



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2017-0018352
(43) 공개일자 2017년02월17일

- (51) 국제특허분류(Int. Cl.)
HO4N 19/167 (2014.01) *HO4N 19/119* (2014.01)
HO4N 19/174 (2014.01) *HO4N 19/85* (2014.01)
HO4N 21/2343 (2011.01) *HO4N 21/4728*
 (2011.01)
HO4N 21/845 (2011.01) *HO4N 21/854* (2011.01)
HO4N 21/858 (2011.01)
- (52) CPC특허분류
HO4N 19/167 (2015.01)
HO4N 19/119 (2015.01)
- (21) 출원번호 10-2016-7036841
- (22) 출원일자(국제) 2015년06월26일
 심사청구일자 2016년12월29일
- (85) 번역문제출일자 2016년12월29일
- (86) 국제출원번호 PCT/EP2015/064511
- (87) 국제공개번호 WO 2015/197815
 국제공개일자 2015년12월30일
- (30) 우선권주장
 14174815.2 2014년06월27일
 유럽특허청(EPO)(EP)

- (71) 출원인
 코닌클리케 케이피엔 엔.브이.
 네덜란드, 엔엘-2516 씨케이 더 하그, 만플레인 55
 네덜란드 오르가니자티에 포오르 토에게파스트-나 투우르베텐샤펠리스크 온데르조에크 테엔오
 네덜란드 엔엘-2595 디에이 ' 에스-그라벤헤이그 안나 반 뷰렌플레인 1
- (72) 발명자
 토마스, 엠마누엘
 네덜란드, 엔엘-2611 케이피 델프트, 트럼펫스트랏 78
 반 브란덴부르크, 레이
 네덜란드, 엔엘-2595 디씨 더 헤이그, 안나 반 뷰렌플레인 144
 반 디벤테르, 마티이스 오스카
 네덜란드, 엔엘-2264 디엘 레이첸담, 한니 샤프트스트랏 62
- (74) 대리인
 김윤배

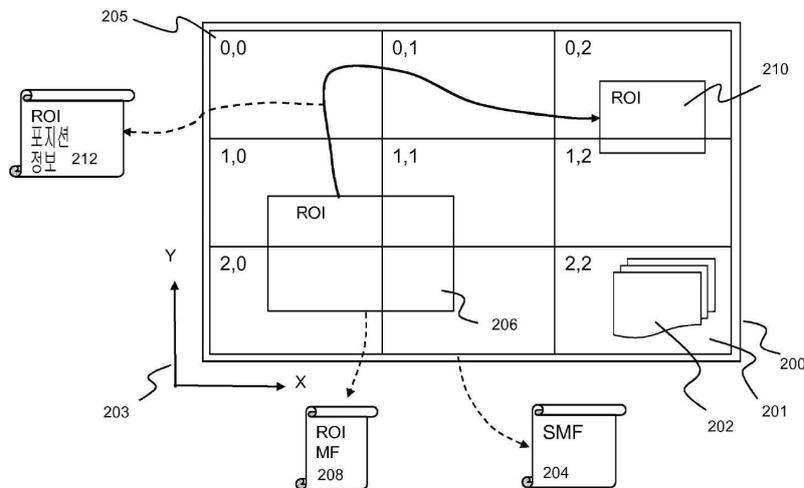
전체 청구항 수 : 총 16 항

(54) 발명의 명칭 HEVC-타일드 비디오 스트림을 기초로 한 관심영역 결정

(57) 요약

적어도 하나의 HEVC-타일드(파노라마) 비디오 스트림을 기초로 하여 클라이언트 장치를 위한 관심영역(ROI)을 결정하는 방법이 설명되고, 본 방법은: 상기 HEVC-타일드 비디오 스트림의 전체 이미지 영역 내에서 제1 서브-영역을 정의하는 제1 ROI를 렌더링하기 위해 ROI 비디오 스트림을 수신하는 단계; 상기 ROI 비디오 스트림의 적어도 (뒷면에 계속)

대표도



하나의 비디오 프레임과 연관된 ROI 포지션 정보를 제공하는 단계; ROI 포지션 정보 및 타일 포지션 정보를 기초로 하여 상기 HEVC-타일드 비디오 스트림의 하나 이상의 HEVC 타일, 바람직하게는 상기 제1 ROI와 오버랩(overlap)하는 하나 이상의 HEVC 타일을 식별하는 단계; 및 상기 HEVC-타일드 비디오 스트림의 전체 이미지 영역 내에서 제2 서브-영역을 정의하는 제2 ROI를 렌더링하기 위해 상기 하나 이상의 식별된 HEVC 타일과 연관된 비디오 데이터를 요청하는 단계;를 포함하고, 상기 HEVC-타일드 파노라마 비디오 내의 HEVC 타일의 포지션은 타일 포지션 정보에 의해 정의되고, 상기 ROI 포지션 정보는 상기 제1 ROI의 적어도 제1 포지션을 포함한다.

(52) CPC특허분류

H04N 19/174 (2015.01)

H04N 19/85 (2015.01)

H04N 21/234327 (2013.01)

H04N 21/234363 (2013.01)

H04N 21/4728 (2013.01)

H04N 21/8455 (2013.01)

H04N 21/85406 (2013.01)

H04N 21/8586 (2013.01)

명세서

청구범위

청구항 1

적어도 하나의 HEVC-타일드(HEVC-tiled) 비디오 스트림을 기초로 하여 클라이언트 장치를 위한 관심영역(region-of-interest: ROI)을 결정하는 방법으로서, 상기 HEVC-타일드 비디오 스트림은 전체 이미지 영역(full image region)을 포함하고, 상기 방법은:

상기 HEVC-타일드 비디오 스트림의 상기 전체 이미지 영역 내에서 제1 서브-영역을 정의하는 제1 ROI를 렌더링하기 위해 ROI 비디오 스트림을, 바람직하게는 상기 클라이언트 장치에서, 수신하는 단계, 상기 HEVC-타일드 비디오 스트림 내의 HEVC 타일의 위치는 타일 위치 정보에 의해 정의됨;

상기 ROI 비디오 스트림의 적어도 하나의 비디오 프레임과 연관된 ROI 위치 정보, 바람직하게는 상기 클라이언트 장치에서, 제공하는 단계, 상기 ROI 위치 정보는 상기 제1 ROI의 적어도 제1 위치를 포함함;

상기 ROI 위치 정보 및 상기 타일 위치 정보를 기초로 하여 상기 HEVC-타일드 비디오 스트림의 하나 이상의 HEVC 타일을, 바람직하게는 상기 클라이언트 장치에 의해, 식별하는 단계, 바람직하게는 상기 제1 ROI와 오버랩(overlap)하는 하나 이상의 HEVC 타일을 식별하는 단계; 및

상기 HEVC-타일드 비디오 스트림의 전체 이미지 영역 내에서 제2 서브-영역을 정의하는 제2 ROI를 렌더링하기 위해 하나 이상의 식별된 HEVC 타일과 연관된 비디오 데이터를 바람직하게는 HEVC 타일 스트림들로서, 바람직하게는 상기 클라이언트 장치에 의해, 요청하는 단계; 및 선택적으로(optionally)

요청된 비디오 데이터 및 상기 ROI 위치 정보를 기초로 하여 상기 제2 ROI를, 바람직하게는 상기 클라이언트 장치에 의해, 렌더링하는 단계;

를 포함하는 것을 특징으로 하는 방법.

청구항 2

청구항 1에 있어서,

상기 ROI 위치 정보를 제공하는 것은:

상기 ROI 비디오 스트림의 적어도 제1 이미지 프레임의 제1 ROI 좌표를 제공하는 것을 더 포함하고,

상기 제1 ROI 좌표는 제2 ROI를 렌더링하기 위한 인스트럭션을 수신하는 제1 순간(time instance)에 상기 제1 ROI의 제1 위치를 정의하고,

바람직하게는 상기 인스트럭션은 사용자-개시 인스트럭션(user-initiated instruction)인 것을 특징으로 하는 방법.

청구항 3

청구항 2에 있어서,

상기 인스트럭션은 상기 제1 ROI의 위치의 변화를 나타내기 위한 ROI 벡터를 포함하고,

상기 제1 ROI 좌표 및 상기 ROI 벡터는 상기 하나 이상의 HEVC 타일을 식별하기 위해 이용되는 것을 특징으로 하는 방법.

청구항 4

청구항 2 또는 3에 있어서,

상기 ROI 위치 정보를 제공하는 것은:

상기 ROI 비디오 스트림의 적어도 제2 이미지 프레임의 제2 ROI 좌표를 제공하는 것을 더 포함하고,

상기 제2 ROI 좌표는 상기 클라이언트 장치에 의해 상기 ROI 비디오 스트림을 처리(processing)하는 제2 순간에

상기 제1 ROI의 제2 포지션을 정의하고, 상기 처리(processing)는 바람직하게는 수신(receiving), 버퍼링(buffering), 및/또는 디코딩(decoding) 중의 하나이고, 상기 제1 ROI 좌표 및 제2 ROI 좌표는 상기 하나 이상의 HEVC 타일을 식별하기 위해 이용되는 것을 특징으로 하는 방법.

청구항 5

청구항 1 내지 4 중 어느 하나의 청구항에 있어서,

상기 요청된 비디오 데이터는 HEVC-인코딩(HEVC-encoded)되고,

상기 제2 ROI의 상기 렌더링은:

요청된 HEVC-인코딩 비디오 데이터를 기초로 하여 디코딩된 이미지 영역(decoded image region)을 형성하는 것; 및 선택적으로(optionally)

상기 디코딩된 이미지 영역으로부터 상기 제2 ROI를 크롭핑하는 것(cropping);

을 포함하는 것을 특징으로 하는 방법.

청구항 6

청구항 1 내지 5 중 어느 하나의 청구항에 있어서,

ROI 포지션 정보를 제공하는 것은:

상기 클라이언트 장치에 의해 아직 수신되지 않은 상기 ROI 비디오 스트림의 적어도 하나의 비디오 프레임과 연관된 룩-어헤드 포지션 정보(look-ahead position information)를 제공하는 것을 더 포함하고,

상기 룩-어헤드 포지션 정보는 상기 전체 이미지 영역 내에서 제2 ROI의 제2 포지션을 포함하는 것을 특징으로 하는 방법.

청구항 7

청구항 6에 있어서,

상기 룩-어헤드 포지션 정보는 상기 클라이언트 장치에 의해 아직 수신되지 않은 상기 ROI 비디오 스트림 내의 하나 이상의 이미지 프레임의 하나 이상의 타임-시프트된(time-shifted) 제1 ROI 좌표를 포함하고,

바람직하게는, 상기 타임-시프트된 제1 ROI 좌표는 상기 하나 이상의 이미지 프레임에 앞서(prior to) 상기 클라이언트 장치에게 제공되고,

더욱 바람직하게는, 상기 타임-시프트된 ROI 좌표는 타임-시프트를 결정하기 위한 타임-시프트 정보, 바람직하게는 타임스탬프와 추가로(further) 연관되어 있는 것을 특징으로 하는 방법.

청구항 8

청구항 1 내지 7 중 어느 하나의 청구항에 있어서,

상기 ROI 포지션 정보의 적어도 일부는 상기 ROI 비디오 스트림의 비트스트림(bitstream) 내에서 상기 클라이언트 장치에 전송되고, 바람직하게는 SEI 메시지 또는 ROI 좌표 플래그(coordinates flag)가 상기 비트스트림 내에서 상기 ROI 좌표의 위치를 정의하는 것; 또는

상기 ROI 포지션 정보의 적어도 일부는 MPEG 트랜스포트 스트림(transport stream) 내에서 상기 클라이언트 장치에 전송되고, 바람직하게는 MPEG 트랜스포트 스트림 내의 기초 트랜스포트 스트림(elementary transport stream) 내에서 또는 MPEG 트랜스포트 스트림 내의 DVB-정의(DVB-defined) SAD(Synchronised Auxiliary Data) 패킷으로서 상기 클라이언트 장치에 전송되고, 선택적으로(optionally) 상기 SAD 패킷은 타임-시프트된 ROI 좌표를 포함하는 것; 또는

상기 ROI 포지션 정보의 적어도 일부는 MPEG-4 스트림 내의 미리 정의된 박스(predefined box) 내에서 전송되고, 바람직하게는 MPEG-4 스트림의 moov, trak, 및/또는 stbl 박스 내에서 전송되는 것; 또는

상기 ROI 포지션 정보의 적어도 일부는 워터마크로서 상기 클라이언트 장치에 전송되고, 바람직하게는 상기 ROI 비디오 스트림 내의 워터마크로서 상기 클라이언트 장치에 전송되는 것;

을 특징으로 하는 방법.

청구항 9

청구항 1 내지 8 중 어느 하나의 청구항에 있어서,

상기 타일 포지션 정보의 적어도 일부는 상기 HEVC-타일드 비디오 스트림 내에서 상기 클라이언트 장치에 제공되고, 바람직하게는 상기 HEVC-타일드 비디오 스트림의 moov 또는 moof 박스 내에서 상기 클라이언트 장치에 제공되는 것; 및/또는,

상기 타일 포지션 정보의 적어도 일부는 공간 매니페스트 파일(spatial manifest file) 내에서 상기 클라이언트 장치에 제공되고, 상기 매니페스트 파일은 상기 HEVC-타일드 비디오 스트림 및/또는 상기 ROI 비디오 스트림을 상기 클라이언트 장치에 전달하도록 구성된 하나 이상의 전송 노드(delivery node)의 위치를 찾아내기(locating) 위한 하나 이상의 스트림 식별자를 포함하는 것;

을 특징으로 하는 방법.

청구항 10

청구항 1 내지 9 중 어느 하나의 청구항에 있어서,

상기 하나 이상의 식별된 HEVC 타일과 연관된 상기 비디오 데이터는, 적어도 하나 이상의 HEVC 타일 식별자, 바람직하게는 하나 이상의 URL, 및 상기 타일 포지션 정보, 및 선택적으로(optionally), 하나 이상의 ROI 비디오 스트림과 연관된 하나 이상의 ROI 비디오 스트림 식별자를 포함하는 공간 매니페스트 파일을 기초로 하여 요청되고, 바람직하게는 상기 하나 이상의 ROI 비디오 스트림 식별자 중의 적어도 하나는 ROI 포지션 정보가 시간에 따라 변한다는 것을 나타내기 위한 동적 인디케이터(dynamic indicator)와 연관되어 있는 것을 특징으로 하는 방법.

청구항 11

관심영역(ROI)을 렌더링하기 위한 클라이언트 장치로서, 상기 클라이언트 장치는:

HEVC-타일드 비디오 스트림의 전체 이미지 영역 내에서 제1 서브-영역을 정의하는 제1 ROI를 렌더링하기 위해 ROI 비디오 스트림을 수신하도록 구성되고;

상기 ROI 비디오 스트림의 적어도 하나의 비디오 프레임과 연관된 ROI 포지션 정보를 제공하도록 구성되고;

ROI 포지션 정보 및 타일 포지션 정보를 기초로 하여 상기 HEVC-타일드 비디오 스트림의 하나 이상의 HEVC 타일을 식별하도록 구성되고;

상기 HEVC-타일드 비디오 스트림의 전체 이미지 영역 내에서 제2 서브-영역을 정의하는 제2 ROI를 렌더링하기 위해, 하나 이상의 식별된 HEVC 타일과 연관된, 바람직하게는 하나 이상의 HEVC 타일 스트림들에 포함된, 비디오 데이터를 요청하도록 구성되고; 그리고, 선택적으로(optionally),

요청된 비디오 데이터 및 상기 ROI 포지션 정보를 기초로 하여 상기 제2 ROI를 렌더링하도록 구성되고;

상기 HEVC-타일드 비디오 스트림 내의 HEVC 타일의 포지션은 타일 포지션 정보에 의해 정의되고,

상기 ROI 포지션 정보는 상기 제1 ROI의 적어도 제1 포지션을 포함하고,

바람직하게는 하나 이상의 HEVC 타일은 상기 제1 ROI와 오버랩하는 것을 특징으로 하는 관심영역(ROI)을 렌더링하기 위한 클라이언트 장치.

청구항 12

비디오 데이터를 저장하기 위한 비-일시적 컴퓨터-판독가능 저장 매체(non-transitory computer-readable storage medium)로서,

상기 저장 매체는 레코딩 영역(recording area)을 포함하고, 상기 레코딩 영역은: ROI 스트림의 비디오 데이터; 및 상기 전체 이미지 영역 내의 상기 서브-영역들의 포지션 정보, 바람직하게는 좌표;를 포함하고,

ROI는 HEVC-타일드 비디오 스트림의 전체 이미지 영역 내의 서브-영역을 정의하고, 바람직하게는, ROI 스트림의

상기 비디오 데이터는 상기 레코딩 영역 내에 하나 이상의 개별 비디오 트랙들로서 저장되고, 포지션 데이터는 상기 레코딩 영역 내의 메타데이터 트랙에 저장되는 것을 포함하는 것을 특징으로 하는 비-일시적 컴퓨터-판독 가능 저장 매체.

청구항 13

청구항 12에 있어서,

상기 포지션 정보는 상기 ROI 스트림의 비트스트림에 삽입되고, 바람직하게는, SEI 메시지 또는 ROI 좌표 플래그가 상기 비트스트림 내의 상기 ROI 좌표의 포지션을 정의하는 것; 또는

상기 비디오 데이터는 MPEG 트랜스포트 스트림으로서 레코딩되고, 상기 ROI 포지션 정보의 적어도 일부는 상기 MPEG 트랜스포트 스트림 내에 들어 있고, 바람직하게는, 상기 ROI 포지션 정보는 기초 트랜스포트 스트림으로서 또는 DVB-정의 SAD(Synchronized Auxiliary Data)로서 상기 MPEG 트랜스포트 스트림에 들어 있고, 선택적으로 (optionally), 상기 SAD 패킷은 타임-시프트된 ROI 좌표를 포함하는 것; 또는

상기 비디오 데이터는 MPEG4 스트림으로서 레코딩되고, 상기 ROI 포지션 정보의 적어도 일부는 미리 정의된 박스 내에서, 바람직하게는 moov, trak, 및/또는 stbl 박스 내에서 상기 MPEG4 스트림 내에 들어 있는 것; 또는

상기 ROI 포지션 정보의 적어도 일부는 워터마크로서 상기 비디오 데이터 내에 삽입되는 것;

을 특징으로 하는 비-일시적 컴퓨터-판독가능 저장 매체.

청구항 14

저장된 데이터 구조(data structure), 바람직하게는 공간 매니페스트 파일을 포함하는, 바람직하게는 청구항 11에 따른 클라이언트 장치에 의해 사용하기 위해 적합한 비-일시적 컴퓨터-판독가능 저장 매체로서, 상기 데이터 구조는:

하나 이상의 HEVC 타일 표현(tile representation),

타일 포지션 정보(tile position information)

를 포함하고,

상기 하나 이상의 HEVC 타일 표현 각각은 HEVC-타일드 비디오 스트림을 나타내고, HEVC 타일 표현은 하나 이상의 HEVC 타일을 식별시키기 위한 하나 이상의 HEVC 타일 식별자를 포함하고,

상기 타일 포지션 정보는 HEVC-타일드 비디오 스트림 내의 HEVC 타일의 포지션을 정의하고,

상기 데이터 구조에 있어서:

- 상기 데이터 구조는 ROI 포지션 정보를 포함하는 ROI 스트림의 적어도 하나의 ROI 스트림 식별자, 바람직하게는 URL을 더 포함하고, 상기 ROI 스트림은 상기 HEVC-타일드 비디오 스트림의 전체 이미지 영역 내에서 서브-영역을 정의하는 제1 ROI를 렌더링하기 위한 비디오 데이터를 포함하는 것; 또는

- 상기 데이터 구조는 ROI 스트림의 적어도 하나의 ROI 스트림 식별자, 바람직하게는 URL, 및 상기 HEVC-타일드 비디오 스트림의 전체 이미지 영역 내의 서브-영역의 포지션을 정의하는 ROI 포지셔닝 정보(positioning information)를 더 포함하고, 상기 ROI 스트림은 상기 HEVC-타일드 비디오 스트림의 전체 이미지 영역 내의 서브-영역을 정의하는 제1 ROI를 렌더링하기 위한 비디오 데이터를 포함하는 것;

중의 적어도 하나인 것을 특징으로 하는 비-일시적 컴퓨터-판독가능 저장 매체.

청구항 15

청구항 14에 있어서,

상기 데이터 구조는 HTTP 적응 스트리밍 프로토콜(adaptive streaming protocol)을 기초로 하여 HEVC 타일드 스트림들로서 타일들을 리트리빙(retrieving)하도록 배열되고,

상기 HEVC 타일드 스트림들 각각은 타일 세그먼트(tile segment)들의 집합이 되도록 시간적으로 세그먼트화되고(temporally segmented), 상기 데이터 구조는 요청될 수 있는 타일 세그먼트들을 식별시키도록 배열되는 것을

특징으로 하는 비-일시적 컴퓨터-판독가능 저장 매체.

청구항 16

컴퓨터의 메모리에서 작동할 때, 청구항 1 내지 10 중의 어느 하나의 청구항에 따른 방법 단계들을 실행하도록 구성된 소프트웨어 코드 부분을 포함하는 컴퓨터 프로그램 제품.

발명의 설명

기술 분야

[0001] 본 발명은 관심영역(region-of-interest: ROI)을 결정하는 것에 관한 것이며, 특히, 전적인 것은 아니지만, 하나 이상의 HEVC-타일드 비디오 스트림을 기초로 하여 관심영역을 결정하기 위한 방법, 상기 관심영역을 결정하고 렌더링(rendering)하기 위한 클라이언트 장치, 관심영역을 결정하고 렌더링하기 위한 비디오 스트림들 및 데이터 구조들을 저장하기 위한 비-일시적 컴퓨터-판독가능 저장 매체(non-transitory computer-readable storage medium), 및 이러한 방법을 이용하는 컴퓨터 프로그램 제품에 관한 것이다.

배경 기술

[0002] 지난 몇 년에 걸쳐서, 카메라 및 이미지 프로세싱 기술들에 있어서의 발전은 더 높은 해상도로 레코딩(recording)하는 것을 가능하게 할 뿐만 아니라, 복수의 카메라들의 출력을 함께 스티칭(stitching)하는 것을 가능하게 해서, 8Kx4K보다 더 높은 해상도로 완전한 360도로 함께 레코딩하는 한 세트의 카메라들을 가능하게 한다.

[0003] 이러한 발전은 사용자가 비디오를 경험하는 방식을 변화시키는 것을 가능하게 한다. 전형적으로, 예컨대 축구 경기의 브로드캐스트는 감독에 의해 주의깊게 정렬되고 제어된 카메라 샷들의 시퀀스(sequence)를 포함한다. 이러한 브로드캐스트 스트림(broadcast stream)에서, 최종 스트림에서의 각각의 카메라 움직임은 카메라 자체의 포지션(position), 앵글(angle), 또는 줌 레벨(zoom level)에 대한 물리적 변화에 상응한다. 하지만, 고해상도 파노라마 비디오들은 사용자(및/또는 감독)로 하여금 물리적 차원에서 카메라를 조작할 필요 없이 사용자(및/또는 감독)가 감상(감독)하고 있는 비디오와의 어느 정도의 상호작용을 할 수 있게 한다. 팬-틸트-줌(pan-tilt-zoom) 상호작용을 이용함으로써, 사용자 또는 감독의 서브-영역(sub-region)을 고해상도 파노라마 비디오로부터 추출하는 것이 가능하다. 이 서브-영역은 관심영역(ROI)이라고 지칭될 수 있다.

[0004] 이러한 특별한 사용의 경우에 있어서, 특정한 사용자는 임의의 주어진 순간에 전체 비디오 파노라마의 서브세트(subset)만을 보고 있기 때문에, 대역폭 요구(bandwidth requirements)는 사용자가 관심있는 비디오의 일부분만을 보냄으로써 감소될 수 있다. 이러한 기능이 달성될 수 있는 다수의 기법들이 존재한다. 이러한 기법들 중의 하나는 소위 타일드 스트리밍 기법(tiled streaming technique)인데, 이 기법을 이용해서 전체 비디오 파노라마는 복수의 독립적으로 인코딩된 비디오들로 분할되고, 클라이언트는 필요하다면 다수의 이러한 독립적인 비디오들을 함께 스티칭함으로써 클라이언트로 하여금 전체 비디오 파노라마의 임의의 부분을 재구성할 수 있게 하는 복수의 디코더(multiple decoder)들을 가진다.

[0005] 하지만, 대부분의 사용자 시나리오들에서, 사용자가 비디오 파노라마와 지속적으로 상호작용하는 것은 필수적인 것이 아니며 심지어 바라던 것도 아니다. 예를 들어, 축구 경기의 경우에, 예컨대 오프-스크린 파울(off-screen foul)이 행하여진 경우에 사용자는 특정 시점에서 비디오와 상호작용하는 데에 관심이 있지만, 감독은 대신 공을 따라가도록 결정하는 것을 상상해 볼 수 있다. 하지만, 대부분의 시간에, 사용자는 감독의 리드(lead)를 따르길 원할 수 있다. 이러한 경우들에 있어서, 타일드 스트리밍을 이용하는 것은 비디오를 스트리밍하는 가장 효율적인 방법이 아니다. 이에 대한 하나의 이유는 타일 사이즈에 의한 제한으로 인해 인코딩 체계에서 이용되는 시공간적 예측들이 최적화되어 있지 않다는 사실이 존재하기 때문이다. 두 번째 이유는 타일들의 임의의 주어진 세트가 거의 절대로 정확하게 ROI를 포함하지 않을 것이기 때문에 주어진 ROI를 재구성하기 위해서 엄밀히 필요하지 않은 몇몇 추가적인 픽셀들/매크로블록(macroblock)들이 보내진다는 점이다.

[0006] Mavlankar, A. 등에 의한, "An interactive region-of-interest video streaming system for online lecture viewing," Packet Video Workshop (PV), 2010 18th International, vol., no., pp.64-71, 13-14 Dec. 2010 논문에서, '트래킹 타일(Tracking Tile)'의 개념이 기술된다. 이 트래킹 타일은 비디오 파노라마의 지속적으로 갱신되는 크롭(crop)으로 이루어진 독립적으로 인코딩된 비디오이다. 트래킹 타일 모드를 선택함으로써, 사용자는 능동적으로(actively) 비디오 파노라마를 탐색(navigate around)할 필요없이, 비디오 파노라마에서 소정의

움직이는 관심지점(예컨대, 축구 경기에서 공의 포지션)을 수동적으로 따라갈(passively) 것이다. 규칙적인 브로드캐스트들이 일어나거나 자동 이미지 인식 시스템에 의해 제어되는 방식과 유사하게 감독에 의해서 제어될 비디오 파노라마 내의 트래킹 타일의 포지션을 상상할 수 있다.

[0007] 트래킹 타일의 이용은 사용자들에게 비디오 파노라마를 탐색할 자유를 제공하고, 이들이 수동적으로 관심지점(point of interest)을 따라가는 것을 가능하게 하지만, Mavlankar 등에 의해 기술된 트래킹 타일 개념은 파노라마를 탐색하는 사용자에게 트래킹 타일을 따라가는 것으로부터 매끄러운 전환(seamless transition)을 수행할 능력을 제공하지 않는다. 만일 비디오 파노라마를 누비며 축구 경기에서 공을 따라가고 있는 트래킹 타일을 보고 있는 사용자를 상상한다면, 사용자는 예를 들어 파울이 행하여진 것을 보기 위하여 갑자기 왼쪽으로 약간 더 움직이길 원할 수 있다.

[0008] Mavlankar 등에 의해서 기술된 바와 같은 트래킹 타일 체계(tracking tile scheme)에 있어서, 매끄러운 전환이 가능하지 않은데, 왜냐하면 시스템은 주어진 시점에서 비디오 파노라마 내의 트래킹 타일의 위치를 알지 못하기 때문이다. 그러므로, 사용자는 파노라마에서 특정 위치를 찾아야 할 것이고, 이로 인하여 연속적인 경험에 지장을 줄 것이다. 사용자가 관심이 없는 비디오 타일이 렌더링되고 디스플레이되기 때문에, 트래킹-타일 모드(tracking-tile mode)와 사용자-제어 모드(user-control mode) 간의 스위칭은 상당한 시간이 걸리고, 자원(resource)의 낭비를 초래한다.

[0009] 그래서, 클라이언트에게 광시야(wide field-of-view) 이미지 영역의 ROI의 효율적인 스트리밍을 하는 것을 가능하게 하는 향상된 방법들 및 시스템들을 위한 필요성이 본 기술분야에 존재한다. 게다가, (논-타일드 모드(non-tiled mode)에서) 클라이언트에게 단일한 스트림을 기초로 한 ROI의 스트리밍과 (타일드 모드(tiled mode)에서) 클라이언트에게 하나 이상의 개별 타일 스트림들을 기초로 한 ROI의 스트리밍 간의 스무딩한(smooth) 또는 심지어 매끈한(seamless) 스위칭을 가능하게 할 필요성이 본 기술분야에 존재한다.

발명의 내용

해결하려는 과제

[0010] 본 발명의 목적은 선행기술에서 알려진 단점들 중의 적어도 하나를 감소시키거나 제거하는 것이다.

과제의 해결 수단

[0011] 제1 관점에 있어서, 본 발명은 적어도 하나의 HEVC-타일드(파노라마) 비디오 스트림을 기초로 하여 클라이언트 장치를 위한 관심영역(ROI)을 결정 및/또는 렌더링하는 방법과 관련이 있을 수 있다.

[0012] 실시예에서, 본 방법은: 상기 HEVC-타일드(HEVC-tiled) 비디오 스트림의 전체 이미지 영역(full image region) 내에서 제1 서브-영역을 정의하는 제1 ROI를 렌더링하기 위해 ROI 비디오 스트림을, 바람직하게는 클라이언트 장치에서, 수신하는 단계; 상기 ROI 비디오 스트림의 적어도 하나의 비디오 프레임과 연관된 ROI 포지션 정보를, 바람직하게는 상기 클라이언트 장치에서, 제공하는 단계; 상기 ROI 포지션 정보 및 타일 포지션 정보를 기초로 하여 상기 HEVC-타일드 비디오 스트림의 하나 이상의 HEVC 타일을, 바람직하게는 상기 클라이언트 장치에 의해, 식별하는 단계, 바람직하게는 상기 제1 ROI와 오버랩하는 하나 이상의 HEVC 타일을 식별하는 단계; 상기 HEVC-타일드 비디오 스트림의 전체 이미지 영역 내의 제2 서브-영역을 정의하는 제2 ROI를 렌더링하기 위해 하나 이상의 식별된 HEVC 타일과 연관된 비디오 데이터를, 바람직하게는 클라이언트 장치에 의해, 요청하는 단계를 포함할 수 있고, 상기 HEVC-타일드 파노라마 비디오 내의 HEVC 타일의 포지션은 타일 포지션 정보에 의해 정의되고, 상기 ROI 포지션 정보는 상기 제1 ROI의 적어도 제1 포지션을 포함한다. 본 방법의 단계들을 실행하는 클라이언트 장치의 관점에서 본 발명의 베스트 모드(best mode)가 정의되지만, 본 발명은 본 발명에 따른 방법의 소정의 단계들이 클라이언트 장치가 아닌 다른 개체들에 의해 실행되는 실시예들도 예상한다는 점이 주목되어야 한다. 예를 들어, 만일 ROI 포지션 정보 및 타일 포지션 정보가 클라이언트 장치가 아닌 다른 개체에 제공된다면(예컨대, 서버 컴퓨터와 같이, 클라이언트 장치에 연결된 장치), 이러한 다른 개체(다른 장치)는 상술한 바와 같은 식별하는 단계를 실행할 수 있다.

[0013] 그래서, ROI 스트림은 HEVC-타일드 비디오 스트림의 전체 이미지 영역의 서브-영역(ROI)의 콘텐츠(예컨대 비디오 데이터)를 포함할 수 있다. HEVC-타일드 비디오는 고해상도(high-resolution) 또는 초-고(ultra-high) 해상도 비디오일 수 있다. 이와 달리, HEVC-타일드 스트림은 다수의 상이한 비디오들, 예컨대 상이한 카메라 앵글들로부터 형성될(구성될/스티칭될) 수 있다. ROI 스트림의 플레이어아웃 동안, ROI의 포지션(HEVC-타일드 스트림의

HEVC-타일드 이미지 영역 내의 서브-영역의 공간적 포지션 또는 지향)은 시간에 따라 변할 수 있다.

- [0014] 논-타일드 ROI 스트림, 예컨대 전형적인 논-타일드 MPEG 스트림을 기초로 하여 ROI를 렌더링하는 것(ROI 스트리밍 모드)으로부터, HEVC 타일이라고도 지칭되는 하나 이상의 HEVC-타일드 스트림을 기초로 하여 ROI를 렌더링하는 것(타일드 스트리밍 모드)으로 스위칭하는 경우에, ROI 비디오 스트림의 적어도 하나의 이미지 프레임의 포지션(즉, ROI 좌표)을 포함하는 ROI 포지션 정보는 클라이언트 장치가 HEVC 타일 스트리밍 모드를 시작할 때 이용(예컨대, 요청(request), 리트리빙(retrieve), 및 렌더링(render))할 수 있는 하나 이상의 HEVC 타일의 HEVC-인코딩 비디오 데이터를 결정하기 위하여 이용될 수 있다. ROI 스트림(ROI 비디오 스트림이라고도 지칭됨)의 ROI의 포지션은 프레임 레벨(frame level), 즉 클라이언트 장치에 의해 수신되었던 비디오 프레임들에서 제공될 수 있어서, 타일드 스트리밍 모드로의 스무딩한 또는 매끄러운 전환이 실현될 수 있다.
- [0015] HEVC 표준은 소위 HEVC 타일을 지원한다.
- [0016] HEVC 인코더는 비디오 내의 비디오 프레임들을 소위 소위 HEVC 타일들로 나누도록 구성될 수 있고, HEVC 타일은 CTU들의 특정한 그룹핑(grouping)에 의해 정의된 사각형 영역이다. HEVC 타일들은 파노라마 비디오의 이미지 영역을 복수의 인접한 사각형 영역들로 나눌 수 있고, HEVC 타일들의 경계들은 칼럼 및 로우 방향으로 HEVC 타일 경계들에 의해 정의된다. 이 애플리케이션에서, HEVC 타일들을 포함하는 HEVC 인코딩 비디오 스트림(HEVC 스트림, HEVC 비디오, 또는 HEVC 비디오 스트림이라고도 지칭됨)은 HEVC-타일드 비디오 스트림이라고 지칭될 수 있다. HEVC 타일들은 개별적으로 요청 및 리트리빙될 수 있고, 그래서 HEVC 타일 스트림들이라고도 지칭된다. 이 HEVC 타일 스트림들/HEVC 타일들은 함께 HEVC 인코딩 비디오를 형성한다.
- [0017] HEVC 타일들은 멀티-코어(multi-core) 프로세서들을 이용해서 비디오 데이터를 인코딩 및 디코딩하기 위해 HEVC 표준에서 원래 도입되었고, 그래서 HEVC-타일드 스트림 내의 타일들은 패럴렐하게(in parallel) 처리(인코딩 및 디코딩)될 수 있다. HEVC-타일들은 또한 HEVC-타일드 스트림의 HEVC 타일들의 서브세트만의 리트리빙 및/또는 플레이아웃을 위해 이용될 수 있다.
- [0018] 이러한 서브세트는 예컨대, 파노라마 비디오(HEVC-타일드 비디오)의 이미지 영역(전체 이미지 영역) 내의 관심 영역(ROI), 예컨대 서브영역과 관련이 있을 수 있다. 이 경우에, HEVC 타일들은 디코더가 복수의 프레임들에 대해 HEVC 타일들의 서브세트만을 디코딩할 수 있도록 시간의 경과에 대해 독립적으로 인코딩되어야 한다. 독립적으로 디코딩가능한 HEVC 타일들의 이러한 세트들을 생성하기 위하여, HEVC 표준은 HEVC 인코더가 하나 이상의 HEVC 타일들의 경계들 내에서 비디오 코딩(예컨대 모션 벡터들 및 인-루프 필터들)에 있어서 시간적 예측(temporal prediction)을 제한하도록 구성되는 것을 가능하게 한다. 단일한 HEVC 디코더에 의한 HEVC 타일들의 처리는 각각의 HEVC 타일 스트림에 대해 별개의(separate) 디코더(또는 디코더 인스턴스(decoder instance))를 필요로 할 수 있는 다른 코덱들을 이용하는 선행 기술에서 기술된 톤링 체계(tiling scheme)들에 비하여 자원들의 관점에서 현저한 이점을 제공한다.
- [0019] 이 공개에서 기술되는 바와 같은 타일들의 개념은 상이한 비디오 코덱들에 의해 지원될 수 있다. 예를 들어, HEVC(High Efficiency Video Coding) 표준은 독립적으로 디코딩가능한 타일들(HEVC 타일들)의 이용을 가능하게 한다. HEVC 타일들은 미디어 스트림의 각각의 비디오 프레임을, CTB(coding tree block)의 단위로 표현되는 미리 정의된 폭 및 높이의 타일들을 정의하는 다수의 로우 및 칼럼("타일들의 그리드(a grid of tiles)")로 나누는 인코더에 의해 생성될 수 있다. HEVC 비트스트림(bitstream)은 비디오 프레임들이 어떻게 타일들로 나누어지는지에 대해서 디코더에게 알려주기 위한 디코더 정보를 포함할 수 있다. 디코더 정보는 상이한 방법으로 비디오 프레임들의 타일 분할에 대해 디코더에게 알려줄 수 있다. 하나의 변형례에 있어서, 디코더 정보는 $n \times m$ 타일들로 이루어진 균일한 그리드(uniform grid)에 관한 정보를 포함할 수 있고, 그리드 내 타일들의 사이즈는 프레임들의 폭 및 CTB 사이즈를 기초로 하여 추론될(deduce) 수 있다. 라운딩 부정확성들(rounding inaccuracies)로 인하여, 모든 타일들이 정확하게 동일한 사이즈를 가질 수 있는 것은 아니다. 다른 변형례에서, 디코더 정보는 (예컨대 코딩 트리 블록(coding tree block) 단위의 관점에서) 타일들의 폭 및 높이에 관한 명시적인 정보(explicit information)를 포함할 수 있다. 이러한 방식으로, 비디오 프레임들은 상이한 사이즈의 타일들로 나누어질 수 있다. 마지막 로우 및 마지막 칼럼의 타일들에 대해서만, 사이즈는 남은 수의 CTB들부터 유도될(derived) 수 있다. 그 이후에, 패킷화기(packetizer)는 원본(raw) HEVC 비트스트림을 전송 프로토콜에 의해 이용되는 적절한 미디어 컨테이너(media container)로 패킷화할(packetize) 수 있다.
- [0020] 독립적으로 디코딩가능한 타일들을 지원하는 다른 비디오 코덱들은 Google의 비디오 코덱 VP9 또는 - 어느 정도는 - MPEG-4 Part 10 AVC/H.264, AVC(Advanced Video Coding) 표준을 포함한다. VP9에서 코딩 의존성들(coding dependencies)은 수직적인 타일 경계들을 따라서 깨지고, 이것은 동일한 타일 로우 내의 두 개의 타일들이 동시

에 디코딩될 수 있다는 것을 의미한다. 유사하게, AVC 인코딩에서, 슬라이스(slice)들은 각각의 프레임의 복수의 로우들로 나누기 위해 이용될 수 있고, 이 로우들 각각은, 미디어 데이터가 독립적으로 디코딩가능(independently decodable)하다는 점에서 타일을 정의한다. 그래서, 본 공개에서, "HEVC 타일(HEVC tile)"이라는 용어는 HEVC 표준에 따른 타일들에만 한정되는 것이 아니라, 비디오 프레임들의 이미지 영역 내에서 임의의 형상 및/또는 차원의 서브영역을 일반적으로 정의하고, 여기서 타일의 경계들 내의 미디어 데이터는 독립적으로 디코딩가능하다. 다른 비디오 코덱들에서, 세그먼트 또는 슬라이스와 같은 다른 용어들이 이러한 독립적으로 디코딩가능한 영역들을 위해 이용될 수 있다. 그래서, 이 코덱들이 비디오를 인코딩하기 위해 적절한 특성을 가지고, 이로써 비디오를 나타내는 이미지들의 상이한 영역들(서브 영역들)이 단일한 인코딩 프로세스에서 독립적으로 인코딩될 수 있고, 이로써 독립적으로 인코딩된 영역들이 단일한 디코딩 프로세스에서 디코딩될 수 있는 한, 본 발명은 HEVC(예컨대 VP9)와 상이하거나 HEVC로부터의 (장래의) 파생물(derivative)들인 비디오 코덱을 가지고 이용하기 위해 동등하게 적합하다는 점이 추가로 주목되어야 한다. 독립적으로(independently)라는 용어는 이 영역들 사이에 아무런 인코딩 의존성들이 존재하지 않는 방식으로 코딩이 수행된다는 개념을 지칭한다.

[0021] 실시예에서, 상기 방법은: 상기 요청된 비디오 데이터 및 상기 ROI 포지션 정보를 기초로 하여 상기 제2 ROI를 렌더링하는 것을 더 포함할 수 있다.

[0022] 실시예에서, 상기 ROI 포지션 정보를 제공하는 것은: 상기 ROI 비디오 스트림의 적어도 제1 이미지 프레임의 제1 ROI 좌표를 제공하는 것을 더 포함할 수 있고, 상기 제1 ROI 좌표는 제2 ROI를 렌더링하기 위한 인스트럭션을 수신하는 제1 순간(time instance)에 상기 제1 ROI의 제1 포지션을 정의하고, 바람직하게는 상기 인스트럭션은 사용자-개시 인스트럭션(user-initiated instruction)이다. 이 실시예에서, 예컨대 사용자-상호작용을 통해서 클라이언트 장치가 지시받는 순간에 클라이언트 장치에 의해 처리되는 프레임의 ROI 좌표는 HEVC 타일 스트리밍 모드(즉, 미리 결정된 ("director-cut") ROI를 렌더링하는 ROI 스트림 대신 HEVC 타일들을 기초로 한 사용자-개시 ROI의 렌더링)로 스위칭하기 위해 필요한 하나 이상의 HEVC 타일들을 결정하는 데에 이용될 수 있다.

[0023] 실시예에서, 상기 인스트럭션은 상기 제1 ROI의 포지션의 변화를 나타내기 위한 ROI 벡터를 포함할 수 있고, 상기 제1 ROI 좌표 및 상기 ROI 벡터는 상기 하나 이상의 HEVC 타일들을 식별하기 위해 이용된다. ROI 벡터는 소정의 사용자-상호작용, 예컨대 패닝(panning), 줌(zooming), 및/또는 틸팅(tilting)과 연관되어 있을 수 있고, 요청된 사용자-상호작용을 위해 적합한 타일 표현(tile representation)을 선택하기 위해 이용될 수 있다.

[0024] 실시예에서, 상기 ROI 포지션 정보를 제공하는 것은: 상기 ROI 비디오 스트림의 적어도 제2 이미지 프레임의 제2 ROI 좌표를 제공하는 것을 더 포함할 수 있고, 상기 제2 ROI 좌표는 상기 클라이언트 장치에 의해 상기 ROI 비디오 스트림을 처리(예컨대 수신, 버퍼링 및/또는 디코딩)하는 제2 순간에 상기 제1 ROI의 제2 포지션을 정의하고; 상기 제1 ROI 좌표 및 제2 ROI 좌표는 상기 하나 이상의 HEVC 타일들을 식별하기 위해 이용된다. 이 실시예에서, 클라이언트 장치에 의한 HEVC 타일 스트림들의 요청과 HEVC 타일 스트림들의 처리 사이의 ROI의 이동(movement)으로 인한 불리한 효과들이 감소될 수 있다.

[0025] 실시예에서, 상기 요청된 비디오 데이터는 HEVC-인코딩(HEVC-encoded)되고, 상기 제2 ROI의 렌더링은: 요청된 HEVC-인코딩 비디오 데이터를 기초로 하여 디코딩된 이미지 영역(decoded image region)을 형성하는 것; 및 선택적으로(optionally), 상기 디코딩된 이미지 영역으로부터 상기 제2 ROI를 크롭핑(cropping)하는 것을 포함할 수 있다. 결정된 HEVC 타일들의 비디오 데이터는 이미지 영역 내의 제2 서브-영역을 형성하기 위하여 이용될 수 있다. ROI 좌표는 제2 서브-영역으로부터 제2 ROI를 생성(크롭핑)하기 위하여 이용될 수 있어서, 타일드 스트리밍 모드로의 스무딩한 전환(smooth transition)이 보장된다.

[0026] 실시예에서, ROI 포지션 정보를 제공하는 것은: 룩-어헤드 포지션 정보가 상기 클라이언트 장치에 제공되는 순간에 상기 클라이언트 장치에 의해 아직 수신되지 않은 상기 ROI 비디오 스트림의 적어도 하나의 비디오 프레임과 연관된 룩-어헤드 포지션 정보(look-ahead position information)를 제공하는 것을 더 포함할 수 있고, 상기 룩-어헤드 포지션 정보는 상기 전체 이미지 영역 내의 제2 ROI의 제2 포지션을 포함한다.

[0027] ROI 포지션 정보는 클라이언트 장치가 미리(ahead in time) ROI 포지션 정보를 수신할 수 있도록 타임-시프트될(time-shifted) 수 있다. 이러한 방식으로, ROI가 매우 동적인 경우(예컨대, ROI와 연관된 서브영역의, HEVC-타일드(파노라마) 비디오의 전체 이미지 영역 내의 공간적 지향(spatial orientation)이 빈번하게 변하는 경우)에도, 클라이언트는 정확한(correct) HEVC 타일 스트림들을 요청할 수 있다. 이 실시예는 동적인 ROI 궤적을 포함하는 ROI 스트림을 다룰 때 특히 유익하다.

- [0028] 타임-시프트는 클라이언트 장치에 의한 HEVC 타일 스트림들의 요청과 HEVC 타일 스트림들의 처리 사이의 시간 기간 동안 ROI의 이동(movement)을 보상한다. 이러한 방식으로, 요청된 HEVC 타일 스트림들에 의해 형성된 서브-영역은 항상 요청된 ROI를 포함한다.
- [0029] 실시예에서, 상기 록-어헤드 포지션 정보는 상기 클라이언트 장치에 의해 아직 수신되지 않은 상기 ROI 비디오 스트림 내의 하나 이상의 이미지 프레임들의 하나 이상의 타임-시프트된 제1 ROI 좌표를 포함할 수 있고, 바람직하게는 상기 타임-시프트된 제1 ROI 좌표는 상기 하나 이상의 이미지 프레임에 앞서(prior to) 상기 클라이언트 장치에 제공되고, 더욱 바람직하게는 상기 타임-시프트된 ROI 좌표는 타임-시프트를 결정하기 위한 타임-시프트 정보, 바람직하게는 타임스탬프와 추가로(further) 연관되어 있다.
- [0030] 실시예에서, 상기 ROI 포지션 정보의 적어도 일부는 상기 ROI 비디오 스트림의 비트스트림 내에서 상기 클라이언트 장치에 전송되고, 바람직하게는 SEI 메시지 또는 ROI 좌표 플래그(coordinates flag)가 상기 비트스트림 내에서 상기 ROI 좌표의 위치(location)를 정의한다.
- [0031] 실시예에서, 상기 ROI 포지션 정보의 적어도 일부는 MPEG 트랜스포트 스트림 내에서 클라이언트 장치에 전송될 수 있고, 바람직하게는 상기 MPEG 트랜스포트 스트림 내의 기초 트랜스포트 스트림 내에서 또는 상기 MPEG 트랜스포트 스트림 내의 DVB-정의(DVB-defined) SAD(Synchronised Auxiliary Data) 패킷으로서 상기 클라이언트 장치에 전송될 수 있고, 선택적으로(optionally) 상기 SAD 패킷은 타임-시프트된 ROI 좌표를 포함한다.
- [0032] 실시예에서, ROI 포지션 정보의 적어도 일부는 MPEG-4 스트림 내의 미리 정의된 박스(predefined box) 내에서 전송될 수 있고, 바람직하게는 MPEG-4 스트림의 moov, trak, 및/또는 stbl 박스 내에서 전송될 수 있다.
- [0033] 실시예에서, 상기 ROI 포지션 정보의 적어도 일부는 워터마크로서 상기 클라이언트 장치에 전송될 수 있고, 바람직하게는 상기 ROI 비디오 스트림 내의 워터마크로서 상기 클라이언트 장치에 전송될 수 있다.
- [0034] 실시예에서, 상기 타일 포지션 정보의 적어도 일부는 상기 HEVC-타일드 비디오 스트림 내에서 상기 클라이언트 장치에 제공되고, 바람직하게는 상기 HEVC-타일드 비디오 스트림의 moov 또는 moof 박스 내에서 상기 클라이언트 장치에 제공된다.
- [0035] 실시예에서, HEVC 타일들의 비디오 데이터는 레코딩 영역(recording area) 내의 별개의(separate) 비디오 트랙으로서 저장될 수 있고; 적어도 하나의 베이스 트랙은 하나 이상의 익스트랙터(extractor)를 포함하고, 익스트랙터는 비디오 트랙을 포인팅한다(point to).
- [0036] 실시예에서, 상기 타일 포지션 정보의 적어도 일부는 공간 매니페스트 파일(spatial manifest file) 내에서 상기 클라이언트 장치에 제공되고, 상기 매니페스트 파일은 상기 HEVC-타일드 비디오 스트림 및/또는 상기 ROI 비디오 스트림을 상기 클라이언트 장치에 전달하도록 구성된 하나 이상의 전송 노드(delivery node)의 위치를 찾아내기(locating) 위한 하나 이상의 스트림 식별자를 포함한다.
- [0037] 실시예에서, 상기 하나 이상의 식별된 HEVC 타일과 연관된 상기 비디오 데이터는 적어도 하나 이상의 HEVC 타일 식별자, 바람직하게는 하나 이상의 URL, 및 상기 타일 포지션 정보, 및 선택적으로(optionally), 하나 이상의 ROI 비디오 스트림과 연관된 하나 이상의 ROI 비디오 스트림 식별자를 포함하는 공간 매니페스트 파일을 기초로 하여 요청될 수 있고, 바람직하게는 상기 하나 이상의 ROI 비디오 스트림 식별자 중의 적어도 하나는 ROI 포지션 정보가 시간에 따라 변한다는 것을 나타내기 위한 동적 인디케이터(dynamic indicator)와 연관되어 있다.
- [0038] 실시예에서, 상기 HEVC-인코딩 비디오 데이터는 공간 매니페스트 파일을 기초로 하여 요청될 수 있고, 상기 공간 매니페스트 파일은 하나 이상의 HEVC 타일 표현들을 포함하고, 하나 이상의 HEVC 타일 표현들 각각은 HEVC-타일드 비디오 스트림과 연관되어 있고, HEVC 타일 표현은 하나 이상의 HEVC-타일드 스트림 및/또는 하나 이상의 HEVC 타일 스트림과 연관되어 있는 하나 이상의 HEVC 타일 식별자, 바람직하게는 하나 이상의 URL의 적어도 일부를 포함한다. 각각의 HEVC 타일 표현은 (예컨대, 각각의 표현은 상이한 줌 레벨과 관련이 있고 각각의 줌 레벨은 동일한 전체 이미지 영역을 포함하지 않는 경우에) 그 자신의 HEVC 타일드 비디오 스트림과 연관이 있을 수 있고, 및/또는 예컨대 만일 줌 레벨들 각각이 동일한 전체 이미지 영역을 포함한다면 단일한 HEVC 비디오(파일 또는 스트림)와 연관이 있을 수 있다.
- [0039] 추가적 관점에 있어서, 본 발명은 관심영역(ROI)을 렌더링하기 위한 클라이언트 장치와 관련이 있을 수 있고, 상기 클라이언트는: HEVC-타일드 비디오 스트림의 전체 이미지 영역 내에서 제1 서브-영역을 정의하는 제1 ROI를 렌더링하기 위해 ROI 비디오 스트림을 수신하도록 구성될 수 있고; 상기 ROI 비디오 스트림의 적어도 하나의 비디오 프레임과 연관된 ROI 포지션 정보를 제공하도록 구성될 수 있고; 상기 ROI 포지션 정보 및 타일 포지션

정보를 기초로 하여 상기 HEVC-타일드 비디오 스트림의 하나 이상의 HEVC 타일을 식별하도록 구성될 수 있고; 상기 HEVC-타일드 비디오 스트림의 전체 이미지 영역 내에서 제2 서브-영역을 정의하는 제2 ROI를 렌더링하기 위해, 하나 이상의 식별된 HEVC 타일과 연관된, 바람직하게는 하나 이상의 HEVC 타일 스트림들에 포함된, 비디오 데이터를 요청하도록 구성될 수 있고; 그리고, 선택적으로(optionally), 요청된 비디오 데이터 및 상기 ROI 포지션 정보를 기초로 하여 상기 제2 ROI를 렌더링하도록 구성될 수 있고, 상기 HEVC-타일드 비디오 스트림 내의 HEVC 타일의 포지션은 타일 포지션 정보에 의해 정의되고, 상기 ROI 포지션 정보는 상기 제1 ROI의 적어도 제1 포지션을 포함하고, 바람직하게는 하나 이상의 HEVC 타일은 상기 제1 ROI와 오버랩한다.

[0040] 다른 관점에서, 본 발명은 비디오 데이터를 저장하기 위한 비-일시적 컴퓨터-관독가능 저장 매체와 관련이 있을 수 있고, 상기 저장 매체는 레코딩 영역을 포함하고, 상기 레코딩 영역은: ROI 스트림의 비디오 데이터; 및 상기 전체 이미지 영역 내의 상기 서브-영역들의 포지션 정보, 바람직하게는 좌표;를 포함하고, ROI는 HEVC-타일드 비디오 스트림의 전체 이미지 영역 내의 서브-영역을 정의한다.

[0041] 실시예에서, 상기 포지션 정보는 상기 ROI 스트림의 비트스트림에 삽입될 수 있고, 바람직하게는, SEI 메시지 또는 ROI 좌표 플래그가 상기 비트스트림 내의 상기 ROI 좌표의 포지션을 정의한다.

[0042] 실시예에서, 상기 비디오 데이터는 MPEG 트랜스포트 스트림으로서 레코딩될 수 있고, 상기 ROI 포지션 정보의 적어도 일부는 상기 MPEG 트랜스포트 스트림 내에 들어 있다.

[0043] 실시예에서, 상기 ROI 포지션 정보는 기초 트랜스포트 스트림으로서 또는 DVB-정의 SAD(Synchronized Auxiliary Data)로서 상기 MPEG 트랜스포트 스트림에 들어 있을 수 있다. 추가적인 실시예에서, 상기 SAD 패킷은 하나 이상의 타임-시프트된 ROI 좌표를 포함할 수 있다.

[0044] 실시예에서, 상기 비디오 데이터는 MPEG4 스트림으로서 레코딩될 수 있고, 상기 ROI 포지션 정보의 적어도 일부는 미리 정의된 박스 내에, 바람직하게는 moov, trak, 및/또는 stbl 박스 내에서 상기 MPEG4 스트림 내에 들어 있다.

[0045] 실시예에 있어서, 상기 ROI 포지션 정보의 적어도 일부는 워터마크로서 상기 비디오 데이터 내에 삽입될 수 있다.

[0046] 추가적인 관점에서, 본 발명은 저장된 데이터 구조(data structure), 바람직하게는 공간 매니페스트 파일의 적어도 일부를 포함하는, 바람직하게는 상술한 바와 같은 클라이언트 장치에 의해 사용하기 위해 적합한 비-일시적 컴퓨터-관독가능 저장 매체와 관련이 있을 수 있고, 상기 데이터 구조는: 하나 이상의 HEVC 타일 표현(tile representation), 타일 포지션 정보(tile position information)를 포함할 수 있고, 상기 하나 이상의 HEVC 타일 표현 각각은 HEVC-타일드 비디오 스트림을 나타내고, HEVC 타일 표현은 하나 이상의 HEVC 타일을 식별시키기 위한 하나 이상의 HEVC 타일 식별자를 포함하고, 상기 타일 포지션 정보는 HEVC-타일드 비디오 스트림 내의 HEVC 타일의 포지션을 정의하고, 상기 데이터 구조는 ROI 포지션 정보를 포함하는 ROI 스트림의 적어도 하나의 ROI 스트림 식별자, 바람직하게는 URL을 더 포함할 수 있고, 상기 ROI 스트림은 상기 HEVC-타일드 비디오 스트림의 전체 이미지 영역 내에서 서브-영역을 정의하는 제1 ROI를 렌더링하기 위한 비디오 데이터를 포함한다. 대안적으로 또는 추가적으로, 상기 데이터 구조는 상기 HEVC-타일드 비디오 스트림의 전체 이미지 영역 내의 상기 서브-영역의 포지션을 정의하는 ROI 포지셔닝 정보(positioning information)를 포함할 수 있다. 이 옵션(option)은 예를 들어 ROI 스트림이 ROI 포지셔닝 정보를 포함하지 않는 경우들에 대해서 예상된다. 바람직한 실시예에서, 데이터 구조는 HTTP 적응 스트리밍 프로토콜(adaptive streaming protocol)을 기초로 하여 HEVC 타일드 스트림들로서 타일들을 리트리빙(retrieving)하도록 배열된다.

[0047] 본 발명은 또한, 컴퓨터의 메모리에서 작동할 때, 청구항들 중의 임의의 청구항에 따른 방법 단계들을 실행하도록 구성된 소프트웨어 코드 부분을 포함하는 컴퓨터 프로그램 제품에 관한 것일 수 있다.

[0048] 본 발명은 본 발명에 따른 실시예들을 개략적으로 도시할 첨부 도면들을 참조하여 더욱 설명될 것이다. 본 발명은 어떤 방식으로든 이러한 특정 실시예들에 한정되지 않는다고 이해되어야 할 것이다.

도면의 간단한 설명

[0049] 도 1은 본 발명의 실시예에 따라 관심영역(ROI) 스트리밍 프로세스를 도시한다.

도 2는 본 발명의 실시예에 따라 ROI 스트리밍 프로세스를 도시한다.

도 3은 본 발명의 실시예에 따라 ROI 스트리밍 프로세스 내에서 이용하기 위한 HEVC 인코딩 MPEG 파일의 데이터

포맷을 도시한다.

도 4는 본 발명의 실시예에 따라 ROI 스트림 및 연관된 타일드 콘텐츠를 생성하기 위한 프로세스를 개략적으로 도시한다.

도 5는 본 발명의 실시예에 따라 ROI 스트리밍 프로세스를 관리하기 위한 데이터 구조를 개략적으로 도시한다.

도 6은 본 발명의 실시예에 따라 ROI 스트리밍 프로세스를 관리하기 위한 공간 매니페스트 파일을 도시한다.

도 7은 본 발명의 실시예에 따라 ROI 스트리밍 프로세스를 관리하기 위한 공간 매니페스트 파일을 도시한다.

도 8a 및 8b는 본 발명의 실시예에 따라 ROI 스트리밍 프로세스를 관리하기 위한 공간 매니페스트 파일을 도시한다.

도 9는 본 발명의 실시예에 따라 공간 매니페스트 파일을 기초로 하여 ROI 스트림을 렌더링하도록 구성된 클라이언트 장치를 도시한다.

도 10a 및 10b는 본 발명의 실시예에 따라 ROI 스트림을 렌더링하는 프로세스의 개략도들을 도시한다.

도 11a 및 11b는 본 발명의 실시예에 따라 ROI 스트리밍 프로세스의 흐름도들을 도시한다.

도 12a 및 12b는 본 발명의 실시예에 따라 논-타일드 ROI 스트리밍 모드와 타일드 ROI 스트리밍 모드 간의 매끄러운 스위칭을 위한 프로세스의 개략도들을 도시한다.

도 13a 및 13b는 본 발명의 다른 실시예에 따라 논-타일드 ROI 스트리밍 모드와 타일드 ROI 스트리밍 모드 간의 매끄러운 스위칭을 위한 프로세스의 개략도들을 도시한다.

도 14는 본 발명의 다양한 실시예들에 따라 ROI 데이터 포맷들을 도시한다.

도 15는 본 발명의 다양한 실시예들에 따라 ROI 데이터를 MPEG 트랜스포트 스트림 내의 기초 스트림으로서 전송하기 위한 데이터 포맷을 도시한다.

도 16a 및 16b는 MPEG 스트림의 인코딩 비트스트림(encoded bitstream)으로 ROI 데이터를 전송하기 위한 데이터 포맷을 도시한다.

도 17a 및 17b는 MPEG-4 스트림의 새로운 박스(box)로 ROI 데이터를 전송하기 위한 데이터 포맷을 도시한다.

도 18은 MPEG-4 스트림의 메타데이터 트랙(metadata track)으로서 ROI 데이터를 전송하기 위한 데이터 포맷을 도시한다.

도 19a 및 19b는 본 발명의 실시예에 따라 타일드 스트리밍을 위한 공간적 버퍼(spatial buffer)의 이용을 도시한다.

도 20은 본 발명의 실시예에 따라 3개-모드 ROI 스트리밍 모델을 도시한다.

도 21은 도 1 내지 도 20을 참조하여 기술되는 시스템들 및 방법들에서 이용될 수 있는 예시적인 데이터 처리 시스템을 도시하는 블록도이다.

발명을 실시하기 위한 구체적인 내용

[0050] 도 1은 본 발명의 실시예에 따라 관심영역(ROI) 스트리밍 프로세스를 도시한다. 이 애플리케이션에서, ROI는 소스 비디오(source video)의 이미지 영역(100), 예컨대 광시야(파노라마) 이미지 영역의 서브-영역(106, 110)으로서 정의된다. 소스 비디오는: 광시야(wide field-of-view), 상이한 카메라 포지션들, 상이한 카메라 앵글들, 3D 등을 가질 수 있는 하나 이상의 고해상도 비디오 파일들(예컨대, 이들의 조합)을 기초로 하여 형성될 수 있다. 이미지 영역(100) 내의 ROI의 포지션 및 크기는 ROI 좌표(103)를 기초로 하여 정의될 수 있다. 도 1에서 도시된 바와 같이, ROI는 ROI 궤적(trajjectory)을 따를 수 있고, 제1 ROI는 제1 순간(time instance)에는 이미지 영역 내에서 제1 포지션(106)에 있고, 제2 순간에는 제2 포지션(110)에 있다.

[0051] ROI는 트래킹 알고리즘(tracking algorithm) 또는 카메라 오퍼레이터(camera operator)를 이용해서 미리 선택될 수 있다. 스트림은, 소스 비디오의 디코딩된 프레임(decoded frame)들의 이미지 영역으로부터 ROI를 크롭핑하고, 크롭핑된 영역(cropped region)들을 별개의 스트림, 예컨대 MPEG 및/또는 DVB 인코딩 스트림(MPEG and/or DVB encoded stream)으로 인코딩함으로써 생성될 수 있다. 이 스트림은 ROI 스트림(107)이라고 지칭될 수 있다. 크롭핑된 영역의 좌표는 ROI 좌표라고 지칭될 수 있다. 예를 들어, 소스 비디오는 축구 경기의 고해상

도, 광시야(파노라마) 비디오와 관련이 있을 수 있고, ROI 스트림은 소스 비디오의 디코딩된 프레임들로부터 공과 연관된 ROI를 크롭핑함으로써 소스 비디오를 기초로 하여 생성될 수 있다.

[0052] 이와 달리, ROI 스트림은, 예컨대 공을 따라가는 브로드캐스트 카메라로부터 유래하는 별개의 스트림으로서 생성될 수 있다. 이 경우에, 이미지 프로세서는 브로드캐스트 비디오의 (제한된) 시야를 소스 비디오의 광시야 이미지 영역의 특정 서브-영역(ROI)과 상관시키기(correlate) 위하여 잘 알려져 있는 이미지 인식 및 트래킹 알고리즘들을 이용할 수 있다. 이러한 방식으로, 이미지 프로세서는 소위 ROI 좌표를 ROI 스트림 내의 프레임들에 연관시킬(relate) 수 있다.

[0053] 실시예에서, ROI 스트림은 DVB 또는 멀티캐스트 기술과 같은 브로드캐스트 기술을 이용해서 전송될 수 있다. 다른 실시예에서, ROI 스트림은 HTTP 적응 스트리밍(HAS)과 같은 적응 스트리밍 프로토콜(adaptive streaming protocol)을 기초로 하여 클라이언트에 전송될 수 있다. 이러한 목적을 위하여, 매니페스트 파일(manifest file)(108)은 세그먼트들을 클라이언트에 전달하도록 구성된 하나 이상의 전송 노드의 위치를 찾아내기 (locating) 위한 세그먼트 식별자, 예컨대 URL(의 일부)를 정의할 수 있다. 적응 스트리밍 프로토콜의 예들은 Apple HTTP Live Streaming [<http://tools.ietf.org/html/draft-pantos-http-live-streaming-13>], Microsoft Smooth Streaming [<http://www.iis.net/download/SmoothStreaming>], Adobe HTTP Dynamic Streaming [<http://www.adobe.com/products/httpdynamicstreaming>], 3GPP-DASH [TS 26.247 Transparent end-to-end Packet-switched Streaming Service (PSS); Progressive Download and Dynamic Adaptive Streaming over HTTP], 및 MPEG Dynamic Adaptive Streaming over HTTP [MPEG DASH ISO/IEC 23001-6]를 포함한다. HTTP는 비디오 스트림들(및 세그먼트들)을 클라이언트들에 전달하기 위한 효율적인, 방화벽-친화적인(firewall-friendly), 그리고 확장가능한(scalable) 체계를 가능하게 한다. 실시예에서, ROI 스트림은 HEVC 코덱을 이용해서 인코딩될 수 있다.

[0054] 이 예들에서, ROI 스트림이 형성될 때 ROI 포지션 정보(112)가 결정될 수 있다. ROI 포지션 정보는 소스 비디오의 광시야 이미지 영역 내의 ROI 비디오의 (제한된 시야) 이미지 영역의 포지션들을 정의한다. 그래서, ROI 비디오의 비디오 프레임들은 소스 비디오의 비디오 프레임들의 이미지 영역 내의 서브-영역을 정의할 수 있다. ROI 포지션 정보는 소스 비디오의 이미지 영역 내의 ROI의 포지션 및 사이즈를 정의하는 ROI 좌표를 포함할 수 있다. 실시예에서, ROI 스트림 내의 각각의 프레임(또는 프레임들의 세트)에 대해, ROI 포지션 정보(ROI 좌표)가 생성될 수 있다. ROI 스트림의 플레이아웃(playlist) 동안, ROI의 포지션은 소스 비디오의 전체 이미지 영역 내에서 이동할 수 있다. ROI 스트림의 플레이아웃 동안, ROI 좌표는 시간에 따라 변할 수 있고, 이로써 ROI 좌표의 스트림을 형성한다. ROI 좌표의 스트림은 소스 비디오의 이미지 영역 내에서 공간 궤적(spatial trajectory)을 정의할 수 있다.

[0055] 실시예에서, ROI 포지션 정보는 ROI 스트림을 가지고 (동기화하여(in sync)) 클라이언트에 보내질 수 있는 ROI 좌표의 스트림을 포함할 수 있다. ROI 포지션 정보는 별개의 통신 채널(separate communication channel)로 클라이언트에 보내질 수 있다. 이와 달리, ROI 포지션 정보는 비디오 프레임의 전송 컨테이너(transport container)의 일부로서 클라이언트에 전송될 수 있다. 예들은 이하에서 더욱 상세하게 설명될 것이다.

[0056] 소스 비디오의 이미지 영역 내의 시간에 따른 ROI의 궤적이 도 1에서 도시된다. 제1 서브-영역(representing the ROI)은 제1 순간에 제1 ROI 사이즈 및 제1 ROI 좌표와 연관된 제1 ROI 포지션(106)에서 시작할 수 있고, 소정의 시간 후에 서브-영역은 제2 ROI 사이즈 및 제2 ROI 좌표와 연관된 제2 ROI 포지션(110)으로 이동할 수 있다. 궤적 동안, ROI는 소스 비디오의 이미지 영역의 미리 정해진 부분으로 줌인(zoom into)할 수 있다. 그래서, 도 1에 도시된 ROI는 소스 파일(source file)의 이미지 영역 내의 그것의 포지션 및 소스 파일의 이미지 영역에 대한 그것의 사이즈가 시간에 따라 변할 수 있다는 점에서 동적(dynamic) ROI이다.

[0057] 도 1에 도시된 스트림들, 제1 (소스) 스트림 및 제2 (ROI) 스트림은 HTTP 적응 스트리밍(HAS) 프로토콜과 같은 적응 스트리밍 프로토콜을 이용해서 플레이아웃을 위한 클라이언트에 전달될 수 있다. 이러한 경우에, 제1 (ROI) 비디오 스트림과 제2 (광시야) 비디오 스트림 간의 공간적 관계가 공간 데이터 구조(spatial data structure) 내에서 기술될 수 있으며, 이것은 이하에서 공간 매니페스트 파일(spatial manifest file: SMF)이라고 지칭될 것이다. SMF는 클라이언트가 스트림을 클라이언트에 전송할 수 있는 하나 이상의 전송 노드(예컨대 미디어 서버들)의 포지션을 찾아내고 액세스하기 위하여 이용할 수 있는 스트림 식별자(예컨대 URL(의 일부))를 포함할 수 있다. 스트림 식별자 외에도, 매니페스트 파일은 기준 시스템(reference system) 내에서 비디오 스트림들의 이미지 영역들의 공간적 관계를 기술하는 포지션 정보를 포함할 수 있다.

[0058] 전형적으로, 클라이언트 장치들은 제한된 렌더링 및 대역폭 능력(rendering and bandwidth capabilities)을 가

지고, 그래서 소스 비디오의 전체 이미지 영역을 고해상도 품질로 수신 및 렌더링하지 못할 수 있다. 이 문제의 적어도 일부는 소위 타일드 스트리밍 기법을 이용함으로써 해결될 수 있고, 이 기법에서 소스 스트림(전형적으로, 파노라마 비디오 스트림)은 한 세트의 타일 스트림들로 공간적으로(spatially) 나뉜다. 이후, 타일 스트림들은 전형적으로 개별적으로 인코딩되고, 소스 비디오의 단일한 인코딩 프로세스(single encoding process)로부터 얻어지지 않는다. 클라이언트는 파노라마 비디오의 타일 스트림들의 서브세트(subset)를 요청함으로써 소스 스트림의 임의의 부분을 재구성할(reconstruct) 수 있다. 선행기술의 인코딩 체계에서 이 타일 스트림들은 독립적으로 인코딩되기 때문에, 이러한 (인코딩된) 타일 스트림들의 수신시에 클라이언트는 비디오 데이터가 사용자에게 디스플레이될 수 있기 전에 복수의 디코더 인스턴스(instance)들을 이용해서 타일 스트림들을 디코딩해서 타일 스트림들의 디코딩된 비디오 데이터를 함께 스티칭할 필요가 있다.

- [0059] 하지만, 하나의 클라이언트를 위한 복수의 디코더 인스턴스들의 이용은 다량의 프로세서 자원들을 희생시킨다. 특히 다량의 타일들을 이용할 때, 클라이언트 프로세스는 매우 복잡하게 되고, 극심하게 리소스가 많이 소모(resource intensive)되게 된다. 이 문제는 HEVC 비디오 압축 표준을 이용함으로써 해결될 수 있다.
- [0060] 도 2는 본 발명의 실시예에 따라 ROI 스트리밍 프로세스를 도시하고, 소스 파일(비디오)은 HEVC-타일드 스트림으로서 인코딩될 수 있다.
- [0061] HEVC에 있어서, 비디오 프레임은 소위 CTU(coding tree unit)들로 파티셔닝되는데(partitioned), 이것은 디코딩 프로세스를 위해 HEVC 표준 내에서 이용되는 기본 처리 단위이다. 게다가, HEVC 인코더(encoder)는 HEVC 스트림 내의 비디오 프레임(200)을 소위 HEVC 타일들(201)로 분할하도록 구성될 수 있고, HEVC 타일은 특정한 그룹의 CTU들에 의해서 정의되는 직사각형 영역이다. HEVC 타일들은 예컨대 파노라마 비디오의 이미지 영역을 복수의 인접한 사각형 영역들(이들은 상이한 사이즈로 이루어질 수 있음)로 분할할 수 있고, HEVC 타일들의 경계들은 HEVC 타일 경계들에 의해서 칼럼(column) 및 로우(row) 방향으로 정의된다. 이 애플리케이션에서, 하나 이상의 HEVC 타일들을 포함하는 HEVC 인코딩 비디오 스트림은 HEVC-타일드 스트림이라고 지칭될 수 있다.
- [0062] HEVC 타일들은 (멀티-코어 프로세서들을 이용해서) 비디오 데이터의 병렬 인코딩 및 디코딩을 위한 HEVC 코덱에서 유입되고, 그래서 비디오 데이터의 효율적인 처리가 달성될 수 있다. 이러한 방식으로, HEVC-타일드 비디오는 단일한 HEVC 디코더를 이용해서 효율적으로 디코딩될 수 있다. 게다가, HEVC 타일들은 또한 HEVC-타일드 스트림에 들어 있는 HEVC 타일들의 서브세트만의 플레이아웃을 위해 이용될 수 있다.
- [0063] 이러한 목적을 위하여, HEVC 표준은 인코딩 및 디코딩 프로세스의 전체 독립성을 보장하기 위하여 각각의 HEVC 타일과 연관된 공간적 및 시간적 제약사항들(constraints)의 시그널링(signalling)을 허용한다. 특히, HEVC 표준은 HEVC 타일의 경계들 내에서 비디오 코딩(예컨대 모션(motion) 벡터들 및 인-루프(in-loop) 필터들) 내의 잘 알려진 공간적 및 시간적 예측(spatial and temporal prediction)들을 제한하는 것을 허용한다. 이러한 공간적 및 시간적 제약사항들을 이용해서, HEVC 인코더는 독립적으로 디코딩가능한(decodable) HEVC-타일들을 생성하도록 구성될 수 있고, 이들은 스트리밍 클라이언트에 의해서 개별적으로 저장 및 액세스될 수 있다.
- [0064] HEVC-타일드 스트림 내의 HEVC 타일들의 포지션은 타일 포지션 정보(205)에 의해서 결정될 수 있고, 타일 포지션 정보(205)는 좌표계(203)를 기초로 하여 정의될 수 있다. 그래서 HEVC 타일의 포지션은 HEVC-타일드 스트림의 전체 이미지 영역(area) 내의 타일의 이미지 영역(서브영역)의 공간적 위치 또는 지향(orientation)을 지칭한다.
- [0065] HEVC 스트림의 이미지 영역 내의 하나 이상의 이동하는 ROI들(206, 210)의 위치는 도 1을 참조하여 설명된 것과 유사한 방식으로 ROI 포지션 정보(212)에 의해 정의될 수 있다. ROI 포지션 정보는 시간에 따라 변할 수 있고, 이로 인하여 소스 파일의 이미지 영역 내에서 이동하는 ROI의 궤적을 정의할 수 있다. 타일 포지션 정보를 정의하기 위해 이용되는 좌표계는 또한 도 1을 참조하여 기술된 바와 같이 ROI 스트림에서의 ROI의 ROI 좌표를 결정하기 위하여 이용될 수 있다. 이미지 영역이 2D 또는 3D 이미지 영역과 관련이 있는 경우에, 데카르트(Cartesian) 좌표계가 이용될 수 있다. 이와 달리, 이미지 영역들이 굽은(curved) 이미지 영역과 관련이 있는 경우에, 다른 논-데카르트(non-Cartesian), 곡선(curvilinear) 좌표계들, 예컨대, 원통형, 구형, 또는 극(polar) 좌표계가 이용될 수 있다.
- [0066] 실시예에서, HEVC 인코더는, HEVC 타일이 HEVC 디코더에 의해서 독립적으로 저장되고 독립적으로 디코딩될 수 있도록 하기 위하여, HEVC-타일드 비디오 스트림과 연관된 HEVC 타일들(202)을 생성하도록 구성될 수 있다. 예를 들어, 도 3은 본 발명의 실시예에 따라 ROI 스트리밍 프로세스에서 이용될 수 있는 HEVC-타일드 MPEG 파일 또는 스트림(302)을 위한 데이터 포맷(300)의 예를 도시한다. 실시예에서, (HEVC 타일 스트림에 포함된 바와 같

이) HEVC 타일과 연관된 독립적으로 디코딩가능한 비디오 데이터는 트랙(304-312)으로서 저장될 수 있다. 특히, HEVC 타일과 연관된 비디오 데이터(314)는 소위 (HEVC) 타일 트랙(306-312)으로서 저장될 수 있다.

- [0067] 타일 트랙을 위한 비디오 데이터를 생성할 때, 비디오 코딩(예컨대, 모션 벡터 및 인-루프 필터들)을 위한 공간적 및 시간적 예측들은 하나의 HEVC 타일의 경계들 내에서 완전히 유지된다. 이러한 방식으로, 비디오 프레임 내의 하나의 HEVC 타일의 비디오 데이터와 비디오 프레임 내의 다른 HEVC 타일들 사이에 아무런 공간적 디코딩 의존성(spatial decoding dependency)이 존재하지 않는다. 게다가, 비디오 프레임 내의 하나의 HEVC 타일의 비디오 데이터와 이전의 비디오 프레임들 내의 다른 HEVC 타일들 사이에 아무런 공간적 디코딩 의존성이 존재하지 않는다.
- [0068] 실시예에서, 타일 트랙은 타일 포지션 정보(316)를 더 포함할 수 있다. 디코더는 디코딩을 위한 HEVC 스트림으로부터 어느 트랙을 추출할 필요가 있는지를 결정하기 위하여 타일 포지션 정보를 이용할 수 있다. 실시예에서, 트랙 내의 타일 포지션 정보는 도 1 및 도 2를 참조하여 기술된 바와 같은 좌표계에 의해 정의된 기준 공간(reference space)에 타일을 배치하기 위하여 타일 기원(tile origin) 및 타일 사이즈 정보(예컨대 폭 및 높이 파라미터들)를 포함할 수 있다.
- [0069] 실시예에서, 도 3의 데이터 포맷(300)은 소위 베이스 트랙(base track)(304)을 포함할 수 있다. 베이스 트랙은 익스트랙터(extractor)들(318)을 포함할 수 있고, 익스트랙터는 하나 이상의 상응하는 타일 트랙들에 대한 참조(reference)를 정의한다. 베이스 트랙을 파싱(parsing)함으로써, 디코더는 익스트랙터를 그것이 참조하는 타일 트랙의 오디오 및/또는 비디오 데이터로 대체할 수 있다. 만일 특정 비디오 애플리케이션이 특정 타일을 필요로 하지 않는다면, 디코더는 단순히 그것의 상응하는 익스트랙터를 무시할 수 있다. 만일 특정 트랙의 비디오 데이터가 수신되지 않거나 리트리빙되지(retrieved) 않는다면, 이러한 트랙의 부존재는 HEVC 디코더에 의해 "유실된 데이터(missing data)"로 해석될 수 있다. HEVC 타일은 다른 HEVC 타일들과는 독립적으로 디코딩될 수 있어서, 하나 이상의 트랙으로부터의 데이터의 부존재("유실된 데이터")는 리트리빙될 수 있는 다른 트랙들을 디코더가 디코딩하는 것에 지장을 주지 않는다.
- [0070] 실시예에서, 도 3의 데이터 포맷은 HEVC-타일드 스트림의 전체 이미지 영역 내의 ROI 스트림의 (제한된 시야) 이미지 영역의 ROI 포지션 정보, 예컨대 ROI 좌표를 포함하는 ROI 트랙을 더 포함할 수 있다. ROI 포지션 정보는 ROI 좌표를 포함할 수 있고, 이것은 HEVC-타일드 스트림의 전체 이미지 영역과 연관된 좌표계(103)를 기초로 하여 정의될 수 있다. 추가적인 실시예에서, ROI 트랙은 ROI의 비디오 데이터를 포함할 수 있다.
- [0071] 도 3에 도시된 데이터 포맷은 또한 HEVC-타일드 스트림의 HEVC 타일들을 독립적인 HEVC 타일 스트림들로서 저장하기 위하여 이용될 수 있고, 그래서 클라이언트 장치 내의 클라이언트는 전송 노드(예컨대 미디어 서버)로부터 HEVC 타일 스트림을 개별적으로 요청하기 위하여 SMF를 이용할 수 있다. 이러한 목적을 위하여, SMF는 하나 이상의 타일 식별자(예컨대, URL 또는 URI의 적어도 일부)를 포함할 수 있다. 타일 식별자는, HEVC-타일드 비디오 데이터를 저장 및 전달하도록 구성된 하나 이상의 전송 노드(예컨대, 하나 이상의 미디어 서버 또는 CDN)의 위치를 찾아내기 위해 클라이언트에 의하여 이용될 수 있다. 하나의 HEVC 타일을 포함하는 HEVC-타일드 스트림은 HEVC 타일 스트림이라고 지칭될 수 있다.
- [0072] 도 2에 도시된 HEVC 타일 스트림들 및 ROI 스트림은 HTTP 적응 스트리밍(HAS) 프로토콜과 같은 적응 스트리밍 프로토콜을 이용해서 플레이아웃을 위한 클라이언트 장치에 전달될 수 있다. 이러한 목적을 위하여, HEVC 타일 스트림은 미리 결정된 기간(period)의 시간적 세그먼트(temporal segment)들로 시간적으로 분할될 수 있다. 그래서, 이러한 경우에, SMF는 HEVC-타일드 스트림과 연관된 시간적 세그먼트 식별자들을 더 포함할 수 있고, 그래서 클라이언트 장치는 네트워크 내의 하나 이상의 전송 노드로부터 HEVC-타일드 스트림의 시간적 세그먼트들을 순차적으로 요청할 수 있다. 제1 (ROI) 비디오 스트림과 하나 이상의 HEVC-타일드 비디오 스트림 간의 시간적 및 공간적 관계는 공간 매니페스트 파일(SMF)(204)에서 정의될 수 있다. MPEG-DASH 스트리밍 프로토콜의 경우에, 매니페스트 파일은 MPD(media presentation description)라고 지칭될 수 있다.
- [0073] SMF는 하나 이상의 HEVC 스트림 식별자, 하나 이상의 ROI 스트림 식별자 및/또는 하나 이상의 HEVC 타일 식별자를 포함할 수 있다. SMF는 특정 좌표계 내에서 하나 이상의 ROI 스트림과 하나 이상의 HEVC 스트림 간의 공간적 관계를 기술하는 포지션 정보를 더 포함할 수 있다. 다른 실시예에서, 포지션 정보는 좌표계 내에서 하나 이상의 HEVC 타일들 간의 공간적 관계를 기술하는 정보를 더 포함할 수 있다. HEVC 타일들에 관한 공간적 정보는 타일 포지션 정보라고 지칭될 수 있다. 본 발명의 다양한 실시예들에 따른 공간 매니페스트 파일의 더욱 상세한 예들은 이하에서 도 5-8을 참조하여 설명된다.

- [0074] 하나의 소스 파일의 (예컨대 타일 사이즈, 2D 및 3D 포맷들, 상이한 비디오 및/또는 오디오 품질들(예컨대 SD/HD/UHD, 비트레이트들 등), 상이한 해상도들, 상이한 시야들, 카메라 앵글들 등을 나타내는) 상이한 HEVC 타일 표현들 또는 - 줄여서 - 타일 표현(tile representation)들이 생성될 수 있다. 각각의 타일 표현에 대해서, 소스 파일의 비디오 프레임들은 미리 결정된 수의 (바람직하게는, 독립적으로) 디코딩가능한 HEVC 타일들을 포함하는 HEVC-타일드 비디오 파일(HEVC-타일드 비디오 파일은 클라이언트 장치로 스트리밍될 수 있는 HEVC 인코딩 비디오 데이터를 본질적으로 포함하기 때문에 HEVC-타일드 비디오 스트림이라고도 지칭됨)로 인코딩될 수 있고, 각각의 HEVC 타일(201)은 HEVC-타일드 비디오 파일의 전체 이미지 영역의 서브영역을 나타낼 수 있다. HEVC 인코더는 소스 파일의 비디오 프레임들을 기초로 하여 HEVC 타일들을 포함하는 HEVC-인코딩 MPEG 스트림을 생성하기 위하여 이용될 수 있다.
- [0075] (HEVC 타일들의 좌표를 포함하는) 타일 포지션 정보 및 (ROI 스트림의 ROI 좌표를 포함하는) ROI 포지션 정보를 기초로 하여, 클라이언트 장치는 제1 서브-영역(ROI)을 포함하는 서브영역을 구성하기 위하여 클라이언트 장치에 의해서 이용될 수 있는 HEVC 타일 스트림 또는 복수의 HEVC 타일 스트림들을 요청할 수 있다. 특히, 바람직하게는 ROI 스트림의 ROI를 포함하는 서브-영역을 구성하기 위하여 디코더에 의해서 HEVC 타일들을 공간적으로 인접하게 컷(spatially adjoining)이 이용될 수 있다.
- [0076] 예를 들어, 도 2에서 도시된 바와 같이, 제1 순간에, ROI 좌표는 4개의 HEVC 타일들(이 경우에, 포지션들 (1,0), (1,1), (2,0), 및(2,1)에서의 HEVC 타일들)과 오버랩(overlap)하는 제1 ROI(206)를 정의한다. 이 4개의 HEVC-타일들은 제1 순간에 ROI를 포함하는 HEVC-타일드 파노라마 스트림의 이미지 영역 내의 제2 서브-영역을 형성할 수 있다. 유사하게, 제2 순간에, ROI 좌표는 2개의 HEVC 타일들(이 경우에, 포지션들 (0,2) 및 (1,2)에서의 HEVC 타일들)과 오버랩하는 제2 ROI(210)를 정의할 수 있다. 이 2개의 HEVC 타일들은 제2 순간에 ROI를 포함하는 HEVC-타일드 파노라마 스트림 내에 제2 서브-영역을 형성할 수 있다. 이 체계(scheme)는 ROI 스트림으로부터 타일드 스트리밍으로 스무딩하게(smoothly) 스위칭하기 위하여 이용될 수 있다.
- [0077] 이하에서 더욱 상세하게 기술되는 바와 같이, 사용자는 사용자 인터페이스(예컨대, 터치 스크린 또는 포인팅(pointing) 장치)를 이용하여 디스플레이된 이미지 영역과 상호작용할 수 있고, (예컨대, 패닝(panning), 줌링(zooming), 틸팅(tilting) 조작들을 기초로 하여) 디스플레이된 콘텐츠를 조작할 수 있다. 이러한 사용자 상호작용에 대한 반응으로, 인터페이스는 클라이언트 장치가 HEVC-타일드 스트림을 기초로 하여 사용자-요청(user-requested) ROI의 렌더링을 시작하도록 하는 인스트럭션을 생성할 수 있다. 그래서, 사용자는 ROI를 이동시키거나 및/또는 확대시킬 수 있고, - 응답으로 - ROI가 위치해 있는 인접한 HEVC 타일들의 영역을 형성하기 위하여, 적절한 타일 표현, 그 타일 표현 내의 HEVC-타일드 스트림, 및 HEVC-타일드 스트림의 하나 이상의 HEVC 타일들이 선택될 수 있다. 만일 타일들과 연관된 비디오 데이터가 이미 클라이언트 장치에 존재한다면, ROI를 렌더링하기 위해서 비디오 데이터가 이용될 수 있기 전에 타일들이 HEVC 타일 스트림들로서 클라이언트 장치에 스트리밍될 필요가 없다. 만일 이 데이터가 존재하지 않는다면, 이들은 ROI가 렌더링되고 있기 전에 하나 이상의 HEVC 타일 스트림들로서 우선 스트리밍된다. 사용자-선택(user-selected) ROI를 렌더링하는 이 모드는 타일드 스트리밍 모드라고 지칭될 수 있다.
- [0078] 그래서, 이상으로부터, 이하의 결론이 된다: 1) 사용자는 (도 1을 참조하여 설명된 바와 같이) ROI 스트림을 기초로 하여 제1 모드((논-타일드) ROI 스트리밍 모드)에서 특정 장면(scene)을 감상 수 있고, ROI 스트림은 예컨대, MPEG 스트림으로서 클라이언트 장치에 전달될 수 있다; 그리고, 2) 사용자는 (도 2를 참조하여 설명된 바와 같이) HEVC-타일드 스트림을 기초로 하여 제2 모드(타일드 스트리밍 모드)에서 동일한 장면, 장면의 디테일(detail)들, 또는 장면 주변의 영역들을 감상할 수 있다. 도 1 및 도 2에서 도시된 바와 같이, ROI 스트리밍 모드로부터 타일드 스트리밍 모드로의 전환 (및 타일드 스트리밍 모드로부터 ROI 스트리밍 모드로의 전환)은 ROI 스트림과 연관된 ROI 포지션 정보, 즉, ROI 좌표를 기초로 하여 실현된다.
- [0079] ROI 스트림은 디폴트 스트리밍 모드(default streaming mode), (논-타일드) ROI 스트리밍 모드에서 클라이언트에 의해 이용될 수 있는데, 브로드캐스트, 멀티캐스트, 또는 유니캐스트(unicast) ROI 스트림은 클라이언트 장치에 의해 렌더링된다. HEVC-타일드 스트림은 클라이언트 장치의 UI(user interface)를 이용해서 사용자에게 의해 조작될 수 있는 ROI를 렌더링하기 위하여 클라이언트에 의해 이용될 수 있다. 클라이언트가 ROI 스트리밍 모드에서 ROI 스트림을 렌더링하고 있을 때, 클라이언트는 ROI 포지션 정보를 기초로 하여 소스 비디오의 전체 이미지 영역 내의 ROI의 포지션을 알 것이다. ROI 스트림과 연관된 ROI 포지션 정보는 타일드 스트리밍 모드로 스무딩하게 또는 매끈하게 스위칭하기 위하여 클라이언트에 의해 이용될 수 있다.
- [0080] 도 4는 본 발명의 실시예에 따라서 ROI 스트림 및 연관된 HEVC-타일드 스트림들을 생성하기 위한 프로세스를 개

략적으로 도시한다. 이 예에서, 하나 이상의 카메라(402), 예컨대 하나 이상의 고해상도, 광시야 카메라는 소스 비디오, 전형적으로 파노라마 비디오를 생성 또는 구성하기 위하여 이용될 수 있다. HEVC 인코더(404)는 소스 비디오를 기초로 하여 하나 이상의 HEVC-타일드 스트림들(408)을 생성하기 위하여 이용될 수 있다. HEVC-타일드 스트림들은 소스 비디오의 상이한 타일 표현들(409)을 정의할 수 있고, HEVC-타일드 스트림들, HEVC-타일드 스트림들 각각에서의 HEVC 타일들, 및 몇몇 실시예들에서 HEVC-타일드 스트림들 내의 세그먼트들의 시간적 관계(temporal relation)에 관한 정보가 SMF(410)에서 정의될 수 있다.

[0081] 소스 비디오는 ROI 스트림 생성기(412)에도 전달될 수 있다. ROI 스트림 생성기는, 소스 프레임들의 프레임들 중의 프레임들의 이미지 영역을 통해서 패닝(panning) 및 주밍(zooming)함으로써 ROI(예컨대, 특정한 활동(activity)을 가진 영역)을 선택하도록 구성되고, 그리고 비디오 프레임들로부터 ROI를 크롭핑하고 크롭핑된 이미지 영역들을 기초로 하여 새로운 비디오 스트림(ROI 스트림(414))을 만들어냄으로써 ROI 스트림을 생성하도록 구성된다. ROI는 알고리즘, 예컨대 이미지들 내의 특정 객체(object)를 추적하기 위한 트래킹 알고리즘, 또는 ROI를 선택하는 인간 오퍼레이터(human operator)를 이용해서 자동으로 선택될 수 있다. ROI 스트림의 생성 동안, ROI의 좌표가 ROI 포지션 정보(418)로서 수집될 수 있다. ROI 포지션 정보는 ROI 스트림 내의 메타데이터로서 또는 별개의 데이터 파일로서 클라이언트에 보내질 수 있다. 만일 ROI 포지션 정보가 별개의 스트림에서 보내진다면, ROI 포지션 정보는 ROI 포지션 정보, 예컨대 ROI 좌표를 ROI 스트림 내의 프레임들에 링크시키기(linking) 위한 정보(예컨대 프레임 번호들)를 포함할 수 있다. 추가적인 실시예에서, ROI 스트림은 적응 스트리밍(adaptive streaming)을 위해 시간적으로 세그먼트화된 스트림(temporally segmented stream)으로 구성될 수 있다. 이 경우에, ROI 스트림의 세그먼트들은 ROI 매니페스트 파일(416)에서 정의될 수 있다.

[0082] 실시예에서, ROI 스트림은 소스 파일을 기초로 하여 (예컨대 크롭핑함으로써) 형성되지 않고, 별개의 브로드캐스트 카메라의 직접적인 출력, 즉 도 1을 참조하여 기술된 바와 같이 소스 비디오에 들어 있는 것과 동일한 장면의 일부를 캡처하는(capture) 브로드캐스트 스트림이다.

[0083] 결합기(combiner)(420)는 ROI 스트림이 전송 네트워크(delivery network)(422)에 보내지기 전에 ROI 스트림을 하나 이상의 HEVC-타일드 스트림들(408)과 결합시킬 수 있다. ROI 스트림을 HEVC-타일드 스트림과 결합시키는 것은 HEVC-타일드 스트림들 및 ROI 스트림에 대한 정보를 포함하는 SMF를 형성하는 것을 포함할 수 있다. 실시예에서, 결합기는 ROI 스트림 식별자(예컨대, URL 또는 URI의 적어도 일부) 및 몇몇 실시예들에서 ROI 스트림의 시간적 세그먼트들과 연관된 매니페스트 파일을 SMF에 삽입할 수 있다.

[0084] 하나 이상의 HEVC-타일드 스트림들(선택적으로(optionally), 독립적인 HEVC 타일 스트림들의 집합으로서) 및 ROI 스트림들은 네트워크 내의 하나 이상의 전송 노드들(426_{1,2})에 저장될 수 있고, 전송 노드는 하나 이상의 HEVC-타일드 스트림들 및/또는 ROI 스트림들의 적어도 일부(예컨대 하나 이상의 HEVC 타일 스트림들)를 (모바일) 클라이언트 장치 내의 클라이언트(428)에 전달하도록 구성된다.

[0085] 실시예에서, 전송 노드는 미디어 서버일 수 있다. 다른 실시예에서, 전송 노드들(때때로 서rogate 노드(surrogate node)라고도 지칭됨)의 적어도 일부는 전용 콘텐츠 전송 네트워크(content delivery network: CDN)의 일부일 수 있다. 이 경우에, HEVC-타일드 스트림들 및 ROI 스트림들은 CDNCF(Content Delivery Network Control Function)(424)에 의해 입수될(ingested) 수 있다. 이후, CDNCF는 스트림들의 효율적인 배포(distribution)가 보장되도록 하기 위하여 상이한 전송 노드들을 통해서 HEVC-타일드 스트림들 및 ROI 스트림들을 배포한다. 실시예에서, CDN은 클라이언트가 (타일드) 콘텐츠의 전달을 요청하기 위하여 CDN의 전송 노드들에 효율적으로 액세스할 수 있도록 타일(및 세그먼트) 식별자들(URL들)을 업데이트할 수 있다.

[0086] 클라이언트 장치의 클라이언트(428)가 ROI 및/또는 HEVC-타일드 스트림에 액세스하길 원하는 경우에, 콘텐츠 제공자 또는 CDN로부터 SMF를 수신할 수 있고, ROI 스트림 및/또는 HEVC-타일드 스트림의 하나 이상의 연관된 HEVC 타일들을 요청 및 플레이아웃하기 위하여 SMF를 이용할 수 있다. 클라이언트 장치는 일반적으로 전자 태블릿(electronic tablet)과 같은 (모바일) 콘텐츠 플레이-아웃 장치, 스마트-폰(smart-phone), 노트북(notebook), 미디어 플레이어(미디어 플레이어), 홈 게이트웨이(home gateway) 또는 DASH-이네이블(DASH-enabled) HbbTV 디스플레이 장치와 같은 DASH 이네이블 장치와 관련이 있을 수 있다. 이와 달리, 클라이언트 장치는 저장된 콘텐츠에 대한 접근권(access)을 가진 콘텐츠 플레이-아웃 장치에 의한 장래의 소비를 위한 콘텐츠를 처리하고 임시로 저장하도록 구성된 셋톱 박스 또는 콘텐츠 저장 장치일 수 있다.

[0087] 도 5는 본 발명의 하나의 실시예에 따라서 ROI 스트리밍 프로세스를 관리하기 위한 데이터 구조(500)를 개략적으로 도시한다. 특히, 도 5는 공간 매니페스트 파일(patial manifest file: SMF)을 도시하고, 이것은 여러 개의

계층적 데이터 레벨(hierarchical data level)들(502, 508, 518_{1,2})을 포함할 수 있다. 제1 데이터 레벨(502)은 소스 비디오(예컨대, source1.mp4)의 하나 이상의 타일 표현(tile representation)들(506₁₋₃)을 정의하는 타일 구성(tile composition) 정보와 관련이 있을 수 있다. 전형적으로, 소스 비디오는 하나 이상의 고해상도 및 중 종 광시야 HD 또는 심지어 UHD 비디오 스트림들 또는 파일들을 기초로 하여 형성될 수 있다.

[0088] 다음 데이터 레벨(508)은 타일 표현들(506_{1,3})을 정의할 수 있다. 타일 표현(508)은 소정의 품질의 HEVC 타일들의 세트(511)를 정의할 수 있다. HEVC 타일 인스턴스(512₁₋₄)는 HEVC 타일을 정의할 수 있다. 실시예에서, HEVC 타일들의 세트는 HEVC-타일드 스트림 내의 HEVC 타일 트랙들로서 정의될 수 있다. 본 발명의 실시예에서, HEVC 타일들의 세트는 개별적으로 저장가능한(storable) 및 액세스가능한(accessible) HEVC 타일 스트림들의 세트로서 정의될 수 있다.

[0089] 타일 포지션 정보는 소정의 타일 표현의 HEVC-타일드 스트림 내의 HEVC 타일들의 포지션을 정의할 수 있다. 예를 들어, 2개의 HEVC 타일들(타일 포지션들(1,1) 및 (1,2))은 소스 비디오(연관된 HEVC-타일드 비디오) 내의 이미지 영역의 상부 절반(top half)과 연관된 콘텐츠를 포함한다.

[0090] 타일 표현은 공간 맵(spatial map)에서 참조되는 HEVC 타일들을 생성하기 위해 이용되는, 소스 비디오의 해상도 버전을 나타내는 소스 해상도(510)에 대한 정보를 더 포함할 수 있다. 예를 들어, 도 5에서, 공간 맵에서 참조되는 HEVC 타일들은 소스 비디오의 4096 × 2160 해상도 버전을 기초로 하여 생성될 수 있다.

[0091] 실시예에서, 하나의 HEVC 타일과 연관된 데이터는 타일 정보(518_{1,2}) 즉, 개별적인 HEVC 타일들에 관한 정보로서 구성될 수 있다. 타일 정보는 포지션 정보(520_{1,2}), 예컨대 HEVC 타일의 타일 좌표를 포함할 수 있다. 타일 좌표는 절대(absolute) 또는 상대(relative) 좌표계를 기초로 할 수 있고, HEVC 타일들을, 디스플레이를 위한 매끄러운(seamless) 비디오 이미지로 공간적으로 배치하기 위하여 HEVC 디코더에 의해 이용된다.

[0092] 타일 정보는 하나 이상의 타일 식별자들(522_{1,2})(예컨대, URI 또는 URL의 일부)을 더 포함할 수 있다. 하나의 실시예에서, 타일 정보는 HEVC 타일 스트림 내의 HEVC 타일의 비디오 데이터를 클라이언트 장치에 전송하도록 구성된 하나 이상의 전송 노드의 위치를 찾아내기 위해 이용될 수 있고, 및/또는 하나 이상의 특정 파일들, 바람직하게는 전송 노드들 중의 하나에서 스트리밍가능한(streamable) 파일들의 위치를 찾아내기 위해 이용될 수 있다.

[0093] 실시예에서, SMF는 하나 이상의 ROI 스트림의 데이터 구조(505)(예컨대, 매니페스트 파일 또는 매니페스트 파일에 대한 참조(reference))를 더 포함할 수 있고, ROI 스트림은 HEVC 타일들과 동일한 소스 비디오를 기초로 하여 생성될 수 있다. 실시예에서, ROI 스트림은 ROI 포지션 정보를 포함할 수 있다. 다른 실시예에서, SMF는 ROI(507)의 포지션 정보를 포함하는 스트림 또는 파일에 대한 참조(reference)(예컨대, URL)를 포함할 수 있다.

[0094] 도 6은 본 발명의 실시예에 따라서 ROI 스트리밍 프로세스를 관리하기 위한 공간 매니페스트 파일(600)을 도시한다. 특히, 도 6은 식별자, 예컨대 URL 또는 URI(의 일부)와 연관이 있는 상이한 MPD 비디오 엘리먼트(element)들을 정의하는 MPEG-DASH MPD의 예를 도시한다. (클라이언트 장치상에 상주하는) DASH 클라이언트는 MPD 비디오 엘리먼트들과 연관된 비디오 데이터에 액세스해서 리트리빙하기 위하여 식별자를 이용할 수 있다. 예를 들어, 이 예에서, 제1 MPD 비디오 엘리먼트(602)는 적어도 하나의 파노라마 비디오(URI "full_panorama.mp4"(606)에 의해 정의된 광시야 비디오)와 연관이 있을 수 있다. 제2 MPD 엘리먼트(604)는 ROI 비디오(URI "zoomed_part.mp4"(608)에 의해 정의된 ROI 비디오)와 연관이 있을 수 있고, 여기서 MPD 비디오 엘리먼트들 간의 공간적 관계의 적어도 일부는 이하에서 더욱 상세하게 설명될 포지션 정보를 기초로 하여 정의된다.

[0095] MPD 비디오 엘리먼트는 하나 이상의 표현들(차이가 하나 이상의 인코딩 파라미터들에 의해 정의될 수 있는 동일 또는 연관된 콘텐츠의 상이한 버전들)을 포함하는 "AdaptationSet" 속성으로서 정의될 수 있다.

[0096] DASH 클라이언트는 MPD 비디오 엘리먼트와 연관된 비디오 데이터를 네트워크로부터 요청하기 위하여 MPD 내의 정보를 이용할 수 있다. 나아가, DASH 클라이언트는 비디오 데이터가 수신되자마자 HEVC 인코딩 스트림의 비디오 데이터의 디코딩을 시작할 수 있도록 HEVC 디코더를 구성하기 위하여 MPD 내의 정보(메타데이터)를 이용할 수 있다. HEVC 디코더를 구성하기 위한 정보(메타데이터)는 MPD 비디오 엘리먼트들 간의 공간적 관계, 이 경우에 있어서는, HEVC 타일드 파노라마 비디오 스트림과 ROI 스트림 간의 공간적 관계를 포함할 수 있다. 이러한 목적을 위하여, MPD 오써(author)는 MPD에 포지션 정보를 포함시킬 수 있다. 포지션 정보는 하나 이상의 공간적

관계 디스크립터(spatial relationship descriptor: SRD)들(610, 612)에 의해서 정의될 수 있다. SRD는 MPD 비디오 엘리먼트들 사이에 공간적 관계가 존재한다는 것을 디코더, 특히 HEVC 디코더에게 알리기 위하여 EssentialProperty 속성(디스크립터를 처리할 때 클라이언트에 의해 이해될 필요가 있는 정보) 또는 SupplementalProperty 속성(디스크립터를 처리할 때 클라이언트에 의해 폐기될 수 있는 정보)에서 이용될 수 있다. 실시예에서, 공간적 관계 디스크립터(schemeldUri "urn:mpeg:dash:srd:2014")가 이용될 수 있다.

- [0097] 실시예에서, 포지션 정보는 다음을 포함하되 이에 한정되지 않는 복수의 파라미터들을 포함할 수 있는 @value 속성(614,616)을 기초로 하여 정의될 수 있다:
- [0098] - source_id 파라미터(618,619)는 서로 공간적 관계를 갖는 MPD 엘리먼트들(AdaptationSet 또는 SubRepresentation)의 세트를 정의할 수 있다.
- [0099] - 포지션 파라미터들(620,621) x,y,w,h는 MPD 비디오 엘리먼트의 포지션을 정의할 수 있는데, 좌표 x,y는 MPD 비디오 엘리먼트의 이미지 영역의 원점을 정의하고, w 및 h는 이미지 영역의 폭 및 높이를 정의한다. 포지션 파라미터들은 주어진 임의의 단위, 예컨대 픽셀 단위로 표현될 수 있다.
- [0100] - 튜플(tuple) W 및 H(622,623)는 x,y,w, 및 h와 동일한 임의의 단위로 표현된 기준 공간의 차원을 정의한다.
- [0101] - spatial_set_id(624,625)는 코히어런트 그룹(coherent group) 내의 MPD 엘리먼트들의 그룹핑을 가능하게 한다. MPD 엘리먼트들의 이러한 그룹은 예컨대 해상도 레이어 인디케이터(resolution layer indicator)로서 이용될 수 있다.
- [0102] - dynamic 파라미터(626)는 공간적 표현 디스크립터(spatial representation descriptor)(612)의 포지션이 시간의 경과에 따라 변한다는 신호를 보내기 위해 이용될 수 있다. 이 경우에, x 및 y 값들은 0으로 설정되는데, 이는 MPD가 이들을 제공하지는 않지만 이들은 MPD 외부로부터 신호를 받는다(signalled)는 것을 의미한다. 묵시적으로(implicitly), Representation의 포지션들은 비디오 컨테이너(video container) 내에서 발견된다.
- [0103] 그래서, 도 6에 도시된 바와 같은 매니페스트 파일은 서로 공간적 관계를 가진 하나 이상의 동적인 ROI 스트림들 및 하나 이상의 HEVC-타일드 스트림들(예컨대, HEVC-타일드 파노라마 스트림)을 정의할 수 있다. 상이한 MPD 비디오 엘리먼트들의 포지션 정보 내의 source 파라미터 "1"은 상이한 MPD 비디오 엘리먼트들이 공간적으로 서로 관련이 있다는 것을 나타낸다.
- [0104] 스트림들 간의 초기의(initial) 공간적 관계는 ROI 스트림 및 HEVC-타일드 스트림의 포지션 파라미터들을 기초로 하여 정의될 수 있다. 그 이후에, (dynamic 파라미터에 의해 표시된 바와 같이) MPD 내의 ROI 스트림이 "dynamic" ROI 스트림으로서 정의되기 때문에 스트림들 간의 공간적 관계는 시간에 따라 지속적으로 변할 수 있다. ROI의 움직임은 ROI 포지션 정보에 의해 클라이언트에게 제공될 수 있다. ROI 포지션 정보는 ROI 포지션 데이터(ROI 좌표)를 포함할 수 있는데, 이것은 ROI 스트림을 가지고 클라이언트 장치에게 전송될 수 있다.
- [0105] 그래서, MPD 비디오 엘리먼트들과 연관된 포지션 정보는 비디오 데이터가 클라이언트에 의해서 실제로 수신되기 전에 스트림들의 공간적 관계에 대해 HEVC 디코더에게 알려주기 위하여 DASH 클라이언트에 의해 이용될 수 있다. 이러한 방식으로, HEVC 디코더는 데이터가 수신되자마자 초기화되고 비디오 데이터의 디코딩을 시작할 수 있다.
- [0106] 도 7은 본 발명의 다른 실시예에 따라 ROI 스트리밍 프로세스를 관리하기 위한 공간 매니페스트 파일(700)을 도시한다. 특히, 도 7은 적어도 하나의 HEVC-타일드 비디오 스트림(이 구체적인 예에서, URL "full_panorama.mp4"에 의해 식별되는 광시야 비디오)와 연관된 적어도 하나의 제1 MPD 비디오 엘리먼트(702) 및 ROI 비디오 스트림(이 구체적인 예에서, URL "roi.mp4"에 의해 정의되는 ROI 비디오)와 연관된 적어도 하나의 제2 MPD 엘리먼트(704)를 포함하여 상이한 MPD 비디오 엘리먼트들을 정의하는 MPEG-DASH MPD의 예를 도시하고, ROI 스트림 및 HEVC-타일드 비디오 스트림 내의 상이한 HEVC 타일들의 공간적 관계는 (예컨대 하나 이상의 SRD의 형태로) 포지션 정보를 기초로 하여 정의된다.
- [0107] 제1 MPD 비디오 엘리먼트는 Representation(703) 및 하나 이상의 SubRepresentation들(즉, 도 3을 참조하여 기술되는 바와 같이 컨테이너 레벨(container level)에서 트랙들의 개념(concept)에 링크될 수 있는 이 Representation을 구성하는 부분들)을 포함하는 AdaptationSet으로 정의될 수 있다. 이것은 콘텐츠(오디오 및 비디오)가 개별 트랙 내의 동일한 파일 컨테이너(file container)에서 믹싱되는(muxed) 예시적인 경우에 대한 것일 수 있다. 실시예에서, 제1 MPD 비디오 엘리먼트는 HEVC-타일드 비디오 스트림을 정의할 수 있고, Representation(703)은 (이 경우에 4개의 HEVC 타일들로 이루어지는) 타일드 광시야 파노라마 스트림을 정의한

다. SubRepresentation들(716₁₋₄)은 개별 HEVC 타일들을 정의하고, 각각의 HEVC 타일은 도 3에서 도시된 바와 같이 MPEG 스트림 내에서 개별 트랙으로서 저장될 수 있다. 트랙 내의 비디오 데이터는 디코더에 의한 HEVC 타일(의 시간적 세그먼트들)의 독립적인 플레이아웃이 가능하도록 인코딩될 수 있다.

- [0108] 인코더 속성 "codecs"(719₁₋₄)는 비디오 데이터가 HEVC 타일에 속한다는 것을 나타내는 "hvt1" 타입 코덱(hvt1에서 t는 타일을 가리킴)을 지칭한다. 게다가, SubRepresentation 내의 각각의 HEVC 타일은 HEVC 타일의 포지션을 정의하기 위한 하나의 포지션 파라미터들(720₁₋₄)을 포함하는 SRD(718₁₋₄)와 연관이 있을 수 있다.
- [0109] 제2 MPD 비디오 엘리먼트는 ROI 스트림의 포지션을 정의하기 위한 하나 이상의 포지션 파라미터들(708)을 가지고 ROI 비디오의 EssentialProperty 속성에서 ROI 비디오 스트림 SRD(706)를 정의할 수 있다.
- [0110] 실시예에서, SRD 체계는 schemeIdUri "urn:mpeg:dash:srd:2014"에 의해서 정의될 수 있고, 공간적 관계는 @value 속성값들(708,710,714,720₁₋₄)을 기초로 하여 정의될 수 있다. ROI 스트림의 ROI 좌표가 시간에 따라 변하기 때문에, ROI 비디오의 SRD는 도 6을 참조하여 상세하게 기술되는 바와 같은 "dynamic" 파라미터(710)를 포함할 수 있다.
- [0111] 제1 및 제2 MPD 비디오 엘리먼트들의 SRD들(712,706,718₁₋₄)은 상이한 스트림들의 공간적 관계에 대해 HEVC 디코더에게 알려주기 위하여 이용될 수 있다.
- [0112] 도 7의 예에서, HEVC-타일드 스트림은 하나의 스트림으로서 클라이언트 장치에게 보내질 수 있다. 플레이아웃을 위한 HEVC 타일들의 선택은 디코딩 단계에서 실현될 수 있다.
- [0113] 실시예에서, 트랙 비디오 데이터는 순차적으로 저장될 수 있고, 이로써 DASH 클라이언트가 하나의 SubRepresentation에 상응하는 SubSegment들의 일부를 리트리빙하는 것을 가능하게 한다. 이러한 HTTP 요청들은 바이트 범위 시그널링(byte range signalling)을 이용할 것이고, 여기에서 각각의 SubRepresentation에 대한 바이트 범위들이 ISO/IEC 14496 part 12에 따라 소위 "ssix" 박스에서 제공된다.
- [0114] 트랙으로부터 추출되는 시간적 세그먼트들이 자급자족적(self-sufficient)이지 않은 경우에, 이 문제를 극복하기 위하여 초기화 세그먼트(Initialization Segment) URL이 Representation의 속성으로서 제공될 수 있다. HEVC 타일의 세그먼트들 내의 비디오 데이터를 렌더링하기 위해 필요한 메타데이터를 포함하는 소위 초기화 세그먼트가 정의될 수 있다. 초기화 세그먼트에서, HEVC 타일의 타일 구조 및/또는 포지션들이 제공될 수 있다.
- [0115] 도 8a는 본 발명의 다른 실시예에 따라서 ROI 스트리밍 프로세스를 관리하기 위한 공간 매니페스트 파일(800)을 정의한다. 특히, 도 8a는 상이한 MPD 비디오 엘리먼트들(802,803_{1,2},804)을 정의하는 MPEG-DASH MPD의 예를 도시한다. HEVC-타일드 파노라마 비디오와 연관된 제1 MPD 비디오 엘리먼트(802)는 URI "full_panorama-base.mp4"(814₁)에 의해 정의된다. 여기서, HEVC-타일드 파노라마 비디오는 MPD 비디오 엘리먼트들(803₁₋₂)에 의해 정의되는 (도 3을 참조하여 기술된 것과 유사한 방식으로) 하나 이상의 개별적으로 어드레싱가능한(addressable) 타일 트랙들을 참조하는 "익스트랙터(extractor)들"을 포함하는 베이스 트랙으로서 정의된다.
- [0116] MPD 비디오 엘리먼트(803₁)는 제1 타일 식별자(814₂)("full_panorama-tile1.mp4")에 의해 식별되는 제1 HEVC 타일의 비디오 데이터를 포함하는 제1 타일 트랙을 정의할 수 있고, MPD 비디오 엘리먼트(803₂)는 제2 타일 식별자(814₃)("full_panorama-tile2.mp4")에 의해 식별되는 제2 HEVC 타일의 비디오 데이터를 포함하는 제2 타일 트랙을 정의할 수 있다. 마지막 MPD 비디오 엘리먼트(804)는 ROI 비디오 스트림(이 특정한 예에서는, ROI 스트림 식별자 "zoomed_part.mp4"(806)에 의해서 정의된 ROI 비디오)을 정의할 수 있다.
- [0117] ROI 비디오의 EssentialProperty 속성에서의 SRD(808) 및 MPD 비디오 엘리먼트들의 SupplementalProperty 속성들에서의 SRD들(816₁₋₃)을 포함하는 포지션 정보는 MPD 비디오 엘리먼트들 간의 공간적 관계에 대해 HEVC 디코더에게 알려주기 위하여 이용될 수 있다. 실시예에서, 공간적 관계 디스크립터 schemeIdUri "urn:mpeg:dash:srd:2014"는 포지션 정보를 정의하기 위해 이용될 수 있고, 공간적 관계는 도 6을 참조하여 상세하게 설명된 바와 같은 @value 속성값들(810,818₁₋₃) 및 "dynamic" 파라미터를 기초로 하여 정의될 수 있다.
- [0118] 이러한 특정 실시예에서, HEVC 타일들은 독립적으로 저장가능 및 어드레싱가능한 HEVC 타일 스트림들로서 포맷팅된다(formatted). @id "파노라마"에 의해 식별되는 Representation(802)은 도 3을 참조하여 설명된 바와 같은 베이스 트랙을 나타내고, @id, "tile1" 및 "tile2"에 의해 식별되는 Representation들(803₁₋₂)에 의해 정의된

HEVC 타일들과 조합하여 이용될 수 있다.

- [0119] 그래서, 클라이언트는 (베이스 트랙 내의 정보를 이용해서) HEVC-타일드 파노라마 비디오 스트림 및 전송 노드로부터의 HEVC 타일 스트림들을 요청하기 위하여 SMF를 이용할 수 있다. 이후, 요청된 트랙들을 수신할 때, 디코더는 베이스 트랙을 파싱해서(parse), 베이스 트랙 내의 익스트랙터를 그것이 참조하는 타일 트랙의 오디오 및/또는 비디오 데이터와 교체한다. 이와 달리, 다른 실시예에서, 클라이언트는 개별 HEVC 타일 트랙들을 개별적으로 요청하기 위하여 SMF를 이용할 수 있고, HEVC 디코더는 렌더링을 위해 타일 트랙들을 디코딩할 수 있다.
- [0120] 도 8b는 본 발명의 다른 실시예에 따라서 ROI 스트리밍 프로세스를 관리하기 위한 공간 매니페스트 파일(800)을 도시한다. 특히, 도 8b는 도 8a를 참조하여 기술된 것과 유사한 방식으로 2개의 HEVC 타일들을 포함하는 HEVC-타일드 파노라마 비디오를 정의하는 MPEG-DASH MPD의 예를 도시한다. 하지만, 이 실시예에서, ROI 스트림(822)은 ROI 비디오 데이터와 연관이 있는 제1 SubRepresentation(824₁) 및 ROI 좌표와 연관이 있는 SubRepresentation(824₂)를 포함하는 Representation으로서 기술된다는 점에서 상이하게 정의된다.
- [0121] 그래서, 이러한 구체적인 실시예에서, ROI 비디오 데이터 및 ROI 좌표는 개별(separate) 트랙들, 즉, (코덱 속성 "hvc1"(826)으로 표시된 바와 같은) HEVC 인코딩 비디오 데이터를 포함하는 비디오 트랙 및 (코덱 속성 "coor"(828)으로 표시된 바와 같은) 좌표를 포함하는 메타데이터 트랙에 저장된다.
- [0122] 그래서, 이상으로부터, 클라이언트는 원하는 비디오 데이터를 클라이언트에게 전달할 수 있는 네트워크 노드 내의 전송 노드들의 위치를 찾아내기 위하여 SMF 내의 정보를 이용할 뿐만 아니라, HEVC-타일드 비디오 데이터 및/또는 ROI(비디오) 데이터가 클라이언트에 의해서 수신되기 전에 디코더를 구성하기 위하여 SMF 내에 정의된 ROI 스트림 및 HEVC-타일드 스트림의 메타데이터를 이용한다는 결론이 된다. 이 메타데이터는 예를 들어 다음을 포함할 수 있다:
 - [0123] - 선택된 공간적 세그먼트(spatial segment)가 HEVC-타일드 비디오 데이터(예컨대, 코덱 속성 "hvt1")와 관련이 있는지를 결정하기 위한 정보;
 - [0124] - (HEVC 타일을 정의하기 위한 SubRepresentation을 이용해서) HEVC 타일의 비디오 데이터가 개별 트랙들에 저장되어 있다는 신호를 보내기 위한 정보;
 - [0125] - HEVC-타일들의 수 및/또는 사이즈를 결정하기 위한 정보(예컨대, SubRepresentation 및/또는 SRD와 연관된 포지션 정보의 일부로서 표현되는 HEVC-타일들의 수);
 - [0126] - HEVC 타일들 또는 ROI 스트림의 포지션을 결정하기 위한 정보(예컨대, SRD와 연관된 포지션 정보의 일부);
 - [0127] - (예컨대, "dynamic" 파라미터를 이용해서) ROI 포지션 및/또는 ROI의 사이즈가 시간에 따라 변한다는 신호를 보내기 위한 정보;
 - [0128] - (예컨대, 메타데이터 트랙을 정의하기 위한 SubRepresentation을 이용해서) ROI 포지션 정보가 ROI 비디오 데이터와는 별개의 트랙에 저장되어 있다는 신호를 보내기 위한 정보.
- [0129] 도 9는 본 발명의 하나의 실시예에 따라 클라이언트 장치를 도시한다. 특히, 도 9는 HEVC 디코더(922) 및 크롭퍼(cropper)(924)를 포함할 수 있는 미디어 플레이어(906)에 의해 처리되는 (타일드) 콘텐츠와의 사용자 상호작용을 해석하기(interpreting) 위한 사용자 내비게이션 기능(user navigation function)(904)을 포함하는 클라이언트 장치를 도시한다. 사용자 내비게이션 기능은 디스플레이된 콘텐츠의 조작(예컨대, 패닝(panning), 줌링(zooming), 및/또는 틸팅(tilting))을 가능하게 하는 터치-스크린, 카메라, 키보드, 마우스, 트랙커볼(trackerball), 조이스틱, 마이크로폰, 헤드 트래킹 센서(head tracking sensor), 아이 및 게이즈 트래킹(eye and gaze tracking), 버튼, 또는 임의의 다른 인간-기계 인터페이스를 포함할 수 있는 사용자 인터페이스에 연결될 수 있다.
- [0130] 클라이언트 장치는 네트워크(예컨대, 미디어 서버 또는 CDN) 내의 콘텐츠 제공자 또는 콘텐츠 소스로부터 하나 이상의 공간 매니페스트 파일(SMF)들을 수신하기 위한 매니페스트 캐시(manifest cache)(922)를 더 포함할 수 있다. 도 5-8을 참조하여 상세하게 설명된 바와 같이, SMF(916)는 하나 이상의 HEVC-타일드 스트림의 타일 식별자들(URL들)(의 일부)를 정의할 수 있다. 나아가, SMF는 HEVC 타일들의 포지션을 포함하는 타일 포지션 정보를 포함할 수 있다.
- [0131] 몇몇 실시예들에서, SMF는 HTTP 적응 스트리밍 프로토콜을 기초로 하여 포맷팅될 수 있는 ROI 스트림을 정의하는 데이터 구조를 더 포함할 수 있다. 이와 달리, ROI 스트림을 정의하는 데이터 구조는 별개의 매니페스트 파

일로서 매니페스트 캐시에 저장될 수 있다.

- [0132] 실시예에서, ROI 스트림은 유니캐스트 스트림일 수 있다. 다른 실시예에서, ROI 스트림은 멀티캐스트 또는 브로드캐스트 스트림일 수 있고, 여기에 ROI 좌표가 내장된다(embedded). 이 경우에, 클라이언트 장치는 브로드캐스트 스트림 또는 멀티캐스트 스트림을 각각 요청(튠인(tune in) 또는 조인(join))하도록 구성될 수 있다.
- [0133] 매니페스트 캐시는 스트림 리트리버(stream retriever)(910) 및 스트림 선택기(912)에 연결될 수 있다. 스트림 선택기는 하나 이상의 스트림(HEVC-타일드 스트림, 하나 이상의 HEVC 타일 스트림들 및/또는 ROI 스트림)을 선택할 수 있고, 소정의 체계에 따라 네트워크(920)로부터 하나 이상의 선택된 스트림을 요청하라고 스트림 리트리버에게 지시할 수 있다.
- [0134] 아무런 사용자 상호작용이 없는 디폴트(default) 상황에서, SMF와 연관된 콘텐츠는 도 1 및 도 2를 참조하여 상세하게 설명된 바와 같은 ROI 포지션 정보를 포함하는(또는 이와 연관이 있는) ROI 스트림을 기초로 하여 클라이언트 장치에 스트리밍될 수 있다. 그래서, ROI 스트림이 HAS 스트림인 경우에, 스트림 선택기는 (디폴트) ROI 스트림의 시간적 세그먼트들을 요청하라고 스트림 리트리버에게 지시할 수 있다. ROI 스트림의 세그먼트들은 스트림 프로세서에 의해 처리되지 않지만, 미디어 플레이어의 버퍼(908)에 직접 포워딩된다(forwarded). 유사하게, ROI 스트림이 브로드캐스트 또는 멀티캐스트 스트림인 경우에, 스트림 선택기는 요청된 브로드캐스트 스트림에 조인(join)하거나 튠인(tune into)할 수 있고, 스트림 수신기는 패킷을 직접(directly) 버퍼(908)에 포워딩할 수 있다.
- [0135] 사용자 내비게이션 기능이 사용자 상호작용에 의해 트리거링될(triggered) 때, 사용자 내비게이션 기능은 ROI 스트림으로부터 ROI 포지션 정보를 추출해서 이 정보를 ROI 생성기에 포워딩하라고 스트림 프로세서에게 지시할 수 있다. ROI 생성기는 ROI 좌표를 생성하는 것을 시작할 수 있고, 제1 좌표는 사용자 상호작용이 검출되었던 순간에 클라이언트 장치에 의해 처리되었던 ROI 스트림의 하나 이상의 프레임들의 ROI 좌표에 상응할 수 있다.
- [0136] 다른 실시예에서, 스트림 프로세서는 항상 ROI 스트림으로부터 ROI 포지션 정보를 추출할 수 있고, 이 정보를 ROI 생성기에 제공할 수 있다. 이 경우에, ROI 스트림의 스트리밍 및 렌더링 동안, 렌더링된 프레임들의 ROI 좌표는 클라이언트 장치에게 항상 입수가능하다. 몇몇 실시예들에서, ROI 포지션 정보는 프레임들을 디코딩하지 않고도 ROI 스트림으로부터 추출될 수 있다. 다른 실시예들에서, ROI 포지션 정보는 프레임들을 디코딩한 후에 ROI 스트림으로부터 추출될 수 있다. 이 경우에, 스트림 프로세서는 인코딩된 프레임들을 디코딩하기 위한 디코더를 포함하는 미디어 플레이어의 일부일 수 있다.
- [0137] ROI 생성기는 사용자 내비게이션 출력 및 ROI 스트림의 ROI 좌표를 기초로 하여 ROI 좌표를 계산하는 것을 시작할 수 있다. 실시예에서, 사용자 내비게이션 기능은 사용자-상호작용을 해석할 수 있고, ROI 생성기가 특정 사용자 상호작용과 상관이 있는 새로운 ROI 좌표를 계산하는 것을 가능하게 하는 ROI 벡터를 생성할 수 있다. 예를 들어, 사용자 내비게이션 기능은 한 손가락 슬라이딩 제스처(one-finger sliding gesture)를 ROI의 패닝(panning)(병진이동(translation))으로 해석해서 그 ROI 생성기를 위한 병진이동 벡터를 생성할 수 있다. 유사한 방식으로, 두 손가락 핀치 제스처(two-finger pinch gesture)는 ROI의 주밍 동작(zooming action)으로서 해석될 수 있고, 버튼-홀드(button-hold)는 ROI의 프리징(freezing)으로서 해석될 수 있고, 구두의 사용자 명령(spoken user command) "전체 화면(fullscreen)"은 전체 화면으로의 주밍(zooming)으로서 해석될 수 있다. ROI 생성기는 사용자 입력이 수신될 때 계속해서 ROI 좌표를 계산한다.
- [0138] 스트림 선택기가 ROI 생성기로부터 계산된 ROI 좌표를 수신하는 경우에, 전체 이미지 영역 내의 서브-영역을 형성하기 위해 이용될 수 있는 인접한 HEVC 타일들의 세트를 선택하기 위하여 SMF 내의 타일 포지션 정보 및 ROI 생성기로부터 유래하는 계산된 ROI 좌표를 이용할 수 있고, HEVC 타일들 각각의 비디오 데이터는 독립적으로 저장가능 및 디코딩가능하다. 이후, 이 HEVC 타일들과 연관된 비디오 데이터는 도 1-9를 참조하여 설명된 바와 같은 SMF를 기초로 하여 요청될 수 있고, 개별 HEVC 타일 스트림들의 세트로서 클라이언트 장치에 스트리밍될 수 있다. HEVC 타일 스트림들의 수신 후에, 스트림 프로세서는 ROI를 포함하는 이미지 영역의 디코딩된 이미지 데이터를 생성하기 위한 HEVC 디코더(924)에 비디오 데이터를 보내기 전에 상이한 HEVC 타일 스트림들의 비디오 데이터를 동기화 및 믹싱(mux)할 수 있다. 이미지 영역은 ROI 좌표를 기초로 하여 크롭퍼(926)에 의해 크롭핑될 수 있다. 그래서, 중국적인(eventual) ROI는 디코딩 프로세스의 결과로 생기는 이미지 영역보다 더 작을 수 있다. 크롭핑된 이미지들(각각은 ROI를 포함함)은 디스플레이(도시되지 않음)상에 버퍼링 및 렌더링될 수 있다.
- [0139] 그래서, 이상으로부터, 사용자 상호작용은 클라이언트 장치가, 미리 결정된 ROI가 디스플레이되는 (디폴트) 눈-타일드 ROI 스트리밍 모드로부터, 사용자에게 디스플레이될 ROI가 사용자 상호작용을 통해서 조작될 수 있는 타

일드 스트리밍 모드로 스위칭하는 것을 초래할 수 있다는 결론이 된다. 용어 "미리 결정된(predetermined)"은 시간의 기간에 대하여 이미지 영역 내의 이 ROI의 공간적 포지셔닝이 클라이언트 장치의 컨트롤(control) 밖에서 미리 결정되어 있다는 사실과 관련이 있다. 실시예에서, ROI 스트리밍 모드로부터 타일드 스트리밍 모드로 스위칭하는 경우에, 클라이언트 장치는 스위칭시 이러한 미리 결정된 ROI 스트림의 ROI의 ROI 좌표를 결정하고, 실질적으로 동일한 ROI(예컨대, 콘텐츠의 렌더링 동안 스무딩한 전환이 달성되도록 하기 위하여 타일드 스트리밍 모드에서 동일한 사이즈 및 실질적으로 동일한 포지션을 가진 것)를 형성하기 위하여 이러한 ROI 좌표를 이용한다.

- [0140] 실시예에서, 외부 신호는 클라이언트 장치가 논-타일드 ROI 스트리밍 모드로부터 타일드 스트리밍 모드로 스위칭하는 것을 초래할 수 있다. 예를 들어, ROI 생성기는 ROI 스트림 내의 신호 또는 다른 외부 신호를 기초로 하여 ROI 좌표를 계산하는 것을 시작하도록 트리거링될 수 있다.
- [0141] ROI 스트리밍 모드로부터 타일드 스트리밍 모드로 스위칭하는 경우에, ROI 스트림 내의 프레임들 및 선택된 HEVC 타일 스트림들의 프레임들은 클라이언트에 의해 수신되고, 두 모드들 간의 스무딩 또는 심지어 매끄러운 전환이 가능하도록 하기 위해서 공동의 콘텐츠 플레이아웃 타임라인(common content playout timeline)을 기초로 하여 동기화될 수 있다.
- [0142] 도 10a 및 10b는 본 발명의 실시예에 따라 클라이언트 프로세스의 개략도를 도시한다. 특히, 도 10a는 사용자 내비게이션 기능으로부터 유래하는 정보를 기초로 하여 ROI들의 생성을 제어하기 위한 클라이언트 프로세스를 도시한다. 본 프로세스는 (디폴트) ROI 스트리밍 모드에서 시작할 수 있다(단계 1000). 아무런 사용자 상호작용이 검출되지 않는 한, 클라이언트 장치("클라이언트"라고도 지칭된)는 ROI 스트림(ROI 스트리밍 모드)을 기초로 하여 ROI의 렌더링을 계속할 수 있다. 만일 사용자-상호작용이 클라이언트에 의해 검출된다면(단계 1002), 스트림 선택기 및 스트림 프로세서를 위한 ROI 좌표를 계산하기 위하여(단계 1006), ROI 생성기가 시작될 수 있다(단계 1004). ROI 생성기에 의해 결정된 제1 ROI 좌표는 사용자가 상호작용하기(interacting)를 시작한 ROI 스트림 내의 하나 이상의 렌더링된 프레임들의 ROI 좌표에 상응할 수 있다. 이 ROI 좌표는 타일드 스트리밍 모드(단계 1008; 도 10b를 참조하여 더욱 상세하게 설명됨)로 스위칭하기 위하여 스트림 선택기에 포워딩될 수 있다(단계 1006).
- [0143] 사용자 내비게이션 기능은 특정 사용자 상호작용을 해석할(interpret) 수 있고(단계 1010), 클라이언트에 의해 렌더링되는 ROI 프레임들과 연관된 ROI 포지션 정보를 수신하기를 계속하는 ROI 생성기에 이 정보를 (ROI 벡터로서) 포워딩한다. ROI 생성기는 ROI 스트림 및 ROI 벡터의 ROI 포지션 정보를 기초로 하여 추가적인 ROI 좌표를 계산할 수 있고(단계 1012), 타일드 스트리밍 모드에서 ROI를 렌더링하기 위하여 계산된 ROI 좌표를 스트림 선택기 및 스트림 프로세서에 포워딩할 수 있다. 이 프로세스는 사용자 상호작용이 검출되는 한(단계 1014) 계속될 수 있다. 사용자 활동이 더 이상 존재하지 않는다는 것을 ROI 생성기가 검출하는 경우(단계 1014)에, ROI 생성기는 ROI 좌표의 생성을 중지할 수 있다.
- [0144] 특히, 도 10b는 ROI 스트리밍 모드에서 ROI를 렌더링하는 것으로부터 타일드 스트리밍 모드에서 ROI를 렌더링하는 것으로 스위칭하기 위한 클라이언트 프로세스를 도시한다. 본 프로세스는 스트림 선택기가 ROI 생성기로부터 계산된 ROI 좌표를 수신하는 것에 의해서 시작될 수 있다(단계 1020). 스트림 선택기에 의해 수신된 제1 ROI 좌표는 사용자가 상호작용하기를 시작한 ROI 스트림 내의 하나 이상의 렌더링된 프레임들의 ROI 좌표에 상응할 수 있다. 스트림 선택기는 프레임 내의 HEVC 타일들을 결정하기 위하여 타일 포지션 정보 및 ROI 좌표를 이용할 수 있고, 이로부터 계산된 ROI를 포함하는 이미지 영역이 형성될 수 있다(단계 1022). 클라이언트는 매니페스트 캐시에 저장된 SMF로부터 타일 포지션 정보를 리트리빙할 수 있다.
- [0145] 이와 달리, 다른 실시예에서, 클라이언트 장치(예컨대, 클라이언트 장치 내의 스트림 프로세서)는 (저장된) HEVC-타일드 비디오로부터 직접 타일 포지션 정보를 획득할 수 있다. 예를 들어, 도 3을 참조하여 상세하게 설명된 바와 같이, (파일로서 저장된) HEVC-타일드 비디오 스트림은 타일 트랙들을 포함할 수 있고, 각각의 타일 트랙은 타일 포지션 정보를 포함한다. 이 경우에, 어느 HEVC 타일들이 원하는 ROI와 오버랩하는지를 판단하기 위하여 아무런 SMF가 필요하지 않다.
- [0146] 스트림 선택기는 선택된 HEVC 타일들과 연관된 비디오 데이터를 네트워크로부터 요청하라고 스트림 리트리버에게 지시할 수 있다(단계 1024). 클라이언트가 요청된 HEVC 타일 스트림들을 수신하는 경우에, 공동의 콘텐츠 플레이아웃 타임라인을 기초로 하여 스트림들을 동기화할 수 있고, 비디오 데이터를 HEVC-타일드 비디오 스트림으로 멀싱할(mux) 수 있고, 디코딩을 위한 비디오 데이터를 HEVC 디코더에 보낼 수 있다(단계 1026). ROI를 포함하는 서브-영역을 나타내는 디코딩된 이미지 데이터는 계산된 ROI 좌표를 기초로 하여 크롭핑될 수 있다(단계

1028). 크롭핑된 이미지들은 디스플레이상에서 버퍼링 및 렌더링될 수 있다(단계 1030). 상술한 프로세스는 타일 스트리밍 모드로 스위칭할 때, ROI 스트리밍 모드에서 마지막으로 렌더링된 ROI가 타일 스트리밍 모드에서 최초로 렌더링된 ROI와 일치하는 것을 보장한다. 이러한 방식으로, 클라이언트는 ROI 스트리밍 모드에서 ROI들을 렌더링하는 것으로부터 타일 스트리밍 모드에서 ROI들을 렌더링하는 것으로 스무딩하게 스위칭할 수 있다.

[0147] 도 11a 및 도 11b는 본 발명의 실시예에 따라 ROI 스트리밍 프로세스를 도시한다. 이러한 구체적인 예에서, HEVC 타일 스트림들은 HTTP 적응 스트리밍 프로토콜과 같은 적응 스트리밍 프로토콜을 이용해서 CDN에 의해 클라이언트 장치들("클라이언트들"이라고도 지칭됨)에 배포될 수 있다. 본 프로세스는 클라이언트가 콘텐츠 제공자 CP로부터 공간 매니페스트 파일 SMF를 요청 및 수신하는 것으로 시작할 수 있다(단계들 1100 및 1102). 클라이언트는 SMF를 파싱할 수 있고(단계 1104), ROI 스트리밍 모드로 스트리밍 프로세스를 시작할 수 있다. 이러한 목적을 위하여, 클라이언트는 SMF를 기초로 하여 ROI 스트림의 제1 시간적(ROI) 세그먼트를 선택할 수 있고, 요청 메시지, 예컨대 HTTP GET 메시지를 CDN의 요청 라우팅 RR 노드(request routing RR node)에 보낼 수 있다(단계 1106). 요청 라우팅 노드는 전송 노드의 위치를 찾아낼 수 있고, 요청된 시간적 ROI 세그먼트들은 여기에 저장되고, 리다이렉트(redirect) 메시지 내의 국지화된(localized) 전송 노드의 URL을 클라이언트에게 도로(back) 보낸다(단계 1108). '시간적(temporal)'이라는 용어는 세그먼트가 세그먼트의(컨텐츠의) 플레이-아웃 기간과 연관된 소정의 기간을 가진다는 개념을 지칭한다. 그래서, 시간적 세그먼트는 시간 차원에서의 소스 비디오 콘텐츠의 세그먼트화의 결과이다. 이 애플리케이션에서 지칭되는 바와 같은 임의의 세그먼트는 소정의 플레이-아웃 기간을 가진다. 게다가, 소정 카테고리의 세그먼트들(SMF 또는 공간 매니페스트 파일에서 참조된 것들)은 추가적인 공간적 세그먼트화 프로세스(spatial segmentation process)일 수 있고, 비디오 콘텐츠는 세그먼트가 소스비디오 콘텐츠의 이미지 영역의 서브셋(예컨대, 일부)인 이미지 영역을 포함하도록 세그먼트화된다. 이 세그먼트들은 경우에 따라 때때로 공간적 세그먼트들 또는 타일 세그먼트(tile segment)들이라고도 지칭된다. 타일 세그먼트는 타일/(HEVC) 타일 스트림과 연관된 시간적 세그먼트의 등가물이라는 점이 주목되어야 한다. ROI 세그먼트는 시간적 세그먼트이지만, 공간적 세그먼트화의 결과가 아니고, 그래서 '정적(static)' 포지션 정보를 가질 필요가 없다(다만, 가질 수는 있음)는 점에서 상이하다. 이에 반해 타일 세그먼트는 시간의 흐름에 대해 정적이고 공간적 세그먼트화를 반영하는 타일 포지션 정보와 연관이 있다. 클라이언트는 SMF에서 식별된 ROI 세그먼트들을 요청하기 위해 URL을 이용할 수 있다.

[0148] 그래서, 리다이렉트 메시지(redirect message) 후에, 클라이언트는 요청 메시지를, 요청된 시간적 ROI 세그먼트를 포함하는 전송 노드에 도로(back) 보낼 수 있다(단계 1110). 전송 노드는 요청된 제1 ROI 세그먼트를 클라이언트에 도로(back) 보낼 수 있다(단계 1112). 이 특정한 예에서, ROI 세그먼트는 ROI의 ROI 좌표를 포함할 수 있다. ROI 좌표는 소스 비디오의 공간적 표현 내의 궤적을 정의할 수 있다. 스트림 내의 ROI 좌표를 전송하는 상이한 방식들이 더욱 상세하게 이하에서 기술된다. 제1 시간적 ROI 세그먼트가 클라이언트에 의해 버퍼링되고 플레이-아웃될(played-out) 수 있는 한편(단계 1114), 추가의 시간적 세그먼트들에 대한 리트리벌 프로세스(retrieval process)가 계속될 수 있다.

[0149] 소정의 시간 이후에, 예컨대 6번째 ROI 세그먼트의 플레이아웃 동안(단계들 1116-1118), 클라이언트는 타일 스트리밍 모드로 스위칭하도록 트리거링될 수 있다(단계 1120). 예를 들어, 클라이언트는 패닝 액션(panning action)으로서 사용자 내비게이션 기능에 의해 해석될 수 있는 사용자 상호작용을 검출할 수 있다. 타일 스트리밍 모드로 스위칭하도록 트리거링할 때, 클라이언트는 우선, 검출된 사용자 상호작용 동안 렌더링되었던 하나 이상의 프레임들과 연관된 ROI 좌표를 결정할 수 있다(단계 1122). 그 이후에, 클라이언트는 도 9 및 도 10a 및 도 10b를 참조하여 기술된 것과 유사한 방식으로 ROI 스트림의 ROI 포지션 정보를 기초로 하여 추가적인 ROI 좌표를 결정하는 것을 계속할 수 있다.

[0150] 실시예에서, 클라이언트는 ROI를 포함하는 이미지 영역을 형성하기 위해 이용될 수 있는 공간적으로 인접한(spatially adjoining) HEVC 타일들의 세트를 결정하기 위하여 SMF 내의 제1 ROI 좌표 및 타일 포지션 정보를 이용할 수 있다. 이 예에서, 선택된 HEVC 타일과 연관된 HEVC 타일 스트림은 HAS(HTTP 적응 스트리밍) 프로토콜이 타일 세그먼트들을 클라이언트에 전달하기 위해 이용될 수 있도록 하기 위하여 시간적으로 세그먼트화될 수 있다. 그 이후에, 클라이언트는 SMF를 기초로 하여 선택된 HEVC 타일 스트림들의 이 타일 세그먼트들을 요청하는 것을 시작할 수 있다. 특히, 클라이언트는 HEVC 타일들의 선택된 시간적 세그먼트들의 URL들을 포함하는 요청 메시지들을 보낼 수 있고(이 경우에, HEVC 타일들 7, 11, 및 12(즉, video7_7, video7_11 및 video7_12)의 7번째 시간적 세그먼트)(단계들 1124-1126), - 응답으로 - 선택된 HEVC 타일 스트림들의 요청된 시간적 세그먼트들을 포함하는 전송 노드로부터 응답 메시지들을 수신한다(단계들 1128-1130).

- [0151] HEVC 타일들의 비디오 데이터가 도 3을 참조하여 상세하게 기술된 바와 같은 개별 비디오 트랙들에 저장된 경우에, 베이스 트랙도 클라이언트 장치(도시되지 않음)에 의해 요청되어야 한다. 익스택트들을 포함하는 베이스 트랙은 선택된 HEVC 타일들의 비디오 데이터를 연속적인 이미지로 디코딩하기 위하여 HEVC 디코더에 의해 이용될 수 있다.
- [0152] HEVC 타일 스트림들을 결정 및 요청하는 상술한 프로세스 동안, 클라이언트는 ROI 세그먼트들(도시되지 않음)의 플레이아웃을 계속한다.
- [0153] 요청된 HEVC 타일 스트림들의 수신시, 클라이언트는 공동의 콘텐츠 플레이아웃 타임라인을 기초로 하여 HEVC 타일 스트림들과 렌더링된 ROI 스트림을 동기화할 수 있다. 인접한 HEVC 타일 스트림들의 동기화된 프레임들(즉, 특정한 공동의 플레이아웃 시간과 연관이 있는 프레임들)은 HEVC 디코더에 의해서 ROI를 포함하는 이미지 영역으로 디코딩될 수 있다(단계 1132). 만일 필요하다면, ROI 좌표는 클라이언트가 요청된 HEVC 타일 스트림들을 수신했을 때 플레이아웃되었던 ROI 스트림의 ROI와 실질적으로 유사한 이미지 영역으로 이미지 영역을 크롭핑하기 위해 이용될 수 있다.
- [0154] HEVC 타일 스트림들을 기초로 한 이미지 영역들의 형성이 ROI 스트림 내의 프레임들의 플레이아웃과 동기(in sync)를 이루는 경우에, 클라이언트는 ROI 스트림을 기초로 한 ROI의 렌더링으로부터 HEVC 타일 스트림들을 기초로 한 사용자-생성(user-generated) ROI의 렌더링으로 스위칭할 수 있다(단계 1134). 타일드 스트리밍 모드에서, 클라이언트는 계속해서 HEVC 타일 스트림들의 시간적 세그먼트들을 요청할 수 있고, 공동의 플레이아웃 시간과 연관된 프레임들을 이미지 영역으로 디코딩할 수 있고, (선택적으로(optionally)) 디스플레이상에서 이를 렌더링하기 전에 이미지 영역을 크롭핑할 수 있다.
- [0155] 소정의 시간 이후에, 예컨대 선택된 HEVC 타일 스트림들의 8번째 및 9번째 시간적 세그먼트들을 기초로 하여 생성되었던 크롭핑된 이미지 영역들의 플레이아웃(단계들 1136 및 1138) 이후에, 사용자는 디폴트 ROI 스트리밍 모드로 도로 스위칭하기 위하여 클라이언트 장치와 상호작용할 수 있다(단계 1140). 이 모드에서, 콘텐츠의 렌더링은 10번째 ROI 세그먼트를 기초로 하여 계속될 수 있다(단계 1142).
- [0156] 도 12a 및 도 12b는 본 발명의 실시예에 따라 논-타일드 스트리밍 모드와 타일드 스트리밍 모드 간의 매끄러운 스위칭을 위한 프로세스를 도시한다. 도 12a 및 도 12b에서 도시된 바와 같이, 타일드 스트리밍 모드를 시작하기 위해 선택되고 있는 HEVC 타일 스트림들은 (예컨대, 사용자 상호작용에 의해서) 클라이언트가 타일드 모드로 스위칭하기 위한 트리거(trigger)를 수신하는 시점 T에서, ROI 스트림의 하나 이상의 ROI 프레임들과 연관된 ROI 포지션 정보를 기초로 하여 결정된다.
- [0157] 이 실시예는 ROI 스트림들 또는 ROI 스트림의 특정 부분들을 위해 이용될 수 있고, 여기서 ROI 궤적은 타일드 스트리밍 모드로 스위칭하기 위한 신호가 수신되는 시점에 클라이언트에 의해 선택된 인접한 HEVC 타일 스트림들에 의해 형성될 수 있는 이미지 영역의 밖에 존재하지 않을 것이다.
- [0158] 도 12a는 클라이언트 장치("클라이언트"라고도 지칭됨)가 시점 T에서 ROI 스트리밍 모드로부터 타일드 스트리밍 모드로 스위칭하기 위한 신호를 수신하는 상황을 도시한다. 이 시점에, 클라이언트는 ROI 스트림의 프레임 f를 렌더링하고 있다. 프레임 f와 연관이 있는 ROI 포지션 정보는 ROI 좌표 프레임 f(도 12a에서의 pos0)를 포함할 수 있다. 이 ROI 좌표는 프레임 f의 ROI를 포함하는 이미지 영역을 형성하기 위해 이용될 수 있는 공간적으로 인접한 (동기화된) HEVC 타일 스트림들의 세트를 리트리빙하기 위해 이용된다.
- [0159] 선택된 HEVC 타일 스트림들을 리트리빙하기 위해 필요한 시간 기간 Δt 동안, n개의 ROI 프레임들이 클라이언트에 의해 렌더링될 수 있고, 여기서 $n = \Delta t * R$ 이고, R은 ROI 스트림을 클라이언트에 전송하기 위하여 전송 노드에 의해 이용되는 (평균) 프레임 전송률(frame transmission rate)이다. 시점 T+n에서, 클라이언트는 요청된 인접한 HEVC 타일 스트림들을 수신할 것이고, 타일 프레임들을 ROI 스트림의 ROI 프레임들과 동기화시킬 것이고, HEVC 타일 스트림들의 동기화된 프레임들을, ROI를 포함하는 이미지 영역으로 디코딩할 것이다. 필요하다면, 클라이언트는 T+n에서 ROI 프레임의 ROI와 실질적으로 일치하는 이미지로 이미지 영역을 크롭핑하기 위하여, 시점 T+n에서 ROI 스트림 내의 ROI의 ROI 좌표를 이용할 수 있다. 그 이후에, 클라이언트는 ROI 스트림을 기초로 한 ROI의 렌더링으로부터 크롭핑된 이미지를 기초로 한 ROI의 렌더링으로 스위칭할 수 있고, 그래서 타일드 스트리밍 모드로의 스무딩한 전환이 실현된다.
- [0160] 도 12b는 도 12a에 도시된 바와 같이 타일드 스트리밍 모드로의 스무딩한 전환을 달성하기 위한 단계들의 프로세스 흐름을 도시한다. 우선, 클라이언트는 타일드 스트리밍 모드로 스위칭하기 위한 신호를 수신할 수 있다(단계 1202). 이후, 프레임 f와 연관된 ROI 좌표가 클라이언트에게 제공된다(단계 1204). 클라이언트는 나중 시점

에서 ROI 스트림의 ROI를 포함하는 이미지 영역을 형성하기 위해 이용될 수 있는 인접한 HEVC 타일 스트림들의 세트를 결정하기(단계 1208) 위하여 SMF 내의 타일 포지션 정보 및 ROI 좌표를 이용할 수 있다(단계 1206). 즉, T+n 시점에서, 선택된 타일 스트림들이 클라이언트에 의해 수신된다.

[0161] 선택된 HEVC 타일 스트림들은 리트리빙될 수 있고(단계 1210), 시점 T+n에서 클라이언트에 의해 수신된다. 그 이후에, 인접한 HEVC 타일 스트림들의 프레임들 및 렌더링된 ROI 스트림의 ROI 프레임들이 동기화될 수 있다. ROI 스트림의 프레임 f+n의 ROI를 포함하는 이미지 영역을 형성하기(단계 1214) 위하여 HEVC 디코더에 의해 컨텐츠 프리젠테이션 타임라인(content presentation timeline)상에서 동일한 포지션을 갖는 공간적으로 인접한 HEVC 타일 스트림들의 타일 프레임들 f+n이 디코딩될 수 있다(단계 1212). 이후, ROI 프레임 f+n과 연관된 ROI 좌표는 이미지 영역으로부터 ROI를 크롭핑하기 위해 이용될 수 있다(단계 1216). 그 이후에, 컨텐츠의 디스플레이(displaying)은 ROI 스트림 내의 ROI 프레임들을 렌더링하는 것으로부터 HEVC 타일 스트림들의 비디오 데이터에 의해 제공되는 크롭핑된 이미지 영역을 렌더링하는 것으로 스위칭될 수 있다. 이러한 방식으로, ROI 스트리밍 모드로부터 타일드 스트리밍으로의 스무딩한 전환이 실현된다.

[0162] 만일 임의의 특정한 이유로, 타일들이 아직 수신되지 않았기 때문에 및/또는 ROI가 요청된 인접한 HEVC 타일 스트림들에 의해 정의된 이미지 영역 밖에 (부분적으로) 있기 때문에, 디스플레이되는 영역이 입수가능하지(available) 않다면, 폴-백 메커니즘(fall-back mechanism)이 구현될 수 있다. 실시예에서, 소스 컨텐츠의 저해상도 비디오 스트림은 전체 스트리밍 프로세스 동안 배경 프로세스(background process)로서 클라이언트에 스트리밍될 수 있다. 소정의 컨텐츠가 입수가능하지 않은 경우에, 유실된 HEVC 타일 스트림들이 리트리빙되는 동안, 유실된 컨텐츠를 임시로 크롭핑하고 이것을 렌더링하기 위해서 저해상도 비디오가 이용될 수 있다. 이러한 방식으로, 고품질 비디오가 다소 나중에 렌더링될 수 있지만, 타일드 스트리밍 모드로의 스무딩한 전환이 달성될 수 있다.

[0163] 도 13a 및 13b는 본 발명의 다른 실시예에 따라 ROI 스트리밍 모드와 타일드 스트리밍 모드 간의 매끄러운 스위칭(seamless switching)을 위한 프로세스를 도시한다. 몇몇 상황들에서, 타일 세그먼트들을 리트리빙하기 위한 시간 Δt 는 (특히, 데이터가 최선 네트워크(best-effort network)를 통해서 전달되는 경우에) 대략 수 초일 수 있다. 게다가, 몇몇 상황들에서, ROI 스트림 내의 ROI의 궤적은 매우 동적일(dynamic) 수 있는데, 예컨대 짧은 시간 기간을 가지고 소스 파일의 이미지 영역 내의 큰 병진이동(translation)들 및/또는 줌링 액션(zooming action)들을 포함할 수 있다. 그래서, HEVC 타일의 시간적 세그먼트들의 리트리빙 시간(retrieval time) 동안, ROI의 궤적은 시점 T에서 원래 선택되었던 타일들의 세트의 이미지 영역 밖으로 시프트될 수 있어서, ROI 스트리밍 모드에서의 렌더링의 타일드 스트리밍 모드로의 스무딩한 전환이 방해될 것이다.

[0164] 이러한 문제를 처리하기 위하여, 본 발명에 따른 소정의 실시예들에서 ROI 포지션 정보는, 록-어헤드 포지션 정보가 수신되는 경우에, 클라이언트에 의해 수신되지 않은 ROI 비디오 스트림의 하나 이상의 비디오 프레임들과 연관된 록-어헤드 포지션 정보(예컨대, 록-어헤드 포지션)를 포함할 수 있다. 실시예에서 록-어헤드 포지션 정보는, 록-어헤드 포지션 정보가 수신되는 경우에, 클라이언트 장치에 의해 아직 수신되지 않은 ROI 비디오 스트림 내의 하나 이상의 이미지 프레임들의 하나 이상의 타임-시프트된 제1 ROI 좌표를 포함할 수 있다. 그래서, 이 타임-시프트된 ROI 좌표는 연관된 이미지 프레임들의 수신 전에 클라이언트 장치에 제공될 수 있다. 추가적인 실시예에서, 타임-시프트된 ROI 좌표는 타임-시프트를 결정하기 위한 타임-시프트 정보, 바람직하게는 타임스탬프와 연관이 있을 수 있다. 타임 시프트는 타임-시프트된 ROI 좌표의 수신과 관련 비디오 컨텐츠, 바람직하게는 하나 이상의 이미지 프레임들의 수신 사이의 기간을 나타낸다. 이 실시예에서, ROI 포지션 정보는 타임-시프트되고(time-shifted), 그래서 클라이언트 장치는 미리(ahead of time) ROI 포지션 정보를 수신할 수 있다. 이러한 방식으로, ROI가 소스 비디오의 이미지 영역(예컨대, HEVC-타일드 비디오의 이미지 영역) 내에서 시간에 맞추어 매우 동적인 궤적(dynamic trajectory)을 따라갈 때에도, 클라이언트는 올바른 HEVC 타일 스트림들을 요청할 수 있다.

[0165] 도 13a 및 13b에서, ROI 포지션 정보 내의 ROI 좌표는 ROI 스트림 내의 연관된 프레임들과 관련하여 미리(ahead of time) 클라이언트에게 제공될 수 있다. 실시예에서, ROI 포지션 정보는 시간 기간 (다수의 프레임들로 표현됨) $n = \Delta t * R$ 에 대하여 타임-시프트되어 있을 수 있고, R은 ROI 스트림을 클라이언트에 전송하기 위하여 전송 노드에 의해 이용되는 (평균) 프레임 전송률이다. 타임-시프트된 ROI 포지션 정보는 네트워크 내의 CDN 또는 다른 개체에 의해 준비될 수 있고, Δt 는 CDN이 세그먼트를 클라이언트에 전달하기 위한 최대 시간 기간일 수 있다.

[0166] 타임-시프트 Δt 는 네트워크의 특징적 전달 시간(characteristic delivery times)을 기초로 하여 결정될 수 있다. 예를 들어, 최선 네트워크에서, Δt 는 비교적 큰, 예컨대 수 초 이상일 수 있고, 네트워크에서 Δt 가 더 작

을 수 있는, 예컨대 수 초보다 또는 심지어 1초보다 더 작을 수 있는 소정의 QoS 규칙들에 따라서 데이터가 전달된다.

- [0167] 상술한 프로세스는 도 13a에서 개략적으로 도시되고, 클라이언트는 시점 T에서 ROI 스트리밍 모드로부터 타일드 스트리밍 모드로 스위칭하기 위한 신호를 수신한다. 이 시점에서, 클라이언트는 ROI 스트림의 프레임 f를 렌더링하고 있다. 프레임 f와 연관되어 있는 ROI 포지션 정보는 미리(ahead in time) 시프트되어 있는 ROI 좌표일 수 있다. 특히, 프레임 f에 대한 ROI 포지션 정보는 장래의 ROI 프레임 f+n (여기서, $n = \Delta t * R$)의 ROI 좌표를 포함할 수 있다. 클라이언트는 SMF 내의 타일 포지션 정보 및 타임-시프트된 ROI 좌표를 기초로 하여 HEVC 타일 스트림들의 세트를 결정할 수 있다.
- [0168] 기간(period) Δt (n개의 프레임들) 이후에, 선택된 HEVC 타일 스트림들의 제1 프레임들은 시점 T+n에 클라이언트에 도착할 것이다. 클라이언트는 공동의 콘텐츠 플레이아웃 타임라인을 기초로 하여 HEVC 타일 스트림들 및 렌더링된 ROI 스트림을 동기화할 수 있다. 인접한 HEVC 타일 스트림들의 동기화된 프레임들(특정한 공동의 플레이아웃 시간과 연관이 있는 선택된 HEVC 타일 스트림들의 프레임들)은 HEVC 디코더에 의해서 ROI를 포함하는 이미지 영역으로 디코딩될 수 있다. 이후, 클라이언트는 프레임 f+n의 ROI에 상응하는 이미지를 크롭핑하기 위하여 타임-시프트된 ROI 좌표 pos f+n을 이용해서 ROI를 포함하는 이미지 영역을 형성할 수 있다. 이러한 방식으로, ROI가 HEVC 타일 스트림들의 리트리벌 동안 움직이고 있을 때에도 타일드 스트리밍 모드로의 스무딩한 전환이 보장된다.
- [0169] 도 13b는 도 13a에 도시된 바와 같이 타일드 스트리밍 모드로의 스무딩한 전환을 달성하기 위한 단계들의 프로세스 흐름을 도시한다. 우선, 클라이언트는 타일드 스트리밍 모드로 스위칭하기 위한 신호를 수신할 수 있다(단계 1302). 이후, 프레임 f+n과 연관된 타임-시프트된 ROI 좌표가 클라이언트에게 제공된다(단계 1304). 클라이언트는 ROI 스트림의 프레임 f+n의 ROI를 포함하는 이미지 영역을 형성하기 위해 이용될 수 있는 인접한 HEVC 타일 스트림들의 세트를 결정하기(단계 1308) 위하여, SMF 내의 타일 포지션 정보 및 타임-시프트된 ROI 좌표를 이용할 수 있다(단계 1306). 선택된 HEVC 타일 스트림들은 리트리빙될 수 있고(단계 1310), 시점 T+n에서 클라이언트에 의해 수신된다. 그 이후에, 인접한 HEVC 타일 스트림들의 타일 프레임들과 렌더링된 ROI 스트림의 ROI 프레임들이 동기화될 수 있다. ROI 스트림의 프레임 f+n의 ROI를 포함하는 이미지 영역을 형성하기(단계 1314) 위하여, 콘텐츠 프리젠테이션 타임라인상에서 동일한 포지션을 갖는 공간적으로 인접한 HEVC 타일 스트림들의 프레임들 f+n이 디코딩될 수 있다(단계 1312). 타임-시프트된 ROI 좌표(pos f+n)는 이미지 영역으로부터 ROI를 크롭핑하기 위해 이용될 수 있고(단계 1316), 콘텐츠의 디스플레이는 ROI 스트림 내의 ROI 프레임들을 렌더링하는 것으로부터 HEVC 타일 스트림들의 크롭핑된 이미지 영역들을 렌더링하는 것으로 스위칭될 수 있다(단계 1318). 이러한 방식으로, ROI 스트리밍 모드의 플레이-아웃된 콘텐츠가 매우 동적인 ROI 궤적과 관련이 있다고 하더라도 ROI 스트리밍 모드로부터 타일드 스트리밍 모드로의 스무딩한 전환이 이루어질 수 있다.
- [0170] 도 13a 및 13b는 룩-어헤드 포지션 정보의 이용의 단지 하나의 실시예를 제공하고, 다른 구현들도 가능하다는 것이 제안된다. 예를 들어, 실시예에서, 룩-어헤드 포지션 정보는 클라이언트 장치의 버퍼에서 이용가능한 ROI 포지션 정보를 이용하는 예컨대 공지의 외삽 알고리즘(extrapolation algorithm)을 이용해서, 소스 비디오의 이미지 영역 내의 ROI의 궤적을 예측하는 알고리즘에 의해 생성될 수 있다.
- [0171] 도 14의 (A) 내지 (C)는 본 발명의 다양한 실시예들에 따라 ROI 포지션 정보를 도시한다. 도 14의 (A)는 특정한 관심영역을 기술하는 XML로 인코딩된 ROI 포지션 정보의 예를 도시한다. 도 14의 (A)에서의 ROI 포지션 정보는 3개의 연속적인 ROI 태그(tag)들을 도시한다.
- [0172] ROI 태그는 ROI 태그를 ROI 스트림 내의 ROI 프레임에 코릴레이팅시키기 위한 정보를 포함할 수 있다. 예를 들어, 도 14a에서, <AbsoluteTime> 태그는 스트림에서 프레임을 식별하기 위해 이용될 수 있는 타임스탬프 또는 다른 (고유한) 값을 포함할 수 있다. 타임스탬프는 ROI와 연관된 절대 시간(absolute time)을 정의할 수 있다. 실시예에서, 타임스탬프는 연관이 있는 ROI 프레임의 레코딩 시간을 정의하는 벽시계 시간(wall-clock time)을 포함할 수 있다. 절대 시간은 부동 소숫점 표기(floating point notation)로 1-1-1970 이후의 초의 숫자로서 인코딩될 수 있다(예컨대, 유닉스 에폭 시간(Unix epoch time)). 타임스탬프를 이용하는 경우에, 타임스탬프는 또한, 프레임들이 올바른 ROI 포지션 정보(ROI 태그)에 코릴레이팅될 수 있도록 비디오 프레임(내)에 내장되어야 한다.
- [0173] 나아가, ROI 태그는 ROI의 포지션 정보를 포함할 수 있다. 예를 들어, <PixelX1X2Y1Y2> 태그는 픽셀 단위(pixel unit)의 관점에서 ROI의 포지션을 기술할 수 있다. 스크린의 꼭대기-왼쪽(top-left) 픽셀은 (X,Y)=(0,0)에 상응할 수 있다. 이러한 방식으로, X1은 ROI의 왼쪽편(left side)을 정의하고; X2는 ROI의 오른쪽편(right side)을

정의하고; Y1은 ROI의 맨 위편(topside)을 정의하고; Y2는 ROI의 맨 아래편(bottom side)을 정의한다. 풀 HD 1920x1080 스크린을 위한 이러한 체계에서, 하부-오른쪽(lower-right) 픽셀은 (X,Y)=(1919, 1079)에 상응한다.

- [0174] 이 예에서의 제1 ROI는 풀 스크린(full screen)을 정의할 수 있다. 이 예에서의 제2 ROI는 80%만큼 중앙으로의 주밍(zooming)을 나타낼 수 있다. 이 예에서의 제3 ROI는 오른쪽으로의 20 픽셀의 병진이동을 나타낼 수 있다.
- [0175] 도 14의 (B)는 XML로 인코딩된 ROI 포지션 정보의 다른 실시예를 도시한다. ROI 태그는 ROI에 적용되는 상대 시간(relative time)을 기술하는 <RelativeTime> 태그를 포함할 수 있다. 상대 시간은 비디오 콘텐츠의 시작점(starting point)에서부터 카운팅될 수 있고, 부동 소숫점 표기로 초로서 인코딩될 수 있다. 상대 시간은 ROI 태그를 ROI 프레임에 관련시키기 위하여 코릴레이션 정보(correlation information)로서 이용될 수 있다. 상대 시간을 이용하는 경우에, ROI 스트림의 명확하게 정의된 시작점(starting point)이 정의되어야 한다.
- [0176] 나아가, ROI 태그는 정규화된 포맷(normalized format)(X,Y,zoom)으로 기술된, ROI의 위치를 기술하는 <NormalizedXYzoom> 태그를 포함할 수 있다. 여기서, 좌표 "X,Y"는 ROI의 꼭대기-왼쪽(top-left) 픽셀의 포지션을 부동 소숫점 표기로 나타낼 수 있고, X 및 Y 양쪽 모두는 0-1 범위 내에 있다. 이 예에서, 스크린의 꼭대기-왼쪽 픽셀은 (X,Y)=(0,0)에 상응하고, 풀 스크린의 하부 오른쪽(lower right)은 (X,Y)=(1,1)에 상응한다. 파라미터 "zoom"은 줌 레벨을 정의할 수 있고, 이것은 또한 부동 소숫점 표기로 존재할 수 있다. zoom=1은 풀 스크린에 해당할 수 있다.
- [0177] 도 14의 (B)의 제1 ROI 태그는 풀 스크린을 나타낼 수 있다. 제2 ROI는 50%만큼의 꼭대기 왼쪽(top left)으로의 주밍(zooming)을 나타낼 수 있다. 예에서의 제3 ROI는 ROI의 꼭대기 왼쪽의 스크린의 중앙으로의 이동을 나타낼 수 있다.
- [0178] 도 14의 (C)는 ROI 태그에서 이용될 수 있는 태그들의 예들을 도시한다. 만일 프레임 번호들이 콘텐츠로부터 도출될 수 있다면, FrameNumber 태그는 코릴레이션 정보로서 이용될 수 있다. PixelXYWH 태그는 X,Y 좌표뿐만 아니라 픽셀 단위로 표현된 ROI의 폭 및 높이를 정의할 수 있다. RadialNormalizedXYR 태그는 원형 ROI의 X,Y 좌표 및 그것의 반지름을 정의할 수 있다. 실시예에서, X,Y 좌표는 0-1 스케일(scale)로 정규화된(normalized) 수 있다. SpeedPixelXdYdWdH 태그는 초-당-픽셀수(pixels-per-second)로 표시된, ROI의 X,Y, 폭 및 높이의 시간 미분(time derivative)을 정의할 수 있다.
- [0179] 도 15는 ROI 포지션 정보가 MPEG 트랜스포트 스트림(MPEG TS) 내의 기초 스트림으로서 전송되는 실시예를 도시한다. 특히, ROI 스트림의 ROI 좌표는 Private Data Stream ID를 가진 별개의 PES(Packetized Elementary Stream)로 패키징될(packaged) 수 있다. 이후, ROI 좌표 스트림이라고 지칭될 수 있는 이 PES 스트림은 ROI 스트림을 또한 포함하는 MPEG 트랜스포트 스트림으로 멀티플렉싱될(multiplexed) 수 있다.
- [0180] 도 15는 비디오 스트림(Type=0x2, PID=0x1B63), 오디오 스트림(Type=0x3, PID=0x1B64), 및 ROI 좌표 스트림(새로운 Type=0x20, PID=0x1B66)을 포함하는 PMT의 XML 표현을 도시한다. 새로운 Stream Type은 ROI 좌표 정보가 들어 있는 PES 스트림을 나타낼 수 있다. 이 Stream Type은 이후 ROI 좌표 스트림이 들어 있는 PID를 나타내기 위하여 PMT(Program Map Table) 테이블에서 이용될 수 있다.
- [0181] 클라이언트 측에서, 비디오와 ROI 좌표 간의 동기화는 (단일한 PMT에 의해 기술되는, 그리고) 특정 프로그램에 속하는 모든 PES 패킷들에 공통인, PTS 타임스탬프의 이용을 통해서 실현될 수 있다. 예로서, 만일 클라이언트가 특정 시점에 ROI 스트림의 ROI 좌표를 체크(check)할 필요가 있다면, 현재 디스플레이된 비디오 프레임의 PTS 값을 체크할 수 있다. 그 이후에, 동일한 PTS를 가지고 패킷에 대한 ROI 좌표(ROI 좌표 스트림)를 포함하는 PES를 검색할 수 있다. 이 패킷 내에서, 특정 비디오 프레임에 대한 ROI 좌표 정보를 찾아낼 것이다.
- [0182] 새로운 Stream Type을 정의하는 대신 (MPEG 트랜스포트 스트림을 가지고 또한 이용될 수 있는) 추가적인 실시예에서, DVB-정의 SAD(SynchronisedAuxiliary Data) 구조(TS 102 823 참조)를 이용하여 ROI 좌표 데이터(예컨대 (선택적으로(optionally) 타임-시프트된) ROI 좌표)가 전송될 수 있다. 이 경우에, PMT 테이블(table)은 프라이빗(private) 데이터 PES 스트림을 나타내기 위하여 0x06 Stream Type을 공시할(advertise) 수 있다. 특정 SAD가 ROI 좌표를 포함한다는 것을 나타내기 위하여, 새로운 Auxiliary 데이터 디스크립터가 정의될 수 있다(TS 102 823의 5.1 섹션 참조).
- [0183] 새로운 PES Stream Type을 정의하는 대신 ROI 좌표 스트림을 운반하고 나타내기 위하여 SAD 구조를 이용하는 것을 제외하고, 이 실시예의 모든 다른 관점들은 도 15를 참조하여 기술된 바와 같은 실시예를 위해 설명된 것과 유사하다. SAD 패킷들은 또한 PTS 값들을 포함하기 때문에, 이것들은 ROI 좌표 스트림과 ROI 스트림 간의 동기

화를 제공하기 위해 이용될 수 있다.

- [0184] 추가적인 실시예에서, 타임-시프트된 ROI 좌표를 전송하기 위하여, SAD 패킷은 ROI 좌표 데이터를 포함할 뿐만 아니라 그 좌표가 관련되어 있는 타임스탬프를 포함한다. 이러한 타임스탬프는 프레임 번호(frame number), 콘텐츠 타임스탬프(content timestamp), 또는 PTS/DTS 값의 형태일 수 있다. 후자의 경우에, 패킷 헤더(packet header) 내의 PTS 값은 TS 스트림에 패킷을 배치하기 위하여 TS 멀티플렉서(multiplexer)에 의해 이용되는 한편, 페이로드(payload)에 포함된 PTS 패킷은 좌표를 특정 비디오 프레임에 상관시키기 위하여 이용된다.
- [0185] 도 16a 및 16b는 ROI 데이터가 MPEG 스트림의 비트스트림 내에서 전송되는 본 발명의 실시예를 도시한다. 특히, 도 16a는, H.264/MPEG-4 기반 코덱을 이용하여 인코딩된 MPEG 스트림의 비트스트림 내에 SEI(supplemental enhancement information) 메시지로서 ROI 데이터가 삽입되는 실시예를 도시한다.
- [0186] 이 체계에서, NAL Unit은 비트스트림 내의 기초 블록(elementary block)을 정의할 수 있다. SEI 메시지는 SEI(supplemental enhancement information)(ISO/IEC 14496-10 AVC에서 7.4.1 NAL Units semantics 참조)와 연관된 NAL Unit으로서 정의된다. 여기서, SEI 메시지는 type 5 message: user data unregistered(타입 5 메시지: 등록되지 않은 사용자 데이터)로서 정의된다. SEI 메시지는 ROI 데이터의 파라미터들에 대한 정수로 이루어진 미리 결정된 숫자(이 경우에는 4개의 정수들)를 포함할 수 있다.
- [0187] SEI 메시지는 디코딩의 프로세스를 지원하기 위하여 추가 정보(extra information)를 운반할 수 있다. 그럼에도 불구하고, 이것의 존재는 디코딩된 신호를 구성하기 위하여 강제적이지 않아서, 합치하는 디코더들(conforming decoders)은 이 추가 정보를 고려하는 것이 요구되지 않는다. 다양한 SEI 메시지 및 그 시맨틱스(semantics)의 정의는 ISO/IEC 14496-10:2012의 D.2에서 정의된다.
- [0188] 등록되지 않은 사용자 데이터라고 지칭되는 SEI 메시지 타입은 임의의 데이터가 비트스트림 내에서 운반되는 것을 가능하게 한다. ROI 좌표의 경우에, 이 SEI 메시지는 ROI 좌표를 운반하기 위하여 이용될 수 있다. 4개의 파라미터들, 즉, 꼭대기 왼쪽 코너(top left corner)의 수평 포지션 및 수직 포지션과 폭 및 높이는 소스 비디오의 디코딩된 프레임들의 이미지 영역에서 ROI를 정의하기 위해 이용될 수 있다.
- [0189] 도 16b는 코딩된 비트스트림(coded bitstream) 내에서 ROI 데이터가 전송되는 본 발명의 다른 실시예를 도시한다. 프레임 헤더 내의 부울리언 플래그(Boolean flag)는 이러한 정보가 존재하는지 여부를 나타낼 수 있다. 플래그가 설정되어(set) 있는 경우에, 플래그에 이어지는 비트들은 ROI 좌표, 예컨대, 수평 및 수직 축에서의 꼭대기-왼쪽 코너(top-left corner)의 포지션뿐만 아니라 소스 비디오의 디코딩된 프레임들의 이미지 영역의 내에서의 ROI의 폭 및 높이를 나타낼 수 있다.
- [0190] 도 17a 및 17b는 ROI 데이터가 비디오 컨테이너 내에서 전송되는 본 발명의 실시예들을 도시한다.
- [0191] 특히, 이 도면들은 ROI 좌표가 MP4 파일 포맷(ISO/IEC 14496-14)과 같은 비디오 컨테이너 내에서 전송되는 실시예들과 관련이 있다. 이러한 파일의 일반적인 구조(structure)는 도 17a에서 기술된다. MP4 파일 포맷은 미디어 데이터 및 이와 연관된 메타데이터를 저장하고 액세스하기 위하여 계층 구조(hierarchical structure)를 이루는 박스(box)들의 세트를 명시한다. 예를 들어, 콘텐츠와 관련된 메타데이터를 위한 루트(root) 박스는 "moov" 박스인 반면에, 미디어 데이터는 "mdat" 박스에 저장된다. 더욱 구체적으로, "stbl" 박스 또는 "Sample Table Box"는 트랙의 미디어 샘플들을 인덱싱해서(index), 추가적 데이터를 각각의 샘플과 연관시키는 것을 가능하게 한다. 비디오 트랙의 경우에, 샘플은 비디오 프레임이다. 결과적으로, "Sample ROI coordinates" 또는 "stro"라고 지칭되는 새로운 박스를 박스 "stbl"에 추가하는 것은 비디오 트랙의 모든 프레임을 위한 ROI 좌표를 저장하기 위하여 이용될 수 있다. 도 17b는 ROI 좌표가 새로운 박스 "stro" 내에서 전송되는 본 발명의 실시예를 도시하고, 이것은 비디오 트랙의 모든 비디오 샘플들 또는 비디오 샘플들의 일부를 위한 ROI 좌표를 포함할 수 있다.
- [0192] 도 18은 ROI 포지션 정보가 MPEG-4 스트림의 전용 메타데이터 트랙으로서 전송되는 본 발명의 실시예를 도시한다. ROI 포지션 정보 트랙은 (오디오 또는 비디오와 같이) 정규 트랙(regular track)의 구조를 따르지만, 그 안의 데이터는 비디오 또는 오디오 프레임들이 아니라 (특정한 샘플 엔트리 코드(specific sample entry code)를 가진) ROI 좌표를 포함하는 메타데이터이다.
- [0193] 표준 ISO/IEC 14496 part 12는 ISO Base Media File Format 내에 (그리고 MP4 파일 포맷에 대한 상속에 의해) 타임드 메타데이터 트랙(timed metadata track)을 포함하기 위한 메커니즘을 제공한다. 이 실시예에서, 메타데이터 트랙 및 비디오 트랙은 ROI 데이터 트랙(ROI 좌표)의 샘플이 비디오 트랙(비디오 프레임)의 샘플에 맵핑되도록 하기 위하여 동기화되어야 한다. 실제로, 메타데이터 트랙은 트랙-참조(track-reference) 'tref' 박스들

이용해서 비디오 트랙에 링크되고, 이것의 참조 타입은 (reference type)은 'cdsc'이고, track_id는 참조된 트랙 (referenced track)의 id이다.

- [0194] 추가적인 실시예에서, ROI 좌표는 대역 밖의(out-of-band) 파일 또는 스트림을 통해서 클라이언트 장치에 전달될 수 있다. 동기화는 타임스탬프들을 이용해서 달성될 수 있다. 실시예에서, ROI 좌표는 (예컨대, VoD 스트림들의 경우에) 플레이백(playback)을 시작하기 전에 또는 (라이브 스트림들의 경우에) WebSocket을 통하여 데이터를 전달하기 전에 리트리빙되는 파일 내에 놓일 수 있다.
- [0195] 이 경우에, ROI 좌표는 실제 비디오 데이터를 가지고 대역 밖으로(out-of-band) 보내지기 때문에, 동기화는 PTS 값들과 같은 비디오 컨테이너-내부 정보(video container-internal information)를 기초로 하지 않을 수 있다. 대신, 콘텐츠 타임스탬프들이 이용될 수 있다. 이러한 타임스탬프들은 라이브 브로드캐스트 스트림들에 대해 콘텐츠 시간(즉, 비디오 스트림의 시작시에 00:00:00에서 시작함)을 기초로 하거나, 또는 벽시계 시간(예컨대 20:31:11)을 기초로 할 수 있다. 유사한 타임스탬프들이 또한 비디오 스트림 내에 존재할 수 있다. MPEG, DASH 또는 HLS 스트림들에 대해서, 이러한 타임스탬프들은 MPD 또는 매니페스트 파일에 의해 제공될 수 있다. MPEG 내에 삽입된 예컨대 DVB, DVB 타임라인 패킷들을 통하여 ROI 스트림이 전달되는 경우에, TS 멀티플렉스(TS Multiplex)(102 823 참조)가 이용될 수 있다.
- [0196] 다른 실시예에서, ROI 좌표는 비디오(워터마크) 내에서 워터마크로서 전송될 수 있다. 추가적인 실시예에서, ROI 좌표는 ROI 스트림 내에 워터마크로서 배치될 수 있다.
- [0197] 도 19a 및 19b는 본 발명의 실시예에 따라 타일드 스트리밍을 위한 소위 "공간적 버퍼(spatial buffer)"의 이용을 도시한다. 도 19a는 ROI 스트림의 ROI 프레임을 도시하고, 여기서 ROI 프레임은 "공간적 버퍼"(1902), 즉, (이미지(1904)의 중심 영역(centred region)의 둘레에 있는) 이미지 데이터의 주변 영역(peripheral region)을 포함하고, 이것은 클라이언트에게 전송되지만, 사용자가 디스플레이된 콘텐츠와 상호작용하지 않거나 사용자가 디스플레이된 콘텐츠와 상호작용하는 것이 허용되지 않는 경우에는 디스플레이되지 않는다. 그래서, 이 경우에, 아무런 사용자-상호작용이 없는 경우에 디스플레이될 수 있는 중심 영역은 ROI 프레임의 밖으로부터 크롭핑되거나 사용자에게 디스플레이될 수 있다. 이후, 사용자가 콘텐츠와 상호작용하는 경우에, ROI(1906)(의 일부)는 공간적 버퍼로서 정의되는 구역 안으로 탐색될(navigated) 수 있다. 하지만, 사용자-상호작용이 존재하기 때문에, 공간적 버퍼로부터의 이미지 데이터의 디스플레이가 허용된다. 사용자 상호작용은, 도 9-14를 참조하여 상세하게 설명된 바와 같은 ROI를 기초로 하여 타일드 스트리밍 모드로 스위칭하고 타일들을 요청하도록 클라이언트를 트리거링할(trigger) 수 있다. HEVC 타일 스트림의 리트리빙 동안, 이동하는(moving) ROI(1908)는 공간적 버퍼의 구역 내에 부분적으로 위치해 있는 이미지 영역으로부터 ROI를 크롭핑함으로써 디스플레이될 수 있다. 본 발명에 따른 실시예에서, 공간적 버퍼(이것은 디스플레이되지 않는 ROI 스트림의 이미지의 영역임)의 이미지 영역의 치수(dimension)들을 나타내기 위하여 SMF에서, 바람직하게는 MPD(Media Presentation Description)에서 하나 이상의 파라미터(들)가 정의된다. 예컨대 @spatialbuffersize라고 불리는 이 파라미터는 예를 들어 다음의 포맷들 중의 하나 이상으로 정의될 수 있다:
- [0198] (ROI의 이미지 영역의) 폭 및 높이의 % 단위의 엣지 사이즈(Edge size)
- [0199] (ROI의 디스플레이된 이미지 영역 둘레에 균일하게 분포된) 픽셀 단위의 엣지 사이즈(Edge size)
- [0200] (ROI의 디스플레이된 이미지 영역 둘레의 왼쪽, 오른쪽, 꼭대기(top), 바닥(bottom)) 픽셀 단위의 엣지 사이즈(Edge size)
- [0201] 이와 달리, 이용가능한 ROI 좌표의 세트는 4(x,y,w,h)에서 8로 확장될 수 있고, 그래서 내부 (디스플레이된) ROI 및 (공간적 버퍼를 포함하는) 외부 ROI 치수들을 각각 나타낼 수 있다.
- [0202] 도 20은 본 발명의 실시예에 따른 3개-모드의 ROI 스트리밍 모델을 도시한다. 이러한 구체적인 실시예에서, 클라이언트는 타일드 스트리밍 모드(제1 모드)에 있을 수 있고, 여기서 사용자-생성 ROI는 4개의 타일들을 기초로 하여 렌더링된다. 만일 사용자가 상호작용하기를 멈추고 ROI 스트리밍 모드(제3 모드) 내의 콘텐츠를 보길 원한다면, 시스템은 ROI 스트리밍 모드로 즉시 스위칭하는 대신 소정의 시간 동안 제2 모드로 스위칭할 수 있다. 이 제2 모드에서, 클라이언트는 ROI 스트림의 ROI 포지션 정보를 수신하는 동안 타일드 스트리밍 모드에 머무른다. 이후, 클라이언트는 도 9-14를 참조하여 상세하게 설명된 바와 같은 HEVC 타일 스트림들의 세트의 동기화된 프레임들을 기초로 하여 타일들을 선택하고 크롭핑된 이미지(cropped image)를 생성하기 위하여 ROI 포지션 정보를 이용한다. 만일 사용자가 디스플레이된 콘텐츠와 상호작용하기를 시작한다면, 사용자-생성 ROI를 디스플레이하는 동안 타일드 스트리밍 모드(제1 모드)로 즉시 스위칭할 수 있다. 이와 달리, 만일 사용자가 곧 상호작용할

것이라고 클라이언트가 예상한다면 또는 사용자가 콘텐츠와 상호작용하길 원할 수 있는 높은 가능성이 존재하는 소정의 기간 동안의 콘텐츠 디스플레이가 시작되거나 곧 시작될 것이라면, 클라이언트 로직(client logic)은 제 2 모드로의 스위칭을 초래할 수 있고, 그래서 사용자-생성 ROI를 디스플레이하기 위한 타일드 스트리밍 모드로의 전환은 신속하고 스무딩한(매끈한) 방식으로 실현될 수 있다.

[0203] 도 21은 도 1-20을 참조하여 설명된 바와 같은 시스템들 및 방법들에서 이용될 수 있는 예시적인 데이터 처리 시스템을 도시하는 블록도이다. 데이터 처리 시스템(1000)은 시스템 버스(system bus)(2106)를 통해서 메모리 엘리먼트들(2104)에 연결될 수 있는 적어도 하나의 프로세서(2102)를 포함할 수 있다. 그래서, 데이터 처리 시스템은 메모리 엘리먼트들(2104) 내에 프로그램 코드를 저장할 수 있다. 나아가, 프로세서(2102)는 시스템 버스(2106)를 통해서 메모리 엘리먼트들(2104)로부터 액세스되는 프로그램 코드를 실행할 수 있다. 하나의 관점에 있어서, 데이터 처리 시스템은 프로그램 코드를 저장 및/또는 실행하기에 적합한 컴퓨터로서 구현될 수 있다. 하지만, 데이터 처리 시스템(2100)은 본 명세서 내에서 설명되는 기능들을 수행할 수 있는 프로세서 및 메모리를 포함하는 임의의 시스템의 형태로 구현될 수 있다고 이해되어야 한다.

[0204] 메모리 엘리먼트들(2104)은 예를 들어, 로컬 메모리(local memory)(2108) 및 하나 이상의 벌크 저장 장치(bulk storage device)(2110)와 같은 하나 이상의 물리적 메모리 장치를 포함할 수 있다. 로컬 메모리는 프로그램 코드의 실제 실행 동안 일반적으로 이용되는 RAM(random access memory) 또는 다른 비-지속성(non-persistent) 메모리 장치(들)을 지칭할 수 있다. 벌크 저장 장치는 하드 드라이브(hard drive) 또는 다른 지속성(persistent) 데이터 저장 장치로서 구현될 수 있다. 처리 시스템(2100)은 또한 프로그램 코드가 실행 동안 벌크 저장 장치(2110)로부터 리트리빙되어야 하는 횟수를 감소시키기 위하여 적어도 몇몇 프로그램 코드의 일시적 저장을 제공하는 하나 이상의 캐시 메모리(도시되지 않음)를 포함할 수 있다.

[0205] 입력 장치(2112) 및 출력 장치(2114)로서 도시된 입력/출력(I/O) 장치들은 데이터 처리 시스템에 선택적으로(optionally) 연결될 수 있다. 입력 장치의 예들은 예를 들어, 키보드, 마우스와 같은 포인팅 장치 등을 포함할 수 있지만 이에 한정되지 않는다. 출력 장치의 예들은 예를 들어, 모니터 또는 디스플레이, 스피커 등을 포함할 수 있지만, 이에 한정되지 않는다. 입력 장치 및/또는 출력 장치는 개재하는(intervening) I/O 제어기(controller)들을 통해서 또는 직접 데이터 처리 시스템에 연결될 수 있다. 데이터 처리 시스템이 개재하는 프라이빗(private) 또는 퍼블릭(public) 네트워크들을 통해서 다른 시스템들, 컴퓨터 시스템들, 원격 네트워크 장치, 및/또는 원격 저장 장치에 연결될 수 있도록 하기 위해서 네트워크 어댑터(network adapter)(2116)도 데이터 처리 시스템에 연결될 수 있다. 네트워크 어댑터는 상술한 데이터에 대해 시스템들, 장치들 및/또는 네트워크들에 의하여 전송되는 데이터를 수신하기 위한 데이터 수신기 및 상술한 데이터를 시스템들, 장치들 및/또는 네트워크들에 전송하기 위한 데이터 송신기를 포함할 수 있다. 모뎀들, 케이블 모뎀들, 및 이더넷 카드(Ethernet card)들은 데이터 처리 시스템(2150)과 함께 이용될 수 있는 상이한 타입의 네트워크 어댑터의 예들이다.

[0206] 도 21에서 도시된 바와 같이, 메모리 엘리먼트들(2104)은 애플리케이션(2118)을 저장할 수 있다. 데이터 처리 시스템(2100)은 애플리케이션의 실행을 용이하게 할 수 있는 운영 체제(도시되지 않음)를 더 실행할 수 있다고 이해되어야 한다. 실행가능한(executable) 프로그램 코드의 형태로 구현되는 애플리케이션은 데이터 처리 시스템(2100)에 의해, 예컨대, 프로세서(2102)에 의해 실행될 수 있다. 애플리케이션을 실행하는 것에 반응하여, 데이터 처리 시스템은 이하에서 더 상세하게 설명될 하나 이상의 오퍼레이션을 수행하도록 구성될 수 있다.

[0207] 하나의 관점에 있어서, 예를 들어, 데이터 처리 시스템(1000)은 클라이언트 데이터 처리 시스템을 나타낼 수 있다. 클라이언트 데이터 처리 시스템은 클라이언트 장치라고도 지칭된다. 이 경우에, 애플리케이션(2118)은 실행될 때 "클라이언트"와 관련하여 본 명세서에서 기술된 다양한 기능들을 수행하도록 데이터 처리 시스템(2100)을 구성하는 클라이언트 애플리케이션을 나타낼 수 있다. 클라이언트의 예들은 퍼스널 컴퓨터, 포터블(portable) 컴퓨터, 모바일 폰 등을 포함하지만 이에 한정되지 않는다.

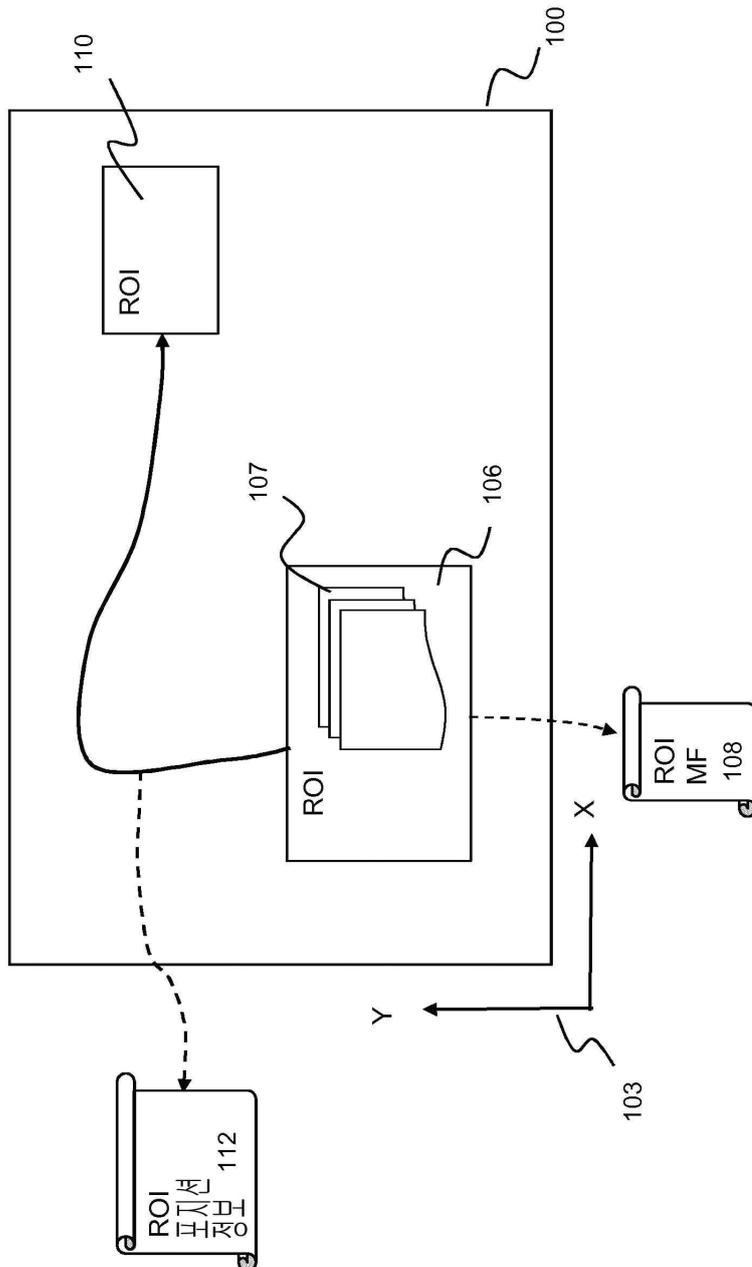
[0208] 본 명세서에서 사용된 용어는 단지 특정 실시예들을 설명할 목적을 위한 것이며, 본 발명의 제한이 되도록 의도된 것이 아니다. 본 명세서에서 사용될 때, 단수 형태들("a," "an," 및 "the")은 문맥에서 명백하게 달리 지정하지 않는다면 복수의 형태도 포함하는 것으로 의도된다. "포함하다(comprises)" 및/또는 "포함하는(comprising)"이라는 용어는 본 명세서에서 사용될 때 언급된 특징들, 정수들, 단계들, 오퍼레이션들, 엘리먼트들, 및/또는 구성요소들의 존재를 특정하지만, 하나 이상의 다른 특징들, 정수들, 단계들, 오퍼레이션들, 엘리먼트들, 구성요소들, 및/또는 이들의 그룹들의 존재 또는 추가를 불가능하게 하는 것은 아니라고 더 이해되어야 할 것이다.

[0209]

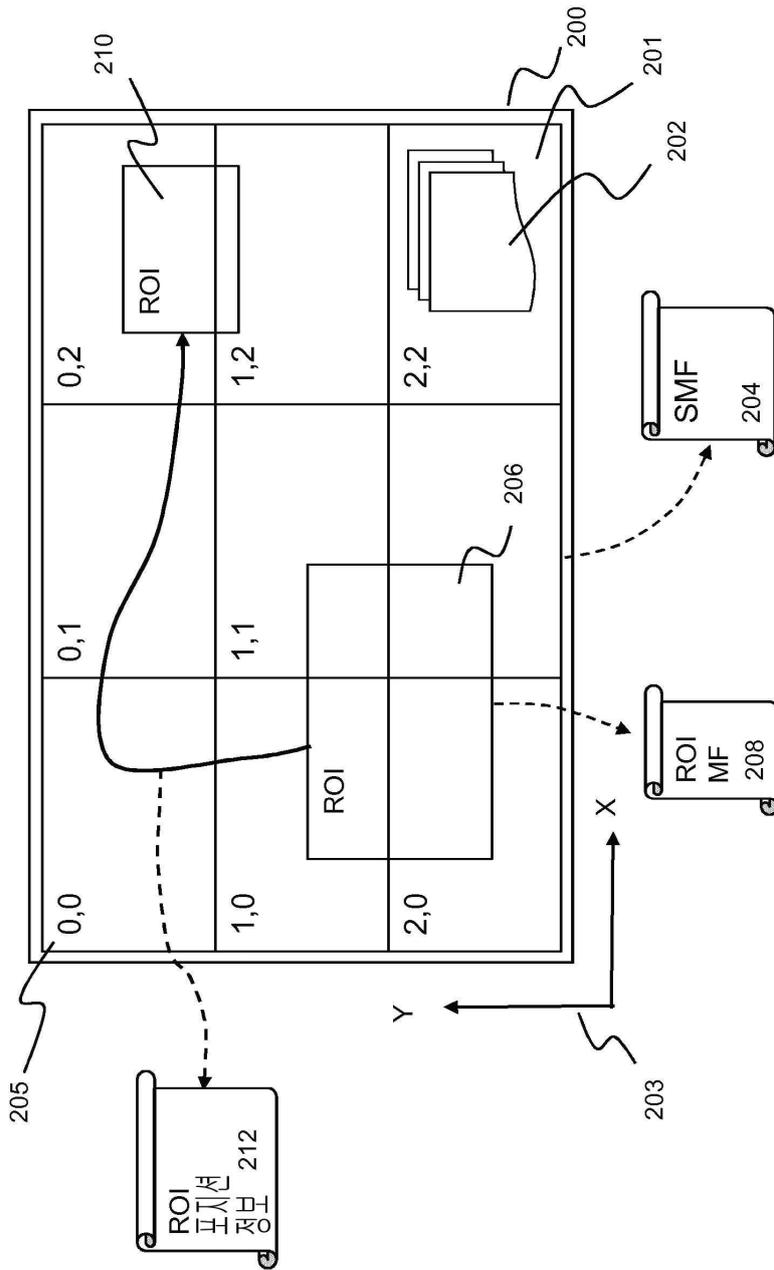
이하의 청구항들에서의 모든 수단 또는 단계 플러스 기능(means or step plus function) 엘리먼트들의 상응하는 구조, 물질, 작동, 및 등가물은 구체적으로 청구된 바와 같은 다른 청구된 엘리먼트들과 조합하여 기능을 수행하기 위한 임의의 구조, 물질, 또는 작동을 포함하는 것으로 의도된다. 본 발명의 설명은 실례 및 설명의 목적을 위해서 제시되었으며, 본 명세서에서 공개된 형태로 본 발명을 한정할 것(exhaustive) 또는 제한하도록 의도된 것이 아니다. 많은 변경들 및 변형들은 본 발명의 범위 및 사상으로부터 벗어나지 않으면서 통상의 기술자에게 분명할 것이다. 실시예는 본 발명의 원리 및 실제적인 애플리케이션을 가장 잘 설명하기 위하여 선택 및 기술되었고, 통상의 기술자들 중의 다른 자들이 심사숙고된 특정 사용에 적합한 다양한 변형들을 가진 다양한 실시예들에 대해 본 발명을 이해할 수 있도록 하기 위하여 선택 및 기술되었다.

도면

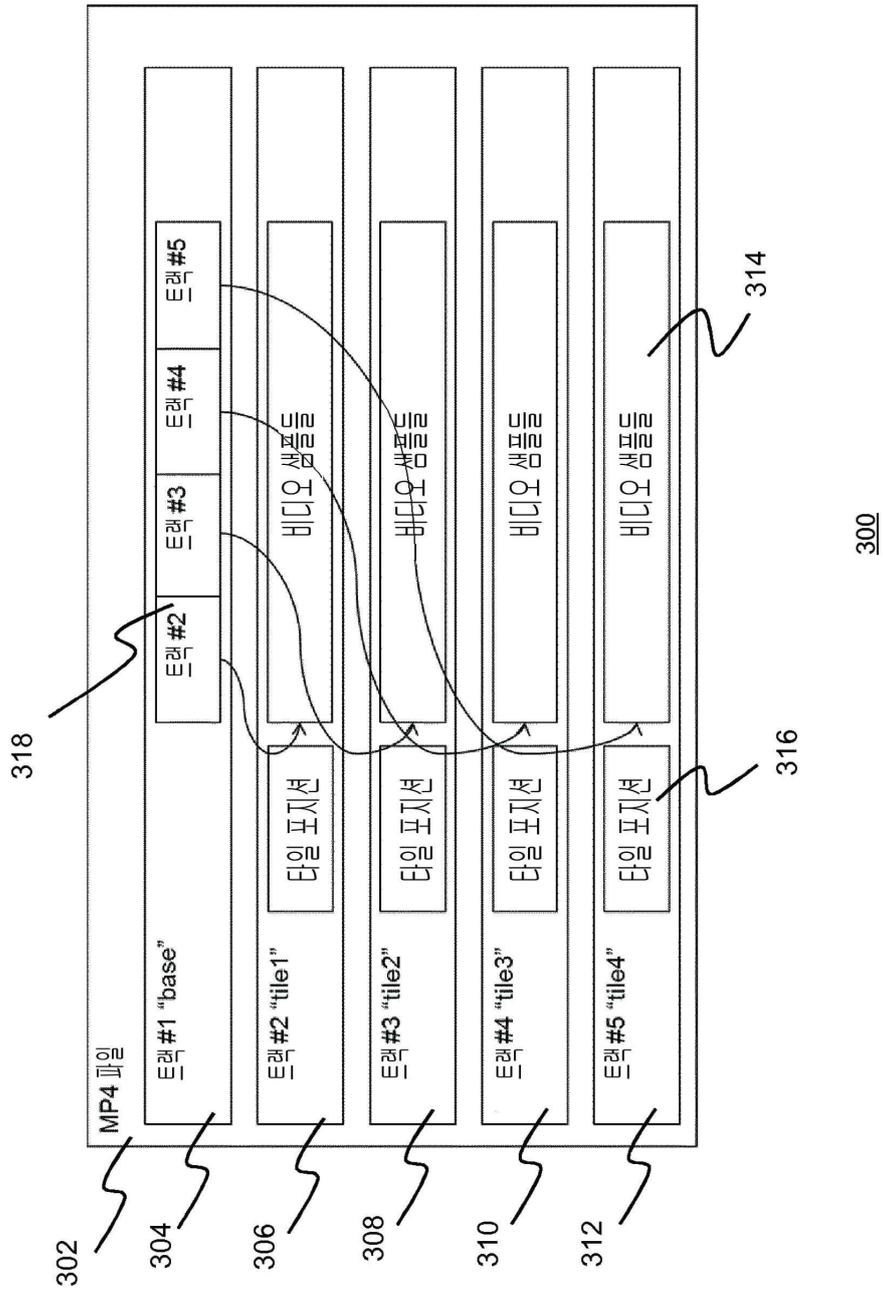
도면1



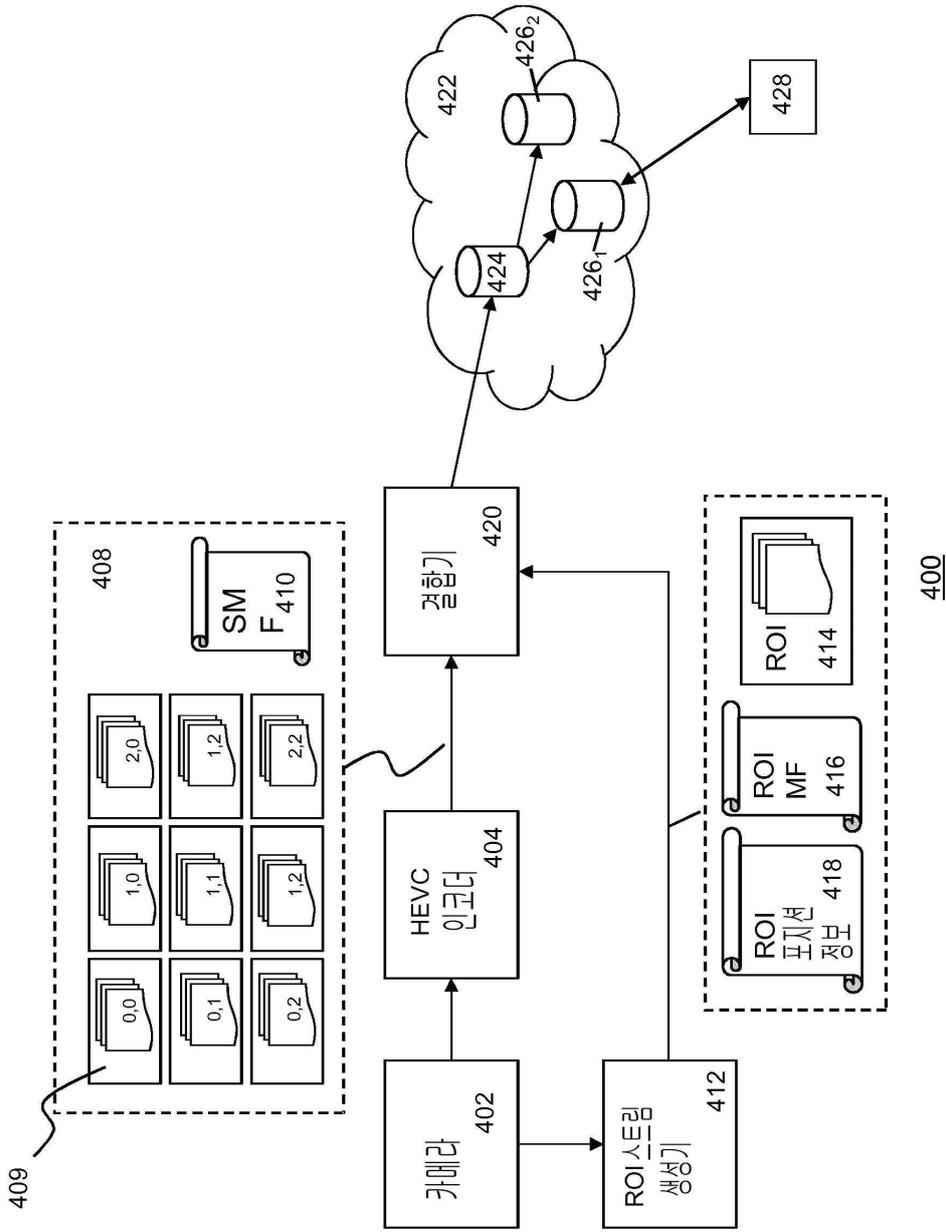
도면2



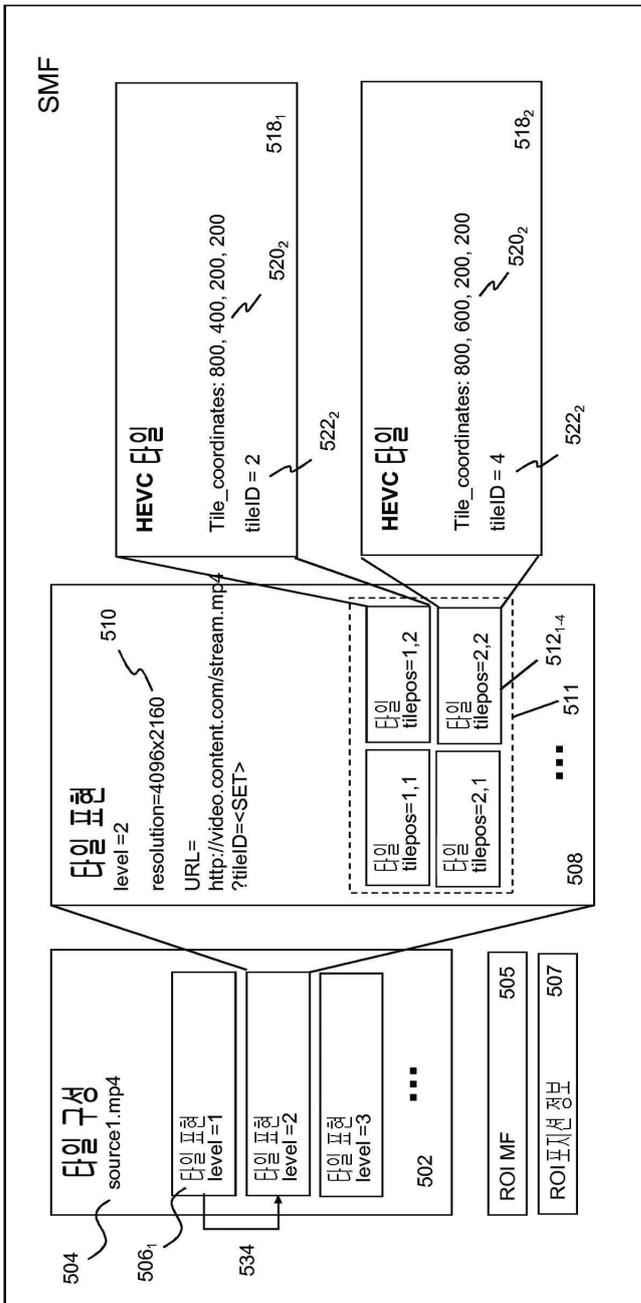
도면3



도면4

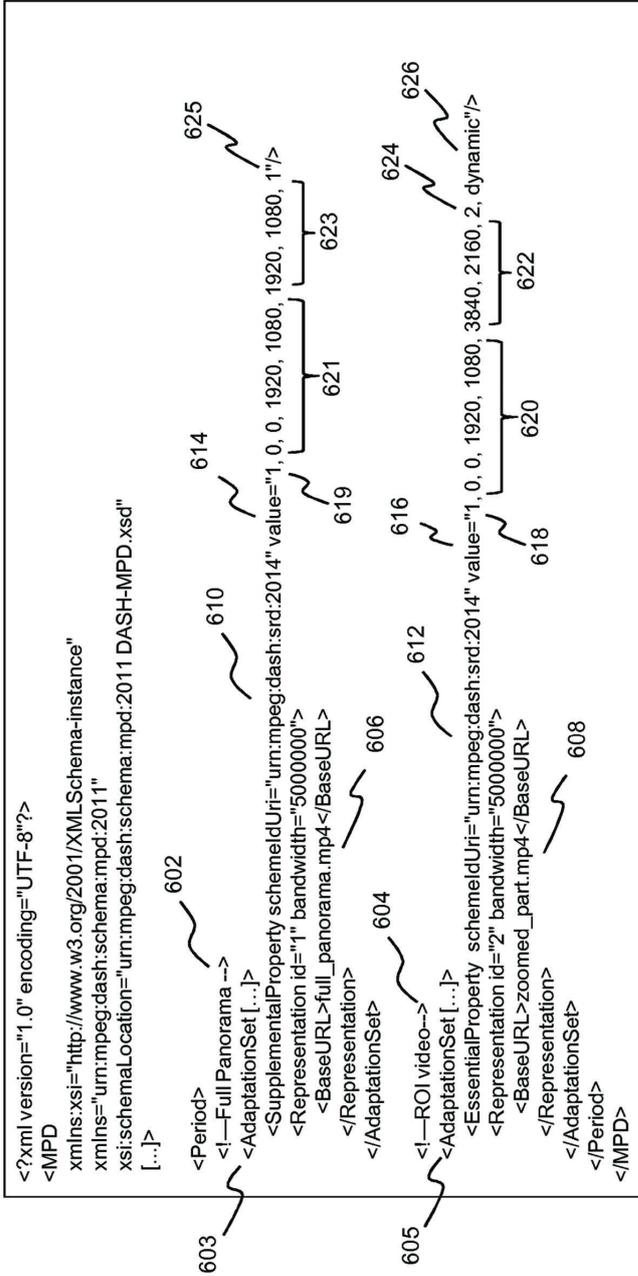


도면5



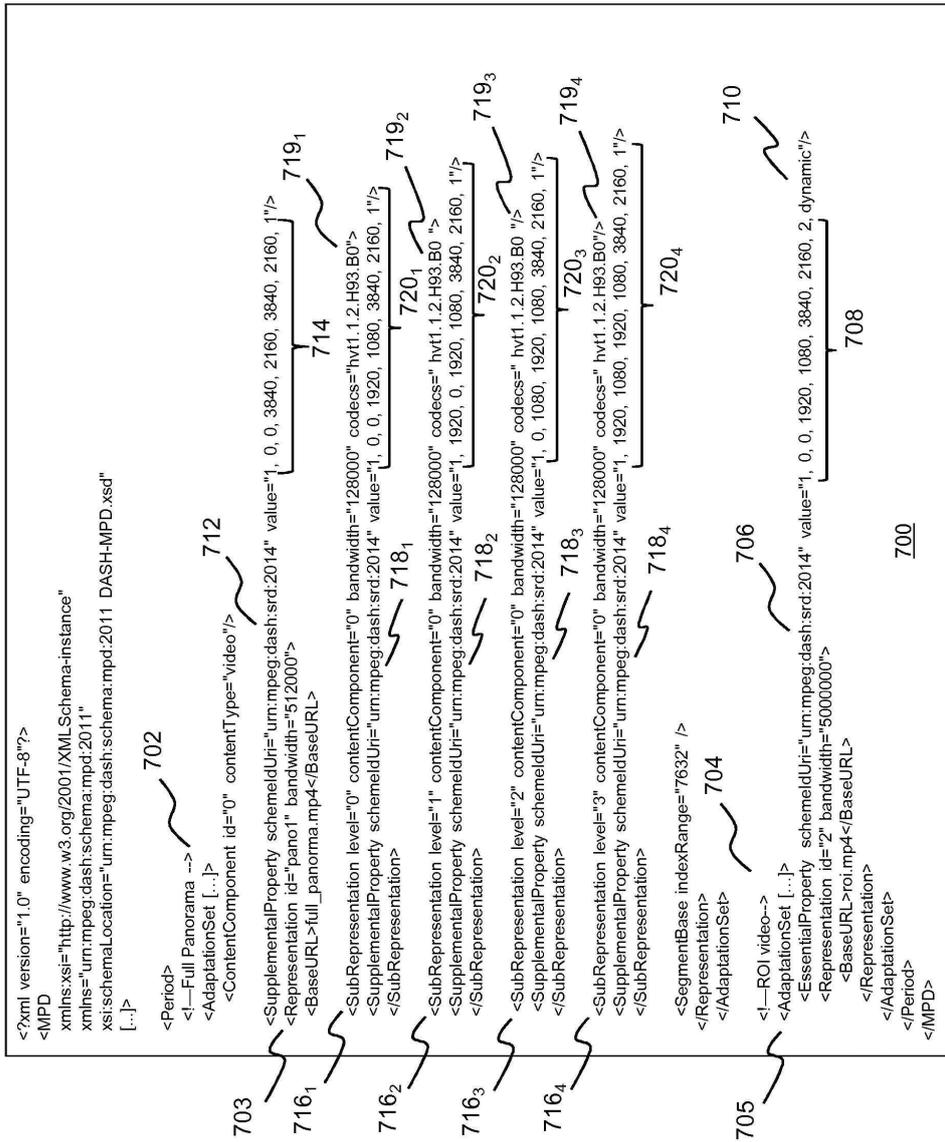
500

도면6

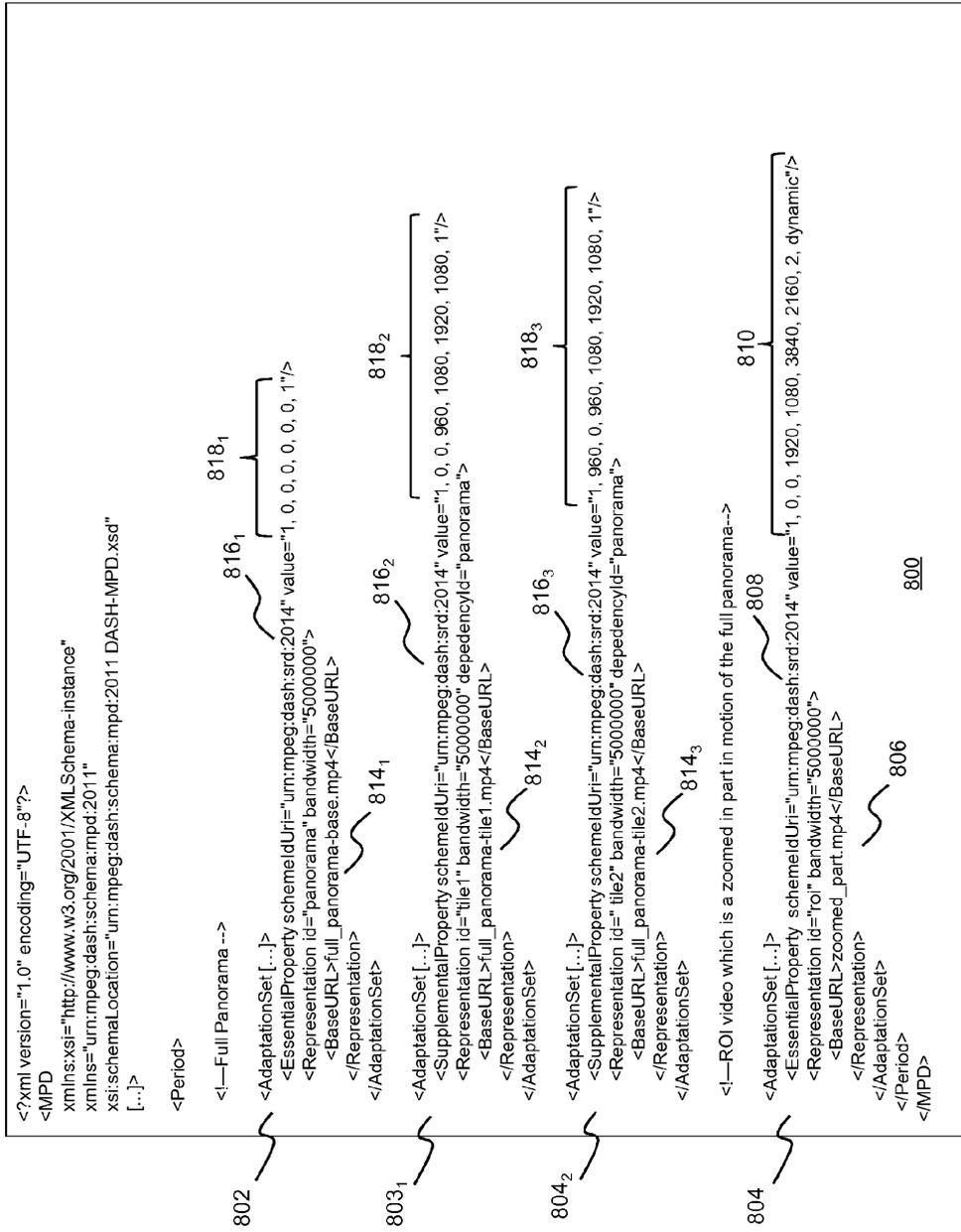


600

도면7



도면8a

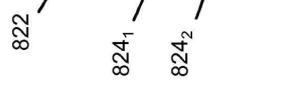


도면8b

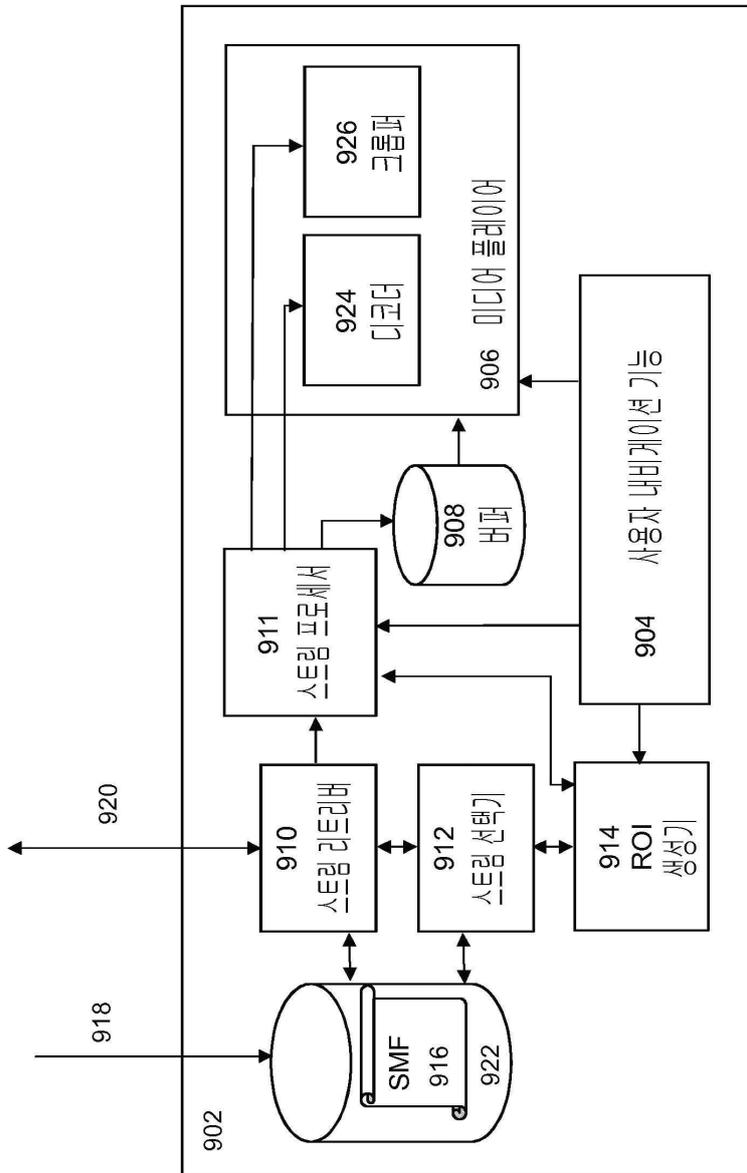
```

<?xml version="1.0" encoding="UTF-8"?>
<MPD
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xmlns="urn:mpeg:dash:schema:mpd:2011"
  xsi:schemaLocation="urn:mpeg:dash:schema:mpd:2011 DASH-MPD.xsd"
  ...>
  <Period>
    <!-- Full Panorama -->
    <AdaptationSet [...]>
      <EssentialProperty schemeIdUri="urn:mpeg:dash:srd:2014" value="1, 0, 0, 0, 0, 0, 1"/>
      <Representation id="panorama" bandwidth="5000000">
        <BaseURL>full_p Panorama-base.mp4</BaseURL>
      </Representation>
    </AdaptationSet>
    <AdaptationSet [...]>
      <SupplementalProperty schemeIdUri="urn:mpeg:dash:srd:2014" value="1, 0, 0, 960, 1080, 1920, 1080, 1"/>
      <Representation id="tile 1" bandwidth="5000000" dependencyId="panorama">
        <BaseURL>full_p Panorama-tile1.mp4</BaseURL>
      </Representation>
    </AdaptationSet>
    <AdaptationSet [...]>
      <SupplementalProperty schemeIdUri="urn:mpeg:dash:srd:2014" value="1, 960, 0, 960, 1080, 1920, 1080, 1"/>
      <Representation id="tile 2" bandwidth="5000000" dependencyId="panorama">
        <BaseURL>full_p Panorama-tile2.mp4</BaseURL>
      </Representation>
    </AdaptationSet>
    <!-- ROI video which is a zoomed in part in motion of the full panorama-->
    <AdaptationSet [...]>
      <ContentComponent id="0" contentType="video"/>
      <ContentComponent id="1" contentType="application"/>
      <EssentialProperty schemeIdUri="urn:mpeg:dash:srd:2014" value="1, 0, 0, 0, 0, 0, 2, dynamic"/>
      <Representation id="roi" bandwidth="5000000">
        <BaseURL>zoomed_part.mp4</BaseURL>
      </Representation>
      <SubRepresentation level="0" contentComponent="0" bandwidth="128000" codecs="hvc1.1.2.L93.B0">
        <SupplementalProperty schemeIdUri="urn:mpeg:dash:srd:2014" value="1, 0, 0, 1920, 1080, 3840, 2160, 1"/>
      </SubRepresentation>
      <SubRepresentation level="1" contentComponent="1" bandwidth="100" codecs="aac">
        </SubRepresentation>
      </AdaptationSet>
    </Period>
  </MPD>

```

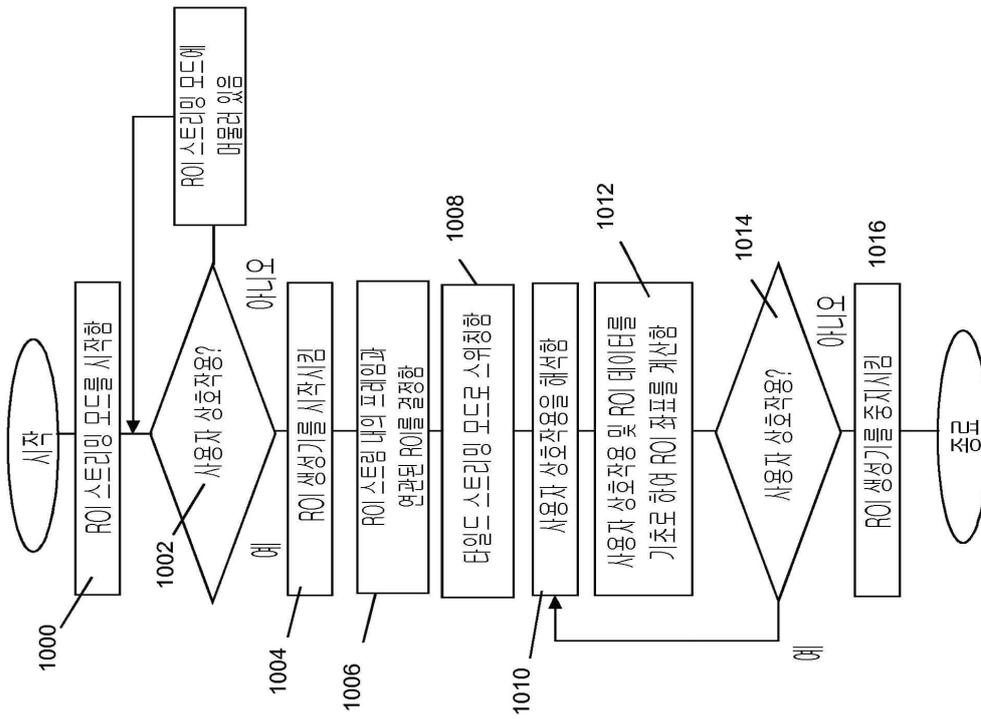


도면9

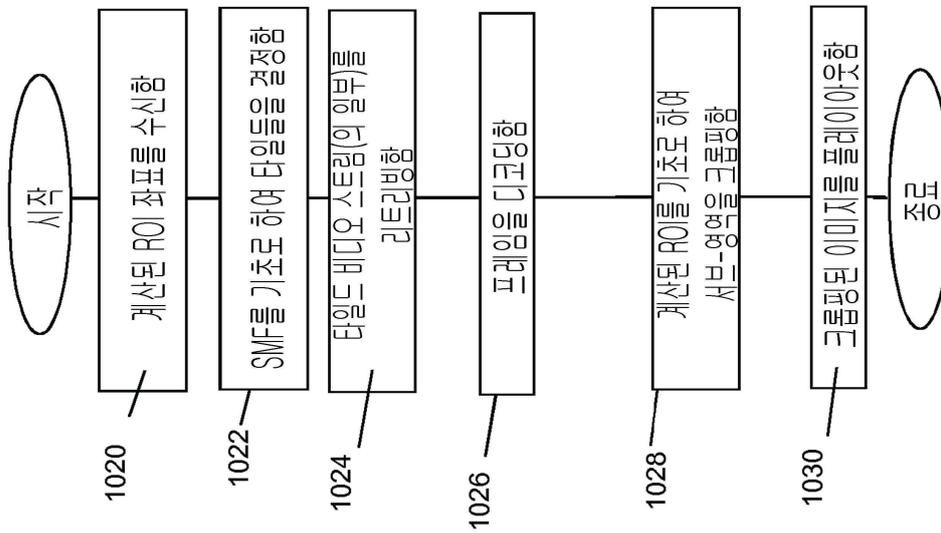


900

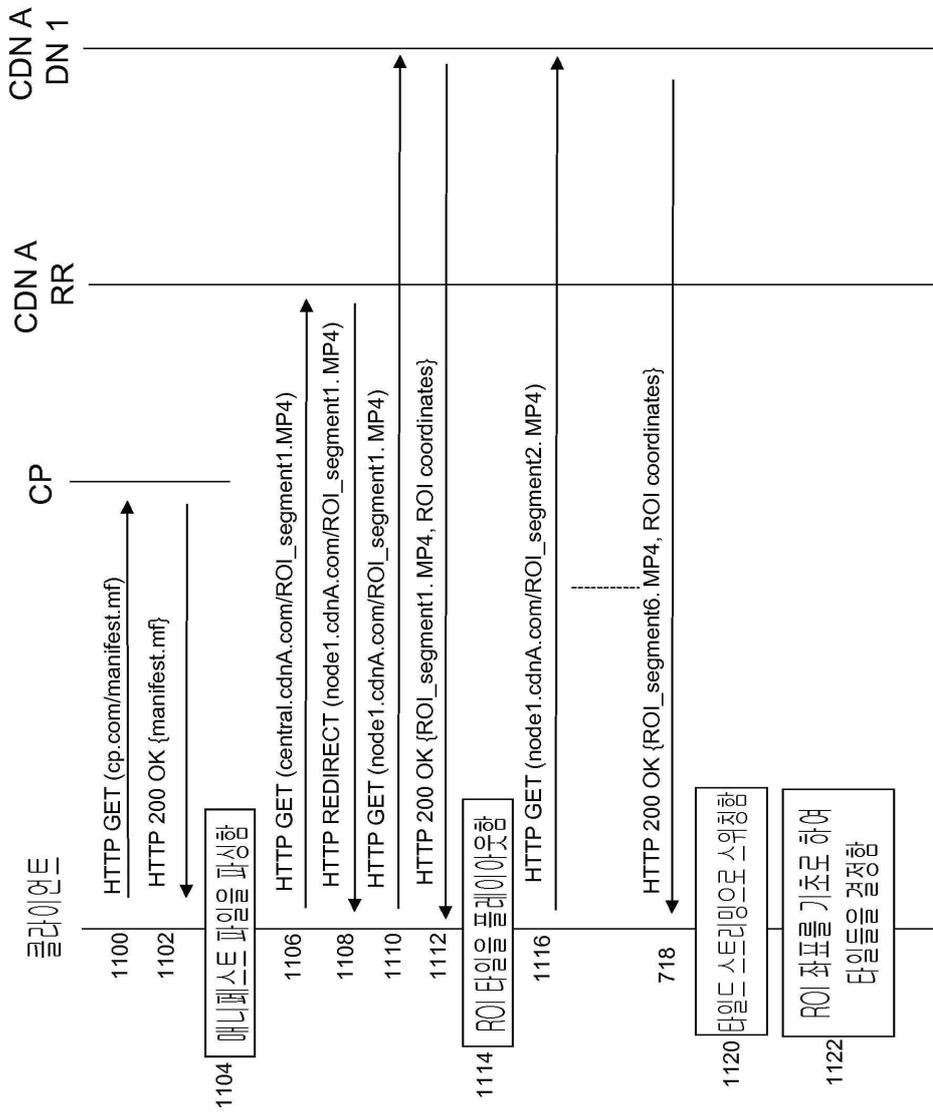
도면10a



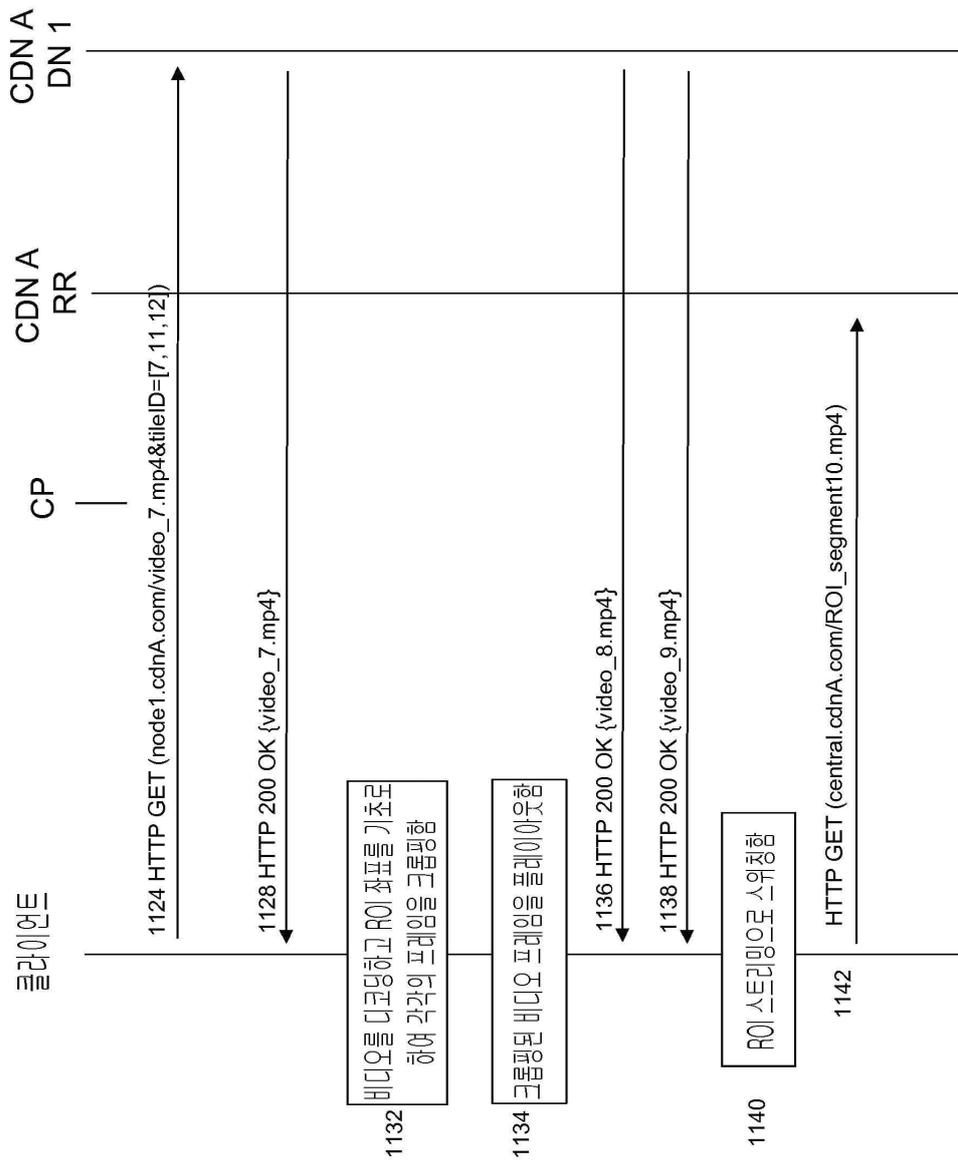
도면10b



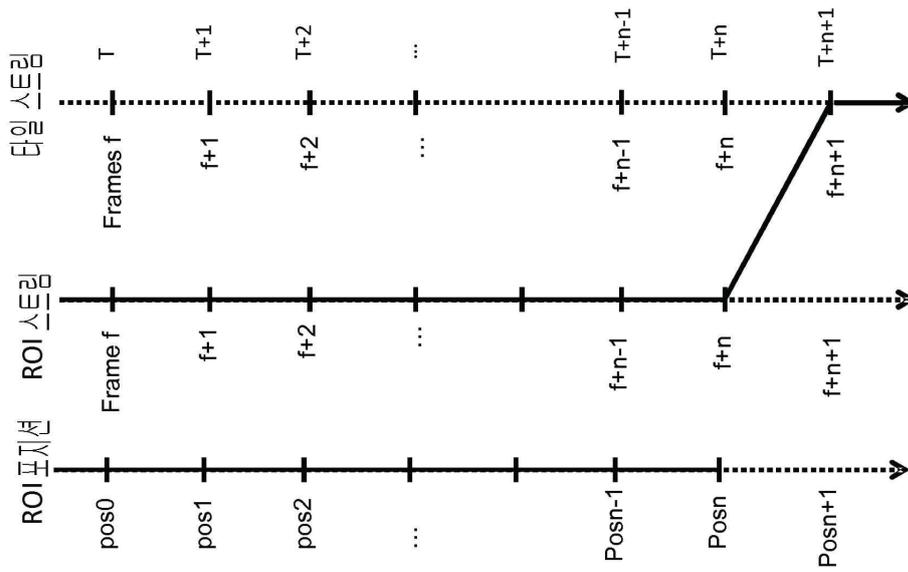
도면11a



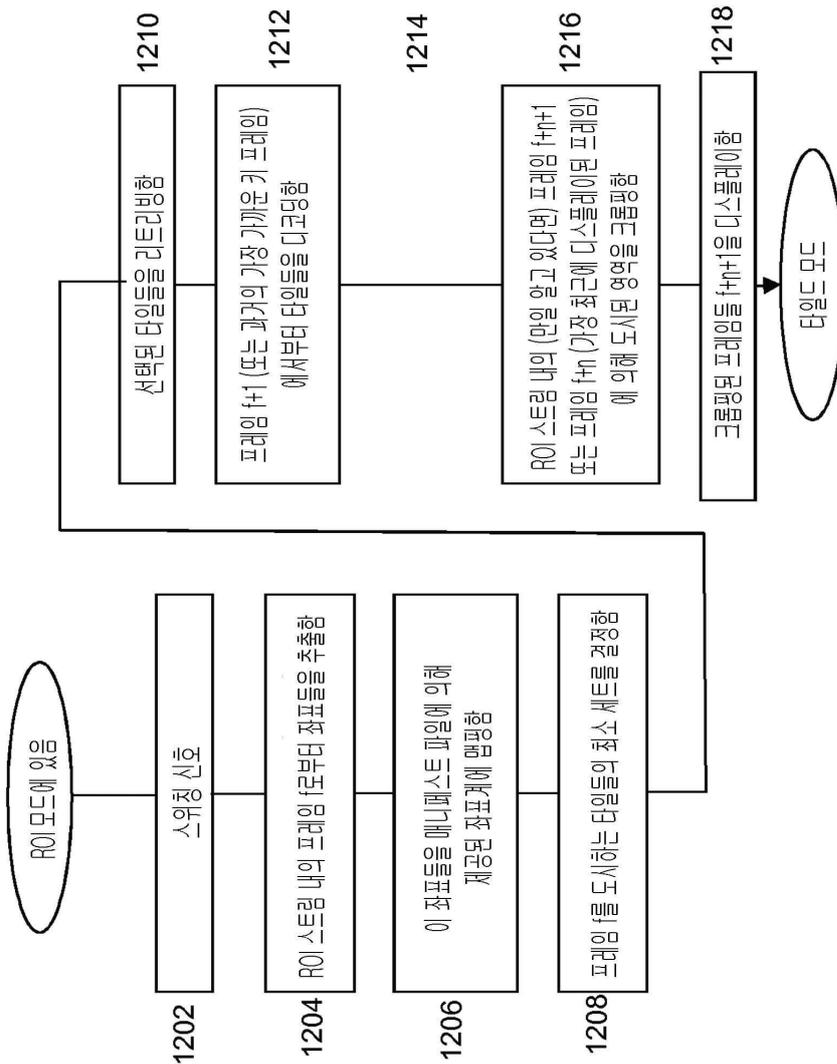
도면11b



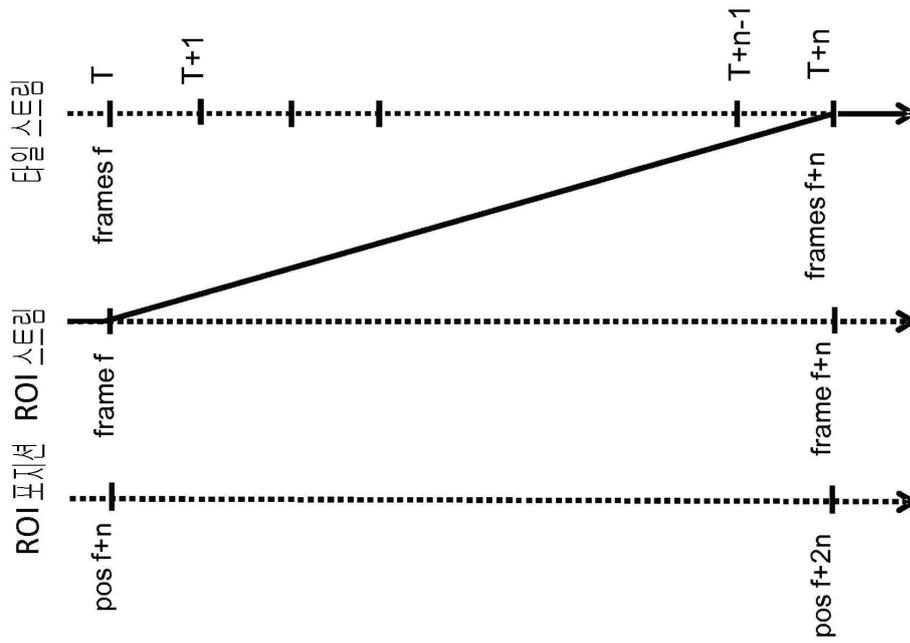
도면12a



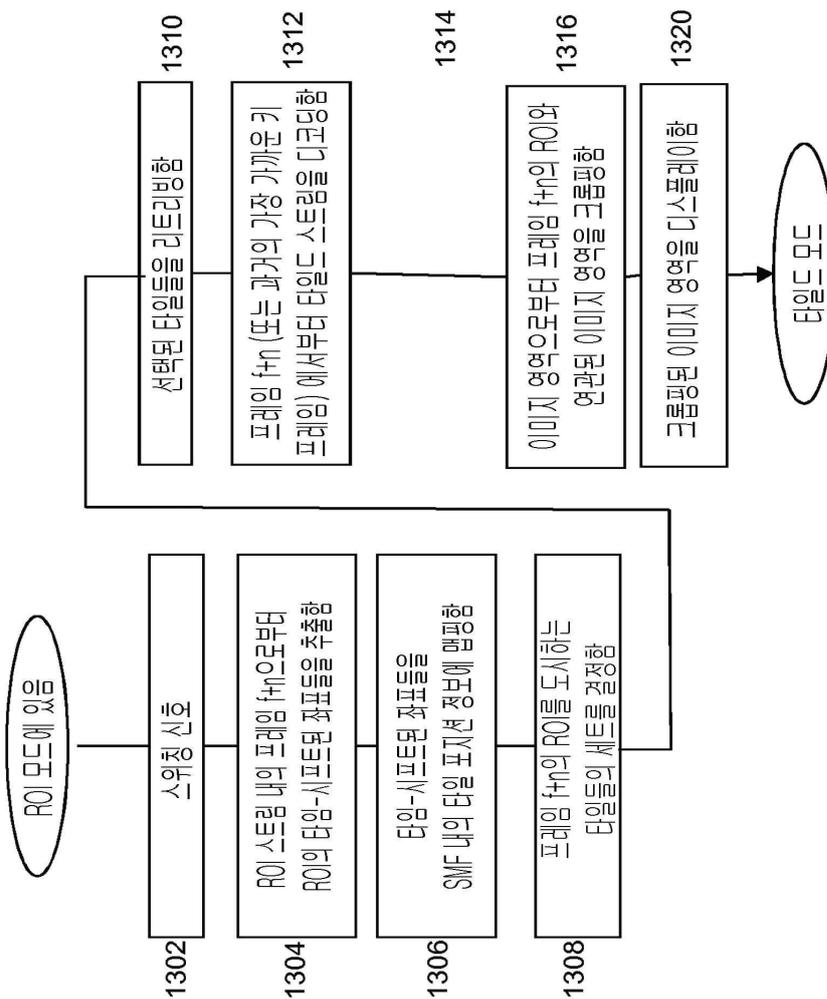
도면12b



도면13a



도면13b



도면14

```

<ROI>
  <RelativeTime 0/>
  <NormalizedXYZzoom 0,0,1/>
</ROI>
<ROI>
  <RelativeTime 1.7523/>
  <NormalizedXYZzoom 0,0,0.5/>
</ROI>
<ROI>
  <RelativeTime 5.89654/>
  <NormalizedXYZzoom 0.5,0.5,0.5/>
</ROI>
    
```

(B)

```

<ROI>
  <AbsoluteTime 1367584123.964324/>
  <PixelX1X2Y1Y2 0,1920,0,1080/>
</ROI>
<ROI>
  <AbsoluteTime 1367584125.516395/>
  <PixelX1X2Y1Y2 192,1728,108,972/>
</ROI>
<ROI>
  <AbsoluteTime 1367584127.234736/>
  <PixelX1X2Y1Y2 212,1748,108,972/>
</ROI>
    
```

(A)

```

<FrameNumber 1752/>
<PixelXYWH 50,70,640,480/>
<RadialNormalizedXYR 0.43,0.85,0.23/>
<SpeedPixelXdYdWdH> 2.3,0,-0.01,-0.01/>
    
```

(C)

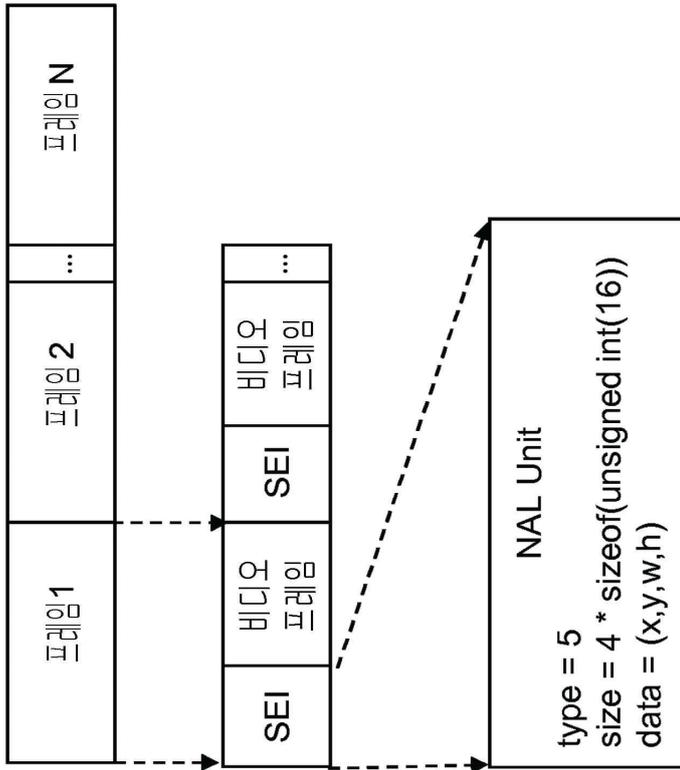
도면15

```

<?xml version="1.0" encoding="iso-8859-1" ?>
<Pmt>
  <Pid>0x1B62</Pid>
  <TableId>0x2</TableId>
  <SectionSyntaxIndicator>0x1</SectionSyntaxIndicator>
  <ProgramNumber>0x44D</ProgramNumber>
  <VersionNumber>0x1</VersionNumber>
  <CurrentNextIndicator>0x1</CurrentNextIndicator>
  <SectionNumber>0x0</SectionNumber>
  <LastSectionNumber>0x0</LastSectionNumber>
  <PcrPid>0x1B63</PcrPid>
  <Stream>
    <StreamType>0x2</StreamType>
    <ElementaryPid>0x1B63</ElementaryPid>
  </Stream>
  <Stream>
    <StreamType>0x3</StreamType>
    <ElementaryPid>0x1B64</ElementaryPid>
  </Stream>
  <Stream>
    <StreamType>0x20</StreamType>
    <ElementaryPid>0x1B66</ElementaryPid>
  </Stream>
</Pmt>

```

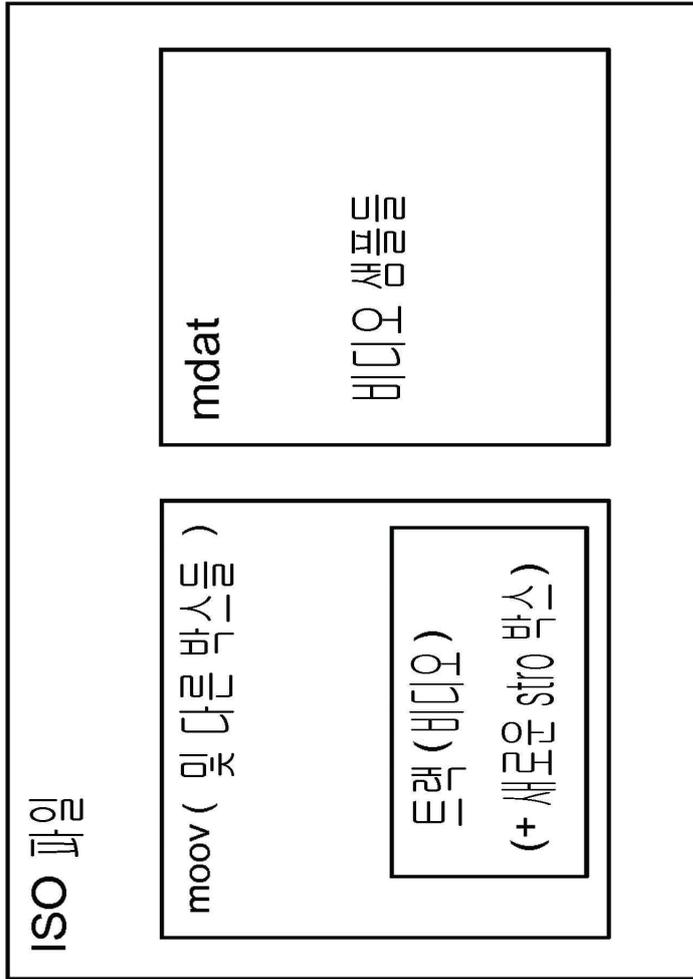
도면16a



도면16b

시맨틱스 (Semantics)	
if(roi_coordinates_flag) {	
roi_top_left_x	u(32)
roi_top_left_y	u(32)
roi_width	u(32)
roi_height	u(32)
}	

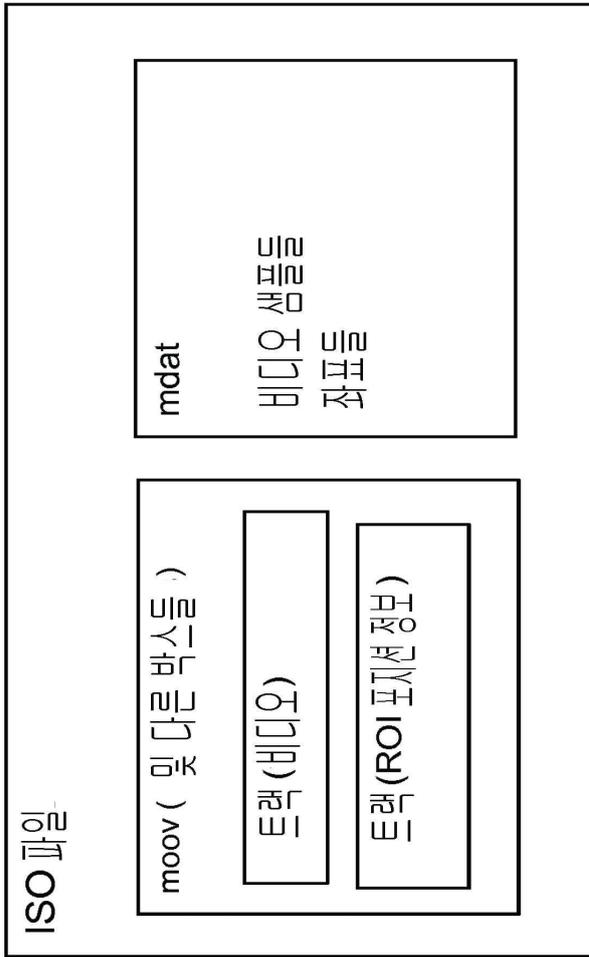
도면17a



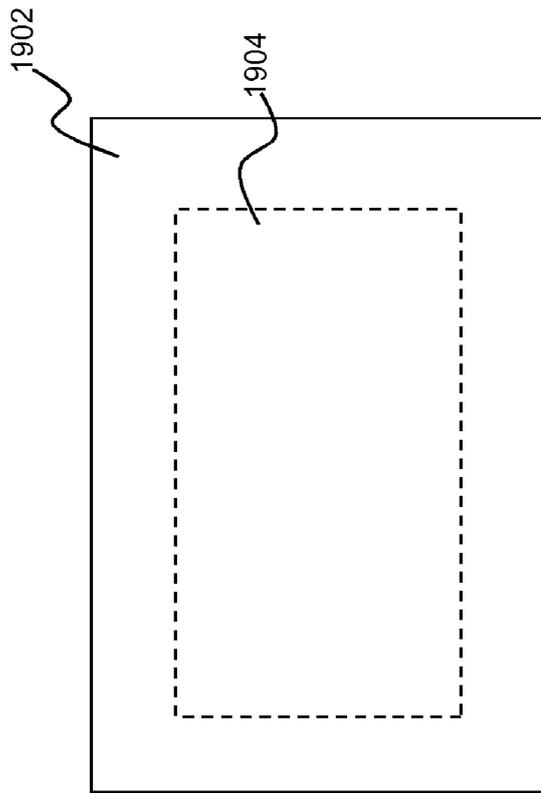
도면17b

aligned(8) class SampleROIcoordinatesBox extends FullBox('stro', version = 0, 0)
{
unsigned int(32) sample_count;
for (i=0; i < sample_count; i++)
{
unsigned int(32) roi_top_left_x;
unsigned int(32) roi_top_left_y;
unsigned int(32) roi_width;
unsigned int(32) roi_height;
}
}

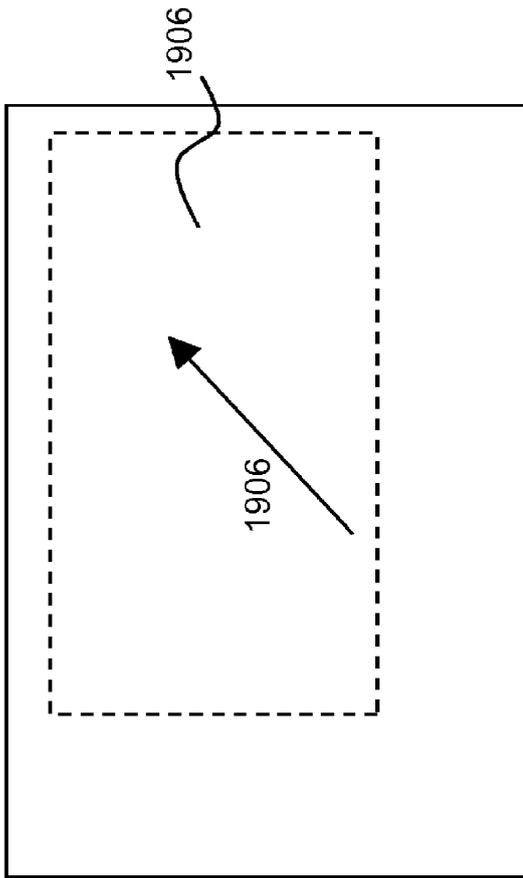
도면18



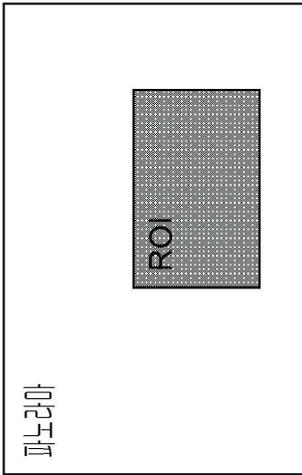
도면19a



도면19b



도면20



모드 3: ROI 스트리밍

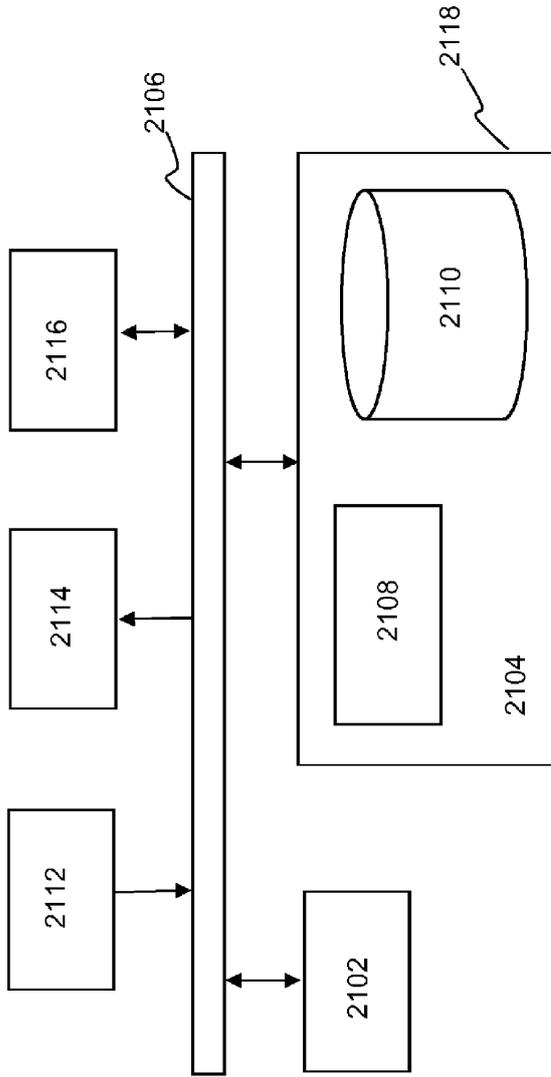
0,0	0,1	0,2
1,0	1,1	1,2
2,0	2,1	

모드1: HEVC-타일드 스트리밍 +
사용자-생성 ROI

0,0	0,1	0,2
1,0	1,1	1,2
2,0	2,1	

모드 2: HEVC-타일드 스트리밍 +
스트리밍된 ROI 좌표들

도면21



1000