



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2024-0111592
(43) 공개일자 2024년07월17일

(51) 국제특허분류(Int. Cl.)

G06V 20/56 (2022.01) G06N 3/094 (2023.01)
G06T 7/11 (2017.01) G06T 7/33 (2017.01)
G06T 7/593 (2017.01) G06T 7/70 (2017.01)
G06V 10/10 (2023.01) G06V 10/46 (2022.01)
G06V 10/764 (2022.01)

(52) CPC특허분류

G06V 20/56 (2023.08)
G06N 3/094 (2023.01)

(21) 출원번호 10-2023-0003614

(22) 출원일자 2023년01월10일

심사청구일자 없음

(71) 출원인

현대자동차주식회사

서울특별시 서초구 현릉로 12 (양재동)

기아 주식회사

서울특별시 서초구 현릉로 12 (양재동)

고려대학교 산학협력단

서울특별시 성북구 안암로 145, 고려대학교 (안암동5가)

(72) 발명자

김영현

서울특별시 강남구 개포로109길 9, 대치아파트
218-207

이서원

서울특별시 구로구 중앙로12길 12, 센츨리아파트
105-1108

(뒷면에 계속)

(74) 대리인

특허법인태평양

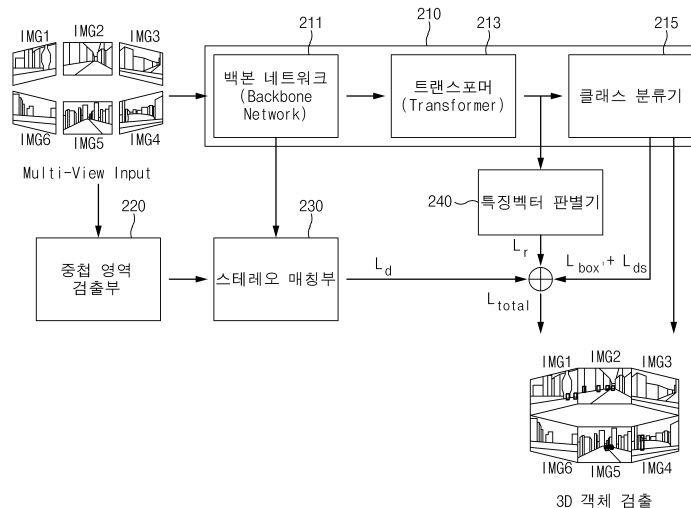
전체 청구항 수 : 총 20 항

(54) 발명의 명칭 객체 검출 장치 및 방법, 그리고 이를 위한 컴퓨터 판독 가능한 기록 매체

(57) 요약

본 발명의 실시 예에 의한 객체 검출 장치는 제1 화각 범위에 대한 제1 이미지를 획득하는 제1 카메라, 제1 화각 범위와 중첩영역을 갖는 제2 화각 범위에 대한 제2 이미지를 획득하는 제2 카메라, 제1 이미지 및 제2 이미지에서 객체를 검출하는 객체 검출 네트워크, 객체 검출 네트워크의 백본 네트워크가 출력하는 특징값을 바탕으로 중첩영역에서 제1 이미지와 제2 이미지 간의 깊이값의 편차를 보정하는 스테레오 매칭부, 및 객체 검출 네트워크가 객체를 검출하는 과정에서 상기 중첩영역의 객체와 비중첩영역의 객체 간의 특징 벡터 편차를 줄일 수 있도록 적대적 학습을 진행하는 특징 벡터 판별기를 포함할 수 있다.

대표도 - 도3



(52) CPC특허분류

G06T 7/11 (2017.01)
 G06T 7/337 (2017.01)
 G06T 7/593 (2017.01)
 G06T 7/70 (2017.01)
 G06V 10/16 (2023.08)
 G06V 10/469 (2023.08)
 G06V 10/764 (2023.08)
 G06T 2210/12 (2013.01)

노원석

서울특별시 동대문구 왕산로 77, 동대문베네스트2
 709호

(72) 발명자

김진규

서울특별시 성북구 안암로 145 고려대학교 내 정보
 대학

김상필

서울특별시 성북구 안암로 145, 우정정보관 606호

이 발명을 지원한 국가연구개발사업

과제고유번호	1711093609
과제번호	2019-0-00079-001
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	인공지능핵심고급인재양성
연구과제명	인공지능대학원지원(고려대학교)
기 여 율	1/1
과제수행기관명	고려대학교산학협력단
연구기간	2019.04.01 ~ 2023.12.31

명세서

청구범위

청구항 1

제1 화각 범위에 대한 제1 이미지를 획득하는 제1 카메라;

상기 제1 화각 범위와 중첩영역을 갖는 제2 화각 범위에 대한 제2 이미지를 획득하는 제2 카메라;

상기 제1 이미지 및 상기 제2 이미지에서 객체를 검출하는 객체 검출 네트워크;

상기 객체 검출 네트워크의 백본 네트워크가 출력하는 특징값을 바탕으로, 상기 중첩영역에서 상기 제1 이미지와 상기 제2 이미지 간의 깊이값의 편차를 보정하는 스테레오 매칭부; 및

상기 객체 검출 네트워크가 상기 객체를 검출하는 과정에서, 상기 중첩영역의 객체와 비중첩영역의 객체 간의 특징 벡터의 편차를 줄일 수 있도록 적대적 학습을 진행하는 특징 벡터 판별기;

를 포함하는 객체 검출 장치.

청구항 2

제 1 항에 있어서,

상기 스테레오 매칭부는

상기 중첩영역에서의 예측된 디스패리티 맵과 실측 디스패리티 맵 간의 편차를 줄여서 상기 깊이값의 편차를 보정하는 것을 특징으로 하는 객체 검출 장치.

청구항 3

제 2 항에 있어서,

상기 스테레오 매칭부는

상기 객체 검출 네트워크의 백본 네트워크가 출력하는 상기 제1 이미지에 대한 제1 이미지 특징값 및 상기 제2 이미지에 대한 제2 이미지 특징값을 매칭하여, 매칭된 특징값을 생성하고,

멀티 스케일 레이어를 이용하여, 상기 매칭된 특징값에 대한 멀티 스케일 특징값들을 획득하며,

상기 멀티 스케일 특징값들을 바탕으로 코스트 볼륨을 생성하는 것을 특징으로 하는 객체 검출 장치.

청구항 4

제 3 항에 있어서,

상기 스테레오 매칭부는

상기 코스트 볼륨과 상기 실측 디스패리티 맵 간의 편차에 대응하는 스테레오 초점 손실을 획득하며,

상기 객체 검출 네트워크는

상기 스테레오 초점 손실을 감소시키도록 상기 객체 검출 모델을 학습하는 것을 특징으로 하는 객체 검출 장치.

청구항 5

제 4 항에 있어서,

상기 특징 벡터 판별기는

상기 객체 검출 네트워크로부터 제공받은 상기 특징 벡터를 바탕으로, 상기 특징 벡터의 편차에 비례하는 지역 분류 손실을 출력하는 것을 특징으로 하는 객체 검출 장치.

청구항 6

제 5 항에 있어서,

상기 객체 검출 네트워크는

상기 백본 네트워크가 출력하는 특징값을 바탕으로 상기 객체의 예측 위치 정보 및 스코어 정보를 출력하는 트랜스포머; 및

상기 예측 위치 정보, 및 상기 스코어 정보, 및 상기 지역 분류 손실을 바탕으로 바운딩 박스 손실 및 클래스 손실을 출력하는 클래스 분류기;를 포함하는 객체 검출 장치.

청구항 7

제 6 항에 있어서,

상기 객체 검출 네트워크는

상기 특징 벡터 판별기가 출력하는 지역 분류 손실을 감소시키도록 학습을 진행하는 것을 특징으로 하는 객체 검출 장치.

청구항 8

제 7 항에 있어서,

상기 객체 검출 네트워크는

상기 바운딩 박스 손실과 상기 클래스 손실 및 상기 스테레오 초점 손실을 합산한 값에서 상기 지역분류 손실을 감산한 총 손실이 감소되도록 학습을 진행하는 것을 특징으로 하는 객체 검출 장치.

청구항 9

제1 카메라가 촬영한 제1 이미지와 제2 카메라가 촬영한 제2 이미지의 중첩 영역을 추출하는 단계;

객체 검출 네트워크를 이용하여 상기 제1 이미지 및 상기 제2 이미지에서 객체를 검출하는 과정에서, 스테레오 매칭부가, 상기 중첩 영역에서 상기 제1 이미지와 상기 제2 이미지 간의 깊이값의 편차를 보정하는 단계; 및

특징 벡터 판별기가, 상기 중첩 영역의 객체와 비중첩영역의 객체 간의 특징 벡터 편차를 줄일 수 있도록 적대적 학습을 진행하는 단계;

를 포함하는 객체 검출 방법.

청구항 10

제 9 항에 있어서,

상기 깊이값의 편차를 보정하는 단계는

상기 중첩 영역에서의 예측된 디스패리티 맵을 획득하는 단계; 및

상기 예측된 디스패리티 맵과 실측 디스패리티 맵 간의 편차를 줄이는 단계를 포함하는 것을 특징으로 하는 객체 검출 방법.

청구항 11

제 10 항에 있어서,

상기 중첩 영역에서의 예측된 디스패리티 맵을 획득하는 단계는

상기 객체 검출 네트워크의 백본 네트워크가 출력하는 상기 제1 이미지에 대한 제1 이미지 특징값 및 상기 제2 이미지에 대한 제2 이미지 특징값을 매칭하여 매칭된 특징값을 생성하는 단계;

멀티 스케일 레이어를 이용하여, 상기 매칭된 특징값에 대한 멀티 스케일 특징값들을 획득하는 단계; 및

상기 멀티 스케일 특징값들을 바탕으로 코스트 볼륨을 획득하는 단계;를 포함하는 것을 특징으로 하는 객체 검출 방법.

청구항 12

제 11 항에 있어서,

상기 깊이값을 보정하는 단계는

상기 코스트 볼륨과 상기 실측 디스패리티 맵 간의 편차에 대응하는 스테레오 초점 손실을 획득하는 단계; 및

상기 스테레오 초점 손실을 감소시키도록 상기 객체 검출 네트워크를 학습하는 단계를 더 포함하는 것을 특징으로 하는 객체 검출 방법.

청구항 13

제 12 항에 있어서,

상기 적대적 학습을 진행하는 단계는

상기 객체 검출 네트워크로부터 제공받은 상기 특징 벡터를 바탕으로, 상기 특징 벡터의 편차에 비례하는 지역 분류 손실을 출력하는 단계를 포함하는 것을 특징으로 하는 객체 검출 방법.

청구항 14

제 13 항에 있어서,

상기 적대적 학습을 진행하는 단계는

네트워크는 바운딩 박스 손실 및 클래스 손실을 출력하는 클래스 분류기에 상기 지역 분류 손실을 제공하는 단계를 더 포함하는 것을 특징으로 하는 객체 검출 방법.

청구항 15

제 14 항에 있어서,

상기 적대적 학습을 진행하는 단계는

상기 지역 분류 손실을 감소시키도록 상기 객체 검출 네트워크를 학습하는 단계를 더 포함하는 것을 특징으로 하는 객체 검출 방법.

청구항 16

제 9 항에 있어서,

상기 중첩영역을 추출하는 단계는

상기 제1 이미지 및 상기 제2 이미지로부터 추출한 2D 정보를 바탕으로 상기 깊이값을 포함하는 4D 정보를 획득하는 단계를 포함하는 것을 특징으로 하는 객체 검출 방법.

청구항 17

단계를 수행하기 위한 컴퓨터 판독 가능한 명령어들을 저장하는 컴퓨터 판독 가능한 기록 매체로써,

상기 단계들은

제1 카메라가 촬영한 제1 이미지와 제2 카메라가 촬영한 제2 이미지의 중첩영역을 추출하는 단계;

상기 중첩영역에서, 상기 제1 이미지와 상기 제2 이미지 간의 깊이값의 편차를 보정하는 단계; 및

객체 검출 네트워크를 이용하여 상기 제1 이미지 및 상기 제2 이미지에서 객체를 검출하는 과정에서, 상기 중첩영역의 객체와 비중첩영역의 객체 간의 특징 벡터의 편차를 줄일 수 있도록 적대적 학습을 진행하는 단계;

를 포함하는 컴퓨터 판독 가능한 기록 매체.

청구항 18

제 17 항에 있어서,

상기 깊이값의 편차를 보정하는 단계는

상기 객체 검출 네트워크의 백본 네트워크가 특징값을 바탕으로 예측된 디스페리티 맵을 획득하는 단계; 및

상기 예측된 디스페리티 맵과 실측 디스페리티 맵 간의 편차를 바탕으로 스테레오 초점 손실을 획득하는 단계;를 포함하는 것을 특징으로 하는 컴퓨터 판독 가능한 기록 매체.

청구항 19

제 18 항에 있어서,

상기 깊이값을 보정하는 단계는

상기 스테레오 초점 손실을 감소시키도록 상기 객체 검출 네트워크를 학습하는 단계를 더 포함하는 것을 특징으로 하는 컴퓨터 판독 가능한 기록 매체.

청구항 20

제 17 항에 있어서,

상기 적대적 학습을 진행하는 단계는

특징 벡터 판별기가 상기 객체 검출 네트워크로부터 제공받은 상기 특징 벡터를 바탕으로 상기 특징 벡터 간의 편차에 비례하는 지역 분류 손실을 출력하는 단계; 및

상기 지역 분류 손실을 감소시키도록 상기 객체 검출 네트워크를 학습하는 단계;를 포함하는 것을 특징으로 하는 컴퓨터 판독 가능한 기록 매체.

발명의 설명

기술 분야

[0001] 본 발명은 객체 검출 장치 및 방법, 그리고 이를 위한 컴퓨터 판독 가능한 기록 매체에 관한 것으로, 보다 상세하게는 차량 외부의 객체 검출 성능을 높일 수 있는 기술에 관한 것이다.

배경 기술

[0002] 자율주행 차량(Autonomous Vehicle)은 주행을 제어하기 위한 일부의 조작 또는 모든 조작이 사람을 대신한 차량 내의 시스템에 의해서 수행되는 차량을 의미한다. 넓은 의미의 자율주행 차량은 운전자 지원 시스템(Advanced Driver Assistance System; ADAS)을 포함할 수 있다.

[0003] 자율주행 시스템은 주행을 제어하기 위해서 도로 이외의 주변 상황을 파악하여야 하며, 이를 위해서 차량 주변의 객체를 인식하는 절차가 필수적이다.

[0004] 차량 주위의 객체를 확인하는 방법들 중에서 복수의 스테레오 카메라를 이용하여, 각각의 카메라의 시차(disparity)를 이용하여 깊이(depth)를 획득하고, 깊이를 이용하여 객체를 획득하는 방식이 이용되기도 한다.

[0005] 스테레오 카메라를 이용하여 객체를 검출하는 과정에서, 두 개의 카메라가 획득한 영상의 중첩(overlap) 영역에 존재하는 동일한 객체의 깊이 정보는 일치하여야 한다. 하지만, 각각의 카메라가 동일한 객체에 대해서 획득한 깊이 정보가 다른 경우가 발생하고 이에 따라 객체 검출의 정확도가 저하되기도 한다.

[0006] 또한, 중첩영역에서는 인접하는 이미지들의 픽셀 데이터들의 편차로 인해서 객체가 가지는 특징 벡터(feature vector)가 편향되는 경향이 발생하고 이에 따라 객체 검출의 정확도가 저하되기도 한다.

발명의 내용

해결하려는 과제

[0007] 본 발명은 서로 다른 카메라가 획득한 영상의 중첩영역에서 인접하는 이미지들 간의 깊이값의 편차를 개선할 수 있는 객체 검출 장치 및 방법, 그리고 이를 위한 컴퓨터 판독 가능한 기록 매체를 제공하기 위한 것이다.

[0008] 또한, 본 발명은 중첩영역 내에 존재하는 객체와 비중첩영역에 존재하는 객체가 가지는 특징 벡터의 편차를 줄일 수 있는 객체 검출 장치 및 방법, 그리고 이를 위한 컴퓨터 판독 가능한 기록 매체를 제공하기 위한 것이다.

[0009] 또한, 본 발명은 객체 검출 속도를 개선할 수 있는 객체 검출 장치 및 방법, 그리고 이를 위한 컴퓨터 판독 가능한 기록 매체를 제공하기 위한 것이다.

[0010] 본 발명의 기술적 과제들은 이상에서 언급한 기술적 과제들로 제한되지 않으며, 언급되지 않은 또 다른 기술적 과제들은 아래의 기재들로부터 당업자에게 명확하게 이해될 수 있을 것이다.

과제의 해결 수단

[0011] 본 발명의 실시 예에 의한 객체 검출 장치는 제1 화각 범위에 대한 제1 이미지를 획득하는 제1 카메라, 제1 화각 범위와 중첩영역을 갖는 제2 화각 범위에 대한 제2 이미지를 획득하는 제2 카메라, 제1 이미지 및 제2 이미지에서 객체를 검출하는 객체 검출 네트워크, 객체 검출 네트워크의 백본 네트워크가 출력하는 특징값을 바탕으로 중첩영역에서 제1 이미지와 제2 이미지 간의 깊이값의 편차를 보정하는 스테레오 매칭부, 및 객체 검출 네트워크가 객체를 검출하는 과정에서 상기 중첩영역의 객체와 비중첩영역의 객체 간의 특징 벡터 편차를 줄일 수 있도록 적대적 학습을 진행하는 특징 벡터 판별기를 포함할 수 있다.

[0012] 실시 예에 의하면, 상기 스테레오 매칭부는 상기 중첩영역에서의 예측된 디스패리티 맵과 실측 디스패리티 맵 간의 편차를 줄여서 상기 깊이값의 편차를 보정할 수 있다.

[0013] 실시 예에 의하면, 상기 스테레오 매칭부는 상기 객체 검출 네트워크의 백본 네트워크가 출력하는 상기 제1 이미지에 대한 제1 이미지 특징값 및 상기 제2 이미지에 대한 제2 이미지 특징값을 매칭하여, 매칭된 특징값을 생성하고, 멀티 스케일 레이어를 이용하여, 상기 매칭된 특징값에 대한 멀티 스케일 특징값들을 획득하며, 상기 멀티 스케일 특징값들을 바탕으로 코스트 볼륨을 생성할 수 있다.

[0014] 실시 예에 의하면, 상기 스테레오 매칭부는 상기 코스트 볼륨과 상기 실측 디스패리티 맵 간의 편차에 대응하는 스테레오 초점 손실을 획득하며, 상기 객체 검출 네트워크는 상기 스테레오 초점 손실을 감소시키도록 상기 객

체 검출 모델을 학습할 수 있다.

- [0015] 실시 예에 의하면, 상기 특징 벡터 판별기는 상기 객체 검출 네트워크로부터 제공받은 클래스 정보 및 위치 정보를 바탕으로, 상기중첩영역의 특징 벡터와 상기 비중첩영역의 특징 벡터 간의 편차에 비례하는 지역 분류 손실을 출력할 수 있다.
- [0016] 실시 예에 의하면, 상기 객체 검출 네트워크는 상기 백본 네트워크가 출력하는 특징값을 바탕으로 상기 객체의 예측 위치 정보 및 스코어 정보를 출력하는 트랜스포머; 및 상기 예측 위치 정보, 및 상기 스코어 정보, 및 상기 지역 분류 손실을 바탕으로 바운딩 박스 손실 및 클래스 손실을 출력하는 클래스 분류기;를 포함할 수 있다.
- [0017] 실시 예에 의하면, 상기 객체 검출 네트워크는 상기 특징 벡터 판별기가 출력하는 지역 분류 손실을 감소시키도록 학습을 진행할 수 있다.
- [0018] 실시 예에 의하면, 상기 객체 검출 네트워크는 상기 바운딩 박스 손실과 상기 클래스 손실 및 상기 스테레오 초점 손실을 합산한 값에서 상기 지역분류 손실을 감산한 총 손실이 감소되도록 학습을 진행할 수 있다.
- [0019] 본 발명의 실시 예에 의한 객체 검출 방법은 제1 카메라가 촬영한 제1 이미지와 제2 카메라가 촬영한 제2 이미지의 중첩 영역을 추출하는 단계, 객체 검출 네트워크를 이용하여 상기 제1 이미지 및 상기 제2 이미지에서 객체를 검출하는 과정에서 스테레오 매칭부가 상기 중첩 영역에서 상기 제1 이미지와 상기 제2 이미지 간의 깊이값의 편차를 보정하는 단계, 및 특징 벡터 판별기가 상기 중첩 영역의 객체와 비중첩영역의 객체 간의 특징 벡터 편차를 줄일 수 있도록 적대적 학습을 진행하는 단계를 포함할 수 있다.
- [0020] 실시 예에 의하면, 상기 깊이값의 편차를 보정하는 단계는 상기 중첩 영역에서의 예측된 디스패리티 맵을 획득하는 단계 및 상기 예측된 디스패리티 맵과 실측 디스패리티 맵 간의 편차를 줄이는 단계를 포함할 수 있다.
- [0021] 실시 예에 의하면, 상기 중첩 영역에서의 예측된 디스패리티 맵을 획득하는 단계는 상기 객체 검출 네트워크의 백본 네트워크가 출력하는 상기 제1 이미지에 대한 제1 이미지 특징값 및 상기 제2 이미지에 대한 제2 이미지 특징값을 매칭하여 매칭된 특징값을 생성하는 단계, 멀티 스케일 레이어를 이용하여, 상기 매칭된 특징값에 대한 멀티 스케일 특징값들을 획득하는 단계 및
- [0022] 상기 멀티 스케일 특징값들을 바탕으로 코스트 볼륨을 획득하는 단계를 포함할 수 있다.
- [0023] 실시 예에 의하면, 상기 깊이값을 보정하는 단계는 상기 코스트 볼륨과 상기 실측 디스패리티 맵 간의 편차에 대응하는 스테레오 초점 손실을 획득하는 단계 및 상기 스테레오 초점 손실을 감소시키도록 상기 객체 검출 네트워크를 학습하는 단계를 더 포함할 수 있다.
- [0024] 실시 예에 의하면, 상기 적대적 학습을 진행하는 단계는 상기 객체 검출 네트워크로부터 제공받은 클래스 정보 및 위치 정보를 바탕으로, 상기 중첩영역의 특징 벡터와 상기 비중첩영역의 특징 벡터 간의 편차에 비례하는 지역 분류 손실을 출력하는 단계를 포함할 수 있다.
- [0025] 실시 예에 의하면, 상기 적대적 학습을 진행하는 단계는 네트워크는 바운딩 박스 손실 및 클래스 손실을 출력하는 클래스 분류기에 상기 지역 분류 손실을 제공하는 단계를 더 포함할 수 있다.
- [0026] 실시 예에 의하면, 상기 적대적 학습을 진행하는 단계는 상기 지역 분류 손실을 감소시키도록 상기 객체 검출 네트워크를 학습하는 단계를 더 포함할 수 있다.
- [0027] 실시 예에 의하면, 상기 중첩영역을 추출하는 단계는 상기 제1 이미지 및 상기 제2 이미지로부터 추출한 2D 정보를 바탕으로 상기 깊이값을 포함하는 4D 정보를 획득하는 단계를 포함할 수 있다.
- [0028] 본 발명의 실시 예에 의한 컴퓨터 판독 가능한 기록 매체는 단계들을 수행하기 위한 컴퓨터 판독 가능한 명령어들을 저장하는 컴퓨터 판독 가능한 기록 매체로써, 상기 단계들은 제1 카메라가 촬영한 제1 이미지와 제2 카메라가 촬영한 제2 이미지의 중첩영역을 추출하는 단계, 상기 중첩영역에서, 상기 제1 이미지와 상기 제2 이미지 간의 깊이값의 편차를 보정하는 단계, 및 객체 검출 네트워크를 이용하여 상기 제1 이미지 및 상기 제2 이미지에서 객체를 검출하는 과정에서, 상기 중첩영역의 객체와 비중첩영역의 객체 간의 특징 벡터 편차를 줄일 수 있도록 적대적 학습을 진행하는 단계를 포함할 수 있다.
- [0029] 실시 예에 의하면, 상기 깊이값의 편차를 보정하는 단계는 상기 객체 검출 네트워크의 백본 네트워크가 특징값을 바탕으로 예측된 디스패리티 맵을 획득하는 단계 및 상기 예측된 디스패리티 맵과 실측 디스패리티 맵 간의 편차를 바탕으로 스테레오 초점 손실을 획득하는 단계를 포함할 수 있다.

[0030] 실시 예에 의하면, 상기 깊이값을 보정하는 단계는 상기 스테레오 초점 손실을 감소시키도록 상기 객체 검출 네트워크를 학습하는 단계를 더 포함할 수 있다.

[0031] 실시 예에 의하면, 상기 적대적 학습을 진행하는 단계는 특징 벡터 판별기가 상기 객체 검출 네트워크로부터 제공받은 클래스 정보 및 위치 정보를 바탕으로 상기 중첩영역의 특징 벡터와 상기 비중첩영역의 특징 벡터 간의 편차에 비례하는 지역 분류 손실을 출력하는 단계 및 상기 지역 분류 손실을 감소시키도록 상기 객체 검출 네트워크를 학습하는 단계를 포함할 수 있다.

발명의 효과

[0032] 본 발명의 실시 예에 의한 객체 검출 장치 및 방법, 그리고 이를 위한 컴퓨터 판독 가능한 기록 매체에 의하면, 서로 다른 카메라가 획득한 영상의 중첩영역에 대한 스테레오 정합을 통해서 깊이값의 편차를 개선할 수 있다.

[0033] 또한, 본 발명의 실시 예에 의한 객체 검출 장치 및 방법, 그리고 이를 위한 컴퓨터 판독 가능한 기록 매체에 의하면, 적대적 학습부를 통해서 객체 검출 네트워크가 비중첩 영역의 객체와 동일한 조건으로 중첩영역의 객체를 검출할 수 있도록 하여, 중첩영역의 객체 검출 성능을 높일 수 있다.

[0034] 또한, 본 발명은 객체 검출 속도를 개선할 수 있는 객체 검출 장치 및 방법, 그리고 이를 위한 컴퓨터 판독 가능한 기록 매체에 의하면, 훈련이 완료된 이후에는 중첩영역 검출부, 스테레오 매칭부 및 적대적 학습부를 이용하지 않고, 객체 검출 네트워크만을 이용하여 객체 검출을 진행하기 때문에 객체 검출 성능을 높이면서 객체 검출 속도를 향상시킬 수 있다.

[0035] 이 외에, 본 문서를 통해 직접적 또는 간접적으로 파악되는 다양한 효과들이 제공될 수 있다.

도면의 간단한 설명

[0036] 도 1은 본 발명의 일 실시 예에 객체 검출 모델을 포함하는 차량을 나타내는 도면이다.

도 2는 본 발명의 일 실시 예에 의한 객체 검출 모델을 포함하는 자율주행 제어 장치의 구성을 나타내는 블록도이다.

도 3은 본 발명의 실시 예에 의한 객체 검출 모델을 나타내는 도면이다.

도 4는 본 발명의 실시 예에 의한 객체 검출 방법을 설명하기 위한 순서도이다.

도 5는 중첩영역 검출부의 동작을 설명하는 도면이다.

도 6은 본 발명의 실시 예에 의한 스테레오 초점 손실을 설명하는 도면이다.

도 7은 본 발명의 실시 예에 의한 스테레오 정합을 설명하는 도면이다.

도 8은 중첩영역에 대한 적대적 학습을 설명하는 도면이다.

도 9는 특징 벡터 판별기의 입력 인자를 설명하는 도면이다.

도 10은 본 발명의 실시 예에 의한 객체 검출 과정을 설명하는 도면이다.

도 11 내지 도 13은 본 발명의 실시 예에 의한 객체 검출 모델의 성능을 나타내는 도면이다.

도 14는 본 발명의 일 실시 예에 따른 컴퓨팅 시스템을 나타내는 도면이다.

발명을 실시하기 위한 구체적인 내용

[0037] 이하, 본 발명의 일부 실시예들을 예시적인 도면을 통해 상세하게 설명한다. 각 도면의 구성요소들에 참조부호를 부가함에 있어서, 동일한 구성요소들에 대해서는 비록 다른 도면상에 표시되더라도 가능한 한 동일한 부호를 가지도록 하고 있음에 유의해야 한다. 또한, 본 발명의 실시예를 설명함에 있어, 관련된 공지 구성 또는 기능에 대한 구체적인 설명이 본 발명의 실시예에 대한 이해를 방해한다고 판단되는 경우에는 그 상세한 설명은 생략한다.

[0038] 본 발명의 실시예의 구성 요소를 설명하는 데 있어서, 제 1, 제 2, A, B, (a), (b) 등의 용어를 사용할 수 있다. 이러한 용어는 그 구성 요소를 다른 구성 요소와 구별하기 위한 것일 뿐, 그 용어에 의해 해당 구성 요소의 본질이나 차례 또는 순서 등이 한정되지 않는다. 또한, 다르게 정의되지 않는 한, 기술적이거나 과학적인 용

어를 포함해서 여기서 사용되는 모든 용어들은 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자에 의해 일반적으로 이해되는 것과 동일한 의미를 가진다. 일반적으로 사용되는 사전에 정의되어 있는 것과 같은 용어들은 관련 기술의 문맥상 가지는 의미와 일치하는 의미를 가진 것으로 해석되어야 하며, 본 출원에서 명백하게 정의하지 않는 한, 이상적이거나 과도하게 형식적인 의미로 해석되지 않는다.

- [0039] 이하, 도 1 내지 도 14를 참조하여, 본 발명의 실시 예들을 구체적으로 설명하기로 한다.
- [0040] 도 1은 본 발명의 일 실시 예에 객체 검출 모델을 포함하는 차량을 나타내는 도면이다. 본 발명의 실시 예에 의한 객체 검출 모델은 자율주행 차량에 차량에 장착되어, 차량의 외부 객체를 검출할 수 있다. 본 발명의 실시 예에 의한 자율주행 차량은 자율주행 레벨1 내지 자율주행 레벨5에 속하는 차량일 수 있으며, 운전자 보조 시스템(Advanced Driver Assistance System; ADAS)을 이용하는 차량을 포함할 수 있다.
- [0041] 도 1을 참조하면, 본 발명의 실시 예에 의한 차량(VEH)은 본체(2), 차륜(61, 62), 도어(71), 전면 유리(80), 사이드미러(81, 82), 센서부(10), 및 프로세서(100)를 포함할 수 있다.
- [0042] 차량(VEH)은 전기자동차(Electric Vehicle, EV), 하이브리드 차량(Hybrid Electric Vehicle, HEV), 플러그인 하이브리드 차량(Plug-in Hybrid Electric Vehicle, PHEV), 또는 수소전기차(Fuel Cell Electric Vehicle, FCEV) 등의 전동화 기반의 차량일 수 있다.
- [0043] 본체(2)는 차량(VEH)의 외형을 이루는 구조물일 수 있다.
- [0044] 차륜(61, 62)은 차량의 전방에 마련되는 전륜(61), 차량의 후방에 마련되는 후륜(62)을 포함하며, 전륜(61) 및 후륜(62)은 구동장치에 의해서 회전하여 차량(VEH)을 이동시킬 수 있다.
- [0045] 도어(71)는 본체(2)의 좌측 및 우측에 회동 가능하게 마련되어 개방 시에 탑승자가 차량(VEH)의 내부에 탑승할 수 있도록 하며, 폐쇄 시에 차량(VEH)의 내부를 외부로부터 차폐시킬 수 있다.
- [0046] 윈드 스크린의 일종인 전면 유리(80)는 본체(2)의 전방 상측에 마련되어 차량(VEH) 내부의 운전자 또는 사용자에게 차량(VEH) 전방의 시야 정보를 제공할 수 있다.
- [0047] 사이드미러(81, 82)는 본체(2)의 좌측에 마련되는 좌측 사이드미러(81) 및 우측에 마련되는 우측 사이드미러(82)를 포함하며, 차량 내부의 운전자에게 차량(VEH)의 측, 후방의 시야 정보를 제공할 수 있다.
- [0048] 센서부(10)는 카메라(11)를 포함할 수 있으며, 카메라(11)가 배치되는 위치는 도 1에 한정되지 않을 수 있다.
- [0049] 프로세서(100)는 본 발명의 실시 예에 의한 객체 검출 모델(200)을 이용하여 차량 외부의 객체를 검출할 수 있다.
- [0050] 도 2는 본 발명의 일 실시 예에 의한 객체 검출 모델을 포함하는 자율주행 제어 장치의 구성을 나타내는 블록도이다. 본 발명의 일 실시 예에 의한 자율주행 제어 장치(190)는 차량의 내부에 구현될 수 있다. 이때, 차량 제어 장치(190)는 차량의 내부 제어 유닛들과 일체로 형성될 수 있으며, 별도의 장치로 구현되어 별도의 연결 수단에 의해 차량의 제어 유닛들과 연결될 수도 있다.
- [0051] 도 2를 참조하면, 본 발명의 실시 예에 의한 객체 검출 모델을 포함하는 자율주행 제어 장치는 센서부(10), 통신부(20), 저장부(30), 프로세서(100), 및 주행 제어 장치(110)를 포함할 수 있다.
- [0052] 센서부(10)는 차량의 외부 객체, 특히 차량의 전방 또는 후방의 위치한 차량들을 검출하기 위한 카메라(11)를 포함할 수 있다.
- [0053] 카메라(11)는 차량 주변의 외부 이미지를 획득하기 위한 것으로, 둘 이상의 카메라를 이용하여 획득한 깊이 정보를 바탕으로 3차원 객체를 검출할 수 있다. 서로 인접하는 복수의 카메라들은 각각의 화각이 적어도 일부는 중첩되도록 배치될 수 있다.
- [0054] 도 1에서와 같이, 제1 카메라(C1)는 차량의 좌측 전방 영역의 이미지를 획득할 수 있고, 제2 카메라(C2)는 차량의 전방 영역의 이미지를 획득할 수 있으며, 제3 카메라(C3)는 차량의 우측 전방 영역의 이미지를 획득할 수 있다.
- [0055] 이와 유사하게, 차량의 후방에는 우측 후방 영역의 이미지를 획득할 수 있는 제4 카메라(C4), 후방 영역의 이미지를 획득할 수 있는 제5 카메라(C5), 좌측 후방 영역의 이미지를 획득할 수 있는 제6 카메라(C6)를 포함할 수 있다.

- [0056] 또한, 센서부(10)는 라이다(Light Imaging Detection and Ranging; LIDAR), 레이다(Radio Detection and Ranging; RADAR), 초음파 센서, 또는 적외선 센서를 포함할 수 있다.
- [0057] 라이다는 레이저 송신 모듈, 수신 모듈을 포함할 수 있다. 라이다는, TOF(Time of Flight) 방식 또는 페이즈 쉬프트(phase-shift) 방식으로 구현될 수 있다. 라이다는 차량의 전방, 후방 또는 측방에 위치하는 객체를 감지하기 위해 차량의 외부에 노출될 수 있다.
- [0058] 레이다는 전자파 송신 모듈, 수신 모듈을 포함할 수 있다. 레이다는 전파 발사 원리상 펄스 레이다(Pulse Radar) 방식 또는 연속파 레이다(Continuous Wave Radar) 방식으로 구현될 수 있다. 레이다는 연속파 레이다 방식 중에서 신호 파형에 따라 FMCW(Frequency Modulated Continuous Wave)방식 또는 FSK(Frequency Shift Keying) 방식으로 구현될 수 있다. 레이다는 전자파를 매개로, TOF(Time of Flight) 방식 또는 페이즈 쉬프트(phase-shift) 방식에 기초하여, 객체를 검출하고, 검출된 객체의 위치, 검출된 객체와의 거리 및 상대 속도를 검출할 수 있다.
- [0059] 초음파 센서는 초음파 송신 모듈, 수신 모듈을 포함할 수 있다. 초음파 센서는, 초음파를 기초로 객체를 검출하고, 검출된 객체의 위치, 검출된 객체와의 거리 및 상대 속도를 검출할 수 있다. 초음파 센서는 차량의 전방, 후방 또는 측방에 위치하는 객체를 감지하기 위해 차량의 외부의 적절한 위치에 배치될 수 있다.
- [0060] 적외선 센서는 적외선 송신 모듈, 수신 모듈을 포함할 수 있다. 적외선 센서는, 적외선 광을 기초로 객체를 검출하고, 검출된 객체의 위치, 검출된 객체와의 거리 및 상대 속도를 검출할 수 있다. 적외선 센서는 차량의 전방, 후방 또는 측방에 위치하는 객체를 감지하기 위해 차량의 외부에 배치될 수 있다.
- [0061] 또한, 센서부(10)는 차량의 변속을 위한 속도 제어 명령을 생성하는 브레이크 페달 위치 센서(Brake-pedal Position Sensor; BPS), 및 가속 페달 위치 센서(Accelerator Position Sensor; APS)를 포함할 수 있다.
- [0062] 브레이크 페달 위치 센서는 차량에 구비된 브레이크 페달의 눌림 정도에 따라 BPS 신호를 출력할 수 있다. 일례로, BPS 신호는 브레이크 페달의 눌림에 따라 0~100의 데이터를 출력할 수 있고, 0의 값은 브레이크 페달의 눌림이 없을 경우이며, 100의 값은 브레이크 페달의 눌림이 최대인 경우일 수 있다.
- [0063] 가속 페달 위치 센서는 차량에 구비된 가속 페달의 눌림 정도에 따라 APS 신호를 출력할 수 있다. 일례로, APS 신호는 가속 페달의 눌림에 따라 0~100의 데이터를 출력할 수 있고, 0의 값은 가속 페달의 눌림이 없을 경우이며, 100의 값은 가속 페달의 눌림이 최대인 경우일 수 있다.
- [0064] 통신부(20)는 사용자 단말기, 다른 차량 또는 외부 서버와 통신을 수행할 수 있다.
- [0065] 근거리 통신(Short range communication), GPS 신호 수신, V2X 통신, 광통신, 방송 송수신 및 ITS(Intelligent Transport Systems) 통신 기능을 수행할 수 있다.
- [0066] 통신부(20)는 블루투스(Bluetooth), RFID(Radio Frequency Identification), 적외선 통신(Infrared Data Association; IrDA), UWB(Ultra Wideband), ZigBee, NFC(Near Field Communication), Wi-Fi(Wireless-Fidelity), Wi-Fi Direct, Wireless USB(Wireless Universal Serial Bus) 기술 중 적어도 하나를 이용하여, 근거리 통신을 지원할 수 있다.
- [0067] 통신부(20)는 위치 정보를 획득하기 위한 GPS(Global Positioning System) 모듈 또는 DGPS(Differential Global Positioning System) 모듈을 포함할 수 있다.
- [0068] 또한 통신부(20)는 V2X 통신 모듈을 포함할 수 있다. V2X 통신 모듈은 서버(Vehicle to Infra; V2I), 타 차량(Vehicle to Vehicle; V2V) 또는 보행자(Vehicle to Pedestrian; V2P)와의 무선 통신 프로토콜을 위한 RF 회로를 포함할 수 있다.
- [0069] 통신부(20)는 이동 통신을 위한 기술 표준들 또는 통신방식에 따라 구축된 이동 통신망 상에서 기지국, 외부의 단말, 센터 중 적어도 하나와 무선 신호를 송수신할 수 있다. 예를 들어, 통신부(20)는 GSM(Global System for Mobile communication), CDMA(Code Division Multi Access), CDMA2000(Code Division Multi Access 2000), EV-DO(Enhanced Voice-Data Optimized or Enhanced Voice-Data Only), WCDMA(Wideband CDMA), HSDPA(High Speed Downlink Packet Access), HSUPA(High Speed Uplink Packet Access), LTE(Long Term Evolution), LTE-A(Long Term Evolution-Advanced) 기반으로 통신을 수행할 수 있다. 무선 신호는 음성 호 신호, 화상 통화 호 신호 또는 문자/멀티미디어 메시지 송수신에 따른 다양한 형태의 데이터를 포함할 수 있다.
- [0070] 저장부(30)는 프로세서(100)의 동작을 위한 알고리즘 및 객체 검출 모델을 저장할 수 있다. 저장부(30)는 하드

디스크 드라이브, 플래시 메모리, EEPROM(Electrically erasable programmable read-only memory), SRAM(Static RAM), FRAM (Ferro-electric RAM), PRAM (Phase-change RAM), MRAM(Magnetic RAM), DRAM(Dynamic Random Access Memory), SDRAM(Synchronous Dynamic Random Access Memory), DDR-SDRAM(Double Date Rate-SDRAM) 등을 이용할 수 있다.

[0071] 프로세서(100)는 카메라들이 획득한 이미지들을 바탕으로 객체를 검출할 수 있다. 이를 위해서 프로세서(100)는 도 3에 도시된 객체 검출 모델(200)을 포함할 수 있다.

[0072] 객체 검출 모델(200)은 카메라(11)가 획득한 이미지를 바탕으로 객체를 검출하기 위해서 인간의 뇌 구조를 컴퓨터 상에서 모의하도록 설계된 신경망을 포함할 수 있다. 객체 검출 모델(200)의 신경망은 인간의 뉴런(neuron)을 모의하는 가중치를 갖는 복수의 네트워크 노드들을 포함할 수 있다. 복수의 네트워크 노드들은 뉴런이 시냅스(synapse)를 통해 신호를 주고 받는 뉴런의 시냅틱 활동을 모의하도록 각각 연결 관계에 따라 데이터를 주고 받을 수 있다. 신경망은 신경망 모델에서 발전한 딥러닝 모델을 포함할 수 있다. 딥 러닝 모델에서 복수의 네트워크 노드들은 서로 다른 레이어에 위치하면서 컨볼루션(convolution) 연결 관계에 따라 데이터를 주고 받을 수 있다. 신경망 모델의 예는 심층 신경망(DNN, deep neural networks), 합성곱 신경망(CNN, convolutional deep neural networks), 순환 신경망(RNN, Recurrent Boltzmann Machine), 제한 볼츠만 머신(RBM, Restricted Boltzmann Machine), 심층 신뢰 신경망(DBN, deep belief networks), 심층 Q-네트워크(Deep Q-Network)와 같은 다양한 딥 러닝 기법들을 포함할 수 있다.

[0073] 이하, 본 발명의 실시 예에서 객체 검출 모델(200)이 객체를 검출하는 동작은 프로세서(100)의 동작으로 이해될 수 있다.

[0074] 또한, 프로세서(100)는 검출된 객체를 바탕으로, 주행 구동부(50)를 제어하여 자율주행 또는 운전 보조 기능을 수행할 수 있다.

[0075] 주행 구동부(50)는 객체를 검출한 프로세서(100)의 제어에 따라 운전 보조 기능을 동작시킬 수 있다.

[0076] 운전 보조 기능은 전방 차량과의 충돌 위험을 미리 감지해 운전자에게 경고하는 전방충돌 경고 기능(Forward Collision Warning), 및 충돌 위험에 대응하여 차량의 제동과 조향을 제어하는 전방 충돌방지 보조 기능(Foward Collision-Avoidance Assist)을 포함할 수 있다.

[0077] 또한, 운전 보조 기능은 차로 이탈방지 보조 시스템(LKA, Lane Keeping Assist), 차선 유지 보조 시스템(Lane Departure Warning System or Lane Keeping Assist System), 후측방 충돌방지 보조(Blind-Spot Collision-Avoidance Assist; BCA), 안전 하차 보조(Safety Exit Assist; SEA), 지능형 속도 제한 보조 (Intelligent Speed Limit Assist; ISLA), 스마트 크루즈(Smart Cruise Control;SCC), 로우빔 보조 시스템(Low Beam Assist), 하이빔 보조 시스템(High Beam Assist), 운전자 주의 경고(Driver Attention Warning; DAW), 후측방 모니터 (BVM, Blind-spot View Monitor), 고속도로 주행 보조(Highway Driving Assist; HDA), 서라운드 뷰 모니터(Surround View Monitor; SVM), 원격 스마트 주차 보조 (Remote Smart Parking Assist; RSPA) 등을 포함할 수 있다.

[0078] 도 3은 본 발명의 실시 예에 의한 객체 검출 모델을 나타내는 도면이다.

[0079] 도 3을 참조하면, 본 발명의 실시 예에 의한 객체 검출 모델은 객체 검출 네트워크(210), 중첩영역 검출부(220), 스테레오 매칭부(230), 및 특징 벡터 판별기(240)를 포함할 수 있다.

[0080] 객체 검출 네트워크(210)는 백본 네트워크(Backbone Network)(211), 트랜스포머(Transformer)(213), 및 클래스 분류기(Class Classification)(215)를 포함할 수 있다.

[0081] 백본 네트워크(211)는 입력 이미지를 바탕으로 특징값을 추출할 수 있다. 입력 이미지는 제1 카메라(C1) 내지 제6 카메라(C6)를 통해서 획득한 제1 이미지(IMG1) 내지 제6 이미지(IMG6)를 포함할 수 있다.

[0082] 트랜스포머(213)는 입력 시퀀스를 제공받는 인코더(Transformer encoder) 및 출력 시퀀스를 출력하는 디코더(Transformer decoder)를 포함할 수 있다. 트랜스포머(213)는 앵커 박스(Anchor Box)를 이용하지 않으며, 백본 네트워크(211)로부터 제공받은 특징값을 바탕으로, 예상되는 객체의 위치 정보, 스코어 정보 및 특징 벡터를 출력할 수 있다. 위치 정보는 예상되는 객체의 좌표 정보를 포함할 수 있고, 스코어 정보는 예상되는 객체의 정확도 또는 유사도를 나타낼 수 있다.

[0083] 클래스 분류기(215)는 객체를 구분하여, 객체의 위치 및 객체 클래스를 출력할 수 있다. 또한, 클래스 분류기

(215)는 바운딩 박스 손실 및 클래스 손실을 출력할 수 있다.

- [0084] 중첩영역 검출부(220)는 인접하는 이미지들 간의 중첩영역을 검출할 수 있다. 중첩영역 검출부(220)는 제1 이미지(IMG1)와 제2 이미지(IMG2)들 간의 제1 중첩영역을 검출할 수 있다. 또한 중첩영역 검출부(220)는 제2 이미지(IMG2)와 제3 이미지(IMG3)들 간의 제2 중첩영역을 검출할 수 있고, 제3 이미지(IMG3)와 제4 이미지(IMG4)들 간의 제3 중첩영역을 검출할 수 있으며, 제4 이미지(IMG4)와 제5 이미지(IMG5)들 간의 제4 중첩영역을 검출할 수 있고, 제5 이미지(IMG5)와 제6 이미지(IMG6)들 간의 제5 중첩영역을 검출할 수 있으며, 제6 이미지(IMG6)와 제1 이미지(IMG1)들 간의 제6 중첩영역을 검출할 수 있다. 각 이미지들에서 중첩영역 이외의 영역은 비중첩영역으로 정의될 수 있다.
- [0085] 스테레오 매칭부(230)는 객체 검출 네트워크(210)가 출력하는 특징값을 바탕으로, 중첩영역에서 이미지들 간의 깊이값 편차를 보정할 수 있다. 예를 들어, 스테레오 매칭부(230)는 제1 중첩영역(OV1)에서 제1 이미지(IMG1)의 깊이값과 제1 중첩영역에서 제2 이미지(IMG2)의 깊이값의 편차를 보정할 수 있다.
- [0086] 이를 위해서, 스테레오 매칭부(230)는 중첩영역에서 중첩영역에서 예측된 디스패리티 맵을 획득할 수 있다. 예측된 디스패리티 맵은 객체 검출 네트워크(210)의 백본 네트워크(211)가 출력하는 특징값을 바탕으로 생성될 수 있다. 스테레오 매칭부(230)는 제1 이미지에 대한 제1 이미지 특징값 및 상기 제2 이미지에 대한 제2 이미지 특징값을 매칭하여, 매칭된 특징값을 생성할 수 있다. 그리고 스테레오 매칭부(230)는 멀티 스케일 레이어를 이용하여 매칭된 특징값에 대한 멀티 스케일 특징값들을 획득하며, 멀티 스케일 특징값들을 바탕으로 코스트 볼륨을 생성할 수 있다.
- [0087] 스테레오 매칭부(230)는 코스트 볼륨과 상기 실측 디스패리티 맵 간의 편차에 대응하는 스테레오 초점 손실을 획득할 수 있다.
- [0088] 특징 벡터 판별기(240)는 중첩영역에서 검출된 객체와 비중첩영역에서 검출된 객체 간의 특징 벡터 편차를 줄일 수 있다. 이를 위해서, 특징 벡터 판별기(240)는 객체 검출 네트워크(210)로부터 제공받은 클래스 정보 및 위치 정보를 바탕으로, 중첩영역에서 검출되는 객체의 특징 벡터와 비중첩영역에서 검출되는 특징 벡터 간의 편차에 비례하는 지역 분류 손실을 출력할 수 있다. 예를 들어, 특징 벡터 판별기(240)는 중첩영역에서 검출되는 객체의 특징 벡터와 비중첩영역에서 검출되는 특징 벡터 간의 편차가 클수록 큰 값으로 설정되는 지역 분류 손실을 결정할 수 있다. 지역 분류 손실은 객체 검출 네트워크(210)의 학습에 이용되고, 객체 검출 네트워크(210)는 지역 분류 손실을 감소시키도록 학습을 진행할 수 있다. 특징 벡터 판별기(240)는 gradient reversal layer(GRL)를 이용하여 구현될 수 있다.
- [0089] 도 4는 본 발명의 실시 예에 의한 객체 검출 방법을 설명하기 위한 순서도이다. 도 3 및 도 4를 참조하여, 본 발명의 실시 예에 의한 객체 검출 방법을 살펴보면 다음과 같다.
- [0090] S410에서, 중첩영역 검출부(220)는 중첩영역을 검출할 수 있다. 중첩영역 검출부(220)는 제1 카메라(C1)가 획득한 제1 이미지와 제2 카메라(C2)가 획득한 제2 이미지의 중첩영역을 획득할 수 있다.
- [0091] S420에서, 스테레오 매칭부(230)는 제1 이미지의 깊이값과 제2 이미지의 깊이값의 편차를 보정할 수 있다. 깊이값의 편차는 객체 검출 네트워크가 객체를 검출하는 과정에서 이용될 수 있다.
- [0092] 스테레오 매칭부(230)는 예측된 디스패리티 맵과 실측 디스패리티 맵 간의 편차 정보를 반영한 스테레오 초점 손실을 출력할 수 있다. 예측된 디스패리티 맵은 코스트 볼륨을 통해서 획득될 수 있다. 코스트 볼륨을 획득하기 위해서, 스테레오 매칭부(230)는 백본 네트워크(211)가 출력하는 제1 이미지에 대한 제1 이미지 특징값 및 제2 이미지에 대한 제2 이미지 특징값을 매칭하여 매칭된 특징값을 생성할 수 있다. 그리고 스테레오 매칭부(230)는 멀티 스케일 레이어를 이용하여, 매칭된 특징값에 대한 멀티 스케일 특징값들을 획득하고, 멀티 스케일 특징값들을 결합하여 코스트 볼륨을 생성할 수 있다.
- [0093] 스테레오 매칭부(230)는 코스트 볼륨과 실측 디스패리티 맵(Disparity Map GT) 간의 편차에 대응하는 스테레오 초점 손실을 출력할 수 있다.
- [0094] 스테레오 초점 손실은 객체 검출 네트워크(210)를 학습하는 과정에 이용되며, 객체 검출 네트워크(210)는 스테레오 초점 손실을 감소시키도록 학습을 진행할 수 있다.
- [0095] S430에서, 특징 벡터 판별기(240)는 중첩 영역에서 검출된 객체의 특징 벡터와 비중첩영역에서 검출된 객체의 특징 벡터 간의 편차를 줄이도록 적대적 학습을 진행할 수 있다.

- [0096] 특징 벡터 판별기(240)는 객체 검출 네트워크(210)의 트랜스포머(213)로부터 클래스 정보, 위치 정보 및 특징 벡터를 제공받을 수 있다. 특징 벡터 판별기(240)는 특징 벡터를 바탕으로, 중첩영역에서 검출되는 객체의 특징 벡터와 비중첩영역에서 검출되는 객체의 특징 벡터의 편차에 비례하는 지역 분류 손실을 출력할 수 있다. 지역 분류 손실은 객체 검출 네트워크(210)의 학습에 이용되며, 객체 검출 네트워크(210)는 지역 분류 손실을 감소시키도록 학습을 진행할 수 있다.
- [0097] 이하, 도 3 및 도 4를 바탕으로 설명된 객체 검출 방법의 세부적인 동작을 살펴보면 다음과 같다.
- [0098] 도 5는 중첩영역 검출부의 동작을 설명하는 도면이다. 도 5는 제1 카메라(C1) 및 제2 카메라(C2)의 중첩영역을 검출하는 실시 예를 도시하고 있다.
- [0099] 도 5를 참조하면, 제1 카메라(C1) 및 제2 카메라(C2)는 각각 차량 외부의 이미지를 획득할 수 있다. 제1 카메라(C1)가 획득한 제1 이미지(IMG1)는 차량의 전방 좌측면 이미지일 수 있고, 제2 카메라(C2)가 획득한 제2 이미지(IMG2)는 차량의 전방 이미지일 수 있다. 제1 카메라(C1)는 제1 화각(FOV1) 범위의 이미지를 획득하며, 제2 카메라(C2)는 제2 화각(FOV2) 범위의 이미지를 획득할 수 있다. 제1 이미지(IMG1) 및 제2 이미지(IMG2)는 제1 화각(FOV1)과 제2 화각(FOV2)이 중첩되는 제1 중첩영역(OV1)을 포함할 수 있다.
- [0100] 프로세서(100)는 다음과 같이 [수학식 1] 내지 [수학식 5]를 이용하여 제1 중첩영역(OV1)을 검출할 수 있다.
- [0101] 제2 카메라(C2)가 획득한 제1 중첩영역(OV1)으로 예상되는 영역의 2D 좌표는 $X2Df=[u, v]$ 으로 표현될 수 있다. 프로세서(100)는 다음의 [수학식 1]에서와 같이 제1 중첩영역(OV1)의 2D 좌표에 깊이값을 무한으로 적용하여 3D 좌표($X3Df$)를 획득할 수 있다.
- [0102] [수학식 1]
- [0103] $X3Di=[u \times d^\infty, v \times d^\infty, d^\infty]$
- [0104] 프로세서(100)는 다음의 [수학식 2]와 같이, 3D 좌표($X3Df$)에 1을 결합(concatenate)하여 4D 좌표($X4Df$)를 획득할 수 있다.
- [0105] [수학식 2]
- [0106] $X4Df=X3Df \oplus 1$
- [0107] 프로세서(100)는 다음의 [수학식 3]와 같이, 4D 좌표($X4Df$)에 변환행렬(T)을 곱하여, 카메라 좌표에 해당하는 4D 좌표를 자기 중심 좌표(Xegof)로 변환할 수 있다.
- [0108] [수학식 3]
- [0109] $Xegof=T \cdot X4Df$
- [0110] 변환행렬(T)은 $T=\{T_f, T_{fr}, \dots, T_{br}\} \in 4 \times 4$ 와 같이 4×4 형태일 수 있다.
- [0111] 프로세서(100)는 다음의 [수학식 4]와 같이, 자기 중심 좌표(Xegof)에 역행렬을 곱하여, 제1 카메라 공간 좌표로 변환할 수 있다.
- [0112] [수학식 4]
- [0113] $X*4Dfl=T^{-1fl} \cdot Xegof$
- [0114] 프로세서(100)는 제1 카메라 공간 좌표를 픽셀의 깊이값(dx)으로 나누어, 제2 카메라(C2)의 좌표를 제1 카메라(C1)의 좌표에 투영할 수 있다.
- [0115] [수학식 5]
- [0116] $X4Dfl=X*4Dfl / dx$
- [0117] 이와 같이, 프로세서(100)는 제1 카메라(C1)와 제2 카메라(C2)가 이미지들 간의 중첩영역을 추출할 수 있다.
- [0118] 도 6은 본 발명의 실시 예에 의한 스테레오 초점 손실을 설명하는 도면이고, 도 7은 본 발명의 실시 예에 의한

스테레오 정합을 설명하는 도면이다.

- [0119] 도 6을 참조하면, 프로세서(100)는 도 5를 바탕으로 설명된 바와 같이, 제1 중첩영역 내지 제6 중첩영역들에 대한 스테레오 정합(stereo matching)을 수행할 수 있다.
- [0120] 도 7을 참조하면, 본 발명의 실시 예에 의한 스테레오 정합부는 백본 네트워크(211)가 출력하는 특징값을 바탕으로 디스패리티(Disparity)를 계산할 수 있다.
- [0121] 백본 네트워크(211)는 백본 네트워크는 이미지 분류를 위해서 미리 훈련된 모델일 수 있으며, 특징값(feature)을 추출할 수 있다. 실시 예에 의한 백본 네트워크(211)는 입력 이미지를 바탕으로 특징맵을 생성하기 위한 합성곱 신경망(Convolutional Neural Networks; CNN)을 이용할 수 있다.
- [0122] 인접하는 카메라들이 획득한 이미지들은 백본 네트워크(211)에 제공될 수 있다. 도 7은 제1 카메라(C1)가 획득한 제1 이미지(IMG1) 및 제2 카메라(C2)가 획득한 제2 이미지(IMG2)는 백본 네트워크(211)에 입력되는 것을 도시하고 있다.
- [0123] 제1 이미지(IMG1) 및 제2 이미지(IMG2)는 해상도가 1/4씩 다운샘플링(downsampling) 되어서 백본 네트워크의 레이어들에 입력될 수 있다. 예를 들어, 제1 이미지(IMG1) 및 제2 이미지(IMG2)의 해상도가 [B, C, H, W]일 경우, $W \times H$ 의 해상도를 갖는 이미지는 제1 레이어에 제공될 수 있고, $(W/2) \times (H/2)$ 의 해상도를 갖는 이미지는 제2 레이어에 제공될 수 있으며, $(W/4) \times (H/4)$ 의 해상도를 갖는 이미지는 제3 레이어에 제공될 수 있다. 이와 같이 각각의 레이어들은 다른 크기의 해상도를 갖는 이미지를 바탕으로 서로 다른 스케일(scale)의 특징값을 추출할 수 있다.
- [0124] 중첩영역에 대한 특징값은 제1 카메라(C1)와 제2 카메라(C2)들로부터 텐서(tensor)의 형태로 획득될 수 있다.
- [0125] 스테레오 매칭부(230)는 디스패리티를 예측하기 위해서 스테레오 초점 손실(Stereo focal loss)를 사용할 수 있다.
- [0126] 초점 손실(focal loss)은 단일 스테이지 검출기(one-stage detector)로 구현된 모델에서 클래스 불균형으로 인하여 객체 검출의 정확도가 낮아지는 현상을 개선할 수 있다. 초점 손실은 배경과 같이 "분류하기 쉬운(Easy Negative)" 케이스 보다 "분류하기 어려운(Hard Positive)" 케이스에 높은 가중치를 부여할 수 있다.
- [0127] 스테레오 초점 손실은 다음의 [수학식 6]과 같이 표현될 수 있다.
- [0128] [수학식 6]
- $$L_d = 1 - |P| \sum_{p \in P} \prod_{d=0}^{D-1} ((1 - P_p(d))^{-\alpha} - (-P_p(d) \log \hat{P}_p(d)))$$
- [0129]
- [0130] [수학식 6]에서, d는 예측된 디스패리티를 의미하고, D는 최대 디스패리티를 의미하며, α 는 초점 가중치를 의미할 수 있다. $P_p(d)$ 는 중첩된 영역에 대한 픽셀들의 예상된 디스패리티 분포를 의미할 수 있다. 또한, $P_p(d)$ 는 다음의 [수학식 7]에서와 같이, 예측된 디스패리티와 실측 디스패리티 간의 편차를 바탕으로 획득된 소프트맥스의 출력값 일 수 있다.
- [0131] [수학식 7]
- [0132] $P_p(d) = \text{Softmax}(2|d - d_{gt}|)$
- [0133] [수학식 7]에서 d_{gt} 는 실측 디스패리티(Ground-Truth disparity)를 의미할 수 있다.
- [0134] 도 8은 중첩영역에 대한 적대적 학습을 설명하는 도면이다. 도 9는 특징 벡터 판별기의 입력 인자를 설명하는 도면이다.
- [0135] 도 8 및 도 9를 참조하면, 본 발명의 실시 예에 의한 객체 검출 모델은 특징 벡터 판별기(240)를 포함할 수 있다.
- [0136] 특징 벡터 판별기(240)는 중첩영역과 비중첩영역의 객체의 특징 벡터들이 서로 상이한 특징값을 표현하는 것을 억제하여, 중첩영역에서 객체 검출 과정이 비중첩영역에서 객체 검출 과정과 동일한 조건으로 수행될 수 있도록 할 수 있다.
- [0137] 특징 벡터 판별기(240)는 다음의 [수학식 8]과 같이, 지역 분류 손실(classification loss)(L_r)을 이용하여 훈

련될 수 있다.

- [0138] [수학식 8]
- [0139] $L_r = E_{q, y_r} \sim D_d[\sum_{r \in R} r \log f_d(q)r]$
- [0140] f_d 는 특징 벡터 판별기(240)의 출력값을 의미할 수 있다. $E_{q, y_r} \sim D$ 는 입력 이미지의 데이터 분포(D_d)에서 추출된 샘플에 대한 기대값을 의미할 수 있다. 예를 들어, q 는 트랜스포머(213)가 출력하는 특징값일 수 있다.
- [0141] y_i 는 특징 벡터의 위치값을 의미할 수 있고, 0 또는 1로 구분될 수 있다. 예를 들어, y_1 은 제1 이미지(IMG1)에서의 특징 벡터의 위치값을 의미할 수 있다. 그리고 제1 이미지(IMG1)와 제2 이미지(IMG2)의 중첩영역에서 y_1 은 1로 결정될 수 있고, 제1 이미지(IMG1)와 제2 이미지(IMG2)의 중첩영역에서 y_1 은 0으로 결정될 수 있다.
- [0142] 특징 벡터 판별기(240)는 트랜스포머(213)가 출력하는 특징값이 중첩영역 또는 중첩영역 이외의 영역에 속하는지에 따라 클래스를 분류하도록 훈련될 수 있다.
- [0143] 또한, 객체 검출 네트워크(210)는 지역 분류 손실(L_r)을 감소시키도록 훈련될 수 있다.
- [0144] 결과적으로, 본 발명의 실시 예에 의한 객체 검출 네트워크(210)는 다음의 [수학식 9]에서와 같이, 총 손실(L_{total})이 최소가 되도록 종단 간(end-to-end) 학습을 수행할 수 있다.
- [0145] [수학식 9]
- [0146] $L_{total} = \lambda_{cls} L_{cls} + \lambda_{box} L_{box} + \lambda_d L_d + \lambda_r L_r$
- [0147] L_{cls} 는 클래스 손실 함수를 의미하고, L_{box} 는 바운딩 박스 손실 함수를 의미할 수 있다. L_d 는 스테레오 초점 손실을 의미할 수 있고, L_r 은 지역 분류 손실을 의미할 수 있다. λ_{cls} , λ_{box} , λ_d , 및 λ_r 들은 하이퍼 파라미터(hyperparameter)를 의미할 수 있다.
- [0148] 도 10은 본 발명의 실시 예에 의한 객체 검출 과정을 설명하는 도면이다.
- [0149] 도 10을 참조하면, 본 발명의 실시 예에 의한 객체 검출 과정은 입력 이미지를 이용하여 백본 네트워크(211)가 특징값을 출력하는 절차를 포함할 수 있다.
- [0150] 트랜스포머(213)는 백본 네트워크(211)가 출력한 특징값을 바탕으로 예측된 위치 정보 및 스코어를 출력할 수 있다.
- [0151] 클래스 분류기(215)는 트랜스포머(213)의 출력을 이용하여, 객체를 검출할 수 있다.
- [0152] 객체 검출 네트워크(210)는 스테레오 매칭부(230)로부터의 스테레오 초점 손실이 최소화되도록 학습된 상태이기 때문에, 중첩영역에서의 깊이값 편차에 의한 객체 검출 성능이 저하되는 것을 개선할 수 있다.
- [0153] 또한, 객체 검출 네트워크(210)는 특징 벡터 판별기(240)로부터의 지역 분류 손실이 최소화되도록 학습된 상태이기 때문에, 중첩영역에서의 객체 분류 정확도를 향상시킬 수 있다.
- [0154] 또한, 도 10에서 보는 바와 같이, 본 발명의 실시 예에 의한 객체 검출 과정은 중첩영역 검출부(220), 스테레오 매칭부(230) 및 특징 벡터 판별기(240)의 동작이 관여되지 않을 수 있다. 따라서, 객체 검출 네트워크(210)는 기존의 객체 검출 속도를 유지하면서, 객체 검출 성능을 향상시킬 수 있다.
- [0155] 이하, 본 발명의 실시 예에 의한 객체 검출 장치를 이용한 객체 검출 성능을 살펴보면 다음과 같다.
- [0156] 아래의 [표 1] 및 [표 2]는 객체 검출 성능 평가를 나타내는 것으로, nuScenes validation/test data를 사용한 실험 결과를 나타내는 표이다. NDS 및 mAP들은 평가지표 파라미터로써, 높을수록 객체 검출 성능이 좋다는 것을 나타내는 지표이다.

표 1

Model	Mono Multi	Backbone	NDS(↑)	mAP(↑)
CenterNet [53]	Mono	DLA34	0.328	0.306
FCOS3D [44]	Mono	ResNet101	0.415	0.343
PGD [43]	Mono	ResNet101	0.428	0.369
DETR3D [47]	Multi	ResNet101	0.425	0.346
DETR3D † [47]	Multi	ResNet101	0.434	0.349
Ours	Multi	ResNet101	0.441	0.364

표 2

Model	Mono Multi	Backbone	NDS(↑)	mAP(↑)
FCOS3D [44]	Mono	ResNet101	0.428	0.358
PGD [43]	Mono	ResNet101	0.448	0.386
DD3D * [31]	Mono	V2-99	0.477	0.418
DETR3D ‡ [47]	Multi	V2-99	0.479	0.412
BEVDet ‡ [17]	Multi	V2-99	0.482	0.422
Ours ‡	Multi	V2-99	0.489	0.423

[표 1] 및 [표 2]에서, Ours는 본 발명의 실시 예에 의한 실험 결과를 나타내고 있다.

[표 1] 및 [표 2]에서 보는 바와 같이, 본 발명의 실시 예에 의한 객체 검출 성능은 다른 객체 검출 모델들에 대비하여 높은 성능을 나타내는 것을 확인할 수 있다.

아래의 [표 3]은 DETR3D 모델에 대비한 클래스별 객체 검출 성능(AP)를 나타내는 것으로, nuScenes validation/test data를 사용한 실험 결과를 나타내는 표이다.

표 3

Model	Car	Truck	Bus	Trailer	C.V	Ped.	Motor.	Bicycle	T.C	Barrier	mAP
DETR3D [47]	0.603	0.333	0.290	0.358	0.170	0.455	0.413	0.308	0.627	0.565	0.412
Ours	0.609	0.338	0.323	0.347	0.174	0.467	0.420	0.311	0.649	0.589	0.423

[표 3]에서, "C.V"는 건설 차량을 의미하며, "Motor."은 모터사이클을 의미하고, "T.C"는 트래픽 콘(Traffic Cone)을 의미할 수 있다.

[표 3]에서, 보는 바와 같이, 본 발명의 실시 예에 의한 객체 검출 모델은 DETR3D 모델보다 객체 검출 성능(AP)이 높은 것을 확인할 수 있다.

아래의 [표 4]는 중첩영역에서 DETR3D 모델에 대비한 클래스별 객체 검출 성능(AP)를 나타내는 것으로, nuScenes validation/test data를 사용한 실험 결과를 나타내는 표이다. [표 4]에서, mATE, mASE, mAOE, mAVE, 및 mAAE들은 각각 평가지표 파라미터를 의미하고,

mATE는 평균 전송 에러(Average Translation Error)를 의미하고, mASE는 평균 스케일 에러(Average Scale Error)를 의미하며, mAOE는 평균 방향 에러(Average Orientation Error)를 의미하고, mAVE는 평균 속도 에러(Average Velocity Error)를 의미하며, mAAE는 평균 속성 에러(Average Attribute Error)를 의미할 수 있다. mATE, mASE, mAOE, mAVE, 및 mAAE들은 낮을수록 객체 검출 성능이 좋다는 것을 나타내는 지표일 수 있다.

표 4

Model	Mono Multi	NDS (↑)	mAP (↑)	mATE (↓)	mASE (↓)	mAOE (↓)	mAVE (↓)	mAAE (↓)
FCOS3D [41]	Mono	0.317	0.213	0.841	0.276	0.604	1.122	0.173
DETR3D [44]	Multi	0.356	0.231	0.825	0.280	0.400	0.863	0.223
Ours	Multi	0.405	0.261	0.676	0.284	0.364	0.751	0.183

[0167]

[0168]

[표 4]에서 보는 바와 같이, 본 발명의 실시 예에 의한 객체 검출 모델은 중첩영역에서 DETR3D 모델보다 객체 검출 성능(AP)이 높은 것을 확인할 수 있다.

[0169]

도 11 내지 도 12는 본 발명의 실시 예에 의한 객체 검출 성능을 실험한 결과를 나타내는 도면이다. 도 11 및 도 12에서 "GT"는 실측 데이터를 의미할 수 있고, "DETR3D"는 비교 예에 의한 객체 검출 모델을 의미할 수 있고, "Ours"는 본 발명의 실시 예에 의한 객체 검출 모델을 의미할 수 있다.

[0170]

도 11 및 도 12에서 확인할 수 있는 바와 같이, 본 발명의 실시 예에 의한 객체 검출 모델은 중첩영역에서 성능이 크게 향상되는 것을 확인할 수 있다.

[0171]

도 13은 본 발명의 실시 예에 의한 객체 검출 성능을 보여주는 도면이다.

[0172]

도 13에서 알 수 있는 바와 같이, 본 발명의 실시 예에 의하면, DETR3D 모델과 동일한 검출 속도를 유지하면서 mAP 및 NDS 지표가 높게 나타나는 것을 알 수 있다.

[0173]

또한, 본 발명의 실시 예에 의하면, BEVDet 모델보다 객체 검출 속도 및 객체 검출 성능이 모두 향상된 것을 확인할 수 있다.

[0174]

도 14는 본 발명의 일 실시 예에 따른 컴퓨팅 시스템을 도시한다.

[0175]

도 14를 참조하면, 컴퓨팅 시스템(1000)은 버스(1200)를 통해 연결되는 적어도 하나의 프로세서(1100), 메모리(1300), 사용자 인터페이스 입력 장치(1400), 사용자 인터페이스 출력 장치(1500), 스토리지(1600), 및 네트워크 인터페이스(1700)를 포함할 수 있다.

[0176]

프로세서(1100)는 중앙 처리 장치(CPU) 또는 메모리(1300) 및/또는 스토리지(1600)에 저장된 명령어들에 대한 처리를 실행하는 반도체 장치일 수 있다. 메모리(1300) 및 스토리지(1600)는 다양한 종류의 휘발성 또는 불휘발성 저장 매체를 포함할 수 있다. 예를 들어, 메모리(1300)는 ROM(Read Only Memory) 및 RAM(Random Access Memory)을 포함할 수 있다.

[0177]

따라서, 본 명세서에 개시된 실시예들과 관련하여 설명된 방법 또는 알고리즘의 단계는 프로세서(1100)에 의해 실행되는 하드웨어, 소프트웨어 모듈, 또는 그 2 개의 결합으로 직접 구현될 수 있다. 소프트웨어 모듈은 RAM 메모리, 플래시 메모리, ROM 메모리, EPROM 메모리, EEPROM 메모리, 레지스터, 하드 디스크, 착탈형 디스크, CD-ROM과 같은 저장 매체(즉, 메모리(1300) 및/또는 스토리지(1600))에 상주할 수도 있다.

[0178]

예시적인 저장 매체는 프로세서(1100)에 커플링되며, 그 프로세서(1100)는 저장 매체로부터 정보를 판독할 수 있고 저장 매체에 정보를 기입할 수 있다. 다른 방법으로, 저장 매체는 프로세서(1100)와 일체형일 수도 있다. 프로세서 및 저장 매체는 주문형 집적회로(ASIC) 내에 상주할 수도 있다. ASIC는 사용자 단말기 내에 상주할 수도 있다. 다른 방법으로, 프로세서 및 저장 매체는 사용자 단말기 내에 개별 컴포넌트로서 상주할 수도 있다.

[0179]

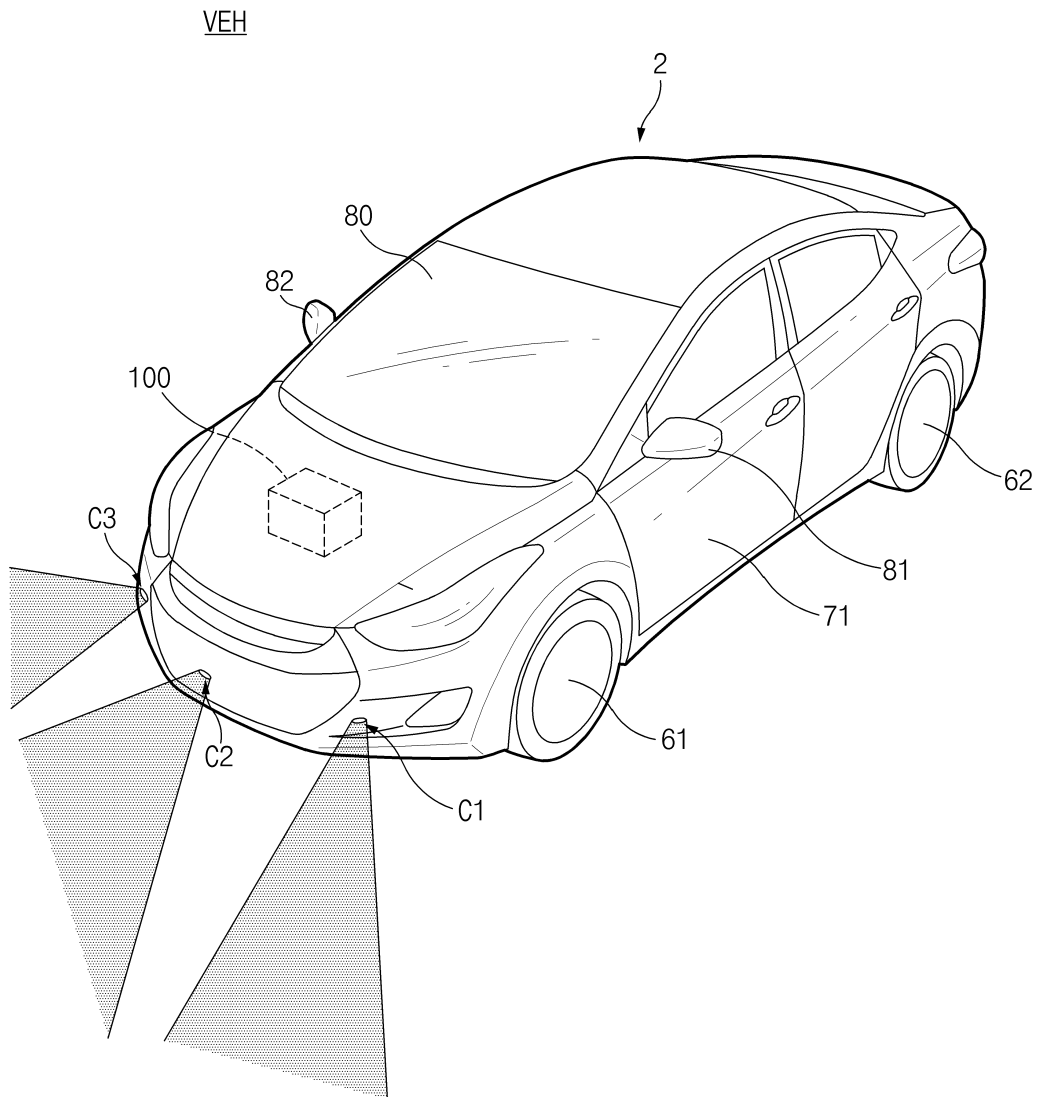
이상의 설명은 본 발명의 기술 사상을 예시적으로 설명한 것에 불과한 것으로서, 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자라면 본 발명의 본질적인 특성에서 벗어나지 않는 범위에서 다양한 수정 및 변형이 가능할 것이다.

[0180]

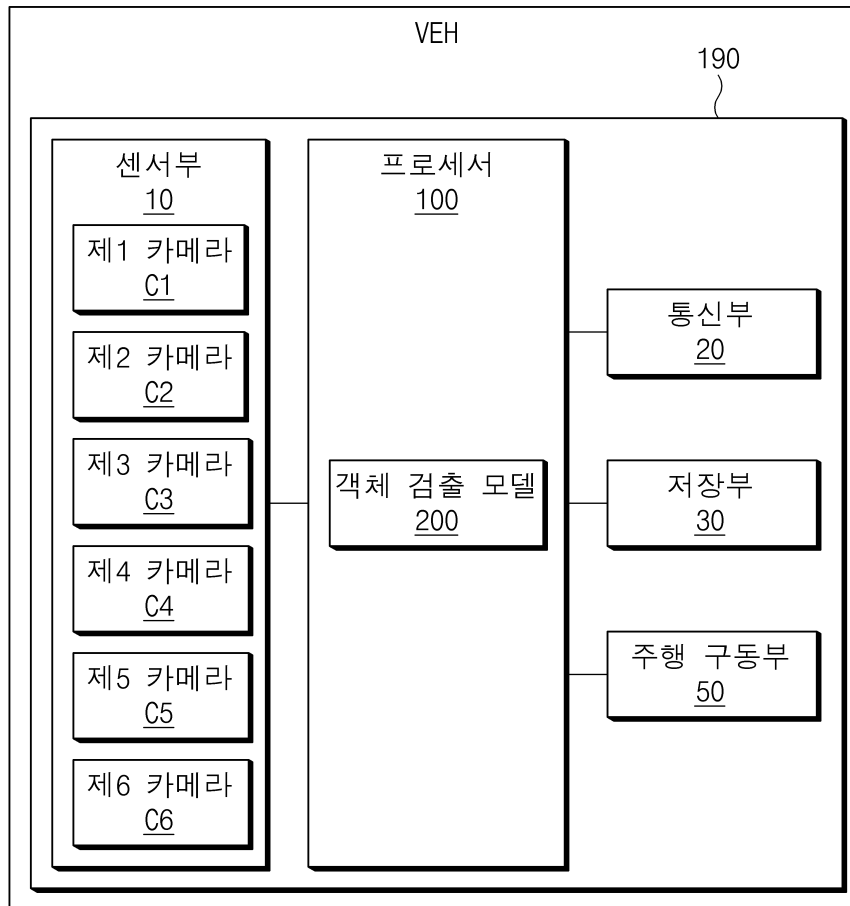
따라서, 본 발명에 개시된 실시예들은 본 발명의 기술 사상을 한정하기 위한 것이 아니라 설명하기 위한 것이고, 이러한 실시예에 의하여 본 발명의 기술 사상의 범위가 한정되는 것은 아니다. 본 발명의 보호 범위는 아래의 청구범위에 의하여 해석되어야 하며, 그와 동등한 범위 내에 있는 모든 기술 사상은 본 발명의 권리범위에 포함되는 것으로 해석되어야 할 것이다.

도면

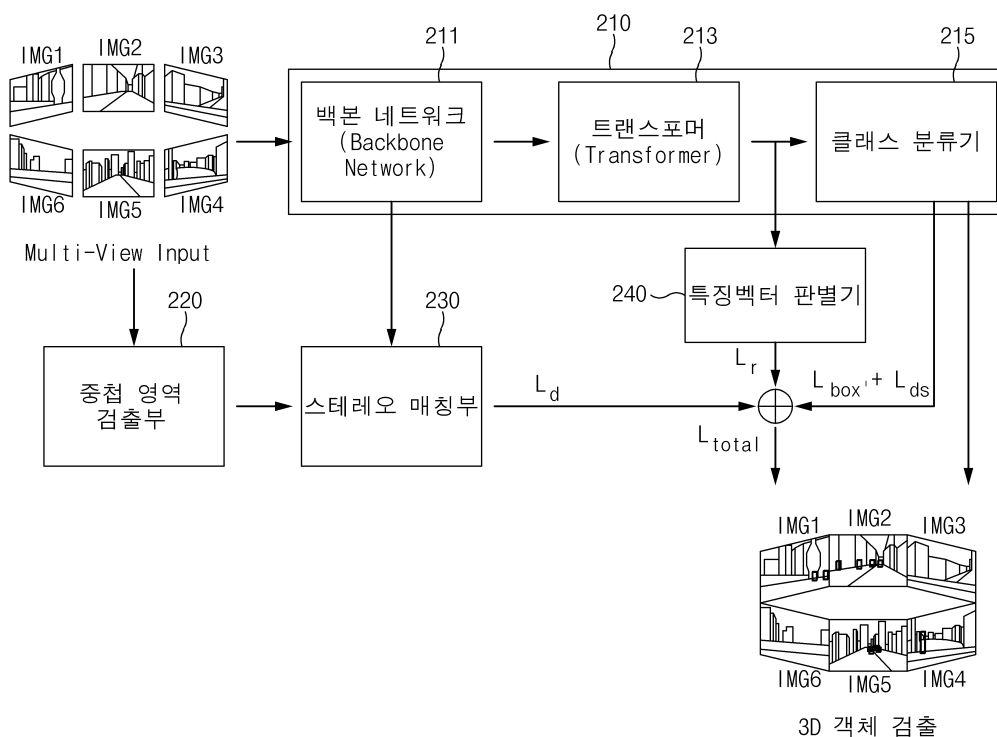
도면1



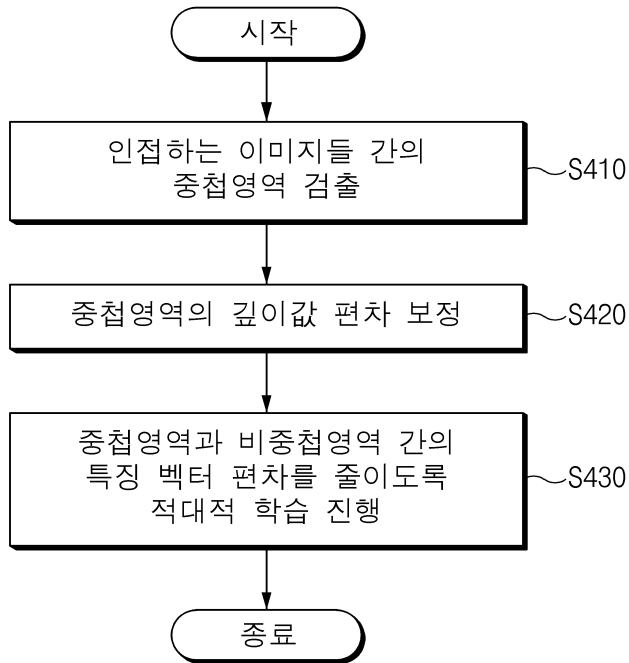
도면2



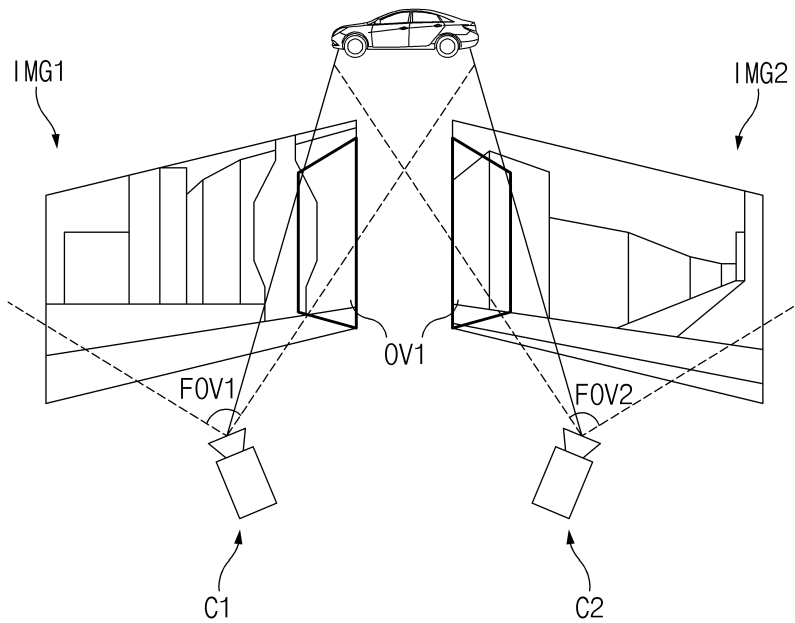
도면3



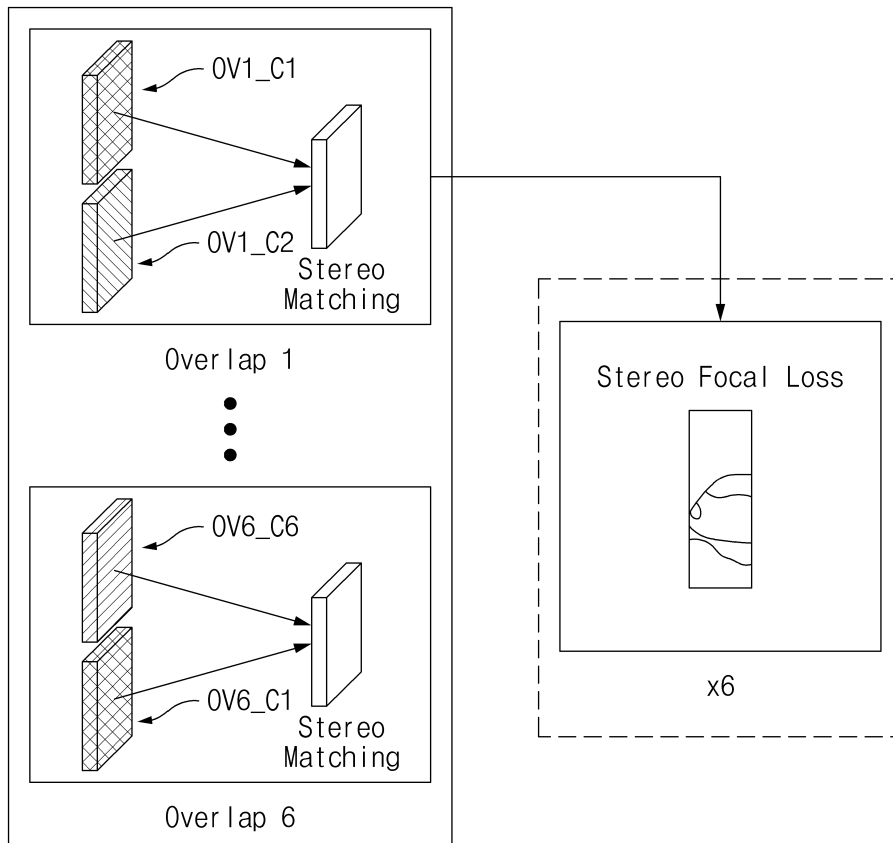
도면4



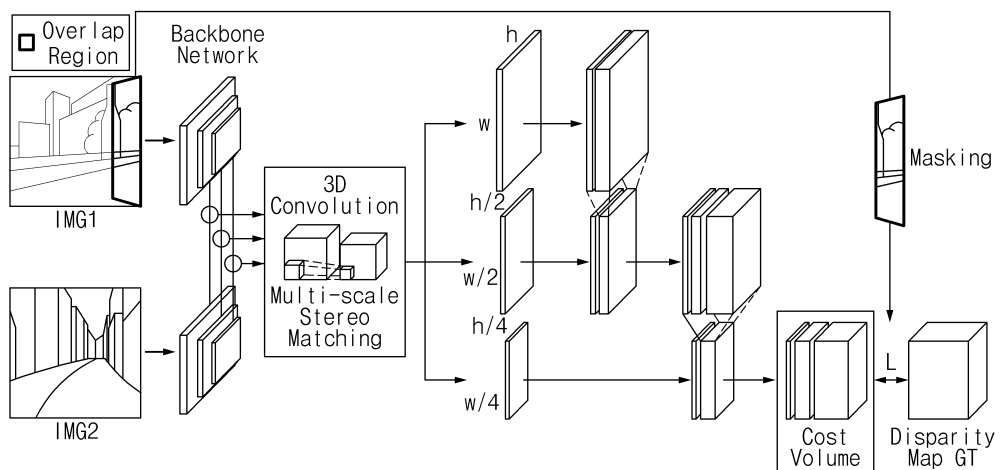
도면5



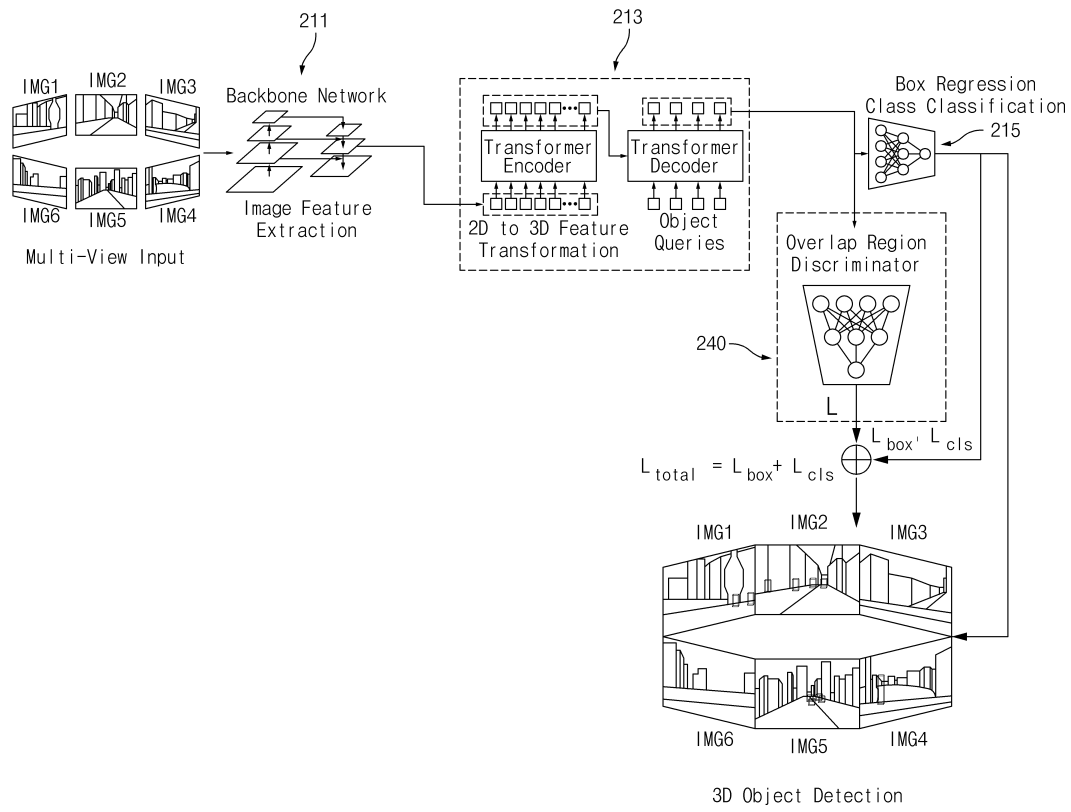
도면6



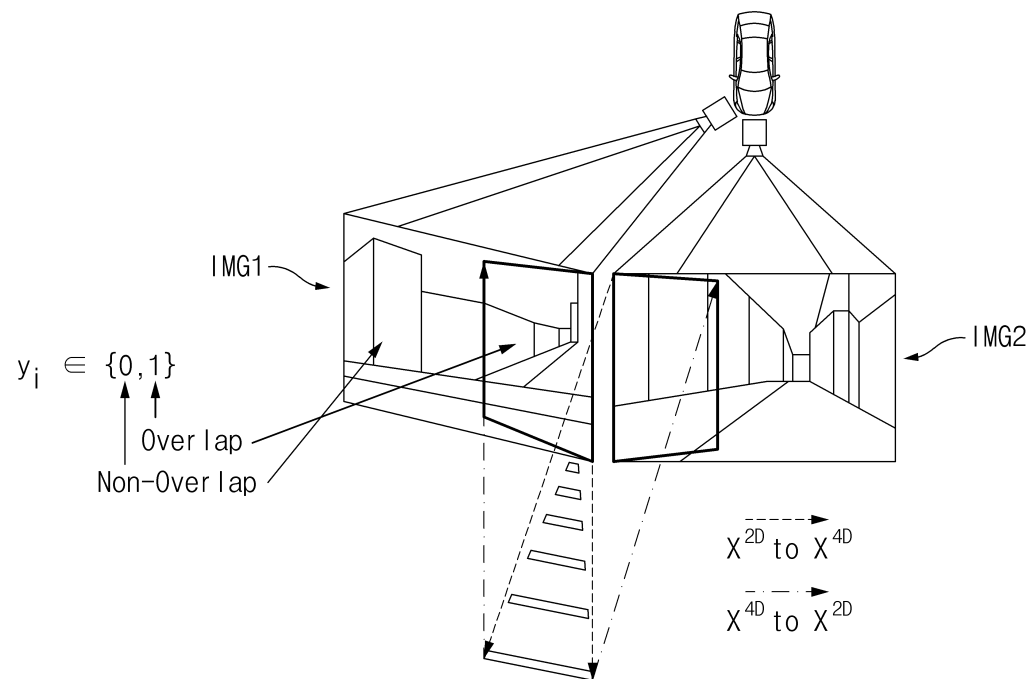
도면7



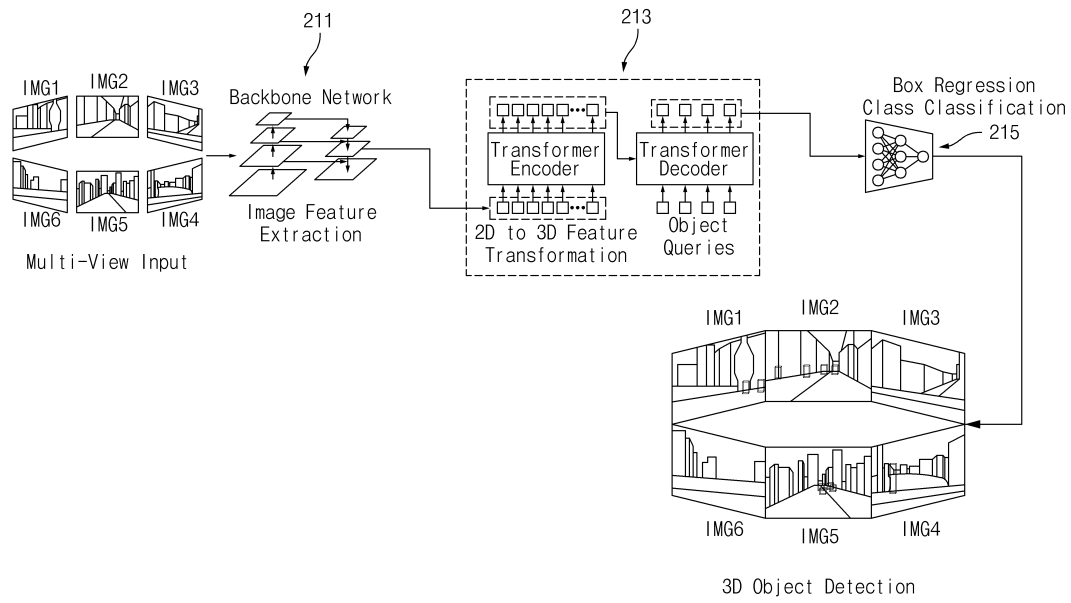
도면8



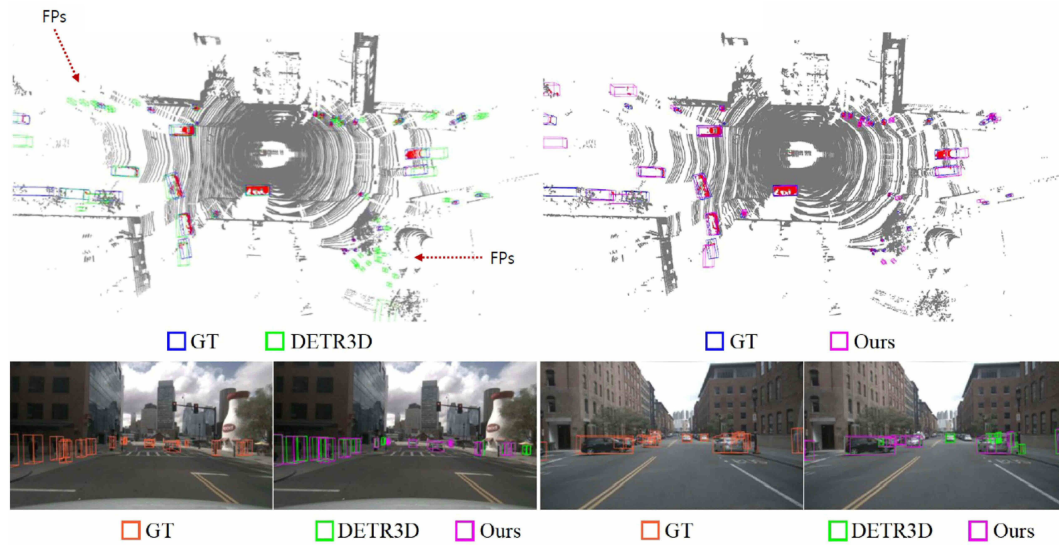
도면9



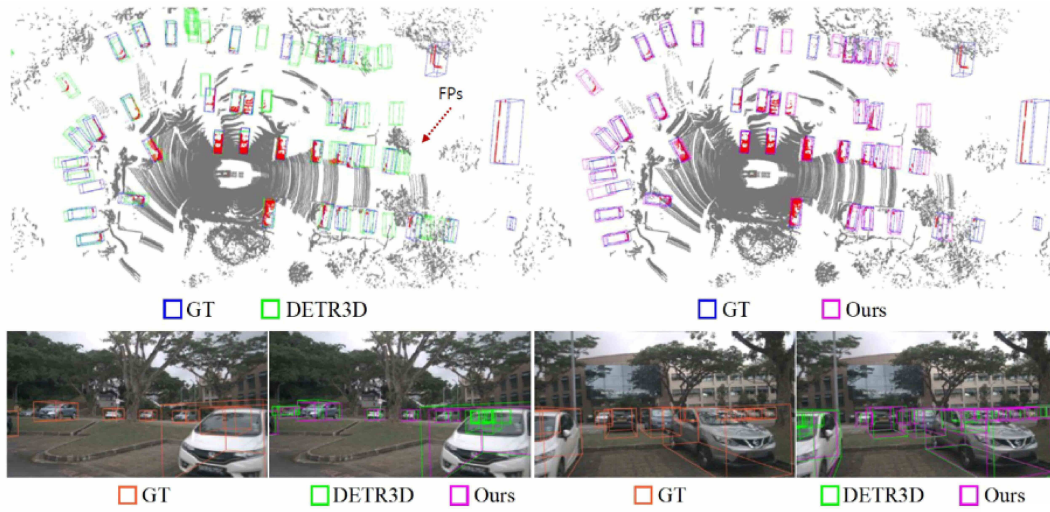
도면10



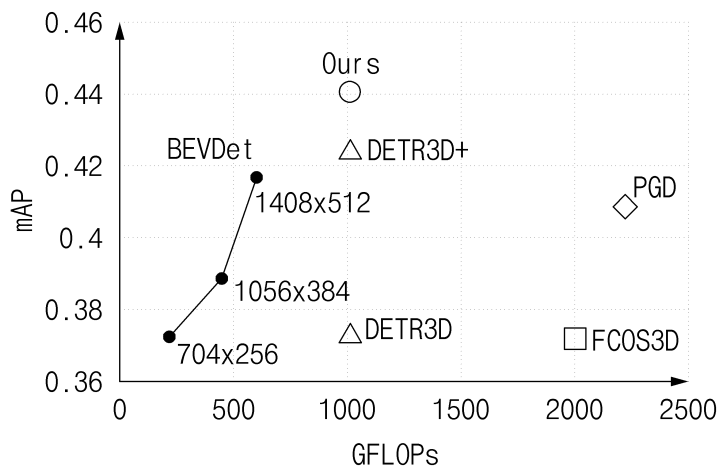
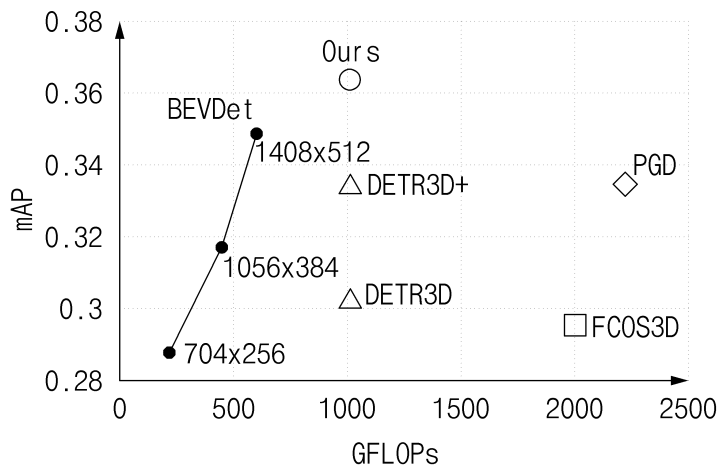
도면11



도면12



도면13



도면14

