



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2024-0146324
(43) 공개일자 2024년10월08일

(51) 국제특허분류(Int. Cl.)

G06T 7/50 (2017.01) G06N 3/0464 (2023.01)
G06N 3/08 (2023.01) G06N 3/09 (2023.01)
G06T 5/00 (2024.01) G06V 10/40 (2022.01)
G06V 10/70 (2022.01)

(52) CPC특허분류

G06T 7/50 (2017.01)
G06N 3/0464 (2023.01)

(21) 출원번호 10-2023-0041146

(22) 출원일자 2023년03월29일

심사청구일자 2023년03월29일

(71) 출원인

충남대학교산학협력단

대전광역시 유성구 대학로 99 (궁동, 충남대학교)

(72) 발명자

조동현

대전광역시 유성구 대학로 99, 충남대 공과대학
2호관 407호 (궁동)

김진욱

대전광역시 유성구 대학로 99, 충남대 공과대학
2호관 472호 (궁동)

(74) 대리인

제일특허법인(유)

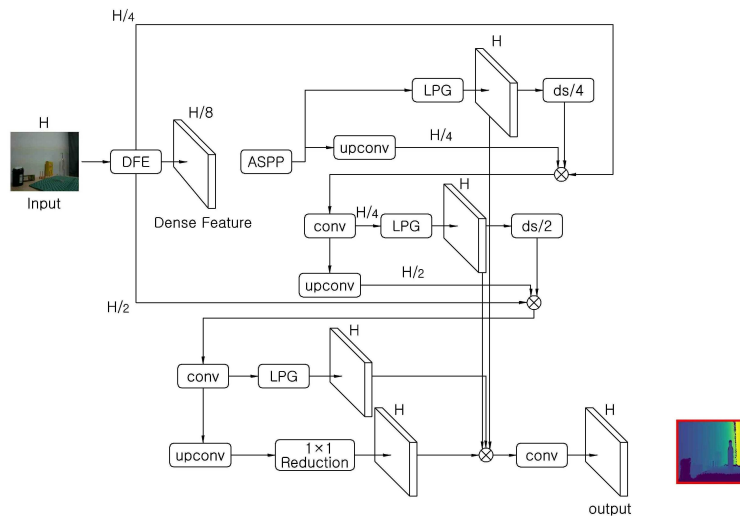
전체 청구항 수 : 총 18 항

(54) 발명의 명칭 투명 객체의 깊이 추정 장치 및 방법, 투명 객체의 깊이 추정을 위한 학습 방법

(57) 요약

본 발명은 투명 객체의 깊이 추정 장치에 관한 것으로, 투명 객체의 깊이 영상을 획득하는 획득부; 상기 투명 객체와 동일한 형태를 가지면서 다른 질감을 갖는 비투명 객체의 깊이 영상을 기초로 상기 투명 객체의 깊이를 추정하도록 학습된 학습 모델을 포함하는 저장부; 및 상기 학습 모델을 기초로 상기 비투명 객체의 깊이 영상으로부터 특징을 추출하고, 상기 특징을 기초로 상기 투명 객체의 깊이 영상의 깊이 맵(map)을 생성하여 상기 투명 객체의 깊이를 추정하는 처리부;를 포함할 수 있다.

대표도



(52) CPC특허분류

G06N 3/084 (2023.01)

G06N 3/09 (2023.01)

G06T 5/00 (2024.01)

G06V 10/40 (2023.08)

G06V 10/768 (2023.08)

G06T 2207/10028 (2013.01)

G06T 2207/20081 (2013.01)

G06T 2207/20084 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711158713
과제번호	2022R1C1C1009334
부처명	과학기술정보통신부
과제관리(전문)기관명	한국연구재단
연구사업명	개인기초연구(과기정통부)
연구과제명	투명 객체의 3차원 정보 추정 및 활용에 관한 연구
기 여 율	1/1
과제수행기관명	충남대학교
연구기간	2022.03.01 ~ 2025.02.28

명세서

청구범위

청구항 1

투명 객체의 깊이 영상을 획득하는 획득부;

상기 투명 객체와 동일한 형태를 가지면서 다른 질감을 갖는 비투명 객체의 깊이 영상을 기초로 상기 투명 객체의 깊이를 추정하도록 학습된 학습 모델을 포함하는 저장부; 및

상기 학습 모델을 기초로 상기 비투명 객체의 깊이 영상으로부터 특징을 추출하고, 상기 특징을 기초로 상기 투명 객체의 깊이 영상의 깊이 맵(map)을 생성하여 상기 투명 객체의 깊이를 추정하는 처리부;를 포함하는

투명 객체의 깊이 추정 장치.

청구항 2

제 1 항에 있어서,

상기 학습 모델은,

상기 비투명 객체의 깊이를 추정하도록 학습되는 제1 학습 모델; 및

상기 제1 학습 모델을 이용하여 상기 투명 객체의 깊이를 추정하도록 학습되는 제2 학습 모델;을 포함하는

투명 객체의 깊이 추정 장치.

청구항 3

제 2 항에 있어서,

상기 제1 학습 모델은 선생 학습 모델이고, 상기 제2 학습 모델은 학생 학습 모델이며, 상기 선생 학습 모델과 상기 학생 학습 모델은 지식 증류(knowledge distillation) 기법을 기초로 학습되는

투명 객체의 깊이 추정 장치.

청구항 4

제 1 항에 있어서,

상기 처리부는,

상기 투명 객체의 깊이 영상에 대해 학습용 데이터 셋을 구성하는 데이터 셋 구성부;

상기 학습 모델을 기초로 상기 학습용 데이터 셋과 상기 비투명 객체의 깊이 영상으로부터 특징을 추출하는 인코딩부; 및

상기 특징을 기초로 상기 투명 객체의 깊이 영상의 깊이 맵을 생성하여 상기 투명 객체의 깊이를 추정하는 디코딩부;를 포함하는

투명 객체의 깊이 추정 장치.

청구항 5

제 4 항에 있어서,

상기 인코딩부는,

상기 특징을 적용하기 위한 태스크에 대응하는 특징 맵을 생성하는 밀집 특징(dense feature) 추출부; 및

상기 특징 맵에 대한 조밀 이미지를 예측하기 위한 글로벌 특징 맵을 추출하는 맥락 정보(contextual information) 추출부;를 포함하는

투명 객체의 깊이 추정 장치.

청구항 6

제 5 항에 있어서,

상기 디코딩부는,

상기 글로벌 특징 맵의 해상도를 출력 해상도로 복원하는 지역 평면(local planar) 지도부; 및

상기 출력 해상도로 복원된 상기 글로벌 특징 맵을 디코더 계층에서 출력되는 동일한 크기의 특징 맵에 연결하여 정보 손실을 최소화하는 스킵(skip) 연결부;를 포함하는

투명 객체의 깊이 추정 장치.

청구항 7

깊이 추정 장치에서 수행되는 투명 객체의 깊이 추정 방법에 있어서,

상기 투명 객체의 깊이 영상을 획득하는 단계;

기 저장된 학습 모델을 기초로 상기 깊이 영상으로부터 특징을 추출하는 단계; 및

상기 특징을 기초로 상기 깊이 영상의 깊이 맵을 생성하여 상기 투명 객체의 깊이를 추정하는 단계;를 포함하되,

상기 기 저장된 학습 모델은,

상기 투명 객체와 동일한 형태를 가지면서 다른 질감을 갖는 비투명 객체의 깊이 영상을 기초로 상기 투명 객체의 깊이를 추정하도록 학습된 학습 모델인

투명 객체의 깊이 추정 방법.

청구항 8

제 7 항에 있어서,

상기 투명 객체의 깊이를 추정하는 단계는,

상기 비투명 객체의 깊이를 추정하도록 제1 학습 모델을 학습시키는 단계; 및

상기 제1 학습 모델을 이용하여 상기 투명 객체의 깊이를 추정하도록 제2 학습 모델을 학습시키는 단계;를 포함하는

투명 객체의 깊이 추정 방법.

청구항 9

제 8 항에 있어서,

상기 제1 학습 모델은 선생 학습 모델이고, 상기 제2 학습 모델은 학생 학습 모델이며, 상기 선생 학습 모델과 상기 학생 학습 모델은 지식 증류 기법을 기초로 학습되는

투명 객체의 깊이 추정 방법.

청구항 10

제 7 항에 있어서,

상기 투명 객체의 깊이를 추정하는 단계는,

상기 투명 객체의 깊이 영상에 대해 학습용 데이터 셋을 구성하는 단계;

상기 학습 모델을 기초로 상기 학습용 데이터 셋과 상기 비투명 객체의 깊이 영상으로부터 특징을 추출하는 단계; 및

상기 특징을 기초로 상기 투명 객체의 깊이 영상의 깊이 맵을 생성하여 상기 투명 객체의 깊이를 추정하는 단계;를 포함하는

투명 객체의 깊이 추정 방법.

청구항 11

제 10 항에 있어서,

상기 특징을 추출하는 단계는,

상기 특징을 적용하기 위한 태스크에 대응하는 특징 맵을 생성하기 위한 밀집 특징을 추출하는 단계; 및

상기 특징 맵에 대한 조밀 이미지를 예측하기 위한 글로벌 특징 맵을 추출하기 위한 맥락 정보를 추출하는 단계;를 포함하는

투명 객체의 깊이 추정 방법.

청구항 12

제 11 항에 있어서,

상기 투명 객체의 깊이를 추정하는 단계는,

상기 글로벌 특징 맵의 해상도를 출력 해상도로 복원하는 단계; 및

상기 출력 해상도로 복원된 상기 글로벌 특징 맵을 디코더 계층에서 출력되는 동일한 크기의 특징 맵에 연결하여 정보 손실을 최소화하는 스킵 연결 단계;를 포함하는

투명 객체의 깊이 추정 방법.

청구항 13

깊이 추정 장치에서 수행되는 투명 객체의 깊이 추정을 위한 학습 방법에 있어서,

투명 객체의 깊이 영상과, 상기 투명 객체와 동일한 형태를 가지면서 다른 질감을 갖는 비투명 객체의 깊이 영상을 획득하는 단계;

상기 비투명 객체의 깊이 영상의 특징을 기초로 상기 비투명 객체의 깊이 추정을 위한 제1 학습 모델을 학습시키는 단계; 및

상기 제1 학습 모델을 이용하여 상기 투명 객체의 깊이 영상의 특징을 기초로 상기 투명 객체의 깊이 추정을 위한 제2 학습 모델을 학습시키는 단계;를 포함하는

투명 객체의 깊이 추정을 위한 학습 방법.

청구항 14

제 13 항에 있어서,

상기 제1 학습 모델은 선생 학습 모델이고, 상기 제2 학습 모델은 학생 학습 모델이며, 상기 선생 학습 모델과 상기 학생 학습 모델은 지식 증류 기법을 기초로 학습되는

투명 객체의 깊이 추정을 위한 학습 방법.

청구항 15

제 14 항에 있어서,

상기 선생 학습 모델과 상기 학생 학습 모델에 대하여 상기 선생 학습 모델의 출력을 소프트 레이블(soft label)로 하여 증류 손실(distillation loss)을 이용한 손실 함수를 계산하는 단계;

상기 학생 학습 모델의 출력을 하드 레이블(hard label)로 하여 학생 손실(student loss)을 이용한 손실 함수를 계산하는 단계; 및

상기 증류 손실을 이용한 손실 함수와 상기 학생 손실을 이용한 손실 함수의 가중 합을 기초로 최종 손실 함수를 정의하는 단계;를 더 포함하는

투명 객체의 깊이 추정을 위한 학습 방법.

청구항 16

제 15 항에 있어서,

상기 최종 손실 함수에 대해 역전파(backpropagation) 알고리즘을 적용하여 상기 제1 및 제2 학습 모델을 갱신하는 단계를 더 포함하는

투명 객체의 깊이 추정을 위한 학습 방법.

청구항 17

컴퓨터 프로그램을 저장하고 있는 컴퓨터 판독 가능 기록매체로서,

상기 컴퓨터 프로그램은,

깊이 추정 장치에서 수행되는 투명 객체의 깊이 추정 방법을 프로세서가 수행하도록 하기 위한 명령어를 포함하고,

상기 투명 객체의 깊이 영상을 획득하는 단계;

기 저장된 학습 모델을 기초로 상기 깊이 영상으로부터 특징을 추출하는 단계; 및

상기 특징을 기초로 상기 깊이 영상의 깊이 맵을 생성하여 상기 투명 객체의 깊이를 추정하는 단계;를 포함하되,

상기 기 저장된 학습 모델은,

상기 투명 객체와 동일한 형태를 가지면서 다른 질감을 갖는 비투명 객체의 깊이 영상을 기초로 상기 투명 객체의 깊이를 추정하도록 학습된 학습 모델인

컴퓨터 판독 가능한 기록매체.

청구항 18

컴퓨터 판독 가능 기록매체에 저장된 컴퓨터 프로그램으로서,

상기 컴퓨터 프로그램은,

깊이 추정 장치에서 수행되는 투명 객체의 깊이 추정을 위한 학습 방법을 프로세서가 수행하도록 하기 위한 명령어를 포함하고,

상기 방법은,

상기 투명 객체의 깊이 영상과, 상기 투명 객체와 동일한 형태를 가지면서 다른 질감을 갖는 비투명 객체의 깊이 영상을 획득하는 단계;

상기 비투명 객체의 깊이 영상의 특징을 기초로 상기 비투명 객체의 깊이 추정을 위한 제1 학습 모델을 학습시키는 단계; 및

상기 제1 학습 모델을 이용하여 상기 투명 객체의 깊이 영상의 특징을 기초로 상기 투명 객체의 깊이 추정을 위한 제2 학습 모델을 학습시키는 단계;를 포함하는

기록매체에 저장된 컴퓨터 프로그램.

발명의 설명

기술 분야

[0001] 본 발명은 영상의 깊이 추정 기술과 관련이 있고, 더 구체적으로 딥러닝(deep-learning)을 이용한 객체의 깊이 추정 기술과 관련이 있다.

배경 기술

[0003] 딥러닝 기술의 도래와 더불어 객체 검출(detection), 이미지 분할(segmentation), 거리 측정(depth estimation) 등 다양한 이미지 신호처리(ISP, image signal processing) 분야에서 획기적인 성장을 거두고 있으며, 최근에는 대부분의 이미지 신호처리에 이러한 딥러닝을 범용적으로 채택하여 사용하고 있는 추세이다.

[0004] 이 중, 거리 측정 기술은 크게 두 가지로 분류되는데 한 장의 이미지에서 깊이를 추정하는 단안 깊이 추정(monocular depth estimation) 모델과 여러 장의 이미지를 활용하여 깊이를 추정하는 스테레오 깊이 추정(stereo depth estimation) 모델로 활발한 연구가 진행되고 있다.

[0005] 하지만, 위의 두 모델 모두 기본적으로 투명한 객체나 액체와 같은 빛의 굴절과 반사에 영향을 받는 환경을 고려하지 않고, 이미지 내의 객체 표면이 램버시안 반사(lambertian reflectance)를 따른다는 가정이 전제되어 있다. 이러한 가정이 특정한 환경에서는 좋은 깊이 추정 성능을 가질 수 있지만, 일반적으로 일상생활에서 자주 볼 수 있는 플라스틱, 유리, 일회용 용기, 액체 등이 포함된 영상에서는 모델이 거리 측정 학습을 하는 데에 어려움을 줄 수 있다.

선행기술문헌

특허문헌

[0007] (특허문헌 0001) 공개특허공보 제10-2022-0080696호 (2022년06월14일 공개)

발명의 내용

해결하려는 과제

- [0008] 본 발명의 실시예에서는, 지도학습(supervision)을 기반으로 하는 컨볼루션 신경망(convolutional neural network) 기반의 영상 깊이 추정 기술을 제안하고자 한다.
- [0009] 본 발명의 실시예에서는, 투명 객체의 특징을 추출하고 깊이를 추정하는데 적합한 딥러닝 네트워크 기반의 영상 깊이 추정 기술을 제안하고자 한다.
- [0010] 본 발명이 해결하고자 하는 과제는 상기에서 언급한 것으로 제한되지 않으며, 언급되지 않은 또 다른 해결하고자 하는 과제는 아래의 기재들로부터 본 발명이 속하는 통상의 지식을 가진 자에 의해 명확하게 이해될 수 있을 것이다.

과제의 해결 수단

- [0012] 본 발명의 실시예에 따르면, 투명 객체의 깊이 영상을 획득하는 획득부; 상기 투명 객체와 동일한 형태를 가지면서 다른 질감을 갖는 비투명 객체의 깊이 영상을 기초로 상기 투명 객체의 깊이를 추정하도록 학습된 학습 모델을 포함하는 저장부; 및 상기 학습 모델을 기초로 상기 비투명 객체의 깊이 영상으로부터 특징을 추출하고, 상기 특징을 기초로 상기 투명 객체의 깊이 영상의 깊이 맵(map)을 생성하여 상기 투명 객체의 깊이를 추정하는 처리부;를 포함하는 투명 객체의 깊이 추정 장치를 제공할 수 있다.
- [0013] 여기서, 상기 학습 모델은, 상기 비투명 객체의 깊이를 추정하도록 학습되는 제1 학습 모델; 및 상기 제1 학습 모델을 이용하여 상기 투명 객체의 깊이를 추정하도록 학습되는 제2 학습 모델;을 포함할 수 있다.
- [0014] 또한, 상기 제1 학습 모델은 선생 학습 모델이고, 상기 제2 학습 모델은 학생 학습 모델이며, 상기 선생 학습 모델과 상기 학생 학습 모델은 지식 증류(knowledge distillation) 기법을 기초로 학습될 수 있다.
- [0015] 또한, 상기 처리부는, 상기 투명 객체의 깊이 영상에 대해 학습용 데이터 셋을 구성하는 데이터 셋 구성부; 상기 학습 모델을 기초로 상기 학습용 데이터 셋과 상기 비투명 객체의 깊이 영상으로부터 특징을 추출하는 인코딩부; 및 상기 특징을 기초로 상기 투명 객체의 깊이 영상의 깊이 맵을 생성하여 상기 투명 객체의 깊이를 추정하는 디코딩부;를 포함할 수 있다.
- [0016] 또한, 상기 인코딩부는, 상기 특징을 적용하기 위한 태스크에 대응하는 특징 맵을 생성하는 밀집 특징(dense feature) 추출부; 및 상기 특징 맵에 대한 조밀 이미지를 예측하기 위한 글로벌 특징 맵을 추출하는 맥락 정보(contextual information) 추출부;를 포함할 수 있다.
- [0017] 또한, 상기 디코딩부는, 상기 글로벌 특징 맵의 해상도를 출력 해상도로 복원하는 지역 평면(local planar) 지도부; 및 상기 출력 해상도로 복원된 상기 글로벌 특징 맵을 디코더 계층에서 출력되는 동일한 크기의 특징 맵에 연결하여 정보 손실을 최소화하는 스킵(skip) 연결부;를 포함할 수 있다.
- [0018] 본 발명의 실시예에 따르면, 깊이 추정 장치에서 수행되는 투명 객체의 깊이 추정 방법에 있어서, 상기 투명 객체의 깊이 영상을 획득하는 단계; 기 저장된 학습 모델을 기초로 상기 깊이 영상으로부터 특징을 추출하는 단계; 및 상기 특징을 기초로 상기 깊이 영상의 깊이 맵을 생성하여 상기 투명 객체의 깊이를 추정하는 단계;를 포함하되, 상기 기 저장된 학습 모델은, 상기 투명 객체와 동일한 형태를 가지면서 다른 질감을 갖는 비투명 객체의 깊이 영상을 기초로 상기 투명 객체의 깊이를 추정하도록 학습된 학습 모델인 투명 객체의 깊이 추정 방법을 제공할 수 있다.
- [0019] 여기서, 상기 투명 객체의 깊이를 추정하는 단계는, 상기 비투명 객체의 깊이를 추정하도록 제1 학습 모델을 학습시키는 단계; 및 상기 제1 학습 모델을 이용하여 상기 투명 객체의 깊이를 추정하도록 제2 학습 모델을 학습시키는 단계;를 포함할 수 있다.
- [0020] 또한, 상기 제1 학습 모델은 선생 학습 모델이고, 상기 제2 학습 모델은 학생 학습 모델이며, 상기 선생 학습 모델과 상기 학생 학습 모델은 지식 증류 기법을 기초로 학습될 수 있다.
- [0021] 또한, 상기 투명 객체의 깊이를 추정하는 단계는, 상기 투명 객체의 깊이 영상에 대해 학습용 데이터 셋을 구성하는 단계; 상기 학습 모델을 기초로 상기 학습용 데이터 셋과 상기 비투명 객체의 깊이 영상으로부터 특징을

추출하는 단계; 및 상기 특징을 기초로 상기 투명 객체의 깊이 영상의 깊이 맵을 생성하여 상기 투명 객체의 깊이를 추정하는 단계;를 포함할 수 있다.

[0022] 또한, 상기 특징을 추출하는 단계는, 상기 특징을 적용하기 위한 태스크에 대응하는 특징 맵을 생성하기 위한 밀집 특징을 추출하는 단계; 및 상기 특징 맵에 대한 조밀 이미지를 예측하기 위한 글로벌 특징 맵을 추출하기 위한 맥락 정보를 추출하는 단계;를 포함할 수 있다.

[0023] 또한, 상기 투명 객체의 깊이를 추정하는 단계는, 상기 글로벌 특징 맵의 해상도를 출력 해상도로 복원하는 단계; 및 상기 출력 해상도로 복원된 상기 글로벌 특징 맵을 디코더 계층에서 출력되는 동일한 크기의 특징 맵에 연결하여 정보 손실을 최소화하는 스킵 연결 단계;를 포함할 수 있다.

[0024] 본 발명의 실시예에 따르면, 깊이 추정 장치에서 수행되는 투명 객체의 깊이 추정을 위한 학습 방법에 있어서, 투명 객체의 깊이 영상과, 상기 투명 객체와 동일한 형태를 가지면서 다른 질감을 갖는 비투명 객체의 깊이 영상을 획득하는 단계; 상기 비투명 객체의 깊이 영상의 특징을 기초로 상기 비투명 객체의 깊이 추정을 위한 제1 학습 모델을 학습시키는 단계; 및 상기 제1 학습 모델을 이용하여 상기 투명 객체의 깊이 영상의 특징을 기초로 상기 투명 객체의 깊이 추정을 위한 제2 학습 모델을 학습시키는 단계;를 포함하는 투명 객체의 깊이 추정을 위한 학습 방법을 제공할 수 있다.

[0025] 여기서, 상기 제1 학습 모델은 선생 학습 모델이고, 상기 제2 학습 모델은 학생 학습 모델이며, 상기 선생 학습 모델과 상기 학생 학습 모델은 지식 증류 기법을 기초로 학습될 수 있다.

[0026] 또한, 상기 방법은, 상기 선생 학습 모델과 상기 학생 학습 모델에 대하여 상기 선생 학습 모델의 출력을 소프트 레이블(soft label)로 하여 증류 손실(distillation loss)을 이용한 손실 함수를 계산하는 단계; 상기 학생 학습 모델의 출력을 하드 레이블(hard label)로 하여 학생 손실(student loss)을 이용한 손실 함수를 계산하는 단계; 및 상기 증류 손실을 이용한 손실 함수와 상기 학생 손실을 이용한 손실 함수의 가중 합을 기초로 최종 손실 함수를 정의하는 단계;를 더 포함할 수 있다.

[0027] 또한, 상기 방법은, 상기 최종 손실 함수에 대해 역전파(backpropagation) 알고리즘을 적용하여 상기 제1 및 제2 학습 모델을 갱신하는 단계를 더 포함할 수 있다.

[0028] 본 발명의 실시예에 따르면, 컴퓨터 프로그램을 저장하고 있는 컴퓨터 판독 가능 기록매체로서, 상기 컴퓨터 프로그램은, 깊이 추정 장치에서 수행되는 투명 객체의 깊이 추정 방법을 프로세서가 수행하도록 하기 위한 명령어를 포함하고, 상기 투명 객체의 깊이 영상을 획득하는 단계; 기 저장된 학습 모델을 기초로 상기 깊이 영상으로부터 특징을 추출하는 단계; 및 상기 특징을 기초로 상기 깊이 영상의 깊이 맵을 생성하여 상기 투명 객체의 깊이를 추정하는 단계;를 포함하되, 상기 기 저장된 학습 모델은, 상기 투명 객체와 동일한 형태를 가지면서 다른 질감을 갖는 비투명 객체의 깊이 영상을 기초로 상기 투명 객체의 깊이를 추정하도록 학습된 학습 모델일 수 있다.

[0029] 본 발명의 실시예에 따르면, 컴퓨터 판독 가능 기록매체에 저장된 컴퓨터 프로그램으로서, 상기 컴퓨터 프로그램은, 깊이 추정 장치에서 수행되는 투명 객체의 깊이 추정을 위한 학습 방법을 프로세서가 수행하도록 하기 위한 명령어를 포함하고, 상기 방법은, 상기 투명 객체의 깊이 영상과, 상기 투명 객체와 동일한 형태를 가지면서 다른 질감을 갖는 비투명 객체의 깊이 영상을 획득하는 단계; 상기 비투명 객체의 깊이 영상의 특징을 기초로 상기 비투명 객체의 깊이 추정을 위한 제1 학습 모델을 학습시키는 단계; 및 상기 제1 학습 모델을 이용하여 상기 투명 객체의 깊이 영상의 특징을 기초로 상기 투명 객체의 깊이 추정을 위한 제2 학습 모델을 학습시키는 단계;를 포함할 수 있다.

발명의 효과

[0031] 본 발명의 실시예에 의하면, 비투명 객체에 대한 특징을 추출하고 영상의 깊이를 추정함으로써, 자동차 자율 주행에서 탐지가 어려운 빛의 반사 및 굴절로 인한 타 운송 수단의 탐지 및 주변 물체 탐지에 활용이 용이하고, 로봇 파지(robotic grasping) 시스템에서 특히 투명하거나 질감이 특이한 객체에 대한 이슈가 많이 존재하고 있는데 해당 문제에 긍정적인 효과를 제공할 수 있다. 모델적인 측면으로는 지식 증류 기법(knowledge distillation)을 적용함으로써 모델의 일반화 능력(generalization ability)을 향상시키며 상대적으로 무거운 모델에 속하는 깊이 추정 모델을 경량화하여 실제 환경에서의 사용에 적합한 작은 모델을 구성하여 휴대폰에 사용되는 카메라 등 접근이 용이한 모델을 구축할 수 있다.

도면의 간단한 설명

[0033]

도 1은 본 발명의 실시예에 따른 투명 객체의 깊이 추정 장치(100)의 기능을 개략적으로 설명하기 위한 블록도이다.

도 2는 도 1의 깊이 추정 장치(100)의 저장부(120) 내의 깊이 추정 프로그램(122)을 나타낸 개념도이다.

도 3은 도 2의 깊이 추정 프로그램(122) 내의 전이 학습 모델(124)의 구체적인 기능 블록도이다.

도 4는 도 1의 처리부(130)의 구체적인 기능 블록도이다.

도 5는 도 4의 처리부(130) 내의 인코딩부(134)의 구체적인 기능 블록도이다.

도 6은 도 4의 처리부(130) 내의 디코딩부(136)의 구체적인 기능 블록도이다.

도 7은 도 3의 전이 학습 모델(124)을 예시적으로 설명하는 개념도이다.

도 8은 도 1의 처리부(130)를 예시적으로 설명하는 개념도이다.

도 9는 도 6의 지역 평면 지도부(136a)의 구성을 예시적으로 설명하는 개념도이다.

발명을 실시하기 위한 구체적인 내용

[0034]

본 발명의 이점 및 특징, 그리고 그것들을 달성하는 방법은 첨부되는 도면과 함께 상세하게 후술되어 있는 실시예들을 참조하면 명확해질 것이다. 그러나 본 발명은 이하에서 개시되는 실시예들에 한정되는 것이 아니라 다양한 형태로 구현될 수 있으며, 단지 본 실시예들은 본 발명의 개시가 완전하도록 하고, 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자에게 발명의 범주를 완전하게 알려주기 위해 제공되는 것이며, 본 발명의 범주는 청구항에 의해 정의될 뿐이다.

[0035]

본 발명의 실시예들을 설명함에 있어서 공지 기능 또는 구성에 대한 구체적인 설명은 본 발명의 실시예들을 설명함에 있어 실제로 필요한 경우 외에는 생략될 것이다. 그리고 후술되는 용어들은 본 발명의 실시예에서의 기능을 고려하여 정의된 용어들로서 이는 사용자, 운용자의 의도 또는 관례 등에 따라 달라질 수 있다. 그러므로 그 정의는 본 명세서 전반에 걸친 내용을 토대로 내려져야 할 것이다.

[0037]

본 발명은 종래의 거리 측정 기술들이 내포하고 있지만 여러 이유로 잘 다루어지지 않은 문제점들을 해결하고자 하는 것으로, 실사 투명 객체 영상과 동일한 비(非)투명 객체 영상의 특징(feature)을 이용하여 다양한 환경에서의 거리 추정과 객체 검출 모델을 생성하는 딥러닝 기반 방법에 관한 것이다.

[0038]

첫째, 동일한 실사 투명 객체 영상과 비투명 객체 영상 확보의 어려움 때문에 대부분의 거리 추정 기술은 합성 영상 생성 도구(synthetic data generator)를 통해 모델이 학습할 정답 영상을 생성한다. 하지만 이와 같은 영상 데이터는 실제 영상이 아닌 가상의 데이터로 모델이 실제 영상에 대한 추론을 진행할 때 명확한 한계점을 가지게 된다. 그러므로, 정확하고 합리적인 거리 측정 모델을 학습시키기 위해 실사 데이터를 학습 데이터로 활용하는 것이 바람직하다.

[0039]

둘째, 위에서 언급한 문제점에 의해 램버시안 반사를 따르지 않는 실사 데이터의 특징 분석(feature analysis)에 대한 연구가 부족하다. 비투명 객체는 저 레벨 특징(low level feature)으로부터 가장자리(edge)나 질감(texture)과 같은 특징을 추출할 수 있고 고 레벨 특징(high level feature)으로 갈수록 점점 더 추상적인 특징을 뽑아낼 수 있다는 것이 증명되었다. 투명 객체 또한 위와 같은 경향을 보이지만 일반 객체와는 확연히 다른 특성을 가지는 성질 때문에 또 다른 의미 있는 특징을 실험 및 의미 있는 특징 추출 방법을 고안하여 목표(target) 영상의 특징 분석이 필요하며, 이를 잘 활용하면 지식 증류(knowledge distillation), 도메인 적응(domain adaptation), 전이학습(transfer learning) 등 여러 분야에 성능의 향상뿐만 아니라 특정 업무(task)에 대한 새로운 방향성을 제시할 수 있다.

[0040]

따라서, 지금까지 다루지 않은 거리 추정 모델에 대한 문제점들을 보완하고 나아가 거리 추정에 국한되지 않고 객체 검출, 이미지 분할 등 다양한 도메인으로 이와 같은 문제를 해결할 신경망 기반 방법 및 장치가 필요하다.

[0041]

이에 본 발명의 실시예에서는 지도학습(supervision)을 기반으로 하는 컨볼루션 신경망(convolutional neural

network) 기반의 영상 깊이 추정 기법을 제안하고자 한다.

- [0042] 거리 측정 모델의 학습 데이터(training data)의 정답 데이터(ground truth)는 키넥트 카메라(Kinect camera)를 통해 획득될 수 있다. 이러한 키넥트 카메라는 물체의 표면에서 적외선이 반사하여 센서로 감지되어야만 거리 측정이 가능하기 때문에 물체 표면에서 빛이 투과되거나 굴절되는 투명한 객체의 거리를 측정하는 데에 어려움이 있다. 그러므로, 투명 객체와 동일한 모양을 가지면서 다른 질감(texture)을 갖는 물체를 같이 준비하여 카메라는 고정된 후 3원색 이미지(RGB image)와 3차원 깊이 이미지(3D depth image)를 따로 촬영하여 투명 객체의 정답 깊이 데이터를 확보할 수 있다.
- [0043] 같은 모양을 갖는 투명 객체와 질감을 살린 객체(textured object)는 서로 다르지만 많은 유사한 특징 정보(feature information)를 가지고 있다. 이를 활용하여 지식 증류(knowledge distillation) 기법을 통해 출력과 특징맵(feature map)을 선생 모델(teacher model)로부터 학생 모델(student model)을 학습시켜 보다 다양한 환경에 강인(robust)하고 가벼우면서 좋은 성능을 가지는 모델을 만들 수 있다. 그리고 이 특징을 활용하여 두 모델의 특징 차이의 분석에 기반하여 이미지 내의 투명하거나 특별한 질감을 갖는 객체의 검출을 좌표 정답(coordinate label) 없이 약 지도학습(weakly supervised-learning)을 적용할 수 있다.
- [0045] 이하, 첨부된 도면을 참조하여 본 발명의 실시예에 대해 상세히 설명하기로 한다.
- [0046] 본 발명의 실시예에서 정의되는 용어들은 아래와 같이 정의될 수 있다.
- [0047] "인코딩"은 입력 영상을 입력으로 받아 밀집 특징 추출부 및 디코딩 부에 입력으로 넣어줄 "인코딩 특징 맵"을 컨볼루션 신경망으로 추출하여 출력하는 부분이다.
- [0048] "디코딩"은 인코딩 특징맵 및 인코딩 부의 컨볼루션 레이어(convolution layer)의 각 출력으로부터의 스킵 연결을 U-Net 구조로 연결하여 "디코딩 특징 맵"을 컨볼루션 신경망으로 추출하는 부분을 의미한다. 디코딩 특징맵은 맥락 정보 추정부, 지역 평면 지도부의 입력을 생성하기 위한 부분으로 작동한다.
- [0049] "스킵 연결(skip-connection)"이란 딥러닝 모델을 학습할 때 모델의 깊이가 깊어질수록 앞부분의 정보가 손실되는데 이를 보완하기 위해 사용되는 구조이다.
- [0050] "U-Net"은 보통 모델의 최종 출력이 픽셀 단위의 분류(이미지 분류, 이미지 세그멘테이션, 깊이 추정 등)를 수행할 때 사용되는 인코더-디코더 구조이다.
- [0051] "전이 학습(transfer learning)"이란 이미지 분류 문제를 해결하기 위해 방대한 데이터를 학습한 네트워크를 다른 데이터셋 혹은 다른 문제(task)에 적용시켜 문제를 해결하는 것을 의미한다.
- [0052] "맥락 정보(contexture information)"란 여러 번의 컨볼루션과 풀링(pooling)을 거치면서 발생하는 정보를 의미하며 위의 단계가 진행될수록 디테일한 정보가 줄어들고 특성이 점점 추상화되는 문제점이 있다. 이를 보완하기 위해 아트로스 공간 피라미드 풀링(atrous spatial pyramid pooling, ASPP)를 사용하여 다양한 시야각을 가지게 된다.
- [0053] "지식 증류(knowledge distillation)"란 미리 잘 학습된 큰 네트워크(teacher network)의 지식을 실제로 사용하고자 하는 작은 네트워크(student network)에게 전달하여 모델 경량화 혹은 모델 성능 향상에 도움을 줄 수 있는 기법이다.
- [0054] 먼저, 도 1은 본 발명의 실시예에 따른 투명 객체의 깊이 추정 장치(100)의 기능을 개략적으로 설명하기 위한 블록도이다.
- [0055] 도 1에 도시한 바와 같이, 투명 객체의 깊이 추정 장치(100)는 획득부(110), 저장부(120) 및 처리부(130)를 포함할 수 있다.
- [0056] 획득부(110)는 투명 객체의 깊이 영상을 획득할 수 있다. 이러한 획득부(110) 투명 객체의 깊이 영상뿐만 아니라, 투명 객체와 동일한 모양을 가지면서 다른 질감(texture)을 갖는 비투명 객체의 깊이 영상을 획득할 수 있다. 획득부(110)를 통해 획득되는 투명 객체의 깊이 영상과 비투명 객체의 깊이 영상은 처리부(130)로 제공될 수 있다. 이를 위해, 본 발명의 실시예에 따른 깊이 추정 장치(100)는 투명 객체 및 비투명 객체의 3원색 이미지와 3차원 깊이 영상을 촬영하기 위한 촬영 수단과 연결될 수 있다. 여기서, 3차원 깊이 영상은, 예를 들어 키넥트 카메라를 통해 획득될 수 있다. 3차원 깊이 정보를 측정하는 키넥트 카메라는 컬러 센서(color sensor),

적외선 방출기(IR emitter), 적외선 깊이 센서(IR depth sensor) 등으로 구성된다. 키넥트 카메라의 동작 원리는 적외선 방출기 부분에서 방출된 적외선을 타겟 물체 표면에서 반사시켜 키넥트 카메라의 적외선 깊이 센서로 감지되는 시간을 측정하여 타겟 물체와 카메라 사이의 거리를 계산하는 비행시간법(time-of-flight, TOF)을 이용한다. 본 발명의 실시예에서는 투명 객체와 동일한 모양을 가지면서 다른 질감(texture)을 갖는 물체를 같이 준비하고, 키넥트 카메라를 고정시킨 후 3원색 이미지(RGB image)와 3차원 깊이 이미지(3D depth image)를 따로 촬영하여 투명 객체의 정답 깊이 데이터를 확보할 수 있다.

[0057] 저장부(120)는 투명 객체와 동일한 형태를 가지면서 다른 질감을 갖는 비투명 객체의 깊이 영상을 기초로 투명 객체의 깊이를 추정하도록 학습된 학습 모델을 포함할 수 있다. 이러한 저장부(120)는 본 발명의 실시예에 따라 객체의 깊이를 추정하기 위한 프로그램과, 이러한 프로그램을 실행시킬 수 있는 정보들을 저장할 수 있다. 이러한 저장부(120)는 컴퓨터 시스템에 의해 읽혀질 수 있는 데이터가 저장되는 모든 종류의 기록 장치를 포함할 수 있으며, 특정 기록 장치에 한정되지 않는다. 저장부(120) 내의 프로그램 및 실행 정보들은 필요에 따라 처리부(130)에 의해 로드될 수 있다.

[0058] 처리부(130)는 학습 모델을 기초로 비투명 객체의 깊이 영상으로부터 특징을 추출하고, 추출되는 특징을 기초로 투명 객체의 깊이 영상의 깊이 맵(map)을 생성하여 투명 객체의 깊이를 추정할 수 있다. 특히, 처리부(130)는 후술하는 깊이 추정 프로그램을 기초로, 획득부(110)를 통해 획득되는 투명 객체의 깊이 영상에 대한 깊이 추정 결과를 출력하도록 학습 모델을 학습시키거나 학습 모델을 통해 투명 객체의 깊이 추정 결과를 출력할 수 있다.

[0060] 도 2는 도 1의 깊이 추정 장치(100)의 저장부(120) 내의 깊이 추정 프로그램(122)을 나타낸 개념도이다.

[0061] 깊이 추정 프로그램(122)은 전이 학습 모델(124)을 포함할 수 있으며, 처리부(130)의 제어 하에 전이 학습 모델(124)을 로드하고, 로드된 전이 학습 모델(124)을 실행시키거나 갱신할 수 있다. 본 발명의 실시예에 따른 깊이 추정 프로그램(122)은 획득부(110)를 통해 투명 객체 및 비투명 객체의 촬영 이미지가 획득되면, 투명 객체의 깊이 영상의 깊이 맵을 생성하여 투명 객체의 깊이를 추정하도록 프로그램된 명령어들을 포함하는 소프트웨어를 의미할 수 있다.

[0062] 이러한 전이 학습 모델(124)은 비투명 객체의 깊이를 추정하기 위한 학습 모델과 투명 객체의 깊이를 추정하기 위한 학습 모델을 포함할 수 있다.

[0064] 도 3은 도 2의 깊이 추정 프로그램(122) 내의 전이 학습 모델(124)의 구체적인 기능 블록도이다.

[0065] 도 3에 도시한 바와 같이, 전이 학습 모델(124)은 비투명 객체 깊이 추정 학습 모델(124a), 투명 객체 깊이 추정 학습 모델(124b) 및 지식 증류부(124c)를 포함할 수 있다.

[0066] 비투명 객체 깊이 추정 학습 모델(124a)은 투명 객체와 동일한 형태를 가지면서 다른 질감을 갖는 비투명 객체의 깊이 영상을 기초로 비투명 객체의 깊이를 추정하도록 학습될 수 있다. 이러한 비투명 객체 깊이 추정 학습 모델(124a)은 투명 객체의 깊이를 추정하기 위한 선행 학습 모델로서, 예컨대 선생 학습 모델(teacher leaning model)로 명명될 수 있다.

[0067] 투명 객체 깊이 추정 학습 모델(124b)은 비투명 객체 깊이 추정 학습 모델(124a)을 이용하여 투명 객체의 깊이를 추정하도록 학습될 수 있다. 이러한 투명 객체 깊이 추정 학습 모델(124b)은 비투명 객체의 깊이를 추정한 결과를 기초로 투명 객체의 깊이를 추정하기 위한 후행 학습 모델로서, 예컨대 학생 학습 모델(student learning model)로 명명될 수 있다.

[0068] 지식 증류부(124c)는 비투명 객체 깊이 추정 학습 모델(124a)과 투명 객체 깊이 추정 학습 모델(124b)에서 구축한 선생 학습 모델과 학생 학습 모델에 대하여, 선생 학습 모델의 출력을 소프트 라벨(soft label)로 증류 손실(distillation loss)을 사용하여 손실 함수를 정의하고, 학생 학습 모델의 출력을 하드 라벨(hard label)로 학생 손실(student loss)을 사용하여 손실 함수를 정의할 수 있다. 이러한 증류 손실은 아래 [수학식 1]로 표현되고, 학생 손실은 아래 [수학식 2]로 각각 표현될 수 있다.

수학식 1

$$2\alpha T^2 \mathcal{L}_{\text{CE}}(\sigma(\frac{\mathbf{z}_S}{T}), \sigma(\frac{\mathbf{z}_T}{T}))$$

수학식 2

$$(1 - \alpha) \mathcal{L}_{\text{CE}}(\mathbf{y}, \sigma(\mathbf{z}_S))$$

이를 통해 지식 증류부(124c)는 정의된 손실함수를 가중합을 통해 최종 손실함수를 정의하여 전이 학습 모델(124)이 학습을 진행하게 된다. 이러한 학습을 위한 최종 손실 함수(L_t)는 역전파(back-propagation) 알고리즘이 적용될 수 있으며, 역전파 알고리즘을 통해 전이 학습 모델(124)이 갱신될 수 있다. 손실 함수는 아래 [수학식 3]과 같이 표현될 수 있다.

수학식 3

$$L = \alpha \sqrt{D(g)}$$

도 4는 도 1의 처리부(130)의 구체적인 기능 블록도이다.

도 4에 도시한 바와 같이, 처리부(130)는 데이터 셋 구성부(132), 인코딩부(134) 및 디코딩부(136)를 포함할 수 있다.

데이터 셋 구성부(132)는 투명 객체의 깊이 영상에 대해 학습용 데이터 셋을 구성할 수 있다. 딥러닝 네트워크를 학습시키기 위한 데이터 셋 제작 과정을 포함하며, 이 과정은 단일 깊이 추정 장비인 키넥트 카메라를 통해 한 장의 이미지 당 한 장의 깊이 영상을 생성할 수 있다.

인코딩부(134)는 전이 학습 모델(124)을 기초로 데이터 셋 구성부(132)를 통해 구성된 학습용 데이터 셋과 비투명 객체의 깊이 영상으로부터 특징을 추출할 수 있다.

즉, 인코딩부(134)는 대용량 데이터 셋 이미지 넷(image-net)에서 이미지 분류를 위해 학습된 사전 학습 모델(pre-trained model)을 고정하여(freeze) 입력 영상의 특징 또는 특징맵을 추출할 수 있다.

디코딩부(136)는 인코딩부(134)에서 추출된 특징 또는 특징맵을 기초로 상기 투명 객체의 깊이 영상의 깊이 맵을 생성하여 투명 객체의 깊이를 추정할 수 있다. 즉, 디코딩부(136)는 인코딩부(134)로부터 추출된 특징맵을 더 풍부한 특징맵으로 만듦과 동시에, 원하고자 하는 최종 출력 깊이맵(depth map)으로 해상도를 키워 조밀 이미지 예측을 수행할 수 있다.

도 5는 도 4의 처리부(130) 내의 인코딩부(134)의 구체적인 기능 블록도이다.

도 5에 도시한 바와 같이, 인코딩부(134)는 밀집 특징(dense feature) 추출부(134a) 및 맥락 정보(contextual information) 추출부(134b)를 포함할 수 있다.

밀집 특징 추출부(134a)는 입력 영상의 특징을 적용하기 위한 태스크에 대응하는 특징 맵을 생성할 수 있다. 이러한 밀집 특징 추출부(134a)는 맥락 정보 추출부(134b)와 유기적으로 작동할 수 있다.

맥락 정보 추출부(134b)는 특징 맵에 대한 조밀 이미지를 예측하기 위한 글로벌 특징 맵을 추출할 수 있다. 이러한 맥락 정보 추출부(134b)는 다양한 시야각을 사용하여 특징맵을 보다 글로벌한 특징맵으로 추출할 수 있게 한다.

[0086] 도 6은 도 4의 처리부(130) 내의 디코딩부(136)의 구체적인 기능 블록도이다.

[0087] 도 6에 도시한 바와 같이, 디코딩부(136)는 지역 평면 지도부(136a) 및 스킵 연결부(136b)를 포함할 수 있다.

[0088] 지역 평면 지도부(136a)는 인코딩부(134)의 맥락 정보 추출부(134b)에서 추출된 글로벌 특징 맵의 해상도를 출력 해상도로 복원할 수 있다. 이러한 지역 평면 지도부(136a)는 기존에 일반적으로 사용하는 업-샘플링(up-sampling) 기법과 다르게 특징맵을 출력 해상도로 복구하기 위해 1x1 컨볼루션과 활성화 함수를 적절하게 사용하여 효율적인 해상도 복원을 진행한다. 해당 과정에서 총 4개의 계수를 추정할 수 있으며, 추정 공식은 n_1 , n_2 및 n_3 와, 원본 및 평면의 직선거리를 통해 n_4 를 추정하여 산출될 수 있다. 이러한 추정 공식의 계수는 아래 [수학식 4]로 표현될 수 있다.

수학식 4

$$\tilde{c}_i = \frac{n_4}{n_1 u_i + n_2 v_i + n_3}$$

[0089]

[0090] [수학식 4]에서 n_1 , n_2 및 n_3 는 아래 [수학식 5]로 표현될 수 있다.

수학식 5

$$n_1 = \sin(\theta) \cos(\phi), n_2 = \sin(\theta) \sin(\phi), n_3 = \cos(\theta)$$

[0091]

[0092] 스킵 연결부(136b)는 지역 평면 지도부(136a)를 통해 출력 해상도로 복원된 글로벌 특징맵을 디코더 계층에서 출력되는 동일한 크기의 특징맵에 연결하여 정보 손실을 최소화할 수 있다. 스킵 연결부(136b)를 통해 출력되는 최종 값이 추정 결과는 아래 [수학식 6]으로 표현될 수 있다.

수학식 6

$$\tilde{d} = f(W_1 \tilde{c}^{1 \times 1} + W_2 \tilde{c}^{2 \times 2} + W_3 \tilde{c}^{4 \times 4} + W_4 \tilde{c}^{8 \times 8})$$

[0093]

[0095] 도 7은 도 3의 전이 학습 모델(124)을 예시적으로 설명하는 개념도이다.

[0096] 도 7에 예시한 바와 같이, 기본적으로 선생 학습 모델의 출력을 소프트 라벨로 증류 손실을 사용하여 손실 함수를 정의하고, 학생 학습 모델의 출력을 하드 라벨로 학생 손실을 사용하여 손실 함수를 정의할 수 있다.

[0098] 도 8은 도 1의 처리부(130)를 예시적으로 설명하는 개념도이다.

[0099] 도 8에서 DFE(dense feature extractor)는 도 4의 인코딩부(134)에 해당되며, ASPP(atrous spatial pyramid pooling) 이후의 구성이 도 4의 디코딩부(136)에 해당될 수 있다.

[0100] 디코딩부(136)는 LPG(local planar guidance), ds(down-sampling), conv(convolution), upconv(up convolution) 등을 포함할 수 있다.

[0102] 도 9는 도 6의 지역 평면 지도부(136a)의 구성을 예시적으로 설명하는 개념도이다.

[0103] 도 9에 예시한 바와 같이, 지역 평면 지도부(136a)는 특징맵을 출력 해상도로 복구하기 위해 1x1 컨볼루션과

활성화 함수를 적절하게 사용하여 효율적인 해상도 복원을 진행한다. 해당 과정에서 총 4개의 계수를 추정할 수 있으며, 추정 공식은 n_1 , n_2 및 n_3 와, 원본 및 평면의 직선거리를 통해 n_4 를 추정하여 산출될 수 있다.

- [0105] 이상 설명한 바와 같은 본 발명의 실시예에 의하면, 딥러닝 분야에서 지금까지 잘 다루어진 일반적인(비투명) 객체에 대한 조밀 이미지 예측(dense image prediction)을 기반으로 일상생활에서 흔하게 볼 수 있는 플라스틱, 유리, 질감이 특별한 객체에 대한 조밀 이미지 예측을 수행한다. 타겟 객체에 대한 관련 연구가 아직 활발하게 수행되지 않아 본 연구에서 해당 주제에 대한 베이스라인(baseline) 구성 및 초기 환경 세팅을 구축하여 새로운 주제에 대한 탐구와 분석을 수행한다.
- [0106] 활용적인 측면에서 해당 발명의 효과로는 자동차 자율주행에서 탐지가 어려운 빛의 반사 및 굴절로 인한 타 운송수단의 탐지 및 주변 물체 탐지에 활용될 수 있고, 로봇과지(robotic grasping) 시스템에서 특히 투명하거나 질감이 특이한 객체에 대한 이슈가 많이 존재하고 있는데 해당 문제에 긍정적인 효과를 가져다 줄 것으로 보인다.
- [0107] 모델적인 측면에서 발명의 효과로는 지식 증류 기법(knowledge distillation)을 적용함으로써 모델의 일반화 능력(generalization ability)을 향상시키며 상대적으로 무거운 모델에 속하는 깊이 추정 모델을 경량화하여 실제 환경에서의 사용에 적합한 작은 모델을 구성하여 휴대폰에 사용되는 카메라 등 접근이 용이한 모델을 구축할 수 있다.
- [0108] 연구적인 측면에서 발명의 효과로는 본 발명의 타겟 객체에 대한 심층적인 분석과 다양한 기법에 대한 특징 변화를 분석함으로써 비(非)투명 객체와 투명 객체의 특징 비교 분석이 가능하다. 관련된 분석으로는 아직 충분한 자료가 확보되지 않았으며, 본 발명을 통해 발견되는 다양한 기법에 대한 특징 변화 분석을 다른 업무(task)에 적용하여 타 연구에도 효율적인 모델을 구축하기 위한 자료로 활용이 가능하다.
- [0110] 한편, 첨부된 블록도의 각 블록과 흐름도의 각 단계의 조합들은 컴퓨터 프로그램 인스트럭션들에 의해 수행될 수도 있다. 이들 컴퓨터 프로그램 인스트럭션들은 범용 컴퓨터, 특수용 컴퓨터 또는 기타 프로그램 가능한 데이터 프로세싱 장비의 프로세서에 탑재될 수 있으므로, 컴퓨터 또는 기타 프로그램 가능한 데이터 프로세싱 장비의 프로세서를 통해 수행되는 그 인스트럭션들이 블록도의 각 블록에서 설명된 기능들을 수행하는 수단을 생성하게 된다.
- [0111] 이들 컴퓨터 프로그램 인스트럭션들은 특정 방식으로 기능을 구현하기 위해 컴퓨터 또는 기타 프로그램 가능한 데이터 프로세싱 장비를 지향할 수 있는 컴퓨터 이용 가능 또는 컴퓨터 판독 가능 기록매체(또는 메모리) 등에 저장되는 것도 가능하므로, 그 컴퓨터 이용 가능 또는 컴퓨터 판독 가능 기록매체(또는 메모리)에 저장된 인스트럭션들은 블록도의 각 블록에서 설명된 기능을 수행하는 인스트럭션 수단을 내포하는 제조 품목을 생산하는 것도 가능하다.
- [0112] 그리고, 컴퓨터 프로그램 인스트럭션들은 컴퓨터 또는 기타 프로그램 가능한 데이터 프로세싱 장비 상에 탑재되는 것도 가능하므로, 컴퓨터 또는 기타 프로그램 가능한 데이터 프로세싱 장비 상에서 일련의 동작 단계들이 수행되어 컴퓨터로 실행되는 프로세스를 생성해서 컴퓨터 또는 기타 프로그램 가능한 데이터 프로세싱 장비를 수행하는 인스트럭션들은 블록도의 각 블록에서 설명된 기능들을 실행하기 위한 단계들을 제공하는 것도 가능하다.
- [0113] 또한, 각 블록은 특정된 논리적 기능(들)을 실행하기 위한 적어도 하나 이상의 실행 가능한 인스트럭션들을 포함하는 모듈, 세그먼트 또는 코드의 일부를 나타낼 수 있다. 또, 몇 가지 대체 실시 예들에서는 블록들에서 언급된 기능들이 순서를 벗어나서 발생하는 것도 가능함을 주목해야 한다. 예컨대, 잇달아 도시되어 있는 두 개의 블록들은 사실 실질적으로 동시에 수행되는 것도 가능하고 또는 그 블록들이 때때로 해당하는 기능에 따라 역순으로 수행되는 것도 가능하다.

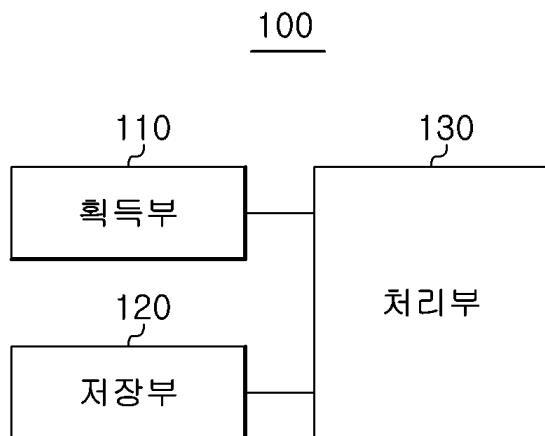
부호의 설명

- [0115] 100: 투명 객체의 깊이 추정 장치

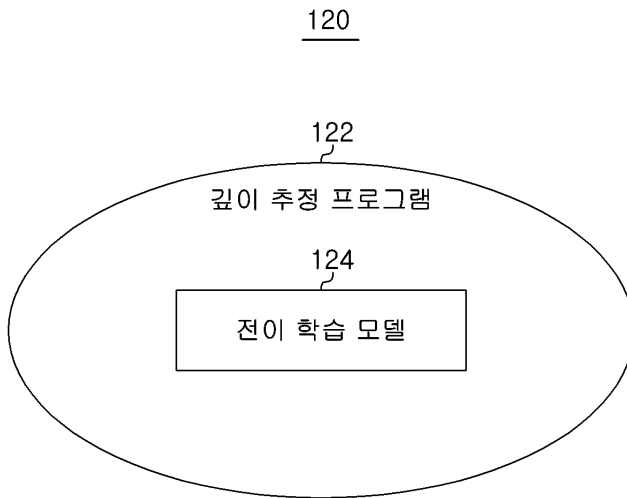
- 110: 획득부
- 120: 저장부
- 122: 깊이 추정 프로그램
- 124: 전이 학습 모델
- 124a: 비투명 객체 깊이 추정 학습 모델
- 124b: 투명 객체 깊이 추정 학습 모델
- 124c: 지식 증류부
- 130: 처리부
- 132: 데이터 셋 구성부
- 134: 인코딩부
- 134a: 밀집 특징 추출부
- 134b: 맥락 정보 추출부
- 136: 디코딩부
- 136a: 지역 평면 지도부
- 136b: 스킵 연결부

도면

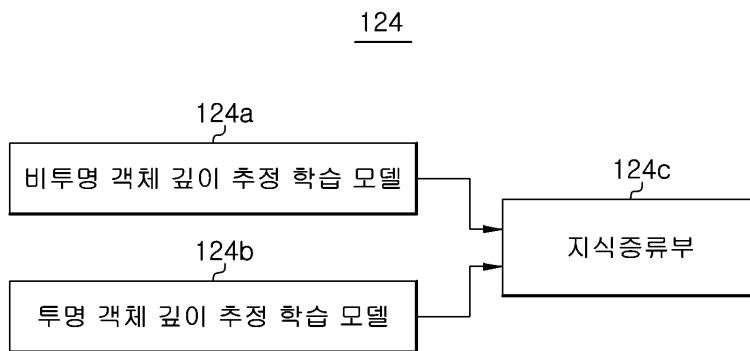
도면1



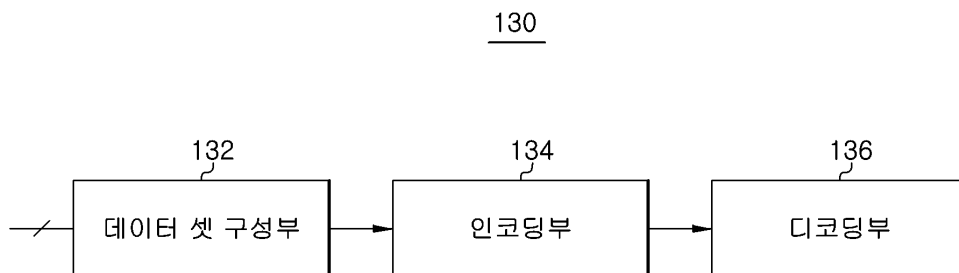
도면2



도면3



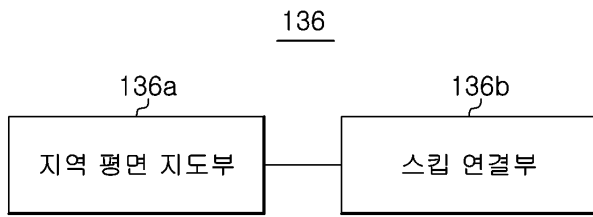
도면4



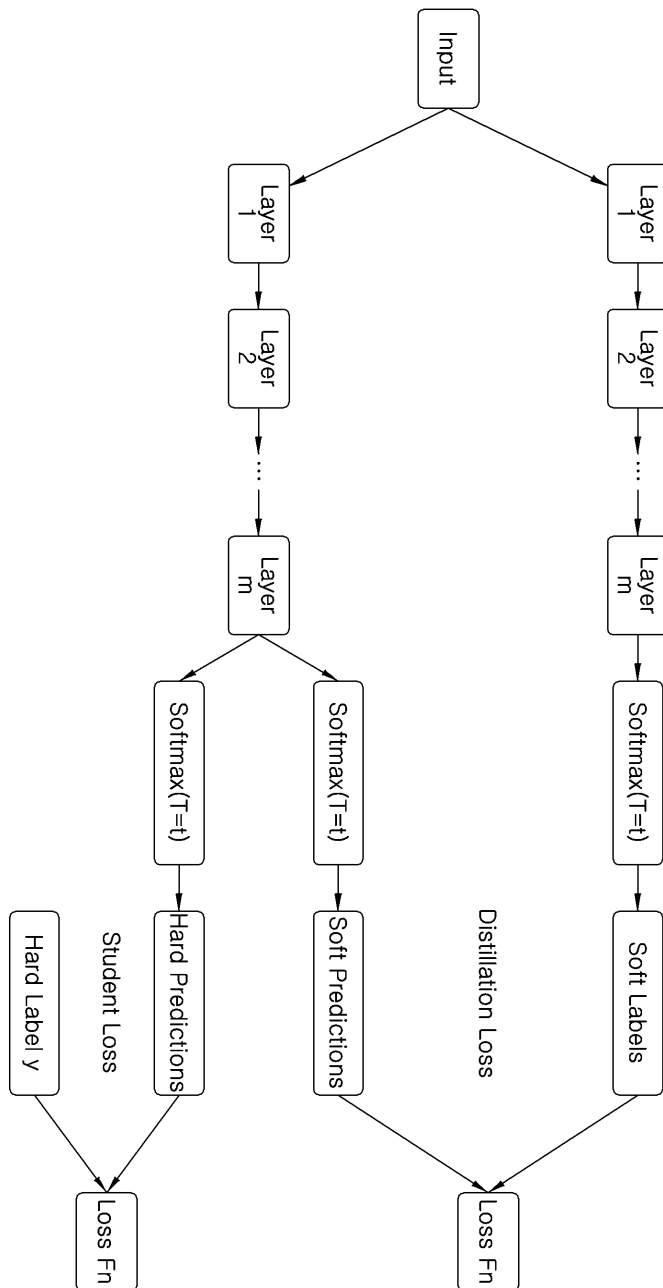
도면5



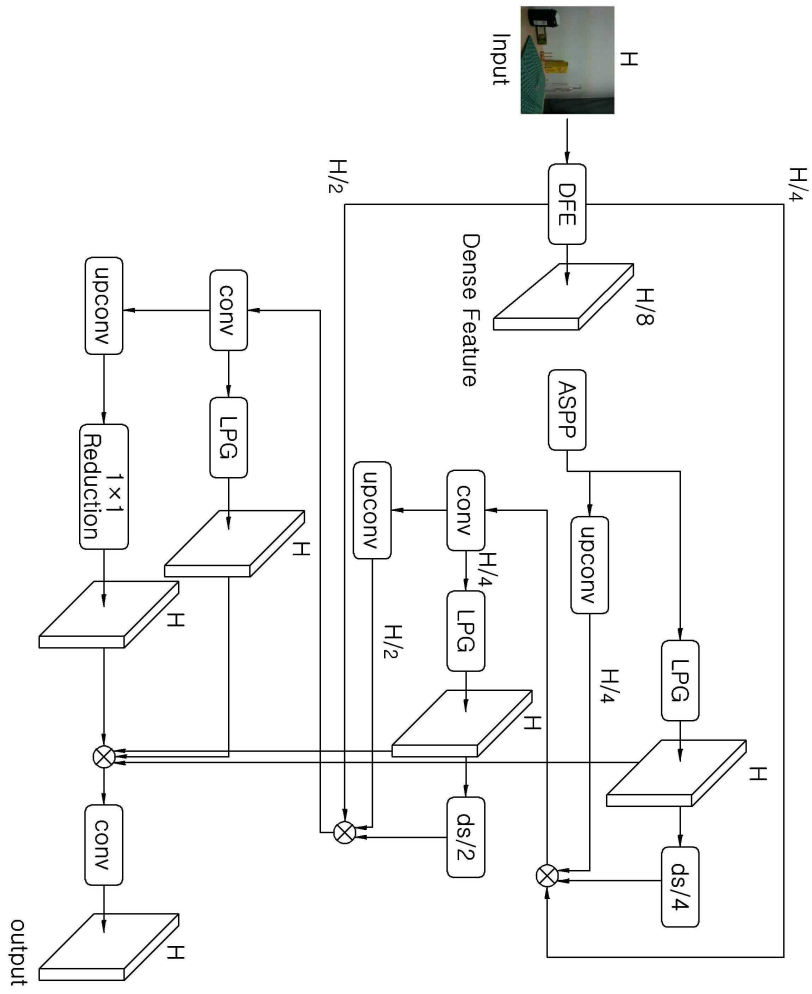
도면6



도면7



도면8



도면9

