

센서융합 기반 의미론적 SLAM 및 내비게이션

이기현¹, 안승현¹, 신수현¹, 류혜선¹, 홍유나¹

¹인하대학교 기계공학과

leekh951@naver.com, asho227@naver.com, ssh@inha.edu, fbgPtjs@inha.edu,

hongyuna64@naver.com

Semantic SLAM & Navigation Based on Sensor Fusion

Gihyeon Lee¹, Seung-hyun Ahn¹, Suhyeon Sin¹, Hyesun Ryu¹, Yuna Hong¹,

¹Dept. of Mechanical Engineering, Inha University

요 약

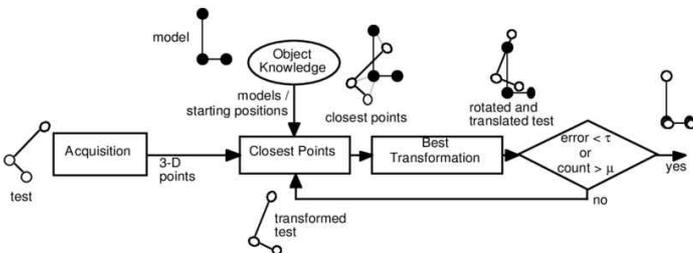
본 논문은 로봇의 실내 환경에서의 자율성을 높이기 위한 SLAM과 내비게이션 방법을 제시한다. 2D LiDAR와 카메라를 이용하여 위치를 인식하고 사람과 장애물을 의미론적으로 검출하며, ICP와 RTAB-map, YOLOv3를 통합하여 Semantic Map을 생성하고 실내 환경에서 자율성을 유지한다. 이 연구를 통해 로봇이 복잡한 환경에서도 높은 수준의 자율성을 유지할 수 있는지 확인하고자 한다.

1. 서론

본 논문에서는 실내 평지 운용에 특화된 바퀴형 로봇을 운용하기 위해 LiDAR, 카메라 센서 등을 이용하여 사람과 장애물을 의미론적으로 파악하고, 로봇이 이용할 수 있는 Semantic Map을 작성하여 장기적으로 로봇의 장기 자율성을 유지할 수 있는 SLAM & Navigation 기술을 구현하였다.

2. 이론적 배경

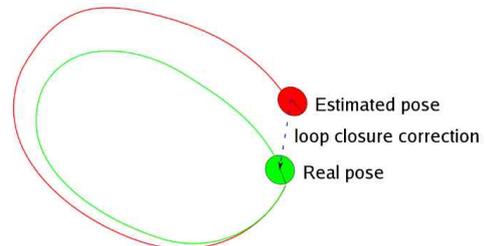
- Iterative Closest Point (ICP)



(그림 1) ICP Algorithm

LiDAR ICP Odometry는 (그림 1)과 같이 ICP 알고리즘을 활용하여 로봇의 위치를 추정하는 기술이다. 이 방법은 LiDAR 센서로부터 얻은 연속된 두 프레임의 포인트 클라우드 데이터를 초기 정렬한 뒤, ICP 알고리즘을 적용하여 더욱 정확하게 정렬한다. 이렇게 얻은 변환 행렬을 통해 로봇의 위치와 방향을 업데이트한다.

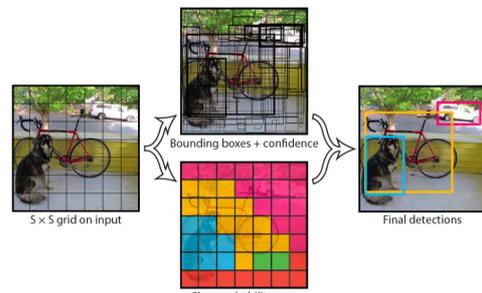
- Odometry Loop Closure



(그림 2) Loop Closure Algorithm

RTAB-map에서는 Visual Odometry와 Loop Closure라는 두 가지 기술이 중요한 역할을 수행한다. Visual Odometry는 연속된 이미지 프레임에서 특징점을 추출하여 로봇의 위치를 실시간으로 추정하고, 이 정보는 맵 데이터베이스에 저장된다. 이와 병행하여 Loop Closure는 (그림 2)처럼 로봇이 이전에 방문한 장소에 다시 도착했을 때 이를 감지하여 Map의 일관성을 유지하고 오차를 보정한다.

- Image Object Detection & Tracking



(그림 3) YOLO Algorithm

YOLOv3(You Only Look Once v3) 기반의 다중 객체 검출 및 추적은 실시간 이미지 처리에서 높은 성능을 보인다. (그림 3)과 같이 YOLOv3는 단일 신경망을 사용하여 이미지를 한 번만 훑으면서 다수의 객체를 빠르고 정확하게 검출하며, 시간에 따른 객체의 움직임을 연속적으로 추적할 수 있다.

3. 제안 방법

- ICP Odometry

실내 환경에서 더욱 강건한 위치 인식을 할 수 있도록 2D LiDAR 기반의 ICP Odometry를 생성하는 알고리즘을 설계한다. 아래와 같이 ICP를 통한 라이더 데이터의 점군 정합을 활용하여 로봇의 위치 추정하고 실시간으로 update할 수 있도록 한다.

$$\forall p \in P, q_{closest} = \arg \min_{q \in Q} \| p - q \|$$

$$(R^*, t^*) = \arg \min_{R, t} \sum_{i=1}^N \| (R \cdot p_i + t) - q_{closest, i} \|^2 \quad (1)$$

ICP(Iterative Closest Point) 알고리즘은 아래와 같이 ICP Cost Function을 정의한다. 두 포인트 셋 P와 Q 사이에서 P의 각 점에 가장 가까운 Q의 점을 찾아, 최적의 회전 R과 이동 T를 계산하여 P를 변환한 후, 비용 함수가 특정 임계값에 도달할 때까지 이 과정을 반복한다.

$$E(R, t) = \sum_{i=1}^N \| (R \cdot p_i + t) - q_{closest, i} \|^2$$

$$|E(R, t)_{new} - E(R, t)_{old}| < Threshold \quad (2)$$

- Visual Loop Closure & Mapping

지도작성 중 카메라의 RGB image를 이용하여 Feature Point Extraction을 수행하고 이 데이터를 활용하여 Visual Loop Closing을 연산하면서 로봇의 위치와 경로를 최적화한다.

$$E_{loop}(X) = \sum_{(i, j) \in C} \| X_i - X_j \|^2 \quad (3)$$

$$E_{SLAM} = E_{odom} + \lambda E_{loop} \quad (4)$$

RTAB-map 알고리즘을 활용하여 실시간으로 최적화되는 로봇의 Pose를 기준으로 하여 주변 환경을 정확하게 3D point cloud (Depth camera) 및 2D Occupancy grid(2D LiDAR)로 mapping(M)을 할 수 있도록 한다. [1]

(K : Intrinsic matrix, D : Depth information)

$$P = K^{-1} D$$

$$M_{updated} = M_{old} \cup T_{odom} \cdot P_{new} \quad (5)$$

- 3D Voxel based Object Detection & Tracking

YOLOv3 (Object Detection) 알고리즘을 사용하여 카메라를 통해 탐지한 객체 주변의 Bounding Box = $(x_{2D}, y_{2D}, w_{2D}, h_{2D})$ 내 픽셀 정보와 깊이 정보 (RGB-D)를 활용하여 해당 객체의 카메라로부터의 위치를 계산하고 Voxel화 하여 3D Bounding Box = $(x_{3D}, y_{3D}, z_{3D}, w_{3D}, h_{3D}, d_{3D})$ 를 생성한다.

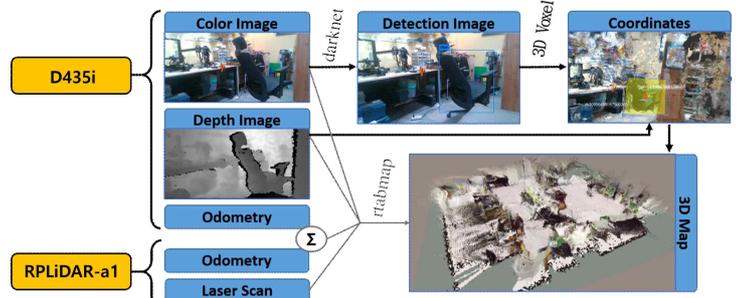
$$(x_{cam}, y_{cam}, z_{cam}) = K^{-1} \times \begin{pmatrix} x_{2D} \\ y_{2D} \\ 1 \end{pmatrix} \times D \quad (6)$$

$$w_{3D} = w_{2D} \times \frac{z_{cam}}{f} \quad x_{3D} = x_{cam}$$

$$h_{3D} = h_{2D} \times \frac{z_{cam}}{f} \quad y_{3D} = y_{cam}$$

$$d_{3D} = D \quad z_{3D} = z_{cam} \quad (7)$$

위 과정들을 통해, 객체의 위치에 3D마커를 등록하면서 최종적으로 Semantic map을 구현한다. 본 논문에서 제안한 전체 프로세스는 (그림 4)와 같다.



(그림 4) Semantic mapping Pipeline

4. 결론

본 논문은 2D LiDAR 및 Depth Camera만을 사용하여 실내 환경에서 로봇의 장기 자율성을 유지하는 SLAM 기술에 집중하였다. 향후 계획으로는 생성한 Semantic Map을 기반으로 물류창고와 같은 복잡한 환경에서의 로봇의 자율성을 평가할 예정이다. 후속 연구가 성공적으로 진행된다면, 물류 및 서비스 분야에서 큰 기대효과를 가져올 것으로 예상된다.

참고문헌

[1] Labbé, Mathieu, and François Michaud. "RTAB Map as an Open source Lidar and Visual Simultaneous Localization and Mapping Library for Large scale and Long term Online Operation." Journal of field robotics 36.2 (2019): 416 - 446.

※ 본 논문은 과학기술정보통신부 정보통신창의인재 양성사업의 지원을 통해 수행한 ICT멘토링 프로젝트 결과물입니다.