



(19) 대한민국특허청(KR)

(12) 등록특허공보(B1)

(45) 공고일자 2024년09월30일

(11) 등록번호 10-2711526

(24) 등록일자 2024년09월25일

(51) 국제특허분류(Int. Cl.)

G06T 7/246 (2017.01) G06T 5/20 (2024.01)

G06T 5/77 (2024.01) G06T 7/13 (2017.01)

G06T 7/223 (2017.01) G06T 7/30 (2017.01)

G06T 7/50 (2017.01) G06T 7/90 (2017.01)

(52) CPC특허분류

G06T 7/246 (2017.01)

G06T 5/20 (2024.01)

(21) 출원번호 10-2024-0028618

(22) 출원일자 2024년02월28일

심사청구일자 2024년02월28일

(56) 선행기술조사문헌

Debaditya Acharya 등, BIM-Tracker: A model-based visual tracking approach for indoor localisation using a 3D building model, ISPRS(2019.02.27.)*

(뒷면에 계속)

(73) 특허권자

고려대학교 산학협력단

서울특별시 성북구 안암로 145, 고려대학교 (안암동5가)

주식회사 이안

서울특별시 강남구 봉은사로 502, 3층(삼성동, 삼하빌딩)

(72) 발명자

한정현

서울특별시 성북구 종암로23길 35, 210동 301호(종암동, 래미안세레니티)

안정현

경기도 고양시 일산서구 고양대로 624, 103동 1802호(일산동, 일산태영테시앙1단지아파트)

(뒷면에 계속)

(74) 대리인

송인호, 윤형근, 최관락

전체 청구항 수 : 총 9 항

심사관 : 이재원

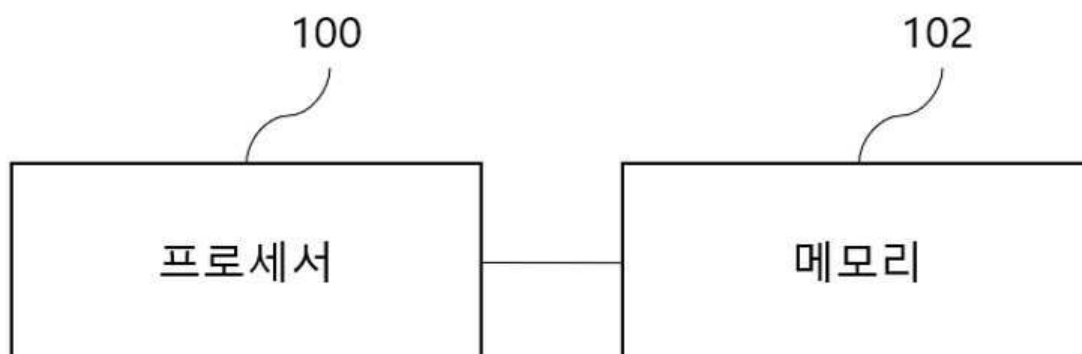
(54) 발명의 명칭 실내 환경에서의 BIM 기반 카메라 트래킹 방법 및 장치

(57) 요약

본 발명은 실내 환경에서의 BIM 기반 카메라 트래킹 방법 및 장치를 개시한다. 본 발명에 따르면, 실내 환경에서의 BIM 기반 분할 지원 카메라 트래킹 장치로서, 프로세서; 및 상기 프로세서에 연결되는 메모리를 포함하되, 상기 메모리는, RGB-D 카메라를 통해 촬영된 기둥을 포함하는 RGB 이미지 및 깊이 맵을 입력 받고, 상기 RGB 이미

(뒷면에 계속)

대표도 - 도1



지 및 상기 깊이 맵을 이용하여 기둥에 대한 RGB 모서리 및 깊이 모서리를 감지하고, 상기 RGB 모서리 및 깊이 모서리를 결합한 결과에서 양 끝점 계산 및 이상치를 제거하여 상기 기둥에 상응하는 장면(scene) 모서리를 감지하고, 상기 장면 모서리를 이용하여 상기 장면 모서리에 대응되는 BIM 모서리를 탐색하고, 평면 감지 알고리즘을 이용하여 바닥과 기둥의 장면 평면을 감지하고, 상기 장면 평면의 중심점을 이용하여 상기 장면 평면에 대응되는 BIM 평면을 탐색하고, 카메라 트래킹을 위해 서로 매칭되는 장면 모서리 및 BIM 모서리의 쌍과, 서로 매칭되는 장면 평면 및 BIM 평면 쌍에 대한 모서리 오차, 평면 오차, 중력 오차 및 카메라 포즈 오차의 가중합으로 정의되는 전체 오차를 최소화하여 잘못된 매칭 결과를 제거하도록, 상기 프로세서에 의해 실행 가능한 프로그램 명령어들을 저장하는 카메라 트래킹 장치가 제공된다.

(52) CPC특허분류

G06T 5/77 (2024.01)
G06T 7/13 (2017.01)
G06T 7/223 (2017.01)
G06T 7/30 (2017.01)
G06T 7/50 (2017.01)
G06T 7/90 (2017.01)
G06T 2207/20061 (2013.01)

(56) 선행기술조사문헌

Walid Darwish 등, AN RGB-D DATA PROCESSING FRAMEWORK BASED ON ENVIRONMENT*
KR1020220054582 A*
KR1020220119664 A
JP2011118724 A
*는 심사관에 의하여 인용된 문헌

(72) 발명자

하정호

경기도 의정부시 정음로 15, 202동 510호(고산동, 정음마을 고산2단지)

손재민

대구광역시 달서구 선원로 171, 105동 703호(이곡동, 성서무지개타운)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711193303
과제번호	2020-0-00861-004
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	SW컴퓨팅산업원천기술개발
연구과제명	(SW 스타랩) 실제 환경과 가상 물체 간 양방향 상호작용을 지원하는 혼합현실 기술

개발

과제수행기관명	고려대학교산학협력단
연구기간	2023.01.01 ~ 2023.12.31

이 발명을 지원한 국가연구개발사업

과제고유번호	1711193663
과제번호	2020-0-01819-004
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	정보통신방송혁신인재양성
연구과제명	ICT명품인재양성(고려대학교)
과제수행기관명	고려대학교산학협력단
연구기간	2023.01.01 ~ 2023.12.31

명세서

청구범위

청구항 1

실내 환경에서의 BIM 기반 분할 지원 카메라 트래킹 장치로서,

프로세서; 및

상기 프로세서에 연결되는 메모리를 포함하되,

상기 메모리는,

RGB-D 카메라를 통해 촬영된 기둥을 포함하는 RGB 이미지 및 깊이 맵을 입력 받고,

상기 RGB 이미지 및 상기 깊이 맵을 이용하여 기둥에 대한 RGB 모서리 및 깊이 모서리를 감지하고,

상기 RGB 모서리 및 깊이 모서리를 결합한 결과에서 양 끝점 계산 및 이상치를 제거하여 상기 기둥에 상응하는 장면(scene) 모서리를 감지하고,

상기 장면 모서리를 이용하여 상기 장면 모서리에 대응되는 BIM 모서리를 탐색하고,

평면 감지 알고리즘을 이용하여 바닥과 기둥의 장면 평면을 감지하고,

상기 장면 평면의 중심점을 이용하여 상기 장면 평면에 대응되는 BIM 평면을 탐색하고,

카메라 트래킹을 위해 서로 매칭되는 장면 모서리 및 BIM 모서리의 쌍과, 서로 매칭되는 장면 평면 및 BIM 평면 쌍에 대한 모서리 오차, 평면 오차, 중력 오차 및 카메라 포즈 오차의 가중합으로 정의되는 전체 오차를 최소화 하여 잘못된 매칭 결과를 제거하도록,

상기 프로세서에 의해 실행 가능한 프로그램 명령어들을 저장하되,

상기 장면 모서리는 복수 개 감지되며

상기 프로그램 명령어들은,

상기 실내 환경에서 검출되는 선분과 평면은 서로 수직이나 수평이라는 가정(Manhattan World Assumption)을 이용하여 복수의 장면 모서리 중 이상치를 제거하고,

상기 복수의 장면 모서리의 허프(Hough) 변환과 각 장면 모서리 각각의 중심점 깊이를 기준으로 k-d 트리를 적용하여 상기 복수의 장면 모서리 각각의 BIM 모서리를 탐색하는 카메라 트래킹 장치.

청구항 2

제1항에 있어서,

상기 프로그램 명령어들은,

상기 깊이 맵을 이용하여 상기 기둥의 표면에서의 노멀 맵을 생성하고,

상기 깊이 맵과 노멀 맵에 라플라시안 필터링을 적용한 결과를 OR 연산하여 상기 기둥에 대한 깊이 모서리를 감지하는 카메라 트래킹 장치.

청구항 3

제1항에 있어서,

상기 프로그램 명령어들은,

상기 RGB 이미지에 캐니 에지 검출기(Canny Edge Detector)를 적용하여 상기 기둥에 대한 RGB 모서리를 감지하는 카메라 트래킹 장치.

청구항 4

제1항에 있어서,

상기 프로그램 명령어들은,

상기 RGB 모서리 및 깊이 모서리를 AND 연산하여 결합하고,

상기 결합한 결과에 확률 허프 라인 변환(Probabilistic Hough Line Transform)을 적용하여 상기 기둥의 양 끝 점을 계산하는 카메라 트래킹 장치.

청구항 5

삭제

청구항 6

삭제

청구항 7

제1항에 있어서,

상기 모서리 오차는 아래의 수학식으로 정의되는 카메라 트래킹 장치.

[수학식]

$$E_{\text{edge}} = \sum_{i,j} (|\mathbf{l}_i \cdot \dot{\mathbf{p}}_{j,1}|^2 + |\mathbf{l}_i \cdot \dot{\mathbf{p}}_{j,2}|^2)$$

여기서, 모서리 이미지에 투영된 i 번째 BIM 모서리에 대한 직선의 방정식이 $a_i x + b_i y + c_i = 0$ 이고, 이를 표현하는 3차원 벡터 $\mathbf{l}_i$ 가 (a_i, b_i, c_i) 이며, 감지된 j 번째 장면 모서리 양 끝점의 동차 좌표를 $\dot{\mathbf{p}}_{j,1}$ 와 $\dot{\mathbf{p}}_{j,2}$ 임

청구항 8

제1항에 있어서,

상기 평면 오차는 아래의 수학식으로 정의되는 카메라 트래킹 장치.

[수학식]

$$E_{\text{face}} = \sum_{i,j} |(\mathbf{c}_i - \mathbf{c}_j) \cdot \mathbf{n}_j|^2$$

여기서, i 번째 BIM 평면의 중심점은 $\mathbf{c}_i$, j 번째 장면 평면의 중심점과 노멀은 $\mathbf{c}_j$, $\mathbf{n}_j$ 임

청구항 9

제1항에 있어서,

상기 중력 오차는 아래의 수학식으로 정의되는 카메라 트래킹 장치.

[수학식]

$$E_{\text{grav}} = |1 - \mathbf{g}_{\text{world}} \cdot \mathbf{R}_{\text{cam}} \mathbf{g}_{\text{imu}}|^2$$

여기서, $\mathbf{g}_{\text{world}} = (0, -1, 0)^T$ 이고 카메라의 회전 행렬은 $\mathbf{R}_{\text{cam}}$, IMU 센서로 측정된 중력 방향 벡터는 $\mathbf{g}_{\text{imu}}$ 임

청구항 10

제1항에 있어서,

상기 포즈 오차는 아래의 수학식으로 정의되는 카메라 트래킹 장치.

[수학식]

$$E_{\text{pose}} = \|(\mathbf{r}_{\text{cam}}, \mathbf{t}_{\text{cam}}) - (\mathbf{r}_A, \mathbf{t}_A)\|_2^2$$

여기서, 추정 카메라 포즈는 6차원 벡터 $(\mathbf{r}_A, \mathbf{t}_A)$ 로 정의됨

청구항 11

실내 환경에서의 BIM 기반 분할 지원 카메라 트래킹 장치로서,

프로세서; 및

상기 프로세서에 연결되는 메모리를 포함하되,

상기 메모리는,

RGB-D 카메라를 통해 촬영된 기둥을 포함하는 RGB 이미지 및 깊이 맵을 입력 받고,

딥러닝 기반 분할 알고리즘을 이용하여 상기 RGB 이미지로부터 복수의 장면(Scene) 기둥 영역을 분할하고,

상기 복수의 장면 기둥 영역 각각의 중심점 및 렌더링을 통해 생성한 복수의 BIM 기둥 영역 각각의 중심점을 기반으로 상기 복수의 장면 기둥 영역과 상기 복수의 BIM 기둥 영역 각각을 매칭하고,

상기 RGB 이미지 및 상기 깊이 맵을 이용하여 기둥에 대한 RGB 모서리 및 깊이 모서리를 감지하고,

상기 장면 기둥 영역에서 계산된 바운딩 박스로 결정되는 로컬 공간에서 계산한 장면 모서리의 허프(Hough) 변환과 BIM 기둥 영역에서 계산된 바운딩 박스로 결정되는 로컬 공간에서 계산한 BIM 모서리의 허프 변환을 기준으로 k-d 트리를 적용하여 복수의 장면 모서리 및 상기 복수의 장면 모서리 각각의 BIM 모서리를 탐색하고,

상기 장면 기둥 영역에서 인접한 픽셀 간의 거리 및 상기 인접한 픽셀 간의 노멀 차이를 이용하여 기둥 평면 및 바닥 평면을 포함하는 장면 평면을 감지하고,

상기 장면 평면의 3차원 좌표 및 노멀의 평균을 각각 중심점과 노멀로 정의하고, 이를 기반으로 k-d 트리를 활용하여 대응되는 BIM 평면을 탐색하고,

매칭된 모든 모서리 쌍과 평면 쌍의 오차를 계산하여 잘못 매칭된 것으로 판단되는 쌍을 제거하고,

상기 제거 후 남은 모서리 쌍과 평면 쌍을 사용하여 최종 카메라 포즈를 계산하도록,

상기 프로세서에 의해 실행 가능한 프로그램 명령어들을 저장하는 카메라 트래킹 장치.

발명의 설명

기술 분야

[0001] 본 발명은 실내 환경에서의 BIM 기반 카메라 트래킹 방법 및 장치에 관한 것이다.

배경 기술

[0002] AR 기반 시각 검사 작업에서 핵심 요소는 카메라 트래킹이다.

[0003] 카메라 위치 및 회전에 관한 카메라 포즈가 정확하게 추정되지 않으면 BIM(Building Information Modeling)이

작업 현장에 올바르게 오버레이되지 않는다.

- [0004] 일반적으로 AEC(Architecture, Engineering & Construction) 산업에서는 실내 환경이 광범위하여 카메라 트래킹이 특히 어려워진다.
- [0005] 예를 들어, 반도체 제조 공장(fab)을 생각해보면, fab은 여러 생산 라인으로 구성되어 있으며 일반적인 생산 라인은 약 10,000제곱미터를 차지한다.
- [0006] 이러한 광범위한 환경에서는 미숙한 카메라 트래킹 방법이 시간이 지남에 따라 드리프트를 누적할 가능성이 높아져 결국 시각 검사를 불가능하게 만든다.

선행기술문헌

특허문헌

- [0007] (특허문헌 0001) KR 등록특허공보 제10-2432164호

발명의 내용

해결하려는 과제

- [0008] 상기한 종래기술의 문제점을 해결하기 위해, 본 발명은 카메라 트래킹의 정확도와 강건성을 높일 수 있는 실내 환경에서의 BIM 기반 카메라 트래킹 방법 및 장치를 제안하고자 한다.

과제의 해결 수단

- [0009] 상기한 바와 같은 목적을 달성하기 위하여, 본 발명의 일 실시예에 따르면, 실내 환경에서의 BIM 기반 분할 지원 카메라 트래킹 장치로서, 프로세서; 및 상기 프로세서에 연결되는 메모리를 포함하되, 상기 메모리는, RGB-D 카메라를 통해 촬영된 기둥을 포함하는 RGB 이미지 및 깊이 맵을 입력 받고, 상기 RGB 이미지 및 상기 깊이 맵을 이용하여 기둥에 대한 RGB 모서리 및 깊이 모서리를 감지하고, 상기 RGB 모서리 및 깊이 모서리를 결합한 결과에서 양 끝점 계산 및 이상치를 제거하여 상기 기둥에 상응하는 장면(scene) 모서리를 감지하고, 상기 장면 모서리를 이용하여 상기 장면 모서리에 대응되는 BIM 모서리를 탐색하고, 평면 감지 알고리즘을 이용하여 바닥과 기둥의 장면 평면을 감지하고, 상기 장면 평면의 중심점을 이용하여 상기 장면 평면에 대응되는 BIM 평면을 탐색하고, 카메라 트래킹을 위해 서로 매칭되는 장면 모서리 및 BIM 모서리의 쌍과, 서로 매칭되는 장면 평면 및 BIM 평면 쌍에 대한 모서리 오차, 평면 오차, 중력 오차 및 카메라 포즈 오차의 가중합으로 정의되는 전체 오차를 최소화하여 잘못된 매칭 결과를 제거하도록, 상기 프로세서에 의해 실행 가능한 프로그램 명령어들을 저장하는 카메라 트래킹 장치가 제공된다.
- [0010] 상기 프로그램 명령어들은, 상기 깊이 맵을 이용하여 상기 기둥의 표면에서의 노멀 맵을 생성하고, 상기 깊이 맵과 노멀 맵에 라플라시안 필터링을 적용한 결과를 OR 연산하여 상기 기둥에 대한 깊이 모서리를 감지할 수 있다.
- [0011] 상기 프로그램 명령어들은, 상기 RGB 이미지에 캐니 에지 검출기(Canny Edge Detector)를 적용하여 상기 기둥에 대한 RGB 모서리를 감지할 수 있다.
- [0012] 상기 프로그램 명령어들은, 상기 RGB 모서리 및 깊이 모서리를 AND 연산하여 결합하고, 상기 결합한 결과에 확률 허프 라인 변환(Probabilistic Hough Line Transform)을 적용하여 상기 기둥의 양 끝점을 계산할 수 있다.
- [0013] 상기 장면 모서리는 복수 개 감지되며 상기 프로그램 명령어들은, 상기 실내 환경에서 검출되는 선분과 평면은 서로 수직이나 수평이라는 가정(Manhattan World Assumption)을 이용하여 복수의 장면 모서리 중 이상치를 제거할 수 있다.
- [0014] 상기 프로그램 명령어들은, 상기 복수의 장면 모서리의 허프(Hough) 변환과 각 장면 모서리 각각의 중심점 깊이를 기준으로 k-d 트리를 적용하여 상기 복수의 장면 모서리 각각의 BIM 모서리를 탐색할 수 있다.
- [0015] 상기 모서리 오차는 아래의 수학식으로 정의될 수 있다.

[0016] [수학식]

$$E_{\text{edge}} = \sum_{i,j} (|\mathbf{l}_i \cdot \dot{\mathbf{p}}_{j,1}|^2 + |\mathbf{l}_i \cdot \dot{\mathbf{p}}_{j,2}|^2)$$

[0017]

[0018] 여기서, 모서리 이미지에 투영된 i 번째 BIM 모서리에 대한 직선의 방정식이 $a_i x + b_i y + c_i = 0$ 이고, 이를 표현하는 3차원 벡터 $\mathbf{l}_i$ 가 (a_i, b_i, c_i) 이며, 감지된 j 번째 장면 모서리 양 끝점의 동차 좌표를 $\dot{\mathbf{p}}_{j,1}$ 와 $\dot{\mathbf{p}}_{j,2}$ 임

[0019] 상기 평면 오차는 아래의 수학식으로 정의될 수 있다.

[0020] [수학식]

$$E_{\text{face}} = \sum_{i,j} |(\mathbf{c}_i - \mathbf{c}_j) \cdot \mathbf{n}_j|^2$$

[0021]

[0022] 여기서, i 번째 BIM 평면의 중심점은 $\mathbf{c}_i$, j 번째 장면 평면의 중심점과 노멀은 $\mathbf{c}_j$, $\mathbf{n}_j$ 임

[0023] 상기 중력 오차는 아래의 수학식으로 정의될 수 있다.

[0024] [수학식]

$$E_{\text{grav}} = |1 - \mathbf{g}_{\text{world}} \cdot \mathbf{R}_{\text{cam}} \mathbf{g}_{\text{imu}}|^2$$

[0025]

[0026] 여기서, $\mathbf{g}_{\text{world}} = (0, -1, 0)^T$ 이고 카메라의 회전 행렬은 $\mathbf{R}_{\text{cam}}$, IMU 센서로 측정된 중력 방향 벡터는 $\mathbf{g}_{\text{imu}}$ 임

[0027] 상기 포즈 오차는 아래의 수학식으로 정의될 수 있다.

[0028] [수학식]

$$E_{\text{pose}} = \|(\mathbf{r}_{\text{cam}}, \mathbf{t}_{\text{cam}}) - (\mathbf{r}_A, \mathbf{t}_A)\|_2^2$$

[0029]

[0030] 여기서, 추정 카메라 포즈는 6차원 벡터 $(\mathbf{r}_A, \mathbf{t}_A)$ 로 정의됨

[0031] 본 발명의 다른 측면에 따르면, 실내 환경에서의 BIM 기반 분할 지원 카메라 트래킹 장치로서, 프로세서; 및 상기 프로세서에 연결되는 메모리를 포함하되, 상기 메모리는, RGB-D 카메라를 통해 촬영된 기둥을 포함하는 RGB 이미지 및 깊이 맵을 입력 받고, 딥러닝 기반 분할 알고리즘을 이용하여 상기 RGB 이미지로부터 복수의 장면(Scene) 기둥 영역을 분할하고, 상기 복수의 장면 기둥 영역 각각의 중심점 및 렌더링을 통해 생성한 복수의 BIM 기둥 영역 각각의 중심점을 기반으로 상기 복수의 장면 기둥 영역과 상기 복수의 BIM 기둥 영역 각각을 매칭하고, 상기 장면 기둥 영역에서 계산된 바운딩 박스로 결정되는 로컬 공간에서 계산한 장면 모서리의 허프(Hough) 변환과 BIM 기둥 영역에서 계산된 바운딩 박스로 결정되는 로컬 공간에서 계산한 BIM 모서리의 허프 변환을 기준으로 k-d 트리를 적용하여 복수의 장면 모서리 및 상기 복수의 장면 모서리 각각의 BIM 모서리를 탐색하고, 상기 장면 기둥 영역에서 인접한 픽셀 간의 거리 및 상기 인접한 픽셀 간의 노멀 차이를 이용하여 기둥 평면 및 바닥 평면을 포함하는 장면 평면을 감지하고, 상기 장면 평면의 3차원 좌표 및 노멀의 평균을 각각 중심점과 노멀로 정의하고, 이를 기반으로 k-d 트리를 활용하여 대응되는 BIM 평면을 탐색하고, 매칭된 모든 모서리 쌍과 평면 쌍의 오차를 계산하여 잘못 매칭된 것으로 판단되는 쌍을 제거하고, 상기 제거 후 남은 모서리 쌍과 평면 쌍을 사용하여 최종 카메라 포즈를 계산하도록, 상기 프로세서에 의해 실행 가능한 프로그램 명령어들을 저장하는 카메라 트래킹 장치가 제공된다.

발명의 효과

- [0032] 본 발명에 따르면, 텍스처가 없는 실내 환경에서 모서리와 평면과 같은 기하학적인 특징을 활용함으로써 카메라 트래킹의 정확도와 강건성을 높일 수 있으며, 이를 통해 광범위한 실내 환경에서도 효과적인 AR 기반 시각 검사가 가능한 장점이 있다.

도면의 간단한 설명

- [0033] 도 1은 본 발명의 바람직한 일 실시예에 따른 카메라 트래킹 장치의 구성을 도시한 도면이다.
 도 2는 본 발명의 제1 실시예에 따른 모서리 감지 과정을 나타낸 도면이다.
 도 3은 본 발명의 제1 실시예에 따른 Hough 변환을 설명하기 위한 도면이다.
 도 4는 본 발명의 제1 실시예에 따른 카메라 트래킹 과정을 도시한 도면이다.
 도 5는 본 발명의 제2 실시예에 따른 기둥 영역 분할 및 중심점을 이용한 매칭 과정을 도시한 도면이다.
 도 6은 본 발명의 제2 실시예에 따른 장면 모서리 감지 및 이상치 제거 과정을 도시한 도면이다.
 도 7은 본 발명의 제2 실시예에 따른 기둥이 로컬 공간에서 정의되는 것을 나타낸 도면이다.
 도 8은 본 발명의 제2 실시예에 따른 평면 매칭 과정을 나타낸 도면이다.
 도 9는 본 발명의 제2 실시예에 따른 장면 기둥 영역을 정의하는 과정을 나타낸 도면이다.
 도 10은 본 발명의 제2 실시예에 따른 기둥 영역이 불완전한 경우를 처리하는 과정을 나타낸 도면이다.

발명을 실시하기 위한 구체적인 내용

- [0034] 본 발명은 다양한 변경을 가할 수 있고 여러 가지 실시예를 가질 수 있는 바, 특정 실시예들을 도면에 예시하고 상세한 설명에 상세하게 설명하고자 한다. 그러나 이는 본 발명을 특정한 실시 형태에 대해 한정하려는 것이 아니며, 본 발명의 사상 및 기술 범위에 포함되는 모든 변경, 균등물 내지 대체물을 포함하는 것으로 이해되어야 한다.
- [0035] 본 명세서에서 사용한 용어는 단지 특정한 실시예를 설명하기 위해 사용된 것으로, 본 발명을 한정하려는 의도가 아니다. 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수의 표현을 포함한다. 본 명세서에서, "포함하다" 또는 "가지다" 등의 용어는 명세서상에 기재된 특징, 숫자, 단계, 동작, 구성요소, 부품 또는 이들을 조합한 것이 존재함을 지정하려는 것이지, 하나 또는 그 이상의 다른 특징들이나 숫자, 단계, 동작, 구성요소, 부품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 미리 배제하지 않는 것으로 이해되어야 한다.
- [0036] 또한, 각 도면을 참조하여 설명하는 실시예의 구성 요소가 해당 실시예에만 제한적으로 적용되는 것은 아니며, 본 발명의 기술적 사상이 유지되는 범위 내에서 다른 실시예에 포함되도록 구현될 수 있으며, 또한 별도의 설명이 생략될지라도 복수의 실시예가 통합된 하나의 실시예로 다시 구현될 수도 있음은 당연하다.
- [0037] 또한, 첨부 도면을 참조하여 설명함에 있어, 도면 부호에 관계없이 동일한 구성 요소는 동일하거나 관련된 참조 부호를 부여하고 이에 대한 중복되는 설명은 생략하기로 한다. 본 발명을 설명함에 있어서 관련된 공지 기술에 대한 구체적인 설명이 본 발명의 요지를 불필요하게 흐릴 수 있다고 판단되는 경우 그 상세한 설명을 생략한다.
- [0039] 도 1은 본 발명의 바람직한 일 실시예에 따른 카메라 트래킹 장치의 구성을 도시한 도면이다.
- [0040] 도 1에 도시된 바와 같이, 본 실시예에 따른 카메라 트래킹 장치는 프로세서(100) 및 메모리(102)를 포함할 수 있다.
- [0041] 여기서, 프로세서(100)는 컴퓨터 프로그램을 실행할 수 있는 CPU(central processing unit)나 그 밖에 가상 머신 등을 포함할 수 있다.
- [0042] 메모리(102)는 고정식 하드 드라이브나 착탈식 저장 장치와 같은 불휘발성 저장 장치를 포함할 수 있다. 착탈식 저장 장치는 콤팩트 플래시 유닛, USB 메모리 스틱 등을 포함할 수 있다. 메모리(102)는 각종 랜덤 액세스 메모리와 같은 휘발성 메모리도 포함할 수 있으며, 컴퓨터 판독 가능한 기록매체로 정의될 수 있다.
- [0043] 본 실시예에 따른 메모리(102)에는 RGB-D 카메라를 통해 촬영된 이미지를 이용하여 실내 환경에서의 장면 모서

리 및 평면을 감지하고, 이를 BIM 모서리 및 평면과 매칭시켜 카메라 트래킹을 수행하기 위한 프로그램 명령어들이 저장된다.

- [0044] 이하에서 설명하는 과정은 본 실시예에 따른 프로세서(100)가 프로그램 명령어들을 실행하는 과정을 정의될 수 있다.
- [0045] 텍스처가 없는 환경에서는 SLAM(Simultaneous localization and mapping)을 위해 기둥 모서리(이하, 모서리라 함)와 평면과 같은 기하학적 특징을 활용하는 것이 효과적인 점을 고려하여, 본 실시예에서는 모서리와 평면에 중점을 두며, 입력 이미지로부터 “장면(Scene) 모서리 및 평면”을 감지한 후 해당 “BIM 모서리 및 평면”과 매칭시킨다.
- [0046] 도 2는 본 발명의 제1 실시예에 따른 모서리 감지 과정을 나타낸 도면이다.
- [0047] 도 2는 RGB-D 카메라를 통해 촬영된 RGB 이미지 및 깊이 이미지(깊이 맵)을 입력으로 하여 제1 실시예에 따른 카메라 트래킹 장치가 수행하는 과정을 나타낸다.
- [0048] 도 2a와 같은 입력 깊이 맵을 이용하여 도 2b와 같은 노멀 맵을 생성하고, 두 맵에 라플라시안 필터링(Laplacian Filtering)을 적용한다.
- [0049] 라플라시안 필터링을 적용한 두 결과를 OR 연산으로 결합하여 도 2c와 같은 깊이 모서리를 감지하며, 동시에 입력 RGB 이미지에 캐니 에지 검출기(Canny Edge Detector)를 적용하여 도 2d와 같이 RGB 모서리를 감지한다.
- [0050] 다음으로 도 2e와 같이 깊이 모서리 및 RGB 모서리 감지 결과를 AND 연산으로 결합하고, 확률 허프 라인 변환(Probabilistic Hough Line Transform)을 이용하여 도 2f와 같이 모서리의 양 끝점을 계산한다.
- [0051] 실내 환경에서 검출되는 선분과 평면은 서로 수직이나 수평이라는 가정인 Manhattan World Assumption을 이용하여 도 2g와 같이 감지된 모서리 중 이상치를 제거하여, 도 2h와 같이 최종적으로 기둥의 장면 모서리를 감지한다.
- [0052] 상기한 바와 같이 감지된 복수의 장면 모서리의 허프(Hough) 변환 (r, θ) 와 중점 깊이를 기준으로 k-d 트리를 활용하여 대응되는 BIM 모서리를 찾는다.
- [0053] 도 3은 본 발명의 제1 실시예에 따른 Hough 변환을 설명하기 위한 도면이다.
- [0054] ARKit와 같은 공지된 평면 감지 기술을 사용하여 바닥과 기둥의 장면 평면을 감지하고 Manhattan World Assumption을 이용하여 이상치를 제거한다.
- [0055] 감지된 각 장면 평면의 3D 중심점을 기준으로 k-d 트리를 활용하여 대응되는 BIM 평면을 찾는다.
- [0056] 도 4는 본 발명의 제1 실시예에 따른 카메라 트래킹 과정을 도시한 도면이다.
- [0057] 도 4에 도시된 바와 같이, 제1 실시예에 따른 카메라 트래킹은 도 2 내지 3을 통해 매칭된 장면 모서리와 BIM 모서리 쌍과 장면 평면과 BIM 평면 쌍이 정합되는 최적의 강체 동작을 찾는 작업으로 정의된다.
- [0058] 도 4a는 장면 모서리(파란색)가 BIM 모서리(빨간색)와 일치하는 것을 나타낸다. 도 4b는 ARKit에 의해 감지된 평면이고, 도 4c는 검은색 원으로 표시된 세 개의 모서리 쌍이 소위 맨해튼 축 제약 조건(Manhattan axis constraint)을 충족하지 않는 경우를 나타내고, 도 4d는 BIM 기둥이 최적화를 통해 장면 기둥에 정합된 상태를 나타낸다.
- [0059] 최적화를 위해 모서리, 평면, 중력 및 포즈의 네 가지 오차가 정의되며, 이들의 가중 합으로 전체 오차가 계산된다.

- [0060] 모서리 이미지에 투영된 i 번째 BIM 모서리에 대한 직선의 방정식은 $a_i x + b_i y + c_i = 0$ 이고, 이를 표현하는 3차원 벡터 $\mathbf{l}_i$ 가 (a_i, b_i, c_i) 이며, 감지된 j 번째 장면 모서리 양 끝점의 동차 좌표를 $\dot{\mathbf{p}}_{j,1}$ 와 $\dot{\mathbf{p}}_{j,2}$ 로 표현할 때, 모서리 오차는 다음과 같다.

수학식 1

$$E_{\text{edge}} = \sum_{i,j} (|\mathbf{l}_i \cdot \dot{\mathbf{p}}_{j,1}|^2 + |\mathbf{l}_i \cdot \dot{\mathbf{p}}_{j,2}|^2)$$

i 번째 BIM 평면의 중심점을 $\mathbf{c}_i$, j 번째 장면 평면의 중심점과 노멀을 $\mathbf{c}_j$, $\mathbf{n}_j$ 로 표현할 때, 평면 오차는 다음과 같다.

수학식 2

$$E_{\text{face}} = \sum_{i,j} |(\mathbf{c}_i - \mathbf{c}_j) \cdot \mathbf{n}_j|^2$$

$\mathbf{g}_{\text{world}} = (0, -1, 0)^T$ 이고 카메라의 회전 행렬을 $\mathbf{R}_{\text{cam}}$, IMU 센서로 측정된 중력 방향 벡터를 $\mathbf{g}_{\text{imu}}$ 로 표현할 때, 중력 오차는 다음과 같다.

수학식 3

$$E_{\text{grav}} = |1 - \mathbf{g}_{\text{world}} \cdot \mathbf{R}_{\text{cam}} \mathbf{g}_{\text{imu}}|^2$$

종래의 ARKit로 추정된 카메라 포즈를 6차원 벡터 $(\mathbf{r}_A, \mathbf{t}_A)$ 로 표현할 때, 포즈 오차는 다음과 같다.

수학식 4

$$E_{\text{pose}} = \|(\mathbf{r}_{\text{cam}}, \mathbf{t}_{\text{cam}}) - (\mathbf{r}_A, \mathbf{t}_A)\|_2^2$$

따라서 전체 오차는 다음과 같으며, 이 최적화 문제는 Levenberg-Marquardt 방법을 사용하여 해결되고 RANSAC(Random Sample Consensus) 알고리즘을 통해 잘못된 매칭 결과를 제거한다.

수학식 5

$$E = w_{\text{edge}} E_{\text{edge}} + w_{\text{face}} E_{\text{face}} + w_{\text{grav}} E_{\text{grav}} + w_{\text{pose}} E_{\text{pose}}$$

카메라 트래킹의 정확도와 강건성을 제고하기 위해, 본 발명의 다른 실시예(제2 실시예)에서는 기동 분할 네트워크를 활용한다.

- [0072] 딥러닝 기반 분할 알고리즘(예를 들어, YOLACT++[Bolya, D., Zhou, C., Xiao, F., Lee, Y.J.: Yolact++: Better real-time instance segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence (2020)])을 사용하여 입력 RGB 이미지로부터 도 5a와 같이 장면 기둥 영역을 분할한다.
- [0073] 장면 기둥과 대응되는 BIM 기둥을 찾기 위해 중심점을 계산한다. 깊이 맵에서 장면 기둥 영역 내의 픽셀을 3차원 공간으로 역투영하고, 이들의 평균을 계산하여 중심점으로 정의한다.
- [0074] BIM 기둥은 렌더링을 통해 장면 기둥과 동일한 방식으로 중심점을 계산한다.
- [0075] 가까운 중심점을 갖는 장면 기둥과 BIM 기둥이 매칭된다.
- [0076] 도 5b와 같이 기둥들 3차원 공간에서 몇 미터 또는 몇 십 미터 정도 떨어져 있기 때문에, 이러한 중심점을 기반으로 한 매칭 과정이 효과적으로 작동한다.
- [0077] 본 발명의 제2 실시예에서는, 장면 기둥 영역에서 계산된 바운딩 박스로 결정되는 로컬 공간에서 계산한 장면 모서리의 허프(Hough) 변환과 BIM 기둥 영역에서 계산된 바운딩 박스로 결정되는 로컬 공간에서 계산한 BIM 모서리의 허프 변환을 기준으로 k-d 트리를 적용하여 복수의 장면 모서리 및 복수의 장면 모서리 각각의 BIM 모서리를 탐색하여 도 6a와 같이 장면 모서리를 감지하고, 기둥 분할 결과에서 장면 기둥에 속하지 않는 이상치를 도 6b와 같이 제거한다.
- [0078] 제1 실시예에서는 감지된 각 장면 모서리의 허프 변환 (r , θ)가 이미지 공간에서 정의되지만, 제2 실시예에서는 모서리를 특정한 기둥 영역에 속하여 (r , θ)를 도 7a와 같이 기둥의 로컬 공간에서 정의할 수 있다.
- [0079] 이를 위해 기둥 영역의 바운딩 박스가 계산되며, 좌측 상단을 로컬 공간의 원점으로 정의한다.
- [0080] 도 7b와 같이 각 BIM 모서리도 렌더링된 BIM 기둥의 로컬 공간에서 정의된다.
- [0081] 기둥 매칭 결과를 기반으로 각 기둥에 대한 모서리 매칭이 수행되며, 제1 실시예에서보다 정확하게 장면 모서리와 BIM 모서리를 일치시킨다.
- [0082] 제2 실시예에서는 기둥 영역을 사용하여 종래기술(ARKit)보다 정확한 장면 평면 감지를 수행한다.
- [0083] 도 8a와 같은 노멀 맵을 사용하여 도 8b와 같이 각 픽셀이 평면에 속하는지를 장면 기둥 영역에서 인접한 픽셀 간의 거리 및 상기 인접한 픽셀 간의 노멀 차이를 이용하여 기둥 평면 및 바닥 평면을 포함하는 장면 평면을 감지한다.
- [0084] 도 8c와 같이 기둥 영역과 겹치는 부분은 기둥 평면으로, 도 8d와 같이 노멀이 수직인 부분을 바닥 평면으로 간주한다.
- [0085] 각 부분의 3차원 좌표와 노멀의 평균을 장면 평면의 중심점과 노멀로 정의하며, 이를 기반으로 k-d 트리를 활용하여 대응되는 BIM 평면을 찾는다.
- [0086] 평면 매칭도 모서리 매칭과 유사하게 기둥 단위로 수행되어 정확도가 향상된다.
- [0087] 제1 실시예에서는 RANSAC을 사용하여 잘못된 매칭 결과를 제거했지만, 제2 실시예에서는 기둥 분할 네트워크를 활용하여 잘못된 매칭 결과가 거의 없다.
- [0088] 따라서 RANSAC 대신 두 단계의 최적화를 수행한다.
- [0089] 먼저 Levenberg-Marquardt 방법을 사용하여 매칭된 모든 장면 모서리와 BIM 모서리 쌍과, 장면 평면과 BIM 평면 쌍이 정합되는 카메라 포즈를 계산하고, 이를 통해 각 쌍의 오차를 계산하여 잘못 매칭된 것으로 판단되는 쌍을 제거한다. 이후에는 남은 쌍을 사용하여 최종 카메라 포즈를 계산한다.
- [0090] 기둥 분할 네트워크(YOLACT++)를 훈련하기 위한 데이터셋은 본 발명이 적용될 환경에서 다양한 시점으로 촬영한 RGB 이미지를 활용하여 생성된다. 기둥 영역은 자동으로 생성되며, 완전하게 분할되지 않은 경우에는 수동으로 후처리된다.
- [0091] 기둥 영역을 자동으로 생성하기 위해 Segment Anything Model (SAM) [Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., Xiao, T., Whitehead, S., Berg, A.C., Lo, W.Y., Dollár, P., Girshick, R.: Segment anything. arXiv:2304.02643 (2023)]을 사용하여 입력 RGB 이미지를 도 9a 및 9b와 같이 세분화한다. 이는 종래기술(ARKit)이나 본 발명의 제1 실시예에서 추정된 카메라 포즈로 렌더링된 BIM 영역

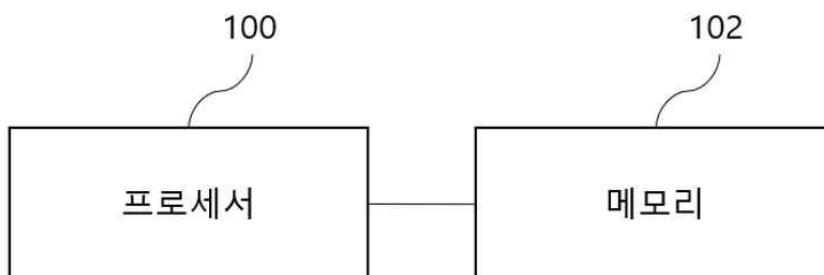
(도 9c)과 비교되어 겹치는 세분화된 영역들을 수집하고, 이를 도 9d와 같이 장면 기둥 영역으로 정의한다.

[0092] RGB 이미지로부터 도 10d와 같이 자동으로 생성된 기둥 영역이 불완전한 경우를 처리하기 위해, 본 발명에서는 사용하기 쉬운 데이터 처리 엔진을 개발하였다. 이 엔진은 SAM의 포인트 프롬프트 입력을 사용하여 새로운 영역을 생성하거나 지우는 기능을 포함하고 있다. 엔진의 인터페이스를 통해 자동으로 생성된 기둥 마스크가 RGB 이미지 위에 표시되며, 사용자는 포인트 프롬프트를 통해 도 10e와 같이 잘못된 기둥 영역을 도 10f와 같이 올바르게 완성한다.

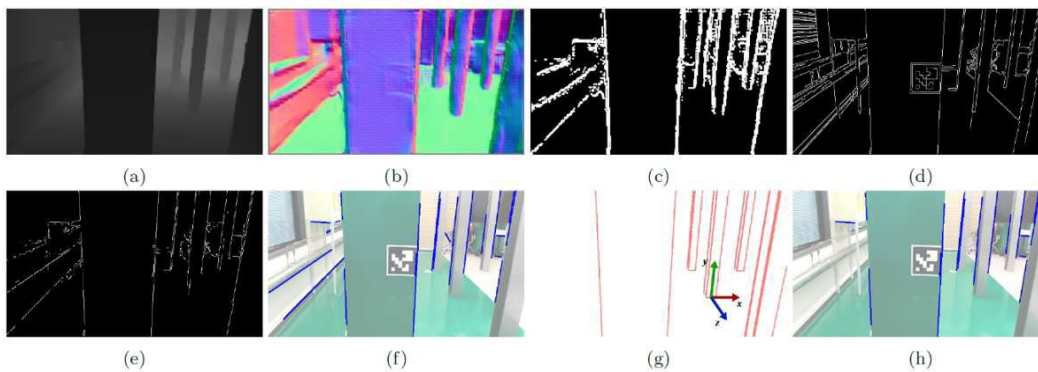
[0093] 상기한 본 발명의 실시예는 예시의 목적을 위해 개시된 것이고, 본 발명에 대한 통상의 지식을 가지는 당업자라면 본 발명의 사상과 범위 안에서 다양한 수정, 변경, 부가가 가능할 것이며, 이러한 수정, 변경 및 부가는 하기의 특허청구범위에 속하는 것으로 보아야 할 것이다.

도면

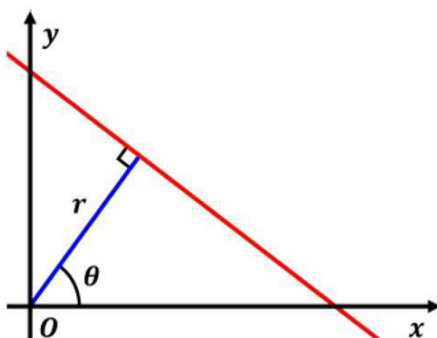
도면1



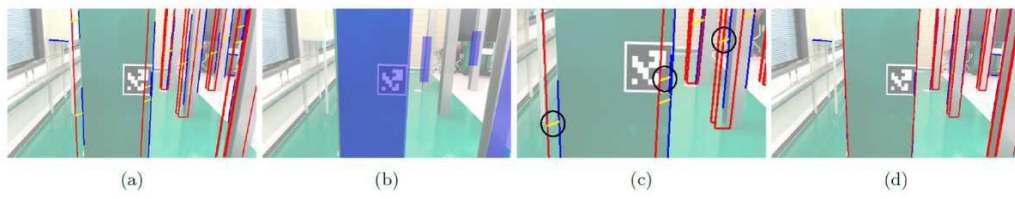
도면2



도면3



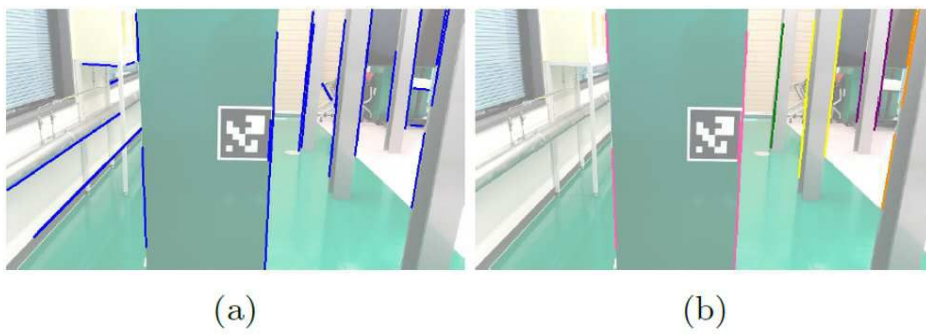
도면4



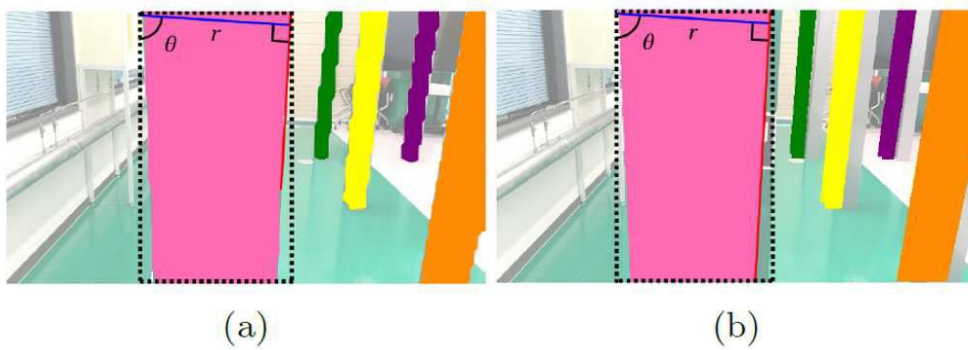
도면5



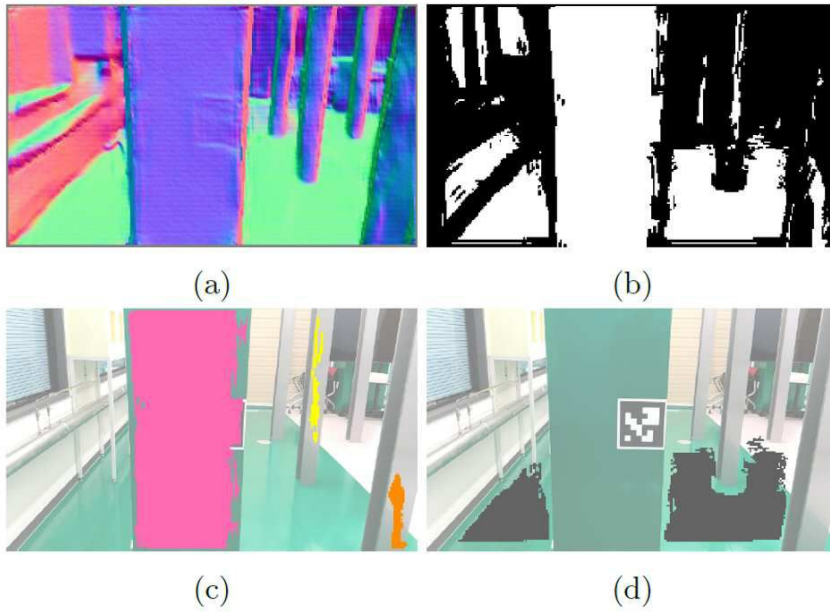
도면6



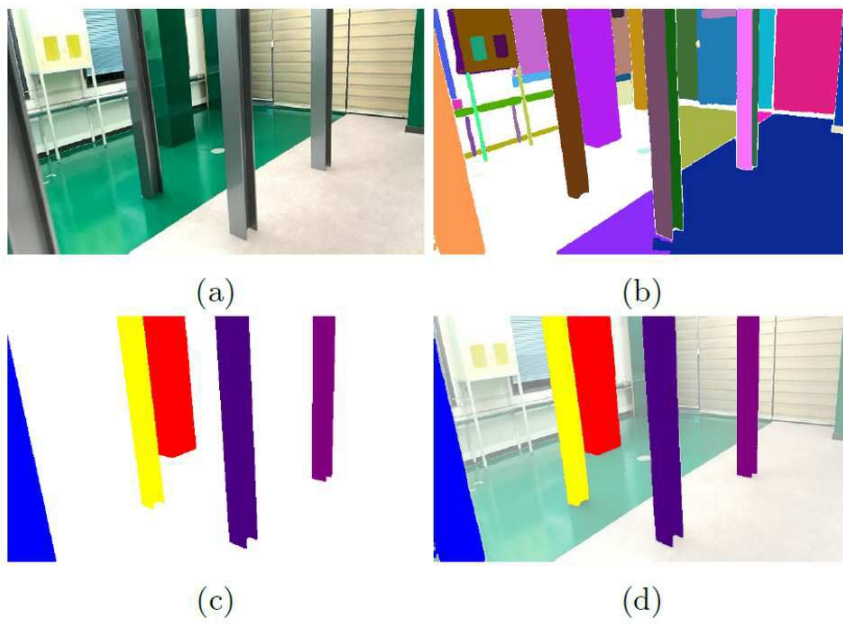
도면7



도면8



도면9



도면10

