



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2022년11월07일
(11) 등록번호 10-2464524
(24) 등록일자 2022년11월03일

(51) 국제특허분류(Int. Cl.)
H04N 13/117 (2018.01) H04N 13/161 (2018.01)
H04N 13/366 (2018.01)
(52) CPC특허분류
H04N 13/117 (2018.05)
H04N 13/161 (2018.05)
(21) 출원번호 10-2019-7031621
(22) 출원일자(국제) 2018년03월02일
심사청구일자 2021년03월02일
(85) 번역문제출일자 2019년10월25일
(65) 공개번호 10-2019-0135502
(43) 공개일자 2019년12월06일
(86) 국제출원번호 PCT/EP2018/055183
(87) 국제공개번호 WO 2018/177681
국제공개일자 2018년10월04일
(30) 우선권주장
17163509.7 2017년03월29일
유럽특허청(EPO)(EP)
(56) 선행기술조사문헌
KR1020170007098 A*
KR1020090065914 A*
*는 심사관에 의하여 인용된 문헌

(73) 특허권자
코닌클리케 필립스 엔.브이.
네덜란드 아인트호벤 5656 에이이 하이 테크 캠퍼스 52
(72) 발명자
바레캄프, 크리스티안
네덜란드 아인트호벤 5656 에이이 하이 테크 캠퍼스 5
크룬 바르트
네덜란드 아인트호벤 5656 에이이 하이 테크 캠퍼스 5
(74) 대리인
장훈

전체 청구항 수 : 총 13 항

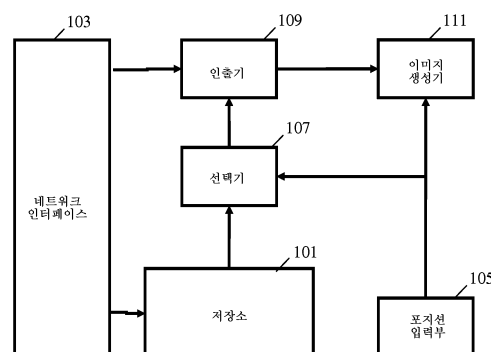
심사관 : 옥윤철

(54) 발명의 명칭 비디오로부터의 이미지 생성

(57) 요약

장치는 N-차원 공간에서의 루트들의 세트에 대한 루트 데이터를 저장하기 위한 저장소(101)를 포함하며, 여기서 루트들의 세트의 각각의 루트는 이미지 및 깊이 정보 둘 모두를 포함하는 프레임들을 포함하는 비디오 아이탬과 연관된다. 입력부(105)는 관찰자 포지션 표시를 수신하고, 선택기(107)는 관찰자 포지션 표시 및 루트들의 세트 (뒷면에 계속)

대표도 - 도1



의 루트들의 포지션들에 의존하는 거리 메트릭에 의존하는 선택 기준에 응답하여 루트들의 세트 중 제1 루트를 선택한다. 인출기(103, 109)는 비디오 소스(203)로부터 제1 루트와 연관된 제1 비디오 아이템을 인출한다. 이미지 생성기(111)는 제1 비디오 아이템으로부터의 프레임들의 제1 세트로부터 관찰자 포지션 표시에 대한 적어도 하나의 뷰 이미지를 생성한다. 이 시스템에서, 선택 기준은 루트들의 세트 중 다른 루트들에 비해 현재 선택된 루트를 향해 편향된다.

(52) CPC특허분류

H04N 13/366 (2018.05)

명세서

청구범위

청구항 1

장치로서,

공간적 포즈들을 표현하는 N-차원 공간에서의 루트(route)들의 세트를 저장하기 위한 저장소(101) - 상기 루트들의 세트는 복수의 루트들을 포함하고, 상기 루트들의 세트의 각각의 루트는 상기 N-차원 공간을 통한 포즈들의 카메라 궤적을 기술하고, 상기 루트의 카메라 포즈들에 대한 이미지 및 깊이 정보 둘 모두를 포함하는 프레임들을 포함하는 비디오 아이템에 링크됨 -;

상기 N-차원 공간에서의 포지션 표시인 관찰자 포지션 표시를 수신하기 위한 입력부(105);

상기 관찰자 포지션 표시 및 상기 루트들의 세트의 상기 루트들의 포지션들에 의존하는 거리 메트릭에 의존하는 선택 기준에 응답하여 상기 루트들의 세트 중 제1 루트를 선택하기 위한 선택기(107);

비디오 소스(203)로부터 상기 제1 루트와 연관된 제1 비디오 아이템을 인출하기 위한 인출기(103, 109); 및

상기 제1 비디오 아이템으로부터의 프레임들의 제1 세트로부터 상기 관찰자 포지션 표시에 대한 적어도 하나의 뷰 이미지를 생성하기 위한 이미지 생성기(111)를 포함하고,

상기 선택 기준은 상기 루트들의 세트 중 다른 루트들에 비해 현재 선택된 루트를 향해 편향(bias)되고,

상기 저장소(101)는 비디오 시퀀스들의 루트들에 대한 교차 포지션들(intersect positions)을 저장하도록 추가로 배열되고, 상기 선택 기준은 상기 교차 포지션들에 의존하고, 상기 현재 선택된 루트를 향한 편향은 상기 관찰자 포지션 표시와 상기 교차 포지션들 사이의 거리에 의존하는, 장치.

청구항 2

제1항에 있어서, 상기 이미지 생성기(111)는 상기 관찰자 포지션 표시에 대한 상기 제1 비디오 아이템의 프레임들에 대한 루트 포지션 데이터에 응답하여 상기 프레임들의 제1 세트를 선택하도록 배열되는, 장치.

청구항 3

제1항 또는 제2항에 있어서, 상기 이미지 생성기(111)는, 상기 제1 루트가 상기 현재 선택된 루트로서 선택되는 경우, 상기 프레임들의 제1 세트의 선택을, 상기 제1 비디오 아이템의 현재 선택된 프레임들의 제1 세트에 대한 더 작은 시간적 거리(temporal distance)를 갖는 프레임들을 향해 편향시키도록 배열되는, 장치.

청구항 4

제1항 또는 제2항에 있어서, 상기 이미지 생성기(111)는, 상기 제1 루트가 상기 현재 선택된 루트로서 선택되는 경우, 상기 프레임들의 제1 세트의 선택을, 상기 제1 비디오 아이템의 현재 선택된 프레임에 대한 감소하는 시간적 거리에 따라 증가하는 프레임 선택 메트릭에 기초하여 수행하도록 배열되는, 장치.

청구항 5

제1항에 있어서, 상기 선택기(107)는 교차 포지션들에 대응하는 관찰자 포지션 표시들에 대한 상기 현재 선택된 루트를 향한 편향을 감소시키도록 배열되는, 장치.

청구항 6

제1항에 있어서, 상기 저장소(101)는 에지들에 의해 반영되는 루트들 및 노드들로서의 교차점들을 갖는 그래프를 저장하도록 배열되는, 장치.

청구항 7

제1항, 제5항 및 제6항 중 어느 한 항에 있어서, 적어도 일부 비디오 시퀀스들에 대한 교차점 포지션들과 연관

된 프레임들은 인트라 코딩된(intra coded) 프레임들인, 장치.

청구항 8

제1항, 제5항 및 제6항 중 어느 한 항에 있어서, 상기 이미지 생성기는, 적어도 상기 프레임들의 제1 세트 중 제1 프레임 및 이전에 선택된 제2 루트에 링크된 비디오 아이템으로부터의 제2 프레임으로부터 상기 뷰 이미지를 생성하도록 배열되는, 장치.

청구항 9

제8항에 있어서, 상기 이미지 생성기는, 적어도 상기 프레임들의 제1 세트 중 제1 프레임 및 이전에 선택된 제2 루트에 링크된 비디오 아이템으로부터의 제2 프레임으로부터 상기 뷰 이미지를 생성하도록 배열되고, 상기 제2 프레임은 상기 제1 루트와 상기 이전에 선택된 제2 루트 사이의 교차점과 연관된 프레임인, 장치.

청구항 10

제1항에 있어서, 상기 저장소(101)는 상기 루트들의 세트의 루트들 사이의 공간적 관계를 나타내는 관계 데이터를 저장하도록 추가로 배열되고, 상기 제1 루트의 선택은 상기 관계 데이터에 응답하는, 장치.

청구항 11

제10항에 있어서, 상기 관계 데이터는, 하기의 그룹으로부터 2개의 루트들 사이의 적어도 하나의 공간적 관계를 나타낼 수 있는, 장치:

- 상기 2개의 루트들의 교차부;
- 상기 2개의 루트들의 접합부;
- 상기 2개의 루트들 중 제1 루트가 상기 2개의 루트들 중 제2 루트와는 반대 이동 방향에 대응하는 것; 및
- 상기 2개의 루트들이 평행인 것.

청구항 12

제1항, 제2항, 제5항 및 제6항 중 어느 한 항에 있어서, 상기 비디오 아이템들은 인트라 코딩된 프레임들 및 인터 코딩된(inter-coded) 프레임들 양쪽 모두를 포함하는 인코딩된 비디오 아이템들이고, 상기 선택기는, 인터 코딩된 프레임들에 대응하는 관찰자 포지션 표시들에 비해 인트라 코딩된 프레임들에 대응하는 관찰자 포지션 표시들에 대한 상기 현재 선택된 루트를 향한 편향을 감소시키도록 배열되는, 장치.

청구항 13

방법으로서,

공간적 포즈들을 표현하는 N-차원 공간에서의 루트들의 세트를 저장하는 단계 - 상기 루트들의 세트는 복수의 루트들을 포함하고, 상기 루트들의 세트의 각각의 루트는 상기 N-차원 공간을 통한 포즈들의 카메라 궤적을 기술하고, 상기 루트의 카메라 포즈들에 대한 이미지 및 깊이 정보 둘 모두를 포함하는 프레임들을 포함하는 비디오 아이템에 링크됨 -;

상기 N-차원 공간에서의 포지션 표시인 관찰자 포지션 표시를 수신하는 단계;

상기 관찰자 포지션 표시 및 상기 루트들의 세트의 상기 루트들의 포지션들에 의존하는 거리 메트릭에 의존하는 선택 기준에 응답하여 상기 루트들의 세트 중 제1 루트를 선택하는 단계;

비디오 소스로부터 상기 제1 루트와 연관된 제1 비디오 아이템을 인출하는 단계; 및

상기 제1 비디오 아이템으로부터의 프레임들의 제1 세트로부터 상기 관찰자 포지션 표시에 대한 적어도 하나의 뷰 이미지를 생성하는 단계를 포함하고,

상기 선택 기준은 상기 루트들의 세트 중 다른 루트들에 비해 현재 선택된 루트를 향해 편향되고,

상기 저장하는 단계는 비디오 시퀀스들의 루트들에 대한 교차 포지션들(intersect positions)을 저장하는 단계를 포함하고, 상기 선택 기준은 상기 교차 포지션들에 의존하고, 상기 현재 선택된 루트를 향한 편향은 상기 관

찰자 포지션 표시와 상기 교차 포지션들 사이의 거리에 의존하는, 방법.

청구항 14

삭제

청구항 15

삭제

발명의 설명

기술 분야

[0001] 본 발명은 비디오 아이템들로부터의 이미지들의 생성에 관한 것으로, 특히, 현실 세계 환경들의 비디오 캡처로부터 비롯될 수 있는 저장된 비디오 아이템들로부터의 가상 현실 환경에 대한 뷰 이미지들의 생성에 관한 것이지만 배타적인 것은 아니다.

배경 기술

[0002] 이미지 및 비디오 애플리케이션들의 다양성과 범위는 최근 몇 년간 크게 증가했으며, 비디오를 이용 및 소비하는 새로운 서비스들 및 방법들이 계속 개발 및 도입되고 있다.

[0003] 예를 들어, 점점 더 대중화되고 있는 하나의 서비스는 관찰자가 렌더링의 파라미터를 변경하기 위해 시스템과 능동적으로 그리고 동적으로 상호작용할 수 있는 방식으로 이미지 시퀀스를 제공하는 것이다. 많은 애플리케이션에서 매우 매력적인 특징은, 예를 들어 관찰자가 제시되는 장면에서 움직이고 "둘러보는" 것을 가능하게 하는 것과 같은, 관찰자의 유효 관찰 포지션 및 관찰 방향을 변경하는 능력이다.

[0004] 그러한 특징은 특히 가상 현실 경험이 사용자에게 제공되는 것을 가능하게 할 수 있다. 이것은 사용자가 가상 환경에서 (비교적) 자유롭게 움직이고 그의 포지션 및 그가 보고 있는 곳을 동적으로 변경하는 것을 가능하게 할 수 있다. 전형적으로, 그러한 가상 현실 애플리케이션은 장면의 3차원 모델에 기초하며, 이러한 모델은 특정한 요청된 뷰를 제공하기 위해 동적으로 평가된다. 이러한 접근법은 예를 들어 컴퓨터 및 콘솔에 대한 1인칭 슈터의 카테고리에서와 같은 게임 애플리케이션으로부터 잘 알려져 있다.

[0005] 특히 가상 현실 애플리케이션의 경우, 제시되는 이미지가 3차원 이미지인 것이 또한 바람직하다. 실제로, 관찰자의 몰입을 최적화하기 위해, 전형적으로 사용자가 제시된 장면을 3차원 장면으로서 경험하는 것이 바람직하다. 실제로, 가상 현실 경험은 바람직하게는 사용자가 가상 세계에 대한 그/그녀 자신의 포지션, 카메라 시점(viewpoint), 및 시간적 순간을 선택할 수 있게 하여야 한다.

[0006] 전형적으로, 가상 현실 애플리케이션들은 이들이 장면의 미리결정된 모델에 기초하고, 전형적으로 가상 세계의 인공 모델에 기초한다는 점에서 본질적으로 제한된다. 가상 현실 경험이 현실 세계 캡처에 기초하여 제공될 수 있다면 바람직할 것이다. 그러나, 많은 경우에, 그러한 접근법은 매우 제한되거나 현실 세계의 가상 모델이 현실 세계 캡처들로부터 구축될 것을 요구하는 경향이 있다. 이어서, 이러한 모델을 평가함으로써 가상 현실 경험이 생성된다.

[0007] 그러나, 현재의 접근법들은 준최적(suboptimal)인 경향이 있고, 종종 높은 계산 또는 통신 리소스 요건을 갖는 경향이 있고/있거나, 예컨대 감소된 품질 또는 제한된 자유를 갖는 준최적의 사용자 경험을 제공하는 경향이 있다.

[0008] 애플리케이션의 일례로서, 가상 현실 안경이 현재 시장에 진입하고 있다. 이러한 안경은 관찰자가 캡처된 360도(파노라마) 비디오를 경험할 수 있게 한다. 이러한 360도 비디오는 종종, 개별 이미지들이 단일 구면 매핑으로 함께 스티칭(stitching)되는 카메라 리그(camera rig)들을 사용하여 미리 캡처된다. 360도 비디오에 대한 일반적인 스테레오 포맷들은 상/하 및 좌/우이다. 눈-파노라마 스테레오 비디오와 유사하게, 좌안 및 우안 픽처(picture)들은 단일 H.264 비디오 스트림의 일부로서 압축된다. 단일 프레임을 디코딩한 후에, 관찰자는 그/그녀 주위의 세계를 보기 위해 그/그녀의 머리를 회전시킨다. 현재의 예는 구글(Google)에 의해 행해진 카네기 홀 콘서트의 레코딩이다[https://performingarts.withgoogle.tabhumblecom/en_us/performances/carnegie-hall1]. 이 예에서, 관찰자는 360도 둘러보기 효과를 경험할 수 있고, 상이한 포지션들로부터 레코딩된 3개의 비디오 스트림들 사이에서 개별적으로 전환할 수 있다. 전환할 때, 다른 비디오 스트림이 로딩되는데, 이는 경

험을 중단시킨다.

- [0009] 스테레오 파노라마 비디오 접근법의 한 가지 결점은 관찰자가 가상 세계에서 포지션을 변경할 수 없다는 점이다. 파노라마 스테레오 비디오 이외의 파노라마 깊이 맵의 인코딩 및 송신은 클라이언트 측에서의 관찰자의 작은 병진 운동의 보상을 허용할 수 있지만, 그러한 보상은 본질적으로 작은 변동 및 움직임으로 제한될 것이며, 몰입적이고 자유로운 가상 현실 경험을 제공할 수 없을 것이다.
- [0010] 관련 기술은 깊이 맵을 갖는 다수의 시점이 단일 비디오 스트림으로 인코딩 및 송신되는 자유 시점 비디오이다. 비디오 스트림의 비트 레이트는 잘 알려진 시간 예측 스킴(temporal prediction scheme)에 더하여 시점들 간의 각도 의존성을 이용함으로써 감소될 수 있다. 그러나, 이러한 접근법은 여전히 높은 비트 레이트를 요구하며, 생성될 수 있는 이미지 면에서 제한적이다. 그것은 3차원 가상 현실 세계에서 완전히 자유로운 움직임의 경험을 사실상 제공할 수 없다.
- [0011] 유감스럽게도, 종래 기술들 중 어느 것도 이상적인 경험을 전달할 수 없으며, 종종 포지션 및 관찰 방향의 변화의 자유에 있어서 제한적인 경향이 있다. 또한, 이러한 기술들은 매우 높은 데이터 레이트를 요구하고 개별 이미지들/뷰들의 생성에 필요한 것보다 더 많은 데이터를 포함하는 데이터 스트림을 제공하는 경향이 있다.
- [0012] 추가로, 종래 기술의 접근법들은 다수의, 예컨대, 개별적으로 캡처된 이미지들을 저장하기 위해 큰 저장 용량을 요구하는 경향이 있다. 또한, 종래 기술의 접근법들은 비교적 복잡한 프로세싱을 요구하고 높은 리소스 계산 요구들을 갖는 경향이 있다. 이것에도 불구하고, 종래 기술의 접근법들은 뷰 이미지들의 준최적의 생성을 야기하는 경향이 있고, 특히, 연결이 안 되고 비교적 일관성이 없는 경험이 종종 가상 세계를 움직이는 사용자에게 제공된다.
- [0013] 예를 들어, KR2017 0007098 A호는 장면의 이미지가 캡처된 캡처 포즈들과는 상이할 수 있는 주어진 사용자 포즈에 대한 이미지들을 생성하기 위한 시스템을 개시한다. 이 접근법은 가장 가까운 캡처 포즈를 선택하고 가장 가까운 캡처 포즈로부터의 이미지로부터 사용자 포즈에 대한 뷰 이미지를 생성할 수 있다. 그러나, 이 접근법은 전술된 모든 단점들을 겪는 경향이 있으며, 특히 높은 계산 리소스를 요구하며, 복잡하며, 복잡하고 어려운 캡처 동작들에 기초하며, 최적의 품질이 아니고 더욱이 사용자가 움직일 때 일관성이 없는 경향이 있는 뷰 이미지들을 생성하는 경향이 있다.
- [0014] 따라서, 개선된 접근법이 유리할 것이다. 특히, 개선된 동작, 증가된 유연성, 상이한 관찰 포지션들 및 방향들에 대응하는 이미지들을 생성하기 위한 증가된 범위, 개선된 가상 현실 경험, 감소된 데이터 레이트, 용이해진 배포, 감소된 복잡성, 용이해진 구현, 감소된 저장 요건들, 및/또는 개선된 성능 및/또는 동작을 가능하게 하는 접근법이 유리할 것이다.
- 발명의 내용**
- [0015] 따라서, 본 발명은 전술한 불리한 점들 중 하나 이상을 단독으로 또는 임의의 조합으로 바람직하게 완화, 경감 또는 제거하고자 한다.
- [0016] 본 발명의 일 태양에 따르면, 장치가 제공되며, 장치는, 공간적 포즈들을 표현하는 N-차원 공간에서의 루트(route)들의 세트를 저장하기 위한 저장소 - 루트들의 세트는 복수의 루트들을 포함하고, 루트들의 세트의 각각의 루트는 N-차원 공간을 통한 포즈들의 카메라 궤적을 기술하고, 루트의 카메라 포즈들에 대한 이미지 및 깊이 정보 둘 모두를 포함하는 프레임들을 포함하는 비디오 아이템에 링크됨 -; 관찰자 포지션 표시를 수신하기 위한 입력부; 관찰자 포지션 표시 및 루트들의 세트의 루트들의 포지션들에 의존하는 거리 메트릭에 의존하는 선택 기준에 응답하여 루트들의 세트 중 제1 루트를 선택하기 위한 선택기; 비디오 소스로부터 제1 루트와 연관된 제1 비디오 아이템을 인출하기 위한 인출기; 및 제1 비디오 아이템으로부터의 프레임들의 제1 세트로부터 관찰자 포지션 표시에 대한 적어도 하나의 뷰 이미지를 생성하기 위한 이미지 생성기를 포함하고, 선택 기준은 루트들의 세트 중 다른 루트들에 비해 현재 선택된 루트를 향해 편향(bias)된다.
- [0017] 이 접근법은 많은 시나리오들에서 그리고 많은 애플리케이션들에 대해 개선된 성능 및/또는 동작을 제공할 수 있다. 특히, 이 접근법은 많은 시나리오들에서 개선된 가상 현실 경험을 제공할 수 있다. 특히, 낮은 레이턴시(latency), 높은 품질, 낮은 통신 리소스 요건, 낮은 저장 리소스 요건, 및/또는 낮은 계산 리소스 요건을 갖는 효율적인 접근법이 종종 달성될 수 있다.
- [0018] 이 접근법은, 예를 들어, 비디오 아이템들을 저장하고 어쩌면 또한 루트 데이터를 제공하는 서버에 기초하여 클라이언트 디바이스에서 가상 현실 경험을 제공하기 위해 클라이언트 서버 배열에 대한 감소된 통신 또는 계산

리소스 요건들을 허용할 수 있다.

- [0019] 이 접근법은, 예를 들어, 개별 이미지들 및 연관된 포지션들에 기초하기보다는 오히려 루트 데이터와 연관된(이에 링크된) 비디오 아이터들에 기초하여 가상 환경을 표현하는 데 사용될 수 있다. 이는, 예를 들어, 효율적인 비디오 인코딩 기법들의 사용을 허용하여, 감소된 저장 및 통신 리소스들을 생성할 수 있다.
- [0020] 이 접근법은 많은 실시예들에서 현실 세계 캡처들로부터 가상 환경을 표현하는 데이터의 용이해진 생성을 허용할 수 있다. 예를 들어, 현실 세계 환경에 대응하는 가상 환경에서의 가상 경험은 현실 세계 환경에서 여기서 기 움직일 때 캡처된 비디오 아이터들에 기초하여 효과적으로 생성될 수 있다.
- [0021] N-차원 공간은 구체적으로 N-차원 공간에서의 위치 및/또는 뷰 배향에 관한 치수들을 포함할 수 있다. N은 가상 환경에서의 위치를 반영하는 1 내지 3개의 공간적 치수들 및 가상 환경에서의 뷰 배향을 반영하는 1 내지 3개의 공간적 치수들을 구체적으로 표현할 수 있는 정수이다.
- [0022] 관찰자 포지션 표시는 구체적으로 N-차원 공간에서의 포지션의 M-차원(이때 M은 전형적으로 N과 동일함) 표시일 수 있다. 관찰자 포지션 표시는 구체적으로 가상 환경에서의 위치를 반영하는 1 내지 3개의 공간적 치수들 및 가상 환경에서의 뷰 배향을 반영하는 1 내지 3개의 공간적 치수들을 표현할 수 있다.
- [0023] 비디오 아이터들은 이미지들 및 깊이 정보 둘 모두를 포함하여, 이에 의해 비디오 아이터들에 의해 캡처된 것들과는 상이할 수 있는 시점들로부터의 뷰 이미지 생성을 허용한다. 이는, 예컨대 사용자가 비디오 아이터들의 이미지들의 카메라 방향에 대하여 회전하게 하는 개선된 경험을 제공할 수 있다.
- [0024] 각각의 루트는 하나의 비디오 아이터에 링크될 수 있으며, 이때 루트는 비디오 아이터에 대응하는 N-차원 공간을 통한 (어쩌면 가상) 카메라 궤적을 기술한다. 루트를 따른 포지션들은 비디오 아이터들의 프레임들의 포지션들과 연관될 수 있다.
- [0025] 선택기는, 관찰자 포지션 표시와 주어진 루트 사이의 감소하는 거리를 나타내는 그 루트에 대한 거리 메트릭에 응답하여 주어진 루트가 선택될 확률을 증가시키도록 배열될 수 있다. 거리 메트릭은 예컨대, 적합한 표준에 기초할 수 있다.
- [0026] 뷰 이미지는, 관찰자 포지션 표시에 의해 주어지는 N-차원 공간에서의 포지션에서 관찰자에 의해 인지될 뷰에 대응하도록 생성될 수 있다.
- [0027] 입력부는 관찰자 포지션 표시들의 시퀀스를 수신할 수 있고, 장치는 뷰 이미지들의 대응하는 시퀀스를 생성하도록 배열될 수 있다. 장치의 동작은 동적으로 반복되어, 이에 의해 동적 사용자 경험을 제공할 수 있다. 예를 들어, 관찰자 포지션 표시들은 N-차원 공간을 통한 경로 또는 루트를 기술할 수 있고, 장치는, 이에 응답하여, 그 경로 또는 루트를 따라 움직이는 관찰자에 의해 경험될 뷰들에 대응하는 이미지들의 시퀀스를 생성할 수 있다.
- [0028] 현재 선택된 루트는 이전의 관찰자 포지션 표시에 대해 선택된 루트일 수 있고, 구체적으로는 마지막으로 수신된 관찰자 포지션 표시에 대해 선택된 루트일 수 있다. 특히, 관찰자 포지션 표시는 관찰자 포지션 표시들의 시퀀스에서의 관찰자 포지션 표시일 수 있고, 현재 선택된 루트는 현재 관찰자 포지션 표시 이전에 수신된 마지막 관찰자 포지션 표시에 대해 선택된 루트일 수 있다.
- [0029] 비디오 아이터들은 인코딩된 비디오 아이터들일 수 있다. 비디오 아이터들은 인트라 코딩된(intra-coded) 프레임들 및 인터 코딩된(inter-coded) 프레임들을 포함할 수 있다. 인트라 코딩된 프레임들은 다른 프레임들과는 독립적으로 인코딩된다. 인터 코딩된 프레임들은 적어도 하나의 다른 프레임에 대해 인코딩된다. 비디오 아이터들은 I 프레임들 및 B 및/또는 P 인코딩된 프레임들을 포함할 수 있다.
- [0030] 현재 선택된 루트를 향한 편향은, 루트들의 세트 중 다른 루트들에 대한 선택 기준과는 상이한 현재 선택된 루트에 대한 선택 기준에 의해 구현될 수 있다. 선택 기준은 현재 선택된 루트 및 루트들의 세트 중 다른 루트들에 대하여 비대칭일 수 있다. 구체적으로, 선택 기준은 루트들의 세트 중 다른 루트들에 대한 거리 메트릭들에 비해 현재 선택된 루트에 대한 거리 메트릭에 대하여 비대칭일 수 있다.
- [0031] 선택 기준은 관찰자 포지션 표시의 상이한 값들에 대해 상이한 루트들이 선택되게 하도록 배열될 수 있다. 편향은, 다른 루트에 대한 거리 메트릭이 현재 루트에 대한 거리 메트릭보다 1 초과인 인자만큼 더 짧지 않다면 현재 선택된 루트가 선택되도록 존재할 것이다.
- [0032] 프레임들의 제1 세트는 하나 이상의 프레임들일 수 있다. 일부 실시예들에서, 이미지 생성기는 하나 초과

디오 아이টে으로부터의 프레임들에 응답하여 뷰 이미지를 추가로 생성하도록 배열될 수 있다. 예를 들어, 새로운 루트가 선택되는 경우, 뷰 이미지는 이전에 선택된 루트 및 새롭게 선택된 루트 둘 모두의 프레임들로부터 생성될 수 있다.

- [0033] 본 발명의 선택적인 특징에 따르면, 이미지 생성기는 관찰자 포지션 표시에 대한 제1 비디오 아이টে의 프레임들에 대한 루트 포지션 데이터에 응답하여 제1 프레임들을 선택하도록 배열된다.
- [0034] 이는 많은 실시예들에서 개선된 성능 및/또는 동작을 제공할 수 있다.
- [0035] 본 발명의 선택적인 특징에 따르면, 이미지 생성기는, 제1 루트가 현재 선택된 루트로서 선택되는 경우, 프레임들의 제1 세트의 선택을, 제1 비디오 아이টে의 현재 선택된 프레임들의 제1 세트에 대한 더 작은 시간적 거리(temporal distance)를 갖는 프레임들을 향해 편향시키도록 배열된다.
- [0036] 이는 많은 실시예들에서 개선된 성능 및/또는 동작을 제공할 수 있다. 그것은 특히 개선된 시간적 일관성을 제공할 수 있고, 시간적으로 변화하는 아티팩트(artefact)들을 경감시킬 수 있다.
- [0037] 본 발명의 선택적인 특징에 따르면, 이미지 생성기는, 제1 루트가 현재 선택된 루트로서 선택되는 경우, 프레임들의 제1 세트의 선택을, 제1 비디오 아이টে의 현재 선택된 프레임에 대한 감소하는 시간적 거리에 따라 증가하는 프레임 선택 메트릭에 기초하여 수행하도록 배열된다.
- [0038] 이는 많은 실시예들에서 개선된 성능 및/또는 동작을 제공할 수 있다. 그것은 특히 개선된 시간적 일관성을 제공할 수 있고, 시간적으로 변화하는 아티팩트들을 경감시킬 수 있다.
- [0039] 본 발명의 선택적인 특징에 따르면, 저장소는 비디오 시퀀스들의 루트들에 대한 교차 포지션들을 저장하도록 추가로 배열되고, 선택 기준은 교차 포지션들에 의존한다.
- [0040] 이는 많은 실시예들에서 개선된 성능을 제공할 수 있고, 특히 적합한 비디오 아이টে들의 개선된 선택을 제공할 수 있다. 그것은, 예를 들어, 일부 실시예들에서, 예컨대, 장치가 현재 루트와 교차하는 잠재적인 미래의 루트들에 대한 비디오 아이টে들을 인출하게 함으로써 레이턴시를 감소시킬 수 있다.
- [0041] 교차 포지션들은 루트들 사이의 교차점들을 나타낼 수 있고, 예컨대 N-차원 공간의 치수들의 (적어도 일부의 치수의) 좌표들로 제공될 수 있거나, 또는 예컨대, 루트 데이터 또는 비디오 아이টে(예컨대, 프레임 수와 같은 것)에 대해 주어질 수 있다.
- [0042] 2개의 루트들의 포지션들이 임계치 미만의 거리에 있는 경우 교차점이 존재하는 것으로 간주될 수 있다.
- [0043] 교차점은, 하나의 루트의 끝 부분과 다른 루트의 시작 부분 사이의 접합부 또는 T-접합부를 포함하는 임의의 접합부일 수 있다.
- [0044] 본 발명의 선택적인 특징에 따르면, 선택기는 교차 포지션들에 대응하는 관찰자 포지션 표시들에 대한 현재 선택된 루트를 향한 편향을 감소시키도록 배열된다.
- [0045] 이는 개선된 성능을 제공할 수 있고, 구체적으로 다른 곳보다 교차점들에서 그러한 전이들을 가능성이 더 크게 함으로써 상이한 비디오 아이টে들 사이의 전이를 개선할 수 있다. 일부 실시예들에서, 새로운 루트로의 전이들은 교차점들에서만 허용될 수 있다.
- [0046] 본 발명의 선택적인 특징에 따르면, 현재 선택된 루트를 향한 편향은 관찰자 포지션 표시와 교차 포지션들 사이의 거리에 의존한다.
- [0047] 이는 개선된 성능을 제공할 수 있고, 구체적으로 다른 곳보다 교차점들에서 그러한 전이들을 가능성이 더 크게 함으로써 상이한 비디오 아이টে들 사이의 전이를 개선할 수 있다. 일부 실시예들에서, 새로운 루트로의 전이들은 교차점들에서만 허용될 수 있다.
- [0048] 본 발명의 선택적인 특징에 따르면, 저장소는 에지들에 의해 반영되는 루트들 및 노드들로서의 교차점들을 갖는 그래프를 저장하도록 배열된다.
- [0049] 이것은 많은 실시예들에서 특히 효율적인 동작을 제공할 수 있다.
- [0050] 본 발명의 선택적인 특징에 따르면, 적어도 일부 비디오 시퀀스들에 대한 교차점 포지션들과 연관된 프레임들은 인트라 코딩된 프레임들이다.
- [0051] 이는 많은 실시예들에서 동작을 용이하게 하고/하거나 개선할 수 있다. 예를 들어, 그것은 디코딩될 필요가 있

는 프레임들의 수를 감소시킬 수 있기 때문에 상이한 루트들/비디오 아이템들 사이의 전이를 용이하게 할 수 있다.

[0052] 본 발명의 선택적인 특징에 따르면, 이미지 생성기는, 프레임들의 제1 세트 중 적어도 제1 프레임 및 이전에 선택된 제2 루트로부터의 제2 프레임으로부터 뷰 이미지를 생성하도록 배열된다.

[0053] 이는 많은 실시예들에서 동작을 용이하게 하고/하거나 개선할 수 있다.

[0054] 제1 프레임은 인트라 코딩된 프레임으로서 선택될 수 있고/있거나 제2 프레임은 인트라 코딩된 프레임으로서 선택될 수 있다. 제1 프레임은 구체적으로 교차점 프레임일 수 있다. 뷰 이미지의 생성은, 이전에 선택된 루트로부터의 인트라 코딩된 (I) 프레임 및 제1 루트로부터의 인트라 코딩된 (I) 프레임(제1 프레임)으로부터 예측된 프레임(종종 양방향(B) 프레임)을 디코딩하는 것을 포함할 수 있다. (이전에 선택된 루트는, 특히 선택기가 (제1 루트의 선택 이전에) 현재 선택된 루트와는 상이한 루트로서 제1 루트를 선택할 때, 제1 루트의 선택 이전의 현재 선택된 루트에 대응할 수 있다). 이는 전형적으로 장치가 하나의 루트로부터 다른 루트로 전환할 때의 시나리오일 수 있다.

[0055] 본 발명의 선택적인 특징에 따르면, 제2 프레임은 제1 루트와 이전에 선택된 제2 루트 사이의 교차점과 연관된 프레임이다.

[0056] 이는 많은 실시예들에서 동작을 용이하게 하고/하거나 개선할 수 있다.

[0057] 본 발명의 선택적인 특징에 따르면, 저장소는 루트들의 세트의 루트들 사이의 공간적 관계를 나타내는 관계 데이터를 저장하도록 추가로 배열되고, 제1 루트의 선택은 관계 데이터에 응답한다.

[0058] 이는 많은 실시예들에서 개선된 성능 및/또는 동작을 제공할 수 있다.

[0059] 본 발명의 선택적인 특징에 따르면, 관계 데이터는, 하기의 그룹으로부터 2개의 루트들 사이의 적어도 하나의 공간적 관계를 나타낼 수 있다: 2개의 루트들의 교차부; 2개의 루트들의 접합부; 2개의 루트들 중 제1 루트가 2개의 루트들 중 제2 루트와는 반대 이동 방향에 대응하는 것; 및 2개의 루트들이 평행인 것.

[0060] 이는 많은 실시예들에서 개선된 성능 및/또는 동작을 제공할 수 있다.

[0061] 본 발명의 선택적인 특징에 따르면, 비디오 아이템들은 인트라 코딩된 프레임들 및 인터 코딩된 프레임들 양쪽 모두를 포함하는 인코딩된 비디오 아이템들이고, 선택기는, 인터 코딩된 프레임들에 대응하는 관찰자 포지션 표시들에 비해 인트라 코딩된 프레임들에 대응하는 관찰자 포지션 표시들에 대한 현재 선택된 루트를 향한 편향을 감소시키도록 배열된다.

[0062] 이는 계산 리소스 사용량을 감소시킬 수 있고, 특히, 뷰 이미지들을 생성하기에 적합한 프레임들을 디코딩하기 위한 리소스 사용 및 복잡성을 감소시킬 수 있다.

[0063] 본 발명의 일 태양에 따르면, 방법이 제공되며, 본 방법은, 공간적 포즈들을 표현하는 N-차원 공간에서의 루트들의 세트를 저장하는 단계 - 루트들의 세트는 복수의 루트들을 포함하고, 루트들의 세트의 각각의 루트는 N-차원 공간을 통한 포즈들의 카메라 궤적을 기술하고, 루트의 카메라 포즈들에 대한 이미지 및 깊이 정보 둘 모두를 포함하는 프레임들을 포함하는 비디오 아이템에 링크됨 -; 관찰자 포지션 표시를 수신하는 단계; 관찰자 포지션 표시 및 루트들의 세트의 루트들의 포지션들에 의존하는 거리 메트릭에 의존하는 선택 기준에 응답하여 루트들의 세트 중 제1 루트를 선택하는 단계; 비디오 소스로부터 제1 루트와 연관된 제1 비디오 아이템을 인출하는 단계; 및 제1 비디오 아이템으로부터의 프레임들의 제1 세트로부터 관찰자 포지션 표시에 대한 적어도 하나의 뷰 이미지를 생성하는 단계를 포함하고, 선택 기준은 루트들의 세트 중 다른 루트들에 비해 현재 선택된 루트를 향해 편향된다.

[0064] 본 발명의 이들 및 다른 태양들, 특징들 및 이점들이 이하에 설명되는 실시예(들)로부터 명백할 것이고 그것을 참조하여 설명될 것이다.

도면의 간단한 설명

[0065] 본 발명의 실시예들이 도면들을 참조하여 단지 예로서 설명될 것이다.

도 1은 본 발명의 일부 실시예들에 따른, 관찰자 포지션 표시에 응답하여 뷰 이미지들을 생성하기 위한 디바이스의 요소들의 일례를 예시한다.

도 2는 클라이언트 디바이스로서 도 1의 디바이스를 포함하는 클라이언트 서버 배열의 일례를 예시한다.

도 3은 도 2의 클라이언트 서버 배열에 대한 서버의 일례를 예시한다.

도 4는 환경을 캡처하기 위한 캡처 루트들의 일례를 예시한다.

도 5는 가상 환경에서 캡처 루트들에 대한 관찰자 루트의 일례를 예시한다.

도 6은 가상 환경을 위한 캡처 루트들의 일례를 예시한다.

도 7은 가상 환경을 표현하는 캡처 루트들/비디오 아이터들의 그래프의 일례를 예시한다.

도 8은 가상 환경을 표현하는 캡처 루트들/비디오 아이터들 사이의 관계들의 그래프의 일례를 예시한다.

발명을 실시하기 위한 구체적인 내용

- [0066] 사용자가 가상 세계에서 여기저기 움직이게 하는 가상 경험들은 점점 더 대중화되고 있고, 그러한 요구를 만족시키기 위한 서비스들이 개발되고 있다. 그러나, 효율적인 가상 현실 서비스들의 제공은, 특히 경험이 완전히 가상으로 생성된 인공 세계에 대해서보다는 오히려 현실 세계 환경의 캡처에 기초하는 것이라면 매우 어렵다.
- [0067] 가상 현실 서비스를 제공하기 위한 접근법은 다수의 이미지들을 저장하는 것이며, 이때 각각의 이미지는 특정 포지션 및 전형적으로 뷰 배향/방향에 링크된다. 가상 세계에서의 사용자 위치를 고려하면, 저장된 포지션들은 그 포지션에 대한 (또는 근거리 포지션에 대한) 이미지를 인출하기 위해 검색될 수 있고, 이는 사용자에게 적절하게 렌더링될 수 있다. 예를 들어, 서버는 사용자들의 포지션들을 수신할 수 있고, 국부 이미지 저장소로부터 관련 이미지들을 추출하고 가상 세계를 통한 사용자의 경로에 대응하는 전용 비디오 시퀀스를 생성할 수 있다.
- [0068] 그러나, 그러한 접근법은 가상 환경의 충분한 커버리지를 보장하기 위해 다량의 데이터가 저장될 것을 요구할 것이다. 요구되는 이미지들의 밀도, 및 그에 따라 저장될 필요가 있는 이미지들의 총 개수는 이미지들과 함께 깊이 정보를 저장함으로써 감소될 수 있다. 이는 저장된 포지션들 및 방향들과 정확하게 정렬되지 않는 뷰 포지션들 및 방향들에 대한 뷰 이미지들의 국부 생성을 허용할 수 있다. 그러나, 다수의 저장된 이미지들이 전형적으로 여전히 필요하다. 추가로, 이미지들 및 연관된 포지션들을 개별적으로 저장하는 데 필요한 데이터 구조는 전형적으로 최적화된 저장 효율을 허용하지 않는다.
- [0069] 추가로, 이 접근법은 전형적으로 비교적 높은 계산 및 통신 리소스 사용량과 함께 비교적 높은 복잡성을 야기한다. 특히, 이 접근법은, 예컨대 인터넷을 통해 공통 서버에 액세스하는 증가하는 수의 사용자들에게 확장되기 어렵다. 그러한 경우에, 서버는 각각의 사용자에게 대한 전용 비디오 신호를 생성하고 그에 따라서 각각의 사용자에게 대한 (전형적으로 복잡한) 비디오 인코딩을 개별적으로 수행할 필요가 있을 것이다.
- [0070] 도 1은 많은 시나리오들에서 개선된 가상 현실 사용자 경험을 제공할 수 있는 장치를 예시한다. 일부 실시예들에서, 장치는, 도 2에 예시된 바와 같이, 예컨대 인터넷과 같은 네트워크(205)를 통해 원격 서버(203)와 협력하는 클라이언트 디바이스(201)일 수 있다. 서버(203)는 잠재적으로 다수의 클라이언트 디바이스들(201)을 동시에 지원하도록 배열될 수 있다.
- [0071] 장치는 주어진 관찰자 뷰 포즈에 대한 장면의 뷰들에 대응하는 이미지들을 생성하도록 배열된다. 많은 실시예들에서, 장치는 주어진 관찰자 포즈에 대한 스테레오 이미지들을 생성하도록 배열되거나, 또는 동등하게는 2개의 이미지들이 관찰자의 2개의 눈들 사이의 오프셋에 대응하는 약간 상이한 관찰자 포즈들에 대해 생성될 수 있다.
- [0072] 이 분야에서, 배치 및 포즈라는 용어들은 포지션 및/또는 방향/배향에 대한 공통 용어로서 사용된다. 예컨대, 객체, 카메라, 또는 뷰의 포지션 및 방향/배향의 조합은 포즈 또는 배치로 지칭될 수 있다. 따라서, 배치 또는 포즈 표시는 6개의 값들/컴포넌트들/자유도를 포함할 수 있으며, 이때 각각의 값/컴포넌트는 전형적으로 대응하는 객체의 포지션/위치 또는 배향/방향의 개별 특성을 기술한다. 물론, 많은 상황에서, 배치 또는 포즈는, 예를 들어, 하나 이상의 컴포넌트들이 고정되거나 무관한 것으로 간주되는 경우 더 적은 컴포넌트들로 표현되거나 더 적은 컴포넌트들을 갖는 것으로 간주될 수 있다(예컨대, 모든 객체들이 동일한 높이에 있고 수평 배향을 갖는 것으로 간주되는 경우, 4개의 컴포넌트들은 객체의 포즈의 전체 표현을 제공할 수 있다). 이하에서, 포즈라는 용어는 1 내지 6개의 값들(최대 가능한 자유도에 대응함)에 의해 표현될 수 있는 포지션 및/또는 배향을 지칭하는 데 사용된다. 본 설명은 포즈가 최대 자유도 - 즉, 포지션 및 배향 각각의 3 자유도가 총 6 자유도를 야기함 - 을 갖는 실시예들 및 예들에 초점을 맞출 것이다. 따라서, 포즈는 6 자유도를 표현하는 6개의 값들의

벡터 또는 세트에 의해 표현될 수 있고, 따라서, 포즈 벡터는 3차원 포지션 및/또는 3차원 방향 표시를 제공할 수 있다. 그러나, 다른 실시예들에서는 포즈가 더 적은 값들에 의해 표현될 수 있다는 것이 이해될 것이다.

[0073] 장치는 N-차원 공간(여기서 N은 정수임)에서 루트들의 세트에 대한 루트 데이터를 저장하도록 배열되는 저장소(101)를 포함한다. N-차원 공간은 가상 세계의 공간일 수 있고, 특히 공간 내의 사용자의 위치들에 대응하는 3차원 공간일 수 있다. 일부 실시예들에서, 치수들 중 하나 이상은 가상 공간에서의 뷰 배향에 대응할 수 있다. 따라서, 일부 실시예들에서, 사용자는 3 자유도(3 DoF)에 의해 표현될 수 있고, 다른 실시예들에서는, 예컨대 3차원 위치 및 3차원 배향에 대응하는 6 자유도(6 DoF)에 의해 표현될 수 있다. N-차원 공간은 공간적 포즈들을 표현하고, 구체적으로 가상 세계에서의 공간적 치수들을 표현하는 N-차원 공간에서의 포즈들을 표현한다.

[0074] 따라서, 저장소(101)는 N-차원 공간에서 루트들(예컨대, 경로들 또는 궤적들)을 저장한다. 루트들 각각은 가상 현실 세계를 통한 경로 또는 궤적을 제공한다. 추가로, 저장되어 있는 루트들의 세트 각각은 비디오 아이템과 연관되고/그에 링크된다. 따라서, 각 루트는 루트의 포지션들 및/또는 배향들에 대응하는 이미지들의 시퀀스를 포함하는 비디오 아이템에 링크된다. 동등하게, 각각의 비디오 아이템은 루트를 따라 뷰들에 대응하는 이미지들의 시퀀스를 제공할 수 있다. 따라서, 주어진 루트에 대한 비디오 아이템은, 루트를 따라 움직일 때의 뷰들에 대응하는 이미지들의 시퀀스를 포함한다. 따라서, 정적 이미지, 포지션, 또는 뷰에 대응하기보다는 오히려, 특정 루트에 대한 비디오 아이템은 루트를 따라 움직일 때 카메라에 의해 캡처되거나, 또는 동등하게는 루트를 따라 움직이는 가상 사용자가 경험하게 될 이미지들을 제공한다. 각각의 비디오 아이템은 링크된 루트의 카메라 포즈들에 대한 이미지 및 깊이 정보 둘 모두를 포함하는 프레임들을 포함하고, 각각의 루트는 N-차원 공간을 통한 포즈들(카메라 포즈들)의 대응하는 카메라 궤적을 기술한다.

[0075] 각각의 비디오 아이템은 루트를 따른 포지션들의 시퀀스에 대응하는 이미지들의 시퀀스를 포함할 수 있다. 비디오 아이템을 재생하는 것은 루트를 따라 움직이는 사람 또는 카메라의 뷰들을 제공할 것이다. 루트들 각각은 루트들을 따른 포지션들의 시퀀스에 의해 표현될 수 있다. 일부 실시예들에서, 명시적으로 정의된 포지션들만이 고려될 수 있는 반면(예를 들어, 하나의 포지션이 각각의 프레임에 대해 포함될 수 있음), 다른 실시예들에서는 저장소에 저장된 명시적인 값들 사이에 있는 포지션들이 보간에 의해 결정될 수 있다. 다른 실시예들에서, 각각의 루트는 예를 들어, 시간 변수의 함수로서, 예컨대 3개의 포지션 값들을 제공하는 수학적 함수에 의해 정의될 수 있다. 시간 변수는, 예컨대, 대응하는 비디오 아이템에 대한 프레임 시간들에 대응하는 시간 세그먼트들로 양자화될 수 있다.

[0076] 이 예에서, 저장소(101)는, 클라이언트 디바이스(201)를 네트워크(205)(이는 하기의 설명에서 인터넷인 것으로 간주될 것임)에 인터페이스하는 네트워크 인터페이스(103)에 커플링된다. 네트워크 인터페이스(103)는 특히 클라이언트 디바이스(201)가 서버(203)에 접속되는 기능을 제공할 수 있다.

[0077] 이 예에서, 서버(203)는 복수의 사용자들을 동시에 서빙할 수 있는 공통 서버이다. 서버(203)의 일부 요소들이 도 3에 예시되어 있다. 이 예에서, 서버(203)는 서버(203)를 네트워크(205)에 인터페이스하도록 배열되는 서버 네트워크 인터페이스(301)를 포함하여, 이에 의해 그것이 클라이언트 디바이스(201)와 통신할 수 있게 한다. 서버 네트워크 인터페이스(301)는 서버(203)의 동작의 대부분을 제어하고 수행하도록 배열되는 서버 제어기(303)에 커플링된다. 서버 제어기(303)는 비디오 저장소(305) 및 루트 저장소(307)에 커플링된다.

[0078] 이 예에서, 루트 저장소(307)는 루트들의 세트에 대한 루트 데이터를 저장하고, 비디오 저장소(305)는 대응하는 링크된 비디오 아이템들을 저장한다. 서버(203)는, 구체적으로, 앞서 기술된 바와 같이 연관된 비디오 아이템들을 갖는 루트들의 세트에 의해 가상 현실 세계를 표현할 수 있다. 예를 들어, 가상 환경은, 가상 환경의 N-차원 공간을 통한 다수의 (어쩌면 비교적 큰 수의) 루트들에 의해 표현될 수 있다. 루트들 각각에 대해, 루트들을 따른 뷰들을 표현하는 비디오 아이템이 저장되는데, 즉, 비디오 아이템은 루트를 따라 움직일 때 카메라 또는 사용자의 뷰들을 표현하도록 생성될 수 있다.

[0079] 많은 실시예들에서, 가상 환경은 현실 환경의 표현일 수 있고, 비디오 아이템들 및 루트들은 현실 환경에서의 기록/캡처에 의해 생성될 수 있다. 예를 들어, 포지션 인식 비디오 카메라가 현실 환경을 통해 운반될 수 있고, 캡처된 비디오는 비디오 아이템으로서 저장될 수 있고, 대응하는 포지션 정보는 루트로서 저장될 수 있다. 일부 실시예들에서, 비디오 아이템 및 루트 데이터는 가상 환경의 가상 모델로부터, 예컨대, 가상 카메라를 가상 환경을 통하여 이동시킴으로써 생성될 수 있다.

[0080] 저장된 비디오 아이템들은 루트들을 따른 뷰들을 표현하는 프레임들/이미지들의 시퀀스를 포함한다. 이 접근법의 특정 이점은, 비디오 인코딩 기법들이 프레임들/이미지들을 인코딩하여 이에 의해 비디오 아이템들을 생성하

는 데 사용될 수 있다는 것이다. 이는 매우 효율적인 인코딩을 허용하여, 주어진 이미지 품질에 대해 더 낮은 데이터 레이트들을 야기하고, 이에 의해 전형적으로 개별 이미지들을 인코딩할 때보다 훨씬 더 효율적인 인코딩을 허용한다. 구체적으로, 모션 추정 및 (전형적으로 양방향) 예측 비디오 인코딩과 같은 비디오 인코딩 기법들이 데이터의 양을 감소시키는 데 사용될 수 있다. 특정 실시예에서, 프레임들은 I-프레임들(인트라-프레임들 - 임의의 다른 프레임들에 의존하지 않고서 인코딩됨)로서 또는

- [0081] B-프레임들(양방향 프레임들 - I-프레임들에 대해, 구체적으로 I-프레임들로부터의 모션 추정에 기초하여 코딩된 잔차 프레임들)로서 인코딩된다. 추가로, 감소된 저장 및 통신 요건들에 더하여, 비디오 인코딩의 사용은 효율적인 기성의(off-the-shelf) 인코딩 및 디코딩 기능(전용 하드웨어를 포함함)이 사용될 수 있게 한다. 사용될 수 있는 비디오 인코딩 기법들은, 예를 들어 H.264 또는 MPEG2 인코딩을 포함한다.
- [0082] 비디오 아이터మ్들은 이미지 및 깊이 정보 둘 모두를 포함한다. 구체적으로, 비디오 아이터మ్들은, 광 세기 이미지들/프레임들이 프레임들에 대한 깊이 정보와 연관되는 3D 비디오 아이터మ్들일 수 있다. 스테레오 이미지들, 명시적인 깊이 또는 시차 맵들 등을 포함하는 깊이 정보를 제공하기 위한 상이한 접근법들이 사용될 수 있다는 것이 이해될 것이다. 깊이 정보는 뷰 생성에 있어서의 유연성을 허용하고, 구체적으로, 루트들/비디오 아이터మ్들이 이미지를 제공하는 포지션들과 함께 정확하게 일렬로 맞춰지지 않는 포지션들(/배향들)에 대응하는 뷰 이미지들의 생성을 허용하거나, 용이하게 하거나, 개선한다.
- [0083] 따라서, 이 예에서, 가상 환경은, 예컨대, 현실 세계 환경을 통해 카메라를 이동시킴으로써 캡처된 공간적 루트들의 전형적으로 비교적 성긴(sparse) 세트에 의해 표현된다. 각각의 루트는 비디오 아이터మ్에 링크될 뿐만 아니라, 비디오 아이터మ్에 대한 포지션 정보를 제공하고 구체적으로 비디오 아이터మ్의 각각의 프레임에 대한 포지션 표시를 제공할 수 있는 루트 데이터의 형태의 메타데이터에 링크된다. 도 4는 2차원 환경이 다수의 (어쩌면 상당히 임의적인) 루트들에 의해 어떻게 커버될 수 있는지의 일례를 예시한다.
- [0084] 이 예에서, 클라이언트 디바이스(201)는 서버(203)로부터 루트 데이터 및 비디오 아이터మ్들을 요청하도록 배열되고, 이에 응답하여, 서버(203)는 요청된 데이터를 제공하도록 진행할 것이다.
- [0085] 그러나, 이 예에서, 클라이언트 디바이스(201)는 단순히 모든 데이터를 다운로드하고 그것을 국부적으로 프로세싱하는 것이 아니라 오히려 초기에 루트 데이터만을 인출하며, 루트 데이터는 이어서 어느 루트(들)가 원하는 포지션들에 가까운지를 동적으로 결정하는 데 사용된다. 이어서, 선택된 루트들에 대한 비디오 아이터మ్들은 서버(203)로부터 선택적으로 인출되고, 적합한 뷰들을 생성하는 데 사용된다.
- [0086] 따라서, 특정 예에서, 클라이언트 디바이스(201)가 가상 현실 애플리케이션을 개시할 때, 그것은 서버(203)로부터 루트 데이터를 요청한다. 이에 응답하여, 서버(203)는 루트 저장소(307)로부터 루트 데이터를 추출하고 이를 클라이언트 디바이스(201)로 송신한다. 그러나, 초기에는, 어떠한 비디오 데이터도 제공되지 않고, 루트 데이터만이 네트워크(205)를 통해 송신된다.
- [0087] 더 상세하게는, 새로운 가상 현실 애플리케이션이 초기화될 때, 네트워크 인터페이스(103)는 루트 저장소(307)에 저장된 루트 데이터를 송신하라는 요청을 서버(203)로 송신한다. 이에 응답하여, 서버(203)는 이러한 루트 데이터를 송신한다. 루트 데이터는 네트워크 인터페이스(103)에 의해 수신되고 저장소(101)에 저장된다. 따라서, 초기 다운로드 후에, 저장소(101)는 N-차원 공간에서의 루트들의 세트에 대한 루트 데이터의 국부 복제본을 저장한다. 루트들 각각이 루트의 뷰들을 표현하는 대응하는 비디오 아이터మ్과 연관(/그에 링크)되지만, 이러한 비디오 아이터మ్들은 다운로드되지 않고 서버에서 유지된다.
- [0088] 클라이언트 디바이스(201)는 관찰자 포지션 표시를 수신하도록 배열되는 포지션 입력부(105)를 추가로 포함한다. 관찰자 포지션 표시는 가상 환경의 N-차원 공간 내의 관찰자의 포지션을 나타내며, 따라서 가상 환경의 원하는 뷰를 나타낸다. 일부 실시예들에서, 관찰자 포지션 표시는, 예컨대 사용자가 직접 포지션을 제공함으로써, 또는 예컨대 조이스틱 또는 마우스를 제어하는 것과 같은, 특정 사용자 입력에 응답하여 생성될 수 있다. 다른 실시예들에서, 관찰자 포지션 표시는, 예컨대 클라이언트 디바이스(201) 상에서 또한 실행되는 애플리케이션에 의해 생성될 수 있다. 예를 들어, 관찰자 포지션 표시는 게임 애플리케이션에 의해 생성될 수 있다.
- [0089] 관찰자 포지션 표시는 관찰자에 대한 3차원 위치 및 3차원 배향을 제공하는 완전한 6차원 포즈 표시일 수 있다. 다른 실시예들에서, 예컨대 2차원 배향과 함께 2차원 위치만이 제공될 수 있다(예컨대, 관찰자가 항상 동일한 높이에 있는 것으로 고려되고 루트들이 본질적으로 2차원 루트들일 때).
- [0090] 클라이언트 디바이스(201)는 저장소(101) 및 포지션 입력부(105)에 커플링되는 선택기(107)를 추가로 포함한다.

선택기(107)는 수신된 관찰자 포지션 표시들에 기초하여 루트들의 저장된 세트로부터 루트들을 선택하도록 배열된다. 구체적으로, 주어진 관찰자 포지션 표시에 대해, 선택기(107)는 루트들의 세트의 루트들의 포지션들 및 관찰자 포지션 표시에 의존하는 거리 메트릭에 의존하는 선택 기준을 평가하도록 배열된다. 따라서, 관찰자 포지션 표시 및 저장된 루트들에 대한 포지션 데이터에 기초하여, 거리 메트릭이 평가되고, 하나 이상의 루트들이 현재 관찰자 포지션 표시에 대한 뷰 이미지들을 생성하기에 적절한 것으로서 선택된다.

[0091] 클라이언트 디바이스(201)는 선택기(107) 및 네트워크 인터페이스(103)에 커플링되는 인출기(109)를 추가로 포함한다. 인출기(109)는 어떤 비디오 아이템(들)이 선택되었는지의 정보를 수신하고, 이에 응답하여, 네트워크 인터페이스(103)를 통해, 선택적 비디오 아이템(들)을 인출하기 위해 서버(203)와 통신한다.

[0092] 클라이언트 디바이스(201)는 제1 비디오 아이템으로부터의 프레임들의 제1 세트로부터 관찰자 포지션 표시에 대한 적어도 하나의 뷰 이미지를 생성하기 위한 이미지 생성기(111)를 추가로 포함한다. 인출된 비디오 아이템은 이미지 생성기(111)로 포워딩되고, 이미지 생성기는 관찰자 포지션 표시에 의해 나타내는 가상 환경에서의 포지션에서 관찰자가 보게 될 것에 대응하는 이미지를 생성하도록 진행된다. 뷰 이미지를 생성하기 위해 사용되는 프레임들의 선택은, 또한, 관찰자 포지션 표시 및 공간적 거리 메트릭에 기초할 수 있다. 실제로, 비디오 아이템에서의 프레임들이 루트들을 따른 포지션들을 갖기 때문에, 관찰자 포지션 표시의 포지션에 가장 가깝게 대응하는 프레임(들)이 (예컨대, 다른 파라미터들 또는 고려사항들을 조건으로) 선택될 수 있다.

[0093] 따라서, 주어진 관찰자 포지션 표시에 대해, 클라이언트 디바이스(201)는 적절한 비디오 아이템(또는 비디오 아이템들) 및 이 비디오 클립의 적절한 프레임(또는 프레임들)을 선택하고 선택된 프레임들을 사용하여 뷰 이미지를 생성하도록 진행할 수 있다. 일부 실시예들 및 시나리오들에서, 선택된 프레임들은 관찰자 포지션 표시에 대한 뷰 이미지에 직접 충분히 가깝게 대응할 수 있고, 생성된 뷰 이미지는 선택된 프레임의 사본으로서 직접 생성될 수 있다.

[0094] 그러나, 대부분의 실시예들 및 시나리오들에서, 프레임(들)의 포지션(배향을 포함함)과 관찰자 포지션 표시 사이의 차이들은 프레임들을 프로세싱함으로써 보상될 수 있다. 이러한 프로세싱은 구체적으로 깊이 정보에 기초한 뷰 변환일 수 있다. 그러한 뷰 변환을 위한 많은 접근법들이 당업자에게 알려져 있을 것이고 본 발명으로부터 벗어남이 없이 사용될 수 있다는 것이 이해될 것이다.

[0095] 따라서, 클라이언트 디바이스(201)는 수신된 관찰자 포지션 표시에 기초하여 뷰 이미지를 생성하도록 배열된다. 실제로, 기술된 프로세스는 반복되고, 따라서 클라이언트 디바이스(201)는 관찰자 포지션 표시들의 시퀀스에 응답하여 뷰 이미지들의 시퀀스를 생성할 수 있다. 클라이언트 디바이스(201)는 구체적으로 가상 환경을 통한 루트를 따라 움직이는 관찰자에 대한 뷰들의 시퀀스를 생성할 수 있다.

[0096] 생성된 이미지들은, 관찰자가 관찰자 포지션 표시들의 시퀀스에 의해 나타내는 궤적을 따라서 가상 환경을 통해 움직일 때 보게 될 뷰들을 제공할 수 있다. 그러나, 이 예에서, 이것은 개별 이미지들에 의존하기보다는 오히려 비디오 아이템 및 루트 기반 접근법으로부터 달성된다. 이는 훨씬 더 효율적인 접근법을 제공하고, 효율적인 인코딩, 용이해진 캡처, 감소된 복잡성, 감소된 저장 요건들 등을 허용한다. 더욱이, 이 시스템은, 요구되는 비디오 아이템들만이 인출되고 다운로드될 필요가 있기 때문에 클라이언트 디바이스(201)에서의 실질적으로 감소된 통신 요건들 및 저장 요건들을 허용할 수 있다. 따라서, 이 접근법은 루트 데이터 및 비디오 아이템 데이터의 처리, 및 구체적으로는 통신을 구별할 수 있다. 루트 데이터가 전형적으로 비디오 데이터보다 적어도 한 자릿수만큼 더 낮을 것이기 때문에, 이는 다운로드 통신 요건들의 상당한 감소를 제공할 수 있다.

[0097] 적절한 루트의 선택은 관찰자 포지션 표시(의 포지션)와 루트들의 포지션들 사이의 거리를 반영하는 거리 메트릭에 기초한다. 구체적으로, 거리 메트릭은 현재 선택되는 비디오 아이템/루트에 대한 현재 프레임의 포지션과 관찰자 포지션 표시 사이의 거리, 및 현재 선택되지 않은 루트들에 대한 루트들의 포지션들과 관찰자 포지션 표시 사이의 최소 거리를 반영할 수 있다. 거리들은 임의의 적합한 표준에 따라, 또는 예컨대, 적합한 표준의 단조 증가 함수에 의해 결정될 수 있다.

[0098] 그러한 거리 메트릭은 어떤 루트가 원하는 포지션에 가장 가까운 포지션을 갖는지를 나타냄으로써 적합한 루트의 양호한 표시를 제공할 수 있다. 그러나, 도 1의 클라이언트 디바이스(201)에서, 선택 기준은 단순히 거리 메트릭에 기초하는 것이 아니며, 단순히 원하는 시점에 가장 가까운 루트를 선택하는 것이 아니다. 오히려, 선택기(107)는 다른 루트들에 비해 현재 선택된 루트를 향해 편향되는 선택 기준을 사용하도록 배열된다. 따라서, 선택 기준은, 현재 루트 상에 유지되는 것, 및 그에 따라서 다른 비디오 아이템으로 점프하기보다는 오히려 동일한 비디오 아이템에 기초하여 뷰 이미지들을 생성하는 것을 향해 편향될 수 있다.

[0099] 도 1의 클라이언트 디바이스(201)에서, 선택기(107)의 선택 기준은, 현재 선택된 루트(및 따라서 비디오 아이템) 및 현재 선택되지 않은 다른 루트들(및 따라서 비디오 아이템들)에 대한 거리 메트릭에 대하여 비대칭이다. 특히, 사용되는 선택 기준은 다른 루트에 대한 거리 메트릭이, 다른 루트가 선택되기 전에, 즉, 클라이언트 디바이스(201)가 다른 루트/비디오 아이템으로 전환하기 전에, 현재 루트에 대한 거리 메트릭보다 소정 마진 또는 임계치만큼 더 작을 것을 요구할 것이다.

[0100] 2개의 평행 루트들(RT1, RT2)을 예시하는 도 5를 참조하여 특정 예가 제공될 수 있으며, 여기서 RT1은 대응하는 비디오 아이템에 기초하여 뷰 이미지들이 현재 생성되고 있는 현재 선택된 루트이다. 환경을 통과하는 관찰자의 루트(RT3)가, 2개의 캡처된 루트들(RT1, RT2) 사이에, 2개의 루트들까지의 다양한 거리들을 갖는 상태로 위치된다.

[0101] 이 예에 대한 선택 기준의 일례는,

$$r_{\text{selected}} = \begin{cases} \operatorname{argmin}_r(d_r) & \text{if } d_r > \Delta \\ r_{\text{current}} & \text{otherwise} \end{cases}$$

[0102] 일 수 있으며,

[0103] 여기서 r_{current} 는 현재 선택된 루트이고, d_r 은 루트(r)까지의 최단 거리이고, r_{selected} 는 선택된 루트이다. 이러한 선택 기준은, 상이한 루트 및 비디오로의 전환이 현재 루트로부터 임계 거리(Δ)보다 더 멀리 벗어날 때에만 일어난다는 결과를 갖는다. 따라서, 루트는 관찰자 포지션 표시까지의 가장 낮은 거리를 갖는 것으로 선택되지만, 현재 선택된 루트까지의 거리가 소정 임계치(Δ)를 초과하는 경우에만 선택된다. 도 5는 이러한 선택 기준이 적용되는 것을 예시한다. 이 예에서, 선택된 루트는, RT2가 일부 포지션들에서 관찰자 포지션 파라미터 벡터에 더 가까우에도 불구하고 전체에 걸쳐 RT1이다.

[0104] 현재 선택된 루트를 향한 선택 기준의 편향은 실제로는 상당한 이익을 제공할 수 있다. 첫째로, 인출되고 다운로드될 필요가 있는 비디오 아이템들의 개수가 전형적으로 실질적으로 감소될 수 있기 때문에, 그것은 감소된 통신 조건들을 갖는 더 효율적인 접근법을 제공하는 경향이 있다.

[0105] 추가로, 그것은 더 적은 인지가능한 아티팩트들을 갖는 실질적으로 더 일관된 출력을 제공할 수 있다. 실제로, 가상 환경의 더 안정적이고 시간적으로 그리고 공간적으로 일관된 경험이 제공되는 것으로 밝혀졌다.

[0106] 일부 실시예들에서, 루트 데이터는 루트들에 대한 포지션 정보로 제한되지 않고, 또한 루트들이 서로 어떻게 관련되는지의 정보를 제공하는 데이터를 포함한다.

[0107] 특히, 루트 데이터는 루트들 사이의 교차점들에 관한 정보를 추가로 포함할 수 있다. 예를 들어, 도 4에 예시된 바와 같이, 루트들은 많은 실시예들에서 비교적 랜덤한 루트들에 기초할 수 있고, 이는 많은 교차점들 및 루트들 사이의 중첩들을 야기할 수 있다. 실제로, 특정 루트들이 환경의 양호한 커버리지를 제공하기 위해 의도적으로 생성될 때 교차점들이 또한 종종 발생할 수 있다. 예를 들어, 루트들의 그리드 시스템을 생성하기 위해 다수의 평행 및 직교 루트들을 의도적으로 걷는 것에 의해 주어진 공간이 커버될 수 있다. 그러한 예에서, 상이한 직교 루트들 사이에 많은 교차점들이 있을 것이다.

[0108] 루트들 사이에 교차점이 존재하는 것을 고려하기 위한 상이한 기준들이 상이한 실시예들에서 사용될 수 있다. 2개의 루트들이 공통 포지션을 가질 때 2개의 루트들 사이에 교차점이 발생하는 것이 고려될 수 있다. 교차점은 T-접합부와 같은, 2개의 루트들 사이의 임의의 접합부, 또는 2개 초과 루트들 사이의 접합부를 포함한다. 일부 실시예들에서, 하나의 루트 상의 포지션과 다른 루트 상의 포지션 사이의 거리가 임계값 미만일 때 교차점/접합부가 존재하는 것으로 간주될 수 있다. 임계값은 전형적으로 매우 낮으며, 따라서 2개의 루트들이 적어도 하나의 포지션에서 실질적으로 중첩하고 있을 때에만 교차점이 존재하는 것으로 전형적으로 간주된다.

[0109] 많은 실시예들에서, 선택 기준은 저장된 교차점 포지션들에 추가로 의존한다. 구체적으로, 선택기(107)는 교차 포지션들에 대응하는 관찰자 포지션 표시들에 대한 현재 선택된 루트를 향한 편향을 감소시키도록 배열될 수 있다. 관찰자 포지션 표시는, 현재 관찰자 포지션 표시에 가장 잘 매칭되는 현재 루트 상의 포지션(예컨대, 최단 거리)이, 현재 루트가 다른 루트와 교차하는 교차 지점일 때, 관찰자 포지션 표시는 교차 포지션에 대응할 수 있다.

- [0110] 따라서, 일부 실시예들에서, 선택 기준은, 현재 루트와 새로운 루트 사이의 교차점에 있을 때(구체적으로, 현재 루트 상으로의 관찰자 포지션 표시의 맵핑이 교차 지점일 때), 교차점에 있지 않을 때보다 새로운 루트를 선택하는 것을 더 용이하게 할 것이다.
- [0111] 실제로, 일부 실시예들에서, 관찰자 포지션 표시가 교차 지점/포지션에 대응할 때 새로운 루트를 선택하는 것이 단지 가능할 수 있다. 그러한 실시예에서, 관찰자 포지션 표시는, 예를 들어 가상 환경을 통한 원하는 경로 또는 궤적에 따라 변경될 수 있다. 이러한 관찰자 포지션 표시는 현재 루트 상에 투영될 것이고, 현재 선택된 비디오의 대응하는 프레임들이 추출될 수 있고 관찰자 포지션 표시에 대응하는 뷰 이미지들을 생성하는 데 사용될 수 있다. 그러나, 교차 지점에 있지 않은 현재 선택된 루트 상의 투영된 포지션들의 경우, 선택기(107)는 새로운 루트를 선택하지 않을 수 있는데, 즉, 이 경우에 현재 선택된 루트를 향한 편향은 임의의 새로운 루트의 선택을 허용하지 않는다는 점에서 극단적이다. 그러나, 관찰자 포지션 표시가, 현재 선택된 루트 상의 투영된 또는 대응하는 포지션이 교차 지점에 매칭되는 값으로 변경될 때, 선택 기준은 이러한 교차점에서 임의의 다른 루트의 선택을 허용할 수 있다. 따라서, 현재 선택된 루트를 향한 극단적인 편향은 감소되고 아마도 완전히 제거되는데, 예컨대, 일부 실시예들에서, 선택기(107)는 단순히 관찰자 포지션 표시에 가장 가까운 교차점의 루트로서 이 루트를 선택할 수 있다.
- [0112] 따라서, 그러한 접근법은 클라이언트 디바이스(201)가 이용가능한 루트들로 "스냅(snap)"하고 교차 지점들에서 다른 루트로 단지 "점프"하게 할 수 있다. 이는 (다른 비디오 아이템들로부터) 다른 이용가능한 프레임들보다 관찰자 포지션 표시로부터 공간적으로 더 멀리 있는 프레임들에 기초하여 뷰 이미지들이 생성되게 할 수 있지만, 이 접근법은 많은 시나리오들에서 바람직한 사용자 경험을 제공할 수 있다. 특히, 그것은 생성된 출력 뷰 이미지 시퀀스에서의 개선된 일관성을 허용하고, 에러들 또는 아티팩트들의 수 및 중요성을 감소시킬 수 있다.
- [0113] 많은 실시예들에서, 다른 루트들의 선택은 교차 지점들로 제한되지 않고, 편향의 더 점진적인 변동이 수행된다. 많은 실시예들에서, 현재 선택된 루트를 향한 편향은 관찰자 포지션 표시와 교차 포지션들 사이의 거리를 나타내는 거리 메트릭에 의존할 수 있다. 거리 메트릭은 많은 실시예들에서 적합한 표준일 수 있거나, 또는 일부 실시예들에서, 거리는 루트를 따라, 그리고 구체적으로는 현재 선택된 루트를 따라 측정될 수 있다. 따라서, 관찰자 포지션 표시는 현재 선택된 루트(예컨대, 관찰자 포지션 표시까지의 최단 거리를 갖는 루트 상의 포지션) 상으로 투영될 수 있고, 이러한 투영된 포지션으로부터 교차 지점까지의 루트를 따른 거리가 결정될 수 있다.
- [0114] 대부분의 실시예들에서, 현재 선택된 루트를 향한 편향은 교차 지점까지의 거리를 감소시키기 위해 감소될 것이다. 그러한 예에서, 클라이언트 디바이스(201)는 여전히 현재 루트로 "스냅"하려고 할 것이지만, 이것이 교차 지점에 있지 않더라도 상이한 루트들 사이에서의 "점프"를 허용할 수 있다. 이러한 접근법은, 예컨대, 적절할 때 루트들 사이의 일부 "코너들의 컷팅(cutting)"을 허용할 수 있는 더 융통성 있는 접근법을 허용할 수 있다.
- [0115] 주어진 관찰자 포지션 표시에 대한 뷰 이미지를 생성하는 데 사용되는 프레임 또는 프레임들의 선택은, 상이한 고려사항들에 기초하여 (선택된 루트에 대응하는) 선택된 비디오 아이템으로부터 선택될 수 있다. 그러나, 프레임들은 전형적으로 관찰자 포지션 표시에 대한 제1 비디오 아이템의 프레임들에 대한 루트 포지션 데이터에 응답하여 이미지 생성기(111)에 의해 선택된다. 구체적으로, 선택된 비디오 아이템에서의 프레임들의 포지션들이 관찰자 포지션 표시와 비교되고 거리 메트릭이 결정될 수 있다. 일부 실시예들에서, 프레임 또는 프레임들은 단순히 관찰자 포지션 표시에 공간적으로 가장 가까운 프레임들로서 선택될 수 있다.
- [0116] 일부 실시예들에서, 프레임(들)의 선택은 시간적 태양을 추가로 고려할 수 있다. 구체적으로, 뷰 이미지를 생성하기 위해 사용될 프레임들이 현재 선택된 루트/비디오로부터의 프레임들일 때, 즉, 시스템이 동일한 루트 상에 유지되고 새로운 루트로 점프하지 않을 때, 선택은 (비디오 아이템에서) 현재 선택된 프레임들에 시간적으로 가까운 프레임들을 향해 프레임 선택을 편향시킬 수 있다. 따라서, 새로운 프레임들이 사용되고 있는 비디오 아이템의 변화 없이 선택될 때, 새로운 프레임들은 비디오 아이템에서 더 작은 시간적 거리를 향한 편향이 있도록 선택된다. 예를 들어, 프레임들의 두 세트가 관찰자 포지션 표시에 공간적으로 동일하게 떨어져 있는 경우, 이미지 생성기(111)는 이전에 선택된 프레임들에 가장 가까운 프레임들을 선택하도록 진행할 것이다.
- [0117] 구체적으로, 비디오 아이템의 현재 선택된 프레임까지의 감소하는 시간적 거리에 대해 증가하는 프레임 선택 메트릭이 사용될 수 있다. 구체적으로, 프레임 선택 메트릭은, 비디오 아이템에서의 프레임 수들 사이의 차이가 낮을수록 증가할 수 있다.

- [0118] 이 접근법은 많은 실시예들에서 개선된 성능을 제공할 수 있다. 특히, 그것은 현재 루트가 실제로 그 자신과 교차하는 많은 시나리오들에서 개선된 동작을 제공할 수 있다. 그러한 루트의 일례가 도 6에 제공되어 있다. 이 예에서, 루트는 자체적으로 중첩되어 이에 의해 교차 지점을 생성한다. 프레임들의 선택이 단순히 공간적 거리에 기초하는 경우, 교차 지점 주위에 유지되고 어쩌면 약간(예컨대, 랜덤하게) 움직이는 관찰자 포지션 표시는 뷰 이미지를 생성하기 위해 사용되는 프레임들의 세트에서의 빈번한 시프트(shift)를 야기할 수 있으며, 즉, 그것은 처음으로 교차 지점에 도달할 때로부터의 프레임들과 두 번째로 교차 지점에 도달할 때로부터의 프레임들 사이에서 전환될 수 있다. 따라서, 선택된 프레임들은 비디오 아이템 내의 매우 상이한 시간들에서 프레임들 사이에서 랜덤하게 그리고 어쩌면 빈번하게 전환될 것이다. 이것은 이미지 생성이 상이한 프레임들에 기초하게 할 것이고, 이것은 상당한 시간적 잡음을 초래할 가능성이 있다.
- [0119] 그러나, 동일한 비디오 아이템의 상이한 부분들로 시간적으로 점프하는 것에 대하여 적합하게 편향시킴으로써, 이것이 회피될 수 있고, 훨씬 더 일관되고 안정적인 출력이 생성될 수 있다. 예컨대, 도 6의 예에 대해, 사용자가 픽처의 저부로부터 교차 지점에 대응하는 영역을 향해 움직이는 경우, 사용자가 영역 내에 있을 때의 뷰 이미지들은 처음으로 교차 지점에 도달했을 때로부터의 프레임들에 기초할 것이고, 계속해서 그렇게 될 것이다. 그러나, 사용자가 픽처의 좌측으로부터 교차 지점에 대응하는 영역에 접근하는 경우, 사용자가 영역 내에 있을 때의 뷰 이미지들은 두 번째로 교차 지점에 도달했을 때로부터의 프레임들에 기초할 것이고, 계속해서 그렇게 될 것이다.
- [0120] 기술된 시스템은 가상 공간을 통한 루트들(/경로들/궤적들)에 대응하는 하나의 (또는 그 이상의) 비디오 아이템들을 사용하여 적합한 뷰 이미지들을 생성하는 것에 기초한다. 그러한 시스템에서의 중요한 고려사항은, 어떤 비디오 아이템을 사용할지 및 구체적으로는 루트들/비디오 아이템들을 언제 전환할지를 선택하는 방법이다.
- [0121] 클라이언트 디바이스(201)에서, 이 접근법은 동일한 비디오로부터 뷰 이미지들을 렌더링하고 새로운 비디오 아이템들로 전환하는 것에 대하여 편향시키려고 한다. 유사한 방식으로, 비디오 아이템 내에서 어떤 프레임들을 사용할지를 선택할 때, 클라이언트 디바이스(201)는 시간적 점프들에 대하여 편향된다.
- [0122] 이는 다수의 이점들을 제공하며, 다양한 고려사항들에 기초한다. 예를 들어:
- [0123] 1. 전형적으로, 뷰 합성에 사용되는 시차/깊이는 스테레오 카메라를 사용하여 추정되고, 시간적 깊이 필터링을 사용하여 시간적 안정성이 강제된다. 깊이 값들이 부정확할 수 있지만, 그들은 전형적으로 하나의 프레임으로부터 다음 프레임까지 일관되며, 이는 시간적 아티팩트들이 종종 회피되는 것을 의미한다.
- [0124] 2. 단일 비디오에서의 인접한 프레임들에 대응하는 가상 카메라 포지션들은 전형적으로 서로 가깝다. 예를 들어, 초당 30 프레임에서, 가상 카메라들은 3 mm 내지 1 cm 분리되어 있을 수 있다. 상이한 비디오들로부터의 가상 카메라 포지션들은 20 cm 이상만큼 멀리 떨어져 있을 수 있는데, 이는 그러한 큰 거리에 걸쳐 기준 시점을 갑자기 시프트시키는 것이 아티팩트들을 초래할 가능성이 있음을 의미한다.
- [0125] 3. 빈번하게 전환하는 것은 코딩 오버헤드를 증가시킬 것인데, 그 이유는 단일 비디오 스트림 내의 각각의 큰 점프 또는 비디오들 사이의 각각의 점프가 인트라 코딩된 프레임의 다운로드 및 디코딩을 요구할 것이기 때문이다.
- [0126] 따라서, 이 접근법은, 환경에서의 특정 포즈로 제한되지 않는 사용자에게 대한 상이한 뷰 포즈들에 대해 이미지들이 생성될 수 있도록 환경을 표현하는 것에 관한 것이다.
- [0127] 이 분야에서 그리고 그러한 애플리케이션들에 대한 주요 과제들 중 하나는, 상이한 포즈들에 대해 고품질 이미지들이 생성되게 하지만 과도한 메모리를 요구하지 않고서 저장될 수 있고, 과도한 복잡성 없이 그리고 과도한 계산 리소스를 요구하지 않고서 적합한 프로세싱 및 이미지 생성을 허용하는 포맷으로 있고, 클라이언트/서버 애플리케이션들에서 통신 요건들을 감소시킬 수 있고, 등등을 하는 방식으로 환경들을 효과적으로 표현하고 캡처하는 방식이다. 다른 매우 실질적인 과제는, 바람직하게는 너무 비싸거나 특화된 회로부를 요구하지 않고, 성가시거나 힘들지 않은 실제적인 캡처 동작을 가능하게 하는 것이다.
- [0128] 종래의 접근법은 단순히, 전형적으로 미리결정된 뷰 포즈들로부터 장면의 개별 픽처들(이미지들 및 깊이)을 촬영하고 이어서 이러한 이미지들 및 깊이 픽처들을 저장하는 것이다. 전형적으로, 캡처 포즈들은 미리결정되고, 사용자는 카메라를 새로운 포즈로 수동으로 움직이고, 픽처를 촬영하고, 그것을 다음 포즈로 움직이고, 다른 픽처를 촬영하고 등등을 해야 한다. 이는 매우 시간 소모적이고 부담이 되는 프로세스이다. 대안적으로, 다수의 카메라들이 사용될 수 있지만, 이는 고가일 뿐만 아니라 매우 번거로운데, 그 이유는 모든 카메라들이 정확한

포즈로 주의깊게 위치설정될 필요가 있기 때문이다. 또한, 정확한 동작을 위해 다수의 픽처들이 요구되고, 따라서 다수의 개별 픽처들이 액세스가능한 방식으로 저장될 필요가 있다.

- [0129] 종래의 접근법들에서, 주어진 뷰 포즈에 대한 뷰 이미지는, 이어서, 잘 알려진 시점 시프트 알고리즘들 중 하나를 사용하여 이용가능하고 저장된 픽처들로부터의 시점 전환에 의해 생성된다.
- [0130] 기술된 접근법에서, 장면은 개별 픽처들에 의해 표현되지 않고, 복수의 루트들 및 링크된 비디오 아이터들에 의해 표현된다. 예를 들어, 이 접근법은, 한 명 이상의 사람들이 단순히 연관된 포지션을 갖는 비디오를 캡처하는 주어진 환경에서 여기저기 움직일 수 있는 캡처 프로세스를 허용한다. 예를 들어, 사람들의 그룹은 단순히 비디오들을 캡처하는 장면 주위를 걸을 수 있는데, 그 이유는 그들이, 예컨대 랜덤 루트들을 따라 걷기 때문이다. 이어서, 연관된 루트 데이터를 갖는 비디오 아이터들은 연관된 비디오 아이터들을 갖는 복수의 루트들의 세트를 제공하도록 컴파일되어, 이에 의해 장면의 표현을 제공할 수 있다.
- [0131] 따라서, 이 접근법은 훨씬 단순화된 캡처 동작을 허용할 수 있고, 예를 들어, 비디오를 이용하여 다수의 랜덤 루트들을 캡처함으로써 캡처가 단순히 수행되게 할 수 있다. 따라서, 장면의 충분히 정확한 표현을 생성하기 위해 개별 이미지들을 사용하여 세심하고 계획된 캡처를 요구하기보다는 오히려, 루트들 및 비디오 스트림들의 사용은 훨씬 더 많은 양의 정보가 캡처될 수 있게 하여, 이에 의해 종종 유리한 포지션들에서 엄격한 제어된 이미지 캡처에 대한 필요성을 제거한다.
- [0132] 구체적으로, 장면 표현을 단순히 개별 이미지들에 기초하기보다는 오히려 비디오 아이터들(및 연관된 루트 정보)에 기초함으로써, 메모리/저장 요건들은 실질적으로 감소될 수 있고/있거나 저장될 수 있는 정보의 양이 훨씬 더 많을 수 있다. 비디오의 인코딩은 개별 이미지들의 인코딩보다 훨씬 더 효율적이다. 비디오 인코딩은 요구되는 데이터의 양을 감소시키기 위해 이미지 내에서 리던던시(redundancy)를 이용할 뿐만 아니라 (그리고 종종 훨씬 더 효과적으로) (예컨대, 모션 추정을 사용하여) 각각의 프레임에 대한 데이터 양을 실질적으로 감소시키기 위해 상이한 프레임들 사이의 리던던시를 이용한다. 예컨대, 가상 현실 애플리케이션들(예컨대, 정적 장면들)에 대한 많은 전형적인 장면들의 경우, 상이한 프레임들 사이의 상관관계는 데이터의 엄청난 감소가 달성될 수 있을 정도로 매우 높다.
- [0133] 따라서, 별개의 이미지들을 사용하는 것과 비교하여 연관된 루트들을 갖는 비디오 아이터들의 사용의 하나의 효과는, 그것이 실제로 완전히 상이한 캡처 동작이 수행되게 하는 훨씬 더 효율적인 압축을 허용한다는 것이며, 여기서 표현은 주의깊게 선택된 포지션들로부터의 이미지들에 기초하지 않고, 오히려 전형적으로 실질적으로 랜덤일 수 있는 복수의(그리고 전형적으로 비교적 큰 수의) 루트들에 기초한다. 이 접근법은 전형적으로 저장 요건들을 감소시키고, 캡처 표현의 복잡성 및 노력을 매우 실질적으로 감소시킨다.
- [0134] 그러나, 기술된 접근법은 단순히 복수의 링크된 루트들 및 비디오 아이터들에 의해 환경의 표현과 관련이 있는 것이 아니다. 오히려, 그것은 또한 그러한 표현으로부터 이미지들이 어떻게 생성되는지에 관련이 있다.
- [0135] 기술된 접근법에서, 뷰 이미지를 생성하는 적합한 베이스의 결정은 루트 기반 선택 프로세스에 의존한다. 선택 기준이 루트를 선택하는 데 사용되고, 이미지는 링크된 비디오의 적어도 하나의 프레임으로부터 생성된다. 선택은 관찰자 포지션으로부터의 거리에 기초하지만, 선택 기준은 현재 선택된 루트를 향한 편향을 추가로 포함한다.
- [0136] 따라서, 단지 가장 가까운 루트를 선택하기보다는 오히려, 선택 기준은 다른 루트들에 비해 현재 선택된 루트를 향한 편향을 추가로 포함한다. 따라서, 시스템은 단순히 가장 가까운 프레임들을 선택하는 것이 아니라, 동일한 루트에 속하는 프레임들의 시간적으로 일관된 선택에 기초하여 이미지들을 계속해서 생성하도록 편향될 수 있다. 시스템은 루트 상으로 효과적으로 "스냅"될 수 있고, 편향이 더 이상 사용자가 이러한 루트에 스냅된 상태를 유지하기에 충분하지 않을 때까지 (이미지들을 생성하기 위한 기초로서) 이러한 루트 상에 유지될 수 있다. 예를 들어, 선택 기준은, 상이한 루트로의 전환이 이러한 다른 루트까지의 거리가, 이를테면, 현재 루트까지의 거리의 절반 미만일 것을 요구하는 것을 필요로 함으로써 현재 루트를 향해 편향될 수 있다.
- [0137] 이 접근법은 복수의 루트들 및 링크된 비디오 아이터들에 의해 장면을 표현하는 특정 접근법에 크게 의존한다. 기술된 접근법의 이미지들(프레임들) 및 캡처 포지션들은 다른 이미지들 및 캡처 포지션들로부터 독립 및 격리되지 않고, 오히려 비디오 아이터들은 밀접하게 상관된 프레임들을 포함하고, 이들은 루트들과 밀접하게 상관된다. 실제로, 비디오 아이터들의 프레임들은 공간적 관계를 가지며 서로 밀접하게 상관되는데, 즉, 비디오 아이터들의 연속적인 프레임들은 전형적으로 높은 상관 정도를 갖는다. 유사하게, 루트들은 비디오 아이터들에 직접 관련되는 다수의 포지션들을 정의하며, 따라서, 루트들은 단순히 공간 내의 포지션들에 대한 정보를 제공하

는 것이 아니라, (본질적으로 시간적 표현인) 비디오 링크에서 프레임들에 대해 포지션들이 어떻게 변화하는지의 정보와 같은 시간적 정보를 제공한다.

[0138] 본 발명자들은 이러한 차이들을 깨달았고, 단지 현재 뷰에 대한 가장 가까운 이미지 정보를 선택하는 접근법이, 단순히 포지션들에 의해서만 정의되는 격리된 캡처들에 대한 것이 아니라, 대신에 밀접하게 관련된 특성들(구체적으로는, 비디오 아이템에서의 시간적 관계를 가짐)을 갖는 밀접하게 상관된 캡처들의 시퀀스를 제공하는 표현들 및 캡처 데이터에 대해 덜 적합하다는 것을 깨달았다.

[0139] 특히, 본 발명자들은, 실제로 가장 가까운 픽처(및 그에 따라 비디오 아이템 기반 캡처 내의 프레임)를 단순히 선택하는 것이 가능할 수 있는 반면, 이는 계산적으로 매우 복잡할 것이고, 프레임들과 캡처 포지션들 사이의 상관관계가 무시됨으로 인해 에러들을 야기하는 경향이 있을 것이라는 것을 깨달았다.

[0140] 구체적으로, 본 발명자들은 비디오 아이템들의 프레임들이 개별 프레임들로서 인코딩되지 않고 비디오 아이템에서의 다른 프레임들을 참조하여 인코딩된다는 것을 깨달았다. 전형적으로, 비디오 아이템은, 다른 프레임들을 참조하여 그리고 인트라 코딩된 프레임들보다 훨씬 더 높은 효율로 인코딩되는 실질적으로 훨씬 더 큰 수의 인터 코딩된 프레임들을 전형적으로 갖는 일부 인트라 코딩된 프레임들을 포함한다.

[0141] 이는 특히 개별 이미지들을 추출하는 것을 어렵게 하는데, 그 이유는 이들이 관련 인트라 코딩된 프레임들 양쪽 모두의 추출, 이들의 디코딩, 원하는 프레임이 예측되는 임의의 예측된 프레임들의 추출, 인트라 코딩된 프레임들을 사용한 이들의 디코딩, 원하는 프레임의 추출, 이의 디코딩, 및 마지막으로 이들 모두의 조합을 요구할 것이기 때문이다. 이는 매우 높은 계산 부담을 갖는다. 또한, 코딩 에러들, 잡음, 및 아티팩트들은 불가피하고, 이들은 결과적인 이미지 품질을 감소시킬 수 있다.

[0142] 본 발명자들은, 포지션뿐만 아니라 관찰자 포지션의 이력도 고려함으로써, 비디오 아이템들에 존재하는 시간적 및 프레임간 상관관계에 적응하고 이를 이용하는 것이 가능하다는 것을 추가로 깨달았다. 이는 단지 개별 포지션들을 고려하기보다는 오히려 루트들에 기초하여 선택을 행함으로써, 그리고 이러한 선택을 동일한 루트 상에 "고정시키는 것(sticking)"을 향해 추가로 편향시킴으로써 행해진다.

[0143] 이는, 특히, 비디오 인코딩의 특성들에 대해 훨씬 더 효율적인 접근법을 야기하고, 그것은 구체적으로 비디오 아이템들로부터 이미지들을 생성하는 문제들을 다룬다. 구체적으로, 실제로는 계산 부담을 매우 실질적으로 감소시키는 것이 밝혀졌다. 동일한 비디오 아이템/루트로부터 (종종 연속적인) 프레임들을 선택하는 루트 기반 선택 및 증가된 일관성은, 이들에 기초하여 인코딩되는 모든 인터 코딩된 프레임들에 대해 한번만 디코딩될 필요가 있는 인트라 코딩된 프레임들의 실질적인 재사용을 야기한다. 유사하게, 다른 프레임에 대한 중간의 예측된 프레임들을 생성하는 것에 대한 필요성은, 이러한 중간 프레임들이 이전 이미지들을 생성하는 데 이미 사용되어 왔다면 오버헤드를 도입하지 않을 것이다.

[0144] 또한, 불가피한 인코딩/디코딩 에러들 및 아티팩트들은 이 접근법에서 연속 이미지들에 대해 훨씬 더 일관되는 경향이 있을 것이다. 예를 들어, 인트라 코딩된 프레임에 대해 발생하는 인코딩/디코딩 아티팩트는, 그 인트라 코딩된 프레임을 사용하여 인코딩되는 동일한 비디오 아이템/루트로부터의 프레임들에 기초하여 생성되는 모든 이미지들에서 동일하게 표현될 것이다. 이는 종종, 잠재적으로 덜 가시적인 에러들 또는 아티팩트들 - 그러나 이들은 연속적인 생성된 이미지들 사이에서 덜 일관됨 - 보다 바람직할 수 있다.

[0145] 따라서, 루트들/비디오 아이템들의 편향된 루트 기반 선택에 대한 접근법은 실질적으로 개선된 성능을 제공하고, 특히, 더 낮은 계산 복잡성 및 리소스 사용량, 감소된 저장 요건들, 뷰 이미지들의 용이해진 생성, 개선된 이미지 품질을 허용한다. 그것은 또한, 예컨대 클라이언트 서버 기반 애플리케이션들에 대한 통신 요건들을 감소시킬 수 있다.

[0146] 일부 실시예들에서, 루트 데이터는 노드들로서의 교차점들 및 에지들로서의 루트들을 가짐으로써 표현되는 그래프를 포함할 수 있다. 집에 대응하는 가상 환경에 대한 일례가 도 7에 제공되어 있다.

[0147] 이 예에서, 각각의 노드는, 따라서, t 점합부들($n2$)을 포함하는, 루트들 사이의 교차점을 표현한다. 그러한 그래프에서, 교차점은 하나의 루트가 끝나고 다른 루트가 시작되는 포지션을 추가로 반영할 수 있다. 이 예에서, 각각의 루트는 따라서 2개의 노드들에 의해 경계지어지며, 이때 노드들은 전형적으로 복수의 루트들에 대해 공통이다. 이 접근법은, 루트들 및 비디오 아이템들이 전용 및 계획된 캡처에 의해 생성되거나 또는 캡처된 루트들 및 비디오 아이템들이 더 작은 루트들 및 비디오 아이템들로 분할되어 있는 시나리오에 대해 특히 적합하다. 이 그래프는, 예컨대, 이진 섹션 또는 텍스트(예컨대, xml)로서 인코딩될 수 있고, 애플리케이션이 초기화될 때

서버(203)로부터 클라이언트 디바이스(201)로 (적어도 부분적으로) 송신될 수 있다.

- [0148] 그래프의 사용은 많은 실시예들에서 효율적일 수 있고, 효율적인 동작을 제공할 수 있다. 예를 들어, 이 그래프 구조는, 주어진 루트의 끝 부분에서, 다수(0,1,...)의 정의된 다른 루트들이 시작될 수 있고 그에 따라 대응하는 비디오들의 인출 및 디코딩이 미리 개시되어 이에 의해 잠재적인 지연 및 중단들을 감소시킬 수 있다는 것이 결정되게 한다.
- [0149] 일부 실시예들에서, 루트 데이터는 루트들의 세트의 루트들 사이의 공간적 관계를 나타내는 관계 데이터를 추가로 포함할 수 있다. 이어서, 관계 데이터는 현재 관찰자 포지션 표시에 대한 루트의 선택에 사용될 수 있다.
- [0150] 특정 예로서, 관계 데이터는, 각각의 루트에 대해, 그것이 어떤 다른 루트들과 교차하는지의 표시를 제공할 수 있다. 이어서, 선택기(107)는, 일부 실시예들에서, 현재 선택된 루트에 대해 관계 데이터를 체크하여, 그것이 교차하고 함께 교차점들을 갖는 어떤 다른 루트들을 찾아낼 수 있다. 이어서, 선택기(107)는 상이한 루트를 선택할지 여부를 결정할 때, 예컨대 이러한 특정 루트들만을 고려하도록 진행할 수 있다.
- [0151] 일부 실시예들에서, 관계 데이터는 어떤 루트들이 서로 접합부들을 형성하는지의 표시를 제공할 수 있다. 그러한 데이터는 뷰 합성을 위해 적절한 비디오 및 프레임들을 선택하기 위해 선택기(107)에 의해 사용될 수 있다. 예를 들어, T-접합부에 근접할 때, 포지션 데이터에 기초하여, 사용자가 이러한 접합부에서 우회전을 취하도록 기울어져 있는지 또는 좌회전을 취하도록 기울어져 있는지 여부를 초기에 예상할 수 있다.
- [0152] 일부 실시예들에서, 관계 데이터는 일부 루트들이 동일한 경로에 대한 반대 이동 방향들에 대응한다는 표시를 제공할 수 있다. 그러한 데이터는 관찰자가 그의 경로 상에서, 예컨대 180도 회전하고 있을 때 방향을 전환하기 위해 선택기(107)에 의해 사용될 수 있다.
- [0153] 일부 실시예들에서, 관계 데이터는 실질적으로 서로 평행한 루트들의 표시를 제공할 수 있다. 그러한 데이터는 사용자가 향하고 있는 방향을 더 잘 표현하는 다른 경로로 전환하기로 결정하기 위해 선택기(107)에 의해 사용될 수 있다.
- [0154] 관계들은, 예를 들어 노드들로서의 루트들, 및 대응하는 2개의 노드들 사이의 관계를 나타내는 에지들을 갖는 그래프의 형태로 제공될 수 있다. 그러한 그래프의 일례가 도 8에 예시되어 있다.
- [0155] 앞서 언급된 바와 같이, 비디오 아이템들의 임의의 적합한 비디오 인코딩이 본 발명으로부터 벗어남이 없이 사용될 수 있다. 많은 실시예들에서, 비디오 아이템들은, 임의의 다른 프레임들을 고려하지 않고서 인코딩되는 인트라 코딩된 프레임들 및 적어도 하나의 다른 프레임에 대해 인코딩되는 예측된 프레임들 양쪽 모두를 사용하여 인코딩될 수 있다. 예측된 프레임들은 하나 이상의 인트라 코딩된 프레임들 및/또는 또는 더 많은 예측된 프레임들에 대해 인코딩될 수 있다. 예측된 프레임들은 하나 이상의 이전 프레임들(즉, 비디오 아이템에서 시간적으로 이전인 것)에 대해서만 인코딩되는 프레임들일 수 있거나, 또는 예컨대, 하나 이상의 이전 프레임들에 대해 그리고 하나 이상의 추후 프레임들에 대해(즉, 현재 프레임 전후 양쪽 모두의 프레임들에 대해) 인코딩되는 양방향 예측된(bi-predicted) 프레임들일 수 있다. 비디오 인코딩은 H.264 또는 H.265 비디오 인코딩일 수 있다.
- [0156] 많은 실시예들에서, 인코딩 프로세스가 B-프레임들(양방향 예측된 프레임들)과 관련하여 I-프레임들(인트라 코딩된 프레임들의 픽처들)만을 사용하는 것이 특히 유리할 수 있다. 이는 많은 실시예들에서 뷰 이미지들의 낮은 레이턴시 생성을 야기할 수 있다.
- [0157] 따라서, 비디오 아이템들은 인트라 코딩된 프레임들 및 예측된/인트라 코딩된 프레임들을 포함하는 인코딩된 비디오 아이템들일 수 있다. 예측된 프레임들은 일부 실시예들에서 오직 양방향으로 예측된/인코딩된 프레임들일 수 있다. 따라서, 뷰 이미지의 생성은, 이미지 생성기(111)가 뷰 이미지를 생성하기 위해 사용되는 프레임(들)을 디코딩하는 것을 포함할 수 있다.
- [0158] 기술된 접근법은, 특히, 복수의 루트들의 세트에 기초하여 가상 환경의 뷰들의 효율적인 캡처, 저장 및 재생을 제공하고자 한다. 개별 이미지들보다는 오히려 비디오 아이템들의 사용은 저장 요건들, 통신 대역폭 요건들 등을 감소시킨다. 이는, 비디오 인코딩이 별개의 이미지들의 인코딩보다 훨씬 더 효율적인 경향이 있기 때문에 달성된다. 특히, 예측된 프레임들의 사용은 프레임을 표현하는 데 요구되는 데이터의 양을 매우 유의하게 감소시킬 수 있다. 그러나, 비디오 캡처 및 코딩의 사용은 사용자가 자유롭게 움직일 수 있는 가상 환경을 표현하기 위해 반직관적이다. 예를 들어, 그러한 상황에서, 사용자는 비디오가 캡처된 방향의 반대 방향으로 움직여서, 이에 의해 비디오 콘텐츠의 공간적 및 시간적 관계들을 따르지 않을 뿐만 아니라 심지어 그의 정반대일

수 있다. 그러나, 본 발명자들은 이 접근법이 매우 효율적인 동작을 제공할 수 있음을 깨달았다.

- [0159] 동일한 루트 상에 유지되는 것을 향한 편향을 갖는 루트 기반 선택의 기술된 접근법은 개선된 동작을 제공한다. 특히, 본 발명자들은, 그러한 편향이 (더 시간적으로 안정적인 뷰 합성 품질을 야기하는 것 이외에) 사용자가 다른 루트로 전환되기(즉, 루트의 선택이 하나의 루트로부터 다른 루트로 변경될 때)보다는 오히려 동일한 루트에 계속해서 링크될 때, 인터 코딩되는(예측되는) 사용자 시점들에 대응하는 프레임들로 인해 훨씬 감소된 대역폭 요건(프레임을 인출/통신하기 위한 감소된 데이터 레이트)을 야기함을 깨달았다.
- [0160] 많은 실시예들에서, 선택은 루트를 선택할 때 비디오 인코딩을 고려할 수 있다. 구체적으로, 많은 실시예들에서, 선택기(107)는 인트라 코딩된 프레임들에 대응하는 관찰자 포지션 표시들에 대한 현재 선택된 루트를 향한 편향을 감소시키도록 배열될 수 있다. 선택기(107)는 인터 코딩된(예측된) 프레임들에 대응하는 관찰자 포지션 표시들에 비해 인트라 코딩된 프레임들에 대응하는 관찰자 포지션 표시들에 대한 현재 선택된 루트를 향한 편향을 감소시키도록 배열될 수 있다. 따라서, 선택기는 관찰자 포지션이 인터 코딩된/예측된 프레임에 대응할 때 동일한 루트에 대해 계속되기보다는 오히려 새로운 루트를 선택할 가능성이 더 클 수 있다.
- [0161] (편향에 영향을 주는) 고려되는 프레임은, 일부 실시예들에서, 현재/이전에 선택된 루트의 프레임일 수 있거나, 또는 새로운 후보 루트의 프레임일 수 있다. 따라서, 프레임은 현재/이전에 선택된 루트(이것이 선택되는 경우)에 대한 뷰 이미지를 생성하는 데 사용될 프레임일 수 있거나, 또는 후보 루트(이것이 선택되는 경우)에 대한 뷰 이미지를 생성하는 데 사용될 프레임일 수 있다. 따라서, 프레임은 현재 선택된 루트 또는 제1 루트(선택되는 루트)의 프레임일 수 있다. 많은 실시예들에서, 편향은, 타겟/후보 프레임이 인트라 코딩된 프레임인 경우 및 현재/이전에 선택된 루트 프레임이 인트라 코딩된 프레임인 경우 양쪽 모두에 대해 감소될 수 있다.
- [0162] 관찰자 포지션 표시에 대한 대응하는 프레임(들)은, 주어진 루트가 선택되는 경우 관찰자 포지션 표시에 대한 뷰 이미지를 생성하는 데 사용될 주어진 루트의 프레임(들)일 수 있다. 대응하는 프레임은 전형적으로, (전형적으로 루트를 선택하는 데 또한 사용되는 거리 메트릭을 사용하여) 관찰자 포지션 표시에 의해 나타내는 포즈/포지션에 가장 가까운 주어진 루트에서의 프레임일 수 있다.
- [0163] 많은 실시예들에서, 제1 프레임은 하나의 루트로부터 다른 루트로의 전환 이후에 인트라 코딩된 프레임으로서 선택될 수 있다. 새로운 루트로 전환될 때, 이미지 생성기(111)는, 그에 따라, 인트라 프레임을 먼저 선택하고 이를 사용하여 뷰 이미지를 생성하도록 진행할 수 있다.
- [0164] 많은 실시예들에서, 이러한 접근법들은 상이한 루트들 사이의 링크들을 확립하는 데 사용될 수 있다. 예를 들어, 도 9는 (작은 원들에 의해 나타낸) 인트라 코딩된 프레임들과 연관된 복수의 포지션들을 갖는 2개의 루트들(901, 903)을 예시한다. 그러한 실시예에서, 동일한 루트 상에 유지되기 위한 편향은 인트라 코딩된 프레임들에 대해 감소될 수 있고, 새로운 루트에 대해 선택되는 프레임은 다른 인트라 코딩된 프레임에 대응할 수 있다. 특히, 현재 루트를 향한 편향은 2개의 인트라 코딩된 프레임들 사이에서 발생하는 루트 전환들에 대해 (전형적으로 매우 실질적으로) 감소될 수 있다. 따라서, "점프들"은 도 9에 화살표로 나타낸 바와 같이 인트라 코딩된 프레임들 사이에서 발생하는 경향이 있을 것이다.
- [0165] 실제로, 이는 앞서 기술된 바와 같이 루트들을 표현하는 그래프에 그러한 고려사항들을 포함시킴으로써 달성될 수 있다. 실제로, 노드들은 그러한 그래프에서 교차부들/교차점들 근처에 생성될 수 있다. 또한, 노드들은 인트라 코딩된 프레임들에 대응하도록 선택될 수 있으며, 이때 그 사이에 있는 세그먼트들은 인터 코딩된 프레임들에 대응한다. 이어서, 그러한 노드들을 링크시키는, 예컨대, (예컨대, "교차부" 또는 "복귀 경로"와 같은) 그들 사이의 관계를 나타내는 추가 데이터가 제공될 수 있다.
- [0166] 후술되는 바와 같이, 시스템은 구체적으로 (예컨대, 비디오 아이템을 트랜스코딩함으로써) 루트들 사이의 교차점들에서 인트라 코딩된 프레임들을 생성할 수 있다.
- [0167] 많은 실시예들에서, 주어진 루트의 I-프레임들 사이에 비교적 큰 수의 B-프레임들을 사용하는 것이 또한 바람직할 수 있다. 이러한 접근법은, B-프레임들(이는 예측을 보상한 후의 잔차 프레임들임)이 종종 I-프레임들보다 더 압축되기 때문에 높은 효율을 가져온다. 높은 수의 B-프레임들의 사용은 전형적으로 정적 장면에서의 인접한 카메라 포지션들 사이의 높은 간섭성으로 인해 허용가능하다. I-프레임들 사이에 다수의 B-프레임들을 갖는 것은 또한 비디오 재생을 위해 매력적인데, 그 이유는 사용자의 주어진 이동에 대해 더 적은 I-프레임들이 디코딩될 필요가 있기 때문이다.
- [0168] 그러나, 새로운 루트의 선택으로 인해 비디오 아이템들을 변경할 때 다수의 B-프레임들을 사용하는 것에 대한

잠재적인 단점들을 경감시키기 위해, 이 접근법은 I-프레임들의 선택적 할당을 사용할 수 있다.

- [0169] 많은 실시예들에서, I-프레임들이 생성될 수 있고 교차점들에 대해 비디오 아이터널들 내에 포함될 수 있다. 따라서, 교차점들과 연관된 프레임들은 I-프레임들로서 인코딩될 수 있다. 이는 많은 실시예들에서 특히 매력적인 성능을 제공할 수 있는데, 그 이유는 그것이 루트들의 전환을 용이하게 할 수 있기 때문이다. 실제로, I-프레임들로서 인코딩된 교차점 프레임들을 갖는 것은, I-프레임으로 시작 및/또는 중지하는 연관된 비디오 아이터널들을 제공하고, 따라서 B-프레임들을 인코딩하기 위해, 추가적인 I-프레임들이 인코딩되어야 하는 것을 회피하는 교차점들에서 전환이 수행되게 하는 경향이 있을 것이다.
- [0170] 실제로, 교차점 프레임이 I-프레임으로서 인코딩되는 경우, 동일한 I-프레임이 잠재적으로, 이전에 선택된 루트 및 새롭게 선택된 루트 양쪽 모두에 대해 사용될 수 있다. 따라서, 이전에 선택된 루트로부터의 I-프레임이 새롭게 선택된 루트의 B-프레임을 디코딩하기 위해 새롭게 선택된 루트에 대한 I-프레임과 함께 사용될 수 있다. 일부 실시예들에서, 이것은, 정확하게 교차 지점에 있지는 않지만 잠재적으로 이에 가깝고 그에 따라 합리적인 추정치를 제공하는 이전 루트의 I-프레임을 사용하여 달성될 수 있다. 그러나, 바람직한 접근법에서, 이전에 선택된 루트의 I-프레임은 교차점에서의 프레임일 수 있고, 실제로 I-프레임은 심지어 이전에 선택된 루트와 연관된 비디오 아이터널 및 새롭게 선택된 루트와 연관된 비디오 아이터널에 대한 공통 프레임일 수 있다.
- [0171] 많은 실시예들에서, B-프레임들 및 I-프레임들의 유리한 구성 및 구조가 구현될 수 있다. 예를 들어, 서버(203)는 먼저 적합한 I-프레임을 제공할 수 있다. 이들은, 현재 관찰자 포지션 표시와 직접 연관되지 않더라도, 클라이언트 디바이스(201)에 의해 디코딩될 수 있다. 그러나, 디코딩된 I-프레임들은 이어서, 이들이 이미지 생성기(111)에 의해 선택되어 뷰 이미지의 생성에 사용될 때 적절한 B-프레임들의 신속한 디코딩을 허용하는 기준 프레임들을 제공할 수 있다.
- [0172] 실제로, 많은 실시예들에서, 계층적 비디오 인코딩 접근법이 사용될 수 있는데, 여기서 B-프레임들은 계층적 구조로 배열되어, 일부 B-프레임들이 다른 B-프레임들에 대한 기준 프레임들로서 추가로 사용될 수 있게 한다. 그러한 경우에, 적절한 B-프레임들은 다른 B-프레임들에 대한 기준 프레임들로서 사용될 준비를 갖추고서 추가로 디코딩될 수 있다.
- [0173] 따라서, 일부 실시예들에서, 인출기(109)는 그에 따라 디코더를 포함할 수 있으며, 디코더는 다수의 I 프레임들 및 잠재적으로 B-프레임들을, 이들이 이러한 프레임들에 대해 인코딩된 다른 B-프레임들의 디코딩에 사용될 준비를 갖추고서 디코딩한다.
- [0174] 따라서, 일부 실시예들에서, 큰 비율의 I-프레임을 회피함으로써 높은 코딩 효율을 유지하면서 비디오로부터의 랜덤 액세스 및 양방향 재생을 허용하는 B-프레임들의 계층적 순서화가 채용될 수 있다. 예를 들어, 항상, 디코더는 적어도 2개의 I-프레임들, 및 B-프레임들을 디코딩하기 위한 기준으로서 사용될 준비가 된 (복호화되고 디코딩된) 다수의 (0개 이상의) B-프레임들을 가질 수 있다.
- [0175] 그러한 접근법들은 H.264/MPEG-4 AVC 및 HEVC와 같은 더 새로운 비디오 인코딩 스킴들에 의해 지원될 수 있다.
- [0176] 위의 설명이 도 2에 의해 예시된 바와 같은 클라이언트 서버 접근법에 초점을 맞추고 있는 반면, 개념들은 이에 제한되지 않고 많은 다른 배열들 및 기능의 분포들에 대해 동등하게 적합하다는 것이 이해될 것이다. 예를 들어, 일부 실시예들에서, 비디오 아이터널들을 저장하는 비디오 저장소는 국부 저장소일 수 있고, 인출기(109)는 국부 저장소로부터 비디오를 인출할 수 있다. 일부 실시예들에서, 이 접근법은 비디오 아이터널들 및 루트 데이터 양쪽 모두를 저장하고 수신된 관찰자 포지션 표시에 응답하여 뷰 이미지들을 생성하는 단일 독립형 디바이스에서 사용될 수 있다.
- [0177] 위의 설명은 명료함을 위해 상이한 기능 회로들, 유닛들 및 프로세서들을 참조하여 본 발명의 실시예들을 설명하였음을 알 것이다. 그러나, 본 발명으로부터 벗어남이 없이 상이한 기능 회로들, 유닛들 또는 프로세서들 간의 기능의 임의의 적합한 분산이 사용될 수 있음이 명백할 것이다. 예를 들어, 별개의 프로세서들 또는 제어기들에 의해 수행되도록 예시된 기능이 동일한 프로세서 또는 제어기에 의해 수행될 수 있다. 따라서, 특정 기능 유닛들 또는 회로들에 대한 언급들은 오로지 엄격한 논리적 또는 물리적 구조 또는 조직을 나타내기보다는 설명된 기능을 제공하기 위한 적합한 수단에 대한 언급들로 간주되어야 한다.
- [0178] 본 발명은 하드웨어, 소프트웨어, 펌웨어 또는 이들의 임의의 조합을 포함한 임의의 적합한 형태로 구현될 수 있다. 본 발명은 선택적으로 하나 이상의 데이터 프로세서 및/또는 디지털 신호 프로세서 상에서 실행되는 컴퓨터 소프트웨어로서 적어도 부분적으로 구현될 수 있다. 본 발명의 실시예의 요소들 및 컴포넌트들은 임의의 적합한 방식으로 물리적으로, 기능적으로 그리고 논리적으로 구현될 수 있다. 실제로, 기능은 단일 유닛에서,

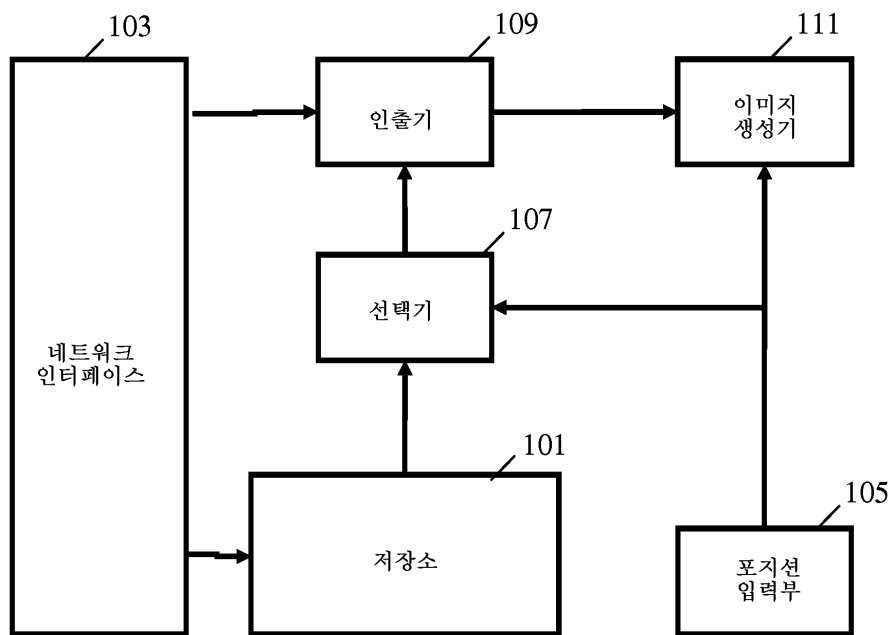
복수의 유닛에서, 또는 다른 기능 유닛들의 일부로서 구현될 수 있다. 그렇기 때문에, 본 발명은 단일 유닛에서 구현될 수 있거나, 상이한 유닛들, 회로들 및 프로세서들 간에 물리적으로 그리고 기능적으로 분산될 수 있다.

[0179] 본 발명이 일부 실시예들과 관련하여 설명되었지만, 본 명세서에 기재된 특정 형태로 제한되도록 의도되지 않는다. 오히려, 본 발명의 범주는 첨부된 청구범위에 의해서만 제한된다. 또한, 소정 특징이 특정 실시예들과 관련하여 설명되는 것처럼 보일 수 있지만, 당업자는 설명된 실시예들의 다양한 특징들이 본 발명에 따라 조합될 수 있음을 인식할 것이다. 청구범위에서, 용어 '포함하는'은 다른 요소들 또는 단계들의 존재를 배제하지 않는다.

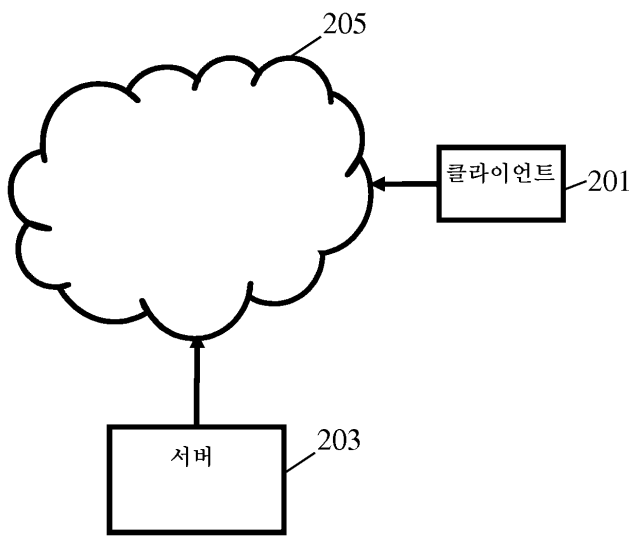
[0180] 또한, 개별적으로 열거되지만, 복수의 수단, 요소, 회로 또는 방법 단계는 예를 들어 단일 회로, 유닛 또는 프로세서에 의해 구현될 수 있다. 또한, 개별 특징들이 상이한 청구항들에 포함될 수 있지만, 이들은 아마도 유리하게 조합될 수 있으며, 상이한 청구항들에의 포함은 특징들의 조합이 실현 가능하지 않고/않거나 유리하지 않다는 것을 암시하지는 않는다. 또한, 하나의 카테고리의 청구항들에의 소정 특징의 포함은 이 카테고리로의 제한을 암시하는 것이 아니라, 오히려 그 특징이 적절한 바에 따라 다른 청구항 카테고리들에 동등하게 적용될 수 있음을 나타낸다. 또한, 청구항들에서의 특징들의 순서는 특징들이 작용되어야 하는 임의의 특정 순서를 암시하지는 않으며, 특히 방법 청구항에서의 개별 단계들의 순서는 단계들이 이 순서대로 수행되어야 함을 암시하지는 않는다. 오히려, 단계들은 임의의 적합한 순서로 수행될 수 있다. 또한, 단수형 언급은 복수를 배제하지 않는다. 따라서, "부정관사(a, an)", "제1", "제2" 등에 대한 언급은 복수를 배제하지 않는다. 청구범위에서의 참조 부호들은 단지 명료화 예로서 제공되며, 어떤 방식으로든 청구범위의 범주를 제한하는 것으로 해석되지 않아야 한다.

도면

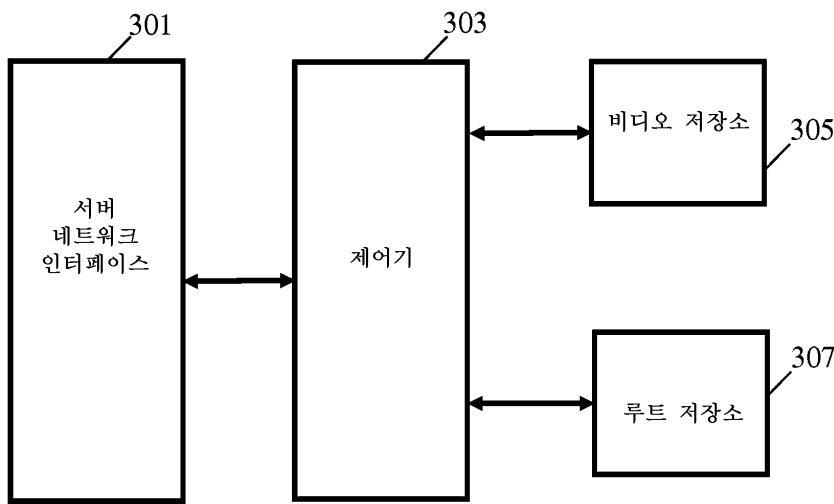
도면1



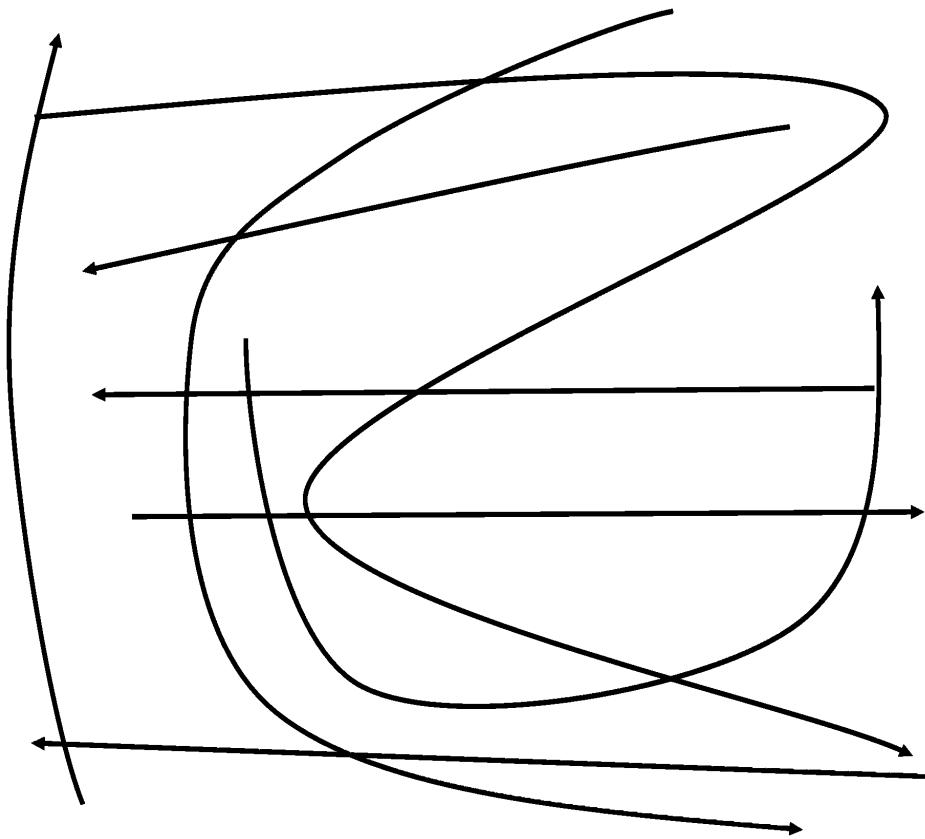
도면2



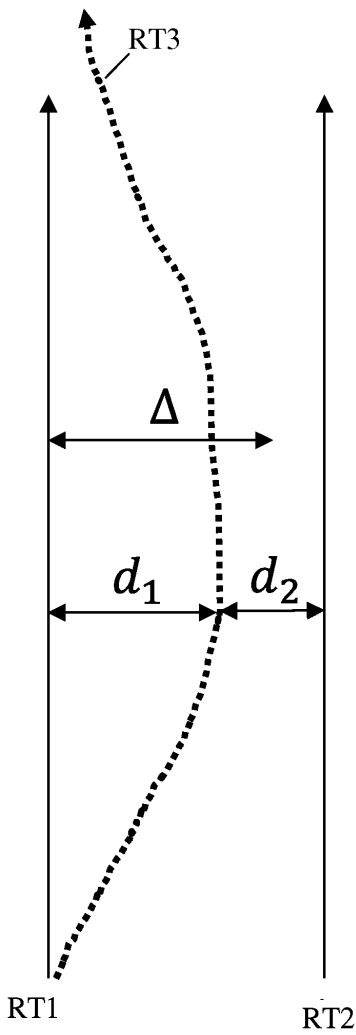
도면3



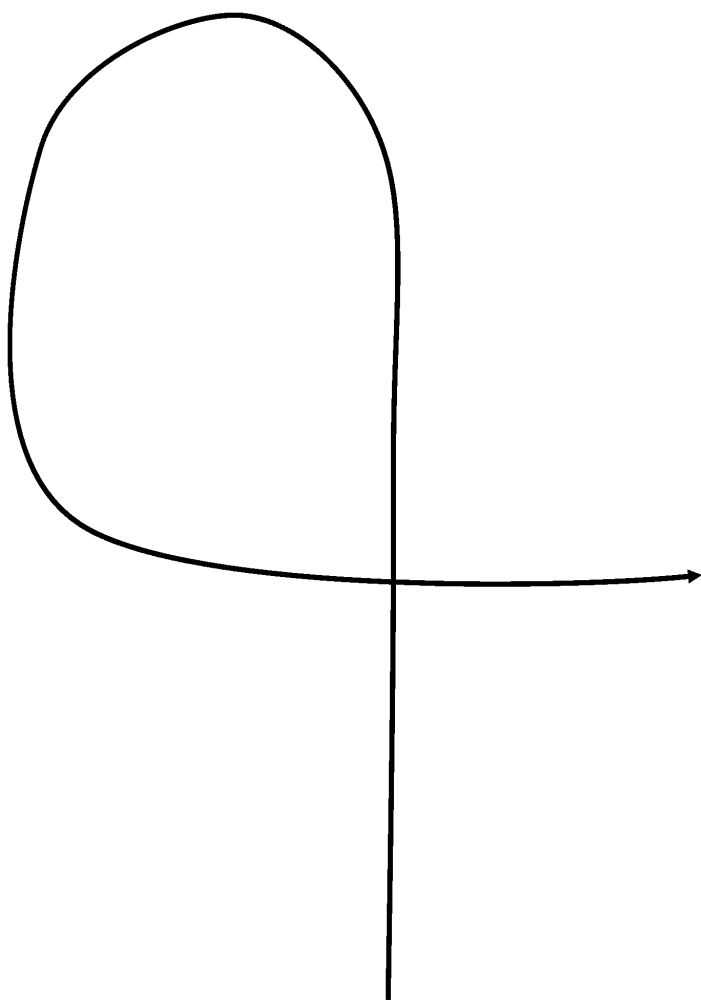
도면4



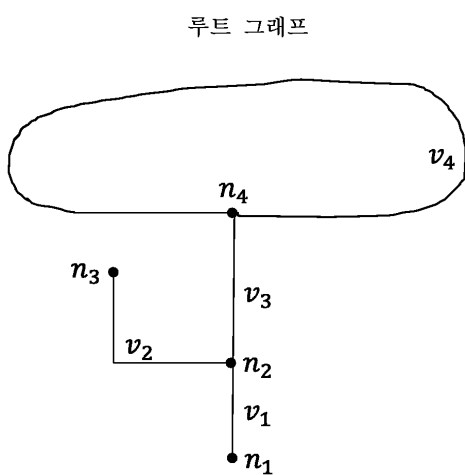
도면5



도면6



도면7



비디오 메타 데이터당

비디오	시작 노드	종료 노드	컨텍스트
v_1	n_1	n_2	복도
v_2	n_2	n_3	계단
v_3	n_2	n_4	복도
v_4	n_4	n_4	거실

프레임 메타 데이터당

각각의 프레임 k 에 대한 R_k, t_k

도면8

