



(19) 대한민국특허청(KR)

(12) 등록특허공보(B1)

(45) 공고일자 2019년10월25일

(11) 등록번호 10-2037009

(24) 등록일자 2019년10월21일

(51) 국제특허분류(Int. Cl.)

H04N 21/44 (2011.01) H04N 19/30 (2014.01)

H04N 21/234 (2014.01) H04N 21/435 (2011.01)

H04N 21/845 (2011.01)

(52) CPC특허분류

H04N 21/44016 (2013.01)

H04N 19/30 (2015.01)

(21) 출원번호 10-2018-7000559

(22) 출원일자(국제) 2016년06월07일

심사청구일자 2018년01월15일

(85) 번역문제출일자 2018년01월08일

(65) 공개번호 10-2018-0018662

(43) 공개일자 2018년02월21일

(86) 국제출원번호 PCT/EP2016/062926

(87) 국제공개번호 WO 2016/202648

국제공개일자 2016년12월22일

(30) 우선권주장

1510610.7 2015년06월16일 영국(GB)

(56) 선행기술조사문헌

KR1020150036299 A*

*는 심사관에 의하여 인용된 문헌

(73) 특허권자

캐논 가부시끼가이샤

일본 도쿄도 오오따꾸 시모마루코 3조메 30방 2고

(72) 발명자

드누알 프랑크

프랑스 35190 생 도미닉 라 빌 에 해 8

마즈 프레데릭

프랑스 35850 랑강 뒤 데 띠엘 6

(뒷면에 계속)

(74) 대리인

장수길, 이중희

전체 청구항 수 : 총 18 항

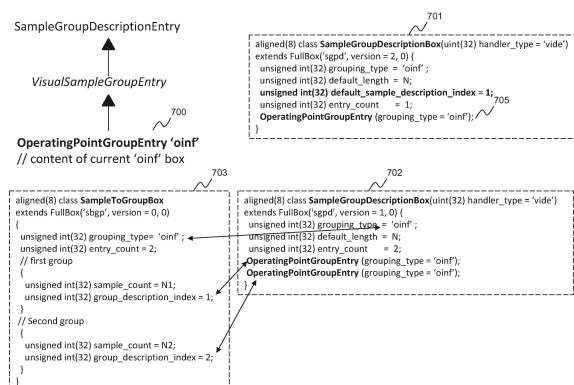
심사관 : 옥윤철

(54) 발명의 명칭 동작 포인트 디스크립터가 동적으로 설정될 수 있는 캡슐화된 비트-스트림으로부터 미디어 데이터 및 메타데이터를 획득하는 방법, 디바이스, 및 컴퓨터 프로그램

(57) 요약

본 발명은 하나 이상의 비트-스트림의 캡슐화로부터 생기는 캡슐화된 비트-스트림으로부터 미디어 데이터 및 메타데이터를 획득하는 방법에 관한 것으로, 상기 하나 이상의 비트-스트림 각각의 미디어 데이터는 하나 이상의 계층으로 조직되고, 각 계층은 시간 설정형 샘플들 또는 하위 샘플들을 포함한다. 상기 미디어 데이터의 시간

(뒷면에 계속)

대표도

설정형 샘플들의 적어도 하나의 그룹의 적어도 하나의 정의를 포함하는 메타데이터를 참조하기 위한 데이터 구조가 획득된다. 상기 하나 이상의 비트-스트림 각각에 대한 적어도 하나의 동작 포인트 디스크립터가 획득되고, 상기 적어도 하나의 동작 포인트 디스크립터는 상기 데이터 구조에 의해 참조된 메타데이터에 정의된 시간 설정형 샘플들의 그룹에 속하는 시간 설정형 샘플들이 독립적으로 디코딩되는 것을 가능하게 하는 기술적 정보에 관련된 메타데이터를 제공한다. 시간 설정형 샘플들의 하나의 그룹과 연관된 각 동작 포인트 디스크립터는 상기 그룹에 속하는 시간 설정형 샘플들의 기술적 정보를 정의한다.

(52) CPC특허분류

H04N 21/23424 (2013.01)

H04N 21/435 (2013.01)

H04N 21/845 (2013.01)

르 웨브르 장

프랑스 94230 까샹 아브뉴 뒤 프레지덩 월슨 7 에스끄 베1

(72) 발명자

우드라오고 나엘

프랑스 35330 모르 드 브르파뉴 꾸뚜즈

콩플라토 씨릴

프랑스 77380 콩 라 빌 뒤 드 라 헤르스 18

명세서

청구범위

청구항 1

베이스 계층과 하나 이상의 인헨스먼트 계층을 포함하는 복수의 계층에 기초하여 계층화된 비트 스트림에 관한 기술 데이터를 생성하는 방법으로서,

상기 계층화된 비트 스트림에 대응하는 복수의 샘플을 복수의 샘플 그룹으로 그룹화하는 단계;

상기 복수의 샘플 그룹 중 하나 이상의 샘플 그룹에 대한 하나 이상의 동작 정보 항목을 결정하는 단계로서, 상기 하나 이상의 동작 정보 항목의 각각이 상기 계층화된 비트 스트림의 독립적으로 디코딩가능한 서브셋을 나타내는, 상기 결정하는 단계; 및

적어도 상기 복수의 샘플 그룹 중 상기 하나 이상의 샘플 그룹에 대한 상기 하나 이상의 동작 정보 항목을 기술하는 기술 데이터를 생성하는 단계를 포함하는, 방법.

청구항 2

제1항에 있어서, 생성된 상기 기술 데이터를 장치로부터의 요청에 응답하여 상기 장치에 출력하는 단계를 더 포함하는, 방법.

청구항 3

제1항에 있어서, 상기 하나 이상의 동작 정보 항목의 각각은, 상기 베이스 계층에 기초하는 비트 스트림이 독립적으로 디코딩가능하고 상기 하나 이상의 인헨스먼트 계층 중 하나와 상기 베이스 계층 모두에 기초하는 비트 스트림이 독립적으로 디코딩가능함을 특정하는, 방법.

청구항 4

제1항에 있어서, 상기 하나 이상의 인헨스먼트 계층 중 하나와 상기 베이스 계층 모두에 기초하는 비트 스트림의 화상 품질은 상기 베이스 계층에 기초하는 비트 스트림의 화상 품질보다 높은, 방법.

청구항 5

제1항에 있어서, 상기 계층화된 비트 스트림의 계층 조직에 관한 메타데이터를 제공하는 하나 이상의 계층 정보 항목을 획득하는 단계를 더 포함하는, 방법.

청구항 6

제5항에 있어서, 하나의 계층 정보 항목이 디폴트 계층 정보 항목으로서 정의되며, 이에 따라 계층 조직에 관한 메타데이터에 연관되지 않은 샘플들의 그룹에 속하는 샘플들은 상기 디폴트 계층 정보 항목에 의해 제공되는 메타데이터에 기초하여 디코딩되는, 방법.

청구항 7

제1항에 있어서, 동작 정보 항목이 디폴트 동작 정보 항목으로서 정의되며, 이에 따라 기술적 정보에 관한 메타데이터에 연관되지 않은 샘플들의 그룹에 속하는 샘플들은 상기 디폴트 동작 정보 항목에 의해 제공되는 메타데이터에 기초하여 디코딩되는, 방법.

청구항 8

하나 이상의 비트 스트림의 캡슐화로부터 발생하는 캡슐화된 비트 스트림으로부터 미디어 데이터 및 메타데이터를 획득하는 방법으로서, 상기 하나 이상의 비트 스트림의 각각의 미디어 데이터는 베이스 계층과 하나 이상의 인헨스먼트 계층을 포함하는 복수의 계층으로 조직되고, 상기 방법은,

상기 하나 이상의 비트 스트림 중 하나 이상에 대한 하나 이상의 동작 정보 항목을 기술하는 기술 데이터를 획득

특하는 단계; 및

기술된 상기 하나 이상의 동작 정보 항목 중 하나 이상을 획득하는 단계를 포함하고,

획득된 상기 하나 이상의 동작 정보 항목의 각각은, 샘플들이 독립적으로 디코딩가능한 하나의 샘플 그룹을 나타내는, 방법.

청구항 9

제8항에 있어서, 상기 기술 데이터를 장치에 요청하는 단계를 더 포함하는, 방법.

청구항 10

제8항에 있어서, 상기 하나 이상의 동작 정보 항목의 각각은, 상기 베이스 계층에 기초하는 비트 스트림이 독립적으로 디코딩가능하고 상기 하나 이상의 인헨스먼트 계층 중 하나와 상기 베이스 계층 모두에 기초하는 비트 스트림이 독립적으로 디코딩가능함을 특징하는, 방법.

청구항 11

제8항에 있어서, 상기 하나 이상의 인헨스먼트 계층 중 하나와 상기 베이스 계층 모두에 기초하는 비트 스트림의 화상 품질은 상기 베이스 계층에 기초하는 비트 스트림의 화상 품질보다 높은, 방법.

청구항 12

제8항에 있어서, 계층화된 비트 스트림의 계층 조직에 관한 메타데이터를 제공하는 하나 이상의 계층 정보 항목을 획득하는 단계를 더 포함하는, 방법.

청구항 13

제12항에 있어서, 하나의 계층 정보 항목이 디폴트 계층 정보 항목으로서 정의되며, 이에 따라 계층 조직에 관한 메타데이터에 연관되지 않은 샘플들의 그룹에 속하는 샘플들은 상기 디폴트 계층 정보 항목에 의해 제공되는 메타데이터에 기초하여 디코딩되는, 방법.

청구항 14

제8항에 있어서, 동작 정보 항목이 디폴트 동작 정보 항목으로서 정의되며, 이에 따라 기술적 정보에 관한 메타데이터에 연관되지 않은 샘플들의 그룹에 속하는 샘플들은 상기 디폴트 동작 정보 항목에 의해 제공되는 메타데이터에 기초하여 디코딩되는, 방법.

청구항 15

제8항에 있어서,

제1 샘플 그룹의 샘플들을 재구성하기 위한 순서를 동작 정보 항목으로부터 획득하는 단계를 더 포함하는, 방법.

청구항 16

제1항 내지 제15항 중 어느 한 항에 따른 방법을 구현하기 위한 컴퓨터 프로그램의 명령어를 저장하는 비일시적 컴퓨터 판독가능 저장매체.

청구항 17

베이스 계층과 하나 이상의 인헨스먼트 계층을 포함하는 복수의 계층에 기초하여 계층화된 비트 스트림에 관한 기술 데이터를 생성하는 장치로서,

마이크로프로세서를 포함하고,

상기 마이크로프로세서는,

상기 계층화된 비트 스트림에 대응하는 복수의 샘플을 복수의 샘플 그룹으로 그룹화하는 단계;

상기 복수의 샘플 그룹 중 하나 이상의 샘플 그룹에 대한 하나 이상의 동작 정보 항목을 결정하는 단계로서, 상

기 하나 이상의 동작 정보 항목의 각각이 상기 계층화된 비트 스트림의 독립적으로 디코딩가능한 서브셋을 나타내는, 상기 결정하는 단계; 및

적어도 상기 복수의 샘플 그룹 중 상기 하나 이상의 샘플 그룹에 대한 상기 하나 이상의 동작 정보 항목을 기술하는 기술 데이터를 생성하는 단계를 실행하도록 구성된, 장치.

청구항 18

하나 이상의 비트 스트림의 캡슐화로부터 발생하는 캡슐화된 비트 스트림으로부터 미디어 데이터 및 메타데이터를 획득하는 장치로서, 상기 하나 이상의 비트 스트림의 각각의 미디어 데이터는 베이스 계층과 하나 이상의 인핸스먼트 계층을 포함하는 복수의 계층으로 조직되고, 상기 장치는,

마이크로프로세서를 포함하고,

상기 마이크로프로세서는,

상기 하나 이상의 비트 스트림 중 하나 이상에 대한 하나 이상의 동작 정보 항목을 기술하는 기술 데이터를 획득하는 단계; 및

기술된 상기 하나 이상의 동작 정보 항목 중 하나 이상을 획득하는 단계를 실행하도록 구성되고,

획득된 상기 하나 이상의 동작 정보 항목의 각각은, 샘플들이 독립적으로 디코딩가능한 하나의 샘플 그룹을 나타내는, 장치.

청구항 19

삭제

청구항 20

삭제

청구항 21

삭제

청구항 22

삭제

청구항 23

삭제

청구항 24

삭제

청구항 25

삭제

청구항 26

삭제

청구항 27

삭제

청구항 28

삭제

청구항 29

삭제

청구항 30

삭제

청구항 31

삭제

청구항 32

삭제

청구항 33

삭제

청구항 34

삭제

청구항 35

삭제

청구항 36

삭제

청구항 37

삭제

청구항 38

삭제

청구항 39

삭제

청구항 40

삭제

청구항 41

삭제

청구항 42

삭제

청구항 43

삭제

청구항 44

삭제

청구항 45

삭제

청구항 46

삭제

청구항 47

삭제

청구항 48

삭제

청구항 49

삭제

청구항 50

삭제

청구항 51

삭제

발명의 설명

기술 분야

[0001] 본 발명은 일반적으로, 미디어 데이터의 교환, 관리, 편집, 및 프레젠테이션을 용이하게 하는 유연하고 확장 가능한 포맷을 제공하기 위해 그리고 특히 압축 비디오 스트림에서 관심있는 사용자 선택 영역의 HTTP(HyperText Transfer Protocol) 및 RTP(Real-time Transport Protocol) 스트리밍에 관한, 스트림 송출을 향상시키기 위해, 예를 들어, MPEG 표준화 기구에 의해 정의된 바와 같은 베이스 미디어 파일 포맷(Base Media File Format)에 따른, 시간 설정형 미디어 데이터의 캡슐화 분야에 관한 것이다. 특히, 본 발명은 동작 포인트 디스크립터들이 동적으로 설정될 수 있는 캡슐화된 비트-스트림들로부터 미디어 데이터 및 메타데이터를 획득하는 방법, 디바이스, 및 컴퓨터 프로그램에 관한 것이다.

배경 기술

[0002] 비디오 코딩은 일련의 비디오 이미지들을 콤팩트한 디지털화된 비트-스트림으로 변환하여, 이러한 비디오 이미지들이 전송되거나 저장될 수 있도록 하는 방식이다. 인코딩 디바이스는 이러한 비디오 이미지들을 코딩하는 데 사용될 수 있고, 연관된 디코딩 디바이스는 디스플레이 및 시청을 위해 이러한 비트-스트림을 재구성하는 데 사용될 수 있다. 일반적인 목적은 이러한 비트-스트림을 본래 비디오 정보보다 더 작은 사이즈가 되도록 형성하는 것이다. 이것은 유리하게도 이러한 비트-스트림을 전송하거나 또는 저장하는데 요구되는 전송 네트워크 또는 스토리지 디바이스의 용량을 감소시켜 준다. 전송되기 위해서, 비디오 비트-스트림은 일반적으로, 통상 헤더들 및 체크 비트들을 추가하는 전송 프로토콜에 따라 캡슐화된다. 비디오 스트리밍 메커니즘은 인터넷 네트워크 및 모바일 네트워크들을 통해 널리 전개되어, 예를 들어, 3GPP의 AHS(Adaptive HTTP Streaming), 마이크로소프트사의 Smooth Streaming 또는 애플사(Apple)의 HTTP 라이브 스트리밍과 같은 HTTP(HyperText Transfer Protocol)를 통해 오디오/비디오 미디어를 스트리밍하는 데 사용된다.

[0003] 최근, 동화상 전문가 그룹(Moving Picture Experts Group, MPEG))은, HTTP를 통한 기존 스트리밍 해결책들을 통합하여 대체하는 새로운 표준을 공개하였다. DASH(Dynamic adaptive streaming over HTTP)라고 지칭되는 이 새로운 표준은, 표준 웹 서버를 기반으로 HTTP를 통한 미디어-스트리밍 모델을 지원하기 위한 것으로, 여기서 인텔리전스(즉, 사용자 선택, 네트워크 조건, 및 클라이언트 능력에 대한 스트리밍할 미디어 데이터 및 비트-스트림의 동적 적응의 선택)는 오로지 클라이언트 선택 및 디바이스에 의존한다.

- [0004] 이러한 모델에서, 미디어 프리젠테이션은, 데이터 세그먼트들로, 그리고 프리젠테이션될 시간 설정형 미디어 데이터의 조직화를 나타내는 "미디어 프리젠테이션 디스크립션(Media Presentation Description)(MPD)"라고 불리는 매니페스트로 조직된다. 특히, 매니페스트는, 데이터 세그먼트들을 다운로드하는 데 사용하는 리소스 식별자들을 포함하고, 이러한 데이터 세그먼트들을 선택 및 조합하여 유효 미디어 프리젠테이션을 획득하기 위한 컨텍스트를 제공한다. 리소스 식별자들은 통상 HTTP-URL(Uniform Resource Locator)들이고, 바이트 범위들과 가능한 조합된다. 매니페스트에 기초하여, 클라이언트 디바이스는, 그의 필요성들, 그의 능력들(예를 들어, 지원되는 코덱들, 디스플레이 사이즈, 프레임 레이트, 품질 레벨 등)에 따라, 그리고 네트워크 조건들(예를 들어, 가용 대역폭)에 의존하여, 어느 미디어 세그먼트들이 미디어 데이터 서버로부터 다운로드될 것인지를 언제든 결정한다.
- [0005] 예를 들어, RTP(Real-time Transport Protocol)와 같은, HTTP에 대한 대안적인 프로토콜들이 존재한다는 점에 주목해야 한다.
- [0006] 또한, 비디오 해상도는, 표준 해상도(standard definition, SD)에서 시작하여 고해상도(high definition, HD)로, 그리고 초고해상도(예를 들어, 4K2K 또는 8K4K, 즉 4,096 x 2,400 화소들 또는 7,680 x 4,320 화소들의 이미지들을 포함하는 비디오)로, 계속 증가하고 있다. 그러나, 수신 및 비디오 디코딩 디바이스들 모두가, 특히 비디오가 초고해상도일 때 전체 해상도의 비디오에 액세스하는 리소스들(예를 들어, 네트워크 액세스 대역폭 또는 CPU(Central Processing Unit))을 갖는 것은 아니며, 모든 사용자들이 그러한 비디오에 액세스할 필요는 없다. 그러한 맥락에서, 비디오 비트-스트림의 일부 부분들에만 액세스하고 그것을 추출하는 능력을 제공하는, 즉, 예를 들어, 전체 비디오 시퀀스의 일부 스케일 가능성 계층, 뷰들, 또는 공간적 하위 파트들에만 액세스하는 것이 특히 유리하다.
- [0007] 비디오에 속하는 프레임들의 스케일 가능성 계층들, 뷰들, 또는 공간적 하위 파트들에 액세스하는 알려진 메커니즘은 비디오의 각 프레임을, 잠재적으로 코딩 종속성들을 갖는, 계층들의 배열로서 조직하는 것에 있다. HEVC(High Efficiency Video Coding)와 같은 일부 비디오 포맷들은 시간적, SNR(품질), 및 공간적 스케일 가능성 계층들에 대한, 다수의 뷰들에 대한 및/또는 타일 인코딩에 대한 지원을 제공한다. 예를 들어, 사용자 정의 ROI는 하나 또는 수 개의 인접한 타일을 커버할 수 있다. 멀티-뷰의 경우, 사용자는 싱글 뷰보다 스테레오를 선호할 수 있다. 스케일 가능성의 경우, 예를 들어 화면 사이즈나 처리 능력에 관하여 사용자의 디바이스에 따라 적절한 계층이 선택될 수 있다.
- [0008] 비디오 비트-스트림(즉, 하위 비트-스트림)의 관련 파트들만의 선택, 추출, 및 전송을 가능하게 하기 위해, 비디오 비트-스트림의 조직(및 더 일반적으로는 비디오뿐만 아니라 오디오, 메타데이터, 자막 등을 포함할 수 있는 미디어 데이터의 조직)이 미디어 플레이어들에 노출되어야 한다. 이 조직은 동작 포인트들의 리스트로서 표현된다.
- [0009] 동작 포인트(operation point)라고도 지칭되는 동작 포인트는 계층화된 HEVC 비트-스트림의 부분 또는 비트-스트림 서브세트를 나타내며, 이는 이 특정 비트-스트림을 디코딩하는 데 필요한 모든 데이터로 구성된 비트-스트림 부분을 추출함으로써 획득될 수 있고 그것은 다른 동작 포인트들과 독립적으로 디코딩될 수 있다. 결과적으로, 동작 포인트는 0부터 선택된 최대 값까지 변화하는 값들을 갖는 시간 식별자들의 범위와 연관된 출력 계층들의 세트이다. 예시를 위해, 30Hz 및 60Hz의 프레임 레이트들에 대응하는 2개의 시간 식별자 값(0 및 1)이 도 5a 및 도 5b에 도시되어 있다.
- [0010] 도 5a 및 도 5b를 포함하는 도 5는 비-출력 계층이 수반되는 계층 구성의 예들을 도시한다.
- [0011] 더 정확하게는, 도 5a는 비디오 시퀀스의 수 개의 표현들 사이의 관계의 예를 도시한다. 이들 표현은 상이한 시간적 해상도들(즉, 30Hz 및 60Hz의 프레임 레이트)를 갖는 표현들을 포함하고, 시간적 해상도들 각각에 대해, 표현들은 상이한 뷰들(즉, 좌측, 우측, 및 공통)을 포함한다.
- [0012] 표현된 바와 같이, 공통 뷰는 좌측 및 우측 뷰들로부터 직접 도출 가능하고 30Hz의 프레임 레이트를 갖는 공통 뷰는 60Hz의 프레임 레이트를 갖는 공통 뷰로부터 직접 도출 가능하다.
- [0013] 표현은 또한 시간적 해상도들 각각에 따라 전체 뷰들의 스케일 가능하지 않은 표현들을 포함한다.
- [0014] 도 5b에 도시된 바와 같이, 도 5a에 도시된 표현들은 3개의 가능한 뷰(즉, 좌측, 우측, 및 공통)에 대응하는 3개의 계층에 따라 그리고 2개의 프레임 레이트(즉, 30Hz 및 60Hz)에 따라 인코딩될 수 있다.
- [0015] 따라서, HTTP 프로토콜에 따라 비디오 시퀀스들 또는 사용자 선택 데이터를 스트리밍하기 위해서는, 사용자 선

택 데이터의 재구성, 디코딩 및 디스플레이를 가능하게 하는 최소 량의 데이터를 전송하기 위해 선택된 데이터에 대한 액세스 및 이들의 코딩 종속성들을 가능하게 하는 방식으로 인코딩된 비디오 비트-스트림의 시간 설정형 미디어 데이터의 캡슐화를 제공하는 것이 중요하다.

- [0016] 적응 HTTP 스트리밍의 전형적인 사용은 비트-스트림 스플라이싱이다.
- [0017] 도 6a, 도 6b 및 도 6c를 포함하는 도 6은 비트-스트림 스플라이싱 애플리케이션의 예를 도시한다.
- [0018] 도 6a에 도시된 바와 같이, 비트-스트림 스플라이싱은 저해상도 비트-스트림(SD)으로부터 고해상도 비트-스트림(HD)으로 스위칭하는 것에 있을 수 있다. 도 6b에 도시된 또 다른 예에 따르면, 비트-스트림 스플라이싱은 재생을 위해 라이브 비트-스트림으로부터 온 디맨드 비트-스트림으로 스위칭하는 것에 있을 수 있다.
- [0019] 그러한 경우들에, 603으로 표시된 스플라이싱된 비트-스트림은 그 각각의 초기화 세그먼트들(즉, 국제 표준 기구 베이스 미디어 파일 포맷(International Standard Organization Base Media File Format)에 따라 캡슐화될 때 'moov' 및 'trak' 박스들)에 그들의 조직이 기술된 601 및 602로 표시된 2개의 대안적인 비트-스트림의 조합에 기인한다.
- [0020] 2개의 스트림(601 및 602)이 상이한 동작 포인트들을 갖는 경우에, 스플라이싱된 비트-스트림(603)은 2개의 상이한 동작 포인트 리스트의 연결(concatenation)을 포함해야 한다. 이는 예를 들어, 상이한 VPS(비디오 파라미터 세트)로 스플라이싱된 HEVC 비트-스트림들을 저장할 때 발생할 수 있다.
- [0021] 그러나, 2개의 상이한 동작 포인트 리스트의 연결을 L-HEVC의 현재 캡슐화 포맷에 따라 스플라이싱된 비트-스트림과 동적으로 연관시키는 것은 불가능하다. 그 계층 조직, 그 스케일 가능성 타입 또는 그 프로파일, 티어 또는 레벨(실제로는 동작 포인트 디스크립션의 임의의 파라미터)이 시간에 따라 변화하는 비디오 비트-스트림들에도 유사한 한계가 있다; 기술적 메타데이터(예를 들어, 'trak' 파일 포맷 박스들의 계층 구조)는 동적으로 업데이트할 수 없다.
- [0022] 그러한 문제들은, 예를 들어, 계층 구성의 변화와 함께 장기간 계속되는 프래그먼트화된 MP4 파일을 스트리밍할 때 직면할 수 있다. 그러한 경우에, 콘텐츠 제작자는 2개의 부류의 디바이스(예를 들어, SD, HD)에 대해 2개의 스케일러블 계층을 정의한다. 일정 기간 후에, 새로운 부류의 디바이스(예를 들어, UHD)가 사용 가능하다면, 2개의 트랙을 재사용하고 추가 계층을 추가하는 것이 가능해야 한다. 무비 프래그먼트들이 사용되는 구성에서는, 프래그먼트들을 제거하는 작업이 정보 손실로 이어지지 않아야 한다. 스트리밍 동안 계층 구성이 변경되면 이것이 캡처되어야 한다.
- [0023] 또 다른 예는 동일한 프로파일 제약들을 준수하는 파일들의 연결을 대상으로 한다. 그러한 연결은 상이한 전략들을 따를 수 있다:
- [0024] - 양쪽 모두의 파일들로부터의 베이스 계층들을 포함하는 트랙들 내의 샘플들은 단순히 연결되어, 상이한 샘플 엔트리들 내의 또는 더 큰 hvcC NALU 어레이들 내의 다수의 VPS/SPS/PPS로 이어질 수 있다;
- [0025] - 비-베이스 계층들로부터의 샘플들은 트랙들을 하나씩 검사하고 이들을, 있다면, 유사한 제약들을 갖는 계층들에 대응하는 트랙들로부터의 샘플들과 연결함으로써 연결될 수 있다;
- [0026] - 제2 파일의 비-베이스 계층 트랙들로부터의 샘플들은 연결된 베이스 트랙과의 동기화를 유지하기 위해 시간상 시프트된 새로운 트랙들의 세트에 추가될 수 있다. 후자의 접근법은 복잡하고 바람직하지 않을 수 있다. 그러한 시나리오에서는, 트랙-계층 구성 변경을 허용하는 것이 유익할 수 있다.
- [0027] 이러한 한계들은 동작 포인트들에 대한 현재 디스크립터가 전체 mp4 파일에 대해 하나의 단일 인스턴스로 선언된다는 사실에서 기인한다.
- [0028] 인코딩된 비디오 비트-스트림들은 일반적으로, 완전한 프레임들에 대응하는 인접한 시간적 샘플들의 세트로서 구성되는 네트워크 추상화 계층(Network Abstraction Layer, NAL) 유닛들로 조직되고, 이 시간적 샘플들은 디코딩 순서의 함수로서 조직된다는 점이 상기되어야 한다. 파일 포맷들은 그러한 인코딩된 비트-스트림들을 캡슐화하고 기술하는 데 사용된다.
- [0029] 예시를 위해, 국제 표준 기구 베이스 미디어 파일 포맷(ISO BMFF)은, 네트워크를 통한 또는 또 다른 비트-스트림 전달 메커니즘을 통한 로컬 저장 또는 전송을 위해 인코딩된 시간 설정형 미디어 데이터 비트-스트림들을 기술하는 잘 알려진 유연하고 확장가능한 포맷이다. 이러한 파일 포맷은 객체 지향형이다. 그것은 순차적 또는 계층적으로 조직되는 그리고 타이밍 및 구조 파라미터들과 같은 인코딩 시간 설정형 미디어 데이터 비트-스트림

의 파라미터들을 정의하는 박스들이라 불리는 빌딩 블록들로 구성된다. 이 파일 포맷에 따르면, 시간 설정형 미디어 데이터 비트-스트림은, 트랙 박스(track box)라고 지칭되는 또 다른 데이터 구조로 정의되는 mdat 박스(mdat box)라고 지칭되는 데이터 구조에 포함된다. 이 트랙은, 하나의 샘플이 단일 타임스탬프와 연관된 모든 데이터, 즉, 단일 프레임과 연관된 모든 데이터 또는 동일 타임스탬프를 공유하는 수 개의 프레임들과 연관된 모든 데이터에 대응하는 샘플들의 시간 설정형 시퀀스를 나타낸다.

[0030] 다중-계층 HEVC 포맷의 비디오와 같은 스케일러블 비디오의 경우, 계층화된 미디어 데이터 조직은 다수의 종속 트랙들을 사용함으로써 효율적으로 표현될 수 있으며, 각 트랙은 특정 레벨의 스케일 가능성에서 비디오를 나타낸다. 트랙들 간의 데이터 중복을 회피하기 위해서, 추출기들이 사용될 수 있다. 표준 파일 포맷에 따르면, 추출기는, 다른 비트-스트림들로부터 네트워크 추상화 계층(NAL) 유닛들의 효과적인 추출을 가능하게 하는, 비트-스트림에 직접 포함되는 특정 종류의 네트워크 추상화 계층(NAL) 데이터 구조이다. 예를 들어, 인핸스먼트 계층 트랙(enhancement layer track)의 비트-스트림은 베이스 계층 트랙(base layer track)으로부터의 NAL 유닛들을 참조하는 추출기들을 포함할 수 있다. 그리고 나중에, 그러한 인핸스먼트 계층 트랙이 파일 포맷으로부터 추출될 때, 추출기들은 그들이 참조하고 있는 데이터로 대체되어야 한다.

[0031] 이러한 메커니즘들을 내포한 ISO BMFF를 사용하여, 하위-정보를 기술하고, 이러한 하위-정보로의 액세스를 용이하게 하거나 또는 비트-스트림들을 다수의 세그먼트들로 효과적으로 조직할 때 수 개의 전략들이 채택될 수 있다.

[0032] 예를 들어, "Implications of the ISO Base Media File Format on Adaptive HTTP Streaming of H.264/SVC"라는 제목의 논문에서, Kofler 등의 저자들은 ISO BMFF의 한계들뿐만 아니라 가능성들을 고려하여 HTTP 스트리밍을 위한 스케일러블 비디오 비트-스트림(H264/SVC)을 조직하기 위한 3개의 상이한 전략을 제시한다:

[0033] a) 모든 ISO BMFF 메타데이터(트랙 정의들을 포함함)를 포함하여 파일 타입 박스 "ftyp" 및 무비 박스 "moov"를 포함하는 특정 파일 헤더를 포함하는 단일 파일 - 이 단일 파일은 전체 인코딩된 비트-스트림을 포함하는 단일 mdat 박스를 또한 포함한다. 이 조직은 로컬 저장에는 적합하지만 클라이언트가 전체 비트-스트림의 일부만을 필요로 할 수 있는 HTTP 스트리밍에는 적응되지 않는다. 그러한 조직은 비트-스트림이 다수의 세그먼트로 프래그먼트화될 때 초기화 파일로서 사용되는 파일에 사용하는 것이 바람직하다. 이 초기화 파일 다음에는 그 조직이 b)에 정의된 다른 하나의 단일 파일이 따라오고, 이 초기화 파일은 모든 세그먼트에 대한 정보를 수집한다;

[0034] b) 프래그먼트화하기에 적합한 다수의 moof/mdat 박스를 포함하는 단일 파일 - moof/mdat의 각 쌍은 비트-스트림들의 다수의 세그먼트 중 하나에 관련됨. 이 포맷은 점진적인 다운로드를 허용한다. 더 상세하게는, moof 박스는 프래그먼트 레벨에서 moov 박스에 해당한다. 이러한 방식에 따라, 프래그먼트화된 미디어 파일을 사용하여, 스케일러블 비트-스트림은, 비디오를 상이한 스케일 가능성 레벨들로 표현하는 다수의 종속 트랙들로 나뉠 수 있다. 추출기들은 다른 트랙(들)으로부터의 NAL 유닛들을 참조하는 데 사용되는 특정 NAL 유닛들이다. 타일 당 트랙이 사용되는 경우, 모든 어드레스 가능형 트랙들은 미리 준비되어야 하고 트랙들은 독립적으로 선택될 수 없다. 수 개의 타일들이 디스플레이되어야 하면, 수 개의 비트-스트림들이 디코딩되어야 하고 베이스 계층이 수 회 디코딩된다. c)에 기술된 마지막 조직은 특히 각 트랙을 독립적으로 선택하는 데 적합하다;

[0035] c) 다수의 세그먼트 파일 - 각 파일은 그 자신의 URL에 의해 액세스 가능하고 독립적으로 다운로드 가능하다. 각 파일은 하나의 프래그먼트와 관련되고, 다수의 세그먼트 파일들 앞에는 전용 초기화 파일이 선행되는 것이 바람직하다. 각 세그먼트는 전형적으로 세그먼트 타입 박스(styp)로 구성되고, 이는 일종의 파일 헤더, 옵션의 세그먼트 인덱스 박스(sidx) 및 하나 또는 다수의 프래그먼트들로서의 역할을 한다. 다시, 각 프래그먼트는 moof 및 mdat 박스로 구성된다. 이러한 방식에 따라, 프래그먼트화된 미디어 파일을 사용하여, 각 트랙은 스케일 가능성의 하나의 레벨에 관한 연관된 비트-스트림과 함께 그 자신의 세그먼트에 저장된다. 필요하다면, 종속 트랙들로부터의 필요한 비트-스트림을 참조하는 데 추출기들이 사용된다. 그러한 코딩 방식은 트랙들을 독립적으로 스트리밍하는 데 특히 적합하다. 이는 DASH 표준에 잘 적응되지만, 수 개의 비트-스트림들이 디코딩되어야 하고, 이에 따라, 트랙 당 하나의 디코더가 요구되기 때문에, 타일 스트리밍에는 적합하지 않다. 더욱이, 둘 이상의 타일을 선택할 때 베이스 계층의 비트-스트림의 잠재적인 중복이 존재한다.

[0036] "Draft text of ISO/IEC DIS 14496-15 4th edition, ISO/IEC JTC1/SC29/WG11, W15182, April 2015, Geneva, Switzerland" (아래에서는 "w15182"로 명명됨)로 알려진 문헌을 참조하여 이루어진 위에서 언급된 박스들의 정의뿐만 아니라 그 박스들에 포함된 하위 박스들의 정의는 ISO BMFF 메타데이터의 복잡하고 덜 효율적인 조직으로 이어질 수 있다.

- [0037] 더욱이, 타일 트랙들은 계층화된 HEVC에 대해 적절하게 정의되어 있지 않아, 그 사용을 제한한다.
- [0038] 이러한 문제들을 해결하고, 특히, 동작 포인트들에 대한 디스크립터들을 동적으로 설정하는 것을 가능하게 하기 위해, 다중-계층 비디오 스트림들에 대한 계층화된 HEVC에서 공간적 타일들, 스케일러블 계층들 및 다수의 뷰를 처리하는 데 특히 적합한 효율적인 데이터 조직 및 트랙 디스크립션 방식이 제공된다. 이것은 ISO BMFF 분석의 결과가 더 효율적이고 계층화된 HEVC에 적용되는 것을 보장한다.

발명의 내용

해결하려는 과제

- [0039] 이러한 제약들과 직면하여, 본 발명자들은, 서버에서 다중-계층 타일형의 시간 설정형 미디어 데이터를 캡슐화 하기 위한, 그리고 복수의 미디어 세그먼트 파일들에 캡슐화된 다중-계층 타일형의 시간 설정형 미디어 데이터로부터 시간 설정형 미디어 데이터 비트-스트림을 제공하기 위한 방법 및 디바이스를 제공한다.
- [0040] 본 발명의 광범위한 목적은 위에 설명된 바와 같은 종래 기술의 단점들을 해결하는 것이다.

과제의 해결 수단

- [0041] 본 발명의 제1 양태에 따르면, 하나 이상의 비트-스트림의 캡슐화로부터 생기는 캡슐화된 비트-스트림으로부터 미디어 데이터 및 메타데이터를 획득하는 방법이 제공되는데, 상기 하나 이상의 비트-스트림 각각의 미디어 데이터는 하나 이상의 계층으로 조직되고, 각 계층은 시간 설정형 샘플들 또는 하위 샘플들을 포함하고, 상기 방법은:
- [0042] 상기 미디어 데이터의 시간 설정형 샘플들의 적어도 하나의 그룹의 적어도 하나의 정의를 포함하는 메타데이터를 참조하기 위한 데이터 구조를 획득하는 단계, 및
- [0043] 상기 하나 이상의 비트-스트림 각각에 대한 적어도 하나의 동작 포인트 디스크립터를 획득하는 단계 - 상기 적어도 하나의 동작 포인트 디스크립터는 상기 데이터 구조에 의해 참조된 메타데이터에 정의된 시간 설정형 샘플들의 그룹에 속하는 시간 설정형 샘플들이 독립적으로 디코딩되는 것을 가능하게 하는 기술적 정보에 관련된 메타데이터를 제공함 - 를 포함하고,
- [0044] 시간 설정형 샘플들의 하나의 그룹과 연관된 각 동작 포인트 디스크립터는 상기 그룹에 속하는 시간 설정형 샘플들의 기술적 정보를 정의한다.
- [0045] 따라서, 본 발명의 방법은 특히 표준 기구 베이스 미디어 파일 포맷에 따라 미디어 데이터를 전송할 때, 동작 포인트 디스크립터들을 동적으로 설정하고, 따라서, 서버와 클라이언트 디바이스 간에 전송되는 데이터의 양을 최적화하는 것을 가능하게 한다.
- [0046] 일 실시예에서, 상기 하나 이상의 비트-스트림 중 적어도 하나의 비트-스트림의 미디어 데이터는 수 개의 계층으로 조직되고, 그 계층들 각각은 시간적 스케일 가능성 레벨과 및/또는 공간적 스케일 가능성 레벨과 및/또는 SNR 스케일 가능성 레벨과 및/또는 멀티뷰 표현의 뷰와 연관된다.
- [0047] 일 실시예에서, 상기 캡슐화된 비트-스트림은 적어도 하나의 트랙을 포함하고, 상기 데이터 구조 및 상기 적어도 하나의 동작 포인트 디스크립터는 상기 적어도 하나의 트랙에서 수신된다.
- [0048] 일 실시예에서, 상기 캡슐화된 비트-스트림은 복수의 트랙을 포함하고, 상기 적어도 하나의 트랙 이외의 상기 복수의 트랙의 각 트랙은 상기 적어도 하나의 트랙을 참조한다.
- [0049] 일 실시예에서, 상기 하나 이상의 비트-스트림 각각의 각 계층은 상기 복수의 트랙 중 특정 트랙과 연관된다.
- [0050] 일 실시예에서, 상기 적어도 하나의 트랙은 상기 데이터 구조 및 상기 적어도 하나의 동작 포인트 디스크립터를 전송하기 위한 전용 트랙이다.
- [0051] 일 실시예에서, 상기 적어도 하나의 트랙은 시간 설정형 샘플들 또는 하위 샘플들을 포함한다.
- [0052] 일 실시예에서, 상기 방법은 적어도 하나의 계층 디스크립터를 획득하는 단계를 추가로 포함하고, 상기 적어도 하나의 계층 디스크립터는 상기 하나 이상의 비트-스트림 중 적어도 하나의 비트-스트림에서 계층 조직에 관련된 메타데이터를 제공한다.
- [0053] 일 실시예에서, 하나의 계층 디스크립터는 디폴트 계층 디스크립터로서 정의되어, 계층 조직에 관련된 메타데이

터와 연관되지 않은 시간 설정형 샘플들의 그룹에 속하는 시간 설정형 샘플들은 상기 디폴트 계층 디스크립터에 의해 제공된 메타데이터에 기초하여 디코딩된다.

- [0054] 일 실시예에서, 하나의 동작 포인트 디스크립터는 디폴트 동작 포인트 디스크립터로서 정의되어, 기술적 정보에 관련된 메타데이터와 연관되지 않은 시간 설정형 샘플들의 그룹에 속하는 시간 설정형 샘플들은 상기 디폴트 동작 포인트 디스크립터에 의해 제공된 메타데이터에 기초하여 디코딩된다.
- [0055] 일 실시예에서, 메타데이터를 참조하기 위한 상기 데이터 구조는 국제 표준 기구 베이스 미디어 파일 포맷에 따른 SampleToGroupBox 타입이다.
- [0056] 일 실시예에서, 상기 데이터 구조에 의해 참조되고 시간 설정형 샘플들의 적어도 하나의 그룹의 정의를 포함하는 메타데이터는 국제 표준 기구 베이스 미디어 파일 포맷에 따른 SampleGroupDescriptionBox 타입의 구조로 정의된다.
- [0057] 본 발명의 제2 양태에 따르면, 하나 이상의 비트-스트림의 캡슐화를 위해 캡슐화될 비트-스트림에 미디어 데이터 및 메타데이터를 제공하기 위한 방법이 제공되는데, 상기 하나 이상의 비트-스트림 각각의 미디어 데이터는 하나 이상의 계층으로 조직되고, 각 계층은 시간 설정형 샘플들 또는 하위 샘플들을 포함하고, 상기 방법은:
- [0058] 상기 미디어 데이터의 시간 설정형 샘플들의 적어도 하나의 그룹의 적어도 하나의 정의를 포함하는 메타데이터를 참조하기 위한 데이터 구조를 제공하는 단계, 및
- [0059] 상기 하나 이상의 비트-스트림 각각에 대한 적어도 하나의 동작 포인트 디스크립터를 제공하는 단계 - 상기 적어도 하나의 동작 포인트 디스크립터는 상기 데이터 구조에 의해 참조된 메타데이터에 정의된 시간 설정형 샘플들의 그룹에 속하는 시간 설정형 샘플들이 독립적으로 디코딩되는 것을 가능하게 하는 기술적 정보에 관련된 메타데이터를 제공함 - 를 포함하고,
- [0060] 시간 설정형 샘플들의 하나의 그룹과 연관된 각 동작 포인트 디스크립터는 상기 그룹에 속하는 시간 설정형 샘플들의 기술적 정보를 정의한다.
- [0061] 따라서, 본 발명의 방법은 특히 표준 기구 베이스 미디어 파일 포맷에 따라 미디어 데이터를 전송할 때, 동작 포인트 디스크립터들을 동적으로 설정하고, 따라서, 서버와 클라이언트 디바이스 간에 전송되는 데이터의 양을 최적화하는 것을 가능하게 한다.
- [0062] 일 실시예에서, 상기 하나 이상의 비트-스트림 중 적어도 하나의 비트-스트림의 미디어 데이터는 수 개의 계층으로 조직되고, 그 계층들 각각은 시간적 스케일 가능성 레벨과 및/또는 공간적 스케일 가능성 레벨과 및/또는 SNR 스케일 가능성 레벨과 및/또는 멀티뷰 표현의 뷰와 연관된다.
- [0063] 일 실시예에서, 상기 캡슐화될 비트-스트림은 적어도 하나의 트랙을 포함하고, 상기 데이터 구조 및 상기 적어도 하나의 동작 포인트 디스크립터는 상기 적어도 하나의 트랙에서 전송된다.
- [0064] 일 실시예에서, 상기 캡슐화된 비트-스트림은 복수의 트랙을 포함하고, 상기 적어도 하나의 트랙 이외의 상기 복수의 트랙의 각 트랙은 상기 적어도 하나의 트랙을 참조한다.
- [0065] 일 실시예에서, 상기 하나 이상의 비트-스트림 각각의 각 계층은 상기 복수의 트랙 중 특정 트랙과 연관된다.
- [0066] 일 실시예에서, 상기 적어도 하나의 트랙은 상기 데이터 구조 및 상기 적어도 하나의 동작 포인트 디스크립터를 전송하기 위한 전용 트랙이다.
- [0067] 일 실시예에서, 상기 적어도 하나의 트랙은 시간 설정형 샘플들 또는 하위 샘플들을 포함한다.
- [0068] 일 실시예에서, 상기 방법은 적어도 하나의 계층 디스크립터를 제공하는 단계를 추가로 포함하고, 상기 적어도 하나의 계층 디스크립터는 상기 하나 이상의 비트-스트림 중 적어도 하나의 비트-스트림에서 계층 조직에 관련된 메타데이터를 제공한다.
- [0069] 일 실시예에서, 하나의 계층 디스크립터는 디폴트 계층 디스크립터로서 정의되어, 계층 조직에 관련된 메타데이터와 연관되지 않은 시간 설정형 샘플들의 그룹에 속하는 시간 설정형 샘플들은 상기 디폴트 계층 디스크립터에 의해 제공된 메타데이터에 기초하여 디코딩된다.
- [0070] 일 실시예에서, 하나의 동작 포인트 디스크립터는 디폴트 동작 포인트 디스크립터로서 정의되어, 기술적 정보에 관련된 메타데이터와 연관되지 않은 시간 설정형 샘플들의 그룹에 속하는 시간 설정형 샘플들은 상기 디폴트 동

작 포인트 디스크립터에 의해 제공된 메타데이터에 기초하여 디코딩된다.

- [0071] 일 실시예에서, 메타데이터를 참조하기 위한 상기 데이터 구조는 국제 표준 기구 베이스 미디어 파일 포맷에 따른 SampleToGroupBox 타입이다.
- [0072] 일 실시예에서, 상기 데이터 구조에 의해 참조되고 시간 설정형 샘플들의 적어도 하나의 그룹의 정의를 포함하는 메타데이터는 국제 표준 기구 베이스 미디어 파일 포맷에 따른 SampleGroupDescriptionBox 타입의 구조로 정의된다.
- [0073] 본 발명의 제3 양태에 따르면, 하나 이상의 비트-스트림의 캡슐화로부터 생기는 캡슐화된 비트-스트림으로부터 미디어 데이터 및 메타데이터를 획득하기 위한 디바이스가 제공되는데, 상기 하나 이상의 비트-스트림 각각의 미디어 데이터는 하나 이상의 계층으로 조직되고, 각 계층은 시간 설정형 샘플들 또는 하위 샘플들을 포함하고, 상기 디바이스는:
- [0074] 상기 미디어 데이터의 시간 설정형 샘플들의 적어도 하나의 그룹의 적어도 하나의 정의를 포함하는 메타데이터를 참조하기 위한 데이터 구조를 획득하는 단계, 및
- [0075] 상기 하나 이상의 비트-스트림 각각에 대한 적어도 하나의 동작 포인트 디스크립터를 획득하는 단계 - 상기 적어도 하나의 동작 포인트 디스크립터는 상기 데이터 구조에 의해 참조된 메타데이터에 정의된 시간 설정형 샘플들의 그룹에 속하는 시간 설정형 샘플들이 독립적으로 디코딩되는 것을 가능하게 하는 기술적 정보에 관련된 메타데이터를 제공함 - 를 수행하도록 구성된 프로세서를 포함하고,
- [0076] 시간 설정형 샘플들의 하나의 그룹과 연관된 각 동작 포인트 디스크립터는 상기 그룹에 속하는 시간 설정형 샘플들의 기술적 정보를 정의한다.
- [0077] 따라서, 본 발명의 디바이스는 특히 표준 기구 베이스 미디어 파일 포맷에 따라 미디어 데이터를 전송할 때, 동작 포인트 디스크립터들을 동적으로 설정하고, 따라서, 서버와 클라이언트 디바이스 간에 전송되는 데이터의 양을 최적화하는 것을 가능하게 한다.
- [0078] 일 실시예에서, 상기 하나 이상의 비트-스트림 중 적어도 하나의 비트-스트림의 미디어 데이터는 수 개의 계층으로 조직되고, 그 계층들 각각은 시간적 스케일 가능성 레벨과 및/또는 공간적 스케일 가능성 레벨과 및/또는 SNR 스케일 가능성 레벨과 및/또는 멀티뷰 표현의 뷰와 연관된다.
- [0079] 일 실시예에서, 상기 캡슐화된 비트-스트림은 적어도 하나의 트랙을 포함하고, 상기 데이터 구조 및 상기 적어도 하나의 동작 포인트 디스크립터는 상기 적어도 하나의 트랙에서 수신된다.
- [0080] 일 실시예에서, 상기 캡슐화된 비트-스트림은 복수의 트랙을 포함하고, 상기 적어도 하나의 트랙 이외의 상기 복수의 트랙의 각 트랙은 상기 적어도 하나의 트랙을 참조한다.
- [0081] 일 실시예에서, 상기 하나 이상의 비트-스트림 각각의 각 계층은 상기 복수의 트랙 중 특정 트랙과 연관된다.
- [0082] 일 실시예에서, 상기 적어도 하나의 트랙은 상기 데이터 구조 및 상기 적어도 하나의 동작 포인트 디스크립터를 전송하기 위한 전용 트랙이다.
- [0083] 일 실시예에서, 상기 적어도 하나의 트랙은 시간 설정형 샘플들 또는 하위 샘플들을 포함한다.
- [0084] 일 실시예에서, 상기 프로세서는 또한 적어도 하나의 계층 디스크립터를 획득하는 단계를 수행하도록 구성되고, 상기 적어도 하나의 계층 디스크립터는 상기 하나 이상의 비트-스트림 중 적어도 하나의 비트-스트림에서 계층 조직에 관련된 메타데이터를 제공한다.
- [0085] 일 실시예에서, 하나의 계층 디스크립터는 디폴트 계층 디스크립터로서 정의되어, 계층 조직에 관련된 메타데이터와 연관되지 않은 시간 설정형 샘플들의 그룹에 속하는 시간 설정형 샘플들은 상기 디폴트 계층 디스크립터에 의해 제공된 메타데이터에 기초하여 디코딩된다.
- [0086] 일 실시예에서, 하나의 동작 포인트 디스크립터는 디폴트 동작 포인트 디스크립터로서 정의되어, 기술적 정보에 관련된 메타데이터와 연관되지 않은 시간 설정형 샘플들의 그룹에 속하는 시간 설정형 샘플들은 상기 디폴트 동작 포인트 디스크립터에 의해 제공된 메타데이터에 기초하여 디코딩된다.
- [0087] 일 실시예에서, 메타데이터를 참조하기 위한 상기 데이터 구조는 국제 표준 기구 베이스 미디어 파일 포맷에 따른 SampleToGroupBox 타입이다.

- [0088] 일 실시예에서, 상기 데이터 구조에 의해 참조되고 시간 설정형 샘플들의 적어도 하나의 그룹의 정의를 포함하는 메타데이터는 국제 표준 기구 베이스 미디어 파일 포맷에 따른 SampleGroupDescriptionBox 타입의 구조로 정의된다.
- [0089] 본 발명의 제4 양태에 따르면, 하나 이상의 비트-스트림의 캡슐화를 위해 캡슐화될 비트-스트림에 미디어 데이터 및 메타데이터를 제공하기 위한 디바이스가 제공되는데, 상기 하나 이상의 비트-스트림 각각의 미디어 데이터는 하나 이상의 계층으로 조직되고, 각 계층은 시간 설정형 샘플들 또는 하위 샘플들을 포함하고, 상기 디바이스는:
- [0090] 상기 미디어 데이터의 시간 설정형 샘플들의 적어도 하나의 그룹의 적어도 하나의 정의를 포함하는 메타데이터를 참조하기 위한 데이터 구조를 제공하는 단계,
- [0091] 상기 하나 이상의 비트-스트림 각각에 대한 적어도 하나의 동작 포인트 디스크립터를 제공하는 단계 - 상기 적어도 하나의 동작 포인트 디스크립터는 상기 데이터 구조에 의해 참조된 메타데이터에 정의된 시간 설정형 샘플들의 그룹에 속하는 시간 설정형 샘플들이 독립적으로 디코딩되는 것을 가능하게 하는 기술적 정보에 관련된 메타데이터를 제공함 - 를 수행하도록 구성된 프로세서를 포함하고,
- [0092] 시간 설정형 샘플들의 하나의 그룹과 연관된 각 동작 포인트 디스크립터는 상기 그룹에 속하는 시간 설정형 샘플들의 기술적 정보를 정의한다.
- [0093] 따라서, 본 발명의 디바이스는 특히 표준 기구 베이스 미디어 파일 포맷에 따라 미디어 데이터를 전송할 때, 동작 포인트 디스크립터들을 동적으로 설정하고, 따라서, 서버와 클라이언트 디바이스 간에 전송되는 데이터의 양을 최적화하는 것을 가능하게 한다.
- [0094] 일 실시예에서, 상기 하나 이상의 비트-스트림 중 적어도 하나의 비트-스트림의 미디어 데이터는 수 개의 계층으로 조직되고, 그 계층들 각각은 시간적 스케일 가능성 레벨과 및/또는 공간적 스케일 가능성 레벨과 및/또는 SNR 스케일 가능성 레벨과 및/또는 멀티뷰 표현의 뷰와 연관된다.
- [0095] 일 실시예에서, 상기 캡슐화될 비트-스트림은 적어도 하나의 트랙을 포함하고, 상기 데이터 구조 및 상기 적어도 하나의 동작 포인트 디스크립터는 상기 적어도 하나의 트랙에서 전송된다.
- [0096] 일 실시예에서, 상기 캡슐화된 비트-스트림은 복수의 트랙을 포함하고, 상기 적어도 하나의 트랙 이외의 상기 복수의 트랙의 각 트랙은 상기 적어도 하나의 트랙을 참조한다.
- [0097] 일 실시예에서, 상기 하나 이상의 비트-스트림 각각의 각 계층은 상기 복수의 트랙 중 특정 트랙과 연관된다.
- [0098] 일 실시예에서, 상기 적어도 하나의 트랙은 상기 데이터 구조 및 상기 적어도 하나의 동작 포인트 디스크립터를 전송하기 위한 전용 트랙이다.
- [0099] 일 실시예에서, 상기 적어도 하나의 트랙은 시간 설정형 샘플들 또는 하위 샘플들을 포함한다.
- [0100] 일 실시예에서, 상기 프로세서는 또한 적어도 하나의 계층 디스크립터를 제공하는 단계를 수행하도록 구성되고, 상기 적어도 하나의 계층 디스크립터는 상기 하나 이상의 비트-스트림 중 적어도 하나의 비트-스트림에서 계층 조직에 관련된 메타데이터를 제공한다.
- [0101] 일 실시예에서, 하나의 계층 디스크립터는 디폴트 계층 디스크립터로서 정의되어, 계층 조직에 관련된 메타데이터와 연관되지 않은 시간 설정형 샘플들의 그룹에 속하는 시간 설정형 샘플들은 상기 디폴트 계층 디스크립터에 의해 제공된 메타데이터에 기초하여 디코딩된다.
- [0102] 일 실시예에서, 하나의 동작 포인트 디스크립터는 디폴트 동작 포인트 디스크립터로서 정의되어, 기술적 정보에 관련된 메타데이터와 연관되지 않은 시간 설정형 샘플들의 그룹에 속하는 시간 설정형 샘플들은 상기 디폴트 동작 포인트 디스크립터에 의해 제공된 메타데이터에 기초하여 디코딩된다.
- [0103] 일 실시예에서, 메타데이터를 참조하기 위한 상기 데이터 구조는 국제 표준 기구 베이스 미디어 파일 포맷에 따른 SampleToGroupBox 타입이다.
- [0104] 일 실시예에서, 상기 데이터 구조에 의해 참조되고 시간 설정형 샘플들의 적어도 하나의 그룹의 정의를 포함하는 메타데이터는 국제 표준 기구 베이스 미디어 파일 포맷에 따른 SampleGroupDescriptionBox 타입의 구조로 정의된다.

[0105] 본 발명은 소프트웨어로 구현될 수 있으므로, 본 발명은 임의의 적합한 캐리어 매체 상에 프로그램 가능 장치에 제공하기 위한 컴퓨터 판독 가능 코드로서 구현될 수 있다. 유형의 캐리어 매체는 플로피 디스크, CD-ROM, 하드 디스크 드라이브, 자기 테이프 디바이스 또는 솔리드 스테이트 메모리 디바이스 등과 같은 저장 매체를 포함할 수 있다. 일시적인 캐리어 매체는 전기 신호, 전자 신호, 광 신호, 음향 신호, 자기 신호 또는 전자기 신호, 예를 들어 마이크로웨이브 또는 RF 신호와 같은 신호를 포함할 수 있다.

도면의 간단한 설명

[0106] 본 발명의 추가의 이점들은 도면들 및 상세한 설명의 검토에 의해 본 기술 분야의 통상의 기술자들에게 명백해질 것이다. 임의의 추가적인 이점들도 본 명세서에 포함될 수 있는 것으로 이해된다.

이제, 본 발명의 실시예들이 단지 예로서, 그리고 다음의 도면들을 참조하여 설명될 것이다:

도 1은 MPEG-4 파일 포맷에 따라 2개의 스케일 가능성 계층을 캡슐화하는 예를 도시한다.

도 2는 타일형의 계층화된 HEVC에 대한 본 발명에 따른 트랙 박스들을 나타내는 실시예를 도시한다.

도 3은 하나 이상의 실시예가 구현될 수 있는 서버 또는 클라이언트 디바이스의 블록도를 나타낸다.

도 4a 및 도 4b를 포함하는 도 4는 대안적인 트랙 및 타일 트랙 종속성 시그널링의 예들을 도시한다.

도 5a 및 도 5b를 포함하는 도 5는 비-출력 계층이 수반되는 계층 구성의 예들을 도시한다.

도 6a, 도 6b 및 도 6c를 포함하는 도 6은 비트-스트림 스플라이싱 애플리케이션의 예를 도시한다.

도 7은 본 발명의 실시예들에 따른 동작 포인트들에 대한 특정하고 동적으로 확장 가능한 디스크립터를 도시한다.

도 8은 표준 동작 포인트 디스크립터를 도시한다.

도 9는 계층 정보 디스크립션을 위한 특정하고 동적으로 확장 가능한 디스크립터를 도시한다.

도 10은 본 발명에 따라 캡슐화된 미디어 파일들에 대한 클라이언트 측의 처리를 도시한다.

도 11은 본 발명에 따른 새로운 디코더 구성 정보를 도시한다.

도 12는 본 발명에 따른 비디오 비트-스트림 캡슐화를 위한 서버 측의 처리를 도시한다.

발명을 실시하기 위한 구체적인 내용

[0107] 아래의 3개의 파트는 3개의 상이한 특징들, 각각 픽처들의 공간적 조직, NAL 유닛들, 및 VisualSampleEntry라 불리는 특정 디스크립터에 관한 알려진 정보를 설명한다. 이들 특징들은 도 1 내지 도 3에 나타난 실시예들을 더 잘 이해하기 위해 여기에 설명된다.

[0108] 픽처들의 공간적 조직

[0109] 비디오는 바람직하게는 스케일러블 또는 멀티뷰 비디오이고 상이한 레벨들의 스케일 가능성/뷰들로 조직된다.

[0110] 특정한 실시예에 따르면, 시간 설정형 샘플들(예를 들어, 이미지들)을 포함하는 다중-계층 타일형의 시간 설정형 미디어 데이터(예를 들어, 스케일러블 또는 멀티뷰 타일형 비디오 데이터)와 같은 다중-계층 파티셔닝된 시간 설정형 미디어 데이터는 수 개의 시간 설정형 미디어 데이터 트랙, 전형적으로 베이스 트랙들 및 타일 트랙들의 세트로서 전송된다. 특정한 실시예에 따르면, 베이스 트랙들은 베이스 계층 베이스 트랙 및 적어도 하나의 인헨스먼트 계층 베이스 트랙을 포함한다. 추가적인 타일 트랙들은 베이스 계층 타일 트랙들 및/또는 인헨스먼트 계층 타일 트랙들일 수 있다. 각 시간 설정형 미디어 데이터 트랙은 수 개의 시간 설정형 샘플들의 하나의 공간적 하위 샘플(예를 들어, 수 개의 NAL 유닛들)을 포함한다. 각 비디오 프레임(시간 설정형 샘플)은 비디오 프레임의 공간적 하위 파트들(공간적 하위 샘플들)에 대응하는 독립적으로 디코딩 가능한 타일들로 구성될 수 있다. 계층화된 HEVC에서, 각 비디오 프레임은 비디오 프레임의 공간적 하위 파트(공간적 하위 샘플들)에 대응하는 종속적으로 디코딩 가능한 계층들로 구성될 수 있다. 또한, 계층화된 HEVC에서, 각 비디오 프레임은 종속적으로 디코딩 가능한 계층들로 구성될 수 있고, 각 계층은 비디오 프레임의 공간적 하위 파트들(공간적 하위 샘플들)에 대응하는 독립적으로 디코딩 가능한 타일들(주어진 계층에 대한)로 구성될 수 있다.

[0111] 리스트는 트랙 종속성들(타일링, 인터-계층 및/또는 인트라-계층 종속성들)을 기술하는 데 사용된다. 시간 설

정형 미디어 데이터 트랙들의 그러한 세트는 다중-계층 공간적 비디오 타일들의 선택, 구성, 및 효율적 스트리밍을 허용한다. 각 트랙은 미디어 세그먼트 파일들의 세트로서 서버 디바이스로부터 클라이언트 디바이스로 전송될 수 있다. 초기화 세그먼트 파일은 미디어 세그먼트 파일들을 디코딩하기 위해 요구되는 메타데이터를 전송하는 데 사용될 수 있다.

[0112] 본 발명의 실시예는, 예를 들어 HEVC 또는 계층화된 HEVC(LHVC 또는 다중-계층 HEVC로도 알려짐)로서 알려진 비디오 포맷에 적용될 수 있다.

[0113] HEVC 표준에 따르면, 이미지들은 타일들, 슬라이스들, 및 슬라이스 세그먼트들로 공간적으로 분할될 수 있다는 점을 상기한다. 이러한 표준에서, 타일은 수평 및 수직 경계들(즉, 행들 및 열들)에 의해 정의되는 이미지의 직사각형 영역에 대응한다. 그것은 정수 개의 코딩 트리 유닛(Coding Tree Unit)(CTU)들을 포함한다. 따라서, 타일들은, 예를 들어 관심 영역들에 대한 위치들 및 사이즈들을 정의함으로써 관심 영역들을 식별하는 데 효율적으로 사용될 수 있다. 그러나, HEVC 비트-스트림의 구조뿐만 아니라 네트워크 추상화 계층(NAL) 유닛들로서의 그것의 캡슐화는 타일들에 관련하여 조직되지 않고 슬라이스들에 기반한다.

[0114] HEVC 표준에서, 슬라이스들은 슬라이스 세그먼트들의 세트들이고, 슬라이스 세그먼트들의 세트의 제1 슬라이스 세그먼트는 독립 슬라이스 세그먼트로서, 즉, 헤더 내에 저장된 일반 정보가 또 다른 슬라이스 세그먼트의 것을 참조하지 않는 슬라이스 세그먼트이다. 슬라이스 세그먼트들의 세트의 다른 슬라이스 세그먼트들은, 있다면, 종속 슬라이스 세그먼트들(즉, 헤더 내에 저장된 일반 정보가 독립 슬라이스 세그먼트의 일반 정보를 참조하는 슬라이스 세그먼트들)이다.

[0115] 슬라이스 세그먼트는 정수 개의 연속적인(래스터 스캔 순서로) 코딩 트리 유닛들을 포함한다. 따라서, 슬라이스 세그먼트는 직사각형 형상일 수도 또는 아닐 수도 있고, 따라서, 관심 영역을 나타내기에 적합하지 않다. 그것은 슬라이스 세그먼트 헤더와 그에 후속하는 슬라이스 세그먼트 데이터에 대해 HEVC 비트-스트림으로 인코딩된다. 독립 및 종속 슬라이스 세그먼트들은 그들의 헤더에 의해 상이하다: 종속 슬라이스 세그먼트는 독립 슬라이스 세그먼트에 종속하고, 그것의 헤더의 정보량은 독립 슬라이스 세그먼트의 정보량보다 작다. 독립 및 종속 슬라이스 세그먼트들 양쪽 모두는 타일들을 정의하는 데 사용되거나 엔트로피 디코딩 동기화 포인트(entropy decoding synchronization point)들로서 사용되는 비트-스트림에서의 엔트리 포인트들의 리스트를 포함한다.

[0116] HEVC 표준에 따라, 슬라이스 세그먼트들은 다음과 같이 요약될 수 있는 규칙들에 따라 타일들에 링크된다(하나 또는 양쪽 모두의 조건들이 만족되어야 함):

[0117] - 슬라이스 세그먼트의 모든 CTU들이 동일한 타일에 속함(즉, 슬라이스 세그먼트는 수 개의 타일에 속할 수 없다); 및

[0118] - 타일의 모든 CTU들이 동일한 슬라이스 세그먼트에 속함(즉, 타일은 수 개의 슬라이스 세그먼트로 분할될 수 있다 - 이 슬라이스 세그먼트들 각각이 해당 타일만에 속한다면).

[0119] NAL 유닛들

[0120] 상술한 바와 같이, 타일들이 관심 영역들에 대한 적절한 지원으로서 간주될 수 있지만, 슬라이스 세그먼트들은 통신 네트워크를 통한 전송을 위해 실제로 NAL 유닛들 내에 삽입되고 액세스 유닛들(즉, 파일 포맷 레벨에서의 코딩된 픽처 또는 샘플들)을 형성하기 위해 집성되는 엔티티들이다.

[0121] HEVC 표준에 따라, NAL 유닛의 타입이 다음과 같이 정의될 수 있는 NAL 유닛 헤더의 2개의 바이트로 인코딩된다는 점을 상기해야 한다:

```
nal_unit_header () {
    forbidden_zero_bit
    nal_unit_type
    nuh_layer_id
    nuh_temporal_id_plus1
}
```

[0122]

[0123]

슬라이스 세그먼트들을 코딩하는 데 사용된 NAL 유닛들은 슬라이스 세그먼트 어드레스 구문 요소 덕분에 슬라이스 세그먼트의 제1 CTU의 어드레스를 나타내는 슬라이스 세그먼트 헤더들을 포함한다. 그러한 슬라이스 세그먼트 헤더들은 다음과 같이 정의될 수 있다:

```
slice_segment_header () {
    first_slice_segment_in_pic_flag
    if(nal_unit_type >= BLA_W_LP && nal_unit_type <= RSV_IRAP_VCL23)
        no_output_of_prior_pics_flag
    slice_pic_parameter_set_id
    if(!first_slice_segment_in_pic_flag){
        if(dependent_slice_segments_enabled_flag)
            dependent_slice_segment_flag
        slice_segment_address
    }
    if(!dependent_slice_segment_flag){
        [...]
    }
```

[0124]

[0125]

타일링 정보가 PPS(Picture Parameter Set) NAL 유닛에 제공된다. 그 후 슬라이스 세그먼트와 타일 간의 관계는 이러한 파라미터들로부터 추론될 수 있다.

[0126]

공간적 예측들이 (정의에 의해) 타일 경계들 상에서 리셋되지만, 아무것도 타일이 참조 프레임(들) 내의 상이한 타일로부터 시간적 예측자들을 이용하는 것을 막지 못한다. 따라서, 독립 타일들을 구축하기 위해, 예측 유닛들을 위한 모션 벡터들이 유리하게도, 인코딩 동안에, 참조 프레임(들) 내의 동위치의 타일에 남아 있도록, 타일 내부에 제약된다. 또한, 인-루프 필터(in-loop filter)(디블로킹(deblocking) 및 샘플 적응 오프셋(sample adaptive offset)(SAO) 필터)들은 바람직하게는 타일 경계들 상에서 비활성화되어, 하나의 타일만을 디코딩하는 경우에 어떠한 에러 드리프트(error drift)도 도입되지 않는다. 그러한 인-루프 필터들의 제어가 HEVC 표준에서 이용 가능하다는 것에 주목해야 한다. 그것은 loop_filter_across_tiles_enabled_flag로서 알려진 플래그를 이용하여 슬라이스 세그먼트 헤더에서 설정된다. 이러한 플래그를 0으로 명시적으로 설정함으로써, 타일 경계들에서의 픽셀들은 이웃 타일들의 경계 상에 속하는 픽셀들에 종속될 수 없다. 모션 벡터 및 인-루프 필터들과 관련된 이 두 조건이 충족되면, 타일들은 "독립적으로 디코딩 가능한 타일들" 또는 "독립 타일들"로 간주될 수 있다.

[0127]

VisualSampleEntry라 불리는 특정 디스크립터

[0128]

MPEG-4 Part 12 표준의 기존 샘플 그룹화 메커니즘이 타일들을 캡슐화하는 데 사용될 수 있다. 따라서, 특정한 샘플 그룹 디스크립션은 특정한 종류의 표준 VisualSampleGroupEntry 디스크립터인 타일 디스크립터를 이용하여 생성된다. 샘플 그룹화 메커니즘들은 트랙들에서의 샘플들의 파티션들을 나타내기 위해 사용된다. 이들은 다음 2개의 박스의 사용에 의존한다: 샘플 그룹들로의 샘플들의 할당을 기술하는 SampleToGroup 박스('sbgp') 및 특정 샘플 그룹 내의 샘플들의 공통 속성들을 기술하는 SampleGroupDescription 박스('sgpd'). 특정 타입의 샘플 그룹화가 타입 필드('grouping_type')를 통해 하나의 SampleToGroup 박스와 하나의 SampleGroupDescription 박스의 조합에 의해 정의된다. 다수의 샘플 그룹화 인스턴스(즉, SampleToGroup 박스와 SampleGroupDescription 박스의 쌍)가 상이한 그룹화 기준들에 기초하여 존재할 수 있다.

[0129]

샘플의 타일링과 관련된 특정한 그룹화 기준이 이용된다. 'trif'로 지칭되는 이러한 특정한 그룹화 타입은 타일의 속성들을 기술하고 표준 VisualSampleGroupEntry로부터 도출된다. 이것은 TileRegionSampleGroupEntry로서 지칭될 수 있으며 다음과 같이 정의된다:

```
class TileRegionGroupEntry () extends VisualSampleGroupEntry ('trif') {
    unsigned int(16) groupID;
    unsigned int(2) independent_flag;
    unsigned int(1) full_picture;
    unsigned int(1) filtering_disabled;
    unsigned int(4) reserved=0;
    if (!full_picture) {
        unsigned int(16) horizontal_offset;
        unsigned int(16) vertical_offset;
    }
    unsigned int(16) region_width;
    unsigned int(16) region_height;
}
```

[0130]

[0131]

이 특정한 타입의 그룹 엔트리에 따르면, groupID 파라미터는 그룹에 의해 기술된 타일에 대한 고유 식별자이다. horizontal_offset 및 vertical_offset 파라미터는 베이스 영역의 루마 샘플들에서, HEVC 프레임의 상부-좌측 픽셀에 대하여, 타일에 의해 표현된 직사각형 영역의 상부-좌측 픽셀의 수평 및 수직 오프셋을 각각 설정하는 데 이용된다. region_width 및 region_height 파라미터는 HEVC 프레임의 루마 샘플들에서, 타일에 의해 표현된 직사각형 영역의 폭 및 높이를 각각 설정하는 데 이용된다.

[0132]

independent_flag 파라미터는 독립 타일들의 정의를 참조하여 상술한 바와 같이, 동일한 타일에만 속하는 샘플들에 관련된 디코딩 종속성들을 타일이 포함하는 것을 명시하는 2-비트 워드이다. 예시를 위해 그리고 타일 조직을 기술하기 위한 SEI 메시지(Supplemental Enhancement Information)의 표준 이용을 참조하여, tile_section_exact_match_flag로 알려진 플래그는 그 의미가 다음과 같이 설정될 수 있는 independent_flag 파라미터의 값을 설정하는 데 사용될 수 있다:

[0133]

- independent_flag 파라미터가 0인 경우에는, 이 타일과 동일한 프레임 또는 이전의 프레임들에서의 다른 타일들과의 사이의 코딩 종속성들은 타일 세트 레벨에서 기술되거나 또는 알려지지 않고;

[0134]

- independent_flag 파라미터가 1인 경우에는, 이 타일과 임의의 참조 프레임들에서의 상이한 groupID를 갖는 다른 타일들과의 사이에는 시간적 코딩 종속성들이 존재할 수 없지만 이 타일과 참조 프레임들에서의 동일한 groupID를 갖는 타일들과의 사이에는 코딩 종속성들이 존재할 수 있고,

[0135]

- independent_flag 파라미터가 2인 경우에는, 이 타일과 동일한 프레임에서의 다른 타일들과의 사이에는 코딩 종속성들이 존재하지 않고, 이 타일과 참조 프레임들에서의 임의의 다른 타일들과의 사이에는 코딩 종속성들이 존재하지 않는다.

[0136]

independent_flag 파라미터 값 3이 예약된다.

[0137]

full_picture 파라미터는, 1로 설정되는 경우, 이 타일 영역이 실제로 완전한 픽처임을 나타낸다. 그러한 경우, region_width 및 region_height는 계층 루마 사이즈로 설정되어야 하고, independent_flag는 1로 설정되어야 한다.

[0138]

filtering_disable 파라미터는, 1로 설정되는 경우, 이 타일 영역에 대한 포스트-디코딩 필터링 연산이 이 타일 영역에 인접한 픽셀들에 대한 액세스를 필요로 하지 않는다는 것, 즉 인접한 타일들을 디코딩하지 않고 타일 영역의 비트-정확한 재구성이 가능하다는 것을 나타낸다.

[0139]

각 타일 트랙에 대해, 'trif' grouping_type 및 TileRegionGroupEntry를 갖는 하나의 SampleGroupDescription 박스('sgpd')를 정의함으로써 각 타일의 속성들이 무비 헤더('moov' 박스)에 한 번 주어진다. 타일 속성들은 또한 트랙 프레임마다 정의될 수 있다. 그러한 mp4 트랙은 비디오 타일 트랙 또는 타일 트랙으로 정의될 수 있다. HEVC 표준에 따르면, HEVC 타일 트랙은 이러한 트랙 내의 타일(들)이 속한 HEVC 계층의 다른 NALU(전형적으로 다양한 파라미터 세트들과 같은 셋업 정보)들을 전달하는 HEVC 트랙에 대한 참조가 있는 비디오 타일

트랙이다. 참조는 'sbas' 4 문자 코드와 같은 MPEG-4 Part 15 표준에 이미 정의된 값들, 또는 'tbas'와 같은 더 특정한 것을 사용하여 타일 베이스 트랙을 나타낼 수 있다.

[0140] 하나의 타일 트랙은, 하나의 유일한 TileRegionGroupEntry(도 8a에 예시됨)를 갖고 TileSetGroupEntry를 갖지 않거나, 또는 이 타일 세트가 만들어진 하나의 유일한 TileSetGroupEntry와 하나 이상의 종속 TileRegionGroupEntry를 가질 것이고, TileSetGroupEntry는 기술된 타일들의 세트에의 TileRegionGroupEntry의 확장이다. 이러한 그룹들 각각이 고유한 식별자를 할당받고, 그것은 그룹에 NALU를 연관시키는 데 사용될 수 있다는 것에 주목해야 한다. 타일 영역들과 타일 세트들은 'tbas' 트랙 참조에 의해 표시되는 바와 같이, 베이스 HEVC 계층에 의해 범위 지정된 groupID에 대해 동일한 네임 스페이스를 공유한다(즉, 동일한 베이스 계층을 갖는 임의의 트랙들에서 동일한 groupID를 가진 2개의 타일 영역 또는 타일 세트는 없을 것이다).

[0141] 여기에 새로운 종류의 트랙을 소개하면, 타일 트랙은, 파일 포맷 디자인을 준수하기 위해 대응하는 샘플 엔트리들을 정의하는 것을 의미한다. 실제로, 각 트랙은 그의 기술적 데이터 중에, 의무사항인 SampleDescriptionBox('stds')를 갖는 SampleTableBox('stbl')를 포함해야 한다. 샘플 디스크립션 테이블은 사용된 코딩 타입에 관한 상세 정보와, 트랙 샘플들의 디코딩에 필요한 임의의 초기화 정보를 제공한다. SampleDescriptionBox에 저장된 정보는 트랙 특징이고 임의의 시각적 샘플 엔트리에 대한 추상화 디스크립션을 전문화함으로써 비디오 트랙들에 대해 기술된다. 전형적으로, 임의의 시각적 샘플 엔트리는 샘플들을 처리하는 데 사용하기 위해 압축 포맷 디코더를 제공하는 "coding name" 파라미터를 포함한다. 이 파라미터는 4 문자 코드로서 인코딩된 고유 식별자이어야 한다. 타일 트랙들에 포함된 샘플들을 기술하기 위해서는, 특정 종류의 VisualSampleEntry로 이러한 샘플들을 기술해야 한다. 예를 들어, 코드 'hvt1'에 의해 표현되어, HEVCTileSampleEntry가 도입되어 타일 능력들을 갖는 HEVC 디코더가 타일 트랙의 샘플들을 처리하는 데 필요하다는 것을 나타낸다. 일반적으로, 샘플 디스크립션 테이블에서는, ConfigurationBox가 파라미터로서 포함되어 디코더 구성 정보를 제공한다. HEVC 타일 트랙들의 특정 경우에, 구성 박스를 복제하지 않고 트랙 헤더에서의 트랙 참조 타입 'tbas'에 표시된 타일 베이스 트랙에 기술된 것을 상속받는다. 옵션으로, 프로파일, 티어 및 레벨 정보뿐만 아니라 HEVCTileSampleEntry에 타일 당 평균 비트레이트를 기술하는 파라미터가 설정될 수 있다. 프로파일은 전형적으로 애플리케이션 도메인을 대상으로 하는 특징들에 관하여 표준의 서브세트를 정의한다. 각각의 프로파일은 티어들과 레벨들을 정의한다. 티어들은 중첩된 복잡도 레벨이라고 볼 수 있으며, 각 레벨은 픽셀 수, 슬라이스 수, 타일 수 등과 같은 일부 값들에 대한 제한들을 정의한다. 증가하는 복잡도에 따라 조직되면, 프로파일의 주어진 레벨에서 최상위 티어를 처리할 수 있는 디코더는 동일한 프로파일에서 동일한 레벨 또는 그 이하의 임의의 하위 티어를 지원할 수 있을 것이다. 대역폭에 기반한 적응을 위해 스트리밍 클라이언트에 제공되도록, 이 박스에 타일 당 비트레이트 정보를 저장하는 것이 유용할 수 있다. mp4 박스들 대부분에 대해, HEVCTileSampleEntry 박스는 애플리케이션 특정 요구를 충족시키기 위해 옵션의 추가 박스들로 확장될 수 있다.

[0142] 도 1은 MPEG-4 파일 포맷에 따라 2개의 스케일 가능성 계층을 캡슐화하는 예를 도시한다. 도시된 바와 같이, 각 계층(인헨스먼트 계층 EL 및 베이스 계층 BL)은 자체 트랙에 캡슐화되어, 효율적인 데이터 어드레싱을 가능하게 하고 비디오를 2개의 트랙으로 캡슐화하게 된다.

[0143] 더 정확하게는, 도 1은 다중-계층 HEVC 비트 스트림으로 인코딩된 미디어 데이터 시퀀스에 대한 모든 ISO BMFF 메타데이터를 포함하고 또한 S 샘플들을 포함하는 무비 박스 "moov"(100)를 나타낸다. 동일한 원칙이 상기 페이지의 b) 및 c)에 정의된 바와 같이 무비 프래그먼트들로 또는 세그먼트들로서 캡슐화된 미디어 데이터에 적용될 것이다.

[0144] 단일 "mdat" 박스(101)는 2개의 청크로 조직된 전체 인코딩된 비트-스트림을 포함하는데, 하나는 베이스 계층(102)에 대한 것이고 다른 하나는 인헨스먼트 계층(103)에 대한 것이며, 각 청크는 S 샘플들(104, 105)을 포함한다. 인헨스먼트 계층 EL의 경우, 청크 EL(103)은 S 샘플들에 대한 인코딩된 비트-스트림의 대응하는 부분을 포함한다. 각 샘플은 하나 이상의 NAL 유닛으로 조직된다. 또한, 베이스 계층 청크에서 대응하는 부분을 참조하기 위해 추출기(106)를 포함하는 부분의 시작 부분에 특정 NAL 유닛이 추가된다. 마지막으로, 인헨스먼트 계층 청크는 주어진 "x" 레벨, 예를 들어 픽처 레벨(PPS) 또는 시퀀스 레벨(SPS)에서 파라미터들을 정의하기 위한 상이한 파라미터 세트들("xPS"(107)로 요약됨)을 포함한다.

[0145] "moov" 박스(100)는 2개의 박스 "track"을 포함하는데, 하나는 (베이스 계층 캡슐화로부터 생기는) 베이스 계층 트랙(110)에 전용이고, 다른 하나는 (인헨스먼트 계층 캡슐화로부터 생기는) 인헨스먼트 계층 트랙(130)에 전용이다.

- [0146] 각 계층 트랙은 mdat 박스(101)에 도시된 그 각각의 S 샘플들을 기술한다.
- [0147] 베이스 계층 트랙(110)은 순차적으로 또는 계층적으로 조직되고 비트-스트림의 위에서 언급된 인코딩된 부분의 파라미터들을 정의하는 수 개의 박스를 포함한다. 명료함을 위해, 선택된 박스들만이 도 1에 도시되어 있다.
- [0148] 트랙 헤더(111)에 대한 'tkhd'로 명명된 박스 또는 하위 박스는 시간적, 공간적 및 식별 정보를 포함한다. 시간적 정보는 S 샘플들(creation_time, modification_time)의 작성 및 수정 시간에 관련된다. 여기서, 식별자(track_ID)는 "BL"과 동일하며, 트랙을 식별할 수 있게 한다. 공간적 정보는 베이스 계층의 디스플레이 사이즈 정보(폭 및 높이)를 포함한다.
- [0149] 'mdia'(112)로 명명된 또 다른 박스 또는 하위 박스는 미디어 정보 디스크립션 박스이고, 비트-스트림의 S 샘플들에 관련된 미디어 데이터에 관한 정보를 포함한다.
- [0150] 'mdia' 박스는 명료함을 위해 표현되지 않은 일부 헤더 박스들과 기술적 정보 자체를 포함하는 미디어 정보 박스 'minf'(113)를 포함한다. 이 예에서, 'minf' 박스는 3개의 상이한 박스 또는 하위 박스로 세분된다.
- [0151] 제1 박스 또는 하위 박스 'oinf'(114)는 동작 포인트를 구성하는 계층들 및 하위 계층들(예를 들어, 시간적 하위 계층들) 및 이들의 조직, 이들 사이의 종속성(존재한다면), 동작 포인트에 대한 HEVC 비트-스트림의 VPS에 포함되는 프로파일, 티어 및 레벨 정보를 반영하는 prof_tier_level 파라미터와 같은 동작 포인트 정보(동작 포인트 디스크립터로도 표시됨)를 포함한다. 더 상세하게는, 박스 'oinf'는 스케일 가능성 구조에 관한 정보를 주는 파라미터(scala_mask), 계층들의 수, 여기서는 2개(max_layer_count=2) 및 각 계층마다, 식별자, 파일에서 동작 포인트들의 수가 후속되는 종속 계층들의 리스트를 포함하고, 각각은 프로파일/티어 및 레벨 정보와 이 동작 포인트를 구성하는 계층들의 세트에 의해 기술되는데, 이는 도 8에서 참조 번호 800으로 도시된 바와 같다.
- [0152] 도 8은 동작 포인트 정보(800)의 표준 동작 포인트 디스크립터를 도시한다.
- [0153] 도시된 바와 같이, 미디어 데이터(801)는 기술적 메타데이터(804)를 포함하는 베이스 계층 트랙(802)에서 인코딩된 베이스 계층을 사용하여 그리고 기술적 메타데이터(805)를 또한 포함하는 인헨스먼트 계층 트랙(803)에서 인코딩된 인헨스먼트 계층을 사용하여 표현될 수 있다.
- [0154] 베이스 계층 트랙 메타데이터(804)는 미디어 데이터(801)에 대한 데이터 참조들을 포함한다. 인헨스먼트 계층 트랙 메타데이터(805)는 예를 들어, 일부 코딩 종속성들을 시그널링하기 위해, 미디어 데이터(801)에 대한 데이터 참조들 및 베이스 계층 트랙에 대한 트랙 참조들을 포함한다.
- [0155] 동작 포인트 정보 'oinf'(800)는 트랙 조직, 특히 스케일러블 HEVC 또는 멀티뷰 HEVC와 같은 다중-계층 비트-스트림의 독립적으로 디코딩 가능한 서브세트인 동작 포인트들의 조직에 관련된 정보를 포함한다.
- [0156] 더 정확하게는, 동작 포인트 정보(800)는 VPS(Video Parameter Set) 및 비디오 비트스트림의 VPS 확장으로부터의 프로파일, 티어 및 레벨 정보(851), 동작 포인트 디스크립터(852), 및 855로 표시된 VPS 확장으로부터의 계층간 종속성들 및 VPS 확장으로부터의 스케일 가능성 차원 식별자를 포함한다. 동작 포인트 디스크립터(852)는 VPS 확장으로부터의 표현 포맷 정보(853)뿐만 아니라, 854로 표시된 VPS VUI(Video Usability Information)로부터의 비트 및 프레임 레이트 정보를 포함한다.
- [0157] 샘플 테이블 박스(Sample Table Box)에 대한 'stbl' 박스(115)라고 명명되는 제2 박스 또는 하위 박스는 각 트랙에 대한 샘플들을 기술하는 정보를 포함한다. HEVC(High Efficiency Video Coding) 방법과 관련된 정보의 일부는 샘플 디스크립션 박스(Sample Description Box)의 'stsd' 박스(116) 또는 하위 박스에 포함된다. 파라미터 "entry_count"는 단 하나의(시각적) 샘플 엔트리만이 포함된다는 것을 나타낸다. 4 바이트 'hvc1'은 고려된 미디어 데이터에 대응하는 비트-스트림이 'hvcC' 박스(117)에서 아래에 정의된 HEVCDecoderConfigurationRecord에 주어진 (프로파일, 티어, 및 레벨을 포함하는) 구성하에서 동작하는 HEVC 디코더에 의해 호환 가능하고 사용 가능하다는 것을 나타낸다. 이 예에서, 버전 구성은 제1 구성이다(configVersion=1). HEVCDecoderConfigurationRecord는 또한 HEVC 비트-스트림의 비디오 파라미터 세트에 포함되는 프로파일, 티어 및 레벨 정보를 제공한다.
- [0158] 'tcon'(118)으로 명명되는 제3 박스 또는 하위 박스는 트랙에 반송(carry)되는 모든 계층들 및 하위 계층을 열거하고, 여기에서는 단 하나(num_layers=1)이다. 예약된 파라미터는 항상 박스의 추가적인 진화를 위해 0 값을 갖는다.

- [0159] 일 실시예에 따르면, 'tcon' 박스는 옵션이고, 이 옵션 특성은 도 1에서 이탤릭체 타입의 사용에 의해 시그널링된다. 예를 들어, 하나의 인헨스먼트 계층만이 있는 경우, 'tcon'(118)이 파일에 존재하지 않는다.
- [0160] 계층들의 조직에 관련된 다른 정보는 상이한 종류의 샘플 엔트리들: 'shv1', 'she1', 'shvC' 및 스케일러블 HEVC에 대한 임의의 4 문자 코드뿐만 아니라 'mhv1', 'mhe1', 'mhvC' 및 멀티뷰 HEVC 샘플 엔트리들에만 전용인 임의의 4 문자 코드와 같이 옵션일 수 있다. 이러한 코드들은 식별된 단일 스케일 가능성 타입을 갖는 트랙들에 대응한다. 둘 이상의 단일 스케일 가능성 타입(예를 들어, SNR 및 뷰 차원들 양쪽 모두)을 갖는 트랙들의 경우 또는 비-출력 계층들인 트랙들의 경우 단지 하나 또는 두 종류의 샘플 엔트리만이 유지될 수 있다: 예를 들어 'lhv1', 'lhvC' 또는 'lhe1' 또는 계층화된 HEVC 샘플 엔트리들을 기술하기 위한 임의의 4 문자 코드.
- [0161] 다른 트랙은 인헨스먼트 계층(130)에 전용된다. 그것은 트랙 헤더 박스 'tkhd'(131) 또는 베이스 계층 트랙의 'tkhd' 박스와 유사한 하위 박스를 포함하며, track_ID는 인헨스먼트 계층에 대해 "EL"로 예상된다.
- [0162] 인헨스먼트 계층을 위한 트랙은 트랙 참조 박스(Track Reference Box) 'tref'(132) 또는 하위 박스를 포함한다. 그것은 포함 트랙(containing track) 여기서는 인헨스먼트 계층 트랙으로부터 또 다른 트랙, 여기서는 베이스 계층 트랙으로의 참조를, 프레젠테이션에서 제공한다.
- [0163] 제1 참조 'sbas'는 베이스 계층(110)의 트랙이 인헨스먼트 트랙(130)에 대한 베이스 트랙임을 나타낸다. (track_ids[]=BL).
- [0164] 또 다른 참조 'oref'는, 여기서는 베이스 계층 트랙에 있는, 'oinf' 박스(동작 포인트 디스크립터)의 참조를 가능하게 한다. 'oref' 참조는 이탤릭체 타입을 사용하여 표현된다. 실제로 아래에 설명된 바와 같이, 'oinf' 박스가 베이스 계층 트랙에 존재하지 않는 경우, 하나의 참조 계층만 있는 경우 'oref' 트랙 참조 타입은 옵션일 수 있다.
- [0165] 베이스 계층 트랙에 관해서, 인헨스먼트 계층 트랙은 'minf' 박스(134)를 포함하는 'mdia 박스'(133)를 포함한다. 이 'minf' 박스는 'stsd' 박스를 포함하는 'stbl' 박스(135) 자체를 포함한다. 이 마지막 박스는 예를 들어 4 바이트 'lhe1'을 포함하고, 이는 고려된 미디어 데이터에 대응하는 비트-스트림이 'lhvC' 박스(137) 내에 아래에 정의된 구성 박스에 주어진 (프로파일, 티어, 및 레벨을 포함하는) 구성 하에서 동작하는 L-HEVC 디코더에 의해 호환가능하고 이용 가능하다는 것을 나타낸다.
- [0166] 이 'lhvC' 박스는 아래에서 더 상세히 설명된다.
- [0167] 마지막으로 베이스 계층 트랙에 관해서, 'mdia' 박스는 옵션인 'tcon' 박스(138)를 포함한다.
- [0168] 전술한 실시예에 따르면, 미리 정의된 조건(예를 들어, 그 자신의 트랙에, 단지 하나의 인헨스먼트 계층)에 기초하여, 'tcon' 박스는 트랙으로부터 제거될 수 있다. 이것은 다수의 계층이 존재하는 경우일 수 있다(이들이 그들 자신의 트랙들에 캡슐화되면).
- [0169] 트랙 당 하나의 계층을 갖는 것은 수 개의 이점들로 이어진다는 점에 주목해야 한다. 우선, 그것은 계층화된 콘텐츠의 필터링을 단순화하는데, 그 이유는 필터링 동안 트랙이 사용되거나 사용되지 않기 때문이고; 트랙 당 다중-계층은 트랙이 원하는 OP에 의해 사용되는 계층들 및 그 OP에 의해 사용되지 않는 계층들의 혼합을 포함하는 경우에 샘플 검사를 부과한다.
- [0170] 또한, 그러한 경우에는 트랙에 단일 스케일 가능성 차원이 존재한다. 이는 트랙의 타입을 나타내기 위해 단순화된 코드-포인트들(shc1/...)을 사용하는 것을 가능하게 하고, 스케일 가능성의 타입을 식별하기 위해 프로파일 및 레벨들에 의존하지 않고, mime 타입 도출을 단순화한다.
- [0171] 마지막으로, 트랙 당 하나의 계층을 갖는 것은 L-HEVC의 MPEG-2 TS 캐리지와의 정렬을 개선하는 것을 가능하게 하고, 여기서 각 타입의 스케일러블 콘텐츠에 대해 스트림 타입들이 정의되고 트랙에서 단일 계층이 허용된다.
- [0172] 그러나, 트랙 당 단일 계층을 갖는 것은 수 개의 단점이 있다.
- [0173] 우선, 비-출력 계층들은 그들 자신의 트랙에 있어야 하기 때문에, 이들은 'stbl' / 'trun' 박스가 중복될 것이므로 스토리지를 낭비한다. 그러나, 그러한 경우(계층이 결코 출력 계층이 아닌 경우)는 매우 제한된다고 생각될 수 있다.
- [0174] 또한 트랙 당 단일 계층을 갖는 것은 베이스 계층이 하나의 트랙에 저장되고 모든 인헨스먼트 계층이 또 다른 트랙에 저장되는 경우 또는 모든 베이스/lhevc가 하나의 트랙에 저장되는 경우들을 허용하지 않는다는 점에 주

목해야 한다. 이것이 트랙 당 하나의 계층이 도 11의 참조 번호 1100로 기술된 바와 같은 LHEVCDecoderConfigurationRecord의 수정 또는 도 9를 참조하여 기술된 바와 같은 동적 'tcon' 박스의 사용으로 이어지는 대안적인 실시예들이 고려되는 이유이다.

- [0175] 'tcon' 박스가 제거되는 경우에, 계층 구조 정보는 다음과 같이 LHEVCDecoderConfigurationRecord(도 11의 참조 번호 1100)에 저장된다: 도 11에서 1101로 표시된 트랙에 의해 전달되는 layerID의 표시. 최소 및 최대 시간적 하위 계층은 'oinf'에 의해 일부 계층들을 선택하는 데 사용될 수 없으므로 제거된다. (단일 계층) HEVC에 대해 정의된 시간적 그룹화 메커니즘(예를 들어 'tscl' 그룹화 타입)을 사용하는 것이 여전히 가능하다는 점에 주목해야 한다.
- [0176] 더 일반적으로, 'tcon' 박스는 각 계층이 베이스 계층을 의미하고 수 개의 인헨스먼트 계층들 각각이 별개의 트랙으로 캡슐화되어 있는 경우 옵션이다. 실제로, 하나의 계층이 하나의 동작 포인트에 대응하는 전형적인 구성의 경우, 이 박스는 유용한 정보를 제공하지 않는다: 'tcon' 박스는 단지 트랙에 하나의 계층이 존재함을 나타낸다. 종속 계층들(즉, 이 경우 트랙들)의 리스트는 트랙 참조 박스(Track Reference Box)를 통해 발견될 수 있다. 'tcon' 박스는 하나의 트랙이 수 개의 계층의 캡슐화로부터 생기는 경우에만 유용할 수 있다.
- [0177] 바람직한 실시예에 따르면, 공통의 'sbas' 트랙 참조를 갖는 트랙들의 세트에 대해, 이 세트 중에는 'oinf' 박스의 정보를 운반하는 최대 하나의 트랙이 존재한다. 공통의 'sbas' 참조되는 트랙을 갖는 모든 트랙들은 또한 'oref' 타입 트랙 참조를 사용하여 'oinf' 박스의 정보를 전달하는 트랙에 링크될 수 있다.
- [0178] 여기에는 두 가지 옵션이 존재한다: 'oref' 타입을 버리고 'sbas' 트랙 참조로 나타낸 바와 같이, 'oinf' 정보가 베이스 트랙에 있도록 명령하는 것(하이브리드 코딩(AVC 베이스 트랙 + HEVC 인헨스먼트 트랙)의 경우, 이는 샘플 그룹 'oinf'가 AVC 트랙에서 허용될 것임을 의미한다); 또는, 가능하지 않은 경우, 'oref'를 유지하여 어떤 트랙이 'oinf' 정보를 나타내는 샘플 그룹 및 샘플 그룹 디스크립션 박스를 포함하는지를 시그널링하는 것.
- [0179] 도 11은 본 발명에 따른 새로운 디코더 구성 정보를 도시한다. 도시된 바와 같이, 계층 구조 정보는 1101로 표시된, 트랙에 의해 전달된 layerID의 표시를 포함하는 LHEVCDecoderConfigurationRecord(1100)에 저장된다.
- [0180] ISOBMFF에서 시간이 지남에 따라 정보가 변하도록 허용하는 것은 전용 트랙, 샘플 엔트리들, 또는 샘플 그룹들을 사용하여 행해질 수 있고, 계층 구성은 현재 'oinf' 및 옵션으로 'tcon' 박스들에 의해 제공된다는 점에 주목한다.
- [0181] 실시예들에 따르면, 이들 박스는 'minf' 박스 밖으로 이동되고, 도 7에 도시된 바와 같이 VisualSampleGroupEntry 클래스로부터 데이터를 상속받는다.
- [0182] 도 7은 본 발명의 실시예들에 따른 동작 포인트에 대한 특정하고 동적으로 확장 가능한 디스크립터를 도시한다.
- [0183] 도 7에 도시된 바와 같이, 동작 포인트 정보는 추상적인 VisualSampleGroupEntry를 전문화함으로써 새로운 OperatingPointGroupEntry(700)를 사용하여 획득될 수 있다.
- [0184] 그 실시예에 따르면, 'oref' 타입 참조되는 트랙은 디폴트 샘플 그룹화의 경우 샘플 그룹 박스(703)와 샘플 그룹 디스크립션 박스들(702) 또는 샘플 그룹 디스크립션 박스(701)만을 포함할 것이다. 샘플 그룹 디스크립션 박스(701)는 임의의 샘플 그룹으로 매핑되지 않은 트랙들의 샘플들이 인덱스 1(default_sample_description_index = 1)을 갖는 샘플 그룹의 샘플 속성들을 상속받는 경우를 도시한다.
- [0185] OperationPointGroupEntry(예를 들어, OperationPointGroupEntry(705))는, 샘플 그룹 박스(703)에 정의된 대응하는 샘플들에 대해, 표준 'oinf' 구조(예를 들어, 도 8을 참조하여 기술된 표준 'oinf' 구조(800))에 주어진 정의에 따라, 사용될 정보의 세트를 제공한다.
- [0186] 샘플 대 그룹 박스(sample to group box)(703)가 'oinf'라 불리는 grouping_type을 통해 샘플 그룹 디스크립션 박스(702)와 연관되는 도 7에 도시된 예(여기서 이 코드는 단지 예시를 위해 주어진 것일 뿐이고, 동작 포인트 그룹 엔트리가 사용될 수 있음을 나타내기 위해 임의의 4 문자 코드가 예약된다는 점에 주목해야 한다)는 도 6(참조 번호 603)을 참조하여 기술된 바와 같은 상이한 동작 포인트들을 갖는 스트림 스플라이싱으로부터 생기는 2개의 샘플 그룹을 도시한다.
- [0187] 당연히, 2개의 샘플 그룹 디스크립션 엔트리 중 하나는 디폴트로 정의된 샘플 그룹 디스크립션이 될 수 있어 하나의 특정 샘플 그룹 디스크립션 엔트리만이 정의될 것이다. 즉, 첫 번째 N1 샘플들을 특정 샘플 그룹 디스크

립션과 연관시키고 다른 샘플들(즉, 'oinf' 그룹에 매핑되지 않은 샘플들)을 디폴트 샘플 그룹 디스크립션에 연관시킬 수 있다.

[0188] 마찬가지로, 트랙 내부의 계층 조직은 시간에 따라 달라질 수 있으므로, 동적 업데이트가 필요한 경우 'tcon' 박스가 유사하게 수정될 수 있다(즉, 'minf' 박스로부터 새로운 grouping_type 및 새로운 VisualSampleGroupEntry로 이동).

[0189] 도 9에 도시된 바와 같이, 새로운 LayerInfoGroupEntry 'tcon'(900)이 그 목적으로 사용될 수 있다. 다시, 4 문자 코드 'tcon'은 단지 예시를 위해 주어지며, 특정 엔트리의 이름(TrackContentGroupEntry일 수도 있음)

[0190] LayerInfoGroupEntry 'tconf'는 다음과 같이 정의할 수 있다:

```
class LayerInfoGroupEntry extends VisualSampleGroupEntry ('tcon') {
    unsigned int (2) reserved;
    unsigned int (6) num_layers_in_track;
    for (i=0; i<num_layers_in_track; i++) {
        unsigned int (4) reserved;
        unsigned int (6) layer_id;
        unsigned int (3) min_sub_layer_id;
        unsigned int (3) max_sub_layer_id;
    }
```

[0191]

[0192] 도 9는 계층 정보 디스크립션을 위한 특정하고 동적으로 확장 가능한 디스크립터를 도시한다.

[0193] 다시, 'oref' 타입 참조되는 트랙은 디폴트 샘플 그룹화의 경우 샘플 그룹 박스(903)와 샘플 그룹 디스크립션 박스들(902) 또는 샘플 그룹 디스크립션 박스들(901)을 포함할 것이다.

[0194] 일정한 계층 조직을 갖는 트랙의 경우, 디폴트 샘플 그룹화가 사용될 수 있다. 도 6을 참조하여 기술된 바와 같이(예를 들어, 비트-스트림 스플라이싱(603)에서), 계층 조직이 시간에 따라 변하는 경우, 샘플 그룹 박스는 조직을 추적하기 위해 샘플들을 그들의 적절한 계층에 매핑하도록 정의될 수 있다. 예시를 위해, 도 6의 스트림(603)에서 스트림(602)의 샘플들과 결합되는 스트림(601)의 샘플들에 대응할 수 있는 그룹(903)에서 정의된 바와 같은 첫 번째 N1 샘플들, 및 도 6의 스트림(603)에서 스트림(601)의 샘플들과 결합되는 스트림(602)의 샘플들에 대응할 수 있는 나머지 N2 샘플들은 샘플 그룹 디스크립션 박스(902) 내의 상이한 계층 정보 그룹 엔트리들과 연관될 수 있다.

[0195] 다시, 2개의 샘플 그룹 디스크립션 엔트리 중 하나는 디폴트로 정의된 샘플 그룹 디스크립션이 될 수 있어 하나의 특정 샘플 그룹 디스크립션 엔트리만이 정의될 것이다. 즉, 첫 번째 N1 샘플들을 특정 샘플 그룹 디스크립션(특정 계층 정보 그룹 엔트리)과 연관시키고 다른 샘플들(즉, 샘플 그룹에 매핑되지 않은 샘플은 'tcon' grouping_type)을 디폴트 샘플 그룹 디스크립션(또 다른 계층 정보 그룹 엔트리)에 연관시킬 수 있다.

[0196] 도 10은 본 발명에 따라 캡슐화된 미디어 파일에 대한 클라이언트 측의 처리를 도시한다.

[0197] 도시된 바와 같이, 클라이언트 디바이스는 제1 단계(단계 1000)에서 mp4(또는 ISOBMF) 파일 또는 mp4(또는 ISOBMF) 세그먼트를 수신한다. 그것은 mp4 파서(또는 mp4 리더)에 의해 해석되거나 파싱(즉, 판독 또는 디코딩)된다. 세그먼트 기반 전달의 경우, 수신은 미디어 프리젠테이션의 제1 초기화 세그먼트로부터 최종 세그먼트까지의 반복 프로세스이다. mp4 파일의 경우, 파일 수신시 모든 데이터가 사용 가능할 수 있다.

[0198] 후속 단계(1001)는 mp4 파서가 동작 포인트 정보를 찾는 것으로 구성된다. 실제로, 하나 이상의 L-HEVC 트랙의 샘플들로부터 액세스 유닛을 재구성하기 위해, 타겟 출력 계층들 및 이들이 속하는 동작 포인트가 먼저 결정될 필요가 있을 수 있다. 이것은 'moov' 박스(또는 mp4 세그먼트들의 경우 'moof' 박스) 아래의 기술적 메타데이터를 파싱하는 것에 의해 수행될 수 있다.

[0199] 바람직한 실시예들에 따르면, 파서는 마주치는 모든 'trak' 박스들로부터 메모리 내의 트랙들의 내부 리스트를 구축한다. 이 트랙들로부터, 적어도 하나의 트랙이 'oref' 트랙으로서 참조되면, 이것은 파일(또는 세그먼트들의 리스트)에서 사용 가능한 동작 포인트들에 대한 정보를 포함해야 하는 것이다. 그 후 이 트랙을 기술하는

메타데이터 박스들, 특히 샘플 테이블 박스를 파싱하여 동작 포인트 정보를 나타내는 그룹화 타입을 갖는 일부 샘플 대 그룹 박스를 찾는다.

- [0200] 이는 예를 들어 바람직한 실시예에 따른 'oinf' 코드일 수 있다(예를 들어, 도 7을 참조하여 설명된 디폴트 샘플 그룹화의 경우 샘플 그룹 박스(703)와 샘플 그룹 디스크립션 박스들(702) 또는 샘플 그룹 디스크립션 박스(701). 그것은 임의의 예약된 고유한(즉, 다른 목적으로 이미 사용되지 않은) 4 문자 코드일 수 있다.
- [0201] 그러한 샘플 그룹(도 7의 참조 번호 703 및 702) 또는 적어도 디폴트 샘플 그룹(도 7의 참조 번호 701)이 발견되면(즉, 테스트 1001이 참), 트랙 선택은 동작 포인트 파라미터들에 기초한다(단계 1002).
- [0202] 반대로, 테스트 1001이 거짓이면, 디폴트 트랙이 선택된다(단계 1003). 그것은, 예를 들어, 파일에서 첫 번째 트랙 또는 트랙 헤더 박스에서 가장 낮은 "계층" 값을 가진 트랙일 수 있다. 대안적으로, 모든 트랙이 사용자에게 프레젠테이션되고 플레이어는 사용자 선택을 기다린다.
- [0203] 선택된 트랙은 단계 1004에서 그의 트랙 참조 박스('tref')로부터 발견될 수 있는 잠재적으로 코딩 종속성들만을 갖는 단순한 트랙인 것으로 가정된다.
- [0204] 동작 포인트 정보로부터 재생할 하나 또는 수 개의 트랙이 선택되는 경우(단계 1002), 동작 포인트 정보 그룹 엔트리로부터 계층들의 리스트가 식별될 수 있다. 그 후 클라이언트 디바이스에 대한 다음 단계는 식별된 계층들이 캡슐화되는 트랙들을 식별하는 것이다(단계 1005). 이것은
- [0205] - w15182에 기술된 계층화된 HEVC 파일 포맷의 현재 버전에 따른 트랙 콘텐츠 정보 박스 'tcon'을 통해,
- [0206] - 샘플 그룹 디스크립션을 통해 도 9를 참조하여 기술된 실시예에 따라, 'tcon' 값(또는 샘플들에 대한 계층 정보 또는 트랙의 샘플들의 그룹을 나타내는 임의의 예약된 고유한 4 문자 코드)을 갖는 그룹화 타입을 찾는 것을 통해, 또는
- [0207] - 하나의 트랙이 하나의 단일 계층에 대응하는 또 다른 실시예에 따라, 계층이 트랙에 대응하는지 여부를 식별하기 위해 트랙들에 대한 'lhvC' 박스 내의 수정된 LHEVCDecoderConfigurationRecord(도 11의 참조 번호 1100)로부터 정보를 파싱하는 것(단계 1101)에 의해 트랙 디스크립션 박스들을 파싱하는 것에 의해 행해질 수 있다. 도 11에서 1101로 표시된 새로운 layer_id 파라미터는 이 샘플 엔트리와 연관된 샘플들에 포함된 VCL NAL 유닛들의 nuh_layer_id의 값을 나타낸다. 그러한 경우 트랙은 추출기들과 집계기들만을 포함하기 때문에, 도 11에서 1102로 표시된 플래그 complete_representation이 0으로 설정되는 경우 이 값은 0일 것이다.
- [0208] 계층들로부터 트랙으로의 매핑은 단계 1005 동안 수행되고, 파서는 (단계 1002 동안) 동작 포인트의 계층들이 종속하는 종속 계층의 리스트를 획득한 것이 관찰된다(도 8의 디스크립터(800)에서 참조 번호 855에 의해 기술된 바와 같이).
- [0209] 동작 포인트들에 대해 획득된 트랙들 및 종속 계층들에 대응하는 종속 트랙들은 단계 1004 동안 상호 종속 트랙들의 리스트를 형성하는 데 사용된다. 이러한 종속성들의 특성을 알기 위해, 파서는 단계 1006에서 획득된 트랙들의 트랙 참조 박스들을 검사한다.
- [0210] 단계 1007에서는 베이스 트랙이라고 불리는 특정 트랙이 발견되어야 한다.
- [0211] 이 특정 트랙은 액세스 유닛 재구성을 시작할 트랙이다. 아무 것도 발견되지 않으면, 이것은 에러이고 플레이어는 무효 미디어 파일을 나타내는 에러 메시지와 함께 종료한다(단계 1008).
- [0212] 'sbas' 트랙 참조 타입을 통해 참조된 베이스 트랙이 발견될 수 있는 경우, 파서는 사용자에게 의해 주어진 시간에 이것으로부터 샘플 재구성을 시작한다(첫 번째 샘플을 식별하기 위해 샘플 테이블 박스에서 일부 탐색이 필요할 수 있음).
- [0213] 먼저, 이 첫 번째 샘플이 추출기들을 포함하는지 여부를 결정하기 위한 테스트가 수행된다(단계 1009). 추출기들이 존재하는 경우, 이것은 발견된 베이스 트랙으로부터 샘플 데이터를 복사할 참조되는 트랙까지 'scal' 트랙 참조 타입으로 표시되어야 하고, 샘플의 재구성은 명시적이고, 단순히 추출기들을 분석하는 것(resolving), 즉 추출기 NAL 유닛을 추출기가 지시하는 NAL 유닛 데이터로 대체하는 것에 있다.
- [0214] 이것은 단계 1010에서 트랙의 모든 샘플에 대해 행해져서 단계 1011에서 트랙을 나타내는 NAL 유닛들의 리스트를 생성하고 마지막으로 단계 1012에서 클라이언트 디바이스에 의해 디코딩되고 디스플레이 될 수 있는 비디오 비트-스트림을 생성한다.

- [0215] 예시를 위해, 이는 도 4a에 나타난 타일 비트-스트림의 베이스 계층(400)에 적용될 수 있다.
- [0216] 반대로, 단계 1009에서 추출기가 발견되지 않으면, 파서는 단계 1013에서 베이스 트랙과 연관된 트랙 참조 리스트에 (예를 들어 나중에 도 4b를 참조하여 기술된 'sabt'와 같은) 일부 특정한 순서화된 트랙 참조들이 있는지 여부를 체크한다.
- [0217] 단계 1013에서 특정한 순서화된 트랙 참조가 식별되면, 단계 1014에서 디코딩 순서 종속 트랙들의 순서화된 리스트가 구축된다.
- [0218] 반대로, 단계 1013에서 특정한 순서화된 트랙 참조가 식별되지 않으면, ISO/IEC 23008-2(HEVC 표준)에 따르는 순서로 그들의 NAL 유닛들을 배열함으로써, 동작 포인트 정보 박스(Operation Point Information Box)에 의해 지시된 바와 같이, 요구된 트랙들 내의 각각의 샘플들로부터 액세스 유닛이 재구성된다. 이것은 단계 1015의 암시적 액세스 유닛 재구성이다. 옵션으로, 암시적 또는 명시적 재구성에 대한 이러한 지시는 도 11에 1100으로 표시된 LHEVCDecoderConfigurationRecord의 예약된 비트 중 하나에, 1101로 표시된 layer_id 바로 앞에 놓일 수 있다.
- [0219] 대안으로서, 추출기들이, 명시적 재구성의 경우, 비디오 디코더에 해상도 없이 제공되지 않고 mp4 파서에 의해 분석될 것임을 보장하기 위해 암시적 재구성을 갖는 트랙들에 대해 그리고 명시적 재구성을 갖는 트랙들에 대해 특정한 샘플 엔트리들이 정의될 수 있다.
- [0220] 재구성이 의미하는 것이 무엇이든 간에, 파서는 단계 1012에서 디코딩 가능한 비트-스트림을 형성하기 위해 단계 1011에서 동작 포인트들에 관련된 트랙들의 모든 샘플들을 고려하고 그 샘플들(NAL 유닛들)을 연결해야 한다.
- [0221] 선택 단계(1001)는 파일에서 재생할 첫 번째 샘플의 시간에 좌우될 수 있다는 점에 주목해야 한다. 예를 들어, 도 6의 참조 번호 603에 의해 기술된 바와 같이 스플라이싱된 스트림의 기록으로부터 생기는 스트림에서, 샘플의 시간은 'oref' 트랙으로 참조된 트랙의 샘플 그룹 디스크립션 박스(예를 들어, 도 7에서 샘플 그룹 디스크립션 박스(702)) 내의 하나 또는 또 다른 OperationPointInfoGroupEntry로 이어질 것이다.
- [0222] 도 12는 본 발명에 따른 비디오 비트-스트림 캡슐화를 위한 서버 측의 처리를 도시한다.
- [0223] 힌트 트랙들을 추가함으로써 RTP 상의 스트리밍 또는 HTTP 상의 적응 스트리밍을 위해 또는 로컬 저장/기록을 위해 스트림을 준비하기 위해 도 12에서 설명된 알고리즘을 사용하여 수행되는 캡슐화가 수행할 수 있다. 캡슐화는 하나 이상의 mp4 파일 또는 mp4 세그먼트들의 리스트를 생성하기 위해 하나 이상의 인코딩된 미디어 스트림을 입력으로 받는 mp4 기록기에 의해 처리된다.
- [0224] 설명이 비디오 스트림 캡슐화에 관한 것이라면, 동작 포인트의 사용 및 종속성들의 시그널링이 다른 계층화된 미디어 타입들에도 적용된다는 점에 주목해야 한다.
- [0225] mp4 기록기의 제1 단계(단계 1200)는 예를 들어, VPS(VideoParameterSet), VPS_VUI(Video Usability Information) 및 HEVC에서의 VPS 확장과 같은 비디오 비트-스트림의 고레벨 디스크립션 파라미터들을 파싱하는 단계로 구성된다. 비디오 파라미터 세트로부터, mp4 기록기는 복잡도의 정도를 추론할 수 있다(특히 스트림이 종속성들을 갖는 많은 계층들 및 다수의 동작 포인트를 포함하는 경우). 이 계층 수로부터, mp4 기록기는, 그 자신의 트랙에 하나의 계층이 있다고 가정하여, 그것이 생성할 트랙 수를 추론한다.
- [0226] 다중-계층 스트림의 경우(즉, 테스트 1201이 참인 경우), mp4 기록기는 VPS 및 VPS 확장 파라미터들을 도 8을 참조하여 기술된 800으로 참조된 디스크립터와 같은 디스크립터에 배치함으로써 단계 1203 동안 동작 포인트 디스크립터를 채운다.
- [0227] 그 동안에, mp4 기록기는 비디오 비트-스트림의 베이스 계층에, 즉 nuh_layer_id가 0인 NAL 유닛들에 대응하는 베이스 트랙의 구성을 개시한다.
- [0228] 그러한 베이스 트랙 구성은 단일 계층 비트-스트림의 경우(즉, 테스트 1201이 거짓인 경우)에도 수행되어 캡슐화되는 단일 트랙을 생성한다(단계 1202).
- [0229] layerID 값, 예를 들어 도 11에 도시된 LHEVCDecoderConfigurationRecord 정보(1100)의 layer_id(1101)를 설정하고, 'sbas' 참조 트랙을 통해 베이스 트랙을 참조함으로써 다른 트랙들(계층 수에 대응)도 개시된다.
- [0230] 동작 포인트 정보 디스크립터에 관련된 단계 1203은 주로 'trak' 박스들의 계층 구조, 특히 VPS 및 VPS 확장 파

라미터들을 갖는 디폴트 OperatingPointGroupEntry(예를 들어, 도 7의 참조 번호 701)를 갖는 샘플 테이블 박스를 생성하는 것으로 구성된다. 디폴트로, 샘플 대 그룹(Sample To Group) 박스(도 7의 참조 번호 703과 같은)는 생성되지 않는다.

- [0231] VPS, VPS_VUI 및 VPS 확장 파라미터들로부터, mp4 기록기는 avg_bit_rate 파라미터를 통해 단계 1204에서 평균 비트레이트를 인식한다(존재하지 않는다면, mp4 기록기는 평균 비트 레이트를 추론하기 위해 전체 스트림의 사이즈를 샘플 수로 나누어 대략 추정할 수 있다).
- [0232] 그 후 그것은 1205에서 추출기들(명시적 재구성)을 사용할지 여부를 결정한다. 예를 들어, 계층 수와 샘플 수를 알고, 그것은 추출기 당, 샘플 당, 및 종속 트랙 당 12 바이트가 입력 비디오 비트-스트림에 관하여 주어진 백분율보다 많이 파일의 사이즈를 증가시킬지를 추정할 수 있다.
- [0233] 예시를 위해, mp4 기록기는 대응하는 오버헤드가 5%보다 많이 비트-스트림의 사이즈를 증가시키지 않을 때 추출기들을 사용하도록 파라미터화될 수 있다.
- [0234] 캡슐화가 대화식인 경우 이 결정과 임계 값은 사용자에게 의해 결정될 수 있다: mp4 기록기는 오버헤드를 추정하고, 그것을 추출기들을 사용하여 비트-스트림을 캡슐화할지 여부를 결정할 수 있는 사용자에게 제공한다.
- [0235] 오버헤드가 수용 가능하다면(또는 사용자가 추출기들을 사용하기로 선택하는 경우), mp4 기록기 단계 1206에서 추출기 생성 모드를 활성화한다. 이 모드에 따르면, 샘플들을 어떻게 재구성할지를 명시적으로 지시하기 위한 표시가 종속성들을 갖는 각 트랙의 각 샘플에 대해 생성된다.
- [0236] 반대로, 오버헤드가 너무 중요한 경우, mp4 기록기는 암시적 재구성 모드를 고려하고, 이는 샘플들의 재구성이 트랙 참조들로부터만 추론되어야 함을 의미한다.
- [0237] 이를 위해, mp4 기록기는 단계 1207에서, 계층 종속성들이 NAL 유닛들의 올바른 순서를 보증하기에 충분한지 여부 또는 모호성이 남아 있는지(예를 들어, 개별 트랙들에서 타일들을 캡슐화할 때)를 체크한다. 순서가 표시되어야 하는 경우, mp4 기록기는 관련 트랙들(예를 들어, 도 4b에서 도시된 바와 같은, 타일들 및 타일 베이스 트랙들)을 연결 순서(예를 들어, 'sabt' 것과 같은)를 암시하는 특정 트랙 참조 타입과 링크한다.
- [0238] 반대로, 순서가 지정되지 않아야 하는 경우(즉, 테스트 1207이 거짓인 경우), mp4 기록기는 다음의 재구성 규칙에 의존하는 임의의 특정 트랙 참조를 지정하지 않는다: 주어진 시간 동안, 증가하는 nuh_layer_id로 NAL 유닛들을 순서화한다.
- [0239] 재구성 방법이 결정되면, mp4 기록기는 NAL 유닛들을 판독하기 시작하고 샘플 테이블 박스에 샘플 정보를 표시한다(예를 들어, 타이밍, 데이터에서의 바이트 오프셋 등).
- [0240] 추출기들이 사용되는 경우 그리고 트랙이 또 다른 것에 종속하는 경우, 이 트랙의 샘플은 추출기들을 사용하여 참조되는 트랙 내의 데이터를 참조한다('scal' 트랙 참조 타입을 통해).
- [0241] mp4 기록기는 모든 샘플이 처리될 때까지(즉, 테스트 1212가 거짓) 샘플 디스크립션 및 캡슐화 프로세스를 수행한다.
- [0242] 하나의 샘플에서 또 다른 샘플로, mp4 기록기가 새로운 VPS, 새로운 VPS 확장자, 또는 계층 구성의 변화(예를 들어, 가장 큰 nuh_layer_id의 변화)을 검출하면, 단계 1213에서, mp4 기록기는 새로운 베이스 트랙이 존재하는지 여부를 체크한다(테스트 1214). 이는 동일한 파일에 다수의 미디어 스트림을 캡슐화할 때의 경우일 수 있다: mp4 파일에 트랙들의 세트가 이미 존재한다, 특히 layer_id가 0인 것과 새로운 비디오 비트-스트림이 캡슐화되어야 한다(예를 들어, nuh_layer_id 값이 0인 NAL 유닛들을 또한 갖는 대안적인 독립 비디오 비트-스트림).
- [0243] 새로운 베이스 트랙은 디폴트 샘플 그룹 디스크립션 박스(예를 들어, 도 7에서 디폴트 샘플 그룹 디스크립션 박스(701))와 함께 1216 동안 새로운 동작 포인트 디스크립터의 생성으로 이어진다. 구성이 변화하지만 새로운 베이스 트랙이 없는 경우(예를 들어, 테스트 1214가 거짓인 경우), 베이스 트랙 샘플 그룹들을 1215 동안 수정하여 새로운 OperationPointGroupEntry(예를 들어, 도 7의 702에서 하나의 OperationPointGroupEntry)를 생성하고 그룹화 타입 'oinf'(또는 예약된 4 문자 코드)를 갖는 연관된 샘플 대 그룹 박스(703)를 업데이트한다(또는 존재하지 않으면 생성한다). 이것은 마지막 샘플에 도달할 때까지 반복적으로 반복된다.
- [0244] 결국, mp4 기록기는 트랙 디스크립션과 그 종속성들을 갖는 mp4 파일을 생성한다.

- [0245] 서버 측에서 수행되는 이 처리는 또한, 예를 들어 mp4 세그먼트들로서, 스트리밍된 미디어 파일을 기록하기 위해 수신기 측에서도 수행될 수 있다. 서버는 일부 mp4 세그먼트들('moof' 박스)의 메타데이터 파트에서 계층 구성의 디스크립션을 준비했을 수 있고, 따라서 하나의 스트림에서 또 다른 스트림으로 스위칭할 때, 수신기가 현재 계층 구성을 통보받고, 단계 1213을 참조하여 기술된 바와 같이, 새로운 샘플 대 그룹 및 그룹화 타입 'oinf'(또는 전용 4 문자 코드)의 연관된 디스크립션을 추가하거나 생성함으로써 그 동작 포인트 디스크립터를 업데이트할 수 있게 된다.
- [0246] 도 7을 참조하여 기술된 것의 대안인 실시예에 따르면, 동작 포인트 정보 디스크립터(예를 들어, 도 8의 동작 포인트 정보 디스크립터(800))로부터 정보를 전달하는 샘플들을 갖는 특정 트랙이 할당된다. 그러면 트랙 샘플들의 대부분은 비어 있을 것이다. 이 트랙은, 예를 들어 Layered-HEVC 파라미터 세트 스트림에 대한 'lhvp'와 같이, 특정 샘플 엔트리들로 선언된다. 계층화된 HEVC 스트림 샘플 엔트리는 다음과 같은 LHEVC 파라미터 스트림 구성을 포함한다:
- ```
class LHEVCParameterSampleEntry() extends VisualSampleEntry('lhvp'){
 LHEVCConfigurationBox config;
 Box extra_boxes[];
}
```
- [0247]
- [0248] 이 트랙은 특정 'oref' 트랙 참조 타입으로 동일한 동작 포인트들을 사용하는 다른 트랙들에 의해 참조될 것이다. 'sbas' 트랙 참조 타입을 통해 베이스 트랙을 공유하는 트랙의 세트에 대해 하나의 그러한 트랙들만이 있을 것이다.
- [0249] 도 7을 참조하여 기술된 것의 또 다른 대안적인 실시예는 'minf' 박스에 정적인 'oinf'(예를 들어, 도 8의 'oinf'(800))를 유지하지만, 예를 들어, 새로운 계층들이 미디어 프레젠테이션에 추가될 때 일부 파라미터들을 변경하는 것을 가능하게 하는 것이다. 이것은 'oinf' 박스의 이 확장 가능 버전에 상이한 것들을 추가하기 위해 프로파일 구성 비교를 위한 추가적인 수단을 mp4 파서와 기록기에 요구한다.
- [0250] 이를 위해, mp4 기록기는 동작 포인트들을 비교하는 특정 모듈을 필요로 한다. VPS로부터 새로운 동작 포인트가 검출될 때, mp4 기록기는, 파라미터마다, 새로운 동작 포인트를 정적인 'oinf' 박스에 이미 저장된 것들과 비교한다. 적어도 하나의 파라미터가 다르면, 이 동작 포인트에 대해 oinf에 새로운 엔트리가 생성되고 동작 포인트의 수가 증가된다.
- [0251] 유사하게, 계층 정보는 업데이트되어 ID 충돌을 피하기 위해 새로운 비트스트림에 있는 것들에 새로운 계층 식별자를 할당한다. 그 후, 최대 계층 수는 새로운 스트림에 존재하는 계층 수로 업데이트된다. 이 새로운 계층들 각각에 대해, 그들의 종속성들이 oinf 박스의 855 디스크립터에 저장된다. 마지막으로, 새로운 스트림이 새로운 스케일 가능성 차원을 가져오면, 스케일 가능성 마스크는 그에 따라 업데이트되어 이 새로운 차원에 대응하는 비트를 1로 설정한다.
- [0252] 위에서 언급한 문헌 w15182는 현재 비트-스트림에 사용된 스케일 가능성의 타입을 표시하지 않고 LHEVCDecoderConfigurationRecord를 정의한다. w15182에서 현재 정의된 대로 일반적인 'lhv1' / 'lhe1'이 사용되는 경우, 클라이언트 측의 파일 판독기는 스케일 가능성 타입을 이해하기 위해 VPS(video parameter set) 확장들을 파싱해야 한다. 이 VPS는 NALU(107)에서 체크될 수 있다. 이것은 복잡한 처리이다.
- [0253] 바람직한 실시예(나타내지 않음)에 따르면, 'scalability\_mask'라고 명명되는 16 비트의 스케일 가능성 마스크를 포함하는 새로운 LHEVCDecoderConfigurationRecord가 제안된다(박스 137 참조). 또 다른 실시예에 따르면, 전체 구조가 정수 바이트로 정렬된 채로 남아있다면, 스케일 가능성 마스크는 n 비트로 표현될 수 있으며, n은 정수이다. 예를 들어, HEVC 사양에서와 같이 n=8이다.
- [0254] LHEVCDecoderConfigurationRecord는 w15182에 정의된 "general\_level\_idc"라고 명명되는 필드를 포함하는데, 이는 명확성을 위해 표현되지 않았다. 필드 "general\_level\_idc"는 최대 픽셀 수와, 가능한 타일들 및 슬라이스들에 대한 정보를 제공한다.
- [0255] 특정 실시예들에 따르면, 예를 들어 "general\_level\_idc" 파라미터 다음에, 샘플들을 디코딩하는 데 요구되는 계층화된 HEVC의 종류를 명확하게 하기 위해 디코더 구성 레코드에 또 다른 파라미터를 추가하는 것이 제안된다.

“  
*unsigned int(8) general\_level\_idc;*  
*unsigned int(16) scalability\_mask;*  
 [...]  
*bit(1) complete\_representation;*  
 ”

[0256]

[0257]

"scalability\_mask"의 값(이 예에서는 '2')은 타입 공간적 또는 품질의 스케일 가능성을 표시한다. 이 필드는 클라이언트가 스케일 가능성 타입이 지원되는지를 검출하고 파일을 재생할 수 있는지 여부를 결정하는 데 도움이 되는 이점이 있다. 파일을 재생할 수 없는 경우, 예를 들어 베이스 계층 트랙과 같이, 더 낮은 동작 포인트를 선택할 수 있다.

[0258]

도 2는 고려되는 픽처들의 인헨스먼트 계층만의 픽처들이 4개의 타일로 분할될 때 ISO 기반 미디어 파일 포맷에 따라 2개의 스케일 가능성 계층을 캡슐화하는 예를 도시한다. 캡슐화는 4개의 추가적인 타일 트랙(140-143) 또는 인헨스먼트 타일 트랙들 ELTT를 전달한다.

[0259]

HEVC 타일 트랙과 유사한 방식으로, 인헨스먼트 계층의 공간적 하위 파트들에 효율적으로 액세스할 수 있도록 계층화된 HEVC 타일 트랙들을 정의하는 것이 가능하다. 그러한 경우, 실시예들에 따르면, LHEVCTileSampleEntry 샘플 디스크립션 포맷을 사용하여 특정 샘플들을 갖는 특정 트랙들이 생성된다.

[0260]

LHEVC 타일 트랙은 이 트랙의 타일(들)이 속하는 HEVC 계층의 비 비디오 코딩 계층의 NALU들을 운반하는 LHEVC 트랙에 대한 'tbas' 참조가 있는 비디오 트랙이다. 본 발명의 일 실시예에 따르면, 새로운 샘플 디스크립션 타입이 정의된다: 'lht1'.

[0261]

실시예들에 따르면, 타일 트랙의 샘플들도 또는 샘플 디스크립션 박스도 VPS, SPS, 또는 PPS NAL 유닛들을 포함하지 않을 것이고, 이들 NAL 유닛은, 트랙 참조 타입 'tbas'에 의해 식별되는 바와 같이, 연관된 계층을 포함하는 트랙의 샘플들 또는 샘플 디스크립션 박스에 있을 것이다(도 2에서의 인헨스먼트 계층 트랙(130)).

[0262]

실시예들에 따르면, 'tbas' 트랙 참조에 의해 표시된 바와 같이, LHEVC 타일 트랙 및 연관된 계층을 포함하는 트랙 또는 계층 트랙 양쪽 모두는, w15182의 부록 B에 정의된 바와 같이, 추출기들을 사용하여 원래의 비트-스트림이 어떻게 재구성되는지를 표시한다. 이 타일 트랙들에 있는 추출기들의 존재는 일부 애플리케이션 도메인들에서 제약이 될 수 있는데, 예를 들어, 특히 디코딩하고 재생할 타일들의 서브세트를 선택이 가능하도록, (도 4a를 참조하여 기술된 바와 같이) 각 타일 트랙보다는 타일 베이스 트랙들에 추출기들을 배치하는 것이 바람직할 수 있다. 또는, 다수의 타일형 계층의 경우에, 추출기들을 타일 베이스 트랙들에만 배치할 때 디스크립션 사이즈가 감소된다.

[0263]

도 4a 및 도 4b를 포함하는 도 4는 대안적인 트랙 및 타일 트랙 종속성 시그널링의 예들을 도시한다.

[0264]

도 4a에 도시된 시그널링 예에 따르면, 시간 설정형 샘플들(401)을 포함하는 베이스 계층 트랙(400)은 존재하는 타일들의 수만큼 베이스 계층의 타일 트랙들에 대한 트랙 참조들(403)을 갖는다(즉, 도 4a에서는 4). 베이스 계층 트랙은 트랙 참조 박스(이것은 메타데이터 파트, 예를 들어 도 8의 메타데이터 파트(804)에서 정의됨)에 선언된 각 타일 트랙에 대한 'scal' 참조 타입을 갖는다.

[0265]

유사하게, 각 타일 트랙은 트랙 참조 박스(이것은 메타데이터 파트, 예를 들어 도 8의 메타데이터 파트(804)에서 정의됨)에 선언된 'tbas' 트랙 참조 타입을 갖는다.

[0266]

또한, 베이스 계층 트랙(400)의 각 샘플은 베이스 계층에 존재하는 타일들의 수만큼 추출기들(402)을 갖는다.

[0267]

샘플들(407)을 포함하는 인헨스먼트 계층(406)은 동일한 디스크립션 및 구성 규칙들을 따른다. 베이스 계층 트랙과 인헨스먼트 계층 트랙은 'sbas' 및 'scal' 트랙 참조 타입들(404 및 405)을 통해 상호 참조된다. 'scal' 참조 타입은 여기서 추출기들이 각각의 참조되는 트랙 내의 샘플들로부터 각 샘플을 재구성하는 데 사용되는 것을 나타낸다.

[0268]

도 1a에 도시된 바와 같이, 타일 트랙들을 사용하는 HEVC 또는 L-HEVC 트랙은 추출기들을 사용하여 그 타일 트랙들로부터의 데이터를 참조할 수 있는데, 그 경우 트랙은 도 4a에 도시된 바와 같이 타일 트랙들에 대한



'scal' 트랙 참조들을 가질 것이다.

- [0269] 타일형 액세스 유닛의 데이터 공유 및 재구성
- [0270] 도 4b는 추출기가 사용되지 않는 종속성 시그널링의 예를 도시한다.
- [0271] 그러한 경우, 타일 트랙들을 사용하는 HEVC 또는 L-HEVC 트랙은 타일 트랙들에 대한 전용 'sabt' 트랙 참조 타입(453)을 사용하여 타일 순서를 표시해야 한다(추출기들의 사용을 위해 예약된, 'scal' 참조 타입은 더 이상 사용하지 않음). 타일 트랙들은 'tbas' 트랙 참조 타입으로 타일 베이스 트랙을 계속 참조한다.
- [0272] 'sabt' 트랙 참조는 HEVC 또는 L-HEVC 트랙으로부터의 타일 트랙들을 참조하는 데에만 사용될 수 있다. 'sabt' 트랙 참조가 트랙에 존재하는 경우, 이 트랙의 샘플들은 추출기들 또는 집계기들을 사용하지 않을 것이다; HEVC 또는 L-HEVC 트랙의 샘플들은, 트랙 참조들의 순서대로, 이 트랙의 'sabt' 트랙 참조들에 표시된 이용 가능한 모든 트랙에서 동일한 디코딩 시간을 갖는 샘플들에 포함된 NAL 유닛들을 샘플 데이터에 추가함으로써(즉, 편집 리스트들을 고려하지 않고 시간 대 샘플 테이블만을 사용하여) 재구성된다. 트랙에 대해 이용 가능한 데이터가 없다면 대응하는 샘플에 대해 데이터가 집계되지 않는다. SEI 접미사 NAL 유닛들은 재구성된 샘플에 있는 모든 비 VCL NAL 유닛들 뒤로 이동될 것이다.
- [0273] 트랙은 그 트랙 참조 박스에 정의된 'sabt' 및 'scal' 트랙 참조들 양쪽 모두를 갖지 않을 것이다. 또한, 추출기들이 사용되지 않는 경우(암시적 AU 재구성), 'scal' 트랙 참조의 사용은 사실상 유용하지 않아 제거될 수 있다. 실제로 'oinf' 박스는 각 동작 포인트에 대한 계층 식별자의 완전한 리스트를 제공하고 계층 식별자들은 각 트랙의 'tcon' 박스에 표시된다(또는 트랙 당 하나의 계층만 있는 경우 디코더 구성 정보 내의 계층 식별자로서 트랙에 직접 표시된다). 그 후 동작 포인트 정보 박스에 의해 표시된 바와 같이(트랙 참조에 의해서는 더 이상 표시되지 않음), 필요한 트랙들 내의 각각의 샘플들로부터, ISO/IEC 23008-2(HEVC 표준)에 부합하는 순서대로 그들의 NAL 유닛들을 배열함으로써, 액세스 유닛이 재구성된다.
- [0274] 타일 트랙에 저장된 LHEVC 샘플은 ISO/IEC 23008-2에 정의된 바와 같이, 하나 이상의 타일에 대한 슬라이스들의 완전한 세트이다. 전형적으로, 타일 트랙이 단일 타일을 참조하는 경우, 이 타일을 코딩하는 데 사용되는 슬라이스(들)만이 샘플에서 발견된다. 타일 트랙은 전형적으로 하나의 TileRegionGroupEntry(단일 타일 트랙), 또는 이미 HEVC에 대해 정의된, 이 타일 세트가 만들어지는 하나의 TileSetGroupEntry와 하나 이상의 종속 TileRegionGroupEntry(멀티 타일 트랙)를 포함한다.
- [0275] 타일 트랙에 저장된 LHEVC 샘플은, 샘플 내의 VCL NAL 유닛들이 샘플에 포함된 코딩된 슬라이스들이 IDR(Instantaneous Decoding Refresh) 슬라이스들이거나, CRA(Clean Random Access) 슬라이스들이거나, 또는 BLA(Broken Link Access) 슬라이스들임을 표시한다면, "sync" 샘플, 예를 들어, 랜덤 액세스형 탐색을 위한 동기 샘플로 간주된다.
- [0276] (w15182에서) 정규 LHEVC 샘플들에 대해 정의된 하위 샘플 및 샘플 그룹화는 LHEVC 타일 샘플에 대해 동일한 정의를 갖는다.
- [0277] 본 발명의 일 실시예에 따르면, HEVC 시퀀스의 완전한 타일들의 서브세트만을 디코딩하기 위한 구현이 결정될 수 있다. 이 경우, TileRegionGroupEntry 및 TileSetGroupEntry 샘플 그룹 디스크립션들 내의 타일 종속성 정보를 사용하여 불필요한 트랙들을 폐기하거나 HEVC 시퀀스를 디코딩하는 동안 일부 추출기들을 무시할 수 있다.
- [0278] 도 2에서, 도 1에서와 동일한 참조를 갖는 엘리먼트들은 유사하다. 또한, 명확성을 위해, 'moov' 박스만이 나타나 있다.
- [0279] 도 2에서, 'moov' 박스는 4개의 타일 트랙 박스(140, 141, 142, 143)인 추가적인 트랙 박스들을 포함한다. 여기서, 타일 트랙(141)만이 기술되어 있다. 다른 타일 트랙 박스들은 쉽게 추론될 수 있다.
- [0280] 타일 트랙 박스는 BL 및 EL 계층 트랙 박스들에 속하는 'tkhd' 박스들(111 또는 131)과 동일한 특성을 갖는, 'tkhd', 트랙 헤더 박스 또는 하위 박스(150)를 포함한다.
- [0281] 타일 트랙 박스는 다음을 제외하고는, BL 및 EL 계층 트랙 박스들에 속하는 'tref' 박스들과 동일한 특성을 갖는, 'tref', 트랙 참조 박스 또는 하위 박스(151)를 포함한다:
- [0282] - 그것은 타일 베이스 트랙과의 관계를 표시하는 4 바이트의 'tbas'를 포함하고,
- [0283] - 식별자 track\_IDs[]는 이 트랙에 대한 타일 베이스 트랙이 식별자 "ELBT"를 갖는 인스턴트 타일 트랙임을

나타낸다.

- [0284] 타일 트랙 박스는 BL 및 EL 트랙으로서 'mdia' 박스(152), 'stbl' 박스 또는 하위 박스(153), 'stds' 박스 또는 하위 박스(154)를 갖는 minf(명료함을 위해 표시하지 않음) 박스를 포함한다.
- [0285] 'stbl' 박스(153)는 'sgpd'에 포함된 'sgpd'(156) 및 'trif'(154)와 같은 트랙 샘플들에 속성들을 연관시키는 2개의 박스 또는 하위 박스를 포함한다. 이 박스들은 w15182에 잘 정의되어 있다.
- [0286] 'sgpd'는 특정 샘플 그룹 내의 샘플들의 공통 속성들을 기술하는 SampleGroupDescription 박스이다. 여기서, 파라미터 "def\_sample\_descr\_index"는 트랙의 모든 샘플에 적용되는 디폴트 속성: 첫 번째('trif' 박스에서 단 하나)를 표시한다.
- [0287] 'trif'는 고려된 타일에 관한 정보를 포함하는 TileRegionGroupEntry 박스이다. 이 경우, 고려된 타일은 값 '1'을 갖는 groupID에 의해 식별되고, 그 위치 및 사이즈는 각각 "horizontal\_offset", "vertical\_offset" 및 "region\_width", "region\_height"에 의해 정의된다. 예약된 파라미터는 의미가 없으며, 독립 플래그는 타일이 자체 포함되어 있는지(즉, 임의의 다른 타일들을 재구성할 필요가 없는지)를 표시한다. 마지막으로, full-frame 파라미터는 타일이 전체 픽처를 커버하고 있는지(1) 또는 그렇지 않은지(0)를 표시한다.
- [0288] 본 발명의 일 실시예에 따르면, 새로운 샘플 엔트리(155)를 정의하는 4 바이트는 LHEVC 타일 트랙의 미디어 데이터 또는 샘플들에 대응하는 비트-스트림이 DecoderConfigurationRecord 또는 더 구체적으로는 아래 'lhvC' 박스(156)에 정의된 LHEVCDecoderConfigurationRecord에 주어진 (프로파일, 티어, 및 레벨을 포함하는) 구성 하에서 동작하는 HEVC 디코더와 호환가능하며 이에 의해 이용 가능하다는 것을 표시한다.
- [0289] 실시예에서, 4 바이트는 'lht1'이다.
- [0290] 이 새로운 샘플 엔트리의 정의는 하기의 것일 수 있다:
- [0291] "
- [0292] 박스 타입: 'lht1'
- [0293] 컨테이너: 샘플 디스크립션 박스('stds')
- [0294] 의무사항(mandatory): 아니요
- [0295] 수량: 0개 이상의 샘플 엔트리가 존재할 수 있다.
- [0296] "
- [0297] 새로운 샘플 엔트리의 정의는 파서가 LHEVC 타일 트랙을 참조한다는 것을 즉시 인식할 수 있게 한다. 이것은 기존 샘플 엔트리에서는 허용되지 않았었다.
- [0298] 또한, 타일 트랙과 관련된 일부 특정 속성들만이 도입될 수 있다.
- [0299] 위에서 설명한 바와 같이, 이 샘플 엔트리는 LHEVC 타일 트랙의 미디어 샘플들을 기술한다. LHEVC 타일 트랙 (샘플 엔트리 타입 'lht1')에 대한 VisualSampleEntry의 폭과 높이는 트랙에 포함된 타일(들) 또는 타일세트(들)의 최대 폭 및 높이로 설정될 것이다. 타일 트랙의 트랙 헤더 내의 레이아웃 정보(즉, 계층, 타일들을 위치 지정하기 위한 매트릭스, 폭 및 높이)는 'tbas' 트랙 참조에 의해 식별되고, 그렇지 않다면 무시되는 바와 같이, 연관된 참조 트랙(타일 베이스 트랙이라고도 함)의 트랙 헤더 정보와 동일할 것이다.
- [0300] 바람직하게는 'lht1' 샘플 디스크립션에서의 임의의 'clap'(Clean Aperture를 의미함) 및 'pasp'(Pixel Aspect Ratio를 의미함) 박스는 무시될 것이다.
- [0301] 따라서, 특정 종류의 표준 VisualSampleGroupEntry 디스크립터인 LHEVC 타일 디스크립터에 대한 특정 샘플 디스크립션이 생성된다.

```

class LHEVCTileSampleEntry () extends VisualSampleEntry ('lht1'){
 MPEG4BitRateBox ();
 extra_boxes boxes;
}

```

[0302]

MPEG4BitRateBox 및 extra\_box들 양쪽 모두는 옵션이다.

[0303]

바람직하게는 LHEVCTileSampleEntry는 임의의 LHE VCCConfigurationBox(또는 LHVCCConfigurationBox 또는 계층화된 HEVC 포맷에 대한 구성 박스를 지정하는 임의의 명칭), HEVCCConfigurationBox 또는 MPEG4ExtensionDescriptorsBox를 포함하지 않는다. 실제로, 이러한 박스들은 'tbas' 트랙 참조 타입으로 표시된 바와 같이 타일 베이스 LHEVC 트랙 샘플 디스크립션에서 발견된다.

[0304]

LHEVCTileSampleEntry에는 다른 옵션 박스들이 포함될 수 있다. 일반적으로, 타일 베이스 LHEVC 트랙의 SampleDescriptionBox에 존재하는 엔트리들의 수만큼 LHEVC 타일 트랙의 SampleDescriptionBox에 엔트리들이 존재한다. SampleDescriptionBox는 HEVC 타일 트랙에 관한 기술적 정보를 포함하는 HEVC 타일 트랙의 잘 알려진 디스크립터이다.

[0305]

LHEVC 타일 트랙에 대한 MIME 타입 'codecs' 파라미터에 대한 하위 파라미터들은 표준을 정의하는 w15182의 부록 E.3에 정의된 규칙을 따른다. 디코더 구성 레코드는 'tbas' 트랙 참조 타입에 의해 표시된 대로, 베이스 트랙 샘플 디스크립션에서 취해진다. 그 후, 이 디코더 구성 레코드를 사용하여 MIME 타입 'codecs'에 대한 하위 파라미터들이 구성된다.

[0306]

바람직하게는, L-HEVC에 대한 코덱 하위 타입 파라미터들은 다음을 제외하고 HEVC와 동일하다: 코덱 타입이 LHVC 타일 트랙의 일반적인 L-HEVC 미디어 샘플들(즉, 'lhv1' 코드 포인트들)을 식별하는 경우, 구성된 HEVC 코덱 파라미터에는 ".SXX"가 추가될 것이고, "S"는 스케일 가능성 타입을 표시하고, "XX"는 이 트랙에 대한 스케일 가능성 마스크 값에 대응하는 바이트들이고; 0인 경우에는 후행 바이트(trailing byte)를 생략할 수 있다. 이것은 예를 들어, DASH 표현들에서, 연관된 비디오를 인코딩하는 데 사용되는 코덱에 대한 정확한 정보를 얻기 위해 유용할 수 있다. 예를 들어, 멀티뷰 스트리밍 애플리케이션에서, 다중-계층 HEVC 디코더를 갖는 DASH 클라이언트는 공간적 또는 품질 스케일 가능성을 표시하는 코덱 파라미터로 선언된 표현을 선택하지 않을 것이다.

[0307]

MIME 타입 코덱들에 대한 하위 파라미터들과 관련된 대안적인 실시예에서, w15182에서 제안된 것보다 더 콤팩트한 형태의 MIME 타입이 DASH에서 유용할 수 있다. 특히 단일 스케일 가능성 타입을 갖는 트랙 당 하나의 계층을 갖는 구성에서; 즉, 옵션의 코덱 MIME 파라미터의 리스트 엔트리 내의 샘플 엔트리가 'shv1', 'she1', 'mhv1', 'mhe1', 'lhv1' 또는 'lhe1'인 경우, 동일한 리스트 엔트리에는 더 이상의 데이터가 제공되지 않는다. 특히, 하나의 단일 동작 포인트가 발표되므로, OlsIdx 및 MaxTid 파라미터들뿐만 아니라 BLInternal 파라미터는 옵션이 될 수 있다. 이는 'lhevcpt1' 옵션을 건너뛰고 예를 들어 인핸스먼트 계층(스케일러블)을 "shv1.A7.1.L120.B0"으로 선언할 수 있게 한다.

[0308]

도 3은 하나 이상의 실시예의 단계들이 구현될 수 있는 서버 또는 클라이언트 디바이스(300)의 블록도를 나타낸다.

[0309]

바람직하게는, 디바이스(300)는 통신 버스(302), 디바이스의 전원 인가 시에는 프로그램 ROM(306)으로부터의 명령어들을, 그리고 전원 인가 후에는 메인 메모리(308)로부터의 소프트웨어 애플리케이션에 관련된 명령어들을 실행할 수 있는 중앙 처리 유닛(CPU)(304)을 포함한다. 메인 메모리(308)는 예를 들어, 통신 버스(302)를 통해 CPU(304)의 작업 영역으로서 기능하는 RAM(Random Access Memory) 타입이고, 그의 메모리 용량은 확장 포트(도시되지 않음)에 연결되는 옵션의 RAM에 의해 확장될 수 있다. 소프트웨어 애플리케이션에 관한 명령어들은, 예를 들어, 하드 디스크(HD)(310) 또는 프로그램 ROM(306)으로부터 메인 메모리(308)에 로딩될 수 있다. 그러한 소프트웨어 애플리케이션은, CPU(304)에 의해 실행될 때, 도 1 및 도 2를 참조하여 기술된 캡슐화 단계가 서버에서 수행되게 한다.

[0310]

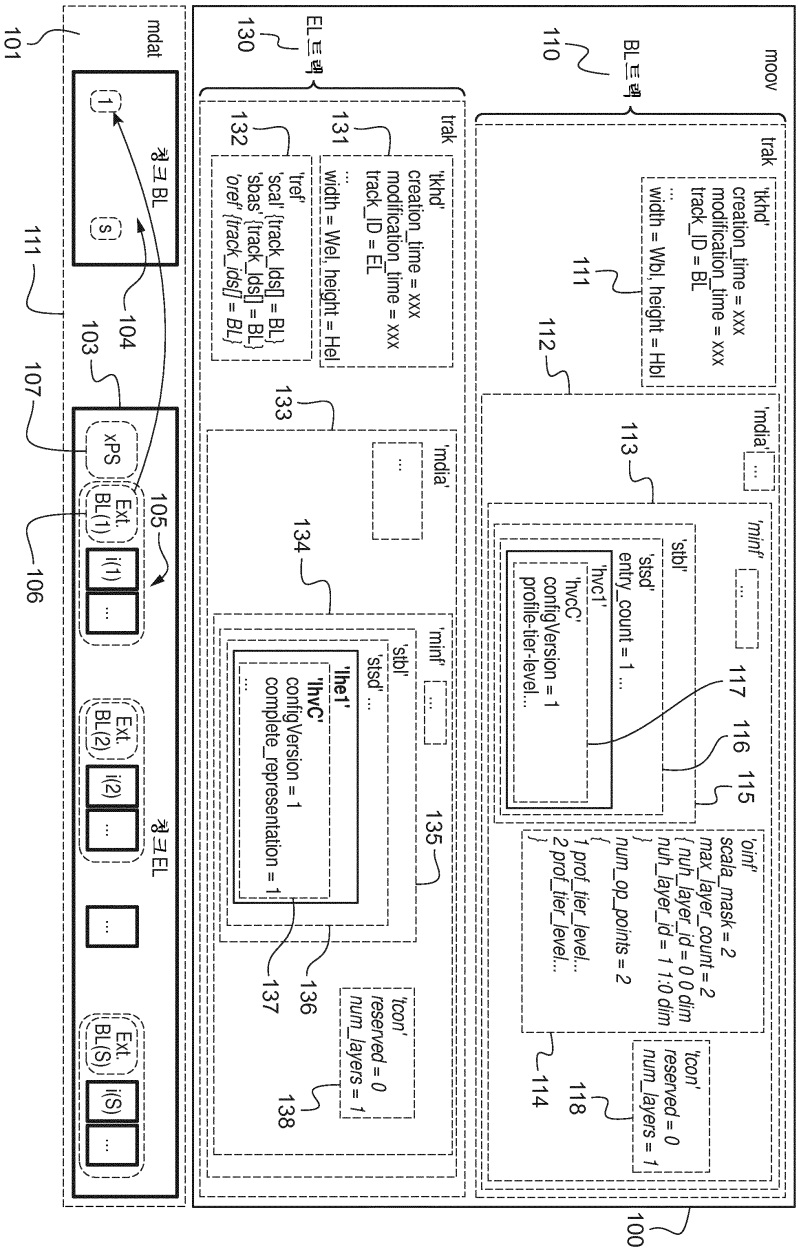
참조 번호 312는 통신 네트워크(314)에 디바이스(300)의 연결을 허용하는 네트워크 인터페이스이다. CPU(304)에 의해 실행되는 경우, 소프트웨어 애플리케이션은 네트워크 인터페이스를 통해 수신된 요청들에 반응하고, 데이터 스트림들 및 요청들을 네트워크를 통해 다른 디바이스들에 제공하도록 구성된다.

[0311]

- [0312] 참조 번호 316은 사용자에게 정보를 디스플레이하고 및/또는 사용자로부터 입력을 수신하기 위한 사용자 인터페이스들을 나타낸다.
- [0313] 변형 예로서, 멀티미디어 비트-스트림들의 수신 또는 송신을 관리하기 위한 디바이스(300)가 도 1, 도 2 및 도 3을 참조하여 기술된 바와 같은 방법들을 구현할 수 있는 하나 이상의 전용 집적 회로(ASIC)로 구성될 수 있다는 것을 여기서 언급해야만 한다. 이러한 집적 회로들은 예를 위한 것이며 비디오 시퀀스들을 생성하거나 디스플레이하고 및/또는 오디오 시퀀스들을 청취하기 위해 장치에 비-제한적으로 통합된다.
- [0314] 본 발명의 실시예들은, 예를 들어 특정한 관심 영역으로 줌인(zoom into)하기 위해서, 예를 들어 TV용의 원격 제어기로서 동작하는 카메라, 스마트폰 또는 태블릿과 같은 디바이스 내에 내장될 수 있다. 그것들은 또한 특정 관심 영역을 선택함으로써 TV 프로그램의 개인화된 브라우징 경험을 갖도록 동일한 디바이스로부터 사용될 수도 있다. 사용자에게 의한 이들 디바이스의 또 다른 사용법은, 그가 선호하는 비디오들의 일부 선택된 하위 파트들을 다른 연결된 디바이스들과 공유하는 것이다. 그것들은 또한 감시 카메라가 본 발명의 생성 파트를 지원하는 경우에는, 감시 하에 놓인 빌딩의 특정 영역에서 발생하는 것을 모니터링하기 위해 스마트폰 또는 태블릿에서 사용될 수도 있다.
- [0315] 물론, 국부적인 그리고 특정 조건들을 만족시키기 위해서, 관련 기술분야의 통상의 기술자라면 전술한 해결책에 많은 수정들 및 변경들을 적용할 수 있는데, 그것들은 모두 다음의 청구항들에 의해 정의된 본 발명의 보호 범위 내에 포함된다.

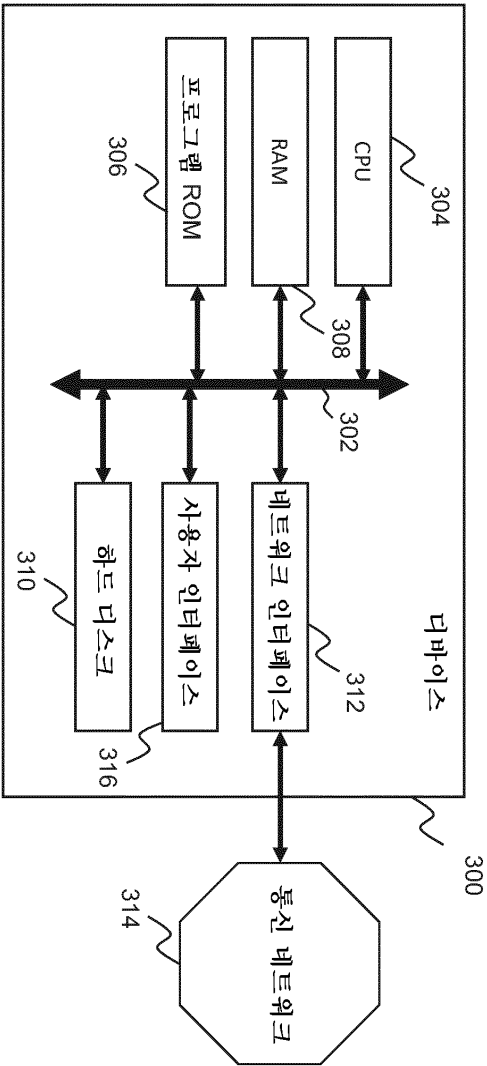


도면  
도면1

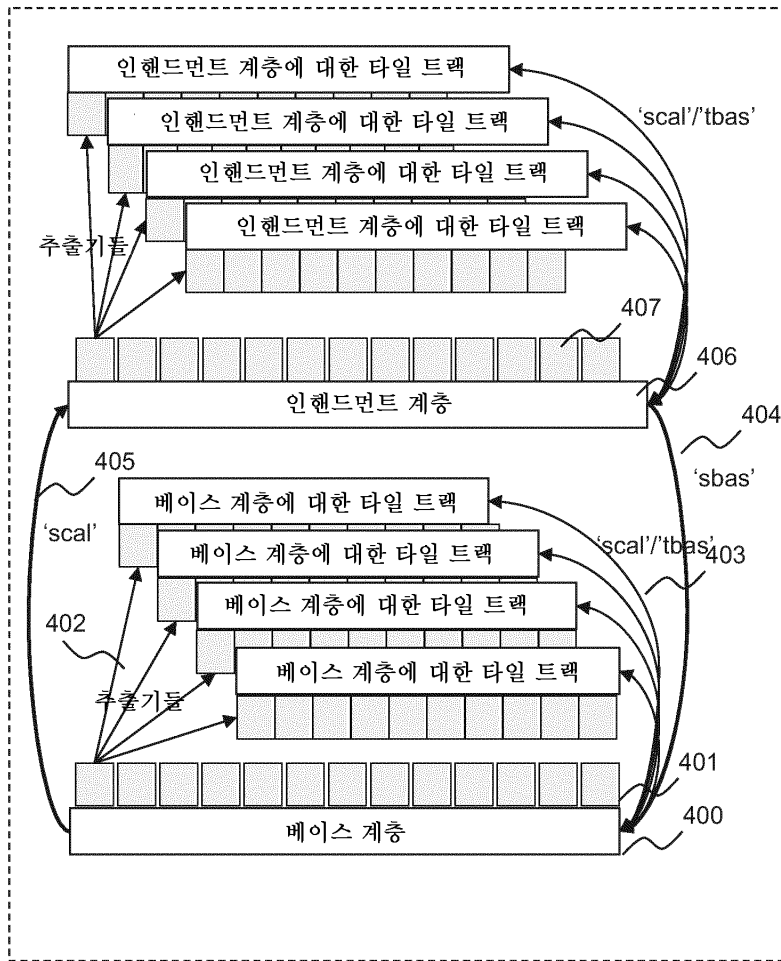




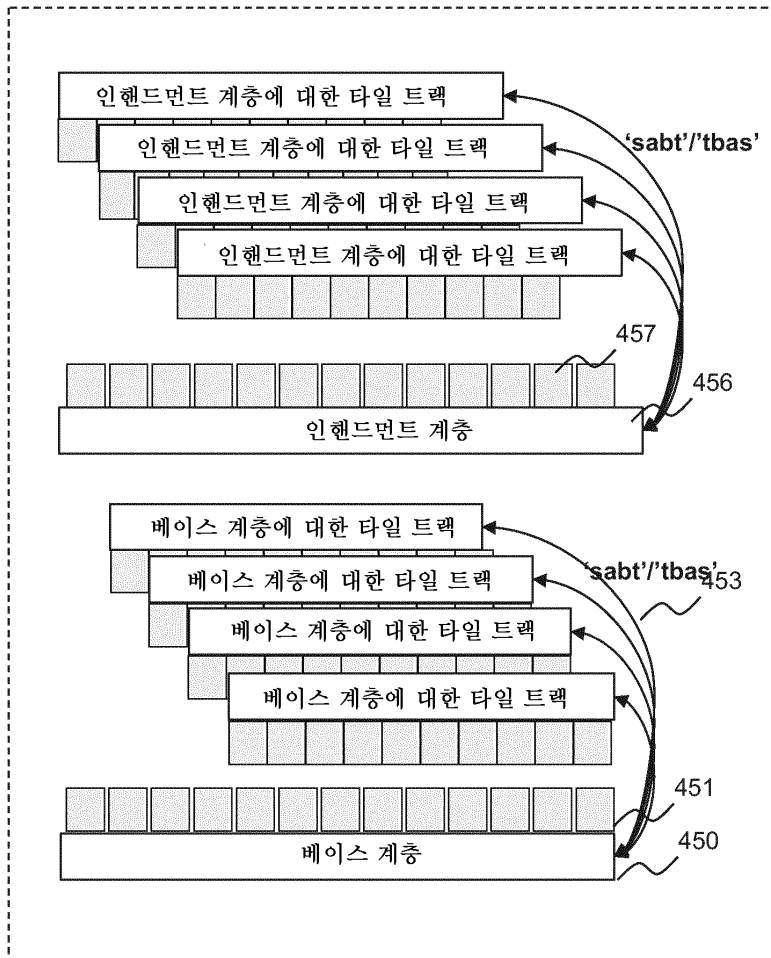
도면3



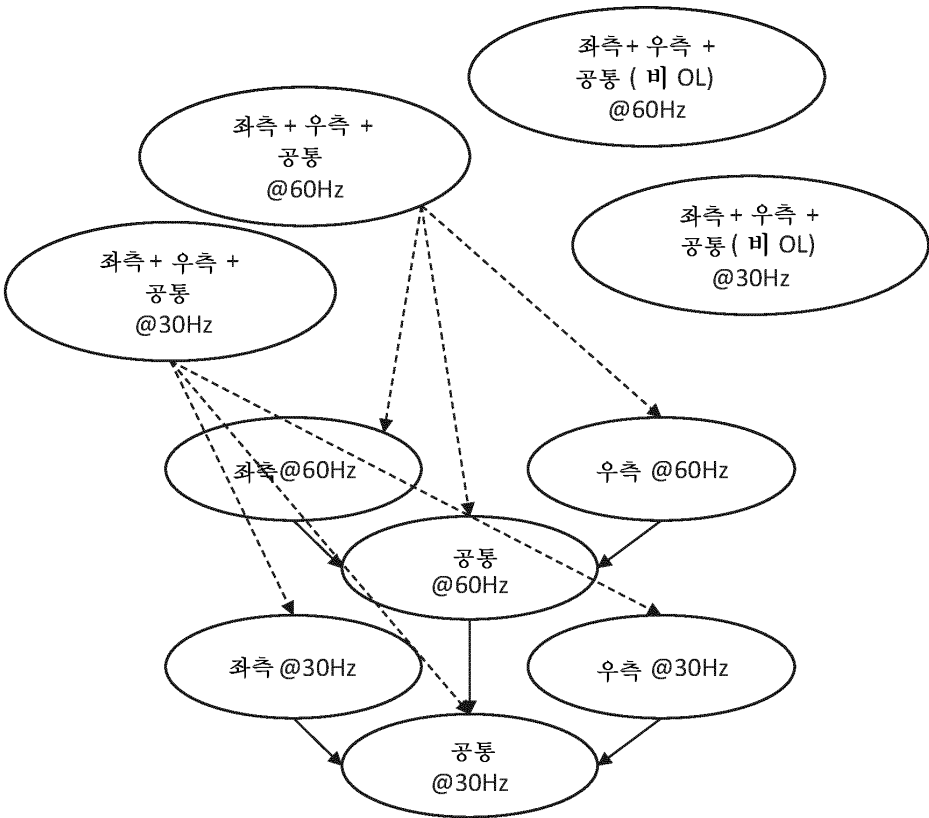
도면4a



도면4b



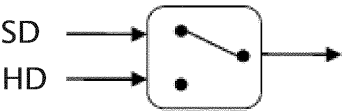
도면5a



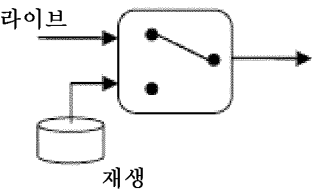
도면5b

|                    |                    |
|--------------------|--------------------|
| HEVC 계층 2 하위 파티션 0 | HEVC 계층 2 하위 파티션 1 |
| HEVC 계층 1 하위 파티션 0 | HEVC 계층 1 하위 파티션 1 |
| HEVC 계층 0 하위 파티션 0 | HEVC 계층 0 하위 파티션 1 |

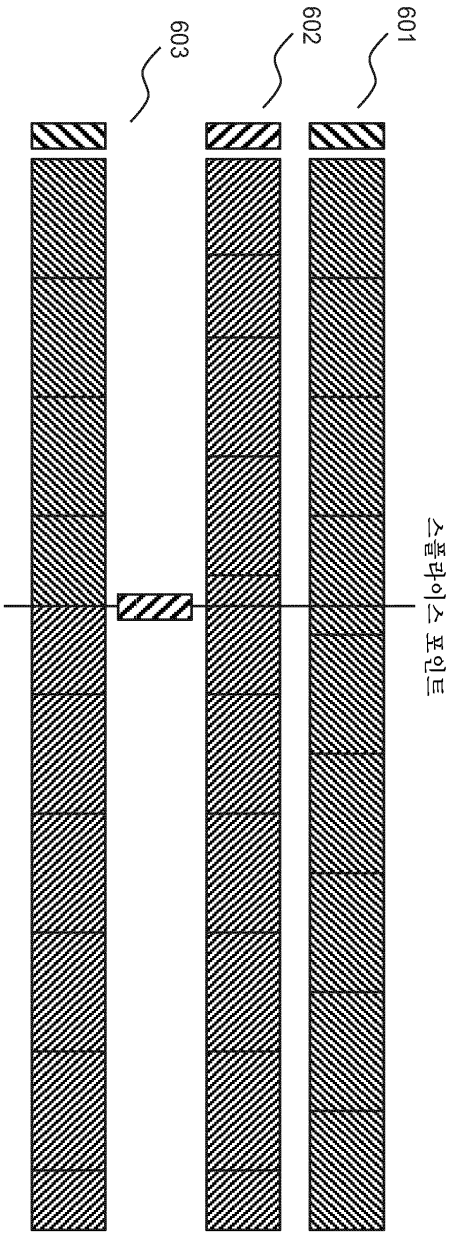
도면6a

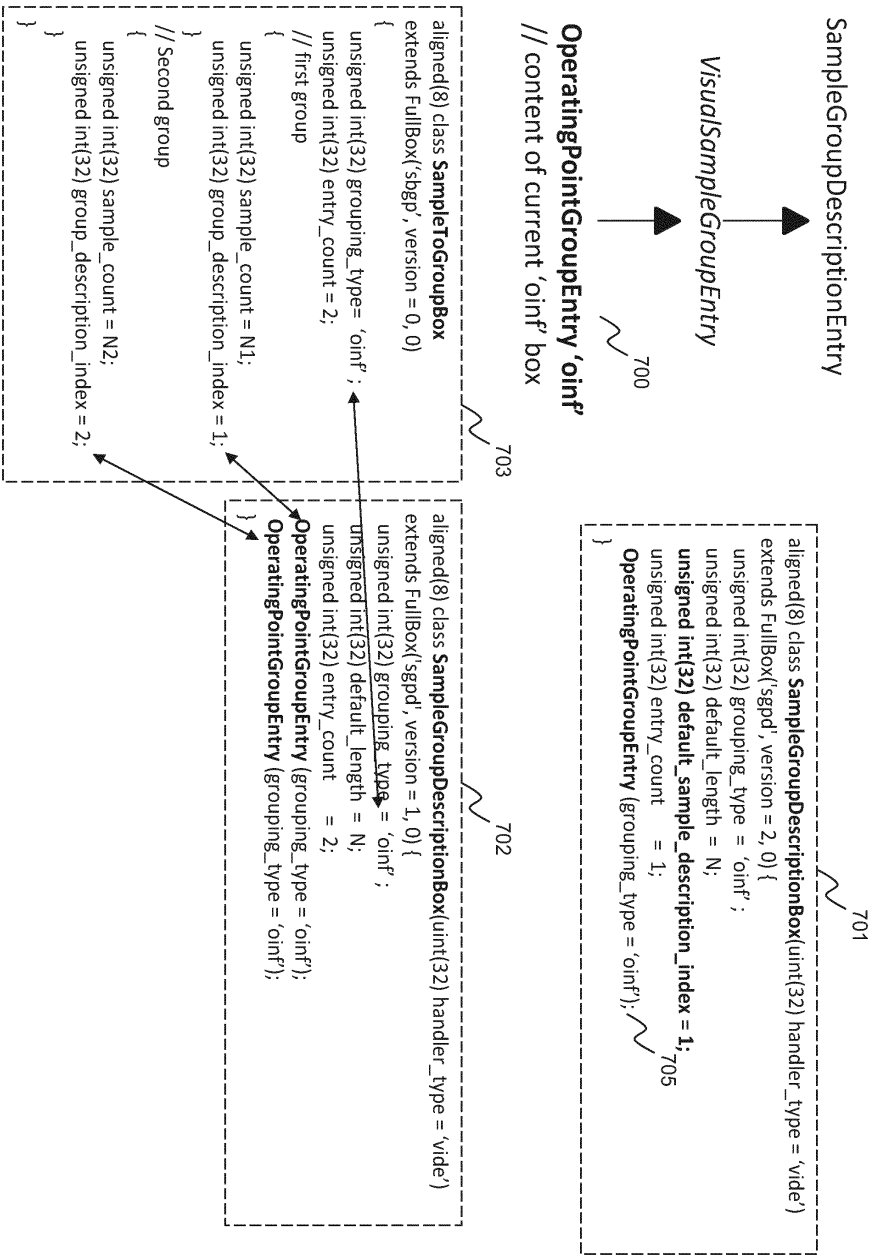


도면6b



도면6c

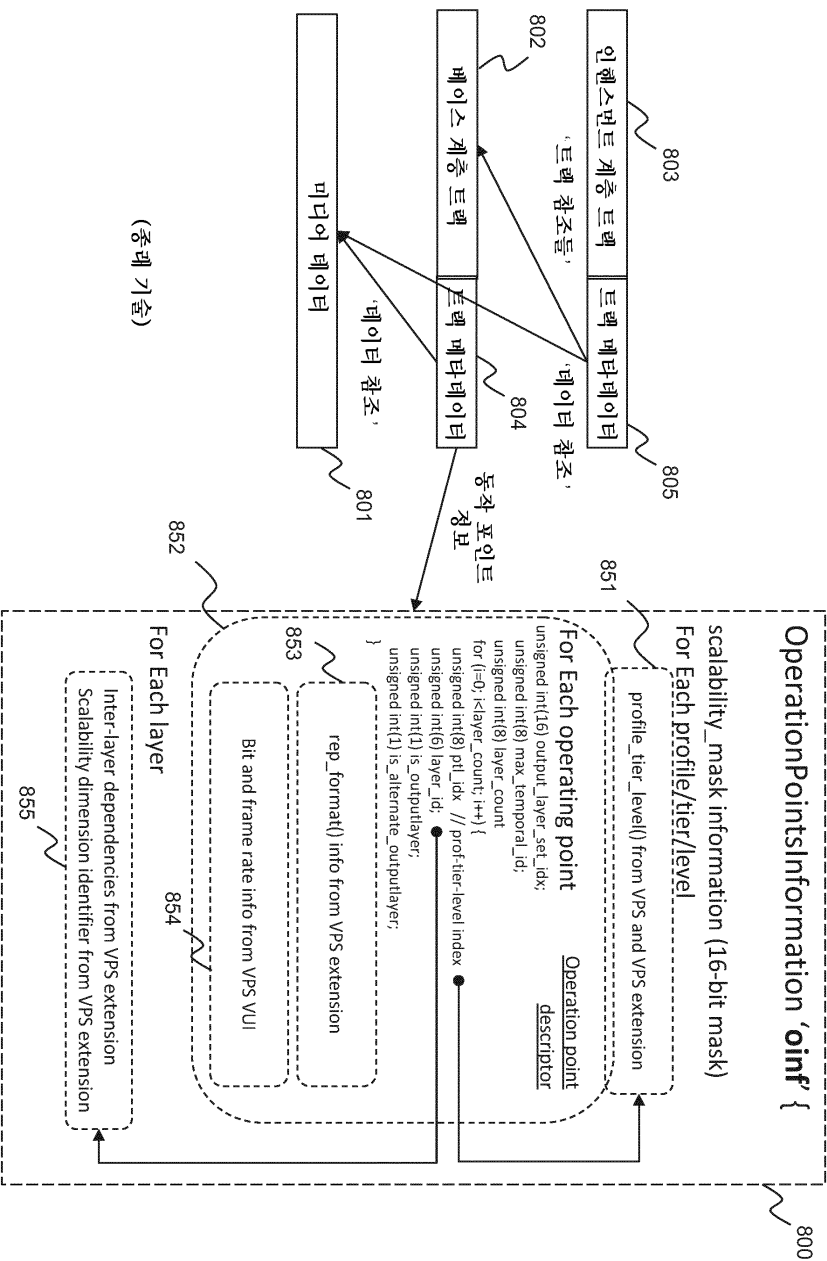


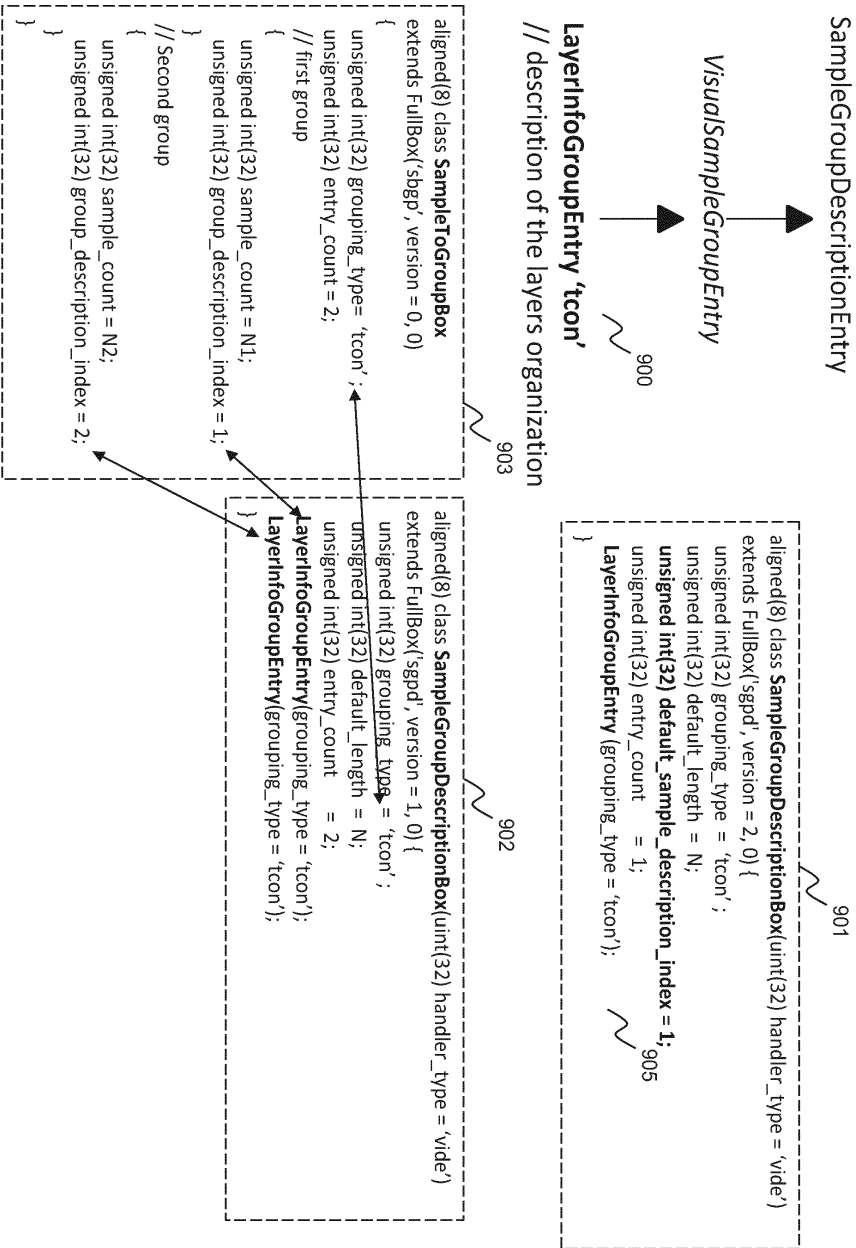


도면 7

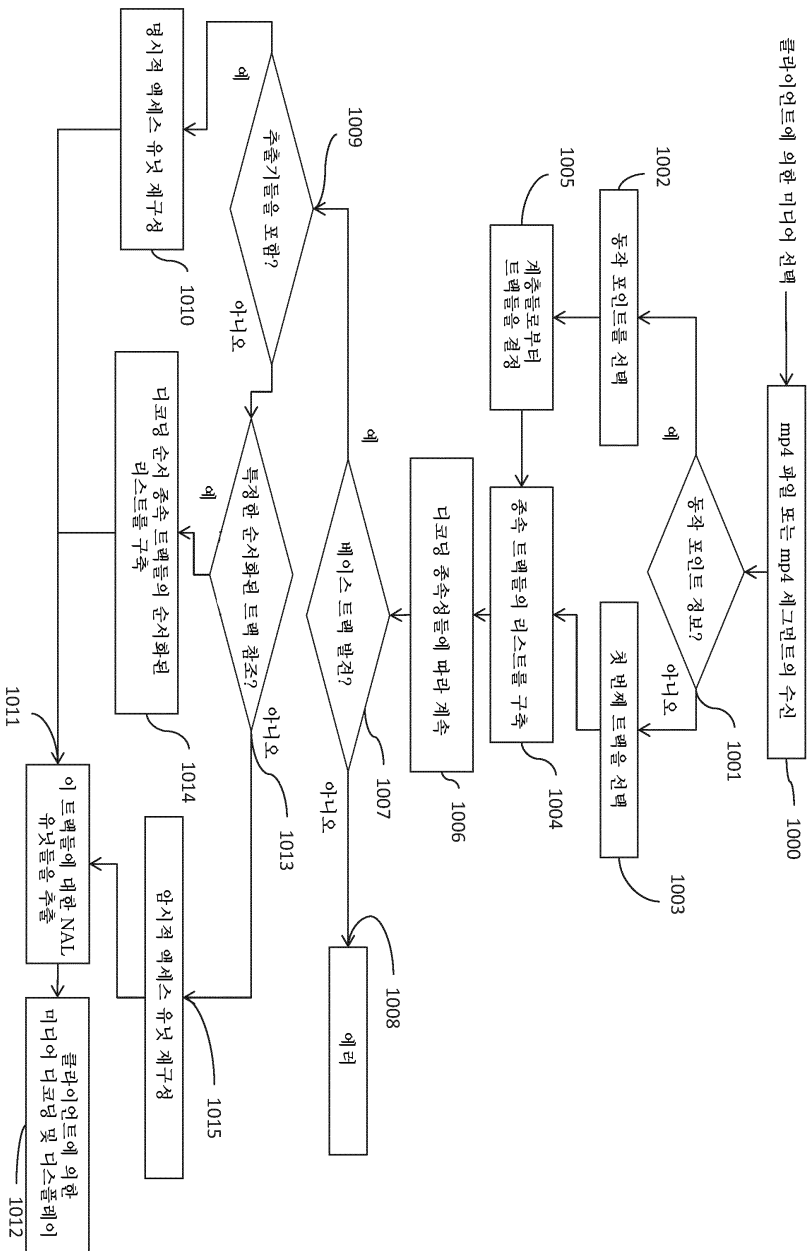


도면8





도면9



도면10

도면11

```

1100
~
aligned(8) class LHEVCDecoderConfigurationRecord {
 unsigned int(8) configurationVersion = 1;
 bit(1) complete_representation; ~ 1102
 bit(3) reserved = '111'b;
 unsigned int(12) min_spatial_segmentation_idc;
 bit(6) reserved = '111111'b;
 unsigned int(2) parallelismType;
 bit(2) reserved = '11'b;
 bit(3) numTemporalLayers;
 bit(1) temporalIdNested;
 bit(2) reserved; ~ 1101
 unsigned int (6) layer_id; ~
 unsigned int(2) lengthSizeMinusOne;
 unsigned int(8) numOfArrays;
 for (j=0; j < numOfArrays; j++) {
 bit(1) array_completeness;
 unsigned int(1) reserved = 0;
 unsigned int(6) NAL_unit_type;
 unsigned int(16) numNalus;
 for (i=0; i < numNalus; i++) {
 unsigned int(16) nalUnitLength;
 bit(8*nalUnitLength) nalUnit;
 }
 }
}

```

도면12

