

RGB-D 카메라 기반의 3D 객체 인식 모델 및 반자동 라벨링 기술 개발

주수빈*, 정덕기

*한국기계연구원

*sbjoo@kimm.re.kr, deokgi@kimm.re.kr

Development of a 3D Object Recognition Model and Semi-Automatic Labeling Technology Based on RGB-D Cameras

Subin Joo*, Deokgi Jeung

*Korea Institute of Machinery and Materials

요 약

본 논문은 3D 객체 인식 알고리즘 개발을 위해 RGB 색상 경계 데이터와 물체간 깊이 정보를 융합하여 물체를 분리할 수 있는 알고리즘을 제안한다. 또한 개발된 알고리즘을 빠르고 쉽게 학습하기 위한 RANSAC 기반 3D 객체 분리 기술을 활용한 반자동 라벨링 방법을 개발하였다.

I. 서 론

영상, 이미지 기반의 2D 물체 인식 알고리즘은 물체 인식 IoT 기술, 스마트폰 어플리케이션, 자율주행 자동차 등 다양한 분야에서 폭 넓게 사용되며 그 실용성과 정확도가 검증되었다. Yolo, SSD, Fast R-CNN 등 강력한 물체인식 알고리즘 기본 모델을 기반으로 하여, 물체 인식 가능한 종류 및 정밀도를 늘리고, 인식 속도를 개선하기 위한 딥러닝 알고리즘이 지금도 다양하게 연구되고 있다. 또한 저가형 LiDAR 및 깊이 카메라의 보급과 대용량 비디오 메모리를 보유한 그래픽 카드의 등장으로 인해 3D 물체 인식 관련 연구도 늘어나고 있다 [1,2].

하지만, 지도학습 기반에 의존이 강한 기존 기술들은, 수많은 클래스에 해당하는 라벨링 데이터를 획득하기 위해 많은 비용이 소모된다는 문제와 라벨링이 되지 않은, 처음 보는 물체는 객체로써 인식을 할 수 없어

오작동 할 수 있다는 문제를 여전히 갖고 있다 [3]. 이러한 문제를 극복하기 위해 self-learning 기법, zero-shot learning, 준 지도 학습 기반의 학습 방식이 제안되고 있으나, 여전히 학습되지 않은 물체에 대해서는 인식률이 떨어진다는 문제와, 학습이 된 물체라도 특정 각도나 조명 상황에서는 잘못 인식될 수 있다는 문제를 안고 있다.

본 연구에서는, 개별 객체에 대해 학습이 되어야만 물체 인식이 가능한 기존 물체인식 알고리즘 기법의 한계를 극복하기 위해, 색 경계 차이와 물체의 깊이 차이를 이용해, 학습되지 않은 물체라도 주변 환경과 분리된 객체로써 인식할 수 있는, 보다 원초적인 물체인식 알고리즘을 개발하고자 한다. 구체적으로는, RGB 데이터의 물체간 색 경계 정보와, 깊이 데이터의 물체간 거리 차이 정보를 융합하여 3D 객체를 인식할 수 있는 알고리즘을 개발하였다.

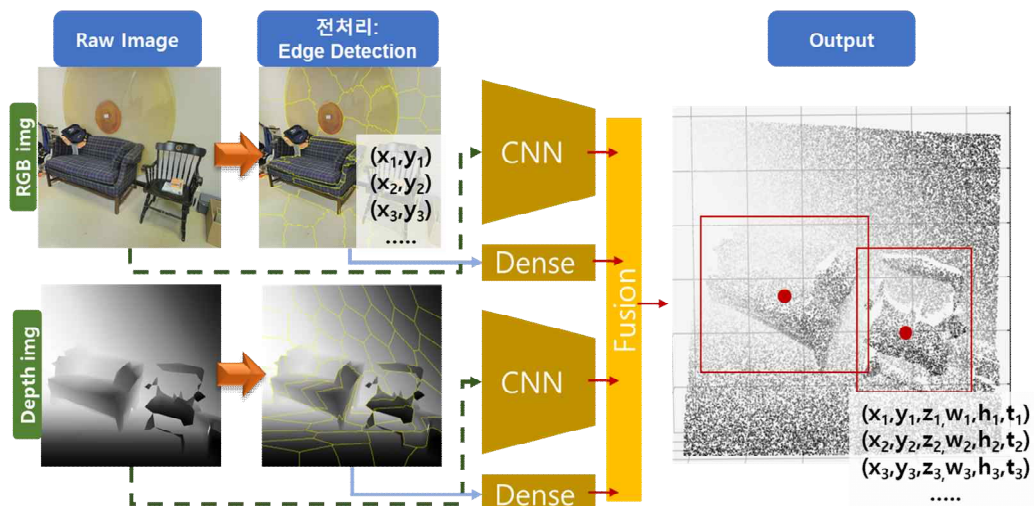


그림 1. RGB-D 기반 3D 객체 인식 모델 구성도

II. RGB-D 기반 3D 객체 인식 모델

본 연구에서는 3 채널 형태의 RGB 이미지와 1 채널 형태의 깊이 이미지, 그리고 두 이미지에 색 경계 분리 알고리즘을 적용하여 획득된 색 타일 좌표 (x_n, y_n)를 입력으로 받은 뒤, 이미지 데이터 내의 3D 객체의 위치를 x, y, z 좌표로 출력해주는 딥러닝 모델 및 체계를 그림 1 과 같이 구성하였다.

먼저 RGB-D 카메라로부터 RGB 이미지 및 깊이 데이터를 획득하였다. RGB 이미지로부터 1 차적 객체 분리를 하기 위해, 전통적인 색 경계 추출 방법인 Superpixel segmentation 알고리즘을 전 처리 방법으로 적용하였다. 전통적인 색 경계 추출 기법은 딥러닝을 이용한 Semantic segmentation 보다 픽셀 정밀도는 낮지만, 학습이 없이도 빠르게 객체를 분리해낼 수 있다는 장점이 있다.

또한, 기존 연구에서는 깊이 데이터에서 객체를 분리해내기 위해 3D point cloud 기반 Segmentation 접근 방법을 사용하는데, 본 연구에서는 정밀도가 다소 낮더라도 빠르면서 대략적인 객체 분리를 위해, RGB 이미지와 동일하게 Superpixel segmentation 기법을 적용하여 전처리를 하였다. 본 연구에서 제안된 3D 객체 인식 딥러닝 모델에는, RGB, Depth 원본 데이터와 함께, 전 처리 과정에서 획득된 객체의 중심 좌표가 입력된다. 모델의 출력 타깃 값은 3D 객체 물체의 중심 좌표로 설정하였다.

III. 반자동 라벨링 기술

딥러닝 기반 물체 인식 알고리즘 개발에 있어 어려움 중 하나는 라벨링 데이터의 획득이다. 본 연구에서는 이러한 문제를 해결하기 위하여 깊이 카메라로부터 얻어진 Point cloud 데이터에 평면 탐지 RANSAC 알고리즘을 적용하여, 라벨링용 3D 객체를 반자동으로 분리하는 방법을 제안한다. RANSAC 기법은 점 군집 데이터로부터 벽, 바닥, 테이블, 책상과 같은 평면을 탐색하는데 유용하기 때문에, 점 군집 데이터에서 벽, 바닥을 쉽게 탐색 후 분리해낼 수 있다.

벽과 바닥을 분리 제거한 원본 데이터에, 남아있는 점 데이터간 서로 일정 기준 값 (역치) 이하의 거리를 갖는 점끼리 클러스터링을 진행한다. 클러스터링 된 독립적인 군집은 수동으로 서로간 구분할 수 있는 번호를 매기는 방식으로 반자동 라벨링 하였다. 제안된 방법으로 정밀하거나 물체의 종류가 식별되는 라벨을 직접 획득할 수는 없으나, 수작업으로 Point cloud 에서 객체를 구분해내는 과정을 단축함으로써 라벨링 과정을 반자동화할 수 있다.

IV. 결론

개발된 3D 객체 알고리즘을 이용하여, RGB 이미지와 깊이 데이터로부터 객체를 인식한 결과는 그림 3 과 같다. 반자동 라벨링 기술을 통해 딥러닝 모델의 학습의 편의성이 증가하였으며, 본 연구의 가설대로 RGB, 깊이 데이터로부터 3D 객체를 분리해낼 수 있음을 확인하였다.

본 연구에서는 개발된 3D 객체 인식 딥러닝 모델에 4 가지 서로 다른 데이터 (RGB 이미지, RGB 이미지로부터 획득된 색 경계 분리 결과의 중심 좌표, 깊이 데이터, 깊이 데이터로부터 획득된 색 경계 분리 결과의 중심 좌표)가 병렬적으로 입력되도록 설계되었는데, 향후 연구에서는 4 가지 입력 데이터의 민감도 분석을 통해 불필요한 데이터는 제거하고, 결과 도출에 중요한 데이터는 보다 보강하는 방식으로 진행할 예정이다.

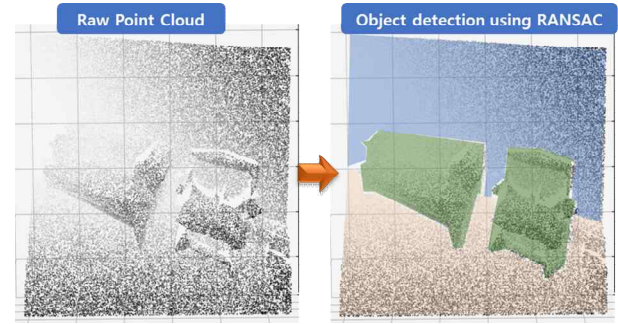


그림 2. 반자동 라벨링을 위한 RANSAC 기반 객체 분리

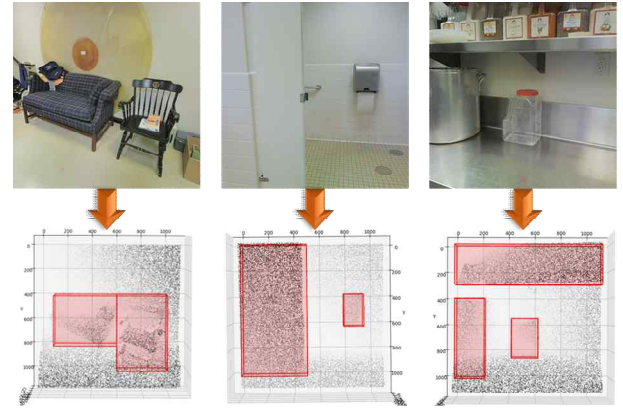


그림 3. 3D 객체 인식 알고리즘의 객체 인식 결과

또한 실시간 3D 객체 인식을 위해 모델의 경량화와 전 처리 과정속도 향상에 대한 연구도 진행할 계획이다. 이렇게 개발된 RGB-D 데이터 기반의 3D 객체 분리 기술은 향후 모든 라벨링 없이도 색상 경계와 물체간 깊이 차이만으로 물체를 인식해내는 알고리즘 개발로 발전되어, 물체 인식기술이 필요한 분야에 도움이 될 것으로 기대된다.

ACKNOWLEDGMENT

This study was supported by the research program of the Korea Institute of Machinery and Materials (NK031K)

참 고 문 헌

- [1] Shehzadi, T., Hashmi, K. A., Stricker, D., & Afzal, M. Z. (2023). 2d object detection with transformers: a review. *arXiv preprint arXiv:2306.04670*.
- [2] Aher, V. A., Jondhale, S. R., Agarkar, B. S., George, S., & Shaikh, S. A. (2024, January). Advances in Deep Learning-Based Object Detection and Tracking for Autonomous Driving: A Review and Future Directions. In *International Conference on Multi-Strategy Learning Environment* (pp. 569-581). Singapore: Springer Nature Singapore.
- [3] Chen, Z., Li, Z., Wang, S., Fu, D., & Zhao, F. (2023). Learning from noisy data for semi-supervised 3d object detection. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 6929-6939).