



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2023-0046038
(43) 공개일자 2023년04월05일

(51) 국제특허분류(Int. Cl.)
G06T 7/73 (2017.01) G06F 18/00 (2023.01)
G06N 20/00 (2019.01) G06T 7/174 (2017.01)
G06T 7/50 (2017.01) G06T 7/90 (2017.01)
(52) CPC특허분류
G06T 7/73 (2017.01)
G06N 20/00 (2021.08)
(21) 출원번호 10-2021-0129023
(22) 출원일자 2021년09월29일
심사청구일자 2021년09월29일

(71) 출원인
한국기술교육대학교 산학협력단
충청남도 천안시 동남구 병천면 충절로 1600 (한
국기술교육대학교내)
(72) 발명자
김덕수
세종특별자치시 남세종로 302, 610동 506호(소담
동, 새샘마을6단지)
노유찬
강원도 춘천시 방송길 70, 106동 1702호(온의동,
온의 롯데캐슬 스카이크래스)
(74) 대리인
특허법인오암

전체 청구항 수 : 총 12 항

(54) 발명의 명칭 기계학습을 기반으로 하는 물체 자세 추정 시스템 및 그 방법

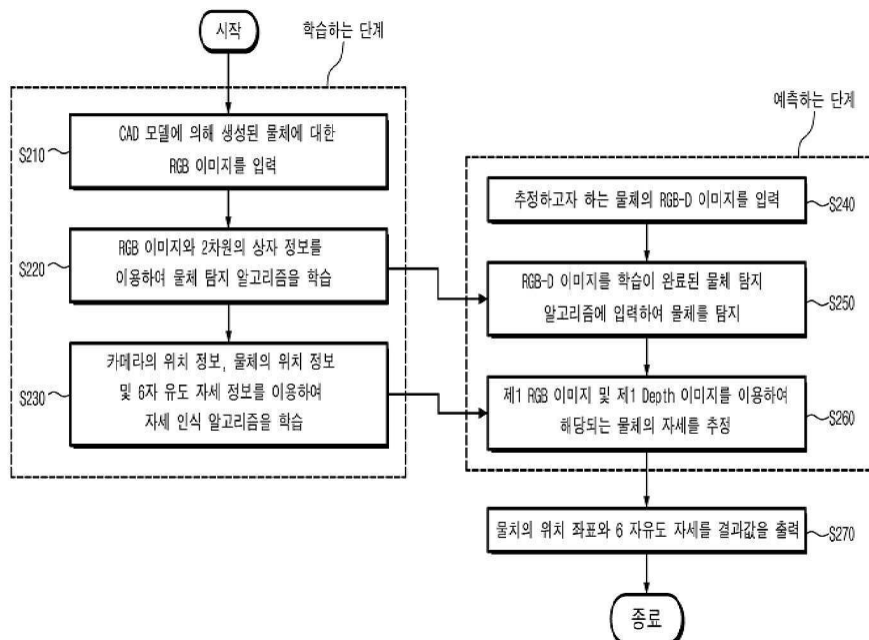
(57) 요약

본 발명은 기계학습을 기반으로 하는 물체 자세 추정 시스템 및 그 방법에 관한 것이다.

본 발명에 따르면, 기계학습을 기반으로 하는 물체 자세 추정 시스템은 RGB-D 카메라로부터 탐지 대상이 되는 물체를 촬영하여 획득한 RGB-D 이미지를 입력받는 이미지 입력부, 상기 RGB-D 이미지를 학습이 완료된 물체 탐지

(뒷면에 계속)

대표도 - 도2



알고리즘에 입력하고, 상기 물체 탐지 알고리즘에 의해 탐지된 물체에 바운딩 박스를 획득하고, 상기 획득한 바운딩 박스에 대응되는 제1 RGB 이미지 및 제1 Depth 이미지를 획득하는 물체 탐지부, 상기 제1 RGB 이미지 및 제1 Depth 이미지를 학습이 완료된 자세 인식 알고리즘에 입력하여 특징맵을 획득하고, 상기 획득한 특징맵으로부터 물체의 자세를 추정하는 자세 추정부, 그리고 탐지된 물체에 대한 이미지상의 좌표와 자세를 추정하여 획득한 좌표를 출력하는 출력부를 포함한다.

이와 같이 본 발명에 따르면, 물체 인식(object detection)과 자세 인식(pose estimation) 단계를 분리하여 제이하므로 데이터 처리 단계의 복잡성을 줄이고, 구현 및 활용의 용이성을 높일 수 있다.

(52) CPC특허분류

G06T 7/174 (2017.01)

G06T 7/50 (2017.01)

G06T 7/90 (2017.01)

G06V 20/00 (2022.01)

G06T 2210/12 (2013.01)

(72) 발명자

김영우

충청북도 청주시 서원구 쌍샘로107번길 12, 101동 1006호(모충동, 금호타운아파트)

김대건

광주광역시 서구 화정로 106-15, 101동 1710호(쌍촌동, 백일대주아파트)

한현규

충청남도 천안시 서북구 쌍용11길 58, 309동 1002호(쌍용동, 주공7단지아파트)

송영기

충청남도 천안시 서북구 공원로 195, 102동 3307호(불당동, 펜타포트)

이 발명을 지원한 국가연구개발사업

과제고유번호	1345332573
과제번호	LINC PLUS-2021-18
부처명	교육부
과제관리(전문)기관명	한국연구재단
연구사업명	LINC+ 가상회의를 위한 컴퓨터 생성 홀로그램 생성 기초기술 개발
연구과제명	사회맞춤형산학협력선도대학(LINC+)육성(0.5)
기 여 율	1/1
과제수행기관명	한국기술교육대학교
연구기간	2021.03.01 ~ 2022.02.28

명세서

청구범위

청구항 1

기계학습을 기반으로 하는 물체 자세 추정 시스템에 있어서,

RGB-D 카메라로부터 탐지 대상이 되는 물체를 촬영하여 획득한 RGB-D 이미지를 입력받는 이미지 입력부,

상기 RGB-D 이미지를 학습이 완료된 물체 탐지 알고리즘에 입력하고, 상기 물체 탐지 알고리즘에 의해 탐지된 물체에 바운딩 박스를 획득하고, 상기 획득한 바운딩 박스에 대응되는 제1 RGB 이미지 및 제1 Depth 이미지를 획득하는 물체 탐지부,

상기 제1 RGB 이미지 및 제1 Depth 이미지를 학습이 완료된 자세 인식 알고리즘에 입력하여 특징맵을 획득하고, 상기 획득한 특징맵으로부터 물체의 자세를 추정하는 자세 추정부, 그리고

탐지된 물체에 대한 이미지상의 좌표와 자세를 추정하여 획득한 좌표를 출력하는 출력부를 포함하는 물체 자세 추정 시스템.

청구항 2

제1항에 있어서,

가상공간에서의 물체에 대한 RGB 이미지와 물체의 위치를 나타내는 2차원의 상자 정보를 획득하고, 상기 획득한 RGB 이미지를 입력데이터로 하고 상기 2차원의 상자 정보를 출력데이터로 하여 기 구축된 물체 탐지 알고리즘을 학습시키는 제1 학습부, 그리고

상기 RGB 이미지를 생성한 카메라의 위치 정보, 물체의 위치 정보 및 6자 유도 자세 정보를 획득하고, 상기 획득한 카메라의 위치 정보 및 물체의 위치 정보를 입력데이터로 하고 6자 유도 자세 정보를 출력데이터로 하여 기 구축된 자세 인식 알고리즘을 학습시키는 제2 학습부를 더 포함하는 물체 자세 추정 시스템.

청구항 3

제2항에 있어서,

상기 2차원의 상자 정보는,

물체에 대한 종류, 물체의 위치, 바운딩 박스에 대한 좌표 중에서 적어도 하나를 포함하는 물체 자세 추정 시스템.

청구항 4

제1항에 있어서,

상기 물체 탐지부는,

입력된 상기 RGB-D 이미지를 RGB 이미지와 Depth 이미지로 분리하고, 상기 분리된 RGB 이미지 및 Depth 이미지를 상기 물체 탐지 알고리즘에 각각 입력하여 바운딩 박스를 획득한 다음,

상기 획득한 바운딩 박스에 대한 좌표를 이용하여 상기 RGB 이미지로부터 상기 바운딩 박스 부분을 크롭핑하여 제1 RGB 이미지를 획득하고, 상기 획득한 바운딩 박스에 대한 좌표를 이용하여 상기 Depth 이미지로부터 상기 바운딩 박스 부분을 크롭핑하여 제1 Depth 이미지를 획득하는 물체 자세 추정 시스템.

청구항 5

제1항에 있어서,

상기 자세 추정부는,

상기 제1 RGB 이미지에 포함된 복수의 픽셀 값을 255의 값으로 나누어 이진화된 값으로 변환하고, 상기 제1

Depth 이미지에 포함된 복수의 픽셀 값을 최대값과 최소값의 평균값으로 변환하는 전처리 과정을 수행하는 물체 자세 추정 시스템.

청구항 6

제5항에 있어서,

상기 자세 추정부는,

전처리가 완료된 상기 제1 RGB 이미지와 상기 제1 Depth 이미지를 자세 인식 알고리즘에 입력하여 커터니언 값을 예측 하는 물체 자세 추정 시스템.

청구항 7

물체 자세 추정 시스템을 이용한 물체 자세 추정 방법에 있어서,

RGB-D 카메라로부터 탐지 대상이 되는 물체를 촬영하여 획득한 RGB-D 이미지를 입력받는 단계,

상기 RGB-D 이미지를 학습이 완료된 물체 탐지 알고리즘에 입력하고, 상기 물체 탐지 알고리즘에 의해 탐지된 물체에 바운딩 박스를 획득하고, 상기 획득한 바운딩 박스에 대응되는 제1 RGB 이미지 및 제1 Depth 이미지를 획득하는 단계,

상기 제1 RGB 이미지 및 제1 Depth 이미지를 학습이 완료된 자세 인식 알고리즘에 입력하여 특징맵을 획득하고, 상기 획득한 특징맵으로부터 물체의 자세를 추정하는 단계, 그리고

탐지된 물체에 대한 이미지상의 좌표와 자세를 추정하여 획득한 좌표를 출력하는 단계를 포함하는 물체 자세 추정 방법.

청구항 8

제7항에 있어서,

가상공간에서의 물체에 대한 RGB 이미지와 물체의 위치를 나타내는 2차원의 상자 정보를 획득하고, 상기 획득한 RGB 이미지를 입력데이터로 하고 상기 2차원의 상자 정보를 출력데이터로 하여 기 구축된 물체 탐지 알고리즘을 학습시키는 단계, 그리고

상기 RGB 이미지를 생성한 카메라의 위치 정보, 물체의 위치 정보 및 6자 유도 자세 정보를 획득하고, 상기 획득한 카메라의 위치 정보 및 물체의 위치 정보를 입력데이터로 하고 6자 유도 자세 정보를 출력데이터로 하여 기 구축된 자세 인식 알고리즘을 학습시키는 단계를 더 포함하는 물체 자세 추정 방법.

청구항 9

제8항에 있어서,

상기 2차원의 상자 정보는,

물체에 대한 종류, 물체의 위치, 바운딩 박스에 대한 좌표 중에서 적어도 하나를 포함하는 물체 자세 추정 방법.

청구항 10

제7항에 있어서,

상기 제1 RGB 이미지 및 제1 Depth 이미지를 획득하는 단계는,

입력된 상기 RGB-D 이미지를 RGB 이미지와 Depth 이미지로 분리하고, 상기 분리된 RGB 이미지 및 Depth 이미지를 상기 물체 탐지 알고리즘에 각각 입력하여 바운딩 박스를 획득하는 단계, 그리고

상기 획득한 바운딩 박스에 대한 좌표를 이용하여 상기 RGB 이미지로부터 상기 바운딩 박스 부분을 크롭핑하여 제1 RGB 이미지를 획득하고, 상기 획득한 바운딩 박스에 대한 좌표를 이용하여 상기 Depth 이미지로부터 상기 바운딩 박스 부분을 크롭핑하여 제1 Depth 이미지를 획득하는 단계를 포함하는 물체 자세 추정 방법.

청구항 11

제7항에 있어서,

상기 물체의 자세를 추정하는 단계는,

상기 제1 RGB 이미지에 포함된 복수의 픽셀 값을 255의 값으로 나누어 이진화된 값으로 변환하고, 상기 제1 Depth 이미지에 포함된 복수의 픽셀 값을 최대값과 최소값의 평균값으로 변환하는 전처리 과정을 수행하는 물체 자세 추정 방법.

청구항 12

제11항에 있어서,

상기 물체의 자세를 추정하는 단계는,

전처리가 완료된 상기 제1 RGB 이미지와 상기 제1 Depth 이미지를 자세 인식 알고리즘에 입력하여 커터니언 값을 예측하는 물체 자세 추정 방법.

발명의 설명

기술 분야

[0001] 본 발명은 기계학습을 기반으로 하는 물체 자세 추정 시스템 및 그 방법에 관한 것으로서, 더욱 상세하게 설명하면 RGB-D 이미지로부터 탐지 대상이 되는 물체의 6자유도 자세를 추정하는 물체 자세 추정 시스템 및 그 방법에 관한 것이다.

배경 기술

[0003] 3차원 공간에서 물체의 6자유도 자세(위치 및 회전) 인식은 로봇 머니플레이션(robot manipulation), 자율 주행(autonomous driving), 증강 현실(augmented reality) 등 다양한 분야에서 중요한 기반 연산 중 하나다. 6자유도 자세 인식은 제조업 등과 같은 산업계 관점에서 로봇을 이용해서 부품이나 물건을 옮기는 pick-and-place 작업 등을 위한 필수 기술 중 하나다. 최근에는 아마존(Amazon)이 Picking challenge를 개최하면서, 정확한 물체 자세 인식에 관한 관심과 요구가 크게 증가하고 있다.

[0004] 물체의 자세 추정은 주로 카메라로 들어온 RGB 이미지를 기반으로 이루어진다. 그리고 종래의 자세 추정 기법은 3차원 모델과 이미지상의 특징점(feature point)을 매칭시키는 방법으로 이루어졌다. 특징점 추출은 SIFT(Scale Invariant Feature Transform) 등과 같이 특징점 추출 알고리즘에 기반을 둔 영상처리 기술을 사용하여 이루어졌다.

[0005] 최근에는 딥 러닝(deep learning) 기반의 인공지능 기술이 발전함에 따라 인공지능망을 이용하는 물체 자세 인식 기술 연구가 활발히 진행되고 있다. 초기 연구들은 인공지능망을 특징 점 검출 정확도 향상을 위해 사용하고, 검출된 특징점을 이용하여 PnP(Perspective-n-Point) 방법으로 자세를 추론하는 2단계 접근법을 사용하였다.

[0006] 근래의 연구들은 한 장의 이미지(RGB 또는 RGB-D)로부터 물체의 6자유도 자세를 바로 추론하는 방법들을 제안하고 있으며, 공개된 데이터 세트들에서 높은 정확도를 보여준다. 하지만 이러한 기술들이 사용하는 인공지능망의 경우, 그 구조가 복잡하고 규모가 크기 때문에 산업현장에서 쉽게 응용 및 적용 하기가 어렵다는 한계점을 가진다.

[0007] 본 발명의 배경이 되는 기술은 한국등록특허 제 10-0926783호(2009.11.06. 공고)에 개시되어 있다.

선행기술문헌

특허문헌

[0009] (특허문헌 0001) 한국등록특허 제10-926783호

발명의 내용

해결하려는 과제

[0010] 이와 같이 본 발명에 따르면, RGB-D 이미지로부터 탐지 대상이 되는 물체의 6자유도 자세를 추정하는 물체 자세 추정 시스템 및 그 방법을 제공하기 위한 것이다.

과제의 해결 수단

[0012] 이러한 기술적 과제를 이루기 위한 본 발명의 실시예에 따르면 기계학습을 기반으로 하는 물체 자세 추정 시스템은 RGB-D 카메라로부터 탐지 대상이 되는 물체를 촬영하여 획득한 RGB-D 이미지를 입력받는 이미지 입력부, 상기 RGB-D 이미지를 학습이 완료된 물체 탐지 알고리즘에 입력하고, 상기 물체 탐지 알고리즘에 의해 탐지된 물체에 바운딩 박스를 획득하고, 상기 획득한 바운딩 박스에 대응되는 제1 RGB 이미지 및 제1 Depth 이미지를 획득하는 물체 탐지부, 상기 제1 RGB 이미지 및 제1 Depth 이미지를 학습이 완료된 자세 인식 알고리즘에 입력하여 특징맵을 획득하고, 상기 획득한 특징맵으로부터 물체의 자세를 추정하는 자세 추정부, 그리고 탐지된 물체에 대한 이미지상의 좌표와 자세를 추정하여 획득한 좌표를 출력하는 출력부를 포함한다.

[0013] 가상공간에서의 물체에 대한 RGB 이미지와 물체의 위치를 나타내는 2차원의 상자 정보를 획득하고, 상기 획득한 RGB 이미지를 입력데이터로 하고 상기 2차원의 상자 정보를 출력데이터로 하여 기 구축된 물체 탐지 알고리즘을 학습시키는 제1 학습부, 그리고 상기 RGB 이미지를 생성한 카메라의 위치 정보, 물체의 위치 정보 및 6자 유도 자세 정보를 획득하고, 상기 획득한 카메라의 위치 정보 및 물체의 위치 정보를 입력데이터로 하고 6자 유도 자세 정보를 출력데이터로 하여 기 구축된 자세 인식 알고리즘을 학습시키는 제2 학습부를 더 포함할 수 있다.

[0014] 상기 2차원의 상자 정보는, 물체에 대한 종류, 물체의 위치, 바운딩 박스에 대한 좌표 중에서 적어도 하나를 포함할 수 있다.

[0015] 상기 물체 탐지부는, 입력된 상기 RGB-D 이미지를 RGB 이미지와 Depth 이미지로 분리하고, 상기 분리된 RGB 이미지 및 Depth 이미지를 상기 물체 탐지 알고리즘에 각각 입력하여 바운딩 박스를 획득한 다음, 상기 획득한 바운딩 박스에 대한 좌표를 이용하여 상기 RGB 이미지로부터 상기 바운딩 박스 부분을 크롭핑하여 제1 RGB 이미지를 획득하고, 상기 획득한 바운딩 박스에 대한 좌표를 이용하여 상기 Depth 이미지로부터 상기 바운딩 박스 부분을 크롭핑하여 제1 Depth 이미지를 획득할 수 있다.

[0016] 상기 자세 추정부는, 상기 제1 RGB 이미지에 포함된 복수의 픽셀 값을 255의 값으로 나누어 이진화된 값으로 변환하고, 상기 제1 Depth 이미지에 포함된 복수의 픽셀 값을 최대값과 최소값의 평균값으로 변환하는 전처리 과정을 수행할 수 있다.

[0017] 상기 자세 추정부는, 전처리가 완료된 상기 제1 RGB 이미지와 상기 제1 Depth 이미지를 자세 인식 알고리즘에 입력하여 커터니언 값을 예측할 수 있다.

[0018] 또한, 본 발명의 실시예에 따르면, 물체 자세 추정 시스템을 이용한 물체 자세 추정 방법은 RGB-D 카메라로부터 탐지 대상이 되는 물체를 촬영하여 획득한 RGB-D 이미지를 입력받는 단계, 상기 RGB-D 이미지를 학습이 완료된 물체 탐지 알고리즘에 입력하고, 상기 물체 탐지 알고리즘에 의해 탐지된 물체에 바운딩 박스를 획득하고, 상기 획득한 바운딩 박스에 대응되는 제1 RGB 이미지 및 제1 Depth 이미지를 획득하는 단계, 상기 제1 RGB 이미지 및 제1 Depth 이미지를 학습이 완료된 자세 인식 알고리즘에 입력하여 특징맵을 획득하고, 상기 획득한 특징맵으로부터 물체의 자세를 추정하는 단계, 그리고 탐지된 물체에 대한 이미지상의 좌표와 자세를 추정하여 획득한 좌표를 출력하는 단계를 포함한다.

발명의 효과

[0020] 이와 같이 본 발명에 따르면, 물체 인식(object detection)과 자세 인식(pose estimation) 단계를 분리하여 제

공하므로 데이터 처리 단계의 복잡성을 줄이고, 구현 및 활용의 용이성을 높일 수 있다.

[0021] 또한, 본 발명에 따르면, 로봇 제어 분야 이외에 가상 시뮬레이션, 가상현실(VR), 증강현실(AR) 등에서 물체의 자세 인식을 위해 사용할 수 있으며, 기존 자세 인식 네트워크에 비해 경량화 되었다는 장점을 고려하여 모바일 분야에서도 활용될 수 있다.

도면의 간단한 설명

[0023] 도 1은 본 발명의 실시예에 따른 물체 탐지 및 자세 추정 시스템을 설명하기 위한 구성도이다.

도 2는 본 발명의 실시예에 따른 물체 자세 추정 시스템을 이용한 물체 자세 추정 방법을 설명하기 위한 순서도이다.

도 3은 도 2의 s210단계를 설명하기 위한 예시도이다.

도 4는 도2의 S250단계를 설명하기 위한 예시도이다.

발명을 실시하기 위한 구체적인 내용

[0024] 이하 첨부된 도면을 참조하여 본 발명에 따른 바람직한 실시예를 상세히 설명하기로 한다. 이 과정에서 도면에 도시된 선들의 두께나 구성요소의 크기 등은 설명의 명료성과 편의상 과장되게 도시되어 있을 수 있다.

[0025] 또한 후술되는 용어들은 본 발명에서의 기능을 고려하여 정의된 용어들로서, 이는 사용자, 운용자의 의도 또는 관례에 따라 달라질 수 있다. 그러므로 이러한 용어들에 대한 정의는 본 명세서 전반에 걸친 내용을 토대로 내려져야 할 것이다.

[0026] 이하에서는 도 1을 이용하여 본 발명의 실시예에 따른 물체 탐지 및 자세 추정 시스템에 대해 더욱 상세하게 설명한다.

[0027] 도 1은 본 발명의 실시예에 따른 물체 탐지 및 자세 추정 시스템을 설명하기 위한 구성도이다.

[0028] 도 1에 도시된 바와 같이, 물체 탐지 및 자세 추정 시스템(100)은 이미지 입력부(110), 제1 학습부(120), 제2 학습부(130), 물체 탐지부(140), 자세 추정부(150) 및 출력부(160)를 포함한다.

[0029] 먼저, 이미지 입력부(110)는 RGB-D 카메라에 의해 획득한 RGB-D 이미지를 입력받는다. 이를 다시 설명하면 이미지 입력부(110)는 RGB-D 카메라로부터 자세를 추정하고자 하는 물체를 촬영한 RGB-D 이미지를 입력받는다.

[0030] 그 다음, 제1 학습부(120)는 물체의 위치를 탐지하기 위하여 인공지능을 기반으로 하는 물체 탐지 알고리즘을 구축하고, 물체의 위치가 다양한 복수의 RGB-D 이미지를 이용하여 생성된 제1 데이터셋을 이용하여 구축된 물체 탐지 알고리즘을 학습시킨다.

[0031] 여기서 제1 데이터셋은 CAD 모델을 이용하여 생성된 물체를 RGB-D 카메라로 촬영하여 생성된 RGB-D 이미지를 통해 생성된다.

[0032] 제2 학습부(130)는 제1 학습부(120)에 의해 탐지된 물체의 위치 정보를 이용하여 인공지능을 기반으로 하는 자세 인식 알고리즘을 학습시킨다. 이를 다시 설명하면, 제2 학습부(130)는 제1 학습부(120)에 입력된 RGB-D 이미지를 촬영한 RGB-D 카메라의 위치 정보와 제1 학습부(120)를 통해 획득한 물체의 위치 정보 및 물체에 대한 6자유도 자세 정보를 이용하여 제2 데이터셋을 생성한다. 그리고, 제2 학습부(130)는 생성된 제2 데이터셋을 이용하여 자세 인식 알고리즘을 학습시킨다.

[0033] 물체 탐지부(140)는 입력된 RGB-D 이미지를 제1 학습부(120)에 의해 학습이 완료된 물체 탐지 알고리즘에 입력한다. 그러면, 물체 탐지 알고리즘은 입력된 RGB-D 이미지에 포함된 물체의 위치 및 종류를 분석하고, 분석된 결과를 바운딩 박스로 출력시킨다. 그리고, 물체 탐지부(140)는 출력된 바운딩 박스에 대응되는 부분을 크롭핑하여 제1 RGB 이미지 및 제1 Depth 이미지를 획득한다.

[0034] 그 다음, 자세 추정부(150)는 획득한 제1 RGB 이미지 및 제1 Depth 이미지를 전처리한 다음, 학습이 완료된 자세 인식 알고리즘에 입력한다. 그러면, 자세 인식 알고리즘은 전처리된 제1 RGB 이미지와 제1 Depth 이미지로부터 커터니언 값을 예측한다.

- [0035] 마지막으로 출력부(160)는 커터니언의 각 원소인 X,Y,Z 및 W를 출력한다.
- [0037] 이하에서는 도 2 내지 도 4를 이용하여 본 발명의 실시예에 따른 물체 자세 추정 시스템(100)을 이용한 물체 자세 추정 방법에 대해 더욱 상세하게 설명한다.
- [0038] 도 2는 본 발명의 실시예에 따른 물체 자세 추정 시스템을 이용한 물체 자세 추정 방법을 설명하기 위한 순서도이다.
- [0039] 먼저, 본 발명의 실시예에 따른 물체 자세 추정 시스템을 이용한 물체 자세 추정 방법은 기계학습을 기반으로 하는 알고리즘을 학습하는 단계와, 구축된 알고리즘을 통하여 물체의 자세를 추정하는 단계를 포함한다.
- [0040] 도 2에 도시된 바와 같이, 학습하는 단계는 S210 내지 S230 단계를 포함하고, 추정하는 단계는 S240 및 S270 단계를 포함한다.
- [0041] 이미지 입력부(110)는 CAD 모델에 의해 생성된 물체에 대한 RGB 이미지를 입력받는다(S210).
- [0042] 도 3은 도 2의 s210단계를 설명하기 위한 예시도이다.
- [0043] 도3에 도시된 바와 같이, 사용자는 시물레이션이 가능한 가상공간을 구성하고, 산업현장 또는 일상 생활에서 사용하는 물체의 CAD 모델로 생성한다. 생성된 CAD 모델은 RGB 이미지로 변환되며, 이미지 입력부(110)는 변환된 RGB 이미지를 입력받는다.
- [0044] 한편, CAD 모델은 다양한 자세를 가지고 있는 것으로 구현되며, 가상의 RGB-D 카메라가 존재하는 것으로 구현될 수 있다.
- [0045] 따라서, 사용자는 CAD 모델을 이용하여 획득한 RGB 이미지와 2차원의 상자 정보, 가상의 카메라에 대한 위치 정보 및 6자 유도 자세 정보를 입력한다.
- [0046] 여기서, 2차원의 상자 정보는 물체에 대한 종류, 물체의 위치, 바운딩 박스에 대한 좌표 중에서 적어도 하나를 포함한다.
- [0047] 이미지 입력부(110)는 입력된 RGB 이미지와 2차원의 상자 정보를 제1 학습부(120)에 전달한다.
- [0048] 그러면, 제1 학습부(120)는 전달받은 RGB 이미지와 2차원의 상자 정보를 이용하여 물체 탐지 알고리즘을 학습시킨다(S220).
- [0049] 이를 다시 설명하면, 제1 학습부(120)는 RGB 이미지를 입력데이터로 하고, 2차원의 상자 정보를 출력데이터로 하는 데이터셋을 생성한다. 그 다음, 제1 학습부(120)는 YOLO 네트워크를 활용하여 구축된 물체 탐지 알고리즘에 생성된 데이터셋을 입력한다. 그러면, 물체 탐지 알고리즘은 입력데이터와 출력데이터를 이용하여 지도 학습된다. 즉, 물체 탐지 알고리즘은 입력된 RGB 이미지에 포함된 물체의 종류 및 위치를 탐지하고, 탐지가 완료되면 물체의 주변을 감싸는 바운딩 박스를 출력하도록 한다.
- [0050] S220 단계가 완료되면, 제2 학습부(130)는 카메라의 위치 정보, 물체의 위치 정보 및 6자 유도 자세 정보를 이용하여 자세 인식 알고리즘을 학습시킨다(S230).
- [0051] 부연하자면, 제2 학습부(130)는 카메라 위치 정보 및 물체의 위치 정보를 입력데이터로 하고, 6자 유도 자세 정보를 출력데이터로 하는 데이터셋을 생성한다. 그리고, 제2 학습부(130)는 생성된 데이터셋을 합성곱 신경망의 구조를 가지는 자세 인식 알고리즘에 입력하여 지도 학습시킨다.
- [0052] S210단계 내지 S230단계를 이용하여 물체 탐지 알고리즘 및 자세 인식 알고리즘에 대한 학습을 완료시킨다.
- [0053] 그 다음, 이미지 입력부(110)는 자세를 추정하고자 하는 물체를 촬영하여 획득한 RGB-D 이미지를 입력받는다(S240).
- [0054] 이를 다시 설명하면, 사용자는 RGB-D 카메라를 이용하여 자세를 추정하고자 하는 물체를 촬영한다. 그러면, RGB-D카메라는 생성된 RGB-D 이미지를 이미지 입력부(110)에 전달한다.
- [0055] S240단계가 완료되면, 물체 탐지부(140)는 입력된 RGB-D 이미지를 학습이 완료된 물체 탐지 알고리즘에 입력하여 물체를 탐지한다(S250).
- [0056] 물체 탐지부(140)는 입력된 RGB 이미지를 물체 탐지 알고리즘에 입력한다. 그러면, 물체 탐지 알고리즘은 입력

된 RGB 이미지에 포함된 물체 각각에 대한 종류 및 위치를 탐지한다. 탐지가 완료되면, 물체 탐지 알고리즘은 물체에 해당되는 바운딩박스를 출력한다.

[0057] 그러면, 물체 탐지부(140)는 출력된 바운딩 박스의 좌표를 이용하여 바운딩 박스에 해당하는 부분만을 크롭핑한다. 그리고, 물체 탐지부(140)는 크롭핑된 부분에 대해 제1 RGB 이미지 및 제1 Depth 이미지를 생성한다.

[0058] 한편, 바운딩 박스는 동일한 비율의 이미지를 자세 인식 알고리즘에 전달하기 위하여 X축과 Y축 중에서 더 긴 축을 기준으로 정사각형 형태로 변형된다.

[0059] 그리고, 물체 탐지부(140)는 생성된 제1 RGB 이미지 및 제1 Depth 이미지와 함께 객체의 종류에 대한 정보를 결과로 출력한다. 이때, 출력되는 객체가 복수 개일 경우에는 각 객체의 정보를 순차적으로 출력한다.

[0060] S250단계가 완료되면, 자세 추정부(150)는 제1 RGB 이미지 및 제1 Depth 이미지를 이용하여 해당되는 물체의 자세를 추정한다(S260).

[0061] 먼저, 자세 추정부(150)는 제1 RGB 이미지 및 제1 Depth 이미지를 자세 인식 알고리즘에 입력한다.

[0062] 도 4는 도2의 S250단계를 설명하기 위한 예시도이다.

[0063] 도 4에 도시된 바와 같이, 본 발명의 실시예에 따른 자세 인식 알고리즘은 두 단계를 이용하여 해당되는 물체의 자세를 추정한다.

[0064] 도 4의 (a)에 도시된 바와 같이, 첫번째 단계는 합성곱 신경망을 이용하여 입력된 제1 RGB 이미지 및 제1 Depth 이미지로부터 특징을 추출한다.

[0065] 부연하자면, 자세 추정부(150)는 제1 RGB 이미지와 제1 Depth 이미지로부터 각각의 특징을 추출하기 위하여, 독립된 합성곱 신경망에 자세 추정부(150)는 제1 RGB 이미지와 제1 Depth 이미지를 각각 입력한다.

[0066] 이때, 입력되는 제1 RGB 이미지 및 제1 Depth 이미지의 크기는 64*64 픽셀로 구성되며, 전처리가 완료된 상태이다. 즉, 제1 RGB 이미지는 각 픽셀의 색상 값을 255로 나누는 방법으로 픽셀 값을 0과 1 사이로 정규화한다. 그리고, 제1 Depth 이미지는 이미지 내의 최대 깊이값과 최소값의 차이를 이용하여 정규화를 수행한다.

[0067] 한편, 합성곱 신경망은 여섯 개의 합성곱 층(convolution layer)으로 구성되며, 필터(filter)의 수는 16에서 512까지 2배씩 증가된다. 그리고, 커널(kernel)의 크기는 3X3으로 모두 동일하며, 제로 패딩(zero-padding)을 사용한다. 활성화함수(activation function)는 SELU(Scaled Exponential Linear Unit)를 사용하였으며, 풀링 층(pooling layer)은 이차원 최대 풀링(2D max pooling)을 사용한다. 그리고, 풀링 크기는 2X2 크기를 가진다. 또한, 합성곱 신경망은 합성곱 층 이후에는 과적합(over-fitting)을 막기 위해 배치 정규화(batch normalization)를 사용한다.

[0068] 그 결과 각각의 합성곱 신경망은 1X1X512의 크기를 가지는 특징맵을 출력한다.

[0069] 그 다음, 도 4의 (b)에 도시된 바와 같이, 자세 인식 알고리즘의 두 번째 단계는 특징 추출 단계에서 추출된 특징을 기반으로 물체의 자세를 나타내는 쿼터니언 값을 예측하는 단계다.

[0070] 쿼터니언이란, 3차원 공간에서 회전을 표현하는 4개의 값으로 구성된 복소수 형태의 체계로, 회전 연산 시 행렬에 비해 연산이 적고 저장공간이 절약된다는 장점을 가진다.

[0071] 합성곱 신경망을 통해 출력되는 특징맵은 원소별(element-wise) 덧셈을 통해 하나의 특징맵으로 결합된다. 그리고, 통합된 특징 맵은 플래튼(flatten) 층을 통해 512차원의 벡터로 변환되어 쿼터니언 예측 완전 연결 네트워크의 입력이 된다. 완전 연결(fully connected, FC) 네트워크는 두 개의 은닉 층(hidden layer)으로 구성되며, 각 층은 256, 128의 크기를 가진다. 이때, 활성화함수로는 SELU를 사용하며, 마지막 출력 층(output layer)은 선형(linear) 활성화함수를 이용한 회귀(regression)를 쿼터니언의 각 원소인 (x, y, z, w)를 추론하고 출력한다.

[0072] 이때, 자세 인식 알고리즘은 하기의 수학적 식 1에 기재된 손실함수를 이용하여 쿼터니언 값을 추론한다.

수학식 1

$$L(q, \hat{q}) = \min \left(\left\| \frac{q}{\|q\|} - \frac{\hat{q}}{\|\hat{q}\|} \right\|_1, \left\| \frac{q}{\|q\|} + \frac{\hat{q}}{\|\hat{q}\|} \right\|_1 \right)$$

[0073]

[0074]

여기서, q 와 $\hat{q}$ 는 각각 정답과 네트워크가 예측한 쿼터니언 값을 의미한다

[0075]

그 다음, 자세 추정부(150)는 자세 인식 알고리즘을 통해 예측된 쿼터니언 값을 출력부(160)에 전달한다.

[0076]

그러면, 출력부(160)는 물체의 위치 좌표와 6 자유도 자세를 결과값을 출력한다(S270).

[0077]

즉 출력부(160)는 물체 탐지 알고리즘을 통해 탐지된 물체의 위치를 이미지상의 x, y, z 좌표로 출력하고, 자세 인식 알고리즘을 통해 예측된 x, y, z 축에 대한 오일러 각도를 결합한 6자유도 자세를 결과로 출력한다.

[0078]

이와 같이 본 발명에 따르면, 물체 인식(object detection)과 자세 인식(pose estimation) 단계를 분리하여 제공하므로 데이터 처리 단계의 복잡성을 줄이고, 구현 및 활용의 용이성을 높일 수 있다.

[0079]

또한, 본 발명에 따르면, 로봇 제어 분야 이외에 가상 시뮬레이션, 가상현실(VR), 증강현실(AR) 등에서 물체의 자세 인식을 위해 사용할 수 있으며, 기존 자세 인식 네트워크에 비해 경량화 되었다는 장점을 고려하여 모바일 분야에서도 활용될 수 있다.

[0081]

본 발명은 도면에 도시된 실시예를 참고로 하여 설명되었으나 이는 예시적인 것에 불과하며, 당해 기술이 속하는 분야에서 통상의 지식을 가진 자라면 이로부터 다양한 변형 및 균등한 타 실시예가 가능하다는 점을 이해할 것이다. 따라서 본 발명의 진정한 기술적 보호범위는 아래의 특허청구범위의 기술적 사상에 의하여 정해져야 할 것이다.

부호의 설명

[0083]

100 : 물체 자세 추정 시스템

110 : 이미지 입력부

120 : 제1 학습부

130 : 제2학습부

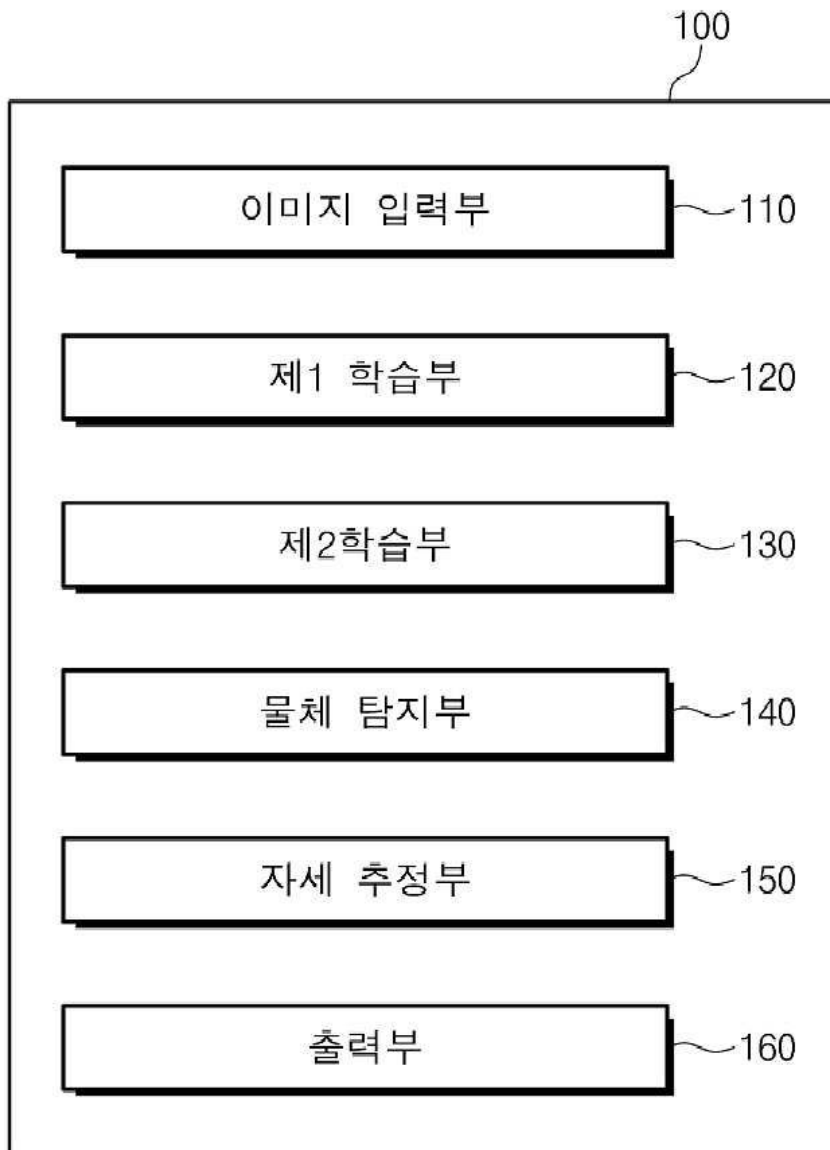
140 : 물체 탐지부

150 : 자세 추정부

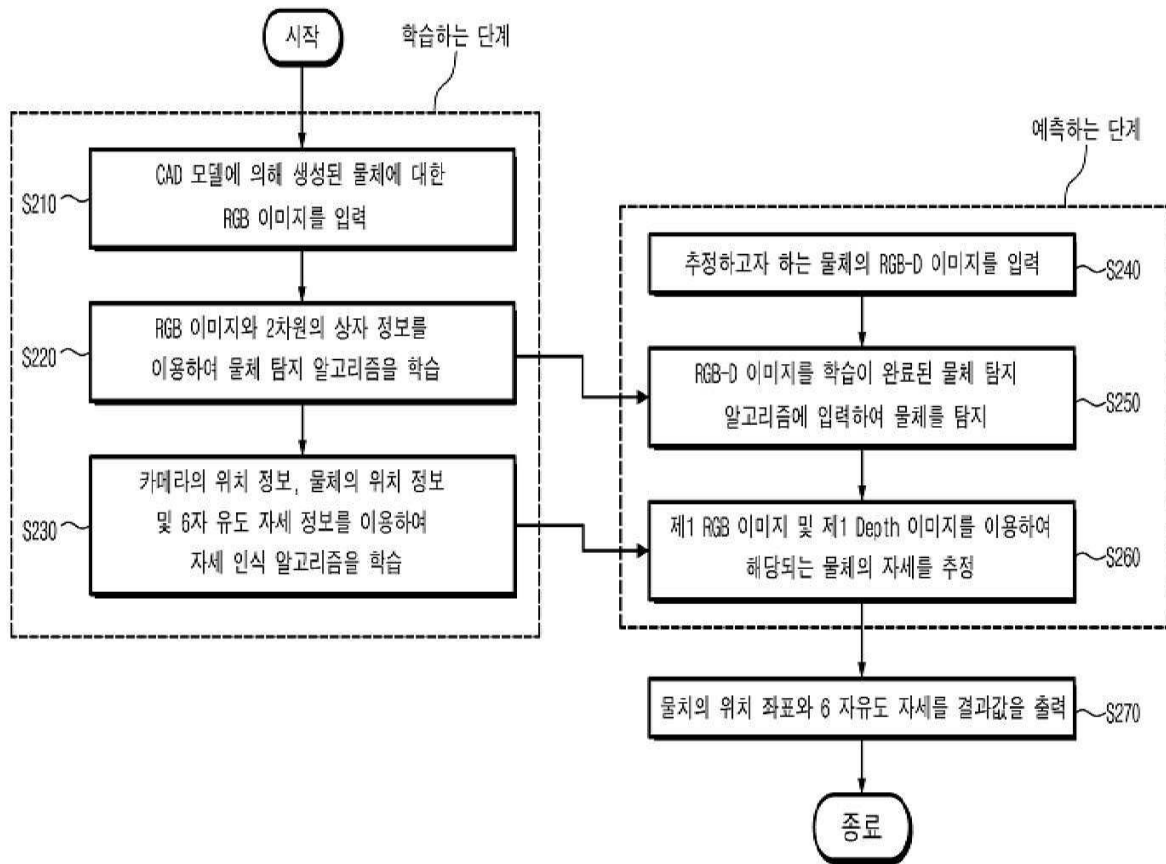
160 : 출력부

도면

도면1



도면2



도면3

모델명	CAD 데이터	생성된 이미지의 예 (별간 상자 물체 영역)	물체 영역 RGB 이미지	물체 영역 깊이 이미지	자세 (회전각도) 가시화
DR-30					
B6000					
LM10					
Chair					
Sofa					

도면4

