



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2024년07월04일
(11) 등록번호 10-2681992
(24) 등록일자 2024년07월02일

(51) 국제특허분류(Int. Cl.)
G01S 17/894 (2020.01) B60W 40/02 (2006.01)
G01S 17/931 (2020.01) G01S 7/48 (2006.01)
G05D 1/20 (2024.01) G06F 18/00 (2023.01)
G06N 3/08 (2023.01) G06T 7/11 (2017.01)
(52) CPC특허분류
G01S 17/894 (2022.01)
B60W 40/02 (2013.01)
(21) 출원번호 10-2021-0108154
(22) 출원일자 2021년08월17일
심사청구일자 2021년08월17일
(65) 공개번호 10-2023-0026130
(43) 공개일자 2023년02월24일
(56) 선행기술조사문헌
JP2017166971 A*
(뒷면에 계속)

(73) 특허권자
충북대학교 산학협력단
충청북도 청주시 서원구 충대로 1 (개신동)
(72) 발명자
김곤우
대전광역시 유성구 은구비남로 56, 1202호(열매마을아파트9단지)
홍록듀이
충청북도 청주시 흥덕구 성봉로271번길 15-11, 102호
(74) 대리인
김정현

전체 청구항 수 : 총 2 항

심사관 : 임일순

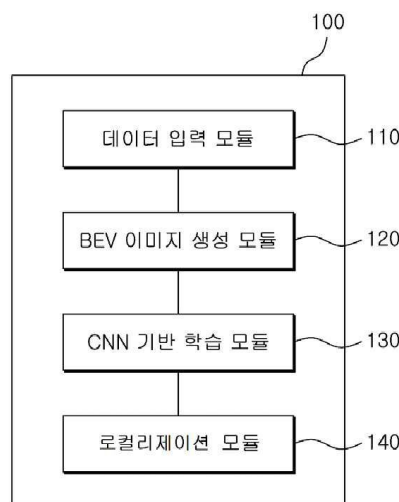
(54) 발명의 명칭 자율 주행을 위한 단일 계층 3차원 다중 객체 검출 장치 및 방법

(57) 요약

본 발명은 라이다 센서를 이용하여 3차원 다중 객체를 검출하기 위한 단일 계층 3차원 다중 객체 검출 장치에 관한 것으로서, 라이다 센서로부터 로(raw) 포인트 클라우드 데이터를 입력받기 위한 데이터 입력 모듈, 상기 로 포인트 클라우드 데이터에서 BEV(Bird's Eye View) 이미지를 생성하기 위한 BEV 이미지 생성 모듈, 상기 BEV 이미지로부터 세분화된 특징 이미지를 추출하기 위해 딥 러닝(deep learning) 알고리즘 기반 학습을 수행하는 학습 모듈 및 상기 세분화된 특징 이미지에서 3차원 객체를 검출하기 위한 3D 후보 박스와 그에 대응하는 클래스를 찾기 위한 회귀(regression) 작업과 로컬리제이션(localization) 작업을 수행하는 로컬리제이션 모듈을 포함한다.

본 발명에 의하면, 자율 주행을 위한 강력한 실시간 3차원 다중 물체 검출 장치를 제안함으로써, 매우 빠른 추론 속도를 유지하면서 3D 물체 감지 작업의 정확성을 향상시킬 수 있는 효과가 있다.

대표도 - 도1



(52) CPC특허분류

G01S 17/931 (2020.01)
G01S 7/4802 (2013.01)
G05D 1/021 (2013.01)
G06N 3/08 (2023.01)
G06T 7/11 (2017.01)
G06V 20/56 (2023.08)
B60W 2420/408 (2024.01)
G06T 2207/10028 (2013.01)

(56) 선행기술조사문헌

JP2019532433 A*
 JP2020042009 A*
 JP2021082296 A*
 W02020253121 A1*
 *는 심사관에 의하여 인용된 문헌

이 발명을 지원한 국가연구개발사업

과제고유번호	1711125911
과제번호	2020-0-01462-002
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	정보통신방송혁신인재양성(R&D)
연구과제명	Grand ICT 연구센터(충북대)
기 여 율	1/1
과제수행기관명	충북대학교산학협력단
연구기간	2021.01.01 ~ 2021.12.31

명세서

청구범위

청구항 1

라이다 센서를 이용하여 3차원 다중 객체를 검출하기 위한 단일 계층 3차원 다중 객체 검출 장치에서,

라이다 센서로부터 로(raw) 포인트 클라우드 데이터를 입력받기 위한 데이터 입력 모듈;

상기 로 포인트 클라우드 데이터에서 BEV(Bird's Eye View) 이미지를 생성하기 위한 BEV 이미지 생성 모듈;

상기 BEV 이미지로부터 세분화된 특징 이미지를 추출하기 위해 딥 러닝(deep learning) 알고리즘 기반 학습을 수행하는 학습 모듈; 및

상기 세분화된 특징 이미지에서 3차원 객체를 검출하기 위한 3D 후보 박스와 그에 대응하는 클래스를 찾기 위한 회귀(regression) 작업과 로컬리제이션(localization) 작업을 수행하는 로컬리제이션 모듈을 포함하며,

상기 BEV 이미지 생성 모듈은 3D 로 포인트 클라우드 데이터를 2D 의사(pseudo) 이미지로 투영하고 이산화하는 방식으로 BEV 이미지를 생성하고,

상기 BEV 이미지 생성 모듈은 3D 로 포인트 클라우드 데이터를 인코딩하여 높이 특징(Height feature), 밀도 특징(Density feature), 강도 특징(Intensity feature) 및 거리 특징(Distance feature)을 포함하는 4개의 특징 맵 이미지를 생성하고,

상기 학습 모듈은 CNN(Convolutional Neural Network) 기반 학습을 수행하며,

상기 높이 특징 맵을 생성함에 있어서, 각 셀 값이 셀 내의 포인트 높이 중 최대 높이로 계산되고, 정규화 단계를 적용하여 정규화된 높이 특징 맵을 얻고,

상기 밀도 특징 맵은 3D 현실 세계에서 다양한 포인트 클라우드 분포와 함께 셀 내의 포인트 밀도를 나타내고,

상기 강도 특징 맵은 라이다(LiDAR) 강도가 기록되고, [0,1] 사이의 물체 표면 값을 반영하는 레이저빔의 반환 강도가 기록되며, 기준 셀에서 최대 높이를 갖는 점의 로(raw) 반사율 값이 기록되고,

상기 거리 특징 맵에서 각 셀의 정규화된 거리 특성 D_{i_norm} 은,

$$D_{i_norm} = \frac{D_{0 \rightarrow P_i}}{D_{max}} = \frac{\sqrt{(x_{P_i}^2 + y_{P_i}^2 + z_{P_i}^2)}}{\sqrt{(x_{max}^2 + y_{max}^2 + z_{max}^2)}} \quad \text{에 의해 계산되고,}$$

여기서 $D_{0 \rightarrow P_i}$ 는 LIDAR 원점(0,0,1.73m)과 현재 점 P_i 사이의 거리이고, D_{max} 는 조사 영역(investigated area) Ψ 내에서 가장 먼 점 P_{max} 까지의 LIDAR 원점이고, $(x_{P_i}, y_{P_i}, z_{P_i})$ 및 $(x_{max}, y_{max}, z_{max})$ 는 각각 점 P_i 및 P_{max} 의 위치이고,

상기 학습 모듈은 백본 네트워크와 헤더 네트워크로 구성된 CNN 기반 학습을 수행하고,

상기 백본 네트워크는 로(raw) BEV 표현에서 일반 정보를 컨볼루션 특징 맵 형태로 회수하는 데 사용되고, 커널(kernel), 다운 샘플링 비(down-sampling ratio) 및 반복양(quantity of repetition)을 나타내는 수정된 Resnet 블록 모듈(Res_Block), 각 스케일의 업 샘플링(up-sampling) 블록을 나타내는 모듈(US_Block) 및 다운 샘플링 모듈(DS_Block)의 3개의 서브 모듈을 포함하고,

상기 헤더 네트워크는 작업별 예측을 학습하도록 설계되고, 객체 중심 포인트(x,y), 오프셋 정보($\Delta x, \Delta y$), 확장 Z 좌표(z), 객체 크기(1,w,h), 객체 회전 각도(yaw)를 포함하는 5개의 하위 작업(sub-task)이 있고, 학습 태스크를 위한 객체 중심(Object Center), 오프셋(offset), Z 차원(Z dimension), 3D 객체 크기(3D Size) 및 회전 각(Rotation angle)의 5개의 특징을 활용하는 헤드 모듈(Head)을 포함하여 이루어지는 것을 특징으로 하는 단일 계층 3차원 다중 객체 검출 장치.

청구항 2

삭제

청구항 3

삭제

청구항 4

삭제

청구항 5

라이다 센서를 이용하여 3차원 다중 객체를 검출하기 위한 단일 계층 3차원 다중 객체 검출 장치에서의 단일 계층 3차원 다중 객체 검출 방법에서,

라이다 센서로부터 로(raw) 포인트 클라우드 데이터를 입력받기 위한 데이터 입력 단계;

상기 로 포인트 클라우드 데이터에서 BEV(Bird's Eye View) 이미지를 생성하기 위한 BEV 이미지 생성 단계;

상기 BEV 이미지로부터 세분화된 특징 이미지를 추출하기 위해 딥 러닝(deep learning) 알고리즘 기반 학습을 수행하는 학습 단계; 및

상기 세분화된 특징 이미지에서 3차원 객체를 검출하기 위한 3D 후보 박스와 그에 대응하는 클래스를 찾기 위한 회귀(regression) 작업과 로컬리제이션(localization) 작업을 수행하는 로컬리제이션 단계를 포함하고,

상기 BEV 이미지 생성 단계에서 3D 로 포인트 클라우드 데이터를 2D 의사(pseudo) 이미지로 투영하고 이산화하는 방식으로 BEV 이미지를 생성하고,

상기 BEV 이미지 생성 단계에서 3D 로 포인트 클라우드 데이터를 인코딩하여 높이, 밀도, 강도 및 거리의 4개의 특징 맵 이미지를 생성하고,

상기 학습 단계에서 CNN(Convolutional Neural Network) 기반 학습을 수행하며,

상기 높이 특징 맵을 생성함에 있어서, 각 셀 값이 셀 내의 포인트 높이 중 최대 높이로 계산되고, 정규화 단계를 적용하여 정규화된 높이 특징 맵을 얻고,

상기 밀도 특징 맵은 3D 현실 세계에서 다양한 포인트 클라우드 분포와 함께 셀 내의 포인트 밀도를 나타내고,

상기 강도 특징 맵은 라이다(LiDAR) 강도가 기록되고, [0,1] 사이의 물체 표면 값을 반영하는 레이저빔의 반환 강도가 기록되며, 기준 셀에서 최대 높이를 갖는 점의 로(raw) 반사율 값이 기록되고,

상기 거리 특징 맵에서 각 셀의 정규화된 거리 특성 D_{i_norm} 은,

$$D_{i_norm} = \frac{D_{O \rightarrow P_i}}{D_{max}} = \frac{\sqrt{(x_{P_i}^2 + y_{P_i}^2 + z_{P_i}^2)}}{\sqrt{(x_{max}^2 + y_{max}^2 + z_{max}^2)}} \quad \text{에 의해 계산되고,}$$

여기서 $D_{O \rightarrow P_i}$ 는 LIDAR 원점(0,0,1.73m)과 현재 점 P_i 사이의 거리이고, D_{max} 는 조사 영역(investigated area) Ψ 내에서 가장 먼 점 P_{max} 까지의 LIDAR 원점이고, $(x_{P_i}, y_{P_i}, z_{P_i})$ 및 $(x_{max}, y_{max}, z_{max})$ 는 각각 점 P_i 및 P_{max} 의 위치이고,

상기 학습 모듈은 백본 네트워크와 헤더 네트워크로 구성된 CNN 기반 학습을 수행하고,

상기 백본 네트워크는 로(raw) BEV 표현에서 일반 정보를 컨볼루션 특징 맵 형태로 회수하는 데 사용되고, 커널(kernel), 다운 샘플링 비(down-sampling ratio) 및 반복양(quantity of repetition)을 나타내는 수정된 Resnet 블록 모듈(Res_Block), 각 스케일의 업 샘플링(up-sampling) 블록을 나타내는 모듈(US_Block) 및 다운 샘플링 모듈(DS_Block)의 3개의 서브 모듈을 포함하고,

상기 헤더 네트워크는 작업별 예측을 학습하도록 설계되고, 객체 중심 포인트(x,y), 오프셋 정보($\Delta x, \Delta y$), 확장 Z 좌표(z), 객체 크기(l,w,h), 객체 회전 각도(yaw)를 포함하는 5개의 하위 작업(sub-task)이 있고, 학습 태

스크를 위한 객체 중심(Object Center), 오프셋(offset), Z 차원(Z dimension), 3D 객체 크기(3D Size) 및 회전 각(Rotation angle)의 5개의 특징을 활용하는 헤드 모듈(Head)을 포함하여 이루어지는 것을 특징으로 하는 단일 계층 3차원 다중 객체 검출 방법.

청구항 6

삭제

청구항 7

삭제

청구항 8

삭제

발명의 설명

기술 분야

[0001] 본 발명은 다중 객체 검출 장치에 관한 것으로서, 더욱 상세하게는 자율 주행을 위한 효율적이고 간결한 단일 계층 3차원 다중 객체 검출 장치에 관한 것이다.

배경 기술

[0002] 무인 차량의 자율 주행을 위해서는 전방의 동적 물체를 탐지하여 물체의 동적 움직임을 추정된 후 자율 주행 경로를 생성하여야 한다. 현재 레이더, 카메라 등을 이용한 동적 물체 탐지 추적 기법에 많은 연구가 진행되고 있으며 최근 레이저 스캐너의 가격 하락에 힘입어 일반 자동차 업체들 또한 운전자 보조 시스템(driver-assistance system) 탑재가 대중화 되고 있다.

[0003] 레이저 스캐너를 이용한 이동물체 탐지를 하기 위해서는 각각의 레이저 포인트를 깊이값으로 환산하여 탑재차량의 주위에 대한 포인트 클라우드(Point cloud)를 생성하게 된다. 생성된 포인트 클라우드에서 각각의 포인트는 어떠한 의미를 가지고 있지 않기 때문에 이동물체 탐지 추적을 위해서는 먼저 클러스터링 기법을 통하여 포인트들을 묶어서 하나의 물체로 표현한다.

[0004] 이처럼, 자율 주행에서 환경 인식은 필수적이며 복잡한 도시 시나리오와 같이 복잡한 환경에서 견고성을 요구한다.

[0005] 운전자가 관리할 수 있는 모든 도로 및 환경 조건에서 모든 운전 업무를 수행할 수 있는 자동화된 주행 시스템은 국제 SAE(Society of Automotive Engineers)에서 최고 수준의 자동화로 분류된다. ADA(Advanced Driving Assists)는 상업적으로 이용 가능하지만 인간의 개입이 필요하거나 특정 환경 조건에서 작동하는 경우도 있다. 이러한 자율성의 실현은 다중 물체 인식 및 추적(Multiple Object Detection and Tracking, MODT)과 같은 관련 연구 영역에 대해 거대한 요구 사항을 제시하고, 주변 환경에서 공존하는 엔티티의 동적 속성을 이해하는 것이 전반적인 자동화를 향상시키는 데 중요하다. 이는 로컬라이제이션(localization), 매핑(mapping) 및 모션 계획(motion planning)의 품질에 직접적인 영향을 미친다.

[0006] 지난 10 년 동안 카메라에 대한 인식을 통해 수많은 MODT 접근법이 전통적으로 연구되어 왔으며, 이에 대한 자세한 검토가 이루어졌다. 이에 의하면, 물체는 2D 좌표계 또는 스테레오 설정의 3D 좌표계에서, 카메라 기준 프레임에 감지되어 각각 2D 또는 3D 궤적을 생성한다. 그러나, 정확하지 않은 카메라 기하학을 이용하여 공간 정보가 산출되며, FOV(Field of view)는 제한적이다. 또한, 카메라 기반 접근법은 객체 절단, 조명 조건, 고속 타겟, 센서 모션 및 타겟 간의 상호 작용을 포함하여 다양한 문제에 직면하고 있다.

[0007] 자율 주행에서 3D 객체 좌표는 위치 정확도와 견고성이 있어야 하며, 대부분의 객체 감지 장치는 자율 차량에 내장되어 동작한다. 이러한 제약을 충족하기 위해서 완전 자율주행 임베디드 시스템의 맥락에서 효율적이고 컴팩트한 3D 감지 프레임워크가 필요하다. 따라서 포인트 클라우드에서 소형 3D 물체를 감지하는 경우, 임베디드 시스템 친화적으로 자율 주행 시스템을 구현하는 것이 중요하다.

[0008] 최근, 넓은 파노라마 배경 정보를 제공하는 라이다(Light Detector and Ranging, LiDAR)기술이 대안 기술로서 점차 대중화되고 있다. 라이다는 10-15 Hz의 합리적인 속도로 최대 100m에 이르는 넓은 파노라마 측정을 제공하기 때문에, MODT 작업을 위한 이상적인 센서이다.

[0009] 다양한 센서 중에서 LIDAR는 3D 객체 검출 작업에 이상적인 센서로 사용되고 있으며, 많은 모바일 로봇 응용 프로그램, 특히 자율 주행에서 로봇 비전에 유비쿼터스한 3D 포인트 클라우드를 제공한다. 또한 LIDAR는 시각 정보와 달리 3D 현실 세계의 불균일한 샘플링, 유효 작동 범위, 폐색, 노이즈 및 시각 센서에서 제한되는 모든 날씨에 대한 상대 포즈와 같은 요인으로 인해 매우 희소한 점 밀도 분포를 제공한다.

선행기술문헌

특허문헌

[0010] (특허문헌 0001) 대한민국 등록특허 10-1655606

발명의 내용

해결하려는 과제

[0011] 본 발명은 상기와 같은 문제점을 해결하기 위하여 안출된 것으로서, 포인트 클라우드 프로젝션 방법과 앵커 프리 방법을 이용하여, 성능 면에서 기존의 포인트 클라우드 프로젝션 방법보다 향상된 작고 효율적인 3D 객체 검출 장치 프레임워크를 제공하는데 그 목적이 있다.

[0012] 본 발명의 목적은 이상에서 언급한 목적으로 제한되지 않으며, 언급되지 않은 또 다른 목적들은 아래의 기재로부터 통상의 기술자에게 명확하게 이해될 수 있을 것이다.

과제의 해결 수단

[0013] 이와 같은 목적을 달성하기 위한 본 발명은 라이다 센서를 이용하여 3차원 다중 객체를 검출하기 위한 단일 계층 3차원 다중 객체 검출 장치에 관한 것으로서, 라이다 센서로부터 로(raw) 포인트 클라우드 데이터를 입력받기 위한 데이터 입력 모듈, 상기 로 포인트 클라우드 데이터에서 BEV(Bird's Eye View) 이미지를 생성하기 위한 BEV 이미지 생성 모듈, 상기 BEV 이미지로부터 세분화된 특징 이미지를 추출하기 위해 딥 러닝(deep learning) 알고리즘 기반 학습을 수행하는 학습 모듈 및 상기 세분화된 특징 이미지에서 3차원 객체를 검출하기 위한 3D 후보 박스와 그에 대응하는 클래스를 찾기 위한 회귀(regression) 작업과 로컬리제이션(localization) 작업을 수행하는 로컬리제이션 모듈을 포함한다.

[0014] 상기 BEV 이미지 생성 모듈은 3D 로 포인트 클라우드 데이터를 2D 의사(pseudo) 이미지로 투영하고 이산화하는 방식으로 BEV 이미지를 생성할 수 있다.

[0015] 상기 BEV 이미지 생성 모듈은 3D 로 포인트 클라우드 데이터를 인코딩하여 높이, 밀도, 강도 및 거리의 4개의 특징 맵 이미지를 생성할 수 있다.

[0016] 상기 학습 모듈은 CNN(Convolutional Neural Network) 기반 학습을 수행할 수 있다.

[0017] 본 발명의 라이다 센서를 이용하여 3차원 다중 객체를 검출하기 위한 단일 계층 3차원 다중 객체 검출 장치에서의 단일 계층 3차원 다중 객체 검출 방법에서, 라이다 센서로부터 로(raw) 포인트 클라우드 데이터를 입력받기 위한 데이터 입력 단계, 상기 로 포인트 클라우드 데이터에서 BEV(Bird's Eye View) 이미지를 생성하기 위한 BEV 이미지 생성 단계, 상기 BEV 이미지로부터 세분화된 특징 이미지를 추출하기 위해 딥 러닝(deep learning) 알고리즘 기반 학습을 수행하는 학습 단계 및 상기 세분화된 특징 이미지에서 3차원 객체를 검출하기 위한 3D 후보 박스와 그에 대응하는 클래스를 찾기 위한 회귀(regression) 작업과 로컬리제이션(localization) 작업을 수행하는 로컬리제이션 단계를 포함한다.

[0018] 상기 BEV 이미지 생성 단계에서 3D 로 포인트 클라우드 데이터를 2D 의사(pseudo) 이미지로 투영하고 이산화하는 방식으로 BEV 이미지를 생성할 수 있다.

[0019] 상기 BEV 이미지 생성 단계에서 3D 로 포인트 클라우드 데이터를 인코딩하여 높이, 밀도, 강도 및 거리의 4개의

특징 맵 이미지를 생성할 수 있다.

[0020] 상기 학습 단계에서 CNN(Convolutional Neural Network) 기반 학습을 수행할 수 있다.

발명의 효과

[0021] 본 발명에 의하면, 자율 주행을 위한 강력한 실시간 3차원 다중 물체 검출 장치를 제안함으로써, 매우 빠른 추론 속도를 유지하면서 3D 물체 감지 작업의 정확성을 향상시킬 수 있는 효과가 있다.

도면의 간단한 설명

[0022] 도 1은 본 발명의 일 실시예에 따른 3차원 다중 객체 검출 장치의 내부 구성을 보여주는 블록도이다.

도 2는 본 발명의 일 실시예에 따른 3차원 다중 객체 검출 방법을 보여주는 흐름도이다.

도 3은 본 발명의 일 실시예에 따른 3차원 다중 객체 검출 장치의 전체 프레임워크를 도시한 것이다.

도 4는 BEV(Bird's Eye View) 특징 맵 생성의 상세 구조를 도시한 것이다.

도 5는 훈련 데이터 세트 샘플에서 관심 영역의 평균 포인트 클라우드 분포를 나타낸 그래프이다.

도 6은 본 발명의 일 실시예에 따른 3차원 다중 객체 검출 장치의 상세한 CNN 아키텍처(architecture)를 도시한 것이다.

발명을 실시하기 위한 구체적인 내용

[0023] 본 명세서에서 개시된 실시 예의 이점 및 특징, 그리고 그것들을 달성하는 방법은 첨부되는 도면과 함께 후술되어 있는 실시 예들을 참조하면 명확해질 것이다. 그러나 본 개시에서 제안하고자 하는 실시 예는 이하에서 개시되는 실시 예들에 한정되는 것이 아니라 서로 다른 다양한 형태로 구현될 수 있으며, 단지 본 실시 예들은 당해 기술분야에서 통상의 지식을 가진 자에게 실시 예들의 범주를 완전하게 알려주기 위해 제공되는 것일 뿐이다.

[0024] 본 명세서에서 사용되는 용어에 대해 간략히 설명하고, 개시된 실시 예에 대해 구체적으로 설명하기로 한다.

[0025] 본 명세서에서 사용되는 용어는 개시된 실시 예들의 기능을 고려하면서 가능한 현재 널리 사용되는 일반적인 용어들을 선택하였으나, 이는 관련 분야에 종사하는 기술자의 의도 또는 관례, 새로운 기술의 출현 등에 따라 달라질 수 있다. 또한, 특정한 경우는 출원인이 임의로 선정한 용어도 있으며, 이 경우 해당되는 명세서의 상세한 설명 부분에서 상세히 그 의미를 기재할 것이다. 따라서 본 개시에서 사용되는 용어는 단순한 용어의 명칭이 아닌, 그 용어가 가지는 의미와 본 명세서의 전반에 걸친 내용을 토대로 정의되어야 한다.

[0026] 본 명세서에서의 단수의 표현은 문맥상 명백하게 단수인 것으로 특정하지 않는 한, 복수의 표현을 포함한다.

[0027] 명세서 전체에서 어떤 부분이 어떤 구성요소를 "포함"한다고 할 때, 이는 특별히 반대되는 기재가 없는 한 다른 구성요소를 제외하는 것이 아니라 다른 구성요소를 더 포함할 수 있음을 의미한다. 또한, 명세서에서 사용되는 "부"라는 용어는 소프트웨어, FPGA 또는 ASIC과 같은 하드웨어 구성요소를 의미하며, "부"는 어떤 역할들을 수행한다. 그렇지만 "부"는 소프트웨어 또는 하드웨어에 한정되는 의미는 아니다. "부"는 어드레싱할 수 있는 저장 매체에 있도록 구성될 수도 있고 하나 또는 그 이상의 프로세서들을 재생시키도록 구성될 수도 있다. 따라서, 일 예로서 "부"는 소프트웨어 구성요소들, 객체지향 소프트웨어 구성요소들, 클래스 구성요소들 및 태스크 구성요소들과 같은 구성요소들과, 프로세스들, 함수들, 속성들, 프로시저들, 서브루틴들, 프로그램 코드의 세그먼트들, 드라이버들, 펌웨어, 마이크로 코드, 회로, 데이터, 데이터베이스, 데이터 구조들, 테이블들, 어레이들 및 변수들을 포함한다. 구성요소들과 "부"들 안에서 제공되는 기능은 더 작은 수의 구성요소들 및 "부"들로 결합되거나 추가적인 구성요소들과 "부"들로 더 분리될 수 있다.

[0028] 또한, 첨부 도면을 참조하여 설명함에 있어, 도면 부호에 관계없이 동일한 구성 요소는 동일한 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.

[0029] 본 발명은 라이다 센서를 이용하여 3차원 다중 객체를 검출하기 위한 단일 계층 3차원 다중 객체 검출 장치 및 방법에 관한 것이다.

[0030] 도 1은 본 발명의 일 실시예에 따른 3차원 다중 객체 검출 장치의 내부 구성을 보여주는 블록도이다.

- [0031] 도 1을 참조하면, 본 발명의 3차원 다중 객체 검출 장치(100)는 데이터 입력 모듈(110), BEV 이미지 생성 모듈(120), 학습 모듈(130), 로컬리제이션 모듈(140)을 포함한다.
- [0032] 데이터 입력 모듈(110)은 라이다 센서로부터 로(raw) 포인트 클라우드 데이터를 입력받는다.
- [0033] BEV 이미지 생성 모듈(120)은 로 포인트 클라우드 데이터에서 BEV(Bird's Eye View) 이미지를 생성한다.
- [0034] 학습 모듈(130)은 BEV 이미지로부터 세분화된 특징 이미지를 추출하기 위해 딥 러닝(deep learning) 알고리즘 기반 학습을 수행한다.
- [0035] 본 발명의 일 실시예에서 학습 모듈(130)은 CNN(Convolutional Neural Network) 기반 학습을 수행할 수 있다.
- [0036] 로컬리제이션 모듈(140)은 세분화된 특징 이미지에서 3차원 객체를 검출하기 위한 3D 후보 박스와 그에 대응하는 클래스를 찾기 위한 회귀(regression) 작업과 로컬리제이션(localization) 작업을 수행한다.
- [0037] BEV 이미지 생성 모듈(120)은 3D 로 포인트 클라우드 데이터를 2D 의사(pseudo) 이미지로 투영하고 이산화하는 방식으로 BEV 이미지를 생성할 수 있다.
- [0038] BEV 이미지 생성 모듈(120)은 3D 로 포인트 클라우드 데이터를 인코딩하여 높이, 밀도, 강도 및 거리의 4개의 특징 맵 이미지를 생성할 수 있다.
- [0039] 도 2는 본 발명의 일 실시예에 따른 3차원 다중 객체 검출 방법을 보여주는 흐름도이다.
- [0040] 도 2를 참조하면, 3차원 다중 객체 검출 방법은 라이다 센서로부터 로(raw) 포인트 클라우드 데이터를 입력받기 위한 데이터 입력 단계(S110), 로 포인트 클라우드 데이터에서 BEV(Bird's Eye View) 이미지를 생성하기 위한 BEV 이미지 생성 단계(S120), BEV 이미지로부터 세분화된 특징 이미지를 추출하기 위해 딥 러닝(deep learning) 알고리즘 기반 학습을 수행하는 학습 단계(S130) 및 세분화된 특징 이미지에서 3차원 객체를 검출하기 위한 3D 후보 박스와 그에 대응하는 클래스를 찾기 위한 회귀(regression) 작업과 로컬리제이션(localization) 작업을 수행하는 로컬리제이션 단계(S140)를 포함한다.
- [0041] BEV 이미지 생성 단계(S120)에서 3D 로 포인트 클라우드 데이터를 2D 의사(pseudo) 이미지로 투영하고 이산화하는 방식으로 BEV 이미지를 생성할 수 있다.
- [0042] BEV 이미지 생성 단계(S120)에서 3D 로 포인트 클라우드 데이터를 인코딩하여 높이, 밀도, 강도 및 거리의 4개의 특징 맵(feature map) 이미지를 생성할 수 있다.
- [0043] 학습 단계(S130)에서 CNN(Convolutional Neural Network) 기반 학습을 수행할 수 있다.
- [0045] 본 발명에서는 실시간 및 안전한 시스템을 위해 효율적이고 간결한 단일 계층 3D 다중 객체 검출 장치를 제안한다. 우선 라이다(LIDAR) 센서 데이터의 간결한 2D 표현이 활용되고, 학습 작업에 대한 세분화된 기능을 추출하기 위해 적합한 CNN(Convolutional Neural Network) 방식이 도입된다. 본 발명은 3D 바운딩 박스(bounding box) 위치뿐만 아니라 헤딩 각도(heading angle)도 추정한다.
- [0046] 이제 본 발명에서 간결한(compact) 입력 생성, 적합한 CNN 아키텍처 및 최종 3D 객체 후보를 로컬리제이션(localization)하기 위한 학습 및 추론 부분 전략을 설명하기로 한다.
- [0047] 도 3은 본 발명의 일 실시예에 따른 3차원 다중 객체 검출 장치의 전체 프레임워크를 도시한 것이다.
- [0048] 도 3을 참조하면, 본 발명의 단일 계층 3차원 다중 객체 장치의 전체 프레임워크(framework)는 라이다(LIDAR) 센서로부터 로(raw) 포인트 클라우드 데이터를 입력받고(a), 압축 방식으로 로 포인트 클라우드 데이터로부터 네 개의 특징 이미지를 포함하는 BEV(Bird's Eye View) 의사(pseudo) 이미지를 생성하고(b), 다중 헤드를 갖는 학습 태스크를 위한 미세 특징 이미지를 추출하는 CNN 기반 학습을 수행하고(c), 3D 후보 박스 및 그 대응 클래스를 찾기 위한 회귀(regression) 및 로컬리제이션(localization)으로 구성된다.
- [0049] BEV(Bird's Eye View) 생성에 대해 설명하면 다음과 같다.
- [0050] 도 4는 BEV(Bird's Eye View) 특징 맵 생성의 상세 구조를 도시한 것이다.
- [0051] 도 4를 참조하면, 본 발명의 3차원 다중 객체 검출 장치에서 높이 특징(Height feature), 강도 특징(Intensity feature), 밀도 특징(Density feature), 거리 특징(Distance feature)을 포함하는 4개의 컴팩트한 특징 맵을 추출한다.

- [0052] 일반적으로 로(raw) 3D 라이다 포인트 클라우드 데이터를 표준 방식의 복셀(voxel) 표현과 3D 그리드 셀(grid cells)로 인코딩하고, 여기서 특징을 추출하기 위해 3D CNN이 사용된다. 그러나 대부분의 3D 공간은 희박하거나 비어 있으므로 이러한 방법은 최적화된 접근 방식으로 간주되지 않아 시간과 하드웨어 모두 비효율적이다. 다른 방법으로, 원시 3D 라이다 포인트 클라우드 데이터는 FV(Front View) 표현으로 인코딩된다. 그러나 이러한 방법은 간결하지만 객체 중첩 문제를 제거할 수 없다.
- [0053] 라이다(LIDAR) 센서는 3D 포인트 위치(x, y, z)와 모든 포인트의 반사율 값 r을 제공하며, 초당 수천에서 수백만 포인트를 얻는다.
- [0054] 본 발명에서는 새로운 간결한 BEV 세대를 제안하며, 로 3D 포인트 클라우드 데이터를 2D 의사(pseudo) 이미지로 투영하고 이산화한다. 이는 시간 효율적인 사전 처리 방식으로 간주되고, 객체의 물리적 모양이 명시적으로 유지된다.
- [0055] 3D 환경의 전체 조사 공간(investigated space)에서 라이다(LIDAR) 스캐너로 획득한 $L \times W \times H$ 는 단일 높이, 밀도, 강도 및 거리 특징 맵(feature map)으로 인코딩된다.
- [0056] 높이 특징(feature)의 각 셀 값은 셀 내의 포인트 높이 중 최대 높이로 계산된다. 그런 다음 정규화 단계를 적용하여 정규화된 높이 특징 맵을 얻는다.
- [0057] 밀도 특징은 3D 현실 세계에서 다양한 포인트 클라우드 분포와 함께 셀 내의 포인트 밀도를 나타낸다. $\min(1, \frac{\log(Q+1)}{\log(64)})$ 공식을 사용하여 밀도 특징을 정규화한다. 여기서 Q는 셀 내 점의 양이다.
- [0058] 강도 특징(intensity feature)에서는 라이다(LiDAR) 강도가 기록되고 [0,1] 사이의 물체 표면 값을 반영하는 레이저빔의 반환 강도가 기록된다. 본 발명에서 강도 특징은 기준 셀에서 최대 높이를 갖는 점의 로(raw) 반사율 값이다.
- [0059] 대부분의 셀은 희소하거나 비어 있으며, 특히 원거리의 경우 훈련 데이터 세트를 조사한 결과, 거의 97%의 포인트 클라우드가 X 방향을 따라 첫 번째 [0,30m] 범위에 위치하는 것을 확인하였다. 훈련 데이터 세트의 거리에 따른 포인트 클라우드 분포는 도 5에 명확하게 도시되어 있다.
- [0060] 도 5는 훈련 데이터 세트 샘플에서 거리에 따른 관심 영역의 평균 포인트 클라우드 분포를 나타낸 그래프이다.
- [0061] 도 5를 참조하면, 물리적으로 포인트 클라우드 분포는 LIDAR 스캔 각도와 장면 시나리오 간의 차이를 기반으로 한다. 근거리에서는 빔 각도가 작아서 LIDAR 센서가 많은 포인트를 획득한 반면, 장거리에서는 빔 각도가 더 커서 적은 수의 포인트를 얻는다. 본 발명에서 BEV 표현을 강화하는 거리 정보를 보완하기 위하여 이 거리 특징 맵을 제안한다. 거리 특징은 학습 작업 외에도 모델이 범위별 포인트 클라우드 분포를 학습하는 데 도움이 된다. 각 셀의 정규화된 거리 특성 D_{i_norm} 은 다음 공식에 의해 계산된다.
- $$D_{i_norm} = \frac{D_{O \rightarrow P_i}}{D_{max}} = \frac{\sqrt{(x_{P_i}^2 + y_{P_i}^2 + z_{P_i}^2)}}{\sqrt{(x_{max}^2 + y_{max}^2 + z_{max}^2)}} \quad (1)$$
- [0062]
- [0063] 여기서 $D_{O \rightarrow P_i}$ 는 LIDAR 원점(0,0,1.73m)과 현재 점 P_i 사이의 거리이고, D_{max} 는 조사 영역(investigated area) Ψ 내에서 가장 먼 점 P_{max} 까지의 LIDAR 원점이고, $(x_{P_i}, y_{P_i}, z_{P_i})$ 및 $(x_{max}, y_{max}, z_{max})$ 는 각각 점 P_i 및 P_{max} 의 위치이다.
- [0064] 본 발명은 빠르고 효율적인 3D 다중 객체 검출 장치에 관한 것으로서, 본 발명에서 제안된 네트워크 아키텍처는 LIDAR 포인트 클라우드의 2D 표현을 강력하게 활용하고 학습하여, 밀집된 2D BEV 의사 이미지 중에서 객체를 감지하고 분류해야 한다. 훈련 데이터 세트의 3D 객체 및 해당 레이블의 인코딩과 관련하여 네트워크는 미리 정의된 객체 앵커에 추가로 의존하지 않고, 직접 추출 및 인코딩하거나, 첫 번째 계층에서 두 번째 계층으로 전달된 지역 제안을 조정한다. 전체 네트워크 아키텍처가 도 6에 도시되어 있다.
- [0065] 본 발명에서 제안된 네트워크 아키텍처는 두 개의 하위 네트워크로 나눌 수 있다.
- [0066] 첫째, 백본(backbone) 네트워크는 로(raw) BEV 표현에서 일반 정보를 컨볼루션 특징 맵 형태로 회수하는 데 사용되며, 컴팩트하고, 강력한 기능 표현을 학습하고 활용할 수 있는 높은 표현 능력을 가지고 있다.
- [0067] 둘째, 헤더(header) 네트워크는 백본 네트워크의 마지막 블록이며, 작업별 예측을 학습하도록 설계되었다. 이

네트워크에는 객체 중심 포인트(x,y), 오프셋 정보($\Delta x, \Delta y$), 확장 Z 좌표(z), 객체 크기(l,w,h), 객체 회전 각도(yaw)를 포함하는 5개의 하위 작업(sub-task)이 있다.

- [0068] 본 발명에서 도면을 참조하여 백본 네트워크와 헤더 네트워크에 대해 상세히 설명하면 다음과 같다.
- [0069] 도 6은 본 발명의 일 실시예에 따른 3차원 다중 객체 검출 장치의 상세한 CNN 아키텍처(architecture)를 도시한 것이다.
- [0070] 도 6을 참조하면, 본 발명의 3차원 다중 객체 검출 장치의 CNN 아키텍처의 전체 네트워크는 2개의 주요 부분으로 나뉜다.
- [0071] 첫 번째는 백본(Backbone) 네트워크로서, 다음과 같은 3개의 서브 모듈(a, b, c)로 구성된다.
- [0072] a) 수정된 Resnet 블록 모듈(Res_Block)로서, 커널(kernel), 다운 샘플링 비(down-sampling ratio) 및 반복량(quantity of repetition)을 나타낸다.
- [0073] b) 각 스케일의 업 샘플링(up-sampling) 블록을 나타내는 모듈(US_Block)이다.
- [0074] c) 다운 샘플링 모듈(DS_Block)이다.
- [0075] 두 번째는 헤드 모듈(Head)을 포함하는 헤더(Header) 네트워크이다. 여기서 헤드 모듈은 학습 태스크를 위한 객체 중심(Object Center), 오프셋(offset), Z 차원(Z dimension), 3D 객체 크기(3D Size) 및 회전 각(Rotation angle)의 5개의 특징을 활용하는 역할을 한다.
- [0076] 딥 러닝 기반 객체 감지 작업에서 CNN은 입력 정보를 컨볼루션 특징 맵 형태로 추출해야 한다. 학습 작업을 위해 고해상도에는 더 적은 레이어를 사용하고, 저해상도에는 더 많은 레이어를 사용하는 기준에 따라 작고 강력한 백본 아키텍처를 설계한다.
- [0077] 도 6의 실시예에서 구체적으로 제안된 네트워크의 블록은 총 10개이다.
- [0078] 첫 번째 블록은 채널 번호 64, 커널 7, 보폭 3, 다운샘플링 계수 2를 갖는 컨볼루션 레이어(convolution layer)이다.
- [0079] 두 번째 블록 내지 다섯 번째 블록은 스킵(skip) 연결 수가 각각 3, 8, 8, 3인 모든 블록에 대해 다운샘플링(down-sampling) 팩터(factor)가 2인 수정된 레지듀얼 레이어(residual layer)로 구성된다.
- [0080] 첫 번째 블록부터 다섯 번째 블록까지의 총 다운 샘플링 팩터는 32이다.
- [0081] 여섯 번째 블록 내지 여덟 번째 블록은 하향식(top-down) 업샘플링(up-sampling) 블록이고, 마지막 두 블록은 상향식(bottom-up) 다운샘플링(down-sampling) 블록이다. 그리고, 마지막 세 블록이 헤더 네트워크의 입력으로 공급되도록 선택된다.
- [0082] 헤더 네트워크는 분류 및 3D 객체 로컬리제이션을 모두 처리하는 다중 특정 작업을 학습하기 위해 작고 효율적으로 설계된다. 헤더 네트워크에는 클래스 $(\hat{x}, \hat{y})$ 에 해당하는 객체 중심 포인트, 오프셋 정보 $(\Delta \hat{x}, \Delta \hat{y})$, $\mathbb{R}^2$ 에서 $\mathbb{R}^3$ 까지의 확장 좌표 $\hat{z}$, 객체 크기 $(\hat{l}, \hat{w}, \hat{h})$, 객체의 회전각 팩터 $(\widehat{\cos \phi_i}, \widehat{\sin \phi_i})$ 를 포함하는 5개의 하위 작업이 있다. 여기서 본 발명의 추론 단계에서 $[-\pi, \pi]$ 범위 내에서 객체 회전 각을 $\tan^{-1}(\widehat{\sin \phi_i}, \widehat{\cos \phi_i})$ 로 용이하게 디코딩할 수 있다.
- [0083] 최종 예측 결과는 이전에 정의된 임계값보다 높은 값으로 설정된 모든 선택된 중심 포인트 $P(\hat{x}, \hat{y})$ 에 대해 $[C, (\hat{x}_i + \Delta \hat{x}_i, \hat{y}_i + \Delta \hat{y}_i, \hat{z}_i), \tan^{-1}(\widehat{\sin \phi_i}, \widehat{\cos \phi_i}), (\hat{l}_i, \hat{w}_i, \hat{h}_i)]$ 로 구성된다.
- [0084] 본 발명에서 학습 및 추론 과정은 컴팩트하고 효율적이며 안전한 임베디드 시스템 친화적으로 구현되며, 이에 대해 설명하기로 한다.
- [0085] 본 발명의 일 실시예에 따른 앵커 없는(anchor-free) 단일 계층 3D 다중 객체 검출 장치는 각 후보에 대해 총 5개의 헤드를 예측하는데, 키포인트 히트맵 헤드, 로컬 오프셋 헤드, 객체 방향 헤드, Z축 위치 헤드 및 3D 개체 치수 헤드가 그것이다. 이러한 헤드는 추론 단계에서 최종 후보를 생성하는 데 필요하다.
- [0086] 중심 회귀는 CNN 아키텍처를 통과한 후 중심 포인트를 출력하며 각 점은 하나의 객체 범주에 해당한다. 중앙 히

트 맵의 모양은 $\hat{H} \in [0,1]^{\frac{L}{S} \times \frac{W}{S} \times C}$ 로 정의된다. 여기서, S는 다운샘플링 비율이고, C는 예측된 클래스의 수를 나타낸다.

[0087] 키포인트 히트맵 H는 요소 R로 나누어지고 BEV에서 객체 중심이 어디에 있는지 찾는 데 사용된다. $\hat{H}_{x,y,c} = 1$ 은 검출된 중심 포인트고, $\hat{H}_{x,y,c} = 0$ 은 백그라운드이다.

[0088] 오프셋 회귀의 주요 역할은 예측된 객체 중심 포인트의 정확도를 강화하고 BEV 생성 프로세스에서 양자화 오류를 완화하는 것이다. 이를 위해 오프셋 회귀를 적용하여 모든 중심 포인트에 대한 오프셋 특징 맵 $\hat{O} \in R^{\frac{L}{S} \times \frac{W}{S} \times 2}$ 를 예측한다. 그리고, 오프셋의 학습 대상으로 L1 손실을 선택한다.

[0089] 안전을 위해 3차원 객체의 위치뿐만 아니라 진행 각도도 정확하게 예측해야 한다. Z축 주위의 헤딩 각도는 요(yaw) 각도로 간주되며, 기울어진 대상의 경우 요 각도 ϕ 를 $(\cos(\phi), \sin(\phi))$ 로 인코딩하고, 추론하는 동안 요 각도 ϕ 를 $\tan^{-1}(\sin(\phi), \cos(\phi))$ 로 디코딩한다.

[0090] 방향 회귀(Orientation Regression)는 특징 맵을 출력한다. 모든 싱글 중심 포인트(single center point)에서 $\hat{Y} \in R^{\frac{L}{S} \times \frac{W}{S} \times 2}$ 일 때, L1 손실 함수가 트레이닝을 위해 다음 수식과 같이 적용된다.

$$[0091] L_{yaw} = \frac{1}{N} \sum_p \sum_{i \in \{\sin(\phi), \cos(\phi)\}} |\sigma(\hat{Y}_{p,i}) - Y_{p,i}| \quad (2)$$

[0093] 객체 중심 포인트는 R^2 에서 예측되며, Z축을 따른 확장이 R^3 에서 중심 포인트를 로컬리제이션하는데 필요하다. Z축 위치 회귀는 예측된 각 중심 포인트에 대해 Z축 특징 맵 $\hat{Z} \in R^{\frac{L}{S} \times \frac{W}{S} \times 1}$ 을 예측한다. Z축 회귀 결과는 3D 바운딩 박스 위치 정확도에 막대한 영향을 미치는데, 다양한 개체 속성 샘플이 있는 Z축의 무한 회귀 대상 때문이다. 따라서 예측은 이상치(outliers), 특히 불균형 훈련 세트에 쉽게 민감해진다는 문제가 있다. 이 문제를 극복하기 위해 균형 1 손실(loss)을 도입하여 불균형 훈련 세트를 최소화하고 모델의 안정성을 향상시킨다. Z축 회귀 학습을 위해 균형 L1 손실이 사용된다.

$$[0094] L_z = \frac{1}{N} \sum_p L_b(|\hat{Z}_p - Z_p|) \quad (3)$$

[0095] 여기서, L_b 는 L1 손실 정의에 따른 균형 L1 손실이다.

$$[0096] L_b(\mu) = \begin{cases} \frac{a}{b}(b|\mu| + 1) \ln(b|\mu| + 1) - a|\mu| & \text{if } |\mu| < 1 \\ \gamma|\mu| + C & \text{otherwise} \end{cases} \quad (4)$$

[0097] 여기서, a, b, γ 는 L1 균형 손실 하이퍼 파라미터(hyper-parameters)이고, $a \ln(b+1) = \gamma$ 의 관계에 있다.

[0098] 크기 회귀(Size Regression) 프로세스는 3D 개체 중심 좌표(x,y,z)를 따라 길이 l, 너비 w 및 높이 인 3D 객체 공간 차원을 생성한다. 크기 회귀 프로세스 작업에서 예측해야 하는 세 가지 값이 있으므로 각 중심 포인트에 대해 크기 회귀 특징 맵 $\hat{S} \in R^{\frac{L}{S} \times \frac{W}{S} \times 3}$ 을 반환한다. 크기 회귀는 Z축 회귀와 동일한 특성을 가지며 또한 제한 없는 회귀 목표로 인해 이상치에 민감하다. 따라서 크기 회귀의 학습 대상으로 균형 L1 손실을 선택한다.

$$[0099] L_{size} = \frac{1}{N} \sum_p \sum_{i \in \{l,w,h\}} L_b|\hat{S}_{p,i} - S_{p,i}| \quad (5)$$

[0100] 본 발명의 단일 계층 3차원 다중 객체 검출 장치의 총 손실 함수는 상기 헤드 회귀 손실의 가중치 합이다.

$$[0101] L_{total} = \chi_{hm} L_{hm} + \chi_{off} L_{off} + \chi_{yaw} L_{yaw} + \chi_z L_z + \chi_{size} L_{size} \quad (6)$$

[0102] 여기서, χ_{hm} , χ_{off} , χ_{yaw} , χ_z 및 χ_{size} 는 각각 히트맵 중심 회귀, 오프셋 회귀, 방향 회귀, Z축 위치 회귀 및 크기 회귀에 대한 균형 계수를 나타낸다.

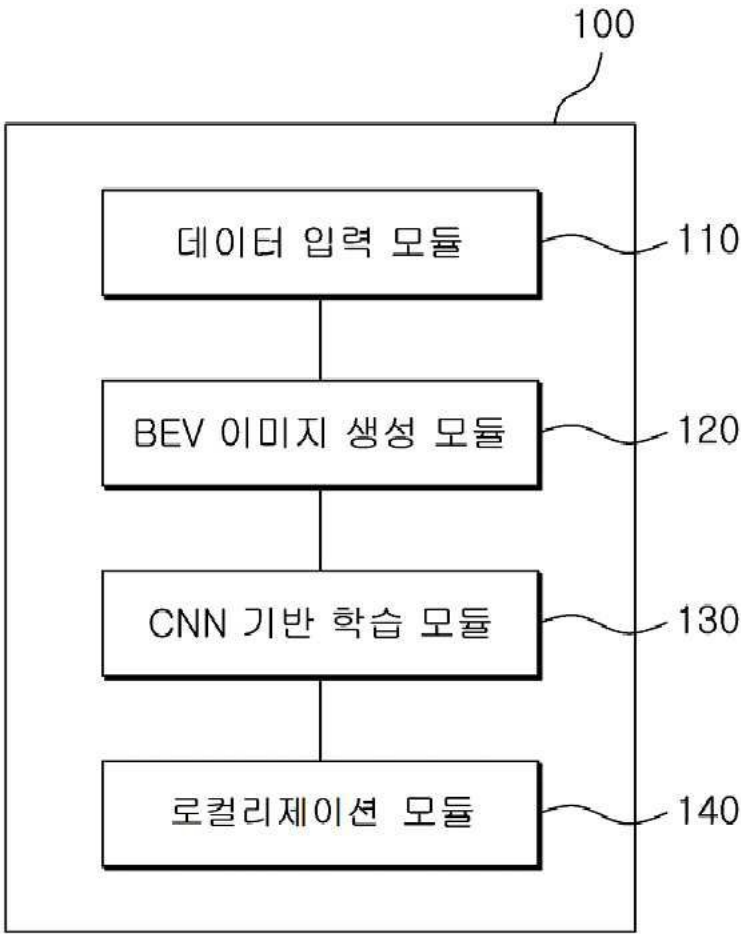
- [0104] 본 발명의 단일 계층 3차원 다중 객체 검출 장치에서 추론 과정에 대해 설명하면 다음과 같다.
- [0105] 3D 바운딩 박스의 정확한 위치 파악을 위해 세분화된 특징 맵을 추출한 후, 연결된 8개의 이웃보다 값이 큰지 비교하여 중심 키포인트의 존재를 확인한다. 여기서, 주변 8개의 이웃과 비교하는 것은 키포인트를 찾는 데 가장 빠르고 정확한 방법이다.
- [0106] 그런 다음 두 가지 기준으로 중심 포인트만 유지한다. 중심 포인트 값은 미리 정의된 임계값보다 높고, 신뢰도 점수는 탐지 범위에서 미리 정의된 개체 번호의 우선 순위에 따라 탐지된 중심 포인트 번호를 필터링한다.
- [0107] $\mathbb{R}^3$ 환경의 객체는 $(c_x, c_y, c_z, r, p, y, l, w, h)$ 로 나타낼 수 있다. 여기서 (c_x, c_y, c_z) 는 3D 객체 중심, (r, p, y) 은 각각 롤, 피치 및 요 회전 각도, (l, w, h) 는 각각 객체의 길이, 너비, 높이를 나타낸다.
- [0108] 객체가 평평한 도로 평면에 있는 것으로 가정하면, $r = p = 0$ 이므로, $\mathbb{R}^3$ 의 객체는 7 자유도($c_x, c_y, c_z, y, l, w, h$)를 갖는다. 추론하는 동안 $\tilde{P}_C = \sum_{i=1}^n \{\tilde{x}_i, \tilde{y}_i\}$ 은 예측 집합이고, 여기서 n 은 클래스 C에서 감지된 중심 포인트의 양이다.
- [0109] 예측 후, 히트맵 중심 포인트, 오프셋, 방향 각도, Z축 위치 및 크기 차원에 대응되는 $(\tilde{x}_i, \tilde{y}_i), (\Delta \tilde{x}_i, \Delta \tilde{y}_i), (\cos \tilde{\phi}_i, \sin \tilde{\phi}_i), \tilde{z}_i, (\tilde{l}_i, \tilde{w}_i, \tilde{h}_i)$ 를 얻는다.
- [0110] 그 다음 모든 후보 타겟은 $(C, \tilde{x}_i + \Delta \tilde{x}_i, \tilde{y}_i + \Delta \tilde{y}_i, \tilde{z}_i, \tan^{-1}(\sin \tilde{\phi}_i, \cos \tilde{\phi}_i), \tilde{l}_i, \tilde{w}_i, \tilde{h}_i)$ 과 같은 형식으로 모두 융합되어, 클래스 C에 대한 정확한 3D 바운딩 박스를 생성한다.
- [0111] 본 발명에서 이러한 작업을 임베디드 시스템 친화적 접근 방식으로 처리한다. 따라서, 본 발명에서는 최대 폴링 연산을 이용하여 객체 중심을 찾으므로, 기존 NMS보다 훨씬 빨리 객체 중심을 찾을 수 있다.
- [0112] 이상 본 발명을 몇 가지 바람직한 실시 예를 사용하여 설명하였으나, 이들 실시 예는 예시적인 것이며 한정적인 것이 아니다. 본 발명이 속하는 기술분야에서 통상의 지식을 지닌 자라면 본 발명의 사상과 첨부된 특허청구범위에 제시된 권리범위에서 벗어나지 않으면서 다양한 변화와 수정을 가할 수 있음을 이해할 것이다.

부호의 설명

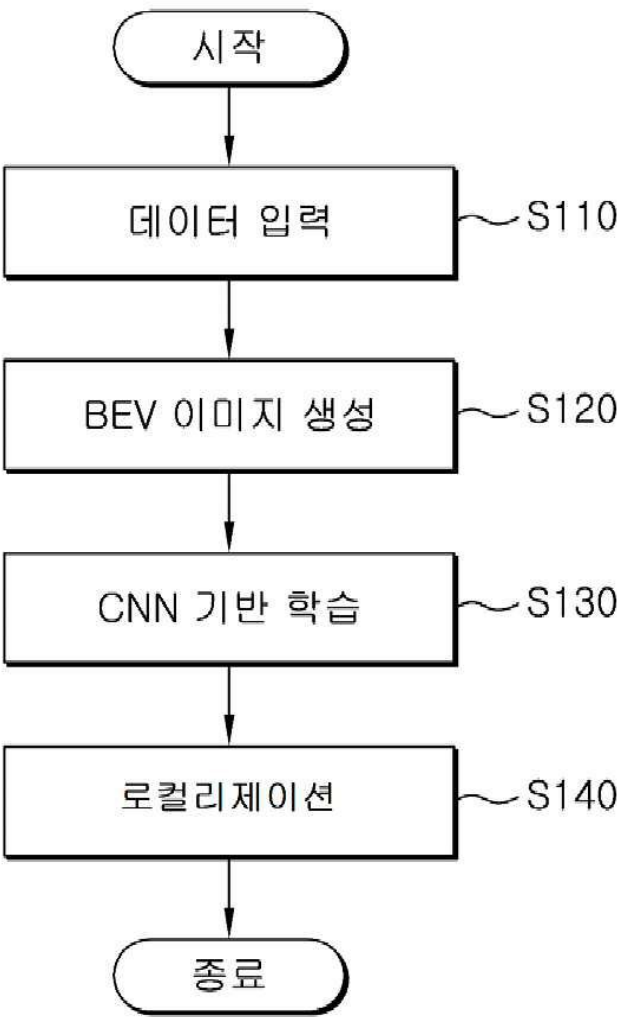
- [0114] 100 단일 계층 3차원 다중 객체 검출 장치
- 110 데이터 입력 모듈
- 120 BEV 이미지 생성 모듈
- 130 CNN 기반 학습 모듈
- 140 로컬리제이션 모듈

도면

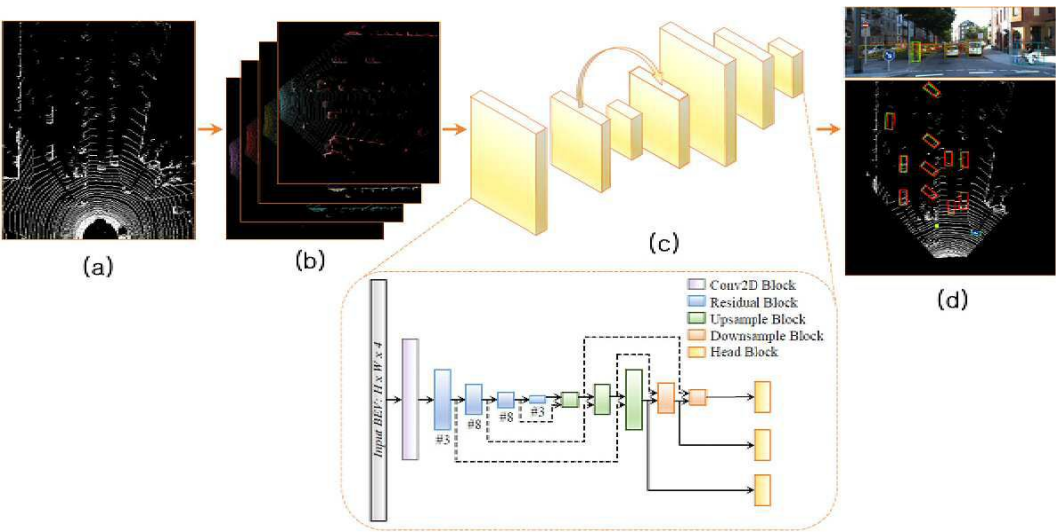
도면1



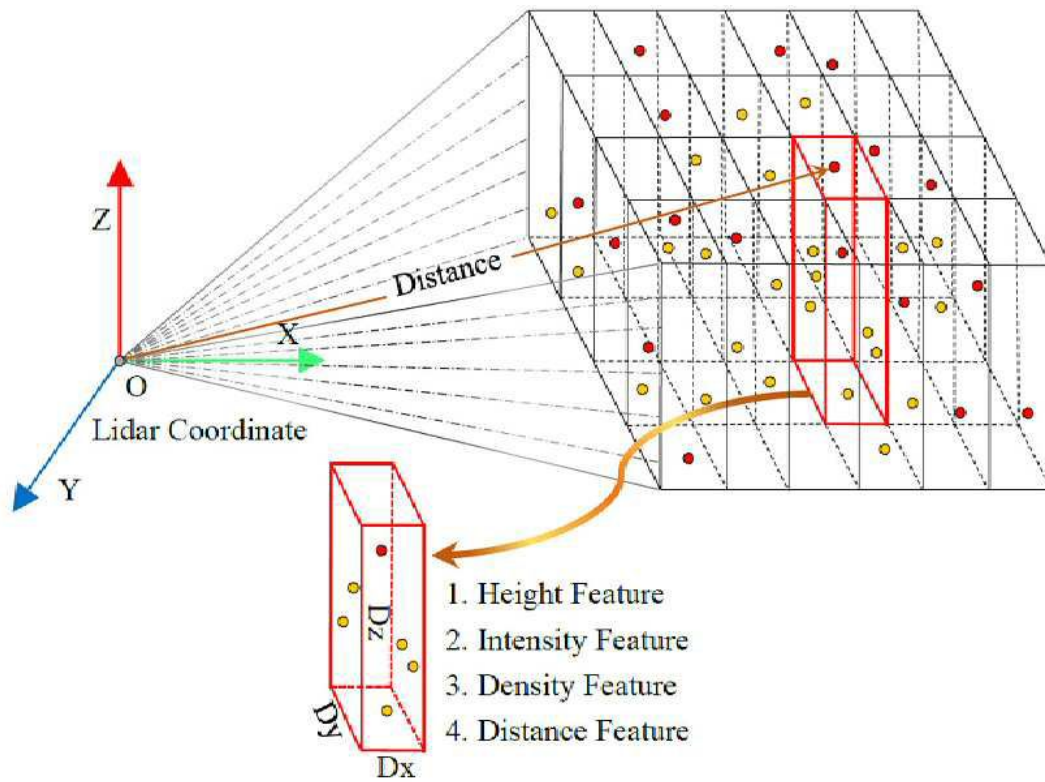
도면2



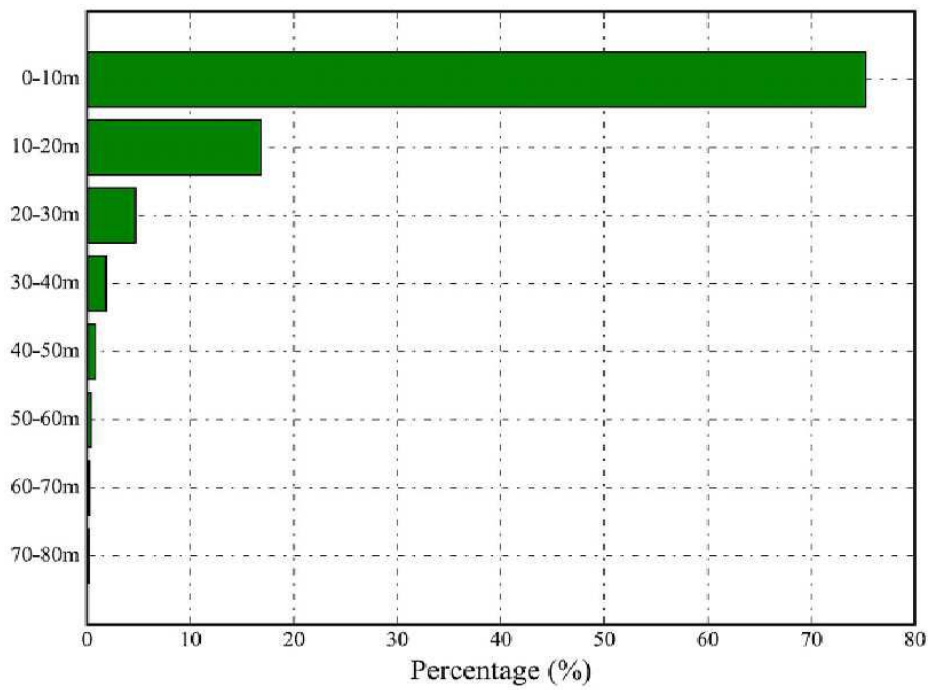
도면3



도면4



도면5



도면6

