



(19) 대한민국특허청(KR)

(12) 등록특허공보(B1)

(45) 공고일자 2024년11월28일

(11) 등록번호 10-2734990

(24) 등록일자 2024년11월22일

(51) 국제특허분류(Int. Cl.)

H04N 13/178 (2018.01) H04N 13/117 (2018.01)

H04N 13/161 (2018.01) H04N 19/46 (2014.01)

H04N 19/597 (2014.01)

(52) CPC특허분류

H04N 13/178 (2018.05)

H04N 13/117 (2018.05)

(21) 출원번호 10-2022-7005458

(22) 출원일자(국제) 2020년07월14일

심사청구일자 2023년07월13일

(85) 번역문제출일자 2022년02월17일

(65) 공개번호 10-2022-0035229

(43) 공개일자 2022년03월21일

(86) 국제출원번호 PCT/US2020/041878

(87) 국제공개번호 WO 2021/015982

국제공개일자 2021년01월28일

(30) 우선권주장

19305968.0 2019년07월22일

유럽특허청(EPO)(EP)

19306721.2 2019년12월20일

유럽특허청(EPO)(EP)

(56) 선행기술조사문헌

W02019055389 A1*

*는 심사관에 의하여 인용된 문헌

(73) 특허권자

인터디지털 브이씨 홀딩스 인코포레이티드

미국 19809 델라웨어주 윌밍턴 스위트 300 벨뷰
파크웨이 200

(72) 발명자

쉬쎌, 베르트랑

프랑스 35576 쉼쉴 세비네 아브뉴 데 상 블랑 975

브리앙, 제라르

프랑스 35576 쉼쉴 세비네 아브뉴 데 상 블랑 975

도레, 르노

프랑스, F-35000 렌느, 튀 이브 마이옥 7

(74) 대리인

양영준, 김연송, 백만기

전체 청구항 수 : 총 22 항

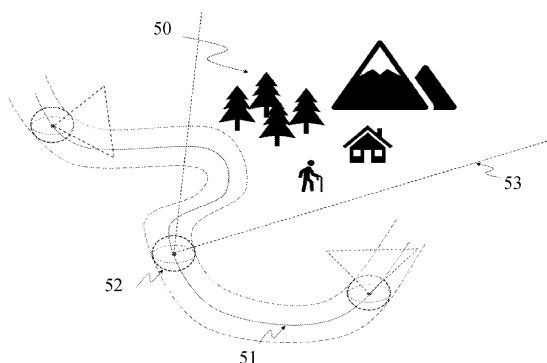
심사관 : 옥윤철

(54) 발명의 명칭 볼류메트릭 비디오 콘텐츠를 전달하기 위한 방법 및 장치

(57) 요약

볼류메트릭 비디오에서의 탐색의 제약들을 표현하는 정보를 시그널링 및 디코딩하기 위한 방법들, 디바이스들 및 데이터 스트림이 제공된다. 데이터 스트림은 볼류메트릭 비디오를 표현하는 비디오 데이터에 연관된 메타데이터를 포함한다. 메타데이터는 뷰잉 경계 상자를 표현하는 데이터, 상기 볼류메트릭 비디오의 3D 공간에서의 곡선 경로를 표현하는 데이터, 및 곡선 경로 상의 포인트에 연관된 적어도 하나의 뷰잉 방향 범위를 표현하는 데이터를 포함한다.

대표도 - 도5



(52) CPC특허분류

H04N 13/161 (2018.05)

H04N 19/46 (2015.01)

H04N 19/597 (2015.01)

명세서

청구범위

청구항 1

볼류메트릭 비디오(volumetric video)에서의 탐색(navigation)의 제약들을 표현하는 정보를 시그널링하기 위한 방법으로서,

상기 볼류메트릭 비디오를 표현하는 비디오 데이터를 포함하는 데이터 스트림 내에서 메타데이터를 인코딩하는 단계

를 포함하고, 상기 메타데이터는:

- 상기 볼류메트릭 비디오의 3차원(3D) 공간에서의 곡선 경로(curvilinear path)를 표현하는 데이터;
- 뷰잉 방향 범위에 연관된 상기 곡선 경로 상의 적어도 2개의 3D 시점들을 표현하는 데이터; 및
- 상기 적어도 2개의 3D 시점들 각각에 대해, 해당 3D 시점을 중심으로 하는 뷰잉 경계 상자를 표현하는 데이터를 포함하는, 방법.

청구항 2

제1항에 있어서, 상기 곡선 경로를 표현하는 데이터는 3D 포인트들의 목록을 포함하고, 상기 목록의 포인트는 적어도 하나의 뷰잉 방향 범위에 연관되는, 방법.

청구항 3

제1항에 있어서, 상기 곡선 경로를 표현하는 데이터는 파라메트릭 3D 곡선들을 표현하는 파라미터들을 포함하고, 상기 적어도 2개의 3D 시점들을 표현하는 데이터는 상기 곡선 경로의 하나의 원점에 연관되는, 방법.

청구항 4

제1항 내지 제3항 중 어느 한 항에 있어서, 상기 메타데이터는 상기 볼류메트릭 비디오의 3D 공간에서의 적어도 2개의 곡선 경로를 표현하는 데이터를 포함하는, 방법.

청구항 5

제1항에 있어서, 상기 곡선 경로를 표현하는 데이터는 적어도 하나의 뷰잉 방향 범위에 연관된 카메라 파라미터들에 대한 참조들의 목록을 포함하는, 방법.

청구항 6

볼류메트릭 비디오에서의 탐색의 제약들을 표현하는 정보를 시그널링하기 위한 디바이스로서,

상기 볼류메트릭 비디오를 표현하는 비디오 데이터를 포함하는 데이터 스트림 내에서 메타데이터를 인코딩하도록 구성되는 프로세서

를 포함하고, 상기 메타데이터는:

- 상기 볼류메트릭 비디오의 3D 공간에서의 곡선 경로를 표현하는 데이터;
- 뷰잉 방향 범위에 연관된 상기 곡선 경로 상의 적어도 2개의 3D 시점들을 표현하는 데이터; 및
- 상기 적어도 2개의 3D 시점들 각각에 대해, 해당 3D 시점을 중심으로 하는 뷰잉 경계 상자를 표현하는 데이터를 포함하는, 디바이스.

청구항 7

제6항에 있어서, 상기 곡선 경로를 표현하는 데이터는 3D 포인트들의 목록을 포함하고, 상기 목록의 포인트는

적어도 하나의 뷰잉 방향 범위에 연관되는, 디바이스.

청구항 8

제6항에 있어서, 상기 곡선 경로를 표현하는 데이터는 파라메트릭 3D 곡선들을 표현하는 파라미터들을 포함하고, 상기 적어도 2개의 3D 시점들을 표현하는 데이터는 상기 곡선 경로의 하나의 원점에 연관되는, 디바이스.

청구항 9

제6항 내지 제8항 중 어느 한 항에 있어서, 상기 메타데이터는 상기 볼류메트릭 비디오의 3D 공간에서의 적어도 2개의 곡선 경로를 표현하는 데이터를 포함하는, 디바이스.

청구항 10

제6항에 있어서, 상기 곡선 경로를 표현하는 데이터는 적어도 하나의 뷰잉 방향 범위에 연관된 카메라 파라미터들에 대한 참조들의 목록을 포함하는, 디바이스.

청구항 11

볼류메트릭 비디오에서의 탐색의 제약들을 표현하는 정보를 디코딩하기 위한 방법으로서,

상기 볼류메트릭 비디오를 표현하는 비디오 데이터를 포함하는 데이터 스트림으로부터 메타데이터를 디코딩하는 단계

를 포함하고, 상기 메타데이터는:

- 상기 볼류메트릭 비디오의 3D 공간에서의 곡선 경로를 표현하는 데이터;
- 뷰잉 방향 범위에 연관된 상기 곡선 경로 상의 적어도 2개의 3D 시점들을 표현하는 데이터; 및
- 상기 적어도 2개의 3D 시점들 각각에 대해, 해당 3D 시점을 중심으로 하는 뷰잉 경계 상자를 표현하는 데이터를 포함하는, 방법.

청구항 12

제11항에 있어서, 상기 곡선 경로를 표현하는 데이터는 3D 포인트들의 목록을 포함하고, 상기 목록의 포인트는 적어도 하나의 뷰잉 방향 범위에 연관되는, 방법.

청구항 13

제11항에 있어서, 상기 곡선 경로를 표현하는 데이터는 파라메트릭 3D 곡선들을 표현하는 파라미터들을 포함하고, 상기 적어도 2개의 3D 시점들을 표현하는 데이터는 상기 곡선 경로의 하나의 원점에 연관되는, 방법.

청구항 14

제11항 내지 제13항 중 어느 한 항에 있어서, 상기 메타데이터는 상기 볼류메트릭 비디오의 3D 공간에서의 적어도 2개의 곡선 경로를 표현하는 데이터를 포함하는, 방법.

청구항 15

제11항에 있어서, 상기 곡선 경로를 표현하는 데이터는 적어도 하나의 뷰잉 방향 범위에 연관된 카메라 파라미터들에 대한 참조들의 목록을 포함하는, 방법.

청구항 16

볼류메트릭 비디오에서의 탐색의 제약들을 표현하는 정보를 디코딩하기 위한 디바이스로서,

상기 볼류메트릭 비디오를 표현하는 비디오 데이터를 포함하는 데이터 스트림 내에서 메타데이터를 디코딩하도록 구성되는 프로세서

를 포함하고, 상기 메타데이터는:

- 상기 볼류메트릭 비디오의 3D 공간에서의 곡선 경로를 표현하는 데이터;
- 뷰잉 방향 범위에 연관된 상기 곡선 경로 상의 적어도 2개의 3D 시점들을 표현하는 데이터; 및
- 상기 적어도 2개의 3D 시점들 각각에 대해, 해당 3D 시점을 중심으로 하는 뷰잉 경계 상자를 표현하는 데이터를 포함하는, 디바이스.

청구항 17

제16항에 있어서, 상기 곡선 경로를 표현하는 데이터는 3D 포인트들의 목록을 포함하고, 상기 목록의 포인트는 적어도 하나의 뷰잉 방향 범위에 연관되는, 디바이스.

청구항 18

제16항에 있어서, 상기 곡선 경로를 표현하는 데이터는 파라메트릭 3D 곡선들을 표현하는 파라미터들을 포함하고, 상기 적어도 2개의 3D 시점들을 표현하는 데이터는 상기 곡선 경로의 하나의 원점에 연관되는, 디바이스.

청구항 19

제16항 내지 제18항 중 어느 한 항에 있어서, 상기 메타데이터는 상기 볼류메트릭 비디오의 3D 공간에서의 적어도 2개의 곡선 경로를 표현하는 데이터를 포함하는, 디바이스.

청구항 20

제16항에 있어서, 상기 곡선 경로를 표현하는 데이터는 적어도 하나의 뷰잉 방향 범위에 연관된 카메라 파라미터들에 대한 참조들의 목록을 포함하는, 디바이스.

청구항 21

기록 매체에 저장된 컴퓨터 프로그램으로서,

프로그램이 컴퓨터에 의해 실행될 때, 상기 컴퓨터로 하여금 제1항의 방법을 수행하게 하는 명령어들을 포함하는, 기록 매체에 저장된 컴퓨터 프로그램.

청구항 22

기록 매체에 저장된 컴퓨터 프로그램으로서,

프로그램이 컴퓨터에 의해 실행될 때, 상기 컴퓨터로 하여금 제11항의 방법을 수행하게 하는 명령어들을 포함하는, 기록 매체에 저장된 컴퓨터 프로그램.

청구항 23

삭제

청구항 24

삭제

청구항 25

삭제

청구항 26

삭제

청구항 27

삭제

청구항 28

삭제

발명의 설명

기술 분야

- [0001] 본 원리들은 일반적으로 3차원(3D) 장면 및 볼류메트릭 비디오 콘텐츠(volumetric video content)의 영역에 관한 것이다. 본 문서는 또한 모바일 디바이스들 또는 헤드 마운트 디스플레이들(Head-Mounted Displays)(HMD)과 같은 최종 사용자 디바이스들에서 볼류메트릭 콘텐츠를 렌더링하기 위한 3D 장면의 지오메트리 및 텍스처를 나타내는 데이터의 인코딩, 포매팅 및 디코딩의 맥락에서 이해된다. 특히, 본 원리들은 볼류메트릭 비디오에서의 탐색(navigation)의 제약들을 나타내는 정보를 시그널링 및 디코딩하는 것에 관한 것이다.

배경 기술

- [0002] 본 섹션은 아래에 설명 및/또는 청구되는 본 원리들의 다양한 양태들에 관련될 수 있는 본 기술분야의 다양한 양태들을 독자에게 소개하도록 의도된 것이다. 이러한 논의는 본 원리들의 다양한 양태들에 대한 더 나은 이해를 촉진하기 위해 독자에게 배경 정보를 제공하는 데 도움이 되는 것으로 여겨진다. 따라서, 이러한 진술들은 이러한 관점에서 읽어야 하며 종래 기술을 인정하는 것이 아님을 이해해야 한다.
- [0003] 최근, 이용가능한 큰 시야 콘텐츠(large field-of-view content)(최대 360°)가 성장했다. 이러한 콘텐츠는 헤드 마운트 디스플레이들(Head Mounted Displays), 스마트 글래스, PC 화면들, 태블릿들, 스마트폰들 및 그와 유사한 것과 같은 몰입형 디스플레이 디바이스들에서 콘텐츠를 보는 사용자가 완전히 볼 수 없을 가능성이 있다. 이는 주어진 순간에 사용자가 콘텐츠의 일부만 보고 있을 수 있음을 의미한다. 그러나, 사용자는 전형적으로 머리 움직임, 마우스 움직임, 터치 스크린, 음성 및 그와 유사한 것과 같은 다양한 수단에 의해 콘텐츠 내에서 탐색할 수 있다. 전형적으로, 이러한 콘텐츠를 인코딩 및 디코딩하는 것이 바람직하다.
- [0004] 360° 평면 비디오라고도 지칭되는 몰입형 비디오는 사용자가 정지된 시점(still point of view)을 중심으로 하는 머리의 회전들을 통해 자신 주변의 모든 것을 보는 것을 허용한다. 회전들은 3 자유도(3 Degrees of Freedom)(3DoF) 경험만을 허용한다. 3DoF 비디오가 예를 들어 헤드 마운트 디스플레이 디바이스(HMD)를 사용하는 최초의 전방향 비디오 경험에는 충분하긴 하지만, 3DoF 비디오는 예를 들어 시차(parallax)를 경험함으로써, 더 많은 자유를 기대하는 뷰어에게 빠르게 좌절감을 줄 수 있다. 추가로, 사용자는 머리를 회전할 뿐만 아니라 머리를 3방향으로 직진이동(translate)시키며, 이러한 직진이동은 3DoF 비디오 경험에서 재현되지 않기 때문에, 3DoF는 또한 어지러움을 유발할 수도 있다.
- [0005] 큰 시야 콘텐츠는 무엇보다도 3차원 컴퓨터 그래픽 이미지 장면(3D CGI 장면), 포인트 클라우드, 또는 몰입형 비디오일 수 있다. 그러한 몰입형 비디오들을 설계하기 위해, 예를 들어 가상 현실(VR), 360, 파노라마, 4π 스테라디안(steradian), 몰입형, 전방향성, 또는 큰 시야와 같은 많은 용어들이 사용될 수 있다.
- [0006] 볼류메트릭 비디오(6자유도(6DoF) 비디오로도 알려짐)는 3DoF 비디오의 대안이다. 6DoF 비디오를 볼 때, 사용자는 회전들에 더하여, 보고 있는 콘텐츠 내에서 자신의 머리, 그리고 심지어는 자신의 신체까지 직진이동시킬 수 있으며, 시차, 그리고 심지어는 볼륨들도 경험할 수 있다. 이러한 비디오들은 머리의 직진이동들 동안 일관된 시각적 피드백을 제공하여 몰입감과 장면 깊이에 대한 인식을 크게 증가시키고 어지러움을 방지한다. 콘텐츠는 관심 장면의 컬러 및 깊이의 동시 기록을 허용하는 전용 센서들을 통해 생성된다. 사진 측량 기술들과 결합된 컬러 카메라 장비의 사용은 기술적인 어려움이 남아 있긴 하지만 그러한 기록을 수행하는 방법이다.
- [0007] 3DoF 비디오들은 텍스처 이미지들(예를 들어, 위도/경도 프로젝션 매핑(latitude/longitude projection mapping) 또는 등장방형 프로젝션 매핑(equirectangular projection mapping)에 따라 인코딩된 구형 이미지들(spherical images))을 역매핑(un-mapping)한 결과인 이미지들의 시퀀스를 포함하는 반면, 6DoF 비디오 프레임들은 여러 시점들로부터의 정보를 내장한다. 그들은 3차원 캡처의 결과인 일시적인 일련의 포인트 클라우드로 볼 수 있다. 뷰잉 조건들에 의존하여 2가지 종류의 볼류메트릭 비디오들이 고려될 수 있다. 첫 번째 종류(즉, 완전 6DoF)는 비디오 콘텐츠 내에서 완전히 자유로운 탐색을 허용하는 반면, 두 번째 종류(일명 3DoF+)는 사용자 뷰잉 공간을 뷰잉 경계 상자라고 지칭되는 제한된 볼륨으로 제약하여, 제한적인 머리 직진이동 및 시차 경험을 허용한다. 이러한 두 번째 맥락은 자유로운 탐색과 앉아 있는 청중의 수동적 뷰잉 조건 사이의 가치 있는 절충안이다. 3DoF+와 6DoF 경험 사이에서 4DoF+ 사례를 3DoF+와 6DoF 사이의 중간으로 정의할 수 있고, 여기서 사용자의 변위는 3D 장면 내에서 곡선(1D) 경로를 따라 제한되며, 각각의 경로 샘플 주위의 수평, 수직 및 깊이

직진이동이 제한된다. 사용자는 양호한 시각적 품질을 보장하면서 경로를 따라 일종의 터널 내를 이동할 수 있다. 사용자가 자신의 가상 카메라를 터널 외부로 이동시키면, 데이터는 양호한 품질의 3D 장면의 재구성을 잃을 것이다.

[0008] 사용자의 가상 카메라에 대해 허용된 위치들의 범위와 주어진 볼류메트릭 비디오 콘텐츠의 탐색 동안의 뷰잉 배향들이 제약되어야 할 수 있다. 그렇지 않으면, 사용자는 이용가능한 시각적 데이터의 부족으로 인해 완전히 렌더링할 수 없는 뷰포트들을 요청할 때 3D 장면을 "떠날" 수 있다. 따라서, 주어진 볼류메트릭 비디오 콘텐츠를 소비할 때 최종 사용자 디바이스에게 적합한, 선호되는 및/또는 허용된 뷰어 (가상) 위치들 및 뷰잉 배향들을 알려주기 위한 솔루션이 필요하다.

발명의 내용

[0009] 이하는 본 원리들의 일부 양태들에 대한 기본적인 이해를 제공하기 위해 본 원리들의 간략한 요약의 제시한다. 이 요약은 본 원리에 대한 광범위한 개요가 아니다. 그것은 본 원리들의 핵심적인 또는 중요한 요소들을 식별하도록 의도된 것이 아니다. 이하의 요약은 아래에 제공된 보다 상세한 설명에 대한 서문으로서 본 원리의 일부 양태들을 단순화된 형태로 제시할 뿐이다.

[0010] 본 원리들은 볼류메트릭 비디오에서의 탐색의 제약들을 표현하는 정보를 시그널링하기 위한 방법 및 디바이스에 관한 것이다. 방법은 상기 볼류메트릭 비디오를 표현하는 비디오 데이터를 포함하는 데이터 스트림 내에서 메타데이터를 인코딩하는 단계를 포함한다. 메타데이터는 이하를 포함한다:

[0011] - 뷰잉 경계 상자를 표현하는 데이터;

[0012] - 상기 볼류메트릭 비디오의 3D 공간에서의 곡선 경로를 표현하는 데이터; 및

[0013] - 상기 곡선 경로 상의 포인트에 연관된 적어도 하나의 뷰잉 방향 범위를 표현하는 데이터.

[0014] 다른 양태에 따르면, 볼류메트릭 비디오에서의 탐색의 제약들을 표현하는 정보를 디코딩하기 위한 제2 방법 및 제2 장치가 제공된다. 제2 방법은 볼류메트릭 비디오를 표현하는 비디오 데이터를 포함하는 데이터 스트림으로부터 메타데이터를 디코딩하는 단계를 포함한다. 메타데이터는 이하를 포함한다:

[0015] - 뷰잉 경계 상자를 표현하는 데이터;

[0016] - 상기 볼류메트릭 비디오의 3D 공간에서의 곡선 경로를 표현하는 데이터; 및

[0017] - 상기 곡선 경로 상의 포인트에 연관된 적어도 하나의 뷰잉 방향 범위를 표현하는 데이터.

[0018] 적어도 하나의 실시예의 다른 일반적인 양태에 따르면, 설명된 인코딩 실시예들 또는 변형들 중 임의의 것에 따라 생성된 비디오 데이터 및 연관된 메타데이터를 포함하는 데이터 스트림이 제공된다.

[0019] 적어도 하나의 실시예의 다른 일반적인 양태에 따르면, 설명된 인코딩 실시예들 또는 변형들 중 임의의 것에 따라 생성된 데이터 콘텐츠를 포함하는 비-일시적 컴퓨터 판독가능한 매체가 제공된다.

[0020] 적어도 하나의 실시예의 다른 일반적인 양태에 따르면, 프로그램이 컴퓨터에 의해 실행될 때 컴퓨터로 하여금 설명된 디코딩 실시예들 또는 변형들 중 임의의 것을 수행하게 하는 명령어들을 포함하는 컴퓨터 프로그램 제품이 제공된다.

도면의 간단한 설명

[0021] 이하의 설명을 읽을 때 본 개시내용이 더 잘 이해될 것이고, 다른 특정한 특징들 및 이점들이 드러날 것이며, 그 설명은 첨부된 도면들을 참조한다:

도 1은 본 원리들의 비제한적인 실시예에 따라 객체의 3차원(3D) 모델, 및 3D 모델에 대응하는 포인트 클라우드의 포인트들을 도시한다.

도 2는 본 원리들의 비제한적인 실시예에 따라 3D 장면들의 시퀀스를 표현하는 데이터의 인코딩, 전송 및 디코딩의 비제한적인 예를 도시한다.

도 3은 본 원리들의 비제한적인 실시예에 따라, 도 7 및 도 8과 관련하여 설명된 방법을 구현하도록 구성될 수 있는 디바이스의 예시적인 아키텍처를 도시한다.

도 4는 본 원리들의 비제한적인 실시예에 따라, 데이터가 패킷 기반 전송 프로토콜을 통해 전송될 때의 스트림

의 선택스의 실시예의 예를 도시한다.

도 5는 본 원리들의 비제한적인 실시예에 따라, 하나의 곡선 경로, 및 이러한 경로를 따른 주어진 위치에서의 구형 경계 상자 및 뷰잉 배향 범위를 묘사함으로써 가상 3D 장면 내에서의 제약된 탐색의 개념을 개략적으로 도시한다.

도 5b는 본 원리들의 비제한적인 실시예에 따라, 취득 카메라들의 위치들 및 방향들로부터 추론된 관형(tuboid) 뷰잉 공간을 갖는, 원호 상에 배치된 5개의 수렴 카메라를 갖는 취득 카메라 장비의 예를 도시한다.

도 6은 본 원리들의 비제한적인 실시예에 따라, 중심 위치에 관심 객체들을 갖는 3D 장면 주위를 탐색하기 위한 원형 경로의 특정 사례를 예시한다.

도 7은 본 원리들의 비제한적인 실시예에 따라, 볼류메트릭 비디오에서의 탐색의 제약들을 표현하는 정보를 시그널링하기 위한 방법(70)을 도식적으로 도시한다.

도 8은 본 원리들의 비제한적인 실시예에 따라, 볼류메트릭 비디오에서의 탐색의 제약들을 표현하는 정보를 디코딩하기 위한 방법(80)을 도식적으로 도시한다.

발명을 실시하기 위한 구체적인 내용

- [0022] 본 원리들은 본 원리들의 예들이 도시되어 있는 첨부 도면들을 참조하여 이하에서 더욱 완전하게 설명될 것이다. 그러나, 본 원리들은 다수의 대안적인 형태들로 구현될 수 있으며, 여기에 제시된 예들에 제한되는 것으로 해석되어서는 안 된다. 따라서, 본 원리들은 다양한 변형들 및 대안적인 형태들이 가능하지만, 그 구체적인 예들이 도면들에 예시적으로 도시되어 있고 여기에서 상세히 설명될 것이다. 그러나, 본 원리들을 개시된 특정 형태들로 제한하려는 의도는 없으며, 반대로 본 개시내용은 청구항들에 의해 정의된 본 원리의 사상 및 범위 내에 속하는 모든 수정, 균등물 및 대안을 포괄하는 것임을 이해해야 한다.
- [0023] 본 명세서에서 사용되는 용어는 단지 특정한 예들을 설명하기 위한 것으로, 본 원리들을 한정하려는 의도가 아니다. 본 명세서에서 사용될 때, 단수 형태들 "a", "an" 및 "the"는 문맥이 명백하게 달리 나타내지 않는 한 복수 형태들도 포함하는 것으로 의도된다. 본 명세서에서 사용될 때, 용어들 "포함한다(comprises)", "포함하는(comprising)", "포함한다(includes)" 및/또는 "포함하는(including)"은 언급된 특징들, 정수들, 단계들, 동작들, 요소들 및/또는 컴포넌트들의 존재를 명시하지만, 하나 이상의 다른 특징들, 정수들, 단계들, 동작들, 요소들, 컴포넌트들 및/또는 이들의 그룹들의 존재 또는 추가를 배제하지 않음이 추가로 이해될 것이다. 더욱이, 요소가 다른 요소에 대해 "응답하는" 또는 "연결되는" 것으로 언급되는 경우, 그것은 다른 요소에 직접 응답하거나 연결될 수도 있고, 또는 중간 요소들이 존재할 수도 있다. 대조적으로, 요소가 다른 요소에 "직접 응답"하거나 "직접 연결"되는 것으로 언급되는 경우에는 중간 요소들이 존재하지 않는다. 본 명세서에 사용될 때, 용어 "및/또는"은 연관된 나열된 항목들 중 하나 이상의 항목의 임의의 및 모든 조합을 포함하고, "/"로 약칭될 수 있다.
- [0024] 본 명세서에서 다양한 요소들을 설명하기 위해 용어들 제1, 제2 등이 사용될 수 있지만, 이러한 요소들은 이러한 용어들에 의해 제한되어서는 안된다는 것이 이해될 것이다. 이러한 용어들은 하나의 요소를 다른 요소로부터 구별하기 위해서만 사용된다. 예를 들어, 본 원리들의 교시를 벗어나지 않으면서 제1 요소는 제2 요소로 명명될 수 있고, 마찬가지로 제2 요소도 제1 요소로 명명될 수 있다.
- [0025] 도면들 중 일부는 통신의 주요 방향을 보여주기 위해 통신 경로들 상에 화살표들을 포함하지만, 통신은 도시된 화살표들과 반대 방향으로 발생할 수 있음을 이해해야 한다.
- [0026] 일부 예들은 각각의 블록이 특정 논리 기능(들)을 구현하기 위한 하나 이상의 실행가능한 명령어를 포함하는 회로 요소, 모듈, 또는 코드 부분을 표현하는 블록도들 및 동작 흐름도들에 관련하여 설명된다. 다른 구현들에서, 블록들에서 언급된 기능(들)은 언급된 순서와 다르게 발생할 수 있다는 점에 또한 주목해야 한다. 예를 들어, 관련된 기능에 따라, 연속적으로 보여진 두 개의 블록이 실제로 실질적으로 동시에 실행되거나 블록들이 때때로 역순으로 실행될 수 있다.
- [0027] 본 명세서에서 "예에 따라" 또는 "예에서"라는 언급은 예와 관련하여 설명된 특정 특징, 구조 또는 특성이 본 원리들의 적어도 하나의 구현에 포함될 수 있음을 의미한다. 명세서 내의 다양한 위치들에 있는 "예에 따른" 또는 "예에서"라는 문구의 출현들이 반드시 모두 동일한 예를 지칭하는 것은 아니며, 별도의 또는 대안적인 예들이 다른 예들과 반드시 상호 배타적인 것은 아니다.

- [0028] 청구항들에 나타난 참조 번호들은 단지 예시를 위한 것이며 청구항들의 범위를 제한하는 효과를 갖지 않는다. 명시적으로 설명되지는 않았지만, 본 예들 및 변형들은 임의의 조합 또는 하위 조합으로 사용될 수 있다.
- [0029] 도 1은 객체의 3차원(3D) 모델(10), 및 3D 모델(10)에 대응하는 포인트 클라우드(11)의 포인트들을 도시한다. 3D 모델(10) 및 포인트 클라우드(11)는 예를 들어 다른 객체들을 포함하는 3D 장면의 한 객체의 가능한 3D 표현에 대응할 수 있다. 모델(10)은 3D 메쉬 표현일 수 있고, 포인트 클라우드(11)의 포인트들은 메쉬의 정점들(vertices)일 수 있다. 포인트 클라우드(11)의 포인트들은 또한 메쉬의 면들(faces)의 표면에 퍼진 포인트들일 수 있다. 모델(10)은 또한 포인트 클라우드(11)의 스플래팅 버전(splatted version)으로서 표현될 수 있으며, 모델(10)의 표면은 포인트 클라우드(11)의 포인트들을 스플래팅함으로써 생성된다. 모델(10)은 복셀들 또는 스플라인들과 같은 다수의 상이한 표현들로 표현될 수 있다. 도 1은 포인트 클라우드가 3D 객체의 표면 표현으로 정의될 수 있고, 3D 객체의 표면 표현이 클라우드의 포인트로부터 생성될 수 있다는 사실을 도시한다. 본 명세서에 사용될 때, (3D 장면의 확장 포인트들에 의해) 3D 객체의 포인트들을 이미지에 투영하는 것은 이 3D 객체의 임의의 표현, 예를 들어 포인트 클라우드, 메쉬, 스플라인 모델 또는 복셀 모델을 투영하는 것과 등가이다.
- [0030] 포인트 클라우드는 예를 들어 벡터 기반 구조로 메모리에서 표현될 수 있고, 여기서 각각의 포인트는 시점의 참조 프레임에서 그 자신의 고유한 좌표들(예를 들어, 3차원 좌표들 XYZ, 또는 시점(viewpoint)로부터의/로의 입체각 및 거리(깊이라고도 함)), 및 성분이라고도 지칭되는 하나 이상의 속성을 갖는다. 성분의 예는 다양한 컬러 공간들, 예를 들어 RGB(적색, 녹색 및 청색) 또는 YUV(Y는 루마 성분이고 UV는 2개의 색차 성분임)로 표현될 수 있는 컬러 성분이다. 포인트 클라우드는 객체들을 포함하는 3D 장면의 표현이다. 3D 장면은 주어진 시점 또는 시점들의 범위로부터 보여질 수 있다. 포인트 클라우드는 예를 들어 다음과 같은 다수의 방식으로 획득될 수 있다:
- [0031] · 임의적으로(optionally) 깊이 능동 감지 디바이스로 보완되는 카메라 장비에 의해 촬영된 실제 객체의 캡처로부터;
 - [0032] · 모델링 도구의 가상 카메라 장비에 의해 촬영된 가상/합성 객체의 캡처로부터;
 - [0033] · 실제 객체와 가상 객체 둘 다의 혼합으로부터.
- [0034] 도 2는 3D 장면들의 시퀀스를 표현하는 데이터의 인코딩, 전송 및 디코딩의 비제한적인 예를 보여준다. 예를 들어, 동시에, 인코딩 포맷은 3DoF, 3DoF+ 및 6DoF 디코딩과 호환될 수 있다.
- [0035] 3D 장면들(20)의 시퀀스가 획득된다. 픽처들의 시퀀스가 2D 비디오인 것과 마찬가지로, 3D 장면들의 시퀀스는 3D(볼류메트릭이라고도 함) 비디오이다. 3D 장면들의 시퀀스는 3DoF, 3DoF+ 또는 6DoF 렌더링 및 디스플레이를 위해 볼류메트릭 비디오 렌더링 디바이스에 제공될 수 있다.
- [0036] 3D 장면들(20)의 시퀀스는 인코더(21)에 제공된다. 인코더(21)는 하나의 3D 장면 또는 3D 장면들의 시퀀스를 입력으로 취하고, 입력을 표현하는 비트 스트림을 제공한다. 비트 스트림은 메모리(22) 및/또는 전자 데이터 매체에 저장될 수 있고, 네트워크(22)를 통해 전송될 수 있다. 3D 장면들의 시퀀스를 표현하는 비트 스트림은 디코더(23)에 의해 메모리(22)로부터 관독될 수 있고/있거나 네트워크(22)로부터 수신될 수 있다. 디코더(23)는 상기 비트 스트림에 의해 입력되고, 예를 들어 포인트 클라우드 포맷으로 3D 장면들의 시퀀스를 제공한다.
- [0037] 인코더(21)는 여러 단계들을 구현하는 여러 회로들을 포함할 수 있다. 제1 단계에서, 인코더(21)는 각각의 3D 장면을 적어도 하나의 2D 픽처에 투영한다. 3D 투영은 3차원 포인트들을 2차원 평면에 매핑하는 임의의 방법이다. 그래픽 데이터를 디스플레이하기 위한 최신 방법들은 평면(여러 비트 평면들로부터의 픽셀 정보) 2차원 매체에 기초하기 때문에, 이러한 유형의 투영의 사용은 특히 컴퓨터 그래픽, 엔지니어링 및 제도 분야에 널리 보급되어 있다. 투영 회로(211)는 시퀀스(20)의 3D 장면에 대해 적어도 하나의 2차원 프레임(2111)을 제공한다. 프레임(2111)은 프레임(2111)에 투영된 3D 장면을 표현하는 컬러 정보 및 깊이 정보를 포함한다. 변형에서, 컬러 정보 및 깊이 정보는 두 개의 개별 프레임(2111 및 2112)에서 인코딩된다.
- [0038] 메타데이터(212)는 투영 회로(211)에 의해 사용되고 업데이트된다. 메타데이터(212)는 투영 동작에 관한 정보(예를 들어, 투영 파라미터들), 및 도 5 내지 7과 관련하여 설명된 바와 같이 컬러 및 깊이 정보가 프레임들(2111 및 2112) 내에서 조직되는 방식에 관한 정보를 포함한다.
- [0039] 비디오 인코딩 회로(213)는 프레임들(2111, 2112)의 시퀀스를 비디오로 인코딩한다. 3D 장면(2111, 2112)의 픽처들(또는 3D 장면의 픽처들의 시퀀스)은 비디오 인코더(213)에 의해 스트림으로 인코딩된다. 다음으로, 비디오 데이터 및 메타데이터(212)는 데이터 캡슐화 회로(214)에 의해 데이터 스트림에 캡슐화된다.

- [0040] 인코더(213)는 예를 들어 다음과 같은 인코더와 호환된다:
- [0041] - JPEG, 사양 ISO/CEI 10918-1 UIT-T Recommendation T.81, <https://www.itu.int/rec/T-REC-T.81/en>;
- [0042] - MPEG-4 AVC 또는 h264라고도 명명되는 AVC. 모두 UIT-T H.264 및 ISO/CEI MPEG-4 Part 10(ISO/CEI 14496-10), <http://www.itu.int/rec/T-REC-H.264/en>, HEVC에 명시된(그것이 사양은 ITU 웹사이트, T recommendation, H series, h265, <http://www.itu.int/rec/T-REC-H.265-201612-I/en>에서 확인할 수 있음);
- [0043] - 3D-HEVC(ITU 웹사이트, T recommendation, H series, h265, <http://www.itu.int/rec/T-REC-H.265-201612-I/en> annex G 및 I에서 사양을 찾을 수 있는 HEVC의 확장);
- [0044] - Google에 의해 개발된 VP9; 또는
- [0045] - Alliance for Open Media에 의해 개발된 AV1(AOMedia Video 1).
- [0046] 데이터 스트림은 디코더(23)에 의해 예를 들어 네트워크(22)를 통해 액세스 가능한 메모리에 저장된다. 디코더(23)는 디코딩의 상이한 단계들을 구현하는 상이한 회로들을 포함한다. 디코더(23)는 인코더(21)에 의해 생성된 데이터 스트림을 입력으로 취하고, 헤드 마운트 디바이스(HMD)와 같은 볼류메트릭 비디오 디스플레이 디바이스에 의해 렌더링 및 디스플레이될 3D 장면들(24)의 시퀀스를 제공한다. 디코더(23)는 소스(22)로부터 스트림을 획득한다. 예를 들어, 소스(22)는 다음을 포함하는 세트에 속한다:
- [0047] - 로컬 메모리, 예를 들어 비디오 메모리 또는 RAM(또는 랜덤 액세스 메모리), 플래시 메모리, ROM(또는 판독 전용 메모리), 하드 디스크;
- [0048] - 저장 인터페이스, 예를 들어 대용량 저장소, RAM, 플래시 메모리, ROM, 광학 디스크 또는 자기 서포트와의 인터페이스;
- [0049] - 통신 인터페이스, 예를 들어 유선 인터페이스(예를 들어, 버스 인터페이스, 광역 네트워크 인터페이스, 근거리 네트워크 인터페이스) 또는 무선 인터페이스(예를 들어, IEEE 802.11 인터페이스 또는 Bluetooth® 인터페이스); 및
- [0050] - 사용자가 데이터를 입력할 수 있도록 하는 그래픽 사용자 인터페이스와 같은 사용자 인터페이스.
- [0051] 디코더(23)는 데이터 스트림에서 인코딩된 데이터를 추출하기 위한 회로(234)를 포함한다. 회로(234)는 데이터 스트림을 입력으로 취하고, 2차원 비디오 및 스트림에 인코딩된 메타데이터(212)에 대응하는 메타데이터(232)를 제공한다. 비디오는 프레임들의 시퀀스를 제공하는 비디오 디코더(233)에 의해 디코딩된다. 디코딩된 프레임들은 컬러 및 깊이 정보를 포함한다. 변형에서, 비디오 디코더(233)는 2개의 프레임 시퀀스를 제공하는데, 하나는 컬러 정보를 포함하고 다른 하나는 깊이 정보를 포함한다. 회로(231)는 3D 장면들(24)의 시퀀스를 제공하기 위해 디코딩된 프레임들로부터 컬러 및 깊이 정보를 역투영(un-project)하기 위해 메타데이터(232)를 사용한다. 3D 장면들(24)의 시퀀스는 3D 장면들(20)의 시퀀스에 대응하며, 2D 비디오로서의 인코딩 및 비디오 압축에 관련된 정밀도 손실의 가능성이 있다.
- [0052] 도 3은 도 7 및 도 8과 관련하여 설명된 방법을 구현하도록 구성될 수 있는 디바이스(30)의 예시적인 아키텍처를 도시한다. 도 2의 인코더(21) 및/또는 디코더(23)는 이 아키텍처를 구현할 수 있다. 대안적으로, 인코더(21) 및/또는 디코더(23)의 각각의 회로는 예를 들어 그들의 버스(31) 및/또는 I/O 인터페이스(36)를 통해 함께 링크된, 도 3의 아키텍처에 따른 디바이스일 수 있다.
- [0053] 디바이스(30)는 데이터 및 어드레스 버스(31)에 의해 함께 링크된 이하의 요소들을 포함한다:
- [0054] - 예를 들어, DSP(또는 디지털 신호 프로세서)인 마이크로프로세서(32)(또는 CPU);
- [0055] - ROM(또는 판독 전용 메모리)(33);
- [0056] - RAM(또는 랜덤 액세스 메모리)(34);
- [0057] - 저장 인터페이스(35);
- [0058] - 애플리케이션으로부터 전송할 데이터를 수신하기 위한 I/O 인터페이스(36); 및
- [0059] - 전력 공급부, 예를 들어 배터리.
- [0060] 예에 따르면, 전력 공급부는 디바이스의 외부에 있다. 언급된 메모리 각각에서, 본 명세서에서 사용되는 <<레

지스터»라는 단어는 작은 용량(일부 비트들)의 영역 또는 매우 큰 영역(예를 들어, 전체 프로그램, 또는 대량의 수신된 또는 디코딩된 데이터)에 대응할 수 있다. ROM(33)은 적어도 프로그램 및 파라미터들을 포함한다. ROM(33)은 본 원리들에 따라 기술들을 수행하기 위한 알고리즘들 및 명령어들을 저장할 수 있다. 스위치 온될 때, CPU(32)는 RAM 내의 프로그램을 업로드하고, 대응하는 명령어들을 실행한다.

[0061] RAM(34)은 레지스터 내에, 디바이스(30)의 스위치-온 후에 CPU(32)에 의해 실행되고 업로드되는 프로그램, 레지스터 내의 입력 데이터, 레지스터 내의 메모드의 상이한 상태들에서의 중간 데이터, 및 레지스터 내의 메모드의 실행을 위해 사용되는 다른 변수들을 포함한다.

[0062] 본 명세서에 설명된 구현들은 예를 들어 방법 또는 프로세스, 장치, 컴퓨터 프로그램 제품, 데이터 스트림, 또는 신호로 구현될 수 있다. 단일 형태의 구현의 맥락에서만 논의되더라도(예를 들어, 방법 또는 디바이스로서만 논의됨), 논의된 특징들의 구현은 또한 다른 형태들(예를 들어, 프로그램)로 구현될 수 있다. 장치는 예를 들어 적절한 하드웨어, 소프트웨어 및 펌웨어로 구현될 수 있다. 방법들은 예를 들어 컴퓨터, 마이크로프로세서, 집적 회로, 또는 프로그래밍가능한 로직 디바이스를 포함하는, 일반적으로 프로세싱 디바이스들을 지칭하는, 예를 들어 프로세서와 같은 장치에서 구현될 수 있다. 프로세서들은 또한 예를 들어 컴퓨터들, 셀폰들, 휴대용/개인용 디지털 단말기들("PDA"), 및 최종 사용자들 간의 정보 통신을 용이하게 하는 다른 디바이스들과 같은 통신 디바이스들을 포함한다.

[0063] 예들에 따르면, 디바이스(30)는 도 7 및 도 8과 관련하여 설명된 방법을 구현하도록 구성되고, 다음을 포함하는 세트에 속한다:

- [0064] - 모바일 디바이스;
- [0065] - 통신 디바이스;
- [0066] - 게임 디바이스;
- [0067] - 태블릿(또는 태블릿 컴퓨터);
- [0068] - 랩톱;
- [0069] - 정지 사진 카메라;
- [0070] - 비디오 카메라;
- [0071] - 인코딩 칩;
- [0072] - 서버(예를 들어, 방송 서버, 주문형 비디오 서버, 또는 웹 서버).

[0073] 도 4는 데이터가 패킷 기반 전송 프로토콜을 통해 전송될 때의 스트림의 선택의 실시예의 예를 도시한다. 도 4는 볼류메트릭 비디오 스트림의 예시적인 구조(4)를 도시한다. 구조는 스트림을 선택의 독립적인 요소들로 구성하는 컨테이너로 이루어진다. 구조는 스트림의 모든 선택 요소들에 공통적인 데이터 세트인 헤더 부분(41)을 포함할 수 있다. 예를 들어, 헤더 부분은 선택 요소들 각각의 본질과 역할을 설명하는, 선택 요소들에 관한 일부 메타데이터를 포함한다. 헤더 부분은 또한 도 2의 메타데이터(212)의 일부, 예를 들어 3D 장면의 포인트를 프레임(2111 및 2112)에 투영하는 데 사용되는 중심 시점의 좌표를 포함할 수 있다. 구조는 선택스(42)의 요소 및 선택스(43)의 적어도 하나의 요소를 포함하는 페이로드를 포함한다. 선택스 요소(42)는 컬러 및 깊이 프레임들을 표현하는 데이터를 포함한다. 이미지들은 비디오 압축 방법에 따라 이미지가 압축되었을 수 있다.

[0074] 선택스(43)의 요소는 데이터 스트림의 페이로드의 일부이고, 선택스(42)의 요소의 프레임들이 인코딩되는 방법에 관한 메타데이터, 예를 들어 프레임들에 3D 장면의 포인트들을 투영하고 패킹하는 데 사용되는 파라미터들을 포함할 수 있다. 이러한 메타데이터는 비디오의 각각의 프레임, 또는 프레임들의 그룹(비디오 압축 표준들에서는 픽처 그룹(GoP)이라고도 함)에 연관될 수 있다.

[0075] 도 5는 하나의 곡선 경로, 및 이러한 경로를 따른 주어진 위치에서의 구형 경계 상자 및 뷰잉 배향 범위를 묘사함으로써, 가상 3D 장면 내에서의 제약된 탐색의 개념을 개략적으로 도시한다. 탐색 제약들은 콘텐츠 생성 단계에서 콘텐츠별로 결정되며, 잠재적으로 3D 지오메트리 객관적 제약들에 주관적(예를 들어, 예술적) 제약들을 더하는 추가 이점이 있다.

[0076] 본 원리들에 따르면, 하이-레벨 선택스 요소들의 배열은 3D 장면(50)을 표현하는 주어진 볼류메트릭 비디오 콘

텐츠에 연관된 뷰잉 위치들 및 배향들의 서브세트를 명시적으로 설명한다. 시그널링된 정보는 이하의 요소들을 포함한다:

- [0077] - 3D 공간 내의 곡선 경로(51)(또는 곡선 경로들의 세트),
- [0078] - 기본 경계 상자 볼륨(52),
- [0079] - 곡선 경로(들)를 따른 위치에 의해 인덱싱된 뷰잉 배향 범위들(53)의 세트.

[0080] 이러한 요소들의 조합은 3D 장면 재구성의 우수한 품질에 대해 적합한, 선호되는, 및/또는 허용되는 경로(들)를 따른 가상 변위들을 설명한다. 경로(들)를 따라 있는 각각의 위치에서, 경계 상자 내에서의 작은 직진이동들이 가능하고, 뷰잉 배향은 주어진 각도 범위 내에서 제약된다. 이것은 4DoF+ 가상 탐색 "터널들"로서 설명될 수 있다.

[0081] 제1 실시예에 따르면, 곡선 탐색 경로들의 세트는 이하의 선택스 및 시맨틱들을 갖는 3D 포인트들의 세트들에 의해 정의된다:

```
aligned(8) class NavigationPathSet()
{
    ViewingBoundingBox;
    unsigned int(8) num_paths;
    for (n=0; n<num_paths; n++) {
        unsigned int(32) num_points[n];
        for (i=0; i<num_points[n]; i++) {
            unsigned int(32) X[n][i];
            unsigned int(32) Y[n][i];
            unsigned int(32) Z[n][i];
            signed int(32) phi_min[n][i];
            signed int(32) phi_max[n][i];
            signed int(32) theta_min[n][i];
            signed int(32) theta_max[n][i];
        }
    }
}
```

```
aligned(8) class ViewingBoundingBox()
{
    unsigned int(8) shape_type;
    unsigned int(32) first_dimension;
    unsigned int(32) second_dimension;
    unsigned int(32) third_dimension;
}
```

[0084] num_paths는 콘텐츠를 탐색하기 위해 정의된 곡선 탐색 경로들의 수를 명시한다. 값 0은 전체 3D 공간이 탐색될 수 있음을 나타낸다.

[0085] num_points는 곡선 탐색 경로를 따른 샘플링된 3D 포인트들의 수를 명시한다.

[0086] X[n][i], Y[n][i], Z[n][i]는 3D 장면의 전역적 좌표계에서 n번째 경로를 따른 i번째 샘플의 3D 좌표들을 정의하는 고정 소수점 값들(예를 들어, 16.16 값들)이다. 포인트들은 경로를 따른 곡선 가로좌표에 따라 정렬된다. 따라서, n번째 경로는 두 개의 3D 엔드포인트 (X[n][0], Y[n][0], Z[n][0])와 (X[n][num_points[n]-1], Y[n][num_points[n]-1], Z[n][num_points[n]-1]) 사이에서 정의된다. 예를 들어, 곡선 경로는 목록 내의 연속적인 포인트들 사이, 및 목록의 마지막 포인트로부터 최초 포인트까지의 직선 경로의 단편 경로(piecewise path)일 수 있다. 변형에서, 곡선 경로는 3개 또는 4개의 연속적인 포인트를 사용하여 2차 또는 3차 베지어(Bezier) 곡선들을 사용하여 결정될 수 있다.

[0087] phi_min[n][i], phi_max[n][i] 및 theta_min[n][i], theta_max[n][i]는 2^{16} 등급(degree)의 단위로, n번째 경

로를 따른 i 번째 포인트에서의 뷰잉 배향의 경계를 정하는 방위각 및 고각(elevation angle)의 최소값 및 최대값이다. 방위각 값들은 예를 들어 양 끝 값을 포함하여, $-180 * 2^{16}$ 내지 $180 * 2^{16} - 1$ 범위 내에 있다. 고도 값들은 예를 들어 양 끝 값을 포함하여, $-90 * 2^{16}$ 내지 $90 * 2^{16}$ 범위 내에 있다. 방위각 및 고도 값들은 예를 들어 라디안과 같은 상이한 측정 단위들로 표현될 수 있다.

[0088] shape_type은 탐색 경로를 따른 주어진 위치에서 사용자가 약간 이동하도록 허용되는 뷰잉 경계 상자의 형상을 명시한다. shape_type이 0인 것은, 반경이 first_dimension에 의해 명시되는 구를 나타낸다. 선택스는 더 복잡한 타원체 또는 직육면체 3D 볼류메트릭 형상들을 정의하는 것을 허용한다.

[0089] 다른 실시예에 따르면, 경량 시청 공간(54) 사양은 볼류메트릭 콘텐츠에 연관된 취득 카메라들의 파라미터들로부터 탐색 경로(51)를 추론함으로써 획득된다. 탐색 경로들을 샘플링하는 3D 포인트들은 취득 카메라의 3D 위치이다. 각각의 샘플 위치에서의 뷰잉 방향(중심 방위각 및 고각들)은 해당 위치에서의 취득 카메라 뷰잉 방향이다. 취득 외부 카메라 파라미터들은 이미 볼류메트릭 비디오 콘텐츠에 연관된 메타데이터의 일부이므로, 시청 공간을 지정하기 위해 전송할 추가 메타데이터는 뷰잉 방향들을 중심으로 한 방위각 및 고각 범위들뿐만 아니라 각각의 샘플 위치에서의 뷰잉 박스의 형상 및 크기로 감소된다. 선택스 및 연관된 시맨틱의 예는 다음과 같다:

```
aligned(8) class NavigationPathFromCameraRig()
{
    ViewingBoundingBox;
    unsigned int(8) num_paths;
    for (n=0; n < num_paths; n++) {
        unsigned int(32) num_cams[n];
        for (i=0; i < num_cams[n]; i++) {
            unsigned int(32) cam_idx[n][i];
            unsigned int(32) phi_range[n][i];
            unsigned int(32) theta_range[n][i];
        }
    }
}
```

[0090]

[0091] num_cams[n]은 n 번째 탐색 경로를 샘플링하는 데 사용되는 취득 카메라들의 서브세트의 크기이다. cam_idx[n][i]는 (모든 취득 카메라들의 목록 중에서) n 번째 탐색 경로를 따른 i 번째 카메라의 인덱스이다. phi_range[n][i] 및 theta_range[n][i]는 cam_idx[n][i] 번째 취득 카메라의 방위각 및 고각 주위의 n 번째 경로를 따른 그러한 i 번째 위치에서의 방위각 및 고각 편차의 범위이다.

[0092] 이러한 뷰잉 공간 사양은 원호를 따라 위치된 카메라 장비로 캡처된 볼류메트릭 콘텐츠에 특히 적합한다. 도 5b는 그러한 장비로부터 캡처된 3D 콘텐츠에 연관된 뷰잉 공간(54)의 예를 도시하는데: 10개의 카메라가 3D 장면을 캡처하는 5쌍의 수렴 카메라(55)로 배열되고; 관 형상의 뷰잉 공간(54)은 하나 걸러 하나씩의 카메라 위치에서 5개의 회전 타원체(spheroid)를 샘플링함으로써 명시된다.

[0093] 도 6은 중심 위치에 관심 객체들이 있는 3D 장면 주위를 탐색하기 위한 원형 경로의 특정 사례를 보여준다. 이러한 사용 사례에서, 사용자는 내향 시야(inward field-of-view)(63)를 사용하여 전역적 좌표계의 주어진 높이 값에서 수평 평면 내의 원형 경로(61)에서 3D 장면(50) 주위를 건도록 허용된다. 경계 상자(52) 내의 변위들은 제1 실시예에서와 같이 허용된다.

[0094] 신택스 및 시맨틱의 예는 다음과 같을 수 있다:

```
aligned(8) class CircularNavigationPath()
{
    ViewingBoundingBox;
    signed int(32) center_x;
    signed int(32) center_y;
    signed int(32) center_z;
    unsigned int(32) radius;
    signed int(32) delta_phi;
    signed int(32) theta_min;
    signed int(32) theta_max;
}
```

[0095]

[0096] center_x, center_y, center_z는 원형 경로 중심의 3D 좌표들을 정의하는 고정 소수점 값들(예를 들어, 16.16 값들)이다.

[0097] radius는 원형 경로의 반경을 정의하는 고정 소수점 값이다.

[0098] phi-range는 2^{16} 등급의 단위로, 원형 경로(반경 방향에 대해 상대적임) 상의 임의의 포인트에서 방위각 뷰잉 방향의 경계를 정하는 각도 범위이다. delta_phi 값들은 양 끝 값을 포함하여, 0 내지 $360 * 2^{16} - 1$ 까지의 범위에 있어야 한다.

[0099] theta_min, theta_max는 2^{16} 등급의 단위로, 원형 경로 상의 임의의 포인트에서의 시야 고각들의 최소값 및 최대값이다. 고도 값들은 양 끝 값을 포함하여, $-90 * 2^{16}$ 내지 $90 * 2^{16}$ 의 범위 내에 있어야 한다.

[0100] 더 일반적으로, 곡선 탐색 경로들은 파라메트릭 3D 곡선들: $X=f(s)$, $Y=g(s)$, $Z=h(s)$ 에 의해 정의될 수 있고, s는 스칼라 값이다. 이러한 기능적 파라미터화들의 다중성(multiplicity)은 본 원리들에 따라 적합하다. 이러한 일반적인 접근방식은 스포츠 이벤트들 또는 콘서트 이벤트들의 볼류메트릭 비디오를 렌더링하는 것과 같은 전형적인 4DoF+ 경험들에 특히 적합하다. 이러한 4DoF+ 비디오들에 대해, 적합한, 선회되는, 및/또는 허용되는 경로는 예를 들어 스타디움의 스탠드의 타원, 또는 무대 또는 스타디움의 필드 주변의 직사각형과 같이, 적은 수의 파라미터를 사용하여 파라미터화할 수 있는 단순한 형상을 갖는다.

[0101] 도 7은 볼류메트릭 비디오에서의 탐색의 제약들을 표현하는 정보를 시그널링하기 위한 방법(70)을 도식적으로 도시한다. 단계(71)에서, 볼류메트릭 비디오 데이터가 획득된다. 곡선 경로를 표현하는 데이터와 곡선 경로의 포인트들에 대한 뷰잉 방향 범위들을 표현하는 데이터가 동시에 획득된다. 뷰잉 경계 상자에 연관된 곡선 경로는 볼류메트릭 비디오의 3D 장면을 포함하는 3D 공간에서의 탐색의 제약들을 표현한다. 본 원리들의 실시예에 따르면, 이러한 데이터들은 도 5 및 6과 관련하여 설명된 바와 같은 상이한 데이터 구조들을 통해 표현될 수 있다. 단계(72)에서, 볼류메트릭 비디오 데이터는 단계(71)에서 획득된 탐색의 제약들을 포함하는 메타데이터에 연관하여 데이터 스트림으로 인코딩된다. 단계(73)에서, 단계(72)에서 인코딩된 데이터 스트림은 비밀시적 매체에 저장되거나 클라이언트 디바이스에 전송될 수 있다.

[0102] 실시예에 따르면, 제안된 제약된 탐색 경로 메시지는 전용 SEI 메시지(Supplemental Enhancement Information)의 비디오 스트림 내에서 인코딩된다. 다른 실시예에 따르면, 제안된 제약된 탐색 경로 메시지는 ISO 기본 미디어 파일 포맷을 사용하여 컨테이너 레벨에서 인코딩된다. 볼류메트릭 비디오 스트림의 메타데이터에 이러한 탐색 경로 메시지를 추가하면, 렌더러 측에서 가상 탐색을 인코딩된 3D 장면 콘텐츠와 일치하는 뷰잉 위치들 및 배향들로 제한하여 몰입형 경험의 품질을 보장하는 것이 허용된다.

[0103] 도 8은 볼류메트릭 비디오에서의 탐색의 제약들을 표현하는 정보를 디코딩하기 위한 방법(80)을 도식적으로 도시한다. 단계(81)에서, 볼류메트릭 비디오의 비디오 데이터를 포함하는 데이터 스트림이 소스로부터 획득된다. 데이터 스트림은 또한 볼류메트릭 비디오에 연관된 메타데이터, 및 볼류메트릭 비디오의 3D 장면을 포함하는 3D 공간에서의 탐색의 제약들의 표현을 포함한다. 단계(82)에서, 비디오 데이터 및 메타데이터가 디코딩된다. 메타데이터는 뷰잉 경계 상자를 표현하는 데이터, 볼류메트릭 비디오의 3D 공간에서의 곡선 경로를 표현하는 데이터, 및 곡선 경로 상의 포인트에 연관된 적어도 하나의 뷰잉 방향 범위를 표현하는 데이터를 포함한다. 이러한

데이터들은 도 5 및 6과 관련하여 도시된 바와 같이 본 원리들의 실시예에 따라 상이한 데이터 구조로 표현될 수 있다. 단계(83)에서, 3D 장면의 3D 공간에서의 탐색의 제약들을 표현하는 정보는 디코딩된 메타데이터를 사용하여 검색(retrieve)되고, 이 정보는 렌더러에 의해 사용된다. 예를 들어, 렌더러는 사용자가 양호한 품질의 렌더링이 보장되는 터널을 떠나려고 할 때 경고할 수 있다. 렌더러는 또한 사용자가 적절한 경로 외부로 이동하고 있을 때 가상 카메라가 설명 볼륨 외부로 이동하는 것을 방지할 수 있거나 (예를 들어, 이미지를 페이딩함으로써) 렌더링을 수정할 수 있다.

[0104] 본 명세서에 설명된 구현들은 예를 들어 방법 또는 프로세스, 장치, 컴퓨터 프로그램 제품, 데이터 스트림, 또는 신호로 구현될 수 있다. 단일 형태의 구현의 맥락에서만 논의되더라도(예를 들어, 방법 또는 디바이스로서만 논의됨), 논의된 특징들의 구현은 또한 다른 형태들(예를 들어, 프로그램)로 구현될 수 있다. 장치는 예를 들어 적절한 하드웨어, 소프트웨어 및 펌웨어로 구현될 수 있다. 방법들은 예를 들어 컴퓨터, 마이크로프로세서, 집적 회로, 또는 프로그래밍가능한 로직 디바이스를 포함하는, 일반적으로 프로세싱 디바이스들을 지칭하는, 예를 들어 프로세서와 같은 장치에서 구현될 수 있다. 프로세서들은 또한 예를 들어 스마트폰들, 태블릿들, 컴퓨터들, 모바일 폰들, 휴대용/개인용 디지털 단말기들("PDA"), 및 최종 사용자들 간의 정보 통신을 용이하게 하는 다른 디바이스들과 같은 통신 디바이스들을 포함한다.

[0105] 본 명세서에 기술된 다양한 프로세스들 및 특징들의 구현들은 다양한 상이한 장비 또는 애플리케이션들, 특히 예를 들어, 데이터 인코딩, 데이터 디코딩, 뷰 생성, 텍스처 프로세싱, 및 이미지들 및 관련된 텍스처 정보 및/또는 깊이 정보의 다른 프로세싱에 연관된 장비 또는 애플리케이션들로 구현될 수 있다. 이러한 장비의 예들은 인코더, 디코더, 디코더로부터의 출력을 프로세싱하는 포스트-프로세서, 인코더에 입력을 제공하는 프리-프로세서, 비디오 코더, 비디오 디코더, 비디오 코덱, 웹 서버, 셋톱 박스, 랩톱, 개인용 컴퓨터, 셀 폰, PDA, 및 다른 통신 디바이스들을 포함한다. 분명히 알 수 있듯이, 장비는 모바일일 수 있으며, 모바일 차량에 설치될 수도 있다.

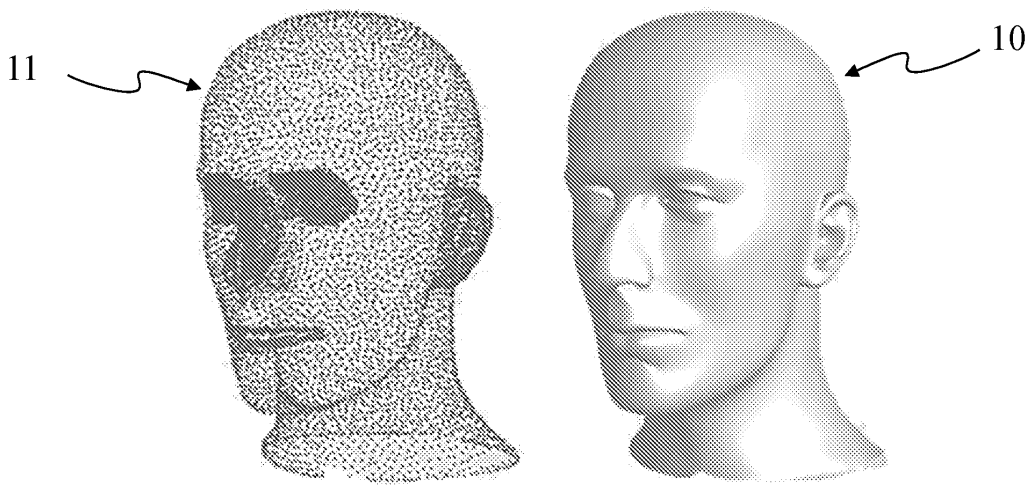
[0106] 추가로, 방법들은 프로세서에 의해 수행되는 명령어들에 의해 구현될 수 있고, 그러한 명령어들(및/또는 구현에 의해 생성된 데이터 값들)은 예를 들어 하드 디스크, 콤팩트 디스켓("CD"), (예를 들어, 종종 디지털 다목적 디스크 또는 디지털 비디오 디스크라고도 지칭되는 DVD와 같은) 광학 디스크, 랜덤 액세스 메모리("RAM"), 또는 판독 전용 메모리("ROM")와 같은, 예를 들어 집적 회로, 소프트웨어 캐리어, 또는 다른 저장 디바이스와 같은 프로세서 판독가능한 매체에 저장될 수 있다. 명령어들은 프로세서 판독가능한 매체에 유형적으로 구현된 애플리케이션 프로그램을 형성할 수 있다. 명령어들은 예를 들어 하드웨어, 펌웨어, 소프트웨어 또는 조합일 수 있다. 명령어들은 예를 들어 운영 체제, 별도의 애플리케이션 또는 이 둘의 조합에서 찾을 수 있다. 따라서, 프로세서는 예를 들어 프로세스를 수행하도록 구성된 디바이스, 및 프로세스를 수행하기 위한 명령어들을 갖는 프로세서 판독가능한 매체(예를 들어, 저장 디바이스)를 포함하는 디바이스 둘 다를 특징지어질 수 있다. 또한, 프로세서 판독가능한 매체는 명령어들에 추가하여 또는 명령어들을 대신하여, 구현에 의해 생성된 데이터 값들을 저장할 수 있다.

[0107] 본 기술분야의 통상의 기술자에게 명백한 바와 같이, 구현들은 예를 들어 저장 또는 전송될 수 있는 정보를 전달하도록 포맷팅된 다양한 신호들을 생성할 수 있다. 정보는 예를 들어 방법을 수행하기 위한 명령어들 또는 설명된 구현들 중 하나에 의해 생성된 데이터를 포함할 수 있다. 예를 들어, 신호는 설명된 실시예의 신택스를 기록 또는 판독하기 위한 규칙들을 데이터로서 전달하거나, 설명된 실시예에 의해 기입된 실제 신택스 값들을 데이터로서 전달하도록 포맷팅될 수 있다. 이러한 신호는 예를 들어 (예를 들어, 스펙트럼의 무선 주파수 부분을 사용하여) 전자기파로, 또는 기저대역 신호로 포맷팅될 수 있다. 포맷팅은 예를 들어 데이터 스트림을 인코딩하고 인코딩된 데이터 스트림으로 캐리어를 변조하는 것을 포함할 수 있다. 신호가 전달하는 정보는 예를 들어 아날로그 또는 디지털 정보일 수 있다. 신호는 알려진 바와 같이 여러 상이한 유선 또는 무선 링크들을 통해 전송될 수 있다. 신호는 프로세서 판독가능한 매체에 저장될 수 있다.

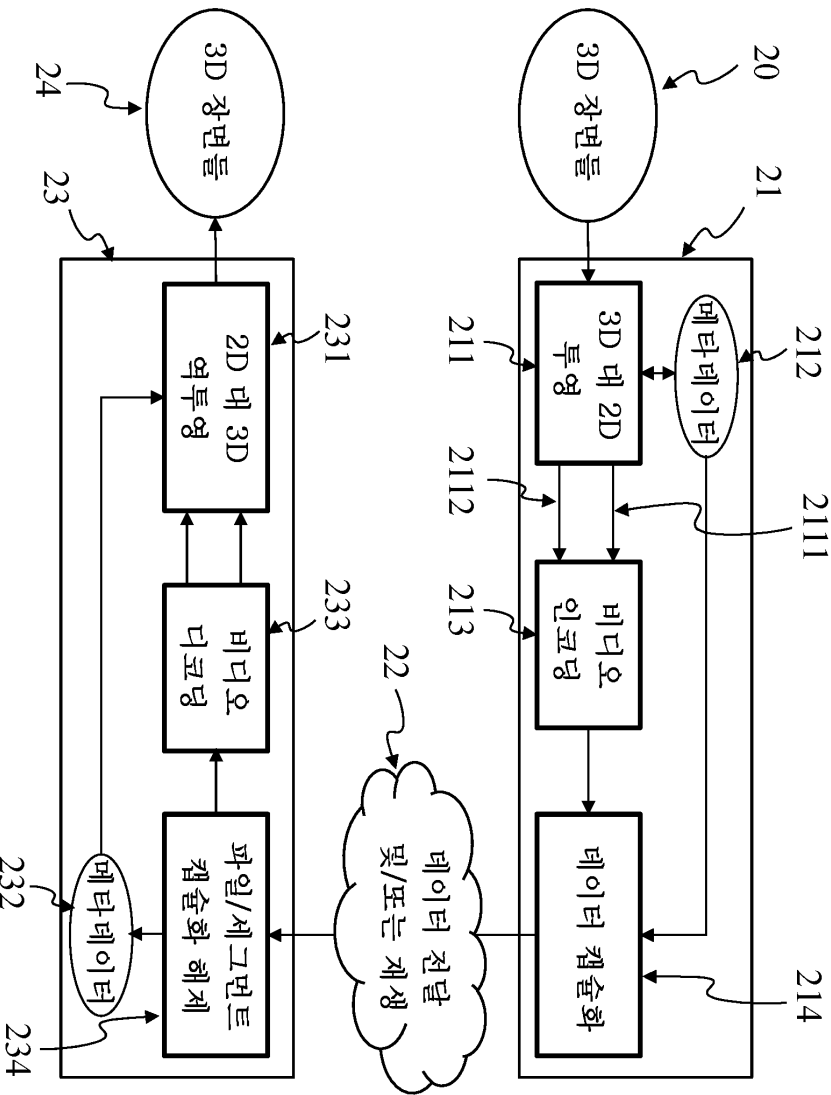
[0108] 다수의 구현이 설명되었다. 그럼에도 불구하고, 다양한 수정들이 이루어질 수 있음을 이해할 것이다. 예를 들어, 다른 구현들을 생성하기 위해 상이한 구현들의 요소들이 결합, 보완, 수정 또는 제거될 수 있다. 추가적으로, 통상의 기술자는 다른 구조들 및 프로세스들이 개시된 것들을 대체할 수 있고, 결과적인 구현들이 적어도 실질적으로 동일한 방식으로 적어도 실질적으로 동일한 기능(들)을 수행하여 개시된 구현들과 적어도 실질적으로 동일한 결과(들)를 달성할 것임을 이해할 것이다. 따라서, 이러한 구현들 및 다른 구현들이 본 출원에 의해 고려된다.

도면

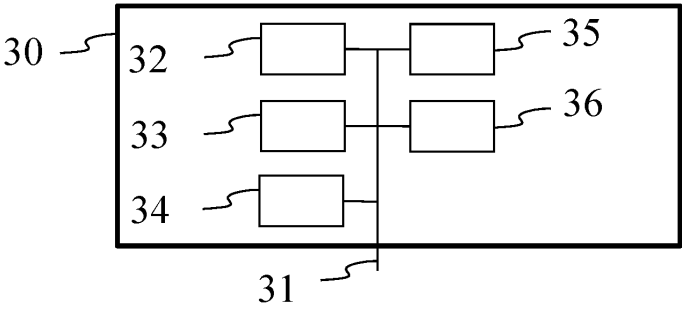
도면1



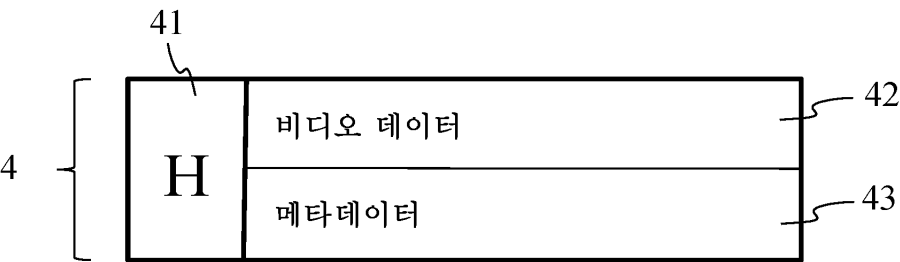
도면2



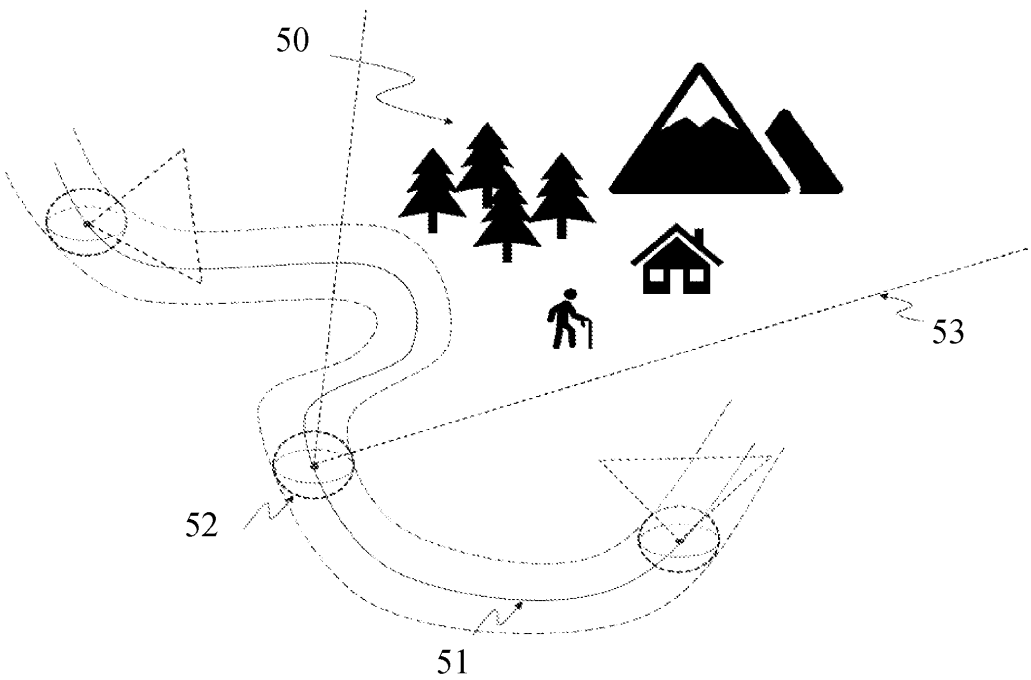
도면3



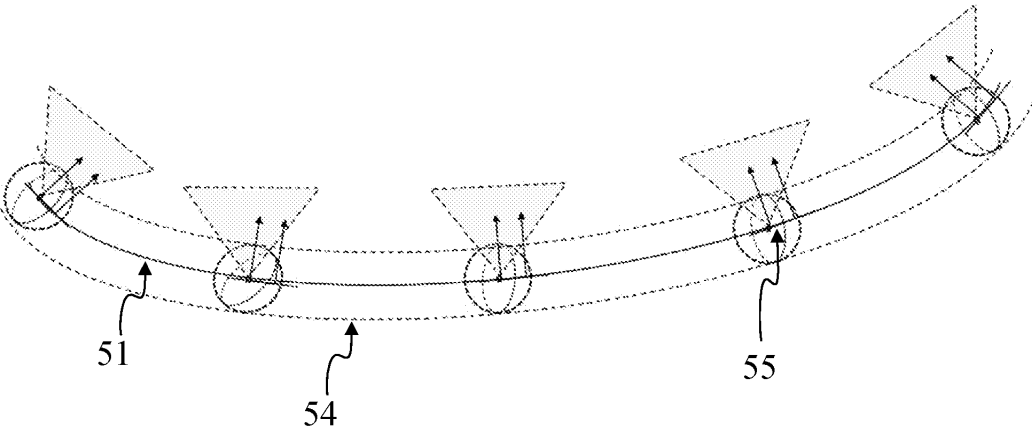
도면4



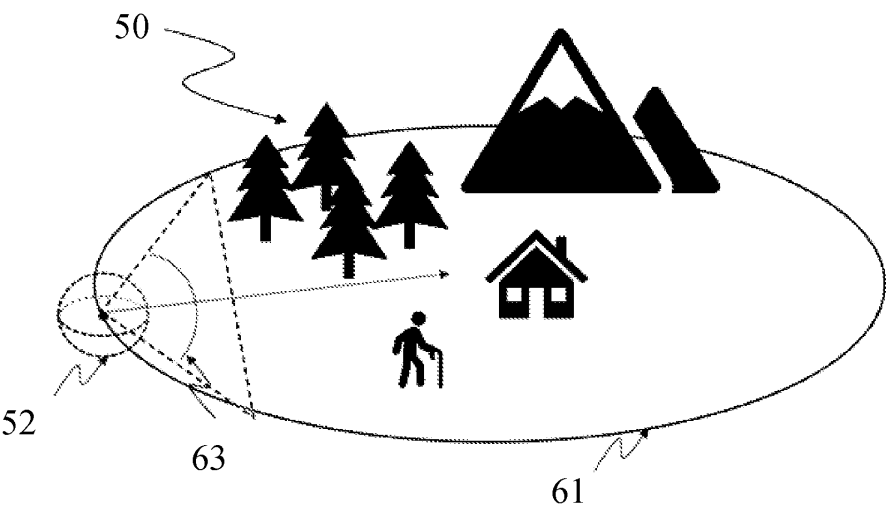
도면5



도면5b

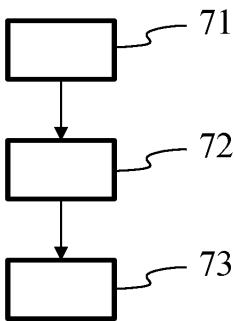


도면6



도면7

70



도면8

80

