



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2020년10월23일
(11) 등록번호 10-2169243
(24) 등록일자 2020년10월19일

(51) 국제특허분류(Int. Cl.)
G06T 7/10 (2017.01) G06T 15/08 (2011.01)
G06T 17/20 (2006.01) G06T 7/50 (2017.01)
(52) CPC특허분류
G06T 7/10 (2017.01)
G06T 15/08 (2013.01)
(21) 출원번호 10-2018-0170999
(22) 출원일자 2018년12월27일
심사청구일자 2018년12월27일
(65) 공개번호 10-2020-0080970
(43) 공개일자 2020년07월07일
(56) 선행기술조사문헌
JP2018515197 A*
(뒷면에 계속)

(73) 특허권자
포항공과대학교 산학협력단
경상북도 포항시 남구 청암로 77 (지곡동)
(72) 발명자
이승용
경상북도 포항시 남구 청암로 77 포항공과대학교
정보통신연구소 333호
전준호
경상북도 포항시 남구 효자동길5번길 23 (효자동)
(뒷면에 계속)
(74) 대리인
김합곤, 안광석, 박영일

전체 청구항 수 : 총 7 항

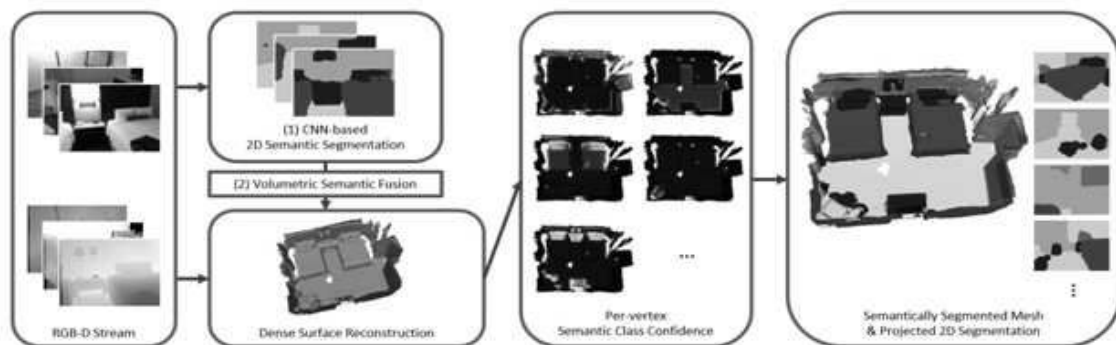
심사관 : 신재철

(54) 발명의 명칭 이차원 의미론적 분할 정보의 점진적인 혼합을 통한 삼차원 복원 모델의 의미론적 분할 방법

(57) 요약

본 발명은 보급용 깊이 영상 카메라에서 온 연속적인 색상 및 깊이 영상 스트림(stream)으로부터 삼차원 복원과 동시에 복원된 모델에 대한 점진적 의미론적 분할을 수행하는 이차원 의미론적 분할 정보의 점진적인 혼합을 통한 삼차원 복원 모델의 의미론적 분할 방법에 관한 것으로, 본 발명에 의한 삼차원 복원 모델의 의미론적 분할 방법은 (a) 다중의 2차원 입력 영상에 대하여, 각 입력 영상의 색상 영상(RGB)과 대응하는 깊이 영상(Depth)을 활용하여 딥러닝 기반 픽셀단위 의미론적 분할을 수행하여 각 픽셀마다 객체 부류에 따른 확률 정보를 획득하는 단계; (b) 각 픽셀마다 획득한 확률 정보를 복셀 그리드에 레이캐스팅으로 업데이트 하는 단계; (c) 마칭큐브 알고리즘에 의해 복셀 그리드에서 메시 모델을 추출하는 단계; 및 (d) 메시 모델에서 각 정점에 대해서 가장 높은 확률을 갖는 부류를 선택함으로써 삼차원 복원 모델의 의미론적 분할을 수행하는 단계;를 포함하여 구성된다.

대표도 - 도1



(52) CPC특허분류

G06T 17/20 (2013.01)

G06T 7/50 (2017.01)

(72) 발명자

정진웅

경상북도 포항시 남구 청암로 77 포항공과대학교
남자기숙사 11동 207호

김준건

경상북도 포항시 북구 양학천로 19-2, 명동빌 304
호 (죽도동)

(56) 선행기술조사문헌

KR101437228 B1*

KR1020140007367 A*

깊이 슈퍼 픽셀을 이용한 실내 장면의 의미론적 분
할 방법, Journal of Korea Multimedia Society
Vol. 19, No. 3, 2016

Incorporating Depth into Semantic Segmentation
via Fusion-based CNN Architecture, ACCV 2016

*는 심사관에 의하여 인용된 문헌

이 발명을 지원한 국가연구개발사업

과제고유번호

1711065143

부처명

과학기술정보통신부

과제관리(전문)기관명

정보통신기술진흥센터

연구사업명

SW컴퓨팅산업원천기술개발

연구과제명

(SW 스타랩) 빅 비주얼 데이터 기반의 고품질 사진 메이크업 SW 개발

기 여 율

40/100

과제수행기관명

포항공과대학교

연구기간

2018.01.01 ~ 2018.12.31

이 발명을 지원한 국가연구개발사업

과제고유번호

1711072842

부처명

과학기술정보통신부

과제관리(전문)기관명

정보통신기술진흥센터

연구사업명

범부처GigaKOREA사업

연구과제명

(4D실감-2세부)초실감 서비스를 위한 동적 객체의 실시간 4D 복원 기술 개발

기 여 율

30/100

과제수행기관명

광주과학기술원

연구기간

2018.01.01 ~ 2018.12.31

이 발명을 지원한 국가연구개발사업

과제고유번호

1711074242

부처명

과학기술정보통신부

과제관리(전문)기관명

한국연구재단

연구사업명

차세대정보·컴퓨팅기술개발

연구과제명

초실감 원격가상 인터랙션을 위한 teleportation 기술 개발

기 여 율

30/100

과제수행기관명

포항공과대학교

연구기간

2018.06.01 ~ 2019.05.31

공지예외적용 : 있음

명세서

청구범위

청구항 1

- (a) 다중의 2차원 입력 영상에 대하여, 각 입력 영상의 색상 영상(RGB)과 대응하는 깊이 영상(Depth)을 활용하여 딥러닝 기반 픽셀단위 의미론적 분할을 수행하여 각 픽셀마다 객체 부류에 따른 확률 정보를 획득하는 단계;
- (b) 각 픽셀마다 획득한 확률 정보를 복셀 그리드에 레이캐스팅으로 업데이트 하는 단계;
- (c) 마칭큐브 알고리즘에 의해 상기 복셀 그리드에서 메시 모델을 추출하는 단계; 및
- (d) 상기 메시 모델에서 각 정점에 대해서 가장 높은 확률을 갖는 부류를 선택함으로써 삼차원 복원 모델의 의미론적 분할을 수행하는 단계;를 포함하고,

상기 (c) 단계의 메시 모델의 각 정점(vertex)의 확률은 양선형 보간(bilinear interpolation)을 통해 결정됨을 특징으로 하는 이차원 의미론적 분할 정보의 점진적인 혼합을 통한 삼차원 복원 모델의 의미론적 분할 방법.

청구항 2

제1항에 있어서, 상기 (b) 단계는

각 픽셀의 객체 부류에 따른 확률 정보에 물체의 카메라로부터의 거리 및 전·배경 경계에 따라 결정된 가중치를 부가하여 해당 픽셀에 대응되는 복셀에 저장된 객체 부류에 따른 확률 정보에 가중 평균하여 업데이트 하는 것을 특징으로 하는 이차원 의미론적 분할 정보의 점진적인 혼합을 통한 삼차원 복원 모델의 의미론적 분할 방법.

청구항 3

삭제

청구항 4

제1항에 있어서, 픽셀의 확률 정보, 가중치 및 정점의 확률은 각각 객체 부류의 개수만큼의 차원을 갖는 벡터임을 특징으로 하는 이차원 의미론적 분할 정보의 점진적인 혼합을 통한 삼차원 복원 모델의 의미론적 분할 방법.

청구항 5

제2항에 있어서,

상기 (b) 단계에서 각 복셀^o의 t번째 프레임까지 통합된 부류 확률은 수학적식

$$C_t(o) = \frac{W_{t-1}(o)C_{t-1}(o) + W_t(p)C_t(p)}{W_{t-1}(o) + W_t(p)}$$

(여기서 $C_{t-1}(o)$ 와 $W_{t-1}(o)$ 는 각각 복셀^o의 t-1번째 프레임까지 통합된 부류 확률과 신뢰도 가중치이고, $C_t(p)$ 와 $W_t(p)$ 는 각각 t번째 프레임에서 픽셀 p의 부류 확률과 신뢰도 가중치이다)

에 의해 산출되고,

신뢰도 가중치 $W_t(p)$ 는 수학적식

$$W_t(p) = W_t^D(p)W_t^B(p)$$

(여기서 $W_t^D(p)$ 는 깊이 기반 정확도 가중치이고, $W_t^B(p)$ 는 전·배경 경계 오정렬 가중치이다)

에 의해 산출되는 것을 특징으로 하는 이차원 의미론적 분할 정보의 점진적인 혼합을 통한 삼차원 복원 모델의 의미론적 분할 방법.

청구항 6

제5항에 있어서, 전·배경 경계 오정렬 가중치 $W_t^B(p)$ 는 수학식

$$W_t^B(p) = \frac{1}{1 + \exp(-\alpha(r(p) - \beta))}$$

$$r(p) = \frac{d(p) - d_{min}}{d_{max} - d_{min}}$$

(여기서 $d(p)$, d_{min} , d_{max} 은 각각 현재 픽셀 p 의 깊이 값, 픽셀 p 위치를 중심으로 한 윈도우 안에서 최소 깊이 값 및 최대 깊이 값이고, α , β 는 양의 상수이다)

에 의해 산출되는 것을 특징으로 하는 이차원 의미론적 분할 정보의 점진적인 혼합을 통한 삼차원 복원 모델의 의미론적 분할 방법.

청구항 7

제1항, 제2항, 제4 내지 제6항 중 어느 한 항의 이차원 의미론적 분할 정보의 점진적인 혼합을 통한 삼차원 복원 모델의 의미론적 분할 방법을 컴퓨터에서 실행시키기 위한 프로그램을 기록한 컴퓨터로 읽을 수 있는 기록 매체.

청구항 8

제1항, 제2항, 제4 내지 제6항 중 어느 한 항의 이차원 의미론적 분할 정보의 점진적인 혼합을 통한 삼차원 복원 모델의 의미론적 분할 방법을 컴퓨터에서 실행시키기 위하여 매체에 저장된 컴퓨터 프로그램.

발명의 설명

기술 분야

[0001] 본 발명은 RGBD 카메라의 색상 및 깊이 영상 스트림으로부터 복원된 삼차원모델의 삼차원 의미론적 분할을 수행할 때 다중의 이차원 영상의 의미론적 분할 정보를 혼합하는 방법에 관한 것이다.

배경 기술

[0002] 삼차원 복원이라 함은 레이저 스캐너(laser scanner), 구조광 기반 깊이 카메라 등 다양한 스캔 장비를 활용하여 관심 객체 혹은 환경에 대한 삼차원 위치 및 색상 정보를 획득함을 의미한다.

[0003] 키넥트(Kinect) 카메라와 같은 보급형 깊이 카메라가 출현되고 다양한 알고리즘들이 개발됨에 따라, 작은 규모의 대상(e.g., 사람)만을 비실시간으로 복원할 수 있었던 기술적 수준에서 비싼 스캔 장비가 없더라도 실시간 복원을 가능케 하는 기술적 성숙도를 보일 수 있게 되었다. 실시간 삼차원 복원 기술로서 대중적인 인기를 불러일으킨 KinectFusion 기술을 시작으로, 복원 공간 크기의 제약을 해결한 Voxel Hashing, BundleFusion 등의 대규모 복원 기술이 잇따라 등장하였다.

[0004] 깊이 카메라를 이용한 삼차원 복원을 하기 위해 수행하는 보편적인 프로세스는 다음과 같다. 우선 카메라의 포즈(pose)(혹은 회전하는 물체의 포즈)를 매 프레임마다 계산한다. 포즈를 구하는 방법으로는 주로 plane-to-point error를 고려한 ICP(Iterative Closest Point)의 이형(variant)이 사용되며, 이 알고리즘의 입력으로는 현재 깊이 영상 정보와 모델의 레이캐스팅(raycasting) 깊이 정보가 이용된다.

[0005] 고가의 스캐너와 달리 보급형 깊이 카메라는 깊이 영상 값에 많은 노이즈가 섞여 있으며 이를 해결하기 위해 TSDF(Truncated Signed Distance Function)의 표현에서의 평균 기법이 사용된다. 깊이 카메라에서 연속적으로 들어오는 영상에서 TSDF 값을 계산하고 이를 미리 구성된 복셀 그리드에 저장하며, 이미 저장되어 있는 TSDF 값이 존재한다면 기존의 값과 가중치 합을 통해 새로운 TSDF 값을 계산한다. 스캔을 완료한 후에 TSDF 값이 저장

된 복셀 그리드에 Marching cubes 알고리즘을 적용하여 최종적인 모델의 메시를 뽑아낸다.

- [0006] 기존에 삼차원 복원 모델에서 객체를 분류하는 몇몇 연구가 진행되었다.
- [0007] Liangliang Nan, Ke Xie, 및 Andrei Sharf의 "A search-classify approach for cluttered indoor scene understanding,"[문헌 1]에서는 포인트 클라우드(point cloud) 형태의 모델에 과분할(over-segmentation)을 수행하여 조각난 패치(patch)를 만들고, 이를 다시 누적하여 구성한 영역에 대한 분류 확률(classification likelihood)을 계산한다. 이 확률을 사용하여 현재 부분 모델을 가장 유사한 깔끔한 모델로 교체를 수행한다. 그러나 포인트 클라우드 상의 특징을 사용하기 때문에 정확성이 떨어지고 상당히 제한된 부류에만 적용될 수 있다.
- [0008] Dai A., Chang A. X., Savva M., Halber M., Funkhouser T., Niessner M의 "Scannet: Richly-annotated 3D reconstructions of indoor scenes,"[문헌 2]는 [문헌 1]과 달리 전체 장면을 복셀화한 데이터를 3D CNN(Convolution Neural Network)의 입력으로 사용하여 출력으로 각 복셀이 레이블링된 결과를 얻게 된다.
- [0009] Tianjia Shao, Weiwei Xu, Kun Zhou, Jingdong Wang, Dongping Li, 및 Baining Guo의 "An interactive approach to semantic modeling of indoor scenes with an RGBD camera,"[문헌 3]은 RGBD 영상에 CRF를 사용한 의미론적 분할을 적용하고 분할된 객체를 random regression forest로 분류한다. 또한 장면 정합(scene registration)에서 SIFT(Scale-Invariant Feature Transform), RANSAC(RANdom SAMple Consensus)과 의미론적 분할 레이블 정보를 활용한다. 그러나 이 기술은 KinectFusion 방식의 대규모 삼차원 복원 알고리즘과 융합되기 어렵다.

선행기술문헌

비특허문헌

- [0010] (비특허문헌 0001) Liangliang Nan, Ke Xie, and Andrei Sharf. "A search-classify approach for cluttered indoor scene understanding," ACM Trans. on Graph., 31(6):137:1-137:10, 2012.
- (비특허문헌 0002) Dai A., Chang A. X., Savva M., Halber M., Funkhouser T., Niessner M. "Scannet: Richly-annotated 3D reconstructions of indoor scenes," IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017.
- (비특허문헌 0003) Tianjia Shao, Weiwei Xu, Kun Zhou, Jingdong Wang, Dongping Li, and Baining Guo. "An interactive approach to semantic modeling of indoor scenes with an RGBD camera," ACM Trans. on Graph., 31(6):136:1-11, 2012.
- (비특허문헌 0004) Seong-Jin Park, Ki-Sang Hong, Seungyong Lee. "RDFNet: RGB-D Multi-Level Residual Feature Fusion for Indoor Semantic Segmentation," The IEEE International Conference on Computer Vision (ICCV), 2017, pp. 4980-4989.
- (비특허문헌 0005) Dai A., Niessner M., Zollhofer M., Izadi S., Theobalt C.: Bundlesfusion: Real-time globally consistent 3D reconstruction using on-the-fly surface reintegration. ACM Transactions on Graphics (TOG) 36, 3 (2017), 24
- (비특허문헌 0006) Lorensen W. E., Cline H. E.: Marching cubes: A high resolution 3D surface construction algorithm. In ACM Transactions on Graphics
- (비특허문헌 0007) (TOG) (1987), vol. 21, ACM, pp. 163-169

발명의 내용

해결하려는 과제

- [0011] 본 발명은 보급용 깊이 영상 카메라에서 온 연속적인 색상 및 깊이 영상 스트림(stream)으로부터 삼차원 복원과 동시에 복원된 모델에 대한 점진적 의미론적 분할을 수행하기 위한 이차원 의미론적 분할 정보의 점진적인 혼합 방법을 제공함을 그 목적으로 한다.

과제의 해결 수단

[0012] 상기의 목적을 달성하기 위하여, 본 발명에 의한 이차원 의미론적 분할 정보의 점진적인 혼합을 통한 삼차원 복원 모델의 의미론적 분할 방법은 (a) 다중의 2차원 입력 영상에 대하여, 각 입력 영상의 색상 영상(RGB)과 대응하는 깊이 영상(Depth)을 활용하여 딥러닝 기반 픽셀단위 의미론적 분할을 수행하여 각 픽셀마다 객체 부류에 따른 확률 정보를 획득하는 단계; (b) 각 픽셀마다 획득한 확률 정보를 복셀 그리드에 레이캐스팅으로 업데이트하는 단계; (c) 마칭큐브 알고리즘에 의해 상기 복셀 그리드에서 메시 모델을 추출하는 단계; 및 (d) 상기 메시 모델에서 각 정점에 대해서 가장 높은 확률을 갖는 부류를 선택함으로써 삼차원 복원 모델의 의미론적 분할을 수행하는 단계;를 포함하여 구성된다.

[0013] 상기의 이차원 의미론적 분할 정보의 점진적인 혼합을 통한 삼차원 복원 모델의 의미론적 분할 방법에 있어서, 상기 (b) 단계는 각 픽셀의 객체 부류에 따른 확률 정보에 물체의 카메라로부터의 거리 및 전·배경 경계에 따라 결정된 가중치를 부가하여 해당 픽셀에 대응되는 복셀에 저장된 객체 부류에 따른 확률 정보에 가중 평균하여 업데이트 하는 것을 특징으로 한다.

[0014] 상기의 이차원 의미론적 분할 정보의 점진적인 혼합을 통한 삼차원 복원 모델의 의미론적 분할 방법에 있어서, 상기 (c) 단계의 메시 모델의 각 정점(vertex)의 20개의 확률은 양선형 보간(bilinear interpolation)을 통해 결정됨을 특징으로 한다.

[0015] 상기의 이차원 의미론적 분할 정보의 점진적인 혼합을 통한 삼차원 복원 모델의 의미론적 분할 방법에 있어서, 픽셀의 확률 정보, 가중치 및 정점의 확률은 각각 객체 부류의 개수만큼의 차원을 갖는 벡터임을 특징으로 한다.

[0016] 상기의 이차원 의미론적 분할 정보의 점진적인 혼합을 통한 삼차원 복원 모델의 의미론적 분할 방법에 있어서, 상기 (b) 단계에서 각 복셀⁰의 t번째 프레임까지 통합된 부류 확률은 수학적

$$C_t(o) = \frac{W_{t-1}(o)C_{t-1}(o) + W_t(p)C_t(p)}{W_{t-1}(o) + W_t(p)}$$

[0017]

[0018] (여기서 $C_{t-1}(o)$ 와 $W_{t-1}(o)$ 는 각각 복셀⁰의 t-1번째 프레임까지 통합된 부류 확률과 신뢰도 가중치이고, $C_t(p)$ 와 $W_t(p)$ 는 각각 t번째 프레임에서 픽셀 p의 부류 확률과 신뢰도 가중치이다)에 의해 산출되고,

[0019] 신뢰도 가중치 $W_t(p)$ 는 수학적

$$W_t(p) = W_t^D(p)W_t^B(p)$$

[0020]

[0021] (여기서 $W_t^D(p)$ 는 깊이 기반 정확도 가중치이고, $W_t^B(p)$ 는 전·배경 경계 오정렬 가중치이다)에 의해 산출되는 것을 특징으로 한다.

[0022] 상기의 이차원 의미론적 분할 정보의 점진적인 혼합을 통한 삼차원 복원 모델의 의미론적 분할 방법에 있어서, 전·배경 경계 오정렬 가중치 $W_t^B(p)$ 는 수학적

$$W_t^B(p) = \frac{1}{1 + \exp(-\alpha(r(p) - \beta))}$$

[0023]

$$r(p) = \frac{d(p) - d_{min}}{d_{max} - d_{min}}$$

[0024]

[0025] (여기서 $d(p)$, d_{min} , d_{max} 은 각각 현재 픽셀 p의 깊이 값, 픽셀 p 위치를 중심으로 한 윈도우 안에서 최소 깊

이 값 및 최대 깊이 값이고, α, β 는 양의 상수이다)에 의해 산출되는 것을 특징으로 한다.

[0026] 상기의 목적을 달성하기 위하여, 본 발명에 의한 컴퓨터로 읽을 수 있는 기록 매체는 상기의 이차원 의미론적 분할 정보의 점진적인 혼합을 통한 삼차원 복원 모델의 의미론적 분할 방법을 컴퓨터에서 실행시키기 위한 프로그램을 기록한다.

[0027] 상기의 목적을 달성하기 위하여, 본 발명에 의한 컴퓨터 프로그램은 상기의 이차원 의미론적 분할 정보의 점진적인 혼합을 통한 삼차원 복원 모델의 의미론적 분할 방법을 컴퓨터에서 실행시키기 위하여 매체에 저장된다.

발명의 효과

[0028] 본 발명에 의하면, RGBD 카메라를 활용한 삼차원 복원에서 점진적으로 기하를 완성해 나가는 것처럼 이차원 영상에서의 의미론적 분할 확률정보를 활용하여 최종적으로 삼차원 모델에 대한 삼차원 의미론적 분할 정보를 얻어내므로, 근본적으로 충분한 계산 능력만 받쳐 준다면 실시간 방법으로 쉽게 이전이 가능할 뿐만 아니라 기존 여타 방법론이 겪고 있는 문제점(메모리 제한으로 인한 대규모 복원 모델에서의 해상도(resolution) 하락 등)를 해결하여 대규모 복원 모델에서도 섬세한 기하표현은 유지함과 동시에 의미론적 분할 결과를 획득할 수 있다.

[0029] 이러한 의미론적 분할 결과는 증강현실 및 가상현실 등에 다양한 사용자와의 상호작용이 필요한 경우에 활용될 수 있다. 간단한 활용 예로 실내 인테리어 구조물 재배치에서 사용자는 의미론적 분할의 결과를 이용하여 자유롭게 원하는 객체(예를 들어 의자, 책상)등을 원 메시에서 분리하여 자유롭게 이동시킬 수 있게 된다.

도면의 간단한 설명

[0030] 도 1은 본 발명에 의한 이차원 의미론적 분할 정보의 점진적인 혼합을 통한 삼차원 복원 모델의 의미론적 분할 방법의 전체적인 흐름을 설명하기 위한 도면이다.

도 2는 의미론적 분할 정보 혼합에 사용하는 가중치 맵을 도시한 것이다.

도 3은 최신 딥 러닝 기반의 의미론적 분할 기술의 거리에 따른 예측 정확도를 도시한 것이다.

도 4는 주요 부류들의 확률 시각화 및 최종 의미론적 분할 결과를 도시한 것이다.

도 5는 실내 환경 복원에서 본 발명에 의한 의미론적 분할 결과를 도시한 것이다.

발명을 실시하기 위한 구체적인 내용

[0031] 이하에서, 첨부된 도면을 참조하면서 본 발명의 바람직한 실시예에 대하여 상세히 설명하기로 한다.

[0032] 본 발명에 의한 이차원 의미론적 분할 정보의 점진적인 혼합을 통한 삼차원 복원 모델의 의미론적 분할 방법의 전체 흐름은 도 1과 같다.

[0033] 우선 색상 및 깊이 입력영상을 활용한 딥 러닝 기반 의미론적 분할을 수행하여 각 픽셀마다 객체 부류에 따른 확률 정보를 획득한다. 만약 20개의 부류를 분류한다고 가정하면, 각 픽셀은 20개의 부류의 발생 확률을 포함하고 있다. 이 정보는 레이캐스팅에 따라 각 픽셀에 대응되는 복셀들에 저장 및 혼합되어진다. 혼합될 때 이차원 의미론적 분할 결과의 가중치를 해당 물체의 카메라로부터의 거리 및 전·배경 경계에 따라 적응적으로 결정한다.

[0034] 모든 영상 스트림(stream) 데이터에 대한 위와 같은 처리가 완료되면 마칭큐브(Marching cubes)(비특허문헌 6 참조) 알고리즘을 활용하여 메시 모델을 추출한다. 이때 마칭큐브 (Marching cubes) 과정에서의 메시의 각 정점(vertex)의 20개의 확률은 양선형 보간(bilinear interpolation)을 통해 결정된다.

[0035] 앞서서 결정된 확률들은 모두 이차원 영상에서 물체의 형태 및 색상 중 일부만을 보고 의미론적 분할을 수행해 얻어낸 결과이다.

[0036] 최종 메시의 각 정점의 부류를 정하기 위해서는 간단히 가장 높은 확률을 갖는 부류를 선택하면 된다.

[0038] 혼합과정에서는 각 복셀의 부류 정보마다 수학식 1과 같은 업데이트가 수행된다.

수학식 1

$$C_t(o) = \frac{W_{t-1}(o)C_{t-1}(o) + W_t(p)C_t(p)}{W_{t-1}(o) + W_t(p)}$$

[0039]

여기서 $C_{t-1}(o)$ 와 $W_{t-1}(o)$ 는 각각 복셀 o 의 $t-1$ 번째 프레임까지 통합된 부류 확률과 신뢰도 가중치를 의미한다. $C_t(p)$ 와 $W_t(p)$ 는 각각 t 번째 프레임에서 픽셀 p 의 부류 확률과 신뢰도 가중치를 나타낸다. 만약 분류할 부류가 20개라면 $C_t(o)$ 는 20차원 벡터가 된다.

[0040]

매 프레임의 픽셀의 부류 확률을 같은 가중치($W_t(p) = 1$)로 처리하게 되면 여전히 부정확한 결과를 얻을 수 있다. 이를 경감하기 위해 본 발명은 수학식 2와 같은 적응적 가중치를 활용한다.

[0041]

수학식 2

$$W_t(p) = W_t^D(p)W_t^B(p)$$

[0042]

여기서 $W_t^D(p)$ 는 깊이 기반 정확도 가중치이고, $W_t^B(p)$ 는 전·배경 경계 오정렬 가중치이다. 관련 가중치 맵은 도 2에서 확인할 수 있다.

[0043]

일반적으로 CNN(convolution neural network)은 고정된 수용영역(receptive field) 크기를 갖고 있고, 의미론적 분할의 정확도는 영상에서의 객체 크기에 따라 다양해진다. 깊이 기반 정확도 가중치 $W_t^D(p)$ 는 이를 반영하기 위한 가중치이다.

[0044]

깊이 영상의 기반 정확도 가중치를 계산하기 위해서 본 발명에서 사용한 딥 러닝 기반 방법론인 RDFNet(비특히 문헌 4 참조)의 이차원 의미론적 분할결과의 성능을 계산해야 한다. 우선 Training set의 깊이 및 색상 영상을 RDFNet의 입력으로 받아 의미론적 분할 결과를 추정한다. 의미론적 분할 영상에는 각 픽셀 마다 추정된 부류 정보(20 개의 부류 중에 하나)가 저장되어 있고, 이를 참 값(Ground truth)과의 비교를 통해 깊이 값에 대한 의미론적 분할 결과의 평균 성능을 계산할 수 있다. 이를 그래프로 그리면 도 3의 파란색 실선과 같다. 이는 이산적인 깊이 값에 대한 그래프며, 따라서 연속된(continuous) 깊이 값에 대한 노이즈 제거된 가중치 그래프를 획득하기 위해 4차 다항식으로 피팅(Fitting)을 수행한다. 그 결과는 도 2에 노란색 실선으로 나타난다. 깊이 기반 정확도 가중치 $W_t^D(p)$ 는 이와 같이 피팅된 다항식 함수를 통해 결정된다.

[0045]

의미론적 분할 결과는 주로 색상 영상에 의존하며 깊이 영상은 보충적으로 사용된다. 그러나 RGBD 카메라의 캘리브레이션(calibration)이 잘 되었을지라도 여전히 색상과 깊이 영상 사이에서의 오정렬이 존재하기 마련이다. 특히, 이런 오정렬은 전경(물체)과 배경(주로 벽과 바닥)사이에서 심화된다. 이러한 오정렬은 결과적으로 경계 주변의 복셀에서의 잘못된 레이블링을 가져오게 된다. 본 발명은 이런 문제를 완화하기 위해 깊이 영상의 에지(edge)를 탐지하고 이를 바탕으로 가중치를 생성한다.

[0046]

에지를 판별하고 가중치를 결정하는 방법은 다음과 같다. 픽셀에서 7x7 윈도우를 씌워 편차가 미리 정해진 상수 값보다 큰 경우 에지라고 탐지하고, 수학식 3을 통해 가중치를 결정한다.

[0047]

수학식 3

$$W_t^B(p) = \frac{1}{1 + \exp(-\alpha(r(p) - \beta))}$$

[0048]

$$r(p) = \frac{d(p) - d_{min}}{d_{max} - d_{min}}$$

[0049]

[0050]

여기서 $d(p)$, d_{min} , d_{max} 은 각각 현재 픽셀 p 의 깊이 값, 픽셀 p 위치를 중심으로 한 윈도우 안에서 최소 깊이 값 및 최대 깊이 값을 의미한다. α , β 는 양의 상수를 의미한다.

[0051]

혼합과정이 끝나면 각 복셀마다 수식식 1을 통해 혼합된 20개의 확률이 존재한다. 현재 장면은 복셀 그리드 형태로 표현되어 있으므로, 이를 보통 그래픽스에서 표현하는 메시 형태로 바꾸기 위해서는 마칭큐브(Marching cubes) 알고리즘이 필요하다. 일반적인 경우에는 복셀에 TSDF값 혹은 색상 값이 포함되어 있는데, 본 발명의 경우 20개의 확률이 존재하므로 마칭큐브(Marching cubes) 과정에서 확률 값을 양선형 보간하여 최종적으로 메시의 각 정점의 20개 확률 값을 결정한다.

[0052]

앞선 과정을 통해 20개 부류 확률을 포함한 메시가 추출되면, 도 5의 왼쪽처럼 주요 부류에 대한 확률맵을 얻을 수 있으며, 최종적인 의미론적 분할 결과는 각 정점에 대해서 가장 높은 확률을 갖는 부류를 선택함으로써 얻을 수 있다(도 5의 오른쪽 참조).

[0053]

도 5는 실내 환경 복원에서 본 발명에 의한 의미론적 분할 결과를 도시한 것이다.

[0055]

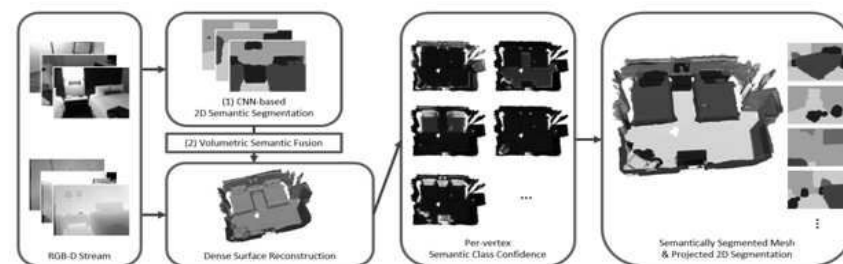
한편, 상술한 본 발명의 실시예는 개인용 컴퓨터를 포함한 범용 컴퓨터에서 사용되는 매체에 기록될 수 있다. 상기 매체는 마그네틱 기록매체(예를 들면, 롬, 플로피 디스크, 하드 디스크 등), 광학적 판독매체(예를 들면, 씨디롬, 디브이디 등) 및 전기적 기록매체(예를 들면, 플래쉬 메모리, 메모리 스틱 등)와 같은 기록매체를 포함한다.

[0056]

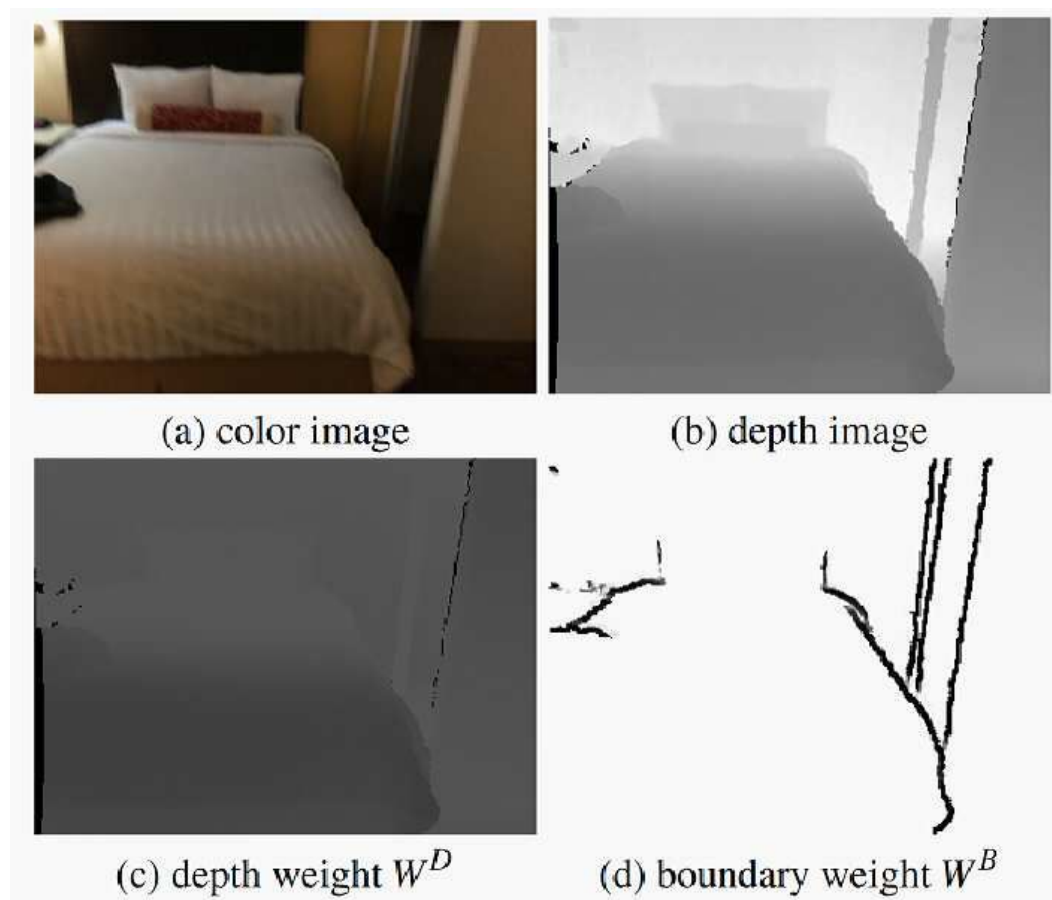
이제까지 본 발명에 대하여 그 바람직한 실시예들을 중심으로 살펴보았다. 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자는 본 발명이 본 발명의 본질적인 특성에서 벗어나지 않는 범위에서 변형된 형태로 구현될 수 있음을 이해할 수 있을 것이다. 그러므로 개시된 실시예는 한정적인 관점이 아니라 설명적인 관점에서 고려되어야 한다. 본 발명의 범위는 전술한 설명이 아니라 특허청구범위에 나타나 있으며, 그와 동등한 범위 내에 있는 모든 차이점은 본 발명에 포함된 것으로 해석되어야 할 것이다.

도면

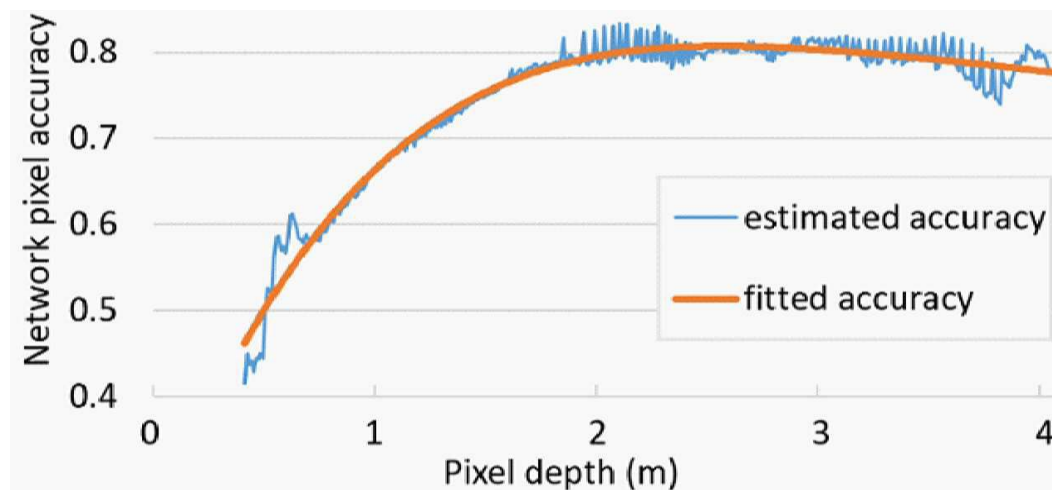
도면1



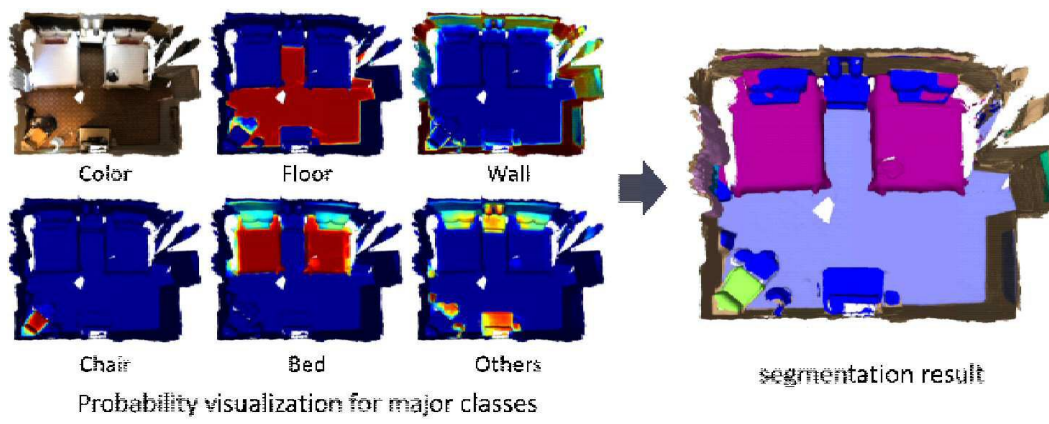
도면2



도면3



도면4



도면5

