



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2019-0101382
(43) 공개일자 2019년08월30일

(51) 국제특허분류(Int. Cl.)
G06T 7/12 (2017.01) G06T 5/00 (2019.01)
G06T 7/50 (2017.01)
(52) CPC특허분류
G06T 7/12 (2017.01)
G06T 5/001 (2013.01)
(21) 출원번호 10-2019-7018567
(22) 출원일자(국제) 2017년12월21일
심사청구일자 없음
(85) 번역문제출일자 2019년06월26일
(86) 국제출원번호 PCT/EP2017/084008
(87) 국제공개번호 WO 2018/122087
국제공개일자 2018년07월05일
(30) 우선권주장
16306836.4 2016년12월28일
유럽특허청(EPO)(EP)

(71) 출원인
인터디지털 씨이 페턴트 홀딩스
프랑스 75017 파리 튀 뒤 꼴로벨 몰 3
(72) 발명자
루오 타오
프랑스 35576 세송-세비네 아브뉴 데 상 블랑 975
씨에스 17616 테크니컬러 씨/오
(74) 대리인
특허법인코리아나

전체 청구항 수 : 총 15 항

(54) 발명의 명칭 장면의 공동 세그먼테이션 및 3D 복원을 위한 방법 및 디바이스

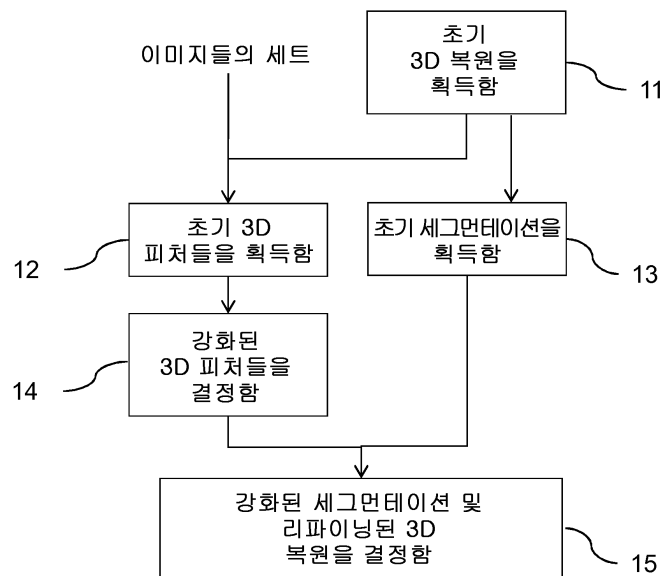
(57) 요약

장면의 적어도 하나의 이미지의 세트로부터, 장면의 공동 세그먼테이션 및 3D 복원을 위한 방법은,

- 장면의 초기 3D 복원을 획득하는 단계 (11);
- 상기 초기 3D 복원과 연관된 초기 3D 피쳐들을 획득하는 단계 (12);

(뒷면에 계속)

대표도 - 도1



상기 초기 3D 복원의 초기 세그멘테이션을 획득하는 단계 (13);

상기 초기 3D 피쳐들로부터 및 상기 세트의 적어도 하나의 이미지에서 결정된 초기 2D 피쳐들로부터, 상기 장면의 상기 초기 3D 복원과 연관된 상기 초기 3D 피쳐들에 대응하는 것으로서, 강화된 3D 피쳐들을 결정하는 단계 (14)로서, 상기 강화된 3D 피쳐들은 적어도 부분적으로 상기 초기 세그멘테이션에 대응하는, 상기 강화된 3D 피쳐들을 결정하는 단계 (14); 및

상기 초기 세그멘테이션 및 상기 강화된 3D 피쳐들 양자로부터, 강화된 세그멘테이션 및 리파이닝된 3D 복원을 결정하는 단계 (15)를 포함한다. 증강 현실로의 적용.

(52) CPC특허분류

G06T 7/50 (2017.01)

G06T 2207/10012 (2013.01)

G06T 2207/10028 (2013.01)

명세서

청구범위

청구항 1

장면의 적어도 하나의 이미지의 세트로부터, 상기 장면의 공동 세그먼테이션 및 3D 복원을 위한 방법으로서, 상기 장면의 상기 세그먼테이션은 상기 장면의 상기 3D 복원을 세그먼트들로 파티셔닝하는 것에 대응하고, 상기 방법은,

- 상기 장면의 초기 3D 복원을 획득하는 단계 (11);
- 상기 초기 3D 복원과 연관된 초기 3D 피쳐들을 획득하는 단계 (12);
- 상기 초기 3D 복원의 초기 세그먼테이션을 획득하는 단계 (13);
- 상기 초기 3D 피쳐들로부터 및 상기 세트의 적어도 하나의 이미지에서 결정된 초기 2D 피쳐들로부터, 상기 장면의 상기 초기 3D 복원과 연관된 상기 초기 3D 피쳐들에 대응하는 것으로서, 강화된 3D 피쳐들을 결정하는 단계 (14)로서, 상기 강화된 3D 피쳐들은 적어도 부분적으로 상기 초기 세그먼테이션에 대응하는, 상기 강화된 3D 피쳐들을 결정하는 단계 (14); 및
- 상기 초기 세그먼테이션 및 상기 강화된 3D 피쳐들 양자로부터, 강화된 세그먼테이션 및 리파이닝된 3D 복원 양자를 결정하는 단계 (15)를 포함하는, 장면의 공동 세그먼테이션 및 3D 복원을 위한 방법.

청구항 2

제 1 항에 있어서,

상기 3D 피쳐들은 3D 피쳐 라인들이고, 상기 2D 피쳐들은 2D 피쳐 라인들인, 장면의 공동 세그먼테이션 및 3D 복원을 위한 방법.

청구항 3

제 1 항 또는 제 2 항에 있어서,

상기 장면의 상기 초기 3D 복원을 획득하는 단계 (11)는 심도 데이터로부터 상기 초기 3D 복원을 구성하는 단계를 포함하는, 장면의 공동 세그먼테이션 및 3D 복원을 위한 방법.

청구항 4

제 1 항 내지 제 3 항 중 어느 한 항에 있어서,

상기 초기 3D 피쳐들을 획득하는 단계 (12)는 지오메트리 특징들 및/또는 로컬 피쳐 디스크립터들을 사용하여 상기 장면의 상기 초기 3D 복원에서 3D 피쳐들을 식별하는 단계를 포함하는, 장면의 공동 세그먼테이션 및 3D 복원을 위한 방법.

청구항 5

제 1 항 내지 제 4 항 중 어느 한 항에 있어서,

상기 장면의 상기 적어도 하나의 이미지의 상기 세트는 적어도 2 개의 이미지들을 포함하고,

상기 방법은,

- 가시적 이미지들로 알려진 상기 초기 3D 피쳐들을 포함하는 상기 세트의 이미지들을 선택하는 것, 및
- 상기 가시적 이미지들에서, 상기 초기 3D 피쳐들과 매칭하는 상기 초기 2D 피쳐들을 식별하는 것으로부터 상기 초기 2D 피쳐들을 결정하는 단계를 포함하고, 그리고

상기 강화된 3D 피쳐들을 결정하는 단계는,

- 적어도 2 개의 가시적 이미지들에 걸쳐 상기 초기 2D 피쳐들을 매칭함으로써 지오메트릭 큐들을 생성하는 단계, 및
- 상기 강화된 3D 피쳐들을 결정하기 위해 상기 지오메트릭 큐들로 상기 초기 3D 피쳐들을 강화시키는 단계를 포함하는, 장면의 공동 세그멘테이션 및 3D 복원을 위한 방법.

청구항 6

제 1 항 내지 제 5 항 중 어느 한 항에 있어서,

- 상기 세트의 상기 적어도 하나의 이미지에서 결정된 강화된 2D 피쳐들로부터 및 상기 강화된 3D 피쳐들로부터, 상기 장면의 상기 리파이닝된 3D 복원과 연관된 상기 강화된 3D 피쳐들에 대응하는 것으로서, 추가로 강화된 3D 피쳐들을 결정하는 단계; 및

상기 강화된 세그멘테이션 및 상기 추가로 강화된 3D 피쳐들로부터, 추가로 강화된 세그멘테이션 및 추가로 리파이닝된 3D 복원을 결정하는 단계

의 적어도 하나의 반복을 포함하는, 장면의 공동 세그멘테이션 및 3D 복원을 위한 방법.

청구항 7

제 6 항에 있어서,

상기 반복들은, 상기 추가로 강화된 3D 피쳐들과 상기 강화된 2D 피쳐들 사이의 적어도 매칭에 대한 미리 결정된 정밀도 임계치가 도달될 때 중지되는, 장면의 공동 세그멘테이션 및 3D 복원을 위한 방법.

청구항 8

제 7 항에 있어서,

상기 미리 결정된 정밀도 임계치는, 상기 장면의 상기 3D 복원을 상기 세그먼트들로 파티셔닝하는 정도, 상기 추가로 리파이닝된 3D 복원에서 측정된 이웃하는 유사한 3D 엘리먼트들 사이의 라벨들의 일관도, 및 상기 세트의 상기 적어도 하나의 이미지와 상기 추가로 리파이닝된 3D 복원 간의 정렬에 의해 제공되는, 세그멘테이션 레벨 중 적어도 하나에 공동으로 적용되는, 장면의 공동 세그멘테이션 및 3D 복원을 위한 방법.

청구항 9

제 6 항 내지 제 8 항 중 어느 한 항에 있어서,

상기 반복들은 미리 결정된 수의 반복들이 도달될 때 중지되는, 장면의 공동 세그멘테이션 및 3D 복원을 위한 방법.

청구항 10

제 1 항 내지 제 9 항 중 어느 한 항에 있어서,

상기 강화된 세그멘테이션을 결정하는 단계는 세그멘테이션 제약들에 의존하는, 장면의 공동 세그멘테이션 및 3D 복원을 위한 방법.

청구항 11

제 10 항에 있어서,

상기 세그멘테이션 제약들은 적어도 하나의 세그먼트 형상과 관련되는, 장면의 공동 세그멘테이션 및 3D 복원을 위한 방법.

청구항 12

제 1 항 내지 제 11 항 중 어느 한 항에 있어서,

상기 초기 3D 복원 및 상기 적어도 하나의 이미지의 세트를 적어도 하나의 입력으로서 수신하는 단계, 적어도

하나의 프로세서로 상기 강화된 3D 피쳐들, 강화된 세그먼테이션 및 리파이닝된 3D 복원을 결정하는 단계, 및 상기 리파이닝된 3D 복원을 사용자에게 디스플레이하고 상기 강화된 세그먼테이션에 의해 상기 리파이닝된 3D 복원을 프로세싱하기 위한 적어도 하나의 출력으로부터 상기 강화된 세그먼테이션 및 상기 리파이닝된 3D 복원을 출력하는 단계를 포함하는, 장면의 공동 세그먼테이션 및 3D 복원을 위한 방법.

청구항 13

통신 네트워크로부터 다운로드가능하고 및/또는 컴퓨터에 의해 판독가능한 매체 상에 기록되고 및/또는 프로세서에 의해 실행가능한 컴퓨터 프로그램 제품으로서, 프로세서에 의해 실행될 경우, 제 1 항 내지 제 12 항 중 어느 한 항에 기재된 방법을 수행하도록 적응된 소프트웨어 코드를 포함하는, 컴퓨터 프로그램 제품.

청구항 14

장면의 적어도 하나의 이미지의 세트로부터, 상기 장면의 공동 세그먼테이션 및 3D 복원을 위한 디바이스로서,

상기 장면의 상기 세그먼테이션은 상기 장면의 3D 복원을 세그먼트들로 파티셔닝하는 것에 대응하고,

상기 디바이스는 적어도 하나의 프로세서를 포함하고, 상기 적어도 하나의 프로세서는,

- 상기 장면의 초기 3D 복원을 획득하고;
- 상기 초기 3D 복원과 연관된 초기 3D 피쳐들을 획득하고;
- 상기 초기 3D 복원의 초기 세그먼테이션을 획득하고;
- 상기 초기 3D 피쳐들로부터 및 상기 세트의 적어도 하나의 이미지에서 결정된 초기 2D 피쳐들로부터, 상기 장면의 상기 초기 3D 복원과 연관된 상기 초기 3D 피쳐들에 대응하는 것으로서, 강화된 3D 피쳐들을 결정하는 것으로서, 상기 강화된 3D 피쳐들은 적어도 부분적으로 상기 초기 세그먼테이션에 대응하는, 상기 강화된 3D 피쳐들을 결정하고; 그리고
- 상기 초기 세그먼테이션 및 상기 강화된 3D 피쳐들 양자로부터, 강화된 세그먼테이션 및 리파이닝된 3D 복원 양자를 결정하도록

적용되고 구성되는, 장면의 공동 세그먼테이션 및 3D 복원을 위한 디바이스.

청구항 15

장치로서,

제 14 항에 기재된 디바이스를 포함하고,

상기 장치는 바람직하게 모바일 전화, 태블릿, 또는 헤드 장착 디스플레이 중에서 선택된 모바일 장치, 또는 바람직하게 로봇, 자율 주행 장치, 또는 스마트 홈 장치 중에서 선택된 자율 장치인 것을 특징으로 하는 장치.

발명의 설명

기술 분야

[0001] 본 개시는 신호 프로세싱 분야에 관한 것으로, 보다 상세하게는 이미지 또는 비디오의 프로세싱에 관한 것이다.

[0002] 보다 상세하게는, 본 개시는 일부 종래 기술에 비해 장면의 세그먼테이션 및 복원을 개선하는 것을 목표로 하는, 장면의 공동 세그먼테이션 및 3D 복원을 위한 방법에 관한 것이다.

[0003] 본 개시는 특히, 3D 복원이 관심있는 모든 애플리케이션들에 적용된다. 이것은 예를 들어, 네비게이션, 자율 로봇, 가상 현실, 증강 및/또는 복합 현실, 스마트 홈 장치 등과 같은 분야의 경우일 수 있다.

배경 기술

[0004] 본 섹션은 본 기술의 다양한 양태들을 독자에게 소개하기 위한 것이고, 이는 하기에서 설명되고 및/또는 청구되는 본 개시의 다양한 양태들에 관련될 수도 있다. 이 논의는, 본 개시의 다양한 양태들의 더 우수한 이해를 촉진하기 위한 배경 정보를 독자에게 제공함에 있어서 도움이 될 것으로 사료된다. 이에 따라, 이들 진술들은 종래 기술의 시인으로서가 아니라 이러한 관점에서 판독되어야 함이 이해되어야 한다.

- [0005] 심도 센서들의 개발로, 점점 더 많은 디바이스들이 3D 데이터를 처리해야 한다. 따라서, 캡처된 데이터를 프로세싱하고 더 양호한 장면 이해를 얻기 위해, 도전 과제들이 발생한다. 특히 세그먼테이션 및 3D 복원 양자는 장면의 정확한 3D 표현을 달성하는데 있어 중요하다.
- [0006] 3D 장면의 세그먼테이션은 3D 장면을 다수의 세그먼트들 또는 컴포넌트들로 파티셔닝하는 것으로 정의되며, 각각의 세그먼트는 이웃하는 픽셀들의 세트를 포함하고 라벨에 의해 유리하게 식별된다.
- [0007] 세그먼테이션 및 3D 복원은 먼저 개별적으로 고려되었다. 그 결과는 만족스럽지 못했다.
- [0008] 3D 메쉬의 형태로, 포인트 클라우드의 3D 복원을 강화시키는 것은 예를 들어, Autodesk, Inc. 의 특허 출원 US 2015/0146971 A1 에 기술된다. 이 문서에 따르면, 포인트 클라우드는 포토 이미지 데이터 및 스캔 데이터의 조합으로부터 생성되고, 초기의 러프한 메쉬는 포인트 클라우드 데이터로부터 추정되며, 그 러프한 메쉬는 3D 메쉬에서 이미지 쌍들 간의 포토-일관성을 최대화하고 3D 메쉬와 포인트 클라우드 사이의 3D 거리를 최소화함으로써 반복적으로 리파이닝된다.
- [0009] 세그먼테이션의 성능은 통상적으로 3D 복원에 의해 영향을 받고, 그 반대로 그러하기 때문에, 세그먼테이션 및 3D 복원이 공동으로 고려되었다. 이를 실행하기 위해, 종래 기술의 일부는 라벨링된 트레이닝 데이터세트에 기초한 공동 시맨틱 세그먼테이션 및 복원에 의존한다.
- [0010] 예를 들어, C. Hane 등은 "Joint 3D Scene Reconstruction and Class Segmentation" (IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2013) 에서, 공동 세그먼테이션 및 밀집된 복원 문제에 대한 해결책을 개시한다. 데이터 이미지 및 대응하는 심도 맵들은 입력으로 취득되고, 정확한 클래스 라벨들을 사용하는 3D 복원이 출력으로 생성된다. 저자는 종래의 체적 복원 방법을 멀티-라벨 체적 세그먼테이션 프레임워크로 확장한다. 이 기술에 따르면, 외형 기반의 큐 및 3D 표면 배향 프라이어어들은 트레이닝 데이터로부터 학습되며, 그 후 클래스별 정규화에 사용된다. 이 프라이어어들은 심도 맵에서 획득된 측정된 증거를 보완하여, 복원 및 라벨링을 함께 개선한다.
- [0011] A. Kundu 등에 의한 "Joint Semantic Segmentation and 3D Reconstruction from Monocular Video" (European Conference on Computer Vision, 2014) 에서, 모노컬러 이미지 스트림에서 시작하여, 시각적 SLAM ("Simultaneous Localization And Mapping") 및 초기 2D 장면 파싱이 수행된다. 이 기술은 3D 구조와 시맨틱 라벨들 양자를 나타내는 3D 맵을 생성한다. 이 기술에 따르면, 카테고리별 센서 모델들은 SLAM 에서 심도 추정들을 향상시키는데 사용되고, 연속적인 카메라 위치에서 비어있는 공간을 인지하는 것은 구조적 모호성을 감소시키는데 도움이 된다.
- [0012] 위에서 언급한 기술들 양자는 시맨틱 세그먼테이션을 고려하고, 3D 복원을 달성하기 위해 오브젝트 카테고리별 큐들을 채용한다. 따라서, 이러한 기술들의 성능은 트레이닝 데이터에 의존하고, 특히 오브젝트 카테고리들의 수와 같은 데이터세트의 확장성 (scalability) 에 의존한다. 또한, 복원은 체적 데이터로 표현되므로, 공간 해상도 측면에서 제한적이다. 추가로, 최종 3D 복원은 지오메트리 측면에서 상대적으로 코오스 (coarse) 하다. 예를 들어, 오브젝트의 날카로운 에지들은 종종 평활하고, 직선들은 종종 잡음 데이터의 영향을 받는다. 따라서, 3D 복원은 일부 애플리케이션들에서 더 세밀한 상호작용들을 수행하기에 충분히 정확하지 않다.

발명의 내용

해결하려는 과제

- [0013] 따라서, 특히 오브젝트 지오메트리의 양호한 복원 품질을 허용하는 장면의 공동 세그먼테이션 및 3D 복원을 위한 방법이 필요하다.

과제의 해결 수단

- [0014] 본 개시는 장면의 적어도 하나의 이미지의 세트로부터, 장면의 공동 세그먼테이션 및 3D 복원을 위한 방법에 관한 것으로, 장면의 세그먼테이션은 장면의 3D 복원을 세그먼트들로 파티셔닝하는 것에 대응하고, 상기 방법은:

- [0015] - 장면의 초기 3D 복원을 획득하는 단계;
- [0016] - 상기 초기 3D 복원과 연관된 초기 3D 피쳐들을 획득하는 단계;

- [0017] - 상기 초기 3D 복원의 초기 세그먼테이션을 획득하는 단계;
- [0018] - 상기 초기 3D 피쳐들로부터 및 상기 세트의 적어도 하나의 이미지에서 결정된 초기 2D 피쳐들로부터, 상기 장면의 상기 초기 3D 복원과 연관된 상기 초기 3D 피쳐들에 대응하는 것으로서, 강화된 3D 피쳐들을 결정하는 단계로서, 상기 강화된 3D 피쳐들은 적어도 부분적으로 상기 초기 세그먼테이션에 대응하는, 상기 강화된 3D 피쳐들을 결정하는 단계; 및
- [0019] - 상기 초기 세그먼테이션 및 상기 강화된 3D 피쳐들 양자로부터, 강화된 세그먼테이션 및 리파이닝된 3D 복원 양자를 결정하는 단계를 포함한다.
- [0020] 따라서, 본 개시는 장면의 공동 세그먼테이션 및 3D 복원을 위한 신규하고 독창적인 해결책을 제안하며 (여기서 장면은 특히 오브젝트일 수도 있음), 기술한 단점들 중 적어도 하나를 극복한다. 특히, 본 개시는 트레이닝 데이터 세트에 의존하지 않는다.
- [0021] 세그먼테이션의 성능은 장면의 3D 복원에 의해 개선될 수 있고, 그 반대도 그러하며, 세그먼테이션 및 복원 양자는 서로 기여할 수 있고 공동으로 고려될 수 있다.
- [0022] 세그먼테이션 및 3D 복원은, 세그먼테이션이 3D 복원에 의해 영향을 받고 3D 복원은 세그먼테이션의 영향을 받는다는 점에서 "공동" 이다. 이것은 오직 초기 세그먼테이션뿐만 아니라 강화된 3D 피쳐들로부터의 강화된 세그먼테이션의 결정에 의해, 그리고 강화된 3D 피쳐들뿐만 아니라 초기 세그먼테이션으로부터의 리파이닝된 3D 복원의 결정에 의해 표현된다.
- [0023] 보다 구체적으로, 본 개시는 초기 3D 모델이라고도 불리는 장면의 초기 3D 복원을 리파이닝하고 이미지 데이터에서 결정된 초기 2D 피쳐들로 인해 세그먼테이션을 강화하기 위한 해결책을 제공한다. 따라서, 리파이닝된 3D 복원 및 강화된 세그먼테이션은 본 개시의 적어도 하나의 실시형태에 따라 공동으로 결정된다.
- [0024] 2D 피쳐들을 고려함으로써, 특히 지오메트리의 관점에서 장면의 정확한 3D 복원이 달성될 수 있다. 또한, 강화된 세그먼테이션이 획득될 수 있다.
- [0025] 그러한 리파이닝된 또는 정확한 3D 복원 및 강화된 세그먼테이션은 텍스처 맵핑, 변형, 증강 현실에서의 충돌 검출 등과 같은 추가의 애플리케이션들에서 사용될 수 있다.
- [0026] 예를 들어, 장면의 3D 복원은 다음을 포함하는 그룹에 속한다:
- [0027] - 포인트 클라우드,
- [0028] - 메쉬 모델,
- [0029] - 체적 모델.
- [0030] 따라서, 세그먼테이션은 리파이닝된 3D 복원 (즉, 강화된 3D 피쳐들로부터 리파이닝되는 초기 3D 복원) 에서 3D 엘리먼트들의 컴포넌트들의 라벨들을 업데이트하여 강화될 수 있다. "3D 엘리먼트" 는 예를 들어, 포인트들의 클라우드 중의 포인트, 다각형 메쉬 모델의 다각형, 체적 모델의 복셀 (voxel) 등이고, "컴포넌트" 는 평면 영역과 같은, 동일한 라벨을 갖는 3D 엘리먼트들의 그룹이다.
- [0031] 구현들에 의존하여, 강화된 3D 피쳐들은 초기 3D 피쳐들을 통해 및/또는 초기 3D 피쳐들로부터 및 초기 2D 피쳐들로부터 강화된 3D 피쳐를 결정하는 것을 통해 초기 세그먼테이션에 적어도 부분적으로 대응한다.
- [0032] 따라서, 특정 구현들에서, 초기 세그먼테이션의 컴포넌트들 사이의 경계들은 초기 3D 피쳐 포인트들 또는 피쳐 라인들을 초기 3D 피쳐들의 적어도 일부로서 제공하고 있다. 그 후, 후자는 강화된 3D 피쳐들을 구성하는데 채용되고, 이들은 자체적으로 강화된 세그먼테이션 및 리파이닝된 3D 복원을 결정하는데 사용된다.
- [0033] 이전의 구현들과 결합될 수 있는 다른 구현들에서, 강화된 3D 피쳐들은 초기 3D 피쳐들 및 초기 2D 피쳐들뿐만 아니라 초기 세그먼테이션으로부터 결정되며, 이에 의해 강화된 3D 피쳐에 직접 기여하고, 따라서 리파이닝된 3D 복원에 직접 기여한다. 일부 관련된 실시형태들에서, 리파이닝된 3D 복원은 초기 3D 피쳐들, 초기 2D 피쳐들 및 초기 세그먼테이션으로부터, 강화된 3D 피쳐들과 함께 결정된다.
- [0034] 강화된 세그먼테이션에 관하여, 강화된 3D 피쳐들을 채용함으로써 초기 세그먼테이션으로부터 도출된다.
- [0035] 일관적으로, 강화된 세그먼테이션은 초기 세그먼테이션 및 강화된 3D 피쳐들 양자로부터 결정되지만, 리파이닝된 3D 복원은 또한 초기 세그먼테이션과 강화된 3D 피쳐들 양자로부터 결정된다 (심지어 위의 특정 구현들에서

럼, 초기 세그먼테이션은 강화된 3D 피쳐들을 통해 고려될 수 있다).

- [0036] 일 실시형태에 따르면, 3D 피쳐들은 3D 피쳐 라인들이고, 2D 피쳐들은 2D 피쳐 라인들이다. 다른 실시형태에 따르면, 3D 피쳐들은 3D 포인트들이고, 2D 피쳐들은 2D 포인트들이다.
- [0037] 따라서, 세그먼테이션은 시맨틱 피쳐들이 아닌 지오메트릭 피쳐들을 기반으로 한다.
- [0038] 따라서, 본 실시형태에 따른 세그먼테이션 및 3D 복원은 시맨틱/라벨링된 트레이닝 데이터의 품질 및/또는 확장성에 의존하지 않는다.
- [0039] 본 개시의 적어도 하나의 실시형태는 예를 들어, RGB-D 데이터로부터 (적색 녹색 청색 및 심도 데이터로부터), 리파이닝된 지오메트리를 갖는 세그먼트화된 영역들의 세트를 결정하는 것을 목적으로 하여, 장면의 세그먼테이션 및 3D 복원의 공동 최적화를 위한 알고리즘을 개시한다. 리파이닝된 지오메트리는 세그먼테이션을 더 정확하게 만들고, 더 정확한 세그먼테이션은 지오메트리의 리파인먼트를 위해 추가의 지오메트릭 큐들을 제공한다.
- [0040] 일 실시형태에 따르면, 장면의 초기 3D 복원을 획득하는 단계는 심도 데이터로부터 초기 3D 복원을 구성하는 단계를 포함한다. 따라서, 장면의 초기 3D 복원은 업스트림으로 결정되어 동작중인 장치에서 직접 수신되거나, 또는 동작중인 장치에서 구성될 수 있다.
- [0041] 일 실시형태에 따르면, 초기 3D 피쳐들을 획득하는 단계는 지오메트리 특징들 및/또는 로컬 피쳐 디스크립터들을 사용하여 장면의 초기 3D 복원에서 3D 피쳐들을 식별하는 단계를 포함한다. 대안적으로, 초기 3D 피쳐들이 업스트림으로 결정되어 동작중인 장치에서 직접 수신될 수도 있다.
- [0042] 장면의 이미지(들)의 세트가 적어도 2 개의 이미지들을 포함하는 일 실시형태에 따르면, 상기 방법은,
- [0043] - 가시적 이미지들로 알려진 초기 3D 피쳐들을 포함하는 세트의 이미지들을 선택하는 것, 및
- [0044] - 상기 가시적 이미지들에서, 초기 3D 피쳐들과 매칭하는 초기 2D 피쳐들을 식별하는 것으로부터 초기 2D 피쳐들을 결정하는 단계를 포함하고,
- [0045] 그리고 강화된 3D 피쳐들을 결정하는 단계는,
- [0046] - 적어도 2 개의 가시적 이미지들에 걸쳐 상기 초기 2D 피쳐들을 매칭함으로써 지오메트릭 큐들을 생성하는 단계, 및
- [0047] - 강화된 3D 피쳐들을 결정하기 위해 지오메트릭 큐들로 초기 3D 피쳐들을 강화시키는 단계를 포함한다.
- [0048] 따라서, 초기 2D 피쳐들은 이미지 데이터로부터 결정되거나 (즉, 세트의 이미지들로부터 도출되거나), 또는 업스트림 프리-프로세싱 이후에 동작중인 장치에서 수신될 수 있다. 특히, 이미지들의 세트 중에서 가시적 이미지들의 선택은, 추가의 프로세싱이 계산적으로 효율적이게 할 수 있다. 또한 부정확한 카메라 포즈 추정들 (예를 들어, 위치 및/또는 방향과 관련하여) 에 의해 생성될 수 있는 에러들의 감소를 초래한다.
- [0049] 강화된 3D 피쳐들은 가시적 이미지들에 걸쳐 초기 2D 피쳐들을 매칭함으로써 결정될 수 있다. 그러한 2D 피쳐들의 매칭은 실제로 예를 들어, 멀티-뷰 스테레오 방법들을 채용함으로써 3D 지오메트릭 큐들을 구성하는데 사용된다.
- [0050] 일 실시형태에 따르면, 방법은:
- [0051] - 세트의 적어도 하나의 이미지에서 결정된 강화된 2D 피쳐들로부터 및 강화된 3D 피쳐들로부터, 상기 장면의 상기 리파이닝된 3D 복원과 연관된 상기 강화된 3D 피쳐들에 대응하는 것으로서, 추가로 강화된 3D 피쳐들을 결정하는 단계; 및
- [0052] - 강화된 세그먼테이션 및 추가로 강화된 3D 피쳐들로부터, 추가로 강화된 세그먼테이션 및 추가로 리파이닝된 3D 복원 양자를 결정하는 단계의 적어도 하나의 반복을 포함한다.
- [0053] 특히, 세트의 상기 이미지는 바람직하게, 선택된 가시적 이미지들이다.
- [0054] 이러한 방식으로, 세그먼테이션을 추가로 강화시키고 3D 복원을 추가로 리파이닝하기 위해 하나 이상의 반복이 구현될 수 있다.
- [0055] 일 실시형태에 따르면, 반복들은 미리 결정된 정밀도 임계치가 도달될 때 중지된다. 이러한 미리 결정된 정

밀도 임계치는 추가로 강화된 3D 피쳐들과 강화된 2D 피쳐들 간의 적어도 매칭에 대한 임계치일 수 있다.

- [0056] 예를 들어, 상기 미리 결정된 정밀도 임계치는 세그멘테이션 레벨 중 적어도 하나에 공동으로 적용되고, 세그먼트 레벨은 장면의 3D 복원을 세그먼트들로 파티셔닝하는 정도, 상기 추가로 리파이닝된 3D 복원에서 측정된 이웃하는 유사한 3D 엘리먼트들 사이의 라벨들의 일관도, 및 상기 세트의 적어도 하나의 이미지 (예를 들어, 가시적 이미지들) 와 상기 추가의 리파이닝된 3D 복원 간의 정렬에 의해 제공된다.
- [0057] 추가로 강화된 3D 피쳐들과 강화된 2D 피쳐들 사이의 매칭은 세트의 이미지들과 추가로 리파이닝된 3D 복원 간의 글로벌 대응으로부터 특히 주목될 수 있으며, 글로벌 대응은 (하기에서 더 상세히 설명되는 것과 같은) 특히 대응하는 에너지 함수의 값으로부터 확립될 수 있다.
- [0058] 일 실시형태에 따르면, 리파이닝된 또는 추가로 리파이닝된 3D 복원 및 강화된 또는 추가로 강화된 세그멘테이션은 동일한 반복에 대하여 적어도 하나의 에너지 함수에서 고려된다. 그러나, 최적화 문제의 해결책은 2 단계로 구현될 수 있다: 제 1 단계에서, 3D 복원은 고정되어 세그멘테이션을 강화시키고, 제 2 단계에서, 강화된 세그멘테이션이 고정되어 3D 복원을 리파이닝한다.
- [0059] 보다 구체적으로, 초기화 시에, 리파이닝된 3D 복원은 초기 3D 복원 및 강화된 3D 피쳐들로부터 결정된다. 리파이닝된 3D 복원은 강화된 세그멘테이션을 결정하기 위해 고정된다. 강화된 세그멘테이션은 일 실시형태에 따라, 추가로 리파이닝된 3D 복원을 결정하기 위해 고정될 수 있다.
- [0060] 후속 반복들 동안, 선행 반복에서 획득된 3D 복원은 추가로 강화된 세그멘테이션을 결정하기 위해 고정된다. 그 후에, 추가로 강화된 세그멘테이션은 추가로 리파이닝된 3D 복원을 결정하기 위해 고정된다.
- [0061] 각 반복에 대하여, 강화된 세그멘테이션 및 리파이닝된 3D 복원이 공동으로 결정된다.
- [0062] 바람직하게는 미리 결정된 정밀도 임계치와 함께 또는 대안적으로, 반복들은 미리 결정된 반복 횟수에 도달할 때 중지된다.
- [0063] 일 실시형태에 따르면, 강화된 세그멘테이션을 결정하는 단계는 세그멘테이션 제약들에 의존한다. 이러한 세그멘테이션 제약들은 또한 "프라이어 (prior) 들" 로 불린다.
- [0064] 특히, 세그멘테이션 제약들은 평면 형상, 볼록 형상, 직육면체 형상, 원통 형상 등과 같은 적어도 하나의 세그먼트 형상에 관련된다.
- [0065] 일 실시형태에 따르면, 상기 방법은 상기 초기 3D 복원 및 상기 적어도 하나의 이미지의 세트를 적어도 하나의 입력으로서 수신하는 단계, 적어도 하나의 프로세서로 상기 강화된 3D 피쳐, 강화된 세그멘테이션 및 리파이닝된 3D 복원을 결정하는 단계, 및 상기 리파이닝된 3D 복원을 사용자에게 디스플레이하고 상기 강화된 세그멘테이션에 의해 상기 리파이닝된 3D 복원을 프로세싱하기 위한 적어도 하나의 출력으로부터 상기 강화된 세그멘테이션 및 상기 리파이닝된 3D 복원을 출력하는 단계를 포함한다.
- [0066] 본 개시의 다른 양태는 통신 네트워크로부터 다운로드 가능하고 및/또는 컴퓨터에 의해 판독가능한 매체 상에 기록되고 및/또는 프로세서에 의해 실행가능한 컴퓨터 프로그램 제품에 관한 것이며, 상기 컴퓨터 프로그램 제품은 컴퓨터 또는 프로세서에 의해 실행될 때, 임의의 실시형태에서 공동 세그멘테이션 및 3D 복원을 위한 전술한 방법을 수행하도록 적응된 소프트웨어 코드를 포함한다.
- [0067] 본 개시의 다른 양태는 컴퓨터 프로그램 제품을 저장하는 비일시적 컴퓨터 판독가능 캐리어 매체에 관한 것이며, 상기 컴퓨터 프로그램 제품은 컴퓨터 또는 프로세서에 의해 실행될 때, 컴퓨터 또는 프로세서로 하여금 임의의 상이한 실시형태들에서 공동 세그멘테이션 및 3D 복원을 위한 전술한 방법을 실행하게 한다.
- [0068] 본 개시는 또한 장면의 적어도 하나의 이미지의 세트로부터, 장면의 공동 세그멘테이션 및 3D 복원을 위한 디바이스에 관한 것으로, 장면의 세그멘테이션은 장면의 3D 복원을 세그먼트들로 파티셔닝하는 것에 대응하고, 상기 디바이스는:
- [0069] - 장면의 초기 3D 복원을 획득하는 수단;
- [0070] - 상기 초기 3D 복원과 연관된 초기 3D 피쳐들을 획득하는 수단;
- [0071] - 상기 초기 3D 복원의 초기 세그멘테이션을 획득하는 수단;
- [0072] - 상기 초기 3D 피쳐들로부터 및 상기 세트의 적어도 하나의 이미지에서 결정된 초기 2D 피쳐들로부터, 상기 장면의 상기 초기 3D 복원과 연관된 상기 초기 3D 피쳐들에 대응하는 것으로서, 강화된 3D 피쳐들을 결정하는 수

단으로서, 상기 강화된 3D 피쳐들은 적어도 부분적으로 상기 초기 세그먼테이션에 대응하는, 상기 강화된 3D 피쳐들을 결정하는 수단; 및

[0073] - 상기 초기 세그먼테이션 및 상기 강화된 3D 피쳐들 양자로부터, 강화된 세그먼테이션 및 리파이닝된 3D 복원 양자를 결정하는 수단을 포함한다.

[0074] 상기 개시는 추가로, 장면의 적어도 하나의 이미지의 세트로부터, 장면의 공동 세그먼테이션 및 3D 복원을 위한 디바이스에 관한 것으로, 장면의 세그먼테이션은 장면의 3D 복원을 세그먼트들로 파티셔닝하는 것에 대응하고, 상기 디바이스는:

[0075] - 장면의 초기 3D 복원을 획득하고;

[0076] - 상기 초기 3D 복원과 연관된 초기 3D 피쳐들을 획득하고;

[0077] - 상기 초기 3D 복원의 초기 세그먼테이션을 획득하고;

[0078] - 상기 초기 3D 피쳐들로부터 및 상기 세트의 적어도 하나의 이미지에서 결정된 초기 2D 피쳐들로부터, 상기 장면의 상기 초기 3D 복원과 연관된 상기 초기 3D 피쳐들에 대응하는 것으로서, 강화된 3D 피쳐들을 결정하는 것으로서, 상기 강화된 3D 피쳐들은 적어도 부분적으로 상기 초기 세그먼테이션에 대응하는, 상기 강화된 3D 피쳐들을 결정하며; 그리고

[0079] - 상기 초기 세그먼테이션 및 상기 강화된 3D 피쳐들 양자로부터, 강화된 세그먼테이션 및 리파이닝된 3D 복원 양자를 결정하도록 적응되고 구성된 적어도 하나의 프로세서를 포함한다.

[0080] 이러한 디바이스는 특히 본 개시에 따른 장면의 공동 세그먼테이션 및 3D 복원을 위한 방법을 구현하기 위해 적응된다. 이는 결합되거나 개별적으로 취해질 수 있는 본 개시의 임의의 실시형태에 따른 방법에 관한 상이한 특징들을 포함할 수 있다. 즉, 이러한 디바이스는 본 개시에 따른 공동 세그먼테이션 및 3D 복원을 위한 방법의 실행 모드들 중 임의의 것을 실행하도록 적응된다.

[0081] 따라서, 이 디바이스의 특징들 및 이점들은 임의의 상이한 실시형태들에서 장면의 공동 세그먼테이션 및 3D 복원을 위한 개시된 방법과 동일하다.

[0082] 본 개시의 다른 양태는 전술한 디바이스와 같은 장면의 공동 세그먼테이션 및 3D 복원을 위한 디바이스를 포함하는 장치에 관한 것이다.

[0083] 따라서, 그러한 장치의 특징들 및 이점들은 임의의 상이한 실시형태들에서 장면의 공동 세그먼테이션 및 3D 복원을 위한 개시된 방법과 동일하다.

[0084] 특히, 그러한 장치는 바람직하게는 모바일 전화, 태블릿 및 헤드 장착 디스플레이 중에서 선택되는, 모바일 장치 일 수 있다.

[0085] 상이한 실시형태들에 따르면, 이러한 장치는 바람직하게는 로봇, 자율 주행 장치 및 스마트 홈 장치 중에서 선택되는, 자율 장치일 수 있다.

[0086] 따라서, 본 개시는 특히, 네비게이션, 자율 로봇, 가상 현실, 증강 및/또는 복합 현실, 스마트 홈 장치 등과 같은 분야의 애플리케이션들에 적합하다.

[0087] 따라서, 본 개시는 또한 그러한 분야들로의 본 개시의 적용에 관한 것이다.

[0088] 개시된 실시형태들과 범위에 있어서 동등한 특정 양태들이 하기에서 설명된다. 이들 양태들은 단지 본 개시가 취할 수도 있는 특정 형태들의 간단한 개요를 독자에게 제공하도록 제시될 뿐이며 이들 양태들이 본 개시의 범위를 한정하도록 의도되지 않음을 이해해야 한다. 실제로, 본 개시는 하기에서 설명되지 않을 수도 있는 다양한 양태들을 포괄할 수도 있다.

도면의 간단한 설명

[0089] 본 개시는, 첨부 도면들을 참조하여, 결코 한정적이지 않은 방식으로 다음의 실시형태 및 실행 예들에 의해 더 잘 이해되고 예시될 것이다.

도 1은 본 개시의 일 실시형태에 따른 공동 세그먼테이션 및 3D 복원을 위한 방법의 주요 단계들을 예시하는 흐름도이다.

도 2 는 2D 및 3D 피쳐들이 피쳐 라인들인, 본 개시의 실시형태를 도시한다.

도 3 은 장면의 초기 3D 복원의 일 예를 도시한다.

도 4 는 도 3 의 장면의 초기 3D 복원과 연관된 초기 3D 피쳐들의 일 예를 도시한다.

도 5 는 도 3 의 장면의 초기 3D 복원과 연관된 초기 세그먼테이션의 일 예를 도시한다.

도 6a 및 도 6b 는 도 3 에 표시된 장면의 멀티-뷰 이미지들의 예들이다.

도 7a 및 도 7b 는 도 6a 및 도 6b 의 멀티-뷰 이미지들에서 결정된 초기 2D 피쳐들의 예들을 도시한다.

도 8 은 강화된 3D 피쳐들의 일 예를 도시한다.

도 9 는 리파이닝된 3D 복원 및 강화된 세그먼테이션의 일 예를 도시한다.

도 10 은 본 개시의 일 실시형태에 따른 공동 세그먼테이션 및 3D 복원을 위한 방법을 구현하는 디바이스의 블록도이다.

도 1, 도 2 및 도 10 에 있어서, 표현된 블록들은, 반드시 물리적으로 분리된 엔터티들에 대응할 필요는 없는 순수하게 기능적인 엔터티들이다. 즉, 이들은 소프트웨어, 하드웨어의 형태로 전개될 수 있거나, 또는 하나 이상의 프로세서들을 포함하는 하나 또는 수개의 집적 회로들에서 구현될 수 있다.

발명을 실시하기 위한 구체적인 내용

- [0090] 본 개시의 도면 및 설명은, 명료함을 위해, 모바일 장치 (예를 들면, 모바일 전화, 태블릿, 헤드 장착 디스플레이, 등) 또는 자율 장치 (예를 들면, 로봇, 자율 주행 장치, 스마트 홈 장치, 등) 와 같은 전형적인 동작중인 장치에서 발견되는 많은 다른 엘리먼트들 제거하면서, 본 개시의 명확한 이해와 관련되는 엘리먼트들을 예시하기 위해 간략화되었음이 이해될 것이다.
- [0091] 본 개시의 일반적인 원리는 장면의 초기 3D 복원으로부터 및 장면의 이미지 세트의 적어도 하나의 이미지에서 결정된 초기 2D 피쳐들로부터, 장면의 초기 3D 복원과 연관된 초기 3D 피쳐들에 대응하는 것으로, 장면의 리파이닝된 3D 복원 및 강화된 세그먼테이션을 결정하는 것에 의존한다.
- [0092] 그 장면은 주목할 만하게, 오브젝트일 수 있다. 장면은 하나 이상의 오브젝트들로 구성될 수 있다.
- [0093] 특히, 장면의 리파이닝된 3D 복원은 초기 2D 및 3D 피쳐들로부터 획득된 강화된 3D 피쳐들로 인해 결정되며, 강화된 세그먼테이션은 리파이닝된 3D 복원에서 결정된다.
- [0094] 본 개시의 일 실시형태에 따른 공동 세그먼테이션 및 3D 복원을 위한 방법의 주요 단계들은 도 1 에 예시된다.
- [0095] 예를 들어, 입력은 장면의 이미지들의 시퀀스 및 그들의 심도 데이터와 같은, RGB-D 데이터이다. 변형하여, 입력은 장면의 초기 3D 복원 (초기 3D 모델이라고도 함) 및 멀티-뷰 이미지들이다.
- [0096] 블록 (11) 에서, 장면의 초기 3D 복원이 획득된다. 이러한 초기 3D 복원은 심도 데이터로부터 또는 장면의 이미지들의 세트로부터 구성될 수 있거나, 또는 업스트림으로 결정되고 동작중인 장치/디바이스에서 직접 수신될 수 있다. 초기 3D 복원은 임의의 공지된 기술에 의해 구성될 수 있다는 것을 알아야 한다. 예를 들어, KinectFusion® 과 같은 기성품의 심도 결합 툴 (off-the-shelf depth fusion tool) 또는 Intel RealSense® 와 같은 심도 센서에 의해 결정될 수 있다.
- [0097] 블록 (12) 에서, 초기 3D 복원과 연관된 초기 3D 피쳐들이 획득된다. 이러한 초기 3D 피쳐들은 초기 3D 복원을 분석함으로써 획득될 수 있거나, 또는 업스트림으로 결정되어 동작중인 장치/디바이스에서 직접 수신될 수 있다.
- [0098] 블록 (13) 에서, 초기 3D 복원의 초기 세그먼테이션이 획득된다. 이러한 초기 세그먼테이션은 장면의 코오스 (coarse) 세그먼테이션일 수 있다. 초기 세그먼테이션은 임의의 공지된 기술에 의해 구성될 수 있다는 것을 알아야 한다. 예를 들어, 랜덤 라벨링 또는 평면 영역 증가에 의해 결정될 수 있다. 초기 세그먼테이션은 또한 본 명세서의 후반부에 설명되는 바와 같은 세그먼테이션 프라이어들에 의해 제한될 수 있다. 특히, 세그먼테이션이 예를 들어 X. Chen 등에 의해 “A Benchmark for 3D Mesh Segmentation” (ACM Transaction on Graphics, 2009) 에 개시된 기술과 같은 진보된 기술에 의해 결정되는 경우, 세그먼테이션 프라

이어들의 사용은 필요하지 않다. 그러나, 세그멘테이션 프라이어들은 타겟화된 애플리케이션에 의존하여, 진보된 세그멘테이션 기술과 함께 사용될 수 있다.

- [0099] 블록 (14) 에서, 강화된 3D 피쳐들은 세트의 적어도 하나의 이미지에서 결정된 초기 3D 피쳐들 및 초기 2D 피쳐들로부터, 장면의 초기 3D 복원과 연관된 초기 3D 피쳐들에 대응하는 것으로 결정된다. 초기 2D 피쳐들은 이미지들의 세트로부터 결정되거나, 또는 동작중인 장치/디바이스에서 직접 수신될 수 있다.
- [0100] 블록 (15) 에서, 장면의 강화된 세그멘테이션 및 리파이닝된 3D 복원은 초기 세그멘테이션 및 강화된 3D 피쳐들로부터 결정된다. 바람직하게, 강화된 세그멘테이션은 세그멘테이션 프라이어들에 의해 제한된다. 초기 세그멘테이션이 통상적으로 3 차원 모델을 평면 영역들로 러프하게 세그먼트화하도록 설계되기 때문에, 강화된 세그멘테이션은 세그멘테이션 프라이어들에 의해 제한되지 않는다면, 정확한 경계들을 갖는 평면 영역을 전달할 수 있는 것을 알아야 한다. 따라서, 장면의 완전한 오브젝트는 평면 컴포넌트들의 시리즈로 세그먼트화된다. 진보된 애플리케이션들의 경우, 세그멘테이션 프라이어들 (예컨대, 블록한 형상) 이 완전한 오브젝트들이 세그먼트화되게 하기 위해 채용된다.
- [0101] 장면의 세그멘테이션 및 3D 복원을 추가로 개선하기 위해, 블록들 (14 및 15) 은 중지 조건이 충족될 때까지 반복적으로 구현될 수 있다. 보다 구체적으로, 매 반복마다, 추가로 강화된 3D 피쳐들은 - 그 세트의 이미지들에서 결정된 강화된 2D 피쳐들 및 강화된 3D 피쳐들로부터, 장면의 리파이닝된 3D 복원과 연관된 강화된 3D 피쳐들에 대응하는 것으로 - 결정될 수 있으며, 추가로 강화된 세그멘테이션 및 추가로 리파이닝된 3D 복원은 그 후에 - 강화된 세그멘테이션 및 추가로 강화된 3D 피쳐들로부터 결정될 수 있다.
- [0102] 지금부터 도 2 를 참조하여, 2D 및 3D 피쳐들이 피쳐 라인들과 같은 지오메트릭 피쳐들인, 본 개시의 실시형태를 도시한다. RGB-D 데이터들, 이미지 데이터 (이미지들의 세트라고도 함) 및 대응하는 심도 데이터를 포함하는, 입력으로 간주한다.
- [0103] 이 실시형태에 따르면, 메인 블록들은 장면의 초기 3D 복원과 연관된 3D 피쳐들과 이미지 데이터로부터 도출된 지오메트릭 큐들 사이의 대응을 확립하고, 컴포넌트 라벨들을 공동으로 최적화하고 3D 오브젝트(들)에 대한 지오메트리를 리파이닝하도록 설계된다. 세그멘테이션은 또한, 형상 제약들, 즉 세그멘테이션 프라이어들을 고려하여 강화된다.
- [0104] 예를 들어, 이미지 데이터 (21), 심도 데이터 (22) 및 카메라 포즈들 (23) 은 (예를 들어, 심도 센서들의 소프트웨어 개발 키트를 통해) Intel RealSense® 같은 심도 센서들에 의해 획득된다. 이미지 및 심도 데이터는 잘 정렬되어 있고, 카메라 포즈들은 큰 오차 없이 계산된다고 가정한다. 이미지와 심도 데이터를 정렬하거나 필요에 따라 카메라 포즈들을 프로세싱하기 위해, 사전 프로세싱이 구현될 수 있다.
- [0105] 블록 (221) 에서, 입력 심도 데이터는 "클린" 데이터, 즉 장면의 3D 복원에 적합한 데이터를 생성하기 위해 사전 프로세싱된다. 예를 들어, 사전 프로세싱 동작은: 아웃라이어 제거, 노이즈 제거, 샘플링, 심도 인페인팅, 오버 세그멘테이션 등의 적어도 하나를 포함한다.
- [0106] 블록 (222) 에서, 프로세싱된 심도 데이터는 병합되어 (심도 결합), 장면의 초기 3D 복원을 생성한다. 예를 들어, KinectFusion® 과 같은 기성품 툴은 장면의 초기 3D 복원을 생성하는데 사용된다. 출력 3D 복원은 포인트들의 클라우드, 메쉬 모델, 체적 모델 등으로 표현될 수 있다.
- [0107] 블록 (223) 에서, 초기 3D 복원과 연관된 초기 3D 피쳐들이 획득된다. 예를 들어, 3D 피쳐들은 곡률, 볼록면/오목면, 또는 로컬 피쳐 디스크립터들과 같은 지오메트리 특징들을 사용하여 장면의 초기 3D 복원에서 추출된 3D 피쳐 라인들이다. 추출된 초기 3D 피쳐 라인들은 장면에서 오브젝트(들)의 형상을 도시한다.
- [0108] 블록 (224) 에서, 초기 세그멘테이션은 하나의 컴포넌트로 세그먼트화될 각각의 3D 엘리먼트를 라벨링하기 위해, 장면의 초기 3D 복원에서 정의된다. 이미 언급된 것과 같이, "3D 엘리먼트" 는 예를 들어, 포인트들의 클라우드 중의 포인트, 다각형 메쉬 모델의 다각형, 체적 모델의 복셀 (voxel) 등이고, "컴포넌트" 는 평면 영역과 같은, 동일한 라벨을 갖는 3D 엘리먼트들의 그룹이다. 초기 세그멘테이션은 상이한 컴포넌트들 중에서 세그먼트 경계들을 도시한다. 또한 세그멘테이션 프라이어들에 의해 제한될 수 있다.
- [0109] 블록 (211) 에서, 블록 (223) 에서 장면의 초기 3D 복원으로부터 추출된 초기 3D 피쳐 라인들에 기초하여, 가시적 이미지들이 이미지 데이터 (21) 중에서 선택된다. 가시적 이미지들을 선택하기 위해, 하나의 해결책은 초기 3D 복원의 투영을 사용하여, 이미지들의 세트의 각 이미지에 3D 피쳐 라인들을 투영하고, 이미지 상의 투영의 가시적 픽셀들의 수를 카운트하여 이미지가 가시적인지의 여부를 결정하는 것 (3D-2D 매칭) 이다. 따

라서, 각각의 3D 피쳐 라인에 대해, 가시적 이미지의 시리즈가 발견될 수 있다.

- [0110] 블록 (212) 에서, 초기 3D 피쳐 라인들과 매칭하는 초기 2D 피쳐 라인들이 선택된 가시적 이미지들에서 추출된다. 가시적 이미지들에서 3D 피쳐 라인들과 2D 피쳐 라인들을 매칭하기 위해, 2D 피쳐 라인과 3D 피쳐 라인의 대응하는 투영된 라인 사이의 배향 및 거리를 고려할 수 있는 측정이 정의된다.
- [0111] 일단 초기 2D 피쳐 라인들이 추출되면, 블록 (213) 에서 상이한 이미지들의 2D 피쳐 라인들 중에서의 2D 매칭이 구축될 수 있다. 예를 들어, 2D 매칭은 선택된 가시적 이미지들에 걸쳐 2D 피쳐 라인에 대해 정의된다.
- [0112] 카메라 포즈들 (23) 이 편차들을 가질 수 있다는 사실 때문에, 가시적 이미지들에 걸친 2D 피쳐 라인들 중에서의 2D 매칭은 잡음이 있는 카메라 포즈들에 대응하여, 부정확한 매칭을 제거하기 위해 블록 (214) 에서 필터링될 수 있다. 예를 들어, 한 쌍의 이미지들 상의 한 쌍의 2D 매칭된 라인들을 고려한다면, 각 2D 라인은 예를 들어 "Incremental Line-based 3D Reconstruction using Geometric Constraints" (M. Hofer 등, British Machine Vision Conference, 2013) 에서 정의된 것과 같은, 에피폴라 매칭 방법을 사용하여 3D 라인을 복원하는데 사용될 수 있다. 2 개의 복원된 3D 라인들의 유사도를 비교함으로써, 이 매칭 쌍의 신뢰도를 추정할 수 있다. 예를 들어, 3D 라인들의 유사도는 길이, 배향, 및/또는 거리를 사용하여 평가될 수 있다. 유사도가 높으면, 대응하는 2D 라인들의 매칭이 신뢰가능하며, 이는 이 이미지 쌍들 간의 카메라 포즈들의 추정이 신뢰할 수 있는 것을 의미한다. 유사도가 낮으면, 이는 카메라 포즈가 큰 오차를 가지며, 이 매칭은 제거되어야 함을 의미한다.
- [0113] 카메라 포즈들이 필터링된 후에, 신뢰성 있는 지오메트릭 큐들이 블록 (215) 에서 가시적 이미지들에 걸친 나머지 2D 매칭으로부터 생성된다. 다시 말해서, 필터링 (214) 후에 남아있는 2D 피쳐 라인들은, 예를 들어 멀티-뷰 스테레오 방법들을 사용함으로써 지오메트릭 큐들로 불리는 3D 피쳐 라인들을 구성하는데 사용된다.
- [0114] 이러한 지오메트릭 큐들은 공동 최적화 블록 (25) 에서 3D 지오메트리를 리파이닝하는데 사용되는 강화된 3D 피쳐들을 정의하는 것을 목표로 하여, 초기 3D 피쳐들에 대한 제약들을 제공할 수 있다.
- [0115] 마지막으로, 블록 (25) 에서, 초기 세그멘테이션 (224) 으로부터의 각 3D 엘리먼트에 대한 컴포넌트 라벨 및 3D 엘리먼트의 지오메트리 양자가 공동으로 최적화되어, 장면의 강화된 세그멘테이션 및 리파이닝된 3D 복원을 획득한다. 특히, 이러한 최적화는 3D 세그멘테이션 프라이어더로도 불리는, 세그멘테이션 제약들 (24) 에 의존한다. 예를 들어, 고전적인 세그멘테이션 프라이어는 평탄도, 연결성, 볼록도/오목도 등을 포함하지만 이에 한정되지 않는다. 세그멘테이션 프라이어들은 공동 최적화를 위해 개별적 또는 조합적 방식으로 셋업될 수 있다. 이러한 세그멘테이션 프라이어들은 디폴트 값으로 셋업되고, 애플리케이션의 타입 또는 사용자에게 의해 선택되고, 명시적으로 또는 암시적으로 사용되는 등일 수 있다.
- [0116] 본 개시의 일 실시형태에 따르면, 미리 구성될 수 있는 세그멘테이션 프라이어들을 불러오기 위해 인터페이스가 제안된다. 예를 들어, 사용자는 세그멘테이션의 상이한 레벨에 대응하는 스크롤 바를 조정할 수 있다. 후자는 3D 장면을 세그먼트들로 파티셔닝하는 정도 (특히 파티셔닝에서의 세그먼트들의 수) 에 의해 제공될 수 있다. 라지-스케일 장면이 고려될 때, 평면 영역과 같은 세그멘테이션 프라이어들에 대응하는 저 레벨의 세그멘테이션이 선택된다. 테이블의 표면 상의 클로즈업과 같은 스몰-스케일 장면이 고려될 때, 직육면체형, 원통형과 같은 세그멘테이션 프라이어들에 대응하는 고 레벨의 세그멘테이션이 선택된다.
- [0117] 공동 최적화를 위해 몇몇 에너지 함수들이 정의될 수 있다. 강화된 세그멘테이션 및 리파이닝된 3D 표현의 결정은 에너지 함수들 중 적어도 하나를 최소화함으로써 구현될 수 있다.
- [0118] 예를 들어, 2 개의 에너지 함수들이 공동 최적화: 세그멘테이션, 평활도 및 지오메트리 리파인먼트를 위해 정의될 수 있다. 각 에너지 함수의 가중치는 예를 들어, 장면의 3D 복원의 원하는 품질에 의존하여 조정될 수 있다.
- [0119] 제 1 반복에서, 세그멘테이션 에너지 함수는 초기 세그멘테이션 (224), 세그멘테이션 프라이어 (24) 및 초기 3D 피쳐들 (223) 을 고려할 수 있다. 이러한 세그멘테이션 에너지 함수는 예를 들어, "A Benchmark for 3D Mesh Segmentation" (X. Chen, 등 ACM Transaction on Graphics, 2009) 에 개시된 기술에 의해 정의될 수 있다.
- [0120] 평활도 에너지 함수는 장면의 초기 복원에서 측정된 이웃하는 유사한 3D 엘리먼트들 사이의 라벨들의 일관도를 고려할 수 있다. 평활도 에너지 함수는 예를 들어, 각 3D 엘리먼트에 대해, 그 레벨과 이웃하는 3D 엘리먼트들의 레벨들 간의 차이를 측정함으로써, 정의될 수 있다.

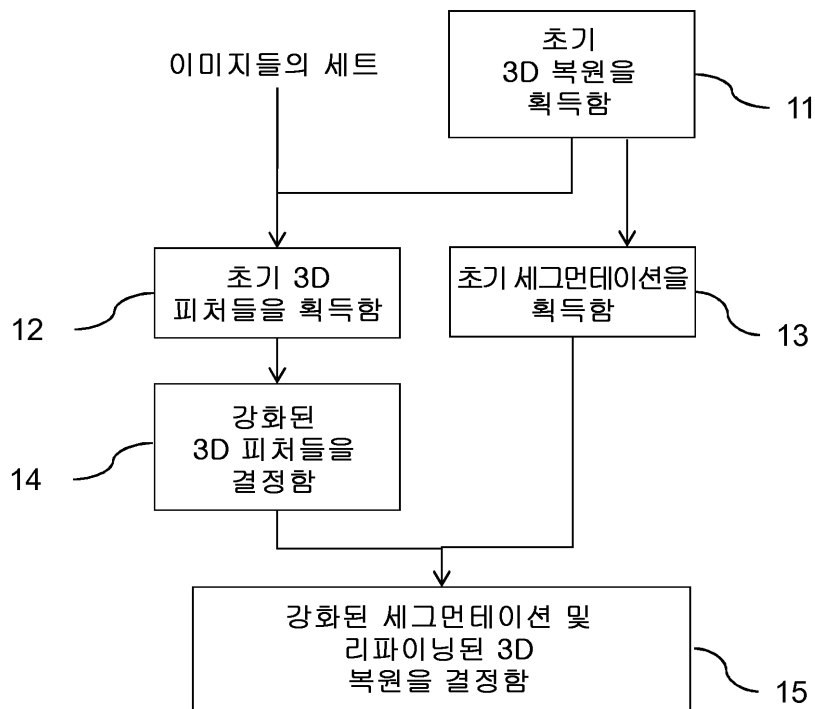
- [0121] 리파인먼트 에너지 함수는 선택된 가시적 이미지들로부터 생성된 지오메트릭 큐들 (215) 과 초기 3D 피쳐 라인들 (223) 간의 정렬을 측정할 수 있다. 리파인먼트 에너지 함수는 예를 들어, 초기 3D 피쳐 라인들 (세그먼테이션 경계들을 포함함) 과 복원된 지오메트릭 큐들 사이의 거리, 배향, 및/또는 길이의 차이를 측정함으로써, 정의될 수 있다.
- [0122] 공동 세그먼테이션 및 리파인먼트는 이들 에너지 함수들을 함께 최소화함으로써 구현될 수 있다. 에너지 함수들의 각각이 다른 에너지 함수들 중 적어도 하나의 변수들에 의해 영향을 받으며, 따라서 이들 에너지 함수들 간의 상호작용이 달성된다. 예를 들어, 초기 세그먼테이션 (224) 이 세그먼테이션 에너지 함수에서 수정되는 경우, 이는 초기 3D 피쳐들 (223) 에 영향을 미치고, 이는 장면의 초기 복원시 라벨들뿐만 아니라 3D 엘리먼트들을 변경하여, 평활도 에너지 함수 및 리파인먼트 에너지 함수 양자에 영향을 준다.
- [0123] 제 4 에너지 함수는 또한, 부정확한 카메라 포즈들의 경우, 이미지와 심도 데이터 양자로부터 정렬 오차를 모델링하는 것으로 정의될 수 있다.
- [0124] 공동 최적화 (25) 의 제 1 반복 후에, 각 3D 엘리먼트에 대한 컴포넌트 라벨 및 피쳐 라인들 주위의 지오메트리가 업데이트될 수 있다. 다시 말해서, 공동 최적화 (25) 의 제 1 반복 후에, 강화된 세그먼테이션 및 리파이닝된 3D 복원은 더욱 강화되고 리파이닝될 수 있다.
- [0125] 예를 들어, 제 2 반복에서, 세그먼테이션 에너지 함수는 장면의 리파이닝된 3D 복원과 연관된, 강화된 세그먼테이션, 세그먼테이션 프라이어 (24) 및 강화된 3D 피쳐들을 고려할 수 있다. 평활도 에너지 함수는 장면의 리파이닝된 3D 복원에서 측정된 이웃하는 유사한 3D 엘리먼트들 사이의 라벨들의 일관도를 고려할 수 있다. 리파인먼트 에너지 함수는 선택된 가시적 이미지들로부터 생성된 지오메트릭 큐들과 강화된 3D 피쳐 라인들 간의 정렬을 측정할 수 있다.
- [0126] 반복들은 미리 결정된 정밀도 임계치가 도달될 때 (예를 들어, 상기 강화된 3D 피쳐들과 상기 강화된 2D 피쳐들 간의 적어도 하나의 매칭에 대한 임계치), 또는 미리 결정된 반복 횟수가 도달될 때, 중지될 수 있다.
- [0127] 개별 세그먼테이션 및 3D 복원과 비교하여, 적어도 하나의 실시형태에 따른 공동 세그먼테이션 및 3D 복원을 위한 방법은 세그먼테이션 및 3D 복원이 서로에게 기여하여 더 양호한 결과들을 달성하게 한다.
- [0128] 도 3 내지 도 9 는 테이블 상의 박스를 포함하는 장면의 예에 대하여, 본 개시의 일 실시형태에 따른 공동 세그먼테이션 및 3D 복원을 위한 알고리즘의 결과를 도시한다.
- [0129] 도 3 은 블록 (222) 에서 예를 들어 KinectFusion® 툴에 의해 획득된, 장면의 초기 3D 복원을 도시한다.
- [0130] 도 4 는 예를 들어 블록 (223) 에서 획득된, 장면의 초기 3D 복원과 연관된 초기 3D 피쳐들을 도시한다.
- [0131] 도 5 는 예를 들어 블록 (224) 에서 획득된, 초기 세그먼테이션을 도시한다. 예를 들어, 장면의 초기 3D 복원이 메시 표면으로 표현되는 경우, 3D 엘리먼트들은 메시의 삼각형면들일 수 있으며, 컴포넌트들은 L1, L2, L3 및 L4 로 라벨링된 세그먼트화된 영역들이다.
- [0132] 도 6a 및 도 6b 는 블록 (211) 에서 입력 이미지들의 세트로부터 선택된, 테이블 상의 박스의 멀티-뷰 이미지들이다.
- [0133] 도 7a 및 도 7b 는 예를 들어, 블록 (212) 에서 획득된, 도 6a 및 도 6b 의 멀티-뷰 이미지들에서 결정된 초기 2D 피쳐들을 도시한다.
- [0134] 도 8 은 지오메트릭 큐들에 의해 정의된 제약들을 초기 3D 피쳐들에 적용함으로써 획득된, 강화된 3D 피쳐들을 도시하며, 여기서 지오메트릭 큐들은 예를 들어, 필터링 (214) 후에 남아있는 2D 피쳐 라인들로부터 블록 (215) 에서 생성된다.
- [0135] 도 9 는 결과적으로, 예를 들어, 공동 최적화 블록 (25) 에서 획득된, 리파이닝된 3D 복원 및 강화된 세그먼테이션을 도시한다.
- [0136] 지금부터 도 10 을 참조하여, 본 개시의 적어도 하나의 실시형태에 따라 장면의 공동 세그먼테이션 및 3D 복원을 위한 방법을 구현하는데 사용될 수 있는 예시적인 디바이스의 구조적인 블록들을 도시한다.
- [0137] 일 실시형태에서, 개시된 방법을 구현하기 위한 디바이스 (100) 는 비휘발성 메모리 (103) (예를 들어, 판독 전용 메모리 (ROM) 또는 하드 디스크), 휘발성 메모리 (101) (예를 들어, 랜덤 액세스 메모리 또는 RAM) 및 프로세서 (102) 를 포함한다. 비휘발성 메모리 (103) 는 비일시적인 컴퓨터 판독가능 캐리어 매체이다. 그

메모리는 다양한 실시형태들에서 전술한 방법의 구현을 가능하게 하기 위해 프로세서 (102) 에 의해 실행되는, 실행가능한 프로그램 코드 명령들을 저장한다.

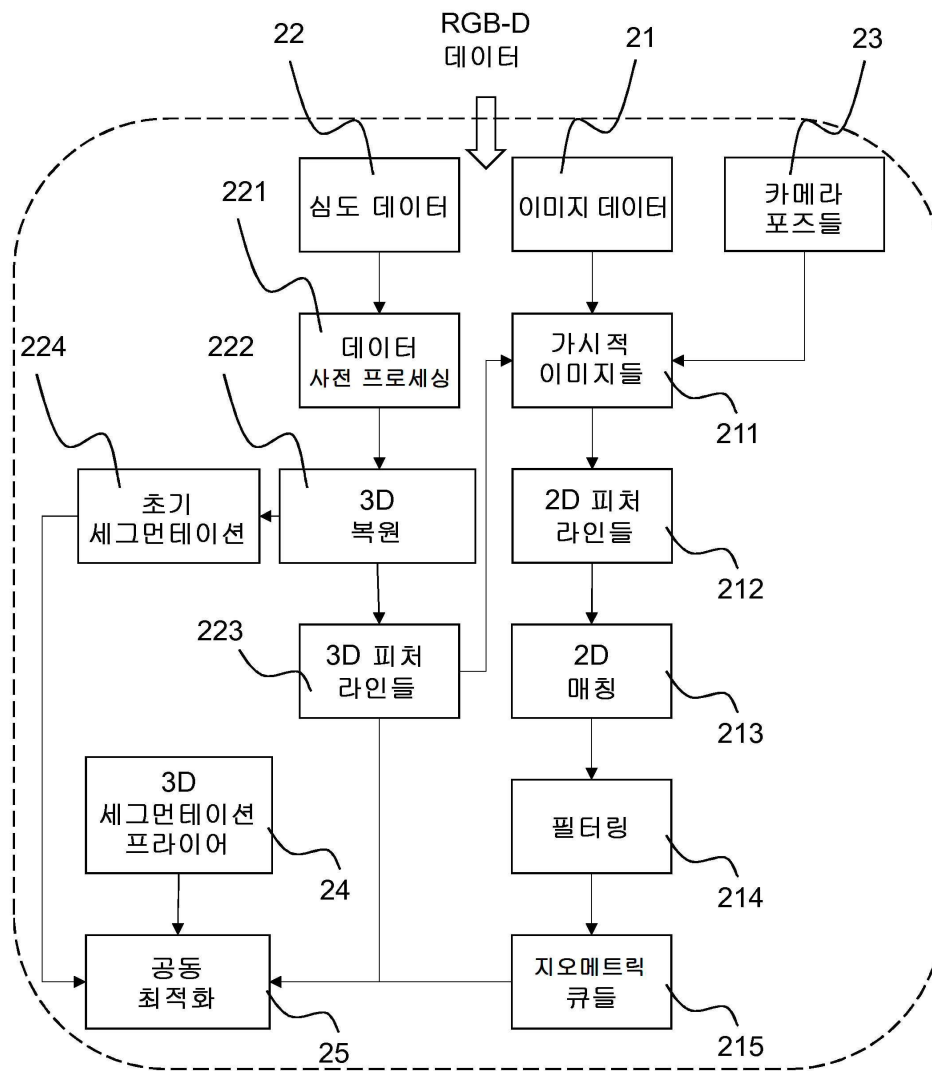
- [0138] 초기화 시에, 전술한 프로그램 코드 명령들은 휘발성 메모리 (103) 로부터 휘발성 메모리 (101) 로 전송되어 프로세서 (102) 에 의해 실행된다. 휘발성 메모리 (101) 는 마찬가지로, 이 실행에 필요한 변수들 및 파라미터들을 저장하기 위한 레지스터들을 포함한다.
- [0139] 본 개시의 적어도 하나의 실시형태에 따른 장면의 공동 세그먼테이션 및 3D 복원을 위한 방법의 단계들은:
- [0140] - PC 타입 장치, DSP (디지털 신호 프로세서) 또는 마이크로제어기와 같은 재프로그램가능한 컴퓨팅 머신에 의해 실행되는 프로그램 코드 명령들의 세트의 실행에 의해. 이 프로그램 코드 명령들은 분리가능하거나 (예를 들어, 플로피 디스크, CD-ROM 또는 DVD-ROM) 또는 분리 불가능한 비일시적 컴퓨터 판독가능 캐리어 매체에 저장될 수 있다; 또는
- [0141] - FPGA (Field Programmable Gate Array), ASIC (Application-Specific Integrated Circuit) 또는 임의의 전용 하드웨어 컴포넌트와 같은 전용 기계 또는 컴포넌트에 의해
- [0142] 동등하게 구현될 수도 있다.
- [0143] 다시 말해서, 본 개시는 컴퓨터 프로그램 명령들의 형태의 순수한 소프트웨어 기반 구현에 한정되지 않고, 하드웨어 형태로 또는 하드웨어 부분과 소프트웨어 부분을 조합한 임의의 형태로 구현될 수도 있다.
- [0144] 적어도 하나의 실시형태에서, 디바이스는 장치 내에 제공된다. 그러한 장치는 모바일 전화, 태블릿, 헤드 장착 디스플레이 등과 같은 모바일 장치, 또는 로봇, 자율 주행 장치 또는 스마트 홈 장치 등과 같은 자율 장치일 수 있다. 이러한 장치는 증강 현실/혼합 현실, 및 자율 로봇/주행의 분야의 적용들을 구현할 수 있다.
- [0145] 설명되지는 않았지만, 그러한 디바이스 또는 장치는 또한 적어도 하나의 카메라, 적어도 하나의 디스플레이 또는 다른 고전적 디바이스들을 또한 포함할 수 있다.

도면

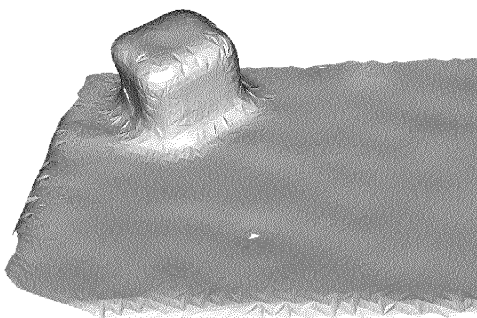
도면1



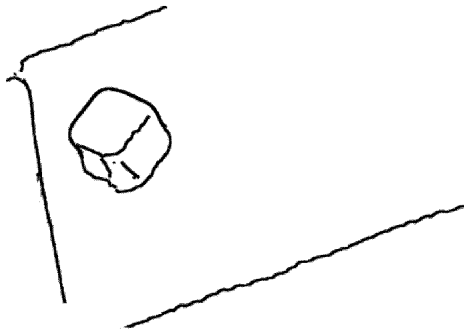
도면2



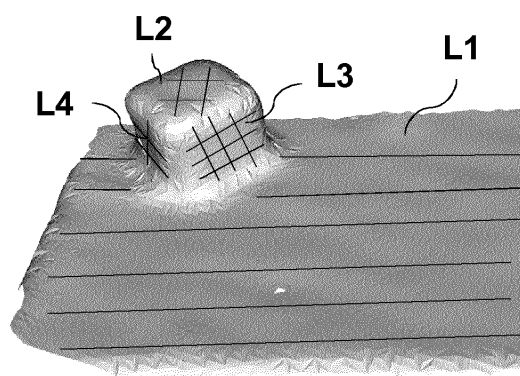
도면3



도면4



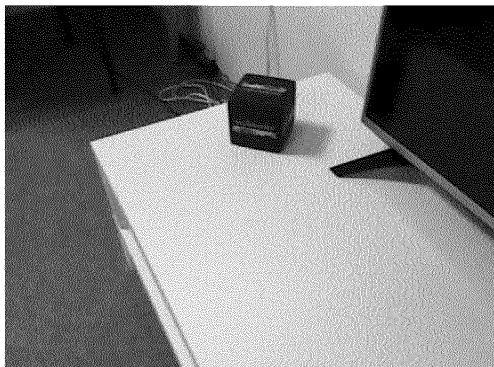
도면5



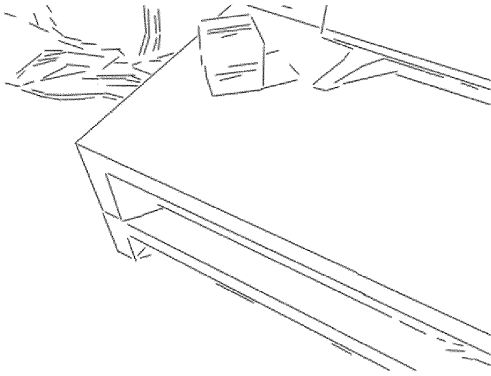
도면6a



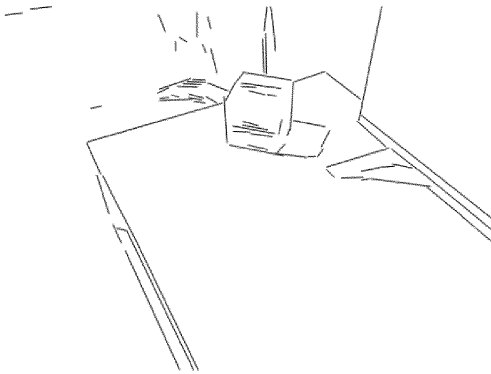
도면6b



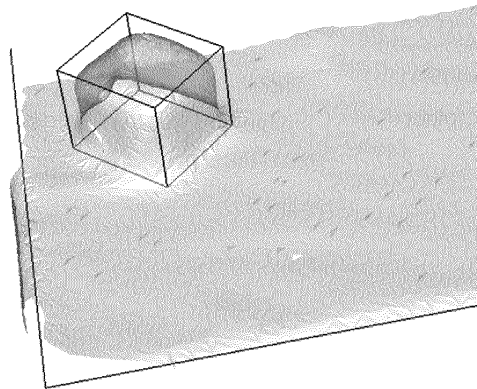
도면7a



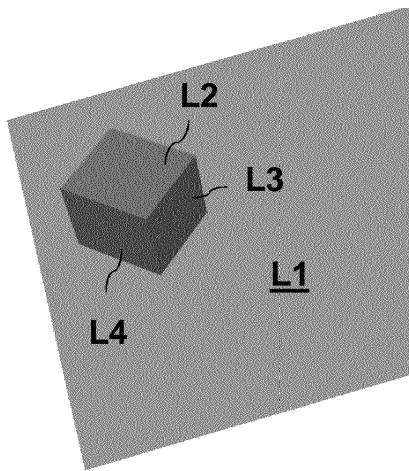
도면7b



도면8



도면9



도면10

