

사용자 추적을 위한 2D LiDAR 데이터 어노테이션 방법

오지용, 정유수, 이준구

한국전자통신연구원 대경권연구센터

{jiyongoh, yoosoojeong, leejg01679}@etri.re.kr

A 2D LiDAR Data Annotation Method for Person-Tracking

Jiyong Oh, Yoosoo Jeong, and Joon-Goo Lee

Daegu-Gyeongbuk Research Division

Electronics and Telecommunications Research Institute (ETRI)

요약

2D LiDAR 기반 사용자 추종 로봇에서 요구되는 인공지능 기반 사용자 추적 기술을 개발하기 위해서는 데이터셋이 필요하지만 최근 공개된 2D LiDAR 데이터셋은 사용자 추적에 적합하지 않아 데이터셋을 직접 구축해야 한다. 2D LiDAR 데이터를 사용자 추적 기술 개발에 활용하기 위해서는 수집된 데이터에 포함된 각각의 점들이 추적 대상인지 아닌지 판단해야 한다. 하지만 그러한 판단이 쉽지 않아 데이터 가공에 많은 시간과 노력이 요구된다. 본 논문에서는 2D LiDAR 기반 사용자 추적 데이터셋을 효과적으로 가공하기 위한 어노테이션 방법을 제안한다. 제안하는 방법은 추가로 사용되는 카메라와 2D LiDAR 사이의 캘리브레이션 정보를 바탕으로 2D LiDAR 데이터의 점들을 영상에 투영한 다음, 영상에 의미론적 영상 분할 기술을 적용하여 영상에 투영된 2D LiDAR 데이터의 점들이 추적 대상인 사용자 영역에 포함되는지를 판단한다. 실제 수집한 데이터를 활용하여 제안하는 방법의 효과를 보여준다.

I. 서론

로봇은 전통적으로 사람들의 노동력을 대체하기 위해 개발되었으며, 중량물은 운송하기 위해 많은 노동력이 요구되는 만큼 중량물을 운송하기 위한 기술도 여전히 개발되고 있다. 중량물을 정해진 목적지까지의 단순한 운송을 위해 이동로봇의 자율주행 기술이 적합할 수 있지만, 적재물을 운반하는 과정에서 사용자의 작업이 빈번하게 요구되는 농작물 수확, 택배 배송, 등과 같은 작업에서는 자율주행 기술보다 로봇이 특정 사용자를 따라다니는 사용자 추종 기술[1]이 더 적합하다. 사용자 추종 기술은 로봇이 사용자를 따라다니게 만드는 주행 제어 기술과 함께 로봇을 기준으로 사용자의 상대 위치를 파악하는 사용자의 위치 추정 기술이 필요하다. 사용자 위치 추정을 위해 다양한 센서들이 활용되고 있다. 대표적으로 RGBD 카메라, 스테레오 카메라, 2D LiDAR가 사용되고 있으며, 제작단가를 극단적으로 낮추기 위해 별도의 비콘 디바이스를 사용하기도 한다. 본 연구에서는 가격과 사용자 편의성 관점에서 매력적인 2D LiDAR 센서를 다루고자 한다.

영상 기반의 객체 추적은 컴퓨터 비전 분야의 전통적인 문제이고 3D

LiDAR를 활용하는 객체 추적을 위한 인공지능 기술도 최근 자율주행 자동차 분야에서 연구가 활발히 수행되고 있다. 그에 비해 2D LiDAR 기반 객체 추적 연구는 상대적으로 덜 주목받아 왔으며 심층 신경망을 활용하는 연구는 아직 발표되지 않았다. 그렇게 된 원인 중 하나는 심층 신경망을 개발하기 위해 요구되는 데이터셋의 부재일 것이다. 최근 2D LiDAR 기반의 DROW 데이터셋[2]이 발표되었지만 DROW 데이터셋은 사람 검출을 위해 가공되어 사용자 추적 기술 개발에는 적합하지 않다.

본 논문에서는 2D LiDAR 기반 사용자 추적을 위한 인공지능 기술 개발을 위해 사용자 추적 데이터셋을 손쉽게 구축하기 위한 방법을 제안한다. 이미지와 비교하면 2D LiDAR 데이터는 수집된 점들의 개수는 픽셀 개수에 비해 상대적으로 적지만 각각의 점들이 추적하고자 하는 사용자에 해당하는지를 판단하는 것이 상대적으로 쉽지 않다. 특히, 사용자가 벽이나 가구와 같은 구조물이나 다른 사람들과 같은 동적 객체 주변에 위치한 경우에는 각 점들이 추적 대상인 사용자인지 아닌지를 판단하는 것은 더욱 어렵다. 그럼에도 불구하고, 2D LiDAR 데이터를 효과적으로 레이블링(labeling)하기 위해 본 논문은 2D LiDAR와 함께 카메라를 활용하는 것

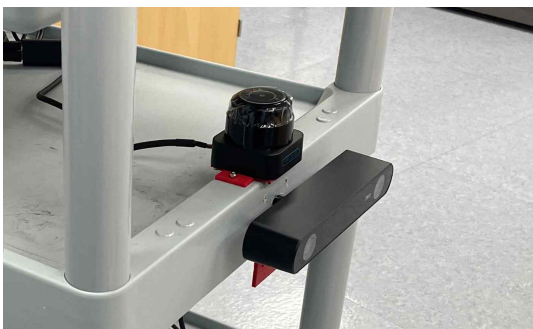


그림 1. 이동 카트에 장착된 2D LiDAR와 카메라

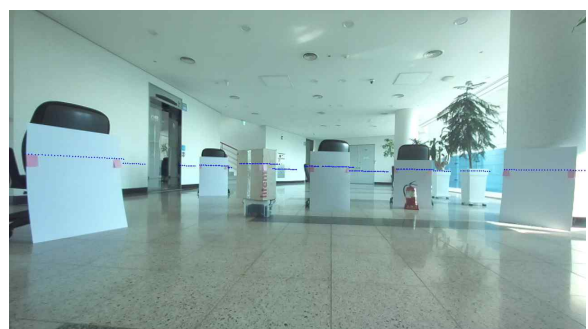


그림 2. 2D LiDAR와 카메라 사이의 캘리브레이션 결과



그림 3. (좌) 수집된 이미지, (중) Segment anything 수행 결과, (우) 추적 대상만 추출한 결과

을 제시한다. 우선, 카메라-2D LiDAR 캘리브레이션[3]과 두 센서 데이터 간의 시간 동기화를 수행한다. 제안하는 방법의 핵심은 카메라로부터 수집된 이미지에 의미론적 분할 (semantic segmentation) 기술을 적용하는 것이다. 이를 통해 추적 대상인 특정 사용자에게 해당하는 픽셀들만을 저장하고, 그런 다음 2D LiDAR에 의해 수집된 점들을 이미지에 투영하면 저장된 픽셀들의 위치 정보를 통해 추적 대상인 사용자에게 해당하는 점들을 다른 점들과 구분할 수 있다.

II. 본 론

1. 카메라와 2D LiDAR의 시간 동기화 및 캘리브레이션

본 연구에서는 데이터 수집을 위해 SLAMTEC 사의 2D LiDAR인 RPLiDAR S1과 StereoLabs 사의 스테레오 카메라인 ZED 2i를 그림 1과 같이 이동 카트에 장착하였다. 두 센서 데이터 간의 시간 동기화를 위해 ROS 2의 message filter를 활용하였으며, 캘리브레이션을 위해 2D LiDAR의 데이터의 11개 점들과 이들에 대응되는 이미지의 11개 픽셀들을 바탕으로 OpenCV의 solvePnP함수를 활용하였다. 이 과정에서 2D LiDAR의 점들에 대응되는 이미지 픽셀들을 찾기 위해 추가적으로 Intel 사의 Realsense D435 카메라의 적외선 이미지를 활용하였다. 그림 2는 두 센서 사이의 캘리브레이션 결과를 보여주고 있다.

2. 영상 분할 및 2D LiDAR 데이터 어노테이션

본 연구에서는 추적 대상인 특정 사용자에게 해당하는 점들과 그렇지 않은 점들을 구분하기 위해 영상 분할 기법을 활용한다. 특히 최근에 우수한 성능으로 주목받고 있는 segment anything[4]을 활용한다. Segment anything은 영상 분할을 위한 기초 모델(foundation model)을 제시했다는 점에서 의미가 있는 연구로 영상의 의미론적 분할을 위해 사용자로부터 마우스 클릭, 객체 검출을 통한 경계 상자(bounding box), 텍스트를 입력받아 원하는 대상의 픽셀들을 추출할 수 있으며, 이미지 전체를 분할하는 zero-shot generalization 기능도 제공한다. 그림 3의 왼쪽 그림과 가운데 그림은 실제 수집된 영상과 segment anything을 적용한 결과를 보여주고 있다.

동기화된 2D LiDAR 데이터와 분할된 영상이 주어지면 앞서 언급한 캘리브레이션 결과를 통해 2D LiDAR 데이터에 포함된 점들을 영상에 투영할 수 있다. 이때 영상 분할을 통해 그림 3의 오른쪽 그림과 같이 사용자에게 해당하는 픽셀들을 알고 있기 때문에, 2D LiDAR 데이터 각 점이 투영되는 영상 좌표값이 추적 대상인 사용자 영역에 포함되어 있는지를 판단할 수 있다. 그림 4는 그림 3의 왼쪽 그림과 함께 수집된 2D LiDAR 데이터의 일부를 보여주고 있다. 제안하는 방법을 통해 사용자에게 해당하는 점으로 판단된 점들은 빨강색 십자가(+)로 표시하였다. 이 결과는 제안하는 방법을 통해 2D LiDAR 데이터의 점들 중에 사용자에게 해당하는 점들을 효과적으로 추출할 수 있는 것을 의미한다.

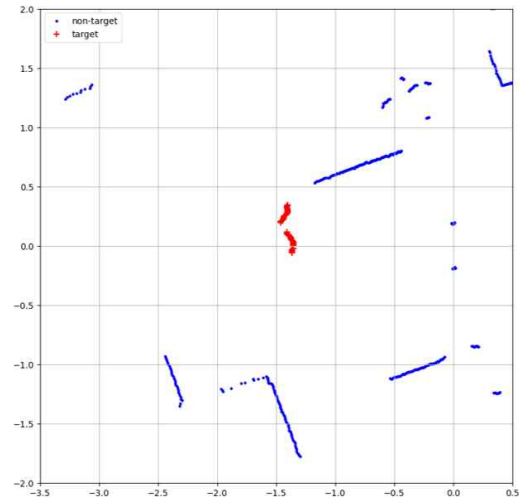


그림 4. 어노테이션 결과

III. 결 론

본 논문에서는 2D LiDAR 기반 사용자 추적 기술 개발을 위해 데이터셋을 구축하는 과정에서 효과적으로 레이블링 하는 방법을 제안하였다. 앞으로 제안하는 방법을 통해 2D LiDAR 기반 사용자 추적 기술 개발을 위한 데이터셋을 구축하고 이를 통해 효과적인 2D LiDAR 기반 사용자 추적을 위한 인공지능 기술을 개발할 예정이다.

ACKNOWLEDGMENT

본 논문은 정부(과학기술정보통신부)의 재원으로 과학기술사업화진흥원의 지원을 받아 수행된 연구임(‘학연협력플랫폼구축 시범사업’ RS-2023-00304776).

참 고 문 헌

[1] M. D. Islam et al, "Person Following by Autonomous Robots: A Categorical Overview," International Journal of Robotic Research, vol. 38, issue 14, pp. 1581-1618, 2019.

[2] L. Beyer et al, "Deep Person Detection in 2D Range Data," IEEE Robotics and Automation Letters (RA-L), vol. 3, no. 3, pp. 2726-2733, 2018.

[3] K. W. Park and S. Y. Park, "Visual LiDAR Odometry Using Tree Trunk Detection and LiDAR Localization," The International Archives of The Photogrammetry, Remote Sensing and Spatial Information Sciences (ISPRS), 2023.

[4] A. Kirillov et al, "Segment Anything," IEEE/CVF International Conference on Computer Vision (ICCV), 2023.