



(19) 대한민국특허청(KR)

(12) 등록특허공보(B1)

(45) 공고일자 2018년05월28일

(11) 등록번호 10-1861979

(24) 등록일자 2018년05월21일

- (51) 국제특허분류(Int. Cl.)
HO4N 21/854 (2011.01) *HO4N 21/4728* (2011.01)
HO4N 21/845 (2011.01)
- (52) CPC특허분류
HO4N 21/85406 (2013.01)
HO4N 21/4728 (2013.01)
- (21) 출원번호 10-2017-7016056(분할)
- (22) 출원일자(국제) 2014년01월17일
 심사청구일자 2017년06월12일
- (85) 번역문제출일자 2017년06월12일
- (65) 공개번호 10-2017-0071612
- (43) 공개일자 2017년06월23일
- (62) 원출원 특허 10-2015-7019368
 원출원일자(국제) 2014년01월17일
 심사청구일자 2015년07월17일
- (86) 국제출원번호 PCT/EP2014/050952
- (87) 국제공개번호 WO 2014/111547
 국제공개일자 2014년07월24일

- (30) 우선권주장
 1300953.5 2013년01월18일 영국(GB)
 1306451.4 2013년04월09일 영국(GB)

(56) 선행기술조사문헌

WO2012168365 A1*

Kofler et al. Implications of the ISO base media file format on adaptive HTTP streaming of H. 264/SVC. IEEE CCNC, 2012, pp. 549–553.*

KR1020120116903 A*

*는 심사관에 의하여 인용된 문헌

전체 청구항 수 : 총 14 항

심사관 : 김창원

(54) 발명의 명칭 미디어 데이터 영역 및 메타 데이터 영역을 포함하는 미디어 파일의 생성 방법 및 디바이스

(57) 요약

본 발명은 서버에서 타일형 시간 설정형 미디어 데이터를 캡슐화하는 것 및 클라이언트 디바이스에서 캡슐화된 타일형 시간 설정형 미디어 데이터로부터 시간 설정형 미디어 데이터를 제공하는 것에 관련된다. 이러한 타일형 시간 설정형 미디어 데이터는 시간 설정형 샘플들을 포함하고, 각각의 시간 설정형 샘플은 복수의 공간 서브샘플들을 포함한다. 시간 설정형 샘플들 중 하나의 복수의 공간 서브샘플들 중에서 적어도 하나의 공간 서브샘플을 선택한 후, 선택된 공간 서브샘플 및 다른 시간 설정형 샘플들 각각의 하나의 대응 공간 서브샘플을 포함하는 하나의 타일 트랙이, 각각의 선택된 공간 서브샘플에 대해 생성된다. 이렇게 생성된 타일 트랙들 중 적어도 하나를 식별하는 적어도 하나의 추출자를 포함하는 하나의 합성 트랙이 생성된다. 각각의 생성된 트랙은 적어도 하나의 미디어 세그먼트 파일에 독립적으로 캡슐화된다.

대표도



(52) CPC특허분류

H04N 21/8451 (2013.01)

H04N 21/8456 (2013.01)

(72) 발명자

드누알 프랑크

프랑스 애프-35190 생 도미누크 라 빌 에스 래 8

꼰꼴라또 씨릴

프랑스 애프-77380 꼼즈 라 빌 뤼 드 라 애르스 18

르 폐브르 장

프랑스 애프-94230 까성 아브뉘 뒤 프레지던 월슨

7 에스끄 베1

명세서

청구범위

청구항 1

ISO BMFF(International Standard Organization Base Media File Format)의 미디어 파일의 생성 방법으로서, 상기 미디어 파일은 ISO BMFF에 따른 메타 데이터 영역 및 미디어 데이터 영역을 포함하고, 각각이 하나 이상의 공간 타일을 포함하는 복수의 샘플을 포함하는 미디어 데이터를 획득하는 단계; 및 상기 획득된 데이터에 기초한 복수의 NAL(Network Abstraction Layer) 유닛이 상기 미디어 데이터 영역에 기술되고, 상기 미디어 데이터 영역에 기술된 상기 복수의 NAL 유닛에 대응하는 상기 하나 이상의 공간 타일에 연관된 정보를 나타내는 타일 정보가 상기 메타 데이터 영역에 기술되는 상기 미디어 파일을 생성하는 단계를 포함하는, 생성 방법.

청구항 2

제1항에 있어서,

상기 타일 정보는

하나의 샘플에 대응하는 복수의 공간 타일들 각각을 식별하기 위한 타일 식별자 정보, 하나의 샘플에 대응하는 복수의 공간 타일들 각각의 폭을 식별하기 위한 타일 폭 정보, 및 하나의 샘플에 대응하는 복수의 공간 타일들 각각의 높이를 식별하기 위한 타일 높이 정보를 포함하는, 생성 방법.

청구항 3

제2항에 있어서,

상기 타일 정보는

하나의 샘플에 대응하는 복수의 공간 타일들 각각의 오프셋을 식별하기 위한 타일 오프셋 정보를 포함하는, 생성 방법.

청구항 4

제2항에 있어서,

상기 타일 정보는

상기 하나 이상의 공간 타일과 다른 공간 타일 간의 디코딩 의존성을 시그널링하기 위한 파라미터를 더 포함하는, 생성 방법.

청구항 5

제4항에 있어서,

상기 파라미터는,

상기 하나 이상의 공간 타일과 동일 샘플이나 이전 샘플 내의 다른 타일 간의 코딩 종속성이 알려지지 않음을 시그널링하는 제1 값,

상기 하나 이상의 공간 타일과 동일 샘플 내의 다른 타일 간에 공간 코딩 종속성이 없지만, 상기 하나 이상의 공간 타일과 이전 샘플 내의 동일 타일 식별자를 갖는 공간 타일 간에 코딩 종속성이 있을 수 있음을 시그널링하는 제2 값,

상기 하나 이상의 공간 타일과 동일 샘플이나 이전 샘플 내의 동일 타일 식별자를 갖는 다른 공간 타일 간의 코

당 종속성이 없음을 시그널링하는 제3 값, 및
예약된 값을 시그널링하는 제4 값
중 선택된 하나의 값으로 설정되는, 생성 방법.

청구항 6

제5항에 있어서,
상기 파라미터는 2 비트 파라미터이고, 상기 제1 내지 제4 값은 각각 "0", "1", "2" 및 "3"인, 생성 방법.

청구항 7

미디어 데이터 영역 및 메타 데이터 영역을 포함하는 미디어 파일의 생성 디바이스로서,
상기 디바이스는
각각이 하나 이상의 공간 타일을 포함하는 복수의 샘플을 포함하는 미디어 데이터를 획득하는 단계; 및
상기 획득된 미디어 데이터에 기초한 복수의 NAL(Network Abstraction Layer) 유닛이 상기 미디어 데이터 영역
에 기술되고, 상기 미디어 데이터 영역에 기술된 상기 복수의 NAL 유닛에 대응하는 상기 하나 이상의 공간 타일
에 연관된 정보를 나타내는 타일 정보가 상기 메타 데이터 영역에 기술되는 상기 미디어 파일을 생성하는 단계
를 수행하도록 구성되는 하나 이상의 마이크로프로세서를 포함하는, 디바이스.

청구항 8

제7항에 있어서,
상기 타일 정보는
하나의 샘플에 대응하는 복수의 공간 타일들 각각을 식별하기 위한 타일 식별자 정보,
하나의 샘플에 대응하는 복수의 공간 타일들 각각의 폭을 식별하기 위한 타일 폭 정보, 및
하나의 샘플에 대응하는 복수의 공간 타일들 각각의 높이를 식별하기 위한 타일 높이 정보를 포함하는, 디바이
스.

청구항 9

제8항에 있어서,
상기 타일 정보는
하나의 샘플에 대응하는 복수의 공간 타일들 각각의 오프셋을 식별하기 위한 타일 오프셋 정보를 포함하는, 디
바이스.

청구항 10

제8항에 있어서,
상기 타일 정보는
상기 하나 이상의 공간 타일과, 동일 샘플 및 이전 샘플 중 하나 또는 양자 내의 다른 공간 타일 간의 디코딩
의존성을 시그널링하기 위한 파라미터를 더 포함하는, 디바이스.

청구항 11

제10항에 있어서,
상기 마이크로프로세서는 상기 파라미터를,
상기 하나 이상의 공간 타일과 동일 샘플이나 이전 샘플 내의 다른 타일 간의 코딩 종속성이 알려지지 않음을
시그널링하는 제1 값,

상기 하나 이상의 공간 타일과 동일 샘플 내의 다른 타일 간에 공간 코딩 종속성이 없지만, 상기 하나 이상의 공간 타일과 이전 샘플 내의 동일 타일 식별자를 갖는 공간 타일 간에 코딩 종속성이 있을 수 있음을 시그널링하는 제2 값,

상기 하나 이상의 공간 타일과 동일 샘플이나 이전 샘플 내의 동일 타일 식별자를 갖는 다른 공간 타일 간의 코딩 종속성이 없음을 시그널링하는 제3 값, 및

예약된 값을 시그널링하는 제4 값

중 선택된 하나의 값으로 설정하도록 구성된, 디바이스.

청구항 12

제11항에 있어서,

상기 파라미터는 2 비트 파라미터이고, 상기 제1 내지 제4 값은 각각 "0", "1", "2" 및 "3"인, 디바이스.

청구항 13

제1항에 있어서,

ISO BMFF에 따른 메타 데이터 영역은 ISO BMFF에 따른 MDAT 박스를 포함하는, 생성 방법.

청구항 14

제1항에 있어서,

ISO BMFF에 따른 메타 데이터 영역은 ISO BMFF에 따른, MOOV 박스 및 MOOF 박스 중 적어도 하나를 포함하는, 생성 방법.

발명의 설명

기술 분야

[0001] 본 발명은 일반적으로, 예를 들어, MPEG 표준화 기구에 의해 정의되는 바와 같은 Base Media File Format에 따라, 스트림 전달을 향상시키는, 시간 설정형 미디어 데이터의 캡슐화의 분야에 관한 것으로, 특히 압축된 비디오 스트림들에서 사용자가 선택한 관심의 영역의 HTTP(HyperText Transfer Protocol) 스트리밍에 관한 것이다. 보다 구체적으로, 본 발명은, 특히 하나 이상의 타일들의, 데이터의 효율적인 스트리밍을 가능하게 하는 공간 타일들과 같은 분할된 데이터를 포함하는 기본 스트림의 캡슐화에 대한 방법, 디바이스 및 컴퓨터 프로그램에 관한 것이다.

배경 기술

[0002] 비디오 코딩은, 일련의 비디오 이미지들을 소형의 디지털화된 비트 스트림으로 변환하여 이러한 비디오 이미지들이 송신되거나 저장될 수 있는 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 기준 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, 스트리밍할 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 세그먼트들에, 및 표현될 시간 설정형 미디어의 조직화를 나타내는 "MPD(Media Presentation Description)"라 불리우는 매니페스트에 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, 바이트 범위로 조합될 수 있다. 매니페스트에 기초하여, 클라이언트 디바이스는, 자신의 필요성들, 자신의 능력들(예를 들어, 지원되는 코덱들, 디스플레이 사이즈, 프레임 레이트, 품질 등)에 따라, 및 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)으로, 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 최대 해상도로 비디오를 액세스하는 리소스들(예를 들어, 네트워크 액세스 대역폭 또는 CPU(Central Processing Unit))을 갖는 것은 아니며, 특히 비디오가 초 고화질일 때 그려하며, 모든 사용자들이 이러한 비디오에 액세스할 필요는 없다. 이러한 정황에서, ROI들(Regions-Of-Interest)만을 액세스하는, 즉 전체 비디오 시퀀스 중 일부 공간 서브 파트들만을 액세스하는 능력을 제공하는 것이 특히 유리하다.

[0006] 비디오에 속하는 프레임들의 공간 서브 파트들을 액세스하는 공지의 메커니즘은, 비디오의 각 프레임을 일반적으로 타일이라고 하는 독립적으로 디코드 가능한 공간 영역들의 배열로서 조직화하는 것으로 이루어진다. SVC(Scalable Video Coding) 또는 HEVC(High Efficiency Video Coding)와 같은 일부 비디오 포맷들은 타일 화질에 대한 지원을 제공한다. 사용자 정의의 ROI는 하나 또는 여러 연속적인 타일들을 커버할 수 있다.

[0007] 따라서, 사용자 선택의 ROI들을 HTTP 프로토콜에 따라 스트리밍하기 위해서는, 하나 이상의 타일들에 대한 공간 액세스들을 가능하게 하고 액세스된 타일들의 조합을 가능하게 하는 방식으로, 인코딩된 비디오 스트림의 시간 설정형 미디어 데이터의 캡슐화를 제공하는 것이 중요하다.

[0008] 인코딩된 비디오 비트 스트림들은 일반적으로, 풀 프레임들에 대응하는 연속적인 시간적 샘플들의 세트로서 구성되고, 이러한 시간적 샘플들은 디코딩 순서의 기능으로서 조직화된다는 점이 상기되어야 한다. 파일 포맷들은 이러한 인코딩된 비트 스트림을 캡슐화하고 설명하는데 사용된다.

[0009] 설명의 목적상, ISO BMFF(International Standard Organization Base Media File Format)은, 네트워크를 통한 또는 다른 비트 스트림 전달 메커니즘을 통한 로컬 스토리지 또는 송신을 위해 인코딩된 시간 설정형 미디어 데이터 비트 스트림들을 설명하는 잘 알려진 유연하고 확장 가능한 포맷이다. 이러한 파일 포맷은 객체 지향형이다. 이는, 순차적으로 또는 계층적으로 조직화되고, 타이밍 및 구조 파라미터들과 같은 인코딩된 시간 설정형 미디어 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 구조에 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, 적층형(layered) 미디어 데이터 구성은 다수의 종속형 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, 중대 레이어 트랙의 비트 스트림은 베이스 레이어 트랙으로부터 NAL 유닛들을 참조하는 추출자들을 포함할 수 있다. 그리고 나중에, 이러한 중대 레이어 트랙이 파일 포맷으로부터 추출될 때, 추출자들은 그들이 참조하고 있는 데이터로 대체되어야 한다.

[0011] ISO BMFF를 사용하여 이러한 메커니즘들을 구현할 때, 서브 정보를 설명하고, 이러한 서브 정보에 대한 액세스를 용이하게 하거나 또는 비트 스트림들을 다수 세그먼트들에 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 가능성들 뿐만 아니라 제한들을 고려하여 HTTP 스트리밍을 위한 스케일 가능한 비디오 비트 스트림(H264/SVC)을 조직화하는 3가지 상이한 전략들을 제시한다:

[0013] a) (트랙 해상도들을 포함하여) 모든 ISO BMFF 메타데이터를 포함하는 파일 타입 박스 "ftyp" 및 영화 박스 "moov"를 포함하는 특정 파일 헤더를 포함하는 싱글 파일, 이러한 싱글 파일은 또한 전체 인코딩된 비트 스트림을 포함하는 단일 mdat 박스를 포함함. 이러한 조직화는 로컬 스토리지에는 적합하지만 클라이언트가 전체 비

트 스트림의 일부만을 필요로 할 수 있는 HTTP 스트리밍에는 적응되지 않는다.

[0014] b) 프레그멘테이션(fragmentation)에 적합한 다수 moof/mdat 박스들을 포함하는 싱글 파일. 이러한 포맷은 프로그레시브 다운로드를 감안한다. moof 박스는 프레그먼트 레벨에서 moov 박스와 등가이다. 이러한 스킴에 따라, 프레그멘테이트된 미디어 파일을 사용하여, 스케일가능형 비트 스트림은, 비디오를 상이한 스케일가능성 레벨들로 표현하는 다수 종속형 트랙들로 나뉜다. 다른 트랙들로부터 NAL 유닛들을 참조하는데 추출자들이 사용된다. 파일 당 트랙이 사용되는 경우, 모든 어드레스가능형 트랙들은 미리 준비되어야 하고 트랙들은 독립적으로 선택될 수 없다. 여러 파일들이 디스플레이되어야 하면, 여러 비트 스트림들이 디코딩되어야 하고 베이스 레이어가 여러 번 디코딩된다.

[0015] c) 각각의 파일이 자기 자신의 URL에 의해 액세스가능하고 독립적으로 다운로드가능한, 다수 세그먼트 파일들. 각각의 세그먼트는 통상적으로 세그먼트 타입 박스(styp)를 포함하고, 이는 일종의 파일 헤더, 음션형 세그먼트 인덱스 박스(sidx) 및 하나 이상의 다수 프레그먼트들로서 역할을 한다. 다시, 각각의 프레그먼트는 moof 및 mdat 박스를 포함한다. 이러한 스킴에 따라, 프레그먼트된 미디어 파일을 사용하여, 각각의 트랙은 스케일가능성의 한 레벨에 관한 관련된 비트 스트림과 함께 자기 자신의 세그먼트에 저장된다. 필요하다면, 종속형 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 스트리밍에 적합하지 않다. 더욱이, 하나보다 많은 파일을 선택할 때 베이스 레이어의 비트 스트림의 잠재적인 중복이 존재한다.

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

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

발명의 내용

해결하려는 과제

[0018] 이러한 제약들과 직면하여, 본 발명자들은, 서버에서 파일형 시간 설정형 미디어 데이터를 캡슐화하고, 복수의 미디어 세그먼트 파일들에 캡슐화된 파일형 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 단점들을 해결하는 것이다.

과제의 해결 수단

[0020] 본 발명의 제1 양태에 따르면, 분할된 시간 설정형 미디어 데이터를 서버에서 캡슐화하는 방법으로서- 분할된 시간 설정형 미디어 데이터는 시간 설정형 샘플들을 포함하고, 각각의 시간 설정형 샘플은 복수의 서브샘플들을 포함함 -,

[0021] 시간 설정형 샘플들 중 하나의 복수의 서브샘플들 중에서 적어도 하나의 서브샘플을 선택하는 단계;

[0022] 각각의 선택된 서브샘플에 대해, 선택된 서브샘플 및 다른 시간 설정형 샘플들 각각의 하나의 대응 서브샘플을 포함하는 하나의 분할 트랙을 생성하는 단계;

[0023] 생성된 분할 트랙들 중 적어도 하나를 식별하는 적어도 하나의 추출자를 포함하는 적어도 하나의 참조 트랙을 생성하는 단계; 및

[0024] 생성된 트랙들 각각을 적어도 하나의 미디어 세그먼트 파일에 독립적으로 캡슐화하는 단계를 포함하는 방법이 제공된다.

[0025] 따라서, 본 발명은, 클라이언트 디바이스에 의해 다운로드되는 부분들 또는 파일들의 선택된 세트가 무엇이든, 상이한 부분들, 특히 상이한 파일들의 조합, 및 분석시 유효 파일 포맷 및 기본 스트림의 생성을 가능하게

한다.

[0026] 따라서, 본 발명의 방법은, 독립형 부분들 또는 타일들의 효율적인 스트리밍에 적합하고, 유용한 데이터만이 클라이언트 디바이스에 보내어질 것을 요구하고, 하나보다 많은 부분 또는 타일 선택(즉, 임의의 ROI)에 적합하고, (바이트 범위 요청들에 비해) 인덱스화 오버헤드를 감소시키며, MPEG 표준에 통합될 수 있다.

[0027] 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 선택적으로 초기화 데이터를 포함할 수 있다.

[0028] 일 실시예에서, 적어도 하나의 추출자는 적어도 하나의 식별된 타일 트랙의 적어도 하나의 공간 서브샘플을 더욱 식별한다.

[0029] 일 실시예에서, 본 방법은, 생성된 타일 트랙들 및 생성된 적어도 하나의 합성 트랙으로부터 타일형 시간 설정형 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 참조대상 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.

[0030] 일 실시예에서, 각각의 선택된 공간 서브샘플에 대해 하나의 타일 트랙을 생성하는 단계에서 적어도 2개의 타일 트랙들이 생성되고, 적어도 하나의 합성 트랙은 동일한 시간 설정형 샘플 내에 포함되는 2개의 공간 서브샘플들 중 임의의 것을 디코딩하는데 사용될 데이터의 적어도 하나의 아이템을 포함한다.

[0031] 일 실시예에서, 타일형 시간 설정형 미디어 데이터는, 시간 설정형 샘플들 각각이 그에 따라 참조 레이어 및 적어도 하나의 중대 레이어를 포함하는, 스케일 가능한 타일형 시간 설정형 미디어 데이터이고, 참조 레이어는 적어도 하나의 참조 공간 서브샘플을 포함하고, 적어도 하나의 중대 레이어는 복수의 중대 공간 서브샘플들을 포함하며, 각각의 선택된 공간 서브샘플에 대해 하나의 타일 트랙을 생성하는 단계에서 생성된 타일 트랙들은 중대 공간 서브샘플들을 포함하는 중대 타일 트랙들이고, 본 방법은 참조 공간 서브샘플들을 포함하는 적어도 하나의 참조 트랙을 생성하는 단계를 더 포함한다.

[0032] 일 실시예에서, 참조 레이어는 복수의 참조 공간 서브샘플들을 포함하고, 복수의 참조 트랙들이 생성되며, 복수의 참조 트랙들의 각 참조 트랙은 참조 타일 트랙을 형성한다.

[0033] 일 실시예에서, 생성된 트랙들 각각을 캡슐화한 결과인 미디어 세그먼트 파일들은 대응 트랙의 해상도들을 포함하는 메타데이터를 저장하는 데이터 구조를 포함한다.

[0034] 일 실시예에서, 생성된 트랙들 각각을 캡슐화한 결과인 미디어 세그먼트 파일들 중 적어도 하나는 대응 트랙의 적어도 하나의 그룹의 서브샘플들의 적어도 하나의 해상도를 포함하는 메타데이터를 저장하는 데이터 구조를 포함한다.

[0035] 일 실시예에서, 생성된 트랙들 각각을 캡슐화한 결과인 미디어 세그먼트 파일들 중 적어도 하나는 대응 트랙의 상이한 그룹들의 서브샘플들의 상이한 해상도들을 포함하는 메타데이터를 저장하는 데이터 구조를 포함한다.

[0036] 일 실시예에서, 메타데이터를 저장하는 데이터 구조는, 적어도 하나의 해상도, 및 적어도 하나의 해상도가 대응 트랙의 모든 서브샘플들에 적용된다는 것을 나타내는 버전 타입을 포함한다.

[0037] 일 실시예에서, 대응 트랙의 적어도 하나의 그룹의 서브샘플들의 적어도 하나의 해상도를 포함하는 메타데이터를 저장하는 데이터 구조 중 적어도 하나의 파라미터는, 서브샘플 디코딩 종속성과 관련하여, 그룹의 각 서브샘플이 다른 그룹의 서브샘플을 사용하지 않고 디코딩될 수 있다는 점을 나타낸다.

[0038] 일 실시예에서, 한 그룹의 서브샘플들은 서브샘플들이 속하는 그룹화 타입에 따라 식별된다.

[0039] 일 실시예에서, 한 그룹의 서브샘플들은 서브샘플들이 속하는 그룹화 타입에 따라 식별되고, 그룹화 타입은 해당 그룹의 각 서브샘플에 대한 해상도를 제공한다.

[0040] 일 실시예에서, 초기화 세그먼트 파일은, 대응하는 적어도 하나의 타일 트랙을 캡슐화한 결과인 미디어 세그먼트 파일에 저장된 임의의 특정 해상도와 관련되지 않는, 적어도 하나의 타일 트랙의 서브샘플들을 처리하는데

디폴트로 사용될 하나의 해상도를 포함한다.

- [0041] 일 실시예에서, 본 방법은, 디폴트로 사용될 초기화 세그먼트 파일에 저장된 해상도를 수정하는 것을 가능하게 하도록, 생성된 트랙들 각각을 캡슐화한 결과인 미디어 세그먼트 파일들 중 적어도 하나에 초기화 데이터를 추가하는 단계를 더 포함한다.
- [0042] 일 실시예에서, 생성된 트랙들 각각을 캡슐화한 결과인 미디어 세그먼트 파일들은 공간 서브샘플들을 저장하는 데이터 구조를 포함한다.
- [0043] 생성된 트랙들 각각을 캡슐화한 결과인 미디어 세그먼트 파일들은 추출자들을 저장하는 데이터 구조를 포함한다.
- [0044] 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.
- [0045] 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.
- [0046] 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.
- [0047] 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.
- [0048] 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.
- [0049] 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.
- [0050] 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.
- [0051] 일 실시예에서, 서버는 HTTP(HyperText Transfer Protocol)와 호환가능하다.
- [0052] 일 실시예에서, 생성된 트랙들 각각을 캡슐화한 결과인 미디어 세그먼트 파일들은 International Standard Organization에 의해 정의되는 바와 같은 Base Media File Format 및 Dynamic Adaptive Streaming over HTTP 포맷과 호환가능하다.
- [0053] 일 실시예에서, 분할된 시간 설정형 미디어 데이터를 서버에서 캡슐화하는 방법- 분할된 시간 설정형 미디어 데이터는 시간 설정형 샘플들을 포함하고, 각각의 시간 설정형 샘플은 복수의 공간 서브샘플들을 포함함 -은,
- [0054] 시간 설정형 샘플들 중 하나의 복수의 서브샘플들 중에서 적어도 하나의 서브샘플을 선택하는 단계;
- [0055] 각각의 선택된 공간 서브샘플에 대해, 선택된 서브샘플 및 다른 시간 설정형 샘플들 각각의 하나의 대응 서브샘플을 포함하는 하나의 분할 트랙을 생성하는 단계;
- [0056] 제1 분할 트랙에 포함되는 제1 공간 서브샘플과 제2 분할 트랙에 포함되는 제2 공간 서브샘플을 관련시키는 정보를 포함하는 적어도 하나의 참조 트랙을 생성하는 단계; 및
- [0057] 생성된 각각의 트랙을 적어도 하나의 미디어 세그먼트 파일에 독립적으로 캡슐화하는 단계를 포함한다.
- [0058] 이러한 일 실시예는 본 발명의 제1 양태를 참조하여 위에 개시된 이전의 실시예들 중 하나 이상과 조합될 수 있다.
- [0059] 본 발명의 제2 양태는, 클라이언트 디바이스에서, 복수의 미디어 세그먼트 파일들에 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 서브샘플들을 포함하고, 복수의 미디어 세그먼트 파일들은 적어도 하나의 참조 트랙 및 적어도 하나의 분할 트랙을 포함하고, 적어도 하

나의 분할 트랙 각각은, 시간 설정형 샘플들 중 하나의 복수의 서브샘플들 중에서 선택되는 하나의 서브샘플을 포함하고, 다른 시간 설정형 샘플들 각각의 하나의 대응 공간 서브샘플을 포함하고, 적어도 하나의 참조 트랙은, 적어도 하나의 분할 트랙을 식별하고 식별된 분할 트랙의 적어도 하나의 서브샘플을 식별하는 적어도 하나의 추출자를 포함함 -,

[0060] 서브샘플을 나타내는 정보의 아이템을 선택하는 단계;

[0061] 선택된 서브샘플을 포함하는 적어도 하나의 분할 트랙을 요청하는 단계- 선택된 서브샘플은 선택된 정보의 아이템에 대응함 -;

[0062] 복수의 미디어 세그먼트 파일들을 수신하는 단계; 및

[0063] 시간 설정형 미디어 비트 스트림을 생성하도록 적어도 하나의 참조 트랙의 추출자들을 대응 식별된 서브샘플들로 대체하는 단계를 포함하는 방법을 제공한다.

[0064] 따라서, 본 발명은, 클라이언트 디바이스에 의해 다운로드되는 부분들 또는 파일들의 선택된 세트가 무엇이든, 상이한 부분들, 특히 상이한 파일들의 조합, 및 분석시 유효 파일 포맷 및 기본 스트림의 생성을 가능하게 한다.

[0065] 따라서, 본 발명의 방법은, 독립형 부분들 또는 파일들의 효율적인 스트리밍에 적합하고, 유용한 데이터만이 클라이언트 디바이스에 보내어질 것을 요구하고, 하나보다 많은 부분 또는 파일 선택(즉, 임의의 ROI)에 적합하고, (바이트 범위 요청들에 비해) 인덱스화 오버헤드를 감소시키며, MPEG 표준에 통합될 수 있다.

[0066] 일 실시예에서, 분할된 시간 설정형 미디어 데이터는 파일형 시간 설정형 미디어 데이터이고, 서브샘플들은 공간 서브샘플들이고, 적어도 하나의 분할 트랙은 적어도 하나의 파일 트랙이며, 적어도 하나의 참조 트랙은 적어도 하나의 합성 트랙이다.

[0067] 일 실시예에서, 분할된 시간 설정형 미디어 데이터는 파일형 시간 설정형 미디어 데이터이고, 서브샘플들은 공간 서브샘플들이고, 적어도 하나의 분할 트랙은 적어도 하나의 파일 트랙이며, 적어도 하나의 참조 트랙은 추출자들 및 선택적으로 초기화 데이터를 포함하는 적어도 하나의 합성 트랙이다.

[0068] 일 실시예에서, 본 방법은,

[0069] 적어도 하나의 합성 트랙의 추출자로부터 트랙 참조를 취득하는 단계;

[0070] 취득된 트랙 참조에 대응하는 트랙이 수신되었는지 여부를 체크하는 단계; 및

[0071] 취득된 트랙 참조에 대응하는 트랙이 수신되지 않았으면, 시간 설정형 미디어 데이터 비트 스트림을 생성하도록 대응 추출자를 제거하는 단계를 더 포함한다.

[0072] 일 실시예에서, 본 방법은, 취득된 트랙 참조에 대응하는 트랙이 파일 타입의 트랙인지 확인하는 단계를 더 포함한다.

[0073] 일 실시예에서, 본 방법은,

[0074] 적어도 하나의 합성 트랙의 추출자로부터 트랙 참조를 취득하는 단계;

[0075] 취득된 트랙 참조에 대응하는 트랙이 수신되었는지 여부를 체크하는 단계; 및

[0076] 취득된 트랙 참조에 대응하는 트랙이 수신되지 않았으면, 시간 설정형 미디어 데이터 비트 스트림을 생성하도록 대응 추출자를 패딩으로 대체하는 단계를 더 포함한다.

[0077] 일 실시예에서, 본 방법은, 취득된 트랙 참조에 대응하는 트랙이 파일 타입의 트랙인지 확인하는 단계를 더 포함한다.

[0078] 일 실시예에서, 본 방법은, 생성된 파일 트랙들 및 생성된 적어도 하나의 합성 트랙으로부터 파일형 시간 설정형 미디어 데이터의 공간 부분의 형성을 가능하게 하는 파라미터들을 포함하는 초기화 세그먼트 파일을 수신하는 단계를 더 포함하다.

[0079] 일 실시예에서, 적어도 2개의 파일 트랙들을 나타내는 적어도 2개의 미디어 세그먼트 파일들이 수신되고, 적어도 하나의 합성 트랙은 동일 세트의 공간 서브샘플들에 속하는 2개의 공간 서브샘플들 중 임의의 것을 디코딩하는데 사용될 데이터의 적어도 하나의 아이템을 포함한다.

- [0080] 일 실시예에서, 타일형 시간 설정형 미디어 데이터는, 시간 설정형 샘플들 각각이 그에 따라 참조 레이어 및 적어도 하나의 중대 레이어를 포함하는, 스케일 가능한 타일형 시간 설정형 미디어 데이터이고, 참조 레이어는 적어도 하나의 참조 공간 서브샘플을 포함하고, 적어도 하나의 중대 레이어는 복수의 중대 공간 서브샘플들을 포함하며, 적어도 2개의 트랙들을 나타내는 적어도 2개의 미디어 세그먼트 파일들이 수신되고, 적어도 2개의 수신된 트랙들 중 하나는 참조 공간 서브샘플들을 포함하는 참조 트랙이고, 적어도 2개의 수신된 트랙들 중 다른 하나는 중대 공간 서브샘플들을 포함하는 중대 타일 트랙이다.
- [0081] 일 실시예에서, 참조 레이어는 복수의 참조 공간 서브샘플들을 포함하고, 각각 참조 타일 트랙을 형성하는, 복수의 참조 트랙들이 수신된다.
- [0082] 일 실시예에서, 미디어 세그먼트 파일들 중 적어도 하나는, 대응 트랙의 적어도 하나의 그룹의 서브샘플들의 적어도 하나의 해상도를 포함하는 메타데이터를 저장하는 데이터 구조를 포함하고, 본 방법은, 대응 트랙의 적어도 하나의 그룹의 서브샘플들을 처리하기 위해 적어도 하나의 해상도를 취득하는 단계를 포함한다.
- [0083] 일 실시예에서, 미디어 세그먼트 파일들 중 적어도 하나는 참조 트랙에 저장된 초기화 데이터를 포함하고, 본 방법은, 세그먼트 파일들 중 적어도 하나를 표준 미디어 세그먼트 파일로서 독립적으로 처리하는 단계를 더 포함한다.
- [0084] 일 실시예에서, 미디어 세그먼트 파일들 중 적어도 하나는 참조 트랙에 저장된 초기화 데이터를 식별하는 적어도 하나의 추출자를 저장하는 데이터 구조를 포함하고, 본 방법은, 미디어 세그먼트 파일들 중 적어도 하나를 처리할 때 참조 트랙에 저장된 데이터를 액세스하는 단계를 더 포함한다.
- [0085] 일 실시예에서, 미디어 세그먼트 파일들 중 적어도 하나는 합성 트랙의 초기화 데이터를 식별하는 적어도 하나의 추출자를 저장하는 데이터 구조를 포함하고, 본 방법은, 세그먼트 파일들 중 적어도 하나를 표준 미디어 세그먼트 파일로서 독립적으로 처리하기 위해 참조 트랙에 저장된 데이터를 액세스하는 단계를 더 포함한다.
- [0086] 일 실시예에서, 본 방법은, 초기화 데이터를 포함하는 적어도 하나의 초기화 데이터 트랙을 취득하는 단계를 더 포함하고, 적어도 하나의 초기화 데이터 트랙은 적어도 하나의 미디어 세그먼트 파일에 독립적으로 캡슐화된다.
- [0087] 일 실시예에서, 미디어 세그먼트 파일들 중 적어도 하나는 초기화 데이터 트랙에 저장된 초기화 데이터를 식별하는 적어도 하나의 추출자를 저장하는 데이터 구조를 포함하고, 본 방법은, 미디어 세그먼트 파일들 중 적어도 하나를 독립적으로 처리하기 위해 참조 트랙에 저장된 데이터를 액세스하는 단계를 더 포함한다.
- [0088] 일 실시예에서, 클라이언트 디바이스는 HTTP(HyperText Transfer Protocol)와 호환가능하다.
- [0089] 일 실시예에서, 수신된 미디어 세그먼트 파일들은 International Standard Organization에 의해 정의되는 바와 같은 Base Media File Format 및 Dynamic Adaptive Streaming over HTTP 포맷과 호환가능하다.
- [0090] 본 발명의 제3 양태에 따르면, 위에 설명된 방법의 각 단계를 수행하도록 적응된 수단들을 포함하는 디바이스가 제공된다.
- [0091] 본 발명의 제4 양태에 따르면, 서버에서 분할된 시간 설정형 미디어 데이터를 캡슐화하는 디바이스로서- 분할된 시간 설정형 미디어 데이터는 시간 설정형 샘플들을 포함하고, 각각의 시간 설정형 샘플은 복수의 서브샘플들을 포함함 -,
- [0092] 시간 설정형 샘플들 중 하나의 복수의 서브샘플들 중에서 적어도 하나의 서브샘플을 선택하는 단계;
- [0093] 각각의 선택된 서브샘플에 대해, 선택된 서브샘플 및 다른 시간 설정형 샘플들 각각의 하나의 대응 서브샘플을 포함하는 하나의 분할 트랙을 생성하는 단계;
- [0094] 생성된 분할 트랙들 중 적어도 하나를 식별하는 적어도 하나의 추출자를 포함하는 적어도 하나의 참조 트랙을 생성하는 단계; 및
- [0095] 생성된 트랙들 각각을 적어도 하나의 미디어 세그먼트 파일에 독립적으로 캡슐화하는 단계를 수행하도록 구성되는 적어도 하나의 마이크로프로세서를 포함하는 디바이스가 제공된다.
- [0096] 따라서, 본 발명은, 클라이언트 디바이스에 의해 다운로드되는 부분들 또는 파일들의 선택된 세트가 무엇이든, 상이한 부분들, 특히 상이한 파일들의 조합, 및 분석시 유효 파일 포맷 및 기본 스트림의 생성을 가능하게 한다.

- [0097] 따라서, 본 발명의 방법은, 독립형 부분들 또는 타일들의 효율적인 스트리밍에 적합하고, 유용한 데이터만이 클라이언트 디바이스에 보내어질 것을 요구하고, 하나보다 많은 부분 또는 타일 선택(즉, 임의의 ROI)에 적합하고, (바이트 범위 요청들에 비해) 인덱스화 오버헤드를 감소시키며, MPEG 표준에 통합될 수 있다.
- [0098] 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.
- [0099] 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 추출자들 및 선택적으로 초기화 데이터를 포함하는 적어도 하나의 합성 트랙이다.
- [0100] 일 실시예에서, 마이크로프로세서는, 생성된 타일 트랙들 및 생성된 적어도 하나의 합성 트랙으로부터 타일형 시간 설정형 미디어 데이터의 공간 부분의 형성을 가능하게 하는 파라미터들을 포함하는 초기화 세그먼트 파일을 생성하는 단계를 수행하도록 더욱 구성된다.
- [0101] 일 실시예에서, 초기화 세그먼트 파일은 적어도 하나의 생성된 합성 트랙과 관련된 참조대상 타일 트랙들의 리스트를 더 포함한다.
- [0102] 일 실시예에서, 초기화 세그먼트 파일은, 트랙이 공간 서브샘플 정보를 포함한다는 것을 나타내는, 생성된 타일 트랙들 각각과 관련된 타일 핸들러를 더 포함한다.
- [0103] 일 실시예에서, 초기화 세그먼트 파일은 생성된 타일 트랙들 각각과 관련된 일반적 프레젠테이션 정보를 더 포함한다.
- [0104] 일 실시예에서, 마이크로프로세서는, 각각의 선택된 공간 서브샘플에 대해 하나의 타일 트랙을 생성하는 단계에서 적어도 2개의 타일 트랙들이 생성되고, 적어도 하나의 합성 트랙은 동일한 시간 설정형 샘플 내에 포함되는 2개의 공간 서브샘플들 중 임의의 것을 디코딩하는데 사용될 데이터의 적어도 하나의 아이템을 포함하도록 더욱 구성된다.
- [0105] 일 실시예에서, 타일형 시간 설정형 미디어 데이터는, 시간 설정형 샘플들 각각이 그에 따라 참조 레이어 및 적어도 하나의 중대 레이어를 포함하는, 스케일 가능한 타일형 시간 설정형 미디어 데이터이고, 참조 레이어는 적어도 하나의 참조 공간 서브샘플을 포함하고, 적어도 하나의 중대 레이어는 복수의 중대 공간 서브샘플들을 포함하며, 마이크로프로세서는, 각각의 선택된 공간 서브샘플에 대해 하나의 타일 트랙을 생성하는 단계에서 생성된 타일 트랙들이 중대 공간 서브샘플들을 포함하는 중대 타일 트랙들이도록 더욱 구성되고, 마이크로프로세서는, 참조 공간 서브샘플들을 포함하는 적어도 하나의 참조 트랙을 생성하는 단계를 수행하도록 더욱 구성된다.
- [0106] 일 실시예에서, 참조 레이어는 복수의 참조 공간 서브샘플들을 포함하고, 마이크로프로세서는, 복수의 참조 트랙들이 생성도록 더욱 구성되며, 복수의 참조 트랙들의 각 참조 트랙은 참조 타일 트랙을 형성한다.
- [0107] 일 실시예에서, 생성된 트랙들 각각을 캡슐화한 결과인 미디어 세그먼트 파일들은 대응 트랙의 해상도들을 포함하는 메타데이터를 저장하는 데이터 구조를 포함한다.
- [0108] 일 실시예에서, 생성된 트랙들 각각을 캡슐화한 결과인 미디어 세그먼트 파일들은 공간 서브샘플들을 저장하는 데이터 구조를 포함한다.
- [0109] 일 실시예에서, 생성된 트랙들 각각을 캡슐화한 결과인 미디어 세그먼트 파일들은 추출자들을 저장하는 데이터 구조를 포함한다.
- [0110] 일 실시예에서, 서버는 HTTP(HyperText Transfer Protocol)와 호환가능하고, 생성된 트랙들 각각을 캡슐화한 결과인 미디어 세그먼트 파일들은 International Standard Organization에 의해 정의되는 바와 같은 Base Media File Format 및 Dynamic Adaptive Streaming over HTTP 포맷과 호환가능하다.
- [0111] 일 실시예에서, 서버에서 분할된 시간 설정형 미디어 데이터를 캡슐화하는 디바이스는- 분할된 시간 설정형 미디어 데이터는 시간 설정형 샘플들을 포함하고, 각각의 시간 설정형 샘플은 복수의 공간 서브샘플들을 포함함-,
- [0112] 시간 설정형 샘플들 중 하나의 복수의 서브샘플들 중에서 적어도 하나의 서브샘플을 선택하도록 구성되는 선택

유닛;

- [0113] 각각의 선택된 공간 서브샘플에 대해, 선택된 서브샘플 및 다른 시간 설정형 샘플들 각각의 하나의 대응 서브샘플을 포함하는 하나의 분할 트랙을 생성하도록 구성되는 제1 생성 유닛;
- [0114] 제1 분할 트랙에 포함되는 제1 공간 서브샘플과 제2 분할 트랙에 포함되는 제2 공간 서브샘플을 관련시키는 정보를 포함하는 적어도 하나의 참조 트랙을 생성하도록 구성되는 제2 생성 유닛; 및
- [0115] 생성된 각각의 트랙을 적어도 하나의 미디어 세그먼트 파일에 독립적으로 캡슐화하도록 구성되는 캡슐화 유닛을 포함한다.
- [0116] 이러한 일 실시예는 본 발명의 제4 양태를 참조하여 위에 개시된 이전의 실시예들 중 하나 이상과 조합될 수 있다.
- [0117] 본 발명의 제5 양태에 따르면, 위에 설명된 바와 같은 디바이스를 포함하는 비디오 인코더가 제공된다.
- [0118] 본 발명의 제6 양태에 따르면, 클라이언트 디바이스에서, 복수의 미디어 세그먼트 파일들에 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 서브샘플들을 포함하고, 복수의 미디어 세그먼트 파일들은 적어도 하나의 참조 트랙 및 적어도 하나의 분할 트랙을 포함하고, 적어도 하나의 분할 트랙 각각은, 시간 설정형 샘플들 중 하나의 복수의 서브샘플들 중에서 선택되는 하나의 서브샘플을 포함하고, 다른 시간 설정형 샘플들 각각의 하나의 대응 서브샘플을 포함하고, 적어도 하나의 참조 트랙은, 적어도 하나의 분할 트랙을 식별하고 식별된 분할 트랙의 적어도 하나의 서브샘플을 식별하는 적어도 하나의 추출자를 포함함 -,
- [0119] 서브샘플을 나타내는 정보의 아이템을 선택하는 단계;
- [0120] 선택된 서브샘플을 포함하는 적어도 하나의 분할 트랙을 요청하는 단계- 선택된 서브샘플은 선택된 정보의 아이템에 대응함 -;
- [0121] 복수의 미디어 세그먼트 파일들을 수신하는 단계; 및
- [0122] 시간 설정형 미디어 비트 스트림을 생성하도록 적어도 하나의 참조 트랙의 추출자들을 대응 식별된 서브샘플들로 대체하는 단계
- [0123] 를 수행하도록 구성되는 적어도 하나의 마이크로프로세서를 포함하는 디바이스가 제공된다.
- [0124] 따라서, 본 발명은, 클라이언트 디바이스에 의해 다운로드되는 부분들 또는 타일들의 선택된 세트가 무엇이든, 상이한 부분들, 특히 상이한 타일들의 조합, 및 분석시 유효 파일 포맷 및 기본 스트림의 생성을 가능하게 한다.
- [0125] 따라서, 본 발명의 방법은, 독립형 부분들 또는 타일들의 효율적인 스트리밍에 적합하고, 유용한 데이터만이 클라이언트 디바이스에 보내어질 것을 요구하고, 하나보다 많은 부분 또는 타일 선택(즉, 임의의 ROI)에 적합하고, (바이트 범위 요청들에 비해) 인덱스화 오버헤드를 감소시키며, MPEG 표준에 통합될 수 있다.
- [0126] 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.
- [0127] 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 추출자들 및 선택적으로 초기화 데이터를 포함하는 적어도 하나의 합성 트랙이다.
- [0128] 일 실시예에서, 마이크로프로세서는,
- [0129] 적어도 하나의 합성 트랙의 추출자로부터 트랙 참조를 취득하는 단계;
- [0130] 취득된 트랙 참조에 대응하는 트랙이 수신되었는지 여부를 체크하는 단계; 및
- [0131] 취득된 트랙 참조에 대응하는 트랙이 수신되지 않았으면, 시간 설정형 미디어 데이터 비트 스트림을 생성하도록 대응 추출자를 제거하는 단계를 수행하도록 더욱 구성된다.

- [0132] 일 실시예에서, 마이크로프로세서는, 취득된 트랙 참조에 대응하는 트랙이 타일 타입의 트랙인지 확인하는 단계를 수행하도록 더욱 구성된다.
- [0133] 일 실시예에서, 마이크로프로세서는,
- [0134] 적어도 하나의 합성 트랙의 추출자로부터 트랙 참조를 취득하는 단계;
- [0135] 취득된 트랙 참조에 대응하는 트랙이 수신되었는지 여부를 체크하는 단계; 및
- [0136] 취득된 트랙 참조에 대응하는 트랙이 수신되지 않았으면, 시간 설정형 미디어 데이터 비트 스트림을 생성하도록 대응 추출자를 패딩으로 대체하는 단계를 수행하도록 더욱 구성된다.
- [0137] 일 실시예에서, 마이크로프로세서는, 취득된 트랙 참조에 대응하는 트랙이 타일 타입의 트랙인지 확인하는 단계를 수행하도록 더욱 구성된다.
- [0138] 일 실시예에서, 마이크로프로세서는, 생성된 타일 트랙들 및 생성된 적어도 하나의 합성 트랙으로부터 타일형 시간 설정형 미디어 데이터의 공간 부분의 형성을 가능하게 하는 파라미터들을 포함하는 초기화 세그먼트 파일을 수신하는 단계를 수행하도록 더욱 구성된다.
- [0139] 일 실시예에서, 마이크로프로세서는, 적어도 2개의 타일 트랙들을 나타내는 적어도 2개의 미디어 세그먼트 파일들이 수신되도록 더욱 구성되고, 적어도 하나의 합성 트랙은 동일 세트의 공간 서브샘플들에 속하는 2개의 공간 서브샘플들 중 임의의 것을 디코딩하는데 사용될 데이터의 적어도 하나의 아이템을 포함한다.
- [0140] 일 실시예에서, 타일형 시간 설정형 미디어 데이터는, 시간 설정형 샘플들 각각이 그에 따라 참조 레이어 및 적어도 하나의 중대 레이어를 포함하는, 스케일 가능한 타일형 시간 설정형 미디어 데이터이고, 참조 레이어는 적어도 하나의 참조 공간 서브샘플을 포함하고, 적어도 하나의 중대 레이어는 복수의 중대 공간 서브샘플들을 포함하며, 마이크로프로세서는, 적어도 2개의 트랙들을 나타내는 적어도 2개의 미디어 세그먼트 파일들이 수신되도록 더욱 구성되고, 적어도 2개의 수신된 트랙들 중 하나는 참조 공간 서브샘플들을 포함하는 참조 트랙이고, 적어도 2개의 수신된 트랙들 중 다른 하나는 중대 공간 서브샘플들을 포함하는 중대 타일 트랙이다.
- [0141] 일 실시예에서, 클라이언트 디바이스는 HTTP(HyperText Transfer Protocol)와 호환가능하고, 수신된 미디어 세그먼트 파일들은 International Standard Organization에 의해 정의되는 바와 같은 Base Media File Format 및 Dynamic Adaptive Streaming over HTTP 포맷과 호환가능하다.
- [0142] 본 발명의 제7 양태에 따르면, 위에 설명된 바와 같은 디바이스를 포함하는 비디오 디코더가 제공된다.
- [0143] 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 고체 상태 메모리 등과 같은 저장 매체를 포함할 수 있다. 일시적 캐리어 매체는, 전기 신호, 전자 신호, 광 신호, 음향 신호, 자기 신호 또는 전자기 신호와 같은 신호, 예를 들어, 마이크로파 또는 RF 신호를 포함할 수 있다.

도면의 간단한 설명

- [0144] 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 (a) 및 (b)를 포함하고, 인코딩된 비디오 비트 스트림에서 코딩 타일들의 일례를 설명한다.
- 도 2는 디스플레이되도록 사용자에 의해 선택된 타일들의 시간적인 파이프를 설명한다.
- 도 3은 특정한 실시예에 따라 스케일 가능한 미디어 데이터 트랙들을 이용하는 초기화 세그먼트 파일에 대한 블록도들의 일례를 설명한다.
- 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 설명한다.
- 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 미

디어 데이터 세그먼트들을 연접시키는(concatenate) 일례를 설명한다.

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 (a), (b), 및 (c)를 포함하고, HEVC 비트 스트림에서 타일들 및 슬라이스 세그먼트들의 예들을 설명한다.

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

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

도 12는 (a) 및 (b)를 포함하고, 풀 비디오에서 타일의 위치, 타일의 사이즈, 및 다양한 타일링 구성을 다루기 위해 적용된, 서브-샘플 레벨에서, 타일 트랙이 임의의 아티팩트 없이 디코딩될 수 있다는 표시에 대한 신호화를 설명한다.

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 설명한다.

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 설명한다.

발명을 실시하기 위한 구체적인 내용

[0145]

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 레이어 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 타입이 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 디코딩하도록 요청된 메타데이터를 송신하는데 사용될 수 있다.

[0146]

도 1은 (a) 및 (b)를 포함하고, 인코딩된 비디오 비트 스트림에서 코딩 타일들의 일례를 도시한다.

[0147]

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 1의 (a)에 도시된 바와 같이, 비디오 프레임(100)은 HD 베이스 레이어(base layer; 102) 및 4K2K 중대 레이어(enhancement layer; 104)를 포함할 수 있다. 계속해서 설명의 목적상, 중대 레이어(104)는 a, b, c 및 d로 표기된 4개의 정규 타일들로 분리될 수 있다. 상이한 형상들의 타일들이 다루어질 수 있다는 것에 주목해야 한다. 마찬가지로, 베이스 레이어(102)는 여러 타일들로 분리될 수 있다. 그러한 경우에, 예를 들어, 베이스 레이어당 하나 및 중대 레이어 당 하나 또는 중대 레이어들 각각에 대해 여러 합성 트랙들이 사용될 수 있다.

[0148]

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.

[0149]

도 1의 (b)는 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 중대 레이어의 NAL 유닛들이 후속하는 베이스 레이어(BL)의 모든 NAL(network abstraction layer) 유닛들을 포함한다. 예를 들어, 제1 비디오 프레임(110)의 베이스 레이어(102-1)의 NAL 유닛들(1BL, 116)에 후속하여 제1 비디오 프레임의 중대 레이어(104-1)의 NAL 유닛들(1common, 1a, 1b, 1c, 1d, 118)이 온다.

[0150] 공간적 타일들을 가진 중대 레이어에 대응하는 비디오 비트 스트림의 파트는 각각의 타일의 NAL 유닛들로 구성된다. 선택적으로, 모든 타일들에 공통이고 타일들 중 임의의 것을 디코딩하도록 요청되는 NAL 유닛들을 또한 포함할 수 있다. 주어진 프레임의 모든 타일들에 공통인 NAL 유닛들은 비디오 비트 스트림의 대응하는 파트에서 어딘가(즉, 비디오 프레임의 타일들의 NAL 유닛들 이전, 사이, 또는 이후)에 위치할 수 있다.

[0151] 도시된 바와 같이, 공간적 타일들 a, b, c 및 d를 포함하는 제1 비디오 프레임(110)의 중대 레이어에 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(1 공통)로 구성된다.

[0152] 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.

[0153] 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 기간 동안 액세스될 때 각각 프레임에 대한 여러 작은 바이트 범위들에 액세스하도록 요청된다. 이는 발생된 요청들의 수의 관점에서 그리고 데이터 오버헤드의 관점에서 HTTP 스트리밍에서는 비효율적이다.

[0154] 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 재조직화될 것이다.

[0155] 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).

[0156] 도 3은 특정한 실시예에 따라 미디어 데이터 트랙들을 사용하는 초기화 세그먼트 파일의 블록도들의 일례를 도시한다.

[0157] 도 3에 도시된 초기화 세그먼트 파일 및 도 4에 도시된 미디어 세그먼트 파일들은 여러 트랙들이 독립적으로 스트리밍되는 동안 ISO BMFF(Base Media File Format)로 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.

[0158] 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.

[0159] 그러나, 타일 트랙이 디코딩 및 재생되기 위해 요청되는 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 유닛들에 관해 언급하기 위해 추출자 대상들(extractor objects)을 이용한다. 또한, 그러한 NAL 유닛들이 존재하면, 모든 타일들에 공통인 NAL 유닛

들을 포함할 수 있다.

- [0160] 특정한 실시예에 따르면, 초기화 세그먼트 파일은 다른 미디어 세그먼트 파일들에서 캡슐화된 시간 설정형 미디어 데이터 비트 스트림들을 정의할 필요가 있는 모든 메타데이터를 송신하는데 사용된다. 도 3에 도시된 바와 같이, 초기화 세그먼트 파일(300)은 파일 타입 박스 "ftyp"(302) 및 영화 박스 "moov"(304)를 포함한다. 파일 타입 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.
- [0161] 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.
- [0162] 트랙 박스(306-1)는 베이스 레이어(track_ID=1)를 나타내고, 4개의 트랙 박스들(306-2 내지 306-5)(트랙 박스 306-3 및 트랙 박스 306-4는 도시되지 않음)은 중대 레이어(track_ID=2 내지 5)의 4개의 파일들 a, b, c 및 d를 나타내고, 트랙 박스 306-6은 중대 레이어를 기술하는 합성 트랙을 나타낸다(트랙_ID=6).
- [0163] 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.
- [0164] 다른 박스들은 시간 설정형 미디어 데이터 비트 스트림을 캡슐화하는데 사용된 ISO BMFF 사양들에 따라 의무적이거나 선택적일 수 있다는 점에 주목해야 한다. 그러나, 본 발명의 실시예들은 이러한 박스들이 적용될지에 의지하지 않으므로, 그들은 여기에서 제시되지 않는다.
- [0165] 트랙 헤더 박스 "tkhd"(308)는 트랙의 특성들을 명시한다. 정보의 여러 아이템들 중에, 그것은 트랙의 식별자 (track_ID), 트랙의 지속시간, 및/또는 트랙의 시작적 프레젠테이션 사이즈(즉, 디스플레이 영역의 폭과 높이)를 제공한다. 트랙이 재생 가능한지 여부를 나타내는 플래그 파라미터를 또한 포함한다.
- [0166] 일 실시예에 따르면, 파일 트랙들에 대한 트랙 헤더 플래그의 디폴트 값은 0(track_enabled = 0, track_in_movie = 0, track_in_preview = 0)이고, 이는 파일 트랙들이 클라이언트 디바이스에 의해 구역 재생 및 미리 보기 위해 무시된다는 것을 의미한다. 또 다른 실시예에서, 새로운 트랙 헤더 플래그는 트랙이 파일 트랙이라는 것을 시그널링하기 위해 만들어질 수 있다.
- [0167] 트랙 미디어 박스 "mdia"(310)는 트랙 내에 시간 설정형 미디어 데이터의 파라미터들을 공표하는데 사용된 모든 대상들을 포함하는 컨테이너로서 보여질 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.
- [0168] 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('참조(vide)'와 동일한 핸들러 타입 속성을 갖는다는 것에 유의). 비디오 샘플은 VisualSampleEntry() 타입의 대상을 이용하여 기술될 수 있다. 특정한 실시예에 따르면, 파일 핸들러('파일'과 동일한 핸들러 타입 속성을 갖는다는 것에 유의)로 불리는 새로운 핸들러 타입은 트랙이 공간적 서브샘플 정보를 포함한다는 것을 나타내기 위해 정의된다. 코딩 포맷에 따라, VisualSampleEntry() 유형의 대상이 파일 트랙에서 샘플을 기술할 수 없으면, 샘플을 기술하기 위해 TileSampleEntry0 유형의 특정 대상을 정의하는 것이 가능하다.
- [0169] 미디어 정보 박스 "minf"(318)는 트랙에서 시간 설정형 미디어 데이터의 특성 정보를 정의하는 모든 대상들을 포함한다. 예를 들어, 합성 트랙에서 정의된 베이스 레이어 및 중대 레이어에 대해, "minf" 박스는 표준 비디오 미디어 헤더 박스 "vmhd"(320)를 포함할 수 있다.
- [0170] 파일 트랙들과 관련하여, 새로운 파일 핸들러에 대응하는 파일 미디어 헤더 박스(Tile Media Header Box; tmhd, 322)로 언급된 특정 박스는 파일들에 대해 코딩에 독립적인 일반적 프레젠테이션 정보를 정의하는데 사용된다. 특히, 그것은 다음과 같이 정의될 수 있는 합성 비디오 트랙에 의해 표현된 비디오 해상도와 관련하여 파일에 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;
}

```

[0171]

이전에 설명된 바와 같이, 합성 트랙은 프레젠테이션에서 또 다른 트랙에게 타이핑된 참조를 제공하는 특정 트랙 참조 박스 "tref"(312)를 포함한다. 특정한 실시예에 따르면, 그러한 타이핑된 참조는 합성 트랙으로부터 그것이 언급한 타일 트랙으로의 링크를 설치하는데 사용될 수 있는 "tile" 참조(324) 및 이 참조를 포함하는 트랙으로부터 그것이 의존하는 시간 설정형 미디어 데이터 트랙(예컨대, 베이스 레이어 트랙(track_ID=1))으로의 링크를 설치하는데 사용될 수 있는 "scal" 참조(326)를 포함할 수 있다.

[0173]

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.

[0174]

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 프레그먼트 박스 "moof", 및 일반적으로 406으로 참조된 적어도 하나의 미디어 데이터 박스 "mdat"를 포함한다. 미디어 세그먼트 파일은 HTTP-URL과 관련된다.

[0175]

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 이러한 박스들이 적용되는지에 의지하지 않으므로, 그들은 여기에서 제시되지 않는다.

[0176]

그러나, 세그먼트 타입 박스 "styp"(402)의 포맷은 도 3에서 파일 타입 박스 "ftyp"(302) 중 하나와 유사하고, 그것의 참조는 파일이 미디어 세그먼트 파일이라는 것을 나타낸다.

[0177]

영화 프레그먼트 박스(404)는 일반적으로 영화 박스 "moov" 내에 저장되는 정보를 제공한다. 그의 헤더("mfhd")는 각각의 영화 프레그먼트에 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 프레그먼트 박스(404)는 관련된 미디어 데이터 박스("mdat", 406)에 데이터를 갖는 각각의 트랙에 대해 트랙 프레그먼트 박스 "trac"(일반적으로 408로 참조됨)를 포함한다. 트랙 프레그먼트 박스(408)는 대응하는 미디어 데이터 박스(("mdat", 406)에 존재하는 트랙의 비트 스트림의 식별자(track_ID)를 저장하는데 사용되는, 일반적으로 410으로 참조된, 트랙 프레그먼트 헤더 박스 "tfhd"를 포함한다.

[0178]

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 중대 레이어의 타일에 대응하는 모든 NAL 유닛들을 포함한다.

[0179]

합성 트랙(도 4에서 track_ID=6)에서, 미디어 데이터 박스(406)는 각각의 타일에 대해 그리고 각각의 종속형 레이어에 대해 추출자들(도 4에서 E로 표기됨)을 포함하고 (존재 한다면) 모든 타일들에 공통인 NAL 유닛들을 포함한다.

[0180]

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

[0181]

- 베이스 레이어 트랙과 관련된 미디어 세그먼트 파일(400-1)의 미디어 데이터 박스(406)에 저장된 베이스 레이

어 트랙 내에 인코딩된 베이스 레이어 데이터(NAL 유닛들(1BL))로의 링크를 제공하는 제1 추출자(412-1);

[0182] 여러 파일들에 공통인 NAL 유닛들(412-2);

[0183] - 증대 레이어의 제1 파일 트랙과 관련된 미디어 세그먼트 파일(400-2)의 미디어 데이터 박스(406) 내에 인코딩된 제1 파일의 증대 레이어 데이터(NAL 유닛들(1a))로의 링크를 제공하는 제2 추출자(412-3);

[0184] - 증대 레이어의 제2 파일 트랙과 관련된 미디어 세그먼트 파일(400-3)(도시되지 않음)의 미디어 데이터 박스(406) 내에 인코딩된 제2 파일의 증대 레이어 데이터(NAL 유닛들(1b))로의 링크를 제공하는 제3 추출자(412-4);

[0185] - 증대 레이어의 제3 파일 트랙과 관련된 미디어 세그먼트 파일(400-4)(도시되지 않음)의 미디어 데이터 박스(406) 내에 인코딩된 제3 파일의 증대 레이어 데이터(NAL 유닛들(1c))로의 링크를 제공하는 제4 추출자(412-5); 및

[0186] 증대 레이어의 제4 파일 트랙과 관련된 미디어 세그먼트 파일(400-5)(도시되지 않음)의 미디어 데이터 파일(406) 내에 인코딩된 제4 파일의 증대 레이어 데이터(NAL 유닛들(1d))로의 링크를 제공하는 제5 추출자(412-6).

[0187] 추출자(412-1) 덕에 획득될 수 있는 NAL 유닛들은 증대 레이어가 추출자들(412-3 내지 412-6) 덕에 획득될 수 있는 NAL 유닛들(412-2) 및 NAL 유닛들을 사용하여 완전히 디코딩될 수 있는 프레임의 베이스 레이어의 디코딩을 허용한다. 도 4로부터 관찰될 수 있는 바와 같이, 단지 프레임의 공간적 파트만 디코딩된다면, 모든 미디어 세그먼트 파일들(400-2 내지 400-5)(즉, 파일 트랙들에 대응하는 비트 스트림들)을 다운로드할 필요는 없다.

[0188] 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;
}
```

[0189] 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(유형 31)으로 설정된다).

[0190] 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.

[0191] 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 파일에 대한 파일 트랙인 식별자 track_ID=2를 갖고, 제1 인덱스는 참조 트랙(예컨대, 베이스 레이어)을 나타냄).

[0192] 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.

[0193] 도 3 및 도 4를 참조하여 설명된 바와 같이, 시간 설정형 파일형 미디어 데이터 비트 스트림은 바람직하게는 하

나의 초기화 세그먼트 파일 및 다수의 미디어 세그먼트 파일들을 포함하는 데이터 세트로서 송신되는데, 후자는 여러 타일 트랙들 및 하나의 합성 트랙을 포함한다.

[0195]

초기화 세그먼트 파일은 각각의 트랙에 대한 일반적 정보, 특히 트랙의 유형(예컨대, 미디어 트랙(오디오 또는 비디오) 또는 타일 트랙), 코딩 포맷, 프레임 해상도 및 트랙들 간의 의존성(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

[0196]

[0197]

도 5는 요청된 미디어 세그먼트 파일들(여기서 베이스 레이어 및 증대 레이어에서의 타일들 a 및 c에 대응)이 서버로부터 다운로드될 때에만 미디어 세그먼트들을 연접시킴으로써 획득된 파일 포맷을 대략적으로 도시한다. 그러한 메커니즘은 단지 요청된 미디어 세그먼트 파일들만의 다운로딩을 허용할 뿐만 아니라 복제 데이터의 다운로딩을 방지한다는 점이 주목되어야 한다.

[0198]

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

[0199]

획득된 파일 포맷은 스케일가능한 파일 포맷 정의를 따른다. 예를 들어, 클라이언트 디바이스는 베이스 레이어 트랙을 선택함으로써 단지 베이스 레이어만을 재생하도록 결정할 수도 있고 또는 합성 트랙을 선택함으로써 선택된 타일들 a 및 c를 갖는 증대 레이어를 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.

[0200]

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.

- [0201] 제1 단계(600 단계)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, 그것이 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.
- [0202] 다음의 단계(단계 602)에서, 서버는 (존재한다면) 하위 레벨 스케일 가능한 트랙들에 링크된 추출자들, 모든 타일들에 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.
- [0203] 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 매체 세그먼트 파일들에 저장된다.
- [0204] 그 다음, 서버는 요청시에 초기화 및 미디어 세그먼트 파일들을 클라이언트 디바이스에게 제공한다(단계 606). 서버는 HTTP 요청들에 응답하는 종래의 HTTP 서버일 수 있다.
- [0205] HTTP 스트리밍의 정황에서, 바람직한 실시예에서는, 클라이언트 디바이스가 서버로부터 이용 가능한 미디어 프레젠테이션을 기술하는 매니페스트 파일에 액세스하는 것으로 가정된다. 매니페스트 파일은 서버로부터 먼저 초기화 세그먼트들을 요청한 다음 미디어 세그먼트 파일들을 요청함으로써 미디어 프레젠테이션을 스트리밍하기 위해 클라이언트 디바이스에게 충분한 정보(미디어 속성들 및 세그먼트들의 리스트)를 제공한다.
- [0206] 클라이언트 디바이스 종단에서 ROI의 선택시, 전형적으로, 포인팅 디바이스와 같은 선택 수단을 갖는 디스플레이 이상에서는, 타일형 비디오의 스트리밍 동안, 선택된 ROI에 대응하는 타일들이 결정된다(단계 608).
- [0207] 다음으로, 각각 시간적 기간 동안, 스케일 가능한 미디어 데이터의 경우에, 클라이언트 디바이스는 종속형 레이어들에 대응하는 세그먼트 파일들을 다운로드하기 위해 서버로 요청을 보낸다(단계 610). 특정한 실시예에 따르면, 종속되는(dependent from) 레이어들은 종속되는 레이어들에 종속하는(depending from) 레이어들 이전에 다운로드된다. 예를 들어, 베이스 레이어 세그먼트 파일들은 중대 레이어 세그먼트 파일들 이전에 다운로드된다.
- [0208] 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).
- [0209] 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).
- [0210] 단계 616은 도 7과 관련하여 상세히 설명된다.
- [0211] 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.
- [0212] 제1 단계(단계 700)에서, 클라이언트 디바이스는 (예를 들어, 도 6의 단계들(612, 614 및 616)에서) 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.
- [0213] 적어도 하나의 미디어 세그먼트 파일이 수신되었다면, 그리고 수신된 미디어 세그먼트 파일들이 합성 트랙 (composite track)을 포함하지 않는다면(즉, 그들이 종속형 트랙들, 예를 들어, 하위 레벨 레이어 또는 타일 트랙들을 포함함), 그들은 추후의 사용을 위해 버퍼링된다.

- [0214] 반대로, 적어도 하나의 미디어 세그먼트가 수신되었다면, 그리고 수신된 미디어 세그먼트 파일이 합성 트랙을 포함한다면, 수신된 미디어 세그먼트의 미디어 데이터 박스 "mdat"으로부터 데이터의 제1 아이템(또는 수신된 미디어 세그먼트의 데이터의 적어도 하나의 아이템이 프로세싱된 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 아이템이 없는 경우, 프로세스는 단계(700)로 복귀하여 새로운 미디어 세그먼트 파일들을 수신한다(단계 706).
- [0215] 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, 비디오 디코더에 의해 더 디코딩되도록 복귀된다(단계 710). 반대로, 데이터의 추출된 아이템이 추출자인 경우, 이것은 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).
- [0216] 참조된 트랙의 식별자가 식별되면, 참조된 트랙이 단계(700) 동안에 베퍼링된 미디어 세그먼트 파일들의 세트에서 이용가능한지 여부를 판정하기 위해 테스트가 수행된다(단계 714). 클라이언트 디바이스는 선택된 ROI에 대응하는 미디어 세그먼트 파일들만을 다운로드하기 때문에 일부 파일 트랙들이 누락된다는 점을 상기한다.
- [0217] 참조된 트랙이 단계(700) 동안에 베퍼링된 미디어 세그먼트 파일들의 세트에서 이용가능하다면, 추출자는 그가 참조하는 데이터에 의해 대체되고(단계 716), 비트 스트림은 디코딩되기 위해 비디오 디코더에 송신된다(단계 710).
- [0218] 참조된 트랙이 단계(700) 동안에 베퍼링된 미디어 세그먼트 파일들의 세트에서 이용가능하지 않다면, 추출자에서 참조된 데이터의 부재가 ISO BMF 표준에 따른 치명적 오류를 야기하기 때문에 특정 단계들이 수행되어야 한다. 참조된 트랙이 파일 트랙인지 여부(참조된 트랙은 종속형 스케일 가능성 레이어(dependent scalability layer)에 대응할 수 있음) 및 추출자가 파일 타입인지 여부를 판정하기 위해 테스트가 수행된다(단계 718).
- [0219] 참조된 트랙이 파일 트랙이 아니거나 추출자가 파일 타입이 아니라면, 표준적인 치명적 오류가 검출된다. 반대로, 참조된 트랙이 파일 트랙이고 추출자가 파일 타입이라면, 추출자는 제거되거나(단계 722), 추출자는, 시간 설정형 미디어 데이터 비트 스트림(timed media data bit-stream)을 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 레이어에 속하거나 다른 스케일 가능성 레이어에 속하는 이전에 디코딩된 이미지로부터 획득된 다른 픽셀 데이터에 의해 대체되는, 현재의 이미지 내에서 누락되는 픽셀 데이터를 나타낸다. '스킵된' 데이터는 일반적으로 적어도 하나의 플래그(flag)에 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.
- [0220] 다음으로, 비트 스트림이 비디오 디코더에 송신되어 디코딩되고(단계 710) 디스플레이되며, 데이터의 후속 아이템을 핸들링하기 위해 단계(704)로 루프된다(loop).
- [0221] 도 8은 하나 이상의 실시 예들의 단계들이 구현되는 서버 또는 클라이언트 디바이스(800)의 블록도를 나타낸다.
- [0222] 바람직하게는, 디바이스(800)는 통신 버스(802), 디바이스의 전원 인가 시에 프로그램 ROM(806)으로부터의 명령어들을 실행할 수 있는 중앙 처리 유닛(CPU)(804), 및 전원 인가 후에 메인 메모리(808)로부터의 소프트웨어 애플리케이션에 관련된 명령어들을 포함한다. 메인 메모리(808)는 예를 들어, 통신 버스(802)를 통해 CPU(804)의 작업 영역으로서 기능하는 RAM(Random Access Memory) 유형이고, 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.
- [0223] 참조 번호(812)는 통신 네트워크(814)에 대한 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.

- [0224] 참조 번호(816)은 사용자에게 정보를 디스플레이 및/또는 사용자로부터 입력을 수신하기 위한 사용자 인터페이스들을 나타낸다.
- [0225] 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 통합 회로들(ASIC)로 이루어질 수 있다는 것을 유의해야 한다. 이러한 통합 회로들은, 예를 들면 그리고 비-제한적으로는, 비디오 시퀀스들을 생성 또는 디스플레이 및/또는 오디오 시퀀스들을 청취하기 위한 장비 내로 통합된다.
- [0226] 전술한 바와 같이, 본 발명의 실시예는, 특히, HEVC로서 알려진 비디오 포맷에 적용할 수 있다.
- [0227] HEVC 표준에 따르면, 이미지들은 타일들, 슬라이스들, 및 슬라이스 세그먼트들로 공간적으로 분할될 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 CTU(코딩 트리 유닛들(*Coding Tree Units*))을 포함한다. 따라서, 타일들은, 예를 들어, ROI에 대한 위치 및 크기를 정의함으로써 ROI를 식별하는데 효율적으로 사용될 수 있다. 그러나, NAL(Network Abstract Layer) 유닛들로서의 HEVC 비트 스트림의 캡슐화뿐 아니라 그의 구조는 타일의 관점에서 조직화되지 않고 슬라이스들에 기반한다.
- [0228] HEVC 표준에서, 슬라이스들은 슬라이스 세그먼트들의 세트들이고, 슬라이스 세그먼트들의 세트의 제1 슬라이스 세그먼트는 독립형 슬라이스 세그먼트이며, 즉, 헤더 내에 저장된 일반 정보가 다른 슬라이스 세그먼트 중 하나를 참조하지 않는 슬라이스 세그먼트이다. 슬라이스 세그먼트들의 세트의 다른 슬라이스 세그먼트들은, 만약에 있다면, 종속형 슬라이스 세그먼트이다(즉, 헤더 내에 저장된 일반 정보가 독립형 슬라이스 세그먼트 중 하나를 참조하는 슬라이스 세그먼트들).
- [0229] 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, ROI를 나타내기에 적절하지 않다. 이것은 슬라이스 세그먼트 데이터가 후속하는 슬라이스 세그먼트 헤더의 형식 하에서 HEVC 비트 스트림으로 인코딩된다. 독립형 및 종속형 슬라이스 세그먼트들은 그들의 헤더에서 차이가 있다: 종속형 슬라이스 세그먼트는 독립형 슬라이스 세그먼트에 종속하고, 그의 헤더의 정보량은 독립형 슬라이스 세그먼트의 것보다 작다. 독립형 및 종속형 슬라이스 세그먼트들은 타일들을 정의하는데 사용되거나 엔트로피 디코딩 동기화 포인트들(*entropy decoding synchronization points*)로서 사용되는 대응하는 비트 스트림에서의 엔트리 포인트들의 리스트를 포함한다.
- [0230] 도 9a, 9b 및 9c를 포함하는 도 9는 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 분할된 이미지(900)를 예시한다. 915-1 내지 915-9로 참조된 9개의 부분들은 특정한 타일을 나타낸다.
- [0231] 도 9b는 수직 경계(905')에 의해 구분된 2개의 수직 타일들을 포함하는 이미지(900')를 예시한다. 이미지(900')는 1개의 독립형 슬라이스 세그먼트(920-1)(빗금(hatched line)으로 나타냄)와 4개의 종속형 슬라이스 세그먼트들(920-2 내지 920-5)의 5개의 슬라이스 세그먼트들을 포함하는 싱글 슬라이스(참조되지 않음)를 포함한다.
- [0232] 도 9c는 수직 경계(905")에 의해 구분된 2개의 수직 타일들을 포함하는 이미지(900")를 예시한다. 좌측 타일은 2개의 슬라이스들을 포함한다: 제1 슬라이스는 1개의 독립형 슬라이스 세그먼트(920'-1) 및 1개의 종속형 슬라이스 세그먼트(920'-2)를 포함하고 제2 슬라이스는 또한 1개의 독립형 슬라이스 세그먼트(920'-3) 및 1개의 종속형 슬라이스 세그먼트(920'-4)를 포함한다. 우측 타일은 1개의 독립형 슬라이스 세그먼트(920'-5) 및 1개의 종속형 슬라이스 세그먼트(920'-6)을 포함하는 하나의 슬라이스를 포함한다.
- [0233] HEVC 표준에 따라, 슬라이스 세그먼트들은 이하와 같이 요약될 수 있는 규칙들에 따라 타일들에 링크된다(하나 또는 양쪽 조건들이 만족되어야 함):
- 슬라이스 세그먼트 내의 모든 CTU들은 동일한 타일에 속한다(즉, 슬라이스 세그먼트는 여러 타일들에 속할 수 없다); 및
 - 타일 내의 모든 CTU들은 동일한 슬라이스 세그먼트에 속한다(즉, 타일은 이들 슬라이스 세그먼트들 각각이 오직 그 타일에만 속하는 경우 여러 슬라이스 세그먼트들로 분할 될 수 있다).
- [0234] 명료함을 위해, 이하에서는 1개의 타일이 1개의 독립형 슬라이스 세그먼트만을 갖는 1개의 슬라이스를 포함하는 것으로 간주된다. 그러나, 본 발명의 실시예들은 도 9b 및 9c에 예시된 것과 같은 다른 구성들로 수행될 수 있

다.

[0237] 전술한 바와 같이, 타일들이 ROI에 대한 적절한 지원으로서 간주될 수 있는 한편, 슬라이스 세그먼트들은 통신 네트워크를 통한 송신을 위해 NAL 유닛들 내에 실제로 삽입되고 액세스 유닛들(즉, 타일 포맷 레벨에서의 코딩된 픽처(coded picture) 또는 샘플들)을 형성하기 위해 집성되는 엔티티들이다.

[0238] HEVC 표준에 따라서, NAL 유닛의 유형은 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
}
```

[0239] [0240] 슬라이스 세그먼트들을 코딩하는데 사용된 NAL 유닛들은 슬라이스 세그먼트 어드레스 구문 요소로 인하여 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){
        [...]
    }
}
```

[0241] [0242] 타일링 정보가 PPS(Picture Parameter Set) NAL 유닛에 제공된다. 그리고 나서, 슬라이스 세그먼트와 타일 간의 관계는 이러한 파라미터들로부터 추론될 수 있다.

[0243] 공간적 예측들이 (정의에 의한) 타일 경계들 상에서 리셋되지만, 타일이 참조 프레임(들) 내의 상이한 타일로부터의 시간적 예측들을 이용하는 것을 막지 못한다. 따라서, 독립형 타일(들)을 구축하기 위해, 인코딩 동안에, 예측 유닛들을 위한 모션 벡터들이 타일 내부에 제약되어 참조 프레임(들) 내에 공동 위치된 타일을 유지한다는 장점이 있다. 또한, 인-루프(in-loop) 필터들(디블로킹(deblocking) 및 샘플 적응형 오프셋 (SAO)(sample adaptive offset) 필터들)은 바람직하게는 타일 경계들 상에서 비활성화되어(deactivated) 오직 하나의 타일을 디코딩하는 경우에 어떤 에러 드리프트(error drift)도 도입되지 않는다. 그러한 인 루프 필터들의 제어가 HEVC 표준에서 이용가능하다는 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 들어오는 픽셀들에 종속될 수 없다. 모션 벡터들 및 인 루프 필터들과 관련된 이러한 2개의 조건들이 만족되는 경우, 타일들은 "독립적으로 디코딩가능한 타일들" 또는 "독립형 타일들"로서 간주될 수 있다.

- [0244] 비디오 비트 스트림이 독립형 파일들의 세트로서 인코딩되는 경우, 이것은 참조 데이터의 누락 또는 재구성 어려들의 전파에 대한 임의의 위험 없이 하나의 프레임으로부터 다른 프레임으로의 파일 기반 디코딩을 가능하게 한다. 그리고 나서, 이러한 구성은, 예를 들어, (파일들 3 및 7을 포함하는) 도 2에 예시된 ROI에 대응할 수 있는 원래 비디오의 공간적 부분만을 재구성하는 것을 가능하게 한다. 그러한 구성은 파일 기반 디코딩이 신뢰할 수 있다는 것을 나타내기 위해 비디오 비트 스트림 내의 보충 정보로서 나타내어질 수 있다.
- [0245] 본 발명의 실시예에 따르면, HTTP 스트리밍의 정황에서의 파일들에 대한 효율적 액세스가 HEVC 표준에 적용된 ISO BMFF 파일 포맷을 이용하여 제공된다. 따라서, 코딩될 독립형 파일들의 각각(예를 들어, 도 2에 나타낸 12개의 파일들의 각각)은 도 10을 참조하여 이하에 설명되는 "파일 트랙"으로 지칭되는 특정 트랙에 의해 나타내어진다.
- [0246] 이러한 파일 트랙들은 도 10에 예시된 바와 같은 풀 프레임(full-frame) 내의 HEVC 비트 스트림에 대응하는 합성 트랙으로 (트랙들 각각에 대한 정의를 포함하는 영화 박스(movie box) 'moov'의 트랙 참조 박스(track reference box) '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.
- [0247] 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;
}
```
- [0248]
- [0249] 추출자는 다른 트랙들로부터의 데이터에 대한 포인터 또는 참조들로서 기능하며 양 트랙들에서의 데이터 복제 대신에 종속형 트랙들을 참조하여 컴팩트 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\_index로 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\_length로 표기됨)은 각각 복사할 곳으로부터의 위치(바람직하게는 바이트로) 및 복사할 데이터량을 제공한다(값 0은 전체 참조된 NAL 유닛의 복사를 나타내기 위해 예약된다).
- [0250] 도 10은 본 발명의 실시예에 따라, 합성 트랙 및 독립형 파일 트랙들을 포함하는 트랙들의 세트로서 HEVC 비트 스트림을 캡슐화하는 예를 예시한다. 예시를 위해, 캡슐화된 비트 스트림은 도 2에 개략적으로 예시된 비디오 시퀀스에 대응한다.
- [0251] 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.

- [0252] 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 프레그먼트 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.
- [0253] 마찬가지로, 타일 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 프레그먼트 박스 "moof", 및 NAL 유닛들(NALUs)로 패키징된 압축된 비디오 데이터를 포함하는 적어도 하나의 미디어 데이터 박스 "mdat"를 포함한다.
- [0254] 식별자(2 내지 13)를 갖는 타일 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.
- [0255] 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.
- [0256] 또한 예시를 위해, 합성 트랙(1015)의 NAL 유닛(1035-1)이 프로세싱되는 경우, 이것은 추출자 타입의 NAL 유닛을 나타낸다고 판정된다(NALUnitHeader는 16진수 값 5E00과 같다). 따라서, 이것은 대응하는 압축된 비디오 데이터를 복구하기 위해 프로세싱된다. 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)).
- [0257] 프로세싱된 후에, 추출자는 그것이 참조하는 데이터에 의해 대체된다. 도 10에서 예시된 예에 따라, 추출자(1035-1)의 파싱 및 프로세싱은 코딩된 슬라이스 세그먼트 NALU(1040)에 의한 그의 대체를 야기하고, 따라서, HEVC 준수 비트 스트림을 형성한다.
- [0258] HEVC 추출자의 파라미터들을 저장하는 것에 사용된 세만틱(semantics)은 SVC 표준에 정의된 것에 가깝다는 것을 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;
  - *nal\_unit\_type*로서 알려진 파라미터는 47(현재 FDIS에서의 예약된 코드)로 설정된다;
  - *nuh\_layer\_id* 및 *nuh\_temporal\_id\_plus1*로서 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
  - *sample\_offset*으로 알려진 파라미터는 0으로 설정된다.
- [0263] 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.
- [0264] 전술한 것과 유사한, 도 11에 예시된 특정 실시예에 따라, 시그널링이 트랙 레벨에서 수행될 수 있다.
- [0265] 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.
- [0266] 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 독립

적으로 스트리밍된다.

[0267] 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 이용가능한 모든 트랙들을 제공한다.

[0268] 영화 박스(110)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.

[0269] 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'가 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.

[0270] 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 이러한 박스들이 적용되는지에 의존하지 않기 때문에, 그들은 여기서 제시되지 않는다.

[0271] 도 11을 참조하여 설명된 실시예에 따르면, 완전한 비디오에서의 타일의 위치, 타일의 크기, 및 타일 트랙이 임의의 결합없이 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.

[0272] 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.

[0273] 고려된 타일의 시작적 프레젠테이션의 크기는 기준의 트랙 헤더의 폭 및 높이 파라미터들에 입력된다. 요구되는 경우, 고려된 파일의 실제 픽셀 치수들이 (*SampleDescriptionBox*로서 알려진 박스 내의 *VisualSampleEntry()*로서 알려진 파라미터를 통해) 샘플 설명들로 문서화될 수 있다

[0274] 마지막으로, 특정 트랙(타일 트랙)을 시그널링하는 것은 핸들러 박스 'hdlr'(1145) 내에 새로운 타입의 트랙 ('tile')을 정의함으로써 명시적으로 행해진다.

[0275] 초기화 세그먼트 파일(1100)의 영화 박스 "moov"(1110)는 'mvex' 박스(1150)를 더 포함한다. 이 박스는 캡슐화된 파일을 액세스하는 클라이언트에게 영화 프로그먼트들이 존재한다는 것을 통지하는데 사용된다. 이것은 프레젠테이션에서 가장 긴 트랙의 지속기간을 초기화 세그먼트 파일 내에 특정하는 것을 가능하게 한다. 이것은 각각의 영화 프로그먼트 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 조각들에 공통인, 예를 들어, 트랙 식별자들 및 트랙에서의 샘플들의 디폴트 크기의 정보의 복제를 회피하기 위해 트랙마다 트랙 화장 박스를 포함한다.

[0276] 타일 트랙들을 시그널링하는 그러한 실시예는 통상의 비디오 트랙 시그널링에 비교하여 현저한 오버헤드를 도입하지 않는다. 또한, 이것은 프레젠테이션이 분할되는 독립적 영화 프로그먼트들의 수를 유지한다. 그러나, 이것은 비디오 시퀀스에 따른 변화하는 타일링 구성으로 구성되지 않는다.

[0277] 비디오 시퀀스를 따른 타일링 구성에 있어서의 변화를 핸들링하도록 구성되는 다른 특정 실시예에 따르면, 시그널링은, ISO BMFF 표준으로부터의 샘플 그룹핑 메커니즘을 이용하여, 샘플 레벨에서 행해진다.

[0278] 그러한 샘플 그룹핑 메커니즘들은 트랙들에서의 샘플들의 분할들을 나타내기 위해 사용된다. 그들은, 샘플 그룹들에 대한 샘플들의 할당을 기술하는 *SampleToGroup* 박스('sbgp') 및 특정 샘플 그룹 내의 샘플들의 공통 특성들을 기술하는 *SampleGroupDescription* 박스('sgpd')의 2개의 박스들의 사용에 의존한다. 특정 타입의 샘플 그룹핑이 유형 필드('grouping\_type')를 통해 하나의 *SampleToGroup* 박스와 하나의 *SampleGroupDescription* 박스의 조합에 의해 정의된다. 다수의 샘플 그룹핑 인스턴스들(즉, *SampleToGroup* 및 *SampleGroupDescription*

박스들의 쌍)은 상이한 그룹핑 범주에 기초하여 존재할 수 있다.

[0279] 본 발명의 실시예들에 따라, 샘플들의 타일링에 관련된 새로운 그룹핑 범주가 정의된다. 'tile'로 지칭되는 이러한 새로운 grouping\_type은 타일의 특성을 기술하고 표준 *VisualSampleGroupEntry*로부터 도출된다. 이것은 *TileRegionSampleGroupEntry* 또는 *HEVCSpatialEntry*로서 지칭될 수 있고, 다음과 같이 정의된다:

```
class HEVCSpatialEntry() extends VisualSampleGroupEntry ('trsg') {
 unsigned int(32) tileID;
 unsigned int(16) horizontal_offset;
 unsigned int(16) vertical_offset;
 unsigned int(16) region_width;
 unsigned int(16) region_height;
 unsigned int(2) independent;
 unsigned int(6) reserved=0;
}
```

[0280] 이러한 새로운 타입의 그룹 엔트리에 따르면, tileID 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, 독립형 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\_flag로서 알려진 플래그는 independent 플래그의 값을 설정하는데 이용될 수 있다. 후자의 의미는 다음과 같이 설정될 수 있다:

[0282] - independent 파라미터가 0과 동등한 경우에는, 이 타일과 동일한 프레임에서의 또는 이전의 프레임들에서의 다른 타일들 사이의 코딩 종속성들은 알려지지 않고;

[0283] - independent 파라미터가 1과 동등한 경우에는, 이 타일과 동일한 프레임에서의 다른 타일들 사이에는 공간적인 코딩 종속성이 존재하지 않지만, 이 타일과 이전의 프레임들에서 동일한 tileID를 갖는 타일 사이에는 코딩 종속성이 존재할 수 있고;

[0284] - independent 파라미터가 2와 동등한 경우에는, 이 타일과 동일한 프레임에서 또는 이전의 프레임들에서 동일한 tileID를 갖는 다른 타일들 사이에는 코딩 종속성이 존재하지 않고;

[0285] independent 파라미터 값 3은 예약된다.

[0286] 옵션으로, 타일당 평균 비트레이트를 설명하는 파라미터는 대역폭에 기초한 적응을 위해 스트리밍 클라이언트에 제공되기 위해서 타일 설명자에 설정될 수 있다.

[0287] 이 실시예에 따르면, 각각의 타일의 특성들은, 각각의 타일 트랙에 대해, '타일' grouping\_type 및 HEVCSpatialEntry를 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 프레그먼트의 각각의 샘플과 그것의 특성들을 관련시키기 위해서 각각의 타일 트랙 프레그먼트에서 SampleToGroup 박스가 정의된다.

[0288] 타일들의 그리드가 시간에 따라 변경되는 경우에, 새로운 HEVCSpatialEntry를 갖는 새로운 SampleGroupDescription 박스('sgpd')가 트랙 프레그먼트 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 프레그먼트마다 정의된다. 이 박스는, 설명의 관점에서, 적어도 28 바이트를 나타낸다. 2초의 지속시간의 프레그먼트들을 갖는 16개의 타일을 가정하면, 이것은, 단지 SampleToGroup 박스에 대해, 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.

[0289] 샘플 그룹 박스들은 개별 샘플들에 메타데이터를 추가하기 위해 효율적이며 확장가능한 틀을 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.

[0290] 이것은 SampleToGroup 박스를 이용하여 그룹의 각각의 샘플을 그에 따라 플래깅함으로써 표시될 수 있다.

[0291] 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.

[0292] 그 때문에, 그룹 설명 타입의 제2 버전(버전=2)이 SampleGroupDescriptionBox로서 알려진 박스에 이용될 수 있는데(traf/stbl 박스당 다수의 SampleGroupDescriptionBox가 존재할 수 있음), 이는 (grouping\_type로서 알려진 파라미터를 통해) 현재의 트랙에서의 또는 현재의 트랙 프레그먼트들에서의 모든 샘플들에 대해 참조된 샘플 그룹이 적용된다는 것을 나타낸다.

[0293] 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;
 }
 }
}
```

[0294]

[0295] 이러한 확장된 타입의 SampleGroupDescription 박스에 따르면, 버전이 2와 동등할 때, 이 그룹 설명 및 임의의 SampleGroupEntry는 현재의 트랙 또는 트랙 프레그먼트의 모든 샘플들에 적용된다. grouping\_type 파라미터는, 이 샘플 그룹 설명과 관련되는 SampleToGroup 박스를 식별하는 정수이다. 트랙 또는 트랙 프레그먼트가 그것의 초기화 데이터에서 버전=2를 갖는 SampleGroupDescription 박스를 포함할 때, 이를 초기화 데이터에 대응하는 grouping\_type을 설명하는 SampleToGroup 박스를 둘 필요가 없다(그 이유는 설명이 정의에 의해 모든 샘플들에 적용되기 때문임)는 점에 유의해야 한다.

- [0296] 새로운 버전의 샘플 그룹 설명 박스의 이러한 정의에서, entry\_count는 다음의 테이블에서의 엔트리들의 개수를 제공하는 정수이고, default\_length는, (길이가 일정한 경우에) 모든 그룹 엔트리의 길이를 나타내거나 또는 그것이 변수인 경우에 0을 나타내고, description\_length는, 엔트리마다 변하고 그에 따라 default\_length가 0인 경우에, 개별 그룹 엔트리의 길이를 나타낸다.
- [0297] 이 실시예에 따르면, grouping\_type은 공간적/타일 그룹화를 나타내는 특정 값을 갖는데: 그것은 예를 들어 '타일'에 대한 ASCII 코드에 대응하는 16진 값일 수 있다(0x74696C65). grouping\_type에 대해 동일한 값을 갖는 이 박스의 최대 하나의 출현이 트랙에 대해 존재할 것이다.
- [0298] 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/mnif/stbl 박스에서 시그널링되는 디폴트 구성으로부터 변경되는 타일링 구성은 타일 트랙들의 트랙 프레그먼트들에서 샘플 그룹 설명 박스만이 업데이트될 필요가 있다. 이것은 적응형 타일들에 대한 신호화 비용을 감소시킨다.
- [0299] 대안적으로, 그리고 여전히 타일 트랙당 초기화 데이터의 양을 감소시키기 위해서(각각의 타일 트랙 프레그먼트에서의 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 박스로부터 단지 SampleTable 박스('stbl')에 포함되도록 정의된다. 이러한 새로운 박스는 트랙 내의 모든 샘플들에 적용될 샘플 그룹 설명들의 세트를 모든 샘플들에 관련시킬 것이다.
- [0300] 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;
    }
}
```
- [0301]
- [0302] 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 공간적/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 하나 내지 SampleGroupDescription 박스에서의 샘플 그룹 엔트리들의 개수의 범위이거나 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.
- [0303] 이것은, 트랙으로부터의 모든 샘플들이, 타일 그룹화만이 이용 중인 경우에(entry_count=1) 영화 프레그먼트들의 개수가 무엇이든지 타일당 최대 32 바이트를 이용하여, 주어진 그룹화 타입에 대해 동일한 그룹 설명을 따른다는 것을 시그널링할 수 있게 한다. 시간에 따라 움직이는 적응형 그리드의 경우에, 새로운 DefaultSampleToGroups 박스 및 새로운 SampleGroupDescription 박스는 트랙 프레그먼트들에서 정의될 수 있다. 새로운 DefaultSampleToGroups 박스는 이전의 정의를 대체할 것이며, 새로운 SampleGroupDescription 박스에서의 새로운 타일 설명을 참조한다. 따라서, SampleToGroup 박스는 타일 그리드 정의가 변경될 때에만 각각의 트랙 프레그먼트에 대해 정의되지 않는다.
- [0304] 여전히 비디오 시퀀스를 따라 타일링 구성에서의 변동을 핸들링하도록 되어 있는 특정 실시예에 따르면, 새로운 샘플 맵 박스를 갖는 서브샘플 레벨에서 시그널링이 행해진다.

- [0305] 타일링 구성에 관한 샘플 설명에 대한 이 실시예는 단일 트랙으로 캡슐화되는 타일형 HEVC 비트 스트림(single track encapsulated tiled HEVC bit-stream)에 적용된다. 그것은, MPEG-4 표준에 따르는 단일 비디오 트랙으로 다운로드된 공간적 부분들을 세이브하기 위해서 합성 트랙의 추출자들의 분해(resolution) 뒤에, 클라이언트에서의 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.
- [0306] 도 12의 (a) 및 도 12의 (b)를 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.
- [0307] 도 12의 (a)는 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.
- [0308] 이들 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).
- [0309] 대응하는 타일 트랙들뿐만 아니라 합성 트랙이 클라이언트 디바이스에 의해 다운로드되거나 판독된다(단계 1215 및 단계 1220). 다음에, 단일 비디오 트랙을 획득하기 위해서 타일 트랙들을 이용하여 합성 트랙들의 추출자들이 분해된다(resolved)(단계 1225). 마지막으로, 클라이언트 디바이스는, 획득된 비디오 트랙에서, 예를 들어 SampleTableBox에서의 타일링 설명을 구축하거나 추가한다(단계 1230).
- [0310] 도 12의 (b)에는 타일링 설명의 예가 도시되어 있다. 도시된 바와 같이, 타일링 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), 타일들과 각각의 샘플의 NAL 유닛들 사이의 맵핑을 설명하는 샘플 그룹 설명 박스(1270), 및 타일 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.
- [0311] TileNALUMapEntry 그룹 엔트리(1270)는 샘플의 NAL 유닛들과 타일들 사이의 맵핑을 정의한다(그 이유는 이러한 실시예가 서브샘플 레벨 시그널링을 참조하기 때문임). grouping_type 파라미터가 'tsgm'과 동등한 이러한 박스는 샘플당 NAL 유닛들의 개수를 포함한다.
- [0312] TileNALUMapEntry 박스는 (도 12의 (b)에 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;
}
```
- [0313] 여기서, entry\_count는 트랙 샘플에서의 NALU들의 개수를 나타내며, tileID는 현재의 트랙에 의해 설명된 공간적 타일에 대한 고유 식별자를 제공한다.

[0315]

또한, TileNALUMapEntry 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;
 }
}
```

[0316]

[0317] 여기서, large\_size 파라미터는 트랙 샘플들에서의 NAL 유닛 엔트리들의 개수가 8 비트로 표현되는지 또는 16 비트로 표현되는지를 나타내고, mode 파라미터는, (mode가 설정되지 않을 때에는) 각각의 NAL 유닛 엔트리가 설명되는지 또는 (mode가 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.

[0318]

[0318] 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.

[0319]

[0319] 주어진 예는 하나의 NAL 유닛이 시퀀스 지속시간을 따라 이것 및 하나의 타일에 대한 데이터를 포함하는 특정 경우라는 점에 유의해야 한다. 타일 데이터가 수개의 NAL 유닛에 걸쳐 분리될 때, 타일에 대응하는 NAL 유닛들의 세트를 설명하기 위해 수개의 어그리게이터가 이용된다. 시간에 따라 NAL 유닛들의 개수가 변할 때, 수개의 tileSampleMap 엔트리가 정의될 수 있고, 대안적으로 그룹으로의 샘플들은 트랙 프레그먼트 헤더에서 프레그먼트마다 그룹화 태입을 통해 적절한 타일 맵 엔트리를 참조한다.

[0320]

[0320] 도 10을 참조하여 설명된 바와 같은 데이터 캡슐화 스킴에서, HEVC 비트 스트림은 실제로 압축된 비디오 데이터를 포함하는 타일 트랙들(1020-1 내지 1020-2)에 대해 포인팅하는 합성 트랙(1015)으로서 캡슐화된다. 합성 트랙은 (도 10에서 PS로 표시된) 상이한 HEVC 파라미터 세트인 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 트랙들로 캡슐화되는 압축된 비디오 데이터에 대해 포인팅한다.

[0321]

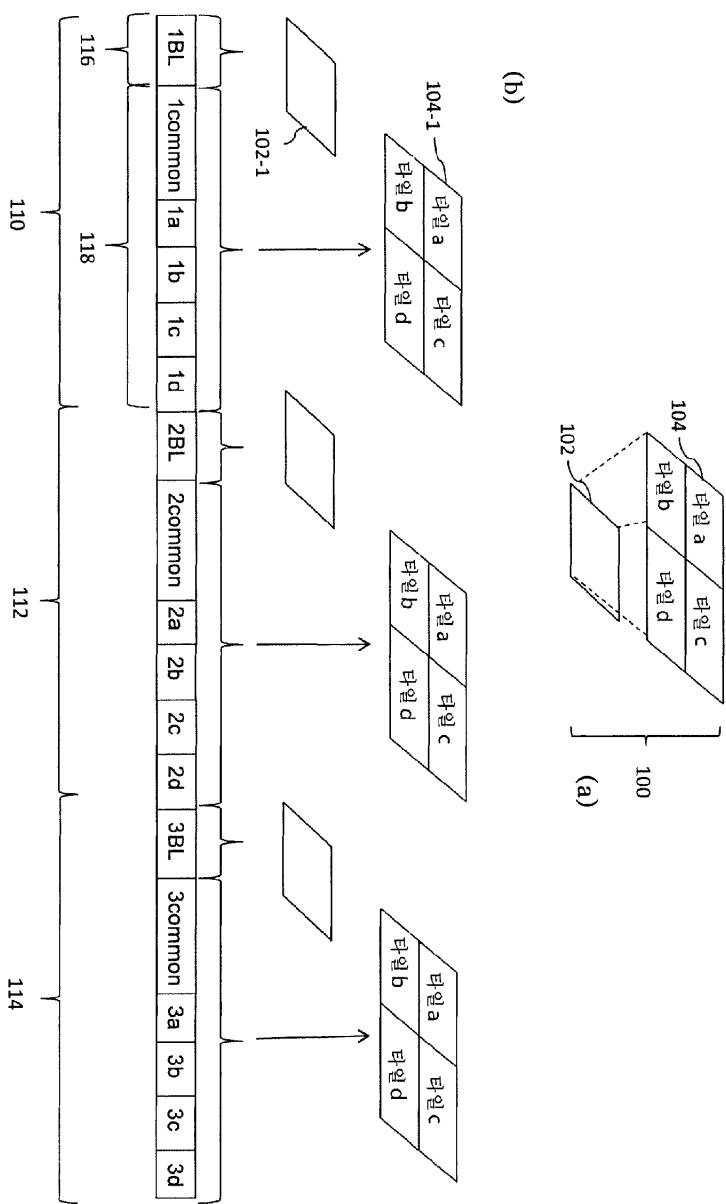
[0321] ISO BMFF 표준(이 표준의 Part 15)에서의 현재의 종속성 시그널링 수단은 초기화 세그먼트 파일(1005)의 moov 박스에서의 트랙 박스들의 일부인 트랙 참조 박스 'tref'에 위치한다. 'tref' 박스는 프레젠테이션에 있어서 포함 트랙으로부터 다른 트랙으로의 참조를 제공한다. 포함 트랙은 프레젠테이션에 있어서 다수의 다른 트랙을 참조할 수 있다. 트랙들 사이의 종속성의 태입은 현재의 표준에서 2개의 값 'scal' 또는 'sbas'를 취할 수 있는 reference\_type 파라미터에 의해 특정된다. 'sbas'는 스케일가능한 베이스(scalable base)를 나타낸다. 그것은, 참조된 트랙이 스케일가능한 프레젠테이션에서 현재의 트랙의 스케일가능한 베이스 트랙임을 나타낸다. 'scal'은 스케일가능성(scalability)을 나타낸다. 그것은 스케일가능한 표현의 상이한 레이어들을 표현하는 트랙들 사이의 관계를 나타낸다. 그것은 포함 트랙이 참조된 트랙에 종속함을 의미한다.

- [0322] 도 10을 참조하여 설명된 실시예에서, 특정 스케일가능성 관련 종속성들은 존재하지 않는다. 스케일가능한 비디오들이 고려될 수 있더라도, 여기서는 합성 트랙과 타일 트랙들 사이의 공간적 종속성들에 대한 포커싱이 이루어진다. 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.
- [0323] 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 레이어 종속성(예를 들어, 'sub1' 값)을 나타낼 수 있다.
- [0324] 따라서, 서브 레이어 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 스케일가능한 HEVC 비트 스트림들에서의 시간적 레이어들에 적용된다. 서브 레이어 트랙에 대응하는 각각의 트랙은, 미리 결정된 값으로 설정될 때에, 이 HEVC 트랙이 서브 레이어 트랙이며 단지 다른 트랙(들)으로부터, 예를 들어 합성 트랙으로부터 참조되는 NAL 유닛들을 포함함(즉, 이 HEVC 트랙은 디스플레이가능하지 않음)을 나타내는 비트(또는 플래그)를 이용하여 HEVCConfiguration 레코드에(즉, SampleTableBox에) 마킹될 수 있다. 이 비트 또는 플래그의 값이 반대 값을 가질 때, 그것은, 이 HEVC 트랙이 초기화 데이터를 또한 포함하는 서브 레이어 트랙임(즉, 이 HEVC 트랙은 디스플레이가능함)을 나타낸다. 예를 들어, 현재의 HEVCDecoderConfigurationRecord 박스에서 예약 비트들을 이용하는 것이 가능하다.
- [0325] 도 12를 참조하여 설명된 특정 실시예에 따르면, 각각의 타일 트랙은 표준 비디오 트랙으로서 플레이가능하다.
- [0326] 도 13은 본 발명의 일 실시예에 따른 표준 비디오 트랙들로서 플레이가능한 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.
- [0327] 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.
- [0328] 도시된 바와 같이, 타일 트랙들(1300-1 내지 1300-12) 각각은, 초기화 및 구성 데이터를 나타내는 합성 트랙(1310)의 HEVC 파라미터 세트인 NAL 유닛들(PS로 표시됨)에 대해 포인팅하는 추출자(1305-1 내지 1305-12)를 포함하는데, HEVC 표준에 따르면, 이를 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 플레이가능하게 한다.
- [0329] 각각의 타일 트랙에 추가된 추출자는 유리하게는 비디오 데이터 샘플들 앞에서 각각의 타일 트랙의 미디어 데이터 박스 "mdat"의 시작에 위치한다.
- [0330] 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 'hbas'로 고려된다(이것은, 스케일가능한 프레젠테이션에서 가장 낮은 동작 포인트를 포함하는 트랙이 '스케일가능한 베이스 트랙'인 'sbas'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 'hbas'를 갖는다.
- [0331] 역시, 타일링 신호화는 트랙 레벨에서, 샘플 레벨에서, 또는 트랙 및 샘플 레벨에서 이루어질 수 있다.
- [0332] 디폴트로 타일 트랙들은 디스플레이가능하지 않은 것으로 고려될 수 있다는 점에 유의해야 한다. 그러나, MPEG-4 표준에 따르는 진보된 파서는 디스플레이가능한 타일 트랙들을 검출하고, 예를 들어 'tref' 박스를 검토함으로써 스트리밍 매니페스트 파일에서 이들을 노출시킬 수 있다(타일 트랙이 'hbas' 타입의 참조 타입을 포함하는 경우, 그것은 디스플레이가능한 것으로 고려될 수 있음). 이것은, 이 타일 트랙이 핸들러 박스에서 '타일' 값으로 마킹되는 경우에도 표준 비디오 트랙으로 고려될 수 있는 것을 의미한다. 타일링 신호화가 샘플 기반일 때, 타일 트랙들 또는 서브 레이어 트랙들은 그들의 핸들러 박스에서 'vide'로서 태깅될 수 있는데, 그 이유는 타일링 정보가 SampleTableBox로서 알려진 박스에 배치되기 때문이다.

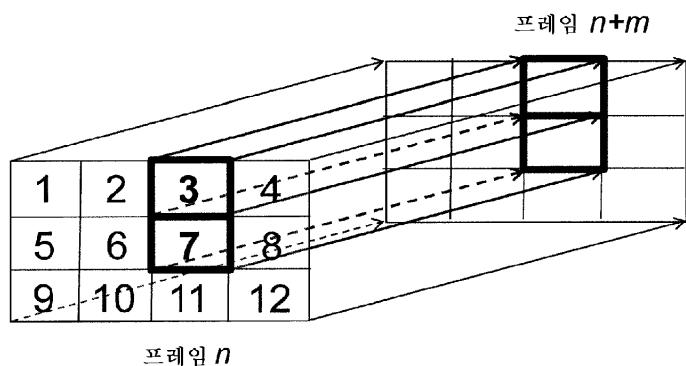
- [0333] 도 14는 본 발명의 다른 실시예에 따른 합성 트랙, 초기화 데이터 트랙, 및 표준 비디오 트랙들로서 플레이가능한 독립적인 타일 트랙들을 포함하는 트랙들의 세트로서 HEVC 비트 스트림을 캡슐화하는 예를 도시한다. 예시를 위해, 캡슐화된 비트 스트림은 도 2에 개략적으로 도시된 비디오 시퀀스에 대응한다.
- [0334] 도 14에 도시된 HEVC 비트 스트림 캡슐화는, 초기화 데이터가 (합성 트랙(1310)에서가 아니라) 전용 초기화 데이터 트랙(1400)에 배치된다는 점에서, 도 13에 도시된 것과는 주로 상이하다.
- [0335] 도 13을 참조하여 설명된 것과 비교하여 볼 때 이러한 실시예에 의해 제공되는 이점들 중 하나는 타일 트랙들이 독립적으로 플레이되어야 할 때 송신될 데이터의 양에 관련된다. 초기화 데이터가 전용 트랙에서 송신되기 때문에, 합성 트랙을 송신하는 것이 요구되지는 않는다.
- [0336] HEVC 파일 포맷의 현재의 사양에 따르면, 파일 포맷에서 파라미터 세트들(PS)을 전달하는 2가지 가능성이 존재한다는 것을 상기해야 한다: 단지 Sample Entry로서 알려진 박스에서 또는 Sample Entry 박스에서 그리고 데이터 샘플들에서. 이를 2가지 구성은 Sample Table로서 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 꾹쳐 파라미터 세트들(PPS) 변경들을 핸들링가능하기 위해서 Sample Entry 박스에서 그리고 (Sample Table 박스에서의 HEVCSampleEntries 파라미터에서 'hev1' 값을 갖는) 데이터 샘플들에서 전달된다.
- [0337] 따라서, 전용 초기화 데이터 트랙(1400)은, 비디오 파라미터 세트, 시퀀스 파라미터 세트 또는 꾹쳐 파라미터 세트에 각각 대응하는 32, 33 또는 34와 동등한 타입의 NAL 유닛들과 같은 낸-VCL HEVC NAL 유닛들만을 데이터로서 포함한다.
- [0338] 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 데이터에 대해 포인팅한다. 마찬가지로, 합성 트랙(1405)의 제1 추출자(1420)는 전용 초기화 데이터 트랙(1400)의 데이터에 대해 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), 후자는 독립적으로 디스플레이가능하지 않은 것으로 고려된다.
- [0339] 비디오 비트 스트림에서 일부 초기화 데이터가 수정될 때(즉, HEVC 비트 스트림에서 꾹쳐 파라미터 세트가 발생할 때), 이들은, 변경이 발생하는 시간적 위치에서, 참조(1425)로 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 이를 새로운 PPS를 참조하는 각각의 타일 트랙에 각각 삽입된다.
- [0340] 캡슐화된 HEVC 비트 스트림의 각각의 트랙에서, 샘플들(및 관련된 NALU들)은 시간적 순서로 구성된다. 마찬가지로, 꾹쳐 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.
- [0341] 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

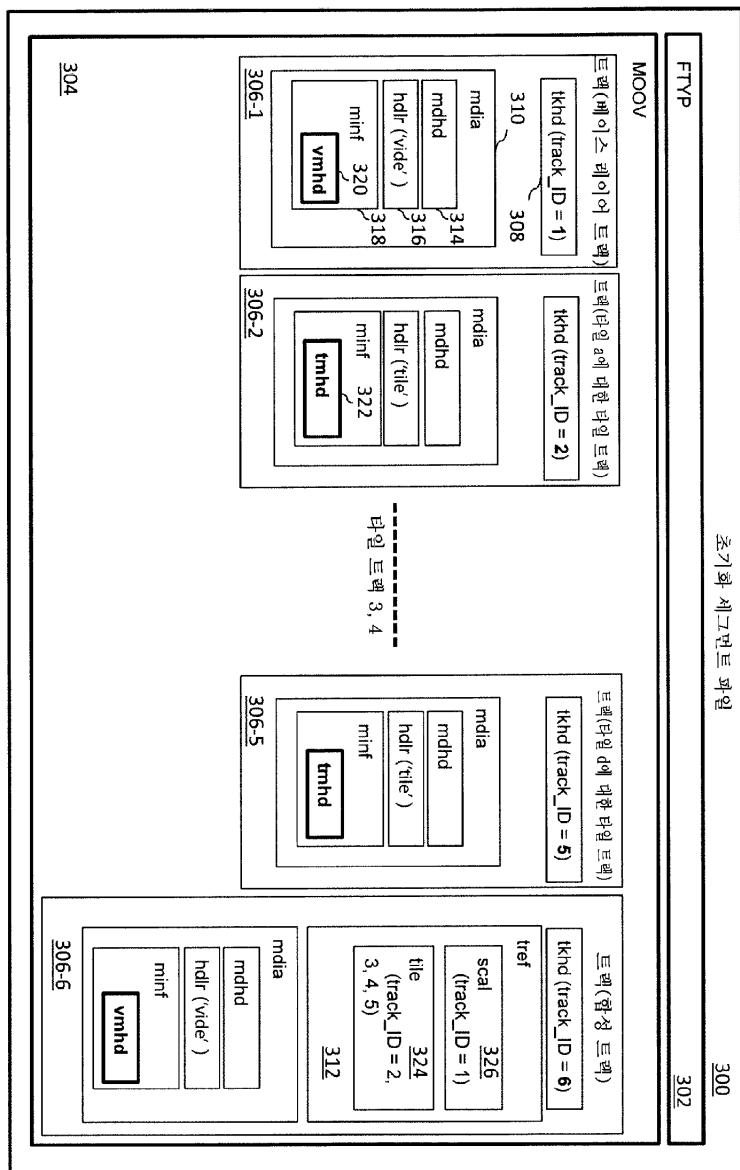
## 도면1



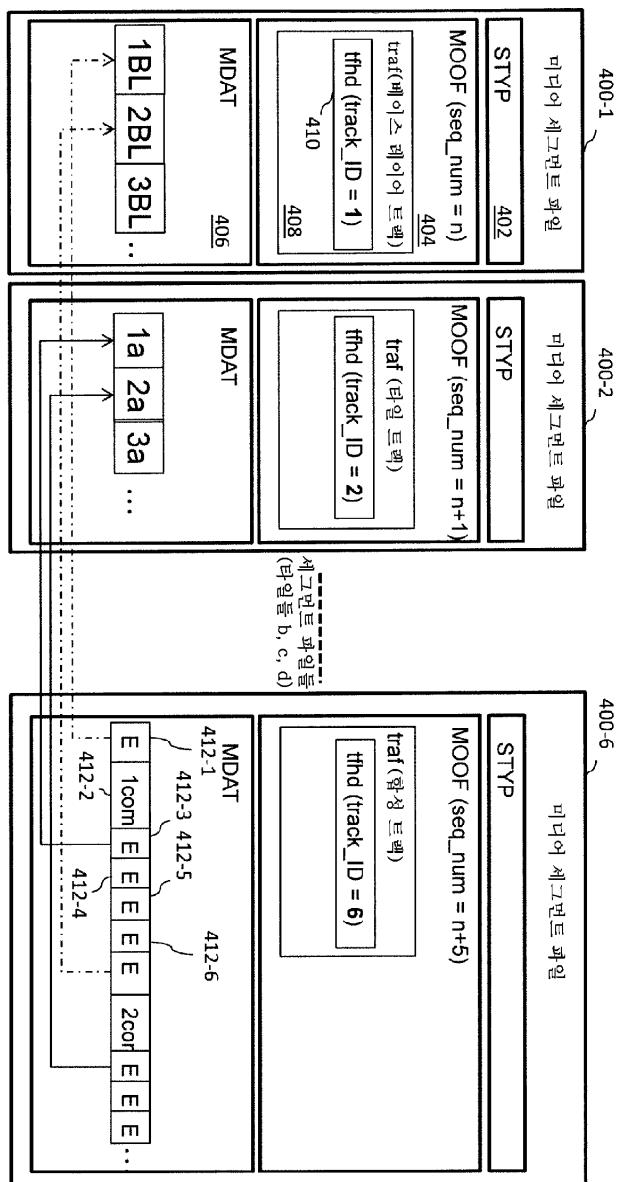
## 도면2



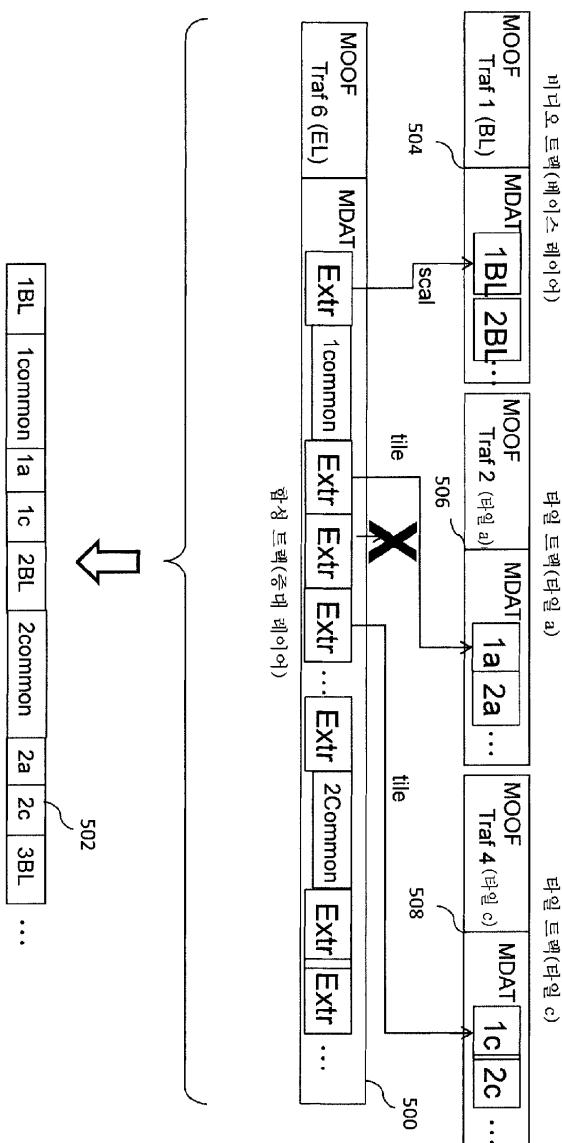
도면3



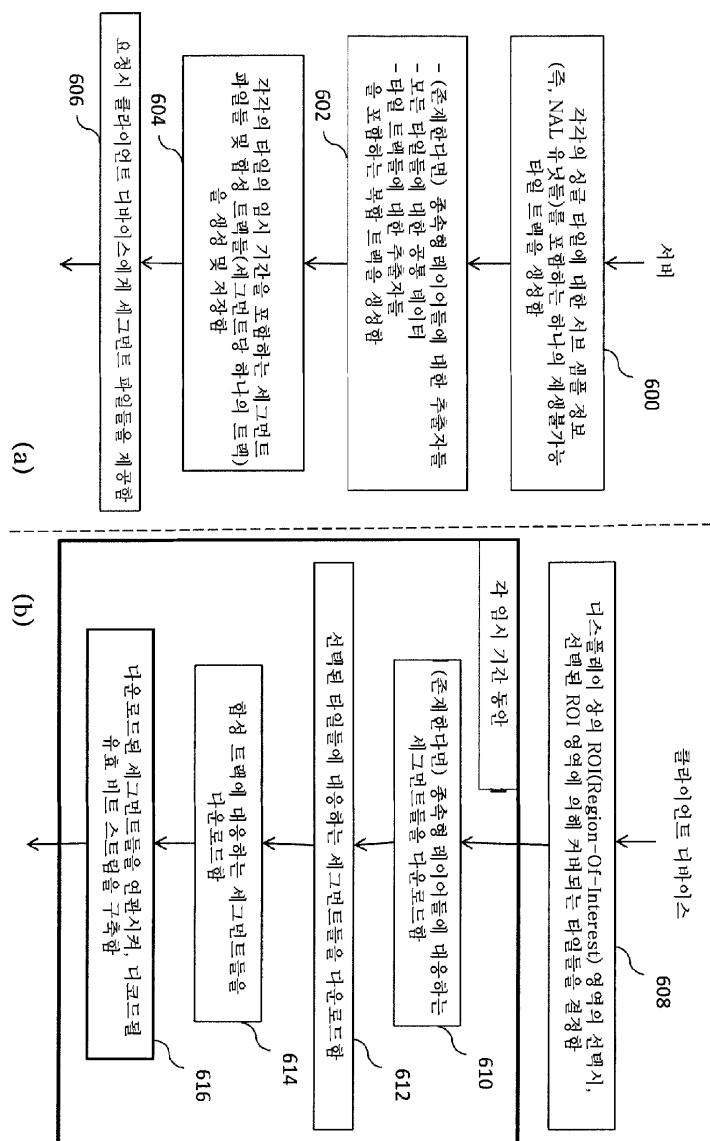
## 도면4



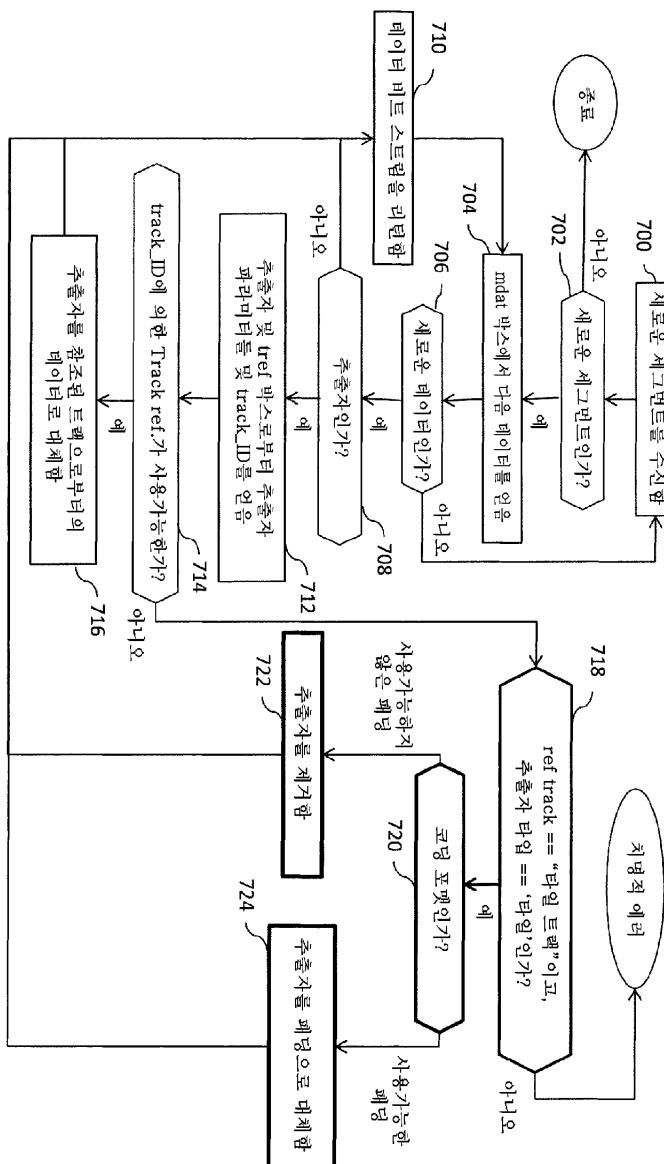
## 도면5



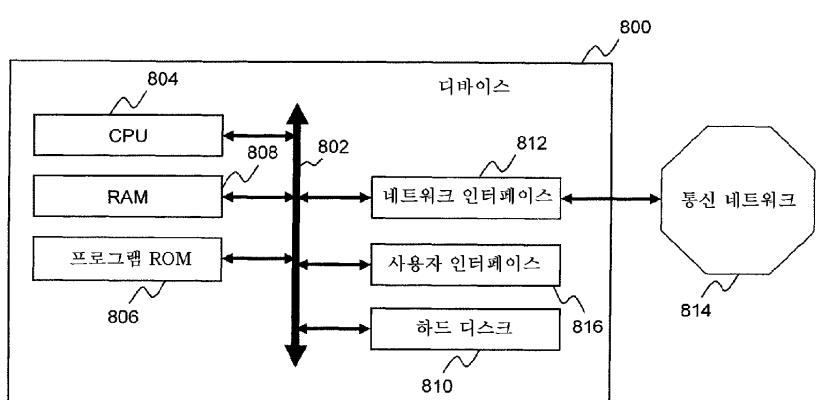
## 도면6



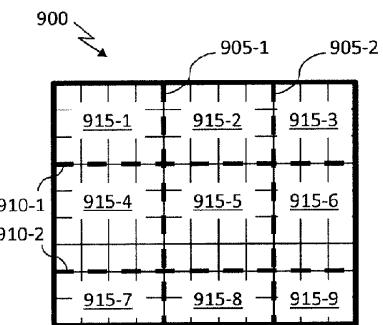
## 도면7



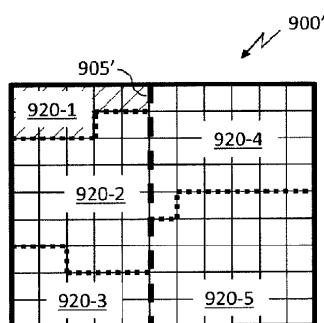
## 도면8



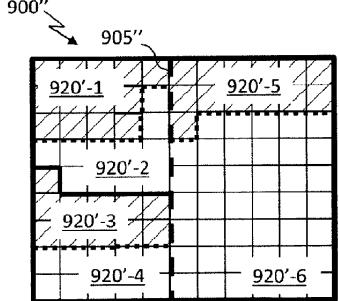
## 도면9



(a)

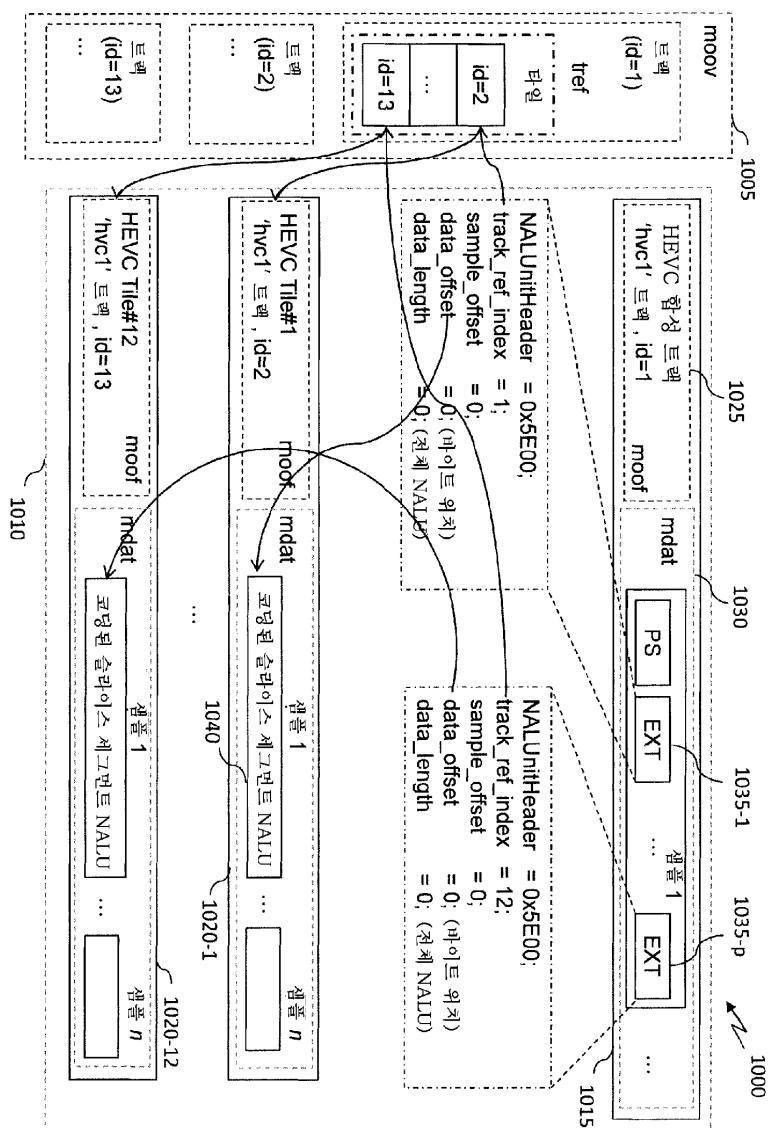


(b)

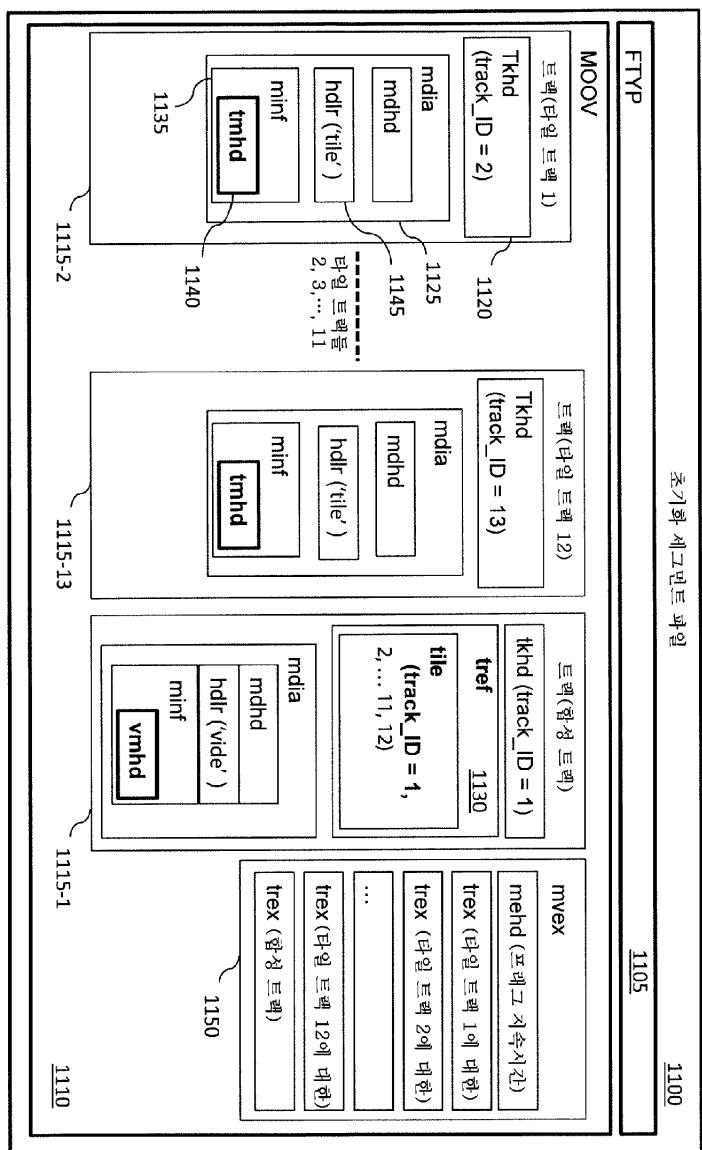


(c)

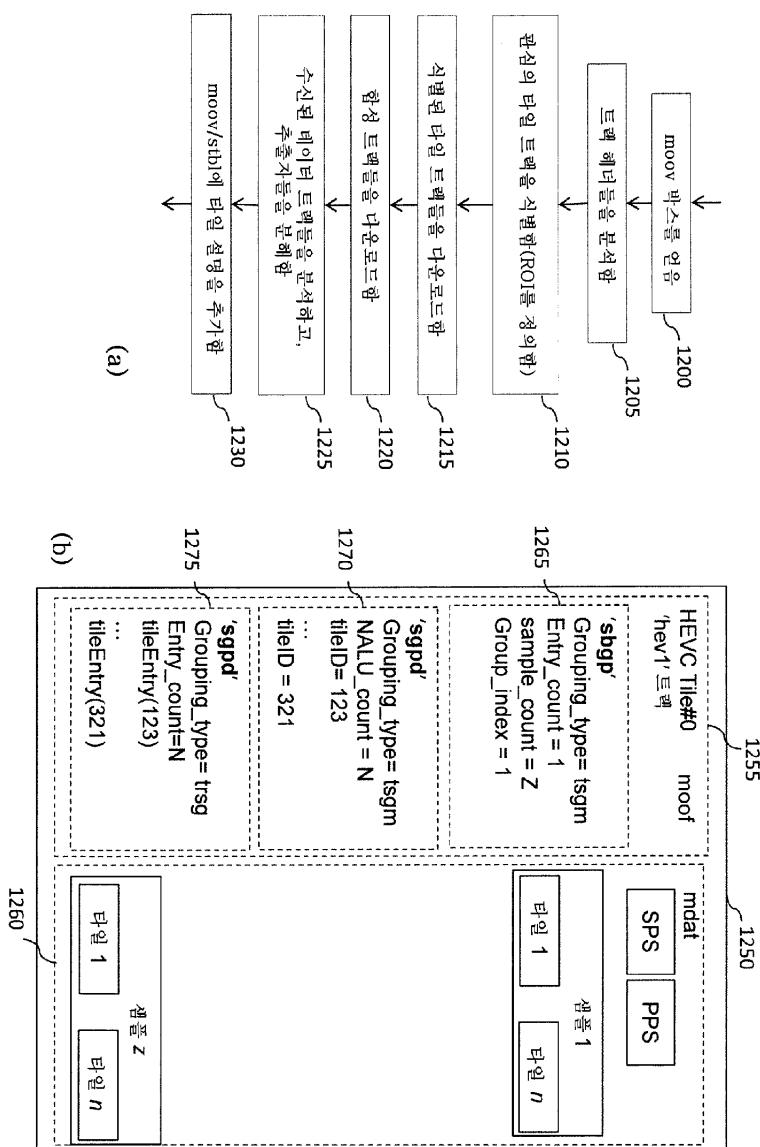
도면10



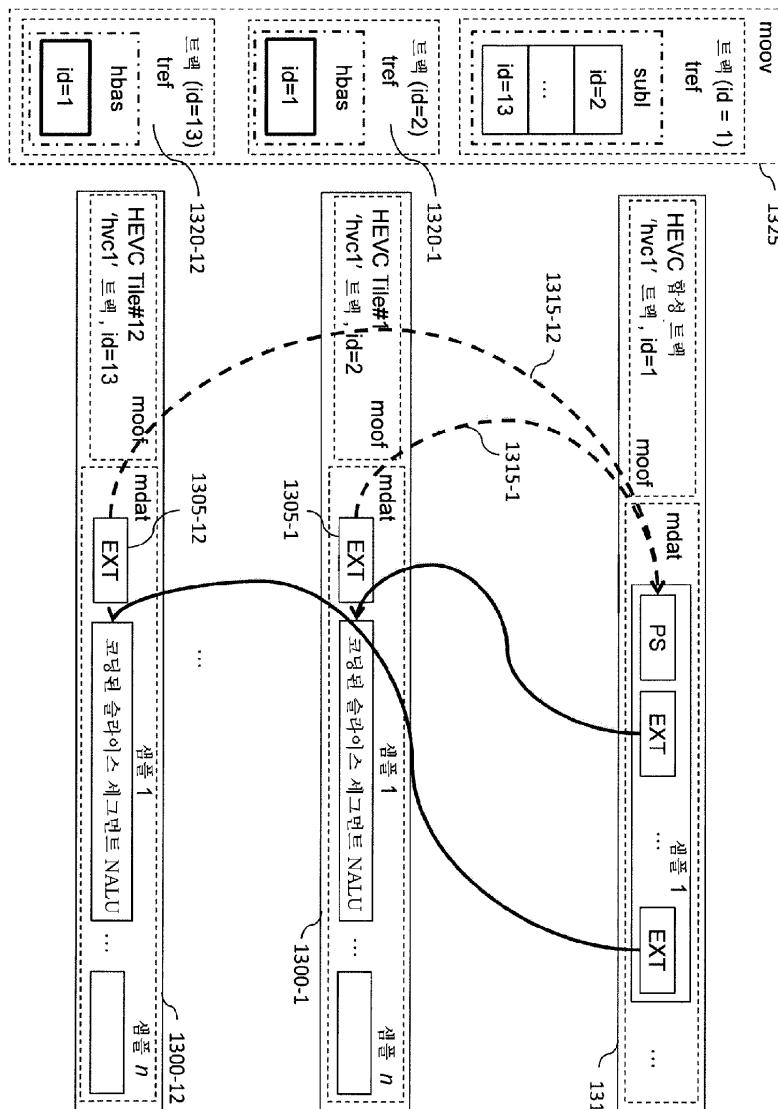
## 도면11



도면12



## 도면13



도면14

