



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2023년07월27일
(11) 등록번호 10-2559862
(24) 등록일자 2023년07월21일

- (51) 국제특허분류(Int. Cl.)
HO4N 21/235 (2011.01) HO4N 13/178 (2018.01)
HO4N 13/194 (2018.01) HO4N 19/597 (2014.01)
HO4N 21/2343 (2011.01) HO4N 21/4402 (2011.01)
HO4N 21/81 (2011.01) HO4N 21/854 (2011.01)
- (52) CPC특허분류
HO4N 21/2353 (2013.01)
HO4N 13/178 (2018.05)
- (21) 출원번호 10-2020-7034288
- (22) 출원일자(국제) 2019년06월05일
심사청구일자 2021년01월06일
- (85) 번역문제출일자 2020년11월27일
- (65) 공개번호 10-2021-0016530
- (43) 공개일자 2021년02월16일
- (86) 국제출원번호 PCT/EP2019/064691
- (87) 국제공개번호 WO 2019/234116
국제공개일자 2019년12월12일
- (30) 우선권주장
1809331.0 2018년06월06일 영국(GB)
- (56) 선행기술조사문헌
US20160057491 A1*
YE-KUI WANG ET AL "WD of ISO/IEC 23090-2 2nd edition OMAF", 122.MPEG MEETING, (2018-05-05)*
FRANCK DENOUAL ET AL "On Coverage information in OMAF" 119.MPEG MEETING, (2017-07-10)*
*는 심사관에 의하여 인용된 문헌

- (73) 특허권자
캐논 가부시끼가이샤
일본 도쿄도 오오따꾸 시모마루쵸 3쵸메 30방 2고
- (72) 발명자
마즈 프레데릭
프랑스 35850 랑강 튀 데 티윌 6
드누알 프랑크
프랑스 35190 생 도미니크 라 빌-에-레 8
우에드라오고 나엘
프랑스 35330 발 다나스트 쿠투스 모르 드 브르타뉴 11
- (74) 대리인
장수길, 이중희

전체 청구항 수 : 총 15 항

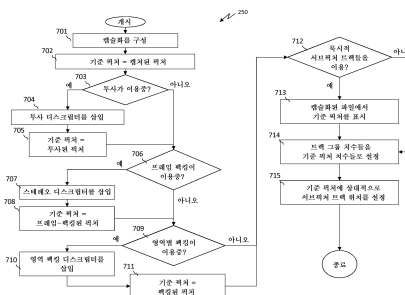
심사관 : 우정훈

(54) 발명의 명칭 미디어 콘텐츠 전송을 위한 방법, 디바이스, 및 컴퓨터 프로그램

(57) 요약

본 발명은 장면의 와이드 뷰에 대응하는 인코딩된 미디어 데이터를 캡슐화하는 방법에 관한 것으로, 이 방법은 : 장면의 와이드 뷰로부터 투사된 픽처를 획득하는 단계; 투사된 픽처를 적어도 하나의 서브픽처로 분할하는 단계; 적어도 하나의 서브픽처를 복수의 트랙으로 인코딩하는 단계; 인코딩된 트랙들과 연관된 설명 메타데이터 - 설명 (뒷면에 계속)

대표도



메타데이터는 트랙으로 인코딩된 적어도 하나의 서브픽처와 기준 픽처 사이의 공간적 관계를 나타내는 각각의 트랙과 연관된 제1 정보 항목을 포함함 - 를 생성하는 단계를 포함하고, 설명 메타데이터는 기준 픽처를 나타내는 제2 정보 항목을 더 포함한다.

(52) CPC특허분류

H04N 13/194 (2018.05)

H04N 19/597 (2015.01)

H04N 21/234327 (2013.01)

H04N 21/234345 (2013.01)

H04N 21/440227 (2013.01)

H04N 21/440245 (2013.01)

H04N 21/816 (2013.01)

H04N 21/85406 (2013.01)

명세서

청구범위

청구항 1

ISO BMFF와 호환가능한 파일 포맷으로 장면의 와이드 뷰에 대응하는 미디어 데이터를 캡슐화하기 위한 방법으로서, 상기 와이드 뷰는 상기 장면의 이미지를 구면(sphere)의 적어도 일부 상으로 투사함으로써 발생하고, 상기 방법은,

상기 장면의 와이드 뷰로부터 2차원 투사된 픽처를 획득하는 단계로서, 상기 2차원 투사된 픽처는 상기 와이드 뷰를 평면 상으로 투사함으로써 발생하는, 단계;

상기 2차원 투사된 픽처를 복수의 서브픽처로 분할하는 단계;

상기 복수의 서브픽처를 복수의 트랙으로 캡슐화하는 단계;

상기 캡슐화된 트랙들과 연관된 설명 메타데이터를 생성하는 단계를 포함하고,

상기 설명 메타데이터는, 상기 트랙으로 캡슐화된 복수의 서브픽처와 합성 픽처 간의 2차원 공간적 관계를 나타내는 각각의 트랙과 연관된 제1 정보를 포함하고, 상기 합성 픽처는 상기 복수의 트랙의 서브픽처 각각의 합성으로부터 발생하고, 상기 복수의 트랙은 트랙들의 동일한 그룹에 속하고,

상기 설명 메타데이터는, 또한, 상기 합성 픽처가 영역별 팩킹된 픽처인지 여부를 나타내기 위한 제2 정보를 포함하고, 상기 영역별 팩킹된 픽처는 상기 2차원 투사된 픽처에 연산을 적용함으로써 획득되는, 방법.

청구항 2

제1항에 있어서, 상기 영역별 팩킹된 픽처는, 상기 투사로부터 발생한 상기 2차원 투사된 픽처를 팩킹된 픽처로 팩킹하고 상기 팩킹된 픽처를 복수의 서브픽처로 분할함으로써 획득되는, 방법.

청구항 3

제1항 또는 제2항에 있어서, 상기 제2 정보는, 상기 합성 픽처가 영역별 팩킹된 픽처인지 여부를 4-캐릭터 코드에 의해 나타내는, 방법.

청구항 4

삭제

청구항 5

제1항에 있어서, 상기 제2 정보는, 각각의 트랙과 연관된 상기 제1 정보에 포함되는, 방법.

청구항 6

제5항에 있어서, 상기 제2 정보는 상기 제1 정보의 파라미터로서 정의되는, 방법.

청구항 7

삭제

청구항 8

삭제

청구항 9

제1항에 있어서, 상기 제2 정보는, 서브픽처들에 대응하는 트랙들의 그룹의 속성들을 기술하기 위해 이용되는 특정한 유형의 그룹 정보로서 정의되는, 방법.

청구항 10

복수의 캡슐화된 트랙 및 연관된 설명 메타데이터를 포함하는, ISO BMFF와 호환가능한 파일 포맷의 미디어 파일로부터, 적어도 하나의 이미지를 생성하기 위한 방법으로서,

상기 복수의 캡슐화된 트랙이, 평면 상으로의 장면의 와이드 뷰의 투사로부터 발생하는 2차원 투사된 픽처의 분할로부터 발생하는 복수의 서브픽처를 캡슐화한 트랙들의 그룹을 포함한다고 결정하는 단계로서, 상기 와이드 뷰는 상기 장면의 이미지들을 구면의 적어도 일부 상으로 투사함으로써 발생하는, 단계; 및

상기 트랙들의 그룹과 연관된 설명 메타데이터를 파싱하는 단계를 포함하고,

상기 트랙들의 그룹과 연관된 설명 메타데이터를 파싱하는 단계는,

상기 트랙으로 캡슐화된 복수의 서브픽처와 적어도 하나의 합성 픽처 간의 2차원 공간적 관계를 나타내는 각각의 트랙과 연관된 제1 정보를 해석하는 단계로서, 상기 합성 픽처는, 상기 복수의 트랙의 서브픽처 각각의 합성으로부터 발생하고, 상기 복수의 트랙은 트랙들의 동일한 그룹에 속하는, 단계; 및

상기 합성 픽처가 영역별 팩킹된 픽처인지 여부를 나타내는 상기 설명 메타데이터 내의 제2 정보를 해석하는 단계로서, 상기 영역별 팩킹된 픽처는 상기 2차원 투사된 픽처에 연산을 적용함으로써 획득되는 단계를 포함하는, 방법.

청구항 11

제10항에 있어서, 상기 영역별 팩킹된 픽처는, 상기 투사로부터 발생한 상기 2차원 투사된 픽처를 팩킹된 픽처로 팩킹하고 상기 팩킹된 픽처를 복수의 서브픽처로 분할함으로써 획득되는, 방법.

청구항 12

제10항에 있어서, 상기 제2 정보는, 상기 합성 픽처가 영역별 팩킹된 픽처인지 여부를 4-캐릭터 코드에 의해 나타내는, 방법.

청구항 13

삭제

청구항 14

제10항에 있어서, 상기 제2 정보는 각각의 트랙과 연관된 상기 제1 정보에 포함되는, 방법.

청구항 15

제14항에 있어서, 상기 제2 정보는 상기 제1 정보의 파라미터로서 정의되는, 방법.

청구항 16

삭제

청구항 17

삭제

청구항 18

제10항에 있어서, 상기 제2 정보는, 서브픽처들에 대응하는 트랙들의 그룹의 속성들을 기술하기 위해 이용되는 특정한 유형의 그룹 정보로서 정의되는, 방법.

청구항 19

ISO BMFF와 호환가능한 파일 포맷으로 장면의 와이드 뷰에 대응하는 미디어 데이터를 캡슐화하기 위한 컴퓨팅 디바이스로서, 상기 와이드 뷰는 상기 장면의 이미지를 구면의 적어도 일부 상으로 투사함으로써 발생하고, 상기 컴퓨팅 디바이스는,

상기 장면의 와이드 뷰로부터의 2차원 투사된 픽처 - 상기 2차원 투사된 픽처는 상기 와이드 뷰를 평면 상으로 투사함으로써 발생함 - 를 획득하고,

상기 2차원 투사된 픽처를 복수의 서브픽처로 분할하고,

상기 복수의 서브픽처를 복수의 트랙으로 캡슐화하고,

상기 캡슐화된 트랙들과 연관된 설명 메타데이터를 생성하도록 구성되고,

상기 설명 메타데이터는, 상기 트랙으로 캡슐화된 복수의 서브픽처와 합성 픽처 간의 2차원 공간적 관계를 나타내는 각각의 트랙과 연관된 제1 정보를 포함하고, 상기 합성 픽처는 상기 복수의 트랙의 서브픽처 각각의 합성으로부터 발생하고, 상기 복수의 트랙은 트랙들의 동일한 그룹에 속하고,

상기 설명 메타데이터는, 또한, 상기 합성 픽처가 영역별 팩킹된 픽처인지 여부를 나타내기 위한 제2 정보를 포함하고, 상기 영역별 팩킹된 픽처는 상기 2차원 투사된 픽처에 연산을 적용함으로써 획득되는, 컴퓨팅 디바이스.

청구항 20

복수의 캡슐화된 트랙 및 연관된 설명 메타데이터를 포함하는, ISO BMFF와 호환가능한 파일 포맷의 미디어 파일로부터, 적어도 하나의 이미지를 생성하기 위한 컴퓨팅 디바이스로서,

상기 복수의 캡슐화된 트랙이, 평면 상으로의 장면의 와이드 뷰의 투사로부터 발생한 2차원 투사된 픽처의 분할로부터 발생하는 복수의 서브픽처를 캡슐화한 트랙들의 그룹을 포함한다고 결정하고 - 상기 와이드 뷰는 상기 장면의 이미지들을 구면의 적어도 일부 상으로 투사함으로써 발생함 - ,

상기 트랙들의 그룹과 연관된 설명 메타데이터를 과싱하도록 구성되고,

상기 트랙들의 그룹과 연관된 설명 메타데이터를 과싱하는 것은,

상기 트랙으로 캡슐화된 복수의 서브픽처와 적어도 하나의 합성 픽처 간의 2차원 공간적 관계를 나타내는 각각의 트랙과 연관된 제1 정보를 해석하는 것 - 상기 합성 픽처는, 상기 복수의 트랙의 서브픽처 각각의 합성으로부터 발생하고, 상기 복수의 트랙은 트랙들의 동일한 그룹에 속함 - ; 및

상기 합성 픽처가 영역별 팩킹된 픽처인지 여부를 나타내는 상기 설명 메타데이터에 내의 제2 정보를 해석하는 것 - 상기 영역별 팩킹된 픽처는 상기 2차원 투사된 픽처에 연산을 적용함으로써 획득됨 - 을 포함하는, 컴퓨팅 디바이스.

청구항 21

프로그래밍가능 장치에 대한, 컴퓨터 판독가능 저장 매체에 저장된 컴퓨터 프로그램으로서,

상기 프로그래밍가능 장치에 로딩되어 상기 프로그래밍가능 장치에 의해 실행되는 경우 제1항 및 제10항 중 어느 한 항에 따른 방법을 구현하기 위한 명령어들의 시퀀스를 포함하는, 컴퓨터 판독가능 저장 매체에 저장된 컴퓨터 프로그램.

청구항 22

삭제

청구항 23

삭제

청구항 24

삭제

청구항 25

삭제

청구항 26

삭제

청구항 27

삭제

발명의 설명

기술 분야

[0001] 본 발명은 미디어 데이터를 캡슐화하고 전송하기 위한 방법 및 디바이스에 관한 것이다.

배경 기술

[0002] 본 발명은, 예를 들어 MPEG 표준화 조직에 의해 정의된 ISO 베이스 미디어 파일 포맷에 따라, 가상 현실 미디어 콘텐츠를 캡슐화, 파싱 및 스트리밍하여, 가상 현실 미디어 콘텐츠의 교환, 관리, 편집 및 프리젠테이션을 용이화하는 유연하고 확장가능한 포맷을 제공하고, 예를 들어 적응형 http 스트리밍 프로토콜을 이용하여 인터넷 등의 IP 네트워크를 통한 전달을 향상시키는 것에 관한 것이다.

[0003] 국제 표준 기구 베이스 미디어 파일 포맷(ISO BMFF, ISO/IEC 14496-12)은, 로컬 저장을 위해 또는 네트워크나 또 다른 비트스트림 전달 메커니즘을 통한 전송을 위해 인코딩된 정시형 미디어 데이터 비트스트림(encoded timed media data bit-stream)을 기술하는 널리 알려진 유연하고 확장가능한 포맷이다. 확장의 예는, 다양한 NAL(Network Abstraction Layer) 유닛 기반 비디오 인코딩 포맷에 대한 캡슐화 도구를 기술하는 ISO/IEC 14496-15이다. 이러한 인코딩 포맷의 예는, AVC(Advanced Video Coding), SVC(Scalable Video Coding), HEVC(High Efficiency Video Coding), 및 L-HEVC(Layered HEVC)이다. 파일 포맷 확장의 또 다른 예는, HEVC Still 이미지 등의 스틸 이미지 또는 스틸 이미지들의 시퀀스에 대한 캡슐화 도구를 기술하는 ISO/IEC 23008-12이다. 이 파일 포맷은 객체 지향형이다. 이것은, 순차적으로 또는 계층적으로 조직화되고 타이밍 및 구조 파라미터 등의 인코딩된 정시형 명시 미디어 데이터 비트스트림의 파라미터를 정의하는 박스(또는 4 캐릭터 코드를 특징으로 하는 데이터 구조)라고 불리는 구축 블록들로 구성된다. 파일 포맷에서, 전체 프리젠테이션은 영화라고 불린다. 영화는, 미디어 또는 프리젠테이션 파일의 최상위 레벨의 (4 캐릭터 코드 'moov'를 갖는) 영화 박스에 의해 기술된다. 이 영화 박스는 프리젠테이션을 기술하는 한 세트의 다양한 박스를 포함하는 초기화 정보 컨테이너를 나타낸다. 이것은, (4개의 캐릭터 코드 'trak'을 갖는) 트랙 박스로 표현되는 트랙으로 논리적으로 분할된다. (트랙 식별자(track_ID)에 의해 고유하게 식별되는) 각각의 트랙은 프리젠테이션(예를 들어, 비디오 프레임)에 속하는 미디어 데이터의 정시형 시퀀스를 나타낸다. 각각의 트랙 내에서, 각각의 정시형 데이터 유닛은 샘플이라고 불린다; 이것은, 비디오, 오디오 또는 정시형 메타데이터의 프레임일 수 있다. 샘플은 순서에 있어서 묵시적으로 넘버링된다. 실제 샘플 데이터는 영화 박스와 동일한 레벨의 (4 캐릭터 코드 'mdat'를 갖는) Media Data Box라 불리는 박스에 저장된다. 영화는, 전체 프리젠테이션에 대한 정보를 포함하는 영화 박스와 후속되는 커플 영화 단편 및 Media Data box 목록으로서 일시적으로 조직화될 수 있다. 영화 단편(4 캐릭터 코드 'moof'를 갖는 박스) 내에는 영화 단편 당 0개 이상의 한 세트의 트랙 단편(4 캐릭터 코드 'traf'를 갖는 박스)가 있다. 결국, 트랙 단편은 0개 이상의 트랙 실행 박스('trun')를 포함하며, 그 각각의 박스는 그 트랙 단편에 대한 샘플들의 연속 실행을 문서화한다.

[0004] 파일 포맷에서, 미디어 또는 프리젠테이션 파일은 또한, 영화 박스와 동일한 레벨의 메타 박스('meta') 내에 기술된 하나 이상의 정적 항목(예를 들어, 하나 이상의 스틸 이미지)를 포함할 수 있다. 이 메타 박스는, 정적 항목을 기술하는 설명 정보 -이 설명 정보는 수개의 박스(예를 들어, 항목 정보 박스('iinf') 내의 항목들의 목록 및 항목 위치 박스('iloc') 내의 데이터 항목의 (데이터 박스 내에서의) 위치)로 조직화됨- 를 포함할 수 있고, 각각의 항목은 항목 식별자(item_ID)에 의해 고유하게 식별된다. 실제 항목 데이터는, 메타 박스 내의 항목 데이터 박스('idat')에 또는 파일 최상위 레벨의 미디어 데이터 박스('mdat')에 저장된다.

[0005] ISOBMFF 파일은 복수의 트랙(비디오 콘텐츠에 대한 서브픽처 트랙이라고도 함) 및/또는 복수의 정적 항목을 형성하는 복수의 인코딩된 정시형 미디어 데이터 비트스트림 또는 인코딩된 정시형 미디어 데이터 비트스트림의 서브파트를 포함할 수 있다. ISOBMFF 및 그 확장은, 트랙, 정적 항목 또는 샘플을 함께 그룹화하는 수개의 그룹화 메커니즘을 포함한다. 그룹은 전형적으로 공통 의미 및/또는 특성을 공유한다.

- [0006] 예를 들어, ISOBMFF는, 엔티티 그룹 메커니즘, 트랙 그룹 메커니즘, 및 샘플 그룹화 메커니즘을 포함한다. 엔티티 그룹화 메커니즘은, 표시된 그룹화 유형 또는 의미에 따라 트랙 및/또는 정적 항목이 그룹화됨을 나타내는데 이용될 수 있다. 트랙 그룹화 메커니즘은, 표시된 그룹화 유형 또는 의미에 따라 트랙이 그룹화됨을 나타내는데 이용될 수 있다. 샘플 그룹화 메커니즘은, 표시된 그룹화 유형 또는 의미와 연관된 소정의 속성이 트랙 내의 샘플들의 표시된 그룹에 적용됨을 나타내는데 이용될 수 있다.
- [0007] 사용자 경험을 개선하고 특히 몰입형 경험을 제공하기 위해, 정시형 미디어 데이터 비트스트림(비디오 및 오디오까지)은 전방향(또는 다중방향 또는 복수방향)일 수 있다. 360° 파노라마 비디오라고도 알려진 비디오에 적용될 때, 사용자는 디스플레이되는 장면에 위치해 있다고 느낀다.
- [0008] 전방향 비디오는, 360° 카메라로부터, 및/또는, 예를 들어 모든 카메라가 공통 노드 지점을 갖도록 특별한 장비에 장착된, 수개의 카메라로부터 획득된 비디오 스트림의 이미지를 결합함으로써 획득될 수 있다. 이러한 이미지 조합은 이미지 스티칭(image stitching) 또는 카메라 스티칭(camera stitching)이라고 알려져 있다.
- [0009] 이러한 전방향 비디오는, 사용자의 시정 방향에 따라, 머리 장착형 디스플레이를 통해 또는 사용자를 둘러싼 만곡형 스크린 상으로의 투사를 통해 렌더링될 수 있다. 이것은 또한, (뷰포트라고도 알려진) 전방향 비디오의 사용자가 원하는 부분에 따라 전방향 비디오로 이동하기 위한 내비게이션 사용자 인터페이스를 갖춘 전통적인 2D 스크린 상에 디스플레이될 수 있다. 이것은, 사용자가 가상 세계에 있는 것처럼 느끼기 때문에 종종 가상 현실(VR)이라고 한다. 전방향 비디오에 가상 객체가 추가되면, 이것은 증강 현실(AR)이라고 부른다.
- [0010] 본 발명자들은, 전송할 미디어 데이터에 관한 정보를 기술하고 시그널링할 때, 특히 전방향 미디어 콘텐츠가 복수의 트랙에 의해 운반되는 수개의 서브파트로 분할될 때 몇 가지 문제점에 주목했다.
- [0011] 한 예는, 클라이언트에게 특정한 파싱 프로세스를 요청하는 트랙의 시그널링을 수반하며, 이것은 오버헤드를 생성하고 복잡하다.
- [0012] 또 다른 예는, 트랙들의 그룹의 시그널링, 특히 원본 전방향 미디어 콘텐츠와 복수의 서브픽처 트랙에 임베딩된 2차원(2D) 미디어 콘텐츠(투사되거나, 캡핑되거나, 어안 코딩됨) 사이의 맵핑에 관한 것이다.
- [0013] 또 다른 예는, 디스플레이할 준비가 된 전방향 미디어 콘텐츠를 재구축하기 위해 결합이 허용되거나 허용되지 않는 서브픽처 트랙들의 시그널링을 수반한다. 기존의 솔루션은 복잡하거나 명확하게 정의되지 않았으며 2차원 복수의 트랙 캡슐화 프로세스에 대한 기존 메커니즘과 완전히 호환되지 않는다.

발명의 내용

해결하려는 과제

- [0014] 본 발명은 전송된 문제 중 하나 이상을 해결하기 위해 고안되었다.

과제의 해결 수단

- [0015] 이러한 맥락에서, 예를 들어 http 프로토콜을 이용하는 인터넷 등의 IP 네트워크를 통해, 미디어 콘텐츠(예를 들어, 전방향 미디어 콘텐츠를)를 스트리밍하기 위한 솔루션이 제공된다.
- [0016] 본 발명의 제1 양태에 따르면, 장면의 와이드 뷰(wide view)에 대응하는 인코딩된 미디어 데이터를 캡슐화하기 위한 방법이 제공되며, 이 방법은 다음과 같은 단계를 포함한다 :
- [0017] 장면의 와이드 뷰로부터 투사된 픽처를 획득하는 단계;
- [0018] 획득된 투사된 픽처를 적어도 하나의 팩킹된 픽처로 팩킹하는 단계;
- [0019] 적어도 하나의 팩킹된 픽처를 적어도 하나의 서브픽처로 분할하는 단계;
- [0020] 적어도 하나의 서브픽처를 복수의 트랙으로 인코딩하는 단계;
- [0021] 인코딩된 트랙들과 연관된 설명 메타데이터를 생성하는 단계,
- [0022] 여기서, 설명 메타데이터는 트랙으로 인코딩된 적어도 하나의 서브픽처와 적어도 하나의 투사된 픽처 사이의 공간적 관계를 나타내는 각각의 트랙과 연관된 정보 항목을 포함한다.
- [0023] 본 발명의 또 다른 양태에 따르면, 장면의 와이드 뷰에 대응하는 인코딩된 미디어 데이터를 캡슐화하기 위한 방

법이 제공되며, 이 방법은 다음과 같은 단계를 포함한다 :

- [0024] 장면의 와이드 뷰로부터 투사된 픽처를 획득하는 단계;
- [0025] 투사된 픽처를 적어도 하나의 서브픽처로 분할하는 단계;
- [0026] 적어도 하나의 서브픽처를 복수의 트랙으로 인코딩하는 단계;
- [0027] 인코딩된 트랙들과 연관된 설명 메타데이터 - 설명 메타데이터는 트랙으로 인코딩된 적어도 하나의 서브픽처와 기준 픽처 사이의 공간적 관계를 나타내는 각각의 트랙과 연관된 제1 정보 항목을 포함함 - 를 생성하는 단계;
- [0028] 설명 메타데이터는 기준 픽처를 나타내는 제2 정보 항목을 더 포함한다.
- [0029] 실시예들에 따르면, 투사된 픽처를 복수의 서브픽처로 분할하는 단계는, 투사된 픽처를 팩킹된 픽처로 팩킹하는 단계 및 팩킹된 픽처를 복수의 서브픽처로 분할하는 단계를 포함한다.
- [0030] 실시예들에 따르면, 제2 정보 항목은 기준 픽처가 투사된 픽처임을 나타내는 브랜드 값(brand value)이다.
- [0031] 실시예들에 따르면, 제2 정보 항목은 기준 픽처가 팩킹된 픽처임을 나타내는 브랜드 값이다.
- [0032] 실시예들에 따르면, 제2 정보 항목은 각각의 트랙과 연관된 제1 정보 항목에 포함된다.
- [0033] 실시예들에 따르면, 제2 정보 항목은 제1 정보 항목의 파라미터로서 정의된다.
- [0034] 실시예들에 따르면, 파라미터의 존재는 제1 정보 항목에 제공된 플래그에 의해 표시된다.
- [0035] 실시예들에 따르면, 제2 정보 항목은 제1 정보 항목에 제공되는 플래그로서 정의된다.
- [0036] 실시예들에 따르면, 제2 정보 항목은 서브픽처들에 대응하는 트랙들의 그룹의 속성들을 기술하기 위해 이용되는 특정한 유형의 그룹 정보로서 정의된다.
- [0037] 본 발명의 또 다른 양태에 따르면, 복수의 인코딩된 트랙 및 연관된 설명 메타데이터를 포함하는 미디어 파일로부터 적어도 하나의 이미지를 생성하는 방법이 제공되며, 이 방법은 다음과 같은 단계를 포함한다 :
- [0038] 복수의 인코딩된 트랙이 장면의 와이드 뷰의 투사된 픽처를 팩킹함으로써 획득된 팩킹된 픽처의 분할로부터 발생하는 적어도 하나의 서브픽처를 인코딩하는 트랙들의 그룹을 포함한다고 결정하는 단계;
- [0039] 트랙들의 그룹과 연관된 설명 메타데이터를 파싱하는 단계;
- [0040] 여기서, 트랙들의 그룹과 연관된 설명 메타데이터를 파싱하는 단계는 :
- [0041] 트랙으로 인코딩된 적어도 하나의 서브픽처와 적어도 하나의 투사된 픽처 사이의 공간적 관계를 나타내는 각각의 트랙과 연관된 정보 항목을 해석하는 단계를 포함한다.
- [0042] 본 발명의 또 다른 양태에 따르면, 복수의 인코딩된 트랙 및 연관된 설명 메타데이터를 포함하는 미디어 파일로부터 적어도 하나의 이미지를 생성하는 방법이 제공되며, 이 방법은 다음과 같은 단계를 포함한다 :
- [0043] 복수의 인코딩된 트랙이 장면의 와이드 뷰의 투사된 픽처의 분할로부터 발생하는 적어도 하나의 서브픽처를 인코딩하는 트랙들의 그룹을 포함한다고 결정하는 단계;
- [0044] 트랙들의 그룹과 연관된 설명 메타데이터를 파싱하는 단계;
- [0045] 여기서, 트랙들의 그룹과 연관된 설명 메타데이터를 파싱하는 단계는 다음과 같은 단계를 포함한다 :
- [0046] 트랙으로 인코딩된 적어도 하나의 서브픽처와 적어도 하나의 기준 픽처 사이의 공간적 관계를 나타내는 각각의 트랙과 연관된 제1 정보 항목을 해석하는 단계; 및
- [0047] 기준 픽처를 나타내는 제2 정보 항목을 해석하는 단계.
- [0048] 실시예들에 따르면, 장면의 와이드 뷰의 투사된 픽처의 분할은 투사된 이미지를 팩킹함으로써 획득된 팩킹된 픽처를 분할함으로써 획득된다.
- [0049] 실시예들에 따르면, 제2 정보 항목은 기준 픽처가 투사된 픽처임을 나타내는 브랜드 값이다.
- [0050] 실시예들에 따르면, 제2 정보 항목은 기준 픽처가 팩킹된 픽처임을 나타내는 브랜드 값이다.
- [0051] 실시예들에 따르면, 제2 정보 항목은 각각의 트랙과 연관된 제1 정보 항목에 포함된다.

- [0052] 실시예들에 따르면, 제2 정보 항목은 제1 정보 항목의 파라미터로서 정의된다.
- [0053] 실시예들에 따르면, 파라미터의 존재는 제1 정보 항목에 제공된 플래그에 의해 표시된다.
- [0054] 실시예들에 따르면, 제2 정보 항목은 제1 정보 항목에 제공되는 플래그로서 정의된다.
- [0055] 실시예들에 따르면, 제2 정보 항목은 서브픽처에 대응하는 트랙들의 그룹의 속성들을 기술하기 위해 이용되는 특정한 유형의 그룹 정보로서 정의된다.
- [0056] 본 발명의 또 다른 양태에 따르면, 장면의 와이드 뷰에 대응하는 인코딩된 미디어 데이터를 캡슐화하기 위한 컴퓨팅 디바이스가 제공되며, 이 컴퓨팅 디바이스는 다음을 위해 구성된다 :
 - [0057] 장면의 와이드 뷰로부터 투사된 픽처를 획득하는 단계;
 - [0058] 획득된 투사된 픽처를 적어도 하나의 팩킹된 픽처로 팩킹하는 단계;
 - [0059] 적어도 하나의 팩킹된 픽처를 적어도 하나의 서브픽처로 분할하는 단계;
 - [0060] 적어도 하나의 서브픽처를 복수의 트랙으로 인코딩하는 단계;
 - [0061] 인코딩된 트랙들과 연관된 설명 메타데이터를 생성하는 단계,
 - [0062] 여기서, 설명 메타데이터는 트랙으로 인코딩된 적어도 하나의 서브픽처와 적어도 하나의 투사된 픽처 사이의 공간적 관계를 나타내는 각각의 트랙과 연관된 정보 항목을 포함한다.
- [0063] 본 발명의 또 다른 양태에 따르면, 장면의 와이드 뷰에 대응하는 인코딩된 미디어 데이터를 캡슐화하기 위한 컴퓨팅 디바이스가 제공되며, 이 컴퓨팅 디바이스는 다음을 위해 구성된다 :
 - [0064] 장면의 와이드 뷰로부터 투사된 픽처를 획득하는 단계;
 - [0065] 투사된 픽처를 적어도 하나의 서브픽처로 분할하는 단계;
 - [0066] 적어도 하나의 서브픽처를 복수의 트랙으로 인코딩하는 단계;
 - [0067] 인코딩된 트랙들과 연관된 설명 메타데이터 - 설명 메타데이터는 트랙으로 인코딩된 적어도 하나의 서브픽처와 기준 픽처 사이의 공간적 관계를 나타내는 각각의 트랙과 연관된 제1 정보 항목을 포함함- 를 생성하는 단계;
 - [0068] 설명 메타데이터는 기준 픽처를 나타내는 제2 정보 항목을 더 포함한다.
- [0069] 본 발명의 또 다른 양태에 따르면, 복수의 인코딩된 트랙 및 연관된 설명 메타데이터를 포함하는 미디어 파일로부터 적어도 하나의 이미지를 생성하기 위한 컴퓨팅 디바이스가 제공되며, 이 컴퓨팅 디바이스는 다음을 위해 구성된다 :
 - [0070] 복수의 인코딩된 트랙이 장면의 와이드 뷰의 투사된 픽처를 팩킹함으로써 획득된 팩킹된 픽처의 분할로부터 발생하는 적어도 하나의 서브픽처를 인코딩하는 트랙들의 그룹을 포함한다고 결정하는 단계;
 - [0071] 트랙들의 그룹과 연관된 설명 메타데이터를 파싱하는 단계;
 - [0072] 여기서, 트랙들의 그룹과 연관된 설명 메타데이터를 파싱하는 단계는 다음과 같은 단계를 포함한다 :
 - [0073] 트랙으로 인코딩된 적어도 하나의 서브픽처와 적어도 하나의 투사된 픽처 사이의 공간적 관계를 나타내는 각각의 트랙과 연관된 정보 항목을 해석하는 단계를 포함한다.
- [0074] 본 발명의 또 다른 양태에 따르면, 복수의 인코딩된 트랙 및 연관된 설명 메타데이터를 포함하는 미디어 파일로부터 적어도 하나의 이미지를 생성하기 위한 컴퓨팅 디바이스가 제공되며, 이 컴퓨팅 디바이스는 다음을 위해 구성된다 :
 - [0075] 복수의 인코딩된 트랙이 장면의 와이드 뷰의 투사된 픽처의 분할로부터 발생하는 적어도 하나의 서브픽처를 인코딩하는 트랙들의 그룹을 포함한다고 결정하는 단계;
 - [0076] 트랙들의 그룹과 연관된 설명 메타데이터를 파싱하는 단계;
 - [0077] 여기서, 트랙들의 그룹과 연관된 설명 메타데이터를 파싱하는 단계는 다음과 같은 단계를 포함한다 :
 - [0078] 트랙으로 인코딩된 적어도 하나의 서브픽처와 적어도 하나의 기준 픽처 사이의 공간적 관계를 나타내는 각각의

트랙과 연관된 제1 정보 항목을 해석하는 단계; 및

[0079]

기준 픽처를 나타내는 제2 정보 항목을 해석하는 단계.

[0080]

본 발명의 또 다른 양태에 따르면, 프로그램가능한 장치를 위한 컴퓨터 프로그램 제품이 제공되며, 이 컴퓨터 프로그램 제품은, 프로그램가능한 장치에 로딩되고 실행될 때, 제1항 내지 제20항 중 어느 한 항에 따른 방법을 구현하기 위한 명령어 시퀀스를 포함한다.

[0081]

본 발명의 또 다른 양태에 따르면, 제1항 내지 제20항 중 어느 한 항에 따른 방법을 구현하기 위한 컴퓨터 프로그램의 명령어를 저장한 컴퓨터 판독가능한 저장 매체가 제공된다.

[0082]

본 발명의 또 다른 양태에 따르면, 실행시 제1항 내지 제20항 중 어느 한 항의 방법이 수행되게 하는 컴퓨터 프로그램이 제공된다.

도면의 간단한 설명

[0083]

본 발명의 추가 이점은 도면 및 상세한 설명을 검토할 때 본 기술분야의 통상의 기술자에게 명백하게 될 것이다. 임의의 추가 이점이 여기서 포함되도록 의도된다.

본 발명의 실시예들이, 단지 예로서, 및 다음과 같은 도면들을 참조하여 이하에서 설명된다 :

도 1은 서버로부터 클라이언트로의 전방향 비디오를 캡처, 처리, 캡슐화, 전송 및 렌더링하기 위한 데이터 흐름의 예를 나타낸다;

도 2는 본 발명의 실시예들에 따른 캡슐화의 예를 나타내는 블록도를 나타낸다;

도 3은 본 발명의 하나 이상의 실시예의 구현을 위한 컴퓨팅 디바이스의 개략적인 블록도이다;

도 4는 수개의 트랙 및 그룹 내의 상이한 미디어 소스들로부터의 미디어 데이터를 포함하는 서브픽처들을 코딩하는 예를 나타낸다;

도 5는 본 발명의 실시예들에 따른 SpatialRelationship2DdescriptionBox 및 source_id의 이용 예를 나타낸다;

도 6은 본 발명의 실시예들에 따른 SpatialRelationship2DdescriptionBox 및 source_id의 제2 이용 예를 나타낸다;

도 7은 본 발명의 실시예들에 따른 서브픽처 캡슐화를 나타낸다;

도 8은 본 발명의 실시예들에 따른 파싱 프로세스를 나타낸다;

도 9는 본 발명의 실시예들에 따른 시스템을 나타낸다;

도 10a, 10b, 10c 및 10d는 본 발명의 실시예들에 따른, 투사, 선택사향적 팩킹 및 서브픽처 트랙들로의 분할의 전체 프로세스의 여러 예를 나타낸다.

발명을 실시하기 위한 구체적인 내용

[0084]

도 1은 서버 디바이스(101)로부터 클라이언트 디바이스(170)(170'로서도 나타냄)로의 전방향 미디어를 캡처, 전송 및 렌더링하기 위한 데이터 흐름의 예를 나타낸다.

[0085]

예시된 바와 같이, 이 미디어는 카메라 시스템(100)으로부터 취득되고 머리 장착형 디스플레이(HMD) 170 및 170'에 전달되는 비디오 콘텐츠를 갖는다. 카메라 시스템(100)은 광각 렌즈를 갖는 하나의 카메라 또는 함께 조립된 복수의 카메라 세트(예를 들어, 가상 현실용 카메라 장비)를 포함할 수 있다. 전달(160)은 예를 들어 스트리밍 서버(161) 및 스트리밍 클라이언트(162)를 통해 적응형 http 스트리밍 프로토콜을 이용하는 인터넷 등의 IP 네트워크(163)를 통해 수행될 수 있다.

[0086]

예시의 목적을 위해, 이용된 카메라 시스템(100)은, 입방체의 각각의 면과 연관된 6개의 표준 카메라 세트에 기초한다. 이것은, 카메라 시스템을 둘러싼 실제 장면을 나타내는 이미지를 캡처(단계 110)하는데 이용된다. 이 배열에 따르면, 한 카메라는 전면 이미지를 제공하고, 한 카메라는 후면 이미지를 제공하고, 한 카메라는 좌측 이미지를 제공하고, 한 카메라는 우측 이미지를 제공하고, 한 카메라는 하단 이미지를 제공하고, 한 카메라는 상단 이미지를 제공한다.

[0087]

카메라 시스템(100)으로부터 획득된 이미지는 서버(101)에서 처리되어(단계 120) 360 비디오 스트림 또는 가상

현실 미디어 데이터 스트림이라고도 불리는 전방향 비디오 스트림을 형성하는 360 이미지를 생성한다.

- [0088] 처리 단계(120)는 동일한 시간 인스턴스의 캡처된 이미지를 스티칭 및 투사하는 것으로 구성된다. 이미지는 먼저 스티칭되어, 수평 및 수직 차원 양쪽 모두에서 360° 뷰를 형성하는 구면(121)을 나타내는 3차원 투사 구조물 상으로 투사된다. 투사 구조물 상의 360 이미지 데이터는 예를 들어 등장방형 투사(equirectangular projection)(https://en.wikipedia.org/wiki/Equirectangular_projection)을 이용하여 2차원 투사된 이미지(122)(캡처 투사라고도 함)로 추가로 변환된다. 투사된 이미지는 전체 구면을 덮는다.
- [0089] 대안으로서, 전방향 미디어가 양안 360도 비디오(stereoscopic 360-degree video)라면, 카메라 시스템(100)은 좌측 뷰와 우측 뷰를 나타내는 이미지 시퀀스를 단계 110에서 캡처하는 복수의 카메라로 구성될 수 있으며, 좌측 뷰와 우측 뷰는 나중에 클라이언트에 의해 3차원 360도 장면을 렌더링하는데 이용될 수 있다. 이러한 경우, 전술된 처리 단계(120)는 좌측 뷰 및 우측 뷰 이미지 시퀀스 양쪽 모두에 별개로 적용된다. 선택사항으로서, 단계 125에서, 프레임 팩킹이 적용되어 동일한 시간 인스턴스의 각각의 좌측 뷰 이미지 및 우측 뷰 이미지를 동일한 투사된 이미지로 팩킹하여 하나의 단일 좌측 + 우측 투사된 이미지 시퀀스를 생성할 수 있다. 예를 들어, 병렬, 상하, 열 기반 인터리빙, 행 기반 인터리빙, 교대하는 좌측 및 우측 뷰의 시간적 인터리빙 등의 수개의 양안 프레임 팩킹 구성이 가능하다. 대안으로서, 양안 프레임 팩킹 구성은 또한, 인코딩 단계(140) 후에 독립적인 비디오 비트스트림을 생성하는 별개의 독립적인 투사된 이미지 시퀀스로 좌측 및 우측 뷰를 유지하는 것으로 구성될 수 있다. 예를 들어, 하나의 비디오 비트스트림은 좌측 뷰 이미지를 나타내고 다른 하나는 우측 뷰 이미지를 나타낸다.
- [0090] 선택사항으로서, 그 다음, 영역별 팩킹(130)이 적용되어 투사된 이미지(122)를 팩킹된 이미지(131)에 맵핑한다. 영역별 팩킹은, 예를 들어 사용자에게 대한 구면의 가장 유용한 부분에 관한 신호 정보를 최대화하기 위해 투사된 이미지 영역의 변환(예를 들어, 회전, 미러링, 픽셀 블록의 복사 또는 이동 ...) 적용, 크기조정, 및 재배치로 구성된다. 팩킹된 이미지는 전체 구면의 일부만 덮을 수 있다는 점에 유의한다. 영역별 팩킹이 적용되지 않는 경우, 팩킹된 이미지(131)는 투사된 이미지(122)와 동일하다. 양안 전방향 미디어의 경우, 영역별 팩킹은, 단계 125에서 선택한 프레임 팩킹 구성에 따라, 좌측 + 우측 투사된 이미지 시퀀스에 적용되거나, 좌측 뷰 및 우측 뷰 투사된 이미지 시퀀스에 별개로 적용된다.
- [0091] 투사된 이미지(122) 또는 팩킹된 이미지(131)는 단계 140에서 하나 또는 수개의 비디오 비트스트림으로 인코딩된다. 양안 전방향 미디어의 경우, 인코딩 단계는, 단계 125에서 선택된 프레임 팩킹 구성에 따라, 좌측 + 우측 팩킹된 이미지 시퀀스에 적용되거나, 좌측 뷰 및 우측 뷰 팩킹된 이미지 시퀀스에 별개로 적용된다. 대안으로서, 멀티 뷰 인코딩은 좌측 뷰 및 우측 뷰 팩킹된 이미지 시퀀스에서 이용될 수 있다.
- [0092] 인코딩 포맷의 예는, AVC(Advanced Video Coding), SVC(Scalable Video Coding), HEVC(High Efficiency Video Coding) 또는 L-HEVC(Layered HEVC)이다. 이하에서, HEVC는 HEVC와 그 계층화된 확장(L-HEVC) 양쪽 모두를 참조하는데 이용된다.
- [0093] HEVC 및 유사한 비디오 인코딩 포맷은, 샘플의 상이한 공간적 분할, 예를 들어, 픽처 : 타일, 슬라이스 및 슬라이스 세그먼트를 정의한다. 타일은, 수평 및 수직 경계(즉, 행 및 열)에 의해 정의되고, 이하에서는 모두 코딩 유닛이라고 지칭되는, 정수 개수의 코딩 트리 유닛(Coding Tree Unit; CTU) 또는 코딩 블록을 포함하는 픽처의 직사각형 영역을 정의한다. 따라서, 타일은 픽처의 공간적 서브파트를 나타내는 양호한 후보이다. 그러나, 구문 및 그 NAL 유닛(또는 NALU)으로의 캡슐화의 측면에서 코딩된 비디오 데이터(비트스트림) 조직화는(AVC에서와 같이) 오히려 슬라이스 및 슬라이스 세그먼트에 기초한다.
- [0094] HEVC에서의 슬라이스는 슬라이스 세그먼트 세트이며, 여기서, 적어도 첫 번째 슬라이스 세그먼트는 독립 슬라이스 세그먼트이고, 있다면, 다른 것들은 종속 슬라이스 세그먼트이다. 슬라이스 세그먼트는 정수 개수의 (래스터 스캔 순서에서) 연속된 CTU를 포함한다. 슬라이스는 반드시 직사각형 형상을 가질 필요는 없다(따라서 공간적 서브파트 표현을 위해 타일보다 적합하지 않음). 슬라이스 세그먼트는, *slice_segment_header*에 후속되는 *slice_segment_data*로서 HEVC 비트스트림으로 인코딩된다. 독립 슬라이스 세그먼트(ISS)와 종속 슬라이스 세그먼트(DSS)는 헤더에 의해 구분된다 : 종속 슬라이스 세그먼트는 독립 슬라이스 세그먼트 헤더로부터의 정보를 재사용하기 때문에 더 짧은 헤더를 갖는다. 독립 및 종속 슬라이스 세그먼트 양쪽 모두는 비트스트림에서의 엔트리 포인트 목록을 포함한다.
- [0095] 비디오 비트스트림이 타일로 인코딩된 경우, 타일이 동일한 픽처 내의 이웃 타일에 의존(공간적 종속성)하지 않고 이전 기준 픽처 내의 이웃 타일에 의존(시간적 종속성)하지 않도록 보장하기 위해 타일은 움직임-제약된다.

따라서, 움직임-제한된 타일은 독립적으로 디코딩될 수 있다.

- [0096] 대안으로서, 투사된 이미지(122) 또는 팩킹된 이미지(131)는 인코딩 전에 수개의 공간적 서브픽처로 분할될 수 있고, 각각의 서브픽처는 독립적으로 인코딩되어 예를 들어 독립 인코딩된 HEVC 비트스트림을 형성한다.
- [0097] 대안으로서, 영역별 팩킹 단계(130) 및 수개의 공간적 서브픽처로 분할하는 단계는 메모리에서 완전한 중간 팩킹된 이미지(131)를 생성하지 않고 동시에 수행될 수 있다. 투사된 이미지(122)(또는 선택사항적인 단계 125 이후의 결과적인 양안 투사된 이미지)는 서브파트들로 분할될 수 있고, 각각의 서브파트는 단계 140에서 인코딩될 공간적 서브픽처로 직접 팩킹될 수 있다.
- [0098] 도 10a, 10b, 10c 및 10d는 본 발명의 실시예들에 따른, 투사, 선택사항적 팩킹 및 서브픽처 트랙들로의 분할 전체 프로세스의 여러 예를 나타낸다. 투사된 픽처(1001)(1, 2, 3 및 4로 표시됨)로부터의 하나 이상의 영역은 수개의 변환 연산(향등, 업 또는 다운 스케일링, 회전, 미러링, 재배치 ...)을 적용함으로써 팩킹된 영역(1002(1', 2', 3' 및 4'로 표시))으로 재배열된 다음, 하나 이상의 서브픽처 트랙(1003)으로 분할 및 재조직화된다. 분할은 팩킹된 영역(1', 2', 3' 또는 4') 당 하나의 서브픽처 트랙으로 이어질 수도 있다. 팩킹 및 분할 동작은 또한, 투사된 픽처(1011)로부터 하나 이상의 서브픽처 트랙(1012)으로 직접, 한번에 수행될 수도 있다. 도 10c 및 도 10d는 전방향 콘텐츠가 스테레오 콘텐츠인 경우 상이한 가능한 캡슐화의 예를 제공한다. 이러한 경우, 캡처 단계(110)는 양안 레코딩을 허용하는 카메라 장비를 이용하며, 양안 레코딩은 일반적으로 육안 당 하나의 비디오이다.
- [0099] 도 10c는 프레임 팩킹이 없는 양안 전방향 콘텐츠의 예를 묘사한다(도 1의 선택사항적인 단계 125). 그 다음, 각각의 투사된 뷰(1021)는, (1022에서) 영역별 팩킹이 각각의 뷰에 적용될 때 아마도 1023처럼 복수의 서브픽처 트랙으로 독립적으로 캡슐화된다. 이 예에서는, 각각의 뷰의 영역 당 하나의 서브픽처 트랙이 있다. 동일한 서브픽처 트랙 내의 동일한 영역의 양쪽 모두의 뷰를 캡슐화하기로 결정할 수도 있다. 그 다음, 서브픽처 트랙은 이용된 프레임 팩킹을 나타내는 샘플 설명 레벨의 스테레오 비디오 박스를 포함할 것이다.
- [0100] 도 10d는, 2개의 투사된 뷰(1031)를 단일 프레임-팩킹된 픽처(1032)에서 팩킹하기 위해 적용된 프레임 팩킹(도 1의 단계 125)이 있는 양안 전방향 콘텐츠의 예를 도시한다. 그 다음, 결과적인 프레임-팩킹된 픽처(1032)는, 아마도 1033처럼 복수의 서브픽처 트랙으로 캡슐화된다. 이 예에서, 각각의 서브픽처 트랙은, 주어진 공간적 영역에 대한 양쪽 모두의 뷰를 기술한다. 팩킹이 뒤따르는 투사의 경우, 하나의 서브픽처 트랙은 (도 10에 도시된 바와 같이) 하나의 영역 또는 많은 영역을 캡슐화할 수 있다. 캡슐화 모듈은, 예를 들어 복수의 팩킹된 영역을 포함하는 서브픽처 트랙들로 콘텐츠를 캡슐화하기 위해 설명 비용 대 액세스 세분성 절충(description cost versus access granularity trade-off)을 결정할 수 있다. 이것은, 팩킹된 영역의 역투사 계산에 의한 캡슐화가 팩킹된 프레임 내의 연속된 영역들의 역투사에서 갭이 없음을 발견한 경우일 수 있다. 이것은, 팩킹된 픽처로부터의 이들 영역을 단일 서브픽처 트랙으로 그룹화하는 결정 기준일 수 있다. 도 10a, 도 10b, 도 10c 및 도 10d는 동일한 서브픽처 트랙 내의 수개의 영역의 이러한 수집을 나타낸다. 캡슐화 모듈이, 투사된 픽처에서 갭, 홀 또는 덮이지 않은 픽셀을 생성하는 서브픽처 트랙 내의 복수의 영역을 수집하는 경우, 서브픽처 트랙 위치 및 크기를 이들 복수의 영역의 경계 박스의 위치 및 크기와 동일하게 설정할 수 있다.
- [0101] 따라서, 인코딩 단계(140)의 결과로서, 투사된 이미지(122) 또는 팩킹된 이미지(131)는 하나 이상의 독립적인 인코딩된 비트스트림에 의해 또는 하나 이상의 독립적으로 인코딩된 서브비트스트림으로 구성된 적어도 하나의 인코딩된 비트스트림에 의해 표현될 수 있다.
- [0102] 그 다음, 이들 인코딩된 비트스트림 및 서브비트스트림은, 단계 150에서 파일로, 또는 캡슐화 파일 포맷에 따라, 예를 들어 MPEG 표준화 기구에 의해 정의된 ISO Base Media File Format and Omnidirectional Media Format(OMAF - ISO/IEC 23090-2)에 따라 작은 시간 세그먼트 파일(165)로 캡슐화된다. 결과적인 파일 또는 세그먼트 파일은 mp4 파일 또는 mp4 세그먼트일 수 있다. 캡슐화 동안, 오디오 스트림은 비디오 또는 오디오 스트림에 관한 정보를 제공하는 메타데이터 트랙뿐만 아니라 비디오 비트스트림에 추가될 수 있다.
- [0103] 그 다음, 캡슐화된 파일 또는 세그먼트 파일은, 전달 메커니즘(160)을 통해, 예를 들어 http(HyperText Transfer Protocol) 프로토콜을 이용하는 인터넷을 통해 또는 예를 들어 디스크 등의 착탈식 디지털 매체를 통해 클라이언트(170)에 전달된다. 예시의 목적을 위해, 전달(160)은 MPEG 표준화 위원회로부터의 DASH(Dynamic Adaptive Streaming over HTTP)("ISO/IEC 23009-1, Dynamic adaptive streaming over HTTP (DASH), Part1: Media presentation description and segment formats") 등의 HTTP를 통한 적응 스트리밍을 이용하여 수행된다.

- [0104] 이 표준은, 미디어 프리젠테이션의 미디어 콘텐츠에 대한 간략한 설명을 HTTP URL(Uniform Resource Location)들과 연관시키는 것을 가능케한다. 이러한 연관, 일반적으로 매니페스트 파일(manifest file) 또는 설명 파일 164라 불리는 파일에서 기술된다. DASH의 맥락에서, 이 매니페스트 파일은 MPD 파일(Media Presentation Description)이라고도 하는 XML 파일이다.
- [0105] MPD 파일을 수신함으로써, 클라이언트 디바이스(170)는 각각의 미디어 콘텐츠 컴포넌트의 설명을 획득한다. 따라서, 이것은, 미디어 프리젠테이션에서 제안된 미디어 콘텐츠 컴포넌트의 종류를 알고 있고 스트리밍 클라이언트(162)를 통해 스트리밍 서버(161)로부터 연관된 미디어 세그먼트(165)를 다운로드하는데 이용될 HTTP URL을 안다. 따라서, 클라이언트(170)는, (HTTP 요청을 통해) 어떤 미디어 콘텐츠 컴포넌트를 다운로드하고 재생할 것인지를(즉, 미디어 세그먼트를 수신한 후 디코딩하고 재생할 것인지를) 결정할 수 있다.
- [0106] 클라이언트 디바이스는, 사용자의 뷰포트에 따라 장면의 와이드 뷰를 나타내는 전체 팩킹된 이미지의 공간적 부분에 대응하는 미디어 세그먼트만(즉, 현재 디스플레이되고 사용자가 보는 구면형 비디오의 일부)을 획득할 수 있다. 장면의 와이드 뷰는 전체 팩킹된 이미지에 의해 표현되는 전체 뷰를 나타낼 수 있다.
- [0107] 수신시, 캡슐화된 가상 현실 미디어 파일 또는 미디어 세그먼트는, 단계 172에서 디코딩되는 하나 이상의 데이터 스트림을 추출하도록 단계 171 동안에 파싱된다. 단계 171에서 수신된 ISOBMFF 파일 또는 세그먼트의 경우, 파싱은 전형적으로, 설명 메타데이터로부터 캡슐화된 비디오 비트스트림 및/또는 비디오 서브비트스트림을 추출할 수 있는 mp4 판독기 또는 mp4 파서에 의해 처리된다.
- [0108] 그 다음, 선택사항으로서, 단계 173에서, 디코딩 단계(172)로부터 발생하는 팩킹된 이미지 또는 팩킹된 서브이미지가 언팩킹되어 투사된 이미지가 획득되고, 이것은 비디오 렌더링을 위해 처리되고(단계 174) 디스플레이된다(단계 175).
- [0109] 대안으로서, 팩킹된 서브이미지는 투사된 픽처로 언팩킹되기 전에 중간 전체 팩킹된 이미지를 구성하도록 재배열될 수 있다.
- [0110] 비디오 렌더링은, 사용자의 관점, 시점 및 투사된 이미지를 생성하는데 이용되는 투사(들) 등의 수개의 파라미터에 의존한다는 점에 유의해야 한다. 예시된 바와 같이, 비디오를 렌더링하는 단계는 디코딩된 투사된 이미지를 구면에 재-투사하는 단계를 포함한다. 이러한 재투사로부터 획득된 이미지는 머리 장착형 디스플레이(170')에서 디스플레이된다.
- [0111] 양안 뷰를 처리하기 위해, 도 1을 참조하여 설명된 프로세스는 중복되거나 부분적으로 중복될 수 있다.
- [0112] UHD(Ultra High Definition) 비디오 스트림의 수개의 이미지를 가상 현실 미디어 데이터 스트림의 파노라마 이미지로 스티칭하는 것은, 매우 높은 비트레이트와 매우 높은 해상도의 가상 현실 미디어 데이터 스트림으로 이어진다는 것이 관찰되었다. 따라서, 시스템의 관점으로부터 및 대역폭 낭비를 방지하고 클라이언트 재생기의 처리 능력과 호환되기 위해, 가상 현실 미디어 데이터에 대한 액세스를 최적화할 필요가 있다.
- [0113] 이러한 필요성은, 가상 현실 미디어 데이터 스트림이 도 1을 참조하여 설명된 것과는 다른 목적으로 이용될 수 있다는 것보다 훨씬 더 중요하다. 특히, 가상 현실 미디어 데이터 스트림은, 360° 프로젝터 어레이 등의 특정한 디스플레이로 360° 이미지를 디스플레이하는데 이용될 수 있다. 이것은 또한, 특정한 시야 및/또는 관점을 디스플레이하거나, 관점, 시야, 및 관점을 변경하는데 이용될 수 있다.
- [0114] 특정한 실시예들에 따르면, 팩킹된 이미지(131)(도 1의 단계 140)의 인코딩으로부터 발생하는 인코딩된 비트스트림 및 서브비트스트림은, 캡슐화 파일 포맷, 예를 들어 ISO 베이스 미디어 파일 포맷(ISO/IEC 14496-12 및 ISO/IEC 14496-15), 전방향 미디어 포맷(Omnidirectional Media Format)(OMAF)(ISO/IEC 23090-2) 및 MPEG 표준화 기구에 의해 정의된 연관된 명세에 따라, 파일 또는 작은 시간 세그먼트 파일로 캡슐화된다.
- [0115] 인코딩된 비트스트림(예를 들어, HEVC) 및 아마도 그 서브비트스트림(예를 들어, 타일형 HEVC, MV-HEVC, 스케일링가능한 HEVC)은 하나의 단일 트랙으로 캡슐화될 수 있다. 대안으로서, 공간적으로 관련된(즉, 투사된 이미지의 서브-공간적 파트인) 복수의 인코딩된 비트스트림은 수개의 서브픽처 트랙으로 캡슐화될 수 있다. 대안으로서, 수개의 서브비트스트림(타일, 뷰, 계층)을 포함하는 인코딩된 비트스트림(예를 들어, 타일형 HEVC, MV-HEVC, 스케일링가능한 HEVC)은 복수의 서브픽처 트랙으로서 캡슐화될 수 있다.
- [0116] 서브픽처 트랙은, 전형적으로 픽처 또는 이미지의 공간적 파트 또는 직사각형 영역인, 서브파트에 대한 데이터를 임베딩하는 트랙이다. 서브픽처 트랙은, 다른 서브픽처 트랙 또는 서브픽처가 추출되는 전체 픽처를 기술하는 트랙과 관련될 수 있다. 예를 들어 서브픽처 트랙은 타일 트랙일 수 있다. 이것은, AVC 트랙, HEVC 트랙,

HEVC 타일 트랙 또는 샘플들의 시퀀스로서 캡슐화된 임의의 압축된 비디오 비트스트림에 의해 표현될 수 있다.

[0117] 타일 트랙은, 이미지의 공간적 부분에 또는 이미지나 픽처의 서브픽처에 대응하는 정사형 비디오 샘플들의 시퀀스이다. 이것은, 예를 들어 이미지 내의 관심 영역 또는 이미지 내의 임의 영역일 수 있다. 타일 트랙에 대응하는 데이터는 비디오 비트스트림으로부터 나오거나 비디오 비트스트림의 서브파트로부터 나올 수 있다. 예를 들어, 타일 트랙은 AVC 또는 HEVC 호환 비트스트림이거나, AVC 또는 HEVC의 서브파트, 또는 예를 들어 HEVC 타일 등의 임의의 인코딩된 비트스트림일 수 있다. 바람직한 실시예에서, 타일 트랙은 독립적으로 디코딩 가능하다(인코더는 "움직임-제약된" 타일을 생성함으로써 다른 타일로부터의 움직임 예측을 제거하도록 주의를 기울였다). 타일 트랙이 타일과 함께 HEVC로 인코딩된 비디오 비트스트림에 대응할 때, ISO/IEC 14496-15 제4 판에 설명된 바와 같이 'hvt1' 트랙으로서 표기되는 HEVC 타일 트랙으로 캡슐화될 수 있다. 그 다음, 파라미터 세트를 획득하기 위해 타일 베이스 트랙을 참조할 수 있고, 비디오 디코더를 셋업하기 위해 상위 레벨 정보를 참조할 수 있다. 이것은 또한, HEVC 트랙 'hvc1' 또는 'hev1' 트랙으로 캡슐화될 수 있다. 타일 트랙은, 더 큰 이미지나 픽처로의 서브픽처의 공간적 합성에 이용될 수 있다.

[0118] 타일 베이스 트랙은, 하나 이상의 타일 트랙에 공통되고 이들 하나 이상의 트랙에 의해 공유되는 데이터 또는 메타데이터를 포함하는 트랙이다. 타일 베이스 트랙은 하나 이상의 타일 트랙으로부터의 이미지들을 합성하는 명령어를 포함할 수 있다. 타일 트랙은, 완전한 디코딩 또는 렌더링을 위해 타일 베이스 트랙에 의존할 수 있다. 타일 베이스 트랙이 타일과 함께 HEVC로 인코딩된 비디오 비트스트림으로부터 유도되면, 이것은 'hvc2' 또는 'hev2' 트랙으로서 표기되는 HEVC 트랙으로 캡슐화된다. 또한 이것은 트랙 기준 'tbas'를 통해 HEVC 타일 트랙에 의해 참조되며 ISO/IEC 14496-15 제4 판에 기술된 바와 같이 HEVC 타일 트랙에 대한 'sabt' 트랙 참조를 이용하여 타일 순서를 표시해야 한다.

[0119] (기준 트랙이라고도 하는) 합성 트랙은 이미지를 합성하기 위해 다른 트랙을 참조하는 트랙이다. 합성 트랙의 한 예는, 비디오 트랙의 경우, 서브픽처 트랙들을 더 큰 이미지로 합성하는 트랙이다. 이것은, 예를 들어, 각각의 비디오 트랙으로부터의 이미지들을 더 큰 이미지로 합성하기 위해 변환 및 변환 파라미터를 제공하는 비디오 트랙으로부터 유도되는 트랙에서, 디코딩 후 동작에 의해 이루어질 수 있다. 합성 트랙은 또한, 다른 비디오 트랙으로부터 NAL 유닛을 추출하는 명령어를 제공하는 추출기 NAL 유닛을 갖는 트랙이거나 서브비트스트림 연결로부터 발생하는 비트스트림을 디코딩하기 전에 형성할 타일 트랙일 수 있다. 합성 트랙은 또한, 예를 들어 다른 트랙에 대한 트랙 참조를 통해, 합성 명령을 명시적으로 제공하는 트랙일 수도 있다.

[0120] ISO/IEC 14496-12는 트랙 레벨(즉, ISOBMMFF 박스 계층 구조의 'trak' 박스 내)에 위치한 'trgr'로 표기된 박스를 제공하여 트랙 그룹들을 기술하고, 여기서, 각각의 그룹은 특정한 특성을 공유하거나, 그룹 내의 트랙들은 특정한 관계를 갖는다. 이 트랙 그룹 박스는 다음과 같이 정의된 빈 컨테이너이다 :

```
Box Type: 'trgr'
Container: TrackBox('trak')
Mandatory: No
Quantity: Zero or one
aligned(8) class TrackGroupBox extends Box('trgr') {
}
```

[0121]

[0122] 이 트랙 그룹 박스는 다음과 같이 정의된 한 세트의 트랙 그룹 유형 박스를 포함할 수 있다 :

```
aligned(8) class TrackGroupTypeBox(unsigned int(32) track_group_type)
    extends FullBox(track_group_type, version = 0, flags = 0)
{
    unsigned int(32) track_group_id;
    // the remaining data may be specified for a particular
    track_group_type
}
```

[0123]

[0124] 이 트랙 그룹 유형 박스의 인스턴스에 의해 선언된 특정한 특성 또는 관계는 박스 유형(track_group_type)에 의해 표시된다. 이 박스는 또한, 동일한 트랙 그룹에 속하는 트랙을 결정하는데 이용할 수 있는 식별자(track_group_id)를 포함한다. 동일한 track_group_type 및 track_group_id 값을 가진 트랙 그룹 유형 박스의 트랙 그룹 박스를 갖는 모든 트랙은 동일한 트랙 그룹의 일부이다. 이 박스는 또한, 특정한 트랙 그룹 유형에

대한 트랙과 연관된 특정한 파라미터의 선언을 허용한다. MPEG ISOBMFF 표준(ISO/IEC 14496-12 7th edition Amendment 1 - 2018년 5월)은, 유형 '2dcc'의 TrackGroupTypeBox로서의 2차원 공간적 관계에 대한 특정한 트랙 그룹 SpatialRelationship2DDescriptionBox를 제안한다.

[0125] '2dcc'와 동일한 track_group_type을 갖는 SpatialRelationship2DDescriptionBox TrackGroupTypeBox는, 이 트랙이 (예를 들어, 비디오 소스의 평면 공간적 파트에 대응하는) 2D 공간적 관계를 갖는 트랙들의 그룹에 속함을 나타낸다. 주어진 track_group_id를 갖는 SpatialRelationship2DDescriptionBox TrackGroupTypeBox는, 임의의 원점(0, 0)과 total_width 및 total_height에 의해 정의된 최대 크기를 갖는 좌표계를 묵시적으로서 정의한다; x 축은 좌측에서 우측으로, y 축은 위에서 아래로 향한다. SpatialRelationship2DDescriptionBox TrackGroupTypeBox 내에서 동일한 값의 source_id를 갖는 트랙들은 동일한 소스로부터 시작되는 것으로 맵핑되고, 그들의 연관된 좌표계는 동일한 원점(0, 0) 및 축 배향을 공유한다. 소스 또는 비디오 소스는 전방향 콘텐츠를 위한 카메라 또는 한 세트의 카메라 세트에 의해 캡처되는 콘텐츠에 대응한다. 예를 들어, 매우 높은 해상도의 비디오는 서브픽처 트랙들로 분할될 수 있다. 그 다음, 각각의 서브픽처 트랙은 소스 비디오에서 그 위치와 크기를 운반한다.

[0126] 동일한 source_ID를 가진 동일한 트랙 그룹 내의 트랙들은 동일한 output_width 및 output_height를 선언해야 한다.

[0127] 유형 '2dcc'의 2차원 공간적 관계 트랙 그룹은 다음과 같이 정의된다 :

```
aligned(8) class SpatialRelationship2DSourceBox
    extends FullBox('2dsr', 0, 0) {
    unsigned int(32) total_width;
    unsigned int(32) total_height;
    unsigned int(32) source_id;
}

aligned(8) class SubPictureRegionBox extends FullBox('sprg',0,0) {
    unsigned int(16) object_x;
    unsigned int(16) object_y;
    unsigned int(16) object_width;
    unsigned int(16) object_height;
}

aligned(8) class SpatialRelationship2DDescriptionBox extends
TrackGroupTypeBox('2dcc')
{
    // track_group_id is inherited from TrackGroupTypeBox;
    SpatialRelationship2DSourceBox(); // mandatory, must be first
    SubPictureRegionBox (); // optional
}
```

[0128] 여기서

[0130] object_x는 둘러싸는 트랙 그룹에 의해 명시된 영역 내에서의 트랙의 좌상귀의 수평 위치를 명시한다. 위치 값은, 0 내지 total_width-1(양쪽 끝값 포함) 범위의, 있다면, 트랙 폭 및 높이에 의해 야기되는 묵시적 리샘플링을 적용하기 전의 값이고, 여기서 total_width는 둘러싸는 트랙 그룹에 의해 정의된다.

[0131] object_y는 둘러싸는 트랙 그룹에 의해 명시된 영역 내에서의 트랙의 좌상귀의 수직 위치를 명시한다. 위치 값은, 0 내지 total_height-1(양쪽 끝값 포함) 범위의, 있다면, 트랙 폭 및 높이에 의해 야기되는 묵시적 리샘플링을 적용하기 전의 값이고, 여기서 total_height는 둘러싸는 트랙 그룹에 의해 정의된다.

[0132] object_width는 둘러싸는 트랙 그룹에 의해 명시된 영역 내에서의 트랙의 폭을 명시한다. 위치 값은, 1 내지 total_width-1(양쪽 끝값 포함) 범위의, 있다면, 트랙 폭 및 높이에 의해 야기되는 묵시적 리샘플링을 적용하기 전의 값이고, 여기서 total_width는 둘러싸는 트랙 그룹에 의해 정의된다.

[0133] object_height는 둘러싸는 트랙 그룹에 의해 명시된 영역 내에서의 트랙의 높이를 명시한다. 위치 값은, 1 내

지 total_height-1(양쪽 끝값 포함) 범위의, 트랙 너비 및 높이에 의해 야기되는 목시적 리샘플링을 적용하기 전의 값이고, 여기서 total_height는 둘러싸는 트랙 그룹에 의해 정의된다.

[0134] total_width는 'srd' 트랙 그룹의 좌표계에서 최대 폭을 픽셀 단위로 명시한다. total_width의 값은 동일한 값의 track_group_id를 갖는 SpatialRelationshipDescriptionBox의 모든 인스턴스에서 동일해야 한다.

[0135] total_height는 'srd' 트랙 그룹의 좌표계에서 최대 높이를 픽셀 단위로 명시한다. total_height의 값은 동일한 값의 track_group_id를 갖는 SpatialRelationshipDescriptionBox의 모든 인스턴스에서 동일해야 한다.

[0136] source_id 파라미터는 소스에 대한 고유 식별자를 제공한다. 이것은 이 소스와 연관된 좌표계를 목시적으로서 정의한다.

[0137] SubPictureRegionBox()는, 둘러싸는 트랙 그룹에 의해 명시된 영역 내에서 트랙의 정적 위치 및 크기를 제공하는 선택사항적인 박스이다.

[0138] SubPictureRegionBox()가 SpatialRelationship2DDescriptionBox에 존재한다면, 연관된 트랙에는 어떠한 연관된 SpatialRelationship2DGroupEntry도 없어야 한다(이 트랙은, 상수의, 정적인, 크기 및 위치를 갖는다).

[0139] SubPictureRegionBox()가 SpatialRelationship2DDescriptionBox에 존재하지 않는다면, 연관된 트랙에는 하나 이상의 연관된 SpatialRelationship2DGroupEntry(들)가 있어야 한다(이 트랙은 아마도 동적인 크기 및/또는 위치를 갖는다).

[0140] '2dcc' 샘플 그룹화를 정의하는 SpatialRelationship2DGroupEntry()는, 2차원 공간적 관계 트랙 그룹 내의 서브픽처 트랙으로부터의 샘플의 위치와 크기를 선언하는 것을 허용한다. grouping_type이 '2dcc'와 같을 때, SampleToGroupBox 버전 1이 이용되어야 한다. grouping_type_parameter의 값은 대응하는 공간적 관계 트랙 그룹의 track_group_id와 같아야 한다.

[0141] SpatialRelationship2DGroupEntry()는 다음과 같이 정의된다 :

```
class SpatialRelationship2DGroupEntry () extends
VisualSampleGroupEntry ('2dcc') {
    unsigned int(16) object_x;
    unsigned int(16) object_y;
    unsigned int(16) object_width;
    unsigned int(16) object_height;
```

[0142] }

[0143] 여기서

[0144] object_x는 대응하는 공간적 관계 트랙 그룹에 의해 명시된 좌표계 내에서 이 그룹의 샘플의 좌상귀의 수평 위치를 명시한다. 위치 값은, 0 내지 total_width-1(양쪽 끝값 포함) 범위의, 있다면, 트랙 폭 및 높이에 의해 야기되는 목시적 리샘플링을 적용하기 전의 값이고, 여기서 total_width는 대응하는 SpatialRelationship2DDescriptionBox에 포함된다.

[0145] object_y는, 대응하는 공간적 관계 트랙 그룹에 의해 명시된 좌표계 내에서 이 그룹의 샘플의 좌상귀의 수직 위치를 명시한다. 위치 값은, 0 내지 total_height-1(양쪽 끝값 포함) 범위의, 있다면, 트랙 폭 및 높이에 의해 야기되는 목시적 리샘플링을 적용하기 전의 값이고, 여기서 total_height는 대응하는 SpatialRelationship2DDescriptionBox에 포함된다.

[0146] object_width는 대응하는 공간적 관계 트랙 그룹에 의해 명시된 좌표계 내에서 이 그룹의 샘플의 폭을 명시한다. 위치 값은, 1 내지 total_width(양쪽 끝값 포함) 범위의, 있다면, 트랙 폭 및 높이에 의해 야기되는 목시적 리샘플링을 적용하기 전의 값이고,

[0147] object_height는 대응하는 공간적 관계 트랙 그룹에 의해 명시된 좌표계 내에서 이 그룹의 샘플의 높이를 명시한다. 위치 값은, 1 내지 total_height(양쪽 끝값 포함) 범위의, 있다면, 트랙 폭 및 높이에 의해 야기되는 목시적 리샘플링을 적용하기 전의 값이다.

[0148] '2dcc' 트랙 그룹 내의 각각의 트랙의 샘플은 더 큰 이미지를 생성하기 위해 이 동일한 그룹 내의 다른 트랙으로부터의 샘플과 함께 (동일한 합성 또는 디코딩 시간에서) 공간적으로 합성될 수 있다.

- [0149] 팩킹된 이미지(131)의 인코딩(도 1의 단계 140)으로부터 발생하는 인코딩된 비트스트림 및 서브비트스트림에 따라, 파일 포맷에서의 캡슐화의 수개의 변형이 가능하다.
- [0150] 도 2는 본 발명의 한 실시예에 따른 파일/세그먼트 캡슐화(도 1의 단계 150)의 예를 나타내는 블록도를 예시한다.
- [0151] 단계 200에서, 서버는, 수개의 공간적으로-관련된 비디오 비트스트림(즉, 팩킹된 이미지의 공간적 서브파트를 나타내며 그 공간적 합성이 더 큰 이미지를 생성할 수 있음)이 있는지 또는 복수의 서브픽처 트랙으로서 클라이언트에게 노출될 수 있는 움직임-제한된 타일 또는 복수의 뷰를 나타내는 비디오 서브비트스트림을 포함하는 비디오 비트스트림이 있는지를 결정한다. 인코딩된 팩킹된 이미지가 단일 비디오 비트스트림으로 인코딩되어 있어서 복수의 트랙으로서 노출될 수 없거나 콘텐츠 생성자가 인코딩된 팩킹된 이미지를 복수의 트랙으로서 노출하기를 원하지 않는다면, 비디오 비트스트림 또는 비디오 서브비트스트림은 하나의 단일 트랙으로 캡슐화된다(단계 210). 그렇지 않고, 캡슐화될 미디어 콘텐츠가 움직임-제한된 타일을 나타내는 비디오 서브비트스트림들로 합성되어 있는지가 단계 220에서 결정된다. 만일 그렇다면, 수개의 타일 트랙의 적어도 하나의 합성을 나타내기 위해 적어도 하나의 합성 트랙이 제공될 필요가 있을 수 있다. 합성은 전체 팩킹된 이미지 또는 전체 팩킹된 이미지의 서브파트만을 나타낼 수 있다. 타일 트랙과 함께 합성 트랙을 이용하면, 클라이언트 측에서의 스트림의 별도의 렌더링 및 디코딩을 요구하는 것을 피할 수 있다. 클라이언트에게 노출될 수 있는 가능한 조합의 수는 콘텐츠 생성자의 선택에 의존한다. 예를 들어, 콘텐츠 생성자는 현재 사용자의 뷰포트에 따라 상이한 시각적 품질을 갖는 타일들을 결합하기를 원할 수도 있다. 이를 위해, 상이한 시각적 품질들을 갖는 팩킹된 이미지를 수차례 인코딩하고 시각적 품질 측면에서 상이한 타일 조합을 포함하는 전체 팩킹된 이미지를 나타내는 수개의 합성 트랙을 제안할 수 있다. 사용자의 뷰포트에 따라 상이한 품질의 타일들을 결합함으로써, 콘텐츠 생성자는 네트워크 자원 소비를 감소시킬 수 있다.
- [0152] 단계 220에서, 합성 트랙이 제공되어야 한다고 결정되면, 합성 트랙에 대해 목시적 재구성이 이용될 수 있는지의 여부가 결정된다(단계 240).
- [0153] 목시적 재구성이란, 예를 들어 ISO/IEC 14496-15 제4 판에 정의된 타일 베이스 및 타일 트랙으로부터의 비트스트림 재구성을 말한다. 추출기 등의 스트림내 구조를 이용하여 합성 트랙의 샘플에서의 추출기를 타일 트랙의 샘플에서 참조하는 데이터로 대체함으로써 타일 트랙의 샘플로부터 합성 트랙의 샘플을 재구축하는 것 대신에, 목시적 재구성은, 트랙 참조(예를 들어, HEVC 목시적 재구성에서 'sabt' 트랙 참조)의 순서로 합성 트랙 및 타일 트랙의 샘플 연결함으로써 합성 트랙의 샘플을 재구축하는 것을 허용한다.
- [0154] 목시적 재구성의 이용은 이용 시나리오에 의존한다. 수개의 타일 트랙의 합성이 인코딩시 타일 순서와 비교하여 디코딩시 타일의 재배열을 요구할 때, 일부 슬라이스 주소가 재가입되어야 한다. 이러한 경우, 목시적 재구성이 가능하지 않고 추출기를 이용한 명시적 재구성이 선택되어야 한다.
- [0155] 목시적 재구성이 가능하다면, 타일 베이스 트랙이 생성되고(단계 241), 비디오 서브비트스트림은 독립적으로 디코딩가능하지 않은 타일 트랙(예를 들어, HEVC 'hvt1' 트랙)으로서 캡슐화된다.
- [0156] 그렇지 않으면, 추출기 트랙이 생성되고(단계 242), 비디오 서브비트스트림은 독립적으로 디코딩가능한 타일 트랙(예를 들어, HEVC 'hvc1' 또는 'hev1' 트랙)으로서 캡슐화된다.
- [0157] 단계 220으로 돌아가서, 미디어 콘텐츠가 타일 서브비트스트림을 포함하지 않거나 콘텐츠 생성자가 합성 트랙을 생성하고 노출하기를 원하지 않는다면, 공간적으로-관련된 비디오 비트스트림 또는 비디오 서브비트스트림(예를 들어, 타일 또는 복수의 뷰)은 별개의 서브픽처 트랙으로 캡슐화된다(단계 230). 이러한 특별한 경우, 타일 서브비트스트림들이 HEVC 타일이라면, 이들은 HEVC 트랙 'hvc1' 또는 'hev1' 트랙으로 캡슐화된다.
- [0158] 단계 250에서, 공간적 합성을 위한 시그널링이 추가되어 공간적으로-관련된 비디오 비트스트림 또는 비디오 서브비트스트림을 함께 그룹화한다. 공간적 합성 시그널링은, 그룹, 예를 들어 앞서 설명된 바와 같이 MPEG ISOBMMF(ISO/IEC 14496-12 7th edition Amendment 1)에 정의된 바와 같이 동일한 그룹에 속하는 모든 트랙에 대해 동일한 track_group_id를 갖는 유형 '2dcc'의 트랙 그룹을 합성하는 각각의 트랙(서브픽처 트랙, 타일 트랙, 합성 트랙)에서 특정한 TrackGroupBox를 정의함으로써 제공될 수 있다.
- [0159] 이 트랙 그룹 박스 '2dcc'는, 합성 내에서 트랙의 상대적인 2차원 좌표 및 합성에 의해 형성된 이미지의 전체 크기를 제공한다. 합성은 전체 팩킹된 이미지 또는 팩킹된 이미지의 서브파트만을 나타낼 수 있다. 예를 들어, 콘텐츠 생성자는 복수의 합성 트랙을 노출하여 전체 팩킹된 이미지 또는 팩킹된 이미지의 서브파트만의

구축을 허용하기를 원할 수 있다.

[0160] 대안으로서, 합성은 전체 투사된 이미지 또는 투사된 이미지의 서브파트만을 나타낼 수 있다.

[0161] '2dcc' 트랙 그룹으로부터의 파라미터(track_group_id, source_id, total_width, total_height, object_x, object_y, object_width, object_height)는, 이들 트랙들을 나타내는 Adaptation Set들의 공간적 관계를 기술하기 위해 DASH 매니페스트에서 이용될 수 있는 (ISO/IEC 23009-1 3rd edition에 정의된) DASH Spatial-Relationship Description(SRD) 디스크립터의 파라미터와 직접 정합된다 :

[0162] track_group_id는 DASH SRD spatial_set_id 파라미터와 정합된다,

[0163] source_id는 DASH SRD source_id 파라미터와 정합된다,

[0164] object_x, object_y, object_width, object_height는, 각각, DASH SRD 파라미터 object_x, object_y, object_width, object_height 파라미터와 정합되고,

[0165] (track_group_id를 통해) 연관된 트랙 그룹의 total_width 및 total_height는, DASH SRD total_width, total_height 파라미터와 정합될 것이다.

[0166] 대안으로서, 합성 트랙이 있는 경우, 이 합성 트랙에 의해 묵시적으로 공간적 합성 시그널링이 제공될 수 있다. 실제로, 합성 트랙이 타일 베이스 트랙인 경우, 타일 베이스 트랙은 유형 'sabt'의 트랙 참조를 통해 한 세트의 타일 트랙을 참조한다. 이 타일 베이스 트랙과 타일 트랙 세트는 합성 그룹을 형성한다. 유사하게, 합성 트랙이 추출기 트랙이라면, 추출기 트랙은 유형 'scal'의 트랙 참조를 통해 한 세트의 타일 트랙을 참조한다. 이 추출기 트랙과 타일 트랙 세트도 역시 합성 그룹을 형성한다. 양쪽 경우 모두에서, ISO/IEC 14496-15 제4 판에 정의된 바와 같이 유형 'trif'의 샘플 그룹화 또는 디폴트 샘플 그룹화를 정의함으로써 합성 내의 각각의 타일 트랙의 상대적인 2차원 좌표가 제공될 수 있다.

[0167] 또 다른 대안으로서, 새로운 엔티티 그룹을 정의함으로써 공간적 합성 시그널링이 제공될 수 있다. 엔티티 그룹은 항목 또는 트랙의 그룹화이다. 엔티티 그룹은 MetaBox 내의 GroupsListBox에서 표시된다. 트랙을 참조하는 엔티티 그룹은, 파일 레벨 MetaBox의 GroupsListBox에서 또는 영화 레벨 MetaBox의 GroupsListBox에서 명시될 수 있다. GroupsListBox('grp1')는, 정의된 그룹화 유형을 나타내는 연관된 4 캐릭터 코드와 함께, 각각이 EntityToGroupBox라고 불리는 한 세트의 전체 박스를 포함한다. EntityToGroupBox는 다음과 같이 정의된다 :

```
aligned(8) class EntityToGroupBox(grouping_type, version, flags)
extends FullBox(grouping_type, version, flags) {
    unsigned int(32) group_id;
    unsigned int(32) num_entities_in_group;
    for(i=0; i<num_entities_in_group; i++)
        unsigned int(32) entity_id;
    // the remaining data may be specified for a particular
    grouping_type
}
```

[0168] }

[0169] 전형적으로, group_id는 그룹의 id를 제공하고 entity_id 세트는 엔티티 그룹에 속하는 트랙의 track_ID를 제공한다. entity_id 세트에 이어, 특정한 grouping_type에 대한 추가 데이터를 정의하여 EntityToGroupBox의 정의를 확장할 수 있다. 한 실시예에 따르면, 예를 들어 (엔티티 그룹 합성에 대한) 'egco'와 동일한 grouping_type을 갖는 새로운 EntityToGroupBox는, 2차원 공간적으로-관련된 비디오 비트스트림 또는 비디오 서브비트스트림의 합성을 기술하기 위해 정의될 수 있다. entity_id 세트는, 그룹을 합성하는 트랙들(서브픽처, 타일 트랙, 합성 트랙)의 track_ID 세트를 포함한다. 합성에 의해 형성된 이미지의 전체 크기는 이 새로운 grouping_type 'egco'와 연관된 추가 데이터의 일부로서 제공될 수 있다.

[0170] EntityToGroupBox('egco')는 다음과 같이 정의될 것이다 :

```
aligned(8) class EntityToGroupBox('egco', version, flags)
extends FullBox('egco', version, flags) {
    unsigned int(32) group_id;
    unsigned int(32) num_entities_in_group;
    for(i=0; i<num_entities_in_group; i++)
        unsigned int(32) entity_id;
    unsigned int(16) total_width;
    unsigned int(16) total_height;
    unsigned int(32) source_id;
}
```

[0171] }

[0172] 여기서, total_width 및 total_height는 합성의 크기를 제공하고 source_id는 소스에 대한 고유 식별자를 제공하고 소스와 연관된 좌표계(즉, 원점(0, 0) 및 축 방향)를 명시적으로서 정의한다.

[0173] DASH와 비교하여, group_id는 DASH SRD spatial_set_id 파라미터와 정합되고 source_id는 DASH SRD source_id 파라미터와 정합되며, total_width 및 total_height는 각각 DASH SRD total_width 및 total_height 파라미터와 정합된다.

[0174] 유형 'egco'의 엔티티 그룹화에 의해 정의된 합성 내의 각각의 트랙의 상대적 2차원 좌표는 아래에 정의된 바와 같이 유형('egco')의 트랙 그룹을 정의함으로써 제공될 수 있다.

```
aligned(8) class SubPictureRegionBox extends FullBox('sprg',0,0) {
    unsigned int(16) object_x;
    unsigned int(16) object_y;
    unsigned int(16) object_width;
    unsigned int(16) object_height;
}
aligned(8) class SpatialRelationship2DDescriptionBox extends
TrackGroupTypeBox('2dcc')
{
    // track_group_id is inherited from TrackGroupTypeBox;
    SubPictureRegionBox ();
}
```

[0175] }

[0176] 여기서, object_x, object_y, object_width 및 object_height는, 합성에서 각각의 트랙의 상대적인 2차원 좌표를 제공한다.

[0177] 유형 'egco'의 주어진 EntityToGroupBox는 group_id가 track_group_id와 같도록 정의함으로써 대응하는 SpatialRelationship2DDescriptionBox와 연관된다.

[0178] 대안으로서, 유형 'egco'의 엔티티 그룹화에 의해 정의된 합성 내의 각각의 트랙의 상대적인 2차원 좌표는 ISO/IEC 14496-15 제4 판에 정의된 바와 같이 각각의 타일 트랙에서 유형 'trif'의 샘플 그룹화 또는 디폴트 샘플 그룹화를 정의함으로써 제공될 수 있다. 대안으로서, 상대적인 2차원 좌표는, 한 그룹에 속하는 각각의 타일 트랙의 VisualSampleEntry에 위치하는 새로운 일반 전체 박스 2DCoordinateForEntityGroupBox('2dco')로서 정의될 수 있다.

```
aligned(8) class 2DCoordinateForEntityGroupBox extends FullBox('2dco',
version, flags)
{
    unsigned int(32) entity_group_id;
    unsigned int(16) object_x;
    unsigned int(16) object_y;
    unsigned int(16) object_width;
    unsigned int(16) object_height;
}
```

[0179]

여기서

[0180]

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

[0181]

object_x 및 object_y는 합성 내에서 이 트랙의 샘플의 좌상귀의 수평 및 수직 위치를 제공한다.

[0182]

object_width 및 object_height는 합성 내에서 이 트랙 샘플의 폭과 높이를 제공한다.

[0183]

대안으로서, 이 새로운 일반 박스 2DCoordinateForEntityGroupBox('2dco')는 다음과 같이 새로운 샘플 그룹화로서 정의될 수 있다:

[0184]

```
class 2DCoordinateForEntityGroupBox extends
VisualSampleGroupEntry('2dco')
{
    unsigned int(32) entity_group_id;
    unsigned int(16) object_x;
    unsigned int(16) object_y;
    unsigned int(16) object_width;
    unsigned int(16) object_height;
}
```

[0185]

다시, 도 2로 돌아가면, 단계 260에서, 트랙에 대한 영역별 팩킹 정보가 비디오 비트스트림 또는 비디오 서브비트스트림의 캡슐화를 기술하는 메타데이터에 추가된다.

[0186]

영역별 팩킹은 팩킹된 영역 내의 루마 샘플 위치를 대응하는 투사 영역의 루마 샘플 위치로 리맵핑하기 위한 정보를 제공한다. MPEG OMAF에서, 영역별 팩킹은 다음 데이터 구조에 따라 기술될 수 있다 :

[0187]

```
aligned(8) class RegionWisePackingStruct() {
    unsigned int(1) constituent_picture_matching_flag;
    bit(7) reserved = 0;
    unsigned int(8) num_regions;
    unsigned int(32) proj_picture_width;
    unsigned int(32) proj_picture_height;
    unsigned int(16) packed_picture_width;
    unsigned int(16) packed_picture_height;
    for (i = 0; i < num_regions; i++) {
        bit(3) reserved = 0;
        unsigned int(1) guard band flag[i];
        unsigned int(4) packing_type[i];
        if (packing_type[i] == 0) {
            RectRegionPacking(i);
            if (guard_band_flag[i])
                GuardBand(i);
        }
    }
}
```

[0188]

- [0189] 여기서
- [0190] proj_picture_width 및 proj_picture_height는 투사된 픽처의 폭과 높이를 상대적인 투사된 픽처 샘플 유닛으로 각각 명시한다.
- [0191] packed_picture_width 및 packed_picture_height는 팩킹된 픽처의 폭과 높이를 상대적인 팩킹된 픽처 샘플 유닛으로 각각 명시한다.
- [0192] num_regions는 constituent_picture_matching_flag가 0일 때 팩킹된 영역의 수를 명시한다. constituent_picture_matching_flag가 1일 때, 팩킹된 영역의 총 수는 $2 * \text{num_regions}$ 와 같은 RectRegionPacking(i) 및 GuardBand(i) 내의 정보는 투사된 픽처와 팩킹된 픽처의 각각의 스테레오 성분 픽처에 적용된다.
- [0193] RectRegionPacking(i)는 i번째 팩킹된 영역과 i번째 투사된 영역 사이의 영역별 팩킹을 명시하고(즉, 선택사항적인 변환(회전, 미러링)을 이용하여 x, y, 폭, 높이 좌표를 팩킹된 영역으로부터 투사된 영역으로 변환),
- [0194] GuardBand(i)는, i번째 팩킹된 영역에 대한 보호 대역을, 있는 경우, 명시한다.
- [0195] 본 발명의 실시예들에 따르면, 영역별 팩킹 정보가 서브픽처 트랙에서 정의될 때, 이 구조는 완전한 투사된 픽처를 참조함으로써 서브픽처 트랙의 팩킹만을 기술한다. 따라서 packed_picture_width 및 packed_picture_height는 서브픽처 트랙의 폭 및 높이와 같다.
- [0196] 단계 270에서, 트랙 및 트랙의 합성에 대한 콘텐츠 커버리지 정보가 비디오 비트스트림 또는 비디오 서브비트스트림의 캡슐화를 기술하는 메타데이터에 추가된다. 이 단계는 선택사항이다.
- [0197] 트랙 커버리지 정보는, 이 트랙에 의해 표현되는 콘텐츠가 커버하는 구면 상의 에어리어(area)에 관한 정보를 제공한다.
- [0198] 합성 커버리지 정보는, 하나 이상의 트랙 조합과 연관된 구면 표면 상의 에어리어에 관한 정보를 제공한다. 예를 들어, 영화 파일이 공간적 관계와 함께 복수의 비디오 트랙을 포함할 때, 합성 커버리지 정보는 이들 복수의 비디오 트랙의 공간적 합성에 의해 커버되는 구면 표면 상의 에어리어이다. 또 다른 예에서, 미디어 파일은 복수의 비디오 트랙과 이 트랙 세트를 렌더링하는 방법을 나타내는 변환 행렬을 포함하고, 합성 커버리지 정보는 조립된 트랙 세트에 의해 커버되는 에어리어에 대응한다. "합성 커버리지 정보"는 또한, "전역적 커버리지 정보" 또는 "트랙 그룹 합성 정보"로도 표기될 수 있다. 합성 또는 전역적 커버리지 정보는 또한, 이들 복수의 비디오 트랙의 서브세트의 합성으로부터 발생하는 구면 표면 상의 에어리어를 기술할 수 있다.
- [0199] 제1 실시예로서, 트랙 커버리지 정보 및 합성 커버리지 정보는 추가 시그널링없이 하나의 공통 CoverageInformationBox를 이용하여 시그널링될 수 있다. 이러한 경우, CoverageInformationBox의 범위는 박스 계층 구조에서 이 박스의 정의의 위치에 의존한다. 클라이언트는, 선언된 위치를 고려함으로써 커버리지 정보가 트랙 콘텐츠와 관련이 있는지 전체 콘텐츠와 관련이 있는지를 결정할 수 있다. 이 실시예에 따르면, CoverageInformationBox는 다음과 같이 정의된다 :

```

Box Type:      'covi'
Container:     Projected omnidirectional video box ('povd') or
SpatialRelationship2DDescriptionBox ('2dcc')
Mandatory:    No
Quantity:     Zero or one
aligned(8) class CoverageInformationBox extends FullBox('covi', 0, 0)
{
    ContentCoverageStruct()
}
    
```

- [0200]
- [0201] 여기서 ContentCoverageStruct는 다음과 같이 SphereRegionStruct()에 의해 기술된 커버된 영역의 수를 명시한다.

```

aligned(8) SphereRegionStruct(range_included_flag) {
    signed int(32) centre_azimuth;
    signed int(32) centre_elevation;
    signed int(32) centre_tilt;
    if (range_included_flag) {
        unsigned int(32) azimuth_range;
        unsigned int(32) elevation_range;
    }
    unsigned int(1) interpolate;
    bit(7) reserved = 0;
}
aligned(8) class ContentCoverageStruct() {
    unsigned int(8) coverage_shape_type;
    unsigned int(8) num_regions;
    unsigned int(1) view_idc_presence_flag;
    if (view_idc_presence_flag == 0) {
        unsigned int(2) default_view_idc;
        bit(5) reserved = 0;
    } else
        bit(7) reserved = 0;
    for ( i = 0; i < num_regions; i++) {
        if (view_idc_presence_flag == 1) {
            unsigned int(2) view_idc[i];
            bit(6) reserved = 0;
        }
        SphereRegionStruct(1);
    }
}

```

[0202]

[0203]

[0204]

[0205]

[0206]

[0207]

[0208]

[0209]

여기서

coverage_shape_type은 콘텐츠 커버리지를 표현하는 구면 영역의 형상을 명시한다.

num_regions는 구면 영역의 수를 명시한다.

view_idc_presence_flag, default_view_idc 및 view_idc[i]는 i번째 구면 영역이 양안 콘텐츠의 좌측, 우측 또는 양쪽 뷰에 있는지를 나타내는데 이용되는 속성이다.

center_azimuth, center_elevation 및 center_tilt는 전역 좌표 축에 상대적인 커버된 에어리어의 뷰포트 방향을 명시하고, azimuth_range 및 elevation_range는, 존재시에, 커버된 구면 영역의 방위각과 고도 범위들을 각각 명시하고, 보간은 현재 이용되지 않는다.

따라서 CoverageInformationBox는 콘텐츠에 의해 커버된 구면 상의 에어리어에 관한 정보를 제공한다. 콘텐츠의 성질은 이 박스의 컨테이너에 의존한다. SpatialRelationship2DDescriptionBox '2dcc'에 존재할 때, 콘텐츠란, 동일한 서브픽처 합성 트랙 그룹에 속하는 모든 트랙에 의해 표현되는 전체 콘텐츠를 말하며, 이들 트랙으로부터 합성된 합성 픽처는 전체 콘텐츠의 팩킹된 픽처라고 지칭된다. 트랙의 샘플 엔트리에 존재할 때, 콘텐츠란, 이 트랙 자체에 의해 표현되는 콘텐츠를 말하고, 이 트랙 내의 샘플의 픽처는 전체 콘텐츠의 팩킹된 픽처라고 지칭된다. 트랙에 대해 어떠한 CoverageInformation Box도 없다면, 이것은 콘텐츠가 전체 구면을 커버한다는 것을 나타낸다.

Projected omnidirectional video box('povd')는 MPEG OMAF에 의해 정의되고 트랙 내의 VisualSampleEntry에 위치한 중간 박스임에 유의해야 한다.

[0210] 또한, SpatialRelationship2DDescriptionBox 트랙 그룹 박스('2dcc')가 다음과 같이 수정된다 :

```
aligned(8) class SpatialRelationship2DDescriptionBox extends
TrackGroupTypeBox('2dcc')
{
    // track_group_id is inherited from TrackGroupTypeBox;
    SpatialRelationship2DSourceBox(); // mandatory, must be
first
    SubPictureRegionBox (); // optional
    CoverageInformationBox(); // optional
}
```

[0211]

[0212] 제2 실시예로서, 트랙 커버리지 정보 및 합성 커버리지 정보는 로컬 표시와 전역적 표시를 구별하기 위해 플래그 값을 갖는 단일 공통 CoverageInformationBox를 이용하여 시그널링될 수 있다. CoverageInformationBox는 ISOBMFF FullBox이므로, 박스의 flags 파라미터를 통해 트랙과 전역적 커버리지의 구분이 표현될 수 있다.

[0213] 이 제2 실시예에 따르면, CoverageInformation Box는 다음과 같이 정의된다 :

```
Box Type: 'covi'
Container: Projected omnidirectional video box ('povd')
Mandatory: No
Quantity: Zero or more
aligned(8) class CoverageInformationBox extends FullBox('covi', 0, 0)
{
    ContentCoverageStruct()
}
```

[0214]

[0215] 박스의 구조는, 로컬 및 합성 커버리지 정보가 동일한 트랙에서 정의되어야 하는 경우 박스의 복수의 인스턴스가 정의될 수 있다는 점을 제외하고는 이전 실시예와 거의 동일하다.

[0216] 그 다음, CoverageInformationBox는 콘텐츠에 의해 커버된 구면의 에어리어에 관한 정보를 제공하는 것으로서 정의된다. 콘텐츠의 성질은 flags 파라미터에 의해 주어진다. Coverage Information 플래그에 대한 디폴트 값은 0이며, 이것은 이 박스가 전체 콘텐츠의 커버리지를 기술함을 의미한다. 이 트랙이 2차원 공간적 관계 트랙 그룹에 속한다면, 전체 콘텐츠란 동일한 2차원 공간적 관계 트랙 그룹에 속하는 모든 트랙에 의해 표현되는 콘텐츠를 말하고, 이들 트랙으로부터 합성된 합성 픽처는 전체 콘텐츠의 팩킹된 또는 투사된 픽처라고 지칭된다. 그렇지 않으면, 전체 콘텐츠란 이 트랙 자체에 의해 표현되는 콘텐츠를 말하고, 이 트랙 내의 샘플의 픽처는 전체 콘텐츠의 팩킹된 또는 투사된 픽처고 지칭된다.

[0217] Coverage Information 플래그에 대한 값이 1일 때, 이 박스는 이 트랙에 의해 표현되는 콘텐츠의 팩킹되거나 투사된 픽처에 의해 커버되는 구면 에어리어를 기술한다.

[0218] 이 박스의 부재는, 콘텐츠가 전체 구면을 커버하고 있다는 것을 나타낸다.

[0219] 또한, 새로운 플래그 값이 다음과 같이 정의된다 :

[0220] Coverage_local : 커버리지 정보가 박스를 포함하는 트랙에 대해 로컬임을 나타낸다. 플래그 값은 0x000001이다. 디폴트로, 이 값은 설정되지 않는다.

[0221] 도 2로 돌아가서, 단계 280에서, 가상 현실 미디어 콘텐츠가 실제로 양안 가상 현실 미디어 콘텐츠인지, 즉, 좌측 및 우측 뷰를 포함하는지가 체크된다.

[0222] 콘텐츠가 단안형(monoscopic)이라면, 프로세스는 단계 290으로 바로 이동한다.

[0223] 콘텐츠가 양안형이라면, 단계 285에서 양안 시그널링이 캡슐화에 추가된다.

[0224] 양안 콘텐츠의 경우, 전통적으로, 좌측 및 우측 뷰 시퀀스 양쪽 모두가 양안 카메라로부터 취득되어 합성 유형에 따라 하나의 비디오 시퀀스 또는 2개의 비디오 시퀀스로 합성된다.

[0225] 양안 콘텐츠의 2개의 상이한 뷰를 나타내는 2개의 프레임을 하나의 단일 프레임으로 결합하는 프로세스는 프레

임 팩킹이라고 불린다(도 1의 단계 125 참조).

[0226] 프레임 팩킹은 스테레오 쌍을 형성하는 2개의 뷰를 단일 프레임으로 팩킹하는 것으로 구성된다. 널리 공지되어 이용되는 수개의 프레임 팩킹 방식이 있다 : 병렬, 상하, 프레임 순차, 수직 라인 인터리브된 유형 예를 들어, MPEG 애플리케이션 포맷 ISO/IEC 23000-11 제1 판("양안 비디오 애플리케이션 포맷") 또는 ISO/IEC 23001-8 제2 판("Coding-independent code-points(CICP)")은 이들 방식들 중 일부를 정의한다. 프레임 팩킹은 또한, 예를 들어 ISO/IEC 23001-8 제2 판("CICP")에 정의된 값 6을 갖는 VideoFramePackingType처럼 각각의 뷰를 별개의 프레임에 유지하는 것으로 구성될 수 있다.

[0227] 예를 들어, 여전히 이 명세에 따르면, 값 3은, 각각의 디코딩된 프레임이 2개의 성분 뷰의 대응하는 프레임들의 병렬 팩킹 구성을 포함하고 있다는 것을 나타내며, 값 4는 각각의 디코딩된 프레임이 2개의 성분 뷰의 대응하는 프레임들의 상하 팩킹 구성을 포함하고 있다는 것을 나타낸다.

[0228] 트랙에 양안 미디어 데이터를 포함하고 있는지를 시그널링하기 위해, 트랙 내의 VisualSampleEntry에서 StereoVideoBox가 정의된다.

[0229] 도 2의 단계 250으로 돌아가서, SpatialRelationship2DDescriptionBox는 DASH(Dynamic Adaptive Streaming over HTTP) 프로토콜(ISO/IEC 23009-1 제3 판)에 정의된 Spatial Relationship Descriptor 'SRD'의 정의와 정합하도록 정의되어 아래 표에 제공된 비디오 트랙들 사이의 공간적 관계를 표현한다 :

ISOBMFF 파라미터	DASH SRD 파라미터
trgr::'2dcc'::track_group_id	spatial_set_id
trgr::'2dcc'::'sprg'::object_x	object_x
trgr::'2dcc'::'sprg'::object_y	object_y
trgr::'2dcc'::'sprg'::object_width	object_width
trgr::'2dcc'::'sprg'::object_height	object_height
trgr::'2dcc'::'2dsr'::total_width	total_width
trgr::'2dcc'::'2dsr'::total_height	total_height
trgr::'2dcc'::'2dsr'::source_id	source_id

[0230] '2dcc' track_grouping_type을 갖는 TrackGroupTypeBox는, 트랙이 비디오의 공간적 파트에 대응하는 트랙들의 그룹에 속함을 나타낸다. track_group_type '2dcc'의 TrackGroupTypeBox 내에서 동일한 값의 source_id를 갖는 트랙은 동일한 소스(즉, 동일한 원점(0, 0) 및 동일한 축 방향)로부터 생성된 것으로 맵핑된다. 더욱 정확하게는, 동일한 source_id를 갖는 2개의 트랙 그룹으로부터의 (크기 total_width 및 total_height를 갖는) 전체 합성 픽처는 인지적으로 또는 시각적으로 동일하다(예를 들어, 2개의 상이한 해상도 또는 2개의 상이한 품질의 동일한 시각적 콘텐츠를 나타내는 2개의 합성 픽처).

[0232] '2dcc' track_grouping_type 및 동일한 track_group_id를 갖는 TrackGroupTypeBox에 속하는 모든 서브픽처 트랙은 동일한 source_id를 가져야 한다.

[0233] '2dcc' track_grouping_type 및 상이한 track_group_id를 갖는 TrackGroupTypeBox에 속하는 트랙들은 호환되며 그들이 동일한 source_id를 갖는다면 함께 결합될 수 있다. 그렇지 않으면, 서브픽처 트랙이 동일한 소스의 서브파트를 나타내지 않거나 및/또는 '2dcc' track_grouping_type 및 상이한 source_id를 갖는 또 다른 TrackGroupTypeBox로부터의 서브픽처 트랙과 결합하기 위한 것이 아니다. 예를 들어, 이 소스를 나타내는 2차원 투사된 픽처가 시각적으로 동일하지 않을 때(예를 들어, 이들이 상이한 투사 포맷 또는 상이한 뷰포트 방향을 가질 때) 2개의 서브픽처 트랙은 동일한 소스의 서브파트를 나타내지 않는다.

[0234] 대안으로서, 이 나중의 규칙은, 상이한 source_id를 갖는 '2dcc' 트랙 그룹으로부터의 서브픽처 트랙들을 그룹화하는 대체 그룹이 있는 경우에도 적용된다. 이것은, 이들 서브픽처 트랙은 대안이긴 하지만(예를 들어, 이들이 AVC 및 HEVC 등의 상이한 코딩 포맷을 갖지만) 상이한 코딩 포맷을 갖는 서브픽처 트랙과 결합되기 위한 것이 아님을 의미한다.

[0235] 도 4는 상기 규칙의 한 예를 나타낸다. 트랙 #1 내지 #4는 track_group_id가 10이고 source_id가 1인 유형 '2dcc'의 트랙 그룹 41에 속한다. 트랙 #5 내지 #8은 track_group_id가 20이지만 동일한 source_id 400이 1인

유형 '2dcc'의 상이한 트랙 그룹 42에 속한다. track_group_id가 30이고 상이한 source_id 401이 2인 유형 '2dcc'의 제3 트랙 그룹 43도 있다. 또한, 수개의 대체 그룹 44 내지 47이 있다. 동일한 대체 그룹에 속하는 (즉, 트랙 헤더 박스 'tkhd'에서 동일한 alternate_group 식별자를 갖는) 모든 트랙은 대체 데이터를 포함하는 트랙 그룹 또는 모음을 명시한다. 대체 데이터는, 대체 비트레이트, 코덱, 언어, 패킷 크기 등에 대응할 수 있다. 대체 그룹 내의 오직 한개 트랙만이 한 번에 재생되거나 스트리밍되어야 한다. 이 예에서, 트랙 #1, #5 및 #9는 식별자가 100인 동일한 대체 그룹 44에 속한다. 예를 들어, 트랙 #1과 트랙 #5는 상이한 품질을 가진 대체 트랙이고, 트랙 #9는 코덱의 측면에서 트랙 #1과 트랙 #5의 대체 트랙이다. 트랙 #2, #6 및 #10은 식별자가 200인 동일한 대체 그룹 45에 속한다, 예를 들어, 트랙 #2와 트랙 #6은 상이한 해상도를 갖는 대체 트랙이고, 트랙 #10은 프레임 레이트의 측면에서 트랙 #2 내지 트랙 #6에 대한 대체 트랙이고, ... 기타 등등이다.

[0236] 트랙 그룹 41 및 42는 동일한 source_id 400을 갖고, 트랙 그룹 43은 상이한 source_id 401을 가지며, 이것은 트랙 그룹 41 및 42에 속하는 서브픽처 트랙들이 함께 결합될 수 있다는 것을 의미한다(다른 제약들에 관하여, 즉, 대체 그룹 당 거의 하나의 서브픽처 트랙). 반대로, 트랙 그룹 43으로부터의 서브픽처 트랙들은, 이들이 동일한 source_id를 갖지 않기 때문에 동일한 대체 그룹에 속할 수 있음에도 불구하고 트랙 그룹 41 및 42로부터의 임의의 서브픽처 트랙과 결합되기 위한 것이 아니다. 그 다음, source_id 파라미터는 동일한 공간적 합성의 일부일 수 있는 서브픽처 트랙의 플레이어에게 표시를 제공한다. 주어진 공간적 위치에 대해, 하나의 서브픽처 트랙은 동일한 주어진 공간적 위치에 있는 또 다른 서브픽처 트랙과 시각적으로 균등한 것으로 간주될 수 있다. 이것은 미디어 콘텐츠가 복수의 트랙 내에 제공될 때 (서브픽처) 트랙 선택에 유용하다. 또한, 이것은 선택된 서브픽처 트랙에 따라, 동일한 공간적 합성을 디스플레이하기 위해 (품질/비트레이트 또는 해상도에서) 동적 적합화를 허용한다. 몇 가지 이용 예가 도 5와 도 6에 따라 설명된다.

[0237] 도 5는 본 발명의 실시예들에 따른 SpatialRelationship2DdescriptionBox 및 source_id의 이용 예를 나타낸다. 동일한 비디오 소스(예를 들어, 동일한 투사 비디오 소스)가 이용되어 품질의 측면에서 2개의 대체 버전(@품질1 및 @품질2)을 생성한다. 각각의 대체 버전은 (투사 영역 또는 팩킹 영역을 포함하는) 8개의 서브픽처 트랙으로 분할된다. 서브픽처 트랙들의 제1 세트는 저품질로 이용가능하다. 서브픽처 트랙들의 제2 세트는 고품질로 이용가능하다. 품질 레벨 당 하나씩 2개의 트랙 그룹이 정의된다. 대응하는 서브픽처 트랙은 ('trak' 박스 계층 구조에서) 도 5의 우측 부분과 같이 기술될 수 있다. 양쪽 트랙 그룹 모두는 동일한 source_id, total_width 및 total_height를 갖는다. 서브픽처 트랙 좌표(object_x, object_y, object_width, object_height)는 각각의 트랙 그룹 합성 내에서 서브픽처 트랙의 공간적 관계 또는 위치를 기술한다. 양쪽 트랙 그룹 모두가 동일한 source_id를 갖기 때문에, 이것은, 이들이 (track_group_id가 10인) 제1 트랙 그룹으로부터의 동일한 소스 및 서브픽처 트랙을 동일한 트랙 그룹으로부터의 서브픽처 트랙과 결합할 수 있지만, 합성에서 그들 각각의 위치와 관련하여 (track_group_id가 20인) 제2 트랙 그룹의 서브픽처 트랙과도 결합될 수 있음을 나타낸다는 것을 의미한다.

[0238] 도 6은 본 발명의 실시예들에 따른 SpatialRelationship2DdescriptionBox 및 source_id의 제2 이용 예를 나타낸다. 동일한 비디오 소스(예를 들어, 동일한 투사된 비디오 소스)가 이용되어 해상도 측면에서 2개의 대체 버전(@해상도1 및 @해상도2)을 생성한다. 2개 세트의 서브픽처 트랙이 있다 : 하나는 고해상도용이고 다른 하나는 저해상도용이다.

[0239] 대응하는 서브픽처 트랙은 ('trak' 박스 계층구조에서) 도 6의 우측 부분과 같이 기술될 수 있다. 양쪽 트랙 그룹 모두는 동일한 source_id를 갖지만 각각의 세트의 서브픽처 트랙의 해상도에 대응하는 상이한 total_width 및 total_height를 갖는다. 서브픽처 트랙 좌표(object_x, object_y, object_width, object_height)는 각각의 트랙 그룹 합성 내에서 서브픽처 트랙의 공간적 관계 또는 위치를 기술한다. 다시 한번, 양쪽 트랙 그룹 모두가 동일한 source_id를 갖기 때문에, 이것은, 이들이 (track_group_id가 10인) 제1 트랙 그룹으로부터의 동일한 소스 및 서브픽처 트랙을 동일한 트랙 그룹으로부터의 서브픽처 트랙과 결합할 수 있지만, 그들 각각의 합성에서 그들 각각의 위치와 관련하여 (track_group_id가 20인) 제2 트랙 그룹의 서브픽처 트랙과도 결합될 수 있음을 나타낸다는 것을 의미한다. 이 경우, 상이한 트랙 그룹들로부터의 서브픽처 트랙들이 만일 함께 결합된다면 이들 트랙에 스케일링이 적용되어야 한다. 스케일링 계수는 각각의 트랙 그룹으로부터의 total_height와 total_width 사이의 비율(예를 들어, H1/H2 및 W1/W2)로부터 추론될 수 있다.

[0240] 이 예에 따르면, track_group_id가 10인 트랙 그룹에 의해 표현되는 합성 픽처는, 각각의 대체 그룹으로부터 하나의 서브픽처를 선택함으로써 합성될 수 있다.

- [0241] 2차원(2D) 비디오 콘텐츠와는 달리, OMAF 미디어 콘텐츠는, 구면의 중심으로부터 구면의 내부 표면을 향해 바깥 쪽으로 바라보는 사용자의 관점을 나타내는 전방향 미디어 콘텐츠를 나타낸다. 그 다음, 이 360° 미디어 콘텐츠는 비디오 투사 포맷을 적용함으로써 2차원 평면에 투사된다. 그 다음, 선택사항으로서, 영역별 팩킹이 적용되어 영역을 투사된 픽처로부터 팩킹된 영역으로 재조직화한다. 360° 미디어 콘텐츠도 역시 어안 렌즈(광각 카메라 렌즈)로 캡처된 수개의 원형 이미지로 표현될 수 있다.
- [0242] 따라서, OMAF의 맥락에서, 2D 픽처는 투사된 픽처 또는 팩킹된 픽처일 수 있고 서브픽처 트랙은 상이한 종류의 콘텐츠들을 포함할 수 있다 :
- [0243] - 투사된 픽처의 서브파트(팩킹 없음),
- [0244] - 프레임-팩킹된 픽처의 서브파트, 예를 들어, 콘텐츠가 양안형일 때,
- [0245] - 투사되고 팩킹된 픽처의 서브파트, 또는
- [0246] - 어안 코딩된 픽처의 서브파트.
- [0247] 본 발명의 실시예들에 따르면, SpatialRelationship2DdescriptionBox의 정의가 개선되어 OMAF 미디어 콘텐츠를 포함하는 서브픽처 트랙의 크기 및 위치 좌표가, 투사된 픽처, 팩킹된 픽처, 또는 또 다른 픽처에 대해 상대적 인지를 나타낸다.
- [0248] 한 실시예에서, SpatialRelationship2DdescriptionBox는, OMAF 미디어 콘텐츠를 포함하는 서브픽처 트랙의 크기 및 위치 좌표가 항상 팩킹된 픽처에 상대적이라도 정의된다. 어떠한 팩킹도 없는 경우, 팩킹된 픽처는 투사된 픽처와 동일하다.
- [0249] 또 다른 실시예에서, SpatialRelationship2DdescriptionBox는, OMAF 미디어 콘텐츠를 포함하는 서브픽처 트랙의 크기 및 위치 좌표가 투사된 픽처 또는 팩킹된 픽처 또는 캡처 단계(110)와 인코딩 단계(140) 사이의 처리 단계에서의 임의의 중간 픽처에 상대적이라도 정의된다. 특히, OMAF(omnidirectional media)를 위한 애플리케이션 포맷의 경우, 2D 공간적 관계로 표현된 위치와 크기가 투사된 픽처를 참조하는지 또는 팩킹된 픽처를 참조하는지가 명확하지 않다.
- [0250] 한 실시예에서, SpatialRelationship2DdescriptionBox는 항상 팩킹된 픽처에 대해 상대적이다. 어떠한 팩킹도 없는 경우, 팩킹된 픽처는 투사된 픽처와 동일하다.
- [0251] 또 다른 실시예에서, 바람직한 접근법은 SpatialRelationship2DdescriptionBox가 항상 투사된 픽처에 대해 상대적임을 정의하는 것이다.
- [0252] 장면의 와이드 뷰에 대응하는 인코딩된 미디어 데이터를 캡슐화하기 위한 방법은 일부 실시예들에서 다음과 같은 단계를 포함할 수 있다 :
- [0253] 장면의 와이드 뷰로부터 투사된 픽처를 획득하는 단계;
- [0254] 획득된 투사된 픽처를 적어도 하나의 팩킹된 픽처로 팩킹하는 단계;
- [0255] 적어도 하나의 팩킹된 픽처를 적어도 하나의 서브픽처로 분할하는 단계;
- [0256] 적어도 하나의 서브픽처를 복수의 트랙으로 인코딩하는 단계;
- [0257] 인코딩된 트랙들과 연관된 설명 메타데이터를 생성하는 단계,
- [0258] 여기서, 설명 메타데이터는 트랙으로 인코딩된 적어도 하나의 서브픽처와 적어도 하나의 투사된 픽처 사이의 공간적 관계를 나타내는 각각의 트랙과 연관된 정보 항목을 포함한다.
- [0259] 따라서, 기준 픽처의 어떠한 특별한 시그널링도 필요하지 않다. 기준 픽처는, 서브픽처가 팩킹된 픽처를 분할함으로써 획득되더라도 투사된 픽처이도록 정의된다.
- [0260] 장면의 와이드 뷰에 대응하는 인코딩된 미디어 데이터를 캡슐화하기 위한 방법은 일부 실시예들에서 다음과 같은 단계를 포함할 수 있다 :
- [0261] 장면의 와이드 뷰로부터 투사된 픽처를 획득하는 단계;
- [0262] 투사된 픽처를 적어도 하나의 서브픽처로 분할하는 단계;

- [0263] 적어도 하나의 서브픽처를 복수의 트랙으로 인코딩하는 단계;
- [0264] 인코딩된 트랙들과 연관된 설명 메타데이터 - 설명 메타데이터는 트랙으로 인코딩된 적어도 하나의 서브픽처와 기준 픽처 사이의 공간적 관계를 나타내는 각각의 트랙과 연관된 제1 정보 항목을 포함함- 를 생성하는 단계;
- [0265] 설명 메타데이터는 기준 픽처를 나타내는 제2 정보 항목을 더 포함한다.
- [0266] 따라서, 메타데이터에서 기준 픽처를 명시함으로써, 분할 동작과는 독립적으로, 투사된 픽처, 팩킹된 픽처 또는 기타 임의의 기준 픽처에 관련된 서브픽처 데이터를 생성하는 것이 가능하다.
- [0267] 아래의 표는, 예를 들어, Equirectangular(ERP) 또는 큐브맵 투사를 이용하여 투사된 콘텐츠, 팩킹되거나 어안 콘텐츠를 포함하는 서브픽처 트랙에 대해 OMAF 맥락에서 투사된 픽처에 대해 상대적인 SpatialRelationship2DdescriptionBox 트랙 그룹 크기 및 좌표 속성의 실제 매핑을 제안한다. 아래의 표에서 "rwpk"는, 영역별 팩킹 구조, 즉, 팩킹된 영역과 각각의 투사된 영역 사이의 매핑을 명시하고, 있다면, 보호 대역의 위치 및 크기를 명시하는 구조의 바로가기(shortcut)이다. 또한, 'fovi'는, OMAF 플레이어에서 어안 이미지들의 스티칭 및 렌더링을 가능케하기 위한 파라미터를 기술하는 구조인 FisheyeVideoEssentialInfoStruct의 바로가기이다.

소스의 유형	total_width/total_height	object_width/object_height	object_x/object_y
투사된 픽처 (팩킹 없음)	투사된 픽처의 루마 픽처 크기와 같아야 함	서브픽처 트랙의 샘플들에 의해 표현된 투사된 영역의 폭 및 높이와 같아야 함. (이러한 경우, 서브픽처 트랙의 트랙 헤더 박스에서 선언된 폭 및 높이와 같아야 함)	투사된 픽처 내의 서브픽처 트랙의 샘플들에 의해 표현된 투사된 영역의 좌상귀의 x, y 좌표와 같아야 함
투사되고 팩킹된 픽처	투사된 픽처의 루마 픽처 크기와 같아야 함 Total_width = rwpk@proj_picture_width Total_height = rwpk@proj_picture_height	서브픽처 트랙의 샘플들을 언팩킹함으로써 표현된 투사된 영역의 폭 및 높이와 같아야 함. (이러한 경우, 서브픽처 트랙의 언팩킹으로부터 발생하는 투사된 영역은 겹들을 포함할 수 있음)	투사된 픽처 내의 서브픽처 트랙의 샘플들을 언팩킹함으로써 표현된 투사된 영역의 좌상귀의 x, y 좌표와 같아야 함

[0268]

<p>어안 투사된 픽처</p>	<p>투사된 이미지의 루마 픽처 크기와 같아야 함, 즉 이미지는 모든 순환 이미지들을 포함함</p> <p>Total_width = 마지막 원형 이미지의 fovi@rect_region_left + fovi@rect_region_width</p> <p>Total_height = 마지막 원형 이미지의 fovi@rect_region_top + fovi@rect_region_height</p>	<p>서브픽처 트랙으로부터의 하나 이상의 원형 이미지를 포함하는 직사각형 투사된 영역의 폭 및 높이가 같아야 함.</p> <p>예를 들어, 서브픽처 트랙이 단 하나의 원형 이미지를 포함하는 경우: Object_width = fovi@rect_region_width Object_height = fovi@rect_region_height</p> <p>서브픽처 트랙이 하나보다 많은 원형 이미지를 포함하는 경우, object_width (대응적으로, object_height)는 포함된 원형 이미지들의 fovi@rect_region_width 의 합 (대응적으로, fovi@rect_region_height 의 합)과 같다</p>	<p>서브픽처 트랙으로부터의 하나 이상의 원형 이미지를 포함하는 직사각형 투사된 영역의 좌상귀의 좌표들과 같아야 함.</p> <p>예를 들어, 단 하나의 원형 이미지의 경우 : Object_x = fovi@rect_region_left Object_y = fovi@rect_region_top</p> <p>서브픽처 트랙이 하나보다 많은 원형 이미지를 포함하는 경우, object_x (대응적으로, object_y)는 포함된 원형 이미지들의 목록 내의 최상단 좌측 원형 이미지의 fovi@rect_region_left (대응적으로, fovi@rect_region_top)와 같다.</p>
------------------	--	--	---

[0269]

[0270]

투사된 픽처에 상대적으로 SpatialRelationship2DdescriptionBox 속성을 정의하는 것은, 팩킹된 픽처에 상대적으로 이들을 정의하는 것에 비해 애플리케이션에 이점을 제공한다. 실제로, 뷰포트-의존 스트리밍의 경우, 애플리케이션은 현재 사용자의 뷰포트에 대응하는(즉, 사용자의 시야 및 방향에 대응하는) 서브픽처 트랙만을 다운로드하기를 원할 수도 있다. SpatialRelationship2DdescriptionBox 속성이 투사된 픽처에 상대적으로 정의된다면, 애플리케이션은 SpatialRelationship2DdescriptionBox 트랙 그룹으로부터의 이 정보를 직접 이용하여 투사된 픽처 내에서 이동하는 동안 적절한 서브픽처 트랙을 선택할 수 있다. 그렇지 않으면, 애플리케이션은, 트랙 그룹 정보 외에도, VisualSampleEntry에 위치한 영역별 팩킹 정보를 파싱하여, 적절한 서브픽처 트랙을 선택하기 전에 서브픽처 팩킹된 콘텐츠를 투사된 픽처로 변환할 필요가 있다.

[0271]

선택사항으로서, 공간적 관계를 기술하는 트랙 그룹(예를 들어, '2dcc' 트랙 그룹)은, 주어진 서브픽처 트랙에 대해, 360° 구면으로의 그 맵핑을 제공하는 추가 디스크립터를 포함할 수 있다. 이 추가 디스크립터는, 미디어 플레이어에 대한 계산없이, 2D 비디오 서브픽처 트랙과 3D 뷰포트 사이의 맵핑을 제공하여, 주어진 사용자의 시점 방향에 대응하는 관련 트랙 또는 트랙 세트의 플레이어에 의한 선택이 더 용이하게 되도록 할 수 있다. 그러면, 공간적 관계를 기술하는 트랙 그룹은 다음과 같이 다시 쓸 수 있다 :

```

aligned(8) class SpatialRelationship2DDescriptionBox extends
TrackGroupTypeBox('2dcc') {
    // track_group_id is inherited from TrackGroupTypeBox;
    SpatialRelationship2DSourceBox(); // mandatory, must be first
    SubPictureRegionBox (); // optional
    SphericalRegionBox (); // optional
}
    
```

[0272]

[0273]

여기서 SpatialRelationship2DSourceBox 및 SubPictureRegionBox는, 각각, 트랙 그룹과 그 위치 및 크기에 관련된 서브픽처 트랙의 2D 좌표계를 기술한다.

[0274]

여기서 SphericalRegionBox는 다음과 같이 정의된 새로운 박스이다(4 캐릭터 코드는 예일 뿐이며, 구면 영역의 표시를 위해 예약된 경우, 임의의 4 캐릭터 코드가 이용될 수 있다) :

```
aligned(8) class SphericalRegionBox extends FullBox('sspr', 0, 0) {
    SphereRegionStruct(1);
}
}
```

[0275]

[0276]

여기서 SphereRegionStruct는 구면 영역을 트리플릿(centre_azimuth, center_elevation, center_pitch)으로서 명시하거나, 때때로 방위각(수직) 및 고도(수평) 치수 범위와 함께 (요(yaw), 피치(pitch), 롤(roll))로서 명시한다.

[0277]

도 7은 본 발명의 실시예들에 따른 서브픽처 캡슐화를 나타낸다. 이것은, 선택사항적인 단계 260 및 280 및 285를 갖는 도 1의 단계 250에 대응한다. 단계 701에서, 사용자는 캡슐화 모듈(예를 들어, 도 1의 단계 150을 담당하는 ISOBMFF 작성기 또는 mp4 팩키저 또는 작성기)을 구성한다. 이것은 캡슐화 소프트웨어를 제어하는 그래픽 사용자 인터페이스를 통해 이루어질 수 있다. 이것은, 예를 들어, 캡슐화할 소스 또는 서브픽처 트랙으로의 분해 등의 캡슐화를 위한 파라미터, 또는 하나의 단일 미디어 파일 또는 많은 세그먼트 파일의 생성에 관한 정보를 명시하는 것으로 구성된다. 대안으로서, 이것은, 장면을 캡처하는 녹화 디바이스(카메라, 네트워크 카메라, 스마트폰...)의 설정으로서 미리 등록될 수 있다. 그 다음, 캡슐화 모듈은 단계 702에서 기준 픽처를 캡처된 이미지로서 초기화한다. 이것은, 캡슐화 모듈을 운영하는 디바이스의 RAM에 캡처된 이미지의 크기를 저장하는 것으로 구성된다. 그 다음, 단계 703에서, 캡슐화 모듈은, 캡슐화 구성이 투사 단계를 포함하는지를 체크한다. 거짓이면, 다음 단계는 706이다. 예를 들어, 캡처된 콘텐츠가 360° 콘텐츠일 때, 이것은, 투사된 픽처라고 불리는 2D 이미지에 투사될 수 있다. 투사가 이용 중이면(테스트 703이 참) 캡슐화 모듈은 이용중인 투사의 설명을 미디어 파일(또는 미디어 세그먼트)의 설명 메타데이터에 삽입한다(단계 704). 이것은, 예를 들어 OMAF 명세에 따라 투사된 전방향 비디오 박스 'povd'일 수 있다. 그 다음(단계 705), 기준 픽처는 투사된 픽처로 설정된다. 이것은, 예를 들어, 이 투사된 픽처의 크기가 메모리에 저장된다는 것을 의미한다. 단계 706은, 캡처된 소스가 양안형인지의 여부 및 뷰가 단일 프레임으로 팩킹되는지를 체크하는 것으로 구성된다. 테스트 706이 참이면, 캡슐화 모듈은 스테레오 콘텐츠에 대한 디스크립터를 미디어 파일에 삽입한다(단계 707). OMAF 또는 ISOBMFF의 경우 이것은 StereoVideoBox이다. 테스트 706이 거짓이면, 다음 단계는 709이다. 단계 707에 이어, 프레임-팩킹된 픽처가 기준 픽처로서 메모리에 저장된다. 테스트(709)는, 캡슐화 구성이, 투사되고 선택 사항으로서 프레임 팩킹된 픽처가 팩킹된 영역으로 더 재배열될 필요가 있음을 나타내는지를 체크하는 것으로 구성된다. 테스트 709가 참이면, 캡슐화 모듈은 영역으로의 이 팩킹의 설명을 삽입한다(단계 710)(도 1의 선택 사항적인 단계 260과 동일). OMAF의 경우, 이것은 'rwpk' 박스 유형에 의해 식별되는 RegionWisePackingBox일 수 있다. 그 다음 711에서, 기준 픽처는 팩킹된 픽처로 설정된다. 테스트 709가 거짓이면, 다음 단계는 712이다. 단계 712의 테스트는 캡슐화 구성을 체크하는 것으로 구성된다 : 서브픽처 트랙에 대한 목시적 시그널링 또는 명시적 시그널링이 사용자 또는 애플리케이션에 의해 선택되었거나 설정되었는지 여부. 목시적 시그널링이 오프이면, 단계 713에서, 캡슐화 모듈은 서브픽처 트랙 생성에 대해 어느 기준 픽처(즉, 각각이 서브픽처 트랙들로 캡슐화되는 공간적 파트들로 분할된 픽처)가 이용되는지를 제공하는 설명 메타데이터를 삽입한다. 목시적 시그널링이 온이면, 다음 단계는 714이다. 단계 714에서, 캡슐화 모듈은 분할된 픽처의 상이한 공간적 파트들 사이의 공간적 관계를 기술하는 트랙 그룹을 삽입한다. 특히, 서브픽처 트랙들의 결과적인 합성의 크기는 (702, 705, 708 또는 711에서) 메모리에 저장된 기준 픽처의 크기로 설정된다. 이것은 예를 들어 SpatialRelationship2DSourceBox에서의 total_width 및 total_height 파라미터일 수 있다. 마지막으로, 단계 715에서, 캡슐화 모듈은 기준 픽처에서 위치 및 크기의 관점에서 각각의 서브픽처 트랙을 기술한다. 이것은, 예를 들어 OMAF 또는 ISOBMFF에서 구성되어 분할로부터 발생하는 값을 SubPictureRegionBox의 파라미터들 -이들 파라미터가 정적일 경우- 에 넣거나, 공간적 관계 기술을 위한 샘플 그룹 설명 박스(예를 들어, SpatialRelationship2DGroupEntry 박스)에 넣는다.

[0278]

단계 713의 명시적 시그널링은, 도 8에 의해 예시된 파싱 프로세스의 설명과 함께 설명된 바와 같이 다양한 방식으로 수행될 수 있다.

[0279]

복수의 인코딩된 트랙 및 연관된 설명 메타데이터를 포함하는 미디어 파일로부터 적어도 하나의 이미지를 생성하기 위한 방법은 일부 실시예들에서 다음을 포함할 수 있다 :

[0280]

복수의 인코딩된 트랙이 장면의 와이드 뷰의 투사된 픽처를 팩킹함으로써 획득된 팩킹된 픽처의 분할로부터 발생하는 적어도 하나의 서브픽처를 인코딩하는 트랙들의 그룹을 포함한다고 결정하는 단계;

[0281]

트랙들의 그룹과 연관된 설명 메타데이터를 파싱하는 단계;

- [0282] 여기서, 트랙들의 그룹과 연관된 설명 메타데이터를 파싱하는 단계는 다음과 같은 단계를 포함한다 :
- [0283] 트랙으로 인코딩된 적어도 하나의 서브픽처와 적어도 하나의 투사된 픽처 사이의 공간적 관계를 나타내는 각각의 트랙과 연관된 정보 항목을 해석하는 단계를 포함한다.
- [0284] 복수의 인코딩된 트랙 및 연관된 설명 메타데이터를 포함하는 미디어 파일로부터 적어도 하나의 이미지를 생성하기 위한 방법은 일부 실시예들에서 다음을 포함할 수 있다 :
- [0285] 복수의 인코딩된 트랙이 장면의 와이드 뷰의 투사된 픽처의 분할로부터 발생하는 적어도 하나의 서브픽처를 인코딩하는 트랙들의 그룹을 포함한다고 결정하는 단계;
- [0286] 트랙들의 그룹과 연관된 설명 메타데이터를 파싱하는 단계;
- [0287] 여기서, 트랙들의 그룹과 연관된 설명 메타데이터를 파싱하는 단계는 다음과 같은 단계를 포함한다 :
- [0288] 트랙으로 인코딩된 적어도 하나의 서브픽처와 적어도 하나의 기준 픽처 사이의 공간적 관계를 나타내는 각각의 트랙과 연관된 제1 정보 항목을 해석하는 단계; 및
- [0289] 기준 픽처를 나타내는 제2 정보 항목을 해석하는 단계.
- [0290] 미디어 플레이어는, ISOBMFF 파서를 이용하여, 801에서 OMAF 파일을 수신한다. 이것은, 미디어 파일, 특히 비디오 트랙 내에 존재하는 상이한 트랙들을 식별한다. 이들 비디오 트랙의 경우, 파서는 이들이 고전적인 2D 비디오인지 또는 2D 픽처에 투사된 전방향 미디어용 비디오 트랙인지를 체크한다. 이것은, 단계 802에서 'ftyp' 박스 내의 주요 브랜드 또는 호환가능한 브랜드들의 목록을 조사함으로써 결정된다. 예를 들어, 'ovdp'로 설정된 브랜드는, 미디어 파일이 OMAF 뷰포트-의존적 베이스라인 프리젠테이션 프로파일을 위한 기술을 이용한 VR 경험을 포함한다는 것을 나타낸다. 본 발명은, 한 실시예에서, OMAF 뷰포트-의존적 프로파일에 따른 VR 경험이 서브픽처 트랙을 더 이용함을 나타내는 명시적 브랜드(major_brand 값으로서 또는 호환가능한 브랜드 목록에 포함)를 정의하는 것을 제안한다. 브랜드(주요 또는 호환가능)에 대해 적어도 2개의 특정한 값이 정의될 수 있다.
- [0291] 전방향 의존적 프로파일에 대해, 예를 들어, 'odpr'이라는 이름의 제1 값이 정의될 수 있다. 이 값은, 전방향 미디어가 투사된 픽처를 참조하는 서브픽처 트랙들로 분할됨을 나타낸다. 이 브랜드를 준수하는 임의의 ISOBMFF 파서 또는 OMAF 플레이어는, 서브픽처 트랙 위치를 투사된 픽처에서의 위치로서 해석해야 한다. 또한, total_width와 total_height는 각각 투사된 픽처의 폭과 높이로서 해석되어야 한다.
- [0292] 전방향 의존적 프로파일에 대해, 예를 들어, 'odpa'라는 이름의 제2 값이 정의될 수 있다. 이 값은, 전방향 미디어가 팩킹된 픽처를 참조하는 서브픽처 트랙들로 분할됨을 나타낸다. 이 브랜드를 준수하는 임의의 ISOBMFF 파서 또는 OMAF 플레이어는, 서브픽처 트랙 위치를 팩킹된 픽처에서의 위치로서 해석해야 한다. 또한, total_width와 total_height는 각각 팩킹된 픽처의 폭과 높이로서 해석되어야 한다.
- [0293] 이 브랜드 중 하나가 존재하면, OMAF 플레이어 또는 미디어 플레이어는 기준 픽처 정보를 획득하는 방법을 즉시 식별한다. 그 다음, 이것은 기준 픽처의 표시를 포함하는 공간적 관계 기술을 위해 명시적 트랙 그룹을 파싱한다. 이것은 단계 803에서 이루어진다.
- [0294] 이들 브랜드 중 어느 것도 'ftyp' 박스에 없을 때, 미디어 파일 파서 또는 미디어 플레이어는 미디어 파일을 추가로 파싱하여 서브픽처 트랙들의 존재와 이들이 투사된 픽처 또는 팩킹된 픽처를 참조하는지를 결정해야 한다 (테스트 802의 목적). 공간적 관계를 기술하는 트랙 그룹이 본 발명의 실시예에 따른 명시적 트랙 그룹이라면, 파서는 803에서 이들 명시적 트랙 그룹을 파싱한다. 이것은, 단계 804에서 (예를 들어, track_group_id를 통해 식별된) 주어진 트랙 그룹에서 서브픽처 트랙을 기술하기 위해 이용중인 기준 픽처를 결정한다. 이것은, 선택을 위해 사용자에게 서브픽처 트랙을 프린팅하거나 서브픽처 트랙을 렌더링할 때 고려되어야 한다. 기준 픽처에서 표현된 서브픽처 트랙으로부터 캡처된 픽처로 이미지를 생성하기 위해 추가적인 변환이 요구될 수 있다. 예를 들어, 기준 픽처가 팩킹된 픽처일 때, 투사된 픽처에서 표현되기 위해서는, 서브픽처 트랙 위치 및 크기가 언팩킹되어야 한다. 이 처리는 단계 812의 목적이다. 이제, 캡슐화 단계(713) 동안 명시적 시그널링이 수행되어 단계 803에서 파서에 의해 이용되는 방법을 설명한다.
- [0295] 새로운 브랜드에 대한 대안적인 실시예에서, 트랙 또는 트랙 그룹 레벨에서 명시적 시그널링을 추가하는 것이 제안된다. 이것은, ISOBMFF에서 2D 공간적 관계 기술을 위해 '2dcc' 트랙 그룹을 이용하여 이루어질 수 있다. 이 추가 시그널링은, 특히 투사된 픽처 또는 팩킹된 픽처에 대해 위치와 크기를 표현하는지를 결정하기 위해 파

서 또는 플레이어가 서브픽처 트랙을 처리하는데 도움이 될 수 있다.

[0296] 이러한 시그널링에 대한 한 실시예는, 공간적 관계 기술을 위해 특정한 트랙 그룹 유형 박스에서 새로운 파라미터를 정의하는 것일 수 있다. 바람직하게는, 이것은, 파서가 정보를 획득될 수 있도록, 공간적 관계 기술을 위해 트랙 그룹 박스의 필수 부분인 SpatialRelationship2DSourceBox에서 정의된다.

[0297] 이 실시예의 한 예는 다음과 같다 :

```
aligned(8) class SpatialRelationship2DDescriptionBox extends
TrackGroupTypeBox('2dcc')
{
    // track_group_id is inherited from TrackGroupTypeBox;
    SpatialRelationship2DSourceBox(); // mandatory, must be first
    SubPictureRegionBox ();          // optional
}

aligned(8) class SpatialRelationship2DSourceBox extends
FullBox('2dsr', 0, 0) {
    unsigned int(32) total_width;
    unsigned int(32) total_height;
    unsigned int(32) source_id;
    unsigned int(1) reference_picture;
    unsigned int(31) reserved
}
}
```

[0298] 여기서 "reference_picture"는, 값 "0"을 취할 때, 이 그룹의 서브픽처 트랙에 대한 위치가 투사된 픽처 좌표계로 표현됨을 나타내는 새로운 파라미터이다. 값 "1"을 취하면, 이 그룹의 서브픽처 트랙이 팩킹된 픽처로 표현됨을 나타낸다. 이 파라미터에 주어진 이름은 한 예이다. 또한, total_width와 total_height는 각각 투사된 픽처의 폭과 높이를 나타낸다.

[0300] 단순히 투사된 픽처 또는 팩킹된 픽처 사이에서 기준 픽처의 선택을 지원하는 것보다 더 일반적으로, reference_picture는 캡처와 인코딩 사이에서 기준으로서 이용할 중간 픽처에 대응하는 값으로서의 수개의 값을 취할 수 있다. 예를 들어, 값 0은 어떠한 투사도 없을 때 캡처된 이미지에 대해 이용될 수 있고(단계 702), 값 1은 투사만 있을 때 이용될 수 있고(단계 705), 값 2는 프레임 팩킹된 픽처에 대해(단계 708), 값 3은 팩킹된 프레임에 대해 이용될 수 있다(711). 이 표시는 투사된 및 팩킹된 프레임만을 지원하는 이전 실시예에 비해 2비트를 요구할 것이다.

[0301] 더욱 명시적인 시그널링인 또 다른 실시예는, (정수 값 대신에) 기준 픽처를 기술하는 4cc 코드를 제공하는 것으로 구성된다. 이것은 설명의 측면에서 비용이 더 많이 든다(서브픽처 트랙 당 4 바이트). 예를 들어, 기준 픽처가 투사된 픽처임을 나타내기 위해, 기준 픽처 값은 'povd'로 설정될 수 있다. 팩킹된 픽처의 경우, 이것은 'rwpk'로 설정될 수 있다; 프레임-팩킹된 픽처의 경우, 이것은 'stvi'일 수 있다. 캡처된 이미지의 경우, 디폴트 케이스는, 전용 4 캐릭터 코드 : 캡처된 이미지를 의미하는, "디폴트"에 대해 'dflt' 로 설정될 수 있다. 바람직하게는, 중간 픽처와 정수 코드 사이의 맵핑이 정의되고, 예를 들어 mp4 등록 기관에 의해 기준 픽처 값에 대한 상호운용가능한 코드를 갖도록 등록된다.

[0302] 추가 reference_picture 파라미터는, 대안으로서, SpatialRelationship2DDescriptionBox의 선택사항적인 부분, 즉, SubPictureRegionBox에서 선언될 수 있다. 단계 712에서 명시적 시그널링이 결정될 때 필수 부분에 이것을 포함하는 것이 바람직할 수 있다. 이것은 파서 또는 플레이어가 정보를 발견할 수 있도록 확실히 하기 위한 것이다.

[0303] 또 다른 대안적인 실시예에서, 공간적 관계 기술에 대한 특정한 트랙 그룹 유형 박스에서의 추가 시그널링은, ISOBMFF 또는 OMAF에서 공간적 관계 기술의 구 버전과의 하위호환성을 보존하는 방식으로 정의된다. 이를 위해, TrackGroupTypeBox의 새로운 버전이 정의된다, 예를 들어, 버전 = 1 또는 동일한 버전 = 0이지만 플래그 값을 가짐. 종래 기술의 TrackGroupTypeBox는 플래그 값을 허용하지 않는다는 점에 유의해야 한다. TrackGroupTypeBox에 플래그 값을 제공하는 것은 본 발명의 이 실시예의 일부이다.

[0304] 예를 들어 값 0x01로 설정된 플래그 값 "Reference_info_is_present"는, 이 트랙 그룹이 공간적 관계 정보의 위치 및 크기를 고려하기 위해 기준 픽처에 관한 정보를 포함한다는 것을 나타내도록 정의될 수 있다. 그러면,

2dcc 트랙 그룹은 다음과 같이 표현할 수 있다 :

```
aligned(8) class SpatialRelationship2DDescriptionBox extends
TrackGroupTypeBox('2dcc', 0, flags)
{
    // track_group_id is inherited from TrackGroupTypeBox;
    SpatialRelationship2DSourceBox(flags); // mandatory, must be
first
    SubPictureRegionBox (); // optional
}
aligned(8) class SpatialRelationship2DSourceBox extends
FullBox('2dsr', 0, flags) {
    unsigned int(32) total_width;
    unsigned int(32) total_height;
    unsigned int(32) source_id;
    if ( (flags & 0x01) == 1) {
        unsigned int(1) reference_picture;
        unsigned int(31) reserved
    }
}
```

[0305]

[0306]

여기서 reference_picture는, 값 "0"을 취할 때, 이 그룹의 서브픽처 트랙에 대한 위치가 투사된 픽처 좌표계로 표현됨을 나타내는 새로운 파라미터이다. 파라미터의 이름은 예로서 주어진 것이다. 또한, total_width와 total_height는 각각 투사된 픽처의 폭과 높이를 나타낸다.

[0307]

플래그를 이용하는 것은, 기준 픽처에 관한 모호성이 없을 때, 예를 들어, 2D 고전적 비디오에 대해, 각각의 서브픽처 트랙의 기술 비용을 감소시킨다. 플래그를 이용하여 기준 픽처의 존재 또는 부재를 표시하는 것은, 2dcc 트랙 그룹화 유형을 재사용하여 전방향 콘텐츠를 서브픽처 트랙들로 분할하는 : 영역별 팩킹 단계 포함하거나 포함하지 않는 양쪽 모두의 경우를 처리하는 것을 허용한다.

[0308]

역시 또 다른 실시예에서, TrackGroupingTypeBox의 플래그 파라미터 또는 SpatialRelationship2DDescriptionBox와 같은 그 상속형 박스들 중 하나는 플래그 값에서 직접 기준 픽처를 제공하는데 이용된다. 예를 들어 flags 파라미터가 0으로 설정된 최하위비트를 가질 때, 이것은 전방향 비디오의 경우 기준 픽처가 투사된 픽처임을 의미한다. flags 파라미터의 1로 설정된 최하위비트를 가질 때, 이것은, 전방향 비디오의 경우 기준 픽처가 팩킹된 픽처임을 의미한다. 디폴트 값은 0으로 설정된 flags 파라미터의 최하위비트이다. 이 실시예에서, SpatialRelationship2DSourceBox에 어떠한 추가 파라미터도 없으므로, 파일 설명이 더욱 간결 해진다(서브픽처 트랙 당 4 바이트 절약).

[0309]

대안적인 실시예에서, 묵시적 또는 명시적 서브픽처 트랙 시그널링 사이의 구별은 2개의 상이한 트랙 그룹화 유형을 이용함으로써 이루어진다. 현재의 그룹화 유형은 묵시적 시그널링에 이용되며, 명시적 공간적 관계 트랙 그룹에 대해 새로운 트랙 그룹화 유형이 정의된다. 예를 들어 4 캐릭터 코드 'edcc'가 이용되고 새로운 TrackGroupingTypeBox가 다음과 같이 생성된다 :


```
aligned(8) class ExplicitSpatialRelationship2DDescriptionBox extends
TrackGroupTypeBox('edcc', 0, flags)
{
    // track_group_id is inherited from TrackGroupTypeBox;
    ExplicitSpatialRelationship2DSourceBox(flags); // mandatory,
must be first
    SubPictureRegionBox (); // optional
}
aligned(8) class ExplicitSpatialRelationship2DSourceBox extends
FullBox('edsr', 0, flags) {
    unsigned int(32) total_width;
    unsigned int(32) total_height;
    unsigned int(32) source_id;
    unsigned int(8) reference_picture;
}
}
```

[0310]

[0311]

캡슐화 구성이 "묵시적"인 것으로 결정되면(테스트 801 및 802 거짓), 이것은 어떠한 특정한 시그널링도 이용되지 않음을 의미하고, 파서는 기준 픽처의 묵시적 결정에 들어간다. 이것은, 변환 또는 사후 디코딩 동작이 수행되어야 하고 잠재적으로 기준 픽처를 제공해야 하는 제약된 정보 박스 'rinf'에서 선언된 방식을 파싱함으로써 구성된다. OMAF에 대한 대부분의 경우, 이것은 팩킹된 픽처이거나 투사된 픽처일 수 있다. 양안형 콘텐츠의 경우, 이것은 또한 프레임 팩킹된 픽처일 수 있다. 그 다음, 파서는 OMAF 디스크립터의 존재를 체크하여 후보 기준 픽처를 결정한다. 파서는, 미디어 파일에 어떠한 영역별 팩킹 표시도 없을 때(테스트 810 거짓) 공간적 관계 기술을 위한 위치 및 크기 파라미터가 투사된 픽처에 대해 표현되는 것으로 가정한다. 영역별 팩킹 박스가 존재할 때, 공간적 관계 기술을 위한 위치 및 크기 파라미터는 팩킹된 픽처에 대해 표현된다(단계 811). 선택사항으로서 파서는 공간적 관계를 기술하는 트랙 그룹의 서브픽처 트랙에서 stvi 박스의 존재를 테스트함으로써 프레임-팩킹된 픽처의 존재 또는 부재를 고려할 수 있다(단계 808). 존재한다면, 파서는 프레임-팩킹된 픽처를 후보 기준 픽처로서 기록한다. 더 일반적으로, 묵시적 시그널링의 경우, 캡처(110)와 인코딩(140) 사이의 상이한 처리 단계들로부터 발생하는 최종 픽처에서 표현되는 서브픽처 트랙의 위치 및 크기가 고려된다. 이들 상이한 처리는 제약된 방식 정보 박스 'rinf'에서 반영된다. 예를 들어, 콘텐츠 준비가 투사(120), 프레임 팩킹(125) 및 영역별 팩킹(130)을 포함할 때, RestrictedSchemeInfoBox 'rinf' 박스는, 투사가 적용되었음을 나타내는 'povd' 박스를 그 SchemeTypeBox에 포함한다. 이 'povd' 박스는 자체적으로, 130에서 이루어진 영역별 팩킹을 기술하는 구조를, 예를 들어, RegionWisePackingBox 'rwpk'으로서 포함할 수 있다. 또한, 스테레오 비디오 박스가 예를 들어 CompatibleSchemeTypeBox에 존재하여 단계 125에서 이용된 프레임 팩킹을 나타낸다.

[0312]

최적화된 묵시적 모드 및 폐쇄형 시스템의 경우, 캡슐화 및 파서는 구성 정보를 교환하거나 설정을 정의하여 서브픽처 트랙 설명에 대해 미리 정의된 디폴트 모드를 선언할 수 있다. 예를 들어, 미디어가 전방향 콘텐츠를 포함할 때 서브픽처 트랙이 항상 투사된 이미지를 참조한다는데 동의할 수 있다.

[0313]

도 9는 본 발명의 실시예들에 따른 인코더(950) 또는 디코더(900) 중 적어도 하나와 통신 네트워크(999)를 포함하는 시스템(991, 995)을 나타낸다. 한 실시예에 따르면, 시스템(995)은, 예를 들어, 디코더(900)를 포함하는 사용자 단말기 또는 디코더(900)와 통신가능한 사용자 단말기의 사용자 인터페이스를 통해 디코더(900)에 액세스하는 사용자에게 콘텐츠(예를 들어, 비디오/오디오 콘텐츠를 디스플레이/출력하거나 스트리밍하기 위한 비디오 및 오디오 콘텐츠를)를 처리하고 제공하기 위한 것이다. 이러한 사용자 단말기는, 컴퓨터, 모바일 전화기, 태블릿 또는(제공/스트리밍 된) 콘텐츠를 사용자에게 제공/디스플레이할 수 있는 기타 임의의 유형의 디바이스일 수 있다. 시스템(995)은 통신 네트워크(999)를 통해 (연속 스트림 또는 신호의 형태로 - 예를 들어 이전 비디오/오디오가 디스플레이/출력되는 동안) 비트스트림(901)을 획득/수신한다. 한 실시예에 따르면, 시스템(991)은 콘텐츠를 처리하고, 처리된 콘텐츠, 예를 들어 나중에 디스플레이/출력/스트리밍하기 위해 처리된 비디오 및 오디오 콘텐츠를 저장하기 위한 것이다. 시스템(991)은 예를 들어 인코더(950)에 의해 수신되고 처리되는 본 발명의 실시예들에서의 와이드 뷰 장면에 대응하는 이미지의 원래 시퀀스(951)를 포함하는 콘텐츠를 획득/수신하고, 인코더(950)는 통신 네트워크(991)를 통해 디코더(900)에 전달될 비트스트림(901)을 생성한다. 그 다음, 비트스트림(901)은 다양한 방식으로 디코더(900)에 전달된다, 예를 들어, 인코더(950)에 의해 미리 생성되어 통신 네트워크(999)의 저장 장치(예를 들어, 서버 또는 클라우드 스토리지)에 데이터로서 저장되되, 사용자가 저장 장치로부터 콘텐츠(즉, 비트스트림 데이터)를 요청할 때까지 저장될 수 있고, 요청의 시점에서 데이터는 저

장 장치로부터 디코더(900)로 전달/스트리밍된다. 시스템(991)은 또한, 저장 장치에 저장된 콘텐츠에 대한 콘텐츠 정보(예를 들어, 콘텐츠의 제목 및 콘텐츠를 식별, 선택 및 요청하기 위한 기타의 메타/저장 위치 데이터)를 사용자에게 (예를 들어, 사용자 단말기에 디스플레이될 사용자 인터페이스를 위한 데이터를 전달함으로써) 제공/스트리밍하고, 요청된 콘텐츠가 저장 장치로부터 사용자 단말기로 전달/스트리밍될 수 있도록 콘텐츠에 대한 사용자 요청을 수신 및 처리하기 위한 콘텐츠 제공 장치를 포함할 수 있다. 유리하게는, 본 발명의 실시예들에서, 사용자 단말기는 머리 장착형 디스플레이이다. 대안으로서, 인코더(950)는 비트스트림(901)을 생성하고 이것을 사용자가 콘텐츠를 요청할 때 디코더(900)에 직접 전달/스트리밍한다. 그 다음, 디코더(900)는, 비트스트림(901)(또는 신호)을 수신하고 본 발명에 따른 서브픽처 트랙의 디코딩을 수행하여 비디오 신호(909) 및/또는 오디오 신호를 획득/생성하고, 그 다음, 이것은 사용자에게 요청된 콘텐츠를 제공하기 위해 사용자 단말기에 의해 이용된다.

- [0314] 도 3은 본 발명의 하나 이상의 실시예의 구현을 위한 컴퓨팅 디바이스(300)의 개략적인 블록도이다. 컴퓨팅 디바이스(300)는, 마이크로컴퓨터, 워크스테이션 또는 경량의 휴대형 디바이스 등의 디바이스일 수 있다. 컴퓨팅 디바이스(300)는 통신 버스를 포함하고, 통신 버스는 다음과 같은 요소들에 접속된다 :
- [0315] -마이크로프로세서 등의 중앙 처리 유닛(CPU)(301);
- [0316] -본 발명의 실시예들의 방법의 실행가능한 코드를 저장하기 위한 랜덤 액세스 메모리(RAM)(302) 뿐만 아니라, 매니페스트를 판독하고 기입하기 위한 방법을 구현하거나 및/또는 비디오를 인코딩하거나 및/또는 주어진 파일 포맷 하에서 데이터를 판독하거나 생성하기 위해 필요한 변수 및 파라미터를 기록하도록 적합화된 레지스터, 그 메모리 용량은 예를 들어 확장 포트에 접속된 선택사항적인 RAM에 의해 확장될 수 있다.
- [0317] -본 발명의 실시예들을 구현하기 위한 컴퓨터 프로그램을 저장하기 위한 읽기 전용 메모리(ROM)(303);
- [0318] -전형적으로, 처리될 디지털 데이터가 전송되거나 수신되는 통신 네트워크에 접속되는 네트워크 인터페이스(304). 네트워크 인터페이스(304)는 단일 네트워크 인터페이스이거나, 한 세트의 상이한 네트워크 인터페이스(예를 들어 유선 및 무선 인터페이스, 또는 상이한 종류의 유선 또는 무선 인터페이스)로 구성될 수 있다. 데이터는 전송을 위해 네트워크 인터페이스에 기입되거나 CPU(301)에서 실행중인 소프트웨어 애플리케이션의 제어 하에 수신을 위해 네트워크 인터페이스로부터 판독된다.
- [0319] -사용자로부터 입력을 수신하거나 사용자에게 정보를 디스플레이하기 위한 사용자 인터페이스(UI)(305);
- [0320] -하드 디스크(HD)(306);
- [0321] -비디오 소스 또는 디스플레이 등의 외부 디바이스와 데이터를 송수신하기 위한 I/O 모듈(307).
- [0322] 실행가능한 코드는, 판독 전용 메모리(303), 하드 디스크(306) 또는 예를 들어 디스크 등의 착탈식 디지털 매체에 저장될 수 있다. 한 변형에 따르면, 프로그램의 실행가능한 코드는, 실행되기 전에, 하드 디스크(306) 등의, 통신 디바이스(300)의 저장 수단 중 하나에 저장되기 위하여, 네트워크 인터페이스(304)를 통해 통신 네트워크에 의해 수신될 수 있다.
- [0323] 중앙 처리 유닛(301)은 본 발명의 실시예들에 따른 프로그램 또는 프로그램들의 소프트웨어 코드의 일부 또는 명령어의 실행을 제어하고 지시하도록 적합화되며, 명령어는 전송된 저장 수단 중 하나에 저장된다. 전원을 켜 후, CPU(301)는 예를 들어 프로그램 ROM(303) 또는 하드 디스크(HD)(306)로부터 이들 명령어가 로딩된 후 소프트웨어 애플리케이션과 관련된 메인 RAM 메모리(302)로부터 명령어를 실행할 수 있다. 이러한 소프트웨어 애플리케이션은, CPU(301)에 의해 실행될 때, 이전 도면들에 도시된 플로차트의 단계들이 수행되게 한다.
- [0324] 이 실시예에서, 장치는 본 발명을 구현하기 위해 소프트웨어를 이용하는 프로그램가능한 장치이다. 그러나, 대안으로서, 본 발명은 하드웨어(예를 들어, 주문형 집적 회로 또는 ASIC의 형태)로 구현될 수 있다.
- [0325] 이상, 본 발명이 특정한 실시예를 참조하여 설명되었지만, 본 발명은 특정한 실시예에 제한되지 않으며, 본 발명의 범위 내에 있는 수정은 본 기술분야의 통상의 기술자에게 명백할 것이다.
- [0326] 예를 들어, 본 발명은, 예를 들어 특정한 관심 영역을 줌인하기 위해 TV 또는 멀티미디어 디스플레이용 원격 제어기로서 역할하는, 카메라, 스마트폰, 머리 장착형 디스플레이 또는 태블릿 등의 디바이스에 내장될 수 있다. 또한, 동일한 디바이스로부터 특정한 관심 에어리어들을 선택함으로써 멀티미디어 프리젠테이션의 개인화된 브라우징 경험을 가질 수 있다. 사용자에게 의한 이들 디바이스 및 방법의 또 다른 용도는 선호하는 비디오의 일부 선택된 서브파트를 다른 접속된 디바이스와 공유하는 것이다. 이것은 또한, 감시 카메라가 본 발명에 따른 데

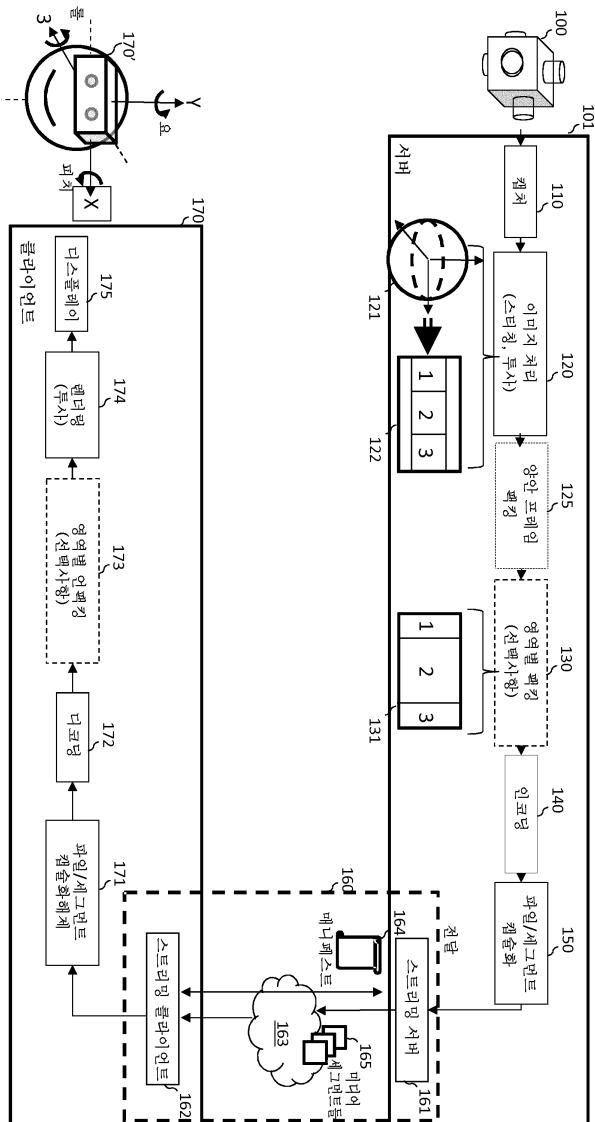
이터를 제공하기 위한 방법을 지원한다면, 감시 중인 건물의 특정한 에어리어에서 발생하는 일을 모니터링하기 위해 스마트폰 또는 태블릿과 함께 이용될 수도 있다.

[0327]

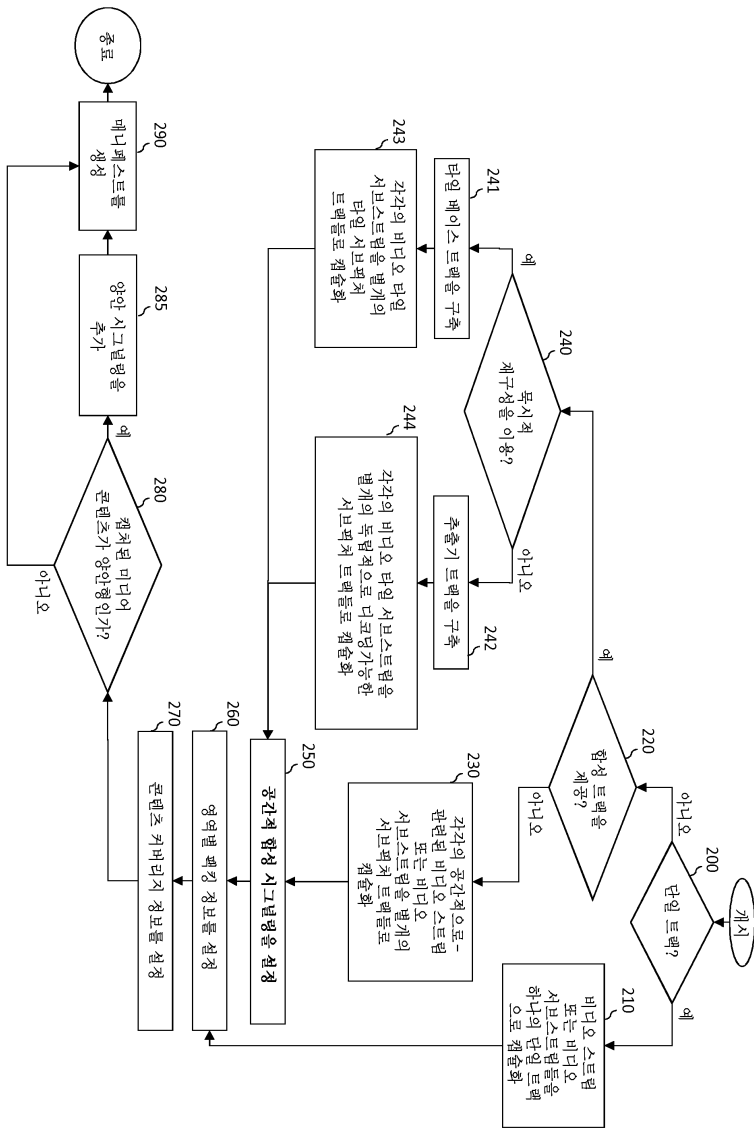
단지 예로서 제공되고 첨부된 청구항들에 의해 결정되는 본 발명의 범위를 제한하기 위한 의도가 아닌 전술된 예시적인 실시예를 참조할 때 많은 추가 수정 및 변형이 본 기술분야의 통상의 기술자에게 명백할 것이다. 특히, 상이한 실시예들로부터의 상이한 피쳐들은 적절한 경우 상호교환될 수 있다.

도면

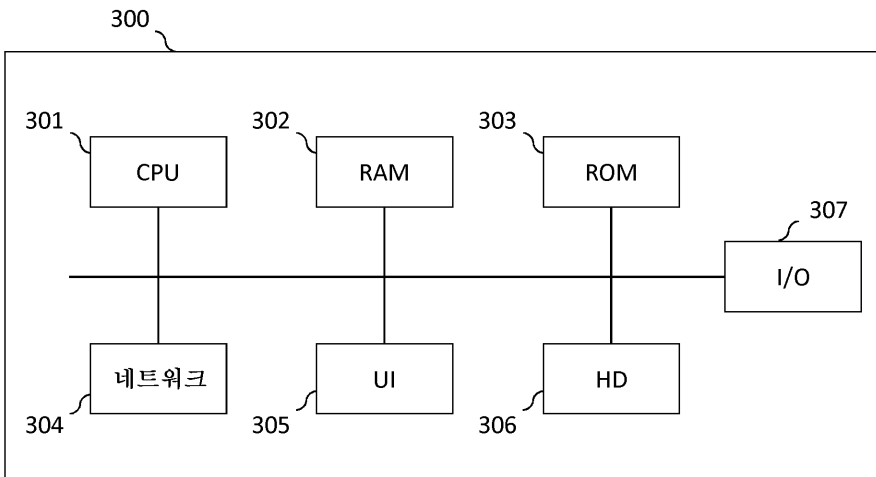
도면1



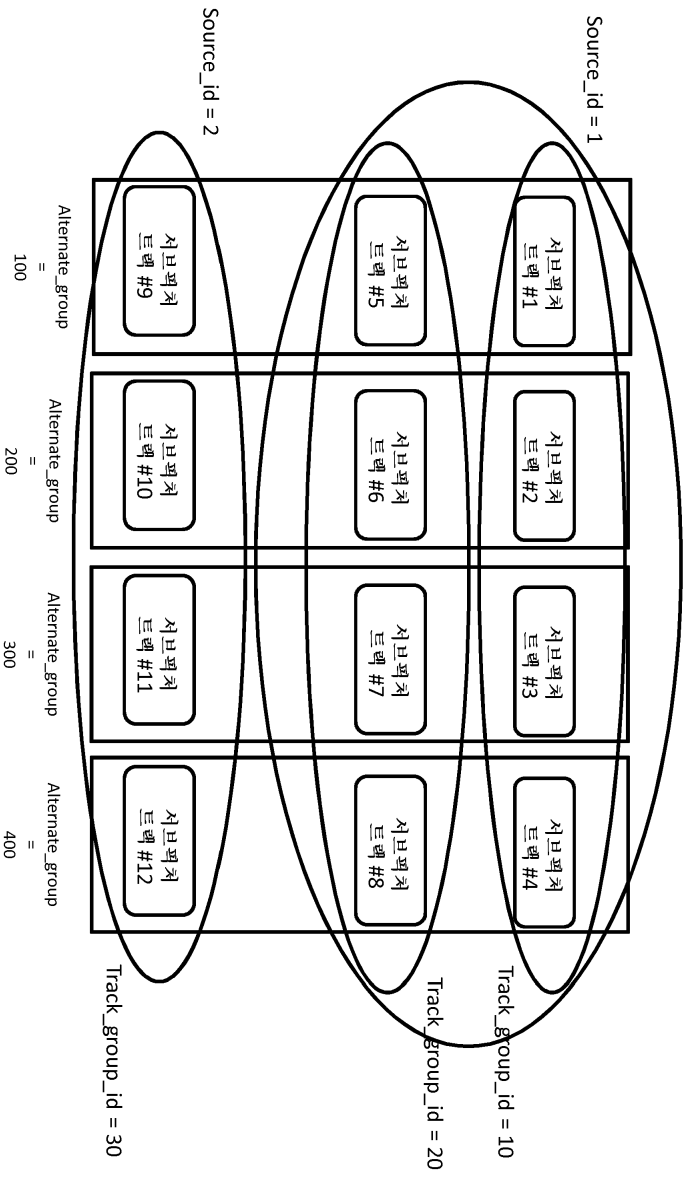
도면2



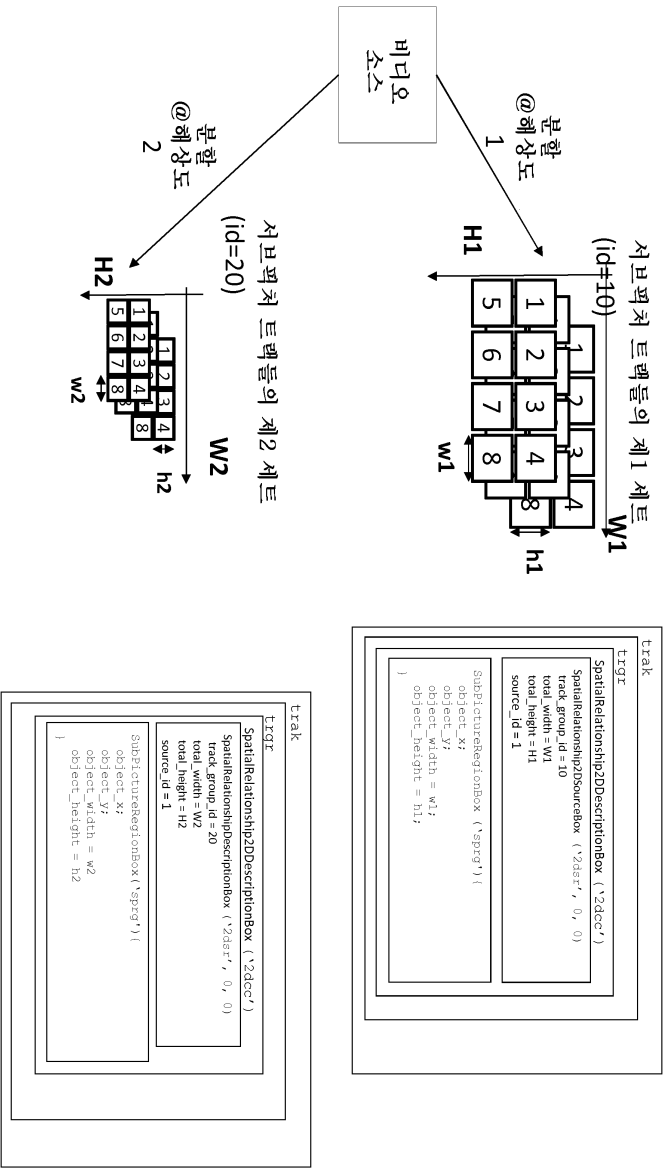
도면3



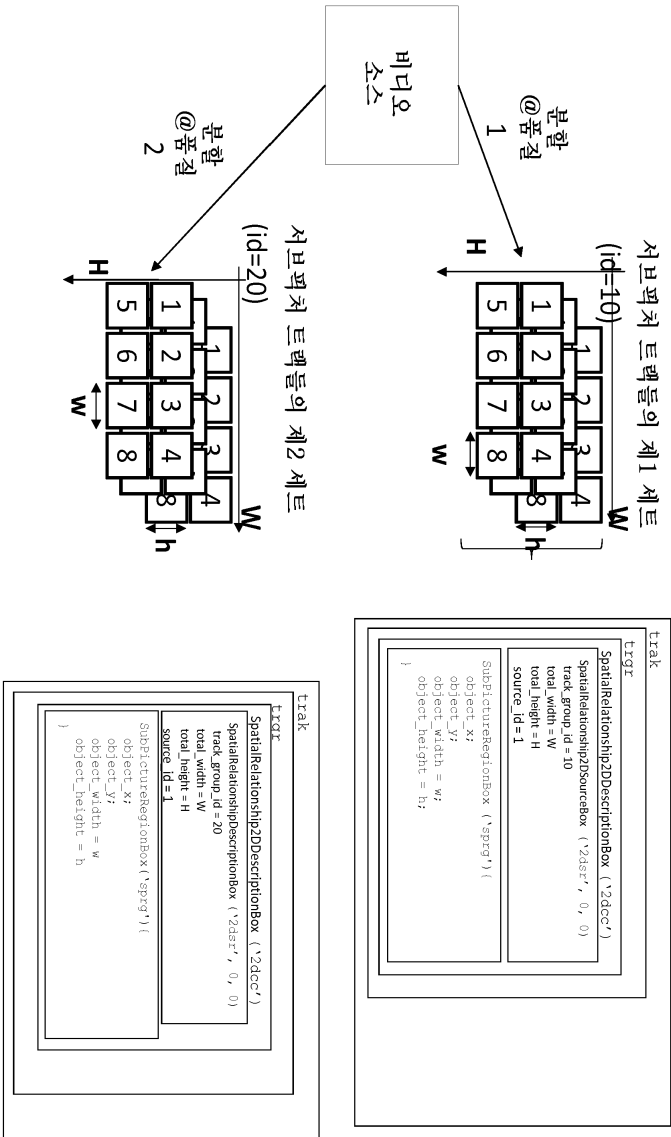
도면4



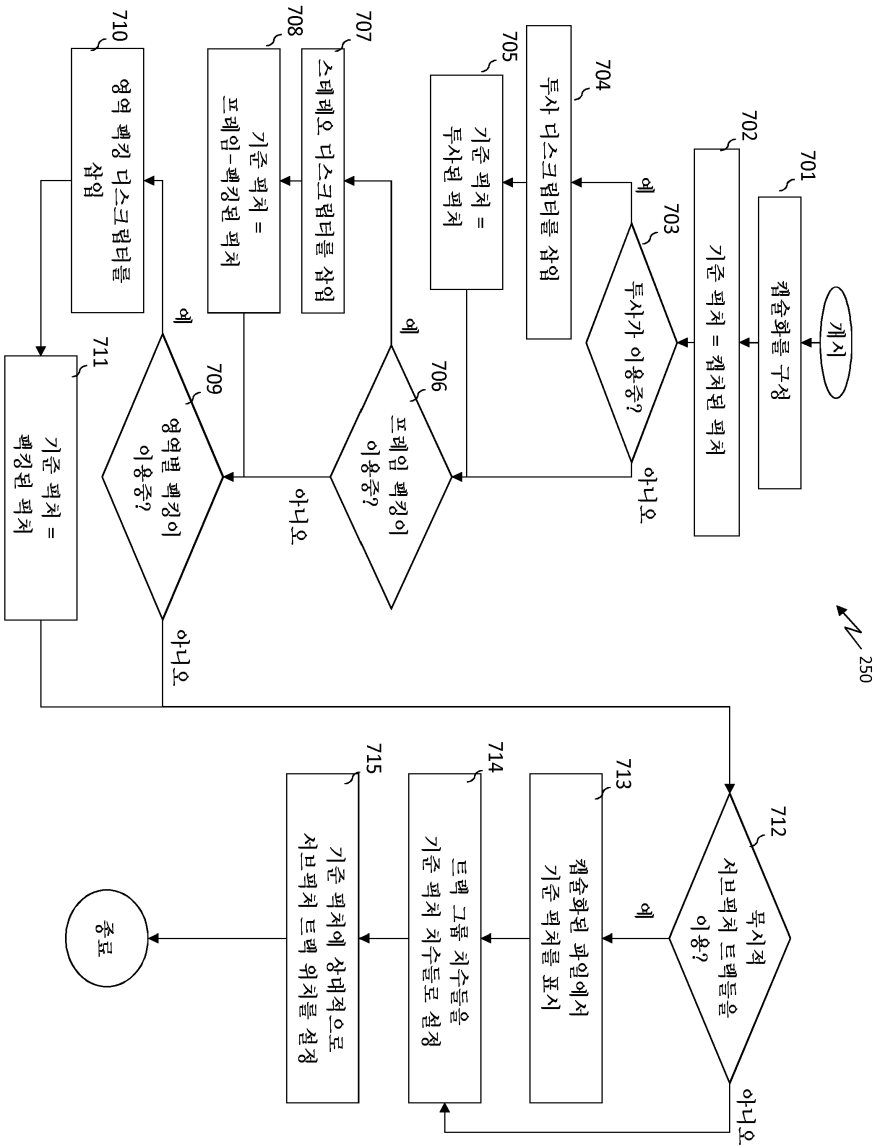
도면5



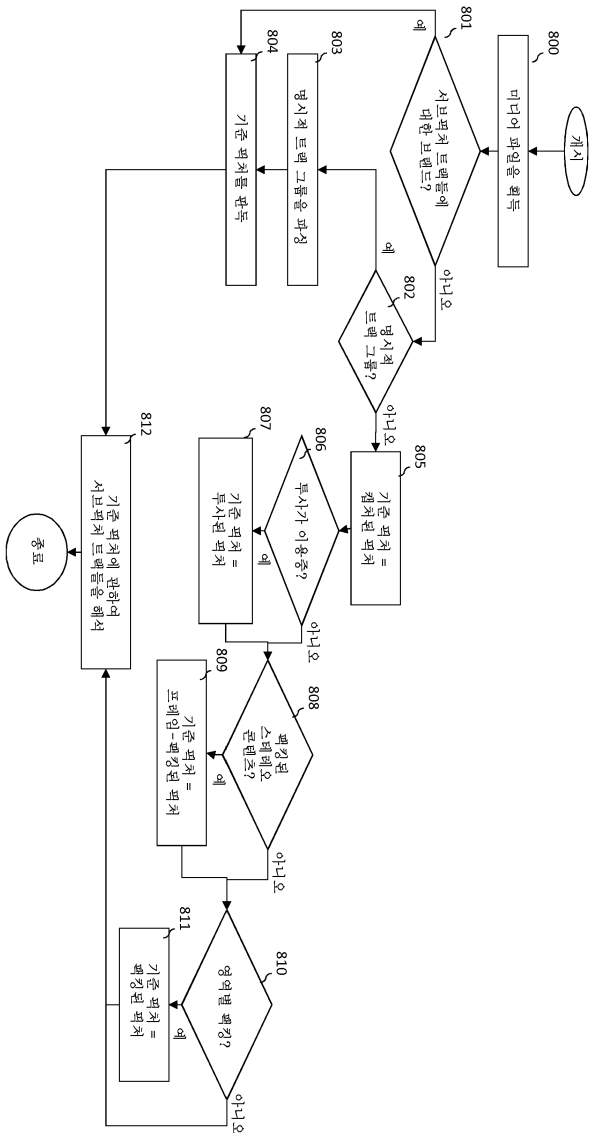
도면6



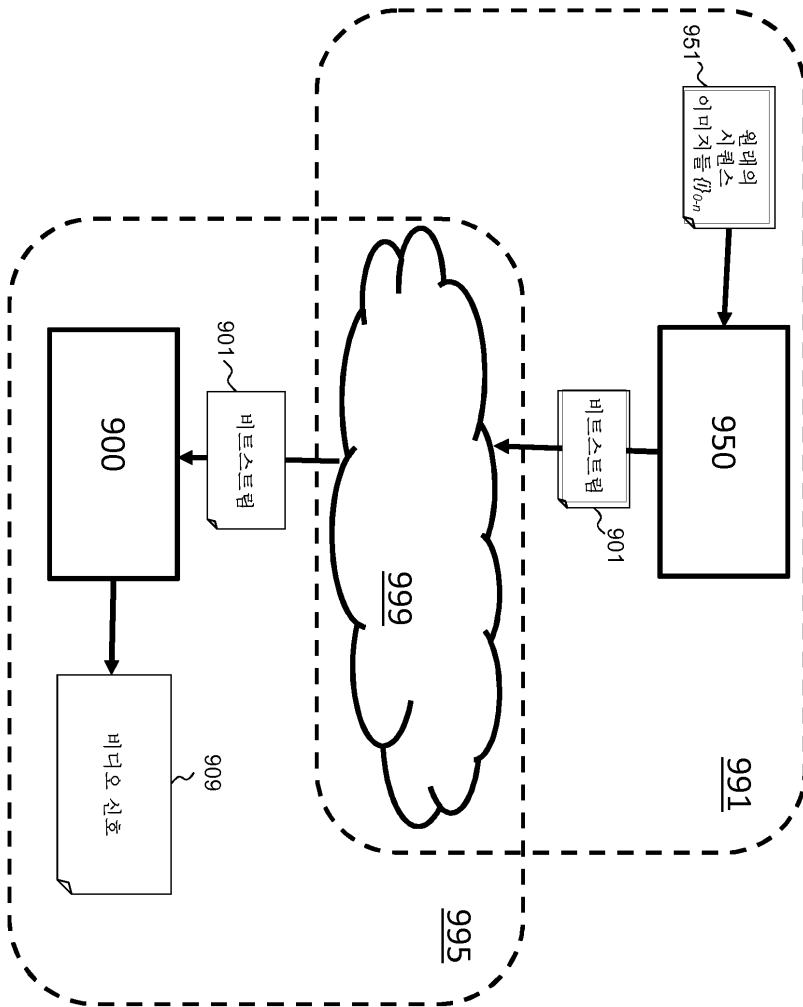
도면7



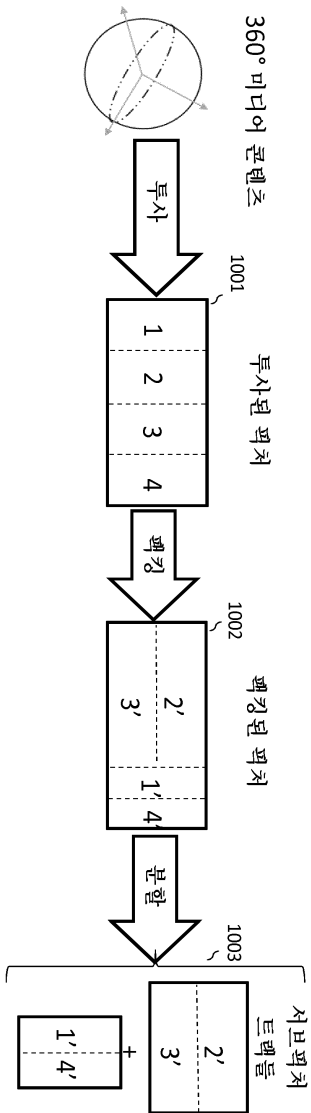
도면8



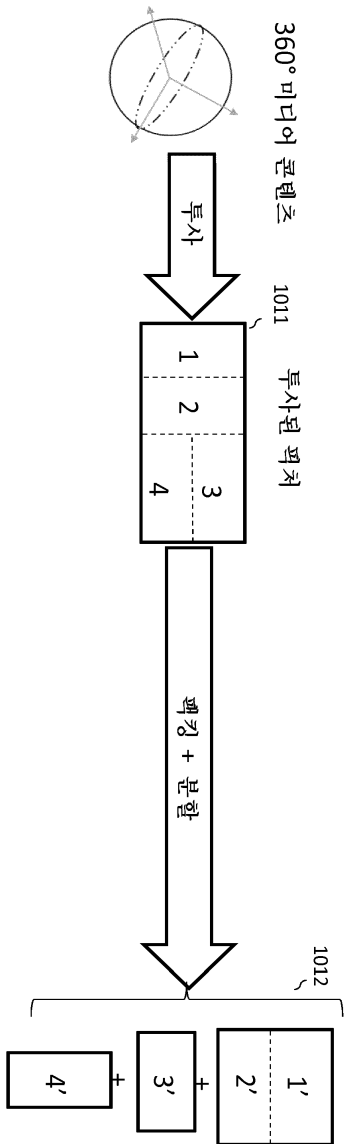
도면9



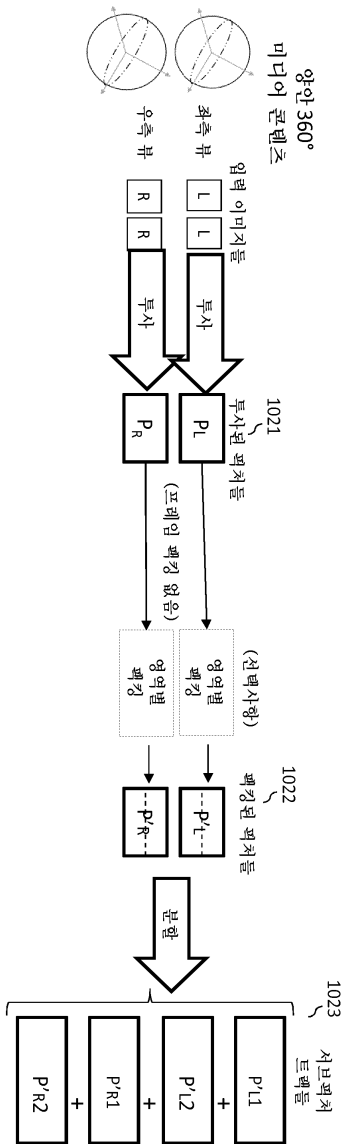
도면10a



도면10b



도면10c



도면10d

