



(19) 대한민국특허청(KR)

(12) 등록특허공보(B1)

(45) 공고일자 2018년09월12일

(11) 등록번호 10-1897945

(24) 등록일자 2018년09월05일

(51) 국제특허분류(Int. Cl.)

H04N 21/2343 (2011.01) H04N 21/4728

(2011.01)

H04N 21/6587 (2011.01) H04N 21/845 (2011.01)

H04N 21/854 (2011.01)

(52) CPC특허분류

H04N 21/2343 (2013.01)

H04N 21/234381 (2013.01)

(21) 출원번호 10-2018-7016192(분할)

(22) 출원일자(국제) 2014년07월21일

심사청구일자 2018년06월08일

(85) 번역문제출일자 2018년06월07일

(65) 공개번호 10-2018-0067713

(43) 공개일자 2018년06월20일

(62) 원출원 특허 10-2016-7003814

원출원일자(국제) 2014년07월21일

심사청구일자 2016년07월25일

(86) 국제출원번호 PCT/EP2014/065654

(87) 국제공개번호 WO 2015/011108

국제공개일자 2015년01월29일

(30) 우선권주장

1313157.8 2013년07월23일 영국(GB)

(56) 선행기술조사문헌

JEAN LE FEUVRE et al., "Support for efficient tile access in the HEVC File Format", 104th MPEG meeting, ISO/IEC JTC1/SC29/WG11 MPEG2012/M29231, April 2013.

WO2012168365 A1

'Text of ISO/IEC FDIS 14496-15:2010 3rd edition Carriage of NAL unit structured video in the ISO BMFF', 104. MPEG Meeting; Incheon(22-4-2013); (Motion Picture Expert Group or ISO/IEC JTC1/SC29/WG11)

전체 청구항 수 : 총 15 항

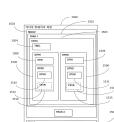
심사관 : 홍기완

(54) 발명의 명칭 서브 트랙 피처를 이용하여 분할된 시간 설정형 미디어 데이터를 캡슐화하는 방법, 디바이스 및 컴퓨터 프로그램

(57) 요 약

본 발명은 클라이언트의 디바이스에 스트리밍되도록 분할된 시간 설정형 미디어 데이터를 캡슐화하는 것과 관련되고, 상기 분할된 시간 설정형 미디어 데이터는 시간 설정형 샘플들을 포함하고, 각각의 시간 설정형 샘플은 복수의 서브샘플을 포함한다. 상기 서브샘플들의 적어도 일부가 서브샘플들의 그룹에 속하도록 몇몇의 상기 서브

(뒷면에 계속)

대 표 도

샘플들을 그룹화하여 적어도 하나의 그룹을 형성한 후에, 상기 시간 설정형 샘플들 중 하나의 상기 복수의 서브 샘플 중에서 동일한 그룹의 적어도 2개의 서브샘플이 선택되고 상기 선택된 서브샘플들을 포함하는 적어도 하나의 트랙이 생성된다. 다음으로, 상기 적어도 하나의 생성된 트랙에 대해, 상기 적어도 하나의 트랙을 생성하기 위해 선택된 상기 서브샘플들 각각에 대해 하나의 서브 트랙 박스가 생성되고, 상기 생성된 서브 트랙 박스들 중 적어도 하나는 상기 선택된 서브샘플들에 공통인 정의들 및 특성들을 포함한다. 상기 생성된 트랙들 각각이 적어도 하나의 미디어 파일에 독립적으로 캡슐화된다.

(52) CPC특허분류

HO4N 21/4728 (2013.01)*HO4N 21/6587* (2013.01)*HO4N 21/8451* (2013.01)*HO4N 21/85406* (2013.01)

(72) 발명자

르 훼브르 장프랑스 애프-94230 까샹 에스끄 베1 아브뉘 뒤 프
레지던 월슨 7**꽁꼴라또 씨릴**

프랑스 애프-77380 꽁 라 빌 뤼 드 라 헤르스 18

명세서

청구범위

청구항 1

미디어 데이터에 기초해서 적어도 하나의 미디어 파일을 서버에서 생성하는 방법이며, -상기 미디어 데이터는 비디오 프레임들을 포함하고, 각각의 비디오 프레임은 복수의 타일 영역들을 포함함-,

동일 그룹에 그룹화된 하나 이상의 타일 영역을 각각 포함하는 서브 트랙들을 포함하는 하나 이상의 트랙을 생성하는 단계;

상기 서브 트랙에 포함된 상기 하나 이상의 타일 영역의 그룹의 식별자를 포함하는 박스를 생성하는 단계; 및

상기 생성된 트랙, 상기 생성된 서브 트랙, 및 상기 생성된 박스 각각에 기초해서 적어도 하나의 미디어 파일을 생성하는 단계를 포함하는, 방법.

청구항 2

제1항에 있어서,

상기 식별자는 ISO/IEC 14496-12에 정의된 영화 박스(moov box)에 기술된, 방법.

청구항 3

제1항에 있어서,

상기 식별자는 샘플의 NAL 유닛들을 특정 TileRegionGroupEntry에 관련시키는데 사용되는, 방법.

청구항 4

제1항에 있어서,

상기 생성된 박스 중 하나 이상 또는 상기 생성된 트랙은 상기 생성된 박스 중 상기 하나 이상 또는 상기 생성된 트랙이 타일링 정보를 제공함을 나타내는 정보를 포함하는, 방법.

청구항 5

제1항에 있어서,

상기 생성된 박스 중 하나 이상과 관련된 하나 이상의 서브 트랙 내에 타일영역들의 그룹이 인코딩되는, 방법.

청구항 6

제5항에 있어서,

상기 생성된 박스 중 하나 이상은 상기 생성된 박스 중 하나 이상과 관련된 서브 트랙 내에 인코딩된 타일 영역들의 수를 나타내는 정보를 포함하는, 방법.

청구항 7

미디어 데이터에 기초해서 적어도 하나의 미디어 파일을 생성하는 장치이며, -상기 미디어 데이터는 비디오 프레임들을 포함하고, 각각의 비디오 프레임은 복수의 타일 영역들을 포함함-,

(i) 동일 그룹에 그룹화된 하나 이상의 타일 영역을 각각 포함하는 서브 트랙들을 포함하는 하나 이상의 트랙 및 (ii) 상기 서브 트랙에 포함된 상기 하나 이상의 타일 영역의 그룹의 식별자를 포함하는 박스를 생성하기 위한 생성 수단; 및

상기 생성된 트랙, 상기 생성된 서브 트랙, 및 상기 생성된 박스 각각에 기초해서 적어도 하나의 미디어 파일을 생성하기 위한 생성 수단을 포함하는, 장치.

청구항 8

제7항에 있어서,

상기 식별자는 ISO/IEC 14496-12에 정의된 영화 박스(moov box)에 기술된, 장치.

청구항 9

제7항에 있어서,

상기 식별자는 샘플의 NAL 유닛들을 특정 TileRegionGroupEntry에 관련시키는데 사용되는, 장치.

청구항 10

미디어 데이터에 기초해서 적어도 하나의 미디어 파일을 디코딩하는 방법이며, -상기 미디어 데이터는 비디오 프레임들을 포함하고, 각각의 비디오 프레임은 복수의 타일 영역들을 포함함-,

상기 적어도 하나의 미디어 파일을 수신하는 단계 -상기 적어도 하나의 미디어 파일은, 동일 그룹에 그룹화된 하나 이상의 타일 영역을 각각 포함하는 서브 트랙들을 포함하는 하나 이상의 트랙, 및 상기 서브 트랙에 포함된 상기 하나 이상의 타일 영역의 그룹의 식별자를 포함하는 박스를 포함함-; 및

상기 식별자를 참조함으로써 식별된 상기 하나 이상의 타일 영역을 디코딩하는 단계를 포함하는, 방법.

청구항 11

제10항에 있어서,

상기 식별자는 ISO/IEC 14496-12에 정의된 영화 박스(moov box)에 기술된, 방법.

청구항 12

제10항에 있어서,

상기 식별자는 샘플의 NAL 유닛들을 특정 TileRegionGroupEntry에 관련시키는데 사용되는, 방법.

청구항 13

미디어 데이터에 기초해서 적어도 하나의 미디어 파일을 디코딩하는 장치이며, -상기 미디어 데이터는 비디오 프레임들을 포함하고, 각각의 비디오 프레임은 복수의 타일 영역들을 포함함-,

상기 적어도 하나의 미디어 파일을 수신하기 위한 수신 수단 -상기 적어도 하나의 미디어 파일은, 동일 그룹에 그룹화된 하나 이상의 타일 영역을 각각 포함하는 서브 트랙들을 포함하는 하나 이상의 트랙, 및 상기 서브 트랙에 포함된 상기 하나 이상의 타일 영역의 그룹의 식별자를 포함하는 박스를 포함함-; 및

상기 식별자를 참조함으로써 식별된 상기 하나 이상의 타일 영역을 디코딩하기 위한 디코딩 수단을 포함하는, 장치.

청구항 14

제13항에 있어서,

상기 식별자는 ISO/IEC 14496-12에 정의된 영화 박스(moov box)에 기술된, 장치.

청구항 15

제13항에 있어서,

상기 식별자는 샘플의 NAL 유닛들을 특정 TileRegionGroupEntry에 관련시키는데 사용되는, 장치.

발명의 설명**기술 분야**

[0001] 본 발명은 일반적으로, 예를 들어, MPEG 표준화 기구에 의해 정의되는 바와 같은 베이스 미디어 파일 포맷(Base

Media File Format)에 따라, 스트림 전달을 향상시키는, 시간 설정형 미디어 데이터의 캡슐화의 분야에 관한 것으로, 특히 압축된 비디오 스트림들에서 사용자가 선택한 관심 영역들의 HTTP(HyperText Transfer Protocol) 스트리밍에 관한 것이다. 보다 구체적으로, 본 발명은, 서브 트랙 피처를 이용하여, 특히 하나 이상의 파일들의, 데이터의 효율적인 스트리밍을 가능하게 하는 공간 파일들과 같은 분할된 데이터를 포함하는 기본 스트림의 캡슐화에 대한 방법, 디바이스 및 컴퓨터 프로그램에 관한 것이다.

배경 기술

[0002]

비디오 코딩은, 일련의 비디오 이미지들을 콤팩트한 디지털화된 비트 스트림으로 변환하여 이러한 비디오 이미지들이 송신되거나 저장될 수 있도록 하는 하나의 방식이다. 인코딩 디바이스는 이러한 비디오 이미지들을 코딩하는데 사용될 수 있고, 관련된 디코딩 디바이스는 디스플레이 및 시청을 위해 이러한 비트 스트림을 재구성하는데 사용될 수 있다. 일반적인 목적은 이러한 비트 스트림을 본래 비디오 정보보다 더 작은 사이즈가 되도록 형성하는 것이다. 이는 유리하게도 이러한 비트 스트림을 송신하거나 또는 저장하는데 요구되는 송신 네트워크 또는 스토리지 디바이스의 용량을 감소시켜 준다. 송신되기 위해서, 비디오 비트 스트림은 일반적으로, 통상 헤더들 또는 체크 비트들을 추가하는 송신 프로토콜에 따라서 캡슐화된다.

[0003]

최근, MPEG(Moving Picture Experts Group)은, HTTP(HyperText Transfer Protocol)을 통한 기존 스트리밍 솔루션들을 통합하여 대체하는 새로운 표준을 공개하였다. "DASH(Dynamic Adaptive Streaming over HTTP)"라 하는, 이러한 새로운 표준은, 표준 웹 서버들을 기반으로 HTTP를 통한 미디어 스트리밍 모델을 지원하는 것으로 의도되어, 인텔리전스(즉, 스트리밍 할 미디어 데이터의 선택, 및 사용자 선택들, 네트워크 조건들 및 클라이언트 능력들에 대한 비트 스트림들의 동적 적응화)는 클라이언스 선택들 및 디바이스들에 배타적으로 의존한다.

[0004]

이러한 모델에서, 미디어 프레젠테이션은, 데이터 세그먼트들로, 그리고 프레젠테이션될 시간 설정형 미디어 데이터의 조직화비디오 비트 스트림의 시간D(Media Presentation Description)"라 불리우는 매니페스트로 조직화된다. 특히, 매니페스트는, 데이터 세그먼트들을 다운로드하는데 사용하는 리소스 식별자들을 포함하고, 이러한 데이터 세그먼트들을 선택 및 조합하여 유효 미디어 프레젠테이션을 얻는 컨텍스트를 제공한다. 리소스 식별자들은 통상 HTTP-URL들(Uniform Resource Locator)이고, 아마 바이트 범위들과 조합된 것이다. 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]

또한, 비디오 해상도는, SD(Standard Definition)에서 시작하여 HD(High Definition)로, 및 초-고화질(예를 들어, 4K2K 또는 8K4K, 즉 4,096 x 2,400 화소들 또는 7,680 x 4,320 화소들의 이미지들을 포함하는 비디오)로, 계속 증대하고 있다. 그러나, 수신 및 비디오 디코딩 디바이스들 모두가, 특히 비디오가 초 고화질일 때 최대 해상도의 비디오에 액세스하는 리소스들(예를 들어, 네트워크 액세스 대역폭 또는 CPU(Central Processing Unit))을 갖는 것은 아니며, 모든 사용자들이 이러한 비디오에 액세스할 필요는 없다. 이러한 정황에서, 관심 영역들(Regions-Of-Interest)(ROIs)에만 액세스하는, 즉 전체 비디오 시퀀스 중 일부 공간 서브 파트들에만 액세스하는 능력을 제공하는 것이 특히 유리하다.

[0006]

비디오에 속하는 프레임들의 공간 서브 파트들에 액세스하는 공지의 메커니즘은, 비디오의 각 프레임을 일반적으로 파일이라고 하는 독립적으로 디코드 가능한 공간 영역들의 배열로서 조직화하는 것에 있다. SVC(Scalable Video Coding) 또는 HEVC(High Efficiency Video Coding)와 같은 일부 비디오 포맷들은 파일 화질에 대한 지원을 제공한다. 사용자 정의의 ROI는 하나 또는 여러 연속적인 파일들을 커버할 수 있다.

[0007]

따라서, 사용자 선택의 ROI들을 HTTP 프로토콜에 따라 스트리밍하기 위해서는, 하나 이상의 파일들에 대한 공간 액세스들을 가능하게 하고 액세스된 파일들의 조합을 가능하게 하는 방식으로, 인코딩된 비디오 비트 스트림의 시간미디어 데이터 비트 스트림의 캡슐화를 제공하는 것이 중요하다.

[0008]

인코딩된 비디오 비트 스트림들은 일반적으로, 완전한 프레임들에 대응하는 연속적인 시간적 샘플들의 세트로서 구성되고, 이러한 시간적 샘플들은 디코딩 순서의 함수로서 조직화된다는 점이 상기되어야 한다. 파일 포맷들은 이러한 인코딩된 비트 스트림을 캡슐화하고 서술하는데 사용된다.

[0009]

설명의 목적상, 국제 표준 기구 베이스 미디어 파일 포맷(International Standard Organization Base Media File Format)(ISO BMFF)은, 네트워크를 통한 또는 다른 비트 스트림 전달 메커니즘을 통한 로컬 저장 또는 송신을 위해 인코딩된 시간 설정형 미디어 데이터 비트 스트림들을 기술하는 잘 알려진 유연하고 확장 가능한 포맷이다. 이러한 파일 포맷은 객체 지향형이다. 이는, 순차적으로 또는 계층적으로 조직화되고, 타이밍 및 구조 파

라미터들과 같은 인코딩된 시간 설정형 미디어 데이터 비트 스트림의 파라미터들을 정의하는, 박스들이라 불리우는 구축 블럭들로 구성된다. 이러한 파일 포맷에 따르면, 시간 설정형 미디어 데이터 비트 스트림은, 트랙 박스(track box)라 하는 다른 데이터 구조로 정의되는 mdat box라 하는 데이터 구조에 포함된다. 이러한 트랙은, 하나의 샘플이 단일 타임스탬프와 관련된 모든 데이터, 즉, 단일 프레임과 관련된 모든 데이터 또는 동일 타임스탬프를 공유하는 여러 프레임들과 관련된 모든 데이터에 대응하는 샘플들의 시간 설정형 시퀀스를 나타낸다.

[0010] SVC 포맷의 비디오와 같은 스케일 가능한 비디오에 대해, 계층화(layered) 미디어 데이터 구성은 다수의 종속 트랙들을 사용함으로써 효과적으로 표현될 수 있고, 각각의 트랙은 특정 레벨의 스케일 가능성으로 해당 비디오를 표현한다. 트랙들 사이의 데이터 중복을 회피하기 위해서, 추출자들이 사용될 수 있다. 표준 파일 포맷에 따르면, 추출자는, 다른 비트 스트림들로부터 NAL(Network Abstraction Layer) 유닛들의 효과적인 추출을 가능하게 하며 비트 스트림에 직접 포함되는 데이터 구조이다. 예를 들어, 항상 계층(enhancement layer) 트랙의 비트 스트림은 베이스 계층(base layer) 트랙으로부터 NAL 유닛들을 참조하는 추출자들을 포함할 수 있다. 그리고 나중에, 이러한 항상 계층 트랙이 파일 포맷으로부터 추출될 때, 추출자들은 그들이 참조하고 있는 데이터로 대체되어야 한다.

[0011] 이러한 메커니즘들을 내포한 ISO BMFF를 사용할 때, 서브 정보를 서술하고, 이러한 서브 정보로의 액세스를 용이하게 하거나 또는 비트 스트림들을 다수 세그먼트들로 효과적으로 조직화하는 여러 전략들이 채택될 수 있다.

[0012] 예를 들어, "Implications of the ISO Base Media File Format on Adaptive HTTP Streaming of H.264/SVC"라는 제목의 문헌에서, Kofler 등의 저자들은, ISO BMFF의 가능성들뿐만 아니라 제한들을 고려하여 HTTP 스트리밍을 위한 스케일 가능한 비디오 비트 스트림(H264/SVC)을 조직화하는 3가지 상이한 전략들을 제시한다:

a) (트랙 정의들을 포함하여) 모든 ISO BMFF 메타데이터를 포함하는 파일 타입 박스 "ftyp" 및 영화 박스 "moov"를 포함하는 특정 파일 헤더를 포함하는 단일 파일, 이러한 단일 파일은 또한 전체 인코딩된 비트 스트림을 포함하는 단일 mdat 박스를 포함함. 이러한 조직화는 로컬 저장에는 적합하지만 클라이언트가 전체 비트 스트림의 일부만을 필요로 할 수 있는 HTTP 스트리밍에는 적응되지 않는다.

b) 프래그먼트화(fragmentation)에 적합한 다수의 moof/mdat 박스들을 포함하는 단일 파일. 이러한 포맷은 프로그레시브 다운로드를 감안한다. moof 박스는 프래그먼트 레벨에서 moov 박스와 등가이다. 이러한 스킴에 따라, 프래그먼트화된 미디어 파일을 사용하여, 스케일 가능한 비트 스트림은, 비디오를 상이한 스케일 가능성 레벨들로 표현하는 다수의 종속 트랙들로 나뉜다. 다른 트랙들로부터의 NAL 유닛들을 참조하는데 추출자들이 사용된다. 파일 당 트랙이 사용되는 경우, 모든 어드레스 가능성 트랙들은 미리 준비되어야 하고 트랙들은 독립적으로 선택될 수 없다. 여러 파일들이 디스플레이되어야 하면, 여러 비트 스트림들이 디코딩되어야 하고 베이스 계층이 여러 번 디코딩된다.

c) 각각의 파일이 그 자신의 URL에 의해 액세스 가능하고 독립적으로 다운로드 가능한, 다수의 세그먼트 파일들. 각각의 세그먼트는 통상적으로 세그먼트 타입 박스(styp)로 구성되고, 이는 일종의 파일 헤더, 옵션형 세그먼트 인덱스 박스(sidx) 및 하나 또는 다수의 프래그먼트들로서 역할을 한다. 다시, 각각의 프래그먼트는 moof 및 mdat 박스로 구성된다. 이러한 스킴에 따라, 프래그먼트화된 미디어 파일을 사용하여, 각각의 트랙은 스케일 가능성의 한 레벨에 관한 관련된 비트 스트림과 함께 그 자신의 세그먼트에 저장된다. 필요하다면, 종속 트랙들로부터 요구된 비트 스트림을 참조하는데 추출자들이 사용된다. 이러한 코딩 스킴은 트랙들을 독립적으로 스트리밍하는데 특히 적합하다. 이는 DASH 표준에 잘 적응되지만, 여러 비트 스트림들이 디코딩되어야 하고, 이에 따라, 트랙 당 하나의 디코더가 요구되기 때문에, 파일 스트리밍에는 적합하지 않다. 더욱이, 둘 이상의 파일을 선택할 때 베이스 계층의 비트 스트림의 잠재적인 중복이 존재한다.

[0016] 공간 파일들에 적용될 때, 이러한 전략들 중 어느 것도 HTTP 스트리밍의 컨텍스트에서 특정 파일들로의 효율적인 액세스를 가능하게 하지 않는다. 실제로 기존 파일 포맷 정의에 의하면, 인코딩된 비트 스트림에서 다수의 비연속 바이트 범위들에 여전히 액세스할 필요가 있거나, 또는 주어진 시간 간격에 대응하는 여러 프레임들의 공간 파일들을 디스플레이하기 위해 비트 스트림 중복을 초래할 것이다.

[0017] 이러한 문제점들을 해결하기 위해, 공간 파일들에 적합한 효율적인 데이터 조직화 및 트랙 서술 스킴이 제공되며, 이는 클라이언트 애플리케이션에 의해 어떠한 트랙 조합이 선택되더라도, ISO BMFF 파일의 결과가 항상 비디오 디코더에 대해 유효한 비디오 기본 비트 스트림에 이르는 것을 보장한다.

발명의 내용

해결하려는 과제

- [0018] 이러한 제약들과 직면하여, 본 발명자들은, 서버에서 분할된 시간 설정형 미디어 데이터를 캡슐화하고, 미디어 파일들에 캡슐화된 분할된 시간 설정형 미디어 데이터로부터 시간 설정형 미디어 데이터 비트 스트림을 제공하는 방법 및 디바이스를 제공한다.
- [0019] 본 발명의 광범위한 목적은 위에 설명된 바와 같은 종래 기술의 단점을 해결하는 것이다.

과제의 해결 수단

- [0020] 본 발명의 양상에 따르면, 서버에서 분할된 시간 설정형 미디어 데이터를 캡슐화하는 방법으로서, 상기 분할된 시간 설정형 미디어 데이터는 시간 설정형 샘플들을 포함하고, 각각의 시간 설정형 샘플은 복수의 서브샘플들을 포함하며, 이 방법은:
- [0021] 상기 시간 설정형 샘플들 중 하나의 상기 복수의 서브샘플들 중에서 적어도 하나의 서브샘플을 선택하는 단계;
- [0022] 각각의 선택된 서브샘플에 대해, 상기 선택된 서브샘플 및 다른 시간 설정형 샘플들 각각의 하나의 대응 서브샘플을 포함하는 하나의 분할 트랙을 생성하는 단계;
- [0023] 적어도 하나의 종속성 박스를 생성하는 단계 - 각각의 종속성 박스는 분할 트랙과 관련이 있고 다른 생성된 분할 트랙들 중 하나 이상의 분할 트랙으로의 적어도 하나의 참조를 포함하고, 적어도 하나의 참조는 다른 분할 트랙들 중 상기 하나 이상의 분할 트랙에 관련하여 디코딩 순서 종속성을 나타냄 -; 및
- [0024] 분할 트랙들 각각을 적어도 하나의 미디어 파일에 독립적으로 캡슐화하는 단계를 포함한다.
- [0025] 따라서, 본 발명은 상이한 부분들의, 특히 상이한 서브샘플들(또는 "타일들")의 결합을 가능하게 한다. 그것은 클라이언트 디바이스에 의해 다운로드된 부분들 또는 타일들의 선택된 세트가 무엇이든 클라이언트에 의한 과정 시에 유효 파일 포맷 및 기본 스트림의 생성과 관련된다.
- [0026] 그러므로, 본 발명의 방법은 유용한 데이터만이 클라이언트 디바이스에 송신될 것을 필요로 한다는 것을 고려할 때, 독립적인 부분들 또는 타일들의 효율적인 스트리밍에 적합하다. 그것은 (샘플 기반의 바이트-범위 요청들과 비교하여) 송신 오버헤드를 감소시키고, MPEG 표준에 통합될 수 있다는 것을 고려할 때, 둘 이상의 부분 또는 타일 선택(즉, 임의의 관심 영역)을 스트리밍하기에 적합하다.
- [0027] 일 실시예에서, 본 방법은 적어도 하나의 다른 종속성 박스를 생성하는 단계를 더 포함하고, 각각의 다른 종속성 박스는 분할 트랙과 관련이 있고 다른 생성된 분할 트랙들 중 하나 이상의 분할 트랙으로의 적어도 하나의 참조를 포함하고, 적어도 하나의 다른 종속성 박스의 참조들 각각은 디코딩 종속성을 나타낸다.
- [0028] 일 실시예에서, 분할된 시간 설정형 미디어 데이터는 타일형 시간 설정형 미디어 데이터이고, 서브샘플들은 공간 서브샘플들이고, 적어도 하나의 분할 트랙은 적어도 하나의 타일 트랙이다.
- [0029] 일 실시예에서, 본 방법은, 고려되는 분할 트랙에 관한 초기화 파라미터들을 포함하는, 각각의 분할 트랙에 관련된 트랙 프래그먼트 헤더를 생성하는 단계를 더 포함하고, 적어도 하나의 트랙 프래그먼트 헤더는 다른 분할 트랙들에 관련하여 고려되는 분할 트랙의 디코딩 순서 종속성을 가진 종속성 박스를 포함한다.
- [0030] 일 실시예에서, 본 방법은 적어도 하나의 생성된 분할 트랙으로의 적어도 하나의 참조를 포함하는 적어도 하나의 참조 트랙을 생성하는 단계 및 적어도 하나의 참조 트랙을 미디어 세그먼트 파일에 독립적으로 캡슐화하는 단계를 더 포함하고, 적어도 하나의 참조 트랙을 캡슐화한 결과인 미디어 세그먼트 파일은 정의들을 포함하는 메타데이터를 저장하기 위한 데이터 구조를 포함하고, 정의들은 적어도 하나의 생성된 합성 트랙과 관련된 참조된 분할 트랙들의 리스트를 포함하고, 참조 타입이 적어도 하나의 생성된 합성 트랙과 관련된 참조된 분할 트랙들 각각과 관련된다.
- [0031] 일 실시예에서, 본 방법은 복수의 분할 트랙에 관한 초기화 데이터를 포함하는 적어도 하나의 초기화 데이터 트랙을 생성하는 단계를 더 포함하고, 적어도 하나의 초기화 데이터 트랙은 적어도 하나의 미디어 세그먼트 파일에 독립적으로 캡슐화되고, 상기 초기화 데이터 트랙은 적어도 하나의 종속성 박스를 포함한다.
- [0032] 일 실시예에서, 생성된 트랙들 각각을 캡슐화한 결과인 적어도 하나의 미디어 파일은 국제 표준 기구

(International Standard Organization)에 의해 정의된 베이스 미디어 파일 포맷(Base Media File Format) 및 DASH(Dynamic Adaptive Streaming over HTTP) 포맷과 호환된다.

[0033] 본 발명의 다른 양상에 따르면, 클라이언트 디바이스에서, 적어도 하나의 미디어 파일에 캡슐화된 분할된 시간 설정형 미디어 데이터로부터 시간 설정형 미디어 데이터 비트 스트림을 제공하는 방법으로서, 분할된 시간 설정형 미디어 데이터는 시간 설정형 샘플들을 포함하고, 각각의 시간 설정형 샘플은 복수의 서브샘플을 포함하고, 적어도 하나의 미디어 파일은 적어도 하나의 종속성 박스 및 적어도 하나의 분할 트랙을 포함하고, 적어도 하나의 분할 트랙 각각은 시간 설정형 샘플들 중 하나의 복수의 서브샘플 중에서 선택된 하나의 서브샘플을 포함하고 다른 시간 설정형 샘플들 각각의 하나의 대응하는 공간 서브샘플을 포함하고, 각각의 종속성 박스는 분할 트랙과 관련이 있고 다른 생성된 분할 트랙들 중 하나 이상의 분할 트랙으로의 적어도 하나의 참조를 포함하고, 적어도 하나의 참조는 다른 분할 트랙들 중 하나 이상의 분할 트랙에 관련하여 디코딩 순서 종속성을 나타내고, 이 방법은:

[0034] 디스플레이될 적어도 하나의 분할 트랙을 선택하는 단계;

[0035] 반복적으로 분할 트랙들을 결정하고 순서화하는 단계 - 분할 트랙을 선택하는 단계는 대응하는 종속성 박스의 적어도 하나의 디코딩 순서 종속성에 기초함 -; 및

[0036] 디스플레이될 선택된 트랙을 역캡슐화하는(de-encapsulating) 단계

[0037] 를 포함하는 방법이 제공된다.

[0038] 따라서, 본 발명은 상이한 부분들의, 특히 상이한 서브샘플들(또는 "타일들")의 결합을 가능하게 한다. 그것은 클라이언트 디바이스에 의해 다운로드된 부분들 또는 타일들의 선택된 세트가 무엇이든 클라이언트에 의한 파싱 시에 유효 파일 포맷 및 기본 스트림의 생성과 관련된다.

[0039] 그러므로, 본 발명의 방법은 유용한 데이터만이 클라이언트 디바이스에 송신될 것을 필요로 한다는 것을 고려할 때, 독립적인 부분들 또는 타일들의 효율적인 스트리밍에 적합하다. 그것은 (샘플 기반의 바이트-범위 요청들과 비교하여) 송신 오버헤드를 감소시키고, MPEG 표준에 통합될 수 있다는 것을 고려할 때, 둘 이상의 부분 또는 타일 선택(즉, 임의의 관심 영역)을 스트리밍하기에 적합하다.

[0040] 일 실시예에서, 적어도 하나의 미디어 파일은 적어도 하나의 다른 종속성 박스를 더 포함하고, 각각의 다른 종속성 박스는 분할 트랙과 관련이 있고 다른 생성된 분할 트랙들 중 하나 이상의 분할 트랙으로의 적어도 하나의 참조를 포함하고, 적어도 하나의 다른 종속성 박스의 참조들 각각은 디코딩 종속성을 나타내고, 반복적으로 분할 트랙들을 결정하고 순서화하는 단계 및 분할 트랙을 선택하는 단계는 대응하는 다른 종속성 박스의 적어도 하나의 디코딩 종속성에 추가로 기초한다.

[0041] 일 실시예에서, 본 방법은, 고려되는 분할 트랙에 관한 초기화 파라미터들을 포함하는, 각각의 분할 트랙에 관련된 트랙 프래그먼트 헤더를 생성하는 단계를 더 포함하고, 적어도 하나의 트랙 프래그먼트 헤더는 다른 분할 트랙들에 관련하여 고려되는 분할 트랙의 디코딩 순서 종속성을 가진 종속성 박스를 포함한다.

[0042] 일 실시예에서, 복수의 분할 트랙에 관한 초기화 데이터를 포함하는 하나의 초기화 데이터 트랙이 적어도 하나의 미디어 세그먼트 파일에 독립적으로 캡슐화되고, 초기화 데이터 트랙은 적어도 하나의 종속성 박스를 포함하고, 본 방법은 초기화 데이터 트랙을 선택하고 그것을 역캡슐화하는 단계를 더 포함한다.

[0043] 일 실시예에서, 적어도 하나의 생성된 분할 트랙으로의 적어도 하나의 참조를 포함하는 적어도 하나의 참조 트랙이 적어도 하나의 미디어 세그먼트 파일에 독립적으로 캡슐화되고, 적어도 하나의 참조 트랙을 캡슐화한 결과인 미디어 세그먼트 파일은 정의들을 포함하는 메타데이터를 저장하기 위한 데이터 구조를 포함하고, 정의들은 적어도 하나의 생성된 합성 트랙과 관련된 참조된 분할 트랙들의 리스트를 포함하고, 참조 타입이 적어도 하나의 생성된 합성 트랙과 관련된 참조된 분할 트랙들 각각과 관련되고, 반복적으로 분할 트랙들을 결정하고 순서화하는 단계 및 분할 트랙을 선택하는 단계는 참조 타입이 적어도 하나의 생성된 합성 트랙과 관련된 참조된 분할 트랙들 각각과 관련되는 것에 추가로 기초한다.

[0044] 본 발명의 다른 양상에 따르면 서버에서 분할된 시간 설정형 미디어 데이터를 캡슐화하는 디바이스로서, 분할된 시간 설정형 미디어 데이터는 시간 설정형 샘플들을 포함하고, 각각의 시간 설정형 샘플은 복수의 서브샘플을 포함하고, 이 디바이스는:

[0045] 상기 시간 설정형 샘플들 중 하나의 상기 복수의 서브샘플들 중에서 적어도 하나의 서브샘플을 선택하는 단계;

- [0046] 각각의 선택된 서브샘플에 대해, 상기 선택된 서브샘플 및 다른 시간 설정형 샘플들 각각의 하나의 대응 서브샘플을 포함하는 하나의 분할 트랙을 생성하는 단계;
- [0047] 적어도 하나의 종속성 박스를 생성하는 단계 - 각각의 종속성 박스는 분할 트랙과 관련이 있고 다른 생성된 분할 트랙들 중 하나 이상의 분할 트랙으로의 적어도 하나의 참조를 포함하고, 적어도 하나의 참조는 다른 분할 트랙들 중 상기 하나 이상의 분할 트랙에 관련하여 디코딩 순서 종속성을 나타냄 -; 및
- [0048] 분할 트랙들 각각을 적어도 하나의 미디어 파일에 독립적으로 캡슐화하는 단계
- [0049] 를 수행하도록 구성된 적어도 하나의 마이크로프로세서를 포함하는 디바이스가 제공된다.
- [0050] 따라서, 본 발명은 상이한 부분들의, 특히 상이한 서브샘플들(또는 "타일들")의 결합을 가능하게 한다. 그것은 클라이언트 디바이스에 의해 다운로드된 부분들 또는 타일들의 선택된 세트가 무엇이든 클라이언트에 의한 파싱 시에 유효 파일 포맷 및 기본 스트림의 생성과 관련된다.
- [0051] 그러므로, 본 발명의 방법은 유용한 데이터만이 클라이언트 디바이스에 송신될 것을 필요로 한다는 것을 고려할 때, 독립적인 부분들 또는 타일들의 효율적인 스트리밍에 적합하다. 그것은 (샘플 기반의 바이트-범위 요청들과 비교하여) 송신 오버헤드를 감소시키고, MPEG 표준에 통합될 수 있다는 것을 고려할 때, 둘 이상의 부분 또는 타일 선택(즉, 임의의 관심 영역)을 스트리밍하기에 적합하다.
- [0052] 일 실시예에서, 적어도 하나의 마이크로프로세서는 적어도 하나의 다른 종속성 박스를 생성하는 단계를 수행하도록 추가로 구성되고, 각각의 다른 종속성 박스는 분할 트랙과 관련이 있고 다른 생성된 분할 트랙들 중 하나 이상의 분할 트랙으로의 적어도 하나의 참조를 포함하고, 적어도 하나의 다른 종속성 박스의 참조들 각각은 디코딩 종속성을 나타낸다.
- [0053] 일 실시예에서, 분할된 시간 설정형 미디어 데이터는 타일형 시간 설정형 미디어 데이터이고, 서브샘플들은 공간 서브샘플들이고, 적어도 하나의 분할 트랙은 적어도 하나의 타일 트랙이다.
- [0054] 일 실시예에서, 적어도 하나의 마이크로프로세서는, 고려되는 분할 트랙에 관한 초기화 파라미터들을 포함하는, 각각의 분할 트랙에 관련된 트랙 프래그먼트 헤더를 생성하는 단계를 수행하도록 추가로 구성되고, 적어도 하나의 트랙 프래그먼트 헤더는 다른 분할 트랙들에 관련하여 고려되는 분할 트랙의 디코딩 순서 종속성을 가진 종속성 박스를 포함한다.
- [0055] 일 실시예에서, 적어도 하나의 마이크로프로세서는 적어도 하나의 생성된 분할 트랙으로의 적어도 하나의 참조를 포함하는 적어도 하나의 참조 트랙을 생성하는 단계 및 적어도 하나의 참조 트랙을 미디어 세그먼트 파일에 독립적으로 캡슐화하는 단계를 수행하도록 추가로 구성되고, 적어도 하나의 참조 트랙을 캡슐화한 결과인 미디어 세그먼트 파일은 정의들을 포함하는 메타데이터를 저장하기 위한 데이터 구조를 포함하고, 정의들은 적어도 하나의 생성된 합성 트랙과 관련된 참조된 분할 트랙들의 리스트를 포함하고, 참조 태입이 적어도 하나의 생성된 합성 트랙과 관련된 참조된 분할 트랙들 각각과 관련된다.
- [0056] 일 실시예에서, 적어도 하나의 마이크로프로세서는 복수의 분할 트랙에 관한 초기화 데이터를 포함하는 적어도 하나의 초기화 데이터 트랙을 생성하는 단계를 수행하도록 추가로 구성되고, 적어도 하나의 초기화 데이터 트랙은 적어도 하나의 미디어 세그먼트 파일에 독립적으로 캡슐화되고, 상기 초기화 데이터 트랙은 적어도 하나의 종속성 박스를 포함한다.
- [0057] 일 실시예에서, 생성된 트랙들 각각을 캡슐화한 결과인 적어도 하나의 미디어 파일은 국제 표준 기구 (International Standard Organization)에 의해 정의된 베이스 미디어 파일 포맷(Base Media File Format) 및 DASH(Dynamic Adaptive Streaming over HTTP) 포맷과 호환된다.
- [0058] 본 발명의 다른 양상에 따르면 전술한 디바이스를 포함하는 비디오 인코더가 제공된다.
- [0059] 본 발명의 다른 양상에 따르면, 클라이언트 디바이스에서, 적어도 하나의 미디어 파일에 캡슐화된 분할된 시간 설정형 미디어 데이터로부터 시간 설정형 미디어 데이터 비트 스트림을 제공하는 디바이스로서, 분할된 시간 설정형 미디어 데이터는 시간 설정형 샘플들을 포함하고, 각각의 시간 설정형 샘플은 복수의 서브샘플을 포함하고, 적어도 하나의 미디어 파일은 적어도 하나의 종속성 박스 및 적어도 하나의 분할 트랙을 포함하고, 적어도 하나의 분할 트랙 각각은 시간 설정형 샘플들 중 하나의 복수의 서브샘플 중에서 선택된 하나의 서브샘플을 포함하고 다른 시간 설정형 샘플들 각각의 하나의 대응하는 공간 서브샘플을 포함하고, 각각의 종속성 박스는 분할 트랙과 관련이 있고 다른 생성된 분할 트랙들 중 하나 이상의 분할 트랙으로의 적어도 하나의 참조를

포함하고, 적어도 하나의 참조는 다른 분할 트랙들 중 하나 이상의 분할 트랙에 관련하여 디코딩 순서 종속성을 나타내고, 이 디바이스는:

- [0060] 디스플레이될 적어도 하나의 분할 트랙을 선택하는 단계;
- [0061] 반복적으로 분할 트랙들을 결정하고 순서화하는 단계 - 분할 트랙을 선택하는 단계는 대응하는 종속성 박스의 적어도 하나의 디코딩 순서 종속성에 기초함 -; 및
- [0062] 디스플레이될 선택된 트랙을 역캡슐화하는 단계
- [0063] 를 수행하도록 구성된 적어도 하나의 마이크로프로세서를 포함하는 디바이스가 제공된다.
- [0064] 따라서, 본 발명은 상이한 부분들의, 특히 상이한 서브샘플들(또는 "타일들")의 결합을 가능하게 한다. 그것은 클라이언트 디바이스에 의해 다운로드된 부분들 또는 타일들의 선택된 세트가 무엇이든 클라이언트에 의한 파싱 시에 유효 파일 포맷 및 기본 스트림의 생성과 관련된다.
- [0065] 그러므로, 본 발명의 방법은 유용한 데이터만이 클라이언트 디바이스에 송신될 것을 필요로 한다는 것을 고려할 때, 독립적인 부분들 또는 타일들의 효율적인 스트리밍에 적합하다. 그것은 (샘플 기반의 바이트-범위 요청들과 비교하여) 송신 오버헤드를 감소시키고, MPEG 표준에 통합될 수 있다는 것을 고려할 때, 둘 이상의 부분 또는 타일 선택(즉, 임의의 관심 영역)을 스트리밍하기에 적합하다.
- [0066] 일 실시예에서, 적어도 하나의 미디어 파일은 적어도 하나의 다른 종속성 박스를 더 포함하고, 각각의 다른 종속성 박스는 분할 트랙과 관련이 있고 다른 생성된 분할 트랙들 중 하나 이상의 분할 트랙으로의 적어도 하나의 참조를 포함하고, 적어도 하나의 다른 종속성 박스의 참조들 각각은 디코딩 종속성을 나타내고, 적어도 하나의 마이크로프로세서는 반복적으로 분할 트랙들을 결정하고 순서화하는 단계 및 분할 트랙을 선택하는 단계가 대응하는 다른 종속성 박스의 적어도 하나의 디코딩 종속성에 추가로 기초하도록 추가로 구성된다.
- [0067] 일 실시예에서, 적어도 하나의 마이크로프로세서는, 고려되는 분할 트랙에 관한 초기화 파라미터들을 포함하는, 각각의 분할 트랙에 관련된 트랙 프래그먼트 헤더를 획득하는 단계를 수행하도록 추가로 구성되고, 적어도 하나의 트랙 프래그먼트 헤더는 다른 분할 트랙들에 관련하여 고려되는 분할 트랙의 디코딩 순서 종속성을 가진 종속성 박스를 포함한다.
- [0068] 일 실시예에서, 복수의 분할 트랙에 관한 초기화 데이터를 포함하는 하나의 초기화 데이터 트랙이 적어도 하나의 미디어 세그먼트 파일에 독립적으로 캡슐화되고, 초기화 데이터 트랙은 적어도 하나의 종속성 박스를 포함하고, 본 방법은 초기화 데이터 트랙을 선택하고 그것을 역캡슐화하는 단계를 더 포함한다.
- [0069] 일 실시예에서, 적어도 하나의 생성된 분할 트랙으로의 적어도 하나의 참조를 포함하는 적어도 하나의 참조 트랙이 적어도 하나의 미디어 세그먼트 파일에 독립적으로 캡슐화되고, 적어도 하나의 참조 트랙을 캡슐화한 결과인 미디어 세그먼트 파일은 정의들을 포함하는 메타데이터를 저장하기 위한 데이터 구조를 포함하고, 정의들은 적어도 하나의 생성된 합성 트랙과 관련된 참조된 분할 트랙들의 리스트를 포함하고, 참조 타입이 적어도 하나의 생성된 합성 트랙과 관련된 참조된 분할 트랙들 각각과 관련되고, 적어도 하나의 마이크로프로세서는 반복적으로 분할 트랙들을 결정하고 순서화하는 단계 및 분할 트랙을 선택하는 단계가 참조 타입이 적어도 하나의 생성된 합성 트랙과 관련된 참조된 분할 트랙들 각각과 관련되는 것에 추가로 기초하도록 추가로 구성된다.
- [0070] 본 발명의 다른 양상에 따르면 전술한 디바이스를 포함하는 비디오 디코더가 제공된다.
- [0071] 본 발명의 다른 양상에 따르면 클라이언트의 디바이스에 스트리밍되도록 서버에서 분할된 시간 설정형 미디어 데이터를 캡슐화하는 방법으로서, 상기 분할된 시간 설정형 미디어 데이터는 시간 설정형 샘플들을 포함하고, 각각의 시간 설정형 샘플은 복수의 서브샘플을 포함하고, 이 방법은:

 - [0072] 상기 서브샘플들의 적어도 일부가 서브샘플들의 그룹에 속하도록 몇몇의 상기 서브샘플들을 그룹화하여 적어도 하나의 그룹을 형성하는 단계;
 - [0073] 상기 시간 설정형 샘플들 중 하나의 상기 복수의 서브샘플 중에서 동일한 그룹의 적어도 2개의 서브샘플을 선택하는 단계;
 - [0074] 상기 적어도 2개의 선택된 서브샘플을 포함하는 적어도 하나의 트랙을 생성하는 단계;
 - [0075] 상기 적어도 하나의 생성된 트랙에 대해, 상기 적어도 하나의 트랙을 생성하기 위해 선택된 상기 서브샘플들 각각에 대해 하나의 서브 트랙 박스를 생성하는 단계 - 상기 생성된 서브 트랙 박스들 중 적어도 하나는 상기 선

택된 서브샘플들에 공통인 정의들 및 특성들을 포함함 -; 및

[0076] 상기 생성된 트랙들 각각을 적어도 하나의 미디어 파일에 독립적으로 캡슐화하는 단계

[0077] 를 포함하는 방법이 제공된다.

[0078] 따라서, 본 발명은 상이한 부분들의, 특히 상이한 서브샘플들(또는 "타일들")의 결합을 가능하게 한다. 그것은 클라이언트 디바이스에 의해 다운로드된 부분들 또는 타일들의 선택된 세트가 무엇이든 클라이언트에 의한 과정 시에 유효 파일 포맷 및 기본 스트림의 생성과 관련된다.

[0079] 그러므로, 본 발명의 방법은 유용한 데이터만이 클라이언트 디바이스에 송신될 것을 필요로 한다는 것을 고려할 때, 독립적인 부분들 또는 타일들의 효율적인 스트리밍에 적합하다. 그것은 (샘플 기반의 바이트-범위 요청들과 비교하여) 송신 오버헤드를 감소시키고, MPEG 표준에 통합될 수 있다는 것을 고려할 때, 둘 이상의 부분 또는 타일 선택(즉, 임의의 관심 영역)을 스트리밍하기에 적합하다.

[0080] 일 실시예에서, 상기 생성된 서브 트랙 박스들은 ISO/IEC 14496-12에 따라 정의된다.

[0081] 일 실시예에서, 상기 생성된 서브 트랙 박스들 중 적어도 하나는 상기 생성된 서브 트랙 박스들 중 상기 적어도 하나가 타일링 정보를 제공하는 특정한 트랙과 관련됨을 나타내는 정보를 포함한다.

[0082] 일 실시예에서, 본 방법은 샘플들의 또는 서브샘플들의 그룹을 적어도 하나의 타일에 링크시키는 단계를 더 포함한다.

[0083] 일 실시예에서, 상기 생성된 서브 트랙 박스들 중 적어도 하나와 관련된 적어도 하나의 서브 트랙 내에 타일들의 그룹이 인코딩된다.

[0084] 일 실시예에서, 상기 생성된 서브 트랙 박스들 중 적어도 하나는 상기 생성된 서브 트랙 박스들 중 적어도 하나와 관련된 서브 트랙 내에 인코딩된 타일들의 수를 나타내는 정보를 포함한다.

[0085] 일 실시예에서, 상기 생성된 트랙들 각각을 캡슐화한 결과인 미디어 파일들은, 대응하는 트랙의 정의들을 포함하는 메타데이터를 저장하기 위한 데이터 구조를 포함하여, 대응하는 트랙 내에 포함된 상기 서브 트랙 박스들의 정의들을 포함한다.

[0086] 본 발명의 다른 양상에 따르면, 클라이언트 디바이스에서, 적어도 하나의 미디어 파일에 캡슐화된 분할된 시간 설정형 미디어 데이터로부터 시간 설정형 미디어 데이터 비트 스트림을 제공하는 방법으로서, 상기 분할된 시간 설정형 미디어 데이터는 시간 설정형 샘플들을 포함하고, 각각의 시간 설정형 샘플은 복수의 서브샘플을 포함하고, 상기 서브샘플들 중 적어도 일부가 서브샘플들의 그룹에 속하도록 몇몇의 상기 서브샘플들이 그룹화되어 적어도 하나의 그룹을 형성하고, 상기 적어도 하나의 미디어 파일은 적어도 하나의 트랙을 포함하고, 상기 적어도 하나의 트랙은 상기 시간 설정형 샘플들 중 하나의 상기 복수의 서브샘플 중에서 선택된 동일한 그룹의 적어도 2개의 서브샘플을 포함하고, 상기 적어도 하나의 트랙을 생성하기 위해 선택된 상기 서브샘플들 각각에 대해 하나의 서브 트랙 박스가 생성되고, 상기 생성된 서브 트랙 박스들 중 적어도 하나는 상기 선택된 서브샘플들에 공통인 정의들 및 특성들을 포함하고, 이 방법은:

[0087] 상기 적어도 하나의 미디어 파일로부터 적어도 하나의 서브 트랙 박스를 선택하는 단계;

[0088] 상기 적어도 하나의 선택된 서브 트랙 박스로부터 정의들 및 특성들의 적어도 하나의 세트를 획득하는 단계; 및

[0089] 디코딩될 복수의 서브샘플을 획득하는 단계 - 상기 복수의 서브샘플은 상기 정의들 및 특성들의 적어도 하나의 세트에 종속함 -

[0090] 를 포함하는 방법이 제공된다.

[0091] 따라서, 본 발명은 상이한 부분들의, 특히 상이한 서브샘플들(또는 "타일들")의 결합을 가능하게 한다. 그것은 클라이언트 디바이스에 의해 다운로드된 부분들 또는 타일들의 선택된 세트가 무엇이든 클라이언트에 의한 과정 시에 유효 파일 포맷 및 기본 스트림의 생성과 관련된다.

[0092] 그러므로, 본 발명의 방법은 유용한 데이터만이 클라이언트 디바이스에 송신될 것을 필요로 한다는 것을 고려할 때, 독립적인 부분들 또는 타일들의 효율적인 스트리밍에 적합하다. 그것은 (샘플 기반의 바이트-범위 요청들과 비교하여) 송신 오버헤드를 감소시키고, MPEG 표준에 통합될 수 있다는 것을 고려할 때, 둘 이상의 부분 또는 타일 선택(즉, 임의의 관심 영역)을 스트리밍하기에 적합하다.

- [0093] 일 실시예에서, 상기 생성된 서브 트랙 박스들은 ISO/IEC 14496-12에 따라 정의된다.
- [0094] 일 실시예에서, 상기 서브 트랙 박스들 중 적어도 하나는 상기 서브 트랙 박스들 중 상기 적어도 하나가 타일링 정보를 제공하는 특정한 트랙과 관련됨을 나타내는 정보를 포함한다.
- [0095] 일 실시예에서, 상기 서브 트랙 박스들 중 적어도 하나와 관련된 적어도 하나의 서브 트랙 내에 타일들의 그룹이 인코딩된다.
- [0096] 일 실시예에서, 상기 서브 트랙 박스들 중 적어도 하나는 상기 생성된 서브 트랙 박스들 중 적어도 하나와 관련된 서브 트랙 내에 인코딩된 타일들의 수를 나타내는 정보를 포함한다.
- [0097] 일 실시예에서, 상기 트랙들 각각을 캡슐화한 결과인 미디어 파일들은, 대응하는 트랙의 정의들을 포함하는 메타데이터를 저장하기 위한 레이터 구조를 포함하여, 대응하는 트랙 내에 포함된 상기 서브 트랙 박스들의 정의들을 포함한다.
- [0098] 본 발명의 다른 양상에 따르면 클라이언트의 디바이스에 스트리밍되도록 서버에서 분할된 시간 설정형 미디어 데이터를 캡슐화하는 방법으로서, 상기 분할된 시간 설정형 미디어 데이터는 시간 설정형 샘플들을 포함하고, 각각의 시간 설정형 샘플은 복수의 서브샘플을 포함하고, 이 방법은:
- [0099] 상기 서브샘플들의 적어도 일부가 서브샘플들의 그룹에 속하도록 몇몇의 상기 서브샘플들을 그룹화하여 적어도 하나의 그룹을 형성하는 단계;
- [0100] 상기 시간 설정형 샘플들 중 하나의 상기 복수의 서브샘플 중에서 동일한 그룹의 적어도 2개의 서브샘플을 선택하는 단계;
- [0101] 상기 적어도 2개의 선택된 서브샘플을 포함하는 적어도 하나의 트랙을 생성하는 단계; 및
- [0102] 상기 생성된 트랙들 각각을 적어도 하나의 미디어 파일에 독립적으로 캡슐화하는 단계
- [0103] 를 포함하는 방법이 제공된다.
- [0104] 따라서, 본 발명은 상이한 부분들의, 특히 상이한 서브샘플들(또는 "타일들")의 결합을 가능하게 한다. 그것은 클라이언트 디바이스에 의해 다운로드된 부분들 또는 타일들의 선택된 세트가 무엇이든 클라이언트에 의한 파싱 시에 유효 파일 포맷 및 기본 스트림의 생성과 관련된다.
- [0105] 그러므로, 본 발명의 방법은 유용한 데이터만이 클라이언트 디바이스에 송신될 것을 필요로 한다는 것을 고려할 때, 독립적인 부분들 또는 타일들의 효율적인 스트리밍에 적합하다. 그것은 (샘플 기반의 바이트-범위 요청들과 비교하여) 송신 오버헤드를 감소시키고, MPEG 표준에 통합될 수 있다는 것을 고려할 때, 둘 이상의 부분 또는 타일 선택(즉, 임의의 관심 영역)을 스트리밍하기에 적합하다.
- [0106] 본 발명의 다른 양상에 따르면 클라이언트의 디바이스에 스트리밍되도록 서버에서 분할된 시간 설정형 미디어 데이터를 캡슐화하는 디바이스로서, 상기 분할된 시간 설정형 미디어 데이터는 시간 설정형 샘플들을 포함하고, 각각의 시간 설정형 샘플은 복수의 서브샘플을 포함하고, 이 디바이스는:
- [0107] 상기 서브샘플들의 적어도 일부가 서브샘플들의 그룹에 속하도록 몇몇의 상기 서브샘플들을 그룹화하여 적어도 하나의 그룹을 형성하는 단계;
- [0108] 상기 시간 설정형 샘플들 중 하나의 상기 복수의 서브샘플 중에서 동일한 그룹의 적어도 2개의 서브샘플을 선택하는 단계;
- [0109] 상기 적어도 2개의 선택된 서브샘플을 포함하는 적어도 하나의 트랙을 생성하는 단계;
- [0110] 상기 적어도 하나의 생성된 트랙에 대해, 상기 적어도 하나의 트랙을 생성하기 위해 선택된 상기 서브샘플들 각각에 대해 하나의 서브 트랙 박스를 생성하는 단계 - 상기 생성된 서브 트랙 박스들 중 적어도 하나는 상기 선택된 서브샘플들에 공통인 정의들 및 특성들을 포함함 -; 및
- [0111] 상기 생성된 트랙들 각각을 적어도 하나의 미디어 파일에 독립적으로 캡슐화하는 단계
- [0112] 를 수행하도록 구성된 적어도 하나의 마이크로프로세서를 포함하는 디바이스가 제공된다.
- [0113] 따라서, 본 발명은 상이한 부분들의, 특히 상이한 서브샘플들(또는 "타일들")의 결합을 가능하게 한다. 그것은 클라이언트 디바이스에 의해 다운로드된 부분들 또는 타일들의 선택된 세트가 무엇이든 클라이언트에 의한 파싱

시에 유효 파일 포맷 및 기본 스트림의 생성과 관련된다.

[0114] 그러므로, 본 발명의 방법은 유용한 데이터만이 클라이언트 디바이스에 송신될 것을 필요로 한다는 것을 고려할 때, 독립적인 부분들 또는 타일들의 효율적인 스트리밍에 적합하다. 그것은 (샘플 기반의 바이트-범위 요청들과 비교하여) 송신 오버헤드를 감소시키고, MPEG 표준에 통합될 수 있다는 것을 고려할 때, 둘 이상의 부분 또는 타일 선택(즉, 임의의 관심 영역)을 스트리밍하기에 적합하다.

[0115] 일 실시예에서, 상기 생성된 서브 트랙 박스들은 ISO/IEC 14496-12에 따라 정의된다.

[0116] 일 실시예에서, 상기 생성된 서브 트랙 박스들 중 적어도 하나는 상기 생성된 서브 트랙 박스들 중 상기 적어도 하나가 타일링 정보를 제공하는 특정한 트랙과 관련됨을 나타내는 정보를 포함한다.

[0117] 일 실시예에서, 상기 적어도 하나의 마이크로프로세서는 샘플들의 또는 서브샘플들의 그룹을 적어도 하나의 타일에 링크시키는 단계를 수행하도록 추가로 구성된다.

[0118] 일 실시예에서, 상기 생성된 서브 트랙 박스들 중 적어도 하나와 관련된 적어도 하나의 서브 트랙 내에 타일들의 그룹이 인코딩된다.

[0119] 일 실시예에서, 상기 생성된 서브 트랙 박스들 중 적어도 하나는 상기 생성된 서브 트랙 박스들 중 적어도 하나와 관련된 서브 트랙 내에 인코딩된 타일들의 수를 나타내는 정보를 포함한다.

[0120] 일 실시예에서, 상기 생성된 트랙들 각각을 캡슐화한 결과인 미디어 파일들은, 대응하는 트랙의 정의들을 포함하는 메타데이터를 저장하기 위한 데이터 구조를 포함하여, 대응하는 트랙 내에 포함된 상기 서브 트랙 박스들의 정의들을 포함한다.

[0121] 본 발명의 다른 양상에 따르면 전술한 디바이스를 포함하는 비디오 인코더가 제공된다.

[0122] 본 발명의 다른 양상에 따르면, 클라이언트 디바이스에서, 적어도 하나의 미디어 파일에 캡슐화된 분할된 시간 설정형 미디어 데이터로부터 시간 설정형 미디어 데이터 비트 스트림을 제공하는 디바이스로서, 상기 분할된 시간 설정형 미디어 데이터는 시간 설정형 샘플들을 포함하고, 각각의 시간 설정형 샘플은 복수의 서브샘플을 포함하고, 상기 서브샘플들 중 적어도 일부가 서브샘플들의 그룹에 속하도록 몇몇의 상기 서브샘플들이 그룹화되어 적어도 하나의 그룹을 형성하고, 상기 적어도 하나의 미디어 파일은 적어도 하나의 트랙을 포함하고, 상기 적어도 하나의 트랙은 상기 시간 설정형 샘플들 중 하나의 상기 복수의 서브샘플 중에서 선택된 동일한 그룹의 적어도 2개의 서브샘플을 포함하고, 상기 적어도 하나의 트랙을 생성하기 위해 선택된 상기 서브샘플들 각각에 대해 하나의 서브 트랙 박스가 생성되고, 상기 생성된 서브 트랙 박스들 중 적어도 하나는 상기 선택된 서브샘플들에 공통인 정의들 및 특성들을 포함하고, 이 디바이스는:

[0123] 상기 적어도 하나의 미디어 파일로부터 적어도 하나의 서브 트랙 박스를 선택하는 단계;

[0124] 상기 적어도 하나의 선택된 서브 트랙 박스로부터 정의들 및 특성들의 적어도 하나의 세트를 획득하는 단계; 및

[0125] 디코딩될 복수의 서브샘플을 획득하는 단계 - 상기 복수의 서브샘플은 상기 정의들 및 특성들의 적어도 하나의 세트에 종속함 -

[0126] 를 수행하도록 구성된 적어도 하나의 마이크로프로세서를 포함하는 디바이스가 제공된다.

[0127] 따라서, 본 발명은 상이한 부분들의, 특히 상이한 서브샘플들(또는 "타일들")의 결합을 가능하게 한다. 그것은 클라이언트 디바이스에 의해 다운로드된 부분들 또는 타일들의 선택된 세트가 무엇이든 클라이언트에 의한 과정 시에 유효 파일 포맷 및 기본 스트림의 생성과 관련된다.

[0128] 그러므로, 본 발명의 방법은 유용한 데이터만이 클라이언트 디바이스에 송신될 것을 필요로 한다는 것을 고려할 때, 독립적인 부분들 또는 타일들의 효율적인 스트리밍에 적합하다. 그것은 (샘플 기반의 바이트-범위 요청들과 비교하여) 송신 오버헤드를 감소시키고, MPEG 표준에 통합될 수 있다는 것을 고려할 때, 둘 이상의 부분 또는 타일 선택(즉, 임의의 관심 영역)을 스트리밍하기에 적합하다.

[0129] 일 실시예에서, 상기 생성된 서브 트랙 박스들은 ISO/IEC 14496-12에 따라 정의된다.

[0130] 일 실시예에서, 상기 서브 트랙 박스들 중 적어도 하나는 상기 서브 트랙 박스들 중 상기 적어도 하나가 타일링 정보를 제공하는 특정한 트랙과 관련됨을 나타내는 정보를 포함한다.

[0131] 일 실시예에서, 상기 서브 트랙 박스들 중 적어도 하나와 관련된 적어도 하나의 서브 트랙 내에 타일들의 그룹

이 인코딩된다.

- [0132] 일 실시예에서, 상기 서브 트랙 박스들 중 적어도 하나는 상기 생성된 서브 트랙 박스들 중 적어도 하나와 관련된 서브 트랙 내에 인코딩된 타일들의 수를 나타내는 정보를 포함한다.
- [0133] 일 실시예에서, 상기 트랙들 각각을 캡슐화한 결과인 미디어 파일들은, 대응하는 트랙의 정의들을 포함하는 메타데이터를 저장하기 위한 데이터 구조를 포함하여, 대응하는 트랙 내에 포함된 상기 서브 트랙 박스들의 정의들을 포함한다.
- [0134] 본 발명의 다른 양상에 따르면 전술한 디바이스를 포함하는 비디오 디코더가 제공된다.
- [0135] 본 발명의 다른 양상에 따르면 클라이언트의 디바이스에 스트리밍되도록 서버에서 분할된 시간 설정형 미디어 데이터를 캡슐화하는 디바이스로서, 상기 분할된 시간 설정형 미디어 데이터는 시간 설정형 샘플들을 포함하고, 각각의 시간 설정형 샘플은 복수의 서브샘플을 포함하고, 이 디바이스는:

 - [0136] 상기 서브샘플들의 적어도 일부가 서브샘플들의 그룹에 속하도록 몇몇의 상기 서브샘플들을 그룹화하여 적어도 하나의 그룹을 형성하는 단계;
 - [0137] 상기 시간 설정형 샘플들 중 하나의 상기 복수의 서브샘플 중에서 동일한 그룹의 적어도 2개의 서브샘플을 선택하는 단계;
 - [0138] 상기 적어도 2개의 선택된 서브샘플을 포함하는 적어도 하나의 트랙을 생성하는 단계; 및
 - [0139] 상기 생성된 트랙들 각각을 적어도 하나의 미디어 파일에 독립적으로 캡슐화하는 단계
 - [0140] 를 수행하도록 구성된 적어도 하나의 마이크로프로세서를 포함하는 디바이스가 제공된다.

- [0141] 따라서, 본 발명은 상이한 부분들의, 특히 상이한 서브샘플들(또는 "타일들")의 결합을 가능하게 한다. 그것은 클라이언트 디바이스에 의해 다운로드된 부분들 또는 타일들의 선택된 세트가 무엇이든 클라이언트에 의한 파싱 시에 유효 파일 포맷 및 기본 스트림의 생성과 관련된다.
- [0142] 그러므로, 본 발명의 방법은 유용한 데이터만이 클라이언트 디바이스에 송신될 것을 필요로 한다는 것을 고려할 때, 독립적인 부분들 또는 타일들의 효율적인 스트리밍에 적합하다. 그것은 (샘플 기반의 바이트-범위 요청들과 비교하여) 송신 오버헤드를 감소시키고, MPEG 표준에 통합될 수 있다는 것을 고려할 때, 둘 이상의 부분 또는 타일 선택(즉, 임의의 관심 영역)을 스트리밍하기에 적합하다.
- [0143] 본 발명의 추가 양상에 따르면, 서버에서 분할된 시간 설정형 미디어 데이터를 캡슐화하는 방법으로서, 상기 분할된 시간 설정형 미디어 데이터는 시간 설정형 샘플들을 포함하고, 각각의 시간 설정형 샘플은 복수의 서브샘플들을 포함하며, 이 방법은:

 - [0144] 상기 시간 설정형 샘플들 중 하나의 상기 복수의 서브샘플들 중에서 적어도 하나의 서브샘플을 선택하는 단계;
 - [0145] 각각의 선택된 서브샘플에 대해, 상기 선택된 서브샘플 및 다른 시간 설정형 샘플들을 각각의 하나의 대응 서브샘플을 포함하는 하나의 분할 트랙을 생성하는 단계;
 - [0146] 상기 생성된 분할 트랙들 중 적어도 하나를 식별하는 적어도 하나의 추출자를 포함하는 적어도 하나의 참조 트랙을 생성하는 단계; 및
 - [0147] 상기 생성된 트랙들 각각을 적어도 하나의 미디어 세그먼트 파일에 독립적으로 캡슐화하는 단계를 포함하는 방법이 제공된다.

- [0148] 따라서, 본 발명은, 클라이언트 디바이스에 의해 다운로드되는 부분들 또는 타일들의 선택된 세트가 무엇이든, 상이한 부분들, 특히 상이한 타일들의 조합, 및 파싱 시 유효 파일 포맷 및 기본 스트림의 생성을 가능하게 한다.
- [0149] 따라서, 본 발명의 방법은, 독립 부분들 또는 타일들의 효율적인 스트리밍에 적합하고, 유용한 데이터만이 클라이언트 디바이스에 보내어질 것을 요구하고, 둘 이상의 부분 또는 타일 선택(즉, 임의의 관심 영역)에 적합하고, (바이트 범위 요청들에 비해) 인덱스화 오버헤드를 감소시키며, MPEG 표준에 통합될 수 있다.
- [0150] 일 실시예에서, 분할된 시간 설정형 미디어 데이터는 타일형 시간 설정형 미디어 데이터이고, 서브샘플들은 공간 서브샘플들이고, 적어도 하나의 분할 트랙은 적어도 하나의 타일 트랙이며, 적어도 하나의 참조 트랙은 적어도 하나의 합성 트랙이다. 적어도 하나의 참조 트랙은 추출자들 및 옵션으로 초기화 데이터를 포함할 수 있다.

- [0151] 일 실시예에서, 적어도 하나의 추출자는 적어도 하나의 식별된 타일 트랙의 적어도 하나의 공간 서브샘플을 더 우 식별한다.
- [0152] 일 실시예에서, 본 방법은, 생성된 타일 트랙들 및 생성된 적어도 하나의 합성 트랙으로부터 타일형 시간 설정 형 미디어 데이터의 공간 부분의 설명을 가능하게 하는 파라미터들을 포함하는 초기화 세그먼트 파일을 생성하는 단계를 더 포함한다. 초기화 세그먼트 파일은, 생성된 타일 트랙들에 대한 참조들, 적어도 하나의 생성된 합성 트랙과 관련된 참조된 타일 트랙들의 리스트, 트랙이 공간 서브샘플 정보를 포함한다는 것을 나타내는, 생성된 타일 트랙들 각각과 관련된 타일 핸들러, 및/또는 생성된 타일 트랙들 각각과 관련된 일반적 프레젠테이션 정보를 더 포함할 수 있다.
- [0153] 일 실시예에서, 각각의 선택된 공간 서브샘플에 대해 하나의 타일 트랙을 생성하는 단계에서 적어도 2개의 타일 트랙들이 생성되고, 적어도 하나의 합성 트랙은 동일한 시간 설정형 샘플내에 포함되는 2개 공간 서브샘플들 중 임의의 것을 디코딩하는데 사용될 데이터의 적어도 하나의 아이템을 포함한다.
- [0154] 일 실시예에서, 타일형 시간 설정형 미디어 데이터는, 시간 설정형 샘플들 각각이 그에 따라 참조 계층 및 적어도 하나의 향상 계층을 포함하는, 스케일 가능한 타일형 시간 설정형 미디어 데이터이고, 참조 계층은 적어도 하나의 참조 공간 서브샘플을 포함하고, 적어도 하나의 향상 계층은 복수의 향상 공간 서브샘플들을 포함하며, 각각의 선택된 공간 서브샘플에 대해 하나의 타일 트랙을 생성하는 단계에서 생성된 타일 트랙들은 향상 공간 서브샘플들을 포함하는 향상 타일 트랙들이고, 본 방법은 참조 공간 서브 샘플들을 포함하는 적어도 하나의 참조 트랙을 생성하는 단계를 더 포함한다.
- [0155] 일 실시예에서, 참조 계층은 복수의 참조 공간 서브샘플들을 포함하고, 복수의 참조 트랙들이 생성되며, 복수의 참조 트랙들의 각 참조 트랙은 참조 타일 트랙을 형성한다.
- [0156] 일 실시예에서, 생성된 트랙들 각각을 캡슐화한 결과인 미디어 세그먼트 파일들은 대응 트랙의 정의들을 포함하는 메타데이터를 저장하는 데이터 구조를 포함한다.
- [0157] 일 실시예에서, 생성된 트랙들 각각을 캡슐화한 결과인 미디어 세그먼트 파일들 중 적어도 하나는 대응 트랙의 적어도 한 그룹의 서브샘플들의 적어도 하나의 정의를 포함하는 메타데이터를 저장하는 데이터 구조를 포함한다.
- [0158] 일 실시예에서, 생성된 트랙들 각각을 캡슐화한 결과인 미디어 세그먼트 파일들 중 적어도 하나는 대응 트랙의 상이한 그룹들의 서브샘플들의 상이한 정의들을 포함하는 메타데이터를 저장하는 데이터 구조를 포함한다.
- [0159] 일 실시예에서, 메타데이터를 저장하는 데이터 구조는, 적어도 하나의 정의, 및 적어도 하나의 정의가 대응 트랙의 모든 서브샘플들에 적용된다는 것을 나타내는 버전 타입을 포함한다.
- [0160] 바람직한 실시예에서, 메타데이터를 저장하는 상기 데이터 구조는 모든 정의들이 대응 트랙의 모든 서브샘플들에 적용된다는 것을 나타내는 파라미터(예를 들어 플래그)를 추가로 포함한다.
- [0161] 이 파라미터는 또한 모든 정의들이 샘플에 적용된다는 것을 나타낼 수 있다(예를 들어 그 정의들이 해당 샘플의 모든 서브샘플들에 적용된다면).
- [0162] 일 실시예에서, 서브샘플 디코딩 종속성과 관련하여, 대응 트랙의 적어도 한 그룹의 서브샘플들의 적어도 하나의 정의를 포함하는 메타데이터를 저장하는 데이터 구조의 적어도 하나의 파라미터는, 해당 그룹의 각 서브샘플이 다른 그룹의 서브샘플을 사용하지 않고 디코딩될 수 있다는 점을 나타낸다.
- [0163] 일 실시예에서, 한 그룹의 서브샘플들은 해당 서브샘플들이 속하는 그룹화 타입에 따라 식별된다.
- [0164] 일 실시예에서, 한 그룹의 서브샘플들은 해당 서브샘플들이 속하는 그룹화 타입에 따라 식별되고, 그룹화 타입은 해당 그룹의 각 서브샘플에 대한 정의를 제공한다.
- [0165] 일 실시예에서, 초기화 세그먼트 파일은, 대응하는 적어도 하나의 타일 트랙을 캡슐화한 결과인 상기 미디어 세그먼트 파일에 저장된 임의의 특정 정의와 관련되지 않는, 적어도 하나의 타일 트랙의 서브샘플들을 처리하는데 디폴트로 사용될 하나의 정의를 포함한다.
- [0166] 일 실시예에서, 본 방법은, 디폴트로 사용될 초기화 세그먼트 파일에 저장된 정의를 수정하는 것을 가능하게 하도록, 생성된 트랙들 각각을 캡슐화한 결과인 미디어 세그먼트 파일들 중 적어도 하나에 초기화 데이터를 추가하는 단계를 더 포함한다.

- [0167] 일 실시예에서, 생성된 트랙들 각각을 캡슐화한 결과인 미디어 세그먼트 파일들은 공간 서브샘플들을 저장하는 데이터 구조를 포함한다.
- [0168] 일 실시예에서, 생성된 트랙들 각각을 캡슐화한 결과인 미디어 세그먼트 파일들은 추출자들을 저장하는 데이터 구조를 포함한다.
- [0169] 일 실시예에서, 생성된 트랙들 각각을 캡슐화한 결과인 미디어 세그먼트 파일들 중 적어도 하나는, 미디어 세그먼트 파일들 중 적어도 하나가 표준 미디어 세그먼트 파일로서 독립적으로 처리될 수 있도록, 참조 트랙에 저장된 초기화 데이터를 포함한다.
- [0170] 일 실시예에서, 생성된 트랙들 각각을 캡슐화한 결과인 미디어 세그먼트 파일들 중 적어도 하나는 참조 트랙에 저장된 초기화 데이터를 식별하는 적어도 하나의 추출자를 저장하는 데이터 구조를 포함한다.
- [0171] 일 실시예에서, 생성된 트랙들 각각을 캡슐화한 결과인 미디어 세그먼트 파일들 중 적어도 하나는, 미디어 세그먼트 파일들 중 적어도 하나가 표준 미디어 세그먼트 파일로서 독립적으로 처리될 수 있도록, 합성 트랙의 초기화 데이터를 식별하는 적어도 하나의 추출자를 저장하는 데이터 구조를 포함한다.
- [0172] 일 실시예에서, 본 방법은, 미디어 세그먼트 파일들 중 적어도 하나가 표준 미디어 세그먼트 파일로서 독립적으로 처리될 수 있다는 것을 나타내는 플래그를 설정하는 단계를 더 포함한다.
- [0173] 일 실시예에서, 본 방법은, 초기화 데이터를 포함하는 적어도 하나의 초기화 데이터 트랙을 생성하는 단계를 더 포함하고, 적어도 하나의 초기화 데이터 트랙은 적어도 하나의 미디어 세그먼트 파일에 독립적으로 캡슐화된다.
- [0174] 일 실시예에서, 생성된 트랙들 각각을 캡슐화한 결과인 미디어 세그먼트 파일들 중 적어도 하나는, 미디어 세그먼트 파일들 중 적어도 하나가 표준 미디어 세그먼트 파일로서 독립적으로 처리될 수 있도록, 초기화 데이터 트랙에 저장된 초기화 데이터를 식별하는 적어도 하나의 추출자를 저장하는 데이터 구조를 포함한다.
- [0175] 일 실시예에서, 본 방법은, 미디어 세그먼트 파일들 중 적어도 하나가 표준 미디어 세그먼트 파일로서 돋립적으로 처리될 수 있다는 것을 나타내는 플래그를 설정하는 단계를 더 포함한다.
- [0176] 일 실시예에서, 서버는 HTTP(HyperText Transfer Protocol)와 호환된다.
- [0177] 일 실시예에서, 생성된 트랙들 각각을 캡슐화한 결과인 미디어 세그먼트 파일들은 국제 표준 기구(International Standard Organization)에 의해 정의되는 바와 같은 베이스 미디어 파일 포맷(Base Media File Format) 및 DASH(Dynamic Adaptive Streaming over HTTP) 포맷과 호환된다.
- [0178] 본 발명의 다른 양상은, 클라이언트 디바이스에서, 복수의 미디어 세그먼트 파일들에 캡슐화된 분할된 시간 설정형 미디어 데이터로부터 시간 설정형 미디어 데이터 비트 스트림을 제공하는 방법으로서, 상기 분할된 시간 설정형 미디어 데이터는 시간 설정형 샘플들을 포함하고, 각각의 시간 설정형 샘플은 복수의 서브샘플들을 포함하고, 상기 복수의 미디어 세그먼트 파일들은 적어도 하나의 참조 트랙 및 적어도 하나의 분할 트랙을 포함하고, 상기 적어도 하나의 분할 트랙 각각은, 상기 시간 설정형 샘플들 중 하나의 상기 복수의 서브샘플들 중에서 선택되는 하나의 서브샘플을 포함하고, 다른 시간 설정형 샘플들 각각의 하나의 대응 공간 서브 샘플을 포함하고, 상기 적어도 하나의 참조 트랙은, 적어도 하나의 분할 트랙을 식별하고 이 식별된 분할 트랙의 적어도 하나의 서브샘플을 식별하는 적어도 하나의 추출자를 포함하고, 상기 방법은:
- [0179] 서브샘플을 나타내는 정보의 아이템을 선택하는 단계;
- [0180] 상기 선택된 서브샘플을 포함하는 상기 적어도 하나의 분할 트랙을 요청하는 단계 - 상기 선택된 서브샘플은 상기 선택된 정보의 아이템에 대응함 -;
- [0181] 상기 복수의 미디어 세그먼트 파일들을 수신하는 단계; 및
- [0182] 상기 시간 설정형 미디어 데이터 비트 스트림을 생성하도록 상기 적어도 하나의 참조 트랙의 추출자들을 대응 식별된 서브샘플들로 대체하는 단계를 포함하는 방법을 제공한다.
- [0183] 따라서, 본 발명은, 클라이언트 디바이스에 의해 다운로드되는 부분들 또는 파일들의 선택된 세트가 무엇이든, 상이한 부분들, 특히 상이한 파일들의 조합, 및 과정시 유효 파일 포맷 및 기본 스트림의 생성을 가능하게 한다.
- [0184] 따라서, 본 발명의 방법은, 독립 부분들 또는 파일들의 효율적인 스트리밍에 적합하고, 유용한 데이터만이 클라

이언트 디바이스에 보내어질 것을 요구하고, 둘 이상의 부분 또는 타일 선택(즉, 임의의 관심 영역)에 적합하고, (바이트 범위 요청들에 비해) 인덱스화 오버헤드를 감소시키며, MPEG 표준에 통합될 수 있다.

[0185] 일 실시예에서, 상기 분할된 시간 설정형 미디어 데이터는 타일형 시간 설정형 미디어 데이터이고, 상기 서브샘플들은 공간 서브샘플들이고, 상기 적어도 하나의 분할 트랙은 적어도 하나의 타일 트랙이며, 상기 적어도 하나의 참조 트랙은 적어도 하나의 합성 트랙이다.

[0186] 일 실시예에서, 상기 분할된 시간 설정형 미디어 데이터는 타일형 시간 설정형 미디어 데이터이고, 상기 서브샘플들은 공간 서브샘플들이고, 상기 적어도 하나의 분할 트랙은 적어도 하나의 타일 트랙이며, 상기 적어도 하나의 참조 트랙은 추출자들 및 옵션으로 초기화 데이터를 포함하는 적어도 하나의 합성 트랙이다.

[0187] 일 실시예에서, 본 방법은,

상기 적어도 하나의 합성 트랙의 추출자로부터 트랙 참조를 취득하는 단계;

[0189] 상기 취득된 트랙 참조에 대응하는 트랙이 수신되었는지 여부를 체크하는 단계; 및

[0190] 상기 취득된 트랙 참조에 대응하는 트랙이 수신되지 않았으면, 상기 시간 설정형 미디어 데이터 비트 스트림을 생성하도록 대응 추출자를 제거하는 단계를 더 포함한다.

[0191] 일 실시예에서, 본 방법은, 상기 취득된 트랙 참조에 대응하는 트랙이 타일 탑입의 트랙인지 확인하는 단계를 더 포함한다.

[0192] 일 실시예에서, 본 방법은,

상기 적어도 하나의 합성 트랙의 추출자로부터 트랙 참조를 취득하는 단계;

[0194] 상기 취득된 트랙 참조에 대응하는 트랙이 수신되었는지 여부를 체크하는 단계; 및

[0195] 상기 취득된 트랙 참조에 대응하는 트랙이 수신되지 않았으면, 상기 시간 설정형 미디어 데이터 비트 스트림을 생성하도록 대응 추출자를 패딩으로 대체하는 단계를 더 포함한다.

[0196] 일 실시예에서, 본 방법은, 상기 취득된 트랙 참조에 대응하는 트랙이 타일 탑입의 트랙인지 확인하는 단계를 더 포함한다.

[0197] 일 실시예에서, 본 방법은, 상기 생성된 타일 트랙들 및 상기 생성된 적어도 하나의 합성 트랙으로부터 상기 타일형 시간 설정형 미디어 데이터의 공간 부분의 형성을 가능하게 하는 파라미터들을 포함하는 초기화 세그먼트 파일을 수신하는 단계를 더 포함하다.

[0198] 일 실시예에서, 적어도 2개의 타일 트랙들을 나타내는 적어도 2개의 미디어 세그먼트 파일들이 수신되고, 상기 적어도 하나의 합성 트랙은 동일 세트의 공간 서브샘플들에 속하는 2개의 공간 서브샘플들 중 임의의 것을 디코딩하는데 사용될 데이터의 적어도 하나의 아이템을 포함한다.

[0199] 일 실시예에서, 상기 타일형 시간 설정형 미디어 데이터는, 상기 시간 설정형 샘플들 각각이 그에 따라 참조 계층 및 적어도 하나의 향상 계층을 포함하는, 스케일 가능한 타일형 시간 설정형 미디어 데이터이고, 상기 참조 계층은 적어도 하나의 참조 공간 서브샘플을 포함하고, 상기 적어도 하나의 향상 계층은 복수의 향상 공간 서브샘플들을 포함하며, 적어도 2개의 트랙들을 나타내는 적어도 2개의 미디어 세그먼트 파일들이 수신되고, 상기 적어도 2개의 수신된 트랙들 중 하나는 참조 공간 서브샘플들을 포함하는 참조 트랙이고, 상기 적어도 2개의 수신된 트랙들 중 다른 하나는 향상 공간 서브샘플들을 포함하는 향상 타일 트랙이다.

[0200] 일 실시예에서, 상기 참조 계층은 복수의 참조 공간 서브샘플들을 포함하고, 각각 참조 타일 트랙을 형성하는, 복수의 참조 트랙들이 수신된다.

[0201] 일 실시예에서, 상기 미디어 세그먼트 파일들 중 적어도 하나는, 대응 트랙의 적어도 한 그룹의 서브샘플들의 적어도 하나의 정의를 포함하는 메타데이터를 저장하는 데이터 구조를 포함하고, 본 방법은, 대응 트랙의 상기 적어도 한 그룹의 서브샘플들을 처리하기 위해 상기 적어도 하나의 정의를 취득하는 단계를 포함한다.

[0202] 일 실시예에서, 상기 미디어 세그먼트 파일들 중 적어도 하나는 상기 참조 트랙에 저장된 초기화 데이터를 포함하고, 본 방법은, 상기 미디어 세그먼트 파일들 중 상기 적어도 하나를 표준 미디어 세그먼트 파일로서 독립적으로 처리하는 단계를 더 포함한다.

[0203] 일 실시예에서, 상기 미디어 세그먼트 파일들 중 적어도 하나는 상기 참조 트랙에 저장된 초기화 데이터를 식별

하는 적어도 하나의 추출자를 저장하는 데이터 구조를 포함하고, 본 방법은, 상기 미디어 세그먼트 파일들 중 적어도 하나를 처리할 때 상기 참조 트랙에 저장된 상기 데이터에 액세스하는 단계를 더 포함한다.

[0204] 일 실시예에서, 상기 미디어 세그먼트 파일들 중 적어도 하나는 상기 합성 트랙의 초기화 데이터를 식별하는 적어도 하나의 추출자를 저장하는 데이터 구조를 포함하고, 본 방법은, 상기 미디어 세그먼트 파일들 중 적어도 하나를 표준 미디어 세그먼트 파일로서 독립적으로 처리하기 위해 상기 참조 트랙에 저장된 상기 데이터에 액세스하는 단계를 더 포함한다.

[0205] 일 실시예에서, 본 방법은, 초기화 데이터를 포함하는 적어도 하나의 초기화 데이터 트랙을 취득하는 단계를 더 포함하고, 상기 적어도 하나의 초기화 데이터 트랙은 적어도 하나의 미디어 세그먼트 파일에 독립적으로 캡슐화된다.

[0206] 일 실시예에서, 상기 미디어 세그먼트 파일들 중 적어도 하나는 상기 초기화 데이터 트랙에 저장된 초기화 데이터를 식별하는 적어도 하나의 추출자를 저장하는 데이터 구조를 포함하고, 본 방법은, 상기 미디어 세그먼트 파일들 중 상기 적어도 하나를 독립적으로 처리하기 위해 상기 참조 트랙에 저장된 데이터에 액세스하는 단계를 더 포함한다.

[0207] 일 실시예에서, 클라이언트 디바이스는 HTTP(HyperText Transfer Protocol)와 호환된다.

[0208] 일 실시예에서, 수신된 미디어 세그먼트 파일들은 국제 표준 기구(International Standard Organization)에 의해 정의되는 바와 같은 베이스 미디어 파일 포맷(Base Media File Format) 및 DASH(Dynamic Adaptive Streaming over HTTP) 포맷과 호환된다.

[0209] 본 발명의 다른 양상에 따르면, 위에 설명된 방법의 각 단계를 수행하도록 적응된 수단들을 포함하는 디바이스가 제공된다.

[0210] 본 발명의 다른 양상에 따르면, 서버에서 분할된 시간 설정형 미디어 데이터를 캡슐화하는 디바이스로서, 상기 분할된 시간 설정형 미디어 데이터는 시간 설정형 샘플들을 포함하고, 각각의 시간 설정형 샘플은 복수의 서브 샘플들을 포함하고, 이 디바이스는:

[0211] 상기 시간 설정형 샘플들 중 하나의 상기 복수의 서브샘플들 중에서 적어도 하나의 서브샘플을 선택하는 단계;

[0212] 각각의 선택된 서브샘플에 대해, 상기 선택된 서브샘플 및 다른 시간 설정형 샘플들 각각의 하나의 대응 서브샘플을 포함하는 하나의 분할 트랙을 생성하는 단계;

[0213] 상기 생성된 분할 트랙들 중 적어도 하나를 식별하는 적어도 하나의 추출자를 포함하는 적어도 하나의 참조 트랙을 생성하는 단계; 및

[0214] 상기 생성된 트랙들 각각을 적어도 하나의 미디어 세그먼트 파일에 독립적으로 캡슐화하는 단계를 수행하도록 구성되는 적어도 하나의 마이크로프로세서를 포함하는 디바이스가 제공된다.

[0215] 따라서, 본 발명은, 클라이언트 디바이스에 의해 다운로드되는 부분들 또는 타일들의 선택된 세트가 무엇이든, 상이한 부분들, 특히 상이한 타일들의 조합, 및 파싱시 유효 파일 포맷 및 기본 스트림의 생성을 가능하게 한다.

[0216] 따라서, 본 발명의 방법은, 독립 부분들 또는 타일들의 효율적인 스트리밍에 적합하고, 유용한 데이터만이 클라이언트 디바이스에 보내어질 것을 요구하고, 둘 이상의 부분 또는 타일 선택(즉, 임의의 관심 영역)에 적합하고, (바이트 범위 요청들에 비해) 인덱스화 오버헤드를 감소시키며, MPEG 표준에 통합될 수 있다.

[0217] 일 실시예에서, 상기 분할된 시간 설정형 미디어 데이터는 타일형 시간 설정형 미디어 데이터이고, 상기 서브샘플들은 공간 서브샘플들이고, 상기 적어도 하나의 분할 트랙은 적어도 하나의 타일 트랙이며, 상기 적어도 하나의 참조 트랙은 적어도 하나의 합성 트랙이다.

[0218] 일 실시예에서, 상기 분할된 시간 설정형 미디어 데이터는 타일형 시간 설정형 미디어 데이터이고, 상기 서브샘플들은 공간 서브샘플들이고, 상기 적어도 하나의 분할 트랙은 적어도 하나의 타일 트랙이며, 상기 적어도 하나의 참조 트랙은 추출자들 및 옵션으로 초기화 데이터를 포함하는 적어도 하나의 합성 트랙이다.

[0219] 일 실시예에서, 상기 마이크로프로세서는, 상기 생성된 타일 트랙들 및 상기 생성된 적어도 하나의 합성 트랙으로부터 상기 타일형 시간 설정형 미디어 데이터의 공간 부분의 형성을 가능하게 하는 파라미터들을 포함하는 초기화 세그먼트 파일을 생성하는 단계를 수행하도록 더 구성되고, 상기 초기화 세그먼트 파일은 상기 생성된 타

일 트랙들에 대한 참조들을 포함한다.

- [0220] 일 실시예에서, 상기 초기화 세그먼트 파일은 상기 적어도 하나의 생성된 합성 트랙과 관련된 참조된 타일 트랙들의 리스트를 더 포함한다.
- [0221] 일 실시예에서, 상기 초기화 세그먼트 파일은, 상기 트랙이 공간 서브샘플 정보를 포함한다는 것을 나타내는, 상기 생성된 타일 트랙들 각각과 관련된 타일 핸들러를 더 포함한다.
- [0222] 일 실시예에서, 상기 초기화 세그먼트 파일은 상기 생성된 타일 트랙들 각각과 관련된 일반적 프레젠테이션 정보를 더 포함한다.
- [0223] 일 실시예에서, 상기 마이크로프로세서는, 각각의 선택된 공간 서브샘플에 대해 하나의 타일 트랙을 생성하는 단계에서 적어도 2개의 타일 트랙들이 생성되도록 더 구성되고, 상기 적어도 하나의 합성 트랙은 동일한 시간 설정형 샘플 내에 포함되는 2개 공간 서브샘플들 중 임의의 것을 디코딩하는데 사용될 데이터의 적어도 하나의 아이템을 포함한다.
- [0224] 일 실시예에서, 상기 타일형 시간 설정형 미디어 데이터는, 상기 시간 설정형 샘플들 각각이 그에 따라 참조 계층 및 적어도 하나의 향상 계층을 포함하는, 스케일 가능한 타일형 시간 설정형 미디어 데이터이고, 상기 참조 계층은 적어도 하나의 참조 공간 서브샘플을 포함하고, 상기 적어도 하나의 향상 계층은 복수의 향상 공간 서브샘플들을 포함하며, 상기 마이크로프로세서는, 각각의 선택된 공간 서브샘플에 대해 하나의 타일 트랙을 생성하는 단계에서 생성된 타일 트랙들이 향상 공간 서브샘플들을 포함하는 향상 타일 트랙들이도록 더 구성되고, 상기 마이크로프로세서는, 참조 공간 서브 샘플들을 포함하는 적어도 하나의 참조 트랙을 생성하는 단계를 수행하도록 더 구성된다.
- [0225] 일 실시예에서, 상기 참조 계층은 복수의 참조 공간 서브샘플들을 포함하고, 상기 마이크로프로세서는, 복수의 참조 트랙들이 생성되도록 더 구성되며, 상기 복수의 참조 트랙들의 각 참조 트랙은 참조 타일 트랙을 형성한다.
- [0226] 일 실시예에서, 상기 생성된 트랙들 각각을 캡슐화한 결과인 미디어 세그먼트 파일들은 대응 트랙의 정의들을 포함하는 메타데이터를 저장하는 데이터 구조를 포함한다.
- [0227] 일 실시예에서, 상기 생성된 트랙들 각각을 캡슐화한 결과인 미디어 세그먼트 파일들은 공간 서브샘플들을 저장하는 데이터 구조를 포함한다.
- [0228] 일 실시예에서, 상기 생성된 트랙들 각각을 캡슐화한 결과인 미디어 세그먼트 파일들은 추출자들을 저장하는 데이터 구조를 포함한다.
- [0229] 일 실시예에서, 서버는 HTTP(HyperText Transfer Protocol)와 호환되고, 상기 생성된 트랙들 각각을 캡슐화한 결과인 상기 미디어 세그먼트 파일들은 국제 표준 기구(International Standard Organization)에 의해 정의되는 바와 같은 베이스 미디어 파일 포맷(Base Media File Format) 및 DASH(Dynamic Adaptive Streaming over HTTP) 포맷과 호환된다.
- [0230] 본 발명의 다른 양상에 따르면, 위에 설명된 바와 같은 디바이스를 포함하는 비디오 인코더가 제공된다.
- [0231] 본 발명의 다른 양상에 따르면, 클라이언트 디바이스에서, 복수의 미디어 세그먼트 파일들에 캡슐화된 분할된 시간 설정형 미디어 데이터로부터 시간 설정형 미디어 데이터 비트 스트림을 제공하는 디바이스로서, 상기 분할된 시간 설정형 미디어 데이터는 시간 설정형 샘플들을 포함하고, 각각의 시간 설정형 샘플은 복수의 서브샘플들을 포함하고, 상기 복수의 미디어 세그먼트 파일들은 적어도 하나의 참조 트랙 및 적어도 하나의 분할 트랙을 포함하고, 상기 적어도 하나의 분할 트랙 각각은, 상기 시간 설정형 샘플들 중 하나의 상기 복수의 서브샘플들 중에서 선택되는 하나의 서브샘플을 포함하고, 다른 시간 설정형 샘플들 각각의 하나의 대응 서브 샘플을 포함하고, 상기 적어도 하나의 참조 트랙은, 적어도 하나의 분할 트랙을 식별하고 이 식별된 분할 트랙의 적어도 하나의 서브샘플을 식별하는 적어도 하나의 추출자를 포함하고, 이 디바이스는:

 - [0232] 서브샘플을 나타내는 정보의 아이템을 선택하는 단계;
 - [0233] 상기 선택된 서브샘플을 포함하는 상기 적어도 하나의 분할 트랙을 요청하는 단계 - 상기 선택된 서브샘플은 상기 선택된 정보의 아이템에 대응함 -;
 - [0234] 상기 복수의 미디어 세그먼트 파일들을 수신하는 단계; 및

- [0235] 상기 시간 설정형 미디어 데이터 비트 스트림을 생성하도록 상기 적어도 하나의 참조 트랙의 추출자들을 대응 식별된 서브샘플들로 대체하는 단계
- [0236] 를 수행하도록 구성되는 적어도 하나의 마이크로프로세서를 포함하는 디바이스가 제공된다.
- [0237] 따라서, 본 발명은, 클라이언트 디바이스에 의해 다운로드되는 부분들 또는 타일들의 선택된 세트가 무엇이든, 상이한 부분들, 특히 상이한 타일들의 조합, 및 과정시 유효 파일 포맷 및 기본 스트림의 생성을 가능하게 한다.
- [0238] 따라서, 본 발명의 방법은, 독립 부분들 또는 타일들의 효율적인 스트리밍에 적합하고, 유용한 데이터만이 클라이언트 디바이스에 보내어질 것을 요구하고, 둘 이상의 부분 또는 타일 선택(즉, 임의의 관심 영역)에 적합하고, (바이트 범위 요청들에 비해) 인덱스화 오버헤드를 감소시키며, MPEG 표준에 통합될 수 있다.
- [0239] 일 실시예에서, 상기 분할된 시간 설정형 미디어 데이터는 타일형 시간 설정형 미디어 데이터이고, 상기 서브샘플들은 공간 서브샘플들이고, 상기 적어도 하나의 분할 트랙은 적어도 하나의 타일 트랙이며, 상기 적어도 하나의 참조 트랙은 적어도 하나의 합성 트랙이다.
- [0240] 일 실시예에서, 상기 분할된 시간 설정형 미디어 데이터는 타일형 시간 설정형 미디어 데이터이고, 상기 서브샘플들은 공간 서브샘플들이고, 상기 적어도 하나의 분할 트랙은 적어도 하나의 타일 트랙이며, 상기 적어도 하나의 참조 트랙은 추출자들 및 옵션으로 초기화 데이터를 포함하는 적어도 하나의 합성 트랙이다.
- [0241] 일 실시예에서, 상기 마이크로프로세서는,
- [0242] 상기 적어도 하나의 합성 트랙의 추출자로부터 트랙 참조를 취득하는 단계;
- [0243] 상기 취득된 트랙 참조에 대응하는 트랙이 수신되었는지 여부를 체크하는 단계; 및
- [0244] 상기 취득된 트랙 참조에 대응하는 트랙이 수신되지 않았으면, 상기 시간 설정형 미디어 데이터 비트 스트림을 생성하도록 대응 추출자를 제거하는 단계를 수행하도록 더 구성된다.
- [0245] 일 실시예에서, 상기 마이크로프로세서는, 상기 취득된 트랙 참조에 대응하는 트랙이 타일 태입의 트랙인지 확인하는 단계를 수행하도록 더 구성된다.
- [0246] 일 실시예에서, 상기 마이크로프로세서는,
- [0247] 상기 적어도 하나의 합성 트랙의 추출자로부터 트랙 참조를 취득하는 단계;
- [0248] 상기 취득된 트랙 참조에 대응하는 트랙이 수신되었는지 여부를 체크하는 단계; 및
- [0249] 상기 취득된 트랙 참조에 대응하는 트랙이 수신되지 않았으면, 상기 시간 설정형 미디어 데이터 비트 스트림을 생성하도록 대응 추출자를 패딩으로 대체하는 단계를 수행하도록 더 구성된다.
- [0250] 일 실시예에서, 상기 마이크로프로세서는, 상기 취득된 트랙 참조에 대응하는 트랙이 타일 태입의 트랙인지 확인하는 단계를 수행하도록 더 구성된다.
- [0251] 일 실시예에서, 상기 마이크로프로세서는, 상기 생성된 타일 트랙들 및 상기 생성된 적어도 하나의 합성 트랙으로부터 상기 타일형 시간 설정형 미디어 데이터의 공간 부분의 형성을 가능하게 하는 파라미터들을 포함하는 초기화 세그먼트 파일을 수신하는 단계를 수행하도록 더 구성된다.
- [0252] 일 실시예에서, 상기 마이크로프로세서는, 적어도 2개의 타일 트랙들을 나타내는 적어도 2개의 미디어 세그먼트 파일들이 수신되도록 더 구성되고, 상기 적어도 하나의 합성 트랙은 동일 세트의 공간 서브샘플들에 속하는 2개의 공간 서브샘플들 중 임의의 것을 디코딩하는데 사용될 데이터의 적어도 하나의 아이템을 포함한다.
- [0253] 일 실시예에서, 상기 타일형 시간 설정형 미디어 데이터는, 상기 시간 설정형 샘플들 각각이 그에 따라 참조 계층 및 적어도 하나의 향상 계층을 포함하는, 스케일 가능한 타일형 시간 설정형 미디어 데이터이고, 상기 참조 계층은 적어도 하나의 참조 공간 서브샘플을 포함하고, 상기 적어도 하나의 향상 계층은 복수의 향상 공간 서브샘플들을 포함하며, 상기 마이크로프로세서는, 적어도 2개의 트랙들을 나타내는 적어도 2개의 미디어 세그먼트 파일들이 수신되도록 더 구성되고, 상기 적어도 2개의 수신된 트랙들 중 하나는 참조 공간 서브샘플들을 포함하는 참조 트랙이고, 상기 적어도 2개의 수신된 트랙들 중 다른 하나는 향상 공간 서브샘플들을 포함하는 향상 타일 트랙이다.
- [0254] 일 실시예에서, 상기 클라이언트 디바이스는 HTTP(HyperText Transfer Protocol)와 호환되고, 상기 수신된 미디

어 세그먼트 파일들은 국제 표준 기구(International Standard Organization)에 의해 정의되는 바와 같은 베이스 미디어 파일 포맷(Base Media File Format) 및 DASH(Dynamic Adaptive Streaming over HTTP) 포맷과 호환된다.

[0255] 본 발명의 다른 양상에 따르면, 위에 설명된 바와 같은 디바이스를 포함하는 비디오 디코더가 제공된다.

[0256] 본 발명이 소프트웨어로 구현될 수 있으므로, 본 발명은 임의의 적합한 캐리어 매체 상의 프로그래머블 장치에 제공하기 위한 컴퓨터 판독가능 코드로서 구현될 수 있다. 유형의 캐리어 매체는, 플로피 디스크, CD-ROM, 하드 디스크 드라이브, 자기 테이프 디바이스 또는 솔리드 스테이트 메모리 등과 같은 저장 매체를 포함할 수 있다. 일시적 캐리어 매체는, 전기 신호, 전자 신호, 광 신호, 음향 신호, 자기 신호 또는 전자기 신호, 예를 들어, 마이크로파 또는 RF 신호와 같은 신호를 포함할 수 있다.

도면의 간단한 설명

[0257] 본 발명의 추가의 이점들은 도면들 및 상세한 설명의 검토에 의해 본 기술 분야의 통상의 기술자에게 명백해질 것이다. 본 명세서에는 임의의 추가적인 이점들이 포함되는 것으로 의도된다.

이제, 본 발명의 실시예들은, 단지 예로서, 다음의 도면들을 참조하여 설명될 것이다.

도 1은 도 1a 및 도 1b를 포함하고, 인코딩된 비디오 비트 스트림에서 코딩 타일들의 일례를 도시한다.

도 2는 디스플레이되도록 사용자에 의해 선택된 타일들의 시간적인 파이프를 도시한다.

도 3은 특정한 실시예에 따라 스케일 가능한 미디어 데이터 트랙들을 이용하는 초기화 세그먼트 파일에 대한 블록도들의 일례를 도시한다.

도 4는 도 3에 도시된 것과 같은 초기화 세그먼트 파일에서 공표된 트랙들에 대응하는, 특정한 실시예에 따른, 타일 트랙들 및 하나의 합성 트랙을 포함하는 미디어 세그먼트 파일들에 대한 블록도들의 일례를 도시한다.

도 5는 주어진 시간 기간 동안 연속적인 비디오 프레임들의 공간 부분을 나타내는 유효 디코드 가능한 시간 설정형 미디어 데이터 비트 스트림(valid decodable timed media data bit-stream)을 구축하기 위해 다운로드된 미디어 데이터 세그먼트들을 연접시키는(concatenate) 일례를 도시한다.

도 6은 (a) 및 (b)를 포함하고, 특정한 실시예에 따라 서버와 클라이언트 디바이스 사이에 시간 설정형 미디어 데이터를 송신하기 위한 단계들을 설명하는 흐름도이다.

도 7은 클라이언트 디바이스에 의해 수신된 연접된 미디어 세그먼트들로부터 유효한 시간 설정형 미디어 데이터 비트 스트림의 생성을 설명하는 흐름도이다.

도 8은 하나 이상의 실시예들의 단계들이 구현될 수 있는 서버 또는 클라이언트 디바이스의 블록도를 나타낸다.

도 9는 도 9a, 도 9b, 및 도 9c를 포함하고, HEVC 비트 스트림에서 타일들 및 슬라이스 세그먼트들의 예들을 도시한다.

도 10은 본 발명의 일 실시예에 따라, 합성 트랙 및 독립적 타일 트랙들을 포함하는 트랙들의 세트로서 HEVC 비트 스트림을 캡슐화하는 일례를 도시한다.

도 11은 HEVC 비트 스트림을 캡슐화하기 위해, 제1 실시예에 따라, 미디어 데이터 트랙들을 이용하는 초기화 세그먼트 파일의 블록도들의 일례를 도시한다.

도 12는 도 12a 및 도 12b를 포함하고, 폴 비디오에서 타일의 위치, 타일의 사이즈, 및 다양한 타일링 구성을 다루기 위해 적응된, 서브-샘플 레벨에서, 타일 트랙이 임의의 아티팩트 없이 디코딩될 수 있다는 표시에 대한 신호화를 도시한다.

도 13은 본 발명의 일 실시예에 따라, 표준 비디오 트랙들로서 재생 가능한 합성 트랙 및 독립적 타일 트랙들을 포함하는 트랙들의 세트로서 HEVC 비트 스트림을 캡슐화하는 일례를 도시한다.

도 14는 본 발명의 또 다른 실시예에 따른, 표준 비디오 트랙들로서 재생 가능한 합성 트랙, 초기화 데이터 트랙, 및 독립적 타일 트랙들을 포함하는 트랙들의 세트로서 HEVC 비트 스트림을 캡슐화하는 일례를 도시한다.

도 15는 타일들을 기술하기 위해 서브 트랙 피처(즉, sub_track 박스들)를 이용하는 일례를 도시한다.

도 16은 서브 트랙들로 타일들을 캡슐화하는 것을 설명하는 흐름도이다.

도 17은 비디오의 다수의 타일들을 커버하는 관심 영역의 일례를 도시한다.

도 18은 코딩 종속성을 위해 일반 시그널링을 이용하여 HEVC 비트 스트림을 캡슐화하는 일례를 도시한다.

도 19는 도 19a 및 도 19b를 포함하고, 코딩 종속성을 위해 일반 시그널링을 이용하여 캡슐화된 트랙들에 대한 트랙 헤더들 및 트랙 참조들의 예들을 도시한다.

도 20은 코딩 종속성을 위해 일반 시그널링을 이용하여 캡슐화된 mp4 파일 또는 세그먼트들을 해석하기 위해 클라이언트 디바이스에 의해 수행되는 프로세스를 설명하는 흐름도이다.

도 21은 도 21a 및 도 21b를 포함하고, 각각 종래 기술과 본 발명의 실시예에 따른 종속성을 설명하는 그림들을 나타낸다.

발명을 실시하기 위한 구체적인 내용

[0258]

특정한 실시예에 따르면, 시간 설정형 샘플들(예컨대, 이미지들)을 포함하는 타일형 시간 설정형 미디어 데이터(예컨대, 비디오 데이터)와 같은 분할된 시간 설정형 미디어 데이터는 여러 시간 설정형 미디어 데이터 트랙들, 전형적으로 베이스 계층 트랙 및 여러 타일 트랙들, 및 시간 설정형 미디어 데이터 트랙들에 대한 참조들을 포함하는 참조 또는 합성 트랙의 세트로서 송신된다. 각각의 타일 트랙은 여러 시간 설정형 샘플들의 하나의 공간적 서브샘플(예컨대, 여러 NAL 유닛들)을 포함한다. 합성 트랙으로부터 시간 설정형 미디어 데이터 트랙들을 참조하기 위해 확장된 추출자 탑입이 정의된다. 시간 설정형 미디어 데이터 트랙들은 디스플레이 불가능한 것으로 표기되고, 타일들에 대한 시간 설정형 미디어 데이터를 운반 및 기술한다. 시간 설정형 미디어 데이터 트랙들과 합성 트랙의 그러한 세트는 공간적 비디오 타일들의 선택, 구성, 및 효율적 스트리밍을 허용한다. 각각의 트랙은 미디어 세그먼트 파일들의 세트로서 서버 디바이스로부터 클라이언트 디바이스로 송신될 수 있다. 초기화 세그먼트 파일은 미디어 세그먼트 파일들을 디코딩하기 위해 요구되는 메타데이터를 송신하는데 사용될 수 있다.

[0259]

도 1은 도 1a 및 도 1b를 포함하고, 인코딩된 비디오 비트 스트림에서 코딩 타일들의 일례를 도시한다.

[0260]

설명의 목적상, 다음의 설명에서는, 각각의 비디오 프레임(시간 설정형 샘플)이 비디오 프레임의 공간적 서브-파트들(공간 서브샘플들)에 대응하는 독립적으로 디코드 가능한 타일들로 구성되는 것으로 간주된다. 비디오는 바람직하게 스케일 가능하고, 상이한 레벨들의 스케일 가능성으로 조직화된다. 도 1a에 도시된 바와 같이, 비디오 프레임(100)은 HD 베이스 계층(base layer; 102) 및 4K2K 향상 계층(enhancement layer; 104)을 포함할 수 있다. 계속해서 설명의 목적상, 향상 계층(104)은 a, b, c 및 d로 표기된 4개의 정규 타일들로 나누어질 수 있다. 상이한 형상들의 타일들이 다루어질 수 있다는 것에 주목해야 한다. 마찬가지로, 베이스 계층(102)은 여러 타일들로 나누어질 수 있다. 그러한 경우에, 예를 들어, 베이스 계층당 하나 및 향상 계층 당 하나 또는 향상 계층들 각각에 대해 여러 합성 트랙들이 사용될 수 있다.

[0261]

본 발명은 스케일 가능한 비디오 포맷으로 제한되지 않는다는 것에 주목해야 한다. 이는 타일들이 독립적으로 디코딩되는 것을 허용하는 모든 비디오 포맷들에 적용될 수 있다. 따라서, MPEG4, AVC, HEVC, SVC 또는 미래의 SHVC와 같은 임의의 비디오 압축 알고리즘들이 본 발명의 일 실시예와 함께 사용될 수 있다.

[0262]

도 1b는 전형적인 인코딩된 비디오 비트 스트림을 디코딩 순서로 나타낸다. 도시된 바와 같이, 인코딩된 비디오 비트 스트림은 시간적 순서로 인코딩된 3개의 비디오 프레임들(110, 112 및 114)을 여기서 포함한다. 각각의 비디오 프레임은 베이스 계층(BL)의 모든 NAL(network abstraction layer) 유닛들과 그에 후속하는 향상 계층의 NAL 유닛들을 포함한다. 예를 들어, 제1 비디오 프레임(110)의 베이스 계층(102-1)의 NAL 유닛들(1BL, 116)에 후속하여 제1 비디오 프레임의 향상 계층(104-1)의 NAL 유닛들(1common, 1a, 1b, 1c, 1d, 118)이 온다.

[0263]

공간적 타일들을 가진 향상 계층에 대응하는 비디오 비트 스트림의 파트는 각각의 타일의 NAL 유닛들로 구성된다. 선택적으로, 그것은 모든 타일들에 공통이고 타일들 중 임의의 것을 디코딩하기 위해 요구되는 NAL 유닛들을 또한 포함할 수 있다. 주어진 프레임의 모든 타일들에 공통인 NAL 유닛들은 비디오 비트 스트림의 대응하는 파트에서 어디든(즉, 비디오 프레임의 타일들의 NAL 유닛들 이전, 사이, 또는 이후) 위치할 수 있다.

[0264]

도시된 바와 같이, 공간적 타일들 a, b, c 및 d를 포함하는 제1 비디오 프레임(110)의 향상 계층에 대응하는 비디오 비트 스트림의 파트는 각각의 타일(1a, 1b, 1c 및 1d)에 대한 NAL 유닛들 및 모든 타일들 a, b, c 및 d에 공통인 NAL 유닛들(1common)로 구성된다.

- [0265] 도 2는 디스플레이되도록 사용자에 의해 선택된 타일들의 시간적인 파이프를 도시한다. 보다 구체적으로, 도 2는 제1 비디오 프레임 n 및 제2 비디오 프레임 n+m(여기서, n 및 m은 정수 값들임)을 나타내고, 제1 비디오 프레임 및 제2 비디오 프레임 각각은 1 내지 12로 넘버링된 12개의 타일들을 포함한다. 이러한 12개의 타일들 중에서, (굵은 선들로 표기된 바와 같이) 단지 제3 타일 및 제7 타일만 디스플레이될 것이다. 비디오 프레임 n 및 n+m은 주어진 시간 기간에 대응하는 일련의 연속적인 프레임들에 속한다. 그러므로, 프레임 n으로부터 프레임 n+m까지의 각각의 프레임의 제3 타일 및 제7 타일은 연속적으로 디스플레이된다.
- [0266] 도 1에 도시된 바와 같이, 비디오 비트 스트림의 데이터는 풀 프레임들에 대응하는 시간적인 샘플들로서 조직화된다. 따라서, 도 2를 참조하여 위에서 설명된 바와 같이, 이러한 프레임들의 특정한 공간 영역들이 주어진 시간 기간 동안 액세스될 때 각각 프레임에 대한 여러 작은 바이트 범위들에 액세스하는 것이 요구된다. 이는 발생된 요청들의 수의 관점에서 그리고 데이터 오버헤드의 관점에서 HTTP 스트리밍에서는 비효율적이다.
- [0267] 그래서, ROI 스트리밍을 위한 압축된 비디오들에서 더 효율적인 액세스를 제공하기 위해, 특정한 타일의 데이터가 주어진 시간 기간(즉, 연속적인 프레임들의 세트) 동안 연속적인 바이트 범위로서 조직화되도록(파이프를 형성), 시간 설정형 미디어 데이터 비트 스트림이 인덱싱되거나 재조직화될 것이다.
- [0268] 따라서, 단지 비디오 프레임들의 공간적 서브-파트만 디스플레이될 때, 단지 선택된 공간 영역에 대응하는 타일들의 파이프들만 파이프당 그리고 시간 기간당 하나의 HTTP 요청을 이용하여 다운로드된다(예를 들어, 도 2에서 타일 3 및 타일 7).
- [0269] 도 3은 특정한 실시예에 따라 미디어 데이터 트랙들을 사용하는 초기화 세그먼트 파일의 블록도들의 일례를 도시한다.
- [0270] 도 3에 도시된 초기화 세그먼트 파일 및 도 4에 도시된 미디어 세그먼트 파일들은 여러 트랙들이 독립적으로 스트리밍되는 동안 ISO 베이스 미디어 파일 포맷(Base Media File Format)으로 비디오 비트 스트림을 재조직화 및 캡슐화하는 본 발명의 일 실시예를 도시한다. 설명의 목적상, 도 3 및 도 4의 내용은 도 1에 도시된 비디오 비트 스트림에 기초한 것이다.
- [0271] 여러 트랙들이 독립적으로 스트리밍되는 동안 ISO BMFF로 비디오 비트 스트림을 재조직화 및 캡슐화하기 위해, 타일 트랙이라 불리는 새로운 타입의 트랙이 정의된다. 타일 트랙은, 정의에 의하면, 샘플이 단일 타임스탬프와 관련된 모든 데이터를 나타내는 관련된 샘플들의 시간 설정형 시퀀스인 트랙이다. 샘플이 전형적으로 개별적인 비디오 프레임인 공지된 비디오 미디어 트랙과 대조적으로, 타일 트랙의 샘플은 풀 비디오 프레임의 공간적으로 정의된 서브-파트를 정의한다. 따라서, 타일 트랙들은 단지 주어진 타일과 관련된 NAL 유닛들만을 포함한다. 이런 방식으로, 각각의 트랙을 독립적인 세그먼트 파일들에 저장함으로써 연속적인 바이트 범위들을 갖는 타일들의 파이프들을 만드는 것이 가능하다.
- [0272] 그러나, 타일 트랙이 디코딩 및 재생되기 위해 요구되는 NAL 유닛들 및 정보 모두를 포함하지는 않기 때문에, "합성 트랙(composite track)"으로 불리는 추가적인 트랙이 사용된다. 특정한 실시예에 따르면, 합성 트랙은 완전한 타일형 프레임(즉, 모든 타일들의 조성물)을 나타내는 미디어 트랙이다. 합성 트랙은 도 1을 참조하여 설명된 바와 같은 그의 전형적인 조직화에서 인코딩된 시간 설정형 미디어 데이터 비트 스트림의 표현이다(프레임별로 디코딩 순서로). 합성 트랙은 그들 각자의 타일 트랙들에서 NAL 유닛들을 언급하기 위해 추출자 객체들(extractor objects)을 이용한다. 또한, 그것은 그러한 NAL 유닛들이 존재하면, 모든 타일들에 공통인 NAL 유닛들을 포함할 수 있다.
- [0273] 특정한 실시예에 따르면, 초기화 세그먼트 파일은 다른 미디어 세그먼트 파일들에서 캡슐화된 시간 설정형 미디어 데이터 비트 스트림들을 정의하기 위해 필요한 모든 메타데이터를 송신하는데 사용된다. 도 3에 도시된 바와 같이, 초기화 세그먼트 파일(300)은 파일 탑 박스 "ftyp"(302) 및 영화 박스 "moov"(304)를 포함한다. 파일 탑 박스(302)는 바람직하게는 세그먼트 파일들이 준수하는 ISO BMF 사양을 식별하고 해당 사양의 버전 넘버를 나타낸다. 영화 박스 "moov"(304)는 미디어 세그먼트 파일들에 저장된 프레젠테이션을 기술하는 모든 메타데이터 및 특히 프레젠테이션에 이용 가능한 모든 트랙들을 제공한다.
- [0274] 영화 박스 "moov"(304)는 도 1에서 일례로서 제공된 스케일 가능한 비디오 비트 스트림에 대응하는 트랙들("트랙" 박스들(306-1 내지 306-6)) 각각에 대한 정의를 포함한다.
- [0275] 트랙 박스(306-1)는 베이스 계층(track_ID=1)을 나타내고, 4개의 트랙 박스들(306-2 내지 306-5)(트랙 박스 306-3 및 트랙 박스 306-4는 도시되지 않음)은 향상 계층의 4개의 타일들 a, b, c 및 d를 나타내고(track_ID=2

내지 5), 트랙 박스 306-6은 향상 계층을 기술하는 합성 트랙을 나타낸다(트랙_ID=6).

[0276] 각각의 트랙 박스는 적어도 일반적으로 308로 참조되는 트랙 헤더 박스 "tkhd" 및 일반적으로 310으로 참조되는 트랙 미디어 박스 "mdia"를 포함한다. 트랙이 다른 트랙들로부터의 데이터에 의존하면, 트랙 참조 박스 "tref" 또한 존재한다. 도시된 바와 같이, 식별자 track_ID=6을 갖는 합성 트랙은, 트랙이 식별자 track_ID=1 내지 6을 갖는 트랙들로부터의 데이터에 의존한다는 것을 나타내는 트랙 참조 박스 "tref"(312)를 포함한다.

[0277] 다른 박스들은 시간 설정형 미디어 데이터 비트 스트림을 캡슐화하는데 사용된 ISO BMFF 사양들에 따라 의무적 이거나 선택적일 수 있다는 점에 주목해야 한다. 그러나, 본 발명의 실시예들은 적용 가능하기 위해 이러한 박스들에 의지하지 않으므로, 그들은 여기에서 제시되지 않는다.

[0278] 트랙 헤더 박스 "tkhd"(308)는 트랙의 특성을 명시한다. 정보의 여러 아이템들 중에, 그것은 트랙의 식별자 (track_ID), 트랙의 지속기간, 및/또는 트랙의 시작적 프레젠테이션 사이즈(즉, 디스플레이 영역의 폭과 높이)를 제공한다. 그것은 트랙이 재생 가능한지 여부를 나타내는 플래그 파라미터를 또한 포함한다.

[0279] 일 실시예에 따르면, 타일 트랙들에 대한 트랙 헤더 플래그의 디폴트 값은 0(track_enabled = 0, track_in_movie = 0, track_in_preview = 0)이고, 이는 타일 트랙들이 클라이언트 디바이스에 의해 로컬 재생 및 미리 보기 위해 무시된다는 것을 의미한다. 또 다른 실시예에서, 새로운 트랙 헤더 플래그가, 트랙이 타일 트랙이라는 것을 시그널링하기 위해 만들어질 수 있다.

[0280] 트랙 미디어 박스 "mdia"(310)는 트랙 내에 시간 설정형 미디어 데이터의 파라미터들을 공표하는데 사용된 모든 객체들을 포함하는 컨테이너로서 간주될 수 있다. 그것은 적어도 일반적으로 314로 참조된 미디어 헤더 박스 "mdhd", 일반적으로 316으로 참조된 핸들러 참조 박스 "hdlr", 및 일반적으로 318로 참조된 미디어 정보 박스 "minf"를 포함한다.

[0281] 핸들러 참조 박스 "hdlr"(316)는 트랙의 시간 설정형 미디어 데이터가 제시되는 프로세스, 및 따라서 트랙에서 시간 설정형 미디어 데이터의 본질을 공표한다. 예를 들어, 비디오 트랙은 비디오 핸들러에 의해 다루어질 것이다('vide'와 동일한 핸들러 탑 속성을 갖는다는 것에 유의). 비디오 샘플은 VisualSampleEntry() 탑의 객체를 이용하여 기술될 수 있다. 특정한 실시예에 따르면, 타일 핸들러('tile'과 동일한 핸들러 탑 속성을 갖는다는 것에 유의)로 불리는 새로운 핸들러 탑이 트랙이 공간적 서브샘플 정보를 포함한다는 것을 나타내기 위해 정의된다. 코딩 포맷에 따라, VisualSampleEntry() 탑의 객체가 타일 트랙에서 샘플을 기술할 수 없으면, 샘플을 기술하기 위해 TileSampleEntry0 탑의 특정 객체를 정의하는 것이 가능하다.

[0282] 미디어 정보 박스 "minf"(318)는 트랙에서 시간 설정형 미디어 데이터의 특성 정보를 정의하는 모든 객체들을 포함한다. 예를 들어, 합성 트랙에서 정의된 베이스 계층 및 향상 계층에 대해, "minf" 박스는 표준 비디오 미디어 헤더 박스 "vmhd"(320)를 포함할 수 있다.

[0283] 타일 트랙들과 관련하여, 새로운 타일 핸들러에 대응하는 타일 미디어 헤더 박스(Tile Media Header Box; tmhd, 322)로 언급된 특정 박스가 타일들에 대해 코딩에 독립적인 일반적 프레젠테이션 정보를 정의하는데 사용된다. 특히, 그것은 다음과 같이 정의될 수 있는 합성 비디오 트랙에 의해 표현된 비디오 해상도와 관련하여 타일에 의해 커버된 공간 영역의 기하학 정보를 포함할 수 있다:

```
aligned(8) class TileMediaHeaderBox
extends FullBox('tmhd', version = 0, 0) {
    unsigned int(16) horizontal_offset;
    unsigned int(16) vertical_offset;
}
```

[0284] [0285] 이전에 설명된 바와 같이, 합성 트랙은 프레젠테이션에서 또 다른 트랙에 대한 타이핑된 참조를 제공하는 특정 트랙 참조 박스 "tref"(312)를 포함한다. 특정한 실시예에 따르면, 그러한 타이핑된 참조는 합성 트랙으로부터 그것이 참조하는 타일 트랙으로의 링크를 설정하는데 사용될 수 있는 "tile" 참조(324) 및 이 참조를 포함하는 트랙으로부터 그것이 의존하는 시간 설정형 미디어 데이터 트랙(예컨대, 베이스 계층(track_ID=1))으로의

링크를 설정하는데 사용될 수 있는 "scal" 참조(326)를 포함할 수 있다.

[0286] 도 4는 도 3에 도시된 것과 같은 초기화 세그먼트 파일에서 공표된 트랙들에 대응하는 특정한 실시예에 따른, 타일 트랙들 및 하나의 합성 트랙을 포함하는 미디어 세그먼트 파일들에 대한 블록도들의 일례를 도시한다. 이전에 설명된 바와 같이, 도 3에 도시된 초기화 세그먼트 파일 및 도 4에 도시된 미디어 세그먼트 파일들은 여러 트랙들이 독립적으로 스트리밍되는 동안 ISO 베이스 미디어 파일 포맷(Base Media File Format)으로 비디오 비트 스트림을 재조직화 및 캡슐화하는 본 발명의 일 실시예를 도시한다.

[0287] 도 4에 도시된 바와 같이, 미디어 세그먼트 파일들(400-1 내지 400-6) 각각(미디어 세그먼트 파일들 400-3 내지 400-5는 도시되지 않음)은 DASH 표준에 명시된 바와 같이, 일반적으로 402로 참조된 세그먼트 타입 박스 "styp", 일반적으로 404로 참조된 적어도 하나의 영화 프래그먼트 박스 "moof", 및 일반적으로 406으로 참조된 적어도 하나의 미디어 데이터 박스 "mdat"를 포함한다. 미디어 세그먼트 파일은 HTTP-URL과 관련된다.

[0288] 다른 박스들은 시간 설정형 미디어 데이터 비트 스트림을 캡슐화하는데 사용된 ISO BMFF 사양에 따라 의무적이거나 선택적일 수 있다는 점이 주목되어야 한다. 그러나, 본 발명의 실시예들은 적용 가능하기 위해 이러한 박스들에 의지하지 않으므로, 그들은 여기에서 제시되지 않는다.

[0289] 세그먼트 타입 박스 "styp"(402)의 포맷은 도 3에서 파일 타입 박스 "ftyp"(302)의 포맷과 유사하지만, 그것의 참조는 파일이 미디어 세그먼트 파일이라는 것을 나타낸다.

[0290] 영화 프래그먼트 박스(404)는 일반적으로 영화 박스 "moov" 내에 저장되는 정보를 제공한다. 그의 헤더 ("mfhd")는 각각의 영화 프래그먼트에 대해 증가하는 시퀀스 넘버(도 4에서 seq_num로 표기됨)를 포함한다. 그러한 시퀀스 넘버는 클라이언트 디바이스로 하여금 수신된 세그먼트 파일들을 오름 차순으로 연접하고 (필요한 경우) 시퀀스의 무결성을 검증하도록 허용한다. 영화 프래그먼트 박스(404)는 관련된 미디어 데이터 박스 ("mdat", 406)에 데이터를 갖는 각각의 트랙에 대해 트랙 프래그먼트 박스 "traf"(일반적으로 408로 참조됨)를 포함한다. 트랙 프래그먼트 박스(408)는 대응하는 미디어 데이터 박스(("mdat", 406)에 존재하는 트랙의 비트 스트림의 식별자(track_ID)를 저장하는데 사용되는, 일반적으로 410으로 참조된, 트랙 프래그먼트 헤더 박스 "tfhd"를 포함한다.

[0291] 미디어 데이터 박스는 일반적으로 시간 설정형 미디어 데이터를 포함한다. 표준 비디오 트랙들에서, 그것은 비디오 프레임들을 포함한다. 타일 트랙들에서, 미디어 데이터 박스(406)는 완전한 비디오 프레임들의 공간적으로 정의된 서브-파트들을 포함한다. 설명의 목적상, 트랙 식별자 track_ID=2와 관련된 미디어 데이터 박스는 항상 계층의 타일에 대응하는 모든 NAL 유닛들을 포함한다.

[0292] 합성 트랙(도 4에서 track_ID=6)에서, 미디어 데이터 박스(406)는 각각의 타일에 대해 그리고 각각의 종속 계층에 대해 추출자들(도 4에서 E로 표기됨)을 포함하고 (존재한다면) 모든 타일들에 공통인 NAL 유닛들을 포함한다.

[0293] 도 4에 도시된 바와 같이, 합성 트랙과 관련된 미디어 세그먼트 파일(400-6)의 미디어 데이터 박스(406)는, 특히, 다음을 포함한다:

[0294] - 베이스 계층 트랙과 관련된 미디어 세그먼트 파일(400-1)의 미디어 데이터 박스(406)에 저장된 베이스 계층 트랙 내에 인코딩된 베이스 계층 데이터(NAL 유닛들(1BL))로의 링크를 제공하는 제1 추출자(412-1);

[0295] - 여러 타일들에 공통인 NAL 유닛들(412-2);

[0296] - 항상 계층의 제1 타일 트랙과 관련된 미디어 세그먼트 파일(400-2)의 미디어 데이터 박스(406) 내에 인코딩된 제1 타일의 항상 계층 데이터(NAL 유닛들(1a))로의 링크를 제공하는 제2 추출자(412-3);

[0297] - 항상 계층의 제2 타일 트랙과 관련된 미디어 세그먼트 파일(400-3)(도시되지 않음)의 미디어 데이터 박스(406) 내에 인코딩된 제2 타일의 항상 계층 데이터(NAL 유닛들(1b))로의 링크를 제공하는 제3 추출자(412-4);

[0298] - 항상 계층의 제3 타일 트랙과 관련된 미디어 세그먼트 파일(400-4)(도시되지 않음)의 미디어 데이터 박스(406) 내에 인코딩된 제3 타일의 항상 계층 데이터(NAL 유닛들(1c))로의 링크를 제공하는 제4 추출자(412-5); 및

[0299] - 항상 계층의 제4 타일 트랙과 관련된 미디어 세그먼트 파일(400-5)(도시되지 않음)의 미디어 데이터 박스(406) 내에 인코딩된 제4 타일의 항상 계층 데이터(NAL 유닛들(1d))로의 링크를 제공하는 제5 추출자(412-6).

[0300] 추출자(412-1) 덕분에 획득될 수 있는 NAL 유닛들은 항상 계층이 추출자들(412-3 내지 412-6) 덕분에 획득될 수 있는 NAL 유닛들 및 NAL 유닛들(412-2)을 사용하여 완전히 디코딩될 수 있는 프레임의 베이스 계층의 디코딩을 허용한다. 도 4로부터 관찰될 수 있는 바와 같이, 단지 프레임의 공간적 파트만 디코딩된다면, 모든 미디어 세그먼트 파일들(400-2 내지 400-5)(즉, 타일 트랙들에 대응하는 비트 스트림들)을 다운로드할 필요는 없다.

[0301] 특정한 실시예에 따르면, 추출자는 다음과 같은 구문을 갖는 파일 포맷 내부 구조이다:

```
class aligned(8) Extractor () {
    NALUnitHeader();
    unsigned int(8) track_ref_index;
    signed int(8) sample_offset;
    unsigned int((lengthSizeMinusOne+1)*8)
        data_offset;
    unsigned int((lengthSizeMinusOne+1)*8)
        data_length;
}
```

[0302]

[0303] 여기서, NALUnitHeader()는 비디오 비트 스트림을 인코딩하는데 사용된 코딩 포맷을 따르는 NAL 유닛의 처음 4 바이트들을 나타낸다. 이러한 4 바이트들은 NAL 유닛을 추출자로서 식별한다(예컨대, SVC에서, 속성 nal_unit_type는 추출자 NAL 유닛 타입(타입 31)으로 설정된다).

[0304] track_ref_index 값은, 데이터가 추출되는 트랙을 찾기 위해, 합성 트랙의 'scal' 또는 'tile' 유형의 트랙 참조 박스 "tref"에서, 이용될 인덱스를 명시한다. sample_offset 값은 정보의 소스로서 사용되는 링크된 트랙에서 샘플의 상대적인 인덱스를 제공한다. data_offset 값 및 data_length 값은 각각 복사하기 위한 참조 샘플 내에서 처음 바이트의 오프셋 및 복사하기 위한 바이트들의 수이다.

[0305] 설명의 목적상, 도 3을 참조하면, 주어진 추출자의 track_ref_index의 값이 2와 동일하면, 이것은 추출자가 tref 박스에서 두 번째 엔트리에 의해 식별된 트랙을 참조한다는 것을 의미한다(즉, 트랙은 타일에 대한 타일 트랙인 식별자 track_ID=2를 갖고, 제1 인덱스는 참조 트랙(예컨대, 베이스 계층)을 나타냄).

[0306] 도 5는 (여기서 2개의 연속적인 프레임에 대응하는) 주어진 시간 기간 동안 연속적인 비디오 프레임들의 공간적 파트를 나타내는 유효 디코드 가능한 시간 설정형 미디어 데이터 비트 스트림을 구축하기 위해 다운로드된 미디어 데이터 세그먼트들을 연접시키는 일례를 도시한다. 다른 시간 기간 동안 동일한 도면이 반복될 수 있다.

[0307] 도 3 및 도 4를 참조하여 설명된 바와 같이, 시간 설정형 타일형 미디어 데이터 비트 스트림은 바람직하게는 하나의 초기화 세그먼트 파일 및 복수의 미디어 세그먼트 파일들을 포함하는 데이터 세트로서 송신되는데, 후자는 여러 타일 트랙들 및 하나의 합성 트랙을 포함한다.

[0308] 초기화 세그먼트 파일은 각각의 트랙에 대한 일반적 정보, 특히 트랙의 타입(예컨대, 미디어 트랙(오디오 또는 비디오) 또는 타일 트랙), 코딩 포맷, 프레임 해상도 및 트랙들 간의 종속성(트랙 참조 박스 "tref"에서 주어짐)을 제공하는 영화 박스("moov")를 포함한다. 이러한 데이터는 다운로드된 미디어 세그먼트 파일들을 처리하는데 사용된다. 도 1, 3 및 4를 참조하여 설명된 예를 참고로 하면, 초기화 세그먼트 파일의 영화 박스의 내용은 특히 다음을 포함할 수 있다:

MOOV

- *track 1: base layer*
- *track 2: tile a*
- *track 3: tile b*
- *track 4: tile c*
- *track 5: tile d*
- *track 6: enhancement layer*
 - *tref (scal): track_ID = 1*
 - *tref(tile): track_ID=2*
track_ID=3
track_ID=4
track_ID=5

[0309]

[0310] 도 5는 요청된 미디어 세그먼트 파일들(여기서 베이스 계층 및 향상 계층에서의 타일들 a 및 c에 대응)만이 서버로부터 다운로드될 때에 미디어 세그먼트들을 연접시킴으로써 획득된 파일 포맷을 대략적으로 도시한다. 그러한 메커니즘은 단지 요청된 미디어 세그먼트 파일들만의 다운로딩을 허용할 뿐만 아니라 복제 데이터의 다운로딩도 방지한다는 점이 주목되어야 한다.

[0311]

도시된 바와 같이, 합성 트랙(500)은, (스케일 가능한 경우에는) 베이스 계층 트랙(504)으로부터의 데이터 및 재생 불가능한 타일 트랙들(506 및 508)로부터의 데이터를 참조하는 것에 의해 그리고 (도 7을 참조하여 설명된 바와 같이) 누락 데이터를 참조하는 추출자들을 적절히 다루는 것에 의해 유효 디코드 가능한 시간 설정형 미디어 데이터 비트 스트림(502)의 구축을 허용한다.

[0312]

획득된 파일 포맷은 스케일 가능한 파일 포맷 정의를 따른다. 예를 들어, 클라이언트 디바이스는 베이스 계층 트랙을 선택함으로써 단지 베이스 계층만을 재생하도록 결정할 수도 있고 또는 합성 트랙을 선택함으로써 선택된 타일들 a 및 c를 갖는 향상 계층을 재생하도록 결정할 수 있다. 클라이언트 디바이스는 또한 그것이 계속 합성 트랙을 재생하는 동안 더 나중의 시간 기간에 상이한 "타일 트랙들"(즉, 미디어 세그먼트 파일들)을 다운로드함으로써 디스플레이될 타일들을 바꿀 수 있다.

[0313]

도 6은 (a) 및 (b)를 포함하고, 특정한 실시예에 따라 서버와 클라이언트 디바이스 사이에 시간 설정형 미디어 데이터를 송신하기 위한 단계들을 도시하는 흐름도이다. 도 6의 (a)에 도시된 단계들은, 도 6의 (b)에 도시된 단계들이 클라이언트 디바이스 내에서 구현되는 동안 타일형 시간 설정형 미디어 데이터 비트 스트림으로부터의 ROI 스트리밍에 적합한 세그먼트 파일들을 만드는 것에 의해 미디어 프레젠테이션을 준비하기 위해 서버 내에 구현된다.

[0314]

제1 단계(단계 600)에서, 서버는 타일들과 관련되는 모든 NAL 유닛들을 식별하고, 각각의 타일에 대해, 주어진 타일에 대응하는 모든 NAL 유닛들로 구성된 서브-샘플들을 포함하는 타일 트랙을 만든다. 예를 들어, 서버는, HEVC 표준화(제안 JCTVC-K0128)에서 제안된 바와 같이, 상이한 영역들과 NAL 유닛들의 관련성을 식별하기 위해 서브-픽처 레벨 SEI 메시지들에 의지할 수 있고, 각각의 ROI의 위치 및 사이즈를 식별하기 위해 시퀀스-레벨 SEI 메시지들에 의지할 수 있다. 따라서, 서버는 주어진 시간 기간 동안 타일들의 파일들을 만들 수 있다.

- [0315] 다음의 단계(단계 602)에서, 서버는 (존재한다면) 하위 레벨 스케일 가능한 트랙들에 링크된 추출자들, 모든 타일 들에 공통인 NAL 유닛들, 및 각각 타일 트랙에 링크된 추출자들을 포함하는 합성 트랙을 만든다. 추출자들 및 공통 NAL 유닛들은, 그들이 참조하고 있는 데이터로 추출자들을 대체하여, (도 1을 참조하여 설명된 바와 같이) 디코딩 순서의 완전한 샘플들로 구성된 유효 시간 설정형 미디어 데이터 비트 스트림을 초래하도록 유리하게 순서화된다.
- [0316] 다음으로, 도 3 및 4를 참조하여 설명된 바와 같이, 단계 604에서, 서버는 ISO BMFF 표현에 따라 시간 기간을 포함하는 초기화 세그먼트 파일 및 미디어 세그먼트 파일들을 생성 및 저장한다. 모든 시간 설정형 미디어 데이터 트랙들(예컨대, 비디오 트랙들), 합성 트랙들, 및 타일 트랙들은 별도의 미디어 세그먼트 파일들에 저장된다.
- [0317] 그 다음, 서버는 요청시에 초기화 및 미디어 세그먼트 파일들을 클라이언트 디바이스에게 제공한다(단계 606). 서버는 HTTP 요청들에 응답하는 종래의 HTTP 서버일 수 있다.
- [0318] HTTP 스트리밍의 정황에서, 바람직한 실시예에서는, 클라이언트 디바이스가 서버로부터 이용 가능한 미디어 프레젠테이션을 기술하는 매니페스트 파일에 액세스하는 것으로 가정된다. 이 매니페스트 파일은 서버에 먼저 초기화 세그먼트들을 요청한 다음 미디어 세그먼트 파일들을 요청함으로써 미디어 프레젠테이션을 스트리밍하기 위해 클라이언트 디바이스에게 충분한 정보(미디어 속성을 및 세그먼트들의 리스트)를 제공한다.
- [0319] 클라이언트 디바이스 종단에서 ROI의 선택시, 전형적으로, 포인팅 디바이스와 같은 선택 수단을 갖는 디스플레이상에서는, 타일형 비디오의 스트리밍 동안, 선택된 ROI에 대응하는 타일들이 결정된다(단계 608).
- [0320] 다음으로, 각각 시간적 기간 동안, 스케일 가능한 미디어 데이터의 경우에, 클라이언트 디바이스는 종속 계층들에 대응하는 세그먼트 파일들을 다운로드하기 위해 서버로 요청을 보낸다(단계 610). 특정한 실시예에 따르면, 종속되는(dependent from) 계층들은 종속되는 계층들에 종속하는(depending from) 계층들 이전에 다운로드된다. 예를 들어, 베이스 계층 세그먼트 파일들은 항상 계층 세그먼트 파일들 이전에 다운로드된다.
- [0321] 다음의 단계에서, 클라이언트 디바이스는 선택된 타일들에 대응하는 미디어 세그먼트 파일들을 다운로드하기 위해 서버로 요청을 보내고(단계 612), 합성 트랙에 대응하는 미디어 세그먼트 파일들을 다운로드하기 위해 서버로 요청을 보낸다(단계 614).
- [0322] 다음으로, 다운로드된 세그먼트 파일들은 선택된 ROI에 대응하는 ISO BMFF 사양에 따르는 유효 (디코드 가능한) 시간 설정형 미디어 데이터 비트 스트림을 구축하기 위해 클라이언트 디바이스에 의해 연접된다(단계 616).
- [0323] 도 6에 도시된 흐름도가 트랙들 사이의 링크들을 설정하기 위해(즉, 종속성을 정의하기 위해) 추출자들을 이용하는 것에 기초한다면, 트랙들 사이의 링크들은 도 18, 19 및 20을 참조하여 설명된 바와 같이 상이하게 시그널링될 수 있다.
- [0324] 단계 616(트랙들 사이의 링크들을 설정하기 위해 추출자들이 이용되는 경우)은 도 7과 관련하여 상세히 설명된다.
- [0325] 도 7은 클라이언트 디바이스에 의해 수신된 연접된 미디어 세그먼트들로부터 유효 시간 설정형 미디어 데이터 비트 스트림의 생성, 바꾸어 말하면, 합성 트랙이 클라이언트 디바이스에 의해 재생될 때 선택된 타일들 및 하나의 합성 트랙으로부터의 디코드 가능한 비트 스트림의 생성을 도시하는 흐름도이다.
- [0326] 제1 단계(단계 700)에서, 클라이언트 디바이스는 (예를 들어, 도 6의 단계들 612, 614 및 616에서) 이전에 요청되었던 미디어 세그먼트 파일들을 수신하고, 적어도 하나의 미디어 세그먼트 파일이 수신되었는지 여부를 판정하도록 테스트가 수행된다(단계 702). 어떤 미디어 세그먼트 파일도 수신되지 않았다면, 프로세스는 종료된다.
- [0327] 적어도 하나의 미디어 세그먼트 파일이 수신되었다면, 그리고 수신된 미디어 세그먼트 파일들이 합성 트랙을 포함하지 않는다면(즉, 그들이 종속 트랙들, 예를 들어, 하위 레벨 계층 또는 타일 트랙들을 포함함), 그들은 추후의 사용을 위해 베퍼링된다.
- [0328] 반대로, 적어도 하나의 미디어 세그먼트가 수신되었다면, 그리고 수신된 미디어 세그먼트 파일이 합성 트랙을 포함한다면, 수신된 미디어 세그먼트의 미디어 데이터 박스 "mdat"로부터 데이터의 제1 아이템(또는 수신된 미디어 세그먼트의 데이터의 적어도 하나의 아이템이 처리된 경우, 데이터의 다음 아이템, 통상적으로 NAL 유닛)을 추출하기 위해 합성 트랙에 대응하는 비트 스트림(통상적으로 NAL 유닛들)이 파싱된다(단계 704). 미디어 데이터 박스에서 파싱할 추가의 데이터 아이템이 없는 경우, 프로세스는 단계 700으로 복귀하여 새로운 미디어

세그먼트 파일들을 수신한다(단계 706).

[0329] 다음으로, 데이터의 추출된 아이템(예를 들어, 추출된 NAL 유닛)이 추출자에 대응하는지 여부를 판정하기 위해 테스트가 수행된다(단계 708). 데이터의 추출된 아이템이 추출자에 대응하지 않는 경우, 그것은 비디오 디코더에 의해 더 디코딩되도록 리턴된다(단계 710). 반대로, 데이터의 추출된 아이템이 추출자인 경우, 그것은 그가 참조하는 데이터의 아이템에 의해 대체되어야만 한다. 그러한 목적을 위해, 추출자의 파라미터들의 값들이 그의 구조로부터 획득된다(단계 712). 전술한 바와 같이, 추출자는 다른 트랙으로부터 데이터를 추출하기 위해 요구되는 모든 파라미터 값들을 포함한다(예를 들어, track_ref_index, sample_offset, data_offset, 및 data_length).

[0330] 참조된 트랙의 식별자가 식별되면, 참조된 트랙이 단계 700 동안에 버퍼링된 미디어 세그먼트 파일들의 세트에서 이용 가능한지 여부를 판정하기 위해 테스트가 수행된다(단계 714). 클라이언트 디바이스는 선택된 관심 영역(Region-of-Interest)에 대응하는 미디어 세그먼트 파일들만을 다운로드하기 때문에 일부 파일 트랙들이 누락 된다는 점을 상기한다.

[0331] 참조된 트랙이 단계 700 동안에 버퍼링된 미디어 세그먼트 파일들의 세트에서 이용 가능하다면, 추출자는 그가 참조하는 데이터에 의해 대체되고(단계 716), 비트 스트림은 디코딩되기 위해 비디오 디코더에 송신된다(단계 710).

[0332] 참조된 트랙이 단계 700 동안에 버퍼링된 미디어 세그먼트 파일들의 세트에서 이용 가능하지 않다면, 추출자에서 참조된 데이터의 부재가 ISO BMF 표준에 따른 치명적 오류를 야기하기 때문에 특정 단계들이 수행되어야 한다. 참조된 트랙이 파일 트랙인지 여부(참조된 트랙은 종속 스케일 가능성 계층에 대응할 수 있음) 및 추출자가 파일 탑입인지 여부를 판정하기 위해 테스트가 수행된다(단계 718).

[0333] 참조된 트랙이 파일 트랙이 아니거나 추출자가 파일 탑입이 아니라면, 표준적인 치명적 오류가 검출된다. 반대로, 참조된 트랙이 파일 트랙이고 추출자가 파일 탑입이라면, 추출자는 제거되거나(단계 722), 추출자는, 시간 설정형 미디어 데이터 비트 스트림을 인코딩(단계 720)하는데 사용된 코딩 포맷에 따라, 누락 파일들에 대한 '스킵된(skipped)' 데이터를 포함하는 대안의 "패딩 트랙(padding track)" 또는 "패딩 박스(padding box)"로부터의 패딩에 의해 대체된다(단계 724). 여기서, '스킵된' 데이터는, 동일한 스케일 가능성 계층에 속하거나 다른 스케일 가능성 계층에 속하는 이전에 디코딩된 이미지로부터 획득된 다른 픽셀 데이터에 의해 대체되는, 현재의 이미지 내에서 누락되는 픽셀 데이터를 나타낸다. '스킵된' 데이터는 일반적으로 적어도 하나의 플래그에 의해 표현된다. 예를 들어, HEVC 비디오 압축 포맷을 고려하는 경우, 패딩 데이터는 1로 설정된 스kip 플래그를 이용하여 인코딩된 코딩 유닛들을 배타적으로 포함하는 하나 이상의 NALU들일 수 있다.

[0334] 다음으로, 비트 스트림이 비디오 디코더에 송신되어 디코딩되고(단계 710) 디스플레이되며, 프로세서는 데이터의 후속 아이템을 핸들링하기 위해 단계 704로 루프된다.

[0335] 도 8은 하나 이상의 실시예들의 단계들이 구현되는 서버 또는 클라이언트 디바이스(800)의 블록도를 나타낸다.

[0336] 바람직하게는, 디바이스(800)는 통신 버스(802), 디바이스의 전원 인가 시에 프로그램 ROM(806)으로부터의 명령어들, 및 전원 인가 후에 메인 메모리(808)로부터의 소프트웨어 애플리케이션에 관련된 명령어들을 실행할 수 있는 중앙 처리 유닛(CPU)(804)을 포함한다. 메인 메모리(808)는 예를 들어, 통신 버스(802)를 통해 CPU(804)의 작업 영역으로서 기능하는 RAM(Random Access Memory) 탑입이고, 그의 메모리 용량은 확장 포트(도시되지 않음)에 접속되는 선택적 RAM에 의해 확장될 수 있다. 소프트웨어 애플리케이션에 관련된 명령어들이, 예를 들어, 하드 디스크(hard-disc)(HD)(810)로부터 또는 프로그램 ROM(806)으로부터 메인 메모리(808)에 로딩될 수 있다. CPU(804)에 의해 실행되는 경우, 그러한 소프트웨어 애플리케이션은 도 6a를 참조하여 설명된 단계들이 서버에서 수행되고 도 6b 및 도 7을 참조하여 설명된 단계들이 클라이언트 디바이스에서 수행되게 한다.

[0337] 참조 번호 812는 통신 네트워크(814)로의 디바이스(800)의 접속을 허용하는 네트워크 인터페이스이다. CPU(804)에 의해 실행되는 경우, 소프트웨어 애플리케이션은 네트워크 인터페이스를 통해 수신된 요청들에 반응하고, 데이터 스트림들 및 요청들을 네트워크를 통해 다른 디바이스들에 제공하도록 구성된다.

[0338] 참조 번호 816은 사용자에게 정보를 디스플레이 및/또는 사용자로부터 입력을 수신하기 위한 사용자 인터페이스들을 나타낸다.

[0339] 여기서, 변형으로서, 멀티미디어 비트 스트림의 수신 또는 송신을 관리하기 위한 디바이스(800)는 도 6a, 6b 및 7을 참조하여 설명한 바와 같은 방법을 구현할 수 있는 하나 이상의 전용 집적 회로들(ASIC)로 이루어질 수 있

다는 것을 유의해야 한다. 이러한 집적 회로들은, 예를 들면 그리고 비-제한적으로는, 비디오 시퀀스들을 생성 또는 디스플레이 및/또는 오디오 시퀀스들을 청취하기 위한 장치 내로 통합된다.

[0340] 전술한 바와 같이, 본 발명의 실시예는, 특히, HEVC로서 알려진 비디오 포맷에 적용할 수 있다.

[0341] HEVC 표준에 따르면, 이미지들은 타일들, 슬라이스들, 및 슬라이스 세그먼트들로 공간적으로 나누어질 수 있다. 이러한 표준에서, 타일은 수평 및 수직 경계들(즉, 행(row)들 및 열(column)들)에 의해 정의되는 이미지의 사각형 영역에 대응한다. 이것은 정수개의 코딩 트리 유닛들(Coding Tree Units)(CTU)을 포함한다. 따라서, 타일들은, 예를 들어, 관심 영역들에 대한 위치 및 크기를 정의함으로써 관심 영역들을 식별하는데 효율적으로 사용될 수 있다. 그러나, HEVC 비트 스트림의 구조뿐 아니라 NAL(Network Abstract Layer) 유닛들로서의 그것의 캡슐화는 타일의 관점에서 조직화되지 않고 슬라이스들에 기반한다.

[0342] HEVC 표준에서, 슬라이스들은 슬라이스 세그먼트들의 세트들이고, 슬라이스 세그먼트들의 세트의 제1 슬라이스 세그먼트는 독립 슬라이스 세그먼트이며, 즉, 헤더 내에 저장된 일반 정보가 다른 슬라이스 세그먼트의 것을 참조하지 않는 슬라이스 세그먼트이다. 슬라이스 세그먼트들의 세트의 다른 슬라이스 세그먼트들은, 만약에 있다면, 종속 슬라이스 세그먼트들이다(즉, 헤더 내에 저장된 일반 정보가 독립 슬라이스 세그먼트의 것을 참조하는 슬라이스 세그먼트들).

[0343] 슬라이스 세그먼트는 정수개의 연속적인(래스터 스캔 순서로) 코딩 트리 유닛들(Coding Tree Units)을 포함한다. 따라서, 슬라이스 세그먼트는 사각형 형상일 수도 또는 아닐 수도 있고, 따라서, 관심 영역을 나타내기에 적절하지 않다. 이것은 슬라이스 세그먼트 헤더와 그에 후속하는 슬라이스 세그먼트 레이터의 형식 하에서 HEVC 비트 스트림으로 인코딩된다. 독립 및 종속 슬라이스 세그먼트들은 그들의 헤더에서 차이가 있다: 종속 슬라이스 세그먼트는 독립 슬라이스 세그먼트에 종속하기 때문에, 그의 헤더의 정보량은 독립 슬라이스 세그먼트의 것보다 작다. 독립 및 종속 슬라이스 세그먼트들은 타일들을 정의하는데 사용되거나 엔트로피 디코딩 동기화 포인트들(entropy decoding synchronization points)로서 사용되는 대응하는 비트 스트림에서의 엔트리 포인트들의 리스트를 포함한다.

[0344] 도 9는 도 9a, 도 9b, 및 도 9c를 포함하고, 타일들 및 슬라이스 세그먼트들의 예들을 예시한다. 보다 정확하게는, 도 9a는 수직 경계들(905-1 및 905-2)과 수평 경계들(910-1 및 910-2)에 의해 9개의 부분들로 나누어진 이미지(900)를 예시한다. 915-1 내지 915-9로 참조된 9개의 부분들은 특정한 타일을 나타낸다.

[0345] 도 9b는 수직 경계(905')에 의해 구분된 2개의 수직 타일들을 포함하는 이미지(900')를 예시한다. 이미지(900')는 1개의 독립 슬라이스 세그먼트(920-1)(빗금으로 나타냄)와 4개의 종속 슬라이스 세그먼트들(920-2 내지 920-5)의 5개의 슬라이스 세그먼트들을 포함하는 싱글 슬라이스(참조되지 않음)를 포함한다.

[0346] 도 9c는 수직 경계(905")에 의해 구분된 2개의 수직 타일들을 포함하는 이미지(900")를 예시한다. 좌측 타일은 2개의 슬라이스들을 포함한다: 제1 슬라이스는 1개의 독립 슬라이스 세그먼트(920'-1) 및 1개의 종속 슬라이스 세그먼트(920'-2)를 포함하고 제2 슬라이스는 또한 1개의 독립 슬라이스 세그먼트(920'-3) 및 1개의 종속 슬라이스 세그먼트(920'-4)를 포함한다. 우측 타일은 1개의 독립 슬라이스 세그먼트(920'-5) 및 1개의 종속 슬라이스 세그먼트(920'-6)를 포함하는 하나의 슬라이스를 포함한다.

[0347] HEVC 표준에 따라, 슬라이스 세그먼트들은 이하와 같이 요약될 수 있는 규칙들에 따라 타일들에 링크된다(하나 또는 양쪽 조건들이 만족되어야 함):

[0348] - 슬라이스 세그먼트 내의 모든 CTU들은 동일한 타일에 속한다(즉, 슬라이스 세그먼트는 여러 타일들에 속할 수 없다); 및

[0349] - 타일 내의 모든 CTU들은 동일한 슬라이스 세그먼트에 속한다(즉, 타일은 이들 슬라이스 세그먼트들 각각이 오직 그 타일에만 속하는 경우 여러 슬라이스 세그먼트들로 나누어질 수 있다).

[0350] 전술한 바와 같이, 타일들이 관심 영역들에 대한 적절한 지원으로서 간주될 수 있는 한편, 슬라이스 세그먼트들은 통신 네트워크를 통한 송신을 위해 NAL 유닛들 내에 실제로 삽입되고 액세스 유닛들(즉, 파일 포맷 레벨에서의 코딩된 팩쳐(coded picture) 또는 샘플들)을 형성하기 위해 집성되는 엔티티들이다.

[0351] HEVC 표준에 따라서, NAL 유닛의 태입은 다음과 같이 정의될 수 있는 2 바이트들의 NAL 유닛 헤더 내에 인코딩 된다는 것을 상기한다:

```

nal_unit_header () {
    forbidden_zero_bit
    nal_unit_type
    nuh_layer_id
    nuh_temporal_id_plus1
}

```

[0352]

슬라이스 세그먼트들을 코딩하는데 사용된 NAL 유닛들은 슬라이스 세그먼트 어드레스 구문 요소 덕분에 슬라이스 세그먼트 내의 제1 CTU의 어드레스를 나타내는 슬라이스 세그먼트 헤더를 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){
        [...]
}

```

[0354]

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

[0356]

공간적 예측들이 (정의에 의해) 타일 경계를 상에서 리셋되지만, 아무것도 타일이 참조 프레임(들) 내의 상이한 타일로부터의 시간적 예측들을 이용하는 것을 막지 못한다. 따라서, 독립 타일(들)을 구축하기 위해, 예측 유닛들을 위한 모션 벡터들이 유리하게도, 인코딩 동안에, 참조 프레임(들) 내에 동위치의 타일에 남아 있도록, 타일 내부에 제약된다. 또한, 인-루프(in-loop) 필터들(디블로킹(deblocking) 및 샘플 적응형 오프셋(sample adaptive offset)(SAO) 필터들)은 바람직하게는 타일 경계를 상에서 비활성화되어 하나의 타일만을 디코딩하는 경우에 어떤 에러 드리프트(error drift)도 도입되지 않는다. 그러한 인-루프 필터들의 제어가 HEVC 표준에서 이용 가능하다는 것을 유의해야 한다. 이것은 *loop_filter_across_tiles_enabled_flag*로서 알려진 플래그를 이용하여 슬라이스 세그먼트 헤더에서 설정된다. 이 플래그를 0으로 명시적으로 설정함으로써, 타일 경계들에서의 픽셀들은 이웃 타일들의 경계 상에 속하는 픽셀들에 종속될 수 없다. 모션 벡터들 및 인 루프 필터들과 관련된 이러한 2개의 조건들이 만족되는 경우, 타일들은 "독립적으로 디코딩 가능한 타일들" 또는 "독립 타일들"로서 간주될 수 있다.

[0357]

비디오 비트 스트림이 독립 타일들의 세트로서 인코딩되는 경우, 이것은 참조 데이터의 누락 또는 재구성 에러들의 전파에 대한 임의의 위험 없이 하나의 프레임으로부터 다른 프레임으로의 타일 기반 디코딩을 가능하게 한다. 그리고 나서, 이러한 구성은, 예를 들어, (타일들 3 및 7을 포함하는) 도 2에 예시된 관심 영역에 대응할 수 있는 원래 비디오의 공간적 부분만을 재구성하는 것을 가능하게 한다. 그러한 구성은 타일 기반 디코딩이 신뢰할 수 있다는 것을 나타내기 위해 비디오 비트 스트림 내의 보충 정보로서 나타내어질 수 있다.

[0358]

HEVC 타일들에 대한 상이한 가능한 구성들, 전형적으로 타일당 슬라이스의 수, 슬라이스당 타일의 수, 코딩 종

속성들, 및 슬라이스 세그먼트들의 탑입(독립 또는 종속)을 다루기 위해서는, 타일 디코딩 순서 종속성들이 서버 측에서 (압축 단계 후에) 캡슐화 단계 동안에 고려되어야 한다. 이것은 클라이언트 디바이스가 수신된 캡슐화된 파일을 파싱한 후에 비디오 디코더에서 처리될 수 있는 유효 비디오 기본 스트림을 구축하는 위치에 있기 위해 중요하다.

[0359] 두 가지 탑입의 종속성이 공존한다는 것에 유의해야 한다: 디코딩 종속성들 및 디코딩 순서 종속성들.

[0360] 첫 번째 탑입의 종속성은 항상 계층이 베이스 계층에 종속하고 단독으로 디코딩될 수 없는 스케일 가능한 비디오 코딩에서 빈번히 사용된다. 또 다른 경우는, 예를 들어 도 10, 13, 및 14를 참조하여 설명된 바와 같이 또는 다수의 뷰를 코딩하기 위해, 상이한 트랙들로 분열되는 데이터와 관련된다. 그러한 경우에, 상이한 샘플들로부터 오는 데이터는 디코딩하고 디스플레이할 하나의 샘플을 구축하기 위해 집성될 수 있다. 그러나, 이 집성은 무작위 순서로 행해질 때 오류를 야기할 수 있다.

[0361] 설명을 위해서, ISO/IEC 14496-15의 "MVC Track structure" 섹션은 상이한 뷰들로부터 샘플을 구축하는 방법을 보여준다. 마찬가지로, ISO/IEC 14496-1에는 단일 객체 설명자에 기본 스트림 설명자들의 집성을 위한 규칙들이 존재한다. 첫 번째 경우에는, 종속성이 'scal' 트랙 참조 탑입으로 시그널링되는 반면, 두 번째 경우에는, 명시적 플래그 *dependsOn_ES_ID* 가 이 기본 스트림이 종속하는 다른 기본 스트림의 식별자를 나타낸다. 이 *dependsOn_ES_ID*의 개념은 mp4 파일 포맷(ISO/IEC 14496-14)에서 그의 정확한 처리 시맨틱들이 MPEG-4 시스템들(ISO/IEC 14496-1)에 의해 주어지는 간단한 'dpnd' 트랙 종속성으로 변환되었다. 그것들은 다중 트랙 데이터 간의 종속성 시그널링을 위한 필요한 피처들 중 일부를 커버하지만, 'dpnd' 및 'scal'은 (상위 계층으로부터 하위 계층으로) 하향 종속성들만을 기술하고 디코딩 순서가 일부 계층들 사이에 요구되지만 다른 계층들 사이에는 결정되지 않은 복잡한 경우들은 다루지 못한다. ISO 베이스 미디어 파일은 일부 트랙들이 디코딩에 관하여 다른 트랙들에 종속하고, 따라서 그것들의 샘플들은 그것들이 종속하는 트랙의 샘플들 뒤에 디코딩되어야 한다는 것을 나타내기 위한 일반 도구가 없다.

[0362] 실제로, 도 21a에 예시된 것과 같은 종래 기술의 종속성 그래프 - 이에 따르면 미디어 파일(2100)은 T1 및 참조 번호 2101 내지 2105로 표시된 5개의 트랙을 포함하고, 트랙 T2(2102), T3(2103) 및 T4(2104)는 트랙 T1(2101)에 종속하지만 서로 종속하지 않고, 트랙 T5(2105)는 T3(2103) 및 T4(2104)에 종속하도록 종속성들을 가짐 - 를 가정하면, 종속성 참조들('scal' 또는 'dpnd')은, 어떤 순서 개념도 없이, 참조 번호들 2112, 2113, 2114, 2115, 및 2116으로 표현되는 것으로 설정된다. 그러한 경우, 트랙 T5(2105)의 디코딩은 다음과 같은 종속성 그래프를 제공한다: T5->[T3, T4] 또는 [T4, T3]-> T1. 알 수 있는 바와 같이, 그러한 메커니즘에 의하면 트랙 T2(2102)는 종속성 그래프에 없을 것이고, 따라서 이를 종속성으로부터 완전한 스트림의 재구성은 문제가 있다. 그것은 T2->T1 종속성(2112)을 발견하기 위해 모든 트랙의 스캔을 암시하고 그럼에도 T2가 처리되어야 하는 순서를 놓친다.

[0363] 두 번째 탑입의 종속성(도 21b에서 숫자 1 내지 3으로 나타내어진, 아래에서 "상향 트랙 참조"라고 불리는 디코딩 순서 종속성을 의미함)에 관하여, 많은 사용 경우들은 각각의 스케일 가능 레벨이 전용 트랙에 저장되고, 각각의 계층이 단일 하위 계층에 종속하는, 예를 들어, 간단한 스케일 가능성(공간, SNR, 시간)과 같이 트랙 처리의 정확한 순서가 명시되는 것을 요구한다. 각각의 계층은 올바른 비디오 기본 스트림 재구성을 위해, 베이스로부터 시작하여, 순서대로 mp4 파일에 전달되어야 한다. 두 번째 사용 경우는 HEVC에서의 독립 타일들과 유사하게, 비디오 스트림에서 공간 영역들의 독립 디코딩이다. 이 경우에, 각 타일은 임의의 순서로 독립적으로 처리될 수 있지만 비-VCL NAL 유닛들을 포함하는 트랙 뒤에 처리되어야 한다. 관심 영역과 관련된 또 다른 사용 경우는 관심 영역들의 SNR 구체화들(refinements)과 같은 고급 스케일 가능성이고, 여기서는 여러 영역들이 정의되고 서로 독립적으로 디코딩될 수 있다. 마지막으로, 베이스 계층과 향상 계층들이 동일한 미디어 포맷을 사용하지 않는 스케일 가능 하이브리드 코딩이다(SHVC 설계와 유사).

[0364] HEVC의 특정한 경우에 대해, 그 사양은, 특히 도 14를 참조하여 설명된 바와 같이 데이터가 상이한 트랙들로 분열되는 경우, 슬라이스 세그먼트들을 나타내는 NAL 유닛들 간의 디코딩 순서를 나타낸다. 따라서, 본 발명의 목적은 공간 부분들의 송신이 클라이언트 측에서의 파싱 후에 유효하고 디코딩 가능한 비디오 기본 스트림으로 이어지도록 캡슐화 계층에서 이 종속성들을 시그널링하는 것과 관련된다.

[0365] 본 발명의 실시예에 따르면, HTTP 스트리밍의 정황에서의 타일들에 대한 효율적 액세스가 HEVC 표준에 적용된 ISO BMFF 파일 포맷을 이용하여 제공된다. 따라서, 코딩될 독립 타일들의 각각(예를 들어, 도 2에 나타낸 12개의 타일들의 각각)은 도 10을 참조하여 이하에 설명되는 "타일 트랙"으로 지칭되는 특정 트랙에 의해 나타내어진다.

[0366] 이러한 파일 트랙들은 도 10에 예시된 바와 같은 풀 프레임 내의 HEVC 비트 스트림에 대응하는 합성 트랙에서 (트랙들 각각에 대한 정의를 포함하는 영화 박스 'moov'의 트랙 참조 박스 'tref'를 통해) 참조된다. 각각의 파일 트랙은 NAL 유닛들로 패키징된, 압축된 비디오 데이터를 포함한다. 합성 트랙은 초기화 데이터에 대응하는 다양한 파라미터 세트들(예를 들어, 비디오 파라미터 세트, 시퀀스 파라미터 세트, 및/또는 픽처 파라미터 세트)을 포함한다. 이것은 또한 특정 타입의 NAL 유닛들인 추출자들을 포함한다.

[0367] 전술한 바와 같이, 추출자는 이하의 구문을 갖는 파일 포맷 내부 구조일 수 있다:

```
class aligned(8) Extractor() {
    NALUnitHeader();
    unsigned int(8) track_ref_index;
    signed int(8) sample_offset;
    unsigned int((lengthSizeMinusOne+1)*8)
        data_offset;
    unsigned int((lengthSizeMinusOne+1)*8)
        data_length;
}
```

[0368]

[0369] 추출자는 다른 트랙들로부터의 데이터에 대한 포인터 또는 참조들로서 기능하며 양 트랙들에서의 데이터 복제 대신에 종속 트랙들을 참조하여 촘촘한 트랙들을 구축하는 것을 가능하게 한다. 추출자는 바람직하게는 NAL 유닛 구문을 사용한다. 따라서, 이것은, 특히, NAL 유닛 타입에 속하는 정보를 포함하는 NAL 유닛 헤더와 동일한 구조를 갖는 헤더를 포함한다. 이 NAL 유닛 타입은, 예를 들어, HEVC의 예약된 NAL 유닛 타입에 현재 대응하는 값 '47'로 설정된다. 추출자에 의해 참조되는 트랙에 대응하는 트랙(*track_id*)의 식별자를 포함하는 tref 박스의 엔트리를 검색하는 것을 허용하는 인덱스(*track_ref_index*로 표기됨)가 트랙 참조 박스(*tref*) 내에서 헤더에 후속한다. 제3 파라미터는 현재 샘플과 비교하여 추출자에 의해 참조되는 샘플의 시간 오프셋 (*sample_offset*)이다. 제4 및 제5 파라미터들(*data_offset* 및 *data_length*로 표기됨)은 각각 복사할 곳으로부터의 위치 (바람직하게는 바이트 단위로) 및 복사할 데이터량을 제공한다(값 0은 전체 참조된 NAL 유닛의 복사를 나타내기 위해 예약된다).

[0370] 도 10은 본 발명의 실시예에 따라, 합성 트랙 및 독립 파일 트랙들을 포함하는 트랙들의 세트로서 HEVC 비트 스트림을 캡슐화하는 예를 예시한다. 예시를 위해, 캡슐화된 비트 스트림은 도 2에 개략적으로 예시된 비디오 시퀀스에 대응한다.

[0371] 예시된 바와 같이, 캡슐화된 비트 스트림(1000)은 합성 트랙(1015) 및 12개의 파일 트랙들(1020-1 내지 1020-12)(파일 트랙들(1020-1 내지 1020-12) 각각은 비디오 시퀀스의 하나의 파일과 관련됨)을 나타내는 미디어 세그먼트 파일(1010) 및 트랙들에 대한 정의를 제공하는 영화 박스('moov')를 포함하는 초기화 세그먼트 파일(1005)을 포함한다.

[0372] 합성 트랙(1015)은, DASH 표준에 특정된 바와 같이, 세그먼트 타입 박스 'styp'(도시되지 않음), 트랙 세그먼트 타입 및 식별자와 같은 메타데이터를 포함하는 적어도 하나의 영화 프래그먼트 박스(movie fragment box) 'moof'(1025), 및 각각의 비디오 데이터 샘플들에 대해, 비디오 데이터에 대한 참조들 및 PPS를 포함하는 적어도 하나의 미디어 데이터 박스 "mdat"(1030)을 포함한다.

[0373] 마찬가지로, 파일 트랙들(1020-1 내지 1020-12)의 각각은 세그먼트 타입 박스 "styp"(도시되지 않음), 트랙 세그먼트 타입 및 식별자와 같은 메타데이터를 포함하는 적어도 하나의 영화 프래그먼트 박스 "moof", 및 NAL 유닛들(NALUs)로 패키징된 압축된 비디오 데이터를 포함하는 적어도 하나의 미디어 데이터 박스 "mdat"를 포함한다.

[0374] 식별자(2 내지 13)를 갖는 파일 트랙들(1020-1 내지 1020-12)은 초기화 세그먼트 파일(1005)의(보다

정확하게는, 식별자 *id*=1을 갖는 합성 트랙의 정의에서, 초기화 세그먼트 파일(1005)의 영화 박스 "moov"의) 트랙 참조 박스 "tref"(1035)에서 참조된다.

[0375] 예시된 바와 같이, 합성 트랙(1015)은 다른 트랙들로부터의 데이터에 대한 포인터들 또는 참조들로서 기능하는 추출자들을 포함한다. 예시를 위해, 합성 트랙(1015)의 추출자들(1035-1 및 1035-p)에 대응하는 타일 트랙의 인덱스(*track_ref_index*), 데이터 오프셋(*data_offset*), 및 데이터 길이(*data_length*) 중 여러 파라미터들이 표현된다.

[0376] 또한 예시를 위해, 합성 트랙(1015)의 NAL 유닛(1035-1)이 처리되는 경우, 이것은 추출자 타입의 NAL 유닛을 나타낸다고 판정된다(*NALUnitHeader*는 16진수 값 5E00과 같다). 따라서, 이것은 대응하는 압축된 비디오 데이터를 복구하기 위해 처리된다. 그러한 목적을 위해, 그의 타일 트랙 인덱스(즉, *track_ref_index*=1)가 획득된다. 이러한 인덱스로부터, 초기화 세그먼트 파일(1005)에 저장된 트랙 정의들로부터 타일 트랙 식별자를 복구하는 것이 가능하다. 주어진 예에서, 인덱스가 1과 같기 때문에, tref 박스의 제1 타일 트랙 식별자가 선택된다(*id*=2). 다음으로, 이러한 식별자가 사용되어 대응하는 타일 트랙을 액세스하고, 그 후, 추출자(1035-1)의 데이터 오프셋(즉, 정보 소스로서 사용되는 식별된 트랙에서의 샘플의 상대적 인덱스) 및 데이터 길이(즉, 복사할 바이트 수, 예를 들어, *data_length*=0인 경우, 전체 NALU) 파라미터를 사용하여, 압축된 비디오 데이터가 타일 트랙(1020-1)으로부터 추출된다(즉, 주어진 예에서 코딩된 슬라이스 세그먼트 NALU(1040)).

[0377] 처리된 후에, 추출자는 그것이 참조하는 데이터에 의해 대체된다. 도 10에서 예시된 예에 따라, 추출자(1035-1)의 파싱 및 처리는 코딩된 슬라이스 세그먼트 NALU(1040)에 의한 그의 대체를 야기하고, 따라서, HEVC 준수 비트 스트림을 형성한다.

[0378] HEVC 추출자의 파라미터들을 저장하기 위해 사용된 시맨틱은 SVC 표준에 정의된 것에 가까울 수 있다는 것에 유의해야 한다. 따라서, HEVC NAL 유닛들을 참조하는 추출자들에 대해, 다음이 적용될 수 있다:

- *forbidden_zero_bit*로서 알려진 파라미터는 ISO/IEC 23008-2에 특정된 바와 같이 설정된다;

[0380] - *nal_unit_type*로서 알려진 파라미터는 47(HEVC에 대한 FDIS(Final Draft International Standard)에서의 예약된 코드)로 설정된다;

[0381] - *nuh_layer_id* 및 *nuh_temporal_id_plus1*으로서 알려진 파라미터들은 추출자에 의해 참조된 제1 NALU로부터 복사된다(HEVC NAL 유닛들을 참조하는 HEVC 트랙에서의 추출자는 상이한 *nuh_layer_id* 및 *nuh_temporal_id_plus1* 값을 갖는 여러 NAL 유닛들을 참조하지 않는다); 및

[0382] - *sample_offset*으로 알려진 파라미터는 0으로 설정된다.

[0383] 본 발명의 실시예들에 따르면, 타일 트랙들은 공간적으로 배열된 데이터의 인덱스화(indexation)을 지원하는데 사용되는 특정 트랙들로서 간주된다(예를 들어, HEVC 타일들). 따라서, 특히, 완전한 비디오에서의 타일의 위치(수평 및 수직 좌표들) 및 타일의 크기(폭 및 높이)를 나타내고, 타일 트랙이 임의의 결함(artifact) 없이 디코딩될 수 있는지 여부를 나타내는 정보를 제공하기 위해 각각의 트랙에 대해 특정 시그널링이 요구된다. 그러한 시그널화(signalization)는 여러 실시예들에 따라 달성될 수 있다.

[0384] 도 10을 참조하여 설명한 것과 상이한, 특정 실시예가 도 15 및 16에 예시되어 있다.

[0385] 도 15는 타일들을 설명하기 위해 서브 트랙 피처(즉, 서브 트랙 박스들)를 이용하는 예를 예시한다. 이 실시예에 따르면, HEVC 기본 비트-스트림으로부터 1502로 참조되는 하나의 비디오 트랙만이 구축된다. 타일 설명 및 그룹화 정보는, 1503으로 참조되는 사용자 데이터와 함께, 바람직하게는 캡슐화된 파일의 사용자(또는 클라이언트)에게 선택들을 제공하는 데 이용되는 1504로 참조되는 트랙 선택 박스와 동일한 곳에 저장될 수 있다.

[0386] ISO 베이스 미디어 파일 포맷은 (전체) 트랙들이 대체 및 스위치 그룹들에 할당될 수 있는 것과 동일한 방식으로 대체 및 스위치 그룹들에 트랙들의 부분들을 할당하는 데 이용될 수 있는 서브 트랙들, 예를 들어 1505 및 1506으로 참조되는 서브 트랙들을 설명하기 위한 박스들을 정의한다는 것을 상기해야 한다. 다시 말해서, 서브 트랙들은 상이한 타일들 또는 상이한 타일들의 상이한 버전들을 설명하는 데 이용될 수 있다.

[0387] 도 15의 박스들을 이용해 예시된 바와 같이 서브 트랙 설명을 위해 일반적 구문이 제공된다(현재의 실시예에 따라 새로운 박스들(Tile Sub Track Group 박스들)인 1512 및 1513은 제외). ISO 베이스 미디어 파일 포맷은 이 일반적 박스들이 각 미디어 포맷(예를 들어 SVC 및 MVC)에 대해 전문화되어야 한다는 것을 나타낸다.

[0388] 이 실시예에서는, 입력 비디오 기본 스트림이 타일형 HEVC 비트-스트림인 것이 고려된다. 그러므로, 타일들을

적절히 캡슐화하기 위해서는(즉, 처음과 마지막 바이트의 위치들뿐만 아니라 위치, 사이즈, 및 디코딩 정보와 같은 타일 파라미터들을 가짐),

[0389] - 각각의 공간 타일에 대응하는 비디오 기본 스트림의 NAL 유닛들로부터 1507로 참조되는 데이터를 식별하는 것이 요구된다. 이것은 처음 바이트와 마지막 바이트의 위치들에 의해 나타내어질 수 있다;

[0390] - 각각의 서브 트랙(1505 및 1506)에 각각의 타일 설명(전형적으로 위치, 사이즈 및 독립 인코딩 같은 타일 파라미터들)을 링크하는 것이 요구된다; 그리고

[0391] - 사용자 타일 선택 정보를 제공하는 것, 특히 타일들과 서브 트랙들 간의 가능한 조합들을 나타내는 것이 요구된다.

[0392] 도 16은 서브 트랙들로 타일들을 캡슐화하는 것을 설명하는 흐름도이다. 그것은 비디오 기본 스트림을 mp4 세그먼트들로 캡슐화하기 위해 (비디오 캡처시에 또는 나중에 스트리밍 서버로의 업로드를 위해 콘텐츠를 준비할 때 차별 없이 캡슐화가 행해지기 때문에 표준 서버 또는, 예를 들어, 비디오 캡처 디바이스일 수 있는) 서버 측에서 수행되는 단계들을 나타낸다.

[0393] mp4 세그먼트는 (예를 들어, 도 11에 예시된 바와 같이) 헤더 부분과 데이터 부분으로 이루어진 어드레스 가능 형 엔티티라는 것을 상기해야 한다. mp4 세그먼트들은 가변적일 수 있는 또는 시퀀스를 따라 동일할 수 있는 지속기간을 가진다. 명확성을 위해, 여기서는 세그먼트 지속기간이 시간을 따라 일정하고 데이터의 캡슐화를 담당하는 서버의 초기화 파라미터로서 제공된다고 가정한다. 그러나, 가변적인 지속기간들을 다루는 것이 도 16에 예시된 단계들을 변화시키지는 않을 것이다.

[0394] 예시된 바와 같이, 캡슐화 프로세스의 제1 단계는 제1 mp4 세그먼트의 초기화를 포함한다(단계 1600).

[0395] 전술한 바와 같이 그리고 이 특정 실시예에 따르면, 캡슐화는 비디오 기본 스트림을 나타내기 위해 (도 12b를 참조하여 설명된 바와 같은) 단일 비디오 트랙을 생성하는 것에 있다.

[0396] 다음 단계에서, 서버는 NAL 유닛 헤더를 찾는다(단계 1601). 어떤 NAL 유닛도 발견되지 않으면(즉, 테스트 1601의 결과가 거짓), 이것은 비디오 기본 스트림의 끝에 도달했다는 것을 의미한다. 이것은 캡슐화 프로세스를 종료한다.

[0397] 그와 반대로, NAL 유닛이 발견되면(즉, 테스트 1601의 결과가 참), 다음 단계(단계 1603)에서 그것의 타입을 체크하여 그것이 하이 레벨 정보(즉, 비-VCL(Video Coding Layer) 데이터)를 포함하는지 아닌지(즉, VCL 데이터를 포함하는지)를 결정한다. 현재의 NAL 유닛이 비-VCL 데이터를 포함한다면, 테스트 1303의 결과는 거짓이지만 그것이 VCL 데이터를 포함한다면, 테스트 1303의 결과는 참이다.

[0398] 서버는, NAL 유닛들을 찾고 처리하는 동안, 처리되는 비디오 프레임에 대응하는 현재의 타임스탬프 값을 유지한다는 것에 유의해야 한다. 초기 타임스탬프 값은 0으로 설정되고 새로운 비디오 프레임이 검출될 때마다 업데이트된다(예를 들어, *PicOrderCntVal*을 계산하는 데 사용되는 구문 파라미터들 같은 HEVC 하이 레벨 구문 요소들로부터).

[0399] 예시된 바와 같이, 비-VCL 및 VCL NAL 유닛들 모두의 데이터가 단계 1604 및 1605에서 비디오 트랙의 데이터 부분에 샘플들로서 저장된다. 비-VCL NAL 유닛들에 대해, 서버는 현재 시간에 대응하는 타임스탬프 값을 가진 트랙의 데이터 부분에 샘플을 생성한다(단계 1604에서). 다음으로, 그것은 이 NAL 유닛의 특정 타입을 보고 그것이 타일링 구성에 관한 정보를 제공할 수 있는지를 결정한다(단계 1606에서). 설명을 위해, 이것은 NAL 유닛의 타입이 타일들의 수평 및 수직 경계들을 명시하는 *PPS_NUT* 와 같은 것이거나 NAL 유닛 타입을 가진 특정 SEI 메시지가 비디오의 타일들 및/또는 관심 영역들 또는 독립적으로 인코딩된 공간 영역들에 관한 정보를 제공하는 *PREFIX_SEL_NUT* 와 같은 것일 수 있다.

[0400] 그러한 NAL 유닛이 이용 가능할 때, 서버는 (아래에 단계 1610 및 1611에 관련하여 설명된) 비디오 데이터의 향후 설명을 위해 단계 1607에서 획득된 타일링 구성을 저장한다. 이것은 타일 설명자에서, 예를 들어, 도 12b(박스 1275)에 관련하여 설명된 바와 같은 *TileRegionSampleGroupEntry*라는 박스에서 행해진다. 그것은 또한 특히, 타일들 또는 독립적으로 인코딩된 공간 부분들을 설명하기 위해 사용되는 대응하는 서브 트랙들(도 15에서 1505 및 1506으로 참조됨) - 각각의 타일 또는 공간 부분에 대해 하나의 서브 트랙이 사용됨 - 을 생성한다(단계 1608).

[0401] 보다 정확하게는, 단계 1608은 현재의 서브 트랙이 타일링 정보를 제공하는 특정 트랙이라는 것을 나타내기 위

한 서브 트랙 정보 박스를 준비하는 것을 목표로 한다. 이것은 도 15에서 1508 및 1509로 참조되는, 기존의 SubtrackInformationBox('stri') 박스의 속성들의 현재의 리스트를 확장하는 것에 의해 행해질 수 있다. 현재, 이 박스는 서브 트랙들 간의 스위칭 및/또는 대체를 설명하지만 원본 비디오의 임의의 공간 영역을 표시하기 위해 함께 구성할 수 있는 비디오의 공간 부분들을 설명하는 것은 가능하지 않다.

[0402] 동시에, 서버는 그것이 비디오의 공간 부분들을 렌더링하기 위한 하나 이상의 타일을 선택할 수 있다는 것을 사용자 디바이스에게 보여주기 위한 트랙 선택 박스(도 15에서 1504로 참조됨)를 생성할 수 있다. 이것은 현재 지원되지 않기 때문에 트랙 선택 박스의 일부 수정들을 요구한다. 특정 실시예에 따르면, 트랙 선택 박스에 대한 attribute_list 내의 값 'ntls' 차별화 속성뿐만 아니라 동일한 attribute_list 내의 새로운 설명 속성으로서 'tile' 값이 부가된다. 여기서의 명칭들은 예시로서 제공되며 ISO 베이스 미디어 파일 포맷 및 그것의 확장들에서의 기존 코드들과 상충하지 않는다면 임의의 4-문자 코드가 선택될 수 있다는 것에 유의한다. 첫 번째 것은 트랙이 타일들의 수에 관하여 스케일될 수 있다는 것을 가능하게 하는 반면 두 번째 값은 서브 트랙 내의 타일들의 수를 나타낸다. 서브 트랙 정보 및 트랙 선택 레벨들에서의 'ntls' 차별화 속성은 미디어 프레젠테이션이 타일 해상도들/수에 관하여 대체 트랙들을 포함한다면 적절할 수 있다. 그러면 원하는 타일들의 해상도/수(공간 액세스 세분성(spatial access granularity)으로도 알려짐)가 선택될 수 있다.

[0403] 유사하게, 대체 그룹의 트랙들이 풀-프레임 비디오(ntls=0) 및 동일한 비디오의 타일형 표현(ntls=N)을 포함할 때, 그 2개의 트랙은 그들의 타일 수에 관하여 구별된다. 서브 트랙 정보 박스 내의 가능한 속성들의 리스트는 또한 현재의 서브 트랙이 비디오 트랙에서 공간 액세스를 제공하는 특정 서브 트랙이라는 것을 나타내기 위한 첫 번째 새로운 값 'tile'로 확장된다.

[0404] 도 16으로 돌아가서, VCL NAL 유닛들에 대해, 단계 1605 동안에 비디오 트랙의 데이터 부분에 데이터가 저장되면, 서버는 단계 1609에서 현재의 비디오 데이터(전형적으로 HEVC에서의 슬라이스 세그먼트)에 대해 트랙의 헤더 부분을 생성한다. 그 목적을 위해, 그것은 현재의 NAL 유닛의 바이트 위치들을 트랙 헤더(도 3에서 참조 번호 308)에 또는 트랙 프래그먼트 헤더(도 4에서 참조 번호 410)에, 예를 들어 샘플 사이즈 박스 'stsz'에 또는 트랙 런(Track Run) 박스('trun')에 저장한다. 그것은 또한 데이터를 참조하기 위해 전용 박스들에 타이밍 정보를 저장한다(예를 들어 'ctts'를 샘플링하기 위한 구성 시간 및 'stts' 박스들을 샘플링하기 위한 디코딩 시간).

[0405] 다음으로, 타일링 정보가 존재하고(즉, 테스트 1606이 참일 때) 저장될 때(단계 1607), 현재의 NAL 유닛이 유래하는 타일이 식별된다(단계 1610). HEVC 비디오 기본 스트림에 대해, 이 식별은 HEVC 슬라이스 세그먼트 헤더에 제공된 슬라이스 세그먼트 어드레스에 기초할 수 있다. 슬라이스 세그먼트 어드레스가 주어지고 타일 구성을 알면, 타일 인덱스가 추정될 수 있다. 식별될 때, 현재의 샘플은 식별된 타일에 대해 단계 1607에서 기록된 타일 설명자와 관련된다(단계 1611).

[0406] 이것은 서브샘플 그룹화 메커니즘에 있고, 서브샘플은 현재의 NAL 유닛에 의해 전달되는 슬라이스 세그먼트이다. 이 슬라이스 세그먼트는 전체 타일을 커버하거나(그러한 경우는 도 9a에 예시된 것에 대응한다) 타일의 부분일 수 있다(즉, 타일은 둘 이상의 슬라이스 세그먼트를 포함할 것이고, 그러한 경우는 도 9b 또는 9c에 예시된 것들에 대응한다). 이 관련 또는 링킹 단계 1611은 또한 단계 1608에서 준비된 서브 트랙을 업데이트한다.

[0407] 단계 1611은 서브 트랙들(도 15에서 1505 및 1506으로 참조됨)과 데이터(도 15에서 1507로 참조됨) 간의 매핑을 제공하는 것을 목표로 한다. 이것은 SubTrackDefinitionBox ('strd', 도 15에서 참조 번호 1514 및 1515)라고 알려진 박스에서 행해질 수 있다. 이 SubTrackDefinitionBox 박스는 포맷 특정 서브 트랙들에 대한(예를 들어, SVCSubTrackLayerBox 또는 MVCSubTrackViewBox) 그리고 다음과 같이 정의되는 옵션의 샘플 그룹화 정보 SubtrackSampleGroupBox (ISO/IEC 14496-12:2012 표준의 섹션 8.14.6에서의 'stsg')에 대한 컨테이너이다:

[0408] `aligned(8) class SubTrackSampleGroupBox extends FullBox('stsg', 0, 0){`

[0409] `unsigned int(32) grouping_type;`

[0410] `unsigned int(16) item_count;`

[0411] `for(i = 0; i < item_count; i++)`

[0412] `unsigned int(32) group_description_index;`

}

[0414] 타일 서브 트랙들의 특정한 경우에, *SubtrackSampleGroupBox* 는 바람직하게는 다음을 나타낸다:

[0415] - *grouping_type*으로서: *TileRegionSampleGroupEntry* 를 나타내는 값(타일 영역 정보(Tile Region InFormation)를 위한 'trif' 또는 도 12b의 참조 번호 1275로 예시된 'trsg'). 이 4-문자 코드가 무엇이든지, 그것은 샘플 그룹화의 종류 즉, 이 특정 실시예에 따라, 샘플들 또는 서브샘플들의 타일로의 그룹화를 식별하기 위해 고유할 것이다;

[0416] - *group_description_index*로서: 현재의 타일에 대한 tileID 또는 groupID의 값. 이 값은 도 12b에 관련하여 설명된 바와 같이, 트랙의 타일들을 설명하기 위해 *SampleGroupDescriptionBox* 에 기록된다.

[0417] 추가적인 그룹화 정보가 서브 트랙을 타일들의 그룹으로서 정의하기 위해 제공될 것이다. 새로운 *TileSubTrackGroupBox* (또는 *HEVCTileSubTrackGroupBox*)가 다음과 같이 정의될 수 있다:

HEVC Tile Sub Track Group box

Definition

[0420] *Box Type*: 'tstb'

[0421] *Container*: Sub Track Definition box ('strd')

[0422] *Mandatory*: No

[0423] *Quantity*: Zero or more

Syntax

```
[0425] aligned(8) class TileSubTrackGroupBox extends FullBox('tstb', 0, 0) {
[0426]     unsigned int(16) item_count;
[0427]     for(i = 0; i < item_count; i++) {
[0428]         unsigned int(16) GroupId;
[0429]     }
[0430] }
```

[0431] 이 박스에서 *GroupId* 들의 결합은 이 박스에 의해 정의된 서브 트랙을 설명하고, *item_count* 는 이 박스에서 리스팅된 타일 그룹들의 수를 카운트하고, *GroupId* 는 이 서브 트랙에 포함된 타일 그룹의 식별자를 나타낸다.

[0432] 도 16으로 돌아가서, 일단 링크되면, 서버는 현재의 타임스탬프가 미리 정의된 세그먼트 지속기간보다 큰지 여부를 결정하는 것(단계 1612)에 의해 현재의 세그먼트가 완료되어야 하는지를 체크하는 것에 의해 캡슐화를 계속한다. 현재의 타임스탬프가 미리 정의된 세그먼트 지속기간보다 크다면, mp4 세그먼트는 서버 상에 스트리밍을 위해 준비된 것으로 저장되고(단계 1613) 1601에서 다음 NAL 유닛들을 처리하기 위해 1600에서 새로운 세그먼트가 생성된다. 세그먼트 지속기간에 도달하지 않았다면, 서버는 현재의 세그먼트의 데이터 부분을 계속 구성하기 위해 1601에서 다음 NAL 유닛을 고려한다.

[0433] 도 17은 비디오의 다수의 타일을 커버하는 관심 영역의 예를 도시한다. 예시된 바와 같이, 새로운 박스 *TileSubTrackGroupBox*(또는 *HEVCTileSubTrackGroupBox*)가 1700으로 참조되는 비디오 시퀀스에서 1701로 참조되는 관심 영역을 설명하는 데 유용할 수 있고, 그 관심 영역은 (1710 내지 1720으로 참조되는) 비디오 프레임들의 세트에서 검출되고 그것은 (1702 내지 1705로 참조되는) 둘 이상의 타일을 커버할 것이다.

[0434] 이 경우, 미리 정의된 ROI를 설명하기 위한 서브 트랙이 생성된다. 관련된 *SubTrackDefinitionBox* 는 (*groupID*의 리스트를 통하여) 어느 타일들이 서브 트랙에 관련된 미리 정의된 ROI에 의해 커버되는지를 나타내기 위한 *TileSubTrackGroupBox*(또는 *HEVCTileSubTrackGroupBox*)를 포함한다. 그러한 관심 영역 및 대응하는 커버된 타일들은 HEVC 비디오 기본 스트림 내의 SEI 메시지에서 제공되고 단계들 1604 내지 1608 및 1610 및 1611에 의해 처리된다.

[0435] 전술한 것과 유사한, 도 11에 예시된 특정 실시예에 따라, 시그널링이 트랙 레벨에서 수행될 수 있다.

- [0436] 도 11은, 제1 실시예에 따라, HEVC 비트 스트림을 캡슐화하기 위해 미디어 데이터 트랙들을 이용하는 초기화 세그먼트 파일의 블록도의 예를 예시한다. 이것은 도 10을 참조하여 전술한 예에 기초한다.
- [0437] 초기화 세그먼트 파일(1100) 및 미디어 세그먼트 파일들(표현되지 않음)은 ISO 베이스 미디어 파일 포맷에서의 HEVC 표준을 준수하는 비디오 비트 스트림을 재조직화(reorganize) 및 캡슐화하는데 사용된다. 트랙들은 독립적으로 스트리밍된다.
- [0438] 전술한 바와 같이, 다른 미디어 세그먼트 파일들 내의 캡슐화된 시간 설정형 미디어 데이터 비트 스트림을 정의하는데 필요한 모든 메타데이터를 송신하는데 초기화 세그먼트 파일이 사용된다. 도 11에 예시된 바와 같이, 초기화 세그먼트 파일(1100)은 파일 타입 박스 'ftyp'(1105) 및 영화 박스 'moov'(1110)을 포함한다. 파일 타입 박스(1105)는 바람직하게는 세그먼트 파일들이 어느 ISO BMF 사양을 준수하는지를 식별하고 그 사양의 버전 넘버를 나타낸다. 영화 박스 'moov'(1110)는 미디어 세그먼트 파일들에 저장된 프레젠테이션을 설명하는 모든 메타데이터 그리고, 특히, 프레젠테이션에서 이용 가능한 모든 트랙들을 제공한다.
- [0439] 영화 박스(1110)는, 주어진 예에서, 1개의 합성 트랙(1115-1) 및 12개의 타일 트랙들(1115-2 내지 1115-13)을 포함하는 트랙들('track' 박스들(1115-1 내지 1115-13)) 각각에 대한 정의를 포함한다.
- [0440] 각각의 트랙 박스는, 일반적으로 1120으로 참조된, 적어도 트랙 헤더 박스 및 일반적으로 1125로 참조된 트랙 미디어 박스 'mdia'를 포함한다. 트랙이 다른 트랙들로부터의 데이터에 종속하면, 트랙 참조 박스 'tref'도 존재한다. 설명된 바와 같이, 식별자 track_ID=1을 갖는 합성 트랙은 트랙이 식별자들 track_ID=2 내지 13을 갖는 타일 트랙들로부터의 데이터에 종속한다는 것을 나타내는 트랙 참조 박스 'tref'(1130)를 포함한다.
- [0441] 전술한 바와 같이, 다른 박스들은 시간 설정형 미디어 데이터 비트 스트림을 캡슐화하는데 사용된 ISO BMFF 사양에 따라 의무적이거나 선택적일 수 있다. 그러나, 본 발명의 실시예들은 적용 가능하기 위해 이러한 박스들에 의존하지 않기 때문에, 그들은 여기서 제시되지 않는다.
- [0442] 도 11을 참조하여 설명된 실시예에 따르면, 풀 비디오에서의 타일의 위치, 타일의 크기, 및 타일 트랙이 어떤 결합도 없이 디코딩될 수 있다는 표시의 시그널화는, 트랙 헤더 박스 'tkhd'(1120) 및 미디어 정보 박스 'mdia'(1125)의 박스들을 이용하여, 각각의 트랙 정의에서, 'moov' 박스(1110)에서, 캡슐화될 전체 HEVC 비트 스트림에 대해 한번 수행된다.
- [0443] 타일들의 위치들은, 수평 및 수직 오프셋들(horizontal_offset 및 vertical_offset)을 정의하는 *TileMediaHandlerEntry* 박스 또는 'tmhd'(1140)으로서 지정되는, 새로운 타입의 미디어 헤더 정보 박스(1135)에 위치된다.
- [0444] 고려된 타일의 시각적 프레젠테이션의 크기는 기존의 트랙 헤더의 폭 및 높이 파라미터들로 표현된다. 요구되는 경우, 고려된 타일의 실제 픽셀 치수들이 (*SampleDescriptionBox*로서 알려진 박스 내의 *VisualSampleEntry()*로서 알려진 파라미터를 통해) 샘플 설명들로 문서화될 수 있다.
- [0445] 마지막으로, 특정 트랙(타일 트랙)을 시그널링하는 것은 핸들러 박스 'hdlr'(1145) 내에 새로운 타입의 트랙 ('tile')을 정의함으로써 명시적으로 행해진다.
- [0446] 초기화 세그먼트 파일(1100)의 영화 박스 "moov"(1110)는 'mvex' 박스(1150)를 더 포함한다. 이 박스는 캡슐화된 파일에 액세스하는 클라이언트에게 영화 프래그먼트들이 존재한다는 것을 통지하는데 사용된다. 이것은 프레젠테이션에서 가장 긴 트랙의 지속기간을 초기화 세그먼트 파일 내에 특정하는 것을 가능하게 한다. 이것은 각각의 영화 프래그먼트 지속기간의 검사를 회피하면서 프레젠테이션 지속기간의 연산을 더 간단하게 한다. 예시된 바와 같이, mvex 박스(1150)는 각각의 트랙(즉, 타일 트랙들 및 합성 트랙)의 모든 프래그먼트들에 공통인, 예를 들어, 트랙 식별자들 및 트랙에서의 샘플들의 디폴트 크기의 정보의 복제를 회피하기 위해 트랙마다 트랙 확장 박스를 포함한다.
- [0447] 타일 트랙들을 시그널링하는 그러한 실시예는 통상의 비디오 트랙 시그널링에 비교하여 현저한 오버헤드를 도입하지 않는다는 점에 유의해야 한다. 또한, 이것은 프레젠테이션이 분열되는 영화 프래그먼트들의 수와 무관함을 유지한다. 그러나, 이것은 비디오 시퀀스에 따른 변화하는 타일링 구성에 적응되지 않는다.
- [0448] 비디오 시퀀스를 따른 타일링 구성에 있어서의 변화를 핸들링하도록 구성되는 다른 특정 실시예에 따르면, 시그널링은, ISO BMFF 표준으로부터의 샘플 그룹화 메커니즘을 이용하여, 샘플 레벨에서 행해진다.
- [0449] 그러한 샘플 그룹화 메커니즘들은 트랙들에서의 샘플들의 분할을 나타내기 위해 사용된다. 그들은, 샘플 그룹

들에 대한 샘플들의 할당을 기술하는 *SampleToGroup* 박스('sbgp') 및 특정 샘플 그룹 내의 샘플들의 공통 특성을 기술하는 *SampleGroupDescription* 박스('sgpd')의 2개의 박스들의 사용에 의존한다. 특정 타입의 샘플 그룹화가 타입 필드('grouping_type')를 통해 하나의 *SampleToGroup* 박스와 하나의 *SampleGroupDescription* 박스의 조합에 의해 정의된다. 다수의 샘플 그룹화 인스턴스들(즉, *SampleToGroup* 및 *SampleGroupDescription* 박스들의 쌍)이 상이한 그룹화 범주에 기초하여 존재할 수 있다.

[0450] 본 발명의 실시예들에 따라, 샘플들의 타일링에 관련된 새로운 그룹화 범주가 정의된다. 'tile'로 지정되는 이러한 새로운 grouping_type은 타일의 특성을 기술하고 표준 *VisualSampleGroupEntry*로부터 도출된다. 이것은 *TileRegionSampleGroupEntry* *HEVC_SpatialEntry* 또는 *TileRegionGroupEntry*라고 지정될 수 있고(명칭이 무엇이든 그것이 기준의 *ScalableGroupEntry* 또는 *MultiviewGroupEntry*와 일관성을 유지하기 위한 고유의 박스 식별자 코드(전형적으로 ISO BMFF에서 4-문자 코드)를 가진다면) 다음과 같이 정의된다:

```
[0451] class TileRegionGroupEntry () extends VisualSampleGroupEntry ('trsg') {
[0452]     unsigned int(16) tileID; (or groupID)
[0453]     unsigned int(2) independent;
[0454]     unsigned int(6) reserved=0;
[0455]     unsigned int(16) horizontal_offset;
[0456]     unsigned int(16) vertical_offset;
[0457]     unsigned int(16) region_width;
[0458]     unsigned int(16) region_height;
[0459]     // other boxes
[0460] }
```

[0461] 이러한 새로운 타입의 그룹 엔트리에 따르면, *tileID* (or *groupID*) 파라미터는 그룹에 의해 설명되는 타일에 대한 고유 식별자이다. *horizontal_offset* 및 *vertical_offset* 파라미터들은, 베이스 영역의 루마 샘플들에서, HEVC 프레임의 상부 좌측 픽셀에 대하여, 타일에 의해 표현되는 직사각형 영역의 상부 좌측 픽셀의 수평 및 수직 오프셋을 설정하는데 각각 이용된다. *region_width* 및 *region_height* 파라미터들은, HEVC 프레임의 루마 샘플들에서, 타일에 의해 표현되는 직사각형 영역의 폭 및 높이를 설정하는데 각각 이용된다. *independent* 파라미터는, 독립 타일들의 정의를 참조하여 전술한 바와 같이, 동일한 타일에만 속하는 샘플들에 관련된 디코딩 종속성들을 타일이 포함하는 것을 명시하는 2 비트 워드이다. 타일 구성을 설명하기 위한 SEI 메시지들의 표준 이용을 참조하여 그리고 예시를 위해, *tile_section_exact_match_flag*로서 알려진 플래그는 *independent* 플래그의 값을 설정하는데 이용될 수 있다. 후자의 의미는 다음과 같이 설정될 수 있다:

[0462] - *independent* 파라미터가 0인 경우에는, 이 타일과 동일한 프레임에서의 또는 이전의 프레임들에서의 다른 타일들 사이의 코딩 종속성들은 알려지지 않고; 잠재적인 디코딩 종속성 신호를 찾아 트랙 참조 박스를 조사하는 것에 의해 정보가 발견될 수 있다.

[0463] - *independent* 파라미터가 1인 경우에는, 이 타일과 동일한 프레임에서의 다른 타일들 사이에는 공간적인 코딩 종속성들이 존재하지 않지만, 이 타일과 이전의 프레임들에서 동일한 *tileID*를 갖는 타일 사이에는 코딩 종속성들이 존재할 수 있고;

[0464] - *independent* 파라미터가 2인 경우에는, 이 타일과 동일한 프레임에서 또는 이전의 프레임들에서 동일한 *tileID*를 갖는 다른 타일들 사이에는 코딩 종속성들이 존재하지 않고;

[0465] *independent* 파라미터 값 3은 다른 타일들로부터의 간단한 *slice_segment_header* 종속성을 나타내는 데 사용될 수 있다. 이것은 하나의 슬라이스가 다수의 타일을 포함하는 경우일 수 있다: 하나의 타일은 따라서 이전의 독립 슬라이스 세그먼트가 인코딩된 타일에 종속하는 종속 슬라이스 세그먼트들만을 이용하여 인코딩될 수 있다. 그러한 경우에, 현재의 것이 종속하는 타일에 관한 정보는 트랙 참조 박스에서 발견될 수 있다.

[0466] 옵션으로, 타일당 평균 비트레이트를 설명하는 파라미터가, 대역폭에 기초한 적응을 위해 스트리밍 클라이언트에 제공되기 위해서 타일 설명자에 설정될 수 있다.

- [0467] 이 실시예에 따르면, 각각의 타일의 특성들은, 각각의 타일 트랙에 대해, '타일' *grouping_type* 및 *HEVCSpatialEntry*(또는 *TileRegionGroupEntry*)를 갖는 하나의 *SampleGroupDescription* 박스('sgpd')를 정의함으로써 영화 헤더('moov' 박스)에서 한번 주어진다. 그러면, ISO BMFF 표준에 따르면, 샘플들의 개수가 미리 알려지지 않으므로, 타일 트랙 프래그먼트의 각각의 샘플과 그것의 특성들을 관련시키기 위해서 각각의 타일 트랙 프래그먼트에서 *SampleToGroup* 박스가 정의된다.
- [0468] 타일들의 그리드가 시간에 따라 변경되는 경우에, 새로운 *HEVCSpatialEntry* (또는 *TileRegionGroupEntry*)를 갖는 새로운 *SampleGroupDescription* 박스('sgpd')가 트랙 프래그먼트 박스('traf')에서 정의되며, *SampleToGroup* 박스('sbgp')에 의해 참조될 수 있다. 그러므로, 그리드가 시간에 따라 정적인 경우에, 적어도 하나의 *SampleToGroup* 박스가 타일 트랙 및 타일 트랙 프래그먼트마다 정의된다. 이 박스는, 설명의 관점에서, 적어도 28 바이트를 나타낸다. 2초의 지속기간의 프래그먼트들을 갖는 16개의 타일을 가정하면, 이것은, *SampleToGroup* 박스에 대해서만, 시간에 따른 타일링 구성을 시그널링하기 위해서 초당 1792 비트를 나타낼 것이다. 그리드가 시간에 따라 변경되는 경우에, 비용은 (데이터 양에 관하여) 더 높을 것이다. 후술되는 바와 같이, 이러한 보충 초기화 데이터의 양은 감소될 수 있다.
- [0469] 샘플 그룹 박스들은 개별 샘플들에 메타데이터를 추가하기 위해 효율적이며 확장가능한 도구를 제공하다는 점에 유의해야 한다. 그러나, 주어진 메타데이터가 주어진 트랙의 모든 샘플들에 대해 유효하다는 것은 아주 일반적 인데, 그 이유는 이것이 각각의 AU에서 고정된 NALU 패턴을 갖는(즉, 시간을 따라 정적인 타일링 구성을 갖는) 타일링 설명자를 이용한 경우일 수 있기 때문이다.
- [0470] 이것은 *SampleToGroup* 박스를 이용하여 그룹의 각각의 샘플을 그에 따라 플래깅함으로써 표시될 수 있다.
- [0471] 대안적으로, 타일 트랙당 초기화 데이터의 양을 감소시키기 위해서, 일부 샘플 그룹들은 '디폴트'(즉, 모든 샘플들에 대해 유효함)로서 마킹될 수 있다. 이러한 디폴트 그룹화는 바이트에 관하여 설명 비용을 제한하는데, 그 이유는, 정적 타일링 구성을 대해, 그것이 전체 시퀀스에 대해 트랙당 한번 'moov' 박스 레벨에서 정의되기 때문이다.
- [0472] 그 목적을 위해, 그룹 설명 타입의 제2 버전(버전=2)이 *SampleGroupDescriptionBox*로서 알려진 박스에 이용될 수 있는데(traf/stbl 박스당 다수의 *SampleGroupDescriptionBox*가 존재할 수 있음), 이는 (*grouping_type*로서 알려진 파라미터를 통해) 현재의 트랙에서의 또는 현재의 트랙 프래그먼트들에서의 모든 샘플들에 대해 참조된 샘플 그룹이 적용된다는 것을 나타낸다.

[0473] 샘플 그룹 설명 박스의 새로운 버전은 다음과 같이 정의될 수 있다:

```
aligned(8) class SampleGroupDescriptionBox (unsigned int (32)
handler_type) extends FullBox('sgpd', version, 0) {
    unsigned int(32) grouping_type;
    if (version ==1) || (version==2) {unsigned int (32) default_length;}
    unsigned int (32) entry_count;
    int i;
    for ( i = 1; i <= entry_count; i++ ) {
        if (version != 0) {
            if (default_length==0) {
                unsigned int(32) description_length;
            }
        }
    }
    switch(handler_type) {
        case 'vide': // for video tracks
        VisualSampleGroupEntry(grouping_type);
        break;
        case 'soun': // for audio tracks
        AudioSampleGroupEntry(grouping_type);
        break;
        case 'hint': // for hint tracks
        HintSampleGroupEntry(grouping_type);
        break;
    }
}
```

[0474]

[0475] 이러한 확장된 타입의 *SampleGroupDescription* 박스에 따르면, *version*이 2일 때, 이 그룹 설명 및 임의의 *SampleGroupEntry*는 현재의 트랙 또는 트랙 프래그먼트의 모든 샘플들에 적용된다. 샘플 그룹 설명 박스가 특정하다는 것을 나타내기 위한 다른 실시예에는 ISO 베이스 미디어 파일 포맷의 일부 박스들에서 이미 사용되는 *flags* 속성의 지원 하에 이 박스를 확장하는 것이다. 이 실시예에서, *SampleGroupDescriptionBox*에 대한 새로운 정의는 다음과 같다:

[0476] aligned(8) class SampleGroupDescriptionBox (unsigned int (32) handler_type) extends FullBox('sgpd', version, flags) {

[0477] unsigned int(32) grouping_type;

[0478] if (version==1) { unsigned int(32) default_length; }

[0479] unsigned int(32) entry_count;

[0480] int i;

[0481] for (i = 1 ; i <= entry_count ; i++) {

[0482] if (version==1) {

[0483] if (default_length==0) {

[0484] unsigned int(32) description_length;

[0485] }

[0486] }

[0487] switch(handler_type) {

```

[0488]     case 'vide': // for video tracks
[0489]         VisualSampleGroupEntry(grouping_type);
[0490]         break;
[0491]     case 'soun': // for audio tracks
[0492]         AudioSampleGroupEntry(grouping_type);
[0493]         break;
[0494]     case 'hint': // for hint tracks
[0495]         HintSampleGroupEntry(grouping_type);
[0496]         break;
[0497]     }
[0498] }
[0499]

```

[0500] 이 시맨틱들은 바뀌지 않고, 유일한 차이는 다음과 같이 정의된 *flags* 속성의 도입이고, 다음의 플래그들은 *sgpd*에 대해 정의되고: *0x000001 sample-group-is-default*는: 이 트랙 내의 또는 현재의 프래그먼트 내의 모든 샘플들이 이 그룹에 할당된다는 것을 나타내고; 따라서, 이 박스에서 정의된 모든 엔트리들(특정한 샘플 그룹 엔트리들)이 트랙 내의 또는 트랙 프래그먼트(*traf*) 내의 모든 샘플들에 적용된다. 디폴트 *sgpd* 박스가 동일한 *grouping_type*을 가지고 디폴트 *sgpd* 박스를 가진(즉, 상기 *flags* 값을 가진 또는 모든 샘플들에 적용되는 것으로 *version=2* 또는 그것을 시그널링하는 임의의 다른 수단을 가진) 트랙에 대한 트랙 프래그먼트(*traf*)에서 정의된다면, *traf*에서 정의된 *sgpd* 박스만이 고려될 것이다. 영화 프래그먼트들의 크기를 변경하거나, 그것들을 제거할 때, 디폴트 관련이 더 이상 적용되지 않는 일이 발생할 수 있고, 그 경우 새로운 관련을 반영하기 위해 *SampleToGroupBox* 박스들과 함께 정규의 *SampleGroupDescriptionBox*가 생성되어야 한다. 이 디폴트 그룹화 메커니즘은 개별 트랙들에서 캡슐화될 때 GDR(Gradual Decoding Refresh) 그룹들 또는 타일들의 경우에 "롤(roll)"과 같은 샘플 설명들, 또는 다른 샘플 그룹 설명 태입들을 위해 유용할 수 있다(감소된 설명 비용).

[0501] *grouping_type* 파라미터는, 이 샘플 그룹 설명과 관련되는 *SampleToGroup* 박스를 식별하는 정수이다. 트랙 또는 트랙 프래그먼트가 그것의 초기화 데이터에서 *version=2*를 갖는 *SampleGroupDescription* 박스를 포함할 때, 이를 초기화 데이터에서 대응하는 *grouping_type*을 설명하는 *SampleToGroup* 박스를 둘 필요가 없다(그 이유는 설명이 정의에 의해 모든 샘플들에 적용되기 때문임)는 점에 유의해야 한다.

[0502] 새로운 버전의 샘플 그룹 설명 박스의 이러한 정의에서, *entry_count*는 다음의 테이블에서의 엔트리들의 개수를 제공하는 정수이고, *default_length*는, (길이가 일정한 경우에) 모든 그룹 엔트리의 길이를 나타내거나 또는 그것이 변수인 경우에 0을 나타내고, *description_length*는, 엔트리마다 변하는 경우에, 개별 그룹 엔트리의 길이를 나타내고 그에 따라 *default_length*는 0이다.

[0503] 이 실시예에 따르면, *grouping_type*은 공간/타일 그룹화를 나타내는 특정 값을 갖는데: 그것은 예를 들어 '타일'에 대한 ASCII 코드에 대응하는 16진 값을 수 있다(0x74696C65). *grouping_type*에 대해 동일한 값을 갖는 이 박스의 최대 하나의 출현이 트랙에 대해 존재할 것이다.

[0504] 시간에 따라 움직이는 적응형 그리드의 경우에, 샘플 투 그룹 박스(*sample to group box*)는 동일하게 유지되며 (즉, '타일' *grouping_type*), 모든 샘플들에 대해 계속 적용되는 점에 유의해야 한다. 이와 같이, *moov/trak/mdia/minf/stbl* 박스에서 시그널링되는 디폴트 구성으로부터 변경되는 타일링 구성을 갖는 타일 트랙들의 트랙 프래그먼트들에서 샘플 그룹 설명 박스만이 업데이트될 필요가 있다. 이것은 적응형 타일들에 대한 신호화 비용을 감소시킨다.

[0505] 대안적으로, 그리고 여전히 타일 트랙당 초기화 데이터의 양을 감소시키기 위해서(각각의 타일 트랙 프래그먼트에서의 *SampleToGroup* 박스의 반복을 회피하기 위해서), 'dsgp'로 지칭되는 새로운 *DefaultSampleToGroups* 박스(또는 그것의 명칭이 무엇이든지 동일한 시맨틱을 갖는 다른 유사한 박스)가 초기화 정보의 일부로서 각각의 *moov/trak* 박스들로부터 *SampleTable* 박스('stbl')에만 포함되는 것으로 정의된다. 이러한 새로운 박스는 트랙

내의 모든 샘플들에 적용될 샘플 그룹 설명들의 세트를 모든 샘플들에 관련시킬 것이다.

[0506] 새로운 DefaultSampleToGroup 박스는 다음과 같이 정의될 수 있는데:

```
aligned(8) class DefaultSampleToGroups extends FullBox('dsgp', version,
0) {
    unsigned int(32) entry_count;
    for (i=1; i <= entry_count; i++) {
        unsigned int(32) grouping_type;
        if (version == 1) {
            unsigned int(32) grouping_type_parameter;
        }
        unsigned int(32) group_description_index;
    }
}
```

[0507]

[0508] 여기서, *entry_count* 파라미터는 각각의 샘플에 관련될 그룹들의 리스트에서의 엔트리들의 개수를 제공하고, *grouping_type* 파라미터는 *SampleGroupDescription* 박스에서 언급되는 그룹화의 타입에 대한 식별자이다. 예를 들어, 특정 실시예에서, 그룹화 타입은 공간/타일 그룹화를 나타내는 특정 값을 가질 수 있다. 그것은 예를 들어 '타일'에 대한 ASCII 코드에 대응하는 16진 값일 수 있다(0x74696C65). *group_description_index* 파라미터는 이 그룹에서의 샘플들을 설명하는 샘플 그룹 엔트리의 인덱스를 제공하는 정수이다. 인덱스는 그 범위가 하나에서 *SampleGroupDescription* 박스에서의 샘플 그룹 엔트리들의 개수까지이거나 또는 이 샘플이 이러한 타입의 임의의 그룹의 멤버가 아닌 것을 나타내기 위해서 값 0을 취한다. 마지막으로, *grouping_type_parameter* 파라미터는 (그룹화 타입에 의해 이용되는 경우) 그룹화의 서브타입에 대한 표시이다.

[0509] 이것은, 트랙으로부터의 모든 샘플들이, 타일 그룹화만이 이용 중인 경우에(*entry_count*=1) 영화 프래그먼트들의 개수가 무엇이든지 타일당 최대 32 바이트를 이용하여, 주어진 그룹화 타입에 대해 동일한 그룹 설명을 따른다는 것을 시그널링하는 것을 가능하게 한다. 시간에 따라 움직이는 적응형 그리드의 경우에, 새로운 *DefaultSampleToGroups* 박스 및 새로운 *SampleGroupDescription* 박스가 트랙 프래그먼트들에서 정의될 수 있다. 새로운 *DefaultSampleToGroups* 박스는 이전의 정의를 대체할 것이며, 새로운 *SampleGroupDescription* 박스에서의 새로운 타일 설명을 참조한다. 따라서, *SampleToGroup* 박스는 각각의 트랙 프래그먼트에 대해 정의되지 않고 타일 그리드 정의가 변경될 때에만 정의된다.

[0510] 여전히 비디오 시퀀스를 따라 타일링 구성에서의 변동을 핸들링하도록 되어 있는 특정 실시예에 따르면, 새로운 샘플 맵 박스를 갖는 서브샘플 레벨에서 시그널링이 행해진다.

[0511] 여전히 특정 실시예에 따르면, 코딩 종속성들을 위해 일반 시그널링이 이용된다. 이 실시예에 따르면, 각각의 타일은 하나의 트랙에서 인코딩되고 (설명 크기에 관하여 많은 비용이 드는(전형적으로 타일마다 샘플 당 14 바이트) 추출자들 외에 참조들이 사용된다.

[0512] 도 18은 코딩 종속성들을 위해 일반 시그널링을 이용하여 HEVC 비트 스트림을 캡슐화하는 예를 도시한다. 이것은 도 17에 도시된 것으로서 단일 계층 스트림에 대한 캡슐화의 예를 제공한다. 이것은 타일형 HEVC 비디오 기본 스트림에 대응한다. 명확성을 위해, 상이한 트랙들의 데이터 부분만이 도 18에 나타내어진다(헤더 부분은 도 19a에 도시된다). 이 데이터 부분들은 다음을 나타낸다:

- 데이터를 집합하는 데 이용되는 합성 또는 가상 트랙(1810);

- 모든 트랙들에 공통인 정보를 저장하는 파라미터 세트 트랙들(1820); 및

- 각각이 주어진 타일에 대한 압축된 비디오 데이터를 포함하는 타일 트랙들(1830 내지 1890)(예시를 위해, 타일 트랙들(1830 및 1890)만이 나타내어진다).

[0516] 도 14와 비교할 때 알 수 있는 바와 같이, 도 18에 예시된 타일 트랙들 내의 샘플들은 추출자 NAL 유닛들을 더 이상 사용하지 않는다. 이에 따라 타일 트랙들의 설명의 크기가 작아진다. 대시 화살들은 디코딩 종속성들을 예시하는 데 반해 연속적인 화살들은 디코딩 순서 종속성들을 나타낸다.

- [0517] 예를 들어, 합성 트랙(1810)에서 1811로 참조되는 첫 번째 샘플은 파라미터 세트 트랙(1820)의 첫 번째 샘플(1821)에, 그리고 타일 트랙들 각각의 첫 번째 샘플들(1831 내지 1891)에 종속한다. 이것은 1819로 참조되는 샘플 *n*에 대해서 동일하다. 이러한 종속성들은 도 14에 예시된 것들과 유사하지만 후자는 추출자 데이터 구조로 표현되어, 샘플마다 종속성 당 14 바이트의 오버헤드를 야기한다.
- [0518] 보다 콤팩트한 설명을 제공하는 바람직한 실시예가 아래에 제안된다. 이 바람직한 실시예는 다음을 가능하게 한다:
- [0519] - 디코딩 순서를 명시한다;
- [0520] - 베이스 및 향상(또는 VCL + 타일)의 고정 패턴의 경우에 각 샘플에 대한 반복 대신에 연결(concatenation)을 위한 간단한 규칙을 제공한다; 그러한 간단한 규칙은 또한 오버헤드의 감소를 가능하게 한다(추출자 NALU들에 대한 필요가 없다);
- [0521] - 추출자들이 추출자들을 지시할 수 없다는 사실(ISO/IEC 14496-15의 하위 조항 A.3.1 참조), 그리고 따라서 완전한 스트림을 집합하기 위해 N번째 스케일 가능 계층이 샘플마다 N개의 추출자를 필요로 한다는 사실을 우회한다.
- [0522] 이 실시예에 따르면, 샘플에 대해 종속성 링크들의 반복이 이용됨에 따라 종속성 링크들의 하나의 설명이 모든 또는 적어도 여러 샘플들을 처리하는 데 이용될 수 있다. 다시 말해서, 종속성 링크들의 그러한 반복은 이들을 각 트랙의 헤더 부분에 인수 분해하고, 따라서 설명 비용을 감소시키는 것을 가능하게 한다.
- [0523] 도 19a 및 도 19b를 포함하는, 도 19에 예시된 바와 같이, 그것은 코딩 종속성들을 위해 일반 시그널링을 이용하여 캡슐화된 트랙들에 대한 트랙 헤더들 및 트랙 참조들의 예들을 예시하고, moov 헤더 박스는 종속성들의 그러한 설명을 저장하는 데 이용될 수 있다.
- [0524] 예시를 위해, 도 19a의 *moov* 헤더(1900)는 도 18에 예시된 트랙들의 헤더 데이터(*trak* 박스)에 대응한다. 보다 정확하게는, 도 18의 합성 또는 가상 트랙(1810)의 헤더 정보는 도 19a에서 참조 번호 1910으로 나타내어지고, 도 18의 파라미터 세트 트랙(1820)에 대한 헤더 데이터는 도 19a에서 참조 번호 1920으로 나타내어지고, 도 18에서 타일 트랙(1830) 또는 타일 트랙(1890) 같은 타일 트랙에 대한 헤더 데이터는 도 19a에서 참조 번호 1940으로 나타내어진다.
- [0525] 예시된 바와 같이, 각각의 트랙 박스는 참조 타입 및 이 특정 참조 타입에 의해 관련되는 트랙 ID들의 리스트로서 주어진 트랙에 대한 종속성들을 리스트하는 *tref*로 표시된 트랙 참조 박스를 포함한다. 예시를 위해, *dpnd* 및 *dodp*로 표시된 2개의 참조 타입이 사용된다. 첫 번째 참조 타입은 디코딩 종속성을 나타내는 데 반해 두 번째 참조 타입은 디코딩 순서 종속성을 나타낸다. 이들은 ISO/IEC 14496 파트 12 또는 파트 15에서 정의된 다른 기준 참조 타입들과 결합될 수 있다. *dpnd* 및 *dodp* 명칭들은 예로서 주어진 것이고, 그것들이 ISO 베이스 미디어 파일 포맷 및 그의 확장들에서의 다른 코드들과 상충하지 않는다면 임의의 다른 4-문자 코드가 사용될 수 있다는 것에 유의해야 한다.
- [0526] 예시된 예에서, 합성 트랙은 파라미터 세트 트랙에 대한 그리고 각각의 타일 트랙에 대한 디코딩 종속성들을 가진다. 따라서, 그것의 *tref* 박스(1911로 참조됨)는 합성 트랙이 종속하는 트랙들로서 모든 다른 트랙들을 리스트한다. 그와 반대로, 파라미터 세트 트랙은 디코딩 종속성을 갖지 않는다(참조 번호 1921로 알 수 있는 바와 같이). 그에 따라, 그것은 도 13을 참조하여 아래에 설명된 바와 같이 '*hbas*' 종속성을 이용하여 베이스 트랙으로서 다른 트랙들로부터 참조될 수도 있다. 이것은 도 20을 참조하여 설명된 바와 같이 클라이언트 측에서 수행되는 파싱 프로세스를 단순화할 수 있다.
- [0527] 이 참조 메커니즘은 각각의 트랙에 대해 디스플레이 가능 트랙을 재구성하기 위해 요구되는 모든 다른 트랙들을 나타내고 추가로 트랙들에 포함된 NAL 유닛들 사이의 올바른 디코딩 순서를 보장한다.
- [0528] 도 20은 코딩 종속성들을 위해 일반 시그널링을 이용하여 캡슐화된 mp4 파일 또는 세그먼트들을 해석하기 위해 클라이언트 디바이스에 의해 수행되는 프로세스를 설명하는 흐름도이다.
- [0529] 예시된 바와 같이, 제1 단계(단계 2000)는 로컬 파일이거나 서버로부터 다운로드되거나 세그먼트들로서 스트리밍된 파일일 수 있는 mp4 파일의 수신에 있다. 다음 단계(단계 2001)에서는 사용자 데이터 박스가 존재하고 트랙 선택 박스(*tsel*)를 포함하는지를 결정하기 위해 *moov* 박스 및 *trak* 박스를 찾는다. 이것은 클라이언트 디바이스 또는 사용자가 디스플레이 할 트랙을 선택하는 데 도움을 줄 수 있다.

- [0530] 트랙 선택 박스가 존재하지 않는다면(즉, 단계 2001의 결과가 거짓), 처음 마주치는 *trak* 박스에 대응하는 트랙이 디폴트로서 선택된다(단계 2002). 그와 반대로, 트랙 선택 박스가 존재한다면(즉, 단계 2001 결과가 참), 클라이언트 디바이스는 트랙들의 속성을 사용자에게 제시하여 그가 선택하는 데에 도움을 줄 수 있다(단계 2003). 대안으로, 클라이언트 디바이스는 트랙 헤더 박스 내의 트랙과 관련된 미디어 핸들러 정보를 파싱할 수 있다(예를 들어, 도 11에서 참조). 이 실시예에 따르면, 비디오 트랙(예를 들어, 도 18에서 합성 트랙(1810)) 만이 선택될 수 있고 타일 트랙(예를 들어, 도 18에서 타일 트랙들(1830 내지 1890) 또는 파라미터 세트 트랙(예를 들어, 도 18에서 파라미터 세트 트랙(1820))은 없다.
- [0531] 다음 단계(단계 2004)에서, 클라이언트 디바이스는 디코딩을 위해 그것이 종속하고 있는 트랙들의 리스트를 구축하기 위하여 선택된 트랙에 대한 트랙 참조 박스의 파싱을 시작한다. 리스트 내의 각각의 트랙에 대해, 그것은 디코딩 종속성을 반복적으로 처리한다(단계 2005).
- [0532] 이 리스트를 구축하는 동안, 클라이언트는 모든 종속 트랙들이 다른 트랙들에의 종속성을 수반하는 '*dpnd*' 참조 타입을 포함하는지를 체크한다(단계 2006). 그러한 경우에, 프로세스는 그러한 경우가 발생하지 않아야 하기 때문에 종료된다(단계 1007). 그와 반대로, 종속 트랙들 중 적어도 하나가 다른 트랙들에의 어떤 '*dpnd*' 참조 타입도 포함하지 않는다면, 그것은 단계 2009에서 샘플들의 생성을 가이드하는 순서화된 트랙들의 리스트를 구축(단계 2008)하기 위한 베이스 트랙으로서 선택된다. 순서는 이 트랙에서 선언된 '*dodp*' 종속성을(도 19에서 참조 번호 1922)을 따르는 것에 의해 주어진다.
- [0533] 이 단계는 단계 2008에서 얻어진 순서화된 트랙들에 대해 그리고 동일한 타임스탬프를 가진 NAL 유닛들에 대해 반복적이다. 주어진 트랙에 대해, 동일한 타임스탬프를 가진 모든 NAL 유닛들은 연결된다. 그렇게 연결된 NAL 유닛들 다음에는 다시 동일한 타임스탬프를 가진 순서화된 리스트 내의 다음 트랙으로부터의 NAL 유닛들이 오고 트랙들의 순서화된 리스트의 끝까지 그렇게 계속된다. 그 후, 입력 세그먼트 또는 파일의 끝까지 동일한 원리에 따라 (다음 타임스탬프에 대응하여) 다음 샘플이 구축된다.
- [0534] 이것은 단계 2010에서 비디오 기본 스트림의 생성으로 이어진다.
- [0535] 클라이언트 디바이스에 의해 수행되는 프로세스, 특히 베이스 트랙을 얻는 단계(즉, 단계 2005 및 2006)는 어느 트랙이 베이스 트랙인지를 나타내는 종속성 정보를 추가하는 것에 의해 단순화될 수 있다는 점에 유의해야 한다. 이것은, 예를 들어, ISO/IEC 14496-95에서와 같은 트랙 참조 타입 '*tbas*' 또는 *sbas* 또는 도 13을 참조하여 설명된 박스 '*hbabs*'를 이용하여 행해질 수 있다.
- [0536] 특히 스트리밍 구성 - 이에 따르면 반드시 모든 데이터가 클라이언트 디바이스에 일반적으로 존재하지 않음 -에서는, 겸중 단계가 단계 2004(이 동안에 클라이언트는 디코딩 종속성을 따른다)와 결합될 수 있다는 점에도 유의해야 한다. 사실, 디코딩 종속성이 클라이언트 디바이스에 의해 관찰될 때, 참조된 트랙은 클라이언트 디바이스에 존재해야 하는데, 그렇지 않다면, 디코딩은 가능하지 않을 것이다. 그것들이 아직 수신되지 않았다면, 그것들은 클라이언트가 샘플들을 생성하는 것(단계 2009)을 가능하게 하기 위해 그 시간에 요청될 수 있다. 반대로, 단계 2008에서 구축된 트랙들의 순서화된 리스트에서, 하나의 트랙이 클라이언트 측에서 이용 가능하지 않다면, 그것은 오류 없이 무시될 수 있다. 이것은 전체 비디오의 공간 서브파트만을 재생하는 것을 가능하게 한다.
- [0537] 명확성을 위해, 도 19b에 예시된 예는 단일 계층 타일형 스트림만이 독립 타일들로 캡슐화되는 단순한 경우이다. 비독립 타일들의 경우에, 일부 디코딩 종속성을 트랙 참조 박스에서 시그널링되고 디코딩 순서 종속성도 트랙에서 시그널링되어 이 트랙으로부터의 NALU가 트랙 내의 것들보다 앞에 발생할 것임을 나타낸다. 예를 들어, 도 19b에 예시된 바와 같이, 타일 i 가 ID가 4인 트랙에 대응하는 타일로부터의 코딩 종속성을 가질 경우(참조 번호 1980), 디코딩 종속성(1981)이 트랙 참조 박스에서 시그널링되고 디코딩 순서 종속성도 ID=4인 트랙(참조 번호 1970)에서 시그널링되어(1971) 이 트랙으로부터의 NALU가 ID= $tile_i$ 인 트랙 내의 것들보다 앞에 발생할 것임을 나타낸다. 이것은 항상 계층에서의 타일들이 하위 계층에서의 동일한 타일에 종속할 때 동일할 것이다.
- [0538] 도 20에 관련하여 설명된 프로세스는 도 21a로부터의 디코딩 종속성 그래프를 고려할 때 도 21b에 예시된다. '*dodp*' 종속성은 마침내 상향으로 진행하는 트랙 참조들을 이용하여 종속성을 설명한다. 예시된 바와 같이, 종속성 그래프는 트랙 T1(2201)으로부터의 종속성들: 2212, 2213, 및 2214 상에 주어진 순서를 포함한다. 이 종속성들은 시간 t 에서의 샘플들 S_i 의 예상 처리 순서가 $S1(t)$, $S2(t)$, $S3(t)$, $S4(t)$, $S5(t)$ 일 것임을 나타낸다. 시간 MT2에서 트랙 T1 상에 샘플이 없고 트랙 T4 상에 샘플이 없다면, 샘플들의 처리 순서는 $S2(MT2)$,

S3(MT2), S5(MT2)일 것이고, 따라서 주어진 시간에 이용 가능한 샘플이 없을 때 디코딩 순서 종속성을 무시한다.

[0539] 마지막으로, 이 새로운 디코딩 순서 또는 상향 종속성('dodp' 또는 4-문자 코드 명칭이 무엇이든)은 다음과 같이 요약될 수 있다: 'dodp' 참조는 참조하는 트랙과 참조 트랙들 사이의 종속성 서브-트리를 정의한다. 참조되는 트랙들도, 다른 서브-트래들을 정의하는, 'dodp'를 가질 수 있다. 'dodp'를 가지고 어떤 'dodp'에 의해서도 참조되지 않는 트랙은 종속성 트리의 루트(root)이다. 임의의 미디어 시간 MT에 대해, 종속성 트리 내의 트랙들 중 하나에 샘플이 있다면, 미디어 시간 MT를 가진 Tracktref[i]로부터의 샘플은 미디어 시간 MT를 가진 Tracktref[i+1]로부터의 샘플(있다면) 앞에 그러나 참조하는 트랙의 미디어 시간 MT를 가진 샘플(있다면) 뒤에 미디어 프로세서에 전달될 것이다. 루트에서 시작하여, 종속성 트리의 하나의 레벨 내의 모든 트랙들이 먼저 다루어지고, 그 후 더 깊은 레벨에 있는 트랙들이 다루어진다. 동일한 트랙을 참조하는 다수의 종속성들이 발견된다면, 참조의 첫 번째 발생에 대응하는 샘플들이 미디어 프로세서에 전달될 것이다.

[0540] 동일한 미디어 핸들러를 갖지 않는 트랙들 사이에 'dodp' 참조들을 갖는 것은 오류이지만, 동일한 샘플 설명 타입들을 갖지 않는 트랙들 사이에 'dodp' 참조들을 갖는 것은 허용된다.

[0541] 'dodp' 트랙 참조를 운반하는 데 사용되는 샘플이 없는 트랙을 갖는 것은 허용되고; 이것은 예를 들어 'dodp'를 운반하는 공백 트랙의 샘플 설명에 모든 파라미터 세트들이 저장된 ISOBMF 파일 내의 분리된 트랙들에 저장된 HEVC 파일들의 코딩 계층 구조를 설명하는 것을 허용한다.

[0542] 'dodp'에 의해 참조되는 트랙들은 개별적으로 처리 가능하지 않을 수 있고, 그 경우, 역방향 호환성의 이유로, 이 트랙들은 사용 불가능한(disabled) 것으로 마킹될 것이다. 'dodp' 인지 미디어 프로세서는 사용 불가능한 트랙들의 일부 또는 전부를 재생하기로 결정할 수 있다.

[0543] 보다 일반적으로, 디코딩 순서 종속성은 코딩 종속 미디어의 유연한 설명을 가능하게 하거나 간단한 샘플 데이터 분할을 제공한다. 사실, 추출자들과 반대로, 이 메커니즘은 종속성 설명을 실제 NAL 유닛 집성 프로세스와 분리시킨다. 그 후 그것은 외부 수단에 의해 전달된 또는 저장된 베이스 미디어 스트림(들)에 대해 ISOBMF 파일들에 저장된 향상 계층들의 코딩 종속성들을 설명하는 것을 허용한다.

[0544] 추출자들의 사용에 기초한 시그널링보다 더 많은 유연성을 제공하는 이 종속성 시그널링의 확장은 디코딩 종속성이 시간을 따라 변화하는 경우와 관련된다. 예를 들어, 여러 파일 트랙들로 이루어진 가상 트랙이 주어진 시간에 관심 영역을 나타낼 수 있다. 이 관심 영역은 시간을 따라 변화하여 가상 트랙과 (ROI에 의해 커버되는) 트랙 파일들 간의 종속성들의 변경들로 이어질 수 있다. 그러한 경우를 다루는 위치에 있기 위해, 새로운 트랙 참조 박스가 트랙 프래그먼트 레벨에서 도입된다: trak 박스 내의 글로벌 트랙 참조 박스에서 선언된 종속성들을 일시적으로 오버로드하는 것을 가능하게 하는 "트랙 프래그먼트 참조 박스(track fragment reference box)".

[0545] 타일링 구성에 관한 샘플 설명에 대한 이 실시예는 단일 트랙 캡슐화된 타일형 HEVC 비트 스트림(single track encapsulated tiled HEVC bit-stream)에 적용된다. 그것은, MPEG-4 표준에 따르는 단일 비디오 트랙으로 다운로드된 공간적 부분들을 저장하기 위해서 합성 트랙의 추출자들의 분해(resolution) 뒤에, 클라이언트 측에서의 또는 타일형 HEVC 비트 스트림의 MPEG-4 캡슐화에 대한 애플리케이션을 찾을 수 있다. 합성 트랙을 MPEG-4에 따르는 표준 단일 트랙으로 변환하는 것에 관한 경우의 다른 애플리케이션은 추출자들을 지원하지 않는 MPEG-4 파서들에 대해 어드레싱되었다.

[0546] 도 12a 및 도 12b를 포함하는 도 12는 다양한 타일링 구성을 핸들링하도록 되어 있는 서브샘플 레벨에서 어떠한 아티팩트 없이 타일 트랙이 디코딩될 수 있다는 표시, 타일의 크기, 및 전체 비디오에서의 타일의 위치의 신호화를 도시한다.

[0547] 도 12a는 클라이언트 디바이스(예를 들어, 비디오 플레이어)에 의해 수행되는 단계들을 도시한다. 제1 단계(단계 1200)에서, 클라이언트 디바이스는 초기화 데이터를 다운로드하거나 또는 파일이 로컬 파일인 경우에는 초기화 데이터, 예를 들어 MPEG-4 표준에 따르는 캡슐화된 비트 스트림의 초기화 데이터, 통상적으로 moov 박스의 콘텐츠를 판독한다.

[0548] 이를 초기화 데이터로부터, 클라이언트 디바이스는 타일 정보가 코딩되는 트랙 헤더 정보를 파싱할 수 있다(단계 1205). 이 타일 정보를 이용하여, 사용자는 하나 이상의 타일에 대응할 수 있는 관심 영역을 클라이언트 디바이스의 그래픽 인터페이스를 통해 선택할 수 있다(단계 1210).

[0549] 대응하는 파일 트랙들뿐만 아니라 합성 트랙이 클라이언트 디바이스에 의해 다운로드되거나 판독된다(단계 1215 및 단계 1220). 다음에, 단일 비디오 트랙을 획득하기 위해서 파일 트랙들을 이용하여 합성 트랙들의 추출자들이 분해된다(resolved)(단계 1225). 마지막으로, 클라이언트 디바이스는, 획득된 비디오 트랙에서, 예를 들어 SampleTableBox에서의 파일링 설명을 구축하거나 추가한다(단계 1230).

[0550] 도 12b에는 파일링 설명의 예가 도시되어 있다. 도시된 바와 같이, 파일링 설명(1250)은 영화 박스 "moof"(1255) 및 데이터 박스 'mdat'(1260)를 포함한다. 'moof' 박스(1255)는, 상이한 샘플 그룹들을 설명하는 SampleToGroup 박스(1265), 파일들과 각각의 샘플의 NAL 유닛들 사이의 맵핑을 설명하는 샘플 그룹 설명 박스(1270), 및 파일 설명들을 포함하는 샘플 그룹 설명 박스(1275)를 포함하는, 트랙당 하나의 SampleTable 박스를 포함한다. 샘플 투 그룹 박스(1265)는 그룹 엔트리 TileSampleMapEntry에 대한 'tsgm' 그룹화 타입을 나타낸다.

[0551] TileNALUMapEntry(또는 더 일반적으로는 NALUMapEntry) 그룹 엔트리(1270)는 샘플의 NAL 유닛들과 파일들 사이의 맵핑을 정의한다(이것이 이러한 실시예가 서브샘플 레벨 시그널링을 참조하는 이유임). grouping_type 파라미터가 'tsgm'과 동등한 이러한 박스는 샘플당 NAL 유닛들의 개수를 포함한다.

[0552] TileNALUMapEntry(또는 NALUMapEntry) 박스는 (도 12b에 도시된 것처럼) 다음과 같이 정의될 수 있는데:

```
class TileNALUMapEntry() extends VisualSampleGroupEntry ('tsgm') {
    unsigned int(8) reserved = 0;
    unsigned int(8) entry_count;
    for (i=1; i<= entry_count; i++)
        unsigned int(32) tileID;
    }
}
```

[0553]

[0554] 여기서, entry_count는 트랙 샘플에서의 NALU들의 개수를 나타내며, tileID는 현재의 트랙에 의해 설명된 공간적 파일에 대한 고유 식별자를 제공한다.

[0555]

또한, TileNALUMapEntry(또는 NALUMapEntry) 박스는 다음과 같이 크기에 관하여 최적화될 수 있는데:

```
class TileNALUMapEntry() extends VisualSampleGroupEntry ('tsgm') {
    unsigned int(6) reserved = 0;
    unsigned int(1) large_size;
    unsigned int(1) mode;
    if (large_size) {
        unsigned int(16) entry_count;
    } else {
        unsigned int(8) entry_count;
    }
    for (i=1; i<= entry_count; i++)
        if (mode) {
            if (large_size) {
                unsigned int(16) NALU_start_number;
            } else {
                unsigned int(8) NALU_start_number;
            }
        }
    }
    unsigned int(16) tileID;
}
```

[0556]

[0557] 여기서, large_size 파라미터는 트랙 샘플들에서의 NAL 유닛 엔트리들의 개수가 8 비트로 표현되는지 또는 16

비트로 표현되는지를 나타내고, *mode* 파라미터는, (*mode*가 설정되지 않을 때에는) 각각의 NAL 유닛 엔트리가 설명되는지 또는 (*mode*가 설정될 때에는) 타일 ID에 대응하는 NAL 유닛 엔트리들만이 변경되는지를 나타낸다. 후자의 경우에, NAL 유닛 개수는 *large_size* 파라미터의 값에 종속하여 16 비트 또는 8 비트로 코딩된다.

[0558] 그것은 이들 NAL 유닛들 각각과 특정 타일 사이의 맵핑을 특정한다. 타일 설명은 본 명세서에서 전술한 바와 같이 샘플 그룹 설명 박스(1275)에 제공되는데, 각각의 타일이 차례로 설명된다.

[0559] 주어진 예는 하나의 NAL 유닛이 시퀀스 지속기간을 따라 이것 및 하나의 타일에 대한 데이터를 포함하는 특정 경우라는 점에 유의해야 한다. 타일 데이터가 수개의 NAL 유닛에 걸쳐 분열될 때, 타일에 대응하는 NAL 유닛들의 세트를 설명하기 위해 수개의 애그리게이터가 이용된다. 시간에 따라 NAL 유닛들의 개수가 변할 때, 수개의 *tileSampleMap* 엔트리가 정의될 수 있고, 대안적으로 그룹으로의 샘플들은 트랙 프래그먼트 헤더에서 프래그먼트마다 그룹화 탑입을 통해 적절한 타일 맵 엔트리를 참조한다.

[0560] 도 10을 참조하여 설명된 바와 같은 데이터 캡슐화 스키마에서, HEVC 비트 스트림은 실제로 압축된 비디오 데이터를 포함하는 타일 트랙들(1020-1 내지 1020-2)을 포인팅하는 합성 트랙(1015)으로서 캡슐화된다. 합성 트랙은 (도 10에서 PS로 표시된) 상이한 HEVC 파라미터 세트 NAL 유닛들로부터 유래하는 구성 데이터를 포함한다. 합성 트랙의 다른 요소들은 주로 샘플마다 그리고 타일 트랙마다 하나씩 추출자들의 리스트에 있는데, 이는 (초기화 세그먼트 파일(1005)의 moov 박스에 포함된 트랙 참조 박스('tref')를 통해) 타일 트랙들에 캡슐화되는 압축된 비디오 데이터를 포인팅한다.

[0561] ISO BMFF 표준(이 표준의 Part 15)에서의 현재의 종속성 수단은 초기화 세그먼트 파일(1005)의 moov 박스에서의 트랙 박스들의 일부인 트랙 참조 박스 'tref'에 위치한다. 'tref' 박스는 프레젠테이션에 있어서 포함 트랙으로부터 다른 트랙으로의 참조를 제공한다. 포함 트랙은 프레젠테이션에 있어서 다수의 다른 트랙을 참조할 수 있다. 트랙들 사이의 종속성의 탑입은 현재의 표준에서 2개의 값 'scal' 또는 'sbas'를 취할 수 있는 *reference_type* 파라미터에 의해 특정된다. 'sbas' 값은 스케일 가능한 베이스를 나타낸다. 그것은 참조된 트랙이 스케일 가능한 프레젠테이션에서 현재의 트랙의 스케일 가능한 베이스 트랙임을 나타낸다. 'scal' 값은 스케일 가능성(*scalability*)을 나타낸다. 그것은 스케일 가능한 표현의 상이한 계층들을 표현하는 트랙들 사이의 관계를 나타낸다. 그것은 포함 트랙이 참조된 트랙에 종속함을 의미한다.

[0562] 도 10을 참조하여 설명된 실시예에서, 특정 스케일 가능성 관련 종속성들은 존재하지 않는다. 스케일 가능한 비디오들이 고려될 수 있더라도, 여기서는 합성 트랙과 타일 트랙들 사이의 공간적 종속성들에 초점이 맞추어진다. 이들 종속성들은, 예를 들어, 합성 트랙(1015)(id=1)에 대응하는 초기화 세그먼트 파일(1005)의 moov 박스의 tref 박스에서 행해진 바와 같이 새로운 '타일' 값으로 명시적으로 표시될 수 있다.

[0563] 하나의 트랙과 다른 트랙들 간의 종속성들의 예시적인 예가 타일들에 관련되지만(즉, 하나의 합성 트랙과 타일 트랙들 간의 종속성들), 하나의 트랙과 하나의 또는 수개의 트랙 간의 다른 타입의 종속성들이 유사하게 핸들링 될 수 있다. 그러므로, *reference_type* 파라미터는, 예를 들어, 트랙이 하나 이상의 트랙에 종속함을 나타내기 위해서 서브 계층 종속성(예를 들어, 'subl' 값)을 나타낼 수 있다.

[0564] 따라서, 서브 계층 트랙들은, 다른 HEVC NAL 유닛들의 디코딩 프로세스에 해를 끼치지 않고 폐기될 수 있는, HEVC 기본 비트 스트림의 부분들을 포함하는 트랙들로서 정의될 수 있다. 이러한 정의는 특히 전술한 바와 같이 타일 트랙들뿐만 아니라 스케일 가능한 HEVC 비트 스트림들에서의 시간적 계층들에 적용된다. 서브 계층 트랙에 대응하는 각각의 트랙은, 미리 결정된 값으로 설정될 때에, 이 HEVC 트랙이 서브 계층 트랙이며 단지 다른 트랙(들)으로부터, 예를 들어 합성 트랙으로부터 참조되는 NAL 유닛들만을 포함함(즉, 이 HEVC 트랙은 디스플레이 가능하지 않음)을 나타내는 비트(또는 플래그)를 이용하여 HEVCConfiguration 래코드에(즉, SampleTableBox에) 마킹될 수 있다. 이 비트 또는 플래그의 값이 반대 값을 가질 때, 그것은, 이 HEVC 트랙이 초기화 데이터를 또한 포함하는 서브 계층 트랙임(즉, 이 HEVC 트랙은 디스플레이 가능함)을 나타낸다. 예를 들어, 현재의 HEVCDriverDecoderConfigurationRecord 박스에서 예약 비트들을 이용하는 것이 가능하다.

[0565] 도 12를 참조하여 설명된 특정 실시예에 따르면, 각각의 타일 트랙은 표준 비디오 트랙으로서 재생 가능하다.

[0566] 도 13은 본 발명의 일 실시예에 따른 표준 비디오 트랙들로서 재생 가능한 독립적인 타일 트랙들 및 합성 트랙을 포함하는 트랙들의 세트로서 HEVC 비트 스트림을 캡슐화하는 예를 도시한다. 예시를 위해, 캡슐화된 비트 스트림은 도 2에 개략적으로 도시된 비디오 시퀀스에 대응한다.

[0567] 도 13에 도시된 HEVC 비트 스트림 캡슐화는, 초기화 및 구성 데이터의 복원을 허용하는 특정 추출자를 각각의

타일 트랙이 포함한다는 점에서, 도 10에 도시된 것과는 주로 상이하다.

[0568] 도시된 바와 같이, 타일 트랙들(1300-1 내지 1300-12) 각각은, 초기화 및 구성 데이터를 나타내는 합성 트랙(1310)의 HEVC 파라미터 세트 NAL 유닛들(PS로 표시됨)에 대해 포인팅하는 추출자(1305-1 내지 1305-12)를 포함하는데, HEVC 표준에 따르면, 이들 초기화 및 구성 데이터는 통상적으로 HEVC 비트 스트림의 다양한 파라미터 세트들에 대응한다고 상기된다. 따라서, 이러한 초기화 및 구성 데이터는 각각의 타일 트랙이 정상 비디오 트랙으로서 재생 가능하게 한다.

[0569] 각각의 타일 트랙에 추가된 추출자는 유리하게는 비디오 데이터 샘플들 앞에서 각각의 타일 트랙의 미디어 데이터 박스 "mdat"의 시작에 위치한다.

[0570] 타일 트랙들(1300-1 내지 1300-12)로부터 합성 트랙(1310)으로의 이들 종속성들(1315-1 내지 1315-12로 표시됨)은 예를 들어 (초기화 세그먼트 파일(1325)의 영화 박스 'moov'에서) 타일 트랙들에 관련된 'tref' 박스들(1320-1 내지 1320-12)의 *reference_type* 파라미터에서 시그널링되어야 한다. 이 실시예에 따르면, 파라미터 세트를 포함하는 트랙은 HEVC 베이스 트랙 'hbabs'로 고려된다(이것은, 스케일 가능한 프레젠테이션에서 가장 낮은 동작 포인트를 포함하는 트랙이 '스케일 가능한 베이스 트랙'인 'sbabs'로 고려되는 SVC 경우에 가깝다). 도시된 바와 같이, 베이스 트랙에 종속하는 트랙들(즉, 식별자 id=1을 갖는 합성 트랙(1310)에 종속하는, 식별자들 id=2 내지 12를 갖는 타일 트랙들(1300-1 내지 1300-12))은 그들의 트랙 참조 박스(1320-1 내지 1320-12)에서 값 'hbabs'를 갖는다.

[0571] 역시, 타일링 신호화는 트랙 레벨에서, 샘플 레벨에서, 또는 트랙 및 샘플 레벨에서 이루어질 수 있다.

[0572] 디폴트로 타일 트랙들은 디스플레이 가능하지 않은 것으로 고려될 수 있다는 점에 유의해야 한다. 그러나, MPEG-4 표준에 따르는 진보된 파서는 디스플레이 가능한 타일 트랙들을 검출하고, 예를 들어 'tref' 박스를 검토함으로써 스트리밍 매니페스트 파일에서 이들을 노출시킬 수 있다(타일 트랙이 'hbabs' 타입의 참조 타입을 포함하는 경우, 그것은 디스플레이 가능한 것으로 고려될 수 있음). 이것은, 이 타일 트랙이 핸들러 박스에서 'title' 값으로 마킹되는 경우에도 표준 비디오 트랙으로 고려될 수 있는 것을 의미한다. 타일링 신호화가 샘플 기반일 때, 타일 트랙들 또는 서브 계층 트랙들은 그들의 핸들러 박스에서 'vide'로서 태깅될 수 있는데, 그 이유는 타일링 정보가 *SampleTableBox*로서 알려진 박스에 배치되기 때문이다.

[0573] 도 14는 본 발명의 다른 실시예에 따른 합성 트랙, 초기화 데이터 트랙, 및 표준 비디오 트랙들로서 재생 가능한 독립적인 타일 트랙들을 포함하는 트랙들의 세트로서 HEVC 비트 스트림을 캡슐화하는 예를 도시한다. 예시를 위해, 캡슐화된 비트 스트림은 도 2에 개략적으로 도시된 비디오 시퀀스에 대응한다.

[0574] 도 14에 도시된 HEVC 비트 스트림 캡슐화는, 초기화 데이터가 (합성 트랙(1310)에서가 아니라) 전용 초기화 데이터 트랙(1400)에 배치된다는 점에서, 도 13에 도시된 것과는 주로 상이하다.

[0575] 도 13을 참조하여 설명된 것과 비교하여 볼 때 이러한 실시예에 의해 제공되는 이점들 중 하나는 타일 트랙들이 독립적으로 재생되어야 할 때 송신될 데이터의 양에 관련된다. 초기화 데이터가 전용 트랙에서 송신되기 때문에, 합성 트랙을 송신하는 것이 요구되지는 않는다.

[0576] HEVC 파일 포맷의 현재의 사양에 따르면, 파일 포맷에서 파라미터 세트들(PS)을 전달하는 2가지 가능성성이 존재한다는 것을 상기해야 한다: 단지 *Sample Entry*로서 알려진 박스에서 또는 *Sample Entry* 박스에서 그리고 데이터 샘플들에서. 이를 2가지 구성은 *Sample Table*로서 알려진 박스에서의 'hvc1' 및 'hev1' 박스들을 이용하여 각각 시그널링된다. 샘플들에 파라미터들을 저장하는 것은 더 복잡하지만, 그것은 파라미터 세트 업데이트의 경우에 더 큰 역동성을 허용한다. 그러므로, 바람직한 실시예에서, 파라미터 세트들은 특히 타일링 구성 변경들에 대한 꽉쳐 파라미터 세트들(PPS) 변경들을 핸들링 가능하기 위해서 *Sample Entry* 박스에서 그리고 (*Sample Table* 박스에서의 *HEVC SampleEntries* 파라미터에서 'hev1' 값을 갖는) 데이터 샘플들에서 전달된다.

[0577] 따라서, 전용 초기화 데이터 트랙(1400)은, 비디오 파라미터 세트, 시퀀스 파라미터 세트 또는 꽉쳐 파라미터 세트에 각각 대응하는 32, 33 또는 34와 동등한 타입의 NAL 유닛들과 같은 비-VCL HEVC NAL 유닛들만을 데이터로서 포함한다.

[0578] 도 14에 도시된 바와 같이, 타일 트랙들(1410-1 내지 1410-12)의 미디어 데이터 박스 "mdat"의 시작에 위치하는 추출자들(1415-1 내지 1415-12)은 전용 초기화 데이터 트랙(1400)의 데이터를 포인팅한다. 마찬가지로, 합성 트랙(1405)의 제1 추출자(1420)는 전용 초기화 데이터 트랙(1400)의 데이터를 포인팅한다. 그러므로, 초기화 데이터 트랙(1400)은 어떠한 다른 트랙도 참조하지 않는 캡슐화된 HEVC 비트 스트림의 유일한 트랙이다. 이와

같이, 초기화 데이터 트랙(1400)(id=2)에 관련되는 tref 박스에 표시된 어떠한 종속성도 존재하지 않기 때문에(tref 박스에서의 'hbas' 종속성 없음), 전자는 독립적으로 디스플레이 가능하지 않은 것으로 고려된다.

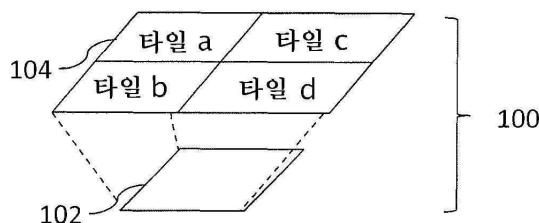
[0579] 비디오 비트 스트림에서 일부 초기화 데이터가 수정될 때(즉, HEVC 비트 스트림에서 픽처 파라미터 세트가 발생할 때), 이들은, 변경이 발생하는 시간적 위치에서, 참조 번호 1425로 도시된 바와 같이 샘플 데이터에 배치된다. 1430 및 1435-1 내지 1435-12로 참조된 대응하는 추출자들은 합성 트랙(1405)에 그리고 타일 트랙들(1410-1 내지 1410-12) 각각에, 즉 이들 새로운 PPS를 참조하는 각각의 타일 트랙에 각각 삽입된다.

[0580] 캡슐화된 HEVC 비트 스트림의 각각의 트랙에서, 샘플들(및 관련된 NALU들)은 시간적 순서로 구성된다. 마찬가지로, 픽처 파라미터 세트들은 전용 초기화 데이터 트랙(1400)에서 시간적 순서로 구성된다. 'trun' 박스(도 14에 표현되지 않음)는 각각의 샘플에 대한 올바른 디코딩 시간의 제공을 허용한다.

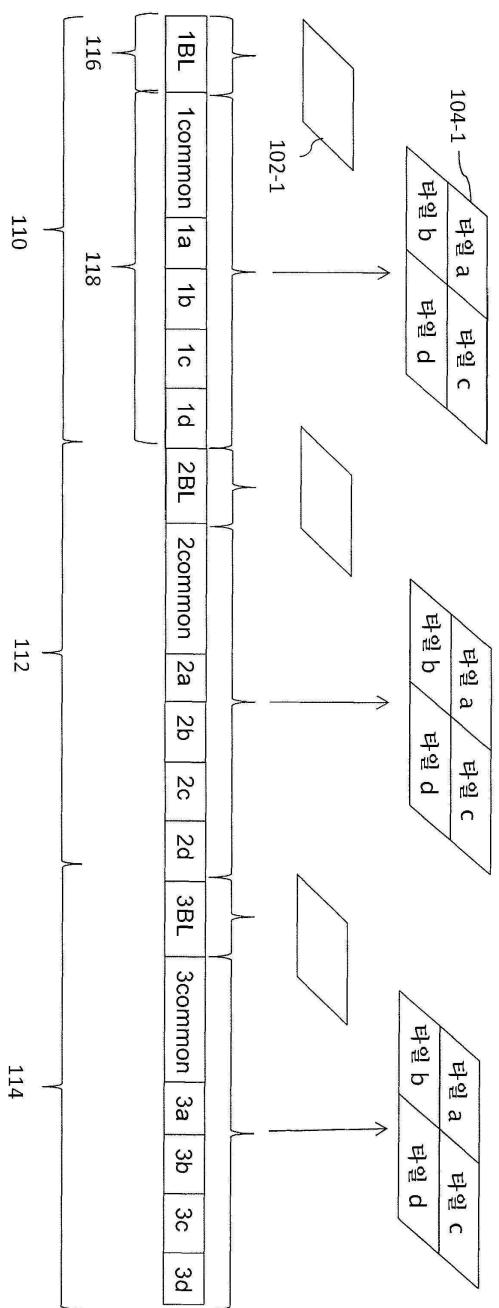
[0581] 물론, 국부적인 그리고 특정 요건들을 만족시키기 위해서, 관련 기술분야의 통상의 기술자라면, 다음의 청구항들에 의해 정의되는 바와 같은 본 발명의 보호 범위 내에 모두 포함되는 많은 수정들 및 변경들을 전술한 솔루션에 적용할 수 있다.

도면

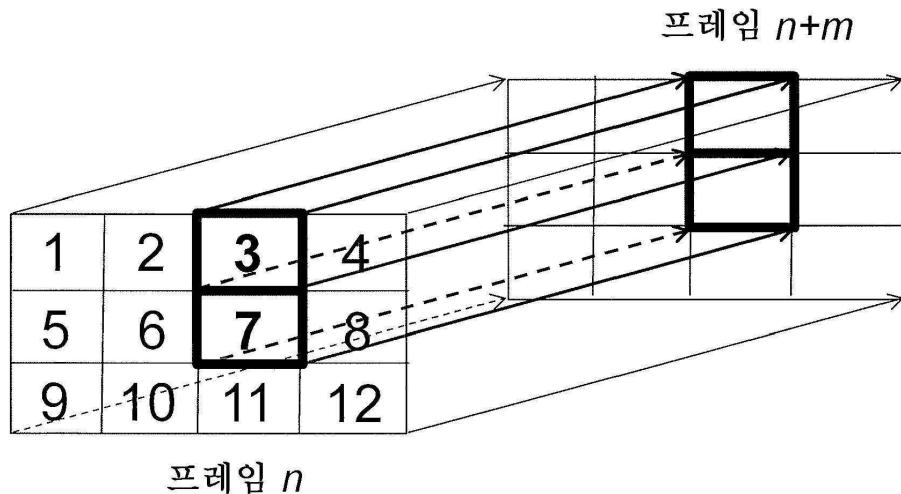
도면 1a



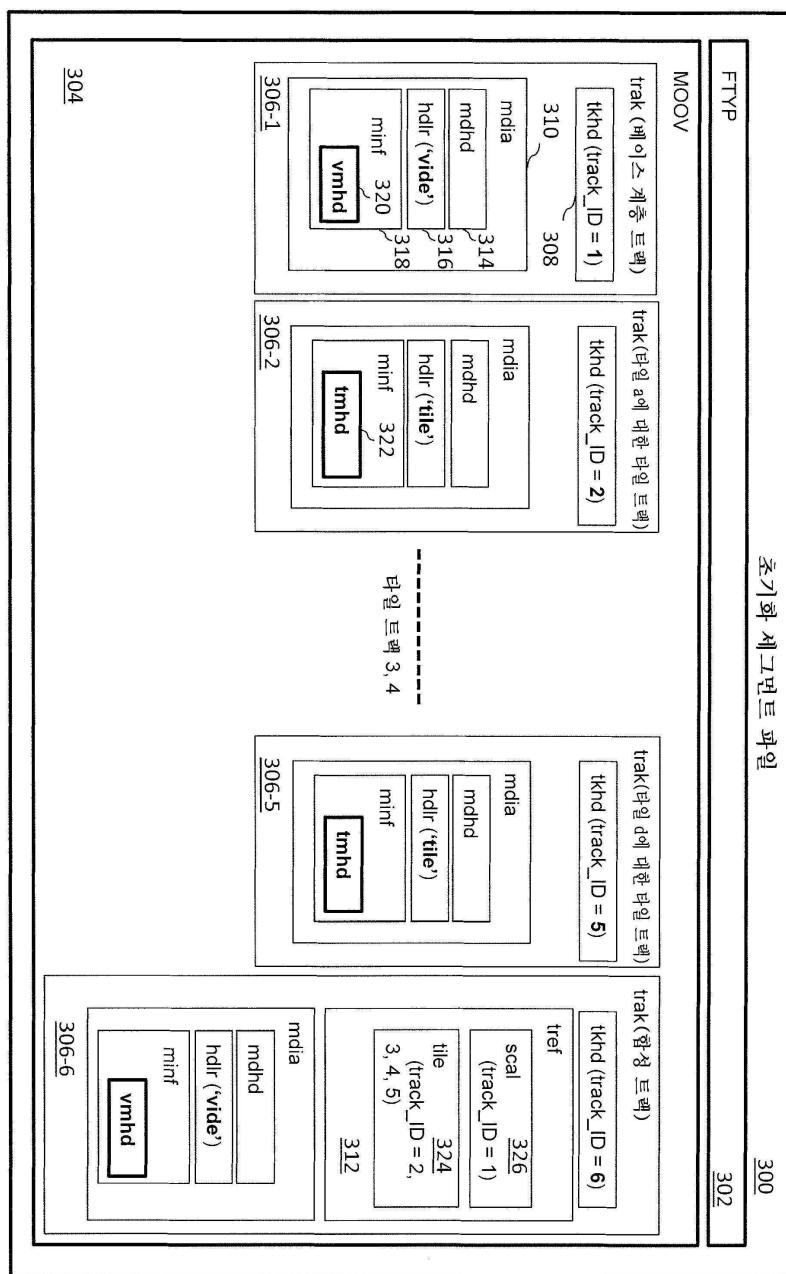
도면1b



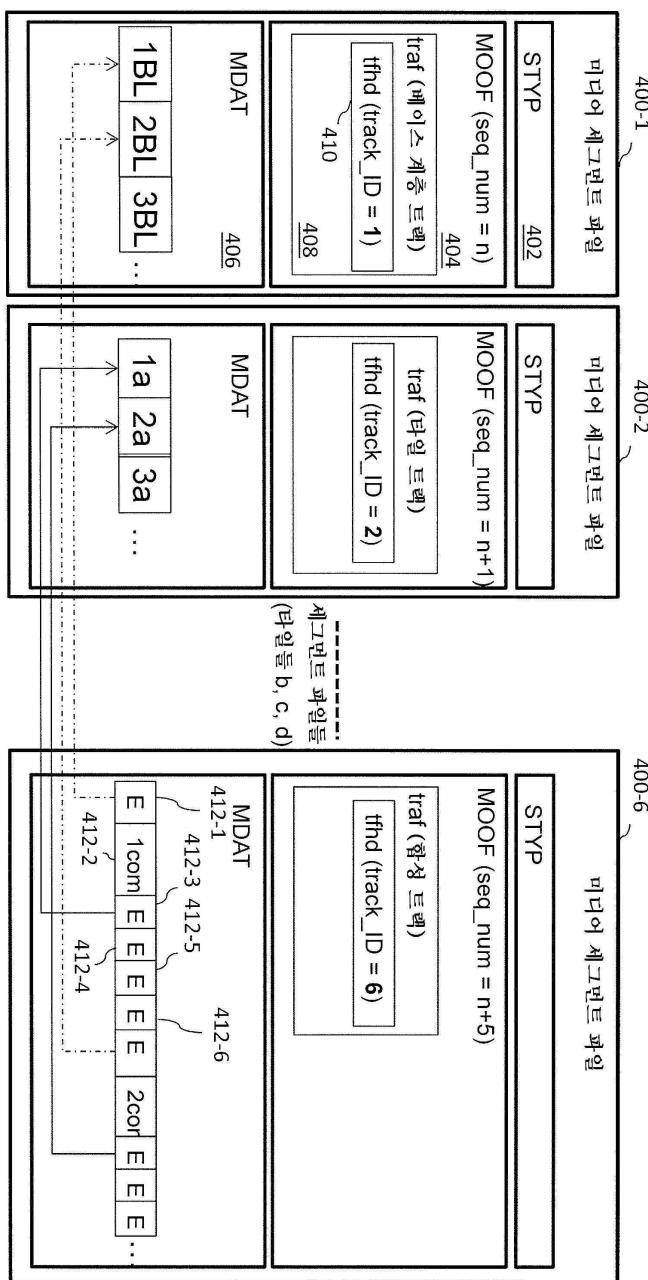
도면2



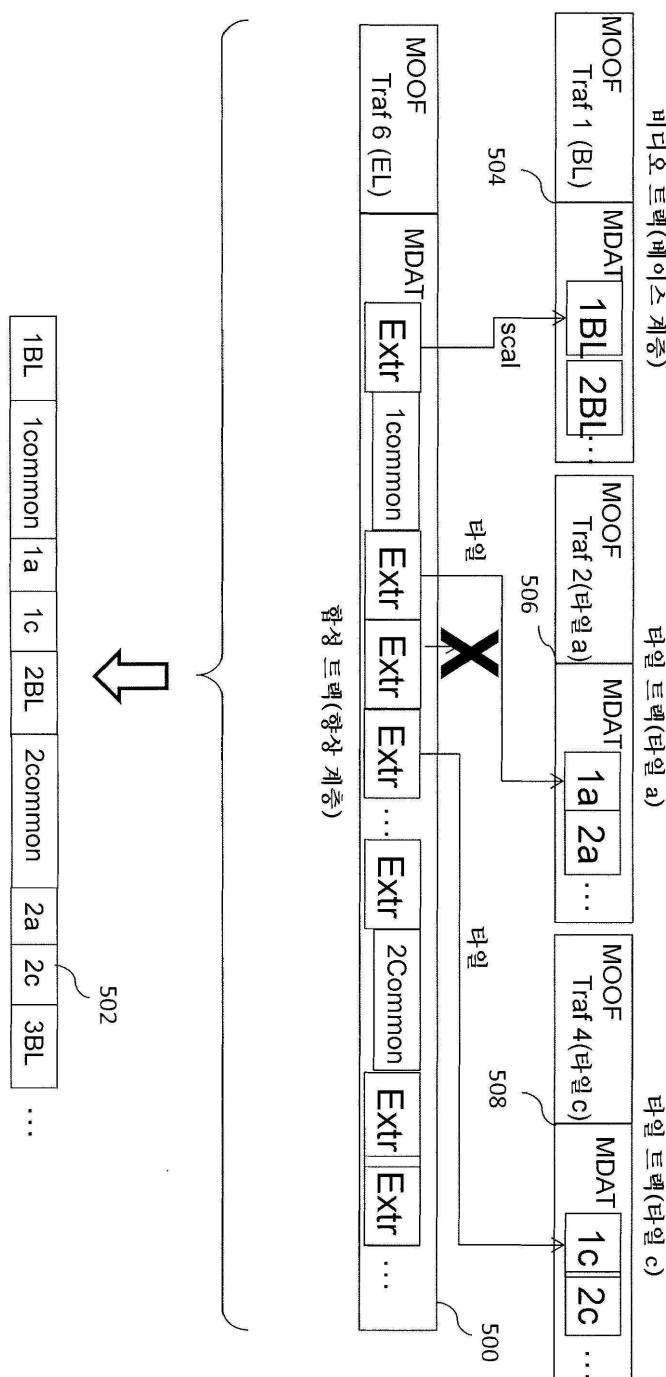
도면3



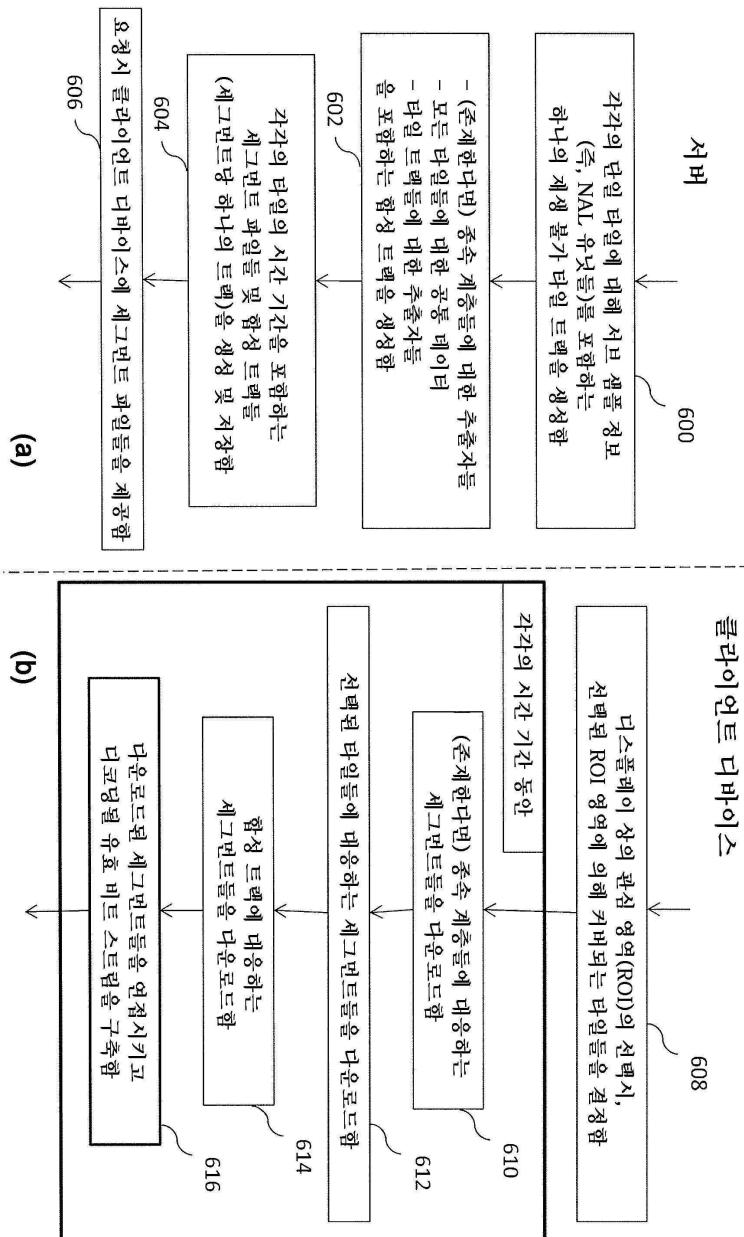
도면4



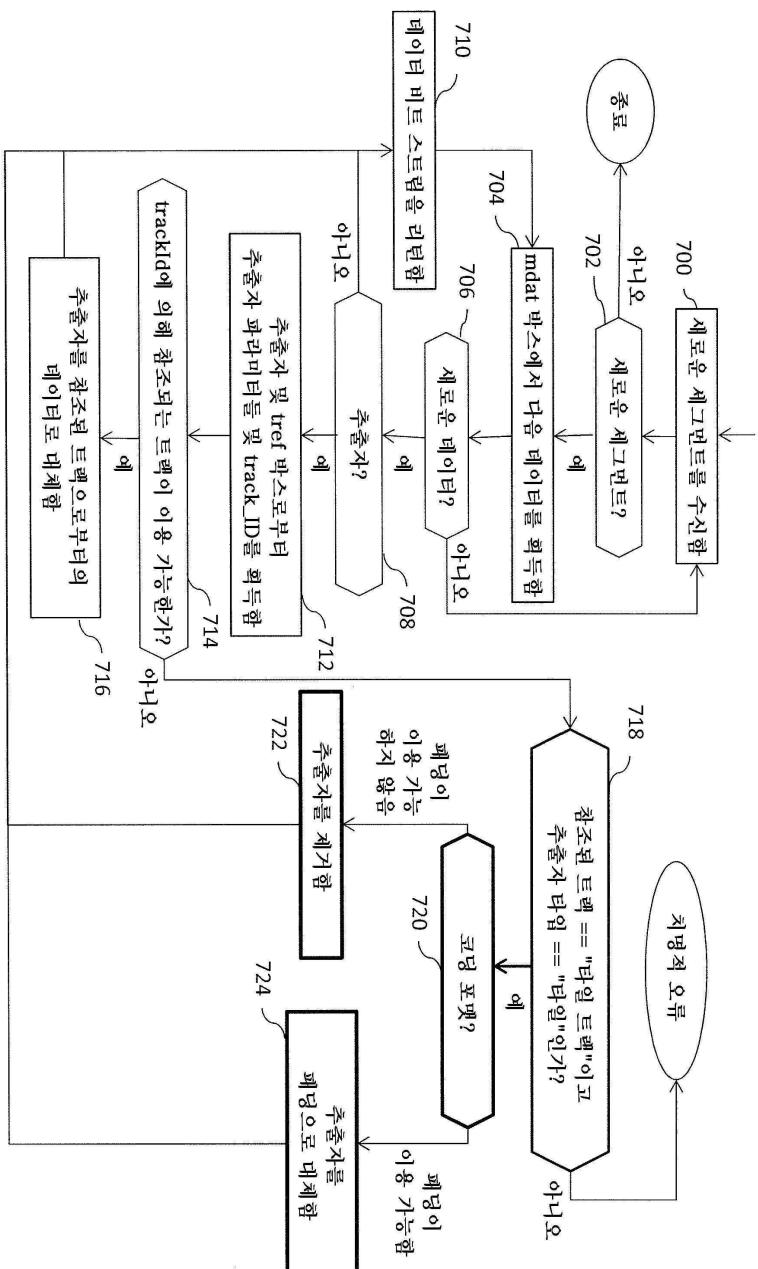
도면5



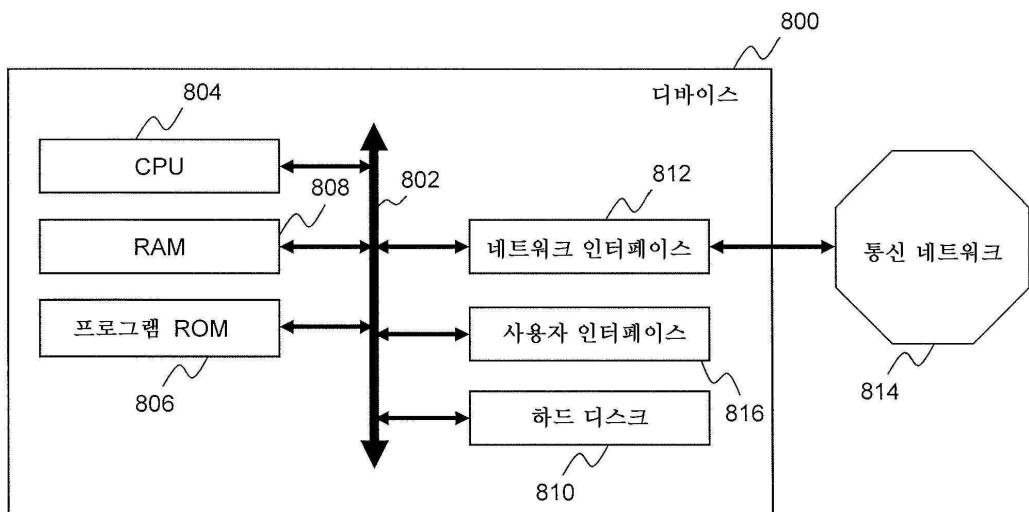
도면6



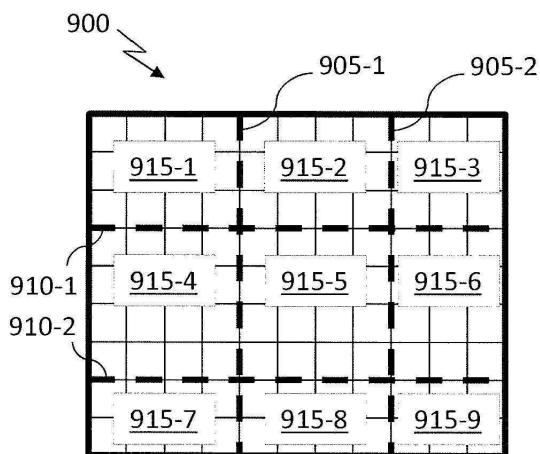
도면7



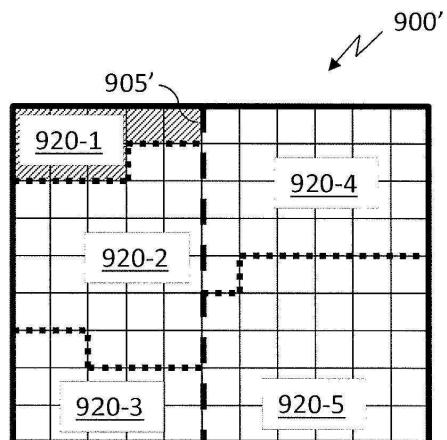
도면8



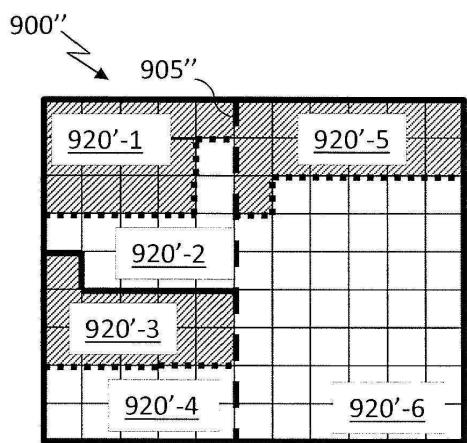
도면9a



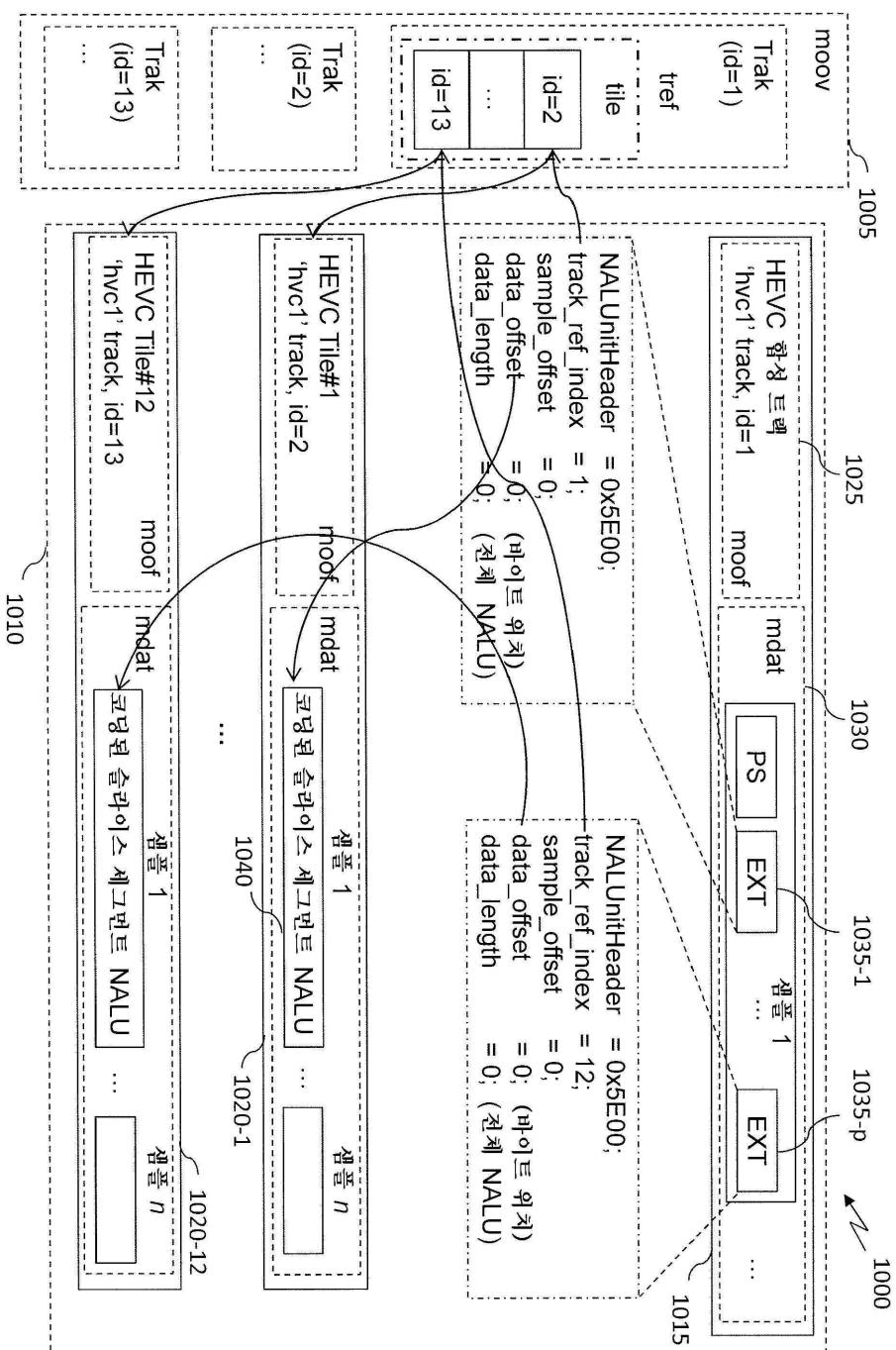
도면9b



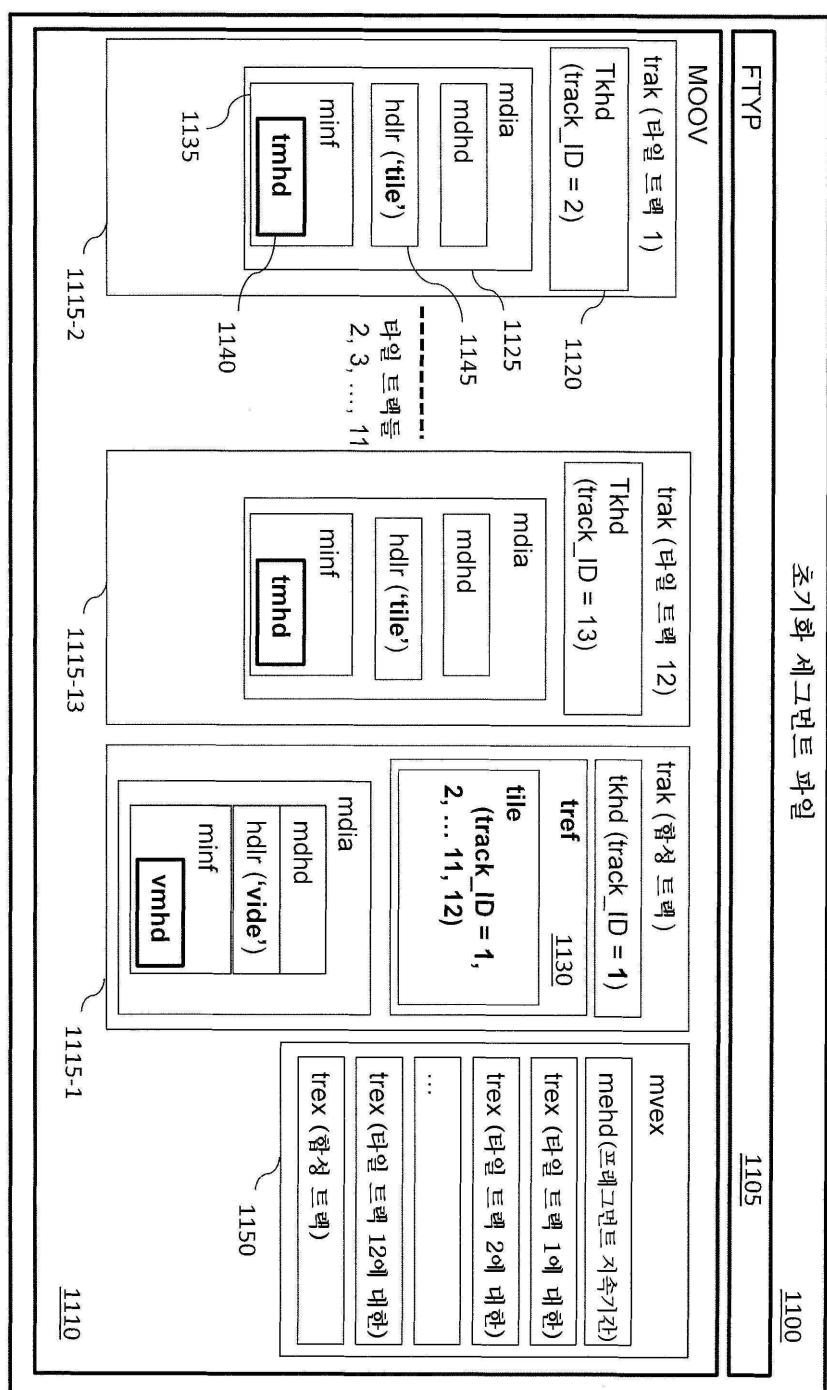
도면9c



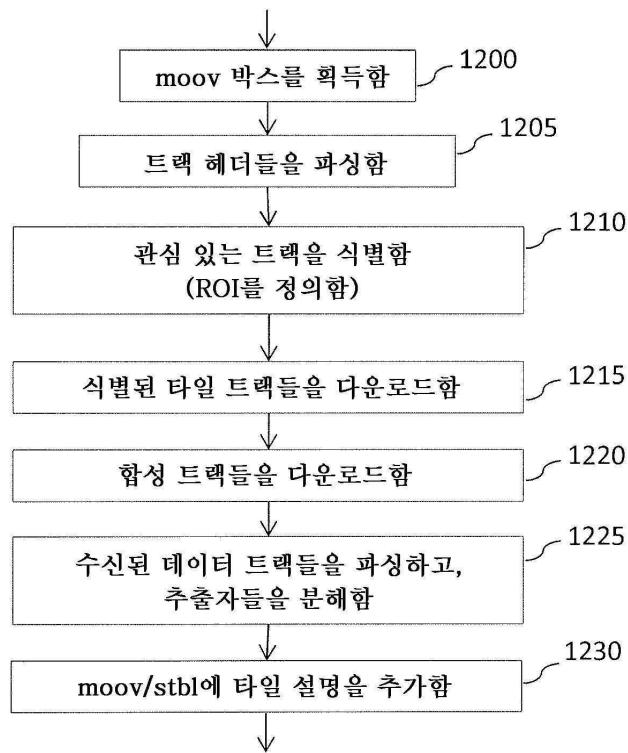
도면10



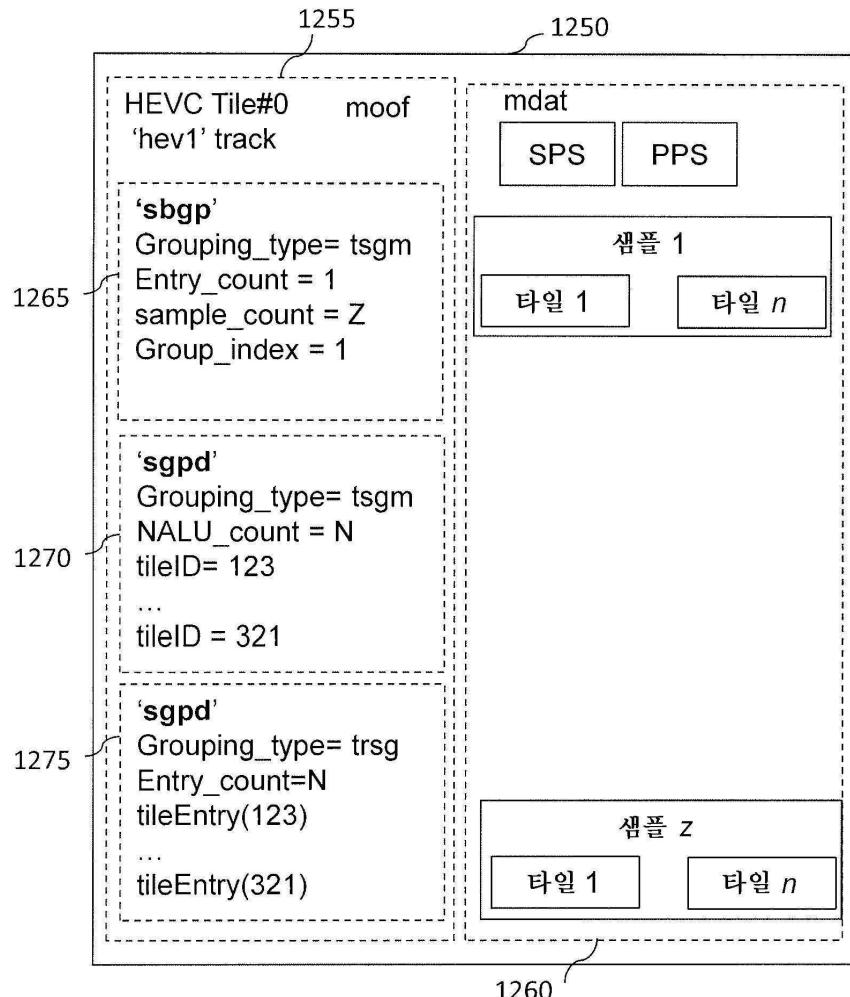
도면11



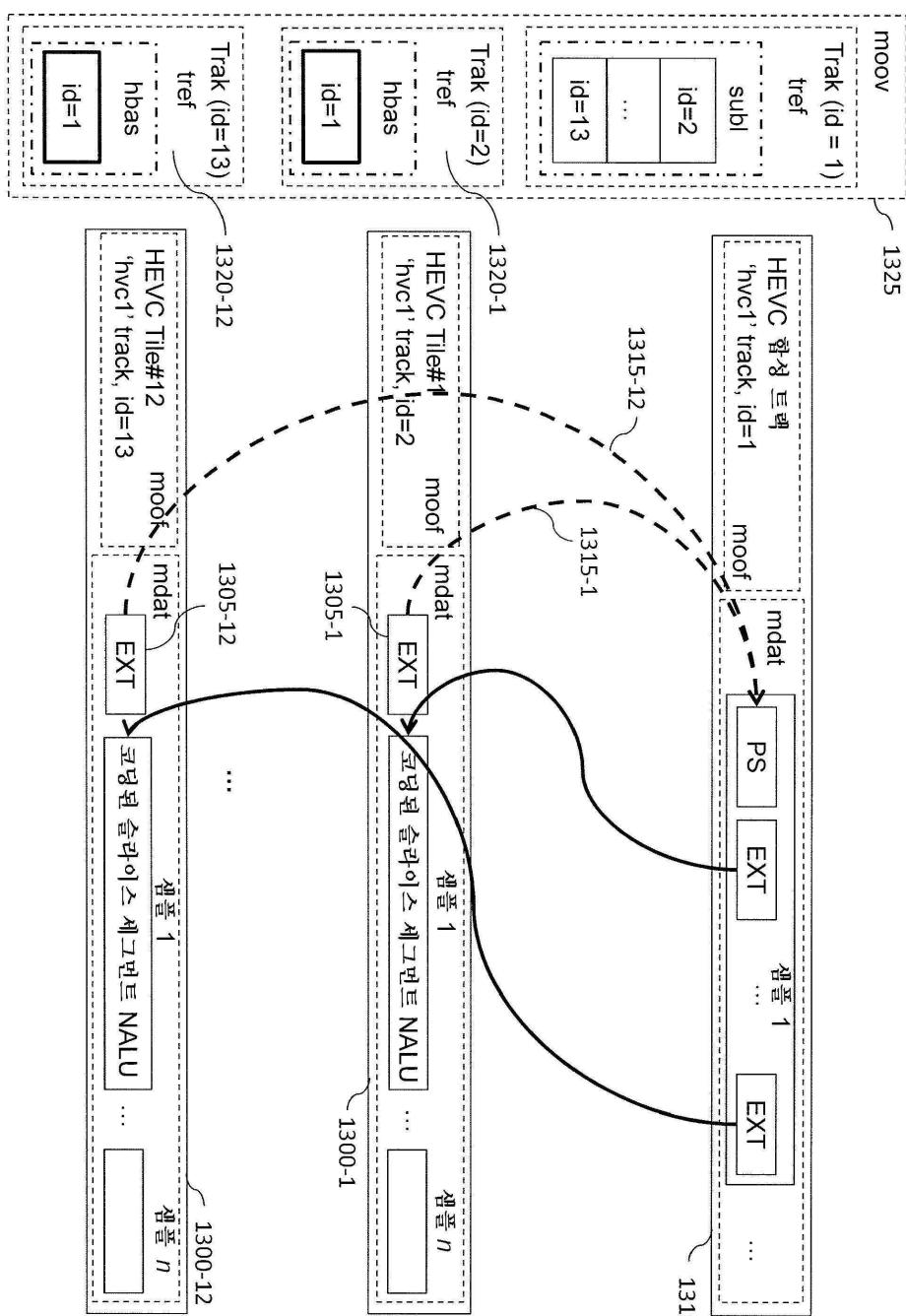
도면12a



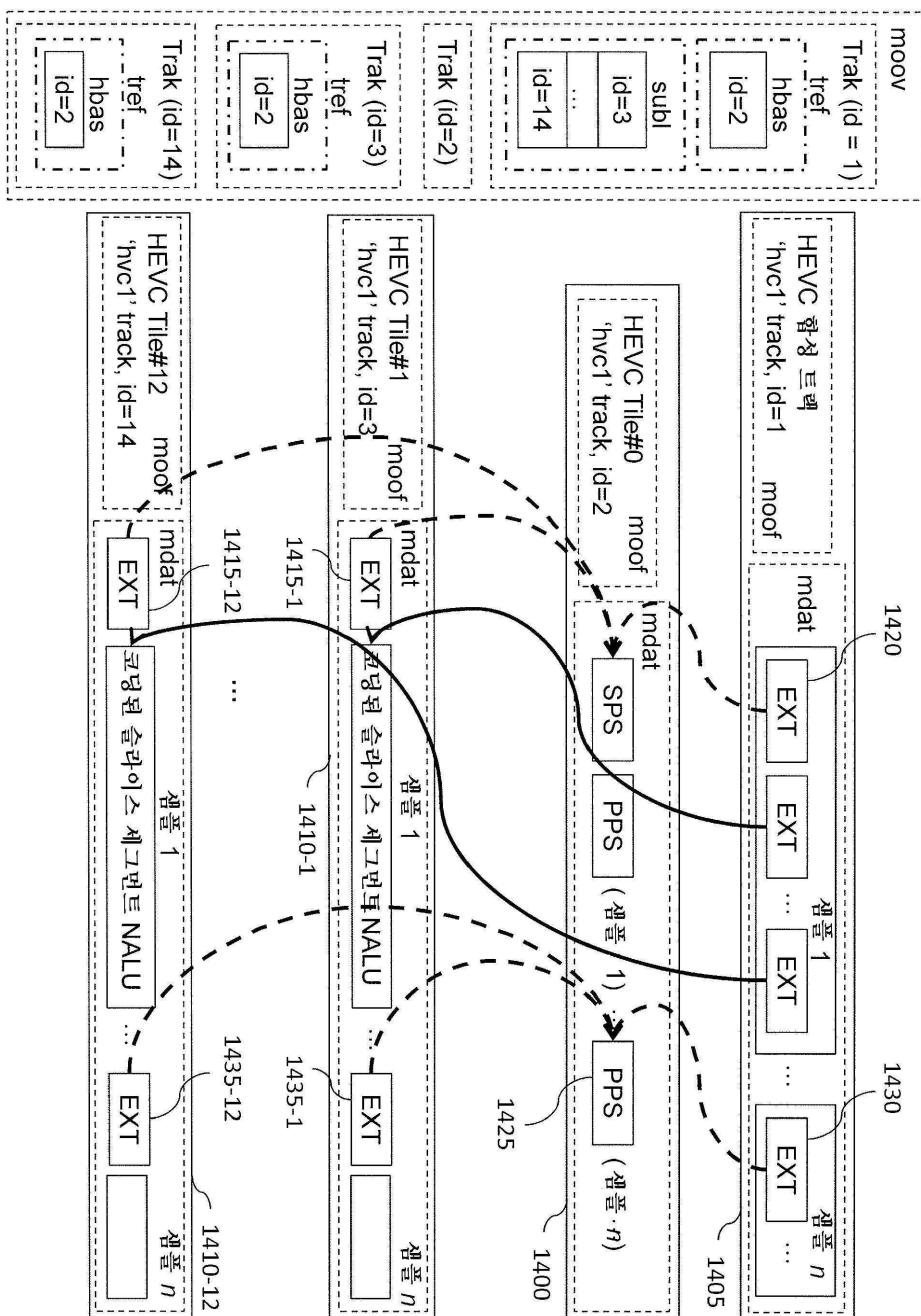
도면12b



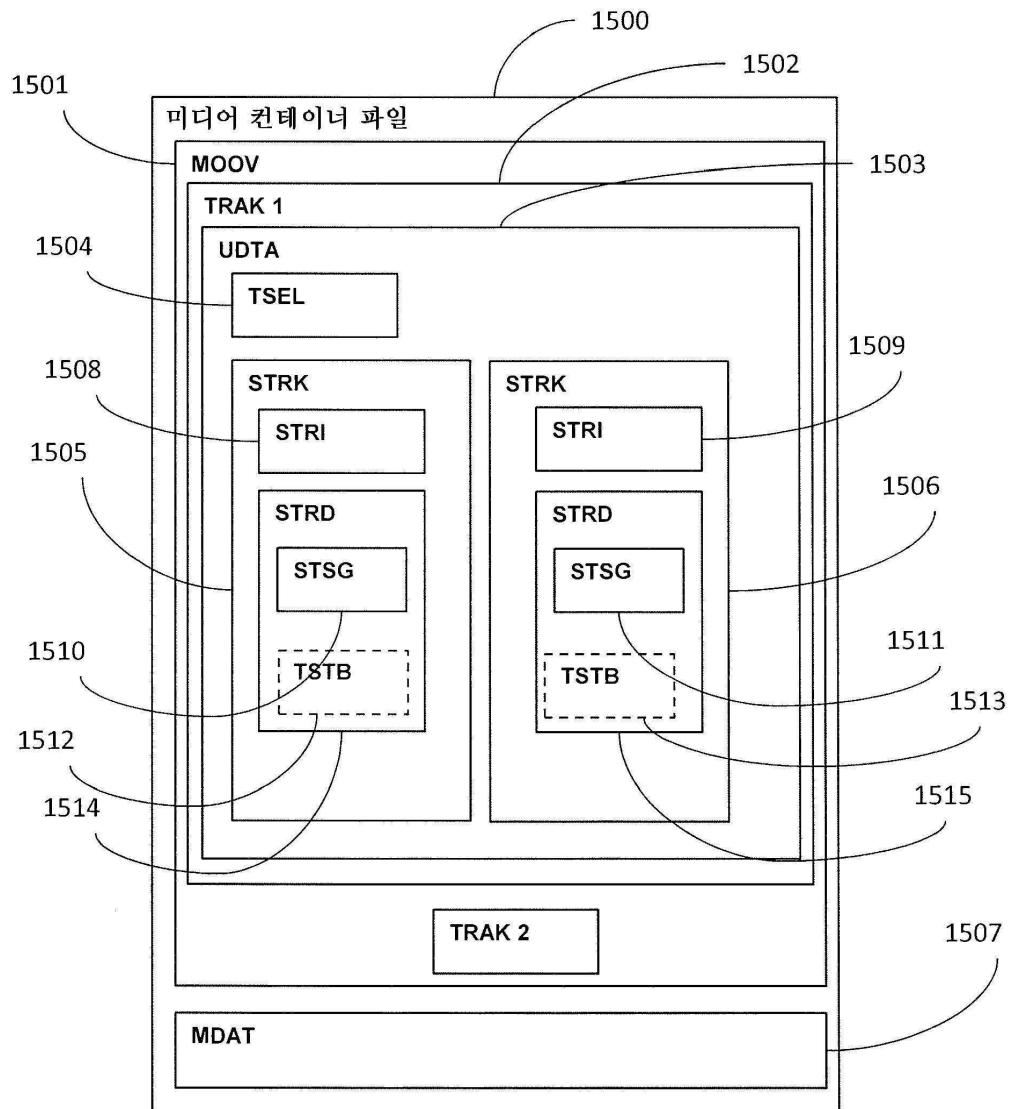
도면13



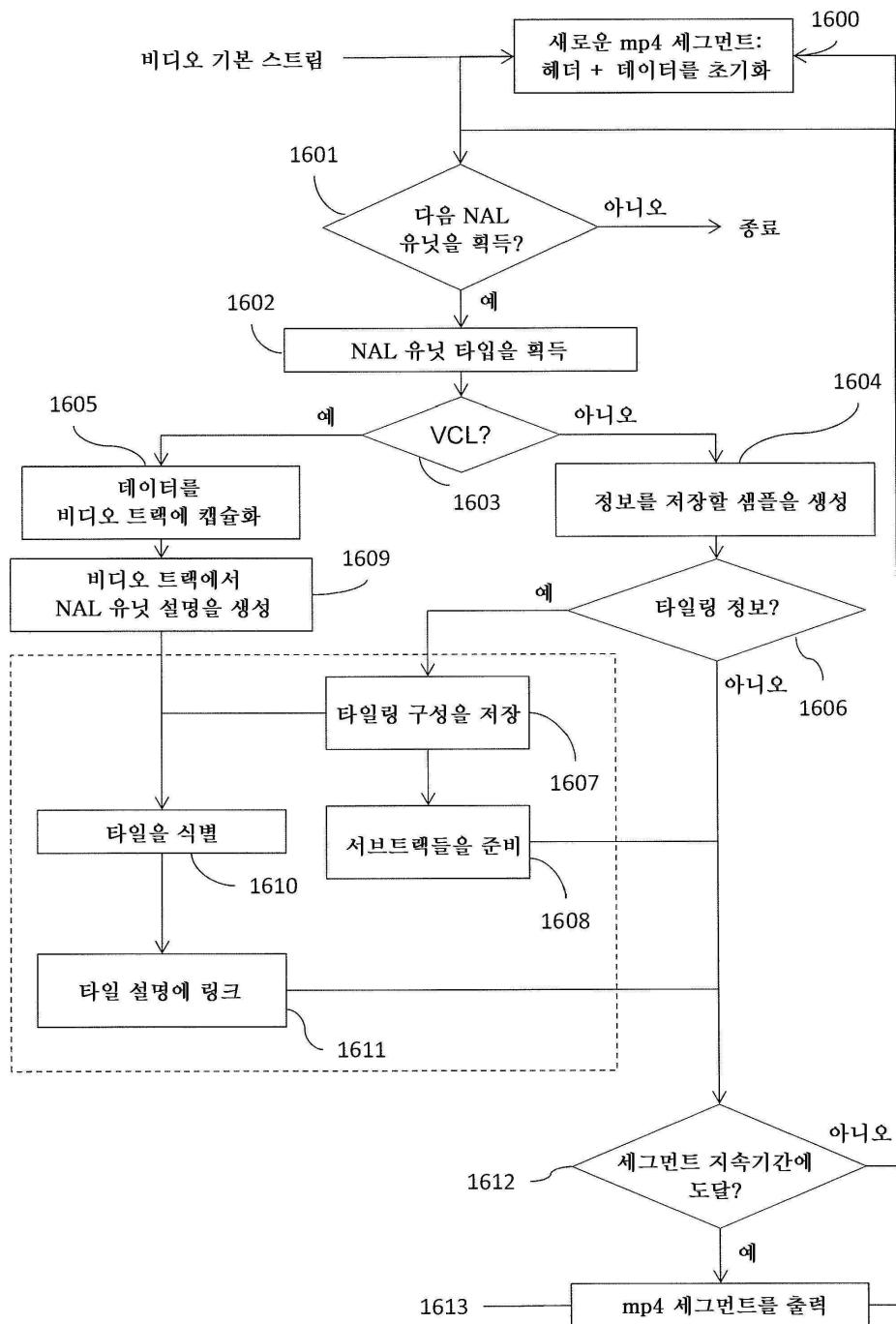
도면14



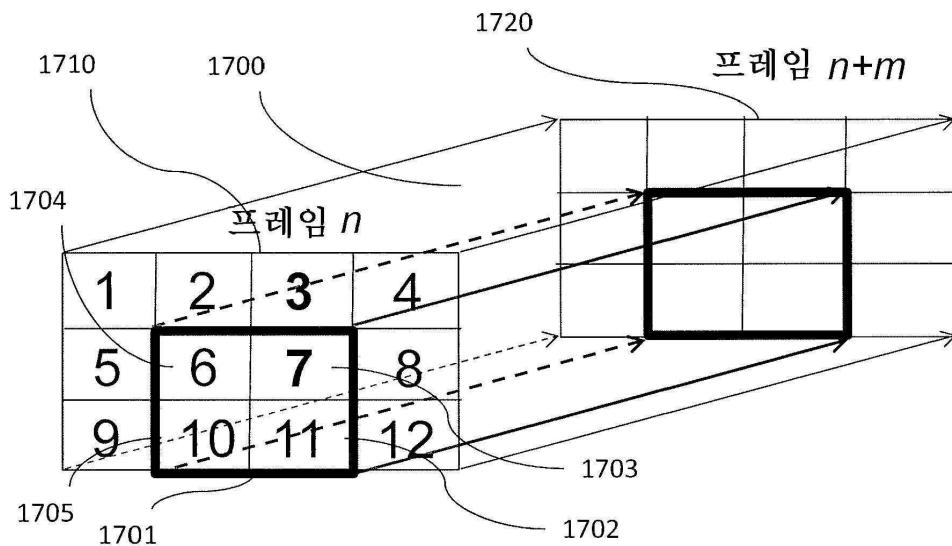
도면15



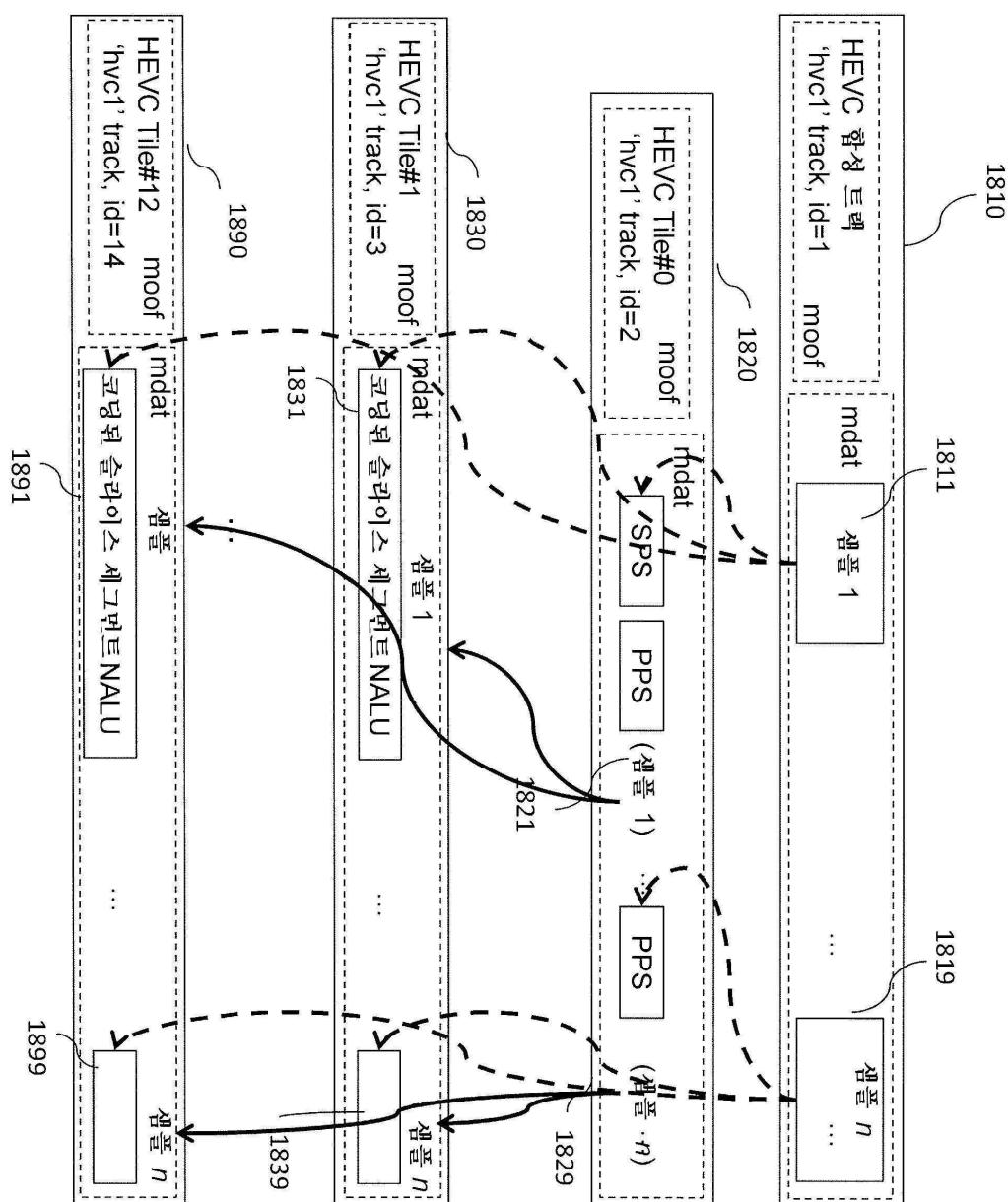
도면16



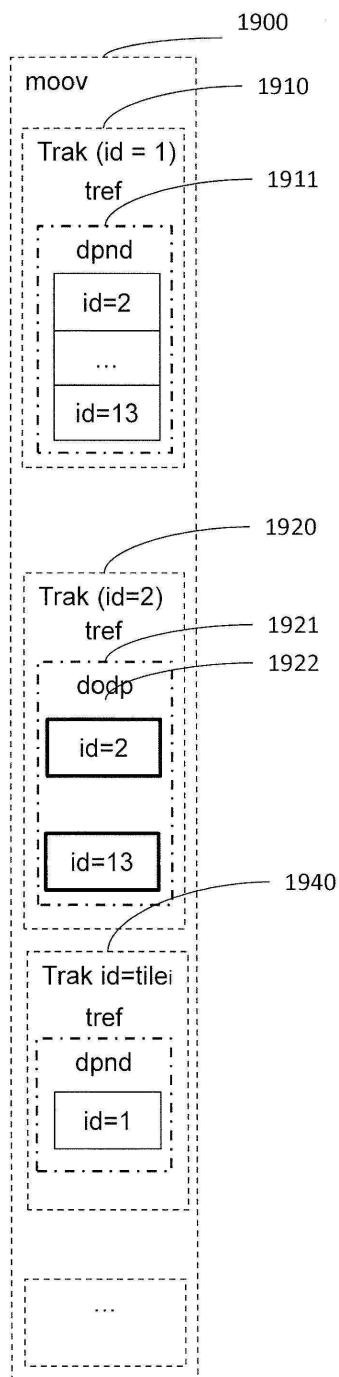
도면17



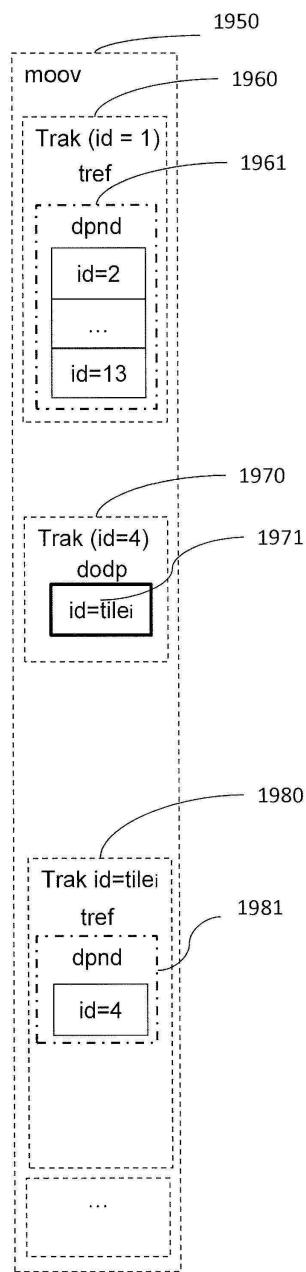
도면18



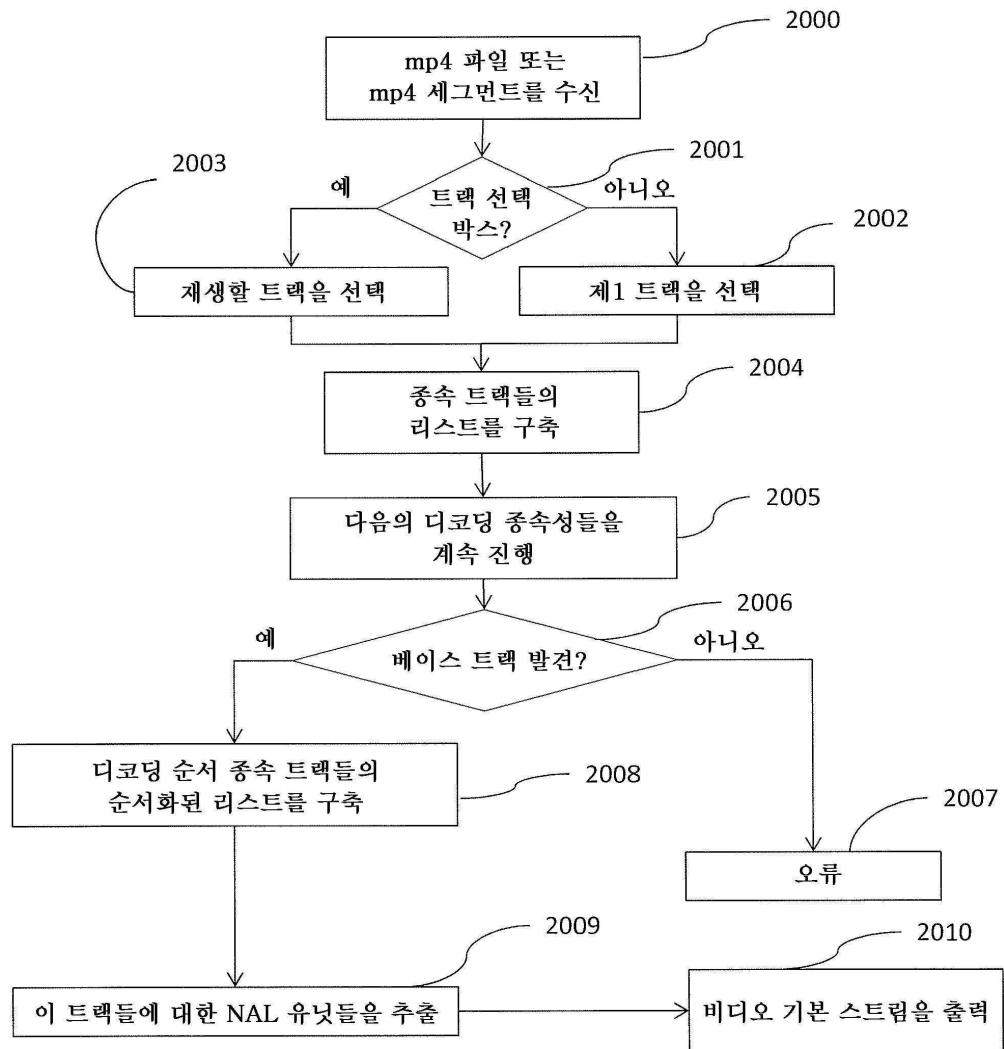
도면19a



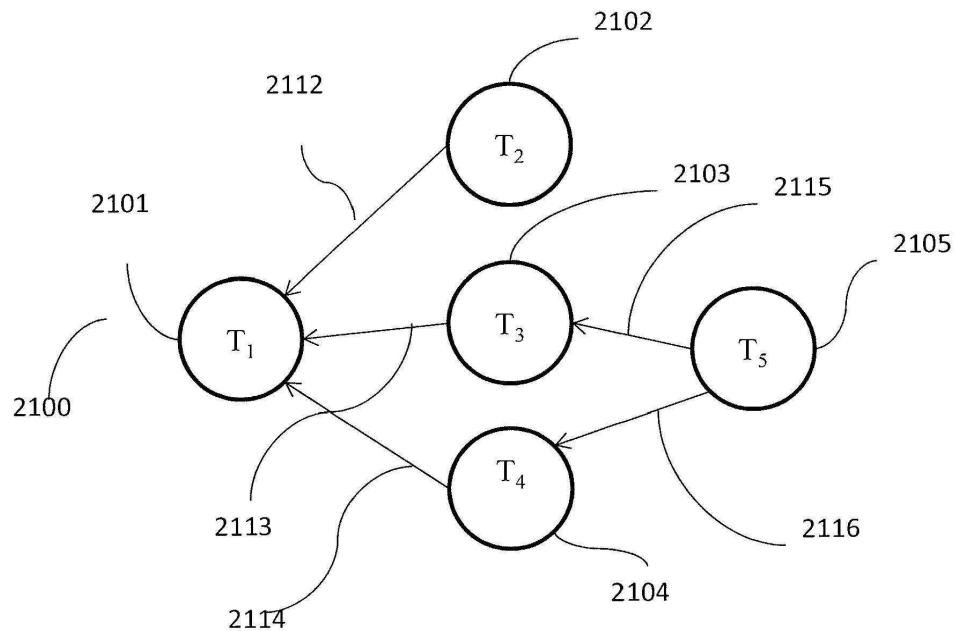
도면19b



도면20



도면21a



도면21b

