



**(19) 대한민국특허청(KR)**  
**(12) 공개특허공보(A)**

(11) 공개번호 10-2020-0032230  
(43) 공개일자 2020년03월25일

- |   |  |
|---|--|
| <p>(51) 국제특허분류(Int. Cl.)<br/> <i>H04N 13/111</i> (2018.01) <i>H04N 13/161</i> (2018.01)<br/> <i>H04N 13/178</i> (2018.01) <i>H04N 13/243</i> (2018.01)<br/> <i>H04N 13/268</i> (2018.01) <i>H04N 13/282</i> (2018.01)<br/> <i>H04N 5/247</i> (2006.01)</p> <p>(52) CPC특허분류<br/> <i>H04N 13/111</i> (2018.05)<br/> <i>H04N 13/161</i> (2018.05)</p> <p>(21) 출원번호 10-2020-7006903</p> <p>(22) 출원일자(국제) 2018년06월08일<br/>                 심사청구일자 2020년03월09일</p> <p>(85) 번역문제출일자 2020년03월09일</p> <p>(86) 국제출원번호 PCT/KR2018/006535</p> <p>(87) 국제공개번호 WO 2019/066191<br/>                 국제공개일자 2019년04월04일</p> <p>(30) 우선권주장<br/>                 62/564,283 2017년09월28일 미국(US)<br/>                 62/576,051 2017년10월23일 미국(US)</p> | <p>(71) 출원인<br/> <b>엘지전자 주식회사</b><br/>                 서울특별시 영등포구 여의대로 128 (여의도동)</p> <p>(72) 발명자<br/> <b>이수연</b><br/>                 서울시 서초구 양재대로11길 19 LG전자 특허센터</p> <p><b>오세진</b><br/>                 서울시 서초구 양재대로11길 19 LG전자 특허센터</p> <p>(74) 대리인<br/> <b>인비전 특허법인</b></p> |
|---|--|

전체 청구항 수 : 총 20 항

**(54) 발명의 명칭 스티칭 및 리프로젝션 관련 메타데이터를 이용한 6DOF 비디오를 송수신하는 방법 및 그 장치**

**(57) 요약**

본 발명에 따른 6DoF 비디오 전송 장치에 의하여 수행되는 6DoF 비디오 데이터 처리 방법은 6DoF 비디오의 픽처들에 대한 정보 및 상기 6DoF 비디오에 대한 메타데이터를 획득하는 단계, 상기 메타데이터 및 상기 픽처들에 대한 정보를 기반으로 상기 픽처들을 디코딩하고, 사용자의 위치에 대한 이미지를 도출하는 단계 및 상기 메타데이터를 기반으로 상기 이미지를 처리하여 상기 사용자의 위치를 중심으로 하는 3D 공간으로 렌더링하는 단계를 포함하되, 상기 메타데이터는 6DoF 비디오를 캡처한 카메라 정보 및 스티칭(stiching) 정보를 포함하는 것을 특징으로 한다.

(52) CPC특허분류

*HO4N 13/178* (2018.05)

*HO4N 13/243* (2018.05)

*HO4N 13/268* (2018.05)

*HO4N 13/282* (2018.05)

*HO4N 5/247* (2018.05)

---

## 명세서

### 청구범위

#### 청구항 1

6DoF(six Degrees of Freedom) 비디오 수신 장치에 의하여 수행되는 6DoF 비디오 데이터 처리 방법으로,

6DoF 비디오의 픽처들에 대한 정보 및 상기 6DoF 비디오에 대한 메타데이터를 획득하는 단계;

상기 메타데이터 및 상기 픽처들에 대한 정보를 기반으로 상기 픽처들을 디코딩하고, 사용자의 위치에 대한 이미지를 도출하는 단계; 및

상기 메타데이터를 기반으로 상기 이미지를 처리하여 상기 사용자의 위치를 중심으로 하는 3D 공간으로 렌더링하는 단계를 포함하되,

상기 메타데이터는 상기 6DoF 비디오를 캡처한 카메라 정보 및 스티칭(stitching) 정보를 포함하는 것으로 하는 6DoF 비디오 데이터 처리 방법.

#### 청구항 2

제1항에 있어서,

상기 메타데이터 및 상기 픽처들에 대한 정보를 기반으로 상기 픽처들을 디코딩하고, 사용자의 위치에 대한 이미지를 도출하는 단계는,

상기 카메라 정보 및 상기 스티칭 정보를 기반으로 적어도 하나의 스티칭 대상 뷰를 도출하는 단계;

상기 디코딩된 픽처들 중 상기 스티칭 대상 뷰에 대한 픽처들을 도출하는 단계; 및

상기 픽처들을 스티칭하여 상기 사용자의 위치에 대한 상기 이미지를 도출하는 단계를 더 포함하고,

상기 카메라 정보는 상기 6DoF 비디오에 대한 카메라 배열 정보, 및 상기 카메라로 캡처된 뷰들의 수를 나타내는 정보, 및 상기 뷰들의 위치 정보를 포함하고,

상기 스티칭 정보는 상기 스티칭 대상 뷰의 수를 나타내는 정보, 상기 스티칭 대상 뷰의 위치 정보를 포함하는 것을 특징으로 하는 6DoF 비디오 데이터 처리 방법.

#### 청구항 3

제2항에 있어서,

상기 스티칭 정보는 상기 이미지를 도출하기 위한 스티칭 과정의 수행 여부를 나타내는 플래그를 포함하는 것을 특징으로 하는 6DoF 비디오 데이터 처리 방법.

#### 청구항 4

제2항에 있어서,

카메라 배열 정보는 상기 6DoF 비디오를 캡처하는데 사용된 카메라 배열(camera array) 타입을 나타내고,

상기 카메라 배열 타입은 HDCA (High Density Camera Array), LDCA(Low Density Camera Array), 마이크로렌즈(microlens), 전방향 카메라(Omnidirectional Camera), 이동 카메라(Movement camera), 사용자 정의(user defined) 배열 중 하나인 것을 특징으로 하는 6DoF 비디오 데이터 처리 방법.

#### 청구항 5

제1항에 있어서,

상기 메타데이터는 상기 6DoF 비디오의 초기 뷰(initial view)에 대한 회전 정보 및 상기 6DoF 비디오의 초기 뷰의 위치 정보를 포함하는 것을 특징으로 하는 6DoF 비디오 데이터 처리 방법.

#### 청구항 6

제1항에 있어서,

렌더링 정보를 기반으로 상기 사용자의 위치에 대한 예비 렌더링 범위를 도출하는 단계; 및

상기 렌더링 정보를 기반으로 상기 예비 렌더링 범위 내 예비 렌더링 위치의 렌더링 타입을 도출하는 단계를 더 포함하고,

상기 메타데이터는 상기 렌더링 정보를 포함하고,

상기 렌더링 정보는 상기 예비 렌더링 범위의 타입을 나타내는 정보 및 상기 렌더링 타입을 나타내는 정보를 포함하는 것을 특징으로 하는 6DoF 비디오 데이터 처리 방법.

#### 청구항 7

제6항에 있어서,

상기 예비 렌더링 범위의 타입은 스피어(sphere) 타입, 탑-셰입(top-shape) 타입, 엘리베이션 디렉션(elevation direction) 타입, 레터럴 디렉션(lateral direction) 타입, 론지튜디널 디렉션(longitudinal direction) 타입 및 사용자 정의(user defined) 타입 중 하나인 것을 특징으로 하는 6DoF 비디오 데이터 처리 방법.

#### 청구항 8

제6항에 있어서,

상기 예비 렌더링 범위의 타입이 상기 스피어 타입인 경우,

상기 렌더링 정보는 상기 예비 렌더링 범위의 반경을 나타내는 정보를 포함하고, 상기 예비 렌더링 범위는 상기 반경을 갖고 상기 사용자의 위치를 중심으로 하는 스피어로 도출되는 것을 특징으로 하는 6DoF 비디오 데이터 처리 방법.

#### 청구항 9

제6항에 있어서,

상기 렌더링 타입은 홀 스피어(whole sphere) 렌더링 타입 및 파셜 뷰포트(patial viewport) 렌더링 타입 중 하나인 것을 특징으로 하는 6DoF 비디오 데이터 처리 방법.

#### 청구항 10

제9항에 있어서,

상기 렌더링 타입이 상기 파셜 뷰포트 렌더링 타입이고, 상기 사용자의 위치가 이동되는 경우,

상기 이동되는 위치의 이미지는 파셜 뷰포트에 렌더링되는 것을 특징으로 하는 6DoF 비디오 데이터 처리 방법.

#### 청구항 11

제10항에 있어서,

상기 렌더링 정보를 기반으로 상기 파셜 뷰포트가 도출되고,

상기 렌더링 정보는 상기 파셜 뷰포트의 중점의 요 값을 나타내는 정보, 상기 파셜 뷰포트의 중점의 피치 값을 나타내는 정보, 상기 파셜 뷰포트의 중점의 롤 값을 나타내는 정보, 상기 파셜 뷰포트의 수평 범위를 나타내는 정보 및 상기 파셜 뷰포트의 수직 범위를 나타내는 정보를 포함하는 것을 특징으로 하는 6DoF 비디오 데이터 처리 방법.

#### 청구항 12

6DoF 비디오 전송 장치에 의하여 수행되는 6DoF 비디오 데이터 처리 방법으로,

적어도 하나의 카메라에 의해 복수의 뷰들에 대한 6DoF 비디오를 획득하는 단계;

상기 6DoF 비디오를 처리하여 상기 뷰들에 대한 픽처들을 획득하는 단계;

상기 6DoF 비디오에 대한 메타데이터를 생성하는 단계;

상기 픽처들을 인코딩하는 단계; 및

상기 인코딩된 픽처들 및 상기 메타데이터에 대하여 저장 또는 전송을 위한 처리를 수행하는 단계를 포함하되, 상기 메타데이터는 상기 6DoF 비디오를 캡처한 카메라 정보 및 스티칭(stitching) 정보를 포함하는 것을 특징으로 하는 6DoF 비디오 데이터 처리 방법.

### 청구항 13

제12항에 있어서,

상기 6DoF 비디오를 처리하여 상기 뷰들에 대한 상기 픽처들을 획득하는 단계는,

적어도 하나의 스티칭 대상 뷰를 도출하는 단계;

상기 스티칭 대상 뷰에 대한 캡처드 씬(captured scene)들을 도출하는 단계;

상기 캡처드 씬들을 스티칭하고, 스티칭된 상기 캡처드 씬들을 처리하여 상기 스티칭 대상 뷰에 대한 픽처를 도출하는 단계; 및

상기 카메라 정보 및 상기 스티칭 정보를 생성하는 단계를 더 포함하고,

상기 카메라 정보는 상기 6DoF 비디오에 대한 카메라 배열 정보, 및 상기 카메라로 캡처된 뷰들의 수를 나타내는 정보, 및 상기 뷰들의 위치 정보를 포함하고,

상기 스티칭 정보는 상기 스티칭 대상 뷰의 수를 나타내는 정보, 상기 스티칭 대상 뷰의 위치 정보를 포함하는 것을 특징으로 하는 6DoF 비디오 데이터 처리 방법.

### 청구항 14

제13항에 있어서,

상기 스티칭 정보는 상기 스티칭 대상 뷰의 위치와 상기 캡처드 씬들 간의 거리를 나타내는 정보 및 상기 캡처드 씬들의 수를 나타내는 정보를 더 포함하는 것을 특징으로 하는 6DoF 비디오 데이터 처리 방법.

### 청구항 15

제13항에 있어서,

상기 스티칭 정보는 상기 스티칭 대상 뷰에 대한 상기 픽처를 도출하기 위한 스티칭 과정의 수행 여부를 나타내는 플래그를 포함하는 것을 특징으로 하는 6DoF 비디오 데이터 처리 방법.

### 청구항 16

제13항에 있어서,

카메라 배열 정보는 상기 6DoF 비디오를 캡처하는데 사용된 카메라 배열(camera array) 타입을 나타내고,

상기 카메라 배열 타입은 HDCA (High Density Camera Array), LDCA(Low Density Camera Array), 마이크로렌즈(microlens), 전방향 카메라(Omnidirectional Camera), 이동 카메라(Movement camera), 사용자 정의(user defined) 배열 중 하나인 것을 특징으로 하는 6DoF 비디오 데이터 처리 방법.

### 청구항 17

제12항에 있어서,

상기 메타데이터는 6DoF 비디오의 초기 뷰(initial view)에 대한 회전 정보 및 상기 6DoF 비디오의 초기 뷰의 위치 정보를 포함하는 것을 특징으로 하는 6DoF 비디오 데이터 처리 방법.

### 청구항 18

제12항에 있어서,

사용자의 위치에 대한 예비 렌더링 범위를 도출하는 단계; 및

상기 예비 렌더링 범위에 대한 렌더링 정보를 생성하는 단계를 더 포함하고,

상기 메타데이터는 상기 렌더링 정보를 포함하고,

상기 렌더링 정보는 상기 예비 렌더링 범위의 타입을 나타내는 정보 및 상기 렌더링 타입을 나타내는 정보를 포함하는 것을 특징으로 하는 6DoF 비디오 데이터 처리 방법.

**청구항 19**

제18항에 있어서,

상기 예비 렌더링 범위의 타입은 스피어(sphere) 타입, 탑-셰입(top-shape) 타입, 엘리베이션 디렉션(elevation direction) 타입, 레터럴 디렉션(lateral direction) 타입, 론지튜디널 디렉션(longitudinal direction) 타입 및 사용자 정의(user defined) 타입 중 하나인 것을 특징으로 하는 6DoF 비디오 데이터 처리 방법.

**청구항 20**

제18항에 있어서,

상기 렌더링 타입은 홀 스피어(whole sphere) 렌더링 타입 및 파셜 뷰포트(patial viewport) 렌더링 타입 중 하나인 것을 특징으로 하는 6DoF 비디오 데이터 처리 방법.

**발명의 설명**

**기술 분야**

[0001] 본 발명은 6DoF 비디오에 관한 것으로, 보다 상세하게는 스티칭 및 리프로젝션 관련 메타데이터를 이용하는 6DoF 비디오를 송수신하는 방법 및 장치에 관한 것이다.

**배경 기술**

[0002] VR(Virtual Reality) 시스템은 사용자에게 전자적으로 투영된 환경 내에 있는 것 같은 감각을 제공한다. VR 을 제공하기 위한 시스템은 더 고화질의 이미지들과, 공간적인 음향을 제공하기 위하여 더 개선될 수 있다. VR 시스템은 사용자가 인터랙티브하게 VR 콘텐츠들을 소비할 수 있도록 할 수 있다.

[0003] 6DoF(six Degrees of Freedom) 콘텐츠는 사용자의 이동에 따라서 새롭게 형성되는 3DoF 혹은 360도 비디오를 제공함으로써 더 다양한 감각적 경험을 소비할 수 있도록 할 수 있다.

**발명의 내용**

**해결하려는 과제**

[0004] 본 발명의 기술적 과제는 6DoF 시스템을 제공하기 위한 6DoF 비디오 데이터 전송의 효율을 높이는 방법 및 장치를 제공함에 있다.

[0005] 본 발명의 다른 기술적 과제는 6DoF 비디오 및 6DoF 비디오에 대한 메타데이터를 전송하는 방법 및 장치를 제공함에 있다.

[0006] 본 발명의 다른 기술적 과제는 6DoF 비디오 및 6DoF 비디오의 캡처 정보 메타데이터를 전송하는 방법 및 장치를 제공함에 있다.

[0007] 본 발명의 다른 기술적 과제는 6DoF 비디오 및 6DoF 비디오의 스티칭 정보 메타데이터를 전송하는 방법 및 장치를 제공함에 있다.

[0008] 본 발명의 다른 기술적 과제는 6DoF 비디오 및 6DoF 비디오의 렌더링 정보 메타데이터를 전송하는 방법 및 장치를 제공함에 있다.

**과제의 해결 수단**

[0009] 본 발명의 일 실시예에 따르면, 6DoF 비디오 전송 장치에 의하여 수행되는 6DoF 비디오 처리 방법이 제공된다.

상기 방법은 적어도 하나의 카메라에 의해 복수의 뷰들에 대한 6DoF 비디오를 획득하는 단계, 상기 6DoF 비디오를 처리하여 상기 뷰들에 대한 픽처들을 획득하는 단계, 상기 6DoF 비디오에 대한 메타데이터를 생성하는 단계, 상기 픽처들을 인코딩하는 단계; 및 상기 인코딩된 픽처들 및 상기 메타데이터에 대하여 저장 또는 전송을 위한 처리를 수행하는 단계를 포함하되, 상기 메타데이터는 6DoF 비디오를 캡처한 카메라 정보 및 스티칭(stiching) 정보를 포함하는 것을 특징으로 한다.

[0010] 본 발명의 다른 일 실시예에 따르면, 6DoF 비디오 데이터를 처리하는 6DoF 비디오 전송 장치가 제공된다. 상기 6DoF 비디오 전송 장치는 적어도 하나의 카메라에 의해 복수의 뷰들에 대한 6DoF 비디오를 획득하는 데이터 입력부, 상기 6DoF 비디오를 처리하여 상기 뷰들에 대한 픽처들을 획득하는 프로젝션 처리부, 상기 6DoF 비디오에 대한 메타데이터를 생성하는 메타데이터 처리부, 상기 픽처들을 인코딩하는 데이터 인코더, 및 상기 인코딩된 픽처들 및 상기 메타데이터에 대하여 저장 또는 전송을 위한 처리를 수행하는 전송 처리부를 포함하되, 상기 메타데이터는 6DoF 비디오를 캡처한 카메라 정보 및 스티칭(stiching) 정보를 포함하는 것을 특징으로 한다.

[0011] 본 발명의 또 다른 일 실시예에 따르면, 6DoF 비디오 수신 장치에 의하여 수행되는 6DoF 비디오 처리 방법이 제공된다. 상기 방법은 6DoF 비디오의 픽처들에 대한 정보 및 상기 6DoF 비디오에 대한 메타데이터를 획득하는 단계, 상기 메타데이터 및 상기 픽처들에 대한 정보를 기반으로 상기 픽처들을 디코딩하고, 사용자의 위치에 대한 이미지를 도출하는 단계, 상기 메타데이터를 기반으로 상기 이미지를 처리하여 상기 사용자의 위치를 중심으로 하는 3D 공간으로 렌더링하는 단계를 포함하되, 상기 메타데이터는 상기 6DoF 비디오를 캡처한 카메라 정보 및 스티칭(stiching) 정보를 포함하는 것을 특징으로 한다.

[0012] 본 발명의 다른 일 실시예에 따르면, 6DoF 비디오 데이터를 처리하는 6DoF 비디오 수신 장치가 제공된다. 6DoF 비디오의 픽처들에 대한 정보 및 상기 6DoF 비디오에 대한 메타데이터를 획득하는 수신 처리부, 상기 메타데이터 및 상기 픽처들에 대한 정보를 기반으로 상기 픽처들을 디코딩하고, 사용자의 위치에 대한 이미지를 도출하는 데이터 디코더, 및 상기 메타데이터를 기반으로 상기 이미지를 처리하여 상기 사용자의 위치를 중심으로 하는 3D 공간으로 렌더링하는 렌더러를 포함하되, 상기 메타데이터는 상기 6DoF 비디오를 캡처한 카메라 정보 및 스티칭(stiching) 정보를 포함하는 것을 특징으로 한다.

### 발명의 효과

[0013] 본 발명에 따르면 지상파 방송망과 인터넷 망을 사용하는 차세대 하이브리드 방송을 지원하는 환경에서 6DoF 콘텐츠를 효율적으로 전송할 수 있다.

[0014] 본 발명에 따르면 사용자의 6DoF 콘텐츠 소비에 있어서, 인터랙티브 경험(interactive experience)을 제공하기 위한 방안을 제안할 수 있다.

[0015] 본 발명에 따르면 사용자의 6DoF 콘텐츠 소비에 있어서, 6DoF 콘텐츠 제작자가 의도하는 바가 정확히 반영되도록 시그널링 하는 방안을 제안할 수 있다.

[0016] 본 발명에 따르면 6DoF 콘텐츠 전달에 있어, 효율적으로 전송 캐패시티를 늘리고, 필요한 정보가 전달될 수 있도록 하는 방안을 제안할 수 있다.

[0017] 본 발명에 따르면 특정 위치에 대한 스티칭 정보를 시그널링하여, 상기 특정 위치에 대한 캡처된 6DoF 비디오가 없는 경우에도 주변 위치의 6DoF 비디오를 기반으로 상기 특정 위치에 대한 6DoF 비디오를 생성 및 사용자에게 제공할 수 있는바, 6DoF 콘텐츠를 보다 효율적으로 제공할 수 있다.

[0018] 본 발명에 따르면 6DoF 콘텐츠에서 제공하는 이동 방향 또는 예측되는 사용자의 위치 이동 방향에 따른 예비 렌더링 정보를 시그널링하여, 사용자에게 제공되어야 하는 6DoF 콘텐츠가 우선적으로 전송 및 렌더링될 수 있는바, 6DoF 콘텐츠를 보다 효율적으로 제공할 수 있다.

### 도면의 간단한 설명

[0019] 도 1은 본 발명에 따른 360도 비디오 제공을 위한 전체 아키텍처를 도시한 도면이다.

도 2 및 도 3은 본 발명의 일 실시예에 따른 미디어 파일의 구조를 도시한 도면이다.

도 4는 DASH 기반 적응형 스트리밍 모델의 전반적인 동작의 일 예를 나타낸다.

도 5는 본 발명이 적용될 수 있는 360도 비디오 전송 장치의 구성을 개략적으로 설명하는 도면이다.

- 도 6은 본 발명이 적용될 수 있는 360도 비디오 수신 장치의 구성을 개략적으로 설명하는 도면이다.
- 도 7a 내지 도 7b는 360도 비디오 전송 장치/360도 비디오 수신 장치를 통하여 수행되는 360도 비디오 제공을 위한 전체 아키텍처를 예시적으로 나타낸다.
- 도 8은 본 발명의 3D 공간을 설명하기 위한 비행기 주축(Aircraft Principal Axes) 개념을 도시한 도면이다.
- 도 9a 내지 도 9b는 본 발명에 따른 프로젝션 스킴들을 예시적으로 나타낸다.
- 도 10a 내지 도 10b는 본 발명의 일 실시예에 따른 타일(Tile)이 도시된 도면을 예시적으로 나타낸다.
- 도 11은 본 발명의 일 실시예에 따른 360도 비디오 관련 메타데이터의 일 예를 나타낸다.
- 도 12는 상기 360도 콘텐츠와 상기 6DoF 콘텐츠를 예시적으로 나타낸다.
- 도 13은 본 발명에 따른 6DoF 비디오 제공을 위한 전체 아키텍처를 도시한 도면이다.
- 도 14는 본 발명이 적용될 수 있는 6DoF 비디오 전송 장치의 구성을 개략적으로 설명하는 도면이다.
- 도 15는 본 발명이 적용될 수 있는 6DoF 비디오 수신 장치의 구성을 개략적으로 설명하는 도면이다.
- 도 16은 6DoF 비디오 전송 장치/6DoF 비디오 수신 장치를 통하여 수행되는 6DoF 콘텐츠 제공을 위한 전체 아키텍처를 예시적으로 나타낸다.
- 도 17은 본 발명의 6DoF 공간을 설명하기 위하여 도시한 도면이다.
- 도 18은 6DoF 비디오 데이터를 획득하기 위하여 사용되는 카메라 배열을 예시적으로 나타낸다.
- 도 19a 내지 도 19e는 본 발명의 상기 예비 렌더링 범위의 형태를 설명하기 위하여 도시한 도면이다.
- 도 20은 본 발명의 상기 예비 렌더링 위치의 상기 렌더링 타입을 설명하기 위하여 도시한 도면이다.
- 도 21은 6DoF 비디오 수신 장치에서 카메라 정보(camera information) 메타데이터와 스티칭 메타데이터가 사용되는 동작을 나타낸다.
- 도 22는 6DoF 비디오 수신 장치에서 리-프로젝션 정보 메타데이터가 사용되는 동작을 나타낸다.
- 도 23은 본 발명에 따른 6DoF 비디오 전송 장치에 의한 6DoF 비디오 데이터 처리 방법을 개략적으로 나타낸다.
- 도 24는 본 발명에 따른 6DoF 비디오 수신 장치에 의한 6DoF 비디오 데이터 처리 방법을 개략적으로 나타낸다.

**발명을 실시하기 위한 구체적인 내용**

- [0020] 본 발명은 다양한 변경을 가할 수 있고 여러 가지 실시예를 가질 수 있는 바, 특정 실시예들을 도면에 예시하고 상세하게 설명하고자 한다. 그러나, 이는 본 발명을 특정 실시예에 한정하려고 하는 것이 아니다. 본 명세서에서 상용하는 용어는 단지 특정한 실시예를 설명하기 위해 사용된 것으로, 본 발명의 기술적 사상을 한정하려는 의도로 사용되는 것은 아니다. 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수의 표현을 포함한다. 본 명세서에서 "포함하다" 또는 "가지다" 등의 용어는 명세서상에 기재된 특징, 숫자, 단계, 동작, 구성 요소, 부품 또는 이들을 조합한 것이 존재함을 지정하려는 것이지, 하나 또는 그 이상의 다른 특징들이나 숫자, 단계, 동작, 구성 요소, 부품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 미리 배제하지 않는 것으로 이해되어야 한다.
- [0021] 한편, 본 발명에서 설명되는 도면상의 각 구성들은 서로 다른 특징적인 기능들에 관한 설명의 편의를 위해 독립적으로 도시된 것으로서, 각 구성들이 서로 별개의 하드웨어나 별개의 소프트웨어로 구현된다는 것을 의미하지는 않는다. 예컨대, 각 구성 중 두 개 이상의 구성이 합쳐져 하나의 구성을 이룰 수도 있고, 하나의 구성이 복수의 구성으로 나뉘어질 수도 있다. 각 구성이 통합 및/또는 분리된 실시예도 본 발명의 본질에서 벗어나지 않는 한 본 발명의 권리범위에 포함된다.
- [0022] 이하, 첨부한 도면들을 참조하여, 본 발명의 바람직한 실시예를 보다 상세하게 설명하고자 한다. 이하, 도면상의 동일한 구성 요소에 대해서는 동일한 참조 부호를 사용하고 동일한 구성 요소에 대해서 중복된 설명은 생략한다.
- [0023] 도 1은 본 발명에 따른 360도 비디오 제공을 위한 전체 아키텍처를 도시한 도면이다.



- [0024] 본 발명은 사용자에게 가상현실 (Virtual Reality, VR)을 제공하기 위하여, 360 콘텐츠를 제공하는 방안을 제안한다. 여기서, 상기 360도 콘텐츠는 3DoF(three Degrees of Freedom) 콘텐츠라고 나타낼 수도 있으며, VR이란 실제 또는 가상의 환경을 복제(replicates) 하기 위한 기술 내지는 그 환경을 의미할 수 있다. VR은 인공적으로 사용자에게 감각적 경험을 제공하며, 이를 통해 사용자는 전자적으로 프로젝션된 환경에 있는 것과 같은 경험을 할 수 있다.
- [0025] 360 콘텐츠는 VR을 구현, 제공하기 위한 콘텐츠 전반을 의미하며, 360도 비디오 및/또는 360 오디오를 포함할 수 있다. 360도 비디오는 VR을 제공하기 위해 필요한, 동시에 모든 방향(360도)으로 캡처되거나 재생되는 비디오 혹은 이미지 콘텐츠를 의미할 수 있다. 이하, 360도 비디오라 함은 360도 비디오를 의미할 수 있다. 360도 비디오는 3D 모델에 따라 다양한 형태의 3D 공간 상에 나타내어지는 비디오 혹은 이미지를 의미할 수 있으며, 예를 들어 360도 비디오는 구형면(Spherical surface) 상에 나타내어질 수 있다. 360 오디오 역시 VR을 제공하기 위한 오디오 콘텐츠로서, 음향 발생지가 3차원의 특정 공간상에 위치하는 것으로 인지될 수 있는, 공간적(Spatial) 오디오 콘텐츠를 의미할 수 있다. 360 콘텐츠는 생성, 처리되어 사용자들로 전송될 수 있으며, 사용자들은 360 콘텐츠를 이용하여 VR 경험을 소비할 수 있다.
- [0026] 본 발명은 특히 360도 비디오를 효과적으로 제공하는 방안을 제안한다. 360도 비디오를 제공하기 위하여, 먼저 하나 이상의 카메라를 통해 360도 비디오가 캡처될 수 있다. 캡처된 360도 비디오는 일련의 과정을 거쳐 전송되고, 수신측에서는 수신된 데이터를 다시 원래의 360도 비디오로 가공하여 렌더링할 수 있다. 이를 통해 360도 비디오가 사용자에게 제공될 수 있다.
- [0027] 구체적으로 360도 비디오 제공을 위한 전체의 과정은 캡처 과정(process), 준비 과정, 전송 과정, 프로세싱 과정, 렌더링 과정 및/또는 피드백 과정을 포함할 수 있다.
- [0028] 캡처 과정은 하나 이상의 카메라를 통하여 복수개의 시점 각각에 대한 이미지 또는 비디오를 캡처하는 과정을 의미할 수 있다. 캡처 과정에 의해 도시된 도 1의 (110)과 같은 이미지/비디오 데이터가 생성될 수 있다. 도시된 도 1의 (110)의 각 평면은 각 시점에 대한 이미지/비디오를 의미할 수 있다. 이 캡처된 복수개의 이미지/비디오를 로(raw) 데이터라 할 수도 있다. 캡처 과정에서 캡처와 관련된 메타데이터가 생성될 수 있다.
- [0029] 이 캡처를 위하여 VR 을 위한 특수한 카메라가 사용될 수 있다. 실시예에 따라 컴퓨터로 생성된 가상의 공간에 대한 360도 비디오를 제공하고자 하는 경우, 실제 카메라를 통한 캡처가 수행되지 않을 수 있다. 이 경우 단순히 관련 데이터가 생성되는 과정으로 해당 캡처 과정이 갈음될 수 있다.
- [0030] 준비 과정은 캡처된 이미지/비디오 및 캡처 과정에서 발생한 메타데이터를 처리하는 과정일 수 있다. 캡처된 이미지/비디오는 이 준비 과정에서, 스티칭 과정, 프로젝션 과정, 리전별 패킹 과정(Region-wise Packing) 및/또는 인코딩 과정 등을 거칠 수 있다.
- [0031] 먼저 각각의 이미지/비디오가 스티칭(Stitching) 과정을 거칠 수 있다. 스티칭 과정은 각각의 캡처된 이미지/비디오들을 연결하여 하나의 파노라마 이미지/비디오 또는 구형의 이미지/비디오를 만드는 과정일 수 있다.
- [0032] 이 후, 스티칭된 이미지/비디오는 프로젝션(Projection) 과정을 거칠 수 있다. 프로젝션 과정에서, 스티칭된 이미지/비디오는 2D 이미지 상에 프로젝션될 수 있다. 이 2D 이미지는 문맥에 따라 2D 이미지 프레임으로 불릴 수도 있다. 2D 이미지로 프로젝션하는 것을 2D 이미지로 맵핑한다고 표현할 수도 있다. 프로젝션된 이미지/비디오 데이터는 도시된 도 1의 (120)과 같은 2D 이미지의 형태가 될 수 있다.
- [0033] 2D 이미지 상에 프로젝션된 비디오 데이터는 비디오 코딩 효율 등을 높이기 위하여 리전별 패킹 과정(Region-wise Packing)을 거칠 수 있다. 리전별 패킹이란, 2D 이미지 상에 프로젝션된 비디오 데이터를 리전(Region) 별로 나누어 처리를 가하는 과정을 의미할 수 있다. 여기서 리전(Region)이란, 360도 비디오 데이터가 프로젝션된 2D 이미지가 나누어진 영역을 의미할 수 있다. 이 리전들은, 실시예에 따라, 2D 이미지를 균등하게 나누어 구분되거나, 임의로 나누어져 구분될 수 있다. 또한 실시예에 따라 리전들은, 프로젝션 스킴에 따라 구분될 수도 있다. 리전별 패킹 과정은 선택적(optional) 과정으로써, 준비 과정에서 생략될 수 있다.
- [0034] 실시예에 따라 이 처리 과정은, 비디오 코딩 효율을 높이기 위해, 각 리전을 회전한다거나 2D 이미지 상에서 재배열하는 과정을 포함할 수 있다. 예를 들어, 리전들을 회전하여 리전들의 특정 변들이 서로 근접하여 위치되도록 함으로써, 코딩 시의 효율이 높아지게 할 수 있다.
- [0035] 실시예에 따라 이 처리 과정은, 360도 비디오상의 영역별로 레졸루션(resolution) 을 차등화하기 위하여, 특정 리전에 대한 레졸루션을 높인다거나, 낮추는 과정을 포함할 수 있다. 예를 들어, 360도 비디오 상에서 상대적으로

로 더 중요한 영역에 해당하는 리전들은, 다른 리전들보다 레졸루션을 높게할 수 있다. 2D 이미지 상에 프로젝션된 비디오 데이터 또는 리전별 패키징된 비디오 데이터는 비디오 코덱을 통한 인코딩 과정을 거칠 수 있다.

- [0036] 실시예에 따라 준비 과정은 부가적으로 에디팅(editing) 과정 등을 더 포함할 수 있다. 이 에디팅 과정에서 프로젝션 전후의 이미지/비디오 데이터들에 대한 편집 등이 더 수행될 수 있다. 준비 과정에서도 마찬가지로, 스티칭/프로젝션/인코딩/에디팅 등에 대한 메타데이터가 생성될 수 있다. 또한 2D 이미지 상에 프로젝션된 비디오 데이터들의 초기 시점, 혹은 ROI (Region of Interest) 등에 관한 메타데이터가 생성될 수 있다.
- [0037] 전송 과정은 준비 과정을 거친 이미지/비디오 데이터 및 메타데이터들을 처리하여 전송하는 과정일 수 있다. 전송을 위해 임의의 전송 프로토콜에 따른 처리가 수행될 수 있다. 전송을 위한 처리를 마친 데이터들은 방송망 및/또는 브로드밴드를 통해 전달될 수 있다. 이 데이터들은 온 디맨드(On Demand) 방식으로 수신측으로 전달될 수도 있다. 수신측에서는 다양한 경로를 통해 해당 데이터를 수신할 수 있다.
- [0038] 프로세싱 과정은 수신한 데이터를 디코딩하고, 프로젝션되어 있는 이미지/비디오 데이터를 3D 모델 상에 리-프로젝션(Re-projection) 하는 과정을 의미할 수 있다. 이 과정에서 2D 이미지들 상에 프로젝션되어 있는 이미지/비디오 데이터가 3D 공간 상으로 리-프로젝션될 수 있다. 이 과정을 문맥에 따라 맵핑, 프로젝션이라고 부를 수도 있다. 이 때 맵핑되는 3D 공간은 3D 모델에 따라 다른 형태를 가질 수 있다. 예를 들어 3D 모델에는 구형(Sphere), 큐브(Cube), 실린더(Cylinder) 또는 피라미드(Pyramid) 가 있을 수 있다.
- [0039] 실시예에 따라 프로세싱 과정은 부가적으로 에디팅(editing) 과정, 업 스케일링(up scaling) 과정 등을 더 포함할 수 있다. 이 에디팅 과정에서 리-프로젝션 전후의 이미지/비디오 데이터에 대한 편집 등이 더 수행될 수 있다. 이미지/비디오 데이터가 축소되어 있는 경우 업 스케일링 과정에서 샘플들의 업 스케일링을 통해 그 크기를 확대할 수 있다. 필요한 경우, 다운 스케일링을 통해 사이즈를 축소하는 작업이 수행될 수도 있다.
- [0040] 렌더링 과정은 3D 공간상에 리-프로젝션된 이미지/비디오 데이터를 렌더링하고 디스플레이하는 과정을 의미할 수 있다. 표현에 따라 리-프로젝션과 렌더링을 합쳐 3D 모델 상에 렌더링한다 라고 표현할 수도 있다. 3D 모델 상에 리-프로젝션된 (또는 3D 모델 상으로 렌더링된) 이미지/비디오는 도시된 도 1의 (130)과 같은 형태를 가질 수 있다. 도시된 도 1의 (130)은 구형(Sphere) 의 3D 모델에 리-프로젝션된 경우이다. 사용자는 VR 디스플레이 등을 통하여 렌더링된 이미지/비디오의 일부 영역을 볼 수 있다. 이 때 사용자가 보게되는 영역은 도시된 도 1의 (140)과 같은 형태일 수 있다.
- [0041] 피드백 과정은 디스플레이 과정에서 획득될 수 있는 다양한 피드백 정보들을 송신측으로 전달하는 과정을 의미할 수 있다. 피드백 과정을 통해 360도 비디오 소비에 있어 인터랙티비티(Interactivity) 가 제공될 수 있다. 실시예에 따라, 피드백 과정에서 헤드 오리엔테이션(Head Orientation) 정보, 사용자가 현재 보고 있는 영역을 나타내는 뷰포트(Viewport) 정보 등이 송신측으로 전달될 수 있다. 실시예에 따라, 사용자는 VR 환경 상에 구현된 것들과 상호작용할 수도 있는데, 이 경우 그 상호작용과 관련된 정보가 피드백 과정에서 송신측 내지 서비스 프로바이더 측으로 전달될 수도 있다. 실시예에 따라 피드백 과정은 수행되지 않을 수도 있다.
- [0042] 헤드 오리엔테이션 정보는 사용자의 머리 위치, 각도, 움직임 등에 대한 정보를 의미할 수 있다. 이 정보를 기반으로 사용자가 현재 360도 비디오 내에서 보고 있는 영역에 대한 정보, 즉 뷰포트 정보가 계산될 수 있다.
- [0043] 뷰포트 정보는 현재 사용자가 360도 비디오에서 보고 있는 영역에 대한 정보일 수 있다. 이를 통해 게이즈 분석(Gaze Analysis) 이 수행되어, 사용자가 어떠한 방식으로 360도 비디오를 소비하는지, 360도 비디오의 어느 영역을 얼마나 응시하는지 등을 확인할 수도 있다. 게이즈 분석은 수신측에서 수행되어 송신측으로 피드백 채널을 통해 전달될 수도 있다. VR 디스플레이 등의 장치는 사용자의 머리 위치/방향, 장치가 지원하는 수직(vertical) 혹은 수평(horizontal) FOV(Field Of View) 정보 등에 근거하여 뷰포트 영역을 추출할 수 있다.
- [0044] 실시예에 따라, 전송한 피드백 정보는 송신측으로 전달되는 것 뿐아니라, 수신측에서 소비될 수도 있다. 즉, 전송한 피드백 정보를 이용하여 수신측의 디코딩, 리-프로젝션, 렌더링 과정 등이 수행될 수 있다. 예를 들어, 헤드 오리엔테이션 정보 및/또는 뷰포트 정보를 이용하여 현재 사용자가 보고 있는 영역에 대한 360도 비디오만 우선적으로 디코딩 및 렌더링될 수도 있다.
- [0045] 여기서 뷰포트(viewport) 내지 뷰포트 영역이란, 사용자가 360도 비디오에서 보고 있는 영역을 의미할 수 있다. 시점(viewpoint) 는 사용자가 360도 비디오에서 보고 있는 지점으로서, 뷰포트 영역의 정중앙 지점을 의미할 수 있다. 즉, 뷰포트는 시점을 중심으로 한 영역인데, 그 영역이 차지하는 크기 형태 등은 후술할 FOV(Field Of View) 에 의해 결정될 수 있다.

- [0046]     전술한 360도 비디오 제공을 위한 전체 아키텍처 내에서, 캡처/프로젝션/인코딩/전송/디코딩/리-프로젝션/렌더링의 일련의 과정을 거치게 되는 이미지/비디오 데이터들을 360도 비디오 데이터라 부를 수 있다. 360도 비디오 데이터라는 용어는 또한 이러한 이미지/비디오 데이터들과 관련되는 메타데이터 내지 시그널링 정보를 포함하는 개념으로 쓰일 수도 있다.
- [0047]     상술한 오디오 또는 비디오 등의 미디어 데이터를 저장하고 전송하기 위하여, 정형화된 미디어 파일 포맷이 정의될 수 있다. 실시예에 따라 미디어 파일은 ISO BMFF (ISO base media file format)를 기반으로 한 파일 포맷을 가질 수 있다.
- [0048]     도 2 및 도 3은 본 발명의 일 실시예에 따른 미디어 파일의 구조를 도시한 도면이다.
- [0049]     본 발명에 따른 미디어 파일은 적어도 하나 이상의 박스를 포함할 수 있다. 여기서 박스(box)는 미디어 데이터 또는 미디어 데이터에 관련된 메타데이터 등을 포함하는 데이터 블록 내지 오브젝트일 수 있다. 박스들은 서로 계층적 구조를 이룰 수 있으며, 이에 따라 데이터들이 분류되어 미디어 파일이 대용량 미디어 데이터의 저장 및/또는 전송에 적합한 형태를 띄게 될 수 있다. 또한 미디어 파일은, 사용자가 미디어 콘텐츠의 특정지점으로 이동하는 등, 미디어 정보에 접근하는데 있어 용이한 구조를 가질 수 있다.
- [0050]     본 발명에 따른 미디어 파일은 ftyp 박스, moov 박스 및/또는 mdat 박스를 포함할 수 있다.
- [0051]     ftyp 박스(파일 타입 박스)는 해당 미디어 파일에 대한 파일 타입 또는 호환성 관련 정보를 제공할 수 있다. ftyp 박스는 해당 미디어 파일의 미디어 데이터에 대한 구성 버전 정보를 포함할 수 있다. 복호기는 ftyp 박스를 참조하여 해당 미디어 파일을 구분할 수 있다.
- [0052]     moov 박스(무비 박스)는 해당 미디어 파일의 미디어 데이터에 대한 메타 데이터를 포함하는 박스일 수 있다. moov 박스는 모든 메타 데이터들을 위한 컨테이너 역할을 할 수 있다. moov 박스는 메타 데이터 관련 박스들 중 최상위 계층의 박스일 수 있다. 실시예에 따라 moov 박스는 미디어 파일 내에 하나만 존재할 수 있다.
- [0053]     mdat 박스(미디어 데이터 박스)는 해당 미디어 파일의 실제 미디어 데이터들을 담는 박스일 수 있다. 미디어 데이터들은 오디오 샘플 및/또는 비디오 샘플들을 포함할 수 있는데, mdat 박스는 이러한 미디어 샘플들을 담는 컨테이너 역할을 할 수 있다.
- [0054]     실시예에 따라 전술한 moov 박스는 mvhd 박스, trak 박스 및/또는 mvex 박스 등을 하위 박스로서 더 포함할 수 있다.
- [0055]     mvhd 박스(무비 헤더 박스)는 해당 미디어 파일에 포함되는 미디어 데이터의 미디어 프리젠테이션 관련 정보를 포함할 수 있다. 즉, mvhd 박스는 해당 미디어 프리젠테이션의 미디어 생성시간, 변경시간, 시간규격, 기간 등의 정보를 포함할 수 있다.
- [0056]     trak 박스(트랙 박스)는 해당 미디어 데이터의 트랙에 관련된 정보를 제공할 수 있다. trak 박스는 오디오 트랙 또는 비디오 트랙에 대한 스트림 관련 정보, 프리젠테이션 관련 정보, 액세스 관련 정보 등의 정보를 포함할 수 있다. Trak 박스는 트랙의 개수에 따라 복수개 존재할 수 있다.
- [0057]     trak 박스는 실시예에 따라 tkhd 박스(트랙 헤더 박스)를 하위 박스로서 더 포함할 수 있다. tkhd 박스는 trak 박스가 나타내는 해당 트랙에 대한 정보를 포함할 수 있다. tkhd 박스는 해당 트랙의 생성시간, 변경시간, 트랙 식별자 등의 정보를 포함할 수 있다.
- [0058]     mvex 박스(무비 익스텐드 박스)는 해당 미디어 파일에 후술할 moof 박스가 있을 수 있음을 지시할 수 있다. 특정 트랙의 모든 미디어 샘플들을 알기 위해서, moof 박스들이 스캔되어야 할 수 있다.
- [0059]     본 발명에 따른 미디어 파일은, 실시예에 따라, 복수개의 프래그먼트로 나뉘어질 수 있다(200). 이를 통해 미디어 파일이 분할되어 저장되거나 전송될 수 있다. 미디어 파일의 미디어 데이터들(mdat 박스)은 복수개의 프래그먼트로 나뉘어지고, 각각의 프래그먼트는 moof 박스와 나뉘어진 mdat 박스를 포함할 수 있다. 실시예에 따라 프래그먼트들을 활용하기 위해서는 ftyp 박스 및/또는 moov 박스의 정보가 필요할 수 있다.
- [0060]     moof 박스(무비 프래그먼트 박스)는 해당 프래그먼트의 미디어 데이터에 대한 메타 데이터를 제공할 수 있다. moof 박스는 해당 프래그먼트의 메타데이터 관련 박스들 중 최상위 계층의 박스일 수 있다.
- [0061]     mdat 박스(미디어 데이터 박스)는 전술한 바와 같이 실제 미디어 데이터를 포함할 수 있다. 이 mdat 박스는 각각의 해당 프래그먼트에 해당하는 미디어 데이터들의 미디어 샘플들을 포함할 수 있다.

- [0062] 실시예에 따라 전술한 moof 박스는 mfhd 박스 및/또는 traf 박스 등을 하위 박스로서 더 포함할 수 있다.
- [0063] mfhd 박스(무비 프래그먼트 헤더 박스)는 분할된 복수개의 프래그먼트들 간의 연관성과 관련한 정보들을 포함할 수 있다. mfhd 박스는 시퀀스 넘버(sequence number) 를 포함하여, 해당 프래그먼트의 미디어 데이터가 분할된 몇 번째 데이터인지를 나타낼 수 있다. 또한, mfhd 박스를 이용하여 분할된 데이터 중 누락된 것은 없는지 여부가 확인될 수 있다.
- [0064] traf 박스(트랙 프래그먼트 박스)는 해당 트랙 프래그먼트에 대한 정보를 포함할 수 있다. traf 박스는 해당 프래그먼트에 포함되는 분할된 트랙 프래그먼트에 대한 메타데이터를 제공할 수 있다. traf 박스는 해당 트랙 프래그먼트 내의 미디어 샘플들이 복호화/재생될 수 있도록 메타데이터를 제공할 수 있다. traf 박스는 트랙 프래그먼트의 개수에 따라 복수개 존재할 수 있다.
- [0065] 실시예에 따라 전술한 traf 박스는 tfhd 박스 및/또는 trun 박스 등을 하위 박스로서 더 포함할 수 있다.
- [0066] tfhd 박스(트랙 프래그먼트 헤더 박스)는 해당 트랙 프래그먼트의 헤더 정보를 포함할 수 있다. tfhd 박스는 전술한 traf 박스가 나타내는 트랙 프래그먼트의 미디어 샘플들에 대하여, 기본적인 샘플크기, 기간, 오프셋, 식별자 등의 정보를 제공할 수 있다.
- [0067] trun 박스(트랙 프래그먼트 런 박스)는 해당 트랙 프래그먼트 관련 정보를 포함할 수 있다. trun 박스는 미디어 샘플별 기간, 크기, 재생시점 등과 같은 정보를 포함할 수 있다.
- [0068] 전술한 미디어 파일 내지 미디어 파일의 프래그먼트들은 세그먼트들로 처리되어 전송될 수 있다. 세그먼트에는 초기화 세그먼트(initialization segment) 및/또는 미디어 세그먼트(media segment) 가 있을 수 있다.
- [0069] 도시된 실시예(210)의 파일은, 미디어 데이터는 제외하고 미디어 디코더의 초기화와 관련된 정보 등을 포함하는 파일일 수 있다. 이 파일은 예를 들어 전술한 초기화 세그먼트에 해당할 수 있다. 초기화 세그먼트는 전술한 ftyp 박스 및/또는 moov 박스를 포함할 수 있다.
- [0070] 도시된 실시예(220)의 파일은, 전술한 프래그먼트를 포함하는 파일일 수 있다. 이 파일은 예를 들어 전술한 미디어 세그먼트에 해당할 수 있다. 미디어 세그먼트는 전술한 moof 박스 및/또는 mdat 박스를 포함할 수 있다. 또한, 미디어 세그먼트는 styp 박스 및/또는 sidx 박스를 더 포함할 수 있다.
- [0071] styp 박스(세그먼트 타입 박스) 는 분할된 프래그먼트의 미디어 데이터를 식별하기 위한 정보를 제공할 수 있다. styp 박스는 분할된 프래그먼트에 대해, 전술한 ftyp 박스와 같은 역할을 수행할 수 있다. 실시예에 따라 styp 박스는 ftyp 박스와 동일한 포맷을 가질 수 있다.
- [0072] sidx 박스(세그먼트 인덱스 박스) 는 분할된 프래그먼트에 대한 인덱스를 나타내는 정보를 제공할 수 있다. 이를 통해 해당 분할된 프래그먼트가 몇 번째 프래그먼트인지가 지시될 수 있다.
- [0073] 실시예에 따라(230) ssix 박스가 더 포함될 수 있는데, ssix 박스(서브 세그먼트 인덱스 박스)는 세그먼트가 서브 세그먼트로 더 나뉘어지는 경우에 있어, 그 서브 세그먼트의 인덱스를 나타내는 정보를 제공할 수 있다.
- [0074] 미디어 파일 내의 박스들은, 도시된 실시예(250)와 같은 박스 내지 풀 박스(FullBox) 형태를 기반으로, 더 확장된 정보들을 포함할 수 있다. 이 실시예에서 size 필드, largesize 필드는 해당 박스의 길이를 바이트 단위로 나타낼 수 있다. version 필드는 해당 박스 포맷의 버전을 나타낼 수 있다. Type 필드는 해당 박스의 타입 내지 식별자를 나타낼 수 있다. flags 필드는 해당 박스와 관련된 플래그 등을 나타낼 수 있다.
- [0075] 한편, 본 발명의 360도 비디오에 대한 필드(속성)들은 DASH 기반 적응형(Adaptive) 스트리밍 모델에 포함되어 전달될 수 있다.
- [0076] 도 4는 DASH 기반 적응형 스트리밍 모델의 전반적인 동작의 일 예를 나타낸다. 도시된 실시예(400)에 따른 DASH 기반 적응형 스트리밍 모델은, HTTP 서버와 DASH 클라이언트 간의 동작을 기술하고 있다. 여기서 DASH(Dynamic Adaptive Streaming over HTTP)는, HTTP 기반 적응형 스트리밍을 지원하기 위한 프로토콜로서, 네트워크 상황에 따라 동적으로 스트리밍을 지원할 수 있다. 이에 따라 AV 콘텐츠 재생이 끊임없이 제공될 수 있다.
- [0077] 먼저 DASH 클라이언트는 MPD를 획득할 수 있다. MPD 는 HTTP 서버 등의 서비스 프로바이더로부터 전달될 수 있다. DASH 클라이언트는 MPD 에 기술된 세그먼트에의 접근 정보를 이용하여 서버로 해당 세그먼트들을 요청할 수 있다. 여기서 이 요청은 네트워크 상태를 반영하여 수행될 수 있다.
- [0078] DASH 클라이언트는 해당 세그먼트를 획득한 후, 이를 미디어 엔진에서 처리하여 화면에 디스플레이할 수 있다.

DASH 클라이언트는 재생 시간 및/또는 네트워크 상황 등을 실시간으로 반영하여, 필요한 세그먼트를 요청, 획득할 수 있다(Adaptive Streaming). 이를 통해 콘텐츠가 끊임없이 재생될 수 있다.

- [0079] MPD (Media Presentation Description) 는 DASH 클라이언트로 하여금 세그먼트를 동적으로 획득할 수 있도록 하기 위한 상세 정보를 포함하는 파일로서 XML 형태로 표현될 수 있다.
- [0080] DASH 클라이언트 컨트롤러(DASH Client Controller) 는 네트워크 상황을 반영하여 MPD 및/또는 세그먼트를 요청하는 커맨드를 생성할 수 있다. 또한, 이 컨트롤러는 획득된 정보를 미디어 엔진 등등의 내부 블록에서 사용할 수 있도록 제어할 수 있다.
- [0081] MPD 파서(Parser) 는 획득한 MPD 를 실시간으로 파싱할 수 있다. 이를 통해, DASH 클라이언트 컨트롤러는 필요한 세그먼트를 획득할 수 있는 커맨드를 생성할 수 있게 될 수 있다.
- [0082] 세그먼트 파서(Parser) 는 획득한 세그먼트를 실시간으로 파싱할 수 있다. 세그먼트에 포함된 정보들에 따라 미디어 엔진 등의 내부 블록들은 특정 동작을 수행할 수 있다.
- [0083] HTTP 클라이언트는 필요한 MPD 및/또는 세그먼트 등을 HTTP 서버에 요청할 수 있다. 또한 HTTP 클라이언트는 서버로부터 획득한 MPD 및/또는 세그먼트들을 MPD 파서 또는 세그먼트 파서로 전달할 수 있다.
- [0084] 미디어 엔진(Media Engine) 은 세그먼트에 포함된 미디어 데이터를 이용하여 콘텐츠를 화면상에 표시할 수 있다. 이 때, MPD 의 정보들이 활용될 수 있다.
- [0085] DASH 데이터 모델은 계층적 구조(410)를 가질 수 있다. 미디어 프리젠테이션은 MPD에 의해 기술될 수 있다. MPD 는 미디어 프리젠테이션을 만드는 복수개의 구간(Period)들의 시간적인 시퀀스를 기술할 수 있다. 피리어드는 미디어 콘텐츠의 한 구간을 나타낼 수 있다.
- [0086] 한 구간에서, 데이터들은 어댑테이션 셋들에 포함될 수 있다. 어댑테이션 셋은 서로 교환될 수 있는 복수개의 미디어 콘텐츠 컴포넌트들의 집합일 수 있다. 어댑테이션은 레프리젠테이션들의 집합을 포함할 수 있다. 레프리젠테이션은 미디어 콘텐츠 컴포넌트에 해당할 수 있다. 한 레프리젠테이션 내에서, 콘텐츠는 복수개의 세그먼트들로 시간적으로 나뉘어질 수 있다. 이는 적절한 접근성과 전달(delivery)를 위할 수 있다. 각각의 세그먼트에 접근하기 위해서 각 세그먼트의 URL 이 제공될 수 있다.
- [0087] MPD는 미디어 프리젠테이션에 관련된 정보들을 제공할 수 있고, 피리어드 요소, 어댑테이션 셋 요소, 레프리젠테이션 요소는 각각 해당 피리어드, 어댑테이션 셋, 레프리젠테이션에 대해서 기술할 수 있다. 레프리젠테이션은 서브 레프리젠테이션들로 나뉘어질 수 있는데, 서브 레프리젠테이션 요소는 해당 서브 레프리젠테이션에 대해서 기술할 수 있다.
- [0088] 여기서 공통(Common) 속성/요소들이 정의될 수 있는데, 이 들은 어댑테이션 셋, 레프리젠테이션, 서브 레프리젠테이션 등에 적용될 수 (포함될 수) 있다. 공통 속성/요소 중에는 에센셜 프로퍼티(EssentialProperty) 및/또는 서플멘탈 프로퍼티(SupplementalProperty) 가 있을 수 있다.
- [0089] 에센셜 프로퍼티는 해당 미디어 프리젠테이션 관련 데이터를 처리함에 있어서 필수적이라고 여겨지는 요소들을 포함하는 정보일 수 있다. 서플멘탈 프로퍼티는 해당 미디어 프리젠테이션 관련 데이터를 처리함에 있어서 사용될 수도 있는 요소들을 포함하는 정보일 수 있다. 실시예에 따라 후술할 디스크립터들은, MPD 를 통해 전달되는 경우, 에센셜 프로퍼티 및/또는 서플멘탈 프로퍼티 내에 정의되어 전달될 수 있다.
- [0090] 도 5는 본 발명이 적용될 수 있는 360도 비디오 전송 장치의 구성을 개략적으로 설명하는 도면이다.
- [0091] 본 발명에 따른 360도 비디오 전송 장치는 전송한 준비 과정 내지 전송 과정에 관련된 동작들을 수행할 수 있다. 360도 비디오 전송 장치는 데이터 입력부, 스티처(Stitcher), 프로젝션 처리부, 리전별 패킹 처리부(도시되지 않음), 메타데이터 처리부, (송신측) 피드백 처리부, 데이터 인코더, 인캡슐레이션 처리부, 전송 처리부 및/또는 전송부를 내/외부 요소로서 포함할 수 있다.
- [0092] 데이터 입력부는 캡처된 각 시점별 이미지/비디오 들을 입력받을 수 있다. 이 시점별 이미지/비디오 들은 하나 이상의 카메라들에 의해 캡처된 이미지/비디오들일 수 있다. 또한 데이터 입력부는 캡처 과정에서 발생된 메타데이터를 입력받을 수 있다. 데이터 입력부는 입력된 시점별 이미지/비디오들을 스티처로 전달하고, 캡처 과정의 메타데이터를 시그널링 처리부로 전달할 수 있다.
- [0093] 스티처는 캡처된 시점별 이미지/비디오들에 대한 스티칭 작업을 수행할 수 있다. 스티처는 스티칭된 360도 비디오

오 데이터를 프로젝션 처리부로 전달할 수 있다. 스티처는 필요한 경우 메타데이터 처리부로부터 필요한 메타데이터를 전달받아 스티칭 작업에 이용할 수 있다. 스티처는 스티칭 과정에서 발생된 메타데이터를 메타데이터 처리부로 전달할 수 있다. 스티칭 과정의 메타데이터에는 스티칭이 수행되었는지 여부, 스티칭 타입 등의 정보들이 있을 수 있다.

[0094] 프로젝션 처리부는 스티칭된 360도 비디오 데이터를 2D 이미지 상에 프로젝션할 수 있다. 프로젝션 처리부는 다양한 스킴(scheme)에 따라 프로젝션을 수행할 수 있는데, 이에 대해서는 후술한다. 프로젝션 처리부는 각 시점별 360도 비디오 데이터의 해당 뎀스(depth)를 고려하여 맵핑을 수행할 수 있다. 프로젝션 처리부는 필요한 경우 메타데이터 처리부로부터 프로젝션에 필요한 메타데이터를 전달받아 프로젝션 작업에 이용할 수 있다. 프로젝션 처리부는 프로젝션 과정에서 발생된 메타데이터를 메타데이터 처리부로 전달할 수 있다. 프로젝션 처리부의 메타데이터에는 프로젝션 스킴의 종류 등이 있을 수 있다.

[0095] 리전별 패킹 처리부(도시되지 않음)는 전술한 리전별 패킹 과정을 수행할 수 있다. 즉, 리전별 패킹 처리부는 프로젝션된 360도 비디오 데이터를 리전별로 나누고, 각 리전들을 회전, 재배열하거나, 각 리전의 레졸루션을 변경하는 등의 처리를 수행할 수 있다. 전술한 바와 같이 리전별 패킹 과정은 선택적(optional) 과정이며, 리전별 패킹이 수행되지 않는 경우, 리전별 패킹 처리부는 생략될 수 있다. 리전별 패킹 처리부는 필요한 경우 메타데이터 처리부로부터 리전별 패킹에 필요한 메타데이터를 전달받아 리전별 패킹 작업에 이용할 수 있다. 리전별 패킹 처리부는 리전별 패킹 과정에서 발생된 메타데이터를 메타데이터 처리부로 전달할 수 있다. 리전별 패킹 처리부의 메타데이터에는 각 리전의 회전 정도, 사이즈 등이 있을 수 있다.

[0096] 전술한 스티처, 프로젝션 처리부 및/또는 리전별 패킹 처리부는 실시예에 따라 하나의 하드웨어 컴포넌트에서 수행될 수도 있다.

[0097] 메타데이터 처리부는 캡처 과정, 스티칭 과정, 프로젝션 과정, 리전별 패킹 과정, 인코딩 과정, 인캡슐레이션 과정 및/또는 전송을 위한 처리 과정에서 발생할 수 있는 메타데이터들을 처리할 수 있다. 메타데이터 처리부는 이러한 메타데이터들을 이용하여 360도 비디오 관련 메타데이터를 생성할 수 있다. 실시예에 따라 메타데이터 처리부는 360도 비디오 관련 메타데이터를 시그널링 테이블의 형태로 생성할 수도 있다. 시그널링 문맥에 따라 360도 비디오 관련 메타데이터는 메타데이터 또는 360도 비디오 관련 시그널링 정보라 불릴 수도 있다. 또한 메타데이터 처리부는 획득하거나 생성한 메타데이터들을 필요에 따라 360도 비디오 전송 장치의 내부 요소들에 전달할 수 있다. 메타데이터 처리부는 360도 비디오 관련 메타데이터가 수신측으로 전송될 수 있도록 데이터 인코더, 인캡슐레이션 처리부 및/또는 전송 처리부에 전달할 수 있다.

[0098] 데이터 인코더는 2D 이미지 상에 프로젝션된 360도 비디오 데이터 및/또는 리전별 패킹된 360도 비디오 데이터를 인코딩할 수 있다. 360도 비디오 데이터는 다양한 포맷으로 인코딩될 수 있다.

[0099] 인캡슐레이션 처리부는 인코딩된 360도 비디오 데이터 및/또는 360도 비디오 관련 메타데이터를 파일 등의 형태로 인캡슐레이션할 수 있다. 여기서 360도 비디오 관련 메타데이터는 전술한 메타데이터 처리부로부터 전달받은 것일 수 있다. 인캡슐레이션 처리부는 해당 데이터들을 ISOBMFF, CFF 등의 파일 포맷으로 인캡슐레이션하거나, 기타 DASH 세그먼트 등의 형태로 처리할 수 있다. 인캡슐레이션 처리부는 실시예에 따라 360도 비디오 관련 메타데이터를 파일 포맷 상에 포함시킬 수 있다. 360도 관련 메타데이터는 예를 들어 ISOBMFF 파일 포맷 상의 다양한 레벨의 박스(box)에 포함되거나 파일 내에서 별도의 트랙내의 데이터로 포함될 수 있다. 실시예에 따라, 인캡슐레이션 처리부는 360도 비디오 관련 메타데이터 자체를 파일로 인캡슐레이션할 수 있다. 전송 처리부는 파일 포맷에 따라 인캡슐레이션된 360도 비디오 데이터에 전송을 위한 처리를 가할 수 있다. 전송 처리부는 임의의 전송 프로토콜에 따라 360도 비디오 데이터를 처리할 수 있다. 전송을 위한 처리에는 방송망을 통한 전달을 위한 처리, 브로드밴드를 통한 전달을 위한 처리를 포함할 수 있다. 실시예에 따라 전송 처리부는 360도 비디오 데이터뿐만 아니라, 메타데이터 처리부로부터 360도 비디오 관련 메타데이터를 전달받아, 이 것에 전송을 위한 처리를 가할 수도 있다.

[0100] 전송부는 전송 처리된 360도 비디오 데이터 및/또는 360도 비디오 관련 메타데이터를 방송망 및/또는 브로드밴드를 통해 전송할 수 있다. 전송부는 방송망을 통한 전송을 위한 요소 및/또는 브로드밴드를 통한 전송을 위한 요소를 포함할 수 있다.

[0101] 본 발명에 따른 360도 비디오 전송 장치의 일 실시예에 의하면, 360도 비디오 전송 장치는 데이터 저장부(도시되지 않음)를 내/외부 요소로서 더 포함할 수 있다. 데이터 저장부는 인코딩된 360도 비디오 데이터 및/또는 360도 비디오 관련 메타데이터를 전송 처리부로 전달하기 전에 저장하고 있을 수 있다. 이 데이터들이 저장

되는 형태는 ISOBMFF 등의 파일 형태일 수 있다. 실시간으로 360도 비디오를 전송하는 경우에는 데이터 저장부가 필요하지 않을 수 있으나, 온 디맨드, NRT (Non Real Time), 브로드밴드 등을 통해 전달하는 경우에는 인캡슐레이션된 360 데이터가 데이터 저장부에 일정 기간 저장되었다가 전송될 수도 있다.

- [0102] 본 발명에 따른 360도 비디오 전송 장치의 다른 실시예에 의하면, 360도 비디오 전송 장치는 (송신측) 피드백 처리부 및/또는 네트워크 인터페이스(도시되지 않음)를 내/외부 엘리먼트로서 더 포함할 수 있다. 네트워크 인터페이스는 본 발명에 따른 360도 비디오 수신 장치로부터 피드백 정보를 전달받고, 이를 송신측 피드백 처리부로 전달할 수 있다. 송신측 피드백 처리부는 피드백 정보를 스티처, 프로젝션 처리부, 리전별 패킹 처리부, 데이터 인코더, 인캡슐레이션 처리부, 메타데이터 처리부 및/또는 전송 처리부로 전달할 수 있다. 실시예에 따라 피드백 정보는 메타데이터 처리부에 일단 전달된 후, 다시 각 내부 엘리먼트들로 전달될 수 있다. 피드백 정보를 전달받은 내부 엘리먼트들은 이 후의 360도 비디오 데이터의 처리에 피드백 정보를 반영할 수 있다.
- [0103] 본 발명에 따른 360도 비디오 전송 장치의 또 다른 실시예에 의하면, 리전별 패킹 처리부는 각 리전을 회전하여 2D 이미지 상에 맵핑할 수 있다. 이 때 각 리전들은 서로 다른 방향, 서로 다른 각도로 회전되어 2D 이미지 상에 맵핑될 수 있다. 리전의 회전은 360도 비디오 데이터가 구형의 면 상에서 프로젝션 전에 인접했던 부분, 스티칭된 부분 등을 고려하여 수행될 수 있다. 리전의 회전에 관한 정보들, 즉 회전 방향, 각도 등은 360도 비디오 관련 메타데이터에 의해 시그널링될 수 있다. 본 발명에 따른 360도 비디오 전송 장치의 또 다른 실시예에 의하면, 데이터 인코더는 각 리전 별로 다르게 인코딩을 수행할 수 있다. 데이터 인코더는 특정 리전은 높은 퀄리티로, 다른 리전은 낮은 퀄리티로 인코딩을 수행할 수 있다. 송신측 피드백 처리부는 360도 비디오 수신 장치로부터 전달받은 피드백 정보를 데이터 인코더로 전달하여, 데이터 인코더가 리전별 차등화된 인코딩 방법을 사용하도록 할 수 있다. 예를 들어 송신측 피드백 처리부는 수신측으로부터 전달받은 뷰포트 정보를 데이터 인코더로 전달할 수 있다. 데이터 인코더는 뷰포트 정보가 지시하는 영역을 포함하는 리전들에 대해 다른 리전들보다 더 높은 퀄리티(UHD 등) 로 인코딩을 수행할 수 있다.
- [0104] 본 발명에 따른 360도 비디오 전송 장치의 또 다른 실시예에 의하면, 전송 처리부는 각 리전 별로 다르게 전송을 위한 처리를 수행할 수 있다. 전송 처리부는 리전 별로 다른 전송 파라미터(모듈레이션 오더, 코드 레이트 등)를 적용하여, 각 리전 별로 전달되는 데이터의 강건성(robustness)을 다르게 할 수 있다.
- [0105] 이 때, 송신측 피드백 처리부는 360도 비디오 수신 장치로부터 전달받은 피드백 정보를 전송 처리부로 전달하여, 전송 처리부가 리전별 차등화된 전송 처리를 수행하도록 할 수 있다. 예를 들어 송신측 피드백 처리부는 수신측으로부터 전달받은 뷰포트 정보를 전송 처리부로 전달할 수 있다. 전송 처리부는 해당 뷰포트 정보가 지시하는 영역을 포함하는 리전들에 대해 다른 리전들보다 더 높은 강건성을 가지도록 전송 처리를 수행할 수 있다.
- [0106] 전술한 본 발명에 따른 360도 비디오 전송 장치의 내/외부 엘리먼트들은 하드웨어로 구현되는 하드웨어 엘리먼트들일 수 있다. 실시예에 따라 내/외부 엘리먼트들은 변경, 생략되거나 다른 엘리먼트로 대체, 통합될 수 있다. 실시예에 따라 부가 엘리먼트들이 360도 비디오 전송 장치에 추가될 수도 있다.
- [0107] 도 6은 본 발명이 적용될 수 있는 360도 비디오 수신 장치의 구성을 개략적으로 설명하는 도면이다.
- [0108] 본 발명에 따른 360도 비디오 수신 장치는 전술한 프로세싱 과정 및/또는 렌더링 과정에 관련된 동작들을 수행할 수 있다. 360도 비디오 수신 장치는 수신부, 수신 처리부, 디캡슐레이션 처리부, 데이터 디코더, 메타데이터 파서, (수신측) 피드백 처리부, 리-프로젝션 처리부 및/또는 렌더러를 내/외부 엘리먼트로서 포함할 수 있다. 한편, 시그널링 파서는 메타데이터 파서라고 불릴 수 있다.
- [0109] 수신부는 본 발명에 따른 360도 비디오 전송 장치가 전송한 360도 비디오 데이터를 수신할 수 있다. 전송되는 채널에 따라 수신부는 방송망을 통하여 360도 비디오 데이터를 수신할 수도 있고, 브로드밴드를 통하여 360도 비디오 데이터를 수신할 수도 있다.
- [0110] 수신 처리부는 수신된 360도 비디오 데이터에 대해 전송 프로토콜에 따른 처리를 수행할 수 있다. 전송측에서 전송을 위한 처리가 수행된 것에 대응되도록, 수신 처리부는 전술한 전송 처리부의 역과정을 수행할 수 있다. 수신 처리부는 획득한 360도 비디오 데이터는 디캡슐레이션 처리부로 전달하고, 획득한 360도 비디오 관련 메타데이터는 메타데이터 파서로 전달할 수 있다. 수신 처리부가 획득하는 360도 비디오 관련 메타데이터는 시그널링 테이블의 형태일 수 있다.
- [0111] 디캡슐레이션 처리부는 수신 처리부로부터 전달받은 파일 형태의 360도 비디오 데이터를 디캡슐레이션할 수 있다. 디캡슐레이션 처리부는 ISOBMFF 등에 따른 파일들을 디캡슐레이션하여, 360도 비디오 데이터 내지 360도 비

디오 관련 메타데이터를 획득할 수 있다. 획득된 360도 비디오 데이터는 데이터 디코더로, 획득된 360도 비디오 관련 메타데이터는 메타데이터 파서로 전달할 수 있다. 디캡슐레이션 처리부가 획득하는 360도 비디오 관련 메타데이터는 파일 포맷 내의 박스 혹은 트랙 형태일 수 있다. 디캡슐레이션 처리부는 필요한 경우 메타데이터 파서로부터 디캡슐레이션에 필요한 메타데이터를 전달받을 수도 있다.

- [0112] 데이터 디코더는 360도 비디오 데이터에 대한 디코딩을 수행할 수 있다. 데이터 디코더는 메타데이터 파서로부터 디코딩에 필요한 메타데이터를 전달받을 수도 있다. 데이터 디코딩 과정에서 획득된 360도 비디오 관련 메타데이터는 메타데이터 파서로 전달될 수도 있다.
- [0113] 메타데이터 파서는 360도 비디오 관련 메타데이터에 대한 파싱/디코딩을 수행할 수 있다. 메타데이터 파서는 획득한 메타데이터를 데이터 디캡슐레이션 처리부, 데이터 디코더, 리-프로젝션 처리부 및/또는 렌더러로 전달할 수 있다.
- [0114] 리-프로젝션 처리부는 디코딩된 360도 비디오 데이터에 대하여 리-프로젝션을 수행할 수 있다. 리-프로젝션 처리부는 360도 비디오 데이터를 3D 공간으로 리-프로젝션할 수 있다. 3D 공간은 사용되는 3D 모델에 따라 다른 형태를 가질 수 있다. 리-프로젝션 처리부는 메타데이터 파서로부터 리-프로젝션에 필요한 메타데이터를 전달받을 수도 있다. 예를 들어 리-프로젝션 처리부는 사용되는 3D 모델의 타입 및 그 세부 정보에 대한 정보를 메타데이터 파서로부터 전달받을 수 있다. 실시예에 따라 리-프로젝션 처리부는 리-프로젝션에 필요한 메타데이터를 이용하여, 3D 공간 상의 특정 영역에 해당하는 360도 비디오 데이터만을 3D 공간으로 리-프로젝션할 수도 있다.
- [0115] 렌더러는 리-프로젝션된 360도 비디오 데이터를 렌더링할 수 있다. 전술한 바와 같이 360도 비디오 데이터가 3D 공간상에 렌더링된다고 표현할 수도 있는데, 이처럼 두 과정이 한번에 일어나는 경우 리-프로젝션 처리부와 렌더러는 통합되어, 렌더러에서 이 과정들이 모두 진행될 수 있다. 실시예에 따라 렌더러는 사용자의 시점 정보에 따라 사용자가 보고 있는 부분만을 렌더링할 수도 있다.
- [0116] 사용자는 VR 디스플레이 등을 통하여 렌더링된 360도 비디오의 일부 영역을 볼 수 있다. VR 디스플레이는 360도 비디오를 재생하는 장치로서, 360도 비디오 수신 장치에 포함될 수도 있고(tethered), 별도의 장치로서 360도 비디오 수신 장치에 연결될 수도 있다(un-tethered).
- [0117] 본 발명에 따른 360도 비디오 수신 장치의 일 실시예에 의하면, 360도 비디오 수신 장치는 (수신측) 피드백 처리부 및/또는 네트워크 인터페이스(도시되지 않음)를 내/외부 요소로서 더 포함할 수 있다. 수신측 피드백 처리부는 렌더러, 리-프로젝션 처리부, 데이터 디코더, 디캡슐레이션 처리부 및/또는 VR 디스플레이로부터 피드백 정보를 획득하여 처리할 수 있다. 피드백 정보는 뷰포트 정보, 헤드 오리엔테이션 정보, 게이즈(Gaze) 정보 등을 포함할 수 있다. 네트워크 인터페이스는 피드백 정보를 수신측 피드백 처리부로부터 전달받고, 이를 360도 비디오 전송 장치로 전송할 수 있다.
- [0118] 전술한 바와 같이, 피드백 정보는 송신측으로 전달되는 것뿐 아니라, 수신측에서 소비될 수도 있다. 수신측 피드백 처리부는 획득한 피드백 정보를 360도 비디오 수신 장치의 내부 요소들로 전달하여, 렌더링 등의 과정에 반영되게 할 수 있다. 수신측 피드백 처리부는 피드백 정보를 렌더러, 리-프로젝션 처리부, 데이터 디코더 및/또는 디캡슐레이션 처리부로 전달할 수 있다. 예를 들어, 렌더러는 피드백 정보를 활용하여 사용자가 보고 있는 영역을 우선적으로 렌더링할 수 있다. 또한 디캡슐레이션 처리부, 데이터 디코더 등은 사용자가 보고 있는 영역 내지 보게 될 영역을 우선적으로 디캡슐레이션, 디코딩할 수 있다.
- [0119] 전술한 본 발명에 따른 360도 비디오 수신 장치의 내/외부 요소들은 하드웨어로 구현되는 하드웨어 요소들일 수 있다. 실시예에 따라 내/외부 요소들은 변경, 생략되거나 다른 요소로 대체, 통합될 수 있다. 실시예에 따라 부가 요소들이 360도 비디오 수신 장치에 추가될 수도 있다.
- [0120] 본 발명의 또 다른 관점은 360도 비디오를 전송하는 방법 및 360도 비디오를 수신하는 방법과 관련될 수 있다. 본 발명에 따른 360도 비디오를 전송/수신하는 방법은, 각각 전술한 본 발명에 따른 360도 비디오 전송/수신 장치 또는 그 장치의 실시예들에 의해 수행될 수 있다.
- [0121] 전술한 본 발명에 따른 360도 비디오 전송/수신 장치, 전송/수신 방법의 각각의 실시예 및 그 내/외부 요소들 각각의 실시예들을 서로 조합될 수 있다. 예를 들어 프로젝트션 처리부의 실시예들과, 데이터 인코더의 실시예들은 서로 조합되어, 그 경우의 수만큼의 360도 비디오 전송 장치의 실시예들을 만들어 낼 수 있다. 이렇게 조합된 실시예들 역시 본 발명의 범위에 포함된다.
- [0122] 도 7a 내지 도 7b는 360도 비디오 전송 장치/360도 비디오 수신 장치를 통하여 수행되는 360도 비디오 제공을



위한 전체 아키텍처를 예시적으로 나타낸다.

- [0123] 도 7a 내지 도 7b에 도시된 것과 같은 아키텍처에 의하여 360 콘텐츠가 제공될 수 있다. 360 콘텐츠는 파일 형태로 제공되거나, DASH 등과 같이 세그먼트(segment) 기반 다운로드 또는 스트리밍 서비스의 형태로 제공될 수 있다. 여기서 360 콘텐츠는 VR 콘텐츠로 불릴 수 있다.
- [0124] 구체적으로 도 7a를 참조하면 전송한 바와 같이 360도 비디오 데이터 및/또는 360 오디오 데이터가 획득될 수 있다(Acquisition). 또한, 360 오디오 데이터는 오디오 프리-프로세싱 과정(Audio Preprocessing), 오디오 인코딩 과정(Audio encoding)을 거칠 수 있다. 이 과정에서 오디오 관련 메타데이터가 생성될 수 있으며, 인코딩된 오디오와 오디오 관련 메타데이터는 전송을 위한 처리(file/segment encapsulation)를 거칠 수 있다.
- [0125] 360도 비디오 데이터는 전송한 것과 같은 과정을 거칠 수 있다. 도 7a를 참조하면 360도 비디오 전송 장치의 스티칭은 360도 비디오 데이터에 스티칭을 수행할 수 있다(Visual stitching). 이 과정은 실시예에 따라 생략되고 수신측에서 수행될 수도 있다.
- [0126] 또한, 도 7a를 참조하면 360도 비디오 전송 장치의 프로젝션 처리부는 360도 비디오 데이터를 2D 이미지 상에 프로젝션할 수 있다(Projection and mapping(packaging)). 프로젝션 처리부는 360도 비디오 데이터(Input Images)를 전달받을 수 있고, 이 경우, 스티칭 및 프로젝션 과정을 수행할 수 있다. 프로젝션 과정은 구체적으로 스티칭된 360도 비디오 데이터를 3D 공간 상으로 프로젝션하고, 프로젝션된 360도 비디오 데이터가 2D 이미지 상으로 배열되는 것으로 볼 수 있다. 본 명세서에서 이 과정을 360도 비디오 데이터를 2D 이미지 상으로 프로젝션한다고 표현할 수도 있다. 여기서 3D 공간은 구(sphere) 또는 큐브(cube) 등일 수 있다. 이 3D 공간은 수신측에서 리-프로젝션에 사용되는 3D 공간과 같을 수도 있다.
- [0127] 2D 이미지는 프로젝션된 프레임(Projected frame) 또는 프로젝션된 픽처(Projected picture)라고 불릴 수도 있다. 또한, 상기 2D 이미지에 리전별 패킹(Region-wise packing) 과정이 선택적으로 더 수행될 수도 있다. 상기 리전별 패킹 과정이 수행되는 경우, 각 리전(Region)의 위치, 형태, 크기를 지시함으로써, 상기 2D 이미지 상의 리전들이 패킹된 프레임(packed frame) 상으로 매핑될 수 있다. 상기 패킹된 프레임은 패킹된 픽처(packed picture)라고 불릴 수 있다. 상기 프로젝션된 프레임에 상기 리전별 패킹 과정이 수행되지 않는 경우, 상기 프로젝션된 프레임은 상기 패킹된 프레임과 같을 수 있다. 리전에 대해서는 후술한다. 프로젝션 과정 및 리전별 패킹 과정을, 360도 비디오 데이터의 각 리전들이 2D 이미지 상에 프로젝션된다고 표현할 수도 있다. 설계에 따라, 360도 비디오 데이터는 중간 과정 없이 패킹된 프레임(packed frame)으로 바로 변환될 수도 있다.
- [0128] 도 7a를 참조하면 360도 비디오 데이터에 관한 패킹된 프레임은 이미지 인코딩 내지 비디오 인코딩될 수 있다. 한편, 같은 360도 비디오 콘텐츠라도 시점(viewpoints)별로 360도 비디오 데이터가 존재할 수 있고, 이 경우 상기 콘텐츠의 각 시점별 360도 비디오 데이터는 서로 다른 비트 스트림으로 인코딩될 수도 있다. 인코딩된 360도 비디오 데이터는 전송한 인캡슐레이션 처리부에 의해 ISOBMFF 등의 파일 포맷으로 처리될 수 있다. 또는 인캡슐레이션 처리부는 인코딩된 360도 비디오 데이터를 세그먼트들로 처리할 수 있다. 세그먼트들은 DASH 에 기반한 전송을 위한 개별 트랙에 포함될 수 있다.
- [0129] 360도 비디오 데이터의 처리와 함께, 전송한 것과 같이 360도 비디오 관련 메타데이터가 생성될 수 있다. 이 메타데이터는 비디오 스트림 혹은 파일 포맷에 포함되어 전달될 수 있다. 이 메타데이터는 인코딩 과정이나 파일 포맷 인캡슐레이션, 전송을 위한 처리 등과 같은 과정에도 쓰일 수 있다.
- [0130] 360 오디오/비디오 데이터는 전송 프로토콜에 따라 전송을 위한 처리를 거치고, 이후 전송될 수 있다. 전송한 360도 비디오 수신 장치는 이를 방송망 또는 브로드밴드를 통해 수신할 수 있다.
- [0131] 한편, 도 7a에 도시된 것과 같이 스피커/헤드폰(Loudspeakers/headphones), 디스플레이(Display), 헤드/아이 트래킹 컴포넌트(Head/eye tracking) 는 360도 비디오 수신 장치의 외부 장치 내지 VR 어플리케이션에 의해 수행될 수 있으나, 실시예에 따라 360도 비디오 수신 장치는 상기 스피커/헤드폰, 상기 디스플레이(Display), 상기 헤드/아이 트래킹 컴포넌트를 모두 포함할 수도 있다. 실시예에 따라 상기 헤드/아이 트래킹 컴포넌트는 전송한 수신측 피드백 처리부에 해당할 수 있다.
- [0132] 360도 비디오 수신 장치는 360 오디오/비디오 데이터에 수신을 위한 처리(File/segment decapsulation)를 수행할 수 있다. 360 오디오 데이터는 오디오 디코딩(Audio decoding), 오디오 렌더링(Audio rendering) 과정을 거쳐 스피커/헤드폰을 통해 사용자에게 제공될 수 있다.
- [0133] 360도 비디오 데이터는 이미지 디코딩 내지 비디오 디코딩, 렌더링(Visual rendering) 과정을 거쳐 디스플레이

를 통해 사용자에게 제공될 수 있다. 여기서 디스플레이는 VR 을 지원하는 디스플레이거나 일반 디스플레이일 수 있다.

- [0134] 전술한 바와 같이 렌더링 과정은 구체적으로, 360도 비디오 데이터가 3D 공간 상에 리-프로젝션되고, 리-프로젝션된 360도 비디오 데이터가 렌더링되는 것으로 볼 수 있다. 이를 360도 비디오 데이터가 3D 공간 상에 렌더링된다고 표현할 수도 있다.
- [0135] 헤드/아이 트래킹 컴포넌트는 사용자의 헤드 오리엔테이션 정보, 게이즈 정보, 뷰포트(Viewport) 정보 등을 획득, 처리할 수 있다. 이와 관련된 내용은 전술한 바와 같을 수 있다.
- [0136] 수신측에서는 전술한 수신측 과정들과 통신하는 VR 어플리케이션이 존재할 수 있다.
- [0137] 또한, 도 7b는 360도 비디오의 처리 과정 및 프로젝션 스킵에 따른 리전별 패키 과정이 적용된 2D 이미지를 예시적으로 나타낸다. 도 7b를 참조하면 입력된 360도 비디오 데이터의 처리 과정을 나타낼 수 있다. 구체적으로, 도 7b를 참조하면 입력된 시점의 360도 비디오 데이터는 다양한 프로젝션 스킵에 따라서 3D 프로젝션 구조에 스티칭 및 프로젝션될 수 있고, 상기 3D 프로젝션 구조에 프로젝션된 360도 비디오 데이터는 2D 이미지로 나타낼 수 있다. 즉, 상기 360도 비디오 데이터는 스티칭될 수 있고, 상기 2D 이미지로 프로젝션될 수 있다. 상기 360도 비디오 데이터가 프로젝션된 2D 이미지는 프로젝션된 프레임(projected frame)이라고 나타낼 수 있다. 또한, 상기 프로젝션된 프레임은 전술한 리전별 패킹 과정이 수행될 수 있다. 즉, 상기 프로젝션된 프레임 상의 프로젝션된 360도 비디오 데이터를 포함하는 영역을 리전들로 나누고, 각 리전들을 회전, 재배열하거나, 각 리전의 레졸루션을 변경하는 등의 처리가 수행될 수 있다. 다시 말해, 상기 리전별 패킹 과정은 상기 프로젝션된 프레임을 하나 이상의 패킹된 프레임(packed frame)으로 맵핑하는 과정을 나타낼 수 있다. 상기 리전별 패킹 과정의 수행은 선택적(optional)일 수 있고, 상기 리전별 패킹 과정이 적용되지 않는 경우, 상기 패킹된 프레임과 상기 프로젝션된 프레임은 동일할 수 있다. 상기 리전별 패킹 과정이 적용되는 경우, 상기 프로젝션된 프레임의 각 리전은 상기 패킹된 프레임의 리전에 맵핑될 수 있고, 상기 프로젝션된 프레임의 각 리전이 맵핑되는 상기 패킹된 프레임의 리전의 위치, 모양 및 크기를 나타내는 메타데이터가 도출될 수 있다.
- [0138] 도 8은 본 발명의 3D 공간을 설명하기 위한 비행기 주축(Aircraft Principal Axes) 개념을 도시한 도면이다. 본 발명에서, 3D 공간에서의 특정 지점, 위치, 방향, 간격, 영역 등을 표현하기 위하여 비행기 주축 개념이 사용될 수 있다. 즉, 본 발명에서 프로젝션 전 또는 리-프로젝션 후의 3D 공간에 대해 기술하고, 그에 대한 시그널링을 수행하기 위하여 비행기 주축 개념이 사용될 수 있다. 실시예에 따라 X, Y, Z 축 개념 또는 구형 좌표계를 이용한 방법이 사용될 수도 있다.
- [0139] 비행기는 3 차원으로 자유롭게 회전할 수 있다. 3차원을 이루는 축을 각각 피치(pitch) 축, 요(yaw) 축 및 롤(roll) 축이라고 한다. 본 명세서에서 이 들을 줄여서 pitch, yaw, roll 내지 pitch 방향, yaw 방향, roll 방향이라고 표현할 수도 있다.
- [0140] Pitch 축은 비행기의 앞코가 위/아래로 회전하는 방향의 기준이 되는 축을 의미할 수 있다. 도시된 비행기 주축 개념에서 pitch 축은 비행기의 날개에서 날개로 이어지는 축을 의미할 수 있다.
- [0141] Yaw 축은 비행기의 앞코가 좌/우로 회전하는 방향의 기준이 되는 축을 의미할 수 있다. 도시된 비행기 주축 개념에서 yaw 축은 비행기의 위에서 아래로 이어지는 축을 의미할 수 있다. Roll 축은 도시된 비행기 주축 개념에서 비행기의 앞코에서 꼬리로 이어지는 축으로서, roll 방향의 회전이란 roll 축을 기준으로 한 회전을 의미할 수 있다. 전술한 바와 같이, pitch, yaw, roll 개념을 통해 본 발명에서의 3D 공간이 기술될 수 있다.
- [0142] 도 9a 내지 도 9b는 본 발명에 따른 프로젝션 스킵들을 예시적으로 나타낸다. 상술한 바와 같이 본 발명에 따른 360도 비디오 전송 장치의 프로젝션 처리부는 스티칭된 360도 비디오 데이터를 2D 이미지 상에 프로젝션할 수 있다. 이 과정에서 다양한 프로젝션 스킵들이 적용될 수 있다. 즉, 상기 프로젝션 처리부는 다양한 프로젝션 스킵들에 따라 360도 비디오 데이터를 2D 이미지 상에 프로젝션할 수 있다. 상기 2D 이미지는 프로젝션된 픽처라고 나타낼 수 있다.
- [0143] 본 발명에 따른 일 실시예에 의하면, 등정방형 프로젝션 스킵(Equirectangular Projection scheme)을 이용하여 프로젝션이 수행될 수 있다. 프로젝션 처리부는 등정방형 프로젝션 스킵을 이용하여 360도 비디오 데이터의 프로젝션을 수행할 수 있다. 도 9a의 (a)는 등정방형 프로젝션 스킵을 나타낼 수 있다. 상기 등정방형 프로젝션 스킵은 ERP(Equirectangular Projection)라고 나타낼 수도 있다. 상기 등정방형 프로젝션 스킵이 적용되는 경우, 구형 면 상의  $(r, \theta_0, 0)$  즉,  $\theta = \theta_0, \phi = 0$  인 점과 2D 이미지의 중앙 픽셀이 매핑될 수 있다.

또한, 앞면 카메라(front camera)의 주점(principal point)을 구형 면의  $(r, 0, 0)$  지점으로 가정할 수 있다. 또한,  $\phi_0 = 0$ 으로 고정될 수 있다. 따라서, XY 좌표계로 변환된 값  $(x, y)$  는 다음의 수학적식을 통하여 2D 이미지 상에  $(X, Y)$  픽셀로 변환될 수 있다.

**수학적식 1**

$$X = K_x * x + X_o = K_x * (\theta - \theta_0) * r + X_o$$

$$Y = -K_y * y - Y_o$$

[0144]

또한, 2D 이미지의 좌상단 픽셀을 XY 좌표계의  $(0,0)$ 에 위치시키는 경우, x축에 대한 오프셋 값 및 y축에 대한 오프셋 값은 다음의 수학적식을 통하여 나타낼 수 있다.

**수학적식 2**

$$X_o = K_x * \pi * r$$

$$Y_o = -K_y * \pi / 2 * r$$

[0146]

이를 이용하여 상술한 수학적식 1에 나타난 XY 좌표계로의 변환식을 다시 쓰면 다음과 같을 수 있다.

**수학적식 3**

$$X = K_x * x + X_o = K_x * (\pi + \theta - \theta_0) * r$$

$$Y = -K_y * y - Y_o = K_y * (\pi / 2 - \phi) * r$$

[0148]

예를 들어  $\theta_0 = 0$  인 경우, 즉 2D 이미지의 중앙 픽셀이 구형 면 상의  $\theta = 0$  인 데이터를 가리키는 경우, 구형 면은  $(0,0)$ 을 기준으로 2D 이미지 상에서 가로길이(width) =  $2K_x \pi r$  이고 세로길이(height) =  $K_x \pi r$  인 영역에 매핑될 수 있다. 구형 면 상에서  $\phi = \pi / 2$  인 데이터는 2D 이미지 상의 윗쪽 변 전체에 매핑될 수 있다. 또한, 구형 면 상에서  $(r, \pi / 2, 0)$  인 데이터는 2D 이미지 상의  $(3\pi K_x r / 2, \pi K_x r / 2)$  인 점에 매핑될 수 있다.

[0149]

수신 측에서는, 2D 이미지 상의 360도 비디오 데이터를 구형 면 상으로 리-프로젝션할 수 있다. 구체적으로 상술한 360도 비디오 수신 장치의 리-프로젝션 처리부는 2D 이미지 상의 360도 비디오 데이터를 구형 면 상으로 리-프로젝션할 수 있다. 상기 2D 이미지는 프로젝션된 픽처라고 나타낼 수 있다. 이를 변환식으로 쓰면 다음의 수학적식과 같을 수 있다.

**수학적식 4**

$$\theta = \theta_0 + X / K_x * r - \pi$$

$$\phi = \pi / 2 - Y / K_y * r$$

[0151]

예를 들어 2D 이미지 상에서 XY 좌표값이  $(K_x \pi r, 0)$  인 픽셀은 구형 면 상의  $\theta = \theta_0, \phi = \pi / 2$  인 점으로 리-프로젝션될 수 있다.

[0152]

본 발명에 따른 또 다른 실시예에 의하면, 큐빅 프로젝션(Cubic Projection) 스킴을 이용하여 프로젝션이 수행될 수 있다. 프로젝션 처리부는 큐빅 프로젝션 스킴을 이용하여 360도 비디오 데이터의 프로젝션을 수행할 수 있다. 상기 큐빅 프로젝션 스킴은 CMP(Cube Map Projection)라고 나타낼 수도 있다. 도 9a의 (b)는 큐빅 프로젝션 스킴을 나타낼 수 있다. 예를 들어 스티칭된 360도 비디오 데이터는 구형의 면 상에 나타내어질 수 있다. 프로젝션 처리부는 이러한 360도 비디오 데이터를 큐브(Cube, 정육면체) 형태로 나누어 2D 이미지 상에 프로젝션할 수 있다. 구형의 면 상의 360도 비디오 데이터는 큐브의 각 면에 대응되어, 2D 이미지 상에 도 9a의 (b) 좌측 또는 (b) 우측에 도시된 것과 같이 프로젝션될 수 있다.

[0153]

본 발명에 따른 또 다른 실시예에 의하면, 실린더형 프로젝션(Cylindrical Projection) 스킴을 이용하여 프로젝

[0154]

선이 수행될 수 있다. 프로젝션 처리부는 실린더형 프로젝션 스킴을 이용하여 360도 비디오 데이터의 프로젝션을 수행할 수 있다. 도 9a의 (c)는 실린더형 프로젝션 스킴을 나타낼 수 있다. 스티칭된 360도 비디오 데이터가 구형의 면 상에 나타내어질 수 있다고 가정할 때, 프로젝션 처리부는 이러한 360도 비디오 데이터를 실린더(Cylinder) 형태로 나누어 2D 이미지 상에 프로젝션할 수 있다. 구형의 면 상의 360도 비디오 데이터는 실린더의 옆면(side)과 윗면(top), 바닥면(bottom)에 각각 대응되어, 2D 이미지 상에 도 9a의 (c) 좌측 또는 (c) 우측에 도시된 것과 같이 프로젝션될 수 있다.

[0155] 본 발명에 따른 또 다른 실시예에 의하면, 타일-기반(Tile-based) 프로젝션 스킴을 이용하여 프로젝션이 수행될 수 있다. 프로젝션 처리부는 타일-기반 프로젝션 스킴을 이용하여 360도 비디오 데이터의 프로젝션을 수행할 수 있다. 도 9a의 (d)는 타일-기반 프로젝션 스킴을 나타낼 수 있다. 타일-기반(Tile-based) 프로젝션 스킴이 쓰이는 경우, 전술한 프로젝션 처리부는 구형 면 상의 360도 비디오 데이터를, 도 9a의 (d)에 도시된 것과 같이 하나 이상의 세부 영역으로 나누어 2D 이미지 상에 프로젝션할 수 있다. 상기 세부 영역은 타일이라고 불릴 수 있다.

[0156] 본 발명에 따른 또 다른 실시예에 의하면, 피라미드 프로젝션(Pyramid Projection) 스킴을 이용하여 프로젝션이 수행될 수 있다. 프로젝션 처리부는 피라미드 프로젝션 스킴을 이용하여 360도 비디오 데이터의 프로젝션을 수행할 수 있다. 도 9b의 (e)는 피라미드 프로젝션 스킴을 나타낼 수 있다. 스티칭된 360도 비디오 데이터가 구형의 면 상에 나타내어질 수 있다고 가정할 때, 프로젝션 처리부는 이러한 360도 비디오 데이터를 피라미드 형태로 보고, 각 면을 나누어 2D 이미지 상에 프로젝션할 수 있다. 구형의 면 상의 360도 비디오 데이터는 피라미드의 바닥면(front), 피라미드의 4방향의 옆면(Left top, Left bottom, Right top, Right bottom)에 각각 대응되어, 2D 이미지 상에 도 9b의 (e) 좌측 또는 (e) 우측에 도시된 것과 같이 프로젝션될 수 있다. 여기서, 상기 바닥면은 정면을 바라보는 카메라가 획득한 데이터를 포함하는 영역일 수 있다.

[0157] 본 발명에 따른 또 다른 실시예에 의하면, 파노라믹 프로젝션(Panoramic Projection) 스킴을 이용하여 프로젝션이 수행될 수 있다. 프로젝션 처리부는 파노라믹 프로젝션 스킴을 이용하여 360도 비디오 데이터의 프로젝션을 수행할 수 있다. 도 9b의 (f)는 파노라믹 프로젝션 스킴을 나타낼 수 있다. 파노라믹 프로젝션 스킴이 사용되는 경우, 전술한 프로젝션 처리부는, 도 9b의 (f)에 도시된 것과 같이 구형 면 상의 360도 비디오 데이터 중 옆면만을 2D 이미지 상에 프로젝션할 수 있다. 이는 실린더형 프로젝션 스킴에서 윗면(top)과 바닥면(bottom)이 존재하지 않는 경우와 같을 수 있다.

[0158] 본 발명에 따른 또 다른 실시예에 의하면, 스티칭없이 프로젝션이 수행될 수 있다. 도 9b의 (g)는 스티칭없이 프로젝션이 수행되는 경우를 나타낼 수 있다. 스티칭없이 프로젝션되는 경우, 전술한 프로젝션 처리부는, 도 9b의 (g)에 도시된 것과 같이, 360도 비디오 데이터를 그대로 2D 이미지 상에 프로젝션할 수 있다. 이 경우 스티칭은 수행되지 않고, 카메라에서 획득된 각각의 이미지들이 그대로 2D 이미지 상에 프로젝션될 수 있다.

[0159] 도 9b의 (g)를 참조하면 두 개의 이미지가 2D 이미지 상에 스티칭없이 프로젝션될 수 있다. 각 이미지는 구형 카메라(spherical camera)에서 각 센서를 통해 획득한 어안(fish-eye) 이미지일 수 있다. 전술한 바와 같이, 수신측에서 카메라 센서들로부터 획득하는 이미지 데이터를 스티칭할 수 있고, 스티칭된 이미지 데이터를 구형면(spherical surface) 상에 맵핑하여 구형 비디오(spherical video), 즉, 360도 비디오를 렌더링할 수 있다.

[0160] 도 10a 내지 도 10b는 본 발명의 일 실시예에 따른 타일(Tile)이 도시된 도면을 예시적으로 나타낸다. 2D 이미지에 프로젝션된 360도 비디오 데이터 또는 리전별 패키징까지 수행된 360도 비디오 데이터는 하나 이상의 타일로 구분될 수 있다. 즉, 프로젝션된 픽처의 360도 비디오 데이터 또는 패키징된 픽처의 360도 비디오 데이터는 하나 이상의 타일로 구분될 수 있다. 도 10a는 16개의 타일로 나뉘어진 하나의 2D 이미지를 나타낼 수 있다. 여기서 상기 2D 이미지는 상술한 프로젝션된 픽처 또는 패키징된 픽처를 나타낼 수 있다. 본 발명에 따른 360도 비디오 전송 장치의 다른 실시예에 의하면, 360도 비디오 전송 장치의 데이터 인코더는 각각의 타일을 독립적으로 인코딩할 수 있다.

[0161] 한편, 상술한 리전별 패키징과 타일링(Tiling)은 구분될 수 있다. 전술한 리전별 패키징은 코딩 효율을 높이기 위해 또는 레졸루션을 조정하기 위하여 2D 이미지상에 프로젝션된 360도 비디오 데이터를 리전으로 구분하여 처리하는 것을 의미할 수 있다. 타일링은 데이터 인코더가 프로젝션된 픽처 또는 패키징된 픽처를 타일이라는 구획별로 나누고, 해당 타일들 별로 독립적으로 인코딩을 수행하는 것을 의미할 수 있다. 360도 비디오가 제공될 때, 사용자는 360도 비디오의 모든 부분을 동시에 소비하지 않을 수 있는바, 타일링은 제한된 밴드위스(bandwidth)상에서 사용자가 현재 보는 뷰포트 등 중요 부분 내지 일정 부분에 해당하는 타일만을 수신측으로 전송 혹은 소비하는 것을 가능케할 수 있다. 즉, 타일링을 통해 제한된 밴드위스가 더 효율적으로 활용될 수 있고, 수신측에서

도 모든 360도 비디오 데이터를 한번에 다 처리하는 것에 비하여 연산 부하를 줄일 수 있다.

- [0162] 리전과 타일은 구분되므로, 두 영역이 같을 필요는 없다. 그러나 실시예에 따라 리전과 타일은 같은 영역을 지칭할 수도 있다. 실시예에 따라 타일에 맞추어 리전별 패키징이 수행되어 리전과 타일이 같아질 수 있다. 또한 실시예에 따라, 프로젝션 스킵에 따른 각 면과 리전이 같은 경우, 프로젝션 스킵에 따른 각 면, 리전, 타일이 같은 영역을 지칭할 수도 있다. 문맥에 따라 리전은 VR 리전, 타일은 타일 리전으로 불릴 수도 있다.
- [0163] ROI (Region of Interest) 는 360 콘텐츠 제공자가 제안하는, 사용자들의 관심 영역을 의미할 수 있다. 360 콘텐츠 제공자는 360도 비디오를 제작할 때, 어느 특정 영역을 사용자들이 관심있어 할 것으로 예측할 수 있고, 이를 고려하여 360도 비디오를 제작할 수 있다. 실시예에 따라 ROI 는 360도 비디오의 콘텐츠 상, 중요한 내용이 재생되는 영역에 해당할 수 있다.
- [0164] 본 발명에 따른 360도 비디오 전송/수신 장치의 또 다른 실시예에 의하면, 수신측 피드백 처리부는 뷰포트 정보를 추출, 수집하여 이를 송신측 피드백 처리부로 전달할 수 있다. 이 과정에서 뷰포트 정보는 양 측의 네트워크 인터페이스를 이용해 전달될 수 있다. 도 10a는 2D 이미지에서의 뷰포트(1000)를 나타낸다. 여기서, 도 10a에 도시된 것과 같이 뷰포트(1000)는 2D 이미지 상의 9 개의 타일에 걸쳐 있을 수 있다.
- [0165] 이 경우, 360도 비디오 전송 장치는 타일링 시스템을 더 포함할 수 있다. 실시예에 따라 상기 타일링 시스템은 데이터 인코더 다음에 위치할 수도 있다. 도 10b는 데이터 인코더 다음에 위치하는 타일링 시스템을 나타낼 수 있다. 또는, 타일링 시스템은 상술한 데이터 인코더 내지 전송 처리부 내에 포함될 수도 있고, 별개의 내/외부 엘리먼트로서 360도 비디오 전송 장치에 포함될 수도 있다.
- [0166] 도 10b를 참조하면 타일링 시스템은 송신측 피드백 처리부로부터 뷰포트 정보를 전달받을 수 있다. 타일링 시스템은 뷰포트 영역이 포함되는 타일만을 선별하여 전송할 수 있다. 예를 들어, 도 10a에 도시된 2D 이미지에서 총 16 개의 타일 중 뷰포트 영역(1000)을 포함하는 9 개의 타일들만이 전송될 수 있다. 여기서 타일링 시스템은 브로드밴드를 통한 유니캐스트 방식으로 타일들을 전송할 수 있다. 사용자에 따라 뷰포트 영역이 다르기 때문이다.
- [0167] 또한 이 경우 송신측 피드백 처리부는 뷰포트 정보를 데이터 인코더로 전달할 수 있다. 데이터 인코더는 뷰포트 영역을 포함하는 타일들에 대해 다른 타일들보다 더 높은 퀄리티로 인코딩을 수행할 수 있다.
- [0168] 또한 이 경우 송신측 피드백 처리부는 뷰포트 정보를 메타데이터 처리부로 전달할 수 있다. 메타데이터 처리부는 뷰포트 영역과 관련된 메타데이터를 360도 비디오 전송 장치의 각 내부 엘리먼트로 전달해주거나, 360도 비디오 관련 메타데이터에 포함시킬 수 있다.
- [0169] 이러한 타일링 방식을 통하여, 전송 밴드위스(bandwidth)가 절약될 수 있으며, 타일 별로 차등화된 처리를 수행하여 효율적 데이터 처리/전송이 가능해질 수 있다.
- [0170] 전술한 뷰포트 영역과 관련된 실시예들은 뷰포트 영역이 아닌 다른 특정 영역들에 대해서도 유사한 방식으로 적용될 수 있다. 예를 들어, 전술한 게이즈 분석을 통해 사용자들이 주로 관심있어 하는 것으로 판단된 영역, ROI 영역, 사용자가 VR 디스플레이를 통해 360도 비디오를 접할 때 처음으로 재생되는 영역(초기 시점, Initial Viewpoint) 등에 대해서도, 전술한 뷰포트 영역과 같은 방식의 처리들이 수행될 수 있다.
- [0171] 또한, 본 발명에 따른 360도 비디오 전송 장치의 또 다른 실시예에 의하면, 전송 처리부는 각 타일 별로 다르게 전송을 위한 처리를 수행할 수 있다. 전송 처리부는 타일 별로 다른 전송 파라미터(모듈레이션 오더, 코드 레이트 등)를 적용하여, 각 타일 별로 전달되는 데이터의 강건성(robustness)을 다르게 할 수 있다.
- [0172] 이 때, 송신측 피드백 처리부는 360도 비디오 수신 장치로부터 전달받은 피드백 정보를 전송 처리부로 전달하여, 전송 처리부가 타일별 차등화된 전송 처리를 수행하도록 할 수 있다. 예를 들어 송신측 피드백 처리부는 수신측으로부터 전달받은 뷰포트 정보를 전송 처리부로 전달할 수 있다. 전송 처리부는 해당 뷰포트 영역을 포함하는 타일들에 대해 다른 타일들보다 더 높은 강건성을 가지도록 전송 처리를 수행할 수 있다.
- [0173] 도 11은 본 발명의 일 실시예에 따른 360도 비디오 관련 메타데이터의 일 예를 나타낸다. 상술한 내용과 같이 360도 비디오 관련 메타데이터는 360도 비디오에 대한 다양한 메타데이터를 포함할 수 있다. 문맥에 따라, 360도 비디오 관련 메타데이터는 360도 비디오 관련 시그널링 정보라고 불릴 수도 있다. 360도 비디오 관련 메타데이터는 별도의 시그널링 테이블에 포함되어 전송될 수도 있고, DASH MPD 내에 포함되어 전송될 수도 있고, ISOBMMF 등의 파일 포맷에 box 형태로 포함되어 전달될 수도 있다. 360도 비디오 관련 메타데이터가 box 형태로 포함되는 경우 파일, 프래그먼트, 트랙, 샘플 엔트리, 샘플 등등 다양한 레벨에 포함되어 해당되는 레벨의 데이

터에 대한 메타데이터를 포함할 수 있다.

- [0174] 실시예에 따라, 후술하는 메타데이터의 일부는 시그널링 테이블로 구성되어 전달되고, 나머지 일부는 파일 포맷 내에 box 혹은 트랙 형태로 포함될 수도 있다.
- [0175] 본 발명에 따른 360도 비디오 관련 메타데이터의 일 실시예에 의하면, 360도 비디오 관련 메타데이터는 프로젝션 스킴 등에 관한 기본 메타데이터, 스테레오스코픽(stereoscopic) 관련 메타데이터, 초기 시점(Initial View/Initial Viewpoint) 관련 메타데이터, ROI 관련 메타데이터, FOV (Field of View) 관련 메타데이터 및/또는 크롭된 영역(cropped region) 관련 메타데이터를 포함할 수 있다. 실시예에 따라 360도 비디오 관련 메타데이터는 전술한 것 외에 추가적인 메타데이터를 더 포함할 수 있다.
- [0176] 본 발명에 따른 360도 비디오 관련 메타데이터의 실시예들은 전술한 기본 메타데이터, 스테레오스코픽 관련 메타데이터, 초기 시점 관련 메타데이터, ROI 관련 메타데이터, FOV 관련 메타데이터, 크롭된 영역 관련 메타데이터 및/또는 이후 추가될 수 있는 메타데이터들 중 적어도 하나 이상을 포함하는 형태일 수 있다. 본 발명에 따른 360도 비디오 관련 메타데이터의 실시예들은, 각각 포함하는 세부 메타데이터들의 경우의 수에 따라 다양하게 구성될 수 있다. 실시예에 따라 360도 비디오 관련 메타데이터는 전술한 것 외에 추가적인 정보들을 더 포함할 수도 있다.
- [0177] 기본 메타데이터에는 3D 모델 관련 정보, 프로젝션 스킴 관련 정보 등이 포함될 수 있다. 기본 메타데이터에는 vr\_geometry 필드, projection\_scheme 필드 등이 포함될 수 있다. 실시예에 따라 기본 메타데이터는 추가적인 정보들을 더 포함할 수도 있다.
- [0178] vr\_geometry 필드는 해당 360도 비디오 데이터가 지원하는 3D 모델의 타입을 지시할 수 있다. 전술한 바와 같이 360도 비디오 데이터가 3D 공간 상에 리-프로젝션되는 경우, 해당 3D 공간은 vr\_geometry 필드가 지시하는 3D 모델에 따른 타입으로 나타낼 수 있다. 실시예에 따라, 렌더링시에 사용되는 3D 모델은 vr\_geometry 필드가 지시하는 리-프로젝션에 사용되는 3D 모델과 다를 수도 있다. 이 경우, 기본 메타데이터는 렌더링시에 사용되는 3D 모델을 지시하는 필드를 더 포함할 수도 있다. 해당 필드가 0, 1, 2, 3 의 값을 가지는 경우, 3D 공간은 각각 구형(Sphere), 큐브(Cube), 실린더(Cylinder), 피라미드(Pyramid)의 3D 모델로 도출될 수 있다. 해당 필드가 나머지 값을 가지는 경우는 향후 사용을 위해 남겨둘 수 있다(Reserved for Future Use). 실시예에 따라 360도 비디오 관련 메타데이터는 해당 필드에 의해 지시되는 3D 모델에 대한 구체적인 정보를 더 포함할 수 있다. 여기서 3D 모델에 대한 구체적인 정보란 예를 들어 구형의 반지름 정보, 실린더의 높이 정보 등을 의미할 수 있다. 본 필드는 생략될 수 있다.
- [0179] 상기 projection\_scheme 필드는 해당 360도 비디오 데이터가 2D 이미지 상에 프로젝션될 때 사용된 프로젝션 스킴을 지시할 수 있다. 예를 들어, 상기 projection\_scheme 필드가 0, 1, 2, 3, 4, 5 의 값을 가지는 경우, 각각 등정방형 프로젝션(Equirectangular Projection) 스킴, 큐빅 프로젝션 스킴, 실린더형 프로젝션 스킴, 타일-베이스드(Tile-based) 프로젝션 스킴, 피라미드 프로젝션 스킴, 파노라믹 프로젝션 스킴을 나타낼 수 있다. 해당 필드가 6 의 값을 가지는 경우는, 360도 비디오 데이터가 스티칭 없이 바로 2D 이미지 상에 프로젝션되는 프로젝션 스킴을 나타낼 수 있다. 해당 필드가 나머지 값을 가지는 경우는 향후 사용을 위해 남겨둘 수 있다(Reserved for Future Use).
- [0180] 실시예에 따라 360도 비디오 관련 메타데이터는 해당 필드에 의해 특정되는 프로젝션 스킴에 의해 발생한 리전(Region)에 대한 구체적인 정보를 더 포함할 수 있다. 여기서 리전에 대한 구체적인 정보란 예를 들어 리전의 회전 여부, 실린더의 윗면(top) 리전의 반지름 정보 등을 의미할 수 있다. 프로젝션 스킴에 의해 발생한 리전(Region)에 대한 구체적인 정보에 대한 메타데이터는 projection\_scheme 필드 관련 메타데이터라고 불릴 수 있다. 상기 projection\_scheme 필드 관련 메타데이터는 지시된 각각의 프로젝션 스킴(등정방형 프로젝션 스킴, 큐빅 프로젝션 스킴, 실린더형 프로젝션 스킴, 피라미드 프로젝션 스킴, 파노라믹 프로젝션 스킴, 스티칭없이 프로젝션되는 경우 등)에 대한 구체적인 정보를 제공할 수 있다.
- [0181] 스테레오스코픽 관련 메타데이터는 360도 비디오 데이터의 3D 관련 속성들에 대한 정보들을 포함할 수 있다. 스테레오스코픽 관련 메타데이터는 is\_stereoscopic 필드 및/또는 stereo\_mode 필드를 포함할 수 있다. 실시예에 따라 스테레오스코픽 관련 메타데이터는 추가적인 정보들을 더 포함할 수도 있다.
- [0182] is\_stereoscopic 필드는 해당 360도 비디오 데이터가 3D 를 지원하는지 여부를 지시할 수 있다. 해당 필드가 1 이면 3D 지원, 0 이면 3D 미지원을 의미할 수 있다. 본 필드는 생략될 수 있다.
- [0183] stereo\_mode 필드는 해당 360도 비디오가 지원하는 3D 레이아웃을 지시할 수 있다. 본 필드만으로 해당 360도

비디오가 3D 를 지원하는지 여부를 지시할 수도 있는데, 이 경우 전술한 is\_stereoscopic 필드는 생략될 수 있다. 본 필드 값이 0 인 경우, 해당 360도 비디오는 모노(mono) 모드일 수 있다. 즉 프로젝션된 2D 이미지는 하나의 모노 뷰(mono view) 만을 포함할 수 있다. 이 경우 해당 360도 비디오는 3D 를 지원하지 않을 수 있다.

[0184] 본 필드 값이 1, 2 인 경우, 해당 360도 비디오는 각각 좌우(Left-Right) 레이아웃, 상하(Top-Bottom) 레이아웃에 따를 수 있다. 좌우 레이아웃, 상하 레이아웃은 각각 사이드-바이-사이드 포맷, 탑-바텀 포맷으로 불릴 수도 있다. 좌우 레이아웃의 경우, 좌영상/우영상이 프로젝션된 2D 이미지들은 이미지 프레임 상에서 각각 좌/우로 위치할 수 있다. 상하 레이아웃의 경우, 좌영상/우영상이 프로젝션된 2D 이미지들은 이미지 프레임 상에서 각각 위/아래로 위치할 수 있다. 해당 필드가 나머지 값을 가지는 경우는 향후 사용을 위해 남겨둘 수 있다 (Reserved for Future Use).

[0185] 초기 시점 관련 메타데이터는 사용자가 360도 비디오를 처음 재생했을 때 보게되는 시점(초기 시점)에 대한 정보를 포함할 수 있다. 초기 시점 관련 메타데이터는 initial\_view\_yaw\_degree 필드, initial\_view\_pitch\_degree 필드 및/또는 initial\_view\_roll\_degree 필드를 포함할 수 있다. 실시예에 따라 초기 시점 관련 메타데이터는 추가적인 정보들을 더 포함할 수도 있다.

[0186] initial\_view\_yaw\_degree 필드, initial\_view\_pitch\_degree 필드, initial\_view\_roll\_degree 필드는 해당 360도 비디오 재생 시의 초기 시점을 나타낼 수 있다. 즉, 재생시 처음 보여지는 뷰포트의 정중앙 지점이, 이 세 필드들에 의해 나타내어질 수 있다. 구체적으로, 상기 initial\_view\_yaw\_degree 필드는 상기 초기 시점에 대한 yaw 값을 나타낼 수 있다. 즉, 상기 initial\_view\_yaw\_degree 필드는 상기 정중앙 지점의 위치를 yaw 축을 기준으로 회전된 방향(부호) 및 그 정도(각도)로 나타낼 수 있다. 또한, 상기 initial\_view\_pitch\_degree 필드는 상기 초기 시점에 대한 pitch 값을 나타낼 수 있다. 즉, 상기 initial\_view\_pitch\_degree 필드는 상기 정중앙 지점의 위치를 pitch 축을 기준으로 회전된 방향(부호) 및 그 정도(각도)로 나타낼 수 있다. 또한, 상기 initial\_view\_roll\_degree 필드는 상기 초기 시점에 대한 roll 값을 나타낼 수 있다. 즉, 상기 initial\_view\_roll\_degree 필드는 상기 정중앙 지점의 위치를 roll 축을 기준으로 회전된 방향(부호) 및 그 정도(각도)로 나타낼 수 있다. 상기 initial\_view\_yaw\_degree 필드, 상기 initial\_view\_pitch\_degree 필드, 상기 initial\_view\_roll\_degree 필드를 기반으로 해당 360도 비디오 재생 시의 초기 시점, 즉, 재생시 처음 보여지는 뷰포트의 정중앙 지점을 나타낼 수 있고, 이를 통하여 상기 360도 비디오의 특정 영역이 사용자에게 초기 시점에 디스플레이되어 제공될 수 있다. 또한, FOV(field of view)를 통하여, 지시된 초기 시점을 기준으로 한, 초기 뷰포트의 가로길이 및 세로길이(width, height) 가 결정될 수 있다. 즉, 이 세 필드들 및 FOV 정보를 이용하여, 360도 비디오 수신 장치는 사용자에게 360도 비디오의 일정 영역을 초기 뷰포트로서 제공할 수 있다.

[0187] 실시예에 따라, 초기 시점 관련 메타데이터가 지시하는 초기 시점은, 장면(scene) 별로 변경될 수 있다. 즉, 360 콘텐츠의 시간적 흐름에 따라 360도 비디오의 장면이 바뀌게 되는데, 해당 360도 비디오의 장면마다 사용자가 처음 보게되는 초기 시점 내지 초기 뷰포트가 변경될 수 있다. 이 경우, 초기 시점 관련 메타데이터는 각 장면별로의 초기 시점을 지시할 수 있다. 이를 위해 초기 시점 관련 메타데이터는, 해당 초기 시점이 적용되는 장면을 식별하는 장면(scene) 식별자를 더 포함할 수도 있다. 또한 360도 비디오의 장면별로 FOV(Field Of View) 가 변할 수도 있으므로, 초기 시점 관련 메타데이터는 해당 장면에 해당하는 FOV를 나타내는 장면별 FOV 정보를 더 포함할 수도 있다.

[0188] ROI 관련 메타데이터는 전술한 ROI 에 관련된 정보들을 포함할 수 있다. ROI 관련 메타데이터는, 2d\_roi\_range\_flag 필드 및/또는 3d\_roi\_range\_flag 필드를 포함할 수 있다. 2d\_roi\_range\_flag 필드는 ROI 관련 메타데이터가 2D 이미지를 기준으로 ROI 를 표현하는 필드들을 포함하는지 여부를 지시할 수 있고, 3d\_roi\_range\_flag 필드는 ROI 관련 메타데이터가 3D 공간을 기준으로 ROI 를 표현하는 필드들을 포함하는지 여부를 지시할 수 있다. 실시예에 따라 ROI 관련 메타데이터는, ROI 에 따른 차등 인코딩 정보, ROI 에 따른 차등 전송처리 정보 등 추가적인 정보들을 더 포함할 수도 있다.

[0189] ROI 관련 메타데이터가 2D 이미지를 기준으로 ROI 를 표현하는 필드들을 포함하는 경우, ROI 관련 메타데이터는 min\_top\_left\_x 필드, max\_top\_left\_x 필드, min\_top\_left\_y 필드, max\_top\_left\_y 필드, min\_width 필드, max\_width 필드, min\_height 필드, max\_height 필드, min\_x 필드, max\_x 필드, min\_y 필드 및/또는 max\_y 필드를 포함할 수 있다.

[0190] min\_top\_left\_x 필드, max\_top\_left\_x 필드, min\_top\_left\_y 필드, max\_top\_left\_y 필드는 ROI 의 좌측 상단 끝의 좌표의 최소/최대값을 나타낼 수 있다. 즉, 상기 필드들은 차례로 좌상단 끝의 최소 x 좌표, 최대 x 좌표, 최소 y 좌표, 최대 y 좌표를 나타낼 수 있다.

- [0191] min\_width 필드, max\_width 필드, min\_height 필드, max\_height 필드는 ROI 의 가로 크기(width), 세로 크기(height)의 최소/최대값을 나타낼 수 있다. 즉, 상기 필드들은 차례로 가로 크기의 최소값, 가로 크기의 최대값, 세로 크기의 최소값, 세로 크기의 최대값을 나타낼 수 있다.
- [0192] min\_x 필드, max\_x 필드, min\_y 필드, max\_y 필드는 ROI 내의 좌표들의 최소/최대값을 나타낼 수 있다. 즉, 상기 필드들은 차례로 ROI 내 좌표들의 최소 x 좌표, 최대 x 좌표, 최소 y 좌표, 최대 y 좌표를 나타낼 수 있다. 이 필드들은 생략될 수 있다.
- [0193] ROI 관련 메타데이터가 3D 랜더링 공간 상의 좌표 기준으로 ROI 를 표현하는 필드들을 포함하는 경우, ROI 관련 메타데이터는 min\_yaw 필드, max\_yaw 필드, min\_pitch 필드, max\_pitch 필드, min\_roll 필드, max\_roll 필드, min\_field\_of\_view 필드 및/또는 max\_field\_of\_view 필드를 포함할 수 있다.
- [0194] min\_yaw 필드, max\_yaw 필드, min\_pitch 필드, max\_pitch 필드, min\_roll 필드, max\_roll 필드는 ROI 가 3D 공간상에서 차지하는 영역을 yaw, pitch, roll 의 최소/최대값으로 나타낼 수 있다. 즉, 상기 필드들은 차례로 yaw 축 기준 회전량의 최소값, yaw 축 기준 회전량의 최대값, pitch 축 기준 회전량의 최소값, pitch 축 기준 회전량의 최대값, roll 축 기준 회전량의 최소값, roll 축 기준 회전량의 최대값을 나타낼 수 있다.
- [0195] min\_field\_of\_view 필드, max\_field\_of\_view 필드는 해당 360도 비디오 데이터의 FOV(Field Of View)의 최소/최대값을 나타낼 수 있다. FOV 는 360도 비디오의 재생시 한번에 디스플레이되는 시야범위를 의미할 수 있다. min\_field\_of\_view 필드, max\_field\_of\_view 필드는 각각 FOV 의 최소값, 최대값을 나타낼 수 있다. 이 필드들은 생략될 수 있다. 이 필드들은 후술할 FOV 관련 메타데이터에 포함될 수도 있다.
- [0196] FOV 관련 메타데이터는 전술한 FOV 에 관련한 정보들을 포함할 수 있다. FOV 관련 메타데이터는 content\_fov\_flag 필드 및/또는 content\_fov 필드를 포함할 수 있다. 실시예에 따라 FOV 관련 메타데이터는 전술한 FOV 의 최소/최대값 관련 정보 등 추가적인 정보들을 더 포함할 수도 있다.
- [0197] content\_fov\_flag 필드는 해당 360도 비디오에 대하여 제작시 의도한 FOV 에 대한 정보가 존재하는지 여부를 지시할 수 있다. 본 필드값이 1인 경우, content\_fov 필드가 존재할 수 있다.
- [0198] content\_fov 필드는 해당 360도 비디오에 대하여 제작시 의도한 FOV 에 대한 정보를 나타낼 수 있다. 실시예에 따라 해당 360도 비디오 수신 장치의 수직(vertical) 혹은 수평(horizontal) FOV 에 따라, 360 영상 중에서 사용자에게 한번에 디스플레이되는 영역이 결정될 수 있다. 혹은 실시예에 따라 본 필드의 FOV 정보를 반영하여 사용자에게 한번에 디스플레이되는 360도 비디오의 영역이 결정될 수도 있다.
- [0199] 크롭된 영역 관련 메타데이터는 이미지 프레임 상에서 실제 360도 비디오 데이터를 포함하는 영역에 대한 정보를 포함할 수 있다. 이미지 프레임은 실제 360도 비디오 데이터 프로젝션된 액티브 비디오 영역(Active Video Area)과 그렇지 않은 영역을 포함할 수 있다. 이 때 액티브 비디오 영역은 크롭된 영역 또는 디폴트 디스플레이 영역이라고 칭할 수 있다. 이 액티브 비디오 영역은 실제 VR 디스플레이 상에서 360도 비디오로서 보여지는 영역으로서, 360도 비디오 수신 장치 또는 VR 디스플레이는 액티브 비디오 영역만을 처리/디스플레이할 수 있다. 예를 들어 이미지 프레임의 종횡비(aspect ratio) 가 4:3 인 경우 이미지 프레임의 윗 부분 일부와 아랫부분 일부를 제외한 영역만 360도 비디오 데이터를 포함할 수 있는데, 이 부분을 액티브 비디오 영역이라고 할 수 있다.
- [0200] 크롭된 영역 관련 메타데이터는 is\_cropped\_region 필드, cr\_region\_left\_top\_x 필드, cr\_region\_left\_top\_y 필드, cr\_region\_width 필드 및/또는 cr\_region\_height 필드를 포함할 수 있다. 실시예에 따라 크롭된 영역 관련 메타데이터는 추가적인 정보들을 더 포함할 수도 있다.
- [0201] is\_cropped\_region 필드는 이미지 프레임의 전체 영역이 360도 비디오 수신 장치 내지 VR 디스플레이에 의해 사용되는지 여부를 나타내는 플래그일 수 있다. 여기서, 360도 비디오 데이터가 매핑된 영역 혹은 VR 디스플레이 상에서 보여지는 영역은 액티브 비디오 영역(Active Video Area)라고 불릴 수 있다. 상기 is\_cropped\_region 필드는 이미지 프레임 전체가 액티브 비디오 영역인지 여부를 지시할 수 있다. 이미지 프레임의 일부만이 액티브 비디오 영역인 경우, 하기의 4 필드가 더 추가될 수 있다.
- [0202] cr\_region\_left\_top\_x 필드, cr\_region\_left\_top\_y 필드, cr\_region\_width 필드, cr\_region\_height 필드는 이미지 프레임 상에서 액티브 비디오 영역을 나타낼 수 있다. 이 필드들은 각각 액티브 비디오 영역의 좌상단의 x 좌표, 액티브 비디오 영역의 좌상단의 y 좌표, 액티브 비디오 영역의 가로 길이(width), 액티브 비디오 영역의 세로 길이(height)를 나타낼 수 있다. 가로 길이와 세로 길이는 픽셀을 단위로 나타내어질 수 있다.



- [0203] 한편, 본 발명은 사용자에게 몰입형 미디어/실감미디어(Immersive media)의 경험을 제공하기 위해, 상술한 360도 콘텐츠를 제공하는 방안이외에 6DoF(6 Degrees of Freedom) 콘텐츠를(contents)를 제공하는 방안을 제안한다.
- [0204] 도 12는 상기 360도 콘텐츠와 상기 6DoF 콘텐츠를 예시적으로 나타낸다.
- [0205] 상기 몰입형 미디어/실감미디어는 상기 360 콘텐츠가 제공하는 가상의 환경에서 확대된 개념을 나타낼 수 있다. 도 12의 (a)는 상기 360 콘텐츠의 일 예를 도시할 수 있다. 도 12의 (a)에 도시된 것과 같이 상기 360도 콘텐츠가 제공되는 경우에는 사용자의 position 위치가 고정되어 있고, 회전에 대한 개념만 제공될 수 있다. 한편, 도 12의 (b) 및 (c)는 상기 6DoF 콘텐츠의 일 예를 나타낼 수 있다. 도 12의 (b) 및 (c)를 참조하면 상기 몰입형 미디어/실감미디어는 사용자에게 콘텐츠를 경험하는동안 이동의 개념을 부여함으로써 가상의 공간에서 사용자의 이동/회전 등 더 다양한 감각적 경험을 제공할 수 있는 환경 혹은 콘텐츠를 의미할 수 있다.
- [0206] 실감미디어 콘텐츠는 해당 콘텐츠를 제공하기 위한 6DoF비디오 및 6DoF오디오를 포함할 수 있으며, 6DoF 비디오는 실감미디어 콘텐츠 제공에 필요한 매 이동 때마다 새롭게 형성되는 3DoF 혹은 360도 비디오로 캡처되거나 재생되는 비디오 혹은 이미지를 의미 할 수 있다. 여기서, 상기 실감미디어 콘텐츠는 상기 6DoF 콘텐츠라고 나타낼 수 있다. 또한, 상기 6DoF 콘텐츠는 3차원 공간 상에 나타내어지는 비디오 혹은 이미지를 의미할 수 있다. 콘텐츠 내에서 이동이 고정된 상태라면 해당 콘텐츠는 기존의 360도 비디오와 같이 다양한 형태의 3차원 공간에서 나타내어질 수 있다. 예를 들어, 구형(Spherical)면 상에 나타내어질 수 있다. 콘텐츠 내에서 이동이 자유로운 상태라면 이동 경로 상에 사용자를 중심으로 3차원 공간이 매번 새롭게 형성되고, 해당 위치의 콘텐츠를 사용자가 경험할 수 있다. 예를 들어 사용자가 처음 보는 위치에서의 구형(spherical)면 상에 나타내어진 영상을 경험하고, 3차원 공간에서 실제 사용자가 이동을 한 경우, 이동한 위치를 중심으로 새로운 구형(spherical)면의 영상이 형성되어 해당 콘텐츠를 소비할 수 있다. 6DoF 오디오도 마찬가지로 실감형 미디어를 경험할 수 있도록 하는 콘텐츠를 제공하기 위한 오디오 콘텐츠로, 음향의 소비하는 위치가 이동함에 따른 공간적(spatial)오디오를 새롭게 형성하고 소비하기 위한 콘텐츠를 의미할 수 있다.
- [0207] 본 발명은 특히 6DoF 비디오를 효과적으로 제공하는 방안을 제안한다. 6DoF 비디오는 서로 다른 위치에서 두 개 이상의 카메라로 캡처될 수 있다. 캡처된 비디오는 일련의 과정을 거쳐 전송될 수 있고, 수신측에서는 수신된 데이터 중 일부를 사용자의 초기 위치를 원점으로 하는 360비디오로 가공하여 렌더링 할 수 있으며 사용자의 위치가 이동된 경우, 이동한 위치를 중심으로 새로운 360도 비디오를 가공하여 렌더링 함으로써 6DoF 비디오가 사용자에게 제공될 수 있다.
- [0208] 도 13은 본 발명에 따른 6DoF 비디오 제공을 위한 전체 아키텍처를 도시한 도면이다.
- [0209] 도 13을 참조하면 6DoF 비디오 데이터 및/또는 6DoF 오디오 데이터가 획득될 수 있다(Acquisition). 구체적으로 상기 6DoF 콘텐츠의 캡처를 위해 HDCA(High Density Camera Array), Lenslet(microlens) camera 등이 사용될 수 있으며, 6DoF 비디오 캡처를 위해 디자인된 새로운 디바이스를 통하여 획득될 수도 있다. 도 13에 도시된 획득된 영상(1310)과 같이 캡처한 카메라의 위치에 따라 생성된 이미지/비디오 데이터 집합이 복수의 개수로 생성될 수 있다. 상기 6DoF 콘텐츠의 캡처 과정에서 카메라의 내부/외부 설정 값 등을 나타내는 메타데이터가 생성될 수 있다. 한편, 카메라가 아닌 컴퓨터로 생성된 영상의 경우 상기 캡처 과정이 같음될 수 있다. 획득된 영상(1310)의 전처리(pre-processing) 과정은 캡처된 이미지/비디오 및 캡처 과정에서 전달된 메타데이터(metadata)를 처리하는 과정일 수 있다. 상기 전처리(pre-processing) 과정은 스티칭(Stitching) 과정, 색보정(color correction)과정, 프로젝션 과정, 코딩 효율을 높이기 위해 주요 시점(primary view)와 부차 시점(secondary view)로 분리 하는 시점 분리(view segmenation)과정 및 인코딩 과정 등 전송 전 콘텐츠를 처리하는 모든 형태의 전처리 단계를 포함할 수 있다.
- [0210] 구체적으로, 상기 스티칭 과정은 각 카메라의 위치에서 360 방향으로 캡처된 영상을 각각의 카메라 위치를 중심으로 하는 파노라마 혹은 구형의 형태로 영상을 잇는 이미지/비디오를 만드는 과정일 수 있다.
- [0211] 이 후, 스티칭된 이미지/비디오는 프로젝션(Projection) 과정을 거칠 수 있다. 프로젝션 과정은 각각의 스티칭된 영상을 2D 이미지로 투영시켜 프로젝션된 픽처(1320)를 도출하는 과정을 의미할 수 있다. 여기서, 상기 프로젝션은 2D 이미지로 맵핑한다고 표현할 수도 있다. 각 카메라 위치에서 맵핑한 영상은 주요시점과 부차 시점으로 분리할 수 있고, 비디오 코딩 효율을 높이기 위해 시점별 다른 해상도(resolution)를 적용할 수 있으며, 주요 시점 내에서도 맵핑 영상의 배치나 해상도(resolution)를 달리함으로써 코딩 시 효율을 높일 수 있다. 상기 부차 시점은 캡처 환경에 따라 없을 수도 있다. 부차 시점은 주요 시점에서 또 다른 주요 시점으로 사용자가 이동하는 경우에 이동 과정에서 재생되어야 하는 이미지/비디오를 의미할 수 있고, 주요 시점에 비해 낮은 해상도

를 가질 수도 있으나 필요에 따라 동일한 해상도를 가질 수도 있다. 또한, 때에 따라서는 수신기에서 부차 시점이 가상의 정보로 새롭게 생성될 수 있다.

- [0212] 실시예에 따라 전처리 과정은 에디팅(editing)과정 등을 더 포함할 수 있다. 상기 전처리 과정에서 프로젝션 전후의 이미지/비디오 데이터들에 대한 편집 등이 더 수행될 수 있으며, 메타메이터가 생성될 수 있다. 또한, 전처리 과정에서 이미지/비디오 제공시 가장 처음 재생해야 하는 초기 시점, 사용자의 초기 위치 및 ROI(Region of Interest)등에 관한 메타메이터가 생성될 수 있다.
- [0213] 도 13에 도시된 미디어 전송 과정(Delivery)은 전처리 과정에서 얻어진 이미지/비디오 데이터 및 메타메이터들을 처리하여 전송하는 과정을 나타낼 수 있다. 상기 이미지/비디오 데이터 및 상기 메타메이터들의 전송을 위하여 임의의 전송 프로토콜에 따른 처리가 수행될 수 있으며, 전처리된 데이터들은 방송망 및/또는 브로드밴드를 통해 전달될 수 있다. 또한, 상기 전처리된 데이터들은 온디맨드(on demand) 방식으로 수신측으로 전달될 수 있다.
- [0214] 프로세싱 과정은 수신된 이미지/비디오 데이터 및 메타메이터를 디코딩하는 과정과 디코딩된 프로젝션된 픽처의 이미지/비디오 데이터를 3차원(3 Dimension, 3D) 모델로 맵핑 혹은 프로젝션하는 리-프로젝션(re-projection) 과정, 가상 시점의 생성 및 합성 과정 등 이미지/비디오를 재생하기 위한 이미지 생성 전 모든 과정을 포함할 수 있다. 맵핑되는 3D 모델 혹은 프로젝션 맵은 기존의 360도 비디오와 같이 구형(sphere), 큐브(cube), 실린더(cylinder), 또는 피라미드(pyramid)가 있을 수 있다. 또는, 상기 3D 모델 혹은 상기 프로젝션 맵은 기존의 360도 비디오의 프로젝션 맵의 변형된 형태가 될 수 있으며, 경우에 따라서는 자유형 형태의 프로젝션 맵이 될 수도 있다.
- [0215] 여기서, 가상 시점의 생성 및 합성 과정은 주요 시점과 부차 시점 사이에 혹은 주요 시점과 주요 시점 사이로 사용자가 이동하는 경우에 재생되어야 하는 이미지/비디오 데이터를 생성하고 합성하는 과정을 나타낼 수 있다. 가상 시점 생성을 위해 캡처 및 전처리 과정에서 전달된 메타메이터를 처리하는 과정이 필요할 수 있고, 경우에 따라서는 가상 시점에서 360 이미지/비디오 전체가 아닌 일부만 생성/합성될 수도 있다.
- [0216] 한편, 실시예에 따라 상기 프로세싱 과정은 부가적으로 에디팅(editing)과정, 업스케일링(up scaling), 다운 스케일링(down scaling) 과정 등을 더 포함할 수도 있다. 에디팅 과정에서 프로세싱 과정 후에 재생 전 필요한 추가 편집 과정이 적용될 수 있다. 필요에 따라서는 전송 받은 이미지/비디오를 업스케일링 혹은 다운 스케일링 하는 작업이 수행될 수도 있다.
- [0217] 렌더링 과정은 전송 혹은 생성되어 리프로젝션된 이미지/비디오를 디스플레이할 수 있도록 렌더링 하는 과정을 나타낼 수 있다. 때에 따라서는 렌더링과 리프로젝션 과정을 렌더링이라고 통칭하기도 한다. 따라서 렌더링 과정 중에 리프로젝션 과정이 포함될 수도 있다. 리프로젝션은 도 13의 (1330)과 같은 형태로 사용자 중심의 360도 비디오/이미지와 사용자가 이동 방향에 따라 이동한 위치 각각을 중심으로 형성되는 360도 비디오/이미지가 형성되는 형태로 다수의 리프로젝션 결과물이 있을 수 있다. 사용자는 디스플레이할 디바이스에 따라 360도 비디오/이미지의 일부 영역을 볼 수 있으며, 이 때 사용자가 보게 되는 영역은 도 13의 (1340)과 같은 형태로 도출 될 수 있다. 또한, 사용자가 이동하는 경우 전체 360도 비디오/이미지가 렌더링 되는 것이 아니라 사용자가 보고 있는 위치에 해당되는 영상만 렌더링될 수 있다. 또한 6DoF 비디오 수신 장치는 사용자의 위치와 이동 방향에 관한 메타메이터를 전달 받아 미리 사용자의 움직임을 예측하고 이동할 위치(즉, 이동할 것으로 예측된 위치)의 비디오/이미지를 추가로 렌더링할 수 있다.
- [0218] 피드백 과정은 디스플레이 과정에서 획득될 수 있는 다양한 피드백 정보들을 송신측으로 전달하는 과정을 나타낼 수 있다. 피드백 과정을 통해 6DoF 콘텐츠와 사용자간의 인터랙티비티(interactivity)가 일어날 수 있으며, 실시예에 따라 피드백 과정에서 사용자의 머리와 포지션 위치 정보(head/position orientation) 및 사용자가 현재 보고 있는 영역(viewport)에 대한 정보 등이 전달될 수도 있다. 해당 정보는 피드백 과정에서 송신측 혹은 서비스 제공자 측에 전달 될 수 있으며, 실시예에 따라 피드백 과정은 수행되지 않을 수도 있다.
- [0219] 사용자의 위치 정보는 사용자의 머리 위치, 각도, 움직임 및 이동 거리 등에 대한 정보를 의미할 수 있으며, 해당 정보를 바탕으로 사용자가 보고 있는 위치(viewport) 정보가 계산될 수 있다.
- [0220] 도 14는 본 발명이 적용될 수 있는 6DoF 비디오 전송 장치의 구성을 개략적으로 설명하는 도면이다. 송신측에서의 본 발명은 6DoF 비디오 전송 장치와 관련될 수 있다. 상기 6DoF 비디오 전송 장치는 상술한 준비 과정 및 동작들을 수행할 수 있다. 구체적으로, 도 14에 도시된 것과 같이 6DoF 비디오/이미지 전송 장치는 데이터 입력부, 스티처(Stitcher), 프로젝션 처리부, 시점 분리 처리부, 시점별 패킹 처리부, 메타메이터 처리부, 피드

백 처리부, 데이터 인코더, 인캡슐레이션 처리부, 전송 처리부 및/또는 전송부를 내/외부 구성 요소로 포함할 수 있고, 또한, 비록 도 14에는 도시되지 않지만 깊이 정보 처리부를 내/외부 구성 요소로 포함할 수 있다.

[0221] 데이터 입력부는 한 군데 이상의 위치에서 한 개 이상의 카메라로 캡처된 각 시점별 이미지/비디오/깊이정보/오디오 데이터를 입력 받을 수 있다. 또한, 데이터 입력부는 캡처 과정에서 발생된 메타데이터를 비디오/이미지/깊이정보/오디오 데이터와 함께 입력 받을 수 있다. 데이터 입력부는 입력된 각 시점별 비디오/이미지 데이터를 스티처로 전달하고, 캡처 과정에서 발생된 메타데이터를 메타데이터 처리부로 전달할 수 있다.

[0222] 스티처는 캡처된 시점별/위치별 이미지/비디오들에 대한 스티칭(stitching)을 수행할 수 있다. 스티처는 스티칭된 6DoF 비디오 데이터를 프로젝션 처리부로 전달할 수 있다. 스티처는 필요한 경우 상기 6DoF 비디오 데이터에 대한 메타데이터를 메타데이터 처리부로부터 전달받아 상기 메타데이터를 기반으로 상기 6DoF 비디오 데이터를 스티칭을 할 수 있다. 스티처는 스티칭 과정에서 발생한 메타데이터를 메타데이터 처리부로 전달 할 수 있다. 또한, 스티처는 깊이(depth) 정보 처리부에서 전달 받은 위치값을 활용하여 비디오/이미지 스티칭 위치를 다르게 할 수 있다. 스티처는 스티칭 과정에서 발생된 메타데이터를 메타데이터 처리부로 전달할 수 있다. 전달되는 메타데이터는 스티칭 수행 여부, 스티칭 타입, 주요 시점(primary view)과 부차 시점(secondary view)의 ID 및 해당 시점의 위치 정보 등이 있을 수 있다.

[0223] 프로젝션 처리부는 스티칭된 6DoF 비디오 데이터를 2D 이미지 프레임에 프로젝션할 수 있다. 프로젝션 처리부는 스킴(scheme)에 따라 다른 형태의 결과물을 도출할 수 있는데, 해당 스킴은 기존의 360도 비디오의 프로젝션 스킴과 유사할 수도 있고, 6DoF를 위해 새롭게 제안된 스킴이 적용될 수도 있다. 또한 각 시점별 서로 다른 스킴이 적용될 수도 있다. 깊이 정보 처리부는 깊이 정보를 프로젝션 처리부로 전달하여 맵핑 결과 값을 다르게 할 수 있다. 프로젝션 처리부는 필요한 경우 메타데이터 처리부로부터 프로젝션에 필요한 메타데이터를 전달받아 6DoF 비디오 데이터의 프로젝션에 이용할 수 있으며, 프로젝션 처리부는 프로젝션 과정에서 발생된 메타데이터를 메타데이터 처리부로 전달할 수 있다. 해당 메타데이터는 스킴의 종류, 프로젝션 수행 여부, 주요시점과 부차 시점의 프로젝션 후의 2D 프레임의 ID 및 시점별 위치 정보 등이 있을 수 있다.

[0224] 시점별 패킹 처리부는 전술한 바와 같이 주요 시점과 부차 시점으로 나누고, 각 시점 내 리전별 패킹 과정을 수행할 수 있다. 즉 시점별 패킹 처리부는 각 시점/위치별 프로젝션된 6DoF 비디오 데이터를 주요 시점과 부차 시점으로 분류하고, 코딩 효율을 높이기 위하여 주요 시점과 부차 시점을 다른 해상도를 갖도록 하거나 각 시점의 비디오 데이터를 복수의 리전들로 나눠 각 리전의 데이터에 대한 회전, 재배열을 달리하여 각 시점 안에서 나누어진 리전별 해상도를 다르게 갖도록 할 수도 있다. 상술한 주요 시점과 부차 시점을 다른 해상도를 갖도록 하는 과정은 시점별 패킹 과정이라고 나타낼 수 있고, 각 시점의 비디오 데이터를 복수의 리전들로 나눠 각 리전의 데이터에 대한 회전, 재배열을 달리하여 각 시점 안에서 나누어진 리전별 해상도를 다르게 갖도록 하는 과정은 리전별 패킹 과정이라고 나타낼 수 있다. 또는 상기 시점별 패킹 및 리전별 패킹을 포함하는 패킹 과정을 시점별 패킹 과정이라고 나타낼 수도 있다.

[0225] 주요 시점과 부차 시점을 분류하는 과정은 생략될 수 있는 선택적인 과정일 수 있으며, 각 시점의 비디오 데이터 내 리전별 다른 해상도를 가지거나 배치를 다르게 하는 것도 선택적으로 수행될 수 있다. 시점별 패킹 처리부에 의한 패킹이 수행되는 경우, 상기 리전별 패킹은 메타데이터 처리부로부터 전달 받은 정보를 활용하여 수행될 수도 있으며, 상기 리전별 패킹 과정에서 발생한 메타데이터는 메타데이터 처리부로 전달될 수도 있다. 시점별 패킹 처리 과정에서 정의되는 메타데이터는 주요 시점과 부차 시점을 분류하기 위한 각 시점의 ID와 시점 내 리전별 적용되는 사이즈, 회전 각 리전별 위치 값 등에 대한 정보를 포함할 수 있다.

[0226] 전술한 스티처, 프로젝션 처리부 및/또는 시점별 패킹 처리부는 실시예에 따라 하나 이상의 하드웨어 컴포넌트 혹은 스트리밍/다운로드 서비스 내의 인제스트 서버(Ingest server)에서 수행될 수도 있다.

[0227] 메타데이터 처리부는 캡처 과정, 스티칭 과정, 프로젝션 과정, 시점별 패킹 과정, 인코딩 과정, 인캡슐레이션 과정 및/또는 전송을 위한 처리 과정에서 발생할 수 있는 메타데이터들을 처리할 수 있다. 메타데이터 처리부는 각 프로세스에서 전달 받은 메타데이터를 활용하여 6DoF 비디오 서비스를 위한 새로운 메타데이터를 생성할 수 있다. 실시예에 따라 메타데이터 처리부는 새롭게 생성된 메타데이터를 시그널링 테이블의 형태로 생성할 수도 있다. 메타데이터 처리부는 전달받거나 메타데이터 처리부에서 새롭게 생성/가공된 메타데이터를 다른 모듈(module)/유닛(unit)들에 전달할 수 있다. 예를 들어, 메타데이터 처리부는 생성되거나 전달 받은 메타데이터를 수신측으로 전송될 수 있도록 데이터 인코더, 인캡슐레이션 처리부 및/또는 전송 처리부에 전달할 수 있다.

[0228] 데이터 인코더는 2D 이미지 상에 프로젝션된 6DoF 비디오 데이터 및/또는 시점별/리전별 패킹된 비디오 데이터

를 인코딩할 수 있다. 인코딩은 다양한 포맷으로 수행될 수 있으며, 상기 비디오 데이터가 시점별 분류가 되어 있는 경우, 시점별 인코딩 결과 값을 분리하여 수신측에 전달할 수도 있다.

- [0229] 인캡슐레이션 처리부는 인코딩된 6DoF 비디오 데이터 및/또는 관련 메타데이터를 파일(file) 등의 형태로 인캡슐레이션할 수 있다. 관련 메타데이터는 전송한 메타데이터 처리부로부터 전달 받을 수 있다. 인캡슐레이션 처리부는 해당 데이터를 ISOBMFF, OMAF 등의 파일 포맷으로 인캡슐레이션 하거나 DASH 세그먼트 등의 형태로 처리할 수 있으며, 새로운 형태의 파일 포맷으로 처리할 수도 있다. 메타데이터는 파일 포맷 내 다양한 레벨에 존재하는 박스(box)에 포함되거나 별도의 트랙내의 데이터로 포함될 수 있고, 또는 메타데이터만 파일로 인캡슐레이션될 수 있다. 시점별 별도의 인캡슐레이션 처리가 가능할 수도 있고, 시점별 필요한 메타데이터와 해당 비디오 정보가 함께 인캡슐레이션될 수도 있다.
- [0230] 전송 처리부는 포맷에 따라 인캡슐레이션된 비디오 데이터에 전송을 위한 추가 처리를 수행할 수 있다. 해당 처리는 메타데이터 처리부에서 전달 받은 메타데이터를 활용하여 작동할 수 있다. 전송부는 전송 처리부로부터 전달 받은 데이터 및/또는 메타데이터를 방송망 및/또는 브로드밴드를 통해 전송할 수 있다. 전송부는 방송망 및/또는 브로드밴드를 통한 전송 시 필요한 구성 요소를 포함할 수 있다.
- [0231] 비록 도면에는 도시되지 않았으나 피드백 처리부(송신측)는 네트워크 인터페이스를 추가로 더 포함할 수 있다. 네트워크 인터페이스는 후술하는 6DoF 비디오 수신 장치로부터 피드백 정보를 전달 받고, 상기 전달받은 피드백 정보를 피드백 처리부(송신측)으로 전달할 수 있다. 피드백 처리부는 수신측에서 전달받은 정보를 스티칭, 프로젝션, 시점별 패키징, 인코더, 인캡슐레이션 처리부 및/또는 전송 처리부로 전달할 수 있고, 또한, 수신측에서 전달받은 정보를 메타데이터 처리부로 전달하여 메타데이터 처리부가 다른 요소들에 전달하거나 메타데이터 처리부에서 새로운 메타데이터를 생성/가공하여 전달하도록 할 수 있다. 또한, 본 발명의 또 다른 실시예에 따르면 피드백 처리부는 네트워크 인터페이스로부터 전달 받은 위치/시점 정보를 메타데이터 처리부로 전달할 수 있고, 메타데이터 처리부는 프로젝션, 시점별 패키징 처리부, 인캡슐레이션 처리부 및/또는 데이터 인코더로 해당 위치/시점 정보를 전달하여 현재 사용자의 시점/위치에 맞는 정보와 주변 정보만을 전송하도록 할 수 있는바, 이를 통하여 코딩 효율을 높일 수 있다.
- [0232] 전송한 6DoF비디오 전송 장치의 구성 요소들은 하드웨어로 구현되는 하드웨어 구성 요소 일 수 있다. 실시예에 따라 각 구성요소들은 변경, 생략 되거나 새로운 구성요소를 추가 혹은 다른 구성요소로 대체, 통합될 수 있다.
- [0233] 도 15는 본 발명이 적용될 수 있는 6DoF 비디오 수신 장치의 구성을 개략적으로 설명하는 도면이다. 수신측에서의 본 발명은 6DoF 비디오 수신 장치와 관련될 수 있다. 상기 6DoF 비디오 수신 장치는 수신부, 수신 처리부, 디캡슐레이션 처리부, 메타데이터 파서, 피드백 처리부, 데이터 디코더, 리-프로젝션 처리부, 가상시점 생성/합성부 및/또는 렌더러를 구성 요소로 포함할 수 있다.
- [0234] 수신부는 전송한 6DoF송신 장치로부터 6DoF 비디오 데이터를 수신할 수 있다. 6DoF 비디오 데이터가 전송되는 채널에 따라 수신부는 방송망 또는 브로드밴드를 통해 6DoF 비디오 데이터를 수신할 수도 있다.
- [0235] 수신 처리부는 수신된 6DoF 비디오 데이터에 대한 전송 프로토콜에 따른 처리를 상기 6DoF 비디오 데이터에 수행할 수 있다. 수신 처리부는 전송 처리부에서 수행된 과정의 역순으로 상기 6DoF 비디오 데이터에 대한 처리를 수행할 수 있고, 또는 프로토콜 처리 방법에 따른 과정을 거쳐 전송 처리부 이전 단계에서 얻은 데이터를 획득할 수 있다. 수신 처리부는 획득한 데이터를 디캡슐레이션 처리부로 전달하고, 수신부로부터 받은 메타데이터를 메타데이터 파서(metadata parser)로 전달할 수 있다.
- [0236] 디캡슐레이션 처리부는 수신 처리부로부터 전달받은 파일 형태의 6DoF 비디오 데이터를 디캡슐레이션할 수 있다. 디캡슐레이션 처리부는 해당 파일 포맷에 맞추어 파일들을 디캡슐레이션하여, 6DoF 비디오 및/또는 메타데이터를 획득할 수 있다. 획득된 6DoF 비디오 데이터는 데이터 디코더로 보낼 수 있고, 6DoF 메타데이터는 메타데이터 파서로 전달할 수 있다. 디캡슐레이션 처리부는 필요에 따라 메타데이터 파서로부터 디캡슐레이션에 필요한 메타데이터를 전달받을 수도 있다.
- [0237] 데이터 디코더는 6DoF 비디오 데이터에 대한 디코딩을 수행할 수 있다. 데이터 디코더는 메타데이터 파서로부터 디코딩에 필요한 메타데이터를 전달 받을 수 있다. 데이터 디코딩 과정에서 획득 된 메타데이터는 메타데이터 파서로 전달되어 처리될 수 있다.
- [0238] 메타데이터 파서는 6DoF 비디오 관련 메타데이터에 대한 파싱/디코딩을 수행할 수 있다. 메타데이터 파서는 획득한 메타데이터를 디캡슐레이션 처리부, 데이터 디코더, 리-프로젝션 처리부, 가상 시점 생성/합성부 및/또는

렌더러로 전달 할 수도 있다.

- [0239] 리-프로젝션 처리부는 디코딩된 6DoF 비디오 데이터에 대하여 리-프로젝션을 수행할 수 있다. 리-프로젝션 처리부는 각 시점/위치별 6DoF 비디오 데이터를 각각 3차원 공간으로 리-프로젝션할 수 있다. 3차원 공간은 사용되는 3차원 모델에 따라 다른 형태를 가질 수도 있고, 변환 과정을 거쳐 동일한 형태의 3차원 모델로 리-프로젝션될 수도 있다. 리-프로젝션 처리부는 메타데이터 파서로부터 필요한 메타데이터를 전달 받을 수 있다. 리-프로젝션 과정에서 정의된 메타데이터를 메타데이터 파서로 전달할 수도 있다. 예를 들어 각 시점/위치 별 6DoF 비디오 데이터의 3차원 모델을 메타데이터 파서로부터 전달 받을 수 있다 또한, 각 시점/위치별 비디오 데이터의 3차원 모델이 다른 경우에 모든 시점의 비디오 데이터를 동일한 3차원 모델로 리-프로젝션한 경우, 어떤 모델이 적용되었는지 메타데이터 파서로 전달할 수 있다. 때에 따라서는 리-프로젝션 처리부는 리-프로젝션에 필요한 메타데이터를 이용하여, 3차원 공간 내에 특정 영역만 리-프로젝션할 수 있으며, 한 개 이상의 특정 영역을 리-프로젝션할 수도 있다.
- [0240] 가상 시점 생성/합성부는 수신된 6DoF 비디오 데이터에 포함되어 있지 않으나 리-프로젝션 된 3차원 공간상에서 재생이 필요한 가상의 시점 영역에서의 비디오 데이터를 주어진 6DoF 비디오 데이터를 활용하여 생성하고, 가상 시점을 중심으로 새로운 시점/위치에서의 6DoF 비디오 데이터를 합성하는 과정을 수행할 수 있다. 비록 도면에는 도시되지 않았으나 새로운 시점의 비디오 데이터를 생성하는 경우, 가상 시점 생성/합성부는 깊이(depth) 정보 처리부의 데이터를 활용할 수 있다. 가상 시점 생성/합성부는 메타데이터 파서로부터 전달 받은 특정 영역과 수신되지 않은 주변 가상 시점 영역의 일부만 생성/합성할 수 있다. 한편, 가상 시점 생성/합성부에 의한 가상 시점에서의 6DoF 비디오 데이터 생성은 선택적으로 수행될 수 있으며, 필요한 시점 및 위치에 해당하는 비디오 정보가 없을 때 수행될 수 있다.
- [0241] 렌더러는 리-프로젝션 혹은 가상 시점 생성/합성부에서 전달된 6DoF 비디오 데이터를 렌더링 할 수 있다. 전술한 바와 같이 3차원 공간상에서 리-프로젝션 혹은 가상 시점 생성/합성부에서 일어나는 모든 과정은 렌더러와 통합되어 렌더러 내에서 이 과정들이 진행될 수 있다. 실시예에 따라 사용자의 시점/위치 정보에 따라 사용자가 보고 있는 부분 및 예상 경로 상의 일부만 렌더링 할 수도 있다.
- [0242] 본 발명에서 피드백 처리부(수신측) 및/또는 도면에는 도시되지 않았으나 네트워크 인터페이스를 추가적인 구성요소로 포함할 수 있다. 수신측 피드백 처리부는 렌더러, 가상 시점 생성/합성부, 리-프로젝션 처리부, 데이터 디코더, 디캡슐레이션 및/또는 VR 디스플레이로부터 피드백 정보를 획득하여 처리할 수 있다. 피드백 정보는 사용자의 뷰포트 정보, 헤드 및 포지션 오리엔테이션 정보, 게이즈(gaze) 정보, 제스처(gesture) 정보 등을 포함할 수 있다. 네트워크 인터페이스는 피드백 정보를 피드백 처리부로부터 전달 받고, 상기 피드백 정보를 전송 장치로 전송할 수 있다. 또는 상기 피드백 정보는 6DoF 비디오 수신 장치의 각 구성 요소에서 소비될 수도 있다. 예를 들면, 디캡슐레이션 처리부는 피드백 처리부로부터 사용자의 위치/시점 정보를 전달 받을 수 있고, 수신된 6DoF 비디오 중에 상기 사용자의 위치의 정보가 있는 경우, 해당 위치 정보만 디캡슐레이션, 디코딩, 리-프로젝션, 렌더링 할 수 있다. 만약 해당 위치의 정보가 없는 경우, 해당 위치 주변에 위치한 6DoF 비디오가 모두 디캡슐레이션, 디코딩, 리-프로젝션, 가상 시점 생성/합성, 렌더링의 과정을 거치도록 할 수 있다.
- [0243] 전술한 6DoF비디오 수신 장치의 구성 요소들은 하드웨어로 구현되는 하드웨어 구성 요소일 수도 있다. 또한, 실시예에 따라 각 구성 요소들은 변경, 생략 될 수 있고, 새로운 구성 요소가 추가되거나 혹은 다른 구성 요소로 대체, 통합될 수도 있다.
- [0244] 도 16은 6DoF 비디오 전송 장치/6DoF 비디오 수신 장치를 통하여 수행되는 6DoF 콘텐츠 제공을 위한 전체 아키텍처를 예시적으로 나타낸다.
- [0245] 도 16에 도시된 것과 같은 아키텍처에 의하여 6DoF 콘텐츠가 제공될 수 있다. 6DoF 콘텐츠는 파일 형태로 제공되거나 DASH 등과 같이 세그먼트(segment) 기반 다운로드 또는 스트리밍 서비스의 형태로 제공될 수 있으며, 또는 새로운 파일 포맷 혹은 스트리밍/다운로드 서비스 방법의 형태로 제공될 수도 있다. 여기서 6DoF 콘텐츠는 실감미디어(immersive media) 콘텐츠 혹은 라이트필드(light field) 콘텐츠, 혹은 포인트 클라우드(point cloud) 콘텐츠로 불릴 수 있다.
- [0246] 도 16에 도시된 6DoF 콘텐츠에 대한 파일 제공 및 스트리밍/다운로드 서비스를 위한 각 과정은 아래와 같이 상세하게 설명될 수 있다.
- [0247] 도 16을 참조하면 획득(Acquisition) 과정이 수행될 수 있다. 상기 획득 과정은 멀티뷰/스테레오/딥스 이미지(multi view/stereo/depth image)를 획득하기 위한 카메라(camera)로부터 캡처(capture) 후 얻어지는 output을

획득하는 과정을 나타낼 수 있다. 즉, 상기 획득 과정을 통하여 카메라(camera)로부터 캡처된 2개 이상의 비디오/이미지(video/image) 및 오디오 데이터가 얻어질 수 있고, 뎁스 카메라(depth camera)가 존재하는 경우, 각 장면(scene)에서의 뎁스 맵(depth map)도 획득(acquisition) 될 수 있다.

- [0248] 또한, 도 16을 참조하면 오디오 인코딩(Audio Encoding) 과정이 수행될 수 있다. 오디오 인코딩(Audio Encoding) 과정은 획득된 6DoF 오디오 데이터에 대하여 오디오 전처리 과정 및 인코딩 과정 수행하는 과정을 나타낼 수 있다. 상기 오디오 인코딩 과정에서 메타데이터가 생성될 수 있으며, 관련 메타데이터는 전송을 위해 인캡슐레이션/인코딩 과정을 거칠 수 있다.
- [0249] 또한, 도 16을 참조하면 스티칭(Stitching), 프로젝션(Projection), 맵핑(mapping) 및 코렉션(correction) 과정이 수행될 수 있다. 상기 스티칭, 상기 프로젝션, 상기 맵핑 및 상기 코렉션 과정은 전술한 바와 같을 수 있다. 획득된 6DoF 비디오 데이터는 전술한 바와 같이 다양한 위치에서 획득된 영상의 에디팅, 스티칭, 프로젝션 과정이 수행될 수 있다. 이 과정은 실시예에 따라 일부만 수행될 수도 있고, 전체가 생략되어 수신기측에서 수행될 수도 있다.
- [0250] 또한, 도 16을 참조하면 스티칭(Stitching), 프로젝션(Projection), 맵핑(mapping) 및 코렉션(correction) 과정이 수행될 수 있다.
- [0251] 또한, 도 16을 참조하면 시점 분리/패킹 과정이 수행될 수 있다. 구체적으로, 전술한 바와 같이 시점 분리/패킹 처리부는 스티칭된 영상을 바탕으로 수신기 측에서 요구 되는 주요 시점, Primary View(PV) 위치의 영상을 분리해 내어 패킹할 수 있다. 주요 시점에 대한 영상이 분리되어 패킹된 후, 나머지 영상을 부차 시점, 즉, Secondary View(SV)에 대한 영상으로 패킹하는 전처리 과정이 수행될 수 있다. 패킹하는 과정에서 코딩 효율을 높이기 위하여 상기 주요 시점에 대한 영상과 상기 부차 시점에 대한 영상의 사이즈, 해상도 등이 조정될 수 있다. 또한, 동일한 성격의 시점 내, 즉, 주요 시점에 대한 영상 또는 부차 시점에 대한 영상 내에서도 리전별 다른 조건으로 해상도를 가지거나 리전에 따라 회전, 재배치될 수 있다.
- [0252] 또한, 도 16을 참조하면 뎁스 센싱 및/또는 추정(Depth sensing and/or estimation) 과정이 수행될 수 있다. 예를 들어, 깊이 캡처 카메라(depth camera)가 존재하지 않는 경우, 획득된 2개 이상의 영상에서 깊이 맵을 추출해 내는 과정이 수행될 수 있다. 상기 과정은 뎁스 추정 과정이라고 나타낼 수 있다. 또한, 깊이 캡처 카메라(depth camera)가 존재하는 경우, 영상 획득 위치에서 영상 내 포함된 각 오브젝트(object)의 깊이(depth)가 얼마나 되는지 위치 정보를 저장하기 위한 과정이 수행될 수 있다. 상기 과정은 뎁스 센싱 과정이라고 나타낼 수 있다. 즉, 상기 영상 내 각 오브젝트의 위치에 따른 깊이값이 저장될 수 있다.
- [0253] 또한, 도 16을 참조하면 포인트 클라우드 퓨전/익스트랙션(Point Cloud Fusion/extraction) 과정이 수행될 수 있다. 상기 포인트 클라우드 퓨전/익스트랙션 과정은 미리 획득된 깊이 맵을 인코딩 가능한 형태의 데이터로 변형하는 과정을 나타낼 수 있다. 예를 들어 상기 깊이 맵을 포인트 클라우드 데이터 타입으로 변형하여 3차원에서 영상의 각 오브젝트의 위치 값을 할당하는 전처리 과정이 수행될 수 있으며, 상기 포인트 클라우드 데이터 타입이 아닌 3차원 공간 정보를 표현할 수 있는 데이터 타입이 대신 적용될 수도 있다.
- [0254] 또한, 도 16을 참조하면 주요 시점(Primary View, PV) 인코딩/부차 시점(Secondary View, SV) 인코딩/라이트 필드/포인트 클라우드 인코딩(PV encoding/SV encoding/light field/point cloud encoding) 과정이 수행될 수 있다. 시점별로 미리 패킹된 경우, 각 시점별 이미지 또는 비디오 인코딩이 수행될 수 있다. 또한, 깊이 정보 및/또는 위치 정보는 각 시점별로 이미지 인코딩 내지 비디오 인코딩될 수 있다. 또한, 동일한 시점의 같은 콘텐츠라도 리전별로 다른 비트 스트림으로 인코딩될 수도 있다. MPEG-I에서 정의될 새로운 코덱(codec) 및 HEVC-3D, OMAF++ 등 다양한 미디어 포맷(media format)이 적용될 수 있다.
- [0255] 또한, 도 16을 참조하면 파일 인캡슐레이션(File-encapsulation) 과정이 수행될 수 있다. 전술한 대로 인코딩된 6DoF 비디오 데이터는 인캡슐레이션 처리부에 ISOBMFF 등의 파일 포맷으로 처리될 수 있다. 즉, 상기 파일 인캡슐레이션(File-encapsulation) 과정은 인코딩된 6DoF 비디오 데이터를 ISOBMFF 등의 파일 포맷으로 변환하는 과정을 나타낼 수 있다. 상기 인코딩된 6DoF 비디오 데이터는 세그먼트들(segements)로 처리될 수 있다.
- [0256] 또한, 도 16을 참조하면 캡처된 6DoF 비디오 데이터를 저장 또는 전송하기 위한 파일로 처리하는 과정에서 6DoF 비디오 데이터에 대한 메타데이터가 생성될 수 있다. 상기 메타데이터는 뎁스 정보를 포함할 수 있다.
- [0257] 구체적으로, 상술한 6DoF 비디오 데이터 처리, 예를 들어, 획득, 스티칭, 프로젝션, 시점별 분리/패킹, 인코딩, 인캡슐레이션 과정 중에 발생한 메타데이터는 메타데이터 처리부로 전달될 수 있다. 또한, 메타데이터 처리부에서 생성된 메타데이터가 각 과정으로 전달될 수도 있다. 또한 송신측에서 생성된 메타데이터는 인캡슐레이션 과

정을 통하여 하나의 트랙 혹은 파일로 생성되어 수신측으로 전달될 수 있다. 수신측에서는 방송망이나 브로드밴드를 통하여 별도의 파일 혹은 파일 내 트랙으로 저장되어 있는 메타데이터를 수신할 수 있다.

- [0258] 또한, 도 16을 참조하면 6DoF 비디오 데이터에 대한 파일 및/또는 세그먼트들이 전달되는 전달(Delivery) 과정이 수행될 수 있다. 파일 및/또는 세그먼트들은 DASH 혹은 유사한 기능을 가진 새로운 모델을 기반으로 전송을 위한 개별 트랙에 포함될 수 있다. 이때 전송을 위해 MPEG DASH, MMT 및/또는 새로운 표준이 적용될 수 있다.
- [0259] 또한, 도 16을 참조하면 6DoF 비디오 수신 장치는 6DoF 비디오/오디오 데이터 수신을 위하여 수신된 파일을 디캡슐레이션하는 파일 디캡슐레이션(File decapsulation) 처리를 수행할 수 있다.
- [0260] 또한, 6DoF 비디오 수신 장치는 6DoF 오디오 데이터에 대한 디코딩 및 렌더링/6DoF 비디오 데이터에 대한 디코딩 및 렌더링을 수행할 수 있다. 디코딩 및 렌더링된 6DoF 오디오 데이터는 스피커(Loudspeakers), 헤드폰(headphones)을 통해 사용자에게 제공될 수 있다.
- [0261] 또한, 도 16을 참조하면 주요 시점(Primary View, PV) 디코딩/부차 시점(Secondary View, SV) 디코딩/라이트 필드/포인트 클라우드 디코딩(PV encoding/SV encoding/light field/point cloud encoding) 과정이 수행될 수 있다. 디코딩에 적용되는 코덱은 HEVC-3D, OMAF++ 및 MPEG에서 6DoF를 위해 새롭게 제안되는 코덱이 적용될 수 있다. 한편, 6DoF 비디오 데이터에 대한 디코딩이 수행되는 경우, 주요 시점(PV)과 부차 시점(SV)이 분리되어 패키징된 경우에는 각 시점에 대한 비디오 혹은 이미지가 각각 디코딩될 수 있고, 또는 시점 분류와 상관없이 비디오 혹은 이미지에 대한 디코딩이 수행될 수 있다. 또한 위치, 깊이 정보를 가지고 있는 라이트필드와 포인트 클라우드에 대한 디코딩이 먼저 수행되고 나서 헤드, 포지션, 시선 트래킹의 피드백이 먼저 전달되고, 상기 피드백으로 도출되는 사용자가 위치한 주변부 시점의 이미지 내지 비디오만 분리해 내어 디코딩될 수도 있다. 여기서, 전술한 바와 같이 헤드, 포지션, 시선 트래킹(head/position/eye tracking)을 통하여 사용자의 헤드, 포지션, 게이즈, 뷰포트 정보 등이 획득, 처리될 수 있다.
- [0262] 또한, 6DoF 비디오 수신 장치는 캡처한 비디오/이미지 데이터를 3차원 공간상에 리-프로젝션할 때에 3차원 공간의 위치를 설정할 수 있고, 수신된 비디오/이미지 데이터에서 확보되지 못하였으나 사용자가 이동 가능한 위치인 가상 시점의 3차원 공간을 생성하는 포인트 클라우드 렌더링(Point Cloud rendering) 과정을 수행할 수 있다.
- [0263] 또한, 6DoF 비디오 수신 장치는 전술한 바와 같이 사용자가 위치한 공간에 6DoF 비디오 데이터가 없는 경우, 사용자 위치/시점 주변에 이미 확보된 6DoF 비디오 데이터를 활용하여 새로운 시점의 비디오 데이터를 생성하고 합성하는 가상 시점 합성(Virtual view synthesis) 과정을 수행할 수 있다. 한편, 실시예에 따라 가상 시점 생성 및/또는 합성 과정은 생략될 수 있다.
- [0264] 또한, 6DoF 비디오 수신 장치는 이미지 합성 및 렌더링(Image composition, and rendering) 과정을 수행할 수 있다. 상기 과정은 전술한 바와 같이 사용자의 위치를 중심으로 한 영상을 렌더링 하기 위한 과정으로 6DoF 비디오 수신 장치는 사용자의 위치 및 시선에 따라 디코딩된 6DoF 비디오 데이터를 이용하거나 가상 시점 생성/합성으로 만들어진 사용자 주변의 비디오 및 이미지를 렌더링할 수 있다.
- [0265] 한편, 상기 6DoF 콘텐츠가 제공되는 6DoF 공간은 후술하는 내용과 같이 설명될 수 있다.
- [0266] 도 17은 본 발명의 6DoF 공간을 설명하기 위하여 도시한 도면이다.
- [0267] 본 발명에서 프로젝션 전 또는 리-프로젝션 후의 6DoF 공간에 대해 기술하고 있는바, 그에 대한 정보의 시그널링을 수행하기 위하여 도 17에 도시된 것과 같은 도면의 개념이 사용될 수 있다.
- [0268] 구체적으로, 6DoF 공간은 360도 비디오 혹은 3DoF 공간이 야(Yaw), 피치(Pitch), 롤(Roll)로 설명될 수 있는 것과 달리 이동의 방향을 레이셔널(rational)과 트랜스레이션(translation) 두 종류로 나눠 설명될 수 있다. 구체적으로, 상기 레이셔널 이동은 도 17의 (a)에 도시된 것과 같이 기존의 3DoF 공간의 방향을 설명한 것과 마찬가지로 야, 피치, 롤로 설명될 수 있으며, 상기 레이셔널 이동은 방향의 이동(orientation movement)으로 불릴 수도 있다. 반면, 상기 트랜스레이션 이동은 도 17의 (b)에 도시된 것과 같이 (사용자 또는 시점의) 포지션의 이동으로 불릴 수 있다. 왼쪽/오른쪽(Left/Right), 앞/뒤(Forward/Backward), 위/아래(Up/down) 방향 중 축이 어디로 이동했는지 알려 줄 수 있는 것으로 한 축 이상의 값을 정의하여 중심축(즉, 사용자 또는 시점)의 이동을 설명할 수 있다. 예를 들어, 상기 중심축의 위치 및 이동을 나타내기 위하여 x축, y축 및 z축이 정의될 수 있고, 중심축의 위치를 나타내는 x성분, y성분, z성분으로 왼쪽/오른쪽(Left/Right), 앞/뒤(Forward/Backward), 위/아래(Up/down) 방향으로의 이동이 설명될 수 있다.

- [0269] 한편, 상술한 6DoF 비디오 처리 과정에서 생성되는 메타데이터는 아래와 같이 정리되어 전달될 수 있다.
- [0270] 예를 들어, 후술하는 6DoF 비디오 관련 메타데이터는 6DoF 비디오에 대한 다양한 메타데이터를 포함할 수 있다. 6DoF 비디오 관련 메타데이터는 별도의 시그널링 테이블에 포함되어 전송될 수도 있고, HEVC SEI message에 포함되어 전달 될 수도 있고, DASH MPD 내에 포함되어 전송될 수도 있고, ISOBMFF 등의 파일 포맷에 박스(box) 형태로 포함되어 전달될 수도 있다. 6DoF 비디오 관련 메타데이터가 박스 형태로 포함되는 경우, 상기 박스는 파일, 프래그먼트, 트랙, 샘플 엔트리, 샘플 등등 다양한 레벨에 포함되어 해당되는 레벨의 데이터에 대한 메타데이터를 포함할 수 있다. 또한, 6DoF 비디오 파일 포맷, 미디어 포맷, 압축 포맷 등에 대한 새로운 포맷이 제안되는 경우, 상기 box 가 포함하는 메타데이터는 각 포맷에서 포함되는 메타데이터에 해당될 수도 있다.
- [0271] 본 발명에 따른 6DoF 비디오 관련 메타데이터의 실시예에 의하면, 6DoF 비디오 관련 메타데이터는 캡처된 비디오를 획득할 때에 전달되어야 하는 캡처 정보에 관한 메타데이터, 프로젝션 스킵이 포함된 스티칭 및/또는 렌더링에 관련된 메타데이터, 사용자 위치 관련 메타데이터, 뷰포트 관련 메타데이터를 포함할 수 있다. 또한, 실시예에 따라 6DoF 비디오 관련 메타데이터는 전술한 것 외에 추가적인 메타데이터를 더 포함할 수 있다.
- [0272] 본 발명의 하나의 실시예로 ISO Base Media File Format (ISO BMFF) 기반으로 6DoF 비디오가 저장되는 경우, 본 발명은 6DoF 비디오 트랙 또는 샘플 등과 관련된 6DoF 비디오 관련 메타데이터를 저장 및 시그널링 하는 방안을 제안한다. 하나의 파일 내 6DoF 비디오에 대한 비디오 트랙 또는 샘플 등과 관련된 메타데이터는 하기와 같은 박스(box) 형태로 저장될 수 있다.

**표 1**

```

aligned(8) class SixDOFVideoConfigurationBox extends FullBox('sdvc', version=0, 0) {
unsigned      int(8)      projection_scheme;
unsigned      int(1)      stitching_flag;
signed        int(8)      initial_view_yaw_degree;
signed        int(8)      initial_view_pitch_degree;
signed        int(8)      initial_view_roll_degree;
signed       int(8)       initial_center_FB_delta;
signed       int(8)       initial_center_UD_delta;
signed       int(8)       initial_center_RL_delta;
unsigned      int(6)      reserved = 0;
unsigned      int(1)      content_fov_flag;
unsigned     int(1)      capture_info_flag;
unsigned     int(1)      region_info_flag;

    if (stitching_flag == 1) {
        unsigned int(8)      num_view_center;
        for (i = 1; i<= num_center;i++){
            signed int(16) position_center_X;
            signed int(16) position_center_Y;
            signed int(16) position_center_Z;
            unsigned int(8)      num_camera;
            unsigned int(8)      stitching_radius;
        }
    }
}
    
```

[0273]



```

if((projection_scheme == '0') ) {
    unsigned int(16)    sphere_radius;
} else if(projection_scheme == '2') { //cylindrical projection
    unsigned int(16)    cylinder_radius;
    unsigned int(16)    cylinder_height;
} else if(projection_scheme == '3'){ //cubic projection
    unsigned int(16)    cube_front_width;
    unsigned int(16)    cube_front_height;
    unsigned int(16)    cube_height;
} else if(projection_scheme == '4'){ //panorama projection
    unsigned int(16)    panorama_height;
}
if (content_fov_flag == 1) {
    unsigned int(16)    content_fov;
}

if (camera_info_flag == 1) {
    string                Capture_type;
    unsigned int(8)    num_view;
    unsigned int(8)    Camera_capture_array_row;
    unsigned int(8)    Camera_capture_array_col;
    unsigned int(8)    Baseline_length;
    unsigned int(3)    Master_camera_ID; // to be origin_view
    for (int i;=0; i< num_view; i++) {
        IntrinsicCameraParametersBox    intrinsic_camera_params;

        ExtrinsicCameraParametersBox    extrinsic_camera_params;
    signed int(8)    camera_center_pitch;
        signed int(8)    camera_center_yaw;
        signed int(8)    camera_center_roll;
        signed int(8)    camera_center_FB; //Forward to Backward
        signed int(8)    camera_center_UD; //Up to Down
        signed int(8)    camera_center_RL; // Right to Left
    }
}
}

```

[0274]

[0275]

표 1과 같이 상기 6DoF 비디오 관련 메타데이터는 SixDOFVideoConfigurationBox 포함되어 저장 및 시그널링될 수 있다. 상기 6DoF 비디오 관련 메타데이터는 캡처 정보 메타데이터, 스티칭 정보 메타데이터, 프로젝션 스킴 메타데이터, 초기 위치 정보 및 FOV(Field of View) 메타데이터를 포함할 수 있다. 상기 6DoF 비디오 관련 메타데이터에 포함된 필드들에 대한 구체적인 설명은 후술하는 바와 같다.

[0276]

또한, 하나의 파일 내 6DoF 비디오에 대한 비디오 트랙 또는 샘플 등과 관련된 메타데이터는 하기와 같은 박스(box) 형태로 저장될 수도 있다.

표 2

```

aligned(8) class SixDOFVideoConfigurationBox extends FullBox('sdvc', version=0, 0) {
    unsigned int(8) projection_scheme;
    unsigned int(1) stitching_flag;
    signed int(8) initial_view_yaw_degree;
    signed int(8) initial_view_pitch_degree;
    signed int(8) initial_view_roll_degree;
    signed int(8) initial_center_FB_delta;
    signed int(8) initial_center_UD_delta;
    signed int(8) initial_center_RL_delta;
    unsigned int(3) reserved = 0;
    unsigned int(1) content_fov_flag;
    unsigned int(1) capture_info_flag;
    unsigned int(1) region_info_flag;
    unsigned int(1) reserved_rendering_range_flag;
    if (stitching_flag == 1) {
        unsigned int(8) num_view_center;
        for (i = 1; i <= num_center; i++){
            signed int(16) position_center_X;
            signed int(16) position_center_Y;
            signed int(16) position_center_Z;
            unsigned int(8) num_camera;
            unsigned int(8) stitching_radius;
        }
    }
}

```

[0277]

```

if (projection_scheme == '0') {
    unsigned int(16) sphere_radius;
} else if (projection_scheme == '2') { //cylindrical projection
    unsigned int(16) cylinder_radius;
    unsigned int(16) cylinder_height;
} else if (projection_scheme == '3') { //cubic projection
    unsigned int(16) cube_front_width;
    unsigned int(16) cube_front_height;
    unsigned int(16) cube_height;
} else if (projection_scheme == '4') { //panorama projection
    unsigned int(16) panorama_height;
}
if (content_fov_flag == 1) {
    unsigned int(16) content_fov;
}
if (camera_info_flag == 1) {
    string Capture_type;
    unsigned int(8) num_view;
    unsigned int(8) Camera_capture_array_row;
    unsigned int(8) Camera_capture_array_col;
    unsigned int(8) Baseline_length;
    unsigned int(3) Master_camera_ID; // to be origin_view
    for (int i;=0; i < num_view; i++) {
        IntrinsicCameraParametersBox
        intrinsic_camera_params;
        ExtrinsicCameraParametersBox extrinsic_camera_params;
        signed int(8) camera_center_pitch;
        signed int(8) camera_center_yaw;
        signed int(8) camera_center_roll;
        signed int(8) camera_center_FB; //Forward to Backward
        signed int(8) camera_center_UD; //Up to Down
        signed int(8) camera_center_RL; // Right to Left
    }
}
if (reserved_rendering_range_flag == 1) {
    unsigned int(8) rendering_type;
    unsigned int(8) reserved_range_type;
    signed int(16) center_X; //user center
    signed int(16) center_Y;
    signed int(16) center_Z;
    unsigned int(8) step size;
    if (reserved_range_type == '0') {
        unsigned int(8) sphere_radius_to_boundary;
    }
}

```

[0278]

```

else if(reserved_range_type =='1'){
    unsigned int(8)    horizontal_range;
    unsigned int(8)    vertical_range;
}
else if(reserved_range_type =='2'){
    unsigned int(8)    distance_to_top_endpoint;
    unsigned int(8)    distance_to_bottom_endpoint;
else if(reserved_range_type =='3'){
    unsigned int(8)    distance_to_left_endpoint;
    unsigned int(8)    distance_to_right_endpoint;
}
else if(reserved_range_type =='4'){
    unsigned int(8)    distance_to_front_endpoint;
    unsigned int(8)    distance_to_back_endpoint;
}
else if(reserved_range_type =='5'){
    signed int(16)    vertex_X[];
    signed int(16)    vertex_Y[];
    signed int(16)    vertex_Z[];
    unsigned int(1)    interpolate;
    unsigned int(7)    reserved;

If(rendering_type=='0'){
    unsigned int(16) sphere_radius;
}
If(rendering_type=='1'){
    signed int(16)    preliminary_viewport_yaw;
    signed int(16)    preliminary_viewport_pitch;
    signed int(16)    preliminary_viewport_roll;
    unsigned int(8)    preliminary_hor_range;
    unsigned int(8)    preliminary_ver_range;
}
}

```

[0279]

[0280]

표 2와 같이 상기 6DoF 비디오 관련 메타데이터는 SixDOFVideoConfigurationBox 포함되어 저장 및 시그널링될 수 있다. 상기 6DoF 비디오 관련 메타데이터는 상술한 캡처 정보 메타데이터, 스티칭 정보 메타데이터, 프로젝션 스킴 메타데이터, 초기 위치 정보 및 FOV 메타데이터와 더불어 리-프로젝션 정보 메타데이터를 더 포함할 수 있다. 상기 6DoF 비디오 관련 메타데이터에 포함된 필드들에 대한 구체적인 설명은 후술하는 바와 같다.

[0281]

구체적으로, 상기 캡처 정보 메타데이터는 해당 6DoF 카메라를 통하여 캡처된 비디오 혹은 이미지의 환경에 대한 정보를 담을 수 있다. 6DoF 비디오를 캡처할 수 있는 디바이스는 전술한 바와 같이 다양한 형태가 있을 수 있고, 캡처된 환경을 깊이 정보나 위치 정보로 변형하여 활용하기 위해서는 캡처 환경에서 카메라 간의 거리나 내부 캡처 설정 환경에 대한 정보가 필요할 수 있다. 구체적으로, 필요한 정보는 기본적으로 캡처한 카메라의 배열의 형태, 카메라 타입, 카메라 간의 거리, 카메라 내부 및/또는 외부 파라미터, 6DoF 비디오를 캡처하기 위하여 사용된 카메라 혹은 시점의 전체 개수 및/또는 각 카메라 혹은 시점의 ID가 있을 수 있고, 상기 캡처 정보 메타데이터는 상기 정보들을 나타내는 필드들을 포함할 수 있다. 또한, 실시예에 따라 캡처 정보 메타데이터는 추가적인 정보들을 더 포함할 수 있고, 또는 상기 정보 중 일부가 포함되지 않을 수 있다.

[0282]

한편, 상술한 표 1에 도시된 것과 같이 상기 캡처 정보 메타데이터는 camera\_info\_flag 필드의 값이 참(true)인 경우에 전송될 수 있다. 즉, 상기 camera\_info\_flag 필드는 상기 캡처 정보 메타데이터의 전송 여부를 나타낼 수 있다. 상기 camera\_info\_flag 필드의 값이 1인 경우, 상기 캡처 정보 메타데이터가 전송될 수 있고, 상기 camera\_info\_flag 필드의 값이 0인 경우, 상기 캡처 정보 메타데이터는 전송되지 않을 수 있다.

[0283]

상기 캡처 정보 메타데이터는 Master\_camera\_ID 필드, Capture\_type 필드, num\_of\_view 필드, Camera\_capture\_array\_row 필드, Camera\_capture\_array\_col 필드 및 Baseline\_length 필드를 포함할 수 있다. 상기 Capture\_type 필드는 camera\_type 필드라고 나타낼 수도 있고, 상기 Camera\_capture\_array\_col 필드는 Camera\_capture\_array\_column 필드라고 나타낼 수도 있다.

[0284]

상기 Master\_camera\_ID 필드는 3D 영상에서 원점이 되는 기준 시점의 카메라 ID를 나타낼 수 있다. 즉, 6DoF 비디오 전송 장치는 Master\_camera\_ID 필드의 값을 지정함으로써 캡처된 모든 영상의 3차원 영상에서의 포지션을 지정되는 경우에 원점이 되는 기준 정보를 전달할 수 있다.

[0285]

상기 Capture\_type 필드는 카메라의 배열 타입 또는 카메라들 또는 렌즈들의 세팅 타입을 나타낼 수 있다. 즉, 상기 Capture\_type 필드는 전술한 바와 같이 캡처시 카메라의 배열 상태 혹은 렌즈의 종류 등과 같은 카메라들 또는 렌즈들의 세팅 타입을 나타낼 수 있다. 예를 들어 카메라 세팅 타입은 사용한 카메라의 수에 따라

HDCA(High Density Camera Array) 와 LDCA(Low Density Camera Array)로 분류될 수 있으며, 해당 카메라 세팅 타입의 분류 기준은 MPEG-I 혹은 JPEG 등 표준 기관에서 분류한 기준을 따른다.

- [0286] 상기 num\_of\_view 필드는 카메라의 수를 나타낼 수 있다. 즉, 상기 6DoF 비디오를 캡처한 카메라의 수는 num\_of\_view 필드의 값과 일치할 수 있다. 또한, 상기 Camera\_capture\_array\_row 필드 및 상기 Camera\_capture\_array\_column 필드는 6DoF 비디오 캡처시 사용된 카메라 배열의 열과 행을 나타낼 수 있다.
- [0287] 도 18은 6DoF 비디오 데이터를 획득하기 위하여 사용되는 카메라 배열을 예시적으로 나타낸다. 도 18의 (a)를 참조하면 복수의 행과 열로 배치되는 카메라들을 사용하여 상기 6DoF 비디오 데이터가 획득될 수 있다. 도 18의 (a)에 도시된 것과 같은 카메라 배열의 경우, 좌상단에 위치한 카메라를 기점으로 하는 행과 열의 순서를 나타내어 카메라 배열을 나타낼 수 있다. 예를 들어, 도 18의 (a)에 도시된 것과 같이 배열된 카메라들을 기반으로 6DoF 비디오가 캡처된 경우, 배열된 카메라들의 수는 20개이므로 상기 num\_of\_view 필드의 값은 20, 상기 Camera\_capture\_array\_row 필드의 값은 4, 상기 Camera\_capture\_array\_column 필드의 값은 5로 도출될 수 있고, 상기 필드들을 기반으로 카메라 배열로 나타낼 수 있다. 따라서, 상기 카메라 배열이 적용되는 경우의 총 캡처한 시점의 수도 20으로 도출될 수 있다.
- [0288] 또한, 도 18의 (b)를 참조하면 원형 혹은 아크 형태로 배치되는 카메라들을 사용하여 상기 6DoF 비디오 데이터가 획득될 수 있다. 이 경우, 시계의 9시 방향과 가장 가까운 위치의 카메라를 기준으로 시계방향으로 배열되는 카메라 배열이 정의될 수 있다. 예를 들어, 도 18의 (b)에 도시된 것과 같이 배열된 카메라들을 기반으로 6DoF 비디오가 캡처된 경우, 상기 Camera\_capture\_array\_row 필드와 상기 Camera\_capture\_array\_column 필드의 값은 각각 1과 10으로 도출될 수 있으며, 상기 num\_of\_view 필드의 값은 10으로 도출될 수 있고, 총 캡처한 시점의 수가 10개임을 알 수 있다.
- [0289] 또한, 도 18의 (c)를 참조하면 360 카메라의 배열로 배치되는 카메라들을 기반으로 상기 6DoF 비디오 데이터가 획득될 수 있다. 이 경우, 캡처될 수 있는 시점의 수는 number\_of\_view 필드 x Camera\_capture\_array\_row 필드 x Camera\_capture\_array\_column 필드로 도출될 수 있다. 즉, 상기 캡처될 수 있는 시점의 수는 상기 number\_of\_view 필드, Camera\_capture\_array\_row 필드, Camera\_capture\_array\_column 필드의 값들을 곱한 값으로 도출될 수 있다. 예를 들어, 도 18의 (c)에 도시된 것과 같이 6DoF 비디오가 캡처된 경우, 상기 number\_of\_view 필드의 값은 6, 상기 Camera\_capture\_array\_row 필드의 값 및 상기 Camera\_capture\_array\_column 필드의 값은 각각 1과 3으로 도출될 수 있다. 따라서, 총 18개의 시점의 캡처 결과를 얻을 수 있다는 것을 알 수 있다.
- [0290] 한편, 상기 capture\_type 필드는 6DoF 비디오 데이터에 대한 카메라 세팅 타입을 나타낼 수 있다. 예를 들어, 상기 capture\_type 필드는 카메라 세팅 타입을 다음의 표와 같이 나타낼 수 있다.

**표 3**

Camera_type	Description
0000	HDCA (High Density Camera Array)
0001	LDCA (Low Density Camera Array)
0010	MicroLens
0011	Omnidirectional Camera
0100	Movement camera
0101	User defined
0110~1111	Reserved

- [0291]
- [0292] 상기 표 3을 참조하면 상기 capture\_type 필드의 값을 기반으로 6DoF 비디오 데이터에 대한 카메라 세팅 타입이 도출될 수 있다. 상기 카메라 세팅 타입은 HDCA(High Density Camera Array), LDCA(Low Density Camera Array), 마이크로렌즈(MicroLens), 전방향 카메라(Omnidirectional Camera) 및 User defined 중 하나로 도출될 수 있다.

- [0293] 예를 들어, 카메라 세팅 타입으로 마이크로렌즈(Microlens)가 있을 수 있는바, 상기 마이크로렌즈는 하나의 카메라 내부에 다양한 시점을 캡처할 수 있는 렌즈가 존재하여 다수의 카메라로 캡처한 결과물을 얻을 수 있는 카메라 세팅 타입일 수 있다. 상기 마이크로렌즈는 6DoF 비디오 캡처에 활용될 수 있고, 하나의 카메라 세팅 타입으로 분류될 수 있다. 카메라 세팅 타입이 상기 마이크로렌즈로 설정되는 경우, 상기 num\_of\_view 필드는 마이크로렌즈 카메라가 가지고 있는 렌즈의 수로 도출될 수 있고, 상기 Camera\_capture\_array\_row 필드 및 상기 Camera\_capture\_array\_col 필드는 렌즈의 배열 값으로 나타낼 수 있고, Baseline\_length 필드는 렌즈 중심 간의 거리를 나타낼 수 있다.
- [0294] 또한, 다른 카메라 세팅 타입으로 전방향 카메라(Omnidirectional Camera)가 있을 수 있다. 상기 전방향 카메라 타입은 기존의 360도 비디오 촬영에 사용되는 전방향 카메라를 이용하여 6DoF 비디오를 촬영하는 세팅 환경을 나타낼 수 있다. 카메라 세팅 타입이 전방향 카메라로 설정되는 경우, 상기 num\_of\_view 필드는 각 360 카메라가 한 위치에서 캡처할 수 있는 시점의 수가 될 수 있다. 또한, 다른 카메라 세팅 타입으로 무브먼트 카메라(Movement camera)가 있을 수 있다. 상기 무브먼트 카메라 타입은 이동 레일 등을 이용하여 하나의 카메라를 이동시켜가며 캡처하는 카메라 세팅 타입을 나타낼 수 있다. 카메라 세팅 타입이 무브먼트 카메라로 설정되는 경우, 상기 num\_of\_view 필드는 캡처한 위치의 수를 나타낼 수 있고, 상기 Camera\_capture\_array\_row 필드와 상기 Camera\_capture\_array\_column 필드는 카메라가 캡처한 위치의 배열을 나타낼 수 있다. 또한, 상기 Baseline\_length 필드는 카메라가 캡처한 위치로 이동한 거리를 나타낼 수 있다. 카메라 세팅 타입은 상술한 타입들이외에도 사용자가 임의로 설정할 수 있는 User defined 타입이 있을 수 있고, 새로운 방식으로 캡처할 수 있는 6DoF 카메라가 있는 경우는 나머지 값으로 설정할 수 있도록 필드를 남겨둘 수 있다.
- [0295] 한편, 캡처 정보 메타데이터는 Baseline\_length 필드를 포함할 수 있다. 상기 Baseline\_length 필드는 카메라 간의 거리를 나타내는 Baseline\_length 값을 나타낼 수 있다. 상기 Baseline\_length 필드는 카메라 간의 거리인 Baseline\_length를 깊이 정보 처리부로 전달하기 위하여 설정될 수 있다.
- [0296] 또한, 각 카메라 혹은 각 렌즈마다 카메라의 초점 거리나 얼라인먼트 값 등 각 카메라의 내부 세팅 정보를 얻을 수 있는 IntrinsicCameraParametersBox 와 각 카메라의 외부 세팅 정보, 즉 각 카메라가 틀어진 상태에서 촬영이 되었는지 수평은 맞는지 등에 대한 정보를 나타내는 ExtrinsicCameraParametersBox 가 캡처 정보 메타데이터에 포함될 수 있다. IntrinsicCameraParameter 및 ExtrinsicCameraParameter는 복수의 필드들을 포함할 수 있는바, 구체적인 설명은 후술하는 내용과 같다.
- [0297] 한편, 캡처 정보 메타데이터는 Master\_camera\_ID 필드를 포함할 수 있고, 각 카메라에 대한 camera\_center\_FB 필드, camera\_center\_UB 필드, camera\_center\_RL 필드, camera\_center\_yaw 필드, camera\_center\_pitch 필드, camera\_center\_roll 필드를 포함할 수 있다. 상기 Master\_camera\_ID 필드가 가리키는 카메라의 중심, 즉, 상기 Master\_camera\_ID 필드가 가리키는 시점의 센터 포지션을 원점으로 두고 캡처된 6DoF 비디오에 대한 나머지 카메라의 위치는 camera\_center\_FB 필드, camera\_center\_UB 필드, camera\_center\_RL 필드로 설정되고, 상기 카메라에 의하여 캡처된 비디오의 방향은 camera\_center\_yaw, camera\_center\_pitch, camera\_center\_roll로 정의될 수 있다. 구체적으로, 상기 camera\_center\_FB 필드는 해당 카메라 위치의 앞뒤 방향 축에 대한 성분, 상기 camera\_center\_UB 필드는 해당 카메라 위치의 위아래 방향 축에 대한 성분, 상기 camera\_center\_RL 필드는 해당 카메라 위치의 좌우 방향 축에 대한 성분을 나타낼 수 있다. 또한, 상기 camera\_center\_yaw 필드는 상기 카메라에 의하여 캡처된 6DoF 비디오의 중점의 yaw 값, camera\_center\_pitch 필드는 상기 카메라에 의하여 캡처된 6DoF 비디오의 중점의 pitch 값, camera\_center\_roll 필드는 상기 카메라에 의하여 캡처된 6DoF 비디오의 중점의 roll 값을 나타낼 수 있다.
- [0298] 한편, 캡처 정보 메타데이터는 각 카메라에 대한 IntrinsicCameraParametersBox 및 ExtrinsicCameraParametersBox 를 포함할 수 있다.
- [0299] 상기 IntrinsicCameraParametersBox 는 다음의 표와 같이 도출될 수 있다.

표 4

```
IntrinsicCameraParameters() {
    unsigned int(6) reserved=0;
    unsigned int(10) ref_view_id;
    unsigned int(32) prec_focal_length;
    unsigned int(32) prec_principal_point;
    unsigned int(32) prec_skew_factor;
    unsigned int(8) exponent_focal_length_x;
    signed int(64) mantissa_focal_length_x;
    unsigned int(8) exponent_focal_length_y;
    signed int(64) mantissa_focal_length_y;
    unsigned int(8) exponent_principal_point_x;
    signed int(64) mantissa_principal_point_x;
    unsigned int(8) exponent_principal_point_y;
    signed int(64) mantissa_principal_point_y;
    unsigned int(8) exponent_skew_factor;
    signed int(64) mantissa_skew_factor;
}
```

[0300]

[0301]

상기 표 4를 참조하면 IntrinsicCameraParametersBox 는 ref\_view\_id 필드, prec\_focal\_length 필드, prec\_principal\_point 필드, prec\_skew\_factor 필드, exponent\_focal\_length\_x 필드, mantissa\_focal\_length\_x 필드, exponent\_focal\_length\_y 필드, mantissa\_focal\_length\_y 필드, exponent\_principal\_point\_x 필드, mantissa\_principal\_point\_x 필드, exponent\_principal\_point\_y 필드, mantissa\_principal\_point\_y 필드, exponent\_skew\_factor 필드 및/또는 mantissa\_skew\_factor 필드를 포함할 수 있다.

[0302]

상기 ref\_view\_id 필드는 해당 카메라의 시점(view)을 식별하는 view\_id 를 지시할 수 있다. 상기 prec\_focal\_length 필드는 focal\_length\_x 및 focal\_length\_y 에 허용되는 최대 절단(truncation) 에러의 익스포넌트(exponent) 를 특정할 수 있다. 상기 최대 절단 에러는  $2^{(-prec\_focal\_length)}$  와 같이 나타내어질 수 있다. 상기 prec\_principal\_point 필드는 principal\_point\_x 및 principal\_point\_y 에 허용되는 최대 절단(truncation) 에러의 익스포넌트(exponent) 를 특정할 수 있다. 상기 최대 절단 에러는  $2^{(-prec\_principal\_point)}$  와 같이 나타내어질 수 있다.

[0303]

상기 prec\_skew\_factor 필드는 skew 팩터에 허용되는 최대 절단(truncation) 에러의 익스포넌트(exponent) 를 특정할 수 있다.  $2^{(-prec\_skew\_factor)}$  와 같이 나타내어질 수 있다.

[0304]

상기 exponent\_focal\_length\_x 필드는 수평 방향의 focal 길이의 익스포넌트(exponent) 파트를 지시할 수 있다. mantissa\_focal\_length\_x 필드는 수평 방향의 i 번째 카메라의 focal 길이의 mantissa 파트를 지시할 수 있다. exponent\_focal\_length\_y 필드는 수직 방향의 focal 길이의 익스포넌트(exponent) 파트를 지시할 수 있다. mantissa\_focal\_length\_y 필드는 수직 방향의 focal 길이의 mantissa 파트를 지시할 수 있다.

[0305]

상기 exponent\_principal\_point\_x 필드는 수평 방향의 주점(principal point)의 익스포넌트 파트를 지시할 수 있다. 상기 mantissa\_principal\_point\_x 필드는 수평 방향의 주점(principal point) 의 mantissa 파트를 지시할 수 있다. 상기 exponent\_principal\_point\_y 필드는 수직 방향의 주점(principal point)의 익스포넌트 파트를 지시할 수 있다. 상기 mantissa\_principal\_point\_y 필드는 수직 방향의 주점(principal point) 의 mantissa 파트를 지시할 수 있다.

[0306]

상기 exponent\_skew\_factor 필드는 skew 팩터의 익스포넌트 파트를 지시할 수 있다. 상기 mantissa\_skew\_factor 필드는 skew 팩터의 mantissa 파트를 지시할 수 있다.

[0307]

또한, 상기 ExtrinsicCameraParametersBox 는 월드 좌표계(known world reference frame) 를 기준으로 카메라 레퍼런스 프레임의 위치, 오리엔테이션을 정의하는 카메라 파라미터들을 포함할 수 있다. 즉, 상기 ExtrinsicCameraParametersBox 는 월드 좌표계를 기준으로 각 카메라의 회전(rotation) 및 평행이동(translation) 등에 관한 내용을 나타내는 파라미터들을 포함할 수 있다.

[0308]

예를 들어, 상기 ExtrinsicCameraParametersBox 는 다음의 표와 같이 도출될 수 있다.

표 5

```

ExtrinsicCameraParameters() {
    unsigned int(6) reserved=0;
    unsigned int(10) ref_view_id;
    unsigned int(8) prec_rotation_param;
    unsigned int(8) prec_translation_param;
    for (j=1; j<=3; j++) { /* row */
        for (k=1; k<=3; k++) { /* column */
            unsigned int(8) exponent_r[j][k];
            signed int(64) mantissa_r [j][k];
        }
        unsigned int(8) exponent_t[j];
        signed int(64) mantissa_t[j];
    }
}
    
```

[0309]

[0310]

상기 표 5를 참조하면 상기 ExtrinsicCameraParametersBox 는 ref\_view\_id 필드, prec\_rotation\_param 필드, prec\_translation\_param 필드, exponent\_r[j][k] 필드, mantissa\_r [j][k] 필드, exponent\_t[j] 필드 및/또는 mantissa\_t[j] 필드를 포함할 수 있다.

[0311]

상기 ref\_view\_id 필드는 내부 카메라 파라미터들과 관계된 시점(view)을 식별하는 view\_id 를 지시할 수 있다.

[0312]

상기 prec\_rotation\_param 필드는 r[j][k] 에 허용되는 최대 절단(truncation) 에러의 익스포넌트 파트를 특정할 수 있다. 이는 2-prec\_rotation\_param와 같이 표현될 수 있다.

[0313]

상기 prec\_translation\_param 필드는 t[j] 에 허용되는 최대 절단(truncation) 에러의 익스포넌트 파트를 특정할 수 있다. 이는 2-prec\_translation\_param와 같이 표현될 수 있다.

[0314]

상기 exponent\_r[j][k] 필드는 로테이션 매트릭스의 (j, k) 컴포넌트의 익스포넌트 파트를 특정할 수 있다. mantissa\_r [j][k] 필드는 로테이션 매트릭스의 (j, k) 컴포넌트의 mantissa 파트를 특정할 수 있다. exponent\_t[j] 필드는 트랜스레이션 벡터의 j 번째 컴포넌트의 익스포넌트 파트를 특정할 수 있다. 이는 0 에서 62 사이의 값을 가질 수 있다. mantissa\_t[j] 필드는 트랜스레이션 벡터의 j 번째 컴포넌트의 mantissa 파트를 특정할 수 있다.

[0315]

또한, 상기 6DoF 비디오 관련 메타데이터는 상기 스티칭 정보 메타데이터를 포함할 수 있다.

[0316]

일 예로, 상기 스티칭 정보 메타데이터는 캡처된 6DoF 비디오를 스티칭할 필요가 있을 때, 여러 개의 캡처 영상 중에 가장 초기 위치의 캡처 영상을 기준으로 나머지 영상의 위치 선정 및 스티칭이 이루어지는데 사용될 수 있다. 또한, 프로젝션 스킵에 따라 스티칭 결과가 다르게 나타날 수 있다. 이 경우, 상기 스티칭을 위하여 필요한 정보는 스티칭이 수행되는지 여부를 나타내는 정보와, 스티칭이 수행되는 경우에 한 개의 3차원 프레임 만들기 위하여 스티칭되는 시점(view)의 수, 스티칭되는 시점을 찾기 위한 기준 거리값과 스티칭이 수행되는 경우에 기준이 되는 시점들의 카메라 정보가 될 수 있다. 또한, 실시예에 따라 스티칭 정보 메타데이터는 추가적인 정보들을 더 포함하거나 상술한 정보가 제외될 수도 있다.

[0317]

예를 들어, 상기 스티칭 정보 메타데이터는 다음의 표와 같이 나타낼 수 있다.

표 6

```

if (stitching_flag == 1) {
    unsigned int(8) num_view_center;
    for (i = 1; i<= num_view_center;i++){
        signed int(16) position_center_X;
        signed int(16) position_center_Y;
        signed int(16) position_center_Z;
        unsigned int(8) num_camera;
        unsigned int(8) stitching_radius;
    }
}
    
```

[0318]

[0319] 여기서, `stiching_flag` 필드는 스티칭이 수행되는지 여부를 나타내는 필드, `num_view_center` 필드는 스티칭되는 기준 시점(view)의 수를 나타내는 필드, `position_center_X` 필드, `position_center_Y` 필드, `position_center_Z` 필드는 해당 기준 시점의 중심 포지션을 나타내는 필드들, `num_camera` 필드는 해당 기준 시점에 대한 스티칭되는 캡처 영상의 수를 나타내는 필드, `stitching_radius` 필드는 해당 기준 시점의 센터 포지션에서 스티칭되는 캡처 영상들간의 거리를 나타내는 필드를 나타낸다. 구체적으로 `stiching_flag` 필드의 값이 1인 경우, 6DoF 비디오 데이터에 대한 스티칭이 수행될 수 있고, `stiching_flag` 필드의 값이 0인 경우, 6DoF 비디오 데이터에 대한 스티칭이 수행되지 않을 수 있다. 스티칭이 일어나지 않고 2차원 이미지로 프로젝션되는 경우가 있을 수 있으므로 스티칭은 `stiching_flag` 필드가 참(true)인 경우에만 실시될 수 있다. 또한, 스티칭이 수행되는 경우, 기준 시점의 포지션 정보가 설정될 수 있다.

[0320] 스티칭이 수행되는 경우, 기준 시점 정보를 바탕으로 기준 시점의 수는 `num_view_center` 필드, 각 기준 시점의 센터 포지션 정보는 `position_center_X`, `position_center_Y`, `position_center_Z`로 정의될 수 있고, 각 시점별 스티칭이 되는 캡처 영상의 수는 `num_camera` 필드로 정의될 수 있다. 또한, `stitching_radius` 필드는 기준 시점의 센터 포지션과 스티칭되는 캡처 영상들간의 거리를 나타내는 필드로 정의될 수 있다. 구체적으로, `num_camera` 필드의 값만큼의 캡처 비디오를 찾기 위해 해당 기준 시점의 센터 포지션에서 `stitching_radius` 필드만큼 떨어진 위치를 중심으로 `num_camera` 필드 값의 수만큼의 캡처 영상이 스티칭되는 영상으로 설정될 수 있다.

[0321] 구체적으로, `position_center_X`, `position_center_Y`, `position_center_Z` 는 대상 기준 시점의 중심 포지션을 나타내는 필드는 각각 상기 대상 기준 시점의 센터 포지션의 x 성분, y 성분, z 성분을 나타낼 수 있다. 상기 기준 시점의 중심 포지션은 3차원 축 상에서 절대 위치 좌표 값으로 나타낼 수 있다. 한편, 상기 기준 시점의 중심 포지션을 나타내는 상기 3차원 공간의 원점이 되는 기준점은 상술한 `Camera_info` 에서 정의한 `Master_camera_ID` 필드가 나타내는 정보를 기준으로 하여 설정될 수 있다.

[0322] 또한, `num_camera` 필드는 대상 기준 시점의 스티칭이 되는 캡처 영상의 수를 나타낼 수 있다. 또한, `stitching_radius` 필드는 대상 기준 시점의 센터 포지션과 스티칭되는 캡처 영상들간의 거리를 나타낼 수 있다. 상기 거리의 단위는 시스템의 요구사항에 맞게 정의될 수 있다. 예를 들어, 상기 거리의 단위는 원점으로부터의 절대 거리 (cm 혹은 mm) 단위로 정의될 수 있고, 또는 상대 거리(normalized distance)로 정의될 수도 있다.

[0323] 또한, 상기 6DoF 비디오 관련 메타데이터는 상기 프로젝션 스킴 메타데이터를 포함할 수 있다. 예를 들어, 상기 프로젝션 스킴 메타데이터는 다음의 표와 같이 도출될 수 있다.

**표 7**

```

if((projection_scheme == '0')) {
    unsigned int(16)    sphere_radius;
} else if(projection_scheme == '1') { //cylindrical projection
    unsigned int(16)    cylinder_radius;
    unsigned int(16)    cylinder_height;
} else if(projection_scheme == '2'){ //cubic projection
    unsigned int(16)    cube_front_width;
    unsigned int(16)    cube_front_height;
    unsigned int(16)    cube_height;
} else if(projection_scheme == '3'){ //panorama projection
    unsigned int(16)    panorama_height;
}
    
```

[0324]

[0325] 표 7을 참조하면 상기 프로젝션 스킴 메타데이터는 `projection_scheme` 필드를 포함할 수 있다. 상기 스티칭이 수행된 이후에 상기 `projection_scheme` 필드는 해당 시점에서의 스티칭된 비디오 데이터가 2D 이미지 상에 프로젝션되는 경우에 사용되는 프로젝션 스킴을 지시할 수 있다. 즉, 상기 `projection_scheme` 필드는 해당 시점에서의 6DoF 비디오 데이터가 2D 이미지 상에 프로젝션될 때 사용된 프로젝션 스킴을 지시할 수 있다. 상기 프로젝션 스킴은 360도 비디오에 대한 프로젝션 스킴과 같거나 유사할 수 있다. 예를 들어, 상기 `projection_scheme` 필드가 0, 1, 2, 3의 값을 가지는 경우, 각각 등정방형 프로젝션(Equirectangular Projection) 스킴, 실린더형 프로젝션 스킴, 큐브 프로젝션 스킴, 파노라믹 프로젝션 스킴을 나타낼 수 있다.



[0326] 상기 projection\_scheme 필드는 해당 시점에서의 6DoF 비디오 데이터에 대한 프로젝션 스킴을 다음의 표와 같이 나타낼 수 있다.

표 8

projection_scheme	Description
0000	ERP(Equi-Rectangular_projection)
0001	Cylindrical
0010	Cube
0011	Panoramic
0100	User defined
0101-1111	Reserved

[0327]

[0328] 표 8을 참조하면 상기 projection\_scheme 필드가 4의 값을 가지는 경우, 상기 등정방향 프로젝션 스킴, 상기 실린더형 프로젝션 스킴, 상기 큐브 프로젝션 스킴 및 상기 파노라믹 프로젝션 스킴 이외의 사용자 임의로 설정한 방법으로 6DoF 비디오 데이터가 2D 이미지 상에 프로젝션된 경우를 나타낼 수 있다. 또한, 상기 projection\_scheme 필드가 0, 1, 2, 3의 값이외의 값을 가지는 경우는 향후 사용을 위해 남겨둘 수 있다 (Reserved for Future Use).

[0329] 상기 프로젝션 스킴 메타데이터는 상기 projection\_scheme 필드에 의하여 지시된 각각의 프로젝션 스킴(등정방향 프로젝션 스킴, 실린더형 프로젝션 스킴, 큐브 프로젝션 스킴, 파노라믹 프로젝션 스킴 등)에 대한 구체적인 정보를 제공할 수 있다.

[0330] 예를 들어, 상기 projection\_scheme 필드의 값이 0인 경우, 즉, 상기 projection\_scheme 필드가 프로젝션 스킴이 등정방향 프로젝션(Equirectangular Projection) 스킴임을 지시하는 경우, 프로젝션 스킴 메타데이터는 sphere\_radius 필드를 포함할 수 있다. 상기 sphere\_radius 필드는 해당 시점의 6DoF 비디오 데이터를 2D 이미지에 맵핑, 즉 프로젝션시에 적용한 구(sphere)의 반지름을 나타낼 수 있다.

[0331] 또한, 다른 예로, 상기 projection\_scheme 필드의 값이 1인 경우, 즉, 상기 projection\_scheme 필드가 프로젝션 스킴이 실린더형 프로젝션 스킴임을 지시하는 경우, 프로젝션 스킴 메타데이터는 cylinder\_radius 필드, 및/또는 cylinder\_height 필드를 포함할 수 있다. 상기 cylinder\_radius 필드는 해당 시점의 6DoF 비디오 데이터를 2D 이미지에 맵핑, 즉 프로젝션시에 적용한 실린더(cylinder)의 윗면(top)/바닥면(bottom)의 반지름을 나타낼 수 있고, 상기 cylinder\_height 필드는 해당 시점의 6DoF 비디오 데이터를 2D 이미지에 맵핑, 즉 프로젝션시에 적용한 실린더의 높이(height)를 나타낼 수 있다.

[0332] 또한, 다른 예로, 상기 projection\_scheme 필드의 값이 2인 경우, 즉, 상기 projection\_scheme 필드가 프로젝션 스킴이 큐브 프로젝션(Cubic Projection) 스킴임을 지시하는 경우, 상기 프로젝션 스킴 메타데이터는 cube\_front\_width 필드, cube\_front\_height 필드 및/또는 cube\_height 필드를 포함할 수 있다. 상기 cube\_front\_width 필드는 해당 시점의 6DoF 비디오 데이터를 2D 이미지에 맵핑, 즉 프로젝션시에 적용한 큐브의 앞면(front)의 가로길이(width)를 나타낼 수 있고, 상기 cube\_front\_height 필드는 상기 6DoF 비디오 데이터를 2D 이미지에 맵핑, 즉 프로젝션시에 적용한 큐브의 앞면의 세로길이(height)를 나타낼 수 있고, 상기 cube\_height 필드는 상기 6DoF 비디오 데이터를 2D 이미지에 맵핑, 즉 프로젝션시에 적용한 큐브의 앞면과 옆면들 사이의 세로길이(height)를 나타낼 수 있다.

[0333] 또한, 다른 예로, 상기 projection\_scheme 필드의 값이 3인 경우, 즉, 상기 projection\_scheme 필드가 프로젝션 스킴이 파노라믹 프로젝션 스킴임을 지시하는 경우, 상기 프로젝션 스킴 메타데이터는 panorama\_height 필드를 포함할 수 있다. 상기 panoramic\_height 필드는 360도 비디오 데이터를 2D 이미지에 맵핑, 즉 프로젝션시에 적용한 파노라마(panorama)의 높이(height)를 나타낼 수 있다.

[0334] 한편, 실시예에 따라 6DoF 비디오 관련 메타데이터는 상기 projection\_scheme 필드에 의해 특정되는 프로젝션 스킴에 의해 발생한 리전(Region)에 대한 구체적인 정보를 더 포함할 수 있다. 여기서 리전에 대한 구체적인 정

보란 예를 들어, 리전의 회전 여부, 실린더의 윗면(top) 리전의 반지름 정보 등을 의미할 수 있다.

[0335] 또한, 상기 6DoF 비디오 관련 메타데이터는 초기 위치 정보 및 FOV(Field of View) 메타데이터를 포함할 수 있다. 예를 들어, 상기 초기 위치 정보 및 FOV 메타데이터는 다음의 표와 같이 도출될 수 있다.

표 9

signed	int(8)	initial_view_yaw_degree;
signed	int(8)	initial_view_pitch_degree;
signed	int(8)	initial_view_roll_degree;
signed	int(8)	initial_center_FB_delta;
signed	int(8)	initial_center_UD_delta;
signed	int(8)	initial_center_RL_delta;
unsigned	int(5)	reserved = 0;
unsigned	int(1)	content_fov_flag;
unsigned	int(1)	capture_info_flag;
unsigned	int(1)	region_info_flag;
if (content_fov_flag == 1) {		
	unsigned int(16)	content_hfov;
	unsigned int(16)	content_vfov;
}		

[0336]

[0337] 표 9를 참조하면 상기 초기 위치 정보 및 FOV 메타데이터는 사용자가 6DoF 비디오를 처음 재생했을 때 보게되는 시점(초기 시점)에 대한 정보와 FOV에 관련한 정보들을 포함할 수 있다. 상기 초기 위치 정보 및 FOV 메타데이터는 initial\_view\_yaw\_degree 필드, initial\_view\_pitch\_degree 필드, initial\_view\_roll\_degree 필드, initial\_center\_FB\_delta 필드, initial\_center\_UD\_delta 필드, initial\_center\_RL\_delta 필드, content\_fov\_flag 필드, capture\_info\_flag 필드, region\_info\_flag 필드, content\_hfov 필드 및/또는 content\_vfov 필드를 포함할 수 있다. 실시예에 따라 상기 초기 위치 정보 및 FOV 메타데이터는 추가적인 정보들을 더 포함할 수도 있다.

[0338] 상기 initial\_view\_yaw\_degree 필드, 상기 initial\_view\_pitch\_degree 필드, 상기 initial\_view\_roll\_degree 필드는 해당 6DoF 비디오 제공시 사용자의 초기 시점(initial view)의 정중앙 지점을 나타낼 수 있다. 즉, 상기 initial\_view\_yaw\_degree 필드, 상기 initial\_view\_pitch\_degree 필드, 상기 initial\_view\_roll\_degree 필드는 상기 초기 시점에 대한 6DoF 비디오 데이터가 맵핑된 3D 공간의 정중앙 지점을 나타낼 수 있다. 따라서, 재생시 처음 보여지는 뷰포트의 정중앙 지점이, 이 세 필드들에 의해 나타내어질 수 있다. 구체적으로, 상기 initial\_view\_yaw\_degree 필드는 상기 초기 시점의 정중앙 지점에 대한 yaw 값을 나타낼 수 있다. 즉, 상기 initial\_view\_yaw\_degree 필드는 상기 정중앙 지점의 위치를 yaw 축을 기준으로 회전된 방향(부호) 및 그 정도(각도)로 나타낼 수 있다. 또한, 상기 initial\_view\_pitch\_degree 필드는 상기 초기 시점의 정중앙 지점에 대한 pitch 값을 나타낼 수 있다. 즉, 상기 initial\_view\_pitch\_degree 필드는 상기 정중앙 지점의 위치를 pitch 축을 기준으로 회전된 방향(부호) 및 그 정도(각도)로 나타낼 수 있다. 또한, 상기 initial\_view\_roll\_degree 필드는 상기 초기 시점의 정중앙 지점에 대한 roll 값을 나타낼 수 있다. 즉, 상기 initial\_view\_roll\_degree 필드는 상기 정중앙 지점의 위치를 roll 축을 기준으로 회전된 방향(부호) 및 그 정도(각도)로 나타낼 수 있다.

[0339] 상기 initial\_center\_FB\_delta 필드, initial\_center\_UD\_delta 필드, initial\_center\_RL\_delta 필드는 해당 6DoF 비디오 제공시 사용자의 초기 시점(initial view)의 위치를 나타낼 수 있다. 구체적으로, 상기 initial\_center\_FB\_delta 필드는 상기 초기 시점의 위치의 앞뒤 방향 축에 대한 성분을 나타낼 수 있다. 즉, 상기 initial\_center\_FB\_delta 필드는 해당 6DoF 비디오 제공시 사용자가 보게될 비디오의 초기 시점(initial view)의 위치가 원점(original point)으로부터 앞뒤 방향(Forward-backward)으로 얼마만큼 떨어져 있는지에 대한 정도를 나타낼 수 있다. 또한, 상기 initial\_center\_UD\_delta 필드는 상기 초기 시점의 위치의 위아래 방향 축에 대한 성분을 나타낼 수 있다. 즉, 상기 initial\_center\_UD\_delta 필드는 해당 6DoF 비디오 제공시 사용자가 보게될 비디오의 초기 시점(initial view)의 위치가 원점(original point)으로부터 위아래 방향(Up-Down)으로 얼마만큼 떨어져 있는지에 대한 정도를 나타낼 수 있다. 또한, 상기 initial\_center\_RL\_delta 필드는 상기 초기 시점의 위치의 좌우 방향 축에 대한 성분을 나타낼 수 있다. 즉, 상기 initial\_center\_RL\_delta 필드는 해당 6DoF 비디오 제공시 사용자가 보게될 비디오의 초기 시점(initial view)의 위치가 원점(original point)으로부터 좌우 방향(Right-Left)으로 얼마만큼 떨어져 있는지에 대한 정도를 나타낼 수 있다.

[0340] 상기 content\_fov\_flag 필드는 해당 6DoF 비디오에 대하여 제작시 의도한 FOV에 대한 정보가 존재하는지 여부를

지시할 수 있다. 예를 들어, 해당 6DoF 비디오 제작시 의도한 FOV 값이 없는 경우, 상기 content\_fov\_flag 필드의 값은 0으로 도출될 수 있고, 해당 6DoF 비디오 제작시 의도한 FOV 값이 있는 경우, 상기 content\_fov\_flag 필드의 값은 1으로 도출될 수 있다. 상기 FOV 값은 content\_hfov 필드 및/또는 content\_vfov 필드로 정의될 수 있다. content\_fov\_flag 필드의 값이 1인 경우, content\_hfov 필드 및/또는 content\_vfov 필드가 존재할 수 있다.

[0341] 상기 content\_hfov 필드는 해당 6DoF 비디오에 대하여 제작시 의도한 콘텐츠(content)의 수평 FOV(horizontal field of view)에 대한 정보를 나타낼 수 있다. 즉, 상기 content\_hfov 필드는 해당 6DoF 비디오에 대하여 제작시 의도한 콘텐츠(content)의 FOV의 수평 방향 값을 나타낼 수 있다. 해당 6DoF 비디오 수신 장치의 수평 FOV에 따라, 6DoF 비디오 중에서 사용자에게 한번에 디스플레이되는 영역이 결정될 수 있다. 혹은 실시예에 따라 본 필드의 수평 FOV 정보를 반영하여 사용자에게 한번에 디스플레이되는 6DoF 비디오의 영역이 결정될 수도 있다. 6DoF 비디오 수신 장치는 상기 수평 FOV 정보를 반영하여 사용자에게 한번에 디스플레이되는 6DoF 비디오의 FOV를 조절하여 보여줄 수 있고, 또는 상기 FOV를 지원해 줄 수 있는 동작을 수행할 수 있다.

[0342] 상기 content\_vfov 필드는 해당 6DoF 비디오에 대하여 제작시 의도한 콘텐츠(content)의 수직 FOV(vertical field of view)에 대한 정보를 나타낼 수 있다. 즉, 상기 content\_vfov 필드는 해당 6DoF 비디오에 대하여 제작시 의도한 콘텐츠(content)의 FOV의 수직 방향 값을 나타낼 수 있다. 해당 6DoF 비디오 수신 장치의 수직 FOV에 따라, 6DoF 비디오 중에서 사용자에게 한번에 디스플레이되는 영역이 결정될 수 있다. 혹은 실시예에 따라 본 필드의 수직 FOV 정보를 반영하여 사용자에게 한번에 디스플레이되는 6DoF 비디오의 영역이 결정될 수도 있다. 6DoF 비디오 수신 장치는 상기 수직 FOV 정보를 반영하여 사용자에게 한번에 디스플레이되는 6DoF 비디오의 FOV를 조절하여 보여줄 수 있고, 또는 상기 FOV를 지원해 줄 수 있는 동작을 수행할 수 있다. 상기 초기 위치 정보 및 FOV 메타데이터는 해당 6DoF 비디오 데이터의 리전(Region) 들과 관련된 메타데이터를 제공할 수 있다. region\_info\_flag 필드는 해당 6DoF 비디오 데이터가 프로젝션된 2D 이미지가 하나 이상의 리전으로 나뉘어져 있는지 여부를 지시할 수 있다. 동시에 본 필드는 6DoF 비디오 관련 메타데이터가 각각의 리전에 대한 세부 정보를 포함하는지 여부를 지시할 수 있다.

[0343] 또한, 상기 초기 위치 정보 및 FOV 메타데이터는 상기 캡처 정보 메타데이터를 제공할 수 있다. capture\_info\_flag 필드의 값이 참(true)인 경우에 전송될 수 있다. 즉, 상기 capture\_info\_flag 필드는 상기 캡처 정보 메타데이터의 전송 여부를 나타낼 수 있다. 상기 capture\_info\_flag 필드의 값이 1인 경우, 상기 캡처 정보 메타데이터가 전송될 수 있고, 상기 capture\_info\_flag 필드의 값이 0인 경우, 상기 캡처 정보 메타데이터는 전송되지 않을 수 있다. 한편, 상기 capture\_info\_flag 필드는 상술한 camera\_info\_flag 필드와 동일할 수 있다.

[0344] 구체적으로, 상기 리-프로젝션 정보 메타데이터는 현재 재생하고 있는 위치, 즉, 현재 사용자에게 제공되고 있는 6DoF 비디오에 대한 위치를 나타내는 사용자 위치 정보와 미리 설정된 진행 방향에 존재하는 예비 렌더링 위치에 대한 정보를 포함할 수 있다. 상기 리-프로젝션 정보 메타데이터는 렌더링 정보 메타데이터라고 나타낼 수도 있다. 6DoF 비디오를 리-프로젝션하기 위해서는 초기 뷰포트 정보와 이동 경로를 예측하여 현재 사용자가 보고 있는 3차원 이미지/비디오의 중심 혹은 초기 위치에 대한 6DoF 비디오/이미지의 중심에서의 사용자의 이동 경로 혹은 제작자의 의도하는 진행 방향을 바탕으로 보여질 6DoF 비디오/이미지 정보를 미리 렌더링 하는 정보가 필요할 수 있다. 상기 정보를 기반으로 가상 시점에서의 영상 생성/합성 및 렌더링 지연을 방지할 수 있다. 이 경우, 상기 리-프로젝션 정보 메타데이터는 미리 생성할 6DoF 콘텐츠의 형태 및 실시예에 따라 사용자 위치 정보를 나타내는 필드들을 포함할 수 있고, 추가적인 정보들에 대한 필드들을 더 포함할 수도 있다.

[0345] 예를 들어, 상기 리-프로젝션 정보 메타데이터는 다음의 표와 같이 도출될 수 있다.

표 10

```

Reserved_rendering_range (){
    unsigned int(8) reserved_range_type;
    signed int(16) center_X; //user center
    signed int(16) center_Y;
    signed int(16) center_Z;
    unsigned int(8) step_size;
    unsigned int(8) rendering_type;
    if (reserved_range_type =='0') {
        unsigned int(8) sphere_radius_to_boundary;
    }
    else if(reserved_range_type =='1'){
        unsigned int(8) horizontal_range;
        unsigned int(8) vertical_range;
    }
    else if(reserved_range_type =='2'){
        unsigned int(8) distance_to_top_endpoint;
        unsigned int(8) distance_to_bottom_endpoint;
    }
    else if(reserved_range_type =='3'){
        unsigned int(8) distance_to_left_endpoint;
        unsigned int(8) distance_to_right_endpoint;
    }
    else if(reserved_range_type =='4'){
        unsigned int(8) distance_to_front_endpoint;
        unsigned int(8) distance_to_back_endpoint;
    }
    else if(reserved_range_type =='5'){
        signed int(16) vertex_X[];
        signed int(16) vertex_Y[];
        signed int(16) vertex_Z[];
        unsigned int(1) interpolate;
        unsigned int(7) reserved;
    }

    If(rendering_type=='0'){
        unsigned int(16) sphere_radius;
    }
    If(rendering_type=='1'){
        signed int(16) preliminary_viewport_yaw;
        signed int(16) preliminary_viewport_pitch;
        signed int(16) preliminary_viewport_roll;
        unsigned int(8) preliminary_horl_range;
        unsigned int(8) preliminary_ver_range;
    }
}
    
```

[0346]

[0347]

표 10을 참조하면 상기 리-프로젝션 정보 메타데이터는 reserved\_range\_type 필드, center\_X 필드, center\_Y 필드, center\_Z 필드, step\_size 필드, rendering\_type 필드, sphere\_radius\_to\_boundary 필드, horizontal\_range 필드, vertical\_range 필드, distance\_to\_top\_endpoint 필드, distance\_to\_bottom\_endpoint 필드, distance\_to\_left\_endpoint 필드, distance\_to\_right\_endpoint 필드, distance\_to\_front\_endpoint 필드, distance\_to\_back\_endpoint 필드, vertex\_X[] 필드, vertex\_Y[] 필드, vertex\_Z[] 필드, sphere\_radius 필드, preliminary\_viewport\_yaw 필드, preliminary\_viewport\_pitch 필드, preliminary\_viewport\_roll 필드, preliminary\_hor\_range 필드 및/또는 preliminary\_ver\_range 필드를 포함할 수 있다. 실시예에 따라 상기 리-프로젝션 정보 메타데이터는 추가적인 정보들을 더 포함할 수도 있다. 한편, 상기 리-프로젝션 정보 메타데이터는 reserved\_rendering\_range\_flag 필드의 값이 참(true)인 경우에 전송될 수 있다. 즉, 상기 reserved\_rendering\_range\_flag 필드는 상기 리-프로젝션 정보 메타데이터의 전송 여부를 나타낼 수 있다. 상기 reserved\_rendering\_range\_flag 필드의 값이 1인 경우, 상기 리-프로젝션 정보 메타데이터가 전송될 수 있고, 상기 reserved\_rendering\_range\_flag 필드의 값이 0인 경우, 상기 리-프로젝션 정보 메타데이터는 전송되지 않을 수 있다.

[0348]

현재 재생하고 있는 위치와 함께 미리 설정된 진행 방향에 존재하는 예비 렌더링 위치의 정보를 정의하기 위해서는 렌더링 타입과 예비 렌더링 범위의 형태, 현재 사용자의 위치 정보, 사용자 중심이 이동하는 경우의 이동 정도(step size), 전체의 이동 범위 등의 정보가 필요할 수 있다. 상기 예비 렌더링 범위는 예비 렌더링 영역이라고 나타낼 수도 있다. 이 경우, 렌더링을 위하여 사용되는 6DoF 비디오 혹은 이미지는 PV(primary view)에 대

한 6DoF 비디오 혹은 이미지를 중심으로 형성될 수도 있고, PV와 SV에 대한 6DoF 비디오 혹은 이미지의 일부 혹은 전체를 중심으로 형성될 수도 있다.

[0349] 한편, 상기 예비 렌더링 범위의 형태는 후술하는 내용과 같이 설명될 수 있다.

[0350] 도 19a 내지 도 19e는 본 발명의 상기 예비 렌더링 범위의 형태를 설명하기 위하여 도시한 도면이다. 상기 예비 렌더링 범위의 형태는 상기 reserved\_range\_type 필드로 정의될 수 있다. 즉, 상기 reserved\_range\_type 필드는 상기 예비 렌더링 범위의 타입을 나타낼 수 있다. 상기 예비 렌더링 범위의 타입은 도 19a에 도시된 것과 같은 사용자 위치를 중심으로 동일한 반경으로 형성되는 스피어(sphere) 타입, 도 19b에 도시된 것과 같은 사용자 위치를 중심으로 가로 세로 방향으로 뻗어나가는 탑-셰입(top-shape) 타입, 도 19c에 도시된 것과 같은 사용자의 위치를 중심으로 위 아래의 이동만 존재하는 엘리베이션 디렉션(elevation direction) 타입, 도 19d에 도시된 것과 같은 사용자 위치를 중심으로 전진하거나 후진하는 방향으로 진행되는 레터럴 디렉션(lateral direction) 타입, 도 19e에 도시된 것과 같은 사용자 위치를 중심으로 왼쪽으로 이동하거나 오른쪽으로 이동하는 경로만 가지는 롱지튜디널 디렉션(longitudinal direction) 타입이 있을 수 있다. 또한 상기 예비 렌더링 범위의 타입으로 자유로운 형태의 예비 렌더링 영역이 정의될 수도 있는바, vertex 기반으로 정의되는 user defined 타입이 있을 수 있다. 그 외 여러 진행 방향의 형태가 발생될 수 있으며, 해당 형태는 상기 예비 렌더링 범위의 타입으로 추가될 수 있다.

[0351] 한편, 상기 reserved\_range\_type 필드는 상기 예비 렌더링 범위의 타입을 나타낼 수 있다. 예를 들어, 상기 reserved\_range\_type 필드는 상기 예비 렌더링 범위의 타입을 다음의 표와 같이 나타낼 수 있다.

표 11

reserved_range_type	Description
0000	sphere
0001	Top-shape
0010	Elevation-direction
0011	Lateral direction
0100	Longitudinal direction
0101	User defined
0110-1111	Reserved

[0352]

[0353] 상기 표 11을 참조하면 상기 reserved\_range\_type 필드의 값을 기반으로 상기 예비 렌더링 범위의 타입이 도출될 수 있다. 구체적으로, 상기 reserved\_range\_type 필드의 값이 0인 경우, 상기 예비 렌더링 범위의 타입은 상기 스피어 타입으로 도출될 수 있고, 상기 reserved\_range\_type 필드의 값이 1인 경우, 상기 예비 렌더링 범위의 타입은 상기 탑-셰입 타입으로 도출될 수 있고, 상기 reserved\_range\_type 필드의 값이 2인 경우, 상기 예비 렌더링 범위의 타입은 상기 엘리베이션 디렉션 타입으로 도출될 수 있고, 상기 reserved\_range\_type 필드의 값이 3인 경우, 상기 예비 렌더링 범위의 타입은 상기 레터럴 디렉션 타입으로 도출될 수 있고, 상기 reserved\_range\_type 필드의 값이 4인 경우, 상기 예비 렌더링 범위의 타입은 상기 롱지튜디널 디렉션 타입으로 도출될 수 있고, 상기 reserved\_range\_type 필드의 값이 5인 경우, 상기 예비 렌더링 범위의 타입은 상기 user defined 타입으로 도출될 수 있다. 또한, 상기 reserved\_range\_type 필드가 0 내지 5의 값이외의 값을 가지는 경우는 향후 사용을 위해 남겨둘 수 있다(Reserved for Future Use).

[0354] 예비 렌더링 범위가 정의된 경우, 즉, 상기 예비 렌더링 범위의 타입이 설정된 경우에는 미리 렌더링할 위치의 비디오 혹은 이미지의 정보가 이동한 위치를 중심으로 어떻게 렌더링되는지를 나타내는 렌더링 타입이 정의될 수 있다. 다시 말해, 예비 렌더링 위치의 비디오 혹은 이미지의 정보가 상기 예비 렌더링 위치에서 어떻게 렌더링되는지를 나타내는 렌더링 타입이 정의될 수 있고, rendering\_type 필드는 상기 예비 렌더링 위치의 렌더링 타입을 나타낼 수 있다.

[0355] 도 20은 본 발명의 상기 예비 렌더링 위치의 상기 렌더링 타입을 설명하기 위하여 도시한 도면이다. 도 20을 참

조하면 화살표 방향으로 진행되는 6DoF 비디오 혹은 이미지인 경우에 사용자의 위치는 화살표 방향으로 이동될 수 있다. 상기 렌더링 타입은 도 20의 (a)에 도시된 것과 이동된 사용자의 위치(상기 예비 렌더링 위치)에서 새로운 360도 비디오 혹은 이미지를 전체 3D 공간(즉, 구(sphere) 전체)에 렌더링하는 홀 스피어(whole sphere) 렌더링 타입과 도 20의 (b)에 도시된 것과 같이 이동된 사용자의 위치에서 현재 사용자가 보고 있는 뷰포트(viewport), 예를 들어, 특정 요(yaw), 피치(pitch), 롤(roll)로 지정되는 중점과 수평 범위(horizontal range) 및 수직 범위(vertical range)로 정의되는 뷰포트(viewport)와 같은 파셜 뷰포트(partial viewport)만 미리 렌더링하는 파셜 뷰포트 렌더링 타입이 있을 수 있다. 또한, 이동된 사용자의 위치에 대한 3D 공간, 즉, 구 상에 다수의 혹은 비정형 형태의 제작자가 의도한 영역을 미리 렌더링하는 user defined 렌더링 타입이 있을 수도 있다.

[0356] 한편, 상기 rendering\_type 필드는 상기 예비 렌더링 위치의 렌더링 타입을 나타낼 수 있다. 예를 들어, 상기 rendering\_type 필드는 상기 렌더링 타입을 다음의 표와 같이 나타낼 수 있다.

표 12

rendering_type	Description
0000	Whole sphere
0001	Partial viewport
0100	User defined
0101~1111	Reserved

[0357]

[0358] 상기 표 12를 참조하면 상기 rendering\_type 필드의 값을 기반으로 상기 렌더링 타입이 도출될 수 있다. 구체적으로, 상기 rendering\_type 필드의 값이 0인 경우, 상기 렌더링 타입은 상기 홀 스피어 렌더링 타입으로 도출될 수 있고, 상기 rendering\_type 필드의 값이 1인 경우, 상기 렌더링 타입은 상기 파셜 뷰포트 렌더링 타입으로 도출될 수 있고, 상기 rendering\_type 필드의 값이 2인 경우, 상기 렌더링 타입은 상기 user defined 타입으로 도출될 수 있다. 또한, 상기 rendering\_type 필드가 0 내지 2의 값이외의 값을 가지는 경우는 향후 사용을 위해 남겨둘 수 있다(Reserved for Future Use).

[0359] 한편, 상기 리-프로젝션 정보 메타데이터에 포함된 필드들의 구체적인 설명은 후술하는 바와 같을 수 있다.

[0360] 상기 center\_X 필드, 상기 center\_Y 필드 및 상기 center\_Z 필드는 현재 6DoF 비디오 혹은 이미지를 보고 있는 사용자의 위치를 나타낼 수 있다. 즉, 상기 center\_X 필드, 상기 center\_Y 필드 및 상기 center\_Z 필드는 사용자의 위치의 3차원 공간 상에서 X, Y, Z 좌표 값을 나타낼 수 있다. 해당 값은 상기 사용자의 위치에 따라 변화할 수 있고, 6DoF 콘텐츠의 범위를 벗어나지 않는 범위에서 정의될 수 있다.

[0361] 상기 step\_size 필드는 상기 center\_X 필드, 상기 center\_Y 필드 및 상기 center\_Z 필드가 나타내는 사용자의 위치로부터 렌더링할 위치(예비 렌더링 위치)의 중심까지의 거리를 나타낼 수 있다. 즉, 상기 step\_size 필드는 사용자의 현재 위치로부터 상기 예비 렌더링 위치까지의 거리를 나타낼 수 있다. 따라서 상기 step\_size 필드는 사용자가 이동할 위치를 중심으로 예비 렌더링할 360도 비디오 혹은 이미지의 중심 위치를 나타낼 수 있다. 이는 예비 렌더링할 360도 비디오 혹은 이미지의 수를 의미할 수 있다. 즉, 상기 step\_size 필드의 값이 작을수록 정의된 예비 렌더링 범위 내에서 렌더링되는 360도 비디오 혹은 이미지의 수가 많아지며, 값이 클수록 예비 렌더링 범위 내에서 듬성듬성하게 예비 렌더링이 수행되는 것을 의미할 수 있다. 사용자의 위치로부터 예비 렌더링 영역의 각 경계까지의 거리를 step\_size 필드 값으로 나눈 값이 예비 렌더링 위치의 중심 개수로 도출될 수 있다.

[0362] 상기 sphere\_radius\_to\_boundary 필드는 예비 렌더링 범위의 형태가 스피어(sphere)인 경우에 전달될 수 있다. 즉, 상기 sphere\_radius\_to\_boundary 필드는 상기 reserved\_range\_type 필드가 예비 렌더링 범위의 타입이 스피어 타입임을 나타내는 경우에 전달될 수 있다. 상기 sphere\_radius\_to\_boundary 필드는 상기 center\_X 필드, 상기 center\_Y 필드 및 상기 center\_Z 필드가 나타내는 사용자의 위치로부터 예비 렌더링 범위의 경계까지의 반경을 나타낼 수 있다. 즉, 상기 sphere\_radius\_to\_boundary 필드는 예비 렌더링 범위의 반경을 나타낼 수 있다.

[0363] 상기 horizontal\_range 필드는 예비 렌더링 범위의 형태가 탑-셰입(top-shape)인 경우에 전달될 수 있다. 즉,

상기 horizontal\_range 필드는 상기 reserved\_range\_type 필드가 예비 렌더링 범위의 타입이 탑-셰입 타입임을 나타내는 경우에 전달될 수 있다. 상기 horizontal\_range 필드는 예비 렌더링 범위의 타입이 탑-셰입 타입인 경우, 상기 예비 렌더링 범위의 수평 범위를 나타낼 수 있다. 상기 예비 렌더링 범위가 상기 center\_X 필드, 상기 center\_Y 필드 및 상기 center\_Z 필드가 나타내는 사용자의 위치를 중심으로 대칭되는 경우, 상기 horizontal\_range 필드가 정의될 수 있고, 상기 horizontal\_range 필드는 사용자의 위치로부터 수평 방향의 경계 점까지의 거리를 나타낼 수도 있다.

[0364] 또는, 상기 reserved\_range\_type 필드가 예비 렌더링 범위의 타입이 탑-셰입 타입임을 나타내는 경우, 상기 horizontal\_range 필드 대신 horizontal\_range\_left\_end 필드 및 horizontal\_range\_right\_end 필드가 전달될 수도 있다. horizontal\_range\_left\_end 필드는 사용자의 위치의 수평 방향 축에 위치하는 예비 렌더링 범위의 좌측 경계 점을 나타낼 수 있고, horizontal\_range\_right\_end 필드는 사용자의 위치의 수평 방향 축에 위치하는 예비 렌더링 범위의 우측 경계 점을 나타낼 수 있다.

[0365] 상기 vertical\_range 필드는 예비 렌더링 범위의 형태가 탑-셰입(top-shape)인 경우에 전달될 수 있다. 즉, 상기 vertical\_range 필드는 상기 reserved\_range\_type 필드가 예비 렌더링 범위의 타입이 탑-셰입 타입임을 나타내는 경우에 전달될 수 있다. 상기 vertical\_range 필드는 예비 렌더링 범위의 타입이 탑-셰입 타입인 경우, 예비 렌더링 범위의 수직 범위를 나타낼 수 있다. 상기 예비 렌더링 범위가 상기 center\_X 필드, 상기 center\_Y 필드 및 상기 center\_Z 필드가 나타내는 사용자의 위치를 중심으로 대칭되는 경우, 상기 vertical\_range 필드가 정의될 수 있고, 상기 vertical\_range 필드는 사용자의 위치로부터 수직 방향의 경계 점까지의 거리를 나타낼 수도 있다.

[0366] 또는, 상기 reserved\_range\_type 필드가 예비 렌더링 범위의 타입이 탑-셰입 타입임을 나타내는 경우, 상기 vertical\_range 필드 대신 vertical\_range\_top\_end 필드 및 vertical\_range\_bottom\_end 필드가 전달될 수도 있다. vertical\_range\_top\_end 필드는 사용자의 위치의 수직 방향 축에 위치하는 예비 렌더링 범위의 상측 경계 점을 나타낼 수 있고, vertical\_range\_bottom\_end 필드는 사용자의 위치의 수직 방향 축에 위치하는 예비 렌더링 범위의 하측 경계 점을 나타낼 수 있다.

[0367] 상기 distance\_to\_top\_endpoint 필드는 예비 렌더링 범위의 형태가 수직축 방향으로만 설정되는 경우에 전달될 수 있다. 즉, 상기 distance\_to\_top\_endpoint 필드는 상기 reserved\_range\_type 필드가 예비 렌더링 범위의 타입이 엘리베이션 디렉션 타입임을 나타내는 경우에 전달될 수 있다. 상기 distance\_to\_top\_endpoint 필드는 상기 center\_X 필드, 상기 center\_Y 필드 및 상기 center\_Z 필드가 나타내는 사용자의 위치의 수직 방향 축에 위치하는 예비 렌더링 범위의 상측 경계 점을 나타낼 수 있다.

[0368] 상기 distance\_to\_bottom\_endpoint 필드는 예비 렌더링 범위의 형태가 수직축 방향으로만 설정되는 경우에 전달될 수 있다. 즉, 상기 distance\_to\_bottom\_endpoint 필드는 상기 reserved\_range\_type 필드가 예비 렌더링 범위의 타입이 엘리베이션 디렉션 타입임을 나타내는 경우에 전달될 수 있다. 상기 distance\_to\_bottom\_endpoint 필드는 상기 center\_X 필드, 상기 center\_Y 필드 및 상기 center\_Z 필드가 나타내는 사용자의 위치의 수직 방향 축에 위치하는 예비 렌더링 범위의 하측 경계 점을 나타낼 수 있다.

[0369] 상기 distance\_to\_left\_endpoint 필드는 예비 렌더링 범위의 형태가 수평축 방향으로만 설정되는 경우에 전달될 수 있다. 즉, 상기 distance\_to\_left\_endpoint 필드는 상기 reserved\_range\_type 필드가 예비 렌더링 범위의 타입이 레터럴 디렉션 타입임을 나타내는 경우에 전달될 수 있다. 상기 distance\_to\_left\_endpoint 필드는 상기 center\_X 필드, 상기 center\_Y 필드 및 상기 center\_Z 필드가 나타내는 사용자의 위치의 수평 방향 축에 위치하는 예비 렌더링 범위의 좌측 경계 점을 나타낼 수 있다.

[0370] 상기 distance\_to\_right\_endpoint 필드는 예비 렌더링 범위의 형태가 수평축 방향으로만 설정되는 경우에 전달될 수 있다. 즉, 상기 distance\_to\_right\_endpoint 필드는 상기 reserved\_range\_type 필드가 예비 렌더링 범위의 타입이 레터럴 디렉션 타입임을 나타내는 경우에 전달될 수 있다. 상기 distance\_to\_right\_endpoint 필드는 상기 center\_X 필드, 상기 center\_Y 필드 및 상기 center\_Z 필드가 나타내는 사용자의 위치의 수평 방향 축에 위치하는 예비 렌더링 범위의 우측 경계 점을 나타낼 수 있다.

[0371] 상기 distance\_to\_front\_endpoint 필드는 예비 렌더링 범위의 형태가 전후 방향으로만 설정되는 경우에 전달될 수 있다. 즉, 상기 distance\_to\_front\_endpoint 필드는 상기 reserved\_range\_type 필드가 예비 렌더링 범위의 타입이 론지튜디널 디렉션 타입임을 나타내는 경우에 전달될 수 있다. 상기 distance\_to\_front\_endpoint 필드는 상기 center\_X 필드, 상기 center\_Y 필드 및 상기 center\_Z 필드가 나타내는 사용자의 위치의 전후 방향 축에

위치하는 예비 렌더링 범위의 전방 경계 점을 나타낼 수 있다.

- [0372] 상기 distance\_to\_back\_endpoint 필드는 예비 렌더링 범위의 형태가 전후 방향으로만 설정되는 경우에 전달될 수 있다. 즉, 상기 distance\_to\_back\_endpoint 필드는 상기 reserved\_range\_type 필드가 예비 렌더링 범위의 타입이 론지튜디널 디렉션 타입임을 나타내는 경우에 전달될 수 있다. 상기 distance\_to\_back\_endpoint 필드는 상기 center\_X 필드, 상기 center\_Y 필드 및 상기 center\_Z 필드가 나타내는 사용자의 위치의 전후 방향 축에 위치하는 예비 렌더링 범위의 후방 경계 점을 나타낼 수 있다.
- [0373] 상기 vertex\_X[] 필드, 상기 vertex\_Y[] 필드 및 상기 vertex\_Z[] 필드는 예비 렌더링 범위의 형태가 제작자가 의도한 영역 혹은 비디오 혹은 이미지가 포함된 영역에서 전술한 형태를 가지는 예비 렌더링 영역으로 설정되는 경우에 전달될 수 있다. 즉, 상기 vertex\_X[] 필드, 상기 vertex\_Y[] 필드 및 상기 vertex\_Z[] 필드는 상기 reserved\_range\_type 필드가 예비 렌더링 범위의 타입이 user defined 타입임을 나타내는 경우에 전달될 수 있다. 상기 vertex\_X[] 필드, 상기 vertex\_Y[] 필드 및 상기 vertex\_Z[] 필드는 상기 예비 렌더링 범위의 꼭지점(vertex)을 나타낼 수 있다. 이 경우, 상기 vertex\_X[] 필드, 상기 vertex\_Y[] 필드 및 상기 vertex\_Z[] 필드를 기반으로 상기 예비 렌더링 범위의 형태는 vertex 기반 비정형 형태로 도출될 수 있다. 즉, 상기 예비 렌더링 범위는 상기 vertex\_X[] 필드, 상기 vertex\_Y[] 필드 및 상기 vertex\_Z[] 필드가 나타내는 각 위치의 조합을 연결한 형태로 정의될 수 있다.
- [0374] 상기 sphere\_radius 필드는 예비 렌더링 타입이 이동된 사용자의 위치(상기 예비 렌더링 위치)에서 새로운 360도 비디오 혹은 이미지를 전체 3D 공간(즉, 구(sphere) 전체)에 렌더링하는 홀 스피어(whole sphere) 렌더링 타입인 경우에 전달될 수 있다. 즉, 상기 sphere\_radius 필드는 상기 rendering\_type 필드가 예비 렌더링 타입이 홀 스피어(whole sphere) 렌더링 타입임을 나타내는 경우에 전달될 수 있다. 상기 sphere\_radius 필드는 이동 가능한 각 위치, 즉, 예비 렌더링 위치를 중심으로 하는 구의 반경을 나타낼 수 있다. 상기 구는 상기 예비 렌더링 위치에 대한 6DoF 비디오 혹은 이미지가 렌더링되는 3D 공간을 의미할 수 있다.
- [0375] 상기 preliminary\_viewport\_yaw 필드, 상기 preliminary\_viewport\_pitch 필드, 상기 preliminary\_viewport\_roll 필드는 예비 렌더링 타입이 특정 요(yaw), 피치(pitch), 롤(roll)로 지정되는 중점과 수평 범위(horizontal range) 및 수직 범위(vertical range)로 정의되는 파셜 뷰포트(partial viewport)만 렌더링하는 파셜 뷰포트 렌더링 타입인 경우에 전달될 수 있다. 즉, 상기 preliminary\_viewport\_yaw 필드, 상기 preliminary\_viewport\_pitch 필드, 상기 preliminary\_viewport\_roll 필드는 상기 rendering\_type 필드가 예비 렌더링 타입이 파셜 뷰포트(partial viewport) 렌더링 타입임을 나타내는 경우에 전달될 수 있다. 상기 preliminary\_viewport\_yaw 필드, 상기 preliminary\_viewport\_pitch 필드, 상기 preliminary\_viewport\_roll 필드는 상기 예비 렌더링 위치에 대한 6DoF 비디오 혹은 이미지가 렌더링되는 상기 파셜 뷰포트의 중점의 요 값, 피치 값, 롤 값을 나타낼 수 있다.
- [0376] 상기 preliminary\_hor\_range 필드, 상기 preliminary\_ver\_range 필드는 예비 렌더링 타입이 특정 요(yaw), 피치(pitch), 롤(roll)로 지정되는 중점과 수평 범위(horizontal range) 및 수직 범위(vertical range)로 정의되는 파셜 뷰포트(partial viewport)만 렌더링하는 파셜 뷰포트 렌더링 타입인 경우에 전달될 수 있다. 즉, 상기 preliminary\_hor\_range 필드, 상기 preliminary\_ver\_range 필드는 상기 rendering\_type 필드가 예비 렌더링 타입이 파셜 뷰포트(partial viewport) 렌더링 타입임을 나타내는 경우에 전달될 수 있다. 상기 preliminary\_hor\_range 필드, 상기 preliminary\_ver\_range 필드는 상기 예비 렌더링 위치에 대한 6DoF 비디오 혹은 이미지가 렌더링되는 상기 파셜 뷰포트의 가로축 방향으로의 길이 및 세로축 방향으로의 길이를 나타낼 수 있다. 상기 Preliminary\_hor\_range 필드의 값은 0에서 720 사이의 값을 가질 수 있으며, 상기 preliminary\_ver\_range 필드의 값은 0에서 180 사이의 값을 가질 수 있다. 상기 예비 렌더링 위치에 대한 6DoF 비디오 혹은 이미지가 렌더링되는 영역(즉, 파셜 뷰포트)은 상기 preliminary\_viewport\_yaw 필드, 상기 preliminary\_viewport\_pitch 필드, 상기 preliminary\_viewport\_roll 필드, 상기 preliminary\_hor\_range 필드, 상기 preliminary\_ver\_range 필드를 기반으로 설정될 수 있다. 또한, 상기 예비 렌더링 위치에 대한 6DoF 비디오 혹은 이미지가 렌더링되는 영역(즉, 파셜 뷰포트)은 사용자의 위치에서 보여질 뷰포트(viewport)의 크기와 같을 수도 있고 혹은 더 크거나 작을 수도 있다.
- [0377] 한편, 상술한 6DoF 비디오 관련 메타데이터는 SixDOFVideoConfigurationBox에 포함될 수 있다. 이 경우 각 실시예들에 따라 상기 SixDOFVideoConfigurationBox는 ISOBMFF, CFF 등의 일부로 포함될 수 있으며, 더 나아가 ISOBMFF 에 포함되는 경우, 본 발명에서 명시하지 않는 다른 박스에 포함될 수 있다. 이러한 방식으로 6DoF 비디오 데이터와 함께 6DoF 비디오 관련 메타데이터가 저장, 시그널링될 수 있다.



[0378] 일 예로, SixDOFVideoConfigureationBox로 정의된 6DoF 비디오 관련 메타데이터는 ISOBMFF의 'moov' 박스에 포함되는 track header(tkhd) 박스에 다음의 표와 같이 포함될 수 있다.

표 13

```

aligned(8) class TrackHeaderBox extends FullBox('tkhd', version, flags){
if (version==1) {
unsigned int(64)          creation_time;
unsigned int(64)          modification_time;
unsigned int(32)          track_ID;
const unsigned int(32)    reserved = 0;
unsigned int(64)          duration;
} else { // version==0
unsigned int(32)          creation_time;
unsigned int(32)          modification_time;
unsigned int(32)          track_ID;
const unsigned int(32)    reserved = 0;
unsigned int(32)          duration;
}

const unsigned int(32)[2] reserved = 0;
template int(16)          layer = 0;
template int(16)          alternate_group = 0;
template int(16)          volume = {if track_is_audio 0x0100 else 0};
const unsigned int(16)    reserved = 0;
template int(32)[9]       matrix=
{ 0x00010000,0,0,0,0,0x00010000,0,0,0,0x40000000 };// unity matrix
unsigned int(32)          width;
unsigned int(32)          height;
unsigned int(1)          sdv_flag;
unsigned int(7)          reserved=0;
if(sdv_flag == 1){
    SixDOFVideoConfigurationBox sdv_config;
}
}
    
```

[0379]

[0380] 표 13을 참조하면 상기 tkhd 박스에는 sdv\_flag 필드 및/또는 SixDOFVideoConfigureationBox 를 가지는 sdv\_config 필드가 포함될 수 있다.

[0381] sdv\_flag 필드는 해당 비디오 트랙 내에 6DoF 비디오가 포함되어 있는지 여부를 나타내는 플래그일 수 있다. 본 필드의 값이 1 인 경우, 6DoF 비디오 데이터가 해당 비디오 트랙 내에 포함되고, 0 인 경우 그렇지 않을 수 있다. 본 필드의 값에 따라 sdv\_config 필드가 존재할 수 있다.

[0382] sdv\_config 필드는 해당 비디오 트랙 내에 포함된 6DoF 비디오 데이터에 대한 전술한 6DoF 비디오 관련 메타데이터를 포함할 수 있다.

[0383] 본 발명의 다른 실시예에 따르면, SixDOFVideoConfigureationBox 로 정의된 6DoF 비디오 관련 메타데이터는 ISOBMFF의 track 박스에 포함되는 video media header(vmhd) 박스에 포함되어 전달될 수 있다. 여기서, vmhd 박스 전술한 trak 박스의 하위 박스로서, 해당 비디오 트랙에 대한 일반적인(general) 프리젠테이션 관련 정보를 제공할 수 있다. 이 경우, vmhd 박스에는 마찬가지로, sdv\_flag 필드 및/또는 SixDOFVideoConfigureationBox 를 가지는 sdv\_config 필드가 포함될 수 있다. 각 필드의 의미는 전술한 바와 같다.

[0384] 실시예에 따라, 상기 tkhd 박스와 상기 vmhd 박스에 동일한 6DoF 비디오 관련 메타데이터(즉, 상술한 SixDOFVideoConfigureationBox)가 포함될 수도 있다. 이 경우 상기 tkhd 박스에 정의된 6DoF 비디오 관련 메타데이터 값은 상기 vmhd 박스에 정의된 6DoF 비디오 관련 메타데이터 값으로 대체될 수 있다.

[0385] 본 발명의 또 다른 실시예에 따르면, SixDOFVideoConfigureationBox 로 정의된 6DoF 비디오 관련 메타데이터는 비주얼 샘플 그룹 엔트리(Visual Sample Group Entry) 에 포함되어 전달될 수 있다.

[0386] 상기 SixDOFVideoConfigureationBox로 정의된 6DoF 비디오 관련 메타데이터는 다음의 표와 같이 상기 비주얼 샘플 그룹 엔트리에 포함될 수 있다.

표 14

```

class VisualSampleEntry(codingname) extends SampleEntry (codingname){
    unsigned int(16) pre_defined = 0;
    const unsigned int(16) reserved = 0;
    unsigned int(32)[3] pre_defined = 0;
    unsigned int(16) width;
    unsigned int(16) height;
    template unsigned int(32) horziresolution = 0x00480000; // 72 dpi
    template unsigned int(32) vertresolution = 0x00480000; // 72 dpi
    const unsigned int(32) reserved = 0;
    template unsigned int(16) frame_count = 1;
    string[32] compressorname;
    template unsigned int(16) depth = 0x0018;
    int(16) pre_defined = -1;
    // other boxes from derived specifications
    CleanApertureBox clap;// optional
    PixelAspectRatioBox pasp;// optional
    unsigned int(1) sdv_flag;
    unsigned int(7) reserved=0;
    if(sdv_flag == 1){
    SixDOFVideoConfigurationBox sdv_config;
    }
}
    
```

[0387]

[0388] 표 14를 참조하면 상기 비주얼 샘플 그룹 엔트리는 sdv\_flag 필드 및/또는 SixDOFVideoConfigureationBox를 가지는 sdv\_config 필드를 포함할 수 있다. 하나의 파일 혹은 무비 프래그먼트 내에 존재하는 하나 이상의 비디오 샘플들에 동일한 6DoF 비디오 관련 메타데이터가 적용될 수 있는 경우, 6DoF 비디오 관련 메타데이터는 비주얼 샘플 그룹 엔트리에 포함되어 전달될 수 있다.

[0389] 본 발명의 또 다른 실시예에 따르면, 비디오 스트림이 하나 이상의 무비 프래그먼트(movie fragment)로 프래그먼테이션(fragmentation)되어 ISOBMFF 기반으로 저장 및 전달되는 경우, SixDOFVideoConfigureationBox로 정의된 6DoF 비디오 관련 메타데이터는 ISOBMFF의 movie extends(mvex) 박스에 포함되는 track extends(trex) 박스에 다음의 표와 같이 포함될 수 있다.

표 15

```

aligned(8) class TrackExtendsBox extends FullBox('trex', 0, 0){
    unsigned int(32) track_ID;
    unsigned int(32) default_sample_description_index;
    unsigned int(32) default_sample_duration;
    unsigned int(32) default_sample_size;
    unsigned int(32) default_sample_flags;
    unsigned int(1) default_sample_sdv_flag;
    unsigned int(7) reserved=0;
    if(default_sample_sdv_flag == 1){
    SixDOFVideoConfigurationBox default_sample_sdv_config;
    }
}
    
```

[0390]

[0391] 표 15를 참조하면 상기 trex 박스는 default\_sample\_sdv\_flag 필드 및/또는 SixDOFVideoConfigureationBox를 가지는 default\_sample\_sdv\_config 필드를 포함할 수 있다.

[0392] 상기 default\_sample\_sdv\_flag 필드는 무비 프래그먼트(movie fragment)에 포함된 비디오 트랙 프래그먼트(video track fragment)에 6DoF 비디오 샘플들이 포함되어 있는지 여부를 나타낼 수 있다. 상기 default\_sample\_sdv\_flag 필드의 값이 1인 경우, 상기 default\_sample\_sdv\_flag 필드는 무비 프래그먼트들의 비디오 트랙 내에 6DoF 비디오 샘플들이 디폴트로 포함될 수 있음을 나타낼 수 있으며, 이 경우, 해당 트랙 프래그먼트(track fragment)의 비디오 샘플들 각각에 적용될 수 있는 6DoF 비디오 관련 메타데이터를 포함하는 SixDOFVideoConfigurationBox 등을 포함할 수 있다.

[0393] 또한, 상기 default\_sample\_sdv\_config 필드는 무비 프래그먼트(movie fragment)의 트랙 프래그먼트(track

fragment) 내에 포함된 샘플들에 디폴트로 적용될 6DoF 비디오 관련 세부 파라미터들을 포함할 수 있다. 이 경우, 상기 SixDOFVideoConfigurationBox 에 포함된 필드들은 앞서 제안한 내용과 동일한 의미를 나타낼 수 있다.

[0394] 본 발명의 또 다른 실시예에 따르면, 비디오 스트림이 하나 이상의 무비 프래그먼트(movie fragment)로 프래그먼트이션(fragmentation)되어 ISOBMFF 기반으로 저장 및 전달되는 경우, SixDOFVideoConfigureationBox로 정의된 6DoF 비디오 관련 메타데이터는 ISOBMFF의 movie fragment(moof) 박스에 포함되는 track fragment header(tfhd) 박스에 다음의 표와 같이 포함될 수 있다.

표 16

```
aligned(8) class TrackFragmentHeaderBox extends FullBox('tfhd', 0, tf_flags){
    unsigned int(32) track_ID;
    unsigned int(1)  sdv_flag;
    unsigned int(7)  reserved=0;
    if(omv_flag == 1){
        SixDOFVideoConfigurationBox    sdv_config;
    }
    // all the following are optional fields
    unsigned int(64) base_data_offset;
    unsigned int(32) sample_description_index;
    unsigned int(32) default_sample_duration;
    unsigned int(32) default_sample_size;
    unsigned int(32) default_sample_flags;
}
```

[0395] 표 16을 참조하면 상기 tfhd 박스는 sdv\_flag 필드 및/또는 SixDOFVideoConfigureationBox를 가지는 sdv\_config 필드를 포함할 수 있다.

[0396] 표 16을 참조하면 상기 tfhd 박스는 sdv\_flag 필드 및/또는 SixDOFVideoConfigureationBox를 가지는 sdv\_config 필드를 포함할 수 있다.

[0397] 상기 sdv\_flag 필드는 무비 프래그먼트(movie fragment)가 포함하는 트랙 프래그먼트(fragment)에 6DoF 비디오가 포함되어 있는지 여부를 나타낼 수 있다. 상기 sdv\_flag 필드의 값이 1인 경우, 상기 sdv\_flag 필드는 해당 트랙 프래그먼트 내에 6DoF 비디오가 포함되어 있음을 나타낼 수 있고, 상기 sdv\_flag 필드의 값이 0인 경우, 상기 sdv\_flag 필드는 해당 트랙 프래그먼트 내에 6DoF 비디오가 포함되어 있지 않음을 나타낼 수 있다.

[0398] 또한, 상기 sdv\_config 필드는 무비 프래그먼트의 트랙 프래그먼트 내에 포함된 6DoF 비디오 관련 세부 파라미터들을 포함할 수 있다. 이 경우, 상기 SixDOFVideoConfigurationBox 에 포함된 필드들은 앞서 제안한 내용과 동일한 의미를 나타낼 수 있다.

[0399] 한편, 본 발명의 또 다른 실시예에 따르면, sdv\_flag 필드가 생략되고, sdv\_config 필드 대신 default\_sample\_sdv\_config 필드가 포함될 수 있다.

[0400] 이 경우, tr\_flags 필드에 의해 6DoF 비디오 관련 메타데이터가 tfhd 박스에 포함되는지가 지시될 수 있다. 구체적으로, 상기 tr\_flags 필드는 다음의 표와 같이 6DoF 비디오 관련 메타데이터가 tfhd 박스에 포함되는지 여부를 나타낼 수 있다.

표 17

tf_flags	Description
0x000001	base-data-offset-present;
0x000002	sample-description-index-present;
0x000008	default-sample-duration-present
0x000010	default-sample-size-present
0x000020	default-sample-flags-present
0x010000	duration-is-empty.
0x020000	default-base-is-moof;
0x400000	default-sample-omv-configuration-present;
<b>0x800000</b>	<b>default-sample-sdv-configuration-present;</b>

[0401] 표 17을 참조하면 상기 tr\_flags 필드가 0x800000을 포함하는 경우, 6DoF 비디오 관련 메타데이터는 tfhd 박스에 포함되어 전달될 수 있다. 즉, 상기 tr\_flags 필드가 0x800000을 포함하는 경우, 무비 프래그먼트(movie fragment) 내 트랙 프래그먼트(track fragment)에 포함된 비디오 샘플과 연관된 6DoF 비디오 관련 메타데이터의

디폴트값이 존재함을 나타낼 수 있다. 이 경우, 하기와 같이 비디오 샘플에 적용되는 디폴트 6DoF 비디오 관련 메타데이터를 저장 및 시그널링 할 수 있다. 한편, 6DoF 비디오 관련 메타데이터가 tfhd 박스에 포함됨을 나타내는 상기 tr\_flags 필드의 값은 0x800000 이외의 값이 할당될 수도 있다.

[0403] 상기 sdv\_flag 필드가 생략되고, 상기 sdv\_config 필드 대신 default\_sample\_sdv\_config 필드가 포함되는 상기 tfhd 박스는 다음의 표와 같이 도출될 수 있다.

표 18

```
aligned(8) class TrackFragmentHeaderBox extends FullBox('tfhd', 0, tf_flags){
    unsigned int(32) track_ID;
    // all the following are optional fields
    unsigned int(64) base_data_offset;
    unsigned int(32) sample_description_index;
    unsigned int(32) default_sample_duration;
    unsigned int(32) default_sample_size;
    unsigned int(32) default_sample_flags;
    SixDOFVideoConfigurationBox default_sample_sdv_config;
}
```

[0404]

[0405] 상기 default\_sample\_sdv\_config 필드는 무비 프래그먼트(movie fragment)의 트랙 프래그먼트(track fragment) 내에 포함된 샘플들에 디폴트로 적용될 6DoF 비디오 관련 세부 파라미터들을 포함할 수 있다. 이 경우, 상기 SixDOFVideoConfigurationBox 에 포함된 필드들은 앞서 제안한 내용과 동일한 의미를 나타낼 수 있다.

[0406] 본 발명의 또 다른 실시예에 따르면, 하나 이상의 무비 프래그먼트(movie fragment) 내 트랙 프래그먼트 박스(track fragment box)에 트랙 런 박스(track run box)가 존재하는 경우, 트랙 프래그먼트(track fragment)에 포함된 비디오 샘플들에 공통적으로 적용될 수 있는 6DoF 비디오 관련 메타데이터를 하기와 같이 저장 및 시그널링 할 수 있다. 상기 6DoF 비디오 관련 메타데이터는 앞서 tfhd 박스에 제안한 SixDOFVideoConfigurationBox 의 필드들에 대한 의미와 같을 수 있다.

[0407] 하나 이상의 무비 프래그먼트(movie fragment) 내 트랙 프래그먼트 박스(track fragment box)에 트랙 런 박스(track run box)가 존재하는 경우, tr\_flags 필드에 의해 6DoF 비디오 관련 메타데이터가 trun 박스에 포함되는 지가 지시될 수 있다. 구체적으로, 상기 tr\_flags 필드는 다음의 표와 같이 6DoF 비디오 관련 메타데이터가 trun 박스에 포함되는지 여부를 나타낼 수 있다.

표 19

tr_flags	Description
0x000001	data-offset-present.
0x000004	first-sample-flags-present;
0x000100	sample-duration-present:
0x000200	sample-size-present:
0x000400	sample-flags-present;
0x000800	sample-composition-time-offsets-present;
0x004000	sample-omv-configuration-present;
0x008000	omv-configuration-present;
<b>0x010000</b>	<b>sample-sdv-configuration-present;</b>
<b>0x020000</b>	<b>sdv-configuration-present;</b>

[0408]

[0409] 표 19를 참조하면 상기 tr\_flags 필드의 값이 0x020000 인 경우, 트랙 프래그먼트에 포함된 비디오 샘플들에 공통적으로 적용될 수 있는 6DoF 비디오 관련 메타데이터가 존재함을 나타낼 수 있다. 또한, 상기 tr\_flags 필드의 값이 0x010000 인 경우, 트랙 프래그먼트 내 비디오 샘플 각각에 적용되는 6DoF 비디오 관련 메타데이터가 존재함을 나타낼 수 있다. 한편, 6DoF 비디오 관련 메타데이터가 trun 박스에 포함됨을 나타내는 상기 tr\_flags 필드의 값은 상술한 값 이외의 값이 할당될 수도 있다.

[0410] 상기 tr\_flags 필드의 값이 0x020000 인 경우, 트랙 프래그먼트에 포함된 비디오 샘플들에 공통적으로 적용되는 SixDOFVideoConfigurationBox 에 대한 상기 sdv\_config 필드는 다음의 표와 같이 상기 trun 박스에 포함될 수

있다.

표 20

```

aligned(8) class TrackFragmentHeaderBox extends FullBox('trun', version,
tr_flags) {
    unsigned int(32) sample_count; // the following are
optional fields
    signed int(32) data_offset;
    unsigned int(32) first_sample_flags;
SixDOFVideoConfigurationBox sdv_config;
    // all fields in the following array are optional
    {
    unsigned int(32) sample_duration;
    unsigned int(32) sample_size;
    unsigned int(32) sample_flags;
    if (version == 0)
    {
    unsigned int(32) sample_composition_time_offset;
    }else {
    signed int(32) sample_composition_time_offset;
    }
    }[ sample_count ]
}
    
```

[0411]

[0412] 상기 tr\_flags 필드의 값이 0x010000 인 경우, 트랙 프래그먼트 내 비디오 샘플 각각에 적용되는 SixDOFVideoConfigurationBox 에 대한 상기 sdv\_config 필드는 다음의 표와 같이 상기 trun 박스에 포함될 수 있다.

표 21

```

aligned(8) class TrackFragmentHeaderBox extends FullBox('trun', version,
tr_flags) {
    unsigned int(32) sample_count; // the following are
optional fields
    signed int(32) data_offset;
    unsigned int(32) first_sample_flags;
    // all fields in the following array are optional
    {
    unsigned int(32) sample_duration;
    unsigned int(32) sample_size;
    unsigned int(32) sample_flags;
    if (version == 0)
    {
    unsigned int(32) sample_composition_time_offset;
    }else {
    signed int(32) sample_composition_time_offset;
    }
    }
SixDOFVideoConfigurationBox sdv_config;
    }[ sample_count ]
}
    
```

[0413]

[0414] 한편, 본 발명의 또 다른 실시예에 따르면, 하나의 파일 혹은 무비 프래그먼트(movie fragment) 내에 존재하는 하나 이상의 비디오 샘플들에 동일한 6DoF 비디오 관련 메타데이터가 적용되는 경우, SixDOFVideoConfigreationBox로 정의된 6DoF 비디오 관련 메타데이터는 다음의 표와 같이 비주얼 샘플 그룹 엔트리(visual sample group entry) 등에 추가될 수 있다.

표 22

```

class VisualSampleGroupEntry (unsigned int(32) grouping_type) extends SampleGroupEntry
(grouping_type)
  unsigned int(1)sdv_flag;
  unsigned int(7) reserved=0;
  if(omv_flag == 1){
    SixDOFVideoConfigurationBox      sdv_config;
  }
}
    
```

- [0415]
- [0416] 표 22를 참조하면 상기 비주얼 샘플 그룹 엔트리는 sdv\_flag 필드 및/또는 SixDOFVideoConfigureationBox를 가지는 sdv\_config 필드를 포함할 수 있다.
- [0417] 상기 sdv\_flag 필드는 샘플 그룹이 6DOF 비디오 샘플 그룹인지 아닌지를 나타낼 수 있다. 예를 들어, 상기 sdv\_flag 필드의 값이 1인 경우, 해당 샘플 그룹이 6DoF 비디오 샘플 그룹임을 나타낼 수 있고, 상기 sdv\_flag 필드의 값이 0인 경우, 해당 샘플 그룹이 6DoF 비디오 샘플 그룹이 아님을 나타낼 수 있다.
- [0418] 상기 sdv\_config 필드는 상기 샘플 그룹 내에 포함되어 있는 6DoF 비디오 샘플들에 공통적으로 적용될 수 있는 6DoF 비디오에 대한 메타 데이터를 포함할 수 있다. 예를 들어, 상술한 내용과 같이 상기 SixDOFVideoConfigurationBox 의 initial\_view\_yaw\_degree 필드, initial\_view\_pitch\_degree 필드, initial\_view\_roll\_degree 필드, initial\_center\_FB\_delta 필드, initial\_center\_UD\_delta 필드, initial\_center\_RL\_delta 필드를 이용하여 각 샘플 그룹과 연관된 비디오에 대한 초기 시점(initial view)와 초기 시점의 초기 센터(initial center), 즉, 상기 초기 시점의 위치(센터 포지션)이 설정될 수 있다.
- [0419] 한편, 본 발명의 또 다른 실시예에 따르면, 하나의 비디오 프레임(frame)이 하나 이상의 리전(region)으로 나뉘어져 코딩되고, 하나 이상의 트랙을 통하여 해당 리전 데이터들이 저장되는 경우, 각 트랙과 연관된 박스(box)에 앞서 제안된 SixDOFVideoConfigurationBox 가 포함될 수 있다. 하지만, 트랙들 중 특정 트랙에만 상기 SixDOFVideoConfigurationBox 가 포함되고, 나머지 트랙들에는 포함되지 않는 경우, 상기 SixDOFVideoConfigurationBox 을 포함하지 않는 트랙들의 TrackReferenceTypeBox 에 'sdvb' track reference type 으로서, SixDOFVideoConfigurationBox 을 포함한 트랙을 가리키는 필드가 포함될 수 있다. 여기서, 'sdvb' track reference type은 하나의 6DoF 비디오를 구성하는 리전 들의 데이터를 포함하는 트랙들이 디코딩 시 기본적인 베이스 레이어(base layer)를 track\_IDs 필드를 통하여 지칭하는 트랙을 지칭할 수 있다. 즉, track\_IDs 필드는 SixDOFVideoConfigurationBox 을 포함한 트랙을 가리킬 수 있다.
- [0420] 상기 track\_IDs 필드를 포함하는 TrackReferenceTypeBox 는 다음의 표와 같이 나타낼 수 있다.

표 23

```

aligned(8) class TrackReferenceBox extends Box('tref'){
}
aligned(8) class TrackReferenceTypeBox (unsigned int(32) reference_type) extends
Box(reference_type){
    unsigned int(32) track_IDs[];
}
    
```

- [0421]
- [0422] 한편, TrackReferenceTypeBox 에 포함된 track\_IDs 필드가 가리키는 트랙 레퍼런스(track reference)의 타입은 다음의 표와 같이 나타낼 수 있다.

표 24

Track Reference	Description
'hint'	The referenced track(s) contain the original media for this hint track
'cdsc'	This track describes the referenced track
'font'	This track uses fonts carried/defined in the referenced track
'hind'	This track depends on the referenced hint track i.e., it should only be used if the referenced hint track is used
'vdep'	This track contains auxiliary depth video information for the referenced video track
'vplx'	This track contains auxiliary parallax video information for the referenced video track
'subt'	This track contains subtitle, timed text or overlay graphical information for the referenced track or any track in the alternate group to which the track belongs, if any.
'sdvb'	<b>This track contains information related to Six DoF contents for the referenced track or any track in the alternate group to which the track belongs, if any</b>

[0423]

[0424] 표 24를 참조하면 상기 track\_IDs 필드가 가리키는 트랙 레퍼런스의 타입이 'hint' 타입인 경우, 상기 track\_id 필드가 지시하는 트랙은 해당 트랙의 원본 미디어(original media)를 포함할 수 있다. 상기 해당 트랙은 힌트(hint) 트랙이라고 불릴 수 있다.

[0425] 또한, 상기 track\_IDs 필드가 가리키는 트랙 레퍼런스의 타입이 'cdsc' 타입인 경우, 해당 트랙은 상기 track\_id 필드가 지시하는 트랙을 기술(describe)할 수 있다.

[0426] 또한, 상기 track\_IDs 필드가 가리키는 트랙 레퍼런스의 타입이 'font' 타입인 경우, 해당 트랙은 상기 track\_id 필드가 지시하는 트랙을 통하여 전송된 폰트(font) 또는 상기 track\_id 필드가 지시하는 트랙에서 정의된 폰트를 사용할 수 있다.

[0427] 또한, 상기 track\_IDs 필드가 가리키는 트랙 레퍼런스의 타입이 'hind' 타입인 경우, 해당 트랙은 상기 힌트 트랙을 참조할 수 있다. 따라서, 상기 힌트 트랙이 사용되는 경우에만 상기 'hind' 타입을 나타낼 수 있다.

[0428] 또한, 상기 track\_IDs 필드가 가리키는 트랙 레퍼런스의 타입이 'vdep' 타입인 경우, 해당 트랙은 TrackReferenceTypeBox의 track\_IDs 필드가 지시하는 트랙을 위한 보조 뎀스 비디오 정보(auxiliary depth video information)를 포함하고 있음이 지시될 수 있다.

[0429] 또한, 상기 track\_IDs 필드가 가리키는 트랙 레퍼런스의 타입이 'vplx' 타입인 경우, 해당 트랙은 TrackReferenceTypeBox의 track\_IDs 필드가 지시하는 트랙을 위한 보조 시차 비디오 정보(auxiliary parallax video information)를 포함하고 있음이 지시될 수 있다.

[0430] 또한, 상기 track\_IDs 필드가 가리키는 트랙 레퍼런스의 타입이 'subt' 타입인 경우, 해당 트랙은 TrackReferenceTypeBox의 track\_IDs 필드가 지시하는 트랙을 위한 subtitle, timed text, overlay graphical information 을 포함하고 있음이 지시될 수 있다. 또는, 해당 트랙은 해당 트랙이 포함된 대체 그룹(alternate group)의 모든 트랙을 위한 subtitle, timed text, overlay graphical information 을 포함하고 있음이 지시될 수 있다.

[0431] 또한, 상기 track\_IDs 필드가 가리키는 트랙 레퍼런스의 타입이 'sdvb' 타입인 경우, 상기 track\_IDs 필드는 진술한 6DoF 비디오 관련 메타데이터를 운반하고 있는 특정 트랙을 가리킬 수 있다. 구체적으로, 각각의 리전들을 포함하는 각 트랙들은 디코딩될 때, 6DoF 비디오 관련 메타데이터 중 기본적인 베이스 레이어 정보가 필요할 수 있다. 상기 track\_IDs 필드는 그 베이스 레이어 정보를 운반하고 있는 특정 트랙을 지시할 수 있다. 상기 베이스 레이어 정보는 베이식 정보(basic information)이라고 불릴 수 있다.

[0432] 6DoF 비디오 서비스에서 사용자는 사용자의 시점(view)에 따라 비디오 중 일부 영역 만을 보게 되기 때문에 앞서 제안한 메타데이터들은 사용자 시점과 위치에 따른 디코딩(decoding) 혹은 렌더링(rendering)을 지원하는데 사용될 수 있다.

- [0433] 또한, 전송한 6DoF 비디오에 대한 비디오 트랙 또는 샘플 등과 관련된 메타데이터에 존재하는 모든 정보들은 HEVC/AVC 의 SEI 혹은 DASH MPD 내에 정의될 수도 있다.
- [0434] 도 21은 6DoF 비디오 수신 장치에서 카메라 정보(camera information) 메타데이터와 스티칭 메타데이터가 사용되는 동작을 나타낸다.
- [0435] 도 21의 (a)를 참조하면 6DoF 비디오 수신 장치는 주요 시점(primary view, PV) 및 부차 시점(secondary view, SV) 각각에 대한 비디오 혹은 이미지를 디코딩할 수 있다. 또는, 시점 분류와 상관없이 비디오 혹은 이미지에 대한 디코딩이 수행될 수 있다.
- [0436] 또한, 6DoF 비디오 수신 장치는 이미지 합성 및 렌더링(Image composition, and rendering) 과정을 수행할 수 있다. 상기 과정은 전송한 바와 같이 사용자의 위치를 중심으로 한 영상을 렌더링 하기 위한 과정으로 6DoF 비디오 수신 장치는 사용자의 위치 및 시선에 따라 디코딩된 6DoF 비디오 데이터를 이용하거나 가상 시점 생성/합성으로 만들어진 사용자 주변의 비디오 및 이미지를 렌더링할 수 있다. 다음으로, 6DoF 비디오 수신 장치는 렌더링된 비디오를 디스플레이할 수 있다. 또한, 전송한 바와 같이 헤드, 포지션, 시선 트래킹(head/position/eye tracking)을 통하여 사용자의 헤드, 포지션, 게이즈, 뷰포트 정보 등이 획득, 처리될 수 있고, 상기 사용자의 헤드, 포지션, 게이즈, 뷰포트 정보 등을 기반으로 상기 사용자의 사용자의 위치 및 시선에 따른 디코딩된 6DoF 비디오 데이터가 이미지 합성, 렌더링 및 디스플레이될 수 있다.
- [0437] 한편, 도 21의 (b)에 도시된 것과 같이 디코딩된 영상은 시점(viewpoint)에 따라 스티칭(Stitching)되는 영상으로 쓰이는 캡처드 씬(captured scene)이 달라질 수 있고, 이 경우 6DoF 비디오 수신 장치는 상기 시점의 선정 후 스티칭할 영상의 수와 이때 쓰여질 캡처드 씬 수 및 캡처 영상의 카메라 배열 타입(camera array type), 카메라 위치(camera position) 정보에 해당하는 메타데이터를 이용하여 적합한 영상을 선정할 수 있다. 만약 스티칭할 영상 즉, 캡처드 씬이 있는 경우, 상기 캡처드 씬이 스티칭될 3D 공간을 생성하는 구 생성(sphere generation) 과정이 수행되는데, 상기 구 생성은 프로젝션 스킴(projection scheme) 정보와 구 반경 사이즈(sphere radius size) 정보 등이 메타데이터로서 사용될 수 있다. 6DoF 비디오 수신 장치는 상기 선정된 시점에 대한 영상을 합성하여(image composition) 디스플레이(display)할 수 있다. 한편, 해당 시점에 스티칭할 캡처드 씬이 없는 경우, 6DoF 비디오 수신 장치는 상기 시점의 주변 시점들에 대한 영상들을 기반으로 가상 시점 합성(virtual view synthesis)한 후에 디스플레이할 수 있다.
- [0438] 이를 통하여 사용자의 위치 이동에 따른 6DoF 콘텐츠를 제공할 수 있는바, 사용자에게 몰입형 미디어/실감미디어(Immersive media)의 경험을 제공할 수 있다.
- [0439] 또한, 상기 해당 시점에 대한 스티칭 정보를 시그널링하여, 상기 해당 시점에 대한 캡처된 6DoF 비디오가 없는 경우에도 상기 해당 시점의 주변 위치의 6DoF 비디오를 기반으로 상기 특정 위치에 대한 6DoF 비디오를 생성 및 사용자에게 제공할 수 있는바, 6DoF 콘텐츠를 보다 효율적으로 제공할 수 있다.
- [0440] 도 22는 6DoF 비디오 수신 장치에서 리-프로젝션 정보 메타데이터가 사용되는 동작을 나타낸다. 상기 리-프로젝션 정보 메타데이터는 렌더링 정보 메타데이터라고 나타낼 수도 있다. 상기 과정은 전송한 바와 같이 사용자의 위치를 중심으로 한 영상을 렌더링 하기 위한 과정으로 6DoF 비디오 수신 장치는 사용자의 위치 및 시선에 따라 디코딩된 6DoF 비디오 데이터를 이용하거나 가상 시점 생성/합성으로 만들어진 사용자 주변의 비디오 및 이미지를 렌더링할 수 있다. 다음으로, 6DoF 비디오 수신 장치는 렌더링된 비디오를 디스플레이할 수 있다. 또한, 전송한 바와 같이 헤드, 포지션, 시선 트래킹(head/position/eye tracking)을 통하여 사용자의 헤드, 포지션, 게이즈, 뷰포트 정보 등이 획득, 처리될 수 있고, 상기 사용자의 헤드, 포지션, 게이즈, 뷰포트 정보 등을 기반으로 상기 사용자의 사용자의 위치 및 시선에 따른 디코딩된 6DoF 비디오 데이터가 이미지 합성, 렌더링 및 디스플레이될 수 있다.
- [0441] 한편, 도 22의 (b)에 도시된 것과 같이 디코딩은 PV에 대한 6DoF 비디오와 SV에 대한 6DoF 비디오로 나뉘어 진행될 수 있다. 이 경우, 6DoF 비디오 수신 장치는 PV 디코딩을 먼저 수행하고, 예비 렌더링 범위에 대한 정보(Reserved\_range\_Type\_info)와 렌더링 타입에 대한 정보의 메타데이터를 기반으로 사용자의 위치를 맵핑하여 예비 렌더링 범위 상에 상기 사용자의 위치가 포함되는지 판단할 수 있다.
- [0442] 상기 사용자의 위치가 상기 예비 렌더링 범위 상에 존재하는 경우, 6DoF 비디오 수신 장치는 상기 시점에 대한 영상을 합성하여(image composition) 디스플레이(display)할 수 있다. 상기 사용자의 위치가 상기 예비 렌더링 범위 상에 위치하지 않고 SV에 해당하는 시점(viewpoint)에 위치하는 경우, 6DoF 비디오 수신 장치는 SV 디코딩을 수행하고 상기 시점에 대한 영상을 합성하여(image composition) 디스플레이(display)할 수 있다. 한편, 상



기 사용자의 위치가 PV와 SV에 해당하는 시점이 아닌 곳에 위치하는 경우, PV 디코딩과 SV 디코딩을 모두 수행하고, 상기 해당하는 시점에 대한 가상 뷰를 상기 디코딩된 PV 및 SV의 6DoF 영상을 합성하여 생성하고, 상기 해당하는 시점에 대한 영상을 디스플레이(display)할 수 있다.

- [0443] 상술한 본 발명에 따르면 6DoF 콘텐츠에서 제공하는 이동 방향 또는 예측되는 사용자의 위치 이동 방향에 따른 예비 렌더링 정보를 시그널링하여, 사용자에게 제공되어야 하는 6DoF 콘텐츠가 우선적으로 전송 및 렌더링될 수 있는바, 6DoF 콘텐츠를 보다 효율적으로 제공할 수 있다.
- [0444] 도 23은 본 발명에 따른 6DoF 비디오 전송 장치에 의한 6DoF 비디오 데이터 처리 방법을 개략적으로 나타낸다. 도 23에서 개시된 방법은 도 14에서 개시된 6DoF 비디오 전송 장치에 의하여 수행될 수 있다. 구체적으로 예를 들어, 도 23의 S2300은 상기 6DoF 비디오 전송 장치의 데이터 입력부에 의하여 수행될 수 있고, S2310은 상기 6DoF 비디오 전송 장치의 프로젝션 처리부에 의하여 수행될 수 있고, S2320은 상기 6DoF 비디오 전송 장치의 메타데이터 처리부에 의하여 수행될 수 있고, S2330은 상기 6DoF 비디오 전송 장치의 데이터 인코더에 의하여 수행될 수 있고, S2340은 상기 6DoF 비디오 전송 장치의 전송 처리부에 의하여 수행될 수 있다. 상기 전송 처리부는 전송부에 포함될 수 있다.
- [0445] 6DoF 비디오 전송 장치는 적어도 하나의 카메라에 의해 복수의 뷰들에 대한 6DoF 비디오를 획득한다(S2300). 6DoF 비디오 전송 장치는 적어도 하나의 카메라에 의해 캡처된 복수의 뷰들에 대한 6DoF 비디오를 획득할 수 있다. 또한, 6DoF 비디오 전송 장치는 적어도 하나의 카메라에 의해 복수의 뷰들에 대한 6DoF 비디오를 획득할 수 있다.
- [0446] 6DoF 비디오 전송 장치는 상기 6DoF 비디오를 처리하여 상기 뷰들에 대한 픽처들을 획득한다(S2310). 6DoF 비디오 전송 장치는 여러 프로젝션 스킴(projection scheme)들 중 상기 6DoF 비디오에 대한 프로젝션 스킴에 따라 2D 이미지에 프로젝션을 수행할 수 있고, 프로젝션된 픽처를 획득할 수 있다. 상기 여러 프로젝션 스킴들은 등정방형 프로젝션(Equirectangular Projection) 스킴, 실린더형 프로젝션 스킴, 큐브 프로젝션 스킴 및 파노라믹 프로젝션 스킴을 포함할 수 있다. 또한, 상기 프로젝션 스킴들은 사용자가 정의한 임의의 유저 디파인드(User defined) 프로젝션 스킴을 포함할 수 있다. 또한, 상기 프로젝션 스킴들은 상술한 프로젝션 스킴들 이외의 프로젝션 스킴을 더 포함할 수도 있다. 상기 프로젝션된 픽처는 프로젝션 스킴의 3D 프로젝션 구조의 면들을 나타내는 리전들을 포함할 수 있다.
- [0447] 또한, 6DoF 비디오 전송 장치는 상기 프로젝션된 픽처의 리전들 각각을 회전, 재배열하거나, 각 리전의 레졸루션을 변경하는 등의 처리를 수행할 수 있다. 상기 처리 과정은 상기 리전별 패킹 과정이라고 불릴 수 있다.
- [0448] 6DoF 비디오 전송 장치는 상기 프로젝션된 픽처에 리전별 패킹 과정을 적용하지 않을 수 있다. 또는, 6DoF 비디오 전송 장치는 상기 프로젝션된 픽처에 리전별 패킹 과정을 적용할 수 있고, 상기 리전별 패킹 과정이 적용된 리전을 포함하는 상기 패킹된 픽처를 획득할 수 있다.
- [0449] 또한, 6DoF 비디오 전송 장치는 특정 뷰에 대한 스티칭 과정을 수행하여 상기 특정 뷰에 대한 상기 픽처들을 도출할 수 있다. 예를 들어, 6DoF 비디오 전송 장치는 적어도 하나의 스티칭 대상 뷰를 도출할 수 있고, 상기 스티칭 대상 뷰에 대한 캡처드 씬(captured scene)들을 도출할 수 있다. 6DoF 비디오 전송 장치는 상기 캡처드 씬들을 스티칭하고, 스티칭된 상기 캡처드 씬들을 처리하여 상기 스티칭 대상 뷰에 대한 픽처를 도출할 수 있고, 상기 카메라 정보 및 상기 스티칭 정보를 생성할 수 있다.
- [0450] 한편, 6DoF 비디오 전송 장치는 사용자의 위치에 대한 예비 렌더링 범위를 도출할 수도 있다. 이 경우, 6DoF 비디오 전송 장치는 상기 예비 렌더링 범위에 대한 렌더링 정보를 생성할 수 있다.
- [0451] 6DoF 비디오 전송 장치는 상기 6DoF 비디오에 대한 메타데이터를 생성한다(S2320). 상기 메타데이터는 상술한 camera\_info\_flag 필드, num\_view 필드, Camera\_capture\_array\_row 필드, Camera\_capture\_array\_col 필드, Baseline\_length 필드, Master\_camera\_ID 필드, camera\_center\_pitch 필드, camera\_center\_yaw 필드, camera\_center\_roll 필드, camera\_center\_FB 필드, camera\_center\_UD 필드, camera\_center\_RL 필드, num\_view\_center 필드, position\_center\_X 필드, position\_center\_Y 필드, position\_center\_Z 필드, num\_camera 필드, stitching\_radius 필드, position\_center\_yaw 필드, position\_center\_roll 필드, position\_center\_pitch 필드, position\_center\_distance 필드, initial\_view\_yaw\_degree 필드, initial\_view\_pitch\_degree 필드, initial\_view\_roll\_degree 필드, initial\_center\_FB\_delta 필드, initial\_center\_UD\_delta 필드, initial\_center\_RL\_delta 필드, content\_fov\_flag 필드, capture\_info\_flag 필드, region\_info\_flag 필드, content\_hfov 필드, content\_vfov 필드, reserved\_rendering\_range\_flag 필드,

rendering\_type 필드, reserved\_range\_type 필드, center\_X 필드, center\_Y 필드, center\_Z 필드, step size 필드, sphere\_radius\_to\_boundary 필드, horizontal\_range 필드, vertical\_range 필드, distance\_to\_top\_endpoint 필드, distance\_to\_bottom\_endpoint 필드, distance\_to\_left\_endpoint 필드, distance\_to\_right\_endpoint 필드, distance\_to\_front\_endpoint 필드, distance\_to\_back\_endpoint 필드, vertex\_X[] 필드, vertex\_Y[] 필드, vertex\_Z[] 필드, sphere\_radius 필드, preliminary\_viewport\_yaw 필드, preliminary\_viewport\_pitch 필드, preliminary\_viewport\_roll 필드, preliminary\_hor\_range 필드 및/또는 preliminary\_ver\_range 필드를 포함할 수 있다. 상기 필드들의 의미는 상술한 바와 같다.

- [0452] 구체적으로, 일 예로, 상기 메타데이터는 6DoF 비디오를 캡처한 카메라 정보 및 스티칭(stitching) 정보를 포함할 수 있다.
- [0453] 상기 카메라 정보는 원점이 되는 뷰에 대한 카메라 ID를 나타내는 정보, 카메라 배열 정보, 상기 카메라로 캡처된 뷰들의 수를 나타내는 정보, 카메라 배열의 열을 나타내는 정보, 카메라 배열의 행을 나타내는 정보, 카메라 간의 거리를 나타내는 정보 및/또는 상기 뷰들의 위치 정보를 포함할 수 있다.
- [0454] 상기 원점이 되는 뷰에 대한 카메라 ID를 나타내는 정보는 상기 Master\_camera\_ID 필드를 나타낼 수 있고, 상기 카메라 배열 정보는 상기 Capture\_type 필드를 나타낼 수 있고, 상기 카메라로 캡처된 뷰들의 수를 나타내는 정보는 상기 num\_view 필드를 나타낼 수 있고, 카메라 배열의 열을 나타내는 정보는 상기 Camera\_capture\_array\_row 필드를 나타낼 수 있고, 카메라 배열의 행을 나타내는 정보는 상기 Camera\_capture\_array\_col 필드를 나타낼 수 있고, 카메라 간의 거리를 나타내는 정보는 Baseline\_length 필드를 나타낼 수 있다.
- [0455] 상기 뷰들의 위치 정보는 해당 뷰에 대한 6DoF 비디오의 중점의 yaw 값을 나타내는 정보, 해당 뷰에 대한 6DoF 비디오의 중점의 pitch 값을 나타내는 정보, 해당 뷰에 대한 6DoF 비디오의 중점의 roll 값을 나타내는 정보, 해당 뷰 위치의 앞뒤 방향 축에 대한 성분을 나타내는 정보, 해당 뷰 위치의 위아래 방향 축에 대한 성분을 나타내는 정보 및/또는 해당 뷰 위치의 좌우 방향 축에 대한 성분을 나타내는 정보를 포함할 수 있다. 여기서, 상기 앞뒤 방향 축은 y축, 상기 좌우 방향 축은 x축, 상기 위아래 방향 축은 z축을 나타낼 수 있다. 상기 해당 뷰에 대한 6DoF 비디오의 중점의 yaw 값을 나타내는 정보는 상기 camera\_center\_yaw 필드를 나타낼 수 있고, 상기 해당 뷰에 대한 6DoF 비디오의 중점의 pitch 값을 나타내는 정보는 상기 camera\_center\_pitch 필드를 나타낼 수 있고, 상기 해당 뷰에 대한 6DoF 비디오의 중점의 roll 값을 나타내는 정보는 상기 camera\_center\_roll 필드를 나타낼 수 있고, 상기 해당 뷰 위치의 앞뒤 방향 축에 대한 성분을 나타내는 정보는 상기 camera\_center\_FB 필드를 나타낼 수 있고, 상기 해당 뷰 위치의 위아래 방향 축에 대한 성분을 나타내는 정보는 상기 camera\_center\_UD 필드를 나타낼 수 있고, 상기 해당 뷰 위치의 좌우 방향 축에 대한 성분을 나타내는 정보는 상기 camera\_center\_RL 필드를 나타낼 수 있다.
- [0456] 상기 카메라 배열 정보는 상기 6DoF 비디오를 캡처하는데 사용된 카메라 배열(camera array) 타입을 나타낼 수 있다. 여기서, 상기 카메라 배열 타입은 HDCA (High Density Camera Array), LDCA(Low Density Camera Array), 마이크로렌즈(microlens), 전방향 카메라(Omnidirectional Camera), 이동 카메라(Movement camera), 사용자 정의(user defined) 배열 중 하나일 수 있다.
- [0457] 구체적으로, 예를 들어, 상기 카메라 배열 정보의 값이 0인 경우, 상기 카메라 배열 정보는 상기 HDCA를 상기 6DoF 비디오의 카메라 배열 타입으로 나타낼 수 있다. 또한, 상기 카메라 배열 정보의 값이 1인 경우, 상기 카메라 배열 정보는 상기 LDCA를 상기 6DoF 비디오의 카메라 배열 타입으로 나타낼 수 있다. 또한, 상기 카메라 배열 정보의 값이 2인 경우, 상기 카메라 배열 정보는 상기 마이크로렌즈를 상기 6DoF 비디오의 카메라 배열 타입으로 나타낼 수 있다. 또한, 상기 카메라 배열 정보의 값이 3인 경우, 상기 카메라 배열 정보는 상기 전방향 카메라를 상기 6DoF 비디오의 카메라 배열 타입으로 나타낼 수 있다. 또한, 상기 카메라 배열 정보의 값이 4인 경우, 상기 카메라 배열 정보는 상기 이동 카메라를 상기 6DoF 비디오의 카메라 배열 타입으로 나타낼 수 있다. 또한, 상기 카메라 배열 정보의 값이 5인 경우, 상기 카메라 배열 정보는 상기 사용자 정의 배열을 상기 6DoF 비디오의 카메라 배열 타입으로 나타낼 수 있다.
- [0458] 또한, 상기 메타데이터는 상기 카메라 정보의 전달 여부를 나타내는 플래그를 포함할 수 있고, 상기 플래그의 값이 1인 경우, 상기 메타데이터는 상기 카메라 정보를 포함할 수 있다. 상기 카메라 정보의 전달 여부를 나타내는 플래그는 상기 camera\_region\_info\_flag 필드를 나타낼 수 있다.
- [0459] 또한, 상기 스티칭 정보는 스티칭 대상 뷰(view)의 수를 나타내는 정보, 스티칭 대상 뷰의 위치 정보를 포함할

수 있다. 상기 스티칭 대상 뷰의 위치 정보는 상기 스티칭 대상 뷰 위치의 x 성분을 나타내는 정보, 상기 스티칭 대상 뷰 위치의 y 성분을 나타내는 정보, 상기 스티칭 대상 뷰 위치의 z 성분을 나타내는 정보, 상기 스티칭 대상 뷰 위치와 스티칭되는 6DoF 비디오 혹은 이미지간의 거리를 나타내는 정보를 포함할 수 있다. 또는, 상기 스티칭 대상 뷰의 위치 정보는 상기 스티칭 대상 뷰 위치의 yaw 성분을 나타내는 정보, 상기 스티칭 대상 뷰 위치의 pitch 성분을 나타내는 정보, 상기 스티칭 대상 뷰 위치의 roll 성분을 나타내는 정보, 상기 스티칭 대상 뷰 위치와 스티칭되는 6DoF 비디오 혹은 이미지간의 거리를 나타내는 정보, 상기 스티칭되는 6DoF 비디오 혹은 이미지의 수를 나타내는 정보를 포함할 수 있다.

[0460] 상기 스티칭 대상 뷰의 수를 나타내는 정보는 상기 num\_view\_center 필드를 나타낼 수 있고, 상기 스티칭 대상 뷰 위치의 x 성분을 나타내는 정보는 상기 position\_center\_X 필드를 나타낼 수 있고, 상기 스티칭 대상 뷰 위치의 y 성분을 나타내는 정보는 상기 position\_center\_Y 필드를 나타낼 수 있고, 상기 스티칭 대상 뷰 위치의 z 성분을 나타내는 정보는 상기 position\_center\_Z 필드를 나타낼 수 있고, 상기 스티칭 대상 뷰 위치와 스티칭되는 6DoF 비디오 혹은 이미지간의 거리를 나타내는 정보는 상기 stitching\_radius 필드를 나타낼 수 있고, 상기 스티칭되는 6DoF 비디오 혹은 이미지의 수를 나타내는 정보는 상기 num\_camera 필드를 나타낼 수 있다. 상기 스티칭 대상 뷰 위치의 yaw 성분을 나타내는 정보는 상기 position\_center\_yaw 필드를 나타낼 수 있고, 상기 스티칭 대상 뷰 위치의 pitch 성분을 나타내는 정보는 상기 position\_center\_pitch 필드를 나타낼 수 있고, 상기 스티칭 대상 뷰 위치의 roll 성분을 나타내는 정보는 상기 position\_center\_roll 필드를 나타낼 수 있다.

[0461] 또한, 상기 스티칭 정보는 상기 이미지를 도출하기 위한 스티칭 과정의 수행 여부를 나타내는 플래그를 포함할 수 있다. 상기 플래그의 값이 1인 경우, 상기 사용자의 위치에 대한 이미지를 도출하기 위한 스티칭 과정이 수행될 수 있고, 상기 플래그의 값이 0인 경우, 상기 이미지를 도출하기 위한 스티칭 과정이 수행되지 않을 수 있다. 상기 이미지를 도출하기 위한 스티칭 과정의 수행 여부를 나타내는 플래그는 상기 stitching\_flag 필드를 나타낼 수 있다.

[0462] 또한, 상기 메타데이터는 상기 6DoF 비디오의 초기 뷰(initial view)에 대한 회전 정보 및 상기 6DoF 비디오의 초기 뷰의 위치 정보를 포함할 수 있다. 여기서, 상기 초기 뷰는 상기 6DoF 비디오의 최초로 제공되는 뷰를 나타낼 수 있고, 상기 6DoF 비디오가 최초로 제공되는 때의 상기 사용자의 위치는 상기 초기 뷰의 위치로 도출될 수 있다.

[0463] 상기 6DoF 비디오의 초기 뷰에 대한 회전 정보는 상기 초기 뷰의 중점에 대한 yaw 값을 나타내는 정보, 상기 초기 뷰의 중점에 대한 pitch 값을 나타내는 정보, 상기 초기 뷰의 중점에 대한 roll 값을 나타내는 정보를 포함할 수 있다. 상기 6DoF 비디오의 초기 뷰의 위치 정보는 상기 초기 뷰의 위치의 앞뒤 방향 축에 대한 성분을 나타내는 정보, 상기 초기 뷰의 위치의 좌우 방향 축에 대한 성분을 나타내는 정보, 상기 초기 뷰의 위치의 위아래 방향 축에 대한 성분을 나타내는 정보를 포함할 수 있다. 상기 초기 뷰의 중점에 대한 yaw 값을 나타내는 정보는 상기 initial\_view\_yaw\_degree 필드를 나타낼 수 있고, 상기 초기 뷰의 중점에 대한 pitch 값을 나타내는 정보는 상기 initial\_view\_pitch\_degree 필드를 나타낼 수 있고, 상기 초기 뷰의 중점에 대한 roll 값을 나타내는 정보는 상기 initial\_view\_yaw\_roll 필드를 나타낼 수 있고, 상기 초기 뷰의 위치의 앞뒤 방향 축에 대한 성분을 나타내는 정보는 상기 initial\_center\_FB\_delta 필드를 나타낼 수 있고, 상기 초기 뷰의 위치의 좌우 방향 축에 대한 성분을 나타내는 정보는 initial\_center\_LR\_delta 필드를 나타낼 수 있고, 상기 초기 뷰의 위치의 위아래 방향 축에 대한 성분을 나타내는 정보는 initial\_center\_UD\_delta 필드를 나타낼 수 있다.

[0464] 또한, 상기 메타데이터는 FOV(Field of View) 정보를 포함할 수 있다. 상기 FOV 정보는 FOV의 수평 방향 값을 나타내는 정보 및/또는 FOV의 수직 방향 값을 나타내는 정보를 포함할 수 있다. 상기 FOV 정보를 기반으로 뷰포트의 사이즈가 도출될 수 있는바, 상기 뷰포트(viewport)는 사용자가 6DoF 비디오에서 보고 있는 영역, 즉, 디스플레이되는 영역을 나타낼 수 있다. 상기 FOV의 수평 방향 값을 나타내는 정보는 상기 content\_hfov 필드를 나타낼 수 있고, 상기 FOV의 수직 방향 값을 나타내는 정보는 상기 content\_vfov 필드를 나타낼 수 있다.

[0465] 또한, 상기 메타데이터는 상기 FOV 정보의 전달 여부를 나타내는 플래그를 포함할 수 있다. 상기 플래그의 값이 1인 경우, 상기 메타데이터는 상기 FOV 정보를 포함할 수 있고, 상기 플래그의 값이 0인 경우, 상기 메타데이터는 상기 FOV 정보를 포함하지 않을 수 있다. 상기 FOV 정보의 전달 여부를 나타내는 플래그는 상기 content\_fov\_flag 필드를 나타낼 수 있다.

[0466] 또한, 상기 메타데이터는 렌더링 정보를 포함할 수 있다. 상기 렌더링 정보는 예비 렌더링 범위의 타입을 나타내는 정보 및 사용자 위치를 나타내는 정보를 포함할 수 있다. 상기 사용자 위치를 나타내는 정보는 상기 사용자 위치의 x 성분을 나타내는 정보, 상기 사용자 위치의 y 성분을 나타내는 정보, 상기 사용자 위치의 z 성분을

나타내는 정보를 포함할 수 있다. 상기 예비 렌더링 범위의 타입을 나타내는 정보는 상기 reserved\_range\_type 필드를 나타낼 수 있고, 상기 사용자 위치의 x 성분을 나타내는 정보는 상기 center\_X 필드를 나타낼 수 있고, 상기 사용자 위치의 y 성분을 나타내는 정보는 상기 center\_Y 필드를 나타낼 수 있고, 상기 사용자 위치의 z 성분을 나타내는 정보는 상기 center\_Z 필드를 나타낼 수 있다.

[0467] 상기 예비 렌더링 범위의 타입은 사용자 위치를 중심으로 동일한 반경으로 형성되는 스피어(sphere) 타입, 사용자 위치를 중심으로 가로 세로 방향으로 형성되는 탑-셰입(top-shape) 타입, 사용자 위치를 중심으로 위 아래의 이동 범위가 형성되는 엘리베이션 디렉션(elevation direction) 타입, 사용자 위치를 중심으로 전진하거나 후진하는 방향으로 진행되는 범위가 형성되는 레터럴 디렉션(lateral direction) 타입, 사용자 위치를 중심으로 왼쪽으로 이동하거나 오른쪽으로 이동하는 범위가 형성되는 론지튜디널 디렉션(longitudinal direction) 타입 및 자유로운 형태의 범위가 형성되는 사용자 정의(user defined) 타입 중 하나일 수 있다.

[0468] 구체적으로, 예를 들어, 상기 예비 렌더링 범위의 타입을 나타내는 정보의 값이 0인 경우, 상기 예비 렌더링 범위의 타입을 나타내는 정보는 상기 스피어 타입을 상기 6DoF 비디오의 상기 예비 렌더링 범위의 타입으로 나타낼 수 있다. 또한, 상기 예비 렌더링 범위의 타입을 나타내는 정보의 값이 1인 경우, 상기 예비 렌더링 범위의 타입을 나타내는 정보는 상기 탑-셰입 타입을 상기 6DoF 비디오의 상기 예비 렌더링 범위의 타입으로 나타낼 수 있다. 또한, 상기 예비 렌더링 범위의 타입을 나타내는 정보의 값이 2인 경우, 상기 예비 렌더링 범위의 타입을 나타내는 정보는 상기 엘리베이션 디렉션 타입을 상기 6DoF 비디오의 상기 예비 렌더링 범위의 타입으로 나타낼 수 있다. 또한, 상기 예비 렌더링 범위의 타입을 나타내는 정보의 값이 3인 경우, 상기 예비 렌더링 범위의 타입을 나타내는 정보는 상기 레터럴 디렉션 타입을 상기 6DoF 비디오의 상기 예비 렌더링 범위의 타입으로 나타낼 수 있다. 또한, 상기 예비 렌더링 범위의 타입을 나타내는 정보의 값이 4인 경우, 상기 예비 렌더링 범위의 타입을 나타내는 정보는 상기 론지튜디널 디렉션 타입을 상기 6DoF 비디오의 상기 예비 렌더링 범위의 타입으로 나타낼 수 있다. 또한, 상기 예비 렌더링 범위의 타입을 나타내는 정보의 값이 5인 경우, 상기 예비 렌더링 범위의 타입을 나타내는 정보는 상기 사용자 정의 타입을 상기 6DoF 비디오의 상기 예비 렌더링 범위의 타입으로 나타낼 수 있다.

[0469] 상기 예비 렌더링 범위의 타입이 상기 스피어 타입으로 도출되는 경우, 상기 렌더링 정보는 상기 예비 렌더링 범위의 반경을 나타내는 정보를 포함할 수 있다. 상기 예비 렌더링 범위의 반경을 나타내는 정보는 상기 sphere\_radius\_to\_boundary 필드를 나타낼 수 있다.

[0470] 또한, 상기 예비 렌더링 범위의 타입이 상기 탑-셰입 타입으로 도출되는 경우, 상기 렌더링 정보는 상기 예비 렌더링 범위의 수평 범위를 나타내는 정보 및/또는 상기 예비 렌더링 범위의 수직 범위를 나타내는 정보를 포함할 수 있다. 상기 예비 렌더링 범위의 수평 범위를 나타내는 정보는 상기 horizontal\_range 필드를 나타낼 수 있고, 상기 예비 렌더링 범위의 수직 범위를 나타내는 정보는 상기 vertical\_range 필드를 나타낼 수 있다. 또는, 상기 예비 렌더링 범위의 타입이 상기 탑-셰입 타입으로 도출되는 경우, 상기 렌더링 정보는 상기 사용자 위치의 수평 방향 축에 위치하는 예비 렌더링 범위의 좌측 경계 점을 나타내는 정보, 상기 사용자의 위치의 수평 방향 축에 위치하는 예비 렌더링 범위의 우측 경계 점을 나타내는 정보, 상기 사용자 위치의 수직 방향 축에 위치하는 예비 렌더링 범위의 상측 경계 점을 나타내는 정보 및/또는 상기 사용자의 위치의 수직 방향 축에 위치하는 예비 렌더링 범위의 하측 경계 점을 나타내는 정보를 포함할 수 있다. 상기 사용자 위치의 수평 방향 축에 위치하는 예비 렌더링 범위의 좌측 경계 점을 나타내는 정보는 상기 horizontal\_range\_left\_end 필드를 나타낼 수 있고, 상기 사용자의 위치의 수평 방향 축에 위치하는 예비 렌더링 범위의 우측 경계 점을 나타내는 정보는 상기 horizontal\_range\_right\_end 필드를 나타낼 수 있고, 상기 사용자 위치의 수직 방향 축에 위치하는 예비 렌더링 범위의 상측 경계 점을 나타내는 정보는 상기 vertical\_range\_top\_end 필드를 나타낼 수 있고, 상기 사용자의 위치의 수직 방향 축에 위치하는 예비 렌더링 범위의 하측 경계 점을 나타내는 정보는 상기 vertical\_range\_bottom\_end 필드를 나타낼 수 있다.

[0471] 또한, 상기 예비 렌더링 범위의 타입이 상기 엘리베이션 디렉션 타입으로 도출되는 경우, 상기 렌더링 정보는 상기 사용자 위치의 수직 방향 축에 위치하는 예비 렌더링 범위의 상측 경계 점을 나타내는 정보 및/또는 상기 사용자의 위치의 수직 방향 축에 위치하는 예비 렌더링 범위의 하측 경계 점을 나타내는 정보를 포함할 수 있다. 상기 사용자 위치의 수직 방향 축에 위치하는 예비 렌더링 범위의 상측 경계 점을 나타내는 정보는 상기 distance\_to\_top\_endpoint 필드를 나타낼 수 있고, 상기 사용자의 위치의 수직 방향 축에 위치하는 예비 렌더링 범위의 하측 경계 점을 나타내는 정보는 상기 distance\_to\_bottom\_endpoint 필드를 나타낼 수 있다.

[0472] 또한, 상기 예비 렌더링 범위의 타입이 상기 레터럴 디렉션 타입으로 도출되는 경우, 상기 렌더링 정보는 상기

사용자 위치의 수평 방향 축에 위치하는 예비 렌더링 범위의 좌측 경계 점을 나타내는 정보 및/또는 상기 사용자 위치의 수평 방향 축에 위치하는 예비 렌더링 범위의 우측 경계 점을 나타내는 정보를 포함할 수 있다. 상기 사용자 위치의 수평 방향 축에 위치하는 예비 렌더링 범위의 좌측 경계 점을 나타내는 정보는 상기 distance\_to\_left\_endpoint 필드를 나타낼 수 있고, 상기 사용자 위치의 수평 방향 축에 위치하는 예비 렌더링 범위의 우측 경계 점을 나타내는 정보는 상기 distance\_to\_right\_endpoint 필드를 나타낼 수 있다.

[0473] 또한, 상기 예비 렌더링 범위의 타입이 상기 론지튜디널 디렉션 타입으로 도출되는 경우, 상기 렌더링 정보는 상기 사용자 위치의 전후 방향 축에 위치하는 예비 렌더링 범위의 전방 경계 점을 나타내는 정보 및/또는 상기 사용자의 위치의 전후 방향 축에 위치하는 예비 렌더링 범위의 후방 경계 점을 나타내는 정보를 포함할 수 있다. 상기 사용자 위치의 전후 방향 축에 위치하는 예비 렌더링 범위의 전방 경계 점을 나타내는 정보는 상기 distance\_to\_front\_endpoint 필드를 나타낼 수 있고, 상기 사용자의 위치의 전후 방향 축에 위치하는 예비 렌더링 범위의 후방 경계 점을 나타내는 정보는 상기 distance\_to\_back\_endpoint 필드를 나타낼 수 있다.

[0474] 또한, 상기 예비 렌더링 범위의 타입이 상기 사용자 정의 타입으로 도출되는 경우, 상기 렌더링 정보는 상기 예비 렌더링 범위의 꼭지점(vertex)을 나타내는 정보를 포함할 수 있다. 상기 예비 렌더링 범위의 꼭지점을 나타내는 정보는 상기 vertex\_X[] 필드, 상기 vertex\_Y[] 필드 및 상기 vertex\_Z[] 필드를 나타낼 수 있다.

[0475] 또한, 상기 렌더링 정보는 상기 사용자의 위치에서 주변의 상기 예비 렌더링 위치까지의 거리를 나타내는 정보를 포함할 수 있다. 상기 사용자의 위치에서 주변의 상기 예비 렌더링 위치까지의 거리를 나타내는 정보는 상기 step\_size 필드를 나타낼 수 있다. 상기 정보의 값이 작을수록 정의된 예비 렌더링 범위 내에서 렌더링되는 비디오 혹은 이미지의 수가 많아지며, 값이 클수록 예비 렌더링 범위 내에서 듬성듬성하게 예비 렌더링이 수행되는 것을 의미할 수 있다. 사용자의 위치로부터 예비 렌더링 영역의 각 경계까지의 거리를 step\_size 필드 값으로 나눈 값이 예비 렌더링 위치의 중심 개수로 도출될 수 있다.

[0476] 또한, 상기 렌더링 정보는 상기 예비 렌더링 위치의 렌더링 타입을 나타내는 정보를 포함할 수 있다. 상기 예비 렌더링 위치의 상기 렌더링 타입을 나타내는 정보는 상기 rendering\_type 필드를 포함할 수 있다.

[0477] 상기 렌더링 타입은 상기 예비 렌더링 위치에서 비디오 혹은 이미지를 전체 3D 공간(즉, 구(sphere) 전체)에 렌더링하는 홀 스피어(whole sphere) 렌더링 타입과 상기 예비 렌더링 위치에서 비디오 혹은 이미지를 파셜 뷰포트(partial viewport)에 렌더링하는 파셜 뷰포트 렌더링 타입 중 하나일 수 있다.

[0478] 구체적으로, 예를 들어, 상기 예비 렌더링 위치의 렌더링 타입을 나타내는 정보의 값이 0인 경우, 상기 예비 렌더링 위치의 렌더링 타입을 나타내는 정보는 상기 홀 스피어 렌더링 타입을 상기 렌더링 타입으로 나타낼 수 있다. 또한, 상기 예비 렌더링 위치의 렌더링 타입을 나타내는 정보의 값이 1인 경우, 상기 예비 렌더링 위치의 렌더링 타입을 나타내는 정보는 상기 파셜 뷰포트 렌더링 타입을 상기 렌더링 타입으로 나타낼 수 있다.

[0479] 상기 렌더링 타입이 상기 홀 스피어 렌더링 타입으로 도출되는 경우, 상기 렌더링 정보는 상기 예비 렌더링 위치의 비디오 혹은 이미지가 렌더링되는 3D 공간(예를 들어 구)의 반경을 나타내는 정보를 포함할 수 있다, 상기 3D 공간의 반경을 나타내는 정보는 상기 sphere\_radius 필드를 나타낼 수 있다.

[0480] 또한, 상기 렌더링 타입이 상기 파셜 뷰포트 렌더링 타입으로 도출되는 경우, 상기 렌더링 정보는 상기 예비 렌더링 위치의 비디오 혹은 이미지가 렌더링되는 상기 파셜 뷰포트의 중심점의 요 값을 나타내는 정보, 상기 파셜 뷰포트의 중심점의 피치 값을 나타내는 정보, 상기 파셜 뷰포트의 중심점의 롤 값을 나타내는 정보, 상기 파셜 뷰포트의 수평 범위를 나타내는 정보, 상기 파셜 뷰포트의 수직 범위를 나타내는 정보를 포함할 수 있다, 상기 파셜 뷰포트의 중심점의 요 값을 나타내는 정보는 상기 preliminary\_viewport\_yaw 필드를 나타낼 수 있고, 상기 파셜 뷰포트의 중심점의 피치 값을 나타내는 정보는 상기 preliminary\_viewport\_pitch 필드를 나타낼 수 있고, 상기 파셜 뷰포트의 중심점의 롤 값을 나타내는 정보는 상기 preliminary\_viewport\_roll 필드를 나타낼 수 있고, 상기 파셜 뷰포트의 수평 범위를 나타내는 정보는 상기 preliminary\_hor\_range 필드를 나타낼 수 있고, 상기 파셜 뷰포트의 수직 범위를 나타내는 정보는 상기 preliminary\_ver\_range 필드를 나타낼 수 있다. 상기 파셜 뷰포트는 상기 정보들을 기반으로 도출된 상기 중심점, 상기 수평 범위 및 상기 수직 범위를 기반으로 도출될 수 있다.

[0481] 또한, 상기 메타데이터는 상기 렌더링 정보의 전달 여부를 나타내는 플래그를 포함할 수 있다. 상기 플래그의 값이 1인 경우, 상기 메타데이터는 상기 렌더링 정보를 포함할 수 있고, 상기 플래그의 값이 0인 경우, 상기 메타데이터는 상기 렌더링 정보를 포함하지 않을 수 있다. 상기 렌더링 정보의 전달 여부를 나타내는 플래그는 상기 reserved\_rendering\_range\_flag 필드를 나타낼 수 있다.

[0482] 한편, 상기 메타데이터는 SEI 메시지를 통하여 전송될 수 있다. 또한, 상기 메타데이터는 MPD(Media

Presentation Description)의 어댑테이션 셋(AdaptationSet), 레프리젠테이션(Representation) 또는 서브 레프리젠테이션(SubRepresentation)에 포함될 수도 있다. 여기서, 상기 SEI 메시지는 2D 이미지의 디코딩 또는 2D 이미지의 3D 공간으로의 디스플레이에 대한 보조를 위하여 사용될 수 있다.

- [0483] 6DoF 비디오 전송 장치는 상기 픽처들을 인코딩한다(S2330). 6DoF 비디오 전송 장치는 상기 픽처들을 인코딩할 수 있다. 또한, 6DoF 비디오 전송 장치는 상기 메타데이터를 인코딩할 수 있다.
- [0484] 6DoF 비디오 전송 장치는 상기 인코딩된 픽처들 및 상기 메타데이터에 대하여 저장 또는 전송을 위한 처리를 수행한다(S2340). 6DoF 비디오 전송 장치는 상기 인코딩된 상기 복수의 뷰들에 대한 6DoF 비디오 및/또는 상기 메타데이터를 파일 등의 형태로 인캡슐레이션(encapsulation)할 수 있다. 6DoF 비디오 전송 장치는 인코딩된 6DoF 비디오 데이터 및/또는 상기 메타데이터를 저장 또는 전송하기 위하여 ISOBMFF, CFF 등의 파일 포맷으로 인캡슐레이션하거나, 기타 DASH 세그먼트 등의 형태로 처리할 수 있다. 6DoF 비디오 전송 장치는 상기 메타데이터를 파일 포맷 상에 포함시킬 수 있다. 예를 들어, 상기 메타데이터는 ISOBMFF 파일 포맷 상의 다양한 레벨의 박스(box)에 포함되거나 파일 내에서 별도의 트랙내의 데이터로 포함될 수 있다. 또한, 6DoF 비디오 전송 장치는 상기 메타데이터 자체를 파일로 인캡슐레이션할 수 있다. 6DoF 비디오 전송 장치는 파일 포맷에 따라 인캡슐레이션된 상기 6DoF 비디오 데이터에 전송을 위한 처리를 가할 수 있다. 6DoF 비디오 전송 장치는 임의의 전송 프로토콜에 따라 상기 6DoF 비디오 데이터를 처리할 수 있다. 전송을 위한 처리에는 방송망을 통한 전달을 위한 처리, 또는 브로드밴드 등의 통신 네트워크를 통한 전달을 위한 처리를 포함할 수 있다. 또한, 6DoF 비디오 전송 장치는 상기 메타데이터에 전송을 위한 처리를 가할 수도 있다. 6DoF 비디오 전송 장치는 전송 처리된 상기 6DoF 비디오 및 상기 메타데이터를 방송망 및/또는 브로드밴드를 통해 전송할 수 있다.
- [0485] 도 24는 본 발명에 따른 6DoF 비디오 수신 장치에 의한 6DoF 비디오 데이터 처리 방법을 개략적으로 나타낸다. 도 24에서 개시된 방법은 도 15에서 개시된 6DoF 비디오 수신 장치에 의하여 수행될 수 있다. 구체적으로 예를 들어, 도 24의 S2400은 상기 6DoF 비디오 수신 장치의 수신부 및 수신 처리부에 의하여 수행될 수 있고, S2410은 상기 6DoF 비디오 수신 장치의 데이터 디코더에 의하여 수행될 수 있고, S2420은 상기 6DoF 비디오 수신 장치의 렌더러에 의하여 수행될 수 있다.
- [0486] 6DoF 비디오 수신 장치는 6DoF 비디오의 픽처들에 대한 정보 및 상기 6DoF 비디오에 대한 메타데이터를 포함하는 신호를 획득한다(S2400).
- [0487] 6DoF 비디오 수신 장치는 방송망을 통하여 6DoF 비디오 전송 장치로부터 시그널링된 상기 6DoF 비디오의 픽처들에 대한 정보 및 상기 메타데이터를 수신할 수 있다. 또한, 6DoF 비디오 수신 장치는 브로드밴드 등의 통신 네트워크, 또는 저장매체를 통하여 상기 픽처들에 대한 정보 및 상기 메타데이터를 수신할 수도 있다.
- [0488] 6DoF 비디오 수신 장치는 수신된 상기 픽처들에 대한 정보 및 상기 메타데이터에 대해 전송 프로토콜에 따른 처리를 수행할 수 있다. 또한, 6DoF 비디오 수신 장치는 전송한 6DoF 비디오 전송 장치의 전송을 위한 처리의 역과정을 수행할 수 있다.
- [0489] 상기 6DoF 비디오의 픽처들은 카메라에 의하여 캡처된 상기 카메라의 뷰(view)에 대한 6DoF 비디오 데이터가 맵핑된 픽처들을 나타낼 수 있다.
- [0490] 또한, 상기 메타데이터는 상술한 camera\_info\_flag 필드, num\_view 필드, Camera\_capture\_array\_row 필드, Camera\_capture\_array\_col 필드, Baseline\_length 필드, Master\_camera\_ID 필드, camera\_center\_pitch 필드, camera\_center\_yaw 필드, camera\_center\_roll 필드, camera\_center\_FB 필드, camera\_center\_UD 필드, camera\_center\_RL 필드, num\_view\_center 필드, position\_center\_X 필드, position\_center\_Y 필드, position\_center\_Z 필드, num\_camera 필드, stitching\_radius 필드, position\_center\_yaw 필드, position\_center\_roll 필드, position\_center\_pitch 필드, position\_center\_distance 필드, initial\_view\_yaw\_degree 필드, initial\_view\_pitch\_degree 필드, initial\_view\_roll\_degree 필드, initial\_center\_FB\_delta 필드, initial\_center\_UD\_delta 필드, initial\_center\_RL\_delta 필드, content\_fov\_flag 필드, capture\_info\_flag 필드, region\_info\_flag 필드, content\_hfov 필드, content\_vfov 필드, reserved\_rendering\_range\_flag 필드, rendering\_type 필드, reserved\_range\_type 필드, center\_X 필드, center\_Y 필드, center\_Z 필드, step size 필드, sphere\_radius\_to\_boundary 필드, horizontal\_range 필드, vertical\_range 필드, distance\_to\_top\_endpoint 필드, distance\_to\_bottom\_endpoint 필드, distance\_to\_left\_endpoint 필드, distance\_to\_right\_endpoint 필드, distance\_to\_front\_endpoint 필드, distance\_to\_back\_endpoint 필드, vertex\_X[] 필드, vertex\_Y[] 필드, vertex\_Z[] 필드, sphere\_radius 필드,

preliminary\_viewport\_yaw 필드, preliminary\_viewport\_pitch 필드, preliminary\_viewport\_roll 필드, preliminary\_hor\_range 필드 및/또는 preliminary\_ver\_range 필드를 포함할 수 있다. 상기 필드들의 의미는 상술한 바와 같다.

- [0491] 구체적으로, 일 예로, 상기 메타데이터는 6DoF 비디오를 캡처한 카메라 정보 및 스티칭(stiching) 정보를 포함할 수 있다.
- [0492] 상기 카메라 정보는 원점이 되는 뷰에 대한 카메라 ID를 나타내는 정보, 카메라 배열 정보, 상기 카메라로 캡처된 뷰들의 수를 나타내는 정보, 카메라 배열의 열을 나타내는 정보, 카메라 배열의 행을 나타내는 정보, 카메라 간의 거리를 나타내는 정보 및/또는 상기 뷰들의 위치 정보를 포함할 수 있다.
- [0493] 상기 원점이 되는 뷰에 대한 카메라 ID를 나타내는 정보는 상기 Master\_camera\_ID 필드를 나타낼 수 있고, 상기 카메라 배열 정보는 상기 Capture\_type 필드를 나타낼 수 있고, 상기 카메라로 캡처된 뷰들의 수를 나타내는 정보는 상기 num\_view 필드를 나타낼 수 있고, 카메라 배열의 열을 나타내는 정보는 상기 Camera\_capture\_array\_row 필드를 나타낼 수 있고, 카메라 배열의 행을 나타내는 정보는 상기 Camera\_capture\_array\_col 필드를 나타낼 수 있고, 카메라 간의 거리를 나타내는 정보는 Baseline\_length 필드를 나타낼 수 있다.
- [0494] 상기 뷰들의 위치 정보는 해당 뷰에 대한 6DoF 비디오의 중점의 yaw 값을 나타내는 정보, 해당 뷰에 대한 6DoF 비디오의 중점의 pitch 값을 나타내는 정보, 해당 뷰에 대한 6DoF 비디오의 중점의 roll 값을 나타내는 정보, 해당 뷰 위치의 앞뒤 방향 축에 대한 성분을 나타내는 정보, 해당 뷰 위치의 위아래 방향 축에 대한 성분을 나타내는 정보 및/또는 해당 뷰 위치의 좌우 방향 축에 대한 성분을 나타내는 정보를 포함할 수 있다. 여기서, 상기 앞뒤 방향 축은 y축, 상기 좌우 방향 축은 x축, 상기 위아래 방향 축은 z축을 나타낼 수 있다. 상기 해당 뷰에 대한 6DoF 비디오의 중점의 yaw 값을 나타내는 정보는 상기 camera\_center\_yaw 필드를 나타낼 수 있고, 상기 해당 뷰에 대한 6DoF 비디오의 중점의 pitch 값을 나타내는 정보는 상기 camera\_center\_pitch 필드를 나타낼 수 있고, 상기 해당 뷰에 대한 6DoF 비디오의 중점의 roll 값을 나타내는 정보는 상기 camera\_center\_roll 필드를 나타낼 수 있고, 상기 해당 뷰 위치의 앞뒤 방향 축에 대한 성분을 나타내는 정보는 상기 camera\_center\_FB 필드를 나타낼 수 있고, 상기 해당 뷰 위치의 위아래 방향 축에 대한 성분을 나타내는 정보는 상기 camera\_center\_UD 필드를 나타낼 수 있고, 상기 해당 뷰 위치의 좌우 방향 축에 대한 성분을 나타내는 정보는 상기 camera\_center\_RL 필드를 나타낼 수 있다.
- [0495] 상기 카메라 배열 정보는 상기 6DoF 비디오를 캡처하는데 사용된 카메라 배열(camera array) 타입을 나타낼 수 있다. 여기서, 상기 카메라 배열 타입은 HDCA (High Density Camera Array), LDCA(Low Density Camera Array), 마이크로렌즈(microlens), 전방향 카메라(Omnidirectional Camera), 이동 카메라(Movement camera), 사용자 정의(user defined) 배열 중 하나일 수 있다.
- [0496] 구체적으로, 예를 들어, 상기 카메라 배열 정보의 값이 0인 경우, 상기 카메라 배열 정보는 상기 HDCA를 상기 6DoF 비디오의 카메라 배열 타입으로 나타낼 수 있다. 또한, 상기 카메라 배열 정보의 값이 1인 경우, 상기 카메라 배열 정보는 상기 LDCA를 상기 6DoF 비디오의 카메라 배열 타입으로 나타낼 수 있다. 또한, 상기 카메라 배열 정보의 값이 2인 경우, 상기 카메라 배열 정보는 상기 마이크로렌즈를 상기 6DoF 비디오의 카메라 배열 타입으로 나타낼 수 있다. 또한, 상기 카메라 배열 정보의 값이 3인 경우, 상기 카메라 배열 정보는 상기 전방향 카메라를 상기 6DoF 비디오의 카메라 배열 타입으로 나타낼 수 있다. 또한, 상기 카메라 배열 정보의 값이 4인 경우, 상기 카메라 배열 정보는 상기 이동 카메라를 상기 6DoF 비디오의 카메라 배열 타입으로 나타낼 수 있다. 또한, 상기 카메라 배열 정보의 값이 5인 경우, 상기 카메라 배열 정보는 상기 사용자 정의 배열을 상기 6DoF 비디오의 카메라 배열 타입으로 나타낼 수 있다.
- [0497] 또한, 상기 메타데이터는 상기 카메라 정보의 전달 여부를 나타내는 플래그를 포함할 수 있고, 상기 플래그의 값이 1인 경우, 상기 메타데이터는 상기 카메라 정보를 포함할 수 있다. 상기 카메라 정보의 전달 여부를 나타내는 플래그는 상기 camera\_region\_info\_flag 필드를 나타낼 수 있다.
- [0498] 또한, 상기 스티칭 정보는 스티칭 대상 뷰(view)의 수를 나타내는 정보, 상기 스티칭 대상 뷰의 위치 정보를 포함할 수 있다. 상기 스티칭 대상 뷰의 위치 정보는 상기 스티칭 대상 뷰 위치의 x 성분을 나타내는 정보, 상기 스티칭 대상 뷰 위치의 y 성분을 나타내는 정보, 상기 스티칭 대상 뷰 위치의 z 성분을 나타내는 정보, 상기 스티칭 대상 뷰 위치와 스티칭되는 6DoF 비디오 혹은 이미지간의 거리를 나타내는 정보를 포함할 수 있다. 또는, 상기 스티칭 대상 뷰의 위치 정보는 상기 스티칭 대상 뷰 위치의 yaw 성분을 나타내는 정보, 상기 스티칭 대상

뷰 위치의 pitch 성분을 나타내는 정보, 상기 스티칭 대상 뷰 위치의 roll 성분을 나타내는 정보, 상기 스티칭 대상 뷰 위치와 스티칭되는 6DoF 비디오 혹은 이미지간의 거리를 나타내는 정보, 상기 스티칭되는 6DoF 비디오 혹은 이미지의 수를 나타내는 정보를 포함할 수 있다.

[0499] 상기 스티칭 대상 뷰의 수를 나타내는 정보는 상기 num\_view\_center 필드를 나타낼 수 있고, 상기 스티칭 대상 뷰 위치의 x 성분을 나타내는 정보는 상기 position\_center\_X 필드를 나타낼 수 있고, 상기 스티칭 대상 뷰 위치의 y 성분을 나타내는 정보는 상기 position\_center\_Y 필드를 나타낼 수 있고, 상기 스티칭 대상 뷰 위치의 z 성분을 나타내는 정보는 상기 position\_center\_Z 필드를 나타낼 수 있고, 상기 스티칭 대상 뷰 위치와 스티칭되는 6DoF 비디오 혹은 이미지간의 거리를 나타내는 정보는 상기 stitching\_radius 필드를 나타낼 수 있고, 상기 스티칭되는 6DoF 비디오 혹은 이미지의 수를 나타내는 정보는 상기 num\_camera 필드를 나타낼 수 있다. 상기 스티칭 대상 뷰 위치의 yaw 성분을 나타내는 정보는 상기 position\_center\_yaw 필드를 나타낼 수 있고, 상기 스티칭 대상 뷰 위치의 pitch 성분을 나타내는 정보는 상기 position\_center\_pitch 필드를 나타낼 수 있고, 상기 스티칭 대상 뷰 위치의 roll 성분을 나타내는 정보는 상기 position\_center\_roll 필드를 나타낼 수 있다.

[0500] 또한, 상기 스티칭 정보는 상기 이미지를 도출하기 위한 스티칭 과정의 수행 여부를 나타내는 플래그를 포함할 수 있다. 상기 플래그의 값이 1인 경우, 상기 사용자의 위치에 대한 이미지를 도출하기 위한 스티칭 과정이 수행될 수 있고, 상기 플래그의 값이 0인 경우, 상기 이미지를 도출하기 위한 스티칭 과정이 수행되지 않을 수 있다. 상기 이미지를 도출하기 위한 스티칭 과정의 수행 여부를 나타내는 플래그는 상기 stitching\_flag 필드를 나타낼 수 있다.

[0501] 또한, 상기 메타데이터는 상기 6DoF 비디오의 초기 뷰(initial view)에 대한 회전 정보 및 상기 6DoF 비디오의 초기 뷰의 위치 정보를 포함할 수 있다. 여기서, 상기 초기 뷰는 상기 6DoF 비디오의 최초로 제공되는 뷰를 나타낼 수 있고, 상기 6DoF 비디오가 최초로 제공되는 때의 상기 사용자의 위치는 상기 초기 뷰의 위치로 도출될 수 있다.

[0502] 상기 6DoF 비디오의 초기 뷰에 대한 회전 정보는 상기 초기 뷰의 중점에 대한 yaw 값을 나타내는 정보, 상기 초기 뷰의 중점에 대한 pitch 값을 나타내는 정보, 상기 초기 뷰의 중점에 대한 roll 값을 나타내는 정보를 포함할 수 있다. 상기 6DoF 비디오의 초기 뷰의 위치 정보는 상기 초기 뷰의 위치의 앞뒤 방향 축에 대한 성분을 나타내는 정보, 상기 초기 뷰의 위치의 좌우 방향 축에 대한 성분을 나타내는 정보, 상기 초기 뷰의 위치의 위아래 방향 축에 대한 성분을 나타내는 정보를 포함할 수 있다. 상기 초기 뷰의 중점에 대한 yaw 값을 나타내는 정보는 상기 initial\_view\_yaw\_degree 필드를 나타낼 수 있고, 상기 초기 뷰의 중점에 대한 pitch 값을 나타내는 정보는 상기 initial\_view\_pitch\_degree 필드를 나타낼 수 있고, 상기 초기 뷰의 중점에 대한 roll 값을 나타내는 정보는 상기 initial\_view\_yaw\_roll 필드를 나타낼 수 있고, 상기 초기 뷰의 위치의 앞뒤 방향 축에 대한 성분을 나타내는 정보는 상기 initial\_center\_FB\_delta 필드를 나타낼 수 있고, 상기 초기 뷰의 위치의 좌우 방향 축에 대한 성분을 나타내는 정보는 initial\_center\_LR\_delta 필드를 나타낼 수 있고, 상기 초기 뷰의 위치의 위아래 방향 축에 대한 성분을 나타내는 정보는 initial\_center\_UD\_delta 필드를 나타낼 수 있다.

[0503] 또한, 상기 메타데이터는 FOV(Field of View) 정보를 포함할 수 있다. 상기 FOV 정보는 FOV의 수평 방향 값을 나타내는 정보 및/또는 FOV의 수직 방향 값을 나타내는 정보를 포함할 수 있다. 상기 FOV 정보를 기반으로 뷰포트의 사이즈가 도출될 수 있는바, 상기 뷰포트(viewport)는 사용자가 6DoF 비디오에서 보고 있는 영역, 즉, 디스플레이되는 영역을 나타낼 수 있다. 상기 FOV의 수평 방향 값을 나타내는 정보는 상기 content\_hfov 필드를 나타낼 수 있고, 상기 FOV의 수직 방향 값을 나타내는 정보는 상기 content\_vfov 필드를 나타낼 수 있다.

[0504] 또한, 상기 메타데이터는 상기 FOV 정보의 전달 여부를 나타내는 플래그를 포함할 수 있다. 상기 플래그의 값이 1인 경우, 상기 메타데이터는 상기 FOV 정보를 포함할 수 있고, 상기 플래그의 값이 0인 경우, 상기 메타데이터는 상기 FOV 정보를 포함하지 않을 수 있다. 상기 FOV 정보의 전달 여부를 나타내는 플래그는 상기 content\_fov\_flag 필드를 나타낼 수 있다.

[0505] 또한, 상기 메타데이터는 렌더링 정보를 포함할 수 있다. 상기 렌더링 정보는 예비 렌더링 범위의 타입을 나타내는 정보 및 사용자 위치를 나타내는 정보를 포함할 수 있다. 상기 사용자 위치를 나타내는 정보는 상기 사용자 위치의 x 성분을 나타내는 정보, 상기 사용자 위치의 y 성분을 나타내는 정보, 상기 사용자 위치의 z 성분을 나타내는 정보를 포함할 수 있다. 상기 예비 렌더링 범위의 타입을 나타내는 정보는 상기 reserved\_range\_type 필드를 나타낼 수 있고, 상기 사용자 위치의 x 성분을 나타내는 정보는 상기 center\_X 필드를 나타낼 수 있고, 상기 사용자 위치의 y 성분을 나타내는 정보는 상기 center\_Y 필드를 나타낼 수 있고, 상기 사용자 위치의 z 성분을 나타내는 정보는 상기 center\_Z 필드를 나타낼 수 있다.



- [0506] 상기 예비 렌더링 범위의 타입은 사용자 위치를 중심으로 동일한 반경으로 형성되는 스피어(sphere) 타입, 사용자 위치를 중심으로 가로 세로 방향으로 형성되는 탑-셰입(top-shape) 타입, 사용자 위치를 중심으로 위 아래의 이동 범위가 형성되는 엘리베이션 디렉션(elevation direction) 타입, 사용자 위치를 중심으로 전진하거나 후진하는 방향으로 진행되는 범위가 형성되는 레터럴 디렉션(lateral direction) 타입, 사용자 위치를 중심으로 왼쪽으로 이동하거나 오른쪽으로 이동하는 범위가 형성되는 론지튜디널 디렉션(longitudinal direction) 타입 및 자유로운 형태의 범위가 형성되는 사용자 정의(user defined) 타입 중 하나일 수 있다.
- [0507] 구체적으로, 예를 들어, 상기 예비 렌더링 범위의 타입을 나타내는 정보의 값이 0인 경우, 상기 예비 렌더링 범위의 타입을 나타내는 정보는 상기 스피어 타입을 상기 6DoF 비디오의 상기 예비 렌더링 범위의 타입으로 나타낼 수 있다. 또한, 상기 예비 렌더링 범위의 타입을 나타내는 정보의 값이 1인 경우, 상기 예비 렌더링 범위의 타입을 나타내는 정보는 상기 탑-셰입 타입을 상기 6DoF 비디오의 상기 예비 렌더링 범위의 타입으로 나타낼 수 있다. 또한, 상기 예비 렌더링 범위의 타입을 나타내는 정보의 값이 2인 경우, 상기 예비 렌더링 범위의 타입을 나타내는 정보는 상기 엘리베이션 디렉션 타입을 상기 6DoF 비디오의 상기 예비 렌더링 범위의 타입으로 나타낼 수 있다. 또한, 상기 예비 렌더링 범위의 타입을 나타내는 정보의 값이 3인 경우, 상기 예비 렌더링 범위의 타입을 나타내는 정보는 상기 레터럴 디렉션 타입을 상기 6DoF 비디오의 상기 예비 렌더링 범위의 타입으로 나타낼 수 있다. 또한, 상기 예비 렌더링 범위의 타입을 나타내는 정보의 값이 4인 경우, 상기 예비 렌더링 범위의 타입을 나타내는 정보는 상기 론지튜디널 디렉션 타입을 상기 6DoF 비디오의 상기 예비 렌더링 범위의 타입으로 나타낼 수 있다. 또한, 상기 예비 렌더링 범위의 타입을 나타내는 정보의 값이 5인 경우, 상기 예비 렌더링 범위의 타입을 나타내는 정보는 상기 사용자 정의 타입을 상기 6DoF 비디오의 상기 예비 렌더링 범위의 타입으로 나타낼 수 있다.
- [0508] 상기 예비 렌더링 범위의 타입이 상기 스피어 타입으로 도출되는 경우, 상기 렌더링 정보는 상기 예비 렌더링 범위의 반경을 나타내는 정보를 포함할 수 있다. 상기 예비 렌더링 범위의 반경을 나타내는 정보는 상기 sphere\_radius\_to\_boundary 필드를 나타낼 수 있다.
- [0509] 또한, 상기 예비 렌더링 범위의 타입이 상기 탑-셰입 타입으로 도출되는 경우, 상기 렌더링 정보는 상기 예비 렌더링 범위의 수평 범위를 나타내는 정보 및/또는 상기 예비 렌더링 범위의 수직 범위를 나타내는 정보를 포함할 수 있다. 상기 예비 렌더링 범위의 수평 범위를 나타내는 정보는 상기 horizontal\_range 필드를 나타낼 수 있고, 상기 예비 렌더링 범위의 수직 범위를 나타내는 정보는 상기 vertical\_range 필드를 나타낼 수 있다. 또는, 상기 예비 렌더링 범위의 타입이 상기 탑-셰입 타입으로 도출되는 경우, 상기 렌더링 정보는 상기 사용자 위치의 수평 방향 축에 위치하는 예비 렌더링 범위의 좌측 경계 점을 나타내는 정보, 상기 사용자의 위치의 수평 방향 축에 위치하는 예비 렌더링 범위의 우측 경계 점을 나타내는 정보, 상기 사용자 위치의 수직 방향 축에 위치하는 예비 렌더링 범위의 상측 경계 점을 나타내는 정보 및/또는 상기 사용자의 위치의 수직 방향 축에 위치하는 예비 렌더링 범위의 하측 경계 점을 나타내는 정보를 포함할 수 있다. 상기 사용자 위치의 수평 방향 축에 위치하는 예비 렌더링 범위의 좌측 경계 점을 나타내는 정보는 상기 horizontal\_range\_left\_end 필드를 나타낼 수 있고, 상기 사용자의 위치의 수평 방향 축에 위치하는 예비 렌더링 범위의 우측 경계 점을 나타내는 정보는 상기 horizontal\_range\_right\_end 필드를 나타낼 수 있고, 상기 사용자 위치의 수직 방향 축에 위치하는 예비 렌더링 범위의 상측 경계 점을 나타내는 정보는 상기 vertical\_range\_top\_end 필드를 나타낼 수 있고, 상기 사용자의 위치의 수직 방향 축에 위치하는 예비 렌더링 범위의 하측 경계 점을 나타내는 정보는 상기 vertical\_range\_bottom\_end 필드를 나타낼 수 있다.
- [0510] 또한, 상기 예비 렌더링 범위의 타입이 상기 엘리베이션 디렉션 타입으로 도출되는 경우, 상기 렌더링 정보는 상기 사용자 위치의 수직 방향 축에 위치하는 예비 렌더링 범위의 상측 경계 점을 나타내는 정보 및/또는 상기 사용자의 위치의 수직 방향 축에 위치하는 예비 렌더링 범위의 하측 경계 점을 나타내는 정보를 포함할 수 있다. 상기 사용자 위치의 수직 방향 축에 위치하는 예비 렌더링 범위의 상측 경계 점을 나타내는 정보는 상기 distance\_to\_top\_endpoint 필드를 나타낼 수 있고, 상기 사용자의 위치의 수직 방향 축에 위치하는 예비 렌더링 범위의 하측 경계 점을 나타내는 정보는 상기 distance\_to\_bottom\_endpoint 필드를 나타낼 수 있다.
- [0511] 또한, 상기 예비 렌더링 범위의 타입이 상기 레터럴 디렉션 타입으로 도출되는 경우, 상기 렌더링 정보는 상기 사용자 위치의 수평 방향 축에 위치하는 예비 렌더링 범위의 좌측 경계 점을 나타내는 정보 및/또는 상기 사용자 위치의 수평 방향 축에 위치하는 예비 렌더링 범위의 우측 경계 점을 나타내는 정보를 포함할 수 있다. 상기 사용자 위치의 수평 방향 축에 위치하는 예비 렌더링 범위의 좌측 경계 점을 나타내는 정보는 상기 distance\_to\_left\_endpoint 필드를 나타낼 수 있고, 상기 사용자 위치의 수평 방향 축에 위치하는 예비 렌더링

범위의 우측 경계 점을 나타내는 정보는 상기 distance\_to\_right\_endpoint 필드를 나타낼 수 있다.

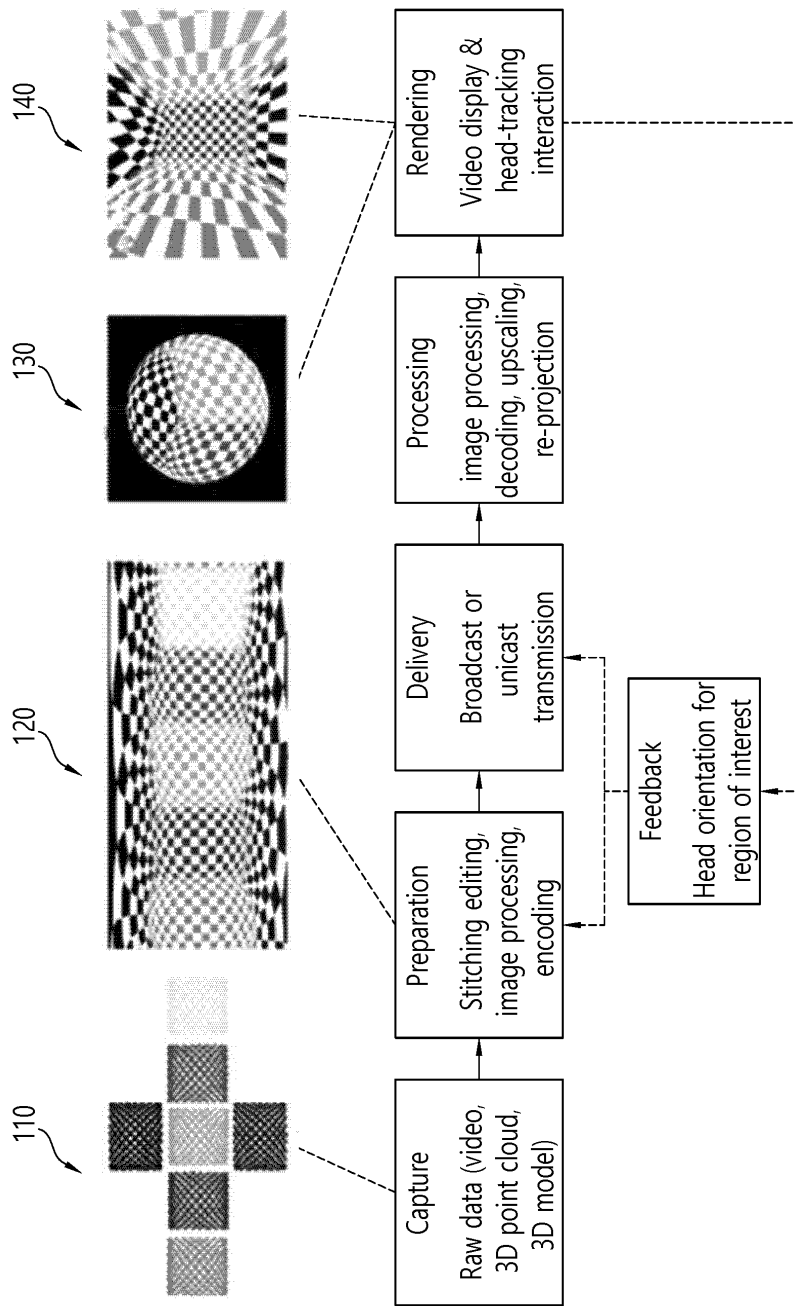
- [0512] 또한, 상기 예비 렌더링 범위의 타입이 상기 론지튜디널 디렉션 타입으로 도출되는 경우, 상기 렌더링 정보는 상기 사용자 위치의 전후 방향 축에 위치하는 예비 렌더링 범위의 전방 경계 점을 나타내는 정보 및/또는 상기 사용자의 위치의 전후 방향 축에 위치하는 예비 렌더링 범위의 후방 경계 점을 나타내는 정보를 포함할 수 있다. 상기 사용자 위치의 전후 방향 축에 위치하는 예비 렌더링 범위의 전방 경계 점을 나타내는 정보는 상기 distance\_to\_front\_endpoint 필드를 나타낼 수 있고, 상기 사용자의 위치의 전후 방향 축에 위치하는 예비 렌더링 범위의 후방 경계 점을 나타내는 정보는 상기 distance\_to\_back\_endpoint 필드를 나타낼 수 있다.
- [0513] 또한, 상기 예비 렌더링 범위의 타입이 상기 사용자 정의 타입으로 도출되는 경우, 상기 렌더링 정보는 상기 예비 렌더링 범위의 꼭지점(vertex)을 나타내는 정보를 포함할 수 있다. 상기 예비 렌더링 범위의 꼭지점을 나타내는 정보는 상기 vertex\_X[] 필드, 상기 vertex\_Y[] 필드 및 상기 vertex\_Z[] 필드를 나타낼 수 있다.
- [0514] 또한, 상기 렌더링 정보는 상기 사용자의 위치에서 주변의 상기 예비 렌더링 위치까지의 거리를 나타내는 정보를 포함할 수 있다. 상기 사용자의 위치에서 주변의 상기 예비 렌더링 위치까지의 거리를 나타내는 정보는 상기 step\_size 필드를 나타낼 수 있다. 상기 정보의 값이 작을수록 정의된 예비 렌더링 범위 내에서 렌더링되는 비디오 혹은 이미지의 수가 많아지며, 값이 클수록 예비 렌더링 범위 내에서 듬성듬성하게 예비 렌더링이 수행되는 것을 의미할 수 있다. 사용자의 위치로부터 예비 렌더링 영역의 각 경계까지의 거리를 step\_size 필드 값으로 나눈 값이 예비 렌더링 위치의 중심 개수로 도출될 수 있다.
- [0515] 또한, 상기 렌더링 정보는 상기 예비 렌더링 위치의 렌더링 타입을 나타내는 정보를 포함할 수 있다. 상기 예비 렌더링 위치의 상기 렌더링 타입을 나타내는 정보는 상기 rendering\_type 필드를 포함할 수 있다.
- [0516] 상기 렌더링 타입은 상기 예비 렌더링 위치에서 비디오 혹은 이미지를 전체 3D 공간(즉, 구(sphere) 전체)에 렌더링하는 홀 스피어(whole sphere) 렌더링 타입과 상기 예비 렌더링 위치에서 비디오 혹은 이미지를 파셜 뷰포트(partial viewport)에 렌더링하는 파셜 뷰포트 렌더링 타입 중 하나일 수 있다.
- [0517] 구체적으로, 예를 들어, 상기 예비 렌더링 위치의 렌더링 타입을 나타내는 정보의 값이 0인 경우, 상기 예비 렌더링 위치의 렌더링 타입을 나타내는 정보는 상기 홀 스피어 렌더링 타입을 상기 렌더링 타입으로 나타낼 수 있다. 또한, 상기 예비 렌더링 위치의 렌더링 타입을 나타내는 정보의 값이 1인 경우, 상기 예비 렌더링 위치의 렌더링 타입을 나타내는 정보는 상기 파셜 뷰포트 렌더링 타입을 상기 렌더링 타입으로 나타낼 수 있다.
- [0518] 상기 렌더링 타입이 상기 홀 스피어 렌더링 타입으로 도출되는 경우, 상기 렌더링 정보는 상기 예비 렌더링 위치의 비디오 혹은 이미지가 렌더링되는 3D 공간(예를 들어 구)의 반경을 나타내는 정보를 포함할 수 있다, 상기 3D 공간의 반경을 나타내는 정보는 상기 sphere\_radius 필드를 나타낼 수 있다.
- [0519] 또한, 상기 렌더링 타입이 상기 파셜 뷰포트 렌더링 타입으로 도출되는 경우, 상기 렌더링 정보는 상기 예비 렌더링 위치의 비디오 혹은 이미지가 렌더링되는 상기 파셜 뷰포트의 중심의 요 값을 나타내는 정보, 상기 파셜 뷰포트의 중심의 피치 값을 나타내는 정보, 상기 파셜 뷰포트의 중심의 롤 값을 나타내는 정보, 상기 파셜 뷰포트의 수평 범위를 나타내는 정보, 상기 파셜 뷰포트의 수직 범위를 나타내는 정보를 포함할 수 있다, 상기 파셜 뷰포트의 중심의 요 값을 나타내는 정보는 상기 preliminary\_viewport\_yaw 필드를 나타낼 수 있고, 상기 파셜 뷰포트의 중심의 피치 값을 나타내는 정보는 상기 preliminary\_viewport\_pitch 필드를 나타낼 수 있고, 상기 파셜 뷰포트의 중심의 롤 값을 나타내는 정보는 상기 preliminary\_viewport\_roll 필드를 나타낼 수 있고, 상기 파셜 뷰포트의 수평 범위를 나타내는 정보는 상기 preliminary\_hor\_range 필드를 나타낼 수 있고, 상기 파셜 뷰포트의 수직 범위를 나타내는 정보는 상기 preliminary\_ver\_range 필드를 나타낼 수 있다. 상기 파셜 뷰포트는 상기 정보들을 기반으로 도출된 상기 중심, 상기 수평 범위 및 상기 수직 범위를 기반으로 도출될 수 있다.
- [0520] 또한, 상기 메타데이터는 상기 렌더링 정보의 전달 여부를 나타내는 플래그를 포함할 수 있다. 상기 플래그의 값이 1인 경우, 상기 메타데이터는 상기 렌더링 정보를 포함할 수 있고, 상기 플래그의 값이 0인 경우, 상기 메타데이터는 상기 렌더링 정보를 포함하지 않을 수 있다. 상기 렌더링 정보의 전달 여부를 나타내는 플래그는 상기 reserved\_rendering\_range\_flag 필드를 나타낼 수 있다.
- [0521] 한편, 상기 메타데이터는 SEI 메시지를 통하여 수신될 수 있다. 또한, 상기 메타데이터는 MPD(Media Presentation Description)의 어댑테이션 셋(AdaptationSet), 레프리젠테이션(Representation) 또는 서브 레프리젠테이션(SubRepresentation)에 포함될 수도 있다. 여기서, 상기 SEI 메시지는 2D 이미지의 디코딩 또는 2D 이미지의 3D 공간으로의 디스플레이에 대한 보조를 위하여 사용될 수 있다.

- [0522] 6DoF 비디오 수신 장치는 상기 메타데이터 및 상기 픽처들에 대한 정보를 기반으로 상기 픽처들을 디코딩하고, 사용자의 위치에 대한 이미지를 도출한다(S2410). 6DoF 비디오 수신 장치는 픽처들에 대한 정보를 기반으로 상기 픽처들을 디코딩할 수 있다.
- [0523] 상기 픽처들은 적어도 하나의 카메라로 캡처된 뷰들의 6DoF 비디오가 맵핑된 픽처들을 나타낼 수 있다. 또한, 상기 뷰들은 주요 뷰(primary view, PV) 및 부차 뷰(secondary view, SV) 를 포함할 수 있고, 상기 픽처들은 PV 에 대한 픽처 및 SV 에 대한 픽처를 포함할 수 있다. 이 경우, 6DoF 비디오 수신 장치는 상기 PV 에 대한 픽처 및 상기 SV에 대한 픽처를 디코딩할 수 있다.
- [0524] 한편, 스티칭 과정의 수행 여부를 나타내는 플래그가 상기 스티칭 과정이 수행됨을 나타내는 경우, 즉, 상기 플래그의 값이 1인 경우, 6DoF 비디오 수신 장치는 스티칭 과정을 수행할 수 있다. 구체적으로, 6DoF 비디오 수신 장치는 상기 카메라 정보 및 상기 스티칭 정보를 기반으로 적어도 하나의 스티칭 대상 뷰를 도출할 수 있고, 상기 디코딩된 픽처들 중 상기 스티칭 대상 뷰에 대한 픽처들을 도출할 수 있고, 상기 픽처들을 스티칭하여 상기 사용자의 위치에 대한 상기 이미지를 도출할 수 있다.
- [0525] 즉, 상기 스티칭 대상 뷰가 도출된 경우, 6DoF 비디오 수신 장치는 상기 스티칭 대상 뷰에 대한 상기 스티칭되는 6DoF 비디오 혹은 이미지의 수를 나타내는 정보, 카메라 배열 정보, 카메라 위치를 나타내는 정보를 기반으로 상기 픽처들 중 상기 스티칭 대상 뷰에 대한 캡처드 씬(captured scene)을 도출할 수 있다. 상기 캡처드 씬이 도출되는 경우, 6DoF 비디오 수신 장치는 프로젝션 스킴(projection scheme) 정보와 구 반경 사이즈(sphere radius size) 정보 등을 기반으로 상기 캡처드 씬이 스티칭될 3D 공간을 생성하는 구 생성(sphere generation) 과정을 수행할 수 있다. 6DoF 비디오 수신 장치는 상기 캡처드 씬을 기반으로 상기 이미지를 도출할 수 있다. 한편, 상기 스티칭 대상 뷰에 대한 캡처드 씬이 없는 경우, 6DoF 비디오 수신 장치는 상기 스티칭 대상 뷰의 주변 뷰들에 대한 픽처들을 기반으로 가상의 이미지를 생성할 수 있다.
- [0526] 한편, 상술한 내용과 같이 상기 메타데이터는 6DoF 비디오의 초기 뷰(initial view)에 대한 회전 정보 및 상기 6DoF 비디오의 초기 뷰의 위치 정보를 포함할 수 있다. 여기서, 상기 초기 뷰는 상기 6DoF 비디오의 최초로 제공되는 뷰를 나타내고, 상기 6DoF 비디오가 최초로 제공되는 때의 상기 현재 위치는 상기 초기 뷰의 위치로 도출될 수 있다.
- [0527] 6DoF 비디오 수신 장치는 상기 메타데이터를 기반으로 상기 이미지를 처리하여 상기 사용자의 위치를 중심으로 하는 3D 공간으로 렌더링한다(S2420).
- [0528] 6DoF 비디오 수신 장치는 상기 메타데이터를 기반으로 상기 이미지의 6DoF 비디오를 3D 공간으로 맵핑할 수 있다. 구체적으로, 6DoF 비디오 수신 장치는 프로젝션 스킴에 따른 3D 모델에 상기 이미지의 6DoF 비디오 데이터를 맵핑할 수 있다. 상기 프로젝션 스킴은 프로젝션 스킴 정보를 기반으로 도출될 수 있고, 상기 메타데이터는 상기 프로젝션 스킴 정보를 포함할 수 있다.
- [0529] 또한, 6DoF 비디오 수신 장치는 상기 렌더링 정보를 기반으로 상기 사용자 위치에 대한 예비 렌더링 범위를 도출할 수 있다. 구체적으로, 상기 렌더링 정보는 상기 예비 렌더링 범위의 타입을 나타내는 정보를 포함할 수 있다. 또한, 상기 렌더링 정보는 상기 타입에 대한 추가적인 정보들을 포함할 수 있다. 6DoF 비디오 수신 장치는 상기 예비 렌더링 범위의 타입을 나타내는 정보 및 상기 추가적인 정보들을 기반으로 상기 예비 렌더링 범위의 형태 및 사이즈를 도출할 수 있다.
- [0530] 또한, 6DoF 비디오 수신 장치는 상기 렌더링 정보를 기반으로 상기 예비 렌더링 범위 내 예비 렌더링 위치의 렌더링 타입을 도출할 수 있다. 상기 렌더링 타입은 상술한 홀 스피어 렌더링 타입과 파셜 뷰포트 렌더링 타입 중 하나로 도출될 수 있다. 상기 렌더링 정보는 상기 렌더링 타입에 대한 추가적인 정보를 포함할 수 있다. 상기 렌더링 타입에 대한 상기 추가적인 정보를 기반으로 렌더링 영역이 도출될 수 있다.
- [0531] 상기 예비 렌더링 범위가 도출되고, 상기 사용자의 위치가 이동되는 경우, 이동되는 위치가 상기 예비 렌더링 범위 내 상기 예비 렌더링 위치인 경우, 상기 이동되는 위치의 이미지는 상기 렌더링 타입에 따라서 렌더링될 수 있다. 예를 들어, 상기 예비 렌더링 범위 내 상기 예비 렌더링 위치의 렌더링 타입이 상기 홀 스피어 렌더링 타입으로 도출되는 경우, 상기 이동되는 위치의 이미지는 구 전체에 렌더링될 수 있고, 상기 예비 렌더링 범위 내 상기 예비 렌더링 위치의 렌더링 타입이 상기 파셜 뷰포트 렌더링 타입으로 도출되는 경우, 상기 이동되는 위치의 이미지는 파셜 뷰포트에 렌더링될 수 있다.
- [0532] 상술한 본 발명에 따르면 사용자의 위치 이동에 따른 6DoF 콘텐츠를 제공할 수 있는바, 사용자에게 몰입형 미디

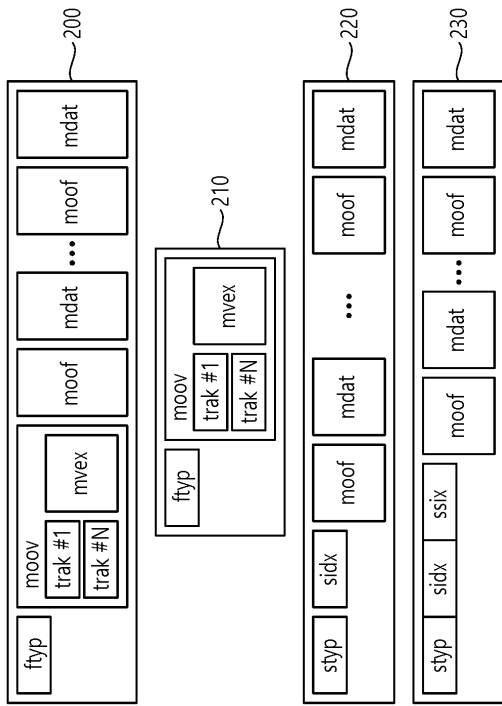
어/실감미디어(Immersive media)의 경험을 제공할 수 있다.

- [0533] 또한, 본 발명에 따르면 특정 위치에 대한 스티칭 정보를 시그널링하여, 상기 특정 위치에 대한 캡처된 6DoF 비디오가 없는 경우에도 주변 위치의 6DoF 비디오를 기반으로 상기 특정 위치에 대한 6DoF 비디오를 생성 및 사용자에게 제공할 수 있는바, 6DoF 콘텐츠를 보다 효율적으로 제공할 수 있다.
- [0534] 또한, 본 발명에 따르면 6DoF 콘텐츠에서 제공하는 이동 방향 또는 예측되는 사용자의 위치 이동 방향에 따른 예비 렌더링 정보를 시그널링하여, 사용자에게 제공되어야 하는 6DoF 콘텐츠가 우선적으로 전송 및 렌더링될 수 있는바, 6DoF 콘텐츠를 보다 효율적으로 제공할 수 있다.
- [0535] 전술한 단계들은 실시예에 따라 생략되거나, 유사/동일한 동작을 수행하는 다른 단계에 의해 대체될 수 있다.
- [0536] 본 발명의 일 실시예에 따른 6DoF 비디오 전송 장치는 전술한 데이터 입력부, 스티처, 시그널링 처리부, 프로젝션 처리부, 데이터 인코더, 전송 처리부 및/또는 전송부를 포함할 수 있다. 각각의 내부 컴포넌트들은 전술한 바와 같다. 본 발명의 일 실시예에 따른 6DoF 비디오 전송 장치 및 그 내부 컴포넌트들은, 전술한 본 발명의 6DoF 비디오를 전송하는 방법의 실시예들을 수행할 수 있다.
- [0537] 본 발명의 일 실시예에 따른 6DoF 비디오 수신 장치는 전술한 수신부, 수신 처리부, 데이터 디코더, 시그널링 파서, 리-프로젝션 처리부 및/또는 렌더러를 포함할 수 있다. 각각의 내부 컴포넌트들은 전술한 바와 같다. 본 발명의 일 실시예에 따른 6DoF 비디오 수신 장치 및 그 내부 컴포넌트들은, 전술한 본 발명의 6DoF 비디오를 수신하는 방법의 실시예들을 수행할 수 있다.
- [0538] 전술한 장치의 내부 컴포넌트들은 메모리에 저장된 연속된 수행과정들을 실행하는 프로세서들이거나, 그 외의 하드웨어로 구성된 하드웨어 컴포넌트들일 수 있다. 이 들은 장치 내/외부에 위치할 수 있다.
- [0539] 전술한 모듈들은 실시예에 따라 생략되거나, 유사/동일한 동작을 수행하는 다른 모듈에 의해 대체될 수 있다.
- [0540] 전술한 각각의 파트, 모듈 또는 유닛은 메모리(또는 저장 유닛)에 저장된 연속된 수행과정들을 실행하는 프로세서이거나 하드웨어 파트일 수 있다. 전술한 실시예에 기술된 각 단계들은 프로세서 또는 하드웨어 파트들에 의해 수행될 수 있다. 전술한 실시예에 기술된 각 모듈/블록/유닛들은 하드웨어/프로세서로서 동작할 수 있다. 또한, 본 발명이 제시하는 방법들은 코드로서 실행될 수 있다. 이 코드는 프로세서가 읽을 수 있는 저장매체에 쓰여질 수 있고, 따라서 장치(apparatus)가 제공하는 프로세서에 의해 읽혀질 수 있다.
- [0541] 상술한 실시예에서, 방법들은 일련의 단계 또는 블록으로써 순서도를 기초로 설명되고 있지만, 본 발명은 단계들의 순서에 한정되는 것은 아니며, 어떤 단계는 상술한 바와 다른 단계와 다른 순서로 또는 동시에 발생할 수 있다. 또한, 당업자라면 순서도에 나타내어진 단계들이 배타적이지 않고, 다른 단계가 포함되거나 순서도의 하나 또는 그 이상의 단계가 본 발명의 범위에 영향을 미치지 않고 삭제될 수 있음을 이해할 수 있을 것이다.
- [0542] 본 발명에서 실시예들이 소프트웨어로 구현될 때, 상술한 방법은 상술한 기능을 수행하는 모듈(과정, 기능 등)로 구현될 수 있다. 모듈은 메모리에 저장되고, 프로세서에 의해 실행될 수 있다. 메모리는 프로세서 내부 또는 외부에 있을 수 있고, 잘 알려진 다양한 수단으로 프로세서와 연결될 수 있다. 프로세서는 ASIC(application-specific integrated circuit), 다른 칩셋, 논리 회로 및/또는 데이터 처리 장치를 포함할 수 있다. 메모리는 ROM(read-only memory), RAM(random access memory), 플래쉬 메모리, 메모리 카드, 저장 매체 및/또는 다른 저장 장치를 포함할 수 있다.

도면  
도면1



도면2



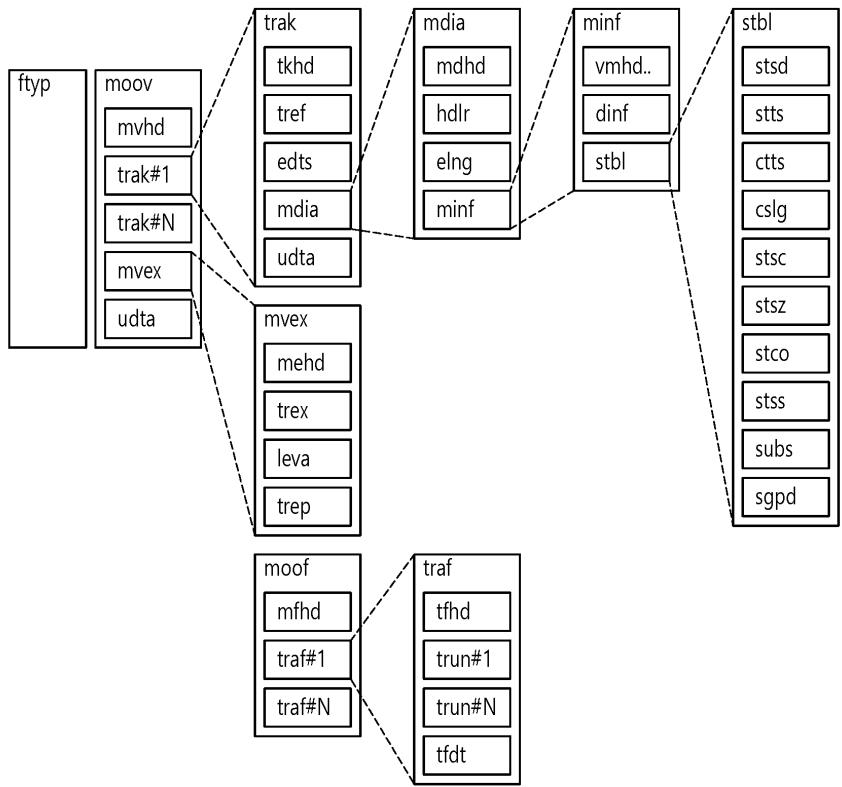
250

```

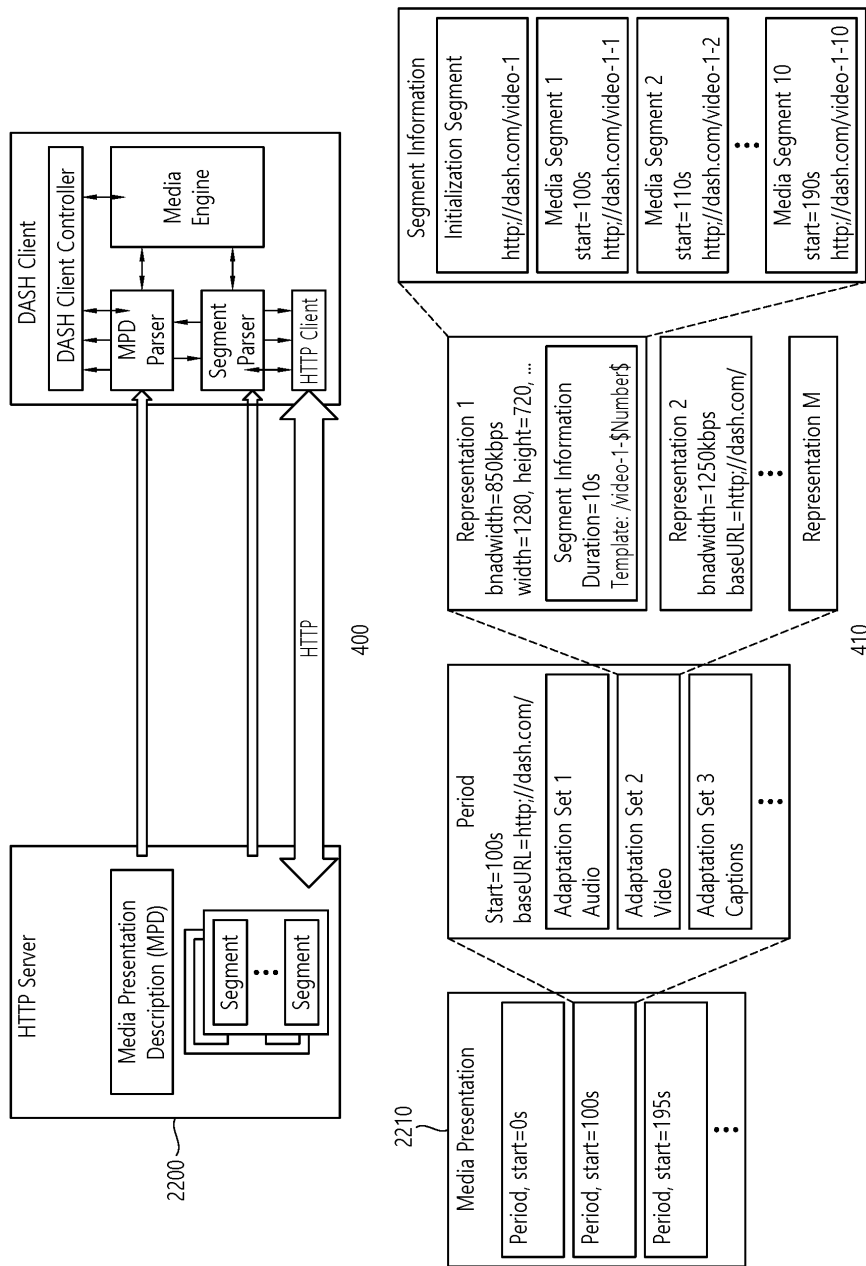
aligned(8) class Box (unsigned int(32) boxtype, optional
unsigned int(8)[16] extended_type)
{
    unsigned int(32) size;
    unsigned int(32) type = boxtype;
    if (size==1) {
        unsigned int(64) largesize;
    } else if (size==0) {
        // box extends to end of file
    }
    if (boxtype=='uuid') {
        unsigned int(8)[16] usertype = extended_type;
    }
}

aligned(8) class FullBox(unsigned int(32) boxtype, unsigned int(8)
v, bit(24) f) extends Box(boxtype) {
    unsigned int(8) version = v;
    bit(24) flags = f;
}
    
```

도면3

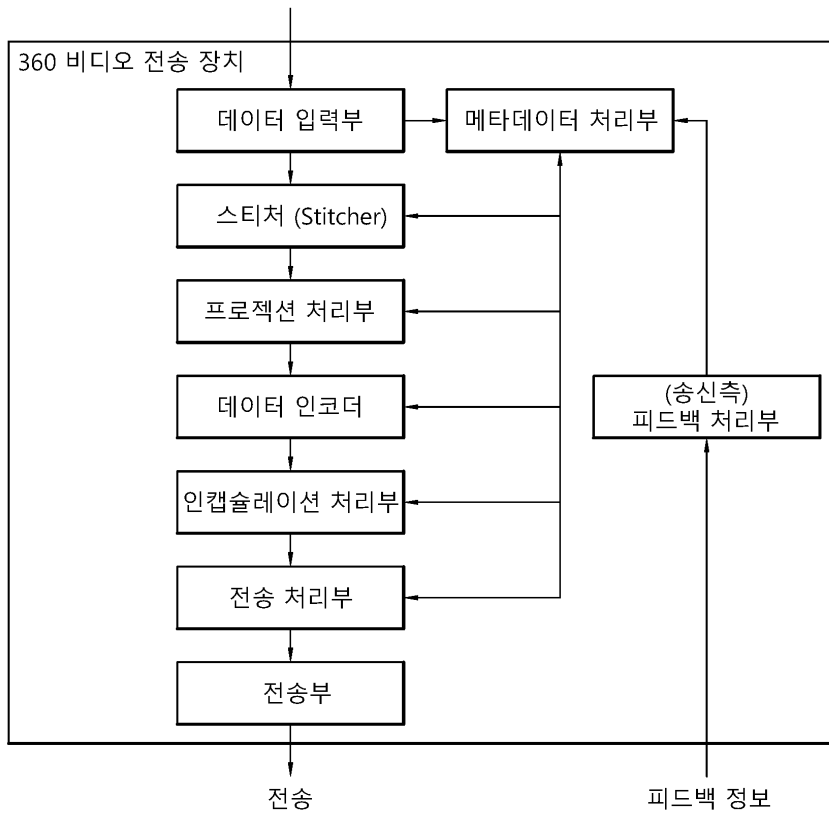


도면4

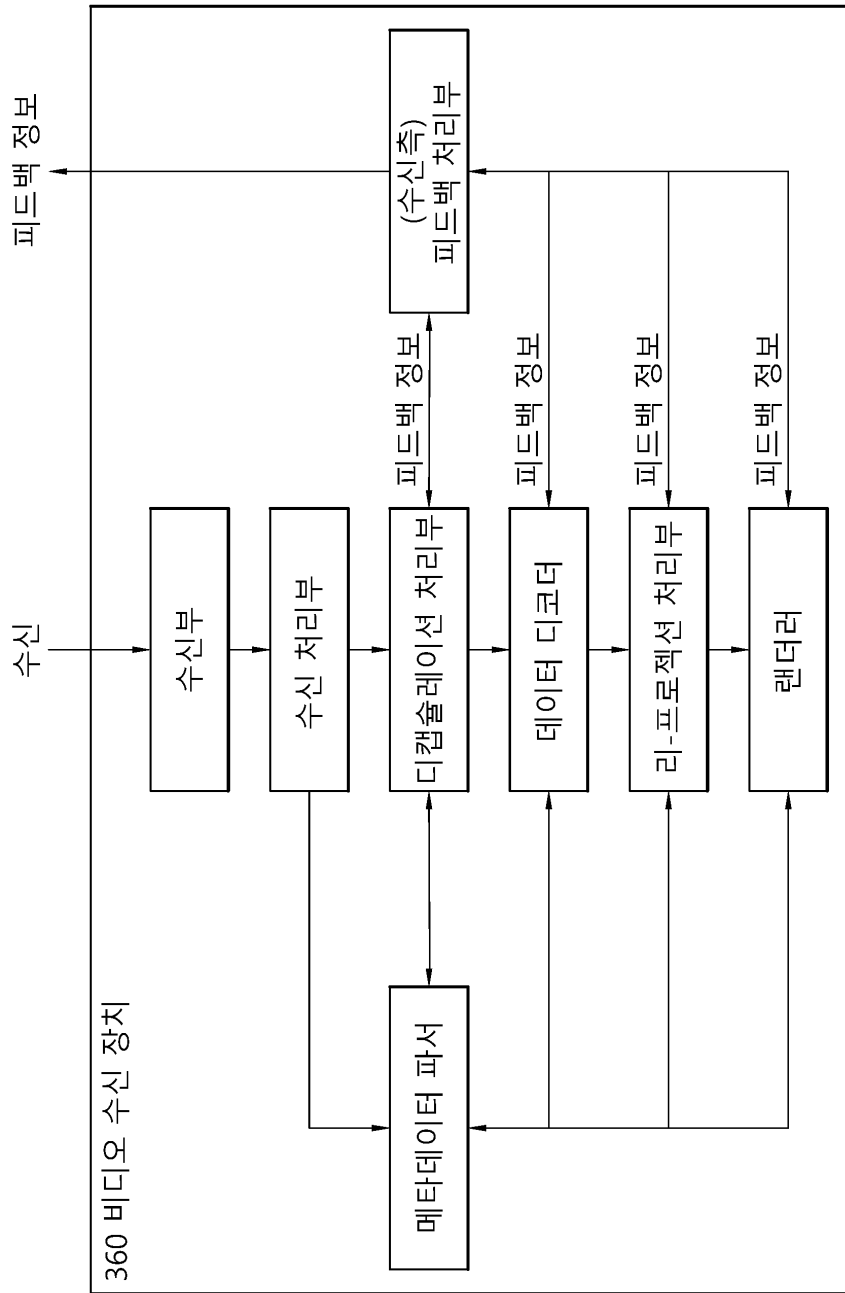




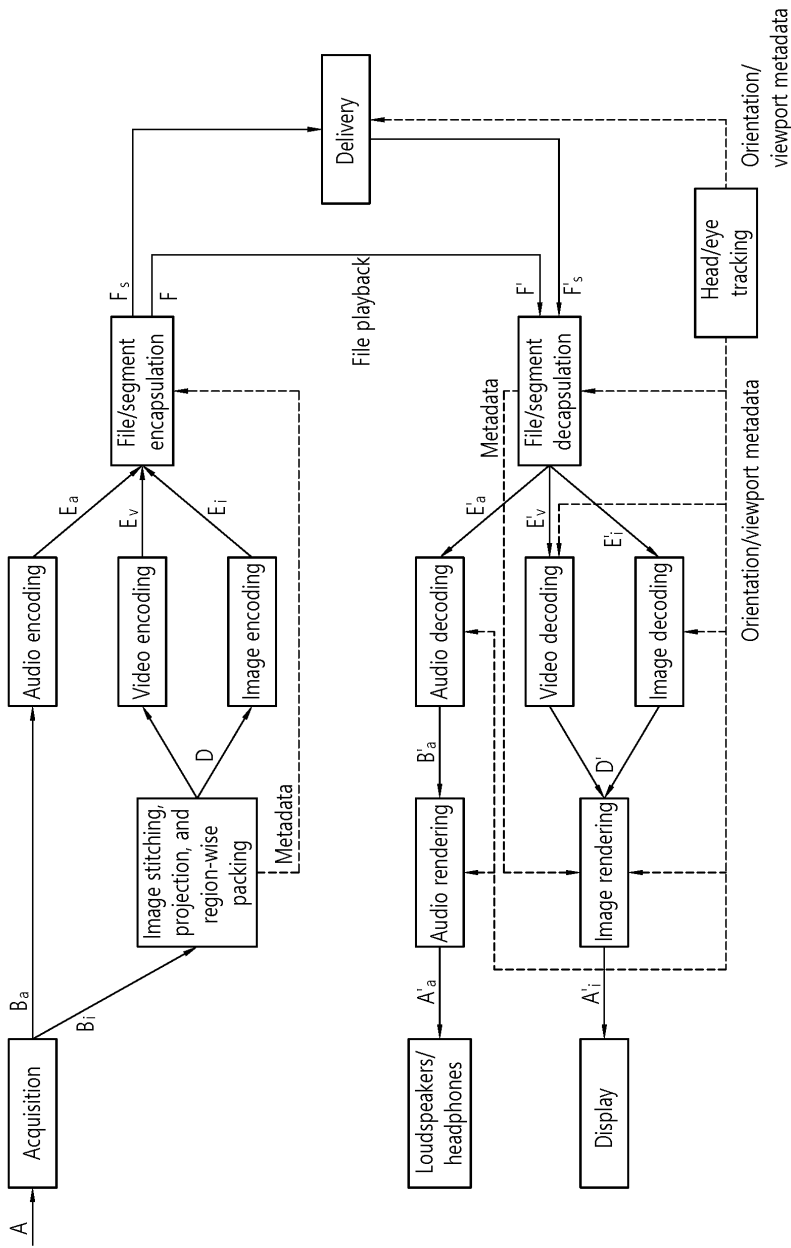
도면5



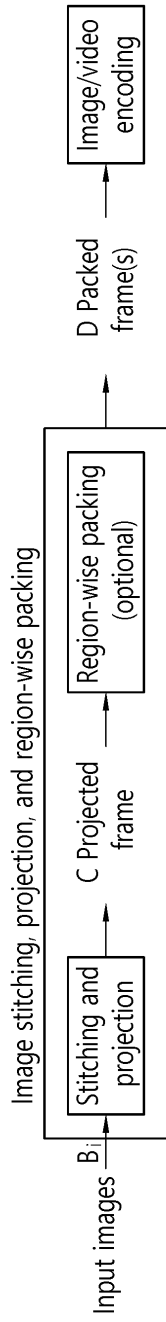
도면6



도면7a

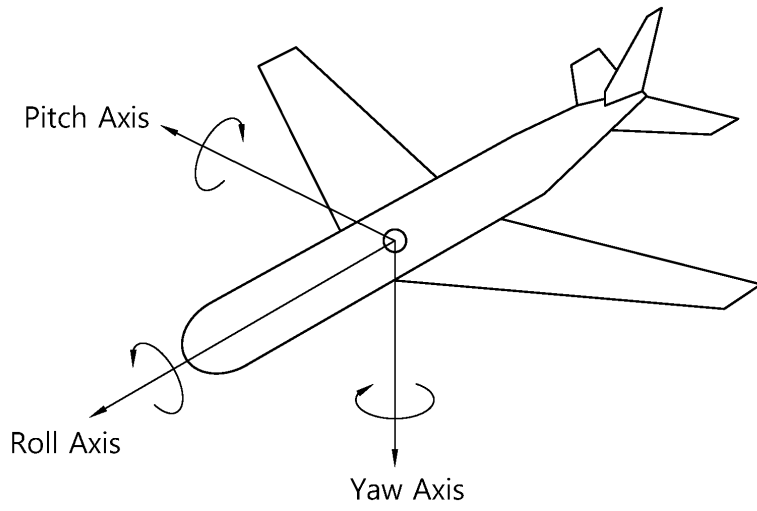


도면7b

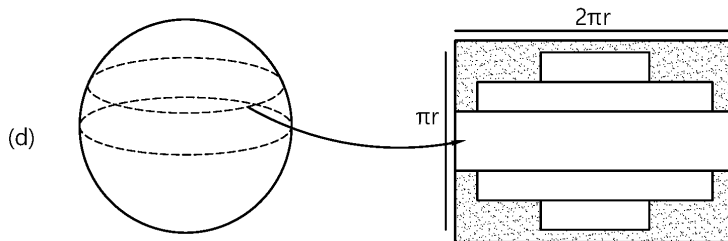
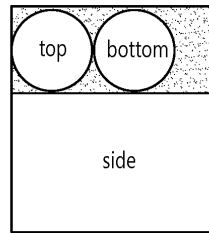
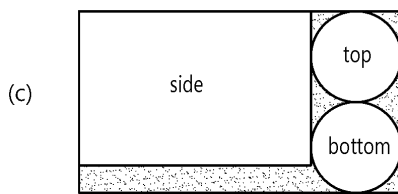
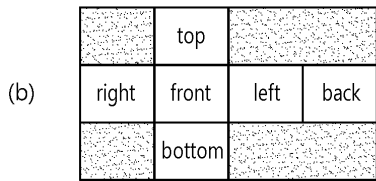
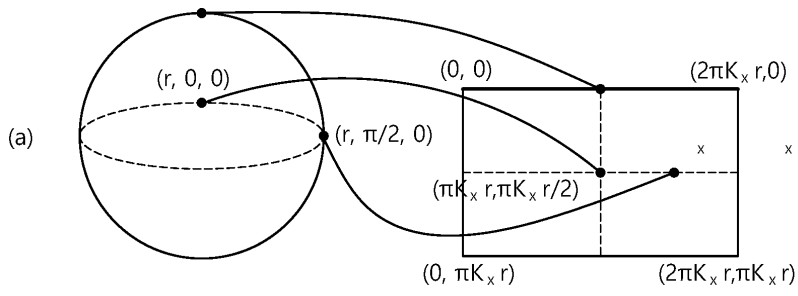


도면8

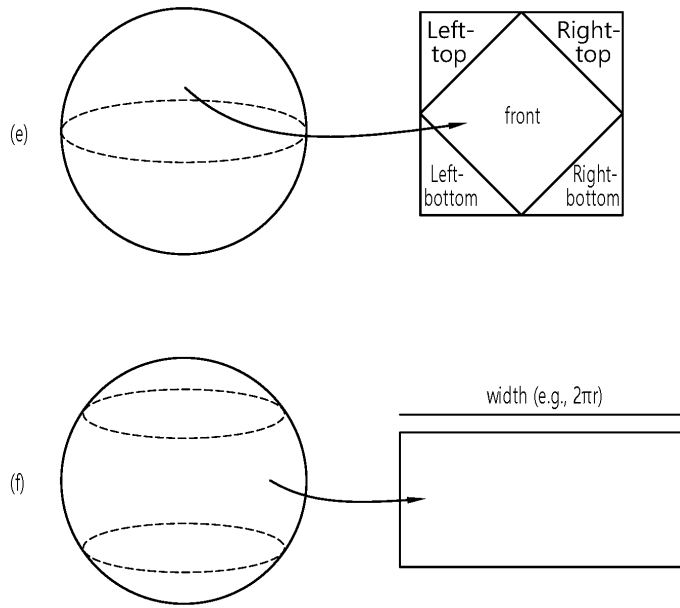
$(3\pi K \ r/2, \pi K \ r/2)$



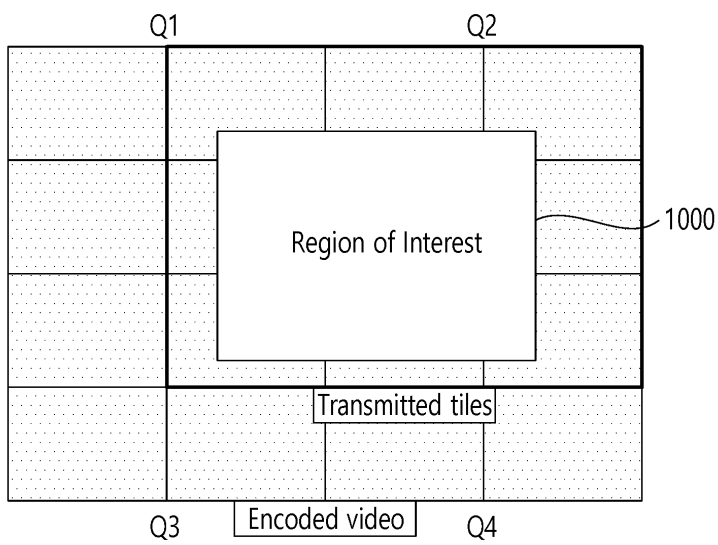
도면9a



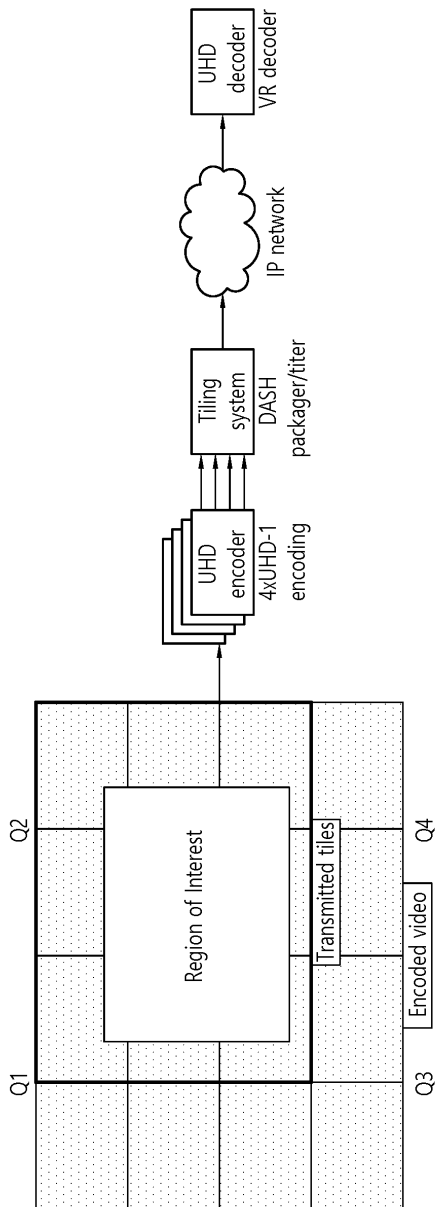
도면9b



도면10a



도면10b





도면11

```

...
    <기본 메타데이터>
    unsigned int(8)    vr_geometry;
    unsigned int(8)    projection_schme;

    <Stereoscopic 관련 메타데이터>
    unsigned int(1)    is_stereoscopic;
    unsigned int(3)    stereo_mode;

    <Initial View 관련 메타데이터>
    signed int(8)      initial_view_yaw_degree;
    signed int(8)      initial_view_pitch_degree;
    signed int(8)      initial_view_roll_degree;

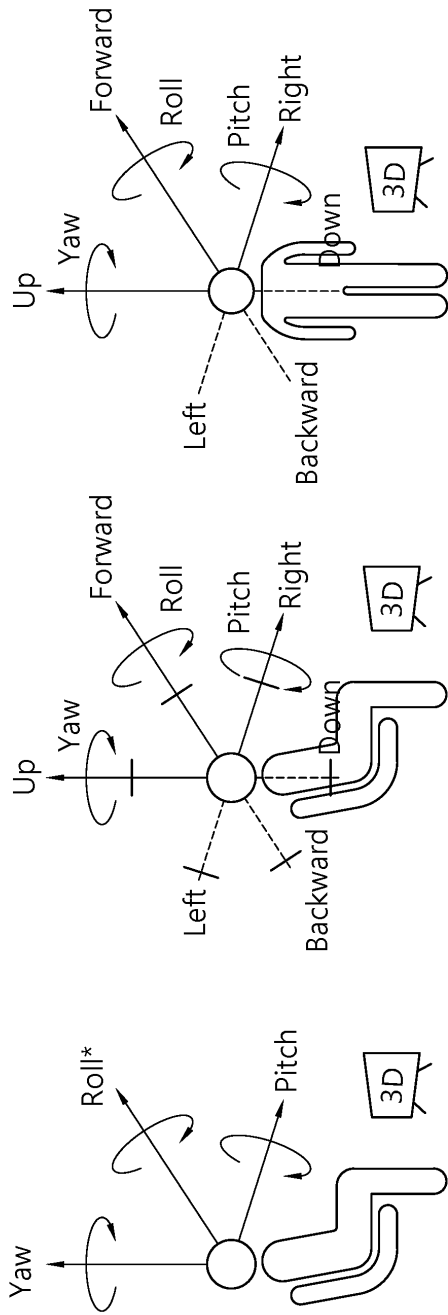
    <ROI 관련 메타데이터>
    unsigned int(1)    2d_roi_range_flag;
    unsigned int(1)    3d_roi_range_flag;
    if (2d_roi_region_flag==1) {
        unsigned int(16) min_top_left_x;
        unsigned int(16) max_top_left_x;
        unsigned int(16) min_top_left_y;
        unsigned int(16) max_top_left_y;
        unsigned int(16) min_width;
        unsigned int(16) max_width;
        unsigned int(16) min_height;
        unsigned int(16) max_height;
        unsigned int(16) min_x;
        unsigned int(16) max_x;
        unsigned int(16) min_y;
        unsigned int(16) max_y;
    }
    if (3d_roi_region_flag==1){
        unsigned int(16) min_yaw;
        unsigned int(16) max_yaw;
        unsigned int(16) min_pitch;
        unsigned int(16) max_pitch;
        unsigned int(16) min_roll;
        unsigned int(16) max_roll;
        unsigned int(16) min_field_of_view;
        unsigned int(16) max_field_of_view;
    }

    <Field Of View 관련 메타데이터>
    unsigned int(1)    content_fov_flag;
    if (content_fov_flag==1) {
        unsigned int(16)    content_fov;
    }

    <Cropped Region 관련 메타데이터>
    unsigned int(1)    is_copped_region;
    if (content_fov_flag==1) {
        unsigned int(16)    cr_region_left_top_x;
        unsigned int(16)    cr_region_left_top_y;
        unsigned int(16)    cr_region_width;
        unsigned int(16)    cr_region_height;
    }
    ...

```

도면12

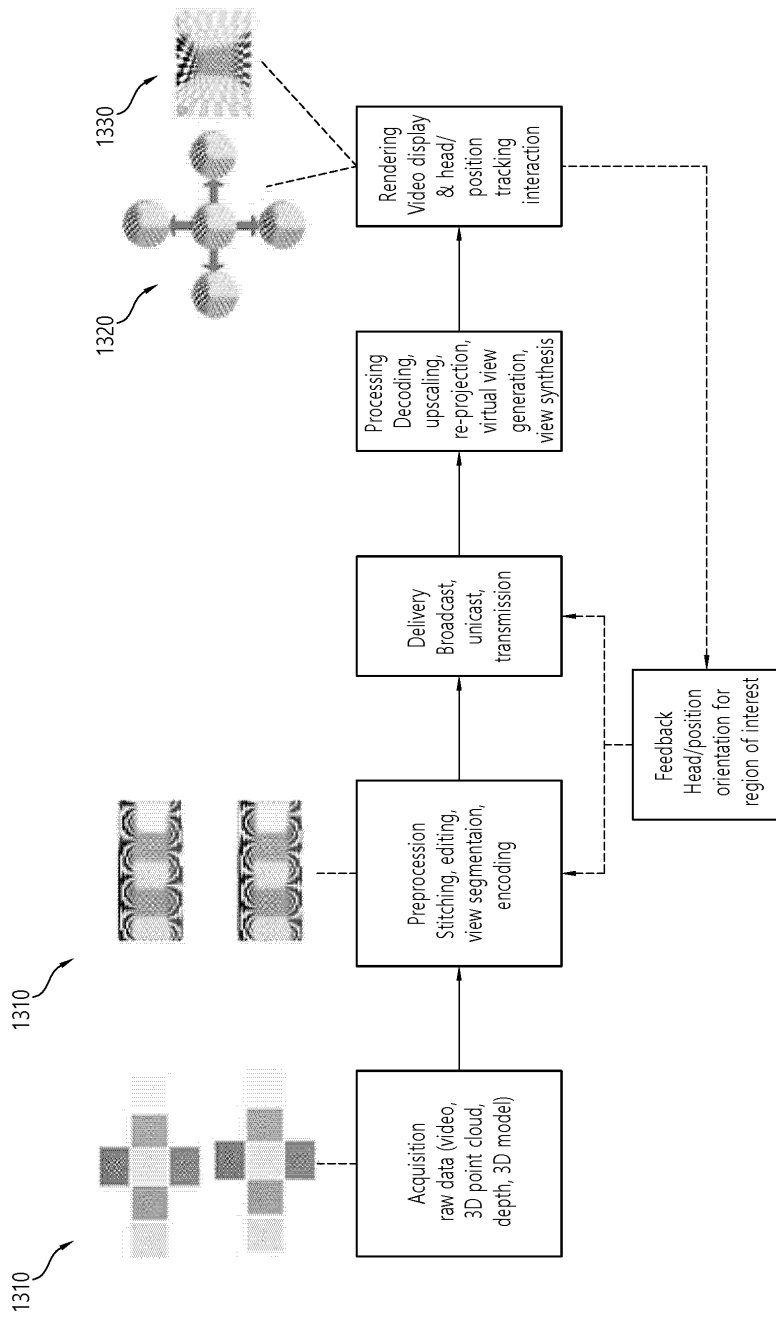


(c)

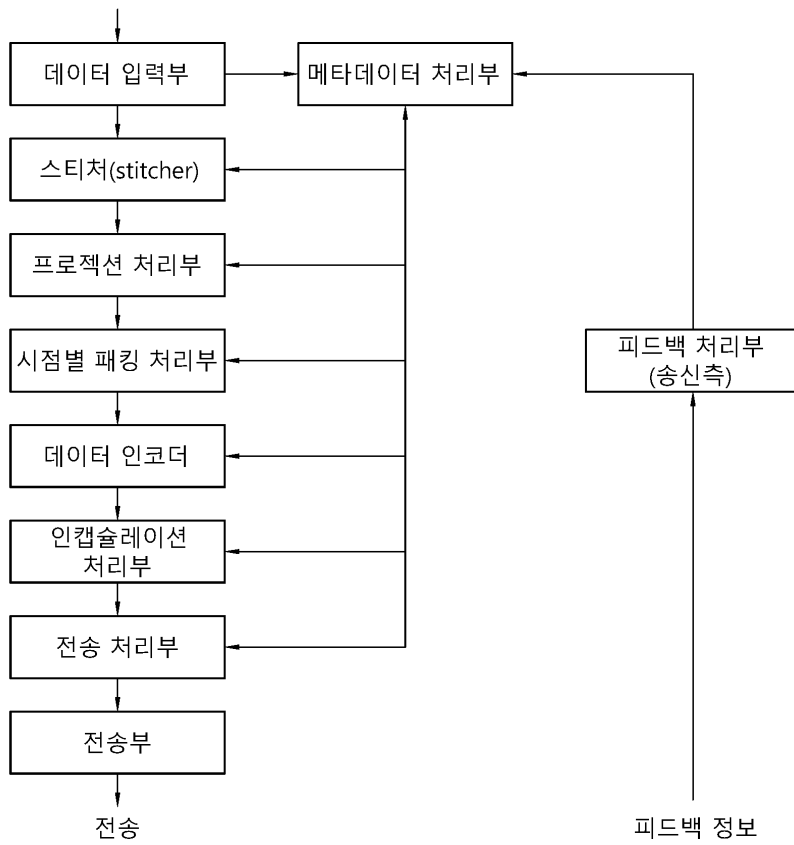
(b)

(a)

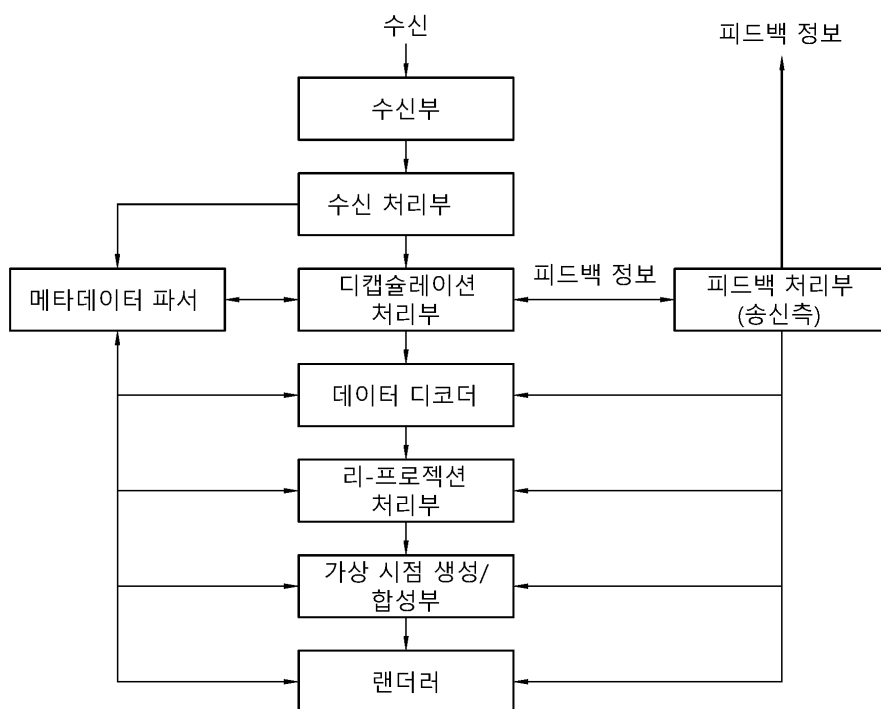
도면13



도면14

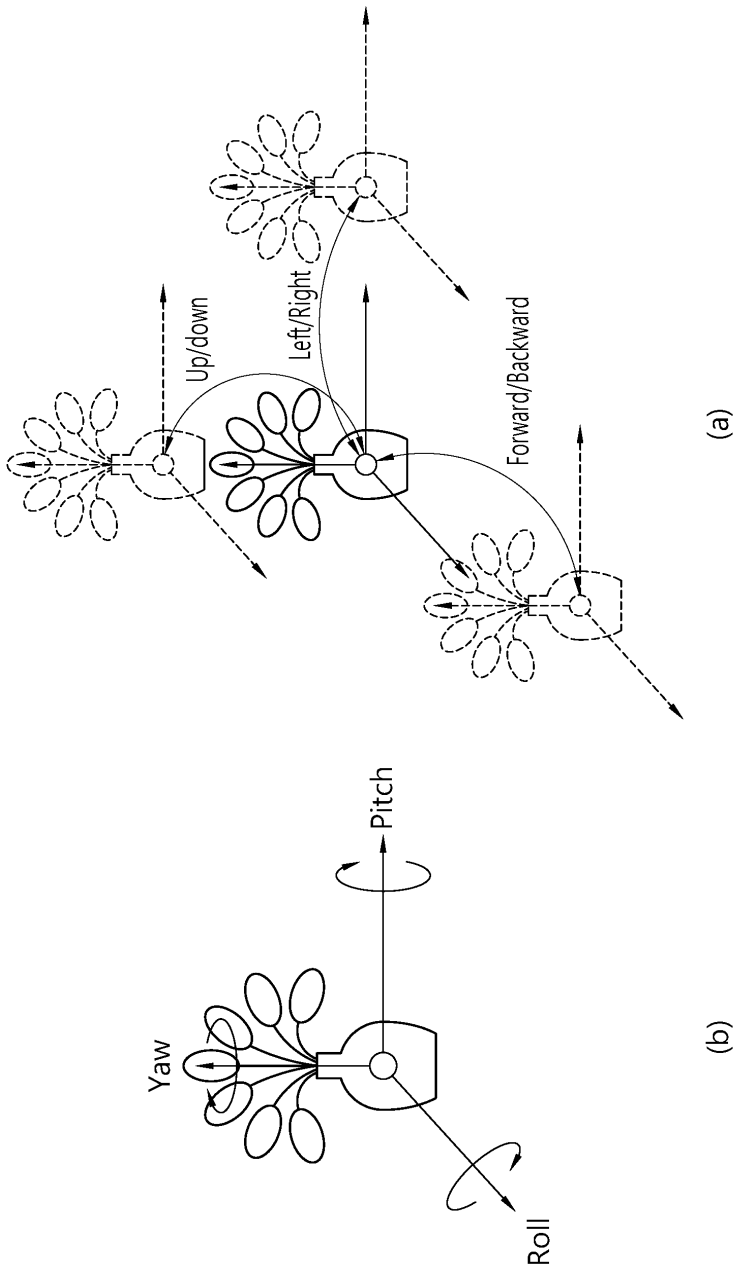


도면15

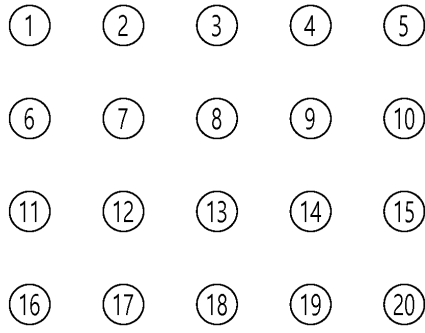




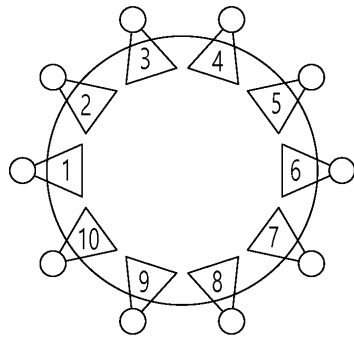
도면17



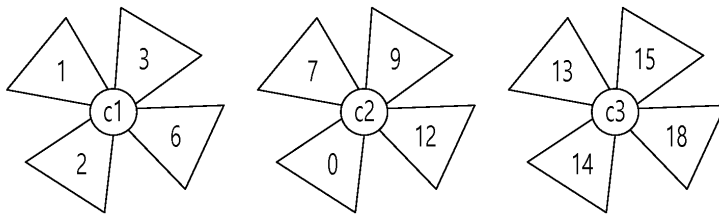
도면18



(a)

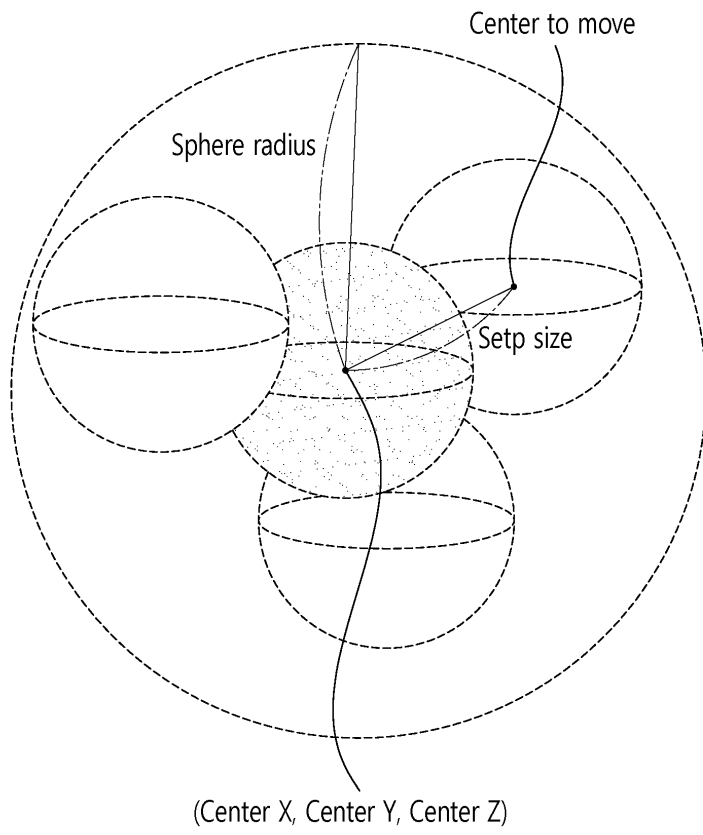


(b)



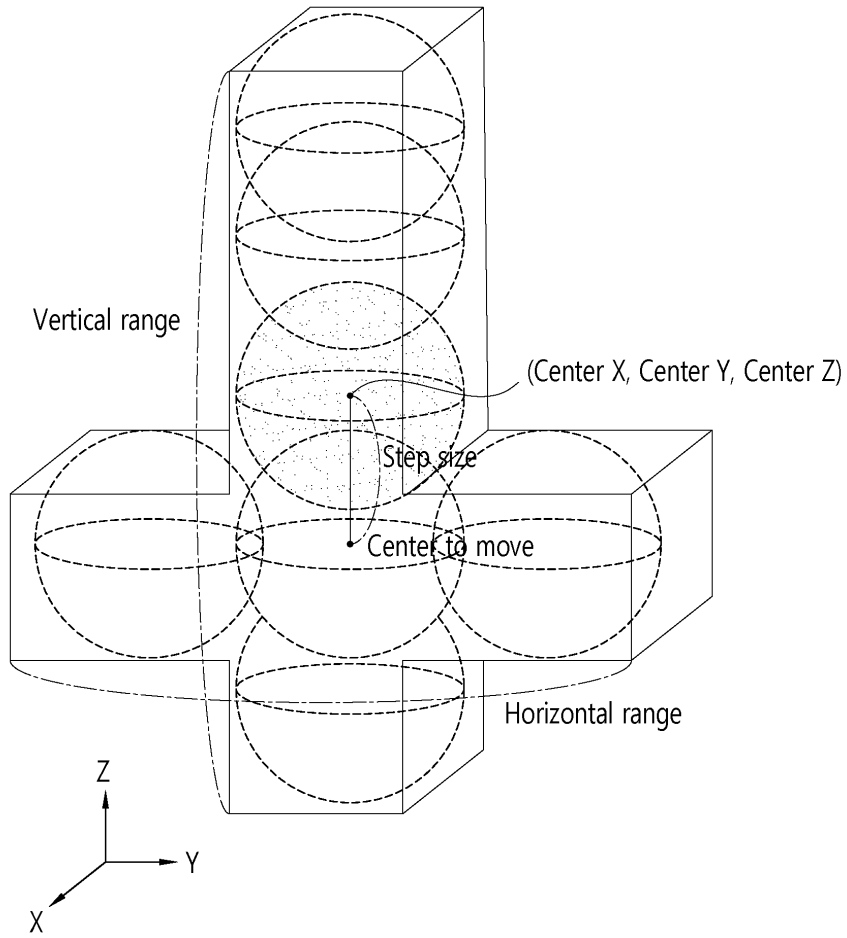
(c)

도면19a

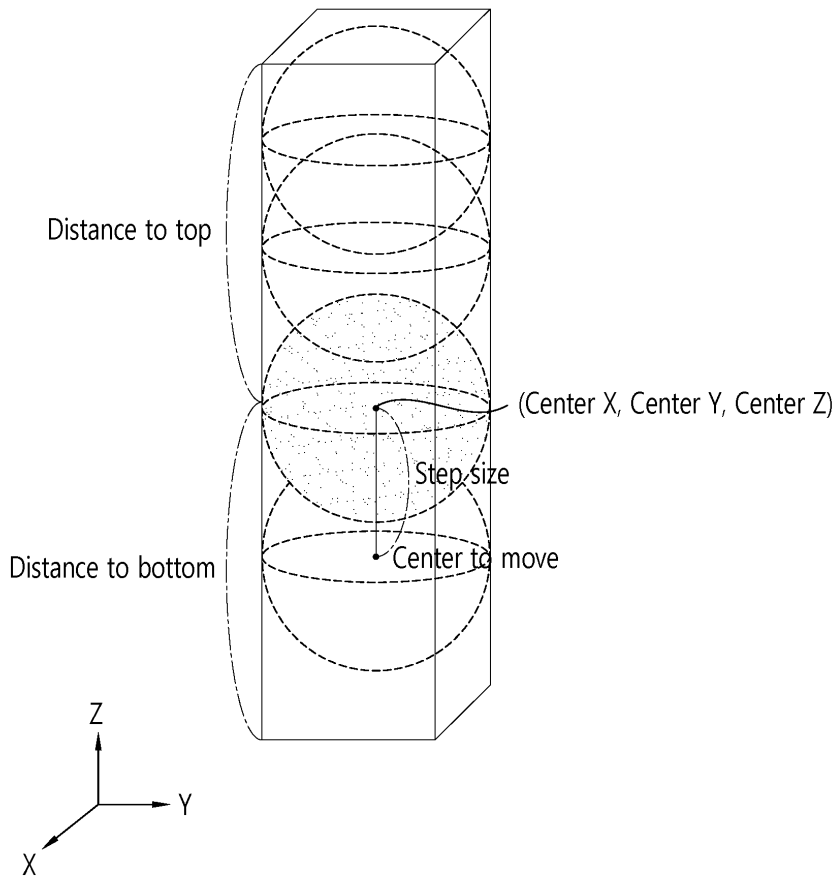




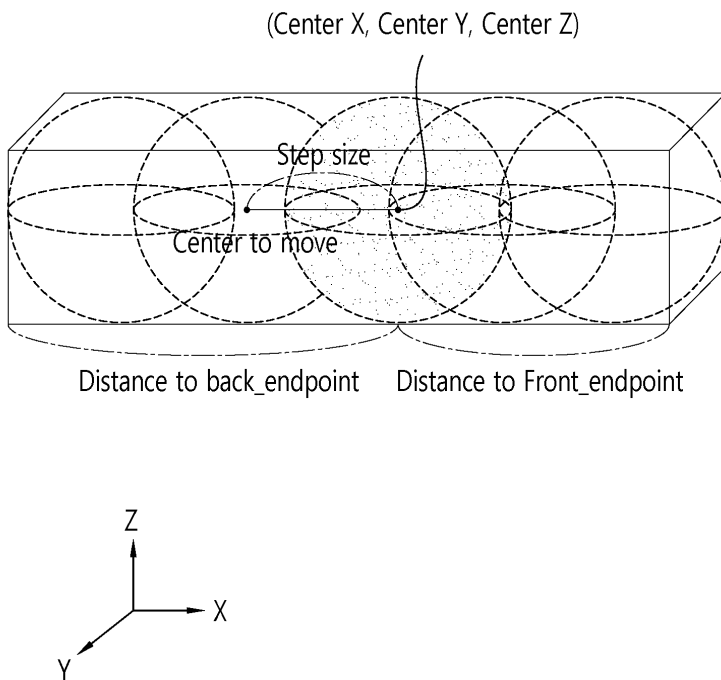
도면19b



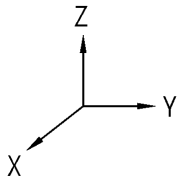
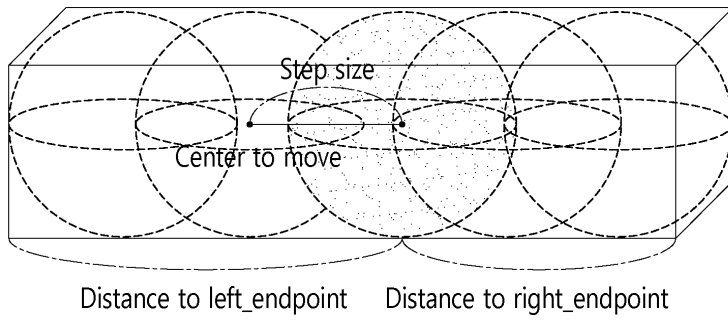
도면19c



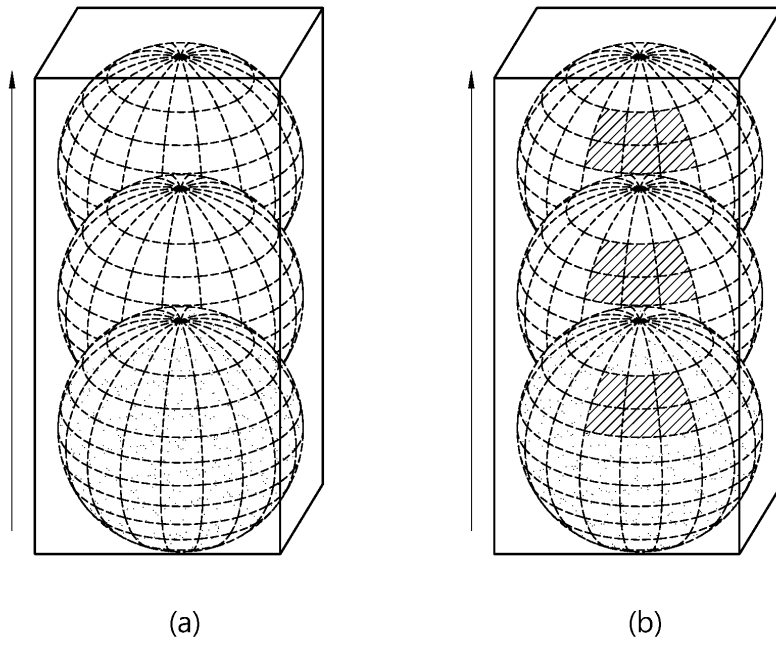
도면19d



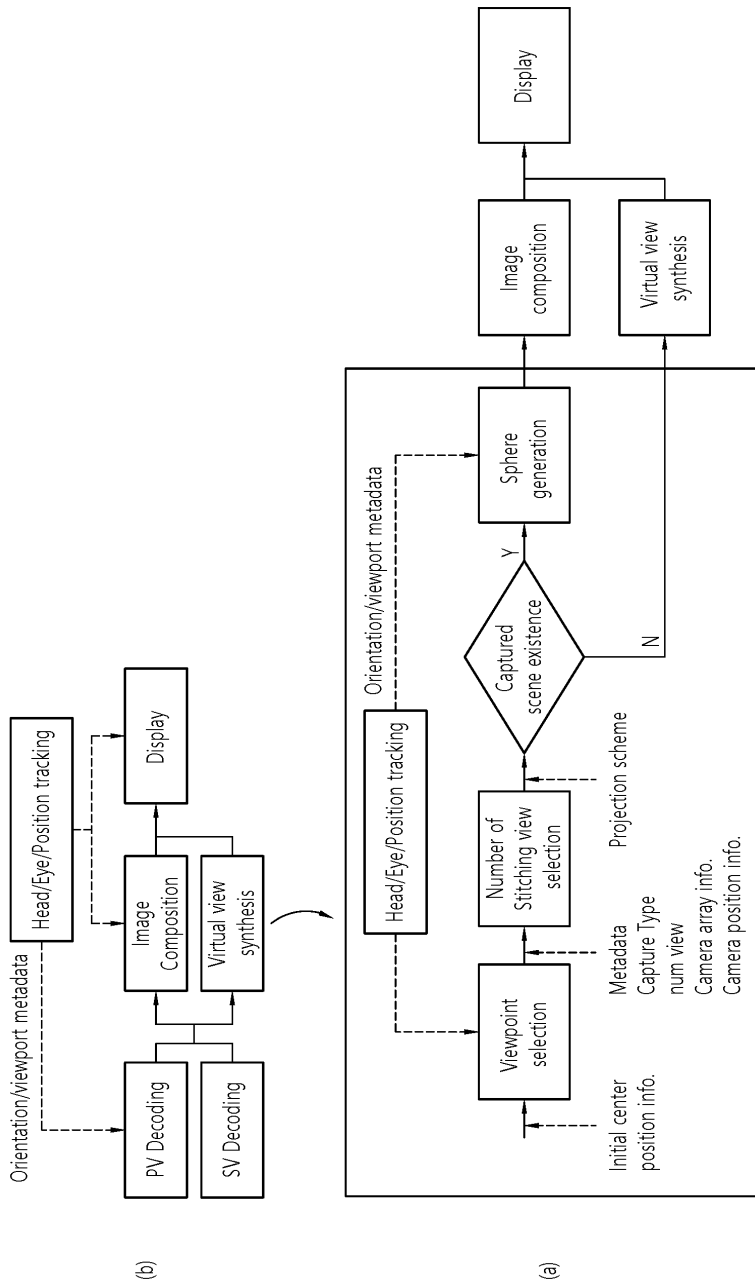
도면19e



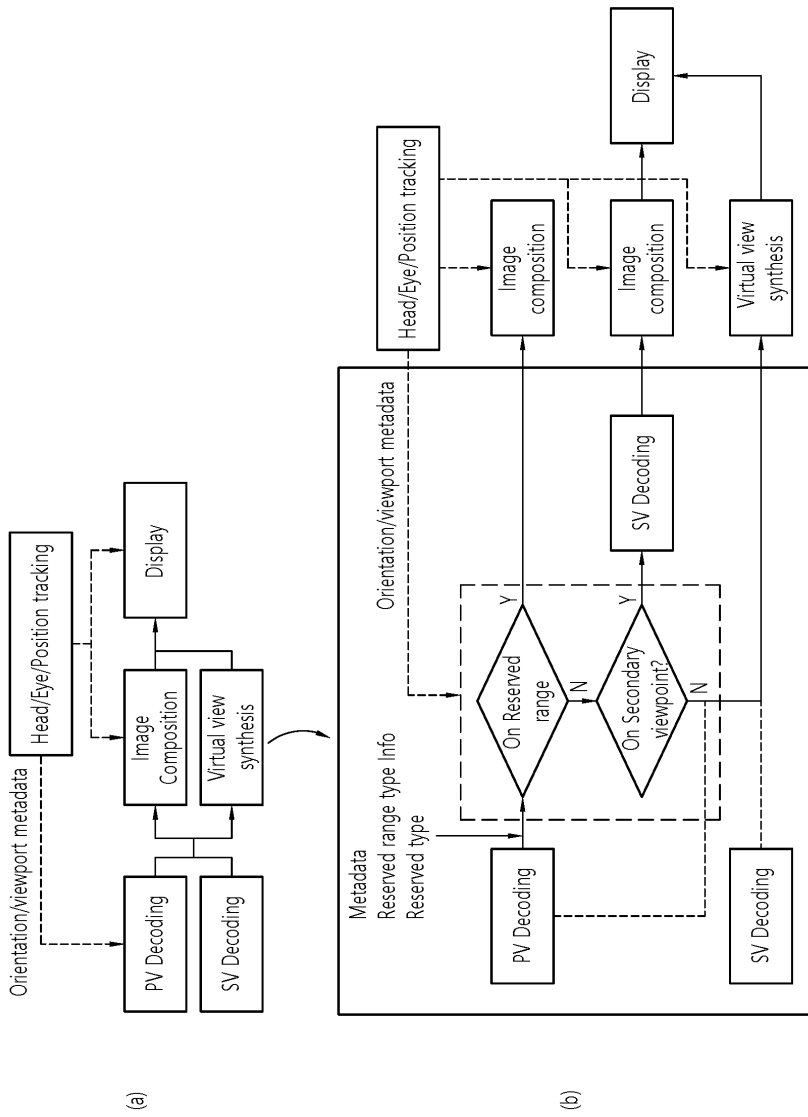
도면20



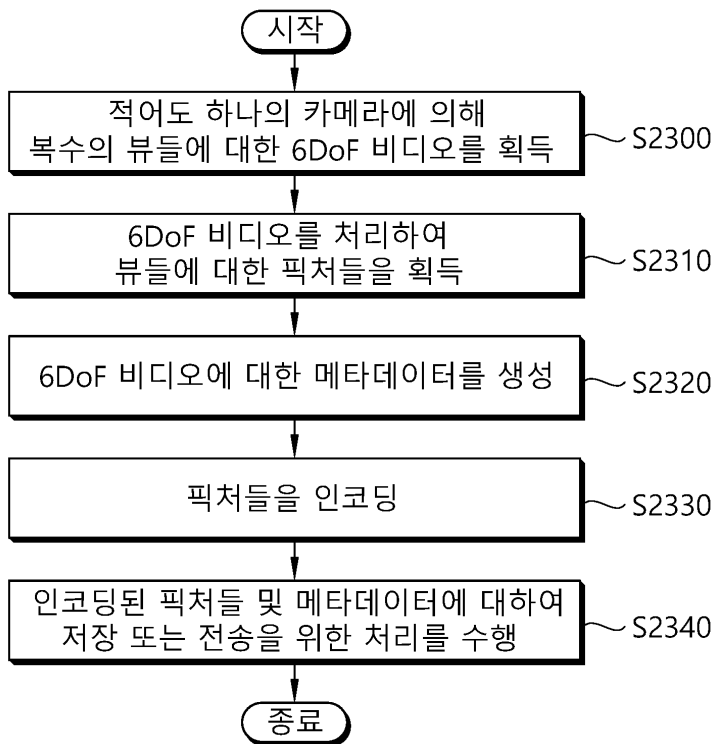
도면21



도면22



도면23



도면24

