



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2020-0024848
(43) 공개일자 2020년03월09일

- (51) 국제특허분류(Int. Cl.)
G06T 7/593 (2017.01) G06K 9/00 (2006.01)
G06K 9/46 (2006.01) G06K 9/62 (2006.01)
G06T 7/33 (2017.01) G06T 7/35 (2017.01)
G06T 7/579 (2017.01)
- (52) CPC특허분류
G06T 7/593 (2017.01)
G06K 9/00201 (2013.01)
- (21) 출원번호 10-2020-7001799
- (22) 출원일자(국제) 2018년06월27일
심사청구일자 없음
- (85) 번역문제출일자 2020년01월20일
- (86) 국제출원번호 PCT/US2018/039804
- (87) 국제공개번호 WO 2019/005999
국제공개일자 2019년01월03일
- (30) 우선권주장
62/526,203 2017년06월28일 미국(US)

- (71) 출원인
매직 립, 인코포레이티드
미국 플로리다 플랜타타운 웨스트 선라이즈 블러바드 7500 (우: 33322)
- (72) 발명자
디튼, 다니엘 엘.
미국 33322 플로리다 플랜테이션 웨스트 선라이즈 블러바드 7500
말리시예비츠, 토마스 잔
미국 33322 플로리다 플랜테이션 웨스트 선라이즈 블러바드 7500
라비노비치, 앤드류
미국 33322 플로리다 플랜테이션 웨스트 선라이즈 블러바드 7500
- (74) 대리인
특허법인 남앤남

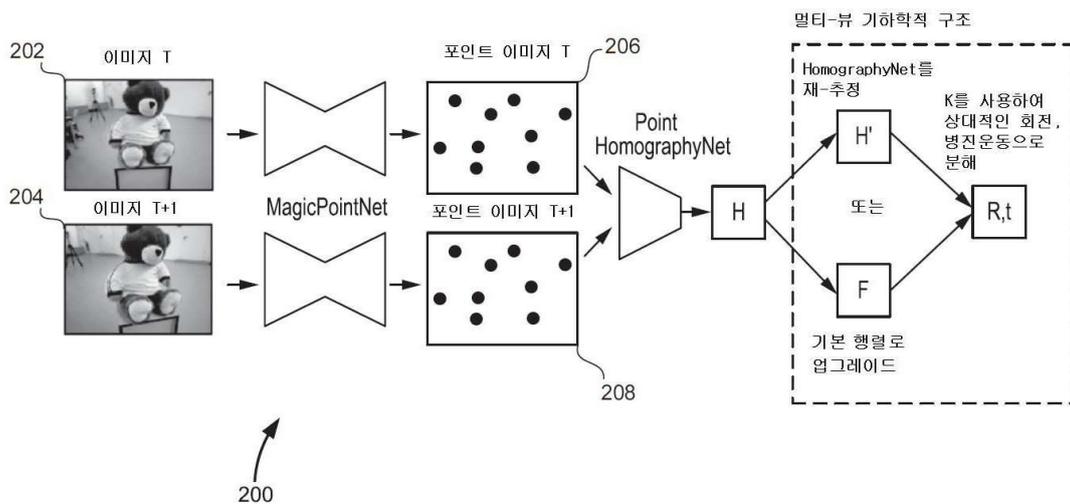
전체 청구항 수 : 총 20 항

(54) 발명의 명칭 **콘블루셔널 이미지 변환을 사용하여 동시 로컬화 및 맵핑을 수행하기 위한 방법 및 시스템**

(57) 요약

본 발명은 2개의 이미지들에 기반하여 호모그래피를 컴퓨팅하기 위한 증강 현실 디바이스들 및 방법들에 관한 것이다. 하나의 방법은 제1 카메라 포즈에 기반한 제1 이미지 및 제2 카메라 포즈에 기반한 제2 이미지를 수신하는 단계, 제1 이미지에 기반한 제1 포인트 클라우드(point cloud) 및 제2 이미지에 기반한 제2 포인트 클라우드(뒷면에 계속)

대표도



를 생성하는 단계, 제1 포인트 클라우드 및 제2 포인트 클라우드를 뉴럴 네트워크(neural network)에 제공하는 단계, 및 뉴럴 네트워크에 의해, 제1 포인트 클라우드 및 제2 포인트 클라우드에 기반하여 호모그래피를 생성하는 단계를 포함할 수 있다. 뉴럴 네트워크는, 복수의 포인트들을 생성하고, 3D 궤적을 결정하고, 포인트들을 보는 카메라 포즈들을 획득하기 위해 3D 궤적을 샘플링하고, 포인트들을 2D 평면들에 투사하고, 투사된 포인트들을 사용하여 생성된 호모그래피와 실측-자료 호모그래피를 비교하고, 그리고 비교에 기반하여 뉴럴 네트워크를 수정함으로써, 트레이닝될 수 있다.

(52) CPC특허분류

G06K 9/00671 (2013.01)
G06K 9/4628 (2013.01)
G06K 9/6211 (2013.01)
G06K 9/627 (2013.01)
G06T 7/337 (2017.01)
G06T 7/35 (2017.01)
G06T 7/579 (2017.01)
G06T 2207/10024 (2013.01)
G06T 2207/20081 (2013.01)

명세서

청구범위

청구항 1

2개의 이미지들에 기반하여 호모그래피(homography)를 컴퓨팅하기 위한 방법으로서,

제1 카메라 포즈에 기반한 제1 이미지 및 제2 카메라 포즈에 기반한 제2 이미지를 수신하는 단계;

상기 제1 이미지에 기반한 제1 포인트 클라우드(point cloud) 및 상기 제2 이미지에 기반한 제2 포인트 클라우드를 생성하는 단계;

상기 제1 포인트 클라우드 및 상기 제2 포인트 클라우드를 뉴럴 네트워크(neural network)에 제공하는 단계; 및

상기 뉴럴 네트워크에 의해, 상기 제1 포인트 클라우드 및 상기 제2 포인트 클라우드에 기반하여 상기 호모그래피를 생성하는 단계를 포함하는,

2개의 이미지들에 기반하여 호모그래피를 컴퓨팅하기 위한 방법.

청구항 2

제1 항에 있어서,

상기 제1 포인트 클라우드 및 상기 제2 포인트 클라우드는 2D(two-dimensional) 포인트 클라우드들인,

2개의 이미지들에 기반하여 호모그래피를 컴퓨팅하기 위한 방법.

청구항 3

제2 항에 있어서,

상기 제1 이미지는 제1 순간의 시간에 제1 카메라에 의해 캡처되었고, 상기 제2 이미지는 상기 제1 순간의 시간 후의 제2 순간의 시간에 상기 제1 카메라에 의해 캡처된,

2개의 이미지들에 기반하여 호모그래피를 컴퓨팅하기 위한 방법.

청구항 4

제2 항에 있어서,

상기 제1 포인트 클라우드 및 상기 제2 포인트 클라우드는 제1 뉴럴 네트워크를 사용하여 생성되고, 상기 뉴럴 네트워크는 제2 뉴럴 네트워크인,

2개의 이미지들에 기반하여 호모그래피를 컴퓨팅하기 위한 방법.

청구항 5

제2 항에 있어서,

상기 뉴럴 네트워크는:

복수의 포인트들을 포함하는 하나 이상의 3D(three-dimensional) 포인트 클라우드들의 각각의 3D 포인트 클라우드에 대해,

상기 복수의 포인트들의 임계 거리 내의 3D 궤적을 결정하고;

특정 제1 카메라 포즈 및 특정 제2 카메라 포즈를 획득하기 위해 상기 3D 궤적을 샘플링하고 - 상기 복수의 포인트들은 상기 특정 제1 카메라 포즈 및 상기 특정 제2 카메라 포즈로부터 적어도 부분적으로 볼 수 있음 - ;

제1 2D 포인트 클라우드를 생성하기 위해, 상기 특정 제1 카메라 포즈에 기반하여, 상기 복수의 포인트들을 제1 2D 평면에 투사하고;

제2 2D 포인트 클라우드를 생성하기 위해, 상기 특정 제2 카메라 포즈에 기반하여, 상기 복수의 포인트들을 제2

2D 평면에 투사하고;

상기 특정 제1 카메라 포즈 및 상기 특정 제2 카메라 포즈에 기반하여, 상기 제1 2D 포인트 클라우드와 상기 제2 2D 포인트 클라우드 사이의 실측-자료 호모그래피(ground-truth homography)를 결정하고;

상기 뉴럴 네트워크에 의해, 상기 제1 2D 포인트 클라우드 및 상기 제2 2D 포인트 클라우드에 기반하여 특정 호모그래피를 생성하고;

상기 특정 호모그래피와 상기 실측-자료 호모그래피를 비교하고; 그리고

상기 비교에 기반하여 상기 뉴럴 네트워크를 수정함으로써 이전에 트레이닝된,

2개의 이미지들에 기반하여 호모그래피를 컴퓨팅하기 위한 방법.

청구항 6

제5 항에 있어서,

상기 복수의 3D 포인트 클라우드들은 하나 이상의 기하학적 구조들(geometries)을 샘플링함으로써 생성되는,

2개의 이미지들에 기반하여 호모그래피를 컴퓨팅하기 위한 방법.

청구항 7

제5 항에 있어서,

상기 특정 제1 카메라 포즈 및 상기 특정 제2 카메라 포즈는 적어도 30% 중첩을 갖는,

2개의 이미지들에 기반하여 호모그래피를 컴퓨팅하기 위한 방법.

청구항 8

카메라;

상기 카메라에 통신 가능하게 커플링되고, 동작들을 수행하도록 구성된 프로세서를 포함하고,

상기 동작들은:

상기 카메라로부터, 제1 카메라 포즈에 기반한 제1 이미지 및 제2 카메라 포즈에 기반한 제2 이미지를 수신하는 동작;

상기 제1 이미지에 기반한 제1 포인트 클라우드 및 상기 제2 이미지에 기반한 제2 포인트 클라우드를 생성하는 동작;

상기 제1 포인트 클라우드 및 상기 제2 포인트 클라우드를 뉴럴 네트워크에 제공하는 동작; 및

상기 뉴럴 네트워크에 의해, 상기 제1 포인트 클라우드 및 상기 제2 포인트 클라우드에 기반하여 호모그래피를 생성하는 동작을 포함하는,

AR(augmented reality) 디바이스.

청구항 9

제8 항에 있어서,

상기 제1 포인트 클라우드 및 상기 제2 포인트 클라우드는 2D(two-dimensional) 포인트 클라우드들인,

AR 디바이스.

청구항 10

제9 항에 있어서,

상기 제1 포인트 클라우드 및 상기 제2 포인트 클라우드는 제1 뉴럴 네트워크를 사용하여 생성되고, 상기 뉴럴 네트워크는 제2 뉴럴 네트워크인,

AR 디바이스.

청구항 11

제9 항에 있어서,

상기 뉴럴 네트워크는:

복수의 포인트들을 포함하는 하나 이상의 3D(three-dimensional) 포인트 클라우드들의 각각의 3D 포인트 클라우드에 대해,

상기 복수의 포인트들의 임계 거리 내의 3D 궤적을 결정하고;

특정 제1 카메라 포즈 및 특정 제2 카메라 포즈를 획득하기 위해 상기 3D 궤적을 샘플링하고 - 상기 복수의 포인트들은 상기 특정 제1 카메라 포즈 및 상기 특정 제2 카메라 포즈로부터 적어도 부분적으로 볼 수 있음 - ;

제1 2D 포인트 클라우드를 생성하기 위해, 상기 특정 제1 카메라 포즈에 기반하여, 상기 복수의 포인트들을 제1 2D 평면에 투사하고;

제2 2D 포인트 클라우드를 생성하기 위해, 상기 특정 제2 카메라 포즈에 기반하여, 상기 복수의 포인트들을 제2 2D 평면에 투사하고;

상기 특정 제1 카메라 포즈 및 상기 특정 제2 카메라 포즈에 기반하여, 상기 제1 2D 포인트 클라우드와 상기 제2 2D 포인트 클라우드 사이의 실측-자료 호모그래피를 결정하고;

상기 뉴럴 네트워크에 의해, 상기 제1 2D 포인트 클라우드 및 상기 제2 2D 포인트 클라우드에 기반하여 특정 호모그래피를 생성하고;

상기 특정 호모그래피와 상기 실측-자료 호모그래피를 비교하고; 그리고

상기 비교에 기반하여 상기 뉴럴 네트워크를 수정함으로써 이전에 트레이닝된,

AR 디바이스.

청구항 12

제11 항에 있어서,

상기 복수의 3D 포인트 클라우드들은 하나 이상의 기하학적 구조들을 샘플링함으로써 생성되는,

AR 디바이스.

청구항 13

제11 항에 있어서,

상기 특정 제1 카메라 포즈 및 상기 특정 제2 카메라 포즈는 적어도 30% 중첩을 갖는,

AR 디바이스.

청구항 14

명령들을 포함하는 비-일시적인 컴퓨터-판독 가능 매체로서,

상기 명령들은, 프로세서에 의해 실행될 때, 상기 프로세서로 하여금, 동작들을 수행하게 하고,

상기 동작들은:

제1 카메라 포즈에 기반한 제1 이미지 및 제2 카메라 포즈에 기반한 제2 이미지를 수신하는 동작;

상기 제1 이미지에 기반한 제1 포인트 클라우드 및 상기 제2 이미지에 기반한 제2 포인트 클라우드를 생성하는 동작;

상기 제1 포인트 클라우드 및 상기 제2 포인트 클라우드를 뉴럴 네트워크에 제공하는 동작; 및

상기 뉴럴 네트워크에 의해, 상기 제1 포인트 클라우드 및 상기 제2 포인트 클라우드에 기반하여 호모그래피를

생성하는 동작을 포함하는,

비-일시적인 컴퓨터-판독 가능 매체.

청구항 15

제14 항에 있어서,

상기 제1 포인트 클라우드 및 상기 제2 포인트 클라우드는 2D(two-dimensional) 포인트 클라우드들인,

비-일시적인 컴퓨터-판독 가능 매체.

청구항 16

제15 항에 있어서,

상기 제1 이미지는 제1 순간의 시간에 제1 카메라에 의해 캡처되었고, 상기 제2 이미지는 상기 제1 순간의 시간 후의 제2 순간의 시간에 상기 제1 카메라에 의해 캡처된,

비-일시적인 컴퓨터-판독 가능 매체.

청구항 17

제15 항에 있어서,

상기 제1 포인트 클라우드 및 상기 제2 포인트 클라우드는 제1 뉴럴 네트워크를 사용하여 생성되고, 상기 뉴럴 네트워크는 제2 뉴럴 네트워크인,

비-일시적인 컴퓨터-판독 가능 매체.

청구항 18

제15 항에 있어서,

상기 뉴럴 네트워크는:

복수의 포인트들을 포함하는 하나 이상의 3D(three-dimensional) 포인트 클라우드들의 각각의 3D 포인트 클라우드에 대해,

상기 복수의 포인트들의 임계 거리 내의 3D 궤적을 결정하고;

특정 제1 카메라 포즈 및 특정 제2 카메라 포즈를 획득하기 위해 상기 3D 궤적을 샘플링하고 - 상기 복수의 포인트들은 상기 특정 제1 카메라 포즈 및 상기 특정 제2 카메라 포즈로부터 적어도 부분적으로 볼 수 있음 - ;

제1 2D 포인트 클라우드를 생성하기 위해, 상기 특정 제1 카메라 포즈에 기반하여, 상기 복수의 포인트들을 제1 2D 평면에 투사하고;

제2 2D 포인트 클라우드를 생성하기 위해, 상기 특정 제2 카메라 포즈에 기반하여, 상기 복수의 포인트들을 제2 2D 평면에 투사하고;

상기 특정 제1 카메라 포즈 및 상기 특정 제2 카메라 포즈에 기반하여, 상기 제1 2D 포인트 클라우드와 상기 제2 2D 포인트 클라우드 사이의 실측-자료 호모그래피를 결정하고;

상기 뉴럴 네트워크에 의해, 상기 제1 2D 포인트 클라우드 및 상기 제2 2D 포인트 클라우드에 기반하여 특정 호모그래피를 생성하고;

상기 특정 호모그래피와 상기 실측-자료 호모그래피를 비교하고; 그리고

상기 비교에 기반하여 상기 뉴럴 네트워크를 수정함으로써 이전에 트레이닝된,

비-일시적인 컴퓨터-판독 가능 매체.

청구항 19

제18 항에 있어서,

상기 복수의 3D 포인트 클라우드들은 하나 이상의 기하학적 구조들을 샘플링함으로써 생성되는, 비-일시적인 컴퓨터-판독 가능 매체.

청구항 20

제18 항에 있어서,
상기 특정 제1 카메라 포즈 및 상기 특정 제2 카메라 포즈는 적어도 30% 중첩을 갖는,
비-일시적인 컴퓨터-판독 가능 매체.

발명의 설명

기술 분야

[0001] [0001] 본 출원은 "METHOD AND SYSTEM FOR PERFORMING SIMULTANEOUS LOCALIZATION AND MAPPING USING CONVOLUTIONAL IMAGE TRANSFORMATION"이란 명칭으로 2017년 6월 28일에 출원된 미국 가특허 출원 제62/526,203 호의 정규 출원이며 상기 출원을 우선권으로 주장하고, 상기 출원의 전체 내용은 본원에 포함된다.

배경 기술

[0002] [0002] 현대 컴퓨팅 및 디스플레이 기술들은 소위 "가상 현실" 또는 "증강 현실" 경험들을 위한 시스템들의 개발을 용이하게 했으며, 여기서 디지털적으로 재생된 이미지들 또는 이미지들의 부분들은, 그들이 실제인 것으로 보이거나, 실제로서 지각될 수 있는 방식으로 사용자에게 제시된다. 가상 현실, 또는 "VR" 시나리오는 통상적으로 다른 실제 세계 시각적 입력에 대한 투명성(transparency) 없는 디지털 또는 가상 이미지 정보의 프리젠테이션(presentation)을 수반하고; 증강 현실, 또는 "AR" 시나리오는 통상적으로 사용자 주위 실제 세계의 시각화에 대한 증강으로서 디지털 또는 가상 이미지 정보의 프리젠테이션을 수반한다.

[0003] [0003] 이러한 디스플레이 기술들에서 이루어진 진보에도 불구하고, 증강 현실 시스템들, 특히, 디스플레이 시스템들에 관련된 향상된 방법들, 시스템들 및 디바이스들에 대한 필요성이 당분야에 존재한다.

발명의 내용

[0004] [0004] 본 개시내용은 일반적으로 SLAM(simultaneous localization and mapping)을 수행하기 위한 시스템들 및 방법들에 관한 것이다. 더 상세하게는, 본 개시내용의 실시예들은 머리-장착 VR(virtual reality), MR(mixed reality) 및/또는 AR(augmented reality) 디바이스들에서 콘볼루션형 이미지 변환을 사용하여 SLAM을 수행하기 위한 시스템들 및 방법들을 제공한다. 본 발명의 실시예들은 사용자가 착용한 디바이스에 의해 캡처된 이미지들을 분석함으로써 사용자/디바이스 움직임의 정확한 검출을 가능하게 하고, 이로써 디스플레이되는 가상 콘텐츠의 정확도를 개선한다. 본 발명이 AR 디바이스를 참조하여 설명될 수 있지만, 본 개시내용은 컴퓨터 비전 및 이미지 디스플레이 시스템들에서의 다양한 애플리케이션들에 적용 가능하다.

[0005] [0005] 본 발명의 제1 양상에서, 2개의 이미지들에 기반하여 호모그래피(homography)를 컴퓨팅하기 위한 방법이 제공된다. 방법은 제1 카메라 포즈에 기반한 제1 이미지 및 제2 카메라 포즈에 기반한 제2 이미지를 수신하는 단계를 포함할 수 있다. 방법은 또한 제1 이미지에 기반한 제1 포인트 클라우드(point cloud) 및 제2 이미지에 기반한 제2 포인트 클라우드를 생성하는 단계를 포함할 수 있다. 방법은 제1 포인트 클라우드 및 제2 포인트 클라우드를 뉴럴 네트워크(neural network)에 제공하는 단계를 더 포함할 수 있다. 방법은 뉴럴 네트워크에 의해, 제1 포인트 클라우드 및 제2 포인트 클라우드에 기반하여 호모그래피를 생성하는 단계를 더 포함할 수 있다. 일부 실시예들에서, 제1 포인트 클라우드 및 제2 포인트 클라우드는 2D(two-dimensional) 포인트 클라우드들이다. 일부 실시예들에서, 제1 이미지는 제1 순간의 시간에 제1 카메라에 의해 캡처되었다. 일부 실시예들에서, 제2 이미지는 제1 순간의 시간 후의 제2 순간의 시간에 제1 카메라에 의해 캡처되었다. 일부 실시예들에서, 제1 포인트 클라우드 및 제2 포인트 클라우드는 제1 뉴럴 네트워크를 사용하여 생성되고, 뉴럴 네트워크는 제2 뉴럴 네트워크이다.

[0006] [0006] 일부 실시예들에서, 뉴럴 네트워크는, 복수의 포인트들을 포함하는 하나 이상의 3D(three-dimensional) 포인트 클라우드들의 각각의 3D 포인트 클라우드에 대해, 복수의 포인트들의 임계 거리 내의 3D 레적을 결정하고, 특정 제1 카메라 포즈 및 특정 제2 카메라 포즈를 획득하기 위해 3D 레적을 샘플링하고 - 복수의 포인트들은 특정 제1 카메라 포즈 및 특정 제2 카메라 포즈로부터 적어도 부분적으로 볼 수 있음 - 제1 2D 포인트 클라

우드를 생성하기 위해, 특정 제1 카메라 포즈에 기반하여, 복수의 포인트들을 제1 2D 평면에 투사하고, 제2 2D 포인트 클라우드를 생성하기 위해, 특정 제2 카메라 포즈에 기반하여, 복수의 포인트들을 제2 2D 평면에 투사하고, 특정 제1 카메라 포즈 및 특정 제2 카메라 포즈에 기반하여, 제1 2D 포인트 클라우드와 제2 2D 포인트 클라우드 사이의 실측-자료 호모그래피(ground-truth homography)를 결정하고, 뉴럴 네트워크에 의해, 제1 2D 포인트 클라우드 및 제2 2D 포인트 클라우드에 기반하여 특정 호모그래피를 생성하고, 특정 호모그래피와 실측-자료 호모그래피를 비교하고, 그리고 비교에 기반하여 뉴럴 네트워크를 수정함으로써 이전에 트레이닝되었다. 일부 실시예들에서, 복수의 3D 포인트 클라우드들은 하나 이상의 기하학적 구조들(geometries)을 샘플링함으로써 생성된다. 일부 실시예들에서, 특정 제1 카메라 포즈 및 특정 제2 카메라 포즈는 적어도 30% 중첩을 갖는다.

[0007] 본 발명의 제2 양상에서, AR 디바이스가 제공된다. AR 디바이스는 카메라를 포함할 수 있다. AR 디바이스는 또한, 카메라에 통신 가능하게 커플링되고, 동작들을 수행하도록 구성된 프로세서를 포함할 수 있고, 동작들은: 카메라로부터, 제1 카메라 포즈에 기반한 제1 이미지 및 제2 카메라 포즈에 기반한 제2 이미지를 수신하는 동작, 제1 이미지에 기반한 제1 포인트 클라우드 및 제2 이미지에 기반한 제2 포인트 클라우드를 생성하는 동작, 제1 포인트 클라우드 및 제2 포인트 클라우드를 뉴럴 네트워크에 제공하는 동작, 및 뉴럴 네트워크에 의해, 제1 포인트 클라우드 및 제2 포인트 클라우드에 기반하여 호모그래피를 생성하는 동작을 포함한다. 일부 실시예들에서, 제1 포인트 클라우드 및 제2 포인트 클라우드는 2D(two-dimensional) 포인트 클라우드들이다. 일부 실시예들에서, 제1 포인트 클라우드 및 제2 포인트 클라우드는 제1 뉴럴 네트워크를 사용하여 생성되고, 뉴럴 네트워크는 제2 뉴럴 네트워크이다.

[0008] 본 발명의 제3 양상에서, 비-일시적인 컴퓨터-관독 가능 매체가 제공된다. 비-일시적인 컴퓨터-관독 가능 매체는 명령들을 포함할 수 있고, 명령들은, 프로세서에 의해 실행될 때, 프로세서로 하여금, 동작들을 수행하게 하고, 동작들은 제1 카메라 포즈에 기반한 제1 이미지 및 제2 카메라 포즈에 기반한 제2 이미지를 수신하는 동작, 제1 이미지에 기반한 제1 포인트 클라우드 및 제2 이미지에 기반한 제2 포인트 클라우드를 생성하는 동작, 제1 포인트 클라우드 및 제2 포인트 클라우드를 뉴럴 네트워크에 제공하는 동작, 및 뉴럴 네트워크에 의해, 제1 포인트 클라우드 및 제2 포인트 클라우드에 기반하여 호모그래피를 생성하는 동작을 포함한다. 일부 실시예들에서, 제1 포인트 클라우드 및 제2 포인트 클라우드는 2D 포인트 클라우드들이다. 일부 실시예들에서, 제1 이미지는 제1 순간의 시간에 제1 카메라에 의해 캡처되었고, 제2 이미지는 제1 순간의 시간 후의 제2 순간의 시간에 제1 카메라에 의해 캡처되었다. 일부 실시예들에서, 제1 포인트 클라우드 및 제2 포인트 클라우드는 제1 뉴럴 네트워크를 사용하여 생성되고, 뉴럴 네트워크는 제2 뉴럴 네트워크이다.

도면의 간단한 설명

[0009] 다양한 실시예들의 성질 및 장점들의 추가의 이해는 다음의 도면들을 참조함으로써 실현될 수 있다. 첨부된 도면들에서, 유사한 컴포넌트들 또는 특징들은 동일한 참조 라벨을 가질 수 있다. 추가로, 동일한 타입의 다양한 컴포넌트들은, 참조 라벨 다음에 대시 기호 및 유사한 컴포넌트들 사이를 구별하는 제2 라벨에 의해 구별될 수 있다. 제1 참조 라벨만이 본 명세서에서 사용되는 경우, 설명은, 제2 참조 라벨과는 관계없이 동일한 제1 참조 라벨을 갖는 유사한 컴포넌트들 중 임의의 하나에 적용 가능하다.

[0010] 도 1은 본 발명의 일부 실시예들에 따른, SLAM(simultaneous localization and mapping) 접근법의 3개의 서브태스크들의 시각적 개요를 예시한다.

[0011] 도 2는 본 발명의 일부 실시예들에 따른 DH-SLAM 추적 시스템을 예시한다.

[0012] 도 3은 본 발명의 일부 실시예들에 따른 MagicPointNet 아키텍처를 예시한다.

[0013] 도 4는 본 발명의 일부 실시예들에 따른 MagicPointNet 합성 트레이닝 데이터를 예시한다.

[0014] 도 5는 본 발명의 일부 실시예들에 따른 포인트 기반 HomographyNet 아키텍처를 예시한다.

[0015] 도 6은 본 발명의 일부 실시예들에 따른 포인트 기반 HomographyNet 데이터 생성을 예시한다.

[0016] 도 7은 본 발명의 일부 실시예들에 따른 스마트 앵커 SLAM 시스템을 예시한다.

[0017] 도 8은 본 발명의 일부 실시예들에 따른, 2개의 이미지들에 기반하여 호모그래피를 컴퓨팅하기 위한 방법을 예시한다.

[0018] 도 9는 본 발명의 일부 실시예들에 따른, 뉴럴 네트워크를 트레이닝시키기 위한 방법을 예시한다.

[0019] 도 10은 본 발명의 일부 실시예들에 따른 웨어러블 증강 현실(AR) 디바이스의 개략도를 예시한다.

발명을 실시하기 위한 구체적인 내용

- [0010] [0020] 이미지 카테고리화 및 객체 검출과 같은 컴퓨터 비전 태스크들에서의 많은 심층 학습 성공은 ImageNet 및 MS-COCO와 같은 대형 주석형 데이터베이스들(large annotated databases)의 가용성에서 비롯된다. 그러나, SLAM(simultaneous localization and mapping)형 포즈 추적 및 재구성 문제들에 대해, 대신에 Microsoft Kinect에 기반한 Freiburg-TUM RGBD 데이터세트, 스테레오 비전 카메라들 및 IMU에 기반한 EuRoC 드론/MAV 데이터세트, 및 4개의 카메라들, GPS 데이터 및 Velodyne 레이저 스캐너가 장착된 차량들의 KITTI 드라이브 데이터세트와 같은 더 작은 디바이스-특정 데이터세트들의 분해된 생태계가 존재한다.
- [0011] [0021] SLAM에 대한 ImageNet이 현재 존재하지 않는다. 실제로, 매우 많은 수의 환경들 및 시나리오들에서 정확한 실측-자료 포즈 측정들(ground-truth pose measurements)을 획득하는 것은 어렵다. 실측-자료 센서들과 Visual SLAM 시스템 사이의 정확한 정렬을 획득하는 것은 상당한 노력을 요구하며, 상이한 카메라들에서의 변동들에 걸쳐 스케일링하는 것이 용이하지 않다. SLAM 태스크들에 대한 모든 관련 기하학적 변수들이 100 % 정확도로 기록될 수 있기 때문에, 사실적인 렌더링(photorealistic rendering)이 유용할 수 있다. 사실적인 시퀀스들에서 SLAM을 벤치마킹하는 것은 장점들을 가질 수 있지만, 이러한 렌더링된 이미지들에 대한 트레이닝은 종종 도메인 적응 이슈들로 어려움을 겪고 있으며, 많은 심층 넷들(deep nets)은 과다 적합(overfit)한 것 같다.
- [0012] [0022] 실제로, SLAM 모델들이 간결한(lean) 경우, 로봇 공학 및 혼합 현실에서의 임베딩된 플랫폼들과 같은 임베딩된 플랫폼들에서 대규모로 실행되는 것이 바람직하다. 본 발명의 실시예들은, 이러한 시스템들이 임베딩된 플랫폼들에 배치될 수 있도록, 풀 프레임 예측과 대조적으로 기하학적 일관성에 초점을 둔다. 심층 학습 시스템에 의해 이루어진 전체 프레임 예측들이 특정 장점들을 갖지만, 많은 예시들에서, 포인트들을 예측/정렬하는 것은 매트릭-레벨 포즈 복구에 충분하다. 심층 넷들의 배치는 일반적으로 넷들을 가능한 작게 만들기 위한 커스텀 오프라인 최적화 절차를 수반하여, 본 발명의 일부 예시들에서, 전체 시각 프레임 예측 문제들이 함께 생략된다.
- [0013] [0023] 본 발명의 실시예들은 미지의 환경에서 6 자유도 단안 카메라 로컬화를 위한, DH-SLAM으로 지칭되는 시스템을 포함할 수 있으며, 해당 시스템은, 높은 센서 잡음, 낮은 조명 및 주요 폐색들의 존재 시에, 호모그래피들을 강건하게 추정하도록 트레이닝된 콘볼루션 뉴럴 네트워크들에 의해 주로 구동된다. 시스템은 3개의 심층 ConvNet들을 포함할 수 있다. 제1 네트워크인 MagicPointNet은 단일 이미지들에 대해 동작하고 이미지에서 핵심적인 포인트들(salient points) - 이 포인트들은, 설계에 의해, 격리되고 이미지에서 잘 분산됨 - 을 추출할 수 있고, 비-최대 억제(non-maximum suppression)와 같은 어떠한 추가적인 사후-프로세싱도 활용하지 않을 수 있다. 제2 네트워크인 PointHomographyNet은 MagicPointNet으로부터의 포인트 응답 맵들의 쌍들에 대해 동작할 수 있고, 2개의 포인트 이미지들에 관련된 호모그래피를 추정할 수 있다. 이어서, 장면이 매우 비-평면형인 경우, 호모그래피 추정은 표준 MVG 기술들을 사용하여 재-추정되거나 기본 행렬로 업그레이드될 수 있다. 제3 네트워크인 RelocalizationNet은 단일 이미지들을, 저차원 임베딩 벡터로 변환할 수 있고, 저차원 임베딩 벡터는 신뢰할 수 있고 빠른 재로컬화에 사용되는 빠른 이미지 비교를 허용한다. 네트워크들은 간단한 합성 데이터로 트레이닝될 수 있어서, 네트워크들을 트레이닝시키는 것을 용이하게 하는데, 왜냐하면 네트워크들이 값비싼 외부 카메라 실측 자료 장비 또는 진보된 그래픽 렌더링 파이프라인들에 의존하지 않기 때문이다. 시스템은 빠르고 간결하며, CPU에서 30+ FPS를 실행할 수 있다.
- [0014] [0024] 본 발명의 다양한 실시예들에서, 3개의 커스텀 SLAM 콘볼루션 뉴럴 네트워크들은 3개의 별개의 트레이닝 도메인들과 함께 제공된다. 그러나, 이러한 특정 구현은 요구되지 않고, 뉴럴 네트워크들 중 하나 이상은 일부 실시예들에서 결합될 수 있다. 호모그래피들이 각각의 시스템의 중요한 컴포넌트일 수 있기 때문에, 호모그래피 방식으로(homographically) 고무된(inspired) 단안 SLAM 시스템이 제공된다. 본 발명의 실시예들은 어떠한 종류의 로컬 특징 포인트 디스크립터들(local feature point descriptors)도 사용하지 않을 수 있고, 수동 데이터 주석 또는 값비싼 오프-라인 그래픽 렌더링 파이프라인들을 요구하지 않을 수 있다.
- [0015] [0025] 도 1은 본 발명의 일부 실시예들에 따른 SLAM 접근법의 3개의 서브태스크들의 시각적 개요를 예시한다. 서브태스크(102)에서, 센서 잡음, 조명, 텍스처들 및 모션 블러(motion blur)와 같은 이미징 효과들을 제거하기 위해, 기하학적 포인트 프리미티브들(geometric point primitives)이 이미지들로부터 추출된다. 서브태스크(104)에서, 이미지들에 관련된 글로벌 포즈 정보를 추출하기 위해 포인트 이미지들의 쌍들이 비교된다. 서브태스크(106)에서, 이미지들은 빠른 이미지 매칭을 위해 저차원 벡터로 압축된다.

- [0016] [0026] 서브태스크(102)는 이미지를 포인트들과 같은 기하학적 엔티티들의 세트에 맵핑하는 것에 관련된다. 서브태스크(102)의 하나의 목표는 조명, 왜곡들 및 전체 글로벌 조명 변동들과 같은 성가신 변수들(nuisance variables)의 손상을 최소화하는 것이다. 코너-검출과 유사한 기술을 사용하여 입력 이미지 내의 신뢰할 수 있는 위치들을 추출하는 것이 또한 바람직할 수 있다. 그러나, 이미지 코너를 분석적으로 정의하는 대신에, 코너 검출기를 트레이닝시키기 위해 커스텀 2D 형상 렌더러(custom 2D shape renderer)가 사용된다. 결과적인 2D 이미지 위치들은 시스템에서 반-조밀하게 출력되는데, 왜냐하면 장면 포인트들의 반-조밀 커버리지가 강건한 변환 추정에서 중요할 수 있기 때문이다.
- [0017] [0027] 서브태스크(104)는 한 쌍의 이미지들 사이의 상대적인 포즈를 찾는 것에 관련된다. 일부 예시들에서, 호모그래피들은 워핑된 자연 이미지들의 큰 데이터셋으로부터 샘플링된 그레이스케일 이미지들의 쌍들에 대해 트레이닝된 CNN으로부터 컴퓨팅될 수 있다. 본 발명의 일부 실시예들에서, CNN은 포인트-이미지들, 즉, MagicPointNet에 의해 출력되는 종류들의 이미지들의 공간을 사용하여 트레이닝된다. 반-조밀 포인트 이미지들의 공간이 전체 RGB 이미지들의 공간보다 상당히 더 작을 수 있기 때문에, 전체 RGB 이미지들을 처리하는 데 필요한 것보다 더 적은 수의 뉴런들을 사용하여 상당히 더 많은 성능이 획득될 수 있다. 본원에 제공된 포즈 추정 네트워크는 포인트-응답 이미지들에 대해 동작할 수 있고, 호모그래피 및 본질적인 모드들 둘 모두에서 실행될 수 있다. 다중-뷰 기하학적 구조의 맥락에서, 장면 평면성(scene planarity)의 가정이 유지될 수 있거나 유지될 수 없기 때문에, 3D 세계를 2개의 상이한 방식들로 처리하는 능력이 중요할 수 있다. 평면성 가정이 유지되면, 호모그래피들이 추정될 수 있고, 나중에 개별 포인트 깊이들에 대한 우려 없이 (R, t) 추정들로 업그레йд될 수 있다. 대신에 솔버(solver)가 본질적인 행렬 추정에 기반하면, 장면 평면성은 퇴보한 E 추정을 제공하고, 모든 다른 추정들이 실패할 수 있다. 따라서, 볼 수 있는 장면 기하학적 구조가 매우 비-평면형일 때, Es, Rs, ts 및 포인트 깊이들이 직접적으로 처리될 수 있다.
- [0018] [0028] 서브태스크(106)는 이미지들의 시각적 임베딩들(visual embeddings)을 생성하는 것에 관련된다. 다른 태스크들에 대해, 포즈-민감도가 중요할 수 있지만, 임베딩들에 대해 포즈-민감도가 요구될 수 있다. 실세계 이미지들을 사용하는 것이 이 태스크에 중요할 수 있는데, 왜냐하면 2D 형상들 또는 포인트 응답 이미지들과 같은 엔지니어링된 표현들의 최상부 상의 임베딩들을 학습하는 것이 도움이 될 가능성이 낮을 수 있기 때문이다. 대신에, 이미지들의 무료로 이용 가능한 대형 데이터셋(즉, ImageNet)가 사용되고, 호모그래피들은 이미지들을 워핑(warp)하는 데 사용된다. 학습 동안, 2개의 이미지들이 호모그래피 방식으로 관련된 경우, 이러한 이미지들은 유사한 임베딩 벡터들을 갖도록 강제될 수 있다.
- [0019] [0029] 도 2는 본 발명의 일부 실시예들에 따른 DH-SLAM 추적 시스템(200)의 일반적인 아키텍처를 예시한다. 일부 예시들에서, 이미지들(예컨대, 이미지(202) 및 이미지(204))의 쌍들은 MagicPointNet에 의해 프로세싱되는데, 이는 이미지들에서 핵심적인 포인트들을 검출하고 한 쌍의 포인트 이미지들(예컨대, 포인트 이미지(206) 및 포인트 이미지(208))을 생성하도록 트레이닝된다. 이어서, 포인트 이미지들은 시간(T) 내지 시간(T+1)에 포인트 이미지들에 관련된 호모그래피(H)를 컴퓨팅하기 위해 PointHomographyNet에 의해 함께 프로세싱된다. 네트워크로부터의 H 추정은 종종 포인트들을 서로의 몇몇의 픽셀들 내로 투사할 수 있고, 따라서 간단한 가장 가까운 이웃 대응성은 대부분 평면형 장면들에 대해 더 정확한 H'를 재-추정하기에 충분하거나, H는 다수의 뷰 기하학적 구조 기술들을 사용하여 비-평면형 장면들을 더 양호하게 설명하는 기본 행렬 F로 업그레йд될 수 있다. 이어서, H' 및 F 둘 모두는 카메라 행렬 K를 사용하여 카메라의 상대적 포즈로 분해될 수 있다.
- [0020] [0030] 일부 예시들에서, SLAM 시스템은 3D 서브시스템들: 이미지들을 2D 포인트 위치들의 표준 서브공간(canonical subspace)으로 운반하는 Magic Point 검출기, 글로벌 변환 추정 포인트-기반 HomographyNet, 및 RelocNet이라 불리는 시각 이미지 임베딩 엔진으로 분해된다. 일부 실시예들에서, 각각의 서브시스템은 별개의 콘볼루션 뉴럴 네트워크이다. MagicPointNet 및 RelocNet 둘 모두가 단일 이미지들에 대해 동작할 수 있지만, PointHomographyNet은 한 쌍의 이미지들에 대해 동작한다. 일부 실시예들에서, 네트워크들은 다음의 문제들: 잡음이 있는 이미지들에서 강건한 2D 위치들을 검출하는 것, 2개의 이미지들 사이의 상대적인 포즈를 컴퓨팅하는 것, 및 재로컬화를 해결할 수 있다. 엔지니어링된 특징 디스크립터들(ORB 또는 SIFT) 둘 모두에 크게 의존하는 종래의 접근법들과 달리, 본 발명의 실시예들은 디스크립터들과 이미지들의 개별 포인트들을 연관시키지 않을 수 있다. 이러한 종래의 특징-기반 SLAM 시스템들과 달리, 상대적인 포즈 추정은 디스크립터가 없는 방식으로 수행될 수 있다. 글로벌 이미지별 디스크립터들과 유사할 수 있는 임베딩들이 또한 사용될 수 있다. 임베딩들은 의사 호모그래피 불변(pseudo homographic invariant)이 되도록 엔지니어링될 수 있다. 설계에 의해, 호모그래피에 의해 관련된 2개의 이미지들이 정해진 매니폴드 상에서 근접할 수 있다.
- [0021] [0031] 일부 예시들에서, 제1 단계는 이미지에서 핵심적이고 로컬화 가능한 2D 위치들을 검출하는 것을 포함할

수 있다. 이 단계는, Harris 또는 FAST와 같은 코너형 응답 맵들을 컴퓨팅하고, 로컬 최대치들을 검출하고, 비-최대 역제를 사용함으로써 수행될 수 있다. 이미지 전반에 걸쳐 이러한 로컬 최대치들을 분배하기 위해 추가적인 단계들이 수행될 수 있다. 이 프로세스는 많은 양의 도메인 전문 지식 및 핸드 엔지니어링(hand engineering)을 수반할 수 있고, 이는 일반화 및 강건성을 제한할 수 있다. SLAM 설정에서 유용성을 증가시키기 위해, 포인트 검출기에 의해 검출된 포인트들은 이미지 전반에 걸쳐 잘 분산되고, 프레임들에 걸친 대응이 용이하도록 서로 격리될 수 있다. 시스템이 높은 센서 잡음 시나리오들 및 낮은 조명에서 포인트들을 검출하는 것이 또한 바람직할 수 있다. 일부 예시들에서, 검출된 각각의 포인트에 대해 신뢰도 스코어가 획득될 수 있으며, 이는 스퍼리어스 포인트들을 거부하는 것을 돕는 데 사용될 수 있다. 이 이미지 포인트들이 이미지의 로컬, 고-그래디언트 에지들에 대응할 필요는 없지만, 대신에, 전통적인 코너 검출기들보다 더 큰 수용장(receptive field)을 활용할 수 있는 블롭들(blobs)의 중심점들과 같은 다른 저레벨 큐들에 대응할 수 있다는 것이 주목되어야 한다.

[0022] [0032] 도 3은 본 발명의 일부 실시예들에 따른 MagicPointNet 아키텍처(300)를 예시한다. 일부 실시예들에서, MagicPointNet은 그레이스케일 이미지들에 대해 동작하고, 입력의 각각의 픽셀에 대한 "포인트-니스(point-ness)" 확률을 출력한다. 이는 명시적인 디코더와 결합된 VGG-스타일 인코더로 이루어질 수 있다. 최종 15x20x65 텐서(tensor)의 각각의 공간 위치는 로컬 8x8 구역 + 단일 디스트리뷰 채널에 걸친 확률 분포를 나타낼 수 있으며, 이는 어떠한 포인트도 검출되지 않는다는 것을 나타낼 수 있다(8 * 8 + 1 = 65). 네트워크는 2D 형상 렌더러의 포인트 감시(point supervision)를 사용하여 표준 크로스 엔트로피 손실을 사용하여 트레이닝된다.

[0023] [0033] 일부 예시들에서, MagicPointNet은 커스텀 ConvNet 아키텍처 및 트레이닝 데이터 파이프라인을 사용하여 구현된다. 일부 예시들에서, 이미지(I)를 등가의 해상도를 갖는 포인트 응답 이미지(P)에 맵핑하는 것이 중요할 수 있으며, 여기서 출력의 각각의 픽셀은 입력의 해당 픽셀에 대한 "포인트-니스"의 확률에 대응한다. 조밀한 예측을 위한 네트워크 설계는 인코더-디코더 쌍을 수반할 수 있고, 여기서 공간 해상도는 풀링(pooling) 또는 스트라이드 컨볼루션(strided convolution)을 통해 감소되고, 이어서 업콘볼루션 연산들을 통해 전체 해상도로 다시 업샘플링된다. 업샘플링 계층들이 더 높은 컴퓨터이셔널 부담을 추가할 수 있기 때문에, MagicPointNet은 모델의 컴퓨터이셔널을 감소시키기 위해 명시적인 디코더로 구현될 수 있다. 컨볼루션 뉴럴 네트워크는, 각각의 공간 포지션에 대해 65개의 채널들이 있는 경우에, 이미지의 치수(dimensionality)를 120x160에서 15x20 셀 그리드로 감소시키기 위해 VGG 스타일 인코더를 사용할 수 있다. 일부 실시예들에서, QQVGA 해상도는 컴퓨터이셔널 부담을 감소시키기 위해 120x160일 수 있다. 65개의 채널들은 픽셀들의 로컬, 비-중첩 8x8 그리드 구역들 + 여분의 디스트리뷰 채널에 대응할 수 있으며, 이는 해당 8x8 구역에서 어떠한 포인트도 검출되지 않는 것에 대응한다. 네트워크는 3x3 컨볼루션들 다음에 BatchNorm 정규화 및 ReLU 비-선형성을 사용하여 완전한 컨볼루션될 수 있다. 최종 convlayer는 1x1 컨볼루션일 수 있다.

[0024] [0034] 도 4는 본 발명의 일부 실시예들에 따른 MagicPointNet 합성 트레이닝 데이터를 예시한다. 일부 실시예들에서, 이미지의 관심 포인트들은 다양한 뷰포인트들, 조명들 및 이미지 잡음 변형들에 걸쳐 안정적인, 이미지 내의 고유하게 식별 가능한 위치들일 수 있다. 스파스(sparse) SLAM 시스템의 사전프로세싱 단계로서 사용될 때, 정해진 SLAM 시스템에 대해 잘 작동하는 포인트들이 검출될 수 있다. 포인트 검출 알고리즘들의 하이퍼 파라미터들을 설계 및 선택하는 것은 전문적인 그리고 도메인 특정 지식을 활용할 수 있다.

[0025] [0035] 오늘날 존재하는 관심 포인트 라벨링된 이미지들의 대형 데이터베이스들이 존재하지 않는다. 값비싼 데이터 수집 노력을 피하기 위해, 렌더러는 컴퓨터 비전 라이브러리를 사용하여 구현할 수 있다. 일부 예시들에서 삼각형들, 사변형들, 체커보드들, 3D 큐브들 및 타원들과 같은 간단한 기하학적 형상들이 렌더링되고, 해당 형상들 각각에 대해 코너들의 실측 자료 위치들이 알려진다. 각각의 객체의 2D면들의 중심점들(centroids)이 또한 알려져 있으며, 이는 추가적인 관심 포인트로 가능하다. 2D 면들의 중심점들을 찾는 간단하고 강건한 검출기들을 설계하는 것은 어려운 것이라는 것이 주목되어야 한다. 형상들이 렌더링된 후, 호모그래픽 워핑들은 트레이닝 예들의 수를 증가시키기 위해 각각의 이미지에 적용될 수 있다. 선택적으로, 잡음, 예컨대, 조명 변화들의 형태의 다량의 잡음, 텍스처 잡음, 가우시안 잡음, 소금 및 후추 잡음(salt and pepper noise), 또는 이들의 조합들 등은 이미지들 중 하나 이상(예컨대, 각각의 이미지)에 적용될 수 있다. 데이터는 즉석에서(on the fly) 생성될 수 있고, 이는, 네트워크가 어떠한 예도 2번 보지 않는다는 것을 의미한다. 15x20 그리드의 각각의 셀에 대한 로짓들(logits)이 softmax 함수를 통해 파이핑된 후, 네트워크는 표준 크로스 엔트로피 손실을 사용하여 트레이닝될 수 있다.

[0026] [0036] 일부 예시들에서 PointHomographyNet은, Magic Point에 의해 생성된 바와 같이, 한 쌍의 포인트 이미지들이 제공되는 호모그래피를 생성한다. 모든 이미지들의 공간 및 상대적인 포즈들의 공간(모든 이미지들의 공

간 x 상대적인 포즈들의 공간) 대신에, 포인트 이미지들의 공간 및 상대적인 포즈들의 공간(포인트 이미지들의 공간 x 상대적인 포즈들의 공간)에 대해 동작하도록 네트워크를 설계함으로써, 조명, 왜곡 및 텍스처들과 같은 몇몇의 고려사항들이 감소된 중요성을 갖거나 무시될 수 있다. 또한, 광도계 일관성 가정(photometric consistency assumption)은 유지에 의존할 필요가 없다.

[0027] [0037] 도 5는 본 발명의 일부 실시예들에 따른 포인트 기반 HomographyNet 아키텍처(500)를 예시한다. 일부 예시들에서, 이진 포인트 이미지들의 쌍들은 연결되고(concatenated), 이어서 VGG-스타일 인코더를 통해 공급된다. 3×3 호모그래피(H)는 완전히 연결된 층에 의해 출력된다. 이어서, 호모그래피(H)는 자신의 최하부 우측 엘리먼트가 1이 되도록 정규화될 수 있다. 손실은, 제2 이미지에서 대응성을 갖는 것으로 알려진 하나의 이미지의 포인트들을 제2 이미지들로 변환하고 이들의 포지션들을 비교함으로써 컴퓨팅된다. PointHomographyNet을 트레이닝시키기 위한 손실 함수가 수학식 1에 도시된다.

수학식 1

$$L_H = \sum_{n=1}^N \|Hx_n - x'_n\|_2$$

[0028]

[0029] [0038] 일부 예시들에서, PointHomographyNet은 (임의의 전통적인 포인트 검출기에 대해 동작할 수는 있지만) MagicPointNet에 의해 출력된 포인트 검출들에 대해 직접 동작하도록 설계된다. 모델은 반-조밀 $15 \times 20 \times 65$ 이미지들의 쌍들에 대해 잘 작동할 수 있다. 이 작은 공간 해상도에서, 네트워크는 매우 적은 컴퓨테이션들을 사용한다. 입력의 채널별 연결 후에, 3×3 컨볼루션들, 최대-풀링, BatchNorm 및 ReLU 활성화들, 다음에 3×3 호모그래피(H)의 9개의 값들을 출력하는 2개의 완전히 연결된 계층들로 구성된 VGG 스타일 인코더가 구현될 수 있다.

[0030] [0039] 도 6은 본 발명의 일부 실시예들에 따른 포인트 기반 HomographyNet 데이터 생성을 예시한다. 일부 실시예들에서, PointHomographyNet을 트레이닝시키기 위해, 2개의 가상 카메라들로 렌더링된 포인트 클라우드들의 수백만의 예들이 생성될 수 있다. 포인트 클라우드들은 평면들, 구체들 및 큐브들을 포함한 간단한 3D 기하학적 구조들로부터 생성될 수 있다. 2개의 가상 카메라들의 포지션들은, 도 6에 도시된 바와 같이, 피스별 선형 병진운동 및 랜덤 축들을 중심으로 한 회전들로 구성된 랜덤 궤적들로부터 샘플링된다. 일부 실시예들에서, 적어도 30 %의 시각적 중첩을 갖는 카메라 쌍들이 랜덤하게 샘플링된다. 일단 포인트들이 2개의 카메라 프레임들에 투사되면, 스퍼리어스 및 누락 포인트 검출들에 대한 네트워크의 강건성을 개선하기 위해 포인트 입력 드롭아웃이 적용된다. 일부 예시들에서, 독립적으로 매치들 중 50 %를 랜덤하게 드롭하고, 포인트들 중 25 %를 랜덤하게 드롭함으로써 성능이 개선된다.

[0031] [0040] 3×3 행렬을 직접 출력하도록 네트워크를 트레이닝시키기 위해 전형적으로 몇몇의 고려사항들이 고려된다. 일부 예시들에서, 최종 FC 계층 바이어스(bias)가 항등 행렬(identity matrix)을 출력하도록 초기화될 때, 호모그래피(H)의 좌표들이 범위 $[-1, 1]$ 로 정규화될 때, 그리고 최하부 우측 엘리먼트가 1이 되도록 H 수량이 정규화될 때, 트레이닝이 가장 잘 작동하는데, 왜냐하면 호모그래피(H)가 8개의 자유도들 및 9개의 엘리먼트들을 갖기 때문이다.

[0032] [0041] 일부 실시예들에서, 임베딩 네트워크의 하나의 목표는 글로벌 128 차원 디스크립터와 입력 이미지를 연관시키는 것일 수 있다. 일부 실시예들에서, 임베딩이 호모그래픽 불변인 것이 바람직하다. 예컨대, 호모그래피에 의해 관련된 2개의 이미지들은 동일한 임베딩 벡터를 가져야 하고, 동일한 장면 콘텐츠를 묘사하지 않는(따라서 동일한 평면이 아닌) 2개의 이미지들은 상이한 임베딩 벡터들을 가져야 한다.

[0033] [0042] 임베딩 네트워크는 128 L2-정규화된 디스크립터를 생성할 수 있다. 이것은 VGG형 인코더 네트워크의 최상부 상에 완전히 연결된 계층에 의해 수행될 수 있다. 일부 예시들에서, 임베딩 네트워크는 ImageNet 데이터 세트로부터의 호모그래픽 방식으로 관련된 이미지들의 쌍들을 사용하여 트레이닝될 수 있다. 동일한 이미지로부터의 이미지 패치들(A, A', B)의 트리플렛들 모두가 샘플링될 수 있고, 여기서 (A, A')는 적어도 30 %의 중첩을 가지며, 호모그래피와 관련되고, 여기서 (A, B)는 어떠한 중첩도 갖지 않는다. 네트워크를 트레이닝시키는데 사용되는 (A, A') 포지티브 쌍 및 (A, B) 네거티브 쌍이 생성될 수 있다. 일부 실시예들에서, 당업자에게 명백할 수 있듯이, 샴 네트워크(siamese network)가 사용될 수 있는 반면에, 다른 실시예들에서, 2개의 타워 점

근법 또는 트리플렛 네트워크가 사용될 수 있다.

- [0034] [0043] MagicPointNet은 FAST 코너 검출기 및 Harris 코너 검출기와 같은 전통적인 코너 검출 베이스라인들에 대해 평가되었다. PointHomographyNet은 실제 센서들로부터의 합성 데이터 및 이미지 시퀀스를 둘 모두에서 고전적인 ORB+RANSAC 기반 호모그래피 추정 엔진과 대조적으로 평가되었다. 합성 도트 세계(synthetic dot world)에서 평가하는 것의 하나의 장점은, 포인트 세트들 사이의 실측 자료 대응성들이 알려져 있다는 것이다. 2개의 알고리즘들이 어떻게 저하되는지를 결정하기 위해 다양한 양들의 잡음이 추가될 수 있다. 임베딩 네트워크를 평가하기 위해, 임베딩 메트릭을 사용하여 가장 가까운 이웃들의 그리드가 컴퓨팅된다. 이것은, 베이스 ResNet 활성화를 사용하고 활성화 공간에서 가장 가까운 이웃들을 획득하는 것과 비교되었다.
- [0035] [0044] 전체 호모그래피 SLAM 시스템을 평가하기 위해, 매치들에 의해 생성된 최종 (R,t) 추정들에 대해 평가가 수행되었다. 추적(최종 프레임에 대한 포즈만을 추정) 및 임베딩-향상된 추적(가장 가까운 임베딩을 갖는 K개의 이미지들에 대한 포즈를 추정) 둘 모두에 대한 수치들이 관찰되었다. 본 발명의 실시예들은 PTAM, ORB-SLAM 및 LSD-SLAM과 같은 몇몇의 개방 소스 SLAM 시스템들에 대해 정성적으로 비교되었다. 위에 설명된 평가들의 결과들은, 본 발명의 다양한 실시예들이 종래의 접근법들과 비교될 때 우수한 성능을 나타낸다는 것을 나타냈다.
- [0036] [0045] 본 발명의 실시예들은 평면들 주위에 설계된 작은 증강 현실 시스템을 포함할 수 있으며, 이 시스템은 PointHomographyNet 및 RelocNet과 함께, 적은 수의 이미지들과 이들과 연관된 포인트 이미지들을 사용한다. 일부 예시들에서, 데이터세트 수집은, 카메라 중심을 향해 직접 가리키는 법선 [0,0,1]을 가진 기준 평면의 헤드-온 구성으로 시작하는 짧은 시퀀스를 생성하는 것을 수반한다. 이러한 기준 평면의 헤드-온 초기 뷰를 제공하는 것은 호모그래피 분해로부터의 2개의 솔루션들의 명확화(disambiguation)뿐만 아니라 각각의 2D 랜드마크에 대한 초기 스케일을 허용할 수 있다.
- [0037] [0046] 도 7은 본 발명의 일부 실시예들에 따른 스마트 앵커 SLAM 시스템(700)을 예시한다. 일부 실시예들에서, 사용자는 먼저 실내 환경에서 대부분 평면형 표면들의 사진을 촬영함으로써 "스마트 앵커들"의 세트를 구축한다. 이것은 포인트 이미지들과 임베딩들(E_0, E_1, E_2, \dots)의 쌍들의 사용자의 키프레임 인덱스들과 과플레이팅(populate)한다. 일단 인덱스가 구축되면, 시스템은 추적 모드에서 실행될 수 있다. MagicPointNet 및 RelocNet은, 포인트 이미지 (P_T) 및 임베딩 벡터(E_T)를 생성하기 위해 시간(T)에서 입력 이미지를 프로세싱할 수 있다. 일부 실시예들에서, 이어서, E_T 의 내적(dot product)은 임베딩 매니폴드에서 가장 가까운 이웃을 찾기 위해 다른(예컨대, 매 2번째) 임베딩들(E_0, E_1, E_2, \dots) 중 하나 이상으로 컴퓨팅된다. 도 7에 도시된 특정 실시예에서, E_2 가 선택된다. 이어서, E_2 에 대응하는 포인트 이미지(P_2)는, P_T 의 포인트를 P_2 로 변환하는 호모그래피(H)를 컴퓨팅하기 위해 P_T 와 함께 PointHomographyNet에 공급될 수 있다. 호모그래피(H)는 마지막으로 회전(R), 병진운동(t) 및 우세 평면(dominant plane)(n)으로 분해될 수 있다. 마지막으로, (P_2, E_2) 앵커에 대응하는 콘텐츠가 워핑되고, 입력 이미지에서 AR 오버레이로서 디스플레이될 수 있다.
- [0038] [0047] 도 8은 2개의 이미지들에 기반하여 호모그래피를 컴퓨팅하기 위한 방법(800)을 예시한다. 방법(800)의 단계들은 도시된 단계와 상이한 순서로 수행될 수 있고, 방법(800)의 수행 동안에 방법(800)의 하나 이상의 단계들이 생략될 수 있다. 방법(800)의 하나 이상의 단계들은 비-일시적 컴퓨터-관독 가능 매체에 포함된 명령들을 실행하도록 구성된 프로세서에 의해 수행 및/또는 개시될 수 있다.
- [0039] [0048] 단계(802)에서, 제1 이미지 및 제2 이미지가 수신된다. 제1 이미지는 제1 카메라 포즈에 기반할 수 있고, 제2 이미지는 제2 카메라 포즈에 기반할 수 있으며, 제2 카메라 포즈는 제1 카메라 포즈와 상이하다. 일부 실시예들에서, 제1 이미지 및 제2 이미지는 동일한 카메라에 의해 (제1 순간의 시간 및 제2 순간의 시간에 각각 - 제2 순간의 시간은 제1 순간의 시간 이후에 발생함 -) 캡처될 수 있고, 다른 실시예들에서, 제1 이미지는 제1 카메라에 의해 캡처될 수 있고, 제2 이미지는 동시에 또는 2개의 순간들의 시간에 제2 카메라에 의해 캡처될 수 있다.
- [0040] [0049] 단계(804)에서, 제1 포인트 클라우드가 제1 이미지에 기반하여 생성되고, 제2 포인트 클라우드가 제2 이미지에 기반하여 생성된다. 일부 실시예들에서, 제1 뉴럴 네트워크는 포인트 클라우드들을 생성하는 데 사용되고 즉, 제1 이미지는 제1 뉴럴 네트워크에 입력으로서 제공되고, 제1 포인트 클라우드는 제1 이미지에 기반하여 제1 뉴럴 네트워크에 의해 생성되고, 제2 이미지는 제1 뉴럴 네트워크에 입력으로서 제공되고, 제2 포인트 클라우드는 제2 이미지에 기반하여 제1 뉴럴 네트워크에 의해 생성된다. 제1 뉴럴 네트워크는, 이미지들에 기반하여 2D 포인트 클라우드들을 생성할 수 있는, 본원에 설명된 MagicPointNet일 수 있다.

- [0041] [0050] 단계(806)에서, 제1 포인트 클라우드 및 제2 포인트 클라우드는 입력으로서 제2 뉴럴 네트워크에 제공된다. 제2 뉴럴 네트워크는, 포인트 클라우드들에 기반하여 호모그래피를 생성할 수 있는, 본원에 설명된 PointHomographyNet일 수 있다. 일부 실시예들에서, 제1 뉴럴 네트워크는 제2 뉴럴 네트워크와 결합될 수 있어서, 2개의 네트워크들의 기능이 단일 시스템에서 결합될 수 있다.
- [0042] [0051] 단계(808)에서, 호모그래피는 제2 뉴럴 네트워크를 사용하여 제1 포인트 클라우드 및 제2 포인트 클라우드에 기반하여 생성된다. 일부 실시예들에서, 생성된 호모그래피는 제1 카메라 포즈와 제2 카메라 포즈 사이의 상대적인 회전 및 상대적인 병진운동(즉, 상대적인 포즈)이 추출될 수 있는 행렬(예컨대, 3x3)을 포함한다.
- [0043] [0052] 도 9는 뉴럴 네트워크를 트레이닝시키기 위한 방법(900)을 예시한다. 방법(900)의 단계들은 도시된 단계와 상이한 순서로 수행될 수 있고, 방법(900)의 수행 동안에 방법(900)의 하나 이상의 단계들이 생략될 수 있다. 방법(900)을 참조하여 설명된 뉴럴 네트워크는 방법(800)을 참조하여 설명된 제2 뉴럴 네트워크일 수 있으며, 이는 본원에 설명된 PointHomographyNet일 수 있다. 방법(900)의 하나 이상의 단계들은 비-일시적 컴퓨터-판독 가능 매체에 포함된 명령들을 실행하도록 구성된 프로세서에 의해 수행 및/또는 개시될 수 있다.
- [0044] [0053] 단계(902)에서, 하나 이상의 3D 포인트 클라우드들이 생성된다. 3D 포인트 클라우드들 각각은 복수의 3D 포인트들을 포함할 수 있다. 일부 실시예들에서, 하나 이상의 3D 포인트 클라우드들은 다른 가능성들 중에서도 평면, 구체, 큐브와 같은 하나 이상의 기하학적 구조들을 랜덤하게 샘플링함으로써 생성된다. 예컨대, 특정 기하학적 구조(예컨대, 구체 또는 큐브)의 표면들은 복수의 3D 포인트들을 생성하도록 랜덤하게 샘플링될 수 있다. 대안적으로, 특정 기하학적 구조의 에지들이 랜덤하게 샘플링될 수 있거나, 표면들 및 에지들 둘 모두가 랜덤하게 샘플링될 수 있다. 일부 실시예들에서, 특정 기하학적 구조의 전체 볼륨이 랜덤하게 샘플링될 수 있다.
- [0045] [0054] 일부 실시예들에서, 단계들(904 내지 916) 각각은 하나 이상의 3D 포인트 클라우드들의 각각의 3D 포인트 클라우드에 대해 수행될 수 있다. 단계(904)에서, 3D 궤적은 복수의 포인트들 부근에서 결정될 수 있다. 예컨대, 모든 3D 궤적은 복수의 포인트들의 임계 거리 내에 있을 수 있다. 일부 실시예들에서, 3D 궤적은 랜덤 시작 위치 및 랜덤 종료 위치를 결정함으로써 형성된 선형 궤적이다. 다른 실시예들에서, 또는 동일한 실시예들에서, 3D 궤적은 비-선형 궤적(예컨대, 곡선)이거나, 랜덤 시작 위치, 랜덤 종료 위치 및 하나 이상의 중간 위치들을 결정함으로써 형성된 일련의 선형 궤적들이다.
- [0046] [0055] 단계(906)에서, 3D 궤적은 특정 제1 카메라 포즈 및 특정 제2 카메라 포즈를 획득하도록 샘플링될 수 있다. 일부 실시예들에서, 복수의 포인트들은 특정 제1 카메라 포즈 및 특정 제2 카메라 포즈로부터 적어도 부분적으로 볼 수 있다. 예컨대, 획득된 카메라 포즈들은 복수의 포인트들 중 적어도 25%, 50%, 75% 또는 100%를 보는 그러한 카메라 포즈들로 제한될 수 있다. 카메라 포즈가 미리 결정된 임계치(예컨대, 복수의 포인트들 중 적어도 50%가 볼 수 있음)를 만족시키지 않는다면, 카메라 포즈가 폐기되고, 다른 카메라 포즈를 획득하기 위해 3D 궤적이 재샘플링된다. 일부 실시예들에서, 획득된 카메라 포즈들은 서로 시각적 중첩의 적어도 일부 임계치(예컨대, 30%)를 갖도록 제한된다. 일부 실시예들에서, 시각적 중첩은, 특정 제1 카메라 포즈 및 특정 제2 카메라 포즈 모두에 의해 볼 수 있는, 복수의 포인트들 중 포인트들의 퍼센티지에 대응할 수 있다. 다른 실시예들에서, 시각적 중첩은 획득된 포즈들 사이의 공유된 시야들에 기반하여 계산될 수 있다.
- [0047] [0056] 단계(908)에서, 제1 2D 포인트 클라우드를 생성하기 위해, 특정 제1 카메라 포즈에 기반하여 복수의 포인트들이 제1 2D 평면 상에 투사되고, 유사하게, 제2 2D 포인트 클라우드를 생성하기 위해, 특정 제2 카메라 포즈에 기반하여 복수의 포인트들이 제2 2D 평면 상에 투사된다. 일부 실시예들에서, 제1 2D 평면은 특정 제1 카메라 포즈의 배향 벡터에 직교할 수 있고, 제2 2D 평면은 특정 제2 카메라 포즈의 배향 벡터에 직교할 수 있다.
- [0048] [0057] 단계(910)에서, 제1 2D 포인트 클라우드와 제2 2D 포인트 클라우드 사이의 실측-자료 호모그래피는 특정 제1 카메라 포즈 및 특정 제2 카메라 포즈에 기반하여 결정된다. 일부 실시예들에서, 실측-자료 호모그래피는, 먼저 특정 제1 카메라 포즈와 특정 제2 카메라 포즈 사이의 상대적인 회전 및 상대적인 병진운동을 결정하고, 이어서 뉴럴 네트워크에 의해 생성된 호모그래피들과 구조적으로 일치하는 호모그래피(행렬)를 형성함으로써 결정된다.
- [0049] [0058] 단계(912)에서, 제1 2D 포인트 클라우드 및 제2 2D 포인트 클라우드는 입력으로서 뉴럴 네트워크에 제공될 수 있고, 특정 호모그래피는 제1 2D 포인트 클라우드 및 제2 2D 포인트 클라우드에 기반하여 뉴럴 네트워크에 의해 생성될 수 있다.

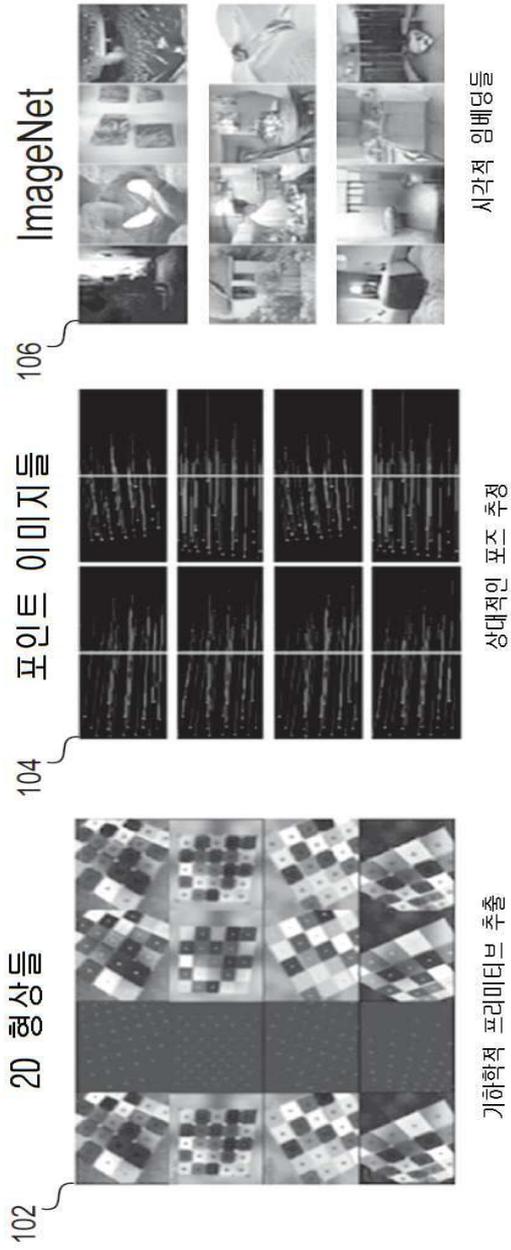
- [0050] [0059] 단계(914)에서, 특정 호모그래피는, 예컨대, 에러 신호를 생성하기 위해 실측-자료 호모그래피와 비교될 수 있다. 일부 실시예들에서, 에러 신호의 크기는 특정 호모그래피와 실측-자료 호모그래피 사이의 차이의 크기에 비례할 수 있다. 하나의 특정 실시예에서, 에러 신호는, 행렬들의 대응하는 엘리먼트들이 서로 감소되는 표준 행렬 감소 방법을 사용하여 계산된다. 다른 실시예들에서, 또는 동일한 실시예들에서, 에러 신호는 상대적인 회전에서의 차이에 대응하는 제1 컴포넌트 및 상대적인 병진운동에서의 차이에 대응하는 제2 컴포넌트를 포함한다. 일부 실시예들에서, 에러 신호는 상대적인 포즈에서의 차이에 대응하는 단일 컴포넌트를 포함한다.
- [0051] [0060] 단계(916)에서, 뉴럴 네트워크는, 예컨대, 뉴럴 네트워크의 하나 이상의 가중치들 또는 계수들을 조정함으로써, 단계(914)에서 수행된 특정 호모그래피와 실측-자료 호모그래피 사이의 비교에 기반하여 수정된다. 일부 실시예들에서, 뉴럴 네트워크는, 더 큰 에러 신호가 뉴럴 네트워크에 대한 더 큰 수정을 발생시키도록, 호모그래피들 사이의 계산된 차이(즉, 에러 신호)에 기반하여 수정될 수 있다. 일반적으로, 뉴럴 네트워크를 수정하는 것은 뉴럴 네트워크가 더 정확해지게 하고, 이로써 특정 호모그래피와 실측-자료 호모그래피 사이의 차이를 감소시킨다.
- [0052] [0061] 도 10은 본원에 설명된 실시예들 중 하나 이상을 사용할 수 있는 웨어러블 AR 디바이스(1000)의 개략도를 예시한다. AR 디바이스(1000)는 좌측 접안렌즈(1002A), 우측 접안렌즈(1002B), 좌측 접안렌즈(1002A) 상에 직접적으로 또는 그 근처에 부착된 좌측 전방향 카메라(1006A), 우측 접안렌즈(1002B) 상에 직접적으로 또는 그 근처에 부착된 우측 전방향 카메라(1006B), 좌측 측방향 카메라(1006C), 우측 측방향 카메라(1006D) 및 프로세싱 모듈(1050)을 포함할 수 있다. AR 디바이스(1000)의 컴포넌트들 중 일부 또는 모두는, 투사된 이미지들이 사용자에게 보여질 수 있도록 머리 장착될 수 있다. 하나의 특정 구현에서, 도 10에 도시된 AR 디바이스(1000)의 모든 컴포넌트들은 사용자가 착용할 수 있는 단일 디바이스(예컨대, 단일 헤드셋)에 장착된다. 다른 구현에서, 프로세싱 모듈(1050)은 AR 디바이스(1000)의 다른 컴포넌트들로부터 물리적으로 분리되고, 유선 또는 무선 연결에 의해 통신 가능하게 커플링된다. 예컨대, 프로세싱 모듈(1050)은 프레임에 고정적으로 부착되거나, 사용자가 착용한 헬멧 또는 모자에 고정적으로 부착되거나, 헤드폰들에 임베딩되거나, 또는 그렇지 않다면 사용자에게 제거 가능하게 부착되는 것(예컨대, 백팩 스타일 구성, 벨트 커플링 스타일 구성 등)과 같은 다양한 구성으로 장착될 수 있다.
- [0053] [0062] 프로세싱 모듈(1050)은 프로세서(1052)뿐만 아니라 디지털 메모리, 이를테면, 비-휘발성 메모리(예컨대, 플래시 메모리)를 포함할 수 있으며, 이들 둘 모두는 데이터의 프로세싱, 캐싱 및 저장을 보조하는 데 활용될 수 있다. 데이터는 이미지 캡처 디바이스들(예컨대, 카메라들(1006)), 마이크로폰들, 관성 측정 유닛들, 가속도계들, 컴퍼스들, GPS 유닛들, 라디오 디바이스들 및/또는 자이로들로부터 캡처된 데이터를 포함할 수 있다. 예컨대, 프로세싱 모듈(1050)은 카메라(1006)로부터 이미지(들)(1020), 또는 더 구체적으로, 좌측 전방향 세계 카메라(1006A)로부터 좌측 전방 이미지(들)(1020A), 우측 전방향 세계 카메라(1006B)로부터 우측 전방 이미지(들)(1020B), 좌측 측방향 세계 카메라(1006C)로부터 좌측 이미지(들)(1020C), 및 우측 측방향 세계 카메라(1006D)로부터 우측 이미지(들)(1020D)를 수신할 수 있다. 일부 실시예들에서, 이미지(들)(1020)는 단일 이미지, 한 쌍의 이미지들, 이미지들의 스트림을 포함하는 비디오, 페어링된 이미지들의 스트림을 포함하는 비디오 등을 포함할 수 있다. 이미지(들)(1020)는, AR 디바이스(1000)에 전력이 공급되는 동안 주기적으로 생성되어 프로세싱 모듈(1050)에 전송될 수 있거나, 프로세싱 모듈(1050)에 의해 카메라들 중 하나 이상에 전송된 명령에 대한 응답으로 생성될 수 있다.
- [0054] [0063] 접안렌즈들(1002A 및 1002B)은, 투사기들(1014A 및 1014B) 각각으로부터의 광을 지향시키도록 구성된 투명 또는 반투명 도파관들을 포함할 수 있다. 구체적으로, 프로세싱 모듈(1050)은 좌측 투사기(1014A)로 하여금 좌측 투사된 이미지(1022A)를 좌측 접안렌즈(1002A)에 출력하게 할 수 있고, 우측 투사기(1014B)로 하여금 우측 투사된 이미지(1022B)를 우측 접안렌즈(1002B)에 출력하게 할 수 있다. 일부 실시예들에서, 접안렌즈들(1002) 각각은, 상이한 컬러들 및/또는 상이한 깊이 평면들에 대응하는 복수의 도파관들을 각각 포함할 수 있다.
- [0055] [0064] 카메라들(1006A 및 1006B)은 사용자의 좌측 및 우측 눈들 각각의 시야와 실질적으로 중첩되는 이미지들을 캡처하도록 포지셔닝될 수 있다. 따라서, 카메라들(1006A 및 1006B)의 배치는 사용자의 눈들 근처에 있을 수 있지만 사용자의 시야를 가릴 만큼 그렇게 근처는 아니다. 대안적으로 또는 부가적으로, 카메라들(1006A 및 1006B)은, 투사된 이미지들(1022A 및 1022B) 각각의 인커플링 위치들과 정렬되도록 포지셔닝될 수 있다. 카메라들(1006C 및 1006D)은 사용자의 측면에 대해, 예컨대, 사용자의 주변 비전에서 또는 사용자의 주변 비전 외부에서 이미지들을 캡처하도록 포지셔닝될 수 있다. 카메라들(1006C 및 1006D)을 사용하여 캡처된 이미지(들)(1020C 및 1020D)는 카메라들(1006A 및 1006B)을 사용하여 캡처된 이미지(들)(1020A 및 1020B)와 반드시

증첩할 필요는 없다.

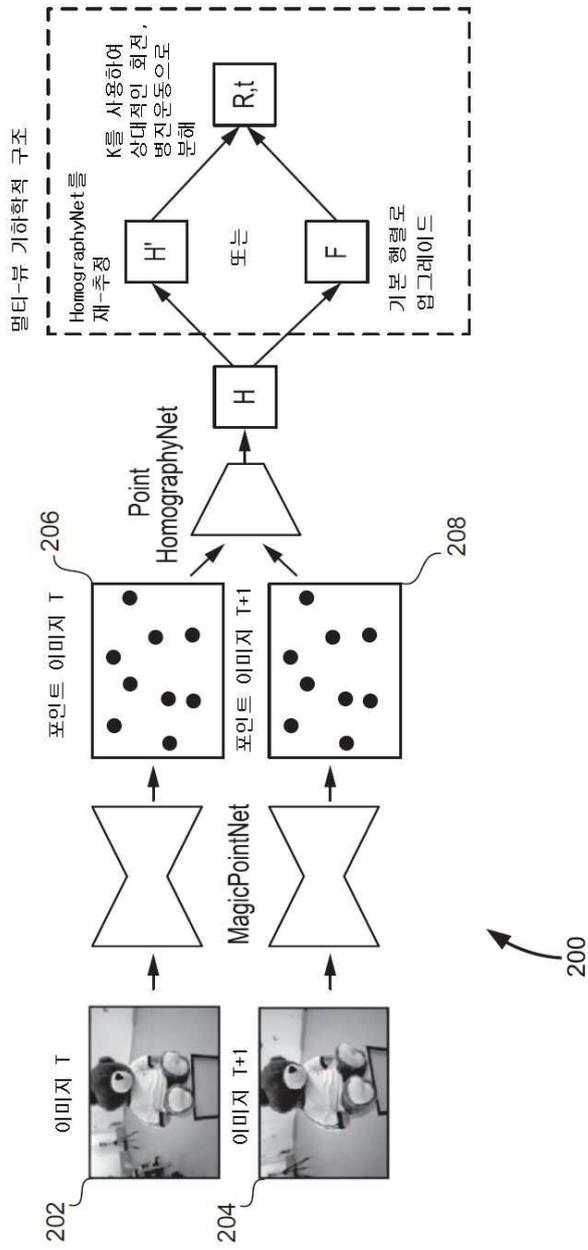
- [0056] [0065] AR 디바이스(1000)의 동작 동안, 프로세싱 모듈(1050)은 카메라들(1006) 중 임의의 카메라에 의한 2개의 캡처된 이미지들에 기반하여 호모그래피를 컴퓨팅하기 위해 트레이닝된 네트워크(1056)를 사용할 수 있다. 추정된 호모그래피는, 사용자의 움직임으로 인한 사용자의 시야에서의 변화들을 더 정확하게 반영하는 가상 콘텐츠를 렌더링하기 위해 프로세서(1052)에 의해 사용될 수 있다. 네트워크(1056)는 인공 뉴럴 네트워크, 콘볼루션 뉴럴 네트워크, 심층 네트워크, 또는 프로세싱 예들에 의해 점진적으로 "학습"할 수 있는 임의의 타입의 네트워크 또는 시스템일 수 있다. 일부 실시예들에서, 네트워크(1056)는 서로 간에 신호들을 송신할 수 있는 연결된 노드들의 집합을 포함한다. 프로세서(1052)는 단일 네트워크(1056)와 통신할 수 있거나, 일부 실시예들에서, 프로세서(1052)는 제1 네트워크(예컨대, MagicPointNet에 대응함), 제2 네트워크(예컨대, PointHomographyNet에 대응함) 및 제3 네트워크(예컨대, RelocNet에 대응함) 등과 같은 다수의 뉴럴 네트워크들과 통신할 수 있다.
- [0057] [0066] 몇몇의 예시적인 구성들을 설명하였지만, 다양한 수정들, 대안적인 구성들, 및 등가물들은 본 개시내용의 사상을 벗어나지 않으면서 사용될 수 있다. 예컨대, 위의 엘리먼트들은 더 큰 시스템의 컴포넌트일 수 있으며, 여기서, 다른 규칙들은 본 기술의 애플리케이션에 우선할 수 있거나 그렇지 않으면 본 발명의 애플리케이션을 수정할 수 있다. 또한, 다수의 단계들은, 위의 엘리먼트들이 고려되기 전에, 그 동안, 또는 그 이후에 착수될 수 있다. 따라서, 위의 설명은 청구항들의 범위를 제한하지 않는다.
- [0058] [0067] 본원에서 그리고 첨부된 청구항들에 사용된 바와 같이, 단수 형태들은, 맥락에서 명확히 다르게 나타내지 않으면 복수의 지시대상들을 포함한다. 따라서, 예컨대, "사용자"에 대한 참조는 복수의 그러한 사용자들을 포함하고, "프로세서"에 대한 참조는 당업자에게 알려진 하나 이상의 프로세서들 및 그의 등가물들에 대한 참조를 포함하는 식이다.
- [0059] [0068] 또한, 본 명세서에서 그리고 다음의 청구항들에서 사용될 때, "포함하는"("comprise", "comprising", "contains", "containing", "include", "including", 및 "includes")이란 단어는 언급된 특징들, 정수들, 컴포넌트들 또는 단계들의 존재를 특정하도록 의도되지만, 이들은 하나 이상의 다른 특징들, 정수들, 컴포넌트들, 단계들, 동작들 또는 그룹들의 존재 또는 부가를 배제하지 않는다.
- [0060] [0069] 또한, 본원에서 설명된 예들 및 실시예들은 단지 예시 목적들을 위한 것이고, 이에 대한 다양한 수정들 또는 변화들이 당업자에게 연상될 것이며 본 출원의 사상 및 범위 및 첨부된 청구항들의 범위 내에 포함될 것이 이해된다.

도면

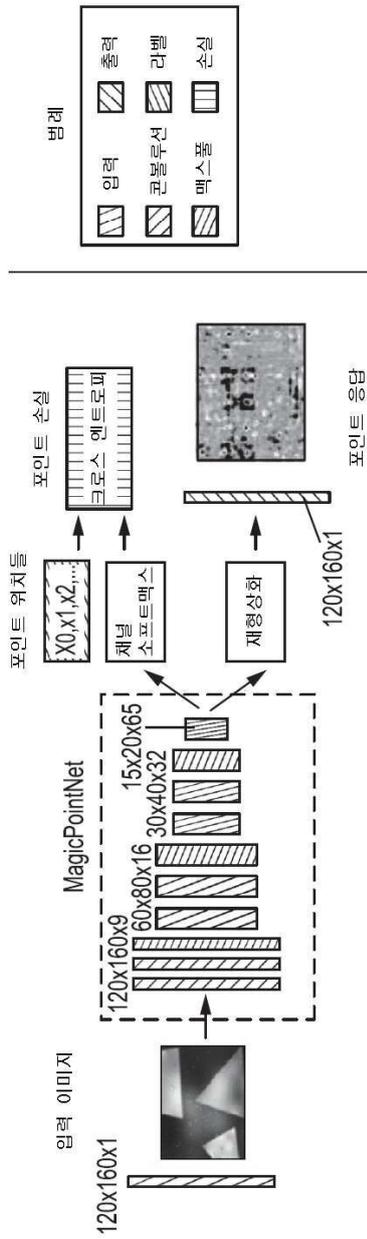
도면1



도면2

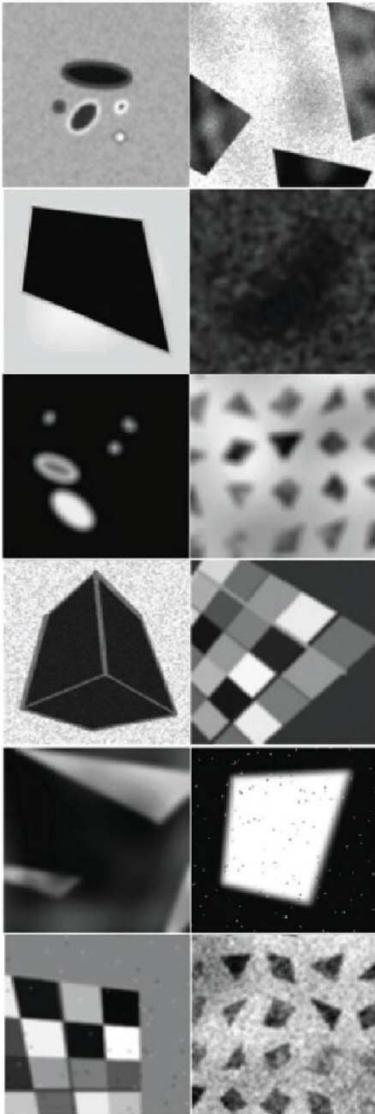


도면3

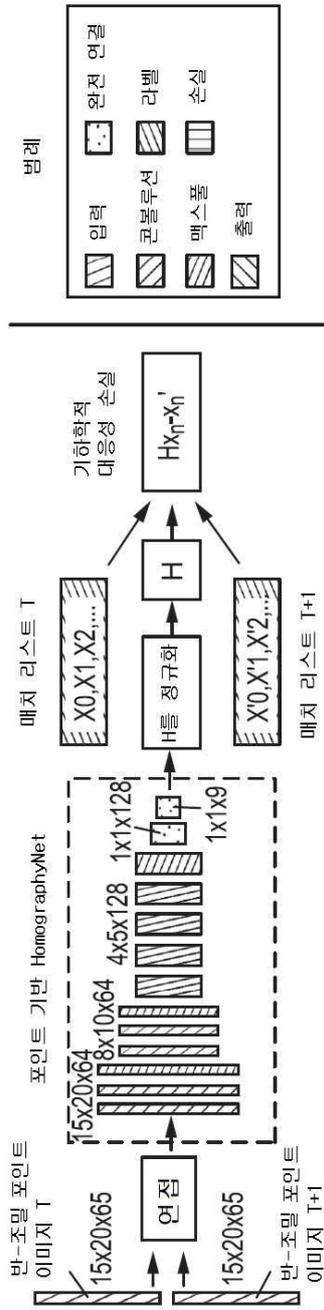


300

도면4



도면5



500

도면6

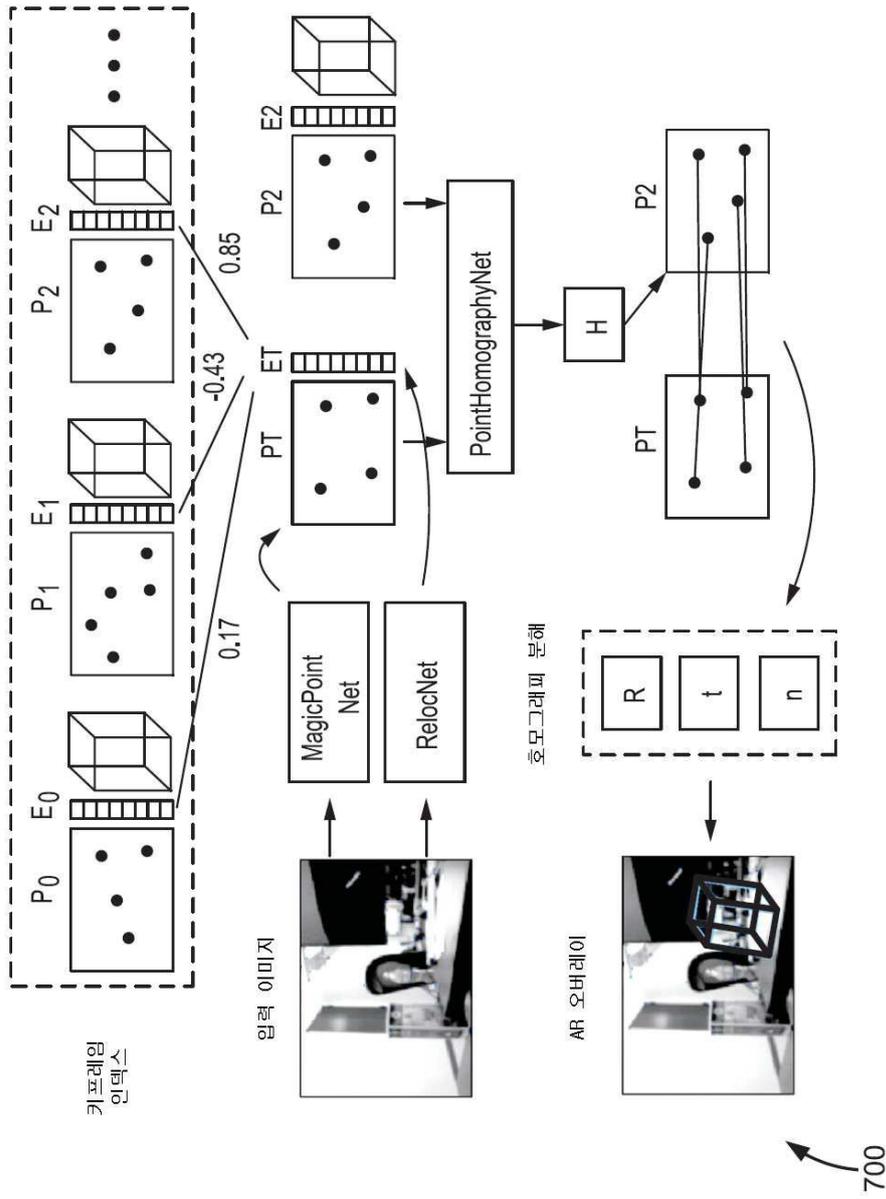
샘플링면 기하학적 구조들



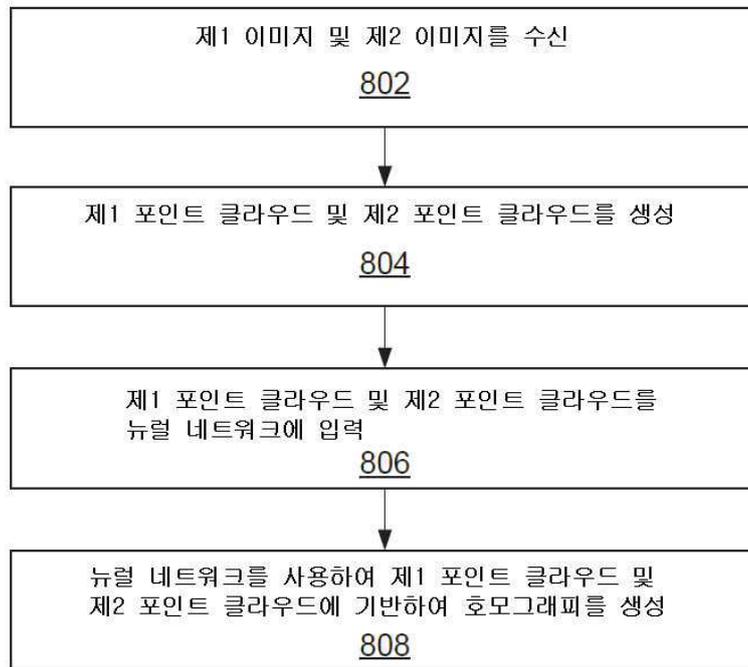
샘플링면 궤적들



도면7

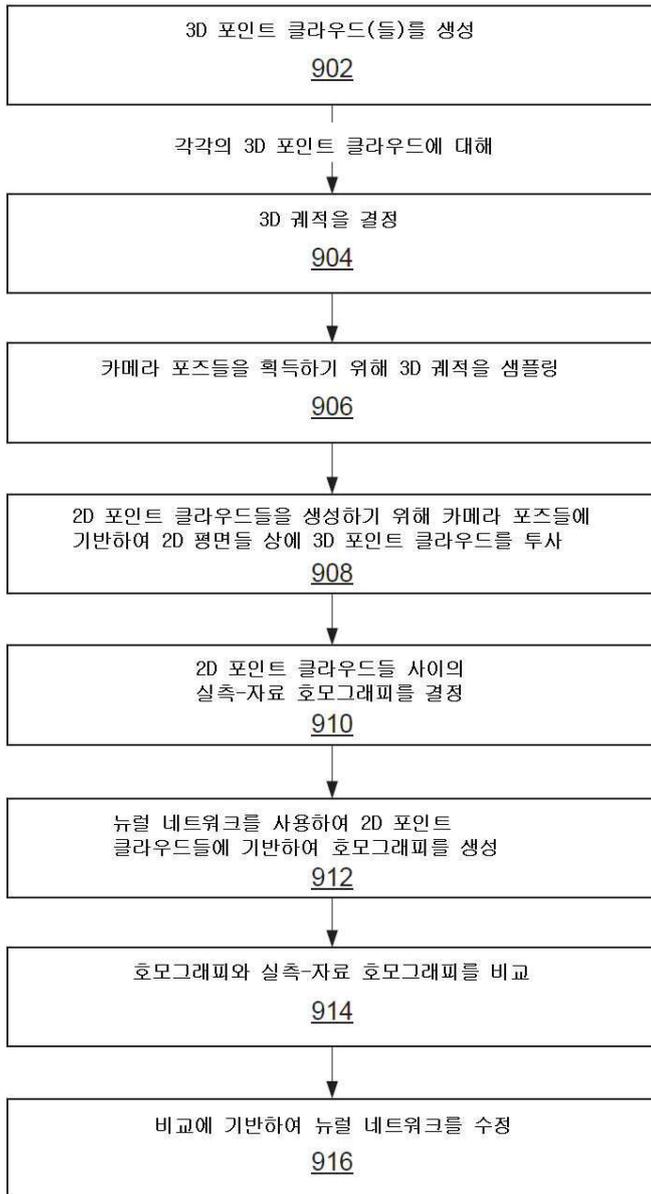


도면8



800 ↗

도면9



900

도면10

