

자율주행차를 위한 실시간 거리예측 기술 연구

한지은, 고석현, 이성진

동서울대학교 전자공학과

haneul2069@du.ac.kr, sungjinlee@du.ac.kr

A Study on the Real-Time Depth Estimation for Autonomous Vehicle

Jieun Han, Sukhyun Ko, Sungjin Lee

Department of Electronic Engineering, Dong Seoul University

요약

자율주행을 위해서는 자동차가 놓인 상황에 대한 3차원 인지 기술이 필수적이다. 이런 3차원 인지를 위해 주로 사용되는 Lidar 기반 3D Object Detection 기술 혹은 3D Reconstruction 기술은 높은 연산 복잡도로 인해 실시간 동작이 힘들다. 본 논문에서는 실시간 성능이 요구되는 자율주행 환경에서 2D Object Detection 기술에 후처리를 더해 3D Object Detection에 준하는 결과를 낼 수 있는 2D-DEPTH (2D object detection based DEPTH estimation) 알고리즘을 제안하였다. 실험 결과 제안하는 2D-DEPTH 알고리즘은 1% 미만의 정확도 손상으로 500% 이상의 가속화를 이룰 수 있었다.

I. 서론

최근 딥러닝 기술의 발전으로 자율주행 기술은 운전 보조 기능을 넘어서 완전 자율주행 단계에 이르고 있다. 이런 자율주행의 핵심적인 기술 중에 하나는 3차원 주변 상황 인지 기술이다. 즉, 3D Object Detection과 3D Reconstruction과 같은 기술로서 이를 정확히 추론해 내기 위해 Google, Intel 등의 진영에서는 Lidar를 사용하고 Tesla 등의 진영에서는 Camera를 기반으로 해당 연산을 해낸다 [1].

하지만 이런 Lidar 기반 연산은 수많은 Point Cloud 데이터를 획득하여 3D Object Detection을 수행해야 하고 Camera 역시 사용해서 Sensor Fusion 단계를 거치기 때문에 실시간 처리에 맞는지 Tesla 진영에서는 의문을 제기한다. 반면 Camera 기반 연산은 여러개의 Camera를 통해 부족한 2D 데이터들로 방대한 3D 데이터를 복원하기 위해 수많은 딥러닝 연산을 수행하므로 이 역시 실시간 처리에 의문이 제기되곤 한다.

본 논문에서는 이런 Lidar와 Multiple Camera에 의존하는 것이 아닌 Mono Camera 만으로 3D Object Detection 전체를 수행하는 것이 아닌 필수 3D 정보 만을 추론해 내는 2D-DEPTH 방식을 제안한다. 실험에 따르면 2D-DEPTH 알고리즘을 사용하여 기존 연산량을 20% 수준으로 감소시킬 수 있고 간단한 도로 상황에서는 정확도 손상이 기존 기술들에 비해 1% 미만일 수 있다.

II. 2D-DEPTH 모델

본 연구에서는 2D Object Detection을 수행하고 그 결과에 기반하여 자율주행에 직접적으로 필요한 해당 객체의 Depth를 예측한다. 이를 위해 사용되는 2D Object Detection 모델로 RetinaNet [2] 모델을 사용하였다.

1. RetinaNet

RetinaNet 기술은 2D Object Detection 모델 중 1-Stage Detector로서 Focal Loss를 사용하여 Class Imbalance 문제, 즉 Easy Example (IoU가 0에 가까운 샘플, 훈련용 데이터의 대부분을 차지하지만 훈련에 악영향

을 미치는 영역)에 낮은 가중치를 두고 Hard Example (IoU가 일정 값보다 큰 샘플, 훈련용 데이터의 작은 부분을 차지하지만 훈련에 유의미한 Gradient를 가지는 부분)에 높은 가중치를 두어 성능을 개선한 모델이다. 본 논문에서는 이 RetinaNet을 사용하여 관심 클래스로 KITTI dataset [3] 기준 Car, Pedestrian, Bicycle 만을 선정하였다.

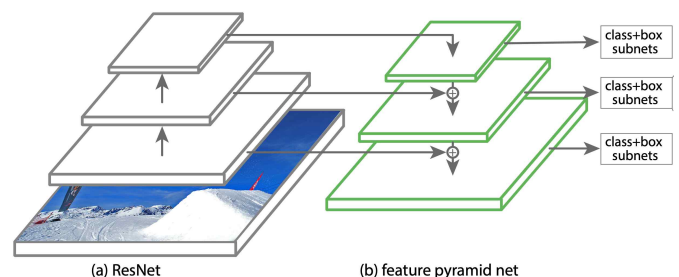


그림 1. RetinaNet 구조 및 동작

2. 2D Object Detection based DEPTH estimation

광학적으로 멀리 있는 물체는 작게 보이고 가까이 있는 물체는 크게 보인다. 이런 단순한 관찰은 Occlusion이 심한 환경에서는 다소 다를 수 있으나 치한 환경이 단순한 환경에서는 꽤 정확할 수 있다. 그래서 본 기술은 치한 환경이 교외지역에 집중해서 Lidar나 3D Reconstruction 기술에 의존하지 않고 2D Object Detection의 결과인 Bounding Box Size 만으로 Depth를 예측하여 3D Object Detection을 가속화 하려 한다. 또한 추론하고자 하는 Depth 역시 정확히 몇 m가 아닌 실 주행을 의미할 수 있도록 다음과 같이 4 단계 깊이로 구분하여 인식한다. 즉,

- **Depth-1.** 1.5m 이내에 있는 아주 가까운 초근접 거리에 위치함. 해당 위치에 관심 객체들이 검출되면 즉시 브레이크를 밟아 충돌을 방지하도록 함
- **Depth-2.** 1.5m 이상 4m 이내 거리에 관심 객체가 위치한 경우. 해당 위치에 관심 객체들이 검출되면 서서히 브레이크를 밟아 속도를 줄이거나 유

지하는 거리

- **Depth-3.** 4m 이상 10m 이내 거리에 관심 객체가 위치한 경우. 해당 위치에 관심 객체들이 검출되면 해당 객체를 주시하며 속도를 유지하는 거리
- **Depth-4.** 10m 이상 거리에 관심 객체가 위치한 경우. 해당 물체 때문에 속도를 줄이거나 유지하지 않으며 자율주행의 원 경로만을 고려하여 주행

이런 4단계 깊이 구분으로 실제 주행에 도움이 되지 않고 불필요한 영역(ex. 인도 영역, 하늘 영역)의 depth 정보를 획득하려 하지 않고, 주행에 도움이 되는 영역(ex. 내 차가 주행하는 차선 영역, 내 차와 근접한 차선 영역)에 집중하여 4단계 depth 정보를 획득하는 방식이다. 또한 이런 4단계 깊이 구분을 위한 수치들은 자동차의 속도, 날씨(비, 눈 등), 브레이크 상태 등에 따라 달라질 수 있다.

그림 2는 해당 2D-DEPTH 알고리즘의 전체 동작과정을 나타내었다.

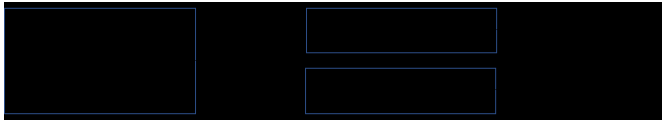


그림 2. 2D-DEPTH 알고리즘의 전체 동작과정

III. 실험데이터 셋 구축

기존의 데이터 셋은 교통상황, 도로 상황에 대한 구분이 없어서 새로운 데이터 셋을 구축하였다. 즉, 제안하는 2D-DEPTH 알고리즘은 상대적으로 occlusion 이 적은 상황 즉, 도로교통 상황이 한적한 상황에 더 적합할 것으로 예상되고 도로교통 상황이 복잡한 상황에서는 상대적으로 부정확한 결과가 나올 것으로 예상되기 때문에 다양한 상황에 따른 정확도를 검증하기 위해 다음 그림 3과 같은 코스에 따라 그림 4의 기기들을 이용하여 실험 데이터 셋을 수집하였다.



(a) Urban Area (b) Downtown Area (c) Highway Area

그림 3. 도로 유형에 따른 3가지 주행 시나리오



그림 4. 데이터 셋 구축을 위해 사용된 기기들

그림 3에서 보듯이 다음의 3가지 상황에 따라 코스를 결정하였다.

- **UrbanArea:** 교외지역으로 교통이 한적하고 간단한 총 2차선 도로
- **DowntownArea:** 도심지역으로 교통이 복잡한 총 8차선 도로
- **HighwayArea:** 주행 속도가 빠른 고속도로

IV. 실험

본 논문에서는 2D-DEPTH 알고리즘의 성능 검증을 위해 다음의 Evaluation Set을 정하였다.

- **Lidar3D :** 라이다 기반의 3D Object Detection 기술
- **Cam3D :** 스테레오 카메라 기반의 3D Object Detection 기술
- **2D-DEPTH :** 제안하는 2D Object Detection 기반 Depth 예측 기술

이런 실험 셋에서 다음과 같은 정확도 및 동작 속도 성능을 얻어내었다.

Evaluation Set	주행 시나리오					
	Urban		Downtown		Highway	
	Acc	Lat	Acc	Lat	Acc	Lat
Lidar3D	100	310	100	350	100	330
Cam3D	94	220	90	240	91	230
2D-DEPTH	94	10	88	10	90	10

Table 1. 세 가지 Evaluation Set에 대해 Accuracy 및 Latency 비교

우선 Urban 지역의 성능을 보면 제안하는 2D-DEPTH 알고리즘은 Lidar3D와 Cam3D에 비해 정확도 성능은 거의 비슷하지만(1% 미만) Latency 성능은 20% 수준으로 감소시킬 수 있음을 알 수 있다. 이는 Lidar3D와 Cam3D가 불필요한 영역의 3차원 복원에 지나치게 많은 자원을 할당하고 있음을 반증한다고 볼 수 있겠다. 하지만 Downtown, Highway 영역에서는 Latency 성능은 물론 20% 수준으로 감소시킬 수 있지만 정확도 성능이 다소 커서(5~10%) 이를 위한 보정 혹은 추가 후처리 기술이 필요하다고 볼 수 있겠다.

IV. 결론

자율주행을 위한 필수 기술인 3D Object Detection의 연산 복잡도를 줄여 실시간 동작이 가능하도록 2D Object Detection 기반의 Depth estimation 알고리즘을 제안하였다. 해당 제안 알고리즘인 2D-DEPTH 알고리즘은 교외 지역 Urban Area에서 기존 Lidar 기반 3D Object Detection이나 Camera 기반 3D Object Detection에 비해 정확도의 손상은 약 1% 미만으로 줄일 수 있었지만 동작 속도는 약 500% 가속화된 성능을 얻을 수 있었다.

ACKNOWLEDGMENT

이 논문은 2021년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기본연구사업(No. NRF-2019R1F1A1062878)

참고 문헌

- [1] Rui Qian, Xin Lai, Xirong Li, "3D Object Detection for Autonomous Driving: A Survey"
- [2] Tsung-Yi Lin, Priya Goyal, Ross Girshick, Kaiming He and Piotr Dollár, "Focal Loss for Dense Object Detection"
- [3] <http://www.cvlibs.net/datasets/kitti/>