



등록특허 10-2320455



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2021년11월03일
(11) 등록번호 10-2320455
(24) 등록일자 2021년10월27일

- (51) 국제특허분류(Int. Cl.)
H04N 21/236 (2011.01) *H04N 19/119* (2014.01)
H04N 19/167 (2014.01) *H04N 21/2343* (2011.01)
H04N 21/4402 (2011.01) *H04N 21/4728*
(2011.01)
H04N 21/6547 (2011.01) *H04N 21/6587*
(2011.01)
H04N 21/845 (2011.01) *H04N 21/854* (2011.01)
H04N 21/858 (2011.01)
- (52) CPC특허분류
H04N 21/23614 (2013.01)
H04N 21/234327 (2013.01)
- (21) 출원번호 10-2019-7037426
- (22) 출원일자(국제) 2018년06월20일
심사청구일자 2020년01월17일
- (85) 번역문제출일자 2019년12월18일
- (65) 공개번호 10-2020-0019881
- (43) 공개일자 2020년02월25일
- (86) 국제출원번호 PCT/EP2018/066457
- (87) 국제공개번호 WO 2019/002055
국제공개일자 2019년01월03일
- (30) 우선권주장
1710264.1 2017년06월27일 영국(GB)
1710463.9 2017년06월29일 영국(GB)

- (56) 선행기술조사문헌
US20160165321 A1
Byeongdoo Choi 등, Text of ISO/IEC DIS
23090-2 Omnidirectional Media Format(w16824),
MPEG hobart meeting, (2017.4.)*

*는 심사관에 의하여 인용된 문헌

전체 청구항 수 : 총 19 항

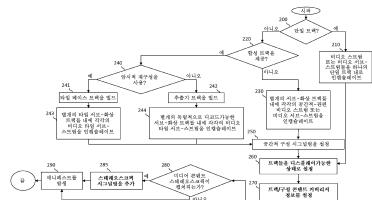
심사관 : 김건우

- (54) 발명의 명칭 미디어 콘텐트를 전송하는 방법, 디바이스, 및 컴퓨터 프로그램

(57) 요 약

본 발명은 장면의 와이드 뷰의 투사에 대응하는 인코드된 미디어 데이터를 서버로부터 클라이언트에 전송하는 방법에 관한 것이고, 이 방법은 인코드된 데이터를 트랙들 내로 인캡슐레이트하는 단계를 포함한다. 적어도 주어진 트랙 내에 인캡슐레이트된 인코드된 데이터를 디스플레이하기 위해 클라이언트가 다른 인코드된 데이터를 파
(뒷면에 계속)

대 표 도



스할 필요가 있다는 것을 시그널하기 위한 정보가 제공된다.

(52) CPC특허분류

H04N 21/234363 (2013.01)

H04N 21/23439 (2013.01)

H04N 21/440263 (2013.01)

H04N 21/4728 (2013.01)

H04N 21/6547 (2013.01)

H04N 21/6587 (2013.01)

H04N 21/8451 (2013.01)

H04N 21/8455 (2013.01)

H04N 21/8456 (2013.01)

(72) 발명자

우드라오고 나엘

프랑스 35330 모르 드 브르따뉴 꾸뚜즈

타웨 조나선

프랑스 35160 탈렌작 라 프로하디에르

사카이 도모야

일본 211-0053 가나가와켄 가와사키시 가미코다나
카 나카하라쿠104 비바체 1-29-1

명세서

청구범위

청구항 1

서버에 의해, 장면의 와이드 뷰에 대응하는 인코드된 미디어 데이터를 인캡슐레이트하는 방법으로서, 상기 방법은

트랙들 내로 인캡슐레이트될 인코드된 미디어 데이터를 획득하는 단계,

주어진 트랙의 트랙 헤더로부터, 상기 주어진 트랙이 단독으로 디스플레이되지 않는다는 것을 시그널하기 위한 정보를 획득하는 단계, 및

상기 인코드된 미디어 데이터 및 상기 정보를 하나 이상의 미디어 파일 내로 인캡슐레이트하는 단계를 포함하는 방법.

청구항 2

제1항에 있어서, 상기 정보는 상기 주어진 트랙이 클라이언트에 의해 선택된 제1 트랙일 수 없다는 것을 시그널하는 방법.

청구항 3

제1항에 있어서, 상기 데이터는 상기 주어진 트랙 내로 인캡슐레이트되지 않는 인코드된 미디어 데이터인 방법.

청구항 4

제1항 내지 제3항 중 어느 한 항에 있어서, 상기 주어진 트랙은, 박스들의 계층 구조에 기초하여 구성된 메타데이터를 포함하는 헤더를 포함하고, 상기 정보는 상기 헤더의 박스 내의 파라미터의 값에 의해 제공되는 방법.

청구항 5

제1항에 있어서, 상기 미디어 데이터는 상기 주어진 트랙이 연관되는 적어도 하나의 다른 트랙의 인코드된 데이터이고, 상기 정보는 상기 적어도 하나의 다른 트랙을 표시하는 방법.

청구항 6

클라이언트에 의해 미디어 파일들을 패스하는 방법으로서, 상기 미디어 파일들은 장면의 와이드 뷰에 대응하는 인코드된 미디어 데이터를 포함하고, 상기 방법은

상기 미디어 파일들 내의, 상기 인코드된 미디어 데이터를 인캡슐레이트하는 적어도 하나의 트랙을 식별하는 단계;

주어진 트랙의 트랙 헤더로부터, 상기 주어진 트랙 내에 인캡슐레이트된 상기 인코드된 미디어 데이터를 프리젠테이션 헤더의 필요성에 관련한 정보를 획득하는 단계, 및

상기 획득된 정보에 따라 상기 주어진 트랙의 상기 인코드된 미디어 데이터를 디코드하는 단계를 포함하는 방법.

청구항 7

제6항에 있어서, 상기 정보는 상기 주어진 트랙이 클라이언트에 의해 선택된 제1 트랙일 수 없다는 것을 시그널하는 방법.

청구항 8

장면의 와이드 뷰의 투사에 대응하는 인코드된 미디어 데이터를 서버에 의해 인캡슐레이트하는 방법으로서, 상기 방법은

상기 인코드된 데이터를 트랙들 내로 인캡슐레이트하는 단계, 및

각각의 트랙에 대한 디스크립티브 메타데이터를 생성하는 단계 - 상기 디스크립티브 메타데이터는 트랙 컨테이너들 내로 구성되고, 트랙 컨테이너는 트랙에 관련됨 - 를 포함하고,

상기 생성하는 단계는

제1 트랙 내에 인캡슐레이트된 인코드된 데이터 및 제1 트랙 컨테이너의 제1 박스에서 정의된 공통 세트에 속하는 하나 이상의 트랙 내에 인캡슐레이트된 인코드된 데이터에 의해 나타내진 상기 와이드 뷰의 커버리지를 표시하는 정보를 포함하는 제1 구조를 제공하는 단계 - 상기 제1 구조는 타일 트랙의 커버리지를 표시하는 정보를 더 포함함 - 를 추가로 포함하는 방법.

청구항 9

제8항에 있어서, 상기 제1 구조의 상기 정보는 상기 제1 박스 내에 제공되는 방법.

청구항 10

제8항에 있어서,

상기 세트는 ISOBMFF 14496-12에 의해 정의된 '트랙 그룹'인 방법.

청구항 11

제10항에 있어서, 상기 제1 박스는 ISOBMFF 14496-12에 의해 정의된 '트랙 그룹' 박스인 방법.

청구항 12

제8항 내지 제11항 중 어느 한 항에 있어서, 상기 트랙들에 관한 단지 디스크립티브 메타데이터를 포함하는 참조 트랙이 발생되는지를 테스트하는 테스트 단계를 추가로 포함하고,

상기 생성하는 단계는

참조 트랙이 발생되면, 풀 와이드 뷰의 커버리지를 표시하는 참조 정보를 포함하는 상기 제1 구조를 제공하는 단계,

그렇지 않은 경우, 트랙 내에 인캡슐레이트된 인코드된 데이터 및 제1 컨테이너의 제1 박스에서 정의된 공통 세트에 속하는 하나 이상의 트랙 내에 인캡슐레이트된 인코드된 데이터에 의해 나타내진 상기 와이드 뷰의 커버리지를 표시하는 상기 정보를 포함하는 상기 제1 구조를 제공하는 단계를 추가로 포함하는 방법.

청구항 13

클라이언트에 의해 미디어 파일들을 파스하는 방법으로서, 상기 미디어 파일들은 장면의 와이드 뷰의 투사에 대응하는 인코드된 미디어 데이터에 대응하고, 상기 방법은

상기 미디어 파일들로부터, 인코드된 데이터를 인캡슐레이트하는 트랙들을 획득하는 단계,

각각의 트랙으로부터 디스크립티브 메타데이터를 획득하는 단계 - 상기 디스크립티브 메타데이터는 트랙 컨테이너들 내로 구성되고, 트랙 컨테이너는 트랙에 관련됨 - 를 포함하고,

상기 디스크립티브 메타데이터를 획득하는 단계는

제1 트랙 내에 인캡슐레이트된 인코드된 데이터 및 제1 트랙 컨테이너의 제1 박스에서 정의된 공통 세트에 속하는 하나 이상의 트랙 내에 인캡슐레이트된 인코드된 데이터에 의해 나타내진 상기 와이드 뷰의 커버리지를 표시하는 정보를 포함하는 제1 구조를 획득하는 단계 - 상기 제1 구조는 타일 트랙의 커버리지를 표시하는 정보를 더 포함함 - 를 추가로 포함하는 방법.

청구항 14

제1항 내지 제3항, 제5항 내지 제11항, 제13항 중 어느 한 항에 있어서, 상기 인코드된 데이터 또는 인코드된 비디오 데이터는 인코드된 전방향 데이터 또는 인코드된 전방향 비디오 데이터인 방법.

청구항 15

컴퓨터 또는 프로세서에 의해 실행될 때, 상기 컴퓨터 또는 프로세서로 하여금 제1항 내지 제3항, 제5항 내지 제11항, 제13항 중 어느 한 항의 방법을 수행하게 하는, 컴퓨터 판독가능 저장 매체에 저장된 컴퓨터 프로그램.

청구항 16

장면의 와이드 뷰에 대응하는 인코드된 미디어 데이터를 서버에 의해 인캡슐레이트하는 디바이스로서, 상기 디바이스는 프로세서를 포함하고, 상기 프로세서는

트랙들 내로 인캡슐레이트될 인코드된 미디어 데이터를 획득하고,

주어진 트랙의 트랙 헤더로부터, 상기 주어진 트랙이 단독으로 디스플레이되지 않는다는 것을 시그널하기 위한 정보를 획득하고,

상기 인코드된 미디어 데이터 및 상기 정보를 하나 이상의 미디어 파일 내로 인캡슐레이트하도록 구성되는 디바이스.

청구항 17

장면의 와이드 뷰의 투사에 대응하는 인코드된 미디어 데이터를 서버에 의해 인캡슐레이트하는 디바이스로서, 상기 디바이스는 프로세서를 포함하고, 상기 프로세서는

상기 인코드된 데이터를 트랙들 내로 인캡슐레이트하고,

각각의 트랙에 대한 디스크립티브 메타데이터를 획득하고 - 상기 디스크립티브 메타데이터는 트랙 컨테이너들 내로 구성되고, 트랙 컨테이너는 트랙에 관련됨 - ,

제1 트랙 내에 인캡슐레이트된 인코드된 데이터 및 제1 트랙 컨테이너의 제1 박스에서 정의된 공통 세트에 속하는 하나 이상의 트랙 내에 인캡슐레이트된 인코드된 데이터에 의해 나타내진 상기 와이드 뷰의 커버리지를 표시하는 정보를 포함하는 제1 구조 - 상기 제1 구조는 타일 트랙의 커버리지를 표시하는 정보를 더 포함함 - 를 제공하도록 구성되는 디바이스.

청구항 18

클라이언트에 의해 미디어 파일들을 패스하는 디바이스로서, 상기 미디어 파일들은 장면의 와이드 뷰의 투사에 대응하는 인코드된 미디어 데이터에 대응하고, 상기 디바이스는 프로세서를 포함하고, 상기 프로세서는

상기 미디어 파일들로부터 트랙들을 획득하고,

각각의 트랙으로부터 디스크립티브 메타데이터를 획득하고 - 상기 디스크립티브 메타데이터는 트랙 컨테이너들 내로 구성되고, 트랙 컨테이너는 트랙에 관련됨 - ,

제1 트랙 내에 인캡슐레이트된 인코드된 데이터 및 제1 트랙 컨테이너의 제1 박스에서 정의된 공통 세트에 속하는 하나 이상의 트랙 내에 인캡슐레이트된 인코드된 데이터에 의해 나타내진 상기 와이드 뷰의 커버리지를 표시하는 정보를 포함하는 제1 구조 - 상기 제1 구조는 타일 트랙의 커버리지를 표시하는 정보를 더 포함함 - 를 획득하도록 구성되는 디바이스.

청구항 19

제16항 내지 제18항 중 어느 한 항에 있어서, 상기 인코드된 데이터 또는 인코드된 비디오 데이터는 인코드된 전방향 데이터 또는 인코드된 전방향 비디오 데이터인 디바이스.

청구항 20

삭제

청구항 21

삭제

청구항 22

삭제

청구항 23

삭제

청구항 24

삭제

청구항 25

삭제

청구항 26

삭제

청구항 27

삭제

청구항 28

삭제

청구항 29

삭제

청구항 30

삭제

청구항 31

삭제

청구항 32

삭제

청구항 33

삭제

청구항 34

삭제

청구항 35

삭제

청구항 36

삭제

청구항 37

삭제

청구항 38

삭제

청구항 39

삭제

청구항 40

삭제

청구항 41

삭제

청구항 42

삭제

청구항 43

삭제

발명의 설명

기술 분야

[0001]

본 발명은 미디어 데이터를 전송하는 방법 및 디바이스에 관한 것이다.

배경 기술

[0002]

본 발명은 예를 들어, MPEG 표준화 기구에 의해 정의된 것과 같은 ISO 베이스 미디어 파일 포맷에 따라 가상 현실 미디어 콘텐트를 인캡슐레이트하고, 패스하고 스트리밍하여, 가상 현실 미디어 콘텐트의 교환, 관리, 편집, 및 프리젠테이션을 용이하게 하는 신축성있고 확장가능한 포맷을 제공하고 예를 들어 적응형 http 스트리밍 프로토콜을 사용하는 인터넷과 같은 IP 네트워크를 통해 그것의 전달을 개선시키는 것에 관한 것이다.

[0003]

국제 표준 기구 베이스 미디어 파일 포맷(ISO BMFF, ISO/IEC 14496-12)은 네트워크를 통해 또는 또 하나의 비트 스트림 전달 메커니즘을 통해 로컬 저장 및 전송을 위한 인코드된 타이밍된 미디어 데이터 비트스트림들을 디스크라이브하는 널리 공지된 신축성있고 확장가능한 포맷이다. 이 파일 포맷은 객체 지향형이다. 그것은 순차적으로 또는 계층적으로 구성되고 타이밍 및 구조 파라미터들과 같은 인코드된 타이밍된 미디어 데이터 비트스트림의 파라미터들을 정의하는 박스들이라고 하는 빌딩 블록들로 구성된다. 파일 포맷에서, 전체적인 프리젠테이션은 영화라고 부른다. 그것은 트랙들로 논리적으로 분할된다. 각각의 트랙은 미디어 데이터(예를 들어, 비디오의 프레임들)의 타이밍된 시퀀스를 나타낸다. 각각의 트랙 내에서, 데이터의 각각의 타이밍된 단위는 샘플이라고 부르고; 이것은 비디오 또는 오디오의 프레임일 수 있다. 샘플들은 암시적으로 시퀀스로 번호가 매겨진다. 영화는 영화 및 트랙 프래그먼트들의 리스트로서 시간적으로 구성될 수 있다. 실제 샘플들은 MediaDataBoxes라고 부르는 박스들 내에 있다. 영화 프래그먼트 내에 트랙 당 제로 이상인, 트랙 프래그먼트들의 세트가 있다. 트랙 프래그먼트들은 결국 제로 이상의 트랙 런들을 포함하고, 그 각각은 그 트랙에 대한 샘플들의 연속적인 런을 도큐먼트한다.

[0004]

사용자 경험을 개선시키고 특히 몰입 경험을 제공하기 위해, 타이밍된 미디어 데이터 비트스트림들(비디오들 및 심지어 오디오)은 전방향(또는 다수 방향 또는 복수 방향)일 수 있다. 360° 파노라마식 비디오라고도 알려진 비디오들에 적용될 때, 사용자는 디스플레이되는 장면에 위치하는 것처럼 느낀다.

[0005]

전방향 비디오는 360° 카메라로부터 및/또는 모든 카메라들이 공통 절점을 갖도록 예를 들어, 특수 장치 상에 장착된 여러 개의 카메라들로부터 획득된 비디오 스트리밍들의 영상들을 조합함으로써 획득될 수 있다. 영상들의 이러한 조합은 영상 스티칭 또는 카메라 스티칭이라고 공지되어 있다.

[0006]

이러한 전방향 비디오는 사용자의 시청 배향에 따라 또는 사용자들을 둘러싸는 커브드 화면 상으로의 투사를 통

해 헤드 마운트 디스플레이들을 통해 렌더링될 수 있다. 그것은 또한 전방향 비디오의 사용자의 원하는 부분(뷰포트라고도 공지됨)에 따라 전방향 비디오 새로 팬하기 위해 내비게이션 사용자 인터페이스를 갖는 전통적인 2D 화면들 상에 디스플레이될 수 있다. 사용자가 가상 세계에 있는 것처럼 느낄 수 있기 때문에 이것은 보통 가상 현실(VR)이라고 한다. 가상 물체들이 전방향 비디오에 추가될 때, 그것은 중장 현실(AR)이라고 한다.

[0007] 도 1은 서버 디바이스(101)로부터 전방향 미디어를 캡처하여 클라이언트 디바이스(170)(또한 170'이라고 표시됨)에 전송하고 렌더링하는 데이터 흐름의 예를 도시한다.

[0008] 도시된 것과 같이, 이 미디어는 카메라 시스템(100)으로부터 취득되고 헤드 마운트 디스플레이(HMD)(170 및 170')에 전달되는 비디오 콘텐트를 갖는다. 전달(160)은 스트리밍 서버(161) 및 스트리밍 클라이언트(162)를 통해 예를 들어 적응형 http 스트리밍 프로토콜을 사용하는 인터넷과 같은 IP 네트워크(163)를 통해 수행될 수 있다.

[0009] 예시의 목적을 위해, 사용된 카메라 시스템(100)은 정육면체의 각각의 면과 연관된 6개의 표준 카메라의 세트에 기초한다. 그것은 카메라 시스템을 둘러싸는 실제 장면을 나타내는 영상들을 캡처하기(단계 110) 위해 사용된다. 이 배열에 따라, 하나의 카메라는 전방 영상들을 제공하고, 하나의 카메라는 후방 영상들을 제공하고, 하나의 카메라는 좌측 영상들을 제공하고, 하나의 카메라는 우측 영상들을 제공하고, 하나의 카메라는 하부 영상들을 제공하고, 하나의 카메라는 상부 영상들을 제공한다.

[0010] 카메라 시스템(100)으로부터 획득된 영상들은 360 비디오 스트림 또는 가상 현실 미디어 데이터 스트림이라고도 하는 전방향 비디오 스트림을 형성하는 360 영상들을 생성하기 위해 서버(101)에서 처리된다(단계 120).

[0011] 처리 단계 120은 동일한 시간 인스턴스의 캡처된 영상들을 스티치하고 투사하는 것으로 구성된다. 영상들은 수평과 수직 차원들 둘 다에서 360° 뷰를 형성하는 구(121)를 나타내는 3차원 투사 구조 상으로 먼저 스티치되고 투사된다. 투사 구조 상의 360 영상 데이터는 예를 들어, 등방형 투사 (https://en.wikipedia.org/wiki/Equirectangular_projection)를 사용하여, 2차원 투사된 영상(122)(캡처링 투사라고도 표시됨) 상으로 추가로 변환된다. 투사된 영상은 전체 구를 커버한다.

[0012] 대안적으로, 전방향 미디어가 스테레오스코픽 360도 비디오이면, 카메라 시스템(100)은 단계 110에서 3차원 360도 장면을 렌더링하기 위해 클라이언트에 의해 나중에 사용될 수 있는 좌측 뷰 및 우측 뷰를 나타내는 영상 시퀀스들을 캡처하는 다수의 카메라로 구성될 수 있다. 이러한 경우에, 위에 설명된 처리 단계 120은 좌측-뷰와 우측-뷰 영상들 시퀀스 둘 다에 별개로 적용된다. 선택적으로, 단계 125에서, 프레임 패킹이 동일한 시간 인스턴스의 각각의 좌측 뷰 영상 및 우측 뷰 영상을 하나의 단일 좌측+우측 투사된 영상들 시퀀스 상에 생기는 동일한 투사된 영상으로 패킹하기 위해 적용될 수 있다. 여러 개의 스테레오스코픽 프레임 패킹 배열들이 예를 들어, 사이드-바이-사이드, 상부-하부, 열 기반 인터리빙, 행 기반 인터리빙, 교대하는 좌측과 우측 뷰들의 시간적 인터리빙에 의해 가능하다. 대안적으로, 스테레오스코픽 프레임 패킹 배열은 인코딩 단계 140 이후에 독립적인 비디오 비트스트림들을 초래하는 별개의 그리고 독립적인 투사된 영상들 시퀀스에 좌측 및 우측 뷰들을 유지하는 것으로 또한 구성될 수 있다. 예를 들어, 하나의 비디오 비트스트림은 좌측 뷰 영상들을 나타내고 다른 하나는 우측 뷰 영상들을 나타낸다.

[0013] 선택적으로, 영역-와이즈 패킹(130)은 다음에 투사된 영상(122)을 패킹된 영상(131) 상에 맵핑하기 위해 적용된다. 영역-와이즈 패킹은 예를 들어 사용자에 대한 구의 가장 유용한 부분들에 관한 신호 정보를 최대화하기 위해 투사된 영상의 영역들의 변환, 리사이즈, 및 재배치를 적용하는 것으로 구성된다. 패킹된 영상은 전체 구의 일부만을 커버할 수 있는 점에 주목할 수 있다. 영역-와이즈 패킹이 적용되지 않으면, 패킹된 영상(131)은 투사된 영상(122)과 동일하다. 스테레오스코픽 전방향 미디어의 경우에, 영역-와이즈 패킹은 단계 125에서 선택된 프레임 패킹 배열에 따라 좌측+우측 투사된 영상들 시퀀스에, 또는 좌측-뷰와 우측-뷰 투사된 영상들 시퀀스들에 따로 적용한다.

[0014] 투사된 영상들(122)은 단계 140에서 하나 또는 여러 개의 비디오 비트스트림들 새로 인코드된다. 스테레오스코픽 전방향 미디어의 경우에, 인코딩 단계는 단계 125에서 선택된 프레임 패킹 배열에 따라 좌측+우측 패킹된 영상들 시퀀스에, 또는 좌측-뷰와 우측-뷰 패킹된 영상들 시퀀스들에 따로 적용한다. 대안적으로, 다수-뷰 인코딩은 좌측-뷰 및 우측-뷰 패킹된 영상들 시퀀스들 상에서 사용될 수 있다.

[0015] 인코딩 포맷들의 예들은 AVC(진보된 비디오 코딩), SVC(스케일가능한 비디오 코딩), HEVC(고효율 비디오 코딩) 또는 L-HEVC(층이 진 HEVC)이다. 다음에서, HEVC는 HEVC와 그것의 층이 진 확장들(L-HEVC) 둘 다를 참조하기 위해 사용된다.

- [0016] HEVC 및 유사한 비디오 인코딩 포맷들은 샘플들, 예를 들어, 화상들: 파일들, 슬라이스들 및 슬라이스 세그먼트들의 상이한 공간적 세분들을 정의한다. 파일은 수평 및 수직 경계들(즉, 행들 및 열들)에 의해 정의되고 이후 모두 코딩 유닛들이라고 하는 정수 개의 코딩 트리 유닛들(CTU들) 또는 코딩 블록들을 포함하는 화상의 직사각형 영역을 정의한다. 이와 같이, 파일들은 화상의 공간적 서브-부분들을 나타내기 위한 좋은 후보들이다. 그러나, NAL 유닛들(또는 NALU들) 내로의 선택스 및 그것의 인캡슐레이션의 관점에서의 코드된 비디오 데이터(비트스트림) 구성은 오히려 (AVC에서와 같이) 슬라이스들 및 슬라이스 세그먼트들에 기초한다.
- [0017] HEVC 내의 슬라이스는 슬라이스 세그먼트들의 세트이고, 적어도 제1 슬라이스 세그먼트는 독립적인 슬라이스 세그먼트이고, 있다면, 다른 것들은 종속적인 슬라이스 세그먼트들이다. 슬라이스 세그먼트는 (래스터 스캔 순서로) 정수 개의 연속적인 CTU들을 포함한다. 슬라이스는 반드시 직사각형 형상을 가질 필요는 없다(그러므로 그것은 공간적 서브-부분 표현들을 위한 파일들보다 덜 적절하다). 슬라이스 세그먼트는 slice_segment_header 이후의 slice_segment_data로서 HEVC 비트스트림에서 인코드된다. 독립적인 슬라이스 세그먼트들(ISS)과 종속적인 슬라이스 세그먼트들(DSS)은 그들의 헤더가 상이하고: 종속적인 슬라이스 세그먼트는 독립적인 슬라이스 세그먼트의 헤더로부터의 정보를 재사용하기 때문에 더 짧은 헤더를 갖는다. 독립적인 슬라이스 세그먼트와 종속적인 슬라이스 세그먼트 둘 다는 비트스트림 내에 엔트리 포인트들의 리스트를 포함한다.
- [0018] 비디오 비트스트림이 파일들로 인코드될 때, 파일들이 동일한 화상 내의 이웃 파일들로부터(공간적 종속) 그리고 이전의 참조 화상들 내의 이웃 파일들로부터(시간적 종속) 종속하지 않는 것을 보장하기 위해, 파일들은 움직임-제한될 수 있다. 그러므로, 움직임-제한된 파일들은 독립적으로 디코드 가능하다.
- [0019] 대안적으로, 패킹된 영상은 인코딩 전에 여러 개의 공간적 서브 화상들로 분리될 수 있고, 각각의 서브 화상은 예를 들어 독립적인 인코드된 HEVC 비트스트림을 독립적으로 형성하게 인코드된다.
- [0020] 그러므로, 인코딩 단계 140의 결과로서, 패킹된 영상(131)은 하나 이상의 독립적인 인코드된 비트스트림에 의해 또는 하나 이상의 독립적으로 인코드된 서브-비트스트림으로 구성된 적어도 하나의 인코드된 비트스트림에 의해 나타내질 수 있다.
- [0021] 그들 인코드된 비트스트림 및 서브-비트스트림은 다음에 인캡슐레이션 파일 포맷에 따라, 예를 들어 MPEG 표준화 기구에 의해 정의된 것과 같은 ISO 베이스 미디어 파일 포맷 및 전방향 미디어 포맷(OMAF)에 따라, 파일 내에 또는 작은 시간적 세그먼트 파일들(165) 내에 단계 150에서 인캡슐레이트된다. 결과적인 파일 또는 세그먼트 파일들은 mp4 파일 또는 mp4 세그먼트들일 수 있다. 인캡슐레이션 동안, 오디오 스트림이 비디오 또는 오디오 스트림들에 관한 정보를 제공하는 메타데이터 트랙들뿐만 아니라 비디오 비트스트림에 추가될 수 있다.
- [0022] 인캡슐레이트된 파일 또는 세그먼트 파일들은 다음에 예를 들어 http(하이퍼텍스트 트랜스퍼 프로토콜) 프로토콜을 사용하는 인터넷을 통해 또는 예를 들어 디스크와 같은 착탈 가능한 디지털 매체 상에서, 전달 메커니즘(160)을 통해 클라이언트(170)에 전달된다. 예시의 목적을 위해, 전달(160)은 MPEG 표준화 위원회("ISO/IEC 23009-1, HTTP를 통한 동적 적응형 스트리밍(DASH), 파트1: 미디어 프리젠테이션 디스크립션 및 세그먼트 포맷들")로부터 DASH(HTTP를 통한 동적 적응형 스트리밍)과 같은 HTTP를 통한 적응형 스트리밍을 사용하여 수행된다.
- [0023] 이 표준은 HTTP 균일한 리소스 위치들(URL들)과의 미디어 프리젠테이션의 미디어 콘텐트의 컴팩트 디스크립션의 연관을 가능하게 한다. 이러한 연관은 전형적으로 매니페스트 파일 또는 디스크립션 파일(164)이라고 하는 파일에서 디스크라이브된다. DASH의 맥락에서, 이 매니페스트 파일은 MPD 파일(미디어 프리젠테이션 디스크립션)이라고도 하는 XML 파일이다.
- [0024] MPD 파일을 수신함으로써, 클라이언트 디바이스(170)는 각각의 미디어 콘텐트 소자의 디스크립션을 얻는다. 따라서, 그것은 미디어 프리젠테이션에서 제안된 미디어 콘텐트 소자들의 종류를 인식하고 스트리밍 클라이언트(162)를 통해, 스트리밍 서버(161)로부터 연관된 미디어 세그먼트들(165)을 다운로드하기 위해 사용될 HTTP URL들을 안다. 그러므로, 클라이언트(170)는 (HTTP 요청들을 통해) 다운로드하고 플레이(즉, 미디어 세그먼트들의 수신 후에 디코드하고 플레이)할 미디어 콘텐트 소자들을 결정할 수 있다.
- [0025] 클라이언트 디바이스는 사용자의 뷰포트에 의존하는 장면의 와이드 뷰를 나타내는 풀 패킹된 영상들의 공간적 부분(즉, 현재 디스플레이되고 사용자에 의해 보여지는 구형의 비디오의 부분)에 대응하는 미디어 세그먼트들을 단지 얻을 수 있다는 점에 주목하여야 한다. 장면의 와이드 뷰는 풀 패킹된 영상에 의해 나타내진 풀 뷰를 나타낼 수 있다.

- [0026] 수신 시에, 인캡슐레이트된 가상 현실 미디어 파일 또는 미디어 세그먼트들은 단계 141에서 디코드된 데이터 스트림을 추출하도록 단계 151 동안 파스된다. 단계 151에서 수신된 ISOBMFF 파일 또는 세그먼트들의 경우에, 파싱은 디스크립티드 메타데이터로부터, 인캡슐레이트된 비디오 비트스트림들 및/또는 비디오 서브-비트스트림들을 추출할 수 있는 mp4 리더 또는 mp4 파서에 의해 전형적으로 취급된다.
- [0027] 다음에, 선택적으로, 디코딩 단계(141)로부터 생긴 패킹된 영상들은 다음에 비디오 렌더링(단계 121)을 위해 처리되고 디스플레이되는(단계 111) 투사된 영상들을 획득하기 위해 언패킹된다. 비디오 렌더링은 그 중에서 사용자의 관점, 시점, 및 투사된 영상들을 생성하기 위해 사용되는 투사(들)인 여러 개의 파라미터들에 의존한다는 점에 주목하여야 한다. 도시된 것과 같이, 비디오를 렌더링하는 것은 디코드된 투사된 영상들을 구상에 재투사하는 단계를 포함한다. 이러한 재투사로부터 획득된 영상들은 헤드 마운트 디스플레이(170')에 디스플레이된다.
- [0028] 스테레오스코픽 뷰들을 취급하기 위해, 도 1을 참조하여 설명된 과정이 중복되거나 부분적으로 중복될 수 있다.
- [0029] UHD(초고화질) 비디오 스트림들의 여러 개의 영상들을 가상 현실 미디어 데이터 스트림의 파노라마 영상들로 스티칭하는 것은 매우 높은 비트레이트 및 매우 높은 해상도 가상 현실 미디어 데이터 스트림에 이르게 한다. 그러므로, 시스템의 관점으로부터 그리고 대역폭을 낭비하는 것을 피하기 위해 그리고 클라이언트 플레이어들의 처리 능력들에 순응하게 남기 위해, 가상 현실 미디어 데이터에의 액세스를 최적화할 필요가 있다.
- [0030] 이러한 필요는 가상 현실 미디어 데이터 스트림이 도 1을 참조하여 설명된 것 이외의 다른 목적들을 위해 사용될 수 있다는 점에서 훨씬 더 중요하다. 특히, 가상 현실 미디어 데이터 스트림은 투사기들의 360° 어레이와 같은 특정한 디스플레이들로 360° 영상들을 디스플레이하기 위해 사용될 수 있다. 그것은 또한 특정한 시야를 디스플레이하고/하거나 관점, 시야, 및 시점을 변경하기 위해 사용될 수 있다.
- [0031] 본 발명자들은 도 1을 참조하여 설명된 과정을 따라, 전송할 미디어 데이터에 관한 정보를 디스크라이브하고 시그널링에 그널할 때 몇가지 문제들을 인식하였다.
- [0032] 한 예는 오버헤드를 발생하고 복잡한, 클라이언트로부터의 특정한 파싱 과정을 요청하는 트랙들의 시그널링에 관한 것이다.
- [0033] 또 하나의 예는 스테레오 뷰들의 시그널링은 특정한 인캡슐레이팅 과정에 제한되고 비교적 비용이 많이 듦다는 문제이다.
- [0034] 또 하나의 예는 트랙들 내의 인코드된 데이터 내의 커버리지의 시그널링에 관한 것이다. 서브-화상 트랙들이 여러 개의 상이한 트랙들 내에 인캡슐레이트될 때, 기존의 해결책들은 복잡하고 다수-트랙들 인캡슐레이션 과정에 완전히 순응하지 못한다.

발명의 내용

해결하려는 과제

- [0035] 본 발명은 전술한 문제들 중 하나 이상을 해결하기 위해 고안되었다.
- [0036] 이 맥락에서, 예를 들어 http 프로토콜을 사용하는 인터넷과 같은 IP 네트워크를 통해, 미디어 콘텐트(예를 들어, 전방향 미디어 콘텐트)를 스트리밍하기 위한 해결책이 제공된다.

과제의 해결 수단

- [0037] 본 발명의 한 실시예에 따르면, 장면의 와이드 뷰의 투사에 대응하는 인코드된 미디어 데이터를 서버로부터 클라이언트에 전송하는 방법이 제공되고, 상기 방법은
- [0038] 인코드된 데이터를 트랙들 내로 인캡슐레이트하는 단계를 포함하고,
- [0039] 주어진 트랙 내에 인캡슐레이트된 인코드된 데이터를 적어도 디스플레이하기 위해 클라이언트가 다른 인코드된 데이터를 파스할 필요가 있다는 것을 시그널하기 위한 정보가 제공된다.
- [0040] 본 발명의 한 실시예는 특정한 트랙들에 대한 더 간단한 시그널링을 제공하고, 특히 트랙들은 OMAF에서 "단독으로 프리젠테일 의도가 없는" 트랙들인 것으로 정의된다. 그것은 OMAF 콘텐트가 서브-화상 트랙들로 분리될 때 시그널링 오버헤드 및 파싱 복잡성을 감소시키는데, 이는 트랙들이 하나 이상의 시간 기간 동안 화상의 부분에

대응하는 데이터를 포함한다는 것을 의미한다.

- [0041] 본 발명의 또 하나의 실시예에 따르면, 클라이언트에 의해 미디어 파일들을 패스하는 방법이 제공되고, 상기 미디어 파일들은 장면의 와이드 뷰의 투사에 대응하는 인코드된 미디어 데이터를 포함하고, 상기 방법은
- [0042] 미디어 파일들로부터, 인코드된 데이터를 인캡슐레이트하는 트랙을 획득하는 단계를 포함하고,
- [0043] 상기 방법은 주어진 트랙 내에 인캡슐레이트된 인코드된 데이터를 적어도 디스플레이하기 위해 클라이언트가 다른 인코드된 데이터를 패스할 필요가 있다는 것을 시그널하기 위한 정보를 획득하는 단계를 포함한다.
- [0044] 본 발명의 또 하나의 실시예에 따르면, 장면의 와이드 뷰의 투사에 대응하는 인코드된 미디어 데이터를 서버로부터 클라이언트에 전송하는 디바이스가 제공되고, 상기 디바이스는
- [0045] 인코드된 데이터를 트랙들 내로 인캡슐레이트하는 수단을 포함하고,
- [0046] 상기 디바이스는 주어진 트랙 내에 인캡슐레이트된 인코드된 데이터를 적어도 디스플레이하기 위해 클라이언트가 다른 인코드된 데이터를 패스할 필요가 있다는 것을 시그널하기 위한 정보를 제공하는 수단을 추가로 포함한다.
- [0047] 본 발명의 또 하나의 실시예에 따르면, 클라이언트에 의해 미디어 파일들을 패스하는 디바이스가 제공되고, 상기 미디어 파일들은 장면의 와이드 뷰의 투사에 대응하는 인코드된 미디어 데이터를 포함하고, 상기 디바이스는
- [0048] 미디어 파일들로부터, 인코드된 데이터를 인캡슐레이트하는 트랙을 획득하는 수단을 포함하고,
- [0049] 상기 디바이스는 주어진 트랙 내에 인캡슐레이트된 인코드된 데이터를 적어도 디스플레이하기 위해 클라이언트가 다른 인코드된 데이터를 패스할 필요가 있다는 것을 시그널하기 위한 정보를 획득하는 수단을 추가로 포함한다.
- [0050] 본 발명의 또 하나의 실시예에 따르면, 장면의 와이드 뷰의 투사에 대응하는 인코드된 미디어 데이터를 서버로부터 클라이언트에 전송하는 방법이 제공되고, 상기 방법은
- [0051] 인코드된 데이터를 트랙들 내로 인캡슐레이트하는 단계,
- [0052] 각각의 트랙에 대한 디스크립티브 메타데이터를 획득하는 단계 - 디스크립티브 메타데이터는 트랙 컨테이너들 내로 구성되고, 트랙 컨테이너는 트랙에 관련됨 - , 및
- [0053] 클라이언트에 트랙들을 전송하는 단계를 포함하고,
- [0054] 획득하는 단계는
- [0055] 제1 트랙 내에 인캡슐레이트된 인코드된 데이터 및 제1 트랙 컨테이너의 제1 박스에서 정의된 공통 세트에 속하는 하나 이상의 트랙 내에 인캡슐레이트된 인코드된 데이터에 의해 나타내진 와이드 뷰의 커버리지를 표시하는 정보를 포함하는 제1 구조를 제공하는 단계를 추가로 포함한다.
- [0056] 본 발명의 상기 다른 실시예들은 미디어 데이터로서 특히 다수-트랙 OMAF 콘텐트에 대한, 와이드 뷰의 커버리지를 표시하는 정보에 대한 더 간단한 시그널링을 채택하는 것을 가능하게 한다.
- [0057] 본 발명의 또 하나의 실시예에 따르면, 클라이언트에 의해 미디어 파일들을 패스하는 방법이 제공되고, 상기 미디어 파일들은 장면의 와이드 뷰의 투사에 대응하는 인코드된 미디어 데이터에 대응하고, 상기 방법은
- [0058] 미디어 파일들로부터, 인코드된 데이터를 인캡슐레이트하는 트랙들을 얻는 단계,
- [0059] 각각의 트랙으로부터 디스크립티브 메타데이터를 획득하는 단계 - 디스크립티브 메타데이터는 트랙 컨테이너들 내로 구성되고, 트랙 컨테이너는 트랙에 관련됨 - 를 포함하고,
- [0060] 획득하는 단계는
- [0061] 제1 트랙 내에 인캡슐레이트된 인코드된 데이터 및 제1 트랙 컨테이너의 제1 박스에서 정의된 공통 세트에 속하는 하나 이상의 트랙 내에 인캡슐레이트된 인코드된 데이터에 의해 나타내진 와이드 뷰의 커버리지를 표시하는 정보를 포함하는 제1 구조를 획득하는 단계를 추가로 포함한다.
- [0062] 본 발명의 또 하나의 실시예에 따르면, 장면의 와이드 뷰의 투사에 대응하는 인코드된 미디어 데이터를 서버로부터 클라이언트에 전송하는 디바이스가 제공되고, 상기 디바이스는

- [0063] 인코드된 데이터를 트랙들 내로 인캡슐레이트하는 수단,
- [0064] 각각의 트랙에 대한 디스크립티브 메타데이터를 획득하는 수단 - 디스크립티브 메타데이터는 트랙 컨테이너들 내로 구성되고, 트랙 컨테이너는 트랙에 관련됨 - , 및
- [0065] 클라이언트에 트랙들을 전송하는 수단을 포함하고,
- [0066] 획득하는 수단은
- [0067] 제1 트랙 내에 인캡슐레이트된 인코드된 데이터 및 제1 트랙 컨테이너의 제1 박스에서 정의된 공통 세트에 속하는 하나 이상의 트랙 내에 인캡슐레이트된 인코드된 데이터에 의해 나타내진 와이드 뷰의 커버리지를 표시하는 정보를 포함하는 제1 구조를 제공하는 수단을 포함한다.
- [0068] 본 발명의 또 하나의 실시예에 따르면, 클라이언트에 의해 미디어 파일들을 파스하는 디바이스가 제공되고, 상기 미디어 파일들은 장면의 와이드 뷰의 투사에 대응하는 인코드된 미디어 데이터에 대응하고, 상기 디바이스는 미디어 파일들로부터 트랙들을 얻는 수단,
- [0069] 각각의 트랙으로부터 디스크립티브 메타데이터를 획득하는 수단 - 디스크립티브 메타데이터는 트랙 컨테이너들 내로 구성되고, 트랙 컨테이너는 트랙에 관련됨 - 을 포함하고,
- [0070] 획득하는 수단은
- [0071] 제1 트랙 내에 인캡슐레이트된 인코드된 데이터 및 제1 트랙 컨테이너의 제1 박스에서 정의된 공통 세트에 속하는 하나 이상의 트랙 내에 인캡슐레이트된 인코드된 데이터에 의해 나타내진 와이드 뷰의 커버리지를 표시하는 정보를 포함하는 제1 구조를 획득하는 다른 수단을 포함한다.
- [0072] 본 발명의 또 하나의 실시예에 따르면, 장면의 와이드 뷰의 투사에 대응하는 인코드된 비디오 데이터를 서버로부터 클라이언트에 전송하는 방법이 제공되고, 상기 방법은
- [0073] 서로 관련된 적어도 제1 및 제2 트랙 내의 스테레오스코픽 비디오 콘텐트에 대응하는 인코드된 비디오 데이터를 인캡슐레이트하는 단계,
- [0074] 각각의 트랙에 대한 디스크립티브 메타데이터를 획득하는 단계 - 디스크립티브 메타데이터는 트랙 컨테이너들 내로 구성되고, 트랙 컨테이너는 트랙에 관련됨 - 를 포함하고,
- [0075] 상기 방법은
- [0076] 트랙 컨테이너 당 하나의 박스에서, 고려된 트랙 내에 각각 인캡슐레이트된 스테레오 비디오 콘텐트에 관한 정보를 제1 트랙 및 제2 트랙 내에 제공하는 단계를 추가로 포함한다.
- [0077] 본 발명의 이 다른 실시예는 특히 ISOBMFF로부터의 기존의 박스를 재사용할 때 그리고 중복된 시그널링을 피하는, 매우 간단한 시그널링을 유지하는 것을 가능하게 한다.
- [0078] 본 발명의 또 하나의 실시예에 따르면, 클라이언트에 의해 미디어 파일들을 파스하는 방법이 제공되고, 상기 미디어 파일들은 장면의 와이드 뷰의 투사에 대응하는 인코드된 비디오 데이터에 대응하고, 상기 방법은
- [0079] 미디어 파일들로부터, 서로 관련된 적어도 제1 및 제2 트랙으로부터의 스테레오스코픽 비디오 콘텐트에 대응하는 인코드된 비디오 데이터를 얻는 단계,
- [0080] 각각의 트랙에 대한 디스크립티브 메타데이터를 획득하는 단계 - 디스크립티브 메타데이터는 트랙 컨테이너들 내로 구성되고, 트랙 컨테이너는 트랙에 관련됨 - 를 포함하고,
- [0081] 상기 방법은
- [0082] 트랙 컨테이너 당 하나의 박스에서, 고려된 트랙 내에 각각 인캡슐레이트된 스테레오 비디오 콘텐트에 관한 정보를 제1 트랙 및 제2 트랙에서 획득하는 단계를 추가로 포함한다.
- [0083] 본 발명의 또 하나의 실시예에 따르면, 장면의 와이드 뷰의 투사에 대응하는 인코드된 비디오 데이터를 서버로부터 클라이언트에 전송하는 디바이스가 제공되고, 상기 디바이스는
- [0084] 서로 관련된 적어도 제1 및 제2 트랙 내의 스테레오스코픽 비디오 콘텐트에 대응하는 인코드된 비디오 데이터를 인캡슐레이트하는 수단,

- [0086] 각각의 트랙에 대한 디스크립티브 메타데이터를 획득하는 수단 - 디스크립티브 메타데이터는 트랙 컨테이너들 내로 구성되고, 트랙 컨테이너는 트랙에 관련됨 - 을 포함하고,
- [0087] 상기 디바이스는
- [0088] 트랙 컨테이너 당 하나의 박스에서, 고려된 트랙 내에 각각 인캡슐레이트된 스테레오 비디오 콘텐트에 관한 정보를 제1 트랙 및 제2 트랙 내에 제공하는 수단을 추가로 포함한다.
- [0089] 본 발명의 또 하나의 실시예에 따르면, 클라이언트에 의해 미디어 파일들을 파스하는 디바이스가 제공되고, 상기 미디어 파일들은 장면의 와이드 뷰의 투사에 대응하는 인코드된 비디오 데이터에 대응하고, 상기 디바이스는
- [0090] 미디어 파일들로부터, 서로 관련된 적어도 제1 및 제2 트랙으로부터의 스테레오스코픽 비디오 콘텐트에 대응하는 인코드된 비디오 데이터를 얻는 수단,
- [0091] 각각의 트랙에 대한 디스크립티브 메타데이터를 획득하는 수단 - 디스크립티브 메타데이터는 트랙 컨테이너들 내로 구성되고, 트랙 컨테이너는 트랙에 관련됨 - 을 포함하고,
- [0092] 상기 디바이스는
- [0093] 트랙 컨테이너 당 하나의 박스에서, 고려된 트랙 내에 각각 인캡슐레이트된 스테레오 비디오 콘텐트에 관한 정보를 제1 트랙 및 제2 트랙에서 획득하는 수단을 추가로 포함한다.

도면의 간단한 설명

- [0094] 본 발명의 추가 장점들이 도면들 및 상세한 설명의 검토 시에 본 기술 분야의 통상의 기술자들에게 분명해질 것이다. 임의의 추가적인 장점들이 여기에 포함되는 것으로 의도된다.
- 본 발명의 실시예들이 단지 예들로서, 그리고 다음의 도면들을 참조하여 아래에 설명된다.
- 도 1은 전방향 비디오를 캡처하고, 처리하고, 인캡슐레이트하고, 서버로부터 클라이언트에 전송하고, 렌더링하는 데이터 흐름의 예를 도시한다.
- 도 2는 본 발명의 실시예들에 따른 인캡슐레이션의 예를 도시한 블록도를 도시하고;
- 도 3은 본 발명의 하나 이상의 실시예의 구현을 위한 컴퓨팅 디바이스의 개략 블록도이다.

발명을 실시하기 위한 구체적인 내용

- [0095] 특정한 실시예들에 따르면, 패킹된 영상(131)의 인코딩(도 1의 단계 140)으로부터 생긴 인코드된 비트스트림들 및 서브-비트스트림들은 인캡슐레이션 파일 포맷에 따라 예를 들어, ISO 베이스 미디어 파일 포맷(ISO/IEC 14496-12 및 ISO/IEC 14496-15), 전방향 미디어 포맷(OMAF)(ISO/IEC 23090-2) 및 MPEG 표준화 기구에 의해 정의된 것과 같은 연관된 규격들에 따라 파일 내로 또는 작은 시간적 세그먼트 파일들 내로 인캡슐레이트된다.
- [0096] 인코드된 비트스트림(예를 들어, HEVC) 및 가능하게는 그것의 서브-비트스트림들(예를 들어, 타일된 HEVC, MV-HEVC, 스케일가능한 HEVC)이 하나의 단일 트랙으로서 인캡슐레이트될 수 있다. 대안적으로 공간적으로 관련된 (즉 투사된 영상의 서브-공간적 부분들인) 다수의 인코드된 비트스트림이 여러 개의 서브-화상 트랙들로서 인캡슐레이트될 수 있다. 대안적으로, 여러 개의 서브-비트스트림들(타일들, 뷰들, 계층들)을 포함하는 인코드된 비트스트림(예를 들어, 타일된 HEVC, MV-HEVC, 스케일가능한 HEVC)이 다수의 서브-화상 트랙으로서 인캡슐레이트될 수 있다.
- [0097] 서브-화상 트랙은 화상 또는 영상의 서브 부분에 대한 데이터를 매립하는 트랙이다. 서브-화상 트랙은 다른 서브-화상 트랙들 또는 그로부터 서브-화상이 추출되는 풀 화상을 디스크라이브하는 트랙에 관련될 수 있다. 예를 들어 서브-화상 트랙은 타일 트랙일 수 있다. 그것은 AVC 트랙, HEVC 트랙, HEVC 타일 트랙 또는 샘플들의 시퀀스로서 인캡슐레이트된 임의의 압축된 비디오 비트스트림에 의해 나타내질 수 있다.
- [0098] 타일 트랙은 영상의 공간적 부분 또는 영상 또는 화상의 서브-화상에 대응하는 타이밍된 비디오 샘플들의 시퀀스이다. 그것은 예를 들어 영상 내의 관심있는 영역 또는 영상 내의 임의의 영역일 수 있다. 타일 트랙에 대응하는 데이터는 비디오 비트스트림으로부터 올 수 있거나 비디오 비트스트림의 서브 부분으로부터 올 수 있다. 예를 들어 타일 트랙은 AVC 또는 HEVC 순응 비트스트림일 수 있거나 AVC 또는 HEVC의 서브-부분 또는 예를 들어 HEVC 타일들과 같은, 임의의 인코드된 비트스트림일 수 있다. 양호한 실시예에서 타일 트랙은 독립적으로 디코

드 가능하다(인코더는 다른 파일들로부터 움직임 예상이 제거하지 않도록 주의한다). 파일 트랙이 파일들로 HEVC에서 인코드된 비디오 비트스트림에 대응할 때, 그것은 ISO/IEC 14496-15 4차 편집에서 설명된 것과 같은 'hvt1' 트랙으로서 표시된 HEVC 파일 트랙 내로 인캡슐레이트될 수 있다. 그것은 다음에 파라미터 세트들을 획득하기 위한 파일 베이스 트랙, 비디오 디코더를 셋업하기 위한 하이 레벨 정보를 참조한다. 그것은 또한 HEVC 트랙 'hvc1' 또는 'hev1' 트랙 내로 인캡슐레이트될 수 있다. 파일 트랙은 서브-화상들의 더 큰 영상 또는 화상 내로의 공간적 구성을 위해 사용될 수 있다.

[0099] 파일 베이스 트랙은 이를 하나 이상의 트랙 간에 공유되는 데이터 또는 메타데이터를 포함하는 하나 이상의 파일 트랙에 공통인 트랙이다. 파일 베이스 트랙은 하나 이상의 파일 트랙으로부터 영상들을 구성하라는 명령들을 포함할 수 있다. 파일 트랙들은 완전한 디코딩 또는 렌더링을 위한 파일 베이스 트랙에 의존할 수 있다. 파일 베이스 트랙이 파일들로 HEVC에서 인코드된 비디오 비트스트림으로부터 유도할 때, 그것은 'hvc2' 또는 'hev2'로서 표시된 HEVC 트랙 내로 인캡슐레이트된다. 또한 트랙 참조 "tbas"를 통해 HEVC 파일 트랙들에 의해 참조되고 그것은 ISO/IEC 14496-15 4차 편집에서 설명된 것과 같은 HEVC 파일 트랙들에의 'sabt' 트랙 참조를 사용하여 파일 정렬을 표시할 것이다.

[0100] (또한 참조 트랙이라고 표시된) 합성 트랙은 영상을 구성하기 위해 다른 트랙들을 참조하는 트랙이다. 합성 트랙의 한 예는 비디오 트랙들의 경우에, 서브-화상 트랙들을 더 큰 영상으로 구성하는 트랙이다. 이것은 예를 들어 각각의 비디오 트랙로부터의 영상들을 더 큰 영상으로 구성하기 위해 변환 또는 변환 파라미터들을 제공하는 비디오 트랙들로부터 유도한 트랙에서, 포스트-디코딩에 의해 이루어질 수 있다. 합성 트랙은 또한 서브-비트스트림 연쇄로부터 생긴 비트스트림을 디코드하기 전에 형성할 다른 비디오 트랙들 또는 파일 트랙들로부터 NAL 유닛들을 추출하라는 명령들을 제공하는 추출기 NAL 유닛들을 갖는 트랙일 수 있다. 합성 트랙은 또한 예를 들어, 다른 트랙들에의 트랙 참조들을 통해, 구성 명령들을 암시적으로 제공하는 트랙일 수 있다.

[0101] ISO/IEC 14496-12는 트랙들의 그룹들을 디스크라이브하기 위해 트랙 레벨에 위치한 박스를 제공하고, 여기서 각각의 그룹은 특정한 특성을 공유하거나 또는 그룹 내의 트랙들은 특정한 관계를 갖지 않는다. 특정한 특성 또는 관계는 포함된 박스들의 박스 타입(track_group_type)에 의해 표시된다. 포함된 박스들은 동일한 트랙 그룹에 속하는 트랙들을 결정하기 위해 사용될 수 있는 식별자(track_group_id)를 포함한다. 동일한 track_group_type 및 track_group_id 값들을 갖는 트랙 그룹 박스를 갖는 모든 트랙들은 동일한 트랙 그룹의 부분이다. THE MPEG OMAF 표준은 타입 'spco'의 TrackGroupTypeBox로서 공간적 구성을 위한 특정한 트랙 그룹을 제안하고 있다. 'spco' 트랙 그룹 내의 각각의 트랙의 샘플들은 더 큰 영상을 발생하기 위해 이 동일한 그룹 내의 다른 트랙들로부터(동일한 구성 또는 디코딩 시간에서)의 샘플들로 공간적으로 구성될 수 있다.

[0102] 패킹된 영상(131)의 인코딩(도 1의 단계 140)으로부터 생긴 인코드된 비트스트림들 및 서브-비트스트림들에 따라, 파일 포맷에서의 인캡슐레이션의 몇 가지 변형들이 가능하다.

[0103] 도 2는 본 발명의 실시예에 따른 파일/세그먼트 인캡슐레이션(도 1의 단계 150)의 예를 도시한 블록도를 도시한다.

[0104] 단계 200에서, 서버는 트랙들 내로 인캡슐레이트될 인코드된 미디어 데이터를 획득한다. 서버는 여러 개의 공간적-관련 비디오 비트스트림들이 있는지(즉 패킹된 영상들의 공간적 서브-부분을 나타내고 그를 위해 공간적 구성이 더 큰 영상을 생성할 수 있는지) 또는 움직임-제한된 파일들 또는 다수의 서브-화상 트랙으로서 클라이언트에게 노출될 수 있는 다수의 뷰를 나타내는 비디오 서브-비트스트림들을 포함하는 비디오 비트스트림들이 있는지를 결정한다. 그것이 단일 비디오 비트스트림으로서 인코드되거나 콘텐트 창작자가 다수의 트랙으로서 인코드된 패킹된 영상을 노출하기를 원하지 않기 때문에 인코드된 패킹된 영상이 다수의 트랙으로서 노출될 수 없으면, 비디오 비트스트림 또는 비디오 서브-비트스트림들은 하나의 단일 트랙 내로 인캡슐레이트된다(단계 210). 그렇지 않으면, 단계 220에서 인캡슐레이트될 미디어 콘텐트가 움직임-제한된 파일들을 나타내는 비디오 서브-비트스트림들로 구성되는지가 결정된다. 예이면, 적어도 하나의 합성 트랙이 여러 개의 파일 트랙들의 적어도 하나의 구성을 나타내기 위해 제공될 필요가 있을 수 있다. 구성은 폴 패킹된 영상들 또는 폴 패킹된 영상들의 서브-부분만을 나타낼 수 있다. 파일 트랙들을 갖는 합성 트랙을 사용하면 클라이언트 측 상의 스트림들의 별도의 렌더링 또는 디코딩의 필요가 피해진다. 클라이언트에게 노출될 가능한 조합들의 수는 콘텐트 창작자의 선택들에 의존한다. 예를 들어, 콘텐트 창작자는 현재의 사용자의 뷰포트에 따라 파일들을 상이한 시각적 품질들과 조합하기를 원할 수 있다. 이를 위해, 그것은 패킹된 영상을 상이한 시각적 품질들로 여러 번 인코드하고 시각적 품질들의 면에서 파일들의 상이한 조합을 포함하는 폴 패킹된 영상을 나타내는 여러 개의 합성 트랙들을 제안할 수 있다. 사용자의 뷰포트에 따라 상이한 품질들로 파일들을 조합함으로써, 콘텐트 창작자는

네트워크 리소스들의 소비를 감소시킬 수 있다.

[0105] 단계 220에서, 합성 트랙들이 제공되어야 한다고 결정되면, 암시적 재구성이 합성 트랙을 위해 사용될 수 있는지 여부가 결정된다(단계 240).

[0106] 예를 들어, 암시적 재구성은 예를 들어 ISO/IEC 14496-15 4차 편집에서 정의된 것과 같은, 타일 베이스 및 타일 트랙들로부터의 비트스트림 재구성이라고 한다. 그들이 타일 트랙들의 샘플 내에서 참조하는 데이터로 합성 트랙의 샘플들 내의 추출기들을 대체함으로써 타일 트랙들의 샘플들로부터 합성 트랙의 샘플들을 리빌드하기 위해 추출기들과 같은 인-스트림 구조를 사용하기 보다는, 암시적 재구성은 트랙 참조들(예를 들어, HEVC 암시적 재구성 내의 'sabt' 트랙 참조들)의 순서로 합성 트랙 및 타일 트랙들의 샘플들을 연쇄시킴으로써 합성 트랙의 샘플들을 리빌드하는 것을 가능하게 한다.

[0107] 암시적 재구성의 사용은 사용의 시나리오에 의존한다. 여러 개의 타일 트랙들의 구성이 인코딩에서의 타일들의 순서와 비교하여 디코딩에서의 타일들의 재배열을 요구할 때, 일부 슬라이스 어드레스들이 재기입되어야 한다. 이러한 경우에, 암시적 재구성이 가능하지 않고 추출기들로의 명시적 재구성이 선택되어야 한다.

[0108] 암시적 재구성이 가능하다면, 타일 베이스 트랙이 발생되고(단계 241), 비디오 서브-비트스트림들이 (예를 들어, HEVC 'hvt1' 트랙들로서) 독립적으로 디코드 가능하지 않은 타일 트랙들로서 인캡슐레이트된다.

[0109] 그렇지 않으면 추출기 트랙이 발생되고(단계 242), 비디오 서브-비트스트림들이 (예를 들어, HEVC 'hvc1' 또는 'hev1' 트랙들로서) 독립적으로 디코드 가능한 타일 트랙들로서 인캡슐레이트된다.

[0110] 다시 단계 220으로 가서, 미디어 콘텐트가 타일 서브-비트스트림들을 포함하지 않거나 콘텐트 창작자가 합성 트랙들을 생성하고 노출하는 것을 원하지 않으면, 공간적-관련 비디오 비트스트림들 또는 비디오 서브-비트스트림들(예를 들어, 타일 또는 다수의 뷰)은 별개의 서브-화상 트랙들 내로 인캡슐레이트된다(단계 230). 이러한 특정한 경우에, 타일 서브-비트스트림들이 HEVC 타일들이면, 그들은 HEVC 트랙 'hvc1' 또는 'hev1' 트랙으로서 인캡슐레이트된다.

[0111] 단계 250에서, 공간적 구성을 위한 시그널링이 공간적-관련 비디오 비트스트림들 또는 비디오 서브-비트스트림들과 함께 그룹에 추가된다. 공간적 구성 시그널링은 MPEG OMAF에서 정의된 것과 같고 아래에 예시되는 것과 동일한 그룹에 관련한 모든 트랙들에 대해 동일한 track_group_id를 갖는 그룹, 예를 들어 타입 'spco'의 트랙 그룹을 구성하는 각각의 트랙(서브-화상 트랙들, 타일 트랙들, 합성 트랙들) 내에 특정한 TrackGroupTypeBox를 정의함으로써 제공될 수 있다:

```

[0112] aligned (8) class TrackGroupTypeBox (unsigned int(32) track_group_type)
[0113] extends FullBox (track_group_type, version = 0, flags = 0)
[0114] {
[0115]     unsigned int(32) track_group_id;
[0116]     // the remaining data may be specified for a particular
[0117]     track_group_type
[0118] }
[0119] aligned (8) class SubPictureCompositionBox extends
[0120] TrackGroupTypeBox ('spco') {
[0121]     unsigned int(16) track_x;
[0122]     unsigned int(16) track_y;
[0123]     unsigned int(16) track_width;
[0124]     unsigned int(16) track_height;
[0125]     unsigned int(16) composition_width;
[0126]     unsigned int(16) composition_height;

```

}

[0128] 이 트랙 그룹 박스는 구성 내의 트랙의 상대적 2차원 좌표들 및 구성에 의해 형성된 영상의 전체 크기를 제공할 것이다. 구성은 전체 패킹된 영상들 또는 패킹된 영상들의 서브-부분만을 나타낼 수 있다. 예를 들어, 콘텐트 창작자는 전체 패킹된 영상들 또는 패킹된 영상들의 서브-부분만을 빌드하는 것을 가능하게 하는 다수의 합성 트랙을 노출하기를 원할 수 있다.

[0129] 대안으로서, SubPictureCompositionBox('spco')는 다음과 같이 구성 화상의 폭 및 높이를 나타내는 파라미터 composition_width 및 composition_height를 단지 정의할 수 있다:

[0130] aligned (8) class SubPictureCompositionBox extends

[0131] TrackGroupTypeBox ('spco') {

[0132] unsigned int(16) composition_width;

[0133] unsigned int(16) composition_height;

[0134] }

[0135] 그리고 구성 내의 트랙의 2차원 좌표들은 다음과 같이 VisualSampleEntry에서 정의된 새로운 fullBox로서 주어질 수 있고:

[0136] aligned (8) class 2DCoordinateInTrackGroupBox extends FullBox ('2dco',

[0137] version, flags)

[0138] {

[0139] unsigned int(32) track_group_id;

[0140] unsigned int(32) track_subgroup_id;

[0141] unsigned int(16) track_x;

[0142] unsigned int(16) track_y;

[0143] unsigned int(16) track_width;

[0144] unsigned int(16) track_height;

[0145] }

[0146] 또는 다음과 같이 새로운 포괄적 샘플 그룹 디스크립션 엔트리로서 주어질 수 있다:

[0147] class 2DCoordinateInTrackGroupBox extends

[0148] VisualSampleGroupEntry ('2dco')

[0149] {

[0150] unsigned int(32) track_group_id;

[0151] unsigned int(32) track_subgroup_id;

[0152] unsigned int(16) track_x;

[0153] unsigned int(16) track_y;

[0154] unsigned int(16) track_width;

[0155] unsigned int(16) track_height;

[0156] }

[0157] track_group_id는 연관된 트랙 그룹의 식별자를 제공한다. 그러므로 트랙은 각각의 트랙 그룹 내의 상이한 위치들에서의 다수의 트랙 그룹에 연관될 수 있다. track_subgroup_id는 서브그룹 식별자를 제공한다. 트랙 그룹 내에 track_subgroup_id를 갖는 모든 트랙은 동일한 트랙 서브-그룹에 관련한다.

- [0158] `track_x`, `track_y`는 구성 내의 이 트랙의 샘플들의 상좌측 모서리의 수평 및 수직 위치를 제공한다.
- [0159] `track_width`, `track_height`는 구성 내의 이 트랙의 샘플들의 폭 및 높이를 제공한다.
- [0160] 그들 파라미터는 그들 트랙을 나타내는 적응 세트들의 공간적 관계를 디스크라이브하기 위해 DASH 매니페스트에서 사용될 수 있는 (ISO/IEC 23009-1 3차 편집에서 정의된) DASH 공간적-관계 디스크립션(SRD) 디스크립터의 파라미터들을 직접 매치한다.
- [0161] `track_group_id`는 DASH SRD `source_id` 파라미터를 매치할 것이다.
- [0162] `track_subgroup_id`는 DASH SRD `spatial_set_id` 파라미터를 매치할 것이다.
- [0163] `track_x`, `track_y`, `track_width`, `track_height`는 각각 DASH SRD 파라미터들 `object_x`, `object_y`, `object_width`, `object_height` 파라미터들을 매치할 것이다.
- [0164] 마지막으로, (`track_group_id`를 통해) 연관된 트랙 그룹으로부터의 `composition_width` 및 `composition_height`는 DASH SRD `total_width`, `total_height` 파라미터들을 매치할 것이다.
- [0165] 대안으로서, 합성 트랙이 있는 경우에, 공간적 구성 시그널링이 이 합성 트랙에 의해 암시적으로 제공될 수 있다. 실제로, 합성 트랙이 타일 베이스 트랙인 경우에, 타일 베이스 트랙은 타입 'sabt'의 트랙 참조를 통해 타일 트랙들의 세트를 참조한다. 이 타일 베이스 트랙 및 타일 트랙들의 세트는 구성 그룹을 형성한다. 유사하게, 합성 트랙이 추출기 트랙이면, 추출기 트랙은 타입 'scal'의 트랙 참조를 통해 타일 트랙들의 세트를 참조한다. 이 추출기 트랙 및 타일 트랙들의 세트는 또한 구성 그룹을 형성한다. 양자의 경우에, 구성 내의 각각의 타일 트랙의 상대적 2차원 좌표들은 ISO/IEC 14496-15 4차 편집에서 정의된 것과 같은 타입 'trif'의 샘플 그룹핑 또는 디폴트 샘플 그룹핑을 정의함으로써 제공될 수 있다.
- [0166] 또 하나의 대안으로서, 공간적 구성 시그널링은 새로운 엔티티 그룹을 정의함으로써 제공될 수 있다. 엔티티 그룹은 아이템들 또는 트랙들의 그룹핑이다. 엔티티 그룹들은 MetaBox에서 GroupsListBox 내에 표시된다. 트랙들을 참조하는 엔티티 그룹들은 file-level MetaBox의 GroupsListBox에서 또는 movie-level MetaBox의 GroupsListBox에서 지정될 수 있다. GroupListBox ('grp1')는 각각 EntityToGroupBox라고 하는 풀 박스들의 세트를 포함하고, 연관된 4-문자 코드들은 정의된 그룹핑 타입을 표시한다. EntityToGroupBox는 다음과 같이 정의된다:
- ```

[0167] aligned (8) class EntityToGroupBox (grouping_type, version, flags)
[0168] extends FullBox (grouping_type, version, flags) {
[0169] unsigned int(32) group_id;
[0170] unsigned int(32) num_entities_in_group;
[0171] for (i=0; i<num_entities_in_group; i++)
[0172] unsigned int(32) entity_id;
[0173] // the remaining data may be specified for a particular
[0174] grouping_type
[0175] }

```
- [0176] 전형적으로 `group_id`는 그룹의 id를 제공하고 `entity_id`의 세트는 엔티티 그룹에 관련한 트랙들의 `track_ID`를 제공한다. `entity_id`의 세트 이후에, 특정한 `grouping_type`을 위한 추가적인 데이터를 정의함으로써 EntityToGroupBox의 정의를 확장하는 것이 가능하다. 실시예에 따라, 예를 들어 (엔티티 그룹 구성에 대해) 'egco'와 동일한 `grouping_type`을 갖는 새로운 EntityToGroupBox는 공간적-관련 비디오 비트스트림들 또는 비디오 서브-비트스트림들의 구성을 디스크라이브하도록 정의될 수 있다. `entity_id`의 세트는 그룹을 구성하는 트랙들(서브-화상들, 타일 트랙들, 합성 트랙들)의 `track_ID`의 세트를 포함할 것이다. 구성에 의해 형성된 영상의 전체 크기는 이 새로운 `grouping_type` 'egco'에 연관된 추가적인 데이터의 부분으로서 제공될 수 있다.
- [0177] EntityToGroupBox('egco')는 다음과 같이 정의될 것이다:
- ```

[0178] aligned (8) class EntityToGroupBox ('egco', version, flags)

```

```
[0179]     extends FullBox ('egco', version, flags) {
[0180]         unsigned int(32) group_id;
[0181]         unsigned int(32) num_entities_in_group;
[0182]         for (i=0; i<num_entities_in_group; i++)
[0183]             unsigned int(32) entity_id;
[0184]             unsigned int(16) composition_width;
[0185]             unsigned int(16) composition_height;
[0186]     }
```

[0187] 타입 'egco'의 엔티티 그룹핑에 의해 정의된 구성 내의 각각의 트랙의 상대적 2차원 좌표들은 ISO/IEC 14496-15 4차 편집에서 정의된 것과 같은 각각의 파일 트랙 내에 타입 'trif'의 샘플 그룹핑 또는 디폴트 샘플 그룹핑을 정의함으로써 제공될 수 있다. 대안으로서, 상대적 2차원 좌표들은 그룹에 관련한 각각의 파일 트랙 내의 VisualSampleEntry에 위치되는 새로운 포괄적 풀 박스 2DCoordinateForEntityGroupBox('2dco')로서 정의될 수 있다.

```
[0188] aligned(8) class 2DCoordinateForEntityGroupBox extends FullBox ('2dco',
```

```
[0189] version, flags)
```

```
[0190] {
```

```
[0191]     unsigned int(32) entity_group_id;
```

```
[0192]     unsigned int(16) track_x;
```

```
[0193]     unsigned int(16) track_y;
```

```
[0194]     unsigned int(16) track_width;
```

```
[0195]     unsigned int(16) track_height;
```

```
[0196] }
```

[0197] entity_group_id는 그룹을 정의하는 연관된 EntityToGroupBox('egco')의 식별자를 제공한다.

[0198] track_x, track_y는 구성 내의 이 트랙의 샘플들의 상좌측 모서리의 수평 및 수직 위치를 제공한다.

[0199] track_width, track_height는 구성 내의 이 트랙의 샘플들의 폭 및 높이를 제공한다.

[0200] 대안으로서, 이 새로운 포괄적 박스 2DCoordinateForEntityGroupBox('2dco')는 다음과 같이 새로운 샘플링 그룹핑으로서 정의될 수 있다:

```
[0201] class 2DCoordinateForEntityGroupBox extends
```

```
[0202] VisualSampleGroupEntry ('2dco')
```

```
[0203] {
```

```
[0204]     unsigned int(32) entity_group_id;
```

```
[0205]     unsigned int(16) track_x;
```

```
[0206]     unsigned int(16) track_y;
```

```
[0207]     unsigned int(16) track_width;
```

```
[0208]     unsigned int(16) track_height;
```

```
[0209] }
```

[0210] 단계 260에서, 주어진 트랙 내에 인캡슐레이트된 인코드된 미디어 데이터를 디코드할 데이터의 필요성에 관련한 정보가 획득된다. 이 정보에 기초한 추가적인 시그널링은 트랙들에 연관된다. 트랙이 프리젠테이션을 위해 총

분하지 않거나 단독으로 프리젠티될 의도가 없으면 시그널링이 클라이언트에게 알리기 위해 추가된다.

[0211] 실제로 일부 트랙들은 단독으로 디코드될 수 없는 부분적 비트스트림들을 단지 포함할 수 있다. 예를 들어, 이 것은 그들의 연관된 타일 베이스 트랙 없이는 디코드될 수 없는 타입 'hvt1'의 일부 타일 트랙들에 대한 경우일 수 있다.

[0212] 또한 콘텐트 창작자는 클라이언트에게 일부 트랙들이 단독으로 프리젠티될 의도가 없고 미디어 프리젠테이션 내에 엔트리 포인트를 구성하지 않는다는 것을 알리기를 원할 수 있다.

[0213] 인코드된 미디어 데이터 및 주어진 트랙 내에 인캡슐레이트된 인코드된 미디어 데이터를 디코드할 데이터의 필요성에 관련한 정보가 하나 이상의 미디어 파일 내에 인캡슐레이트된다. 실제로, ISOBMFF 파일이 다수의 비디오 트랙을 포함할 때, 주 트랙으로서 시그널된 이들 비디오 트랙 중 하나 이상을 갖게 되면 사용자에게 노출하거나 스트리밍 매니페스트에서 노출할 트랙들의 선택에 있어서 미디어 플레이어에게 도움이 된다. 주 트랙 시그널링은 미디어 파일 내로의 미디어 플레이어들을 위한 엔트리 포인트들을 제공한다. 실제로, 동일한 중요성의 레벨을 갖는 트랙들의 긴 리스트를 갖는 것 대신에, 일부는 주석이 달리거나 또는 더 많은 중요성을 갖고 일종의 주된 아이템으로서 플레이어들에 의해 우선순위로 처리되는 것으로 디스크라이브된다.

[0214] 제1 실시예에서, 트랙이 단독으로 디스플레이될 의도가 없다는 정보가 트랙 헤더 내에 시그널될 수 있다. 각각의 트랙은 일부 트랙의 특성들을 지정하는 트랙 헤더 박스 'tkhd'(의무적)를 갖는다. ISOBMFF 풀 박스로서, 이 트랙 헤더 박스는 박스에 연관된 특정한 시그널링을 위해 사용될 수 있는 24비트 플래그 파라미터를 갖는다. 미디어 트랙들에 대한 트랙 헤더의 플래그들의 값은 트랙이 프리젠테이션(예를 들어, Track_enabled, Track_in_movie, Track_in_preview)에서 어떻게 사용되는지에 관한 정보를 제공하기 위해 이미 사용되고 있다. 비디오 트랙이 "단독으로 프리젠티될 의도가 있는지" 여부를 표시하기 위해, 플래그들에 대한 새로운 특정한 값 "Track_non_displayable_alone"이 트랙 헤더 박스에서 정의될 수 있다. 이 새로운 플래그는 다음과 같이 정의된다:

[0215] "Track_non_displayable_alone": = 0x000010은 트랙이 단독으로 디스플레이될 의도가 없고 프리뷰를 위해 사용될 수 없다는 것을 표시한다(Track_in_preview 플래그들 값은 오버라이드함). 디폴트로 이 플래그들 값은 설정되지 않는다. 트랙 헤더 플래그들의 디폴트 값은 여전히 7(track_enabled 0x1, track_in_movie 0x2, track_in_preview 0x4)인 것으로 변화되지 않은 채로 남는다는 점에 주목될 수 있다.

[0216] 제2 실시예에서, 단독으로 디스플레이될 의도가 없는 트랙들은 보조 트랙들로서 정의될 수 있다. 보조 트랙은 비디오 트랙과 동일하게 코드되지만, MediaBox의 HandlerBox 내의 'vide' 대신에 핸들러 타입 값 'auxv'를 사용하고, 시각적으로 디스플레이될 의도가 없다. 대안으로서, 비디오를 위한 새로운 핸들러 타입들 'subv' 및 오디오를 위한 'suba'는 트랙이 각각 비디오 또는 오디오 트랙과 동일하지만 그들이 프리젠테이션을 위해 충분하지 않고 단독으로 프리젠티될 의도가 없다는 것을 시그널하도록 정의될 수 있다.

[0217] 제3 실시예에서, 트랙이 프리젠테이션을 위해 충분하지 않고 단독으로 프리젠티될 의도가 없다는 정보는 트랙 그룹 정보의 부분으로서 시그널될 수 있다. 실제로 서브-화상 구성 트랙 그룹(즉 track_group_type이 'spco'인 TrackGroupTypeBox 내의 track_group_id의 동일한 값을 갖는 트랙들)에 맵핑된 트랙들은 프리젠티될 수 있는 시각적 콘텐트를 총체적으로 나타낸다. 그러나 이 그룹핑에 맵핑된 각각의 개별적인 트랙은 다른 트랙들 없이 단독으로 프리젠티될 의도가 있거나 없을 수 있다. 간단한 대안은 트랙이 "단독으로 프리젠티될 의도가 있는지" 여부를 표시하는 'spco' 박스 내에 새로운 파라미터 "not_output_track"를 정의하는 것으로 구성된다.

[0218] aligned (8) class SubPictureCompositionBox extends

[0219] TrackGroupTypeBox ('spco') {

[0220] unsigned int(16) track_x;

[0221] unsigned int(16) track_y;

[0222] unsigned int(16) track_width;

[0223] unsigned int(16) track_height;

[0224] unsigned int(16) composition_width;

[0225] unsigned int(16) composition_height;

```

[0226]         unsigned int (8) not_output_track;
[0227]     }
[0228] 대안적 실시예에서, 이 파라미터는 단일 비트로 나타내질 수 있고, 다른 7개의 비트는 미래의 사용 또는 다음과
[0229] 대안적 실시예에서, 이 파라미터는 단일 비트로 나타내질 수 있고, 다른 7개의 비트는 미래의 사용 또는 다음과
[0230] 대안적 실시예에서, 이 파라미터는 단일 비트로 나타내질 수 있고, 다른 7개의 비트는 미래의 사용 또는 다음과
[0231] 대안적 실시예에서, 이 파라미터는 단일 비트로 나타내질 수 있고, 다른 7개의 비트는 미래의 사용 또는 다음과
[0232] 대안적 실시예에서, 이 파라미터는 단일 비트로 나타내질 수 있고, 다른 7개의 비트는 미래의 사용 또는 다음과
[0233] 대안적 실시예에서, 이 파라미터는 단일 비트로 나타내질 수 있고, 다른 7개의 비트는 미래의 사용 또는 다음과
[0234] 대안적 실시예에서, 이 파라미터는 단일 비트로 나타내질 수 있고, 다른 7개의 비트는 미래의 사용 또는 다음과
[0235] 대안적 실시예에서, 이 파라미터는 단일 비트로 나타내질 수 있고, 다른 7개의 비트는 미래의 사용 또는 다음과
[0236] 대안적 실시예에서, 이 파라미터는 단일 비트로 나타내질 수 있고, 다른 7개의 비트는 미래의 사용 또는 다음과
[0237] 대안적 실시예에서, 이 파라미터는 단일 비트로 나타내질 수 있고, 다른 7개의 비트는 미래의 사용 또는 다음과
[0238] 대안적 실시예에서, 이 파라미터는 단일 비트로 나타내질 수 있고, 다른 7개의 비트는 미래의 사용 또는 다음과
[0239] 대안적 실시예에서, 이 파라미터는 단일 비트로 나타내질 수 있고, 다른 7개의 비트는 미래의 사용 또는 다음과

```

1로 설정될 때 not_output_track은 트랙이 단독으로 디스플레이될 의도가 없다는 것을 표시한다. 디폴트로 그 것은 제로인 것으로 가정되고 트랙은 사용자에게의 프리젠테이션을 위해 선택가능하다. 역시, 트랙이 임의의 SubPictureCompositionBox를 포함하지 않을 때, 그것은 단독으로 디스플레이가능한 것으로 가정된다.

제4 실시예에서, 트랙이 프리젠테이션을 위해 충분하지 않고 단독으로 프리젠테이션될 의도가 없다는 정보는 다른 트랙 정보 또는 다른 트랙 정보의 조합으로부터 추론될 수 있다.

예를 들어, 트랙의 디스플레이가능한 상태는 트랙의 종속성을 제공하는 트랙 참조 박스('tref') 및 트랙 내의 샘플들에 대한 공통 정의를 제공하는 SampleEntry(비디오에 대한 VisualSampleEntry)에 기초하여 검출될 수 있다.

예를 들어, 트랙이 타입 'sabt'의 트랙 참조를 갖고, 그것이 타입 'tbas'의 트랙 참조를 갖는 일부 다른 트랙들에 의해 참조된다면 트랙은 파일 베이스 트랙으로서 마크되고 플레이가능한/선택가능한 트랙으로서 마크된다. 트랙 기준 'sabt'를 통해 이 트랙으로부터의 모든 참조된 트랙들은 그들이 타입 'hvt1'의 VisualSampleEntry를 갖고 비플레이가능한 트랙으로서 마크되면 파일 트랙('hvt1')으로서 마크될 수 있다. 대안적으로, 트랙이 타입 'tbas'의 트랙 참조 및 타입 'hvt1'의 VisualSampleEntry를 갖는다면, 트랙은 파일 트랙('hvt1')으로서 마크되고 비플레이가능한 트랙으로서 마크된다. 이 트랙으로부터의 모든 참조된 트랙들은 파일 베이스 트랙으로서 마크되고 비플레이가능한 트랙으로서 마크된다. 대안적으로, 트랙이 타입 'scal'의 트랙 참조를 갖는다면, 트랙은 추출기 트랙이고 플레이가능한 트랙으로서 마크된다. 이 트랙으로부터의 모든 참조된 트랙들은 가능하게 플레이가능한 파일 트랙('hvc1')으로서 마크된다. 디폴트로, 그들은 플레이가능한 트랙으로서 마크될 수 있다. 그러나 콘텐트 창작자가 그들을 디폴트로 비플레이가능한 트랙으로 마크하는 것을 선호할 수 있다. 또한, 트랙이 어떤 트랙 참조('tref') 박스도 포함하지 않고 트랙 그룹에 관련하지 않으면, SampleEntry가 체크되어야 한다. 트랙이 'hvc1' 또는 'hev1'로서 검출되면, 트랙은 적어도 플레이가능한 트랙으로서 마크된다.

제3 실시예의 대안으로서, 제5 실시예에서, 2차원 좌표들(track_x, track_y, track_width 및 track_weight)이 SubPictureCompositionBox('spco')에서 정의되지 않으면, 파라미터 not_output_track은 여전히 아래에 예시되는 것과 같이 SubPictureCompositionBox('spco')에서 정의될 수 있고:

```

[0245] aligned (8) class SubPictureCompositionBox extends
[0246] TrackGroupTypeBox ('spco') {
[0247]     unsigned int(16) composition_width;

```

[0248] unsigned int(16) composition_height;

[0249] unsigned int(1) not_output_track;

[0250] unsigned int(7) reserved;

[0251] }

[0252] 또는 파라미터 not_output_track은 2DCoordinateInTrackGroupBox()에서 또는 2DCoordinateForEntityGroupBox()에서 또는 VisualSampleEntry 또는 Sample Group Description 엔트리 레벨들에서 정의된 유사한 박스들에서 정의될 수 있다.

[0253] 여전히 단계 260에서, 유사하게 명시적으로 시그널링이 주 트랙들 또는 부분적 트랙들을 식별하기 위해 추가될 수 있다.

[0254] 미디어 파일 내의 주 트랙은 동일한 미디어 타입을 갖는 트랙들보다 또는 상이한 미디어 타입들을 갖는 관련된 트랙들보다 많은 중요성을 갖는 것으로 고려되는 트랙이다. 예를 들어, 주 비디오 트랙은 미디어 플레이어가 선택 및 플레이백을 위해 사용자에게 노출하여야 하는 트랙이다. 역시, 주 트랙은 미디어 파일이 스트리밍되거나 전송될 때 스트리밍 매니페스트에서 노출되어야 한다. 예를 들어, 서브-화상 또는 타일들의 공간적 구성의 경우에, 주 트랙은 합성 트랙일 수 있다. 또한, 공간적 구성의 경우에, 주 트랙은 (서브-화상 또는 타일 트랙들과 대조적으로) 폴 화상에 대응하는 비디오 트랙일 수 있다. 플레이어에 의해 렌더링될 트랙들의 세트에서, 주 트랙은 우선순위로 렌더링하여야 할 것일 수 있다. 전송 맥락에서, 주 트랙은 우선순위로 폐치하여야 할 것일 수 있다. 예를 들어 미디어 파일 내의 주 트랙은 주 미디어 소자로서 스트리밍 매니페스트에서 디스크라이브될 수 있다. 예를 들어, MPEG DASH 매니페스트에서 주 트랙은 사전 선택의 주 AdaptationSet, 또는 "주" 값 또는 그것이 주 트랙이라는 것을 표시하는 라벨을 갖는 역할 디스크립터를 갖는 AdaptationSet일 수 있다. 본 발명은 미디어 파일 내에 주 트랙을 시그널하는 상이한 방식들을 설명한다.

[0255] 미디어 파일 내의 부분적 트랙은 주 트랙과 조합하여 또는 주 트랙 또는 다른 부분적 트랙들과 조합하여 단지 처리될 수 있는 트랙이다. 타입 'hvt1'의 타일 트랙들은 부분적 트랙들의 예들이다. 그들은 타일 베이스 트랙과 조합하여 단지 처리될 수 있다.

[0256] 주/부분적 트랙 시그널링은 위에 설명된 "프리젠테이션을 위해 충분하지 않거나 단독으로 프리젠테이션될 의도가 없는" 시그널링과 유사하게 시그널될 수 있다. 그것은 트랙 헤더 플래그들 내의 명시적인 플래그 값 (예를 들어, "Is_Main_Track": = 0x000020)에 의해, 또는 아래에 예시되는 것과 같은 서브-화상 구성 트랙 그룹 ('spco') 박스 내의 새로운 특정한 파라미터 "main_track"에 의해 시그널될 수 있다:

[0257] aligned (8) class SubPictureCompositionBox extends

[0258] TrackGroupBox ('spco') {

[0259] unsigned int(16) track_x;

[0260] unsigned int(16) track_y;

[0261] unsigned int(16) track_width;

[0262] unsigned int(16) track_height;

[0263] unsigned int(16) composition_width;

[0264] unsigned int(16) composition_height;

[0265] unsigned int (1) main_track;

[0266] unsigned int (7) reserved;

[0267] }

[0268] 이 파라미터 main_track은 트랙 그룹 내의 트랙이 주 트랙 또는 폴 화상 트랙이라는 것을 표시하기 위해 사용될 수 있다. 이 경우에, 파서들은 트랙 그룹 내의 단지 이 주 트랙 또는 폴 화상 트랙이 렌더링되어야 한다는 것을 고려한다(그룹 내의 다른 트랙들은 값 0으로 설정된 이 파라미터를 갖지 않는다). 바꾸어 말하면, 다른 트랙들은 부분적 트랙들로서 고려된다.

- [0269] 대안으로서, 주 트랙은 트랙 내의 UserDataBox ('udta') 내의 KindBox ('kind')를 사용하여 시그널될 수 있다. KindBox는 그것의 역할 또는 종류로 트랙을 라벨링하는 것을 가능하게 한다. 주 트랙은 특정한 schemeURI, 예를 들어 "urn:mpeg:14496-12:main"을 갖는 KindBox를 정의함으로써 시그널될 것이다.
- [0270] mp4 라이터는 DASH 사전 선택 요소 내의 주 적응 세트로서 주 트랙을 설정하고 DASH MPD 내의 "hidden" 적응 세트들로서 부분적 트랙들을 설정하기 위해 주 트랙 시그널링을 이용할 수 있다. "Hidden" 적응 세트들은 사용자에 의해 선택되는 것이 의도되지 않는 적응 세트이다. 그들은 예를 들어 "urn:mpeg:dash:not-selectable:2016"에의 특정한 @schemeURI 세트로 연관된 보충적 또는 필수적 디스크립터를 정의함으로써 DASH MPD에서 명시적으로 시그널될 수 있다.
- [0271] 단계 270에서, 트랙에 대한 그리고 트랙들의 구성들에 대한 콘텐트 커버리지 정보는 비디오 비트스트림들 또는 비디오 서브-비트스트림들의 인캡슐레이션을 디스크라이브하는 메타데이터에 추가된다.
- [0272] 트랙 커버리지 정보는 이 트랙에 의해 나타내진 콘텐트에 의해 커버된 구 상의 영역에 관한 정보를 제공한다.
- [0273] 구성 커버리지 정보는 하나 이상의 트랙의 조합과 연관된 구 표면 상의 영역에 관한 정보를 제공한다. 예를 들어, 영화 파일이 공간적 관계들을 갖는 다수의 비디오 트랙을 포함할 때, 구성 커버리지 정보는 이들 다수의 비디오 트랙의 공간적 구성에 의해 커버된 구 표면 상의 영역이다. 또 하나의 예에서, 미디어 파일은 다수의 비디오 트랙 및 트랙들의 이 세트를 어떻게 렌더링하는지를 표시하는 변환 매트릭스를 포함하고, 구성 커버리지 정보는 다음에 트랙들의 어셈블된 세트에 의해 커버된 영역에 대응한다. "구성 커버리지 정보"는 또한 "글로벌 커버리지 정보" 또는 "트랙 그룹 구성 정보"로 표시될 수 있다. 구성 또는 글로벌 커버리지 정보는 또한 이들 다수의 비디오 트랙의 서브셋의 구성으로부터 생긴 구 표면 상의 영역을 디스크라이브할 수 있다.
- [0274] 제1 실시예로서, 트랙 커버리지 정보 및 구성 커버리지 정보는 추가적인 시그널링 없이 단일의 공통 CoverageInformationBox를 사용하여 시그널될 수 있다. 이러한 경우에, CoverageInformationBox의 범위는 박스 계층 내의 이 박스의 정의의 위치에 의존한다. 클라이언트들은 커버리지 정보가 그것이 선언되는 장소를 단지 고려함으로써 트랙 콘텐트에 또는 전체 콘텐트에 관련하는지를 결정할 수 있다. 본 실시예에 따라, CoverageInformationBox는 다음과 같이 정의된다:
- [0275] 박스 타입: 'cov1'
- [0276] 컨테이너: 투사된 전방향 비디오 박스 ('povd') 또는 SubPictureCompositionBox ('spco')
- [0277] 의무적: 아님
- [0278] 양: 제로 또는 하나
- ```
[0279] aligned (8) class CoverageInformationBox extends FullBox ('cov1',
[0280] version = 0, flags) {
[0281] unsigned int(8) coverage_shape_type;
[0282] SphereRegionStruct (1);
[0283] }
```
- [0284] 여기서 coverage\_shape\_type은 커버된 구 영역의 형상을 지정하고 SphereRegionStruct()는 다음과 같이 정의된다:
- ```
[0285] aligned (8) SphereRegionStruct (range_included_flag) {
[0286]     signed int(32) center_yaw;
[0287]     signed int(32) center_pitch;
[0288]     signed int(32) center_roll;
[0289]     if (range_included_flag) {
[0290]         unsigned int(32) hor_range;
[0291]         unsigned int(32) ver_range;
```

```

[0292]    }
[0293]    unsigned int(1) interpolate;
[0294]    bit (7) reserved = 0;
[0295] }

[0296] 여기서 center_yaw, center_pitch, 및 center_roll은 글로벌 좌표 축들에 대한 커버된 영역의 뷰포트 배향을
[0296] 지정하고, hor_range 및 ver_range는 존재할 때, 커버된 구 영역의 수평 및 수직 범위들을 각각 지정하고 인터
[0296] 폴레이트는 현재 사용되지 않는다.

[0297] 따라서 CoverageInformationBox는 콘텐트에 의해 커버된 구 상의 영역에 관한 정보를 제공한다. 콘텐트의 성질
[0297] 은 이 박스의 컨테이너에 의존한다. SubPictureCompositionBox 'spco'에 존재할 때, 콘텐트는 동일한 서브-화
[0297] 상 구성 트랙 그룹에 속하는 모든 트랙들에 의해 나타내진 전체 콘텐트를 참조하고 이를 트랙으로부터 구성된
[0297] 구성 화상은 전체 콘텐트의 패킹된 화상으로서 참조된다. 트랙의 샘플 엔트리에 존재할 때, 콘텐트는 이 트랙
[0297] 자체에 의해 나타내진 콘텐트를 참조하고, 이 트랙 내의 샘플들의 화상이 전체 콘텐트의 패킹된 화상으로서 참
[0297] 조된다. CoverageInformation Box가 트랙에 대해 존재하지 않을 때, 그것은 콘텐트가 전체 구를 커버한다는 것
[0297] 을 표시한다.

[0298] 투사된 전방향 비디오 박스('povd')는 MPEG OMAF에 의해 정의되고 트랙 내의 VisualSampleEntry 내로 위치되는
[0298] 중간 박스라는 점에 주목하여야 한다.

[0299] 또한, SubPictureComposition 트랙 그룹 박스('spco')는 다음과 같이 수정된다:

[0300] aligned (8) class SubPictureCompositionBox extends

[0301] TrackGroupTypeBox ('spco') {

[0302]     unsigned int(16) track_x;
[0303]     unsigned int(16) track_y;
[0304]     unsigned int(16) track_width;
[0305]     unsigned int(16) track_height;
[0306]     unsigned int(16) composition_width;
[0307]     unsigned int(16) composition_height;
[0308]     CoverageInformationBox ();
[0309] }

[0310] ISOBMFF fullBox CoverageInformation Box()를 SubPictureCompositionBox 내로 추가하기 보다는 대안으로서,
[0310] 다음과 같이, SphereRegionOnStruct를 직접 포함시키는 것이 또한 가능하다:

[0311] aligned (8) class SubPictureCompositionBox extends

[0312] TrackGroupTypeBox ('spco') {

[0313]     unsigned int(16) track_x;
[0314]     unsigned int(16) track_y;
[0315]     unsigned int(16) track_width;
[0316]     unsigned int(16) track_height;
[0317]     unsigned int(16) composition_width;
[0318]     unsigned int(16) composition_height;
[0319]     SphereRegionStruct (1);
[0320] }

```

[0321] 여전히 대안으로서, 구성을 위한 커버리지 정보의 존재는 예를 들어 아래에 예시되는 것과 같은 is_coverage_info_is_present로 표시된 추가적인 파라미터의 값으로 조절될 수 있다:

[0322] aligned (8) class SubPictureCompositionBox extends

[0323] TrackGroupBox ('spco') {

[0324] unsigned int(16) track_x;

[0325] unsigned int(16) track_y;

[0326] unsigned int(16) track_width;

[0327] unsigned int(16) track_height;

[0328] unsigned int(16) composition_width;

[0329] unsigned int(16) composition_height;

[0330] unsigned int(1) is_coverage_info_is_present

[0331] unsigned int(7) reserved

[0332] if (is_coverage_info_is_present == 1)

[0333] SphereRegionStruct(1); // or full CoverageInformationBox ()

[0334] }

[0335] 실제로 SubPictureCompositionBox가 이 SubPictureCompositionBox에 의해 정의된 그룹에 관련한 모든 트랙들에서 정의될 때, 트랙 그룹 내에 합성 트랙이 있으면, 구성 커버리지 정보는 단지 이 합성 트랙에 대해 정의될 수 있고 각각의 단일 트랙에 대해 정의될 필요는 없다.

[0336] 제2 실시예로서, 트랙 커버리지 정보 및 구성 커버리지 정보는 로컬과 글로벌 표시를 구별하기 위해 플래그 값으로 단일의 공통 CoverageInformationBox를 사용하여 시그널될 수 있다. CoverageInformationBox는 ISOBMFF FullBox이기 때문에, 트랙과 글로벌 커버리지 간의 구별은 박스의 플래그들 파라미터를 통해 표현될 수 있다.

[0337] 제2 실시예에 따라, CoverageInformation Box는 다음과 같이 정의된다:

[0338] 박스 타입: 'cov1'

[0339] 컨테이너: 투사된 전방향 비디오 박스 ('povd')

[0340] 의무적: 아님

[0341] 양: 제로 이상

[0342] aligned (8) class CoverageInformationBox extends FullBox ('cov1',

[0343] version = 0, flags) {

[0344] unsigned int(8) coverage_shape_type;

[0345] SphereRegionStruct (1);

[0346] }

[0347] 박스의 구조는 로컬 및 구성 커버리지 정보가 동일한 트랙에서 정의되어야 하는 경우에 박스의 다수의 예가 정의될 수 있는 것을 제외하고 이전의 실시예와 거의 동일하다.

[0348] CoverageInformationBox는 다음에 콘텐트에 의해 커버된 구상의 영역에 관한 정보를 제공함으로써 정의된다. 콘텐트의 성질은 플래그들 파라미터에 의해 주어진다. 커버리지 정보 플래그들에 대한 디폴트 값은 0이고, 이는 이 박스가 전체 콘텐트의 커버리지를 디스크라이브하는 것을 의미한다. 이 트랙이 서브-화상 구성 트랙 그룹에 속하면, 전체 콘텐트는 동일한 서브-화상 구성 트랙 그룹에 속하는 모든 트랙들에 의해 나타내진 콘텐트를 참조하고, 이를 트랙으로부터 구성된 구성 화상은 전체 콘텐트의 패킹된 화상으로서 참조된다. 그렇지 않으면, 전체 콘텐트는 이 트랙 자체에 의해 나타내진 콘텐트를 참조하고, 이 트랙 내의 샘플의 화상은 전체 콘텐트의

패킹된 화상으로서 참조된다.

[0349] 커버리지 정보 플래그들에 대한 값이 1일 때, 이 박스는 이 트랙에 의해 나타내진 콘텐트의 패킹된 화상들에 의해 커버된 구 영역을 디스크라이브한다.

[0350] 이 박스의 부재는 콘텐트가 전체 구를 커버한다는 것을 표시한다.

[0351] 또한, 새로운 플래그 값이 다음과 같이 정의된다:

[0352] Coverage_local은: 커버리지 정보가 박스를 포함하는 트랙에 로컬이라는 것을 표시한다. 플래그 값은 0x0000001이다. 디폴트로, 이 값은 설정되지 않는다.

[0353] 제2 실시예의 대안으로서, CoverageInformationBox의 정의는 글로벌 커버리지 정보를 갖는 CoverageInformation Box에 의해 나타내진 트랙 그룹(예를 들어, 'spco' 박스의 하나)을 식별하기 위해 track_group_id를 포함할 수 있다.

[0354] CoverageInformation Box는 다음에 다음과 같이 정의된다:

[0355] 박스 타입: 'cov1'

[0356] 컨테이너: 투사된 전방향 비디오 박스 ('povd')

[0357] 의무적: 아님

[0358] 양: 제로 이상

[0359] aligned (8) class CoverageInformationBox extends FullBox ('cov1',

[0360] version = 0, flags) {

[0361] unsigned int(8) coverage_shape_type;

[0362] if (flags & Coverage_local ==0)

[0363] unsigned int(32) track_group_id;

[0364] SphereRegionStruct (1);

[0365] }

[0366] 대안으로서, 제3 실시예에서, 2개의 상이한 박스가 구성 커버리지 정보(TrackCoverageInformationBox) 또는 트랙 커버리지 정보(TrackCoverageInformationBox)를 디스크라이브하기 위해 정의된다. 박스들은 이 트랙이 다수의 트랙 그룹에 관련한다면 CompositionCoverageInformationBox가 트랙에서 다수 번 정의될 수 있는 것을 제외하고 이전의 실시예들과 동일한 시멘틱들로 다음과 같이 정의된다. 파라미터 track_group_id는 CompositionCoverageInformationBox에 의해 디스크라이브된 트랙 그룹(예를 들어, 'spco' box의 하나)을 식별하는 것을 가능하게 한다.

[0367] 박스 타입: 'covt'

[0368] 컨테이너: 투사된 전방향 비디오 박스 ('povd')

[0369] 의무적: 아님

[0370] 양: 제로 또는 하나

[0371] aligned (8) class TrackCoverageInformationBox extends FullBox ('cov1',

[0372] version = 0, flags) {

[0373] unsigned int(8) coverage_shape_type;

[0374] SphereRegionStruct (1);

[0375] }

[0376] 박스 타입: 'cov1'

[0377] 컨테이너: 투사된 전방향 비디오 박스 ('povd')

[0378] 의무적: 아님

[0379] 양: 제로 이상

[0380] aligned (8) class CompositionCoverageInformationBox extends

[0381] FullBox ('cov1', version = 0, flags) {

[0382] unsigned int(8) coverage_shape_type;

[0383] unsigned int(32) track_group_id;

[0384] SphereRegionStruct (1);

[0385] }

[0386] 대안으로서, 제4 실시예에서, 트랙과 구성 커버리지 정보를 구별하기 위해 플래그들을 사용하는 CoverageInformationBox를 갖는 실시예들(실시예 2)과 SubPictureCompositionBox 트랙 그룹('spco') 내에 또는 VisualSampleEntry 내의 투사된 전방향 비디오 박스('povd') 내에 CoverageInformationBox를 정의하는 능력(실시예 1)과 조합하는 것이 가능하다. 2가지 방식을 가능하게 함으로써, 이것은 OMAF 콘텐트에 대한 인캡슐레이션 모드에 의존하는 커버리지 시그널링의 신축성을 제공한다:

[0387] - 단일 트랙 인캡슐레이션: 단일 CoverageInformationBox는 (플래그들 값이 설정되지 않은) 트랙의 'povd' 박스에서 선언될 수 있다.

[0388] - 다수의 트랙 인캡슐레이션:

[0389] o 합성 트랙이 있는 경우: 글로벌 커버리지 정보는 (플래그 값 Coverage_local이 설정되지 않은) 이 합성 트랙의 'povd' 내부의 CoverageInformationBox에서 선언된다. 선택적으로, 서브-화상 트랙들은 (플래그 값 Coverage_local이 설정된) CoverageInformationBox를 선언할 수 있다.

[0390] o 합성 트랙이 없는 경우: 구성 커버리지 정보는 플래그들 값 Coverage_local이 설정되지 않은 'spco' 내부의 CoverageInformationBox에서 선언된다. 선택적으로, 서브-화상 트랙들은 (플래그 값 Coverage_local이 설정된) CoverageInformationBox를 선언할 수 있다.

[0391] 대안으로서, 제5 실시예에서, 트랙들의 구성이 트랙 그룹('trgr') 메커니즘을 사용하기 보다는 새로운 엔티티 그룹을 사용하여, 즉 file-level MetaBox의 GroupsListBox 내에 또는 movie-level MetaBox의 GroupsListBox 내에 특정한 EntityToGroupBox를 정의함으로써 디스크라이브되면, 구성 커버리지 정보는 이 특정한 EntityToGroupBox의 특성으로서 직접 정의될 수 있는데, 즉 위의 제1 실시예에서 설명된 것과 같은 CoverageInformation box가 이 특정한 EntityToGroupBox 내부에서 직접 선언될 수 있다. 트랙-관련된 커버리지 정보는 여전히 트랙 내의 VisualSampleEntry 내의 투사된 전방향 비디오 박스에서 정의될 것이다.

[0392] 이 특정한 엔티티 그룹은 (단계 250을 참조하여 정의된 엔티티 그룹 'egco'에 기초하여:

[0393] aligned (8) class EntityToGroupBox ('egco', version, flags)

[0394] extends FullBox ('egco', version, flags) {

[0395] unsigned int(32) group_id;

[0396] unsigned int(32) num_entities_in_group;

[0397] for (i=0; i<num_entities_in_group; i++)

[0398] unsigned int(32) entity_id;

[0399] unsigned int(16) composition_width;

[0400] unsigned int(16) composition_height;

[0401] CoverageInformationBox ();

[0402] }

[0403] 와 같이 보일 수 있고,

[0404] 또는, 다음과 같이 직접 SphereRegionOnStruct를 포함시키는 것이 또한 가능하다:

[0405] aligned (8) class EntityToGroupBox ('egco', version, flags)

[0406] extends FullBox ('egco', version, flags) {

[0407] unsigned int(32) group_id;

[0408] unsigned int(32) num_entities_in_group;

[0409] for (i=0; i<num_entities_in_group; i++)

[0410] unsigned int(32) entity_id;

[0411] unsigned int(16) composition_width;

[0412] unsigned int(16) composition_height;

[0413] SphereRegionStruct (1);

[0414] }

[0415] 대안으로서, 제6 실시예에서, 커버리지 정보가 트랙 핸들러 타입에 의존함으로써 트랙 그룹 박스 'spco'에 존재하는지 여부를 결정하거나 하지 않는 것이 또한 가능하다. 주 트랙이 'vide' 핸들러 타입을 갖고 서브-화상 트랙들이 'auxv' 또는 'subv' 트랙들을 갖는다고 가정하는 것은: 'spco' 박스의 is_coverage_info_is_present 플래그는 'auxv' 또는 'subv' 트랙들에 대해 0으로 설정되고(즉 커버리지 정보는 존재하지 않음) 'vide' 트랙들에 대해 1로 설정된다(즉 커버리지 정보는 존재함)는 것이다.

[0416] 도 2로 다시 가서, 단계 280에서, 가상 현실 미디어 콘텐트가 실제로 스테레오스코픽 가상 현실 미디어 콘텐트인지, 즉 좌측 및 우측 뷰들을 포함하는지가 체크된다.

[0417] 콘텐트가 단지 모노스코픽이면, 과정은 단계 290으로 직접 간다.

[0418] 콘텐트가 스테레오스코픽이면, 단계 285에서 스테레오스코픽 시그널링이 인캡슐레이션에 추가된다.

[0419] 스테레오스코픽 콘텐트에 대해, 고전적으로, 좌측과 우측 뷰 둘 다의 시퀀스들이 스테레오스코픽 카메라로부터 취득되고 구성 타입에 따라 비디오 시퀀스 또는 2개의 비디오 시퀀스로 구성된다.

[0420] 스테레오스코픽 콘텐트의 2개의 상이한 뷰를 나타내는 2개의 프레임을 하나의 단일 프레임 내로 조합하는 과정은 프레임 패킹(도 1의 단계 125 참조)이라고 한다.

[0421] 프레임 패킹은 스테레오 쌍을 형성하는 2개의 뷰를 단일 프레임 내로 패킹하는 것으로 구성된다. 여러가지의 널리 공지되고 사용되는 패킹 방식들: 사이드-바이-사이드, 상부-하부, 프레임 순차적, 수직 라인 인터리브식 타입...이 존재한다. 예를 들어, MPEG 애플리케이션 포맷 ISO/IEC 23000-11 1차 편집("스테레오스코픽 비디오 애플리케이션 포맷") 또는 ISO/IEC 23001-8 2차 편집("코딩-독립 코드-포인트들(CICP)")는 이를 방식의 일부를 정의한다. 프레임 패킹은 또한 예를 들어 ISO/IEC 23001-8 2차 편집("CICP")에서 정의된 값 6을 갖는 VideoFramePackingType과 같은 별개의 프레임들 내에 각각의 뷰를 유지하는 것으로 구성될 수 있다.

[0422] 예를 들어, 여전히 이 규격에 따라, 값 3은 각각의 디코드된 프레임이 2개의 구성 뷰의 대응하는 프레임들의 사이드-바이-사이드 패킹 배열을 포함하는 것을 시그널하고, 값 4는 각각의 디코드된 프레임이 2개의 구성 뷰의 대응하는 프레임들의 상부-하부 패킹 배열을 포함하는 것을 시그널한다.

[0423] 트랙이 스테레오스코픽 미디어 데이터를 포함하는지를 시그널하기 위해, StereoVideoBox가 트랙 내의 VisualSampleEntry에서 정의된다.

[0424] StereoVideoBox는 스테레오스코픽 콘텐트를 디스크라이브하는 ISOBMFF 구조이다. StereoVideoBox는 비디오 트랙 내의 디코드된 프레임들이 스테레오 쌍을 형성하는 2개의 공간적으로 패킹된 구성 프레임들의 표현을 포함하거나 또는 스테레오 쌍의 2개의 뷰 중 하나를 포함하는 것을 표시하기 위해 사용된다. StereoVideoBox 내의 파라미터들은 뷰들의 프레임 내로의 배열들 또는 패킹에 관한 정보를 제공한다. StereoVideoBox는 미디어 파일을 디코드하고 렌더링할 수 있어야 하는 플레이어에 대한 요구들을 제공하는 미디어 파일의 샘플 디스크립션들 부분에서 디스크라이브된다.

[0425] StereoVideoBox는 (ISO/IEC 14496-12에 따라) 다음과 같이 정의된다:

[0426] 박스 타입: 'stvi'

[0427] 컨테이너: SchemeInformationBox ('schi')

[0428] 의무적: 예 (SchemeType⁰ 'stvi'일 때)

[0429] 양: 하나

[0430] aligned (8) class StereoVideoBox extends FullBox ('stvi',

[0431] version = 0, 0)

[0432] {

[0433] template unsigned int(30) reserved = 0;

[0434] unsigned int(2) single_view_allowed;

[0435] unsigned int(32) stereo_scheme;

[0436] unsigned int(32) length;

[0437] unsigned int(8) [length] stereo_indication_type;

[0438] Box [] any_box; // optional

[0439] }

[0440] 여기서 single_view_allowed는 콘텐트가 단지 스테레오스코픽 디스플레이들 상에서 디스플레이될 수 있거나 또는 어떤 뷰가 모노스코픽 단일-뷰 디스플레이 상에 디스플레이하기 위해 사용될 수 있는지를 표시하고, stereo_scheme은 사용된 방식에 따라 사용된 스테레오 배열 방식 및 스테레오 표시 타입을 표시하는 정수이고, stereo_indication_type은 사용된 스테레오 표시 방식에 따라 스테레오 배열 타입을 표시한다.

[0441] StereoVideoBox가 하나의 단일 트랙을 형성하는 공통 패킹된 프레임들 내에 좌측 뷰 프레임들 및 우측 뷰 프레임들을 패킹하기 위해 사용된 프레임 패킹 방식을 디스크라이브하는 것을 가능하게 한다면, 그것은 좌측 뷰 및 우측 뷰가 ISO/IEC 23090-2(MPEG OMAF)의 맵락에서 별개의 트랙들 내에 패킹될 때 용이한 디스크립션을 가능하게 하지 않는다.

[0442] 또한, MPEG OMAF 규격은 단지 스테레오 뷰들의 사이드-바이-사이드 및 상부-하부 프레임 패킹 각각에 대한 값들 3 및 4를 가능하게 하고 다음과 같이 StereoVideoBox로 스테레오 콘텐트를 디스크라이브하는 것을 추천한다:

[0443] aligned (8) class StereoVideoBox extends FullBox ('stvi', version = 0,

[0444] 0)

[0445] {

[0446] template unsigned int(30) reserved = 0;

[0447] unsigned int(2) single_view_allowed;

[0448] unsigned int(32) stereo_scheme; // = 4 to refer CICP ISO/IEC

[0449] 23001-8;

[0450] unsigned int(32) length; // = 2

[0451] unsigned int(8) [length] stereo_indication_type;

[0452] // = {3, 0} for side-by-side frame packing or

[0453] // = { 4, 0 } for top-bottom frame packing

[0454] Box [] any_box; // optional

[0455] }

- [0456] 그러나 이 규격은 별개의 트랙들 내에 스테레오 뷰들을 디스크라이브하는 것을 가능하게 하지 않는다.
- [0457] 스테레오스코픽 콘텐트의 디스크립션을 간단히 하고 상이한 OMAF 디스크립터들에서 스테레오스코픽 정보를 반복하는 것을 피하기 위해, StereoVideoBox는 뷰들이 단일 프레임 내에 패킹되든 별개의 트랙들 내에 인캡슐레이트 되는지 간에, 뷰 인캡슐레이션 또는 패킹의 임의의 타입을 지원하기 위해 확장될 수 있다.
- [0458] 먼저 일부 제한들이 인캡슐레이션 과정으로 강요될 수 있다: 스테레오 뷰들이 예를 들어 영역-와이즈 품질 랭킹에서 상이한 특성들을 갖는다면 각각의 뷰는 그 자신의 트랙에서 인캡슐레이트되어야 하고 각각의 트랙에 대한 StereoVideoBox는 stereo_scheme = 4(즉 CICP ISO/IEC 23001-8에서 정의된 것과 같은 프레임 패킹을 사용하여야 함), stereo_indication_type = {6,0}을 가져야 하고, 이는 디코드된 프레임이 어떤 프레임 패킹 없이도 완전한 2D 프레임을 구성하는 것을 의미한다.
- [0459] 그렇게 함으로써, SphereRegionQualityRankingBox 또는 2DRegionQualityRankingBox에서와 같은 OMAF 디스크립터들에서 어딘가 다른 곳으로 뷰 식별자들(view_idc)을 반복할 필요가 없다. 트랙을 파스함으로써, 플레이어는
- 트랙이 모노스코픽 콘텐트를 포함하는지(StereoVideoBox 없음)
- [0460] - 트랙이 스테레오스코픽 콘텐트(StereoVideoBox의 존재)를 포함하는지
- [0461] o 스테레오일 때, 그것이 하나의 뷰(tref = 'svdp'를 참조하거나 그에 의해 참조되는) 또는 둘 다의 뷰를 포함하는지
- [0462] o 스테레오이고 단일 뷰를 포함할 때, (아래에 설명되는 것과 같은) StereoVideoBox를 통하는 뷰 식별자
- [0463] o 를 결정할 수 있다.
- [0464] 좌측 뷰 또는 우측 뷰를 포함하는 각각의 트랙에 대해 stereo_scheme = 4, 및 stereo_indication_type = {6, 0}으로 StereoVideoBox를 정의함으로써, 그것은 콘텐트가 스테레오스코픽 콘텐트의 부분인 것을 시그널하는 것을 가능하게 하지만 그것은 어떤 트랙이 우측 뷰의 좌측인지를 식별하는 것을 가능하게 하지 않는다.
- [0465] 좌측 및 우측 뷰들은 다음에 타입 'svdp'의 트랙 참조를 사용하여 식별된다. 참조 트랙 'svdp'를 포함하는 트랙은 참조된 트랙에 종속성을 갖고, 스테레오스코픽 관련 메타 정보를 또한 포함하는 참조 트랙으로서 식별된다.
- [0466] 또한, 트랙에 대응하는 것이 어떤 뷰인지를 표시하기 위해, 파라미터들(single_view_allowed, stereo_indication_type)의 커플이 사용된다.
- [0467] single_view_allowed에 대한 시멘틱들이 다음과 같이 정의된다:
- [0468] "stereo_scheme = 4이고 stereo_indication_type이 "노-패킹"을 표시할 때", 즉 stereo_indication_type = {6, 0}, single_view_allowed &1이 1이라는 것은 트랙이 우측 뷰를 포함하는 것을 표시하고 single_view_allowed &2가 2라는 것은 트랙이 좌측 뷰를 포함하는 것을 표시한다. 이 경우에, 값 0 및 3은 금지된다".
- [0469] 대안으로서, single_view_allowed 파라미터의 기존의 시멘틱을 수정하는 것을 피하기 위해, 트랙이 좌측 뷰 (is_left_view=1) 또는 우측 뷰(is_left_view=0)를 포함하는지를 시그널하기 위해 추가적인 1비트 파라미터 "is_left_view"를 제공하는 StereoVideoBox의 새로운 버전이 정의된다.
- [0470] aligned (8) class StereoVideoBox extends FullBox ('stvi', version, 0)
- [0471] {
- [0472] if (version >=1)
- [0473] {
- [0474] template unsigned int(29) reserved = 0;
- [0475] unsigned int(1) is_left_view;
- [0476] }
- [0477] else

```

[0479]     {
[0480]         template unsigned int(30) reserved = 0;
[0481]     }
[0482]     unsigned int(2) single_view_allowed;
[0483]     unsigned int(32) stereo_scheme; // = 4 to refer CICP ISO/IEC
[0484] 23001-8;
[0485]     unsigned int(32) length; // = 2
[0486]     unsigned int(8) [length] stereo_indication_type;
[0487]         // = {3, 0} for side-by-side frame packing or
[0488]         // = {4, 0} for top-bottom frame packing or
[0489]         // = {6, 0} for no-packing and left/right views signaling
[0490] provided by is_left_view parameter
[0491]     Box [ ] any_box; // optional
[0492] }
[0493] 대안적으로, 추가적인 파라미터는 다음의 시멘틱으로 (아래에 예시되는 것과 같은) 2비트 파라미터 "view_idc"이다: 그것이 0이라고 하는 것이 트랙 내의 미디어 콘텐트가 모노스코픽이라는 것을 표시하면, 1은 트랙 내의 미디어 콘텐트가 스테레오스코픽 콘텐트의 좌측 뷰라는 것을 표시하고, 2는 트랙 내의 미디어 콘텐트가 스테레오스코픽 콘텐트의 우측 뷰라는 것을 표시하고, 3은 트랙 내의 미디어 콘텐트가 스테레오스코픽 콘텐트의 좌측과 우측 둘 다의 뷰를 포함하는 것을 표시한다.
[0494] aligned (8) class StereoVideoBox extends FullBox ('stvi', version, 0)
[0495] {
[0496]     if (version >=1)
[0497]     {
[0498]         template unsigned int(28) reserved = 0;
[0499]         unsigned int(2) view_idc;
[0500]     }
[0501]     else
[0502]     {
[0503]         template unsigned int(30) reserved = 0;
[0504]     }
[0505]     unsigned int(2) single_view_allowed;
[0506]     unsigned int(32) stereo_scheme; // = 4 to refer CICP ISO/IEC
[0507] 23001-8;
[0508]     unsigned int(32) length; // = 2
[0509]     unsigned int(8) [length] stereo_indication_type;
[0510]         // = {3, 0} for side-by-side frame packing or
[0511]         // = {4, 0} for top-bottom frame packing or

```

```

[0512]           // = {6, 0} for no-packing and left/right views signaling
[0513] provided by view_idc parameter
[0514]           Box [ ] any_box; // optional
[0515] }
[0516] 또 하나의 대안으로서, 새로운 파라미터를 추가하고 StereoVideoBox의 새로운 버전을 만들기 보다는, 새로운 프레임 패킹 배열이 (CICP ISO/IEC 23001-8의 확장에 대응하는) stereo_scheme = 4에 대해 정의되는데, 즉 새로운 값이 정의되고, 예를 들어, stereo_scheme = 4일 때 파라미터 stereo_indication_type에 대해 7이다. 새로운 값은 다음과 같이 정의된다:
[0517] VideoFramePackingType = 7은 디코드된 프레임이 2개의 구성 프레임(즉, 스테레오스코픽 시퀀스의 좌측 뷰 또는 우측 뷰)의 대응하는 평면들 중 하나의 단일 평면을 포함한다는 것을 표시한다.
[0518] 이 새로운 VideoFramePackingType 값 및 오점형 샘플링 구조가 프레임 패킹된 비디오 표현에서 사용되는지를 시그널하는 기존의 연관된 플래그 QuincunxSamplingFlag 외에, 예를 들어 ViewldcFlag로 표시된 새로운 연관된 플래그는 프레임 패킹된 비디오 표현에 존재하는 뷰들의 타입을 식별하는 것을 가능하게 정의된다. ViewldcFlag에 대해 존재하지 않거나 특정되지 않거나 또는 값 0이 좌측과 우측 뷰들 둘 다가 존재하는 것을 표시하기 위해 추론될 때, 값 1은 단지 스테레오스코픽 콘텐트의 좌측 뷰가 존재한다는 것을 표시하고, 값 2는 단지 스테레오스코픽 콘텐트의 우측 뷰가 존재한다는 것을 표시하고, ViewldcFlag의 모든 다른 값들은 ISO/IEC에 의한 미래의 사용을 위해 확보된다.
[0519] StereoVideoBox 내의 stereo_scheme = 4의 정의는 다음과 같이 수정된다:
[0520] "stereo_scheme가 4: 길이의 값은 2일 것이고 stereo_indication_type은 부호없는 int(8)의 2개의 신택스 요소를 포함할 것이다. 제1 신택스 요소는 ISO/IEC 23001-8로부터의 VideoFramePackingType을 포함할 것이다. 값 0 내지 6의 VideoFramePackingType에 대해, 제2 신택스 요소의 최하위 비트는 ISO/IEC 23001-8에서 지정된 것과 같은 QuincunxSamplingFlag의 값을 포함할 것이지만, 다른 비트들이 확보되고 0으로 설정될 것이다. 값 7을 갖는 VideoFramePackingType에 대해, 제2 신택스 요소의 최하위 2비트는 좌측 뷰 및 우측 뷰를 식별하고 (위에 정의된 것과 같은) ViewldcFlag의 값을 포함할 것이지만, 다른 비트들이 확보되고 0으로 설정될 것이다".
[0521] 대안으로서 QuincunxSamplingFlag와 ViewldcFlag 둘 다는 다음과 같이 StereoVideoBox에서 stereo_scheme = 4를 정의함으로써 동시에 시그널될 수 있다:
[0522] "stereo_scheme가 4: 길이의 값은 3일 것이고 stereo_indication_type은 부호없는 int(8)의 3개의 신택스 요소를 포함할 것이다. 제1 신택스 요소는 ISO/IEC 23001-8로부터의 VideoFramePackingType을 포함할 것이다. 제2 신택스 요소의 최하위 비트는 ISO/IEC 23001-8에서 지정된 것과 같은 QuincunxSamplingFlag의 값을 포함할 것이지만, 다른 비트들이 확보되고 0으로 설정될 것이다. 제3 신택스 요소의 최하위 2비트는 좌측 뷰 및 우측 뷰를 식별하고 (위에 정의된 것과 같은) ViewldcFlag의 값을 포함할 것이지만, 다른 비트들이 확보되고 0으로 설정될 것이다".
[0523] 한 예로서, 상기 대안에 따라, StereoVideoBox는 다음과 같이 코멘트들에 나타낸 가능한 값으로 변화되지 않은 채로 남을 것이다:
[0524] aligned (8) class StereoVideoBox extends FullBox ('stvi', version = 0,
[0525]           0)
[0526] {
[0527]           template unsigned int(30) reserved = 0;
[0528]           unsigned int(2) single_view_allowed;
[0529]           unsigned int(32) stereo_scheme; // = 4 to refer CICP ISO/IEC
[0530]           23001-8;
[0531]           unsigned int(32) length; // = 3

```

```

[0532]     unsigned int(8) [length] stereo_indication_type;
[0533]         // = {3, 0, 0} for side-by-side frame packing or
[0534]         // = {4, 0, 0} for top-bottom frame packing or
[0535]         // = {7, 0, 1} for single left view packing or
[0536]         // = {7, 0, 2} for single right view packing
[0537]     Box [ ] any_box; // optional
[0538] }

```

[0539] 대안으로서 QuincunxSamplingFlag와 ViewldcFlag 둘 다가 다음과 같이 StereoVideoBox에서 stereo_scheme = 4를 정의함으로써 선택적으로 시그널될 수 있다:

[0540] "stereo_scheme가 4: 길이의 값은 1, 2 또는 3일 것이고 stereo_indication_type은 각각 부호없는 int(8)의 1개, 2개 또는 3개의 신택스 요소를 포함할 것이다. 제1 신택스 요소는 ISO/IEC 23001-8로부터의 VideoFramePackingType을 포함할 것이다. 제2 신택스 요소의 최하위 비트는 존재한다면, ISO/IEC 23001-8에서 지정된 것과 같은 QuincunxSamplingFlag의 값을 포함할 것이지만, 다른 비트들이 확보되고 0으로 설정될 것이다. 제3 신택스 요소의 최하위 2비트는 존재한다면, 좌측 뷰 및 우측 뷰를 식별하고 (위에 정의된 것과 같은) ViewldcFlag의 값을 포함할 것이지만, 다른 비트들이 확보되고 0으로 설정될 것이다". 제3 신택스 요소가 존재한다면 제2 신택스 요소는 존재할 것이다.

[0541] 또 하나의 대안으로서, 별개의 트랙들 내에 좌측 및 우측 뷰들을 구성하는 스테레오스코픽 전방향 미디어에 대한 프레임 패킹 배열들이 4(CICP ISO/IEC 23001-8에서 정의된 것과 같은 사용 프레임 패킹) 대신에 3(ISO/IEC 23000-11 1차 편집("스테레오스코픽 비디오 애플리케이션 포맷")에서 정의된 것과 같은 사용 프레임 패킹)인 stereo_scheme을 사용하여 시그널될 수 있다. ISO/IEC 14496-12 4차 편집에서의 StereoVideoBox 정의에 따라:

[0542] stereo_scheme이 3인 것은 길이의 값은 2일 것이고 stereo_indication_type은 부호없는 int(8)의 2개의 신택스 요소를 포함하는 것을 표시한다. 제1 신택스 요소는 ISO/IEC 23000-11:2009의 테이블 4로부터의 스테레오스코픽 구성 타입을 포함할 것이다. 제2 신택스 요소의 최하위 비트는 ISO/IEC 23000-11:2009의 8.4.3에서 지정된 것과 같은 is_left_first의 값을 포함할 것이지만, 다른 비트들이 확보되고 0으로 설정될 것이다.

[0543] 그러므로 stereo_scheme이 3인 StereoVideoBox를 트랙 내에 정의함으로써 그리고 값 0x3을 갖는 stereo_indication_type의 제1 신택스 요소를 정의함으로써(트랙은 좌측/우측 뷰 시퀀스 타입, 즉 좌측 또는 우측 뷰만을 나타내는 것을 의미) 그리고 좌측 뷰가 2차 뷰라는 것을 시그널하는 0 또는 좌측 뷰가 1차 뷰라는 것을 시그널하는 1로서 제2 신택스 요소를 정의함으로써 그 트랙이 스테레오스코픽 콘텐트의 좌측 뷰 또는 우측 뷰를 포함한다는 것을 시그널하는 것이 가능하다. 1차 뷰 및 2차 뷰는 좌측-뷰 및 우측-뷰 트랙들을 링크하는 트랙 참조 'svdp'의 도움으로 식별된다. 타입 'svdp'의 'tref' 박스를 갖는 트랙은 2차 뷰 시퀀스이고, 참조된 트랙은 1차 뷰 시퀀스이다.

[0544] StereoVideoBox의 새로운 버전(표시된 버전 = 1)을 생성하는 실시예들에서의 StereoVideoBox의 크기는 stereo_scheme 및 stereo_indication_type에 대해 허가된 몇 개의 값들에 적은 바이트들을 할당함으로써 버전 0과 비교하여 감소될 수 있다는 점에 주목할 수 있다.

[0545] 대안에서 새로운 파라미터 view_idc를 도입하는 실시예에 대한 StereoVideoBox의 보다 작은 버전 1이 다음과 같이 디스크라이브될 수 있다(6바이트들을 절약):

```

[0546] aligned (8) class StereoVideoBox extends FullBox ('stvi', version = 1,
[0547] 0)
[0548] {
[0549]     template unsigned int(28) reserved = 0;
[0550]     unsigned int(2) view_idc;
[0551]     unsigned int(2) single_view_allowed;

```

```

[0552]     unsigned int(8) stereo_scheme; // = 4 to refer CICP ISO/IEC
[0553]     23001-8;
[0554]     unsigned int(8) length; // = 2 as in CICP indication
[0555]     unsigned int(8)[length] stereo_indication_type; //={3, 0} or {4,
[0556]     0}
[0557]     Box [ ] any_box; // optional
[0558] }
```

[0559] 유사하게, 추가적인 파라미터가 view_idc 대신에 "is_left_view"이면 동일한 작은 버전이 정의될 수 있다.

[0560] 또한, 프레임 패킹이 뷰 당 하나의 패킹된 프레임을 야기할 때, DASH 다수 뷰들 방식이 스테레오 쌍들을 디스크 라이브하기 위해 적응 세트 레벨에서의 역할 요소에서 사용될 수 있다.

[0561] 모든 상기 실시예들에 따라, SphereRegionQualityRankingBox 및 2DRegionQualityRankingBox 내의 view_idc 및 view_idc_presence_flag 파라미터들은 뷰들이 다음과 같이 상이한 트랙들로 분리될 때 더 이상 필요없는 것으로서 제거된다:

```

aligned(8) class SphereRegionQualityRankingBox extends FullBox('srqr',
version = 0, 0) {
    unsigned int(8) region_definition_type;
    unsigned int(8) num_regions;
    unsigned int(1) remaining_area_flag;
    unsigned int(1) view_idc_presence_flag;
    bit(76) reserved = 0;
    if (view_idc_presence_flag == 0) {
        unsigned int(2) default_view_idc;
        bit(6) reserved = 0;
    }
    for (i = 0; i < num_regions; i++) {
        unsigned int(8) quality_ranking;
        if (view_idc_presence_flag == 1) {
            unsigned int(2) view_idc;
            bit(6) reserved = 0;
        }
        if ((i < (num_regions - 1)) || (remaining_area_flag == 0))
            SphereRegionStruct(1);
    }
}
```

```

aligned(8) class 2DRegionQualityRankingBox extends FullBox('2dqr',
version = 0, 0) {
    unsigned int(8) num_regions;
    unsigned int(1) remaining_area_flag;
    unsigned int(1) view_idc_presence_flag;
    bit(76) reserved = 0;
    for (i = 0; i < num_regions; i++) {
        unsigned int(8) quality_ranking;
        if (view_idc_presence_flag == 1) {
            unsigned int(2) view_idc;
            bit(6) reserved = 0;
        }
        if ((i < (num_regions - 1)) || (remaining_area_flag == 0)) {
            unsigned int(16) left_offset;
            unsigned int(16) top_offset;
            unsigned int(16) region_width;
            unsigned int(16) region_height;
        }
    }
}
```

[0563] 대안으로서, view_idc 및 view_idc_presence_flag 파라미터들은 아래에 예시되는 것과 같이 SphereRegionQualityRankingBox 또는 2DRegionQualityRankingBox의 특정한 버전으로 조절된다:

```

aligned(8) class SphereRegionQualityRankingBox extends FullBox('srqr',
version, 0) {
    unsigned int(8) region_definition_type;
    unsigned int(8) num_regions;
    unsigned int(1) remaining_area_flag;
    if (version >=1)
    {
        unsigned int(1) view_idc_presence_flag;
        if (view_idc_presence_flag == 0) {
            unsigned int(2) default_view_idc;
            bit(4) reserved = 0;
        }else{
            bit(6) reserved = 0;
        }else{
            bit(7) reserved = 0;
        }
        for (i = 0; i < num_regions; i++) {
            unsigned int(8) quality_ranking;
            if (version >= 1)
            {
                if (view_idc_presence_flag == 1) {
                    unsigned int(2) view_idc;
                    bit(6) reserved = 0;
                }
            }
            if ((i < (num_regions - 1)) || (remaining_area_flag == 0))
                SphereRegionStruct(1);
        }
    }
}

```

[0565]

```

aligned(8) class 2DRegionQualityRankingBox extends FullBox('2dqr',
version, 0) {
    unsigned int(8) num_regions;
    unsigned int(1) remaining_area_flag;
    if (version >=1)
    {
        unsigned int(1) view_idc_presence_flag;
        if (view_idc_presence_flag == 0) {
            unsigned int(2) default_view_idc;
            bit(4) reserved = 0;
        }else{
            bit(6) reserved = 0;
        }else{
            bit(7) reserved = 0;
        }
        for (i = 0; i < num_regions; i++) {
            unsigned int(8) quality_ranking;
            if (version >= 1)
            {
                if (view_idc_presence_flag == 1) {
                    unsigned int(2) view_idc;
                    bit(6) reserved = 0;
                }
            }
            if ((i < (num_regions - 1)) || (remaining_area_flag == 0)) {
                unsigned int(16) left_offset;
                unsigned int(16) top_offset;
                unsigned int(16) region_width;
                unsigned int(16) region_height;
            }
        }
    }
}

```

[0566]

실제로, 트랙이 단지 전체 좌측 뷰 또는 전체 우측 뷰를 포함할 때, 이 트랙에서 정의된 각각의 품질 랭킹 영역에 대한 view_idc(스테레오스코픽 뷰를 시그널링)를 포함할 필요가 없다. 이러한 경우에 그들 박스의 버전 == 0이 사용된다. 그렇지 않고, 트랙이 패킹된 뷰들을 포함하면, 그들 박스의 버전 == 1이 사용된다.

[0568]

본 발명의 실시예들에 따라 발생된 미디어 파일들을 파스하는 것은, 미디어 파일들이 장면의 와이드 뷰에 대응하는 인코드된 미디어 데이터를 포함하는 경우에, 클라이언트에 의한 다음의 단계들을 포함할 수 있다. 인코드된 미디어 데이터를 인캡슐레이트하는 적어도 하나의 트랙이 미디어 파일들에서 식별된다. 주어진 트랙 내에

인캡슐레이트된 인코드된 미디어 데이터를 디코드할 데이터의 필요성에 관련한 정보가 미디어 파일들로부터 획득된다. 주어진 트랙의 인코드된 미디어 데이터는 획득된 정보에 따라 디코드된다.

[0569] 도 3은 본 발명의 하나 이상의 실시예의 구현을 위한 컴퓨팅 디바이스(300)의 개략 블록도이다. 컴퓨팅 디바이스(300)는 마이크로-컴퓨터, 워크스테이션 또는 경량 휴대용 디바이스와 같은 디바이스일 수 있다. 컴퓨팅 디바이스(300)는

[0570] 마이크로프로세서와 같은 중앙 처리 장치(CPU)(301);

[0571] 매니페스트들을 판독 및 기입하고/하거나 비디오를 인코드하고/하거나 주어진 파일 포맷 하에서 데이터를 판독 및 발생하는 방법을 구현하기 위해 필요한 변수들 및 파라미터들을 기록하도록 적응된 레지스터들뿐만 아니라 본 발명의 실시예들의 방법의 실행가능한 코드를 저장하는 랜덤 액세스 메모리(RAM)(302) - 그 메모리 용량은 예를 들어 확장 포트에 접속된 선택적인 RAM에 의해 확장될 수 있음 - ;

[0572] 본 발명의 실시예들을 구현하기 위한 컴퓨터 프로그램들을 저장하는 리드 온리 메모리(ROM)(303);

[0573] 결국, 전형적으로 처리될 디지털 데이터가 송신 또는 수신되는 통신 네트워크에 접속된 네트워크 인터페이스(304) - 네트워크 인터페이스(304)는 단일 네트워크 인터페이스일 수 있거나, 상이한 네트워크 인터페이스들(예를 들어, 유선 또는 무선 인터페이스들, 또는 상이한 종류들의 유선 또는 무선 인터페이스들)의 세트로 구성될 수 있고, 데이터는 CPU(301)에서 실행하는 소프트웨어 애플리케이션의 제어 하에서 송신을 위해 네트워크 인터페이스에 기입되거나 수신을 위해 네트워크 인터페이스로부터 판독됨 - ;

[0574] 사용자로부터 입력들을 수신하고 사용자에게 정보를 디스플레이하는 사용자 인터페이스(UI)(305);

[0575] 하드 디스크(HD)(306);

[0576] 비디오 소스 또는 디스플레이와 같은 외부 디바이스들로부터/에 데이터를 수신/송신하는 I/O 모듈(307)

[0577] 예 접속된 통신 버스를 포함한다.

[0578] 실행가능한 코드는 리드 온리 메모리(303) 내에, 하드 디스크(306) 상에 또는 예를 들어 디스크와 같은 착탈가능한 디지털 매체 상에 저장될 수 있다. 변형에 따라, 프로그램들의 실행가능한 코드는 실행되기 전에, 하드 디스크(306)와 같은, 통신 디바이스(300)의 저장 수단 중 하나 내에 저장되기 위해, 네트워크 인터페이스(304)를 통해, 통신 네트워크에 의해 수신될 수 있다.

[0579] 중앙 처리 장치(301)는 본 발명의 실시예들에 따라 명령어들 또는 프로그램 또는 프로그램들의 소프트웨어 코드의 부분들의 실행을 제어 및 지시하도록 적응되고, 그 명령어들은 위에 언급된 저장 수단 중 하나 내에 저장된다. 전력 온할 때, CPU(301)는 그들 명령어가 예를 들어 프로그램 ROM(303) 또는 하드-디스크(HD)(306)으로부터 로드된 후에 소프트웨어 애플리케이션에 관련한 주 RAM 메모리(302)로부터의 명령어들을 실행할 수 있다. 이러한 소프트웨어 애플리케이션은 CPU(301)에 의해 실행될 때, 이전의 도면들에 도시한 플로우차트들의 단계들이 수행되게 한다.

[0580] 본 실시예에서, 장치는 본 발명을 구현하기 위해 소프트웨어를 사용하는 프로그램가능한 장치이다. 그러나, 대안적으로, 본 발명은 하드웨어에서(예를 들어, 주문형 접적 회로 또는 ASIC의 형태로) 구현될 수 있다.

[0581] 본 발명이 특정한 실시예들을 참조하여 위에 설명되었지만, 본 발명은 특정한 실시예들로 제한되지 않고, 수정들이 본 발명의 범위 내에 있는 기술 분야의 통상의 기술자에게 분명할 것이다.

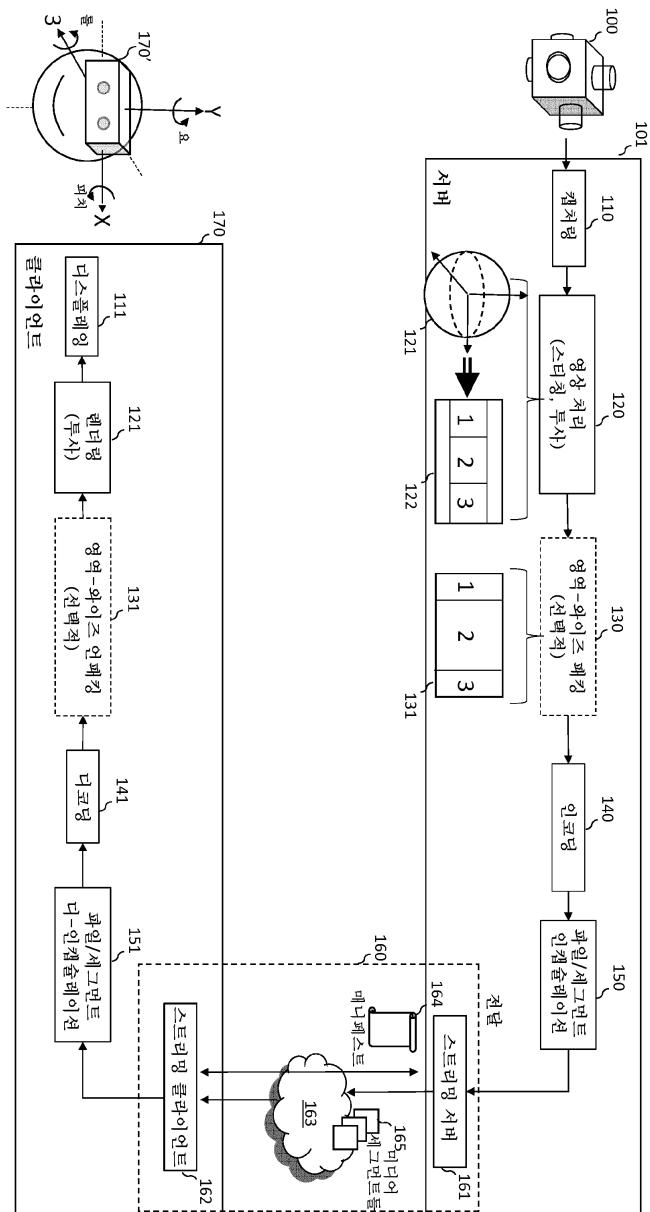
[0582] 예를 들어, 본 발명은 카메라, 스마트폰, 헤드 마운트 디스플레이 또는 예를 들어 관심있는 특정한 영역 상으로 줌 인하기 위해 멀티미디어 디스플레이로서 동작하는 TV용 멀티미디어 디스플레이용 원격 제어기로서 동작하는 태블릿과 같은 디바이스 내에 내장될 수 있다. 그것은 또한 관심있는 특정한 영역들을 선택함으로써 멀티미디어 프리젠테이션의 개인화된 브라우징 경험을 갖기 위해 동일한 디바이스로부터 사용될 수 있다. 사용자에 의한 이들 디바이스 및 방법으로부터의 또 하나의 용도는 그의 선호되는 비디오들의 일부 선택된 서브-부분들을 다른 접속된 디바이스들과 공유하는 것이다. 그것은 또한 감시 카메라가 본 발명에 따른 데이터를 제공하는 방법을 지원한다면 감시 하에 있는 빌딩의 특정한 영역에서 어떤 것이 발생하는지를 모니터하기 위해 스마트폰 또는 태블릿과 함께 사용될 수 있다.

[0583] 추가의 수정들 및 변형들이 전술한 예시적인 실시예들을 참조하여 본 기술 분야의 통상의 기술자들에게 그들을 참조하기를 제안할 것이고, 그것은 본 발명의 범위를 제한하려는 것이 아니고 예로서 주어지고, 그 범위는 첨부

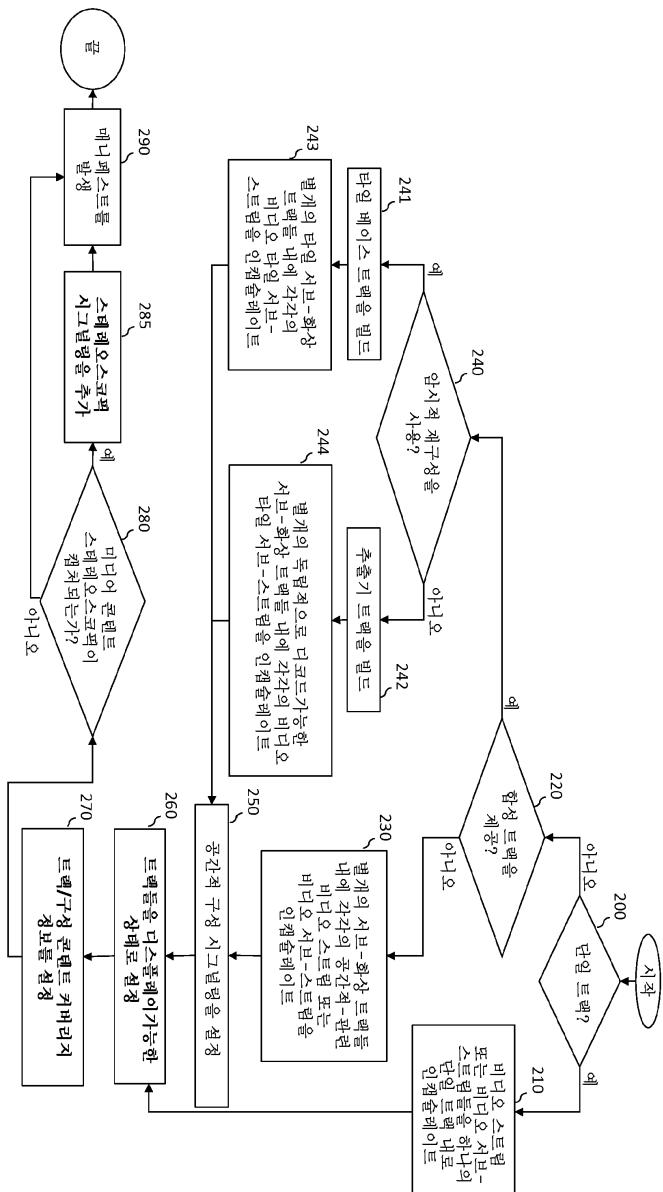
된 청구범위에 의해서만 결정된다. 특히 상이한 실시예들로부터의 상이한 특징들이 적절한 경우에 상호교환될 수 있다.

도면

도면1



도면2



도면3

