매체로부터 미디어 데이터를 취출하기 위한 편리한 랜덤 액세스를 수행하기 위해 본 개시의 기술들을 사용할 수도 있다. 따라서, 도 1 은 네트워크 기반 스트리밍을 포함하는 예를 도시하지만, 본 개시의 기술들은 다른 시나리오들 및 컨텍스트들에도 마찬가지로 적용될 수도 있음을 이해해야 한다.

[0114] 도 2 는 도 1 의 취출 유닛 (52) 의 컴포넌트들의 예시적인 세트를 더 상세히 설명하는 블록도이다. 이 예에서, 취출 유닛 (52) 은 eMBMS 미들웨어 유닛 (100), DASH 클라이언트 (110) 및 미디어 애플리케이션 (112) 을 포함한다.

[0115] 이 예에서, eMBMS 미들웨어 유닛 (100) 은 eMBMS 수신 유닛 (106), 캐시 (104), 및 서버 유닛 (102) 을 더 포함한다. 이 예에서, eMBMS 수신 유닛 (106) 은 예컨대, FLUTE (File Delivery over Unidirectional Transport) 에 따라 eMBMS 를 통해 데이터를 수신하도록 구성되고, 이는 http://tools.ietf.org/html/rfc6726 에서 입수가능한 T. Paila 등의 "FLUTE File Delivery over Unidirectional Transport", Network Working Group, RFC 6726, 2012 년 11월 에 기술된다. 즉, eMBMS 수신 유닛 (106) 은 예를 들어 BM-SC 로서 작용할 수도 있는 서버 디바이스 (60) 로부터 브로드캐스트를 통해 파일들을 수신할 수도 있다.

[0116] eMBMS 미들웨어 유닛 (100) 이 파일들에 대한 데이터를 수신할 때, eMBMS 미들웨어 유닛은 수신된 된 데이터를 캐시 (104) 에 저장할 수도 있다. 캐시 (104) 는 플래시 메모리, 하드 디스크, RAM, 또는 임의의 다른 적절한 저장 매체와 같은 컴퓨터-판독가능 저장 매체를 포함할 수도 있다.

[0117] 로컬 서버 유닛 (102) 은 DASH 클라이언트 (110) 를 위한 서버로서 작용할 수도 있다. 예를 들어, 로컬 서버 유닛 (102) 은 DASH 클라이언트 (110) 에 MPD 파일 또는 다른 매니페스트 파일을 제공할 수도 있다. 로컬 서버 유닛 (102) 은 세그먼트들이 취출될 수 있는 하이퍼링크들뿐만 아니라, MPD 파일 내의 세그먼트들에 대한 가용 시간들을 광고할 수도 있다. 이들 하이퍼링크들은 클라이언트 디바이스 (40) 에 대응하는 로컬호스트 어드레스 프리픽스 (예를 들어, IPv4 의 경우 127.0.0.1) 를 포함할 수도 있다. 이러한 방식으로, DASH 클라이언트 (110) 는 HTTP GET 또는 부분 GET 요청들을 사용하여 로컬 서버 유닛 (102) 으로부터 세그먼트들을 요청할 수도 있다. 예를 들어, 링크 http://127.0.0.1/rep1/seg3 에서 입수가능한 세그먼트의 경우, DASH 클라이언트 (110) 는 http://127.0.0.1/rep1/seg3 에 대한 요청을 포함하는 HTTP GET 요청을 구성하고 그 요청을 로컬 서버 유닛 (102) 에 제출한다. 로컬 서버 유닛 (102) 은 캐시 (104) 로부터 요청된 데이터를 취출하고, 그러한 요청들에 응답하여 데이터를 DASH 클라이언트 (110) 에 제공할 수도 있다.

[0118] 도 3 은 예시적인 멀티미디어 컨텐츠 (120) 의 엘리먼트들을 도시하는 개념도이다. 멀티미디어 컨텐츠 (120) 는 멀티미디어 컨텐츠 (64) (도 1), 또는 저장 매체 (62) 에 저장된 다른 멀티미디어 컨텐츠에 대응할 수도 있다. 도 3 의 예에서는, 멀티미디어 컨텐츠 (120) 는 미디어 프리젠테이션 디스크립션 (MPD) (122) 및 복수의 리프리젠테이션들 (124A-124N) (리프리젠테이션들 (124)) 을 포함한다. 리프리젠테이션 (124A) 은 옵션의 헤더 데이터 (126) 및 세그먼트들 (128A-128N) (세그먼트들 (128)) 을 포함하는 한편, 리프리젠테이션 (124N) 은 옵션의 헤더 데이터 (130) 및 세그먼트들 (132A-132N) (세그먼트들 (132)) 을 포함한다. 문자 N 은 편의를 위하여 리프리젠테이션들 (124) 의 각각에서 최종 무비 프래그먼트를 지정하기 위해 이용된다. 일부 예들에서, 리프리젠테이션들 (124) 사이에는 상이한 개수의 무비 프래그먼트들이 존재할 수도 있다.

[0119] MPD (122) 는 리프리젠테이션들 (124) 로부터 별개인 데이터 구조를 포함할 수도 있다. MPD (122) 는 도 1 의 매니페스트 파일 (66) 에 대응할 수도 있다. 마찬가지로, 리프리젠테이션들 (124) 은 도 2 의 리프리젠테이션들 (68) 에 대응할 수도 있다. 일반적으로, MPD (122) 는 리프리젠테이션들 (124) 의 특성들을 일반적으로 기술하는 데이터, 예컨대 코딩 및 렌더링 특성들, 적응 세트들, MPD (122) 가 대응하는 프로파일, 텍스트 타입 정보, 카메라 각도 정보, 등급 정보, 트릭 모드 정보 (예를 들어, 시간적 서브-시퀀스들을 포함하는 리프리젠테이션들을 표시하는 정보) 및/또는 원격 주기들을 취출하기 위한 (예를 들어, 재생 도중에 미디어 컨텐츠 내로의 타겟화된 광고 삽입을 위한) 정보를 포함할 수도 있다.

[0120] 존재할 경우에, 헤더 데이터 (126) 는 세그먼트들 (128) 의 특성들, 예를 들어, 랜덤 액세스 포인트들 (RAP) 들 (스트림 액세스 포인트 (SAP) 이라고도 지칭됨) 의 시간적 로케이션들, 세그먼트들 (128) 중 어느 것이 랜덤 액세스 포인트들을 포함하는지, 세그먼트들 (128) 내의 랜덤 액세스 포인트들에 대한 바이트 오프셋들, 세그먼트들 (128) 의 유니폼 리소스 로케이터 (URL) 들, 또는 세그먼트들 (128) 의 다른 양태들을 기술할 수도 있다. 존재할 경우에, 헤더 데이터 (130) 는 세그먼트들 (132) 에 대한 유사한 특성들을 기술할 수도 있다. 추

가적으로 또는 대안적으로, 이러한 특성들은 MPD (122) 내에 완전하게 포함될 수도 있다.

[0121] 세그먼트들 (128, 132) 은 하나 이상의 코딩된 비디오 샘플들을 포함하는데, 이들 각각은 비디오 데이터의 프레임들 또는 슬라이스들을 포함할 수도 있다. 세그먼트들 (128) 의 코딩된 비디오 샘플들 각각은 유사한 특성들, 예를 들어, 높이, 폭, 및 대역폭 요건들을 가질 수도 있다. 비록 이러한 데이터가 도 3 의 예에서는 예시되지 않지만, 이러한 특성들은 MPD (122) 의 데이터에 의하여 기술될 수도 있다. MPD (122) 는, 본 개시에서 기술되는 시그널링된 정보 중 임의의 것 또는 전부가 추가된, 3GPP 명세에 의하여 기술되는 바와 같은 특성들을 포함할 수도 있다.

[0122] 세그먼트들 (128, 132) 각각은 고유한 유니폼 리소스 로케이터 (URL) 와 연관될 수도 있다. 따라서, 세그먼트들 (128, 132) 의 각각은 스트리밍 네트워크 프로토콜, 예컨대 DASH 를 이용하여 독립적으로 취출가능할 수도 있다. 이러한 방식으로, 목적지 디바이스, 예컨대 클라이언트 디바이스 (40) 는, HTTP GET 요청을 이용하여 세그먼트들 (128 또는 1324) 을 취출할 수도 있다. 일부 예들에서, 클라이언트 디바이스 (40) 는 HTTP 부분 GET 요청들을 이용하여 세그먼트들 (128 또는 132) 의 특정 바이트 범위들을 취출할 수도 있다.

[0123] 도 4 는 리프리젠테이션의 세그먼트, 이를테면 도 3 의 세그먼트들 (114, 124) 중 하나의 세그먼트에 대응할 수도 있는, 예시적인 비디오 파일 (150) 의 엘리먼트들을 도시하는 블록도이다. 세그먼트들 (128, 132) 의 각각은 도 4 의 예에서 도시된 데이터의 배열에 실질적으로 부합하는 데이터를 포함할 수도 있다. 비디오 파일 (150) 은 세그먼트를 인캡슐레이션하는 것으로 말해질 수도 있다. 위에서 설명된 바와 같이, ISO 기반 미디어 파일 포맷 및 그 확장물들에 따른 비디오 파일들은 데이터를 "박스들"이라고 지칭되는 일련의 오브젝트들에 저장한다. 도 4 의 예에서, 비디오 파일 (150) 은 파일 타입 (FTYP) 박스 (152), 무비 (MOOV) 박스 (154), 세그먼트 인덱스 (sidx) 박스들 (162), 무비 프래그먼트 (MOOF) 박스들 (164), 및 무비 프래그먼트 랜덤 액세스 (MFRA) 박스 (166) 를 포함한다. 도 4 가 비디오 파일의 일 예를 나타내지만, 다른 미디어 파일들이 ISO 기본 미디어 파일 포맷 및 그것의 확장에 따른, 비디오 파일 (150) 의 데이터와 유사하게 구조화된 다른 타입들의 미디어 데이터 (예컨대, 오디오 데이터, 타임드 텍스트 데이터 등) 을 포함할 수도 있다는 것이 이해되어야 한다.

[0124] 파일 타입 (FTYP) 박스 (152) 는 비디오 파일 (150) 에 대한 파일 타입을 일반적으로 기술한다. 파일 타입 박스 (152) 는 비디오 파일 (150) 에 대한 최상의 이용을 서술하는 사양을 식별하는 데이터를 포함할 수도 있다. 파일 타입 박스 (152) 는 대안적으로 MOOV 박스 (154), 무비 프래그먼트 박스들 (164), 및/또는 MFRA 박스 (166) 전에 배치될 수도 있다.

[0125] 일부 예들에서, 세그먼트, 이를테면 비디오 파일 (150) 은, FTYP 박스 (152) 전에 MPD 업데이트 박스 (도시되지 않음) 를 포함할 수도 있다. MPD 업데이트 박스는 비디오 파일 (150) 을 포함하는 리프리젠테이션에 대응하는 MPD가 그 MPD를 업데이트하기 위한 정보와 함께 업데이트될 것임을 나타내는 정보를 포함할 수도 있다. 예를 들어, MPD 업데이트 박스는 MPD를 업데이트하는데 사용될 리소스에 대한 URI 또는 URL을 제공할 수도 있다. 다른 예로서, MPD 업데이트 박스는 MPD를 업데이트하기 위한 데이터를 포함할 수도 있다. 일부 예들에서, MPD 업데이트 박스는 비디오 파일 (150) 의 세그먼트 타입 (STYP) 박스 (도시되지 않음) 를 바로 뒤따를 수도 있으며, 여기서 STYP 박스는 비디오 파일 (150) 에 대한 세그먼트 타입을 정의할 수도 있다. 이하 더 상세히 논의되는 도 7 은 MPD 업데이트 박스에 대한 추가의 정보를 제공한다.

[0126] MOOV 박스 (154) 는, 도 4 의 예에서, 무비 헤더 (MVHD) 박스 (156), 트랙 (TRAK) 박스 (158), 및 하나 이상의 무비 확장 (MVEX) 박스들 (160) 을 포함한다. 일반적으로, MVHD 박스 (156) 는 비디오 파일 (150) 의 일반 특성들을 기술할 수도 있다. 예를 들어, MVHD 박스 (156) 는 비디오 파일 (150) 이 원래 생성되었을 때, 비디오 파일 (150) 이 최종 수정되었을 때, 비디오 파일 (150) 에 대한 타임스케일, 비디오 파일 (150) 에 대한 재생의 지속기간을 기술하는 데이터, 또는 비디오 파일 (150) 을 일반적으로 기술하는 다른 데이터를 포함할 수도 있다.

[0127] TRAK 박스 (158) 는 비디오 파일 (150) 의 트랙에 대한 데이터를 포함할 수도 있다. TRAK 박스 (158) 는 TRAK 박스 (158) 에 대응하는 트랙의 특성들을 기술하는 트랙 헤더 (TKHD) 박스를 포함할 수도 있다. 일부 예들에서, TRAK 박스 (158) 는 코딩된 비디오 픽처들을 포함하는 한편, 다른 예들에서, 트랙의 코딩된 비디오 픽처들은 TRAK 박스 (158) 및/또는 sidx 박스들 (162) 의 데이터에 의해 참조될 수도 있는 무비 프래그먼트들 (164) 에 포함될 수도 있다.

[0128] 일부 예들에서, 비디오 파일 (150) 은 하나를 초과하는 트랙을 포함할 수도 있다. 따라서, MOOV 박스 (154)

는 트랙들의 수와 동일한 수의 TRAK 박스들을 비디오 파일 (150) 내에 포함할 수도 있다. TRAK 박스 (158) 는 비디오 파일 (150) 의 대응하는 트랙의 특성들을 기술할 수도 있다. 예를 들어, TRAK 박스 (158) 는 대응하는 트랙에 대한 시간적 및/또는 공간적 정보를 기술할 수도 있다. MOOV 박스 (154) 의 TRAK 박스 (158) 에 유사한 TRAK 박스가, 인캡슐레이션 유닛 (30) (도 3) 이 비디오 파일, 이를테면 비디오 파일 (150) 에 파라미터 세트 트랙을 포함시키는 경우, 파라미터 세트 트랙의 특성들을 기술할 수도 있다. 인캡슐레이션 유닛 (30) 은 파라미터 세트 트랙을 기술하는 TRAK 박스 내의 파라미터 세트 트랙에서 시퀀스 레벨 SEI 메시지들의 존재를 시그널링할 수도 있다.

[0129] MVEX 박스들 (160) 은, 예컨대, 비디오 파일 (150) 이, 만약 있다면, MOOV 박스 (154) 내에 포함된 비디오 데이터 외에도, 무비 프래그먼트들 (164) 을 포함한다는 것을 시그널링하기 위해, 대응하는 무비 프래그먼트들 (164) 의 특성들을 기술할 수도 있다. 비디오 데이터를 스트리밍하는 맥락에서, 코딩된 비디오 픽처들은 MOOV 박스 (154) 내에 보다는 무비 프래그먼트들 (164) 내에 포함될 수도 있다. 따라서, 모든 코딩된 비디오 샘플들은 MOOV 박스 (154) 내가 아니라, 무비 프래그먼트들 (164) 내에 포함될 수도 있다.

[0130] MOOV 박스 (154) 는 비디오 파일 (150) 에서의 무비 프래그먼트들 (164) 의 수와 동일한 수의 MVEX 박스들 (160) 을 포함할 수도 있다. MVEX 박스들 (160) 의 각각은 무비 프래그먼트들 (164) 중 대응하는 무비 프래그먼트의 특성들을 기술할 수도 있다. 예를 들어, 각각의 MVEX 박스는 무비 프래그먼트들 (164) 중 대응하는 무비 프래그먼트에 대한 시간적 지속기간을 기술하는 MEHD (movie extends header box) 박스를 포함할 수도 있다.

[0131] 위에서 언급된 바와 같이, 인캡슐레이션 유닛 (30) 은 실제 코딩된 비디오 데이터를 포함하지 않는 비디오 샘플에 시퀀스 데이터 세트를 저장할 수도 있다. 비디오 샘플은 특정 시간 인스턴스에서의 코딩된 픽처의 리프리젠테이션인 액세스 유닛에 일반적으로 대응할 수도 있다. AVC 의 맥락에서, 코딩된 픽처는 액세스 유닛의 모든 화소들을 구축하기 위한 정보를 포함하는 하나 이상의 VCL NAL 유닛들과, 다른 연관된 비-VCL NAL 유닛들, 이를테면 SEI 메시지들을 포함한다. 따라서, 인캡슐레이션 유닛 (30) 은 시퀀스 레벨 SEI 메시지들을 포함할 수도 있는 시퀀스 데이터 세트를 무비 프래그먼트들 (164) 중 하나의 무비 프래그먼트에 포함시킬 수도 있다. 인캡슐레이션 유닛 (30) 은, 무비 프래그먼트들 (164) 중 상기 하나의 무비 프래그먼트에 대응하는 MVEX 박스들 (160) 중 하나의 MVEX 박스 내에서 무비 프래그먼트들 (164) 중 하나의 무비 프래그먼트에 존재하는 것으로서의 시퀀스 데이터 세트 및/또는 시퀀스 레벨 SEI 메시지들의 존재를 추가로 시그널링할 수도 있다.

[0132] SIDX 박스들 (162) 은 비디오 파일 (150) 의 옵션적 엘리먼트들이다. 다시 말하면, 3GPP 파일 포맷, 또는 다른 이러한 파일 포맷들에 부합하는 비디오 파일들은 반드시 SIDX 박스들 (162) 을 포함하는 것은 아니다. 3GPP 파일 포맷의 예에 따라, SIDX 박스는 세그먼트 (예컨대, 비디오 파일 (150) 내에 포함된 세그먼트) 의 서브-세그먼트를 식별하는데 사용될 수도 있다. 3GPP 파일 포맷은 서브-세그먼트를 "대응하는 미디어 데이터 박스(들)를 갖는 하나 이상의 연속적인 무비 프래그먼트 박스들의 독립식 세트 및 무비 프래그먼트 박스에 의해 참조되는 데이터를 포함하는 미디어 데이터 박스가 그 무비 프래그먼트 박스를 뒤따라야만 하고 동일한 트랙에 관한 정보를 포함하는 다음 무비 프래그먼트 박스에 선행해야만 한다" 로서 정의한다. 3GPP 파일 포맷은 SIDX 박스가 "박스에 의해 기록된 (서브) 세그먼트의 서브세그먼트들에 대한 참조들의 시퀀스를 포함한다"는 것을 또한 표시한다. 참조된 서브세그먼트들은 프리젠테이션 시간에 연속적이다. 마찬가지로, 세그먼트 인덱스 박스에 의해 참조되는 바이트들은 세그먼트 내에서 항상 연속적이다. 참조된 사이즈는 참조된 자료에서의 바이트들의 수의 카운트를 제공한다.

[0133] SIDX 박스들 (162) 은 비디오 파일 (150) 에 포함된 세그먼트의 하나 이상의 서브-세그먼트들을 나타내는 정보를 일반적으로 제공한다. 예를 들면, 이러한 정보는 서브-세그먼트들이 시작하며 그리고/또는 종료하는 재생 시간들, 서브-세그먼트들에 대한 바이트 오프셋들, 서브-세그먼트들이 스트림 액세스 포인트 (SAP) 를 포함하는지 (예컨대 그 SAP로 시작하는지) 의 여부, SAP에 대한 타입 (예컨대, SAP가 순간적 디코더 리프레시 (IDR) 픽처인지, 클린 랜덤 액세스 (CRA) 픽처인지, 브로큰 링크 액세스 (BLA) 픽처인지 등), 서브-세그먼트에서의 SAP의 포지션 (재생 시간 및/또는 바이트 오프셋 측면에서임) 등을 포함할 수도 있다.

[0134] 무비 프래그먼트들 (164) 은 하나 이상의 코딩된 비디오 픽처들을 포함할 수도 있다. 일부 예들에서, 무비 프래그먼트들 (164) 은 하나 이상의 픽처들의 그룹들 (GOP들) 을 포함할 수도 있으며, 그 GOP들의 각각은 다수의 코딩된 비디오 픽처들, 예컨대, 프레임들 또는 픽처들을 포함할 수도 있다. 덧붙여서, 위에서 설명된 바와 같이, 무비 프래그먼트들 (164) 은 일부 예들에서 시퀀스 데이터 세트들을 포함할 수도 있다. 무비 프래그먼트들 (164) 의 각각은 무비 프래그먼트 헤더 박스 (MFHD, 도 4 에는 도시되지 않음) 를 포함할 수도 있다.

MFHD 박스는 대응하는 무비 프래그먼트 박스들의 특성들, 이를테면 무비 프래그먼트 박스들에 대한 시퀀스 번호를 기술할 수도 있다. 무비 프래그먼트들 (164) 은 비디오 파일 (150) 내에 시퀀스 번호의 순서로 포함될 수도 있다.

[0135] MFRA 박스 (166) 는 비디오 파일 (150) 의 무비 프래그먼트들 (164) 내의 랜덤 액세스 포인트들을 기술할 수도 있다. 이는 비디오 파일 (150) 에 의해 인캡슐레이션된 세그먼트 내에서 특정한 시간적 로케이션들 (즉, 재생 시간들) 에 대한 탐색들을 수행하는 것과 같은, 트릭 모드들을 수행하는 것을 지원할 수도 있다. MFRA 박스 (166) 는 일반적으로 옵션적이고 일부 예들에서, 비디오 파일들에 포함될 필요가 없다. 비슷하게, 클라이언트 디바이스, 이를테면 클라이언트 디바이스 (40) 는, 비디오 파일 (150) 의 비디오 데이터를 정확히 디코딩하고 디스플레이하기 위해 MFRA 박스 (166) 를 반드시 참조할 필요는 없다. MFRA 박스 (166) 는 비디오 파일 (150) 의 트랙들의 수와 동일한, 또는 일부 예들에서는, 비디오 파일 (150) 의 미디어 트랙들 (예컨대, 비-힌트 트랙들) 의 수와 동일한 수의 트랙 무비 프래그먼트 박스들 랜덤 액세스 (TFRA) 박스들 (도시되지 않음) 을 포함할 수도 있다.

[0136] 일부 예들에서, 무비 프래그먼트들 (164) 은 하나 이상의 스트림 액세스 포인트들 (SAP들), 이를테면 IDR 픽처들을 포함할 수도 있다. 비슷하게, MFRA 박스 (166) 는 SAP들의 비디오 파일 (150) 내의 로케이션들의 표시들을 제공할 수도 있다. 따라서, 비디오 파일 (150) 의 시간적 서브-시퀀스가 비디오 파일 (150) 의 SAP들로부터 형성될 수도 있다. 시간적 서브-시퀀스는 SAP들에 의존하는 P-프레임들 및/또는 B-프레임들과 같은 다른 픽처들을 또한 포함할 수도 있다. 시간적 서브-시퀀스의 프레임들 및/또는 슬라이스들은 서브-시퀀스의 다른 프레임들/슬라이스들에 의존하는 시간적 서브-시퀀스의 프레임들/슬라이스들이 적절히 디코딩될 수 있도록 세그먼트들 내에 배열될 수도 있다. 예를 들어, 데이터의 계층적 배열에서, 다른 데이터에 대한 예측을 위해 사용되는 데이터가 시간적 서브-시퀀스에 또한 포함될 수도 있다.

[0137] 비디오 파일 (150) 은 또한, 이 예에서 샘플 디스크립션 박스 (168) 를 포함한다. 특히, 샘플 디스크립션 박스 (168) 는 이 예에서 TRAK 박스 (158) 내에 포함된다. 예시적인 샘플 디스크립션 박스 (168) 는 다음과 같이 정의될 수도 있다:

[0138] 샘플 엔트리 및 박스 타입들: 'hvc2', 'hev2', 'lhv1', 'lhe1', 'lhvC'

[0139] ● 컨테이너: 샘플 디스크립션 박스 ('stsd')

[0140] ● 필수: 'hvc1', 'hev1', 'hvc2', 'hev2', 'lhv1' 또는 'lhe1' 샘플 엔트리는 필수적이다

[0141] ● 수량: 하나 이상의 샘플 엔트리들이 존재할 수도 있다

[0142] 샘플 디스크립션 박스 (168) 에 대한 상기 예시적인 정의에서, 샘플 엔트리 명칭이 'lhv1' 이면, array_completeness 의 디폴트 및 필수 값은 모든 타입들의 파라미터 세트들의 어레이들에 대해 1 이고 모든 다른 어레이들에 대해 0 이다. 샘플 엔트리 명칭이 'lhe1' 이면, 모든 어레이들에 대해 array_completeness 의 디폴트 값이 0 이다.

[0143] 샘플 디스크립션 박스 (168) 에 대한 상기 예시적인 정의에서, 샘플 엔트리 명칭이 'hev2' 이고, 샘플 엔트리가 오직 HEVC 구성만을 포함할 때, 샘플 엔트리 명칭 'hev1' 에 대하여 절 8.4.3 에 명시된 것과 동일한 제약들이 적용된다.

[0144] 샘플 디스크립션 박스 (168) 에 대한 상기 예시적인 정의에서, 샘플 엔트리 명칭이 'lhe1' 일 때, 또는 샘플 엔트리 명칭이 'hev1' 또는 'hev2' 이고 샘플 엔트리가 양자의 HEVC 및 L-HEVC 구성들을 포함할 때, 다음이 적용된다:

[0145] 아래 예시적인 제약들은 적어도 일부 계층들에서 IRAP 픽처들을 포함하는 액세스 유닛들에서 편리한 랜덤 액세스를 인에이블하기 위해, 대역 외 파라미터 세트들 (샘플 엔트리들에서) 및 대역 내 파라미터 세트들 (샘플들에서) 의 배치에 대한 제한들을 부과한다. 이들 제약들로, 모든 픽처들이 IRAP 픽처들인, 샘플 엔트리들로 초기화되고 액세스 유닛으로부터 전진하는 파일 판독기는 필요로 하는 모든 파라미터 세트들을 가질 것이다.

[0146] 샘플 디스크립션 박스 (168) 에 대한 상기 예시적인 정의에서, 특정 트랙의 임의의 특정 샘플에 대해, 다른 트랙에서 시간적으로 병치된 (collocated) 샘플은 이 특정 샘플의 것과 동일한 디코딩 시간을 갖는 것으로 정의된다.

[0147] 샘플 디스크립션 박스 (168) 에 대한 상기 예시적인 정의에서, 주어진 샘플, 트랙 및 계층의 IRAP 픽처의 경우,

IRAP 픽처를 디코딩하는데 필요한 각 파라미터 세트는 다음 중 하나에 포함되어야 한다:

[0148] a. 소정 트랙 내의 소정 샘플에 적용되는 샘플 엔트리

[0149] b. 소정 계층의 참조 계층을 운반하는 트랙의 초기 샘플의 샘플 엔트리, 여기서, 초기 샘플은 시간적으로 병치된 샘플이 참조 계층의 IRAP 픽처를 포함하는 경우, 소정의 샘플의 시간적으로 병치된 샘플, 또는 참조 계층의 IRAP 픽처를 포함하는 이전 샘플임

[0150] c. 가능하면 추출기들을 사용함으로써, 소정 샘플 자체

[0151] d. 존재하는 경우, 가능하면 추출기들을 사용함으로써, 소정 계층의 참조 계층들을 운반하는 트랙들의 임의의 시간적으로 병치된 샘플

[0152] 샘플 디스크립션 박스 (168) 에 대한 상기 예시적인 정의에서, 주어진 샘플, 트랙 및 계층의 비-IRAP 픽처의 경우, 그 픽처를 디코딩하는데 필요한 각 파라미터 세트는 다음 중 하나에 포함되어야 한다:

[0153] a. 소정 트랙 내의 소정 샘플에 적용되는 샘플 엔트리

[0154] b. 소정 계층의 참조 계층을 운반하는 트랙의 초기 샘플의 샘플 엔트리, 여기서, 초기 샘플은 시간적으로 병치된 샘플이 참조 계층의 IRAP 픽처를 포함하는 경우, 소정의 샘플의 시간적으로 병치된 샘플, 또는 참조 계층의 IRAP 픽처를 포함하는 이전 샘플임

[0155] c. 가능하면 추출기들을 사용함으로써, 소정 계층에서 IRAP 픽처를 포함하는 이전 샘플부터 소정 샘플 자체 까지, 소정 트랙에서 임의의 샘플들

[0156] d. 존재하는 경우, 가능하면 추출기들을 사용함으로써, 소정 계층에서 IRAP 픽처를 포함하는 이전 샘플의 시간적으로 병치된 샘플부터 소정 샘플의 시간적으로 병치된 샘플까지, 소정 계층의 참조 계층을 운반하는 트랙에서 임의의 샘플들

[0157] 샘플 디스크립션 박스 (168) 에 대한 상기 예시적인 정의에서, 1 초과의 트랙에서 운반되는 HEVC 또는 L-HEVC 비트스트림에 대해, 기본 트랙의 샘플 엔트리 명칭이 'hvc1' 또는 'hvc2' 일 때, 다른 트랙들의 샘플 엔트리 명칭이 'hev1' 또는 'hev2' 이어야 하고, 기본 트랙의 샘플 엔트리 명칭이 'hev1' 또는 'hev2' 일 때, 다른 트랙들의 샘플 엔트리 명칭은 'hev2' 또는 'lhe1' 이어야 한다. 기본 트랙은 기본적으로 존재하는 기본 계층의 최저 시간 서브 계층 (그 VCL NAL 유닛들이 0 과 동일한 TemporalId 를 가짐) 을 갖는 트랙이다.

[0158] 이 예에서, 위의 제약은 편리한 랜덤 액세스가 HEVC 또는 L-HEVC 비트스트림을 운반하는 모든 트랙들에 대하여 또는 어떤 트랙들도 없이 인에이블되고 표시되는 것을 요구한다.

[0159] 트랙의 샘플들이 HEVC 호환가능 기본 계층을 포함하는 경우, 'hvc1', 'hev1', 'hvc2' 또는 'hev2' 샘플 엔트리는 이 예에서, 사용되어야 한다. 여기서, 엔트리는 초기에 HEVC 구성 박스를 포함해야 하며, 가능하면 이하 정의된 것과 같은 L-HEVC 구성 박스가 뒤따른다. HEVC 구성 박스는 HEVCDecoderConfigurationRecord 에 의해 정의된 바와 같이 HEVC 호환가능한 기본 계층의 프로파일, 티어, 레벨 및 가능하게는 파라미터 세트들을 기록한다. L-HEVC 구성 박스는 가능하게, L-HEVC 구성 박스에 저장된 LHEVCDecoderConfigurationRecord에 의해 정의된 바와 같은 L-HEVC 호환가능 강화 계층들의 파라미터 세트들을 기록한다.

[0160] 트랙의 샘플들이 HEVC 기본 계층을 포함하지 않을 경우, 샘플 엔트리 타입 'lhv1' 또는 'lhe1' 이 사용되어야 하고, 샘플 엔트리는 아래 정의된 바와 같이, L-HEVC 구성 박스를 포함해야 한다. 이것은 예컨대, ISO/IEC 14496-15 에서 정의된 것과 같은 HEVCDecoderConfigurationRecord 를 포함한다.

[0161] 도 5 는 본 개시의 기술들에 따라 미디어 데이터를 형성하고 전송하는 예시적인 방법을 나타내는 흐름도이다. 도 5 의 방법은, 예컨대 도 1 의 컨텐츠 준비 디바이스 (20) 및/또는 도 1 의 서버 디바이스 (60), 컨텐츠 준비 디바이스 (20) 와 서버 디바이스 (60) 에 의한 기능성을 수행하도록 구성된 디바이스, 등에 의해 수행될 수도 있다. 설명의 목적으로, 도 5 의 기술들은 컨텐츠 준비 디바이스 (20) 및 서버 디바이스 (60) 양자에 관하여 논의된다.

[0162] 처음에, 인캡슐레이션 유닛 (30) 은 예를 들어, 오디오 인코더 (26) 및/또는 비디오 인코더 (28) 로부터 인코딩된 미디어 데이터를 수신한다 (200). 이 예에서, 인캡슐레이션 유닛 (30) 은 비디오 데이터를 여러 샘플들로 인캡슐레이션하고, 그 중 일부는 랜덤 액세스 샘플들을 나타낸다. 비디오 데이터는 HEVC 또는 L-HEVC 로서 인코딩될 수도 있다. 편리한 랜덤 액세스를 인에이블하기 위해, 인캡슐레이션 유닛 (30) 은 'hev1' 또는

'hev2' 의 샘플 엔트리 타입 값들을 랜덤 액세스 포인트 미디어 데이터를 인캡슐레이션하는 샘플에 할당하고 (202), 또한 샘플 엔트리 또는 랜덤 액세스 샘플들을 파라미터 세트들에 제공한다 (204). 랜덤 액세스 포인트들은 인트라-예측을 사용하여 인코딩된 프레임들 (I-프레임들) 에 대응할 수도 있는 순시 디코더 리프레시 (IDR) 프레임들에 대응할 수도 있다. 파라미터 세트들은 임의의 또는 모든 비디오 파라미터 세트들 (VPS), 시퀀스 파라미터 세트들 (SPS들) 및/또는 픽처 파라미터 세트들 (SPS들) 을 포함할 수도 있다. 일반적으로, 인캡슐레이션 유닛 (30) 은 편리한 랜덤 액세스를 인에이블하기 위해, 모든 파라미터 세트들에 랜덤 액세스 샘플들 및/또는 랜덤 액세스 샘플들에 대응하는 샘플 엔트리들이 제공되는 것을 보장할 수도 있다.

[0163] 인캡슐레이션 유닛 (30) 은 샘플들 및 파라미터 세트들을 포함하여, 예를 들어 ISO/IEC 14496-15 에 따라 하나 이상의 파일들을 형성할 수도 있다 (206). 특히, 인캡슐레이션 유닛 (30) 은 편리한 랜덤 액세스를 인에이블하기 위해, 파라미터 세트들에 샘플 엔트리 타입들 'hev1 '또는 'hev2' 및/또는 이들 샘플 엔트리 타입들에 대응하는 샘플들이 포함되도록 파일들을 형성할 수도 있다. 컨텐츠 준비 디바이스 (20) 는 그 후에, 랜덤 액세스 샘플에 대한 샘플 엔트리 타입 데이터를 출력할 수도 있다 (208). 일부 예들에서, 이 출력은 디지털 다용도 디스크 (DVD) 또는 Blu-ray 디스크와 같은 고정된 매체에 나머지 파일(들)과 함께 출력될 수도 있다. 다른 예들에서, 이 출력은 서버 디바이스 (60) 에 전송될 수도 있으며, 서버 디바이스 (60) 는 그 후, 샘플 엔트리 타입 데이터를 도 1 의 클라이언트 디바이스 (40) 와 같은 클라이언트 디바이스에 전송할 수도 있다.

[0164] 도 5 의 예에서, 서버 디바이스 (60) 는 샘플 엔트리 타입 데이터를 클라이언트 디바이스 (40) 에 전송하여, 클라이언트 디바이스 (40) 가 편리한 랜덤 액세스를 수행하게 한다. 따라서, 도 1의 서버 디바이스 (60) 의 요청 프로세싱 유닛 (70) 은 클라이언트 디바이스 (40) 로부터, 샘플 엔트리 타입이 'hev1' 또는 'hev2' 인 랜덤 액세스 샘플에 대한 요청을 수신한다 (210). 이에 응답하여, 요청 프로세싱 유닛 (70) 은 요청된 랜덤 액세스 샘플 및 파라미터 세트들 (VPS, SPS 및/또는 PPS) 을 클라이언트 디바이스 (40) 에 전송한다 (212). 특히, 이 데이터는 파라미터 세트들이 샘플 엔트리 데이터 또는 샘플 엔트리 데이터에 대응하는 샘플과 함께 위치되어, 클라이언트 디바이스 (40) 가 랜덤 액세스 샘플보다 이전 샘플들의 파라미터 세트들을 검색하고 페치할 필요가 없도록, 배열될 수도 있다. 이러한 방식으로, 도 5 의 기법들은 클라이언트 디바이스 (40) 에 대한 편리한 랜덤 액세스를 인에이블함으로써 서버 디바이스 (60) 및 클라이언트 디바이스 (40) 에 대한 프로세싱 요건들을 감소시키고, 편리 랜덤 액세스가 인에이블되지 않은 경우와 비교하여 이 데이터 교환에 대한 대역폭 소비를 감소시키며, 그리고 클라이언트 디바이스 (40) 가 미디어 데이터를 요청하는 입력을 사용자로부터 수신하는 시간과 클라이언트 디바이스 (40) 가 요청된 미디어 데이터를 사용자에게 제시할 수 있는 시간 간의 레이턴시를 개선할 수도 있다.

[0165] 도 6 은 본 개시의 기술들에 따라 랜덤 액세스를 수행하는 예시적인 방법을 나타내는 흐름도이다. 예를 들어, 도 6 의 방법은 도 1 의 클라이언트 디바이스 (40) 와 관련하여 설명된다.

[0166] 이 예에서, 클라이언트 디바이스 (40) 의 취출 유닛 (52) 은 초기에 랜덤 액세스 포인트 샘플에 대한 샘플 엔트리 타입 데이터를 요청한다 (220). 예를 들어, 취출 유닛 (52) 은 DASH 에 따라 샘플 엔트리 타입 데이터에 대한 HTTP 요청을 전송할 수도 있다. 샘플 엔트리 타입 데이터를 수신한 후, 클라이언트 디바이스 (40) 의 파일 포맷 프로세싱 유닛 (50) 은 샘플 엔트리 타입 데이터가, 예를 들어, HEVC 또는 L-HEVC 중 하나에 따라 인코딩된 비디오 데이터를 포함하는 연관된 샘플에 대한 'hev1' 또는 'hev2' 의 샘플 엔트리 타입을 표시하는 것으로 결정한다 (222). 따라서, 클라이언트 디바이스 (40) 는 샘플이 편리한 랜덤 액세스를 위해 사용될 수 있다고 결정할 수도 있다. 즉, 클라이언트 디바이스 (40) 는 디코딩 순서로 이전 샘플들의 파라미터 세트들을 검색 및 페치하지 않고, 샘플의 미디어 데이터 또는 샘플에 대한 샘플 엔트리가 포함된 샘플 및 파라미터 세트들을 취출할 수도 있다. 이 예에서, 샘플을 포함하는 비디오 비트스트림은 또한, 디코딩 순서에서 샘플보다 이전의 하나 이상의 다른 샘플들을 포함하는 것이 이해되어야 한다.

[0167] 이에 응답하여, 클라이언트 디바이스 (40) 는 샘플 미디어 데이터 및 대응하는 파라미터 세트들을 취출한다 (226). 클라이언트 디바이스 (40) 는 이전 샘플들의 파라미터 세트들을 검색 및 패치할 필요는 없지만, 대신, 취출된 파라미터 세트들은 편리한 랜덤 액세스로 인해, 취출된 샘플의 미디어 데이터를 디코딩하는데 필요한 모든 파라미터 세트 데이터를 포함한다. 따라서, 파일 포맷 프로세싱 유닛 (50) 은 취출된 샘플을 포함하는 파일로부터 비디오 데이터를 디캡슐레이션하고, 디캡슐레이션된 미디어 데이터를 비디오 디코더 (48) 에 제공한다. 비디오 디코더 (48) 는 파라미터 세트들을 사용하여 비디오 데이터를 디코딩하고, 디코딩된 비디오 데이터를 미디어 데이터를 제시하는 비디오 출력 (44) 에 제공한다 (228). 취출 유닛 (52) 은 디코딩 순서에서 후속 샘플을 더 요청할 수도 있고 (230), 비디오 디코더 (48) 는 그 후, 후속 샘플의 비디오 데이터를 디코딩하고, 비디오 출력 (44) 은 디코딩된 비디오 데이터를 제시할 수도 있다. 이 프로세스는 대응하는 미

디어 프리젠테이션의 종료 까지 또는 예컨대, 사용자에 의해 미디어 데이터의 새로운 세트가 요청될 때까지 계속될 수도 있다.

[0168] 이러한 방식으로, 도 6 의 방법은 비디오 데이터를 프로세싱하는 방법의 일 예를 나타내고, 그 방법은, 비디오 비트스트림의 샘플에 대한 샘플 엔트리 타입을 기술하는 데이터를 수신하는 것으로서, 상기 샘플 엔트리 타입은 'hev1' 또는 'hev2' 중 하나이고, 상기 샘플은 고효율 비디오 코딩 (HEVC) 또는 계층화된 HEVC (L-HEVC) 중 하나에 따라 인코딩된 비디오 데이터를 포함하며, 하나 이상의 다른 샘플들은 디코딩 순서에 있어서 상기 비디오 비트스트림에서의 상기 샘플에 선행하는, 상기 데이터를 수신하는 것, 및 상기 샘플 엔트리 타입이 'hev1' 또는 'hev2' 이고 상기 샘플이 HEVC 또는 L-HEVC 중 하나에 따라 인코딩된 비디오 데이터를 포함하는 것에 응답하여, 디코딩 순서에 있어서 상기 샘플에 선행하는 상기 하나 이상의 다른 샘플들 중 임의의 샘플의 비디오 데이터를 취출하지 않고 상기 비디오 비트스트림의 임의의 이전 샘플들의 파라미터 세트들을 취출하지 않고, 상기 샘플을 사용하여 랜덤 액세스를 수행하도록 상기 샘플을 취출하는 것을 포함한다.

[0169] 도 7 은 비디오 데이터를 포함하는 파일을 생성하는 예시적인 기술을 도시하는 흐름도이다. 도 7 의 방법은, 도 1 의 인캡슐레이션 유닛 (30) 에 의해 수행되는 것으로 설명된다. 그러나, 일반적으로, 도 7 의 방법은, 예컨대 서버 디바이스 (60) 의 컴포넌트들에 의해서뿐만 아니라 다른 디바이스들에 의해 수행될 수도 있다. 또한, 도 7 의 기술들은 앞서 논의된 것과 같은 도 5 의 파일 생성 기술들과 함께 수행될 수도 있다.

[0170] 초기에, 인캡슐레이션 유닛 (30) 은 인코딩된 미디어 데이터를 수신한다 (250). 인캡슐레이션 유닛 (30) 은 예를 들어, HEVC, L-HEVC 또는 다른 비디오 코딩 표준에 따라 미디어 데이터를 인코딩할 수도 있는 비디오 인코더 (28) 로부터 인코딩된 미디어 데이터를 수신할 수도 있다. 특히, 인코딩된 미디어 데이터는, 예를 들어, MV-HEVC, 3D-HEVC, SHVC 등 (예컨대, 다른 비디오 인코딩 표준들에 대한 계층화된 확장물들) 을 위한 복수의 계층들을 포함할 수도 있다.

[0171] 인캡슐레이션 유닛 (30) 은 일반적으로 복수의 트랙들의 형태로 인코딩된 미디어 데이터를 포함하는 파일을 형성할 수도 있다. 인캡슐레이션 유닛 (30) 은 비디오 데이터를 포함하는 각각의 트랙이 다중 계층 비디오 데이터의 하나 이상의 계층들을 포함하도록, 파일의 트랙을 형성할 수도 있다 (252). 트랙 내의 다수의 계층들은 예를 들어, 다양한 시간적 스케일가능 계층들에 대응할 수도 있다. 별개의 계층들을 포함하는 트랙들은 예를 들어, MV-HEVC 또는 3D-HEVC 에 대한 별개의 뷰들 또는 SHVC 에 대한 상이한 스케일가능 계층들 (예컨대, 공간 해상도 스케일가능성, 비트 심도 스케일가능성, PSNR (peak signal to noise ratio) 스케일가능성 등) 에 대응할 수도 있다.

[0172] 도 7 의 예에서, 인캡슐레이션 유닛 (30) 은 편리한 랜덤 액세스를 인에이블할지 여부를 결정한다 (254). 예를 들어, 인캡슐레이션 유닛 (30) 은 관리자와 같은 사용자로부터 구성 데이터를 수신할 수도 있다. 본 개시에서 제안된 제한에 따라, 복수의 트랙들 중 최저 트랙, 즉 비디오 데이터의 최저 서브 계층을 운반하는 비디오 데이터의 기본 계층을 포함하는 최저 트랙에 대하여 편리한 랜덤 액세스가 인에이블될 때, 인캡슐레이션 유닛 (30) 은 비디오 데이터를 포함하는 복수의 트랙들의 서로 다른 트랙에 대하여 편리한 랜덤 액세스를 인에이블한다 (256). 대안적으로, 편리한 랜덤 액세스가 최저 트랙에 대해 디스에이블되면, 인캡슐레이션 유닛 (30) 은 복수의 트랙들 중 다른 트랙들에 대하여 편리한 랜덤 액세스를 디스에이블한다 (258). 다른 예들에서, 인캡슐레이션 유닛 (30) 은 앞서 논의 된 대안적인 제한 (즉, (HEVC, L-HEVC, 또는 임의의 다른 코덱들로 코딩된) 단일 계층 또는 다중 계층 비트스트림을 운반하는 모든 트랙들에 대해, (그 VCL NAL 유닛들이 0 과 동일한 TemporalId 을 갖는) 최저 서브 계층을 운반하는 트랙이 편리한 랜덤 액세스가 인에이블되는 것을 나타내는 샘플 엔트리 타입을 (HEVC 및/또는 L-HEVC 구성들의 존재와 함께) 사용할 때, 모든 다른 트랙들이 (HEVC 및/또는 L-HEVC 구성들의 존재와 함께) 편리한 랜덤 액세스가 인에이블되는 것을 나타내는 샘플 엔트리 타입들을 사용해야만 하는 것) 에 따라 편리한 랜덤 액세스를 인에이블 또는 디스에이블할 수도 있다.

[0173] 인캡슐레이션 유닛 (30) 은 비디오 데이터를 포함하는 트랙들의 샘플들에 대한 샘플 엔트리들에 편리한 랜덤 액세스가 인에이블되는지 여부를 나타내는 샘플 엔트리 타입 값들을 할당한다 (260). 예를 들어, 샘플 엔트리 타입 값들로서 'hev1' 및 'hev2' 는 비디오 데이터를 포함하는 트랙들에 대하여 편리한 랜덤 액세스가 인에이블되는 것을 나타낼 수도 있다. 인캡슐레이션 유닛 (30) 은 또한, 상기 결정된 것과 같은 트랙들 및 샘플 엔트리 타입 값들을 포함하는 파일을 형성한다 (262).

[0174] 이러한 방식으로, 도 7 의 방법은 비디오 데이터를 포함하는 파일을 생성하는 방법의 일 예를 나타내고, 그 방법은 상기 비디오 데이터의 최저 서브 계층을 운반하는 상기 비디오 데이터의 최저 서브 계층을 운반하는 비디오 데이터의 기본 계층을 포함하는 복수의 트랙들 중의 최저 트랙이 편리한 랜덤 액세스가 인에이블되는 것을

나타내는 샘플들에 대한 샘플 엔트리 타입 값들을 포함할 것으로 결정하는 것에 응답하여, 편리한 랜덤 액세스가 인에이블되는 것을 나타내기 위한 비디오 데이터를 포함하는 상기 복수의 트랙들의 다른 트랙들의 각각의 샘플들에 대한 샘플 엔트리 타입 값들을 세팅하는 단계, 및 상기 복수의 트랙들의 트랙들에 대한 상기 샘플 엔트리 타입 값들이 편리한 랜덤 액세스가 인에이블되는 것을 나타내도록, 상기 복수의 트랙들을 포함하는 파일을 생성하는 단계를 포함한다.

[0175] 하나 이상의 예들에 있어서, 설명된 기능들은 하드웨어, 소프트웨어, 펌웨어, 또는 이들의 임의의 조합으로 구현될 수도 있다. 소프트웨어로 구현된다면, 그 기능들은 하나 이상의 명령들 또는 코드로서 컴퓨터 판독가능 매체 상으로 저장 또는 전송되고 하드웨어 기반 프로세싱 유닛에 의해 실행될 수도 있다. 컴퓨터 판독가능 매체들은 데이터 저장 매체들과 같은 유형의 매체에 대응하는 컴퓨터 판독가능 저장 매체들, 또는 예를 들어, 통신 프로토콜에 따라 일 장소로부터 다른 장소로의 컴퓨터 프로그램의 전송을 용이하게 하는 임의의 매체를 포함하는 통신 매체들을 포함할 수도 있다. 이러한 방식으로, 컴퓨터 판독가능 매체는 일반적으로 (1) 비-일시적인 유형의 컴퓨터 판독가능 저장 매체 또는 (2) 신호 또는 캐리어파와 같은 통신 매체에 대응할 수도 있다. 데이터 저장 매체는 본 개시에서 설명된 기술들의 구현을 위한 명령들, 코드 및/또는 데이터 구조들을 취출하기 위해 하나 이상의 컴퓨터들 또는 하나 이상의 프로세서들에 의해 액세스될 수 있는 임의의 가용 매체일 수도 있다. 컴퓨터 프로그램 제품이 컴퓨터 판독가능 매체를 포함할 수도 있다.

[0176] 한정이 아닌 예로서, 그러한 컴퓨터 판독가능 저장 매체는 RAM, ROM, EEPROM, CD-ROM 또는 다른 광학 디스크 저장부, 자기 디스크 저장부 또는 다른 자기 저장 디바이스들, 플래시 메모리, 또는 원하는 프로그램 코드를 명령들 또는 데이터 구조들의 형태로 저장하는데 이용될 수 있고 컴퓨터에 의해 액세스될 수 있는 임의의 다른 매체를 포함할 수 있다. 또한, 임의의 커넥션이 컴퓨터 판독가능 매체로 적절히 명명된다. 예를 들어, 동축 케이블, 광섬유 케이블, 꼬임쌍선, 디지털 가입자 라인 (DSL), 또는 적외선, 무선, 및 마이크로파와 같은 무선 기술들을 이용하여 웹사이트, 서버, 또는 다른 원격 소스로부터 명령들이 송신된다면, 동축 케이블, 광섬유 케이블, 꼬임쌍선, DSL, 또는 적외선, 무선, 및 마이크로파와 같은 무선 기술들은 매체의 정의에 포함된다. 하지만, 컴퓨터 판독가능 저장 매체 및 데이터 저장 매체는 커넥션들, 캐리어파들, 신호들, 또는 다른 일시적 매체들을 포함하지 않지만 대신 비-일시적인 유형의 저장 매체들로 지향됨을 이해해야 한다. 본원에서 이용되는 바와 같은 디스크 (disk) 및 디스크 (disc) 는 CD (compact disc), 레이저 디스크, 광 디스크, DVD (digital versatile disc), 플로피 디스크, 및 블루레이 디스크를 포함하는데, 여기서 디스크 (disk) 는 보통 데이터를 자기적으로 재생하며, 반면 디스크 (disc) 는 레이저들을 이용하여 광학적으로 데이터를 재생한다. 위의 조합들도 컴퓨터 판독가능 매체들의 범위 내에 포함되어야 한다.

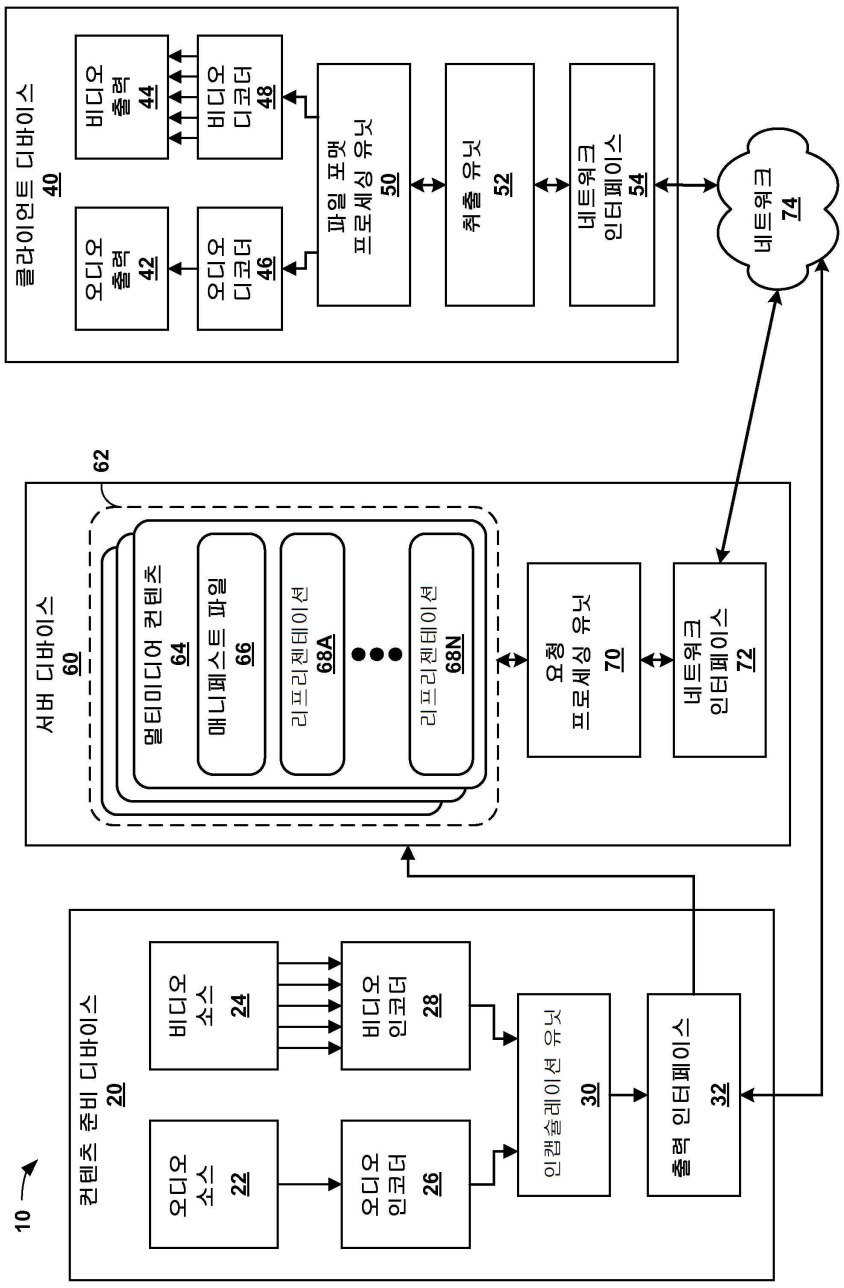
[0177] 명령들은 하나 이상의 디지털 신호 프로세서들 (DSP들), 범용 마이크로프로세서들, 주문형 집적회로들 (ASIC들), 필드 프로그래밍가능 로직 어레이들 (FPGA들), 또는 다른 등가의 집적된 또는 별도의 로직 회로와 같은 하나 이상의 프로세서들에 의해 실행될 수도 있다. 이에 따라, 본 명세서에서 사용된 바와 같은 용어 "프로세서" 는 본 명세서에서 설명된 기술들의 구현에 적합한 전술한 구조 또는 임의의 다른 구조 중 임의의 구조를 지칭할 수도 있다. 추가적으로, 일부 양태들에 있어서, 본 명세서에서 설명된 기능은 인코딩 및 디코딩을 위해 구성되고 결합된 코덱에서 통합된 전용 하드웨어 및/또는 소프트웨어 모듈들 내에 제공될 수도 있다. 또한, 그 기술들은 하나 이상의 회로들 또는 로직 엘리먼트들에서 완전히 구현될 수 있다.

[0178] 본 개시의 기술들은 무선 핸드셋, 집적 회로 (IC) 또는 IC들의 세트 (예를 들어, 칩 세트) 를 포함하여 매우 다양한 디바이스들 또는 장치들에서 구현될 수도 있다. 다양한 컴포넌트들, 모듈들 또는 유닛들이 개시된 기술들을 수행하도록 구성된 디바이스들의 기능적 양태들을 강조하기 위해 본 개시에서 설명되지만, 반드시 상이한 하드웨어 유닛들에 의한 실현을 요구하지는 않는다. 오히려, 상기 설명된 바와 같이, 다양한 유닛들은 적절한 소프트웨어 및/또는 펌웨어와 함께 상기 설명된 바와 같은 하나 이상의 프로세서들을 포함하여 코덱 하드웨어 유닛으로 결합되거나 또는 상호운용식 하드웨어 유닛들의 집합에 의해 제공될 수도 있다.

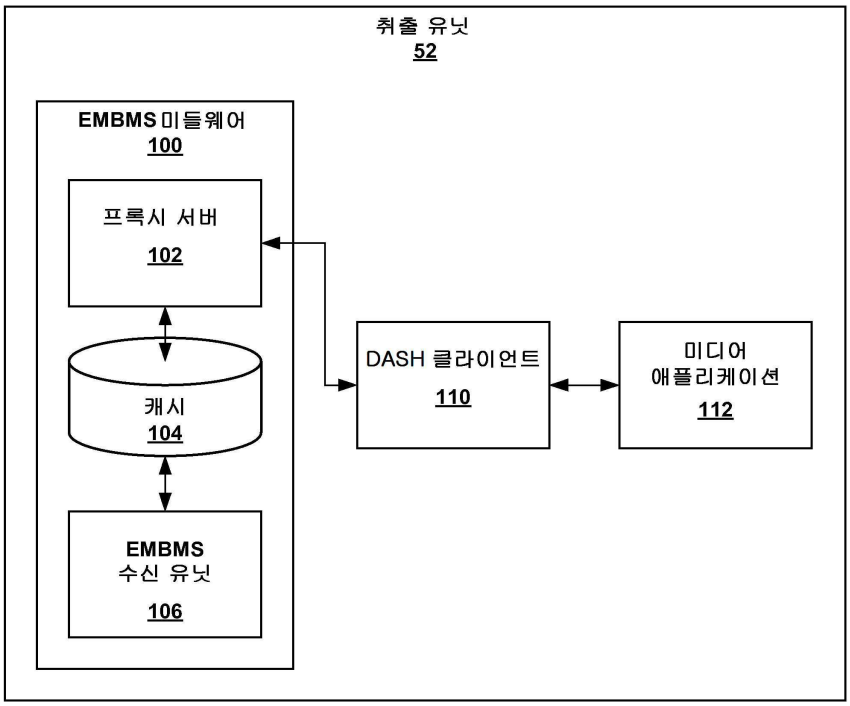
[0179] 다양한 예들이 설명되었다. 이들 및 다른 예들은 다음의 청구항들의 범위 내에 있다.

도면

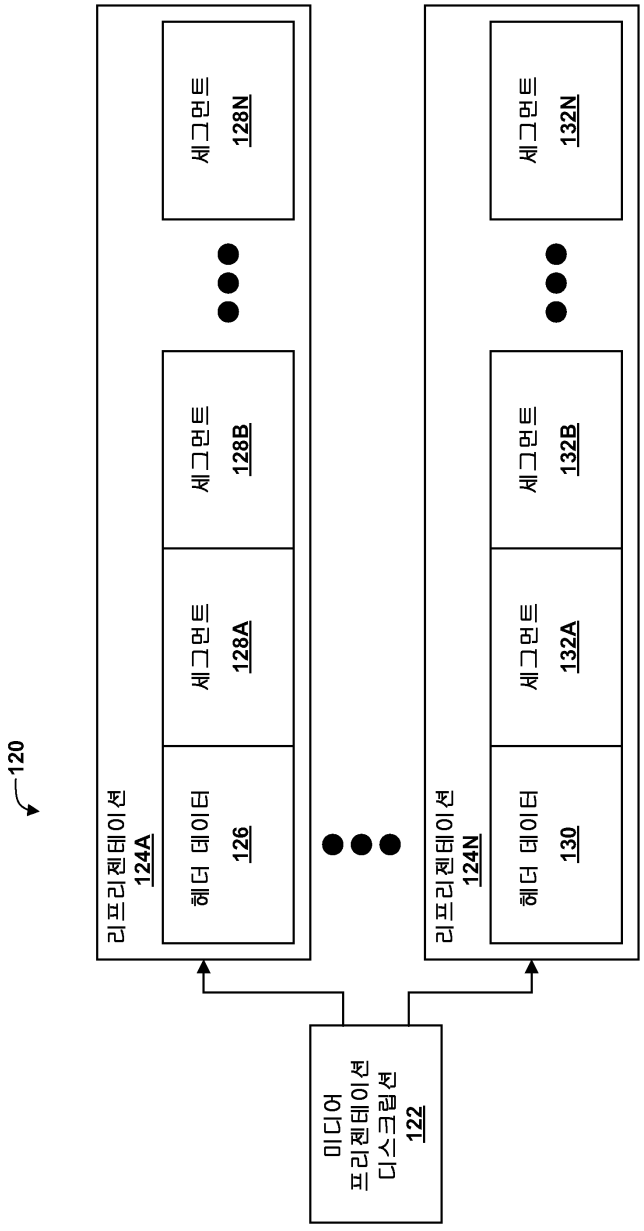
도면1


10
컨텐츠 준비 디바이스
20
오디오 소스
22
비디오 소스
24
오디오 인코더
26
비디오 인코더
28
인캡슐레이션 유닛
30
출력 인터페이스
32
서버 디바이스
60
62
멀티미디어 컨텐츠
64
매니페스트 파일
66
리프리젠테이션
68A
리프리젠테이션
68N
요청 프로세싱 유닛
70
네트워크 인터페이스
72
네트워크
74
클라이언트 디바이스
40
오디오 출력
42
비디오 출력
44
오디오 디코더
46
비디오 디코더
48
파일 포맷 프로세싱 유닛
50
취출 유닛
52
네트워크 인터페이스
54

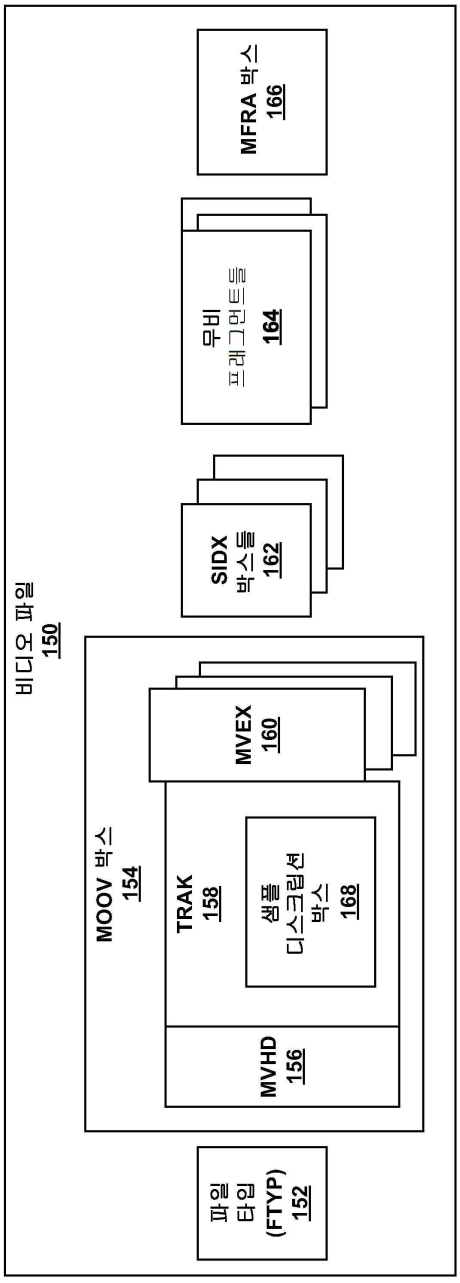
도면2


취출 유닛
52
EMBMS 미들웨어
100
프록시 서버
102
캐시
104
EMBMS
수신 유닛
106
DASH 클라이언트
110
미디어
애플리케이션
112

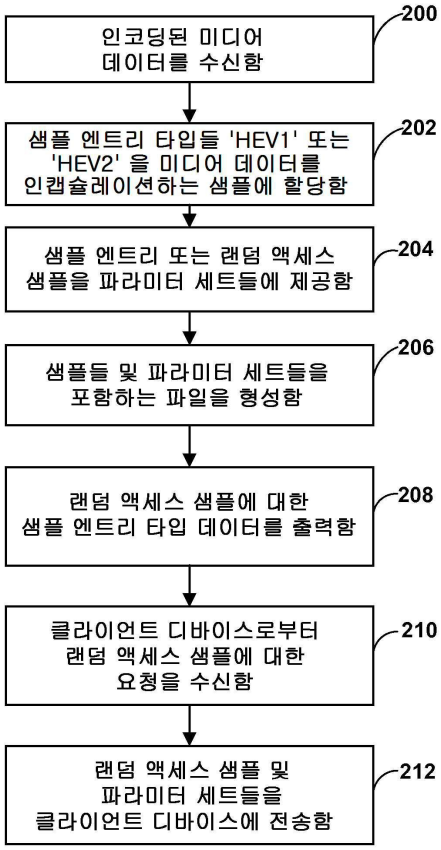
도면3


120
리프리젠테이션
124A
헤더 데이터
126
세그먼트
128A
세그먼트
128B
세그먼트
128N
리프리젠테이션
124N
헤더 데이터
130
세그먼트
132A
세그먼트
132B
세그먼트
132N
미디어
프리젠테이션
디스크립션
122

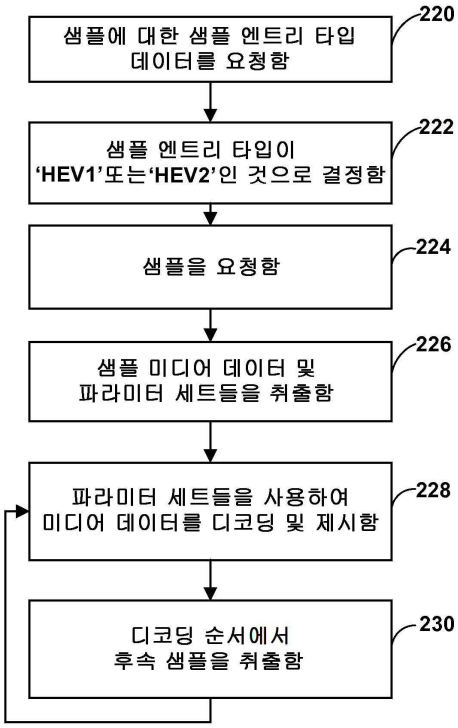
도면4


비디오 파일
150
파일 타입 (FTYP)
152
MOOV 박스
154
MVHD
156
TRAK
158
샘플 디스크립션 박스
168
MVEX
160
SIDX 박스들
162
무비 프래그먼트들
164
MFRA 박스
166

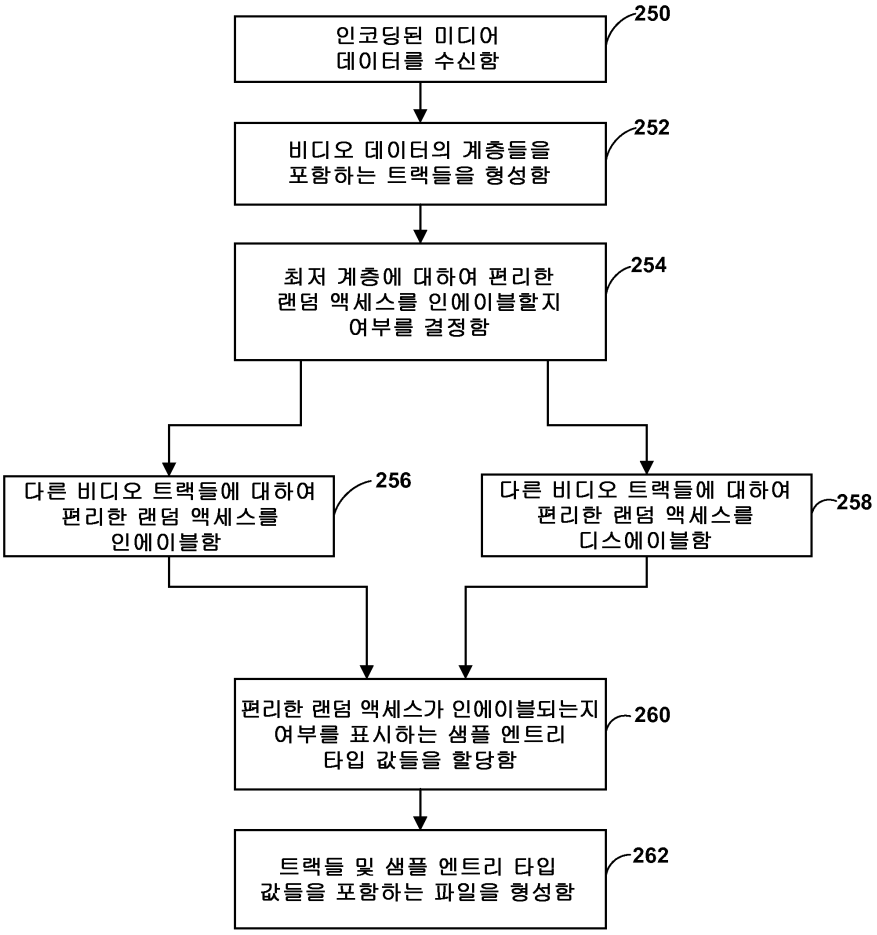
도면5


인코딩된 미디어 데이터를 수신함
200
샘플 엔트리 타입들 'HEV1' 또는 'HEV2' 을 미디어 데이터를 인캡슐레이션하는 샘플에 할당함
202
샘플 엔트리 또는 랜덤 액세스 샘플을 파라미터 세트들에 제공함
204
샘플들 및 파라미터 세트들을 포함하는 파일을 형성함
206
랜덤 액세스 샘플에 대한 샘플 엔트리 타입 데이터를 출력함
208
클라이언트 디바이스로부터 랜덤 액세스 샘플에 대한 요청을 수신함
210
랜덤 액세스 샘플 및 파라미터 세트들을 클라이언트 디바이스에 전송함
212

도면6


샘플에 대한 샘플 엔트리 타입 데이터를 요청함
220
샘플 엔트리 타입이 'HEV1'또는'HEV2'인 것으로 결정함
222
샘플을 요청함
224
샘플 미디어 데이터 및 파라미터 세트들을 취출함
226
파라미터 세트들을 사용하여 미디어 데이터를 디코딩 및 제시함
228
디코딩 순서에서 후속 샘플을 취출함
230

도면7


인코딩된 미디어 데이터를 수신함
250
비디오 데이터의 계층들을 포함하는 트랙들을 형성함
252
최저 계층에 대하여 편리한 랜덤 액세스를 인에이블할지 여부를 결정함
254
다른 비디오 트랙들에 대하여 편리한 랜덤 액세스를 인에이블함
256
다른 비디오 트랙들에 대하여 편리한 랜덤 액세스를 디스에이블함
258
편리한 랜덤 액세스가 인에이블되는지 여부를 표시하는 샘플 엔트리 타입 값들을 할당함
260
트랙들 및 샘플 엔트리 타입 값들을 포함하는 파일을 형성함
262