

도시환경 매핑 시 SLAM 불확실성 최소화를 위한 강화 학습 기반 경로 계획법

RL-based Path Planning for SLAM Uncertainty Minimization in Urban Mapping

조영훈¹, 김아영[†]

Younghun Cho¹, Ayoung Kim[†]

Abstract: For the Simultaneous Localization and Mapping (SLAM) problem, a different path results in different SLAM results. Usually, SLAM follows a trail of input data. Active SLAM, which determines where to sense for the next step, can suggest a better path for a better SLAM result during the data acquisition step. In this paper, we will use reinforcement learning to find where to perceive. By assigning entire target area coverage to a goal and uncertainty as a negative reward, the reinforcement learning network finds an optimal path to minimize trajectory uncertainty and maximize map coverage. However, most active SLAM researches are performed in indoor or aerial environments where robots can move in every direction. In the urban environment, vehicles only can move following road structure and traffic rules. Graph structure can efficiently express road environment, considering crossroads and streets as nodes and edges, respectively. In this paper, we propose a novel method to find optimal SLAM path using graph structure and reinforcement learning technique.

Keywords: SLAM, Path Planning, Reinforcement Learning, Mobile Robot

1. 서 론

SLAM 연구는 대부분의 경우 카메라, 라이다, 레이더 등의 센서를 이용해 취득된 센서 데이터의 순서를 따라 진행된다. 한편, 다른 경로를 통해 얻어진 센서 데이터는 서로 다른 SLAM의 결과를 초래한다. 따라서, SLAM의 결과를 향상시키기 위해서는 데이터 취득 시에 보다 나은 경로를 계획하는 것이 필요하다. 하지만, 지금까지 이루어진 SLAM 관련 선행 연구에서는 데이터 취득 경로에 대한 언급이 없는 경우가 대부분이고 따라서 데이터 취득은 사람의 편의 또는 직관에 의해 이루어지는 경우가 많다. 이러한 문제점은 Active SLAM 이라

는 분야에서 오랜 기간 논의되어 왔으며, SLAM을 개선하는 path를 동시에 구하는 최적화 방법이 다양하게 제시되었다^[1]. 본 논문에서는 SLAM의 결과를 향상시킬 수 있는 데이터 취득 경로를 계획하는 방법을 제시하고자 한다.

SLAM의 결과는 크게 정확도와 범위 두 가지의 기준으로 판단될 수 있다. 정확도를 높이기 위해서는 같은 지역을 재방문하는 횟수를 최소화하는 경로가 필요하고 범위를 높이기 위해서는 지금까지 방문하지 않은 새로운 지역을 방문하는 경로가 필요하다. 따라서 경로의 길이가 길어질수록 정확도 또는 범위 중 한 가지 기준에서는 SLAM의 결과가 향상된다고 생각할 수 있다. 하지만, 센서시스템이 가동될 수 있는 시간, 범위 등이 주어진 자원에 따라 한정되어 있으므로 정해진 경로의 길이 내에서 SLAM의 결과를 향상시킬 수 있는 경로를 계획하는 것이 필요하다.

지금까지 SLAM을 위한 경로 계획에 대한 연구는 실내 환경이나 공중과 같이 센서 시스템이 모든 방향으로 자유롭게 움직일 수 있는 환경에서 실행되어왔다. 하지만, 본 논문에서 대상으로 삼고있는 도시환경에서는 주로 차량에 의한 데이터

Received : Feb. 26. 2021; Revised : Apr. 2. 2021; Accepted : Apr. 5. 2021

※ This research was supported by a grant (21TSRD-B151228-03) from Urban Declining Area Regenerative Capacity-Enhancing Technology Research Program funded by Ministry of Land, Infrastructure and Transport of Korean government

1. Doctoral Student, Dept. of Civil and Environmental Engineering, KAIST, Daejeon, Korea (iriter@kaist.ac.kr)

† Associate Professor, Corresponding author: Dept. of Civil and Environmental Engineering, KAIST, 291 Daehak-ro, Yuseong-gu, Daejeon, Korea (ayoungk@kaist.ac.kr)

취득이 활발히 이루어지고 있다. 플랫폼을 차량으로 가정할 경우 센서 시스템이 움직일 수 있는 방향은 제한되게 된다. 또한, 차량을 이용하는 센서 시스템의 경우, 차도로만 주행할 수 있기 때문에 갈 수 있는 공간의 제약이 있고 교통 법규에 따라서만 움직일 수 있기 때문에 취할 수 있는 행동에도 추가적인 제약이 따른다.

본 논문에서는 이러한 도시환경의 특성을 그래프 구조로 나타낼 수 있다는 점에 착안하여, 도심 내 active SLAM 문제를 해결하려고 한다. 도로 구조는 교차로를 노드, 교차로 사이에 연결된 도로를 간선으로 생각했을 때 그래프 구조로 나타낼 수 있다. 또한, 교차로에서 한번 진행 방향을 정하면 다음 교차로가 나올 때 까지 방향을 바꿀 수 없다는 점이 그래프 상에서 경로를 찾는 것에 부합한다. 각 간선의 방문 횟수를 SLAM의 정확도, 각 간선의 방문 여부를 SLAM의 범위라고 생각하면, SLAM의 성능을 향상시키는 경로를 그래프 구조 상에서 계획할 수 있다.

기준에 그래프 구조에서 최단 경로, 최다 방문 경로 등을 구하는 알고리즘들은 이미 많이 연구되어 왔다. 하지만 기준에 있는 알고리즘들은 계산복잡도가 $O(N^3)$ 으로 SLAM의 대상 지역이 커짐에 따라 경로를 계획하는데 걸리는 시간이 기하급수적으로 늘어나게 된다. 따라서, 본 논문에서는 강화 학습 기법을 적용하여 $O(N)$ 의 계산복잡도를 가지는 경로 계획법을 제시하고자 한다.

다음 항목들은 본 논문에서 제시하고자 하는 논점들이다.

- 기존의 움직임에 제한이 없는 로봇이 아닌 차량을 대상으로 한 문제 정의를 제시한다. 도심 지역을 그래프 구조로 변환하여 도시 환경에서 적용 가능한 경로 계획법을 제시하였다.
- 강화 학습 기법을 사용하여 그래프의 크기와 관계없이 일정한 계산 복잡도를 가지는 경로 계획법을 제시하였다. 이는 도심 환경과 같이 규모가 증가하는 맵핑 문제에서 중요하다.
- 그래프 구조와 강화 학습 기법을 적용하여 SLAM의 정확도와 범위를 모두 최적화하는 경로 계획법을 제시하였다.

2. 선행 연구 조사

2.1 그래프 구조에서의 경로 계획법

그래프 구조에서 경로 계획법은 주로 최단 경로를 구하는 문제를 중심으로 연구되어왔다. 최단 경로를 구하는 경로 계획법은 [2]에서 소개된 다익스트라 알고리즘과 [3]에서 소개된 A* 검색 알고리즘이 대표적이다. 다익스트라 알고리즘은 시작 노드부터 시작하여 인접 노드들의 점수를 현재 노드의 점수와 인접 노드까지 필요한 비용의 합으로 업데이트 하고,

최소의 점수를 가지는 노드로 이동하면서 종결 노드에 도착할 때 까지 반복하여 최단 경로를 찾아낸다. 이 과정에서, 같은 노드를 여러 번 방문하면 해당 노드가 원래 가지고 있던 점수와 새로 구한 점수 중 작은 것을 선택한다. 다익스트라 알고리즘은 최적해를 구할 수 있지만 $O(N^3)$ 의 계산복잡도를 가진다는 단점이 존재한다. A* 검색 알고리즘은 다익스트라 알고리즘에서 사용한 점수에 도착점까지의 거리를 더한 값을 점수로 사용한다. A* 검색 알고리즘도 최적해를 구할 수 있지만 branching factor b와 그래프의 깊이 d에 대해 $O(b^d)$ 의 계산복잡도를 가지기 때문에 그래프의 크기가 커질 수록 사용하기 힘들다는 단점이 있다.

한편, 모든 노드를 방문하는 최단 경로를 구하는 연구로는 여행하는 외판원 문제가 대표적이다. 여행하는 외판원 문제는 N개의 노드를 가지는 그래프에서 노드를 순서대로 방문하는 모든 경우를 비교하므로 $O(N!)$ 의 매우 큰 계산복잡도를 가진다. 동적계획법을 사용하면 계산복잡도를 $O(2^N \times N^2)$ 까지 낮출 수 있지만 여전히 사용하기에는 큰 계산복잡도를 가진다. 지금까지의 경로 계획법 연구를 보면, 계산복잡도를 줄이는 것이 매우 도전적인 과제임을 알 수 있다.

2.2 강화 학습을 사용한 경로 계획법

2.1에서 언급한 경로 계획법들의 계산복잡도를 줄이는 방법으로는 강화 학습을 사용하는 것이 하나의 대안으로 연구되어왔다. [4]에서 소개된 Singh의 논문에서는 Neumann method와 Dirichlet method를 혼합하여 보상을 설계하여 두 방법의 장점을 모두 취하려고 하였다. [5]에서는 parti-game 알고리즘을 제안하여 지연된 보상을 통해 2차원에서 9차원까지의 공간에서 적용 가능한 알고리즘을 제안하였다. [6]에서는 강화 학습 기법을 확률적 로드맵에 적용하여 동적인 환경에서 경로 계획법을 제안하였다. [7]에서는 UAV에 특화된 보상을 적용하여 geometric 강화 학습이라는 개념을 제안하였다.

지금까지 연구된 내용들은 모두 자유롭게 움직일 수 있는 환경을 대상으로 하였다. 하지만 본 연구에서는 도심 환경에서 차량을 통한 SLAM에 집중된 접근을 제시한다. 구체적으로, 본 연구에서는 그래프 구조를 활용하여 도시 환경에 적합한 강화 학습 방법을 제안하고자 한다.

2.3 강화 학습을 SLAM에 적용한 연구

강화 학습을 SLAM에 적용하여 결과를 향상시키기 위한 연구는 다음과 같이 진행되어왔다. [8]에서 Kollar는 강화 학습을 사용하여 야외 환경에서 map exploration을 위한 경로 최적화에 대해 연구하였다. 하지만 이 연구에서는 다음으로 방문할

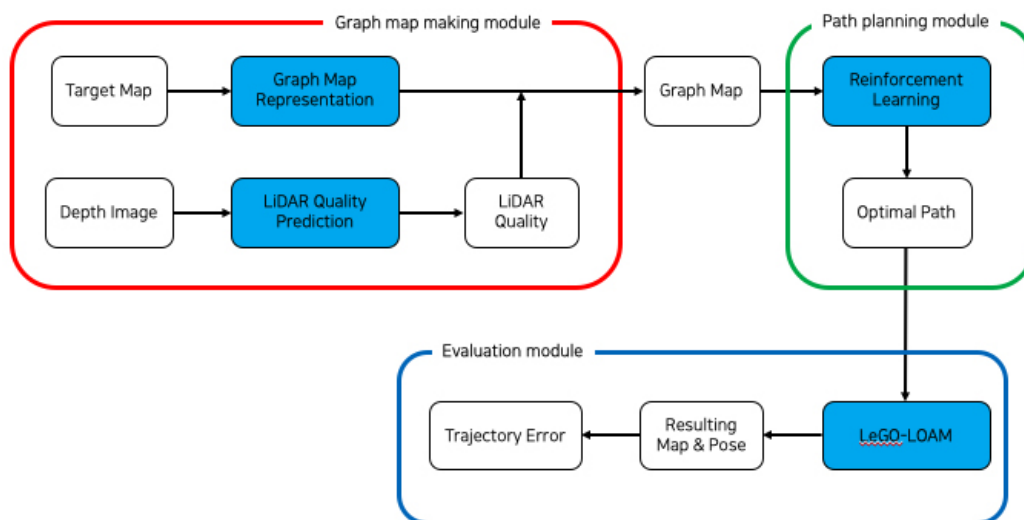
위치가 정해져 있는 상태에서 그 위치까지의 모션을 최적화하는 연구로 다음에 어느 지역을 탐사할지 직접 정하고자하는 본 논문의 목적과는 약간의 차이가 있다. [9]에서는 야외 환경에서 확률적 코스트맵을 사용하여 A* 검색 알고리즘과 D* 검색 알고리즘에 비해 90% 향상된 범위와 60% 향상된 검색 시간을 도출하였다. [10]에서는 항공 사진을 이용하여 다양한 지형이 혼합되어있는 환경에서 진동과 전력 사용을 최소화하는 경로를 계획하는 연구를 진행하였다. 야외 환경에서 실행되었지만 이 연구에서도 로봇이 자유롭게 움직일 수 있는 환경에서 연구되었다. 따라서, 본 연구에서는 야외 환경 중에서도 도시 환경, 특히 도로 구조를 반영하여 경로를 계획하는 방법을 제시하고자 한다.

3. 연구 방법

본 논문에서 제시하는 방법론은 크게 세 부분으로 이루어져 있다. 첫번째로는 그래프 지도를 제작하는 모듈로, 주어진 지역의 위치 정보들과 깊이 이미지를 사용하여 그래프 구조와 각 간선의 가중치를 만들어 그래프 지도를 완성한다. 두번째로는 경로 계획 모듈로, 만들어진 그래프 지도와 강화 학습 기법을 사용하여 SLAM의 결과를 향상시킬 수 있는 최적 경로를 계획한다. 마지막으로는 평가 모듈로, 구해진 최적 경로를 따라 데이터를 취득하고, 취득한 데이터를 사용하여 LeGO-LOAM^[11]을 사용하여 SLAM을 수행하고, 결과로 나오는 경로와 실제 경로를 비교하여 경로의 오차를 계산한다. 방법론에 대한 전체 과정은 [Fig. 1]에 그림으로 설명되어 있다.

3.1 그래프 지도 제작

그래프 지도 제작은 크게 노드 선정, 간선 클러스터링, 그래프 지도 제작, 가중치 부여 의 4단계로 나누어져 있다. 그래프 지도 제작의 전 과정은 아래의 [Fig. 2]에 그림으로 표시되어 있다. 제일 먼저, 주어진 위치 정보들로부터 도로의 교차로에 해당하는 노드를 선정하는 과정을 거친다. 교차로로 판단되는 점은 해당 지점으로 들어오는 도로가 3개 이상이어서 직진 이외의 행동을 선택할 수 있는 지점일 것이다. 따라서, 각 위치에서 해당 지점을 중심으로 하는 내부 반지름 10 m, 외부 반지름 25 m, 높이 10 m의 공동실린더에 해당하는 점들을 클러스터링하여 클러스터의 개수가 3개 이상인 점들을 노드로 판단하였다. 높이를 포함하여 클러스터링 한 이유는 고가도로, 지하차도 등의 도로구조가 있을 경우, 실제로 직진만이 가능한 구조이지만 높이를 고려하지 않으면 교차로로 인식될 가능성이 있기 때문이다. 다음으로 간선 클러스터링은 노드로 선정된 위치와 그 주변 위치들을 제거한 후 남은 위치들을 클러스터링하여 각각의 클러스터를 하나의 간선으로 클러스터링 하였다. 세번째로, 그래프 지도 제작은 선정된 노드들과 간선들을 연결하는 작업이다. 각각의 간선에 연결된 노드와 간선의 끝 점은 모든 점들 중에 가장 가까운 거리에 있을 것이다. 따라서 각각의 간선에서, 간선의 모든 점에서 모든 노드까지의 거리를 구한 뒤, 가장 가까이 위치하고 있는 두 개의 노드를 찾아 간선에 연결된 노드로 지정해 주었다. 가장 가까운 두 개의 노드 j, k 를 찾는 식은 아래의 (1), (2)번 식과 같다.



[Fig. 1] Path planning algorithm on graph structure using reinforcement learning. Graph map making module makes graph map utilizing location information and depth images of target area as input. In path planning module, it plans an optimal path using graph map obtained in graph map making module. Finally, in evaluation module, it performs SLAM following the optimal path and calculates trajectory error using LeGO-LOAM.

$$[idx, j] = \underset{idx, j}{\operatorname{argmin}}(\operatorname{edge}(idx), \operatorname{node}(j)) \quad (1)$$

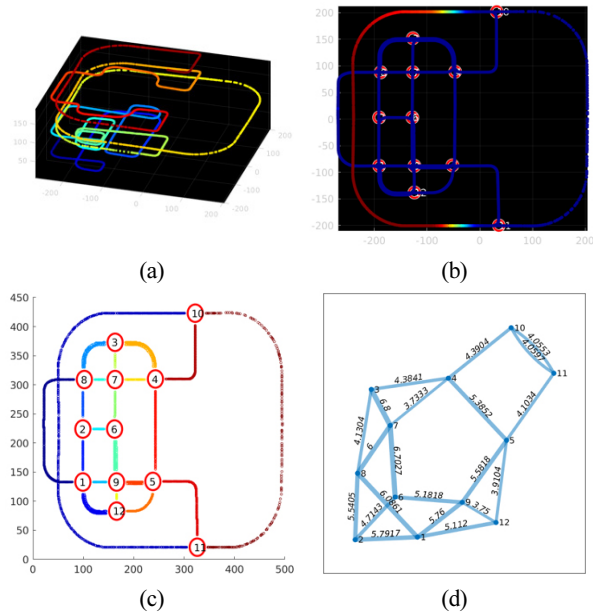
$$[idx, k] = \underset{idx, k}{\operatorname{argmin}}(\operatorname{edge}(idx), \operatorname{node}(k)), k \neq j \quad (2)$$

마지막으로, 깊이 이미지를 사용하여 각 위치의 라이다 센서의 매칭 퀄리티 점수를 구하여 각 간선에 포함되는 위치들의 점수의 평균을 각 간선의 가중치로 사용하였다. 평가 모듈에서 사용하는 LeGO-LOAM은 평면과 모서리를 특징점으로 사용하여 SLAM을 수행하기 때문에 다양한 방향을 향하고 있는 평면과 모서리가 있을수록 라이다의 매칭이 잘 이루어질 것이라고 판단하였고 따라서 각 점들의 법선 벡터가 다양하게 분포하고 있을수록 라이다 매칭 퀄리티 점수를 높게 부여하였다. 가장 먼저, 깊이 이미지의 각 점 (x, y) 들을 그 점에서의 깊이 $D_{x,y}$ 를 사용하여 3차원 공간에 아래의 (3) 식을 사용하여 (X, Y, Z) 로 재투영해주었다. 법선 벡터들의 분포를 구하기 위함이므로 (3) 식과 같이 깊이 방향으로만 재투영하였다. 재투영한 결과는 [Fig. 3]의 (a)와 같다.

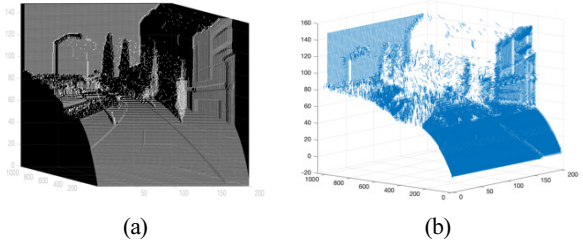
$$X = y \quad (3)$$

$$Y = D_{x,y}$$

$$Z = W - x$$



[Fig. 2] (a): Raw location data of target area. (b): Extracted nodes. Nodes represent crossroad in urban environment. Change of direction only can occur at the crossroad. At the left side of the map, there are elevated road, so they are not considered as crossroads. (c): Clustered Edges. Each color represents each edge. (d) A complete graph map connecting nodes and edges. It also includes calculated weights.



[Fig. 3] (a): Reprojected 3D points using equation (3). Depth values are assigned to each pixel to get 3D coordinates of each pixel. (b): Calculated normal vectors

다음으로는 재투영된 각 점의 법선 벡터를 구해주었다. 각 점에서 주변 14개 점을 선정하여 이를 하나의 타원체로 생각했을 때, 이는 납작한 모양의 거의 타원에 가까운 도형일 것이므로 3개의 축 중에서 가장 짧은 축이 타원의 법선 벡터라고 생각할 수 있다. 따라서, 식 (4)를 사용하여 각 점의 법선 벡터를 구해주었다. C 는 14개의 점으로 이루어진 타원의 분산 행렬, \vec{v}_o 은 가장 작은 아이젠벡터를 뜻한다.

$$\bar{p} = \frac{1}{14} \sum_{i=1}^{14} p_i \quad (4)$$

$$C = \frac{1}{14} \sum_{i=1}^{14} (p_i - \bar{p}) \cdot (p_i - \bar{p})^T$$

$$C \cdot \vec{v}_o = \lambda_0 \cdot \vec{v}_o$$

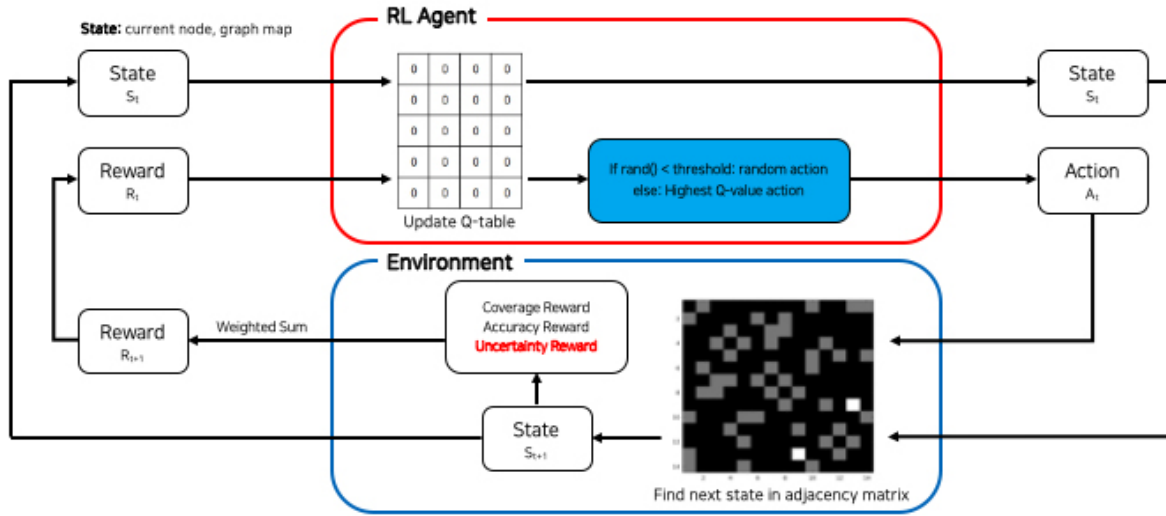
구해진 법선 벡터는 [Fig. 3]의 (b)와 같다.

3.2 그래프 지도에서 강화 학습을 통한 경로 계획

3.1에서 만든 그래프 지도를 바탕으로, 3.2에서는 모든 지점을 방문하면서 불확실성을 최소화하는 경로를 계획한다. 앞에서 언급한대로, 기존의 그래프 경로 계획 알고리즘들은 높은 계산복잡도를 가지기 때문에 이를 해결하기위해 강화 학습을 도입하여 그래프의 크기와 관계없이 일정한 계산복잡도를 가지는 경로 계획법을 소개하고자 한다.

강화 학습을 이용한 경로 계획법의 전체 과정은 [Fig. 4]와 같다. 시간 t 에서의 상태 S_t 와 보상 R_t 가 입력으로 주어지면, 강화 학습 에이전트에서 Q-table이 식 (5)에 따라 업데이트 되고, Q-table의 값에 따라 현재 상태에서 가장 좋다고 판단되는 행동 A_t 를 출력으로 반환한다. 본 논문에서는 γ 의 값을 0.25, 0.5, 0.75, 0.9 중 가장 좋은 성능을 보인 0.25로 설정하여 강화 학습 모델을 학습시켜주었다.

$$Q_{\pi}(S_t, A_t) = R_t + \gamma \cdot \max(S_{t+1}, A_{t+1}) \quad (5)$$



[Fig. 4] Overall pipeline of reinforcement learning. RL agent takes state and reward as input and gives actions as output. Utilizing state and action, environment calculates new state and rewards: coverage reward, accuracy reward, and uncertainty reward. Finally, it calculates its next reward as weighted sum of those three reward terms. While every node is visited it repeats all the processes

상태 S_t 와 행동 A_t 가 환경에 입력으로 주어지면, 환경에서는 그래프 지도의 인접 행렬을 사용하여 다음 시간인 $t+1$ 에서의 상태 S_{t+1} 을 구한다. 상태 S_t 에서 S_{t+1} 로 진행되는 과정에서의 변화를 가지고 환경에서는 다음 시간 $t+1$ 에서의 보상 R_{t+1} 을 구한다. 보상은 범위 보상 R_C , 정확도 보상 R_A , 불확실성 보상 R_U 3개의 보상으로 이루어져 있고 이들의 가중 합을 보상으로 반환한다. 각각의 보상을 구하는 방법은 아래의 식 (6)에 설명되어 있다. 불확실성 보상 R_U 의 경우 EIF (Extended Information Filter)를 통해 얻어진 information matrix L 의 행렬식을 사용하였다. Information matrix는 covariance matrix와 반대로 둘 사이의 관계가 정확 할수록 큰 값을 가진다.

$$\begin{aligned} R_{C,t+1} &= l_{path} - R_{C,t} \\ R_{A,t+1} &= \text{weight}(S_t, S_{t+1}) \\ R_{U,t+1} &= \det(L(1:3, 1:3)) \end{aligned} \quad (6)$$

이를 이용한 최종 보상 R_t 는 아래 식 (7)에 나타나있는 가중 합을 사용하였다. 평균적으로 계산되는 각각의 보상의 크기가 가중치와 곱해져 비슷한 값을 가지도록 가중치를 정해주었다.

$$R_t = R_{C,t} + R_{A,t} + 20 \cdot R_{U,t} \quad (7)$$

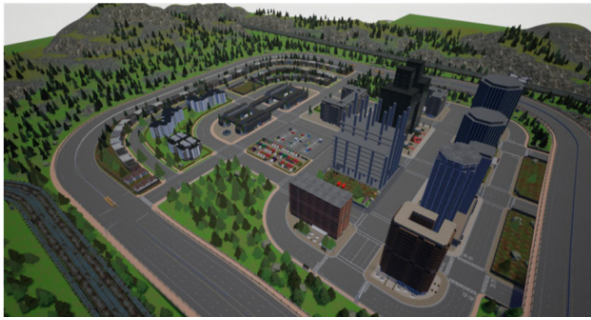
위 과정을 그래프 지도 상의 모든 노드를 방문할 때 까지 반복해주면 한 번의 학습 에피소드가 끝나게 된다. 본 논문에서는 총 1000번의 에피소드를 통해 강화 학습 모델을 학습 시켜주었다.

3.3 LeGO-LOAM을 통한 라이다 지도 제작 및 이동 궤적 오차 계산

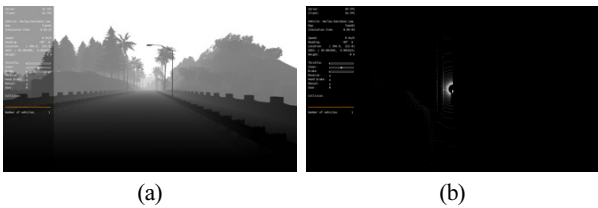
마지막으로, 강화 학습을 통해 얻은 경로가 SLAM의 성능을 얼마나 향상시켰는지 확인하기 위해 강화 학습으로 계산한 경로와 임의의 경로 두 가지 경우에 대해 LeGO-LOAM을 사용하여 라이다 지도를 제작하고 이에 따른 궤적을 얻을 것이다. LeGO-LOAM의 결과로 얻어진 궤적과 시뮬레이션 상에서 얻은 실제 궤적의 차이를 ICP를 이용하여 정량적인 값을 비교해 성능을 향상시킨 정도를 확인할 수 있고 또한 제작된 라이다 지도를 정성적으로 관찰하여도 성능이 어느정도 향상되었는지 확인할 수 있을 것이다. 또한, 강화 학습에서 보상을 정확도 보상만 사용하였을 경우와 본 논문에서 제시한 세 가지 보상의 가중 합을 보상으로 사용하였을 경우 두 가지를 비교하여 보상이 적절하게 설계되었는지 확인해 볼 것이다. 마지막으로, 앞에서 소개한 여행하는 외판원 문제와 실행 시간을 비교하여 실제로 계산복잡도가 향상되었는지 확인할 것이다.

4. 연구 결과

앞서 설명한 방법론을 평가하기 위한 실험 데이터로는 CARLA^[12] 시뮬레이터를 사용하여 취득한 위치 정보와 깊이 이미지를 사용하였다. 다른 변인들이 완전히 통제된 상태에서 경로만 다르게 SLAM을 실행하여 평가하기 위해 시뮬레이션 환경이 더 적합하다고 판단하였기 때문에 시뮬레이션 데이터를 사용하였다. 가장 먼저, 시뮬레이션 데이터를 사용하여 그래프 지도를 제작하고 강화 학습을 통해 경로를 얻어냈다. 또



[Fig. 5] Experimental environment in CARLA simulator



[Fig. 6] (a): Obtained depth image from CARLA simulator. (b): Obtained LiDAR measurement from CARLA simulator

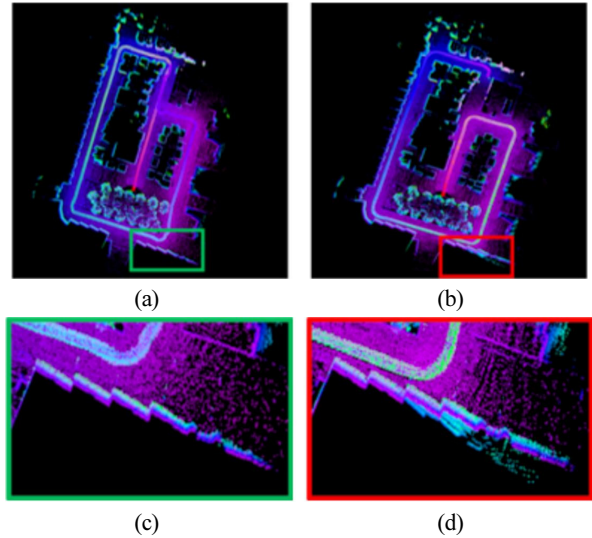
한, 그래프 지도 상에서 임의로 생성한 경로도 얻어낼 수 있다. 그 후, 두 가지 경로를 통해 CARLA 시뮬레이터 상에서 SLAM을 실행하였다. 위에서 설명한대로, 정성적으로는 제작된 라이다 지도의 품질을 비교하였고, 정량적으로는 ICP (Iterative Closest Point)를 통해 얻어낸 궤적의 오차, 실행시간, 그리고 다른 보상을 사용해 얻어낸 궤적의 오차를 비교하여 SLAM의 결과를 향상시켰음을 확인할 수 있었다.

4.1 CARLA 시뮬레이터 실험 환경

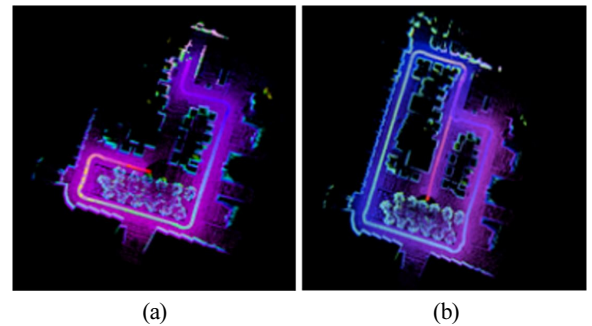
실험은 CARLA 시뮬레이터 상에서 진행되었으며 CARLA 시뮬레이터의 모습은 [Fig. 5]에서 확인할 수 있다. 사용한 센서는 라이다 센서와 깊이 카메라이며, 별도로 위치정보를 사용하였다. 라이다 센서의 경우, 100 m의 센서 범위, 각각 10도, 30도의 위 아래 시야 각을 가지는 64채널 라이다 센서를 사용하였고, 깊이 카메라의 경우, 800 × 600 이미지를 촬영하는 깊이 카메라를 사용하였다. 취득한 깊이 이미지와 라이다 센서 데이터의 예시는 [Fig. 6]과 같다.

4.2 라이다 지도 제작 및 정성적 평가

강화 학습을 통해 얻은 경로와 임의로 생성된 경로를 따라 LeGO-LOAM을 실행해 얻은 SLAM의 결과는 [Fig. 7(a-b)]에 보이는 것과 같다. 앞서 설명한 대로, 다른 경로를 통해 SLAM을 실행했을 때, 다른 결과를 얻는 것을 확인할 수 있다. [Fig. 7(c-d)]는 생성된 지도를 확대한 사진이다. 제안된 경로로 실



[Fig. 7] (a): Result of LeGO-LOAM using path from proposed method. (b): Result of LeGO-LOAM using randomly generated path. (c): Enlarged view of (a). (d): Enlarged view of (b)



[Fig. 8] (a): Result of LeGO-LOAM using accuracy reward only. (b): Result of LeGO-LOAM using proposed weighted sum reward

행한 결과에서는 벽이 하나로 잘 매칭되어 지도가 제작된 것을 볼 수 있지만, 임의로 생성된 경로 실행한 결과에서는 제대로 매칭이 되지않고 번지는 것을 볼 수 있다. 따라서, 정성적으로 평가했을 때는 제안된 방법을 통해 얻은 경로가 더 나은 결과를 얻은 것을 확인할 수 있다. 또한, [Fig. 8]에서 정확도 보상만을 사용해 얻은 경로와 제안된 방법을 통해 얻은 경로를 통해 실행한 LeGO-LOAM의 결과를 볼 수 있는데, 제안된 방법을 통해 얻은 결과가 더 많은 지역을 포함하고 있는 것을 확인할 수 있다.

4.3 정량적 평가

정량적 평가는 총 3가지 기준에 따라 실시되었다. 가장 먼저, [Table 1]에서 임의로 생성한 경로로 얻은 결과와 제안된 방법을 통해 얻은 경로로 얻은 결과를 비교하였다. 결과를 비

[Table 1] Trajectory Error by Path Planning Method

Path	Proposed	Random
Average Trajectory Error (m)	0.4006	0.5745

[Table 2] Trajectory Error by Reward

Reward	Proposed	Accuracy
Average Trajectory Error (m)	0.4006	0.1645

[Table 3] Executing Time by Algorithm

Algorithm	# Nodes	Elapsed Time (s)
Proposed	50	19.28
	100	31.25
	500	163.61
Traveling Salesman Problem	50	3.28
	100	15.58
	500	> 4000

교해보면 임의의 경로에서는 평균 제적 오차가 0.5745 m 임에 비해 제안된 방법으로 얻은 경로에서는 평균 제적 오차가 0.4006 m로 약 30% 감소된 오차를 보임을 확인할 수 있다. [Table 2]에서는 강화 학습에서 각각 다른 보상을 사용하여 얻은 경로를 통해 얻은 결과를 비교하였다. 정확도 보상만을 이용한 결과의 경우, 0.1645 m의 평균 제적 오차를 보이는데 이는 앞에서 언급한 0.4006 m에 비해 59% 정도 감소된 오차이다. 정확도 보상만을 이용하면 평균 제적 오차는 상당히 감소하지만 정성적 평가에서 확인한 것과 같이 범위가 약 반 정도만 포함되는 것을 확인할 수 있다. 마지막으로, [Table 3]에서 기존에 연구되었던 $O(2^N \times N^2)$ 의 계산복잡도를 가지는 여행하는 외판원 문제와 실행 시간을 비교하였다. 그래프 지도의 노드의 개수가 100개까지는 여행하는 외판원 문제가 더 빠른 것을 확인할 수 있지만 세 제곱의 계산복잡도를 가지기 때문에 노드가 500개일 때는 4000초의 시간에도 결과를 계산하지 못했다. 이에 반해, 제안된 강화 학습 방법은 $O(N)$ 의 계산복잡도를 가지기 때문에 163.61초, 약 3분 이내에 결과를 계산하는 것을 확인할 수 있었다. 따라서, 그래프 지도의 크기가 커질수록 제안된 방법이 효과적인 것을 확인할 수 있다.

5. 결 론

본 논문에서는 강화 학습과 그래프 구조를 이용하여 SLAM의 불확실성을 최소화하는 경로 계획법에 대한 방법론을 제시하였다. 앞서 언급한 논점인 그래프 구조로 도시 환경 표현, 강화 학습 기법을 사용한 계산복잡도 감소를 정성적, 정략적으

로 확인했다. 범위 보상은 SLAM의 범위가 향상되도록 작용하고 정확도와 불확실성 보상은 SLAM의 제적 오차가 낮아지도록 작용하도록 보상을 설계하였다. 그 결과, 임의의 경로에 비해 더 작은 오차, 즉 작은 불확실성을 보임을 확인할 수 있었고, 정확도만을 고려한 결과에 비해 더 넓은 범위를 포함하는 것을 확인할 수 있었다. 또한, 기존 알고리즘에 비해 낮은 $O(N)$ 의 계산복잡도를 가지는 것을 확인할 수 있었다.

결과에서 뛰어난 성능을 보였음에도 두 가지의 문제점을 확인할 수 있었다. 첫번째로, [Table 2]에서 확인할 수 있듯이, 정확도 보상만 사용하여 학습한 것 보다 넓은 범위를 포함하지만 정확도만을 비교하였을 때는 오차가 더 큰 것을 확인할 수 있었다. 이러한 점을 개선하기 위해서는 더 정교한 보상의 설계가 필요할 것으로 보인다. 또한, 실제 환경이 아닌 시뮬레이션에서만 실험이 이루어진 것도 발전시켜야 할 점으로 보인다. 이를 위해서는 다른 조건을 통제할 수 있는 실제 실험 환경의 구성에 대해 고민해야 할 것으로 보인다.

추후에 진행될 연구로는, 실제 도로 환경의 데이터를 사용하여 그래프 지도를 제작하는 것과, 이를 이용하여 SLAM을 수행하고 결과를 비교하는 것이 필요할 것이다.

References

- [1] C. Leung, S. Huang, and G. Dissanayake, "Active SLAM using Model Predictive Control and Attractor based Exploration," *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Beijing, China, 2006, DOI: 10.1109/IROS.2006.282530.
- [2] E. W. Dijkstra, "A note on two problems in connexion with graphs," *Numerische Mathematik*, vol. 1, pp. 269-271, 1959, DOI: 10.1007/BF01386390.
- [3] P. E. Hart, N. J. Nilsson, and B. Raphael, "A Formal Basis for the Heuristic Determination of Minimum Cost Paths," *IEEE Transactions on Systems Science and Cybernetics*, vol. 4, no. 2, pp. 100-107, Jul., 1968, DOI: 10.1109/TSSC.1968.300136.
- [4] S. Singh, A. Barto, R. Grupen, and C. Connolly, "Robust Reinforcement Learning in Motion Planning," *Advances in Neural Information Processing Systems 6 (NIPS 1993)*, 1993, [Online], <https://proceedings.neurips.cc/paper/1993/hash/3d8e28caf901313a554cebc7d32e67e5-Abstract.html>.
- [5] A. W. Moore and C. G. Atkeson, "The parti-game algorithm for variable resolution reinforcement learning in multidimensional state-spaces," *Machine Learning*, vol. 21, pp. 199-233, 1995, DOI: 10.1007/bf00993591.
- [6] J. J. Park, J. H. Kim, and J. B. Song, "Path planning for a robot manipulator based on probabilistic roadmap and reinforcement learning," *International Journal of Control, Automation, and Systems*, vol. 5, no. 6, pp. 674-680, 2007, [Online], <https://koreauiniv.pure.elsevier.com/en/publications/path-planning-for-a-robot-manipulator-based-on-probabilistic-road>.

- [7] B. Zhang, Z. Mao, W. Liu, and J. Liu, "Geometric Reinforcement Learning for Path Planning of UAVs," *Journal of Intelligent & Robotic Systems*, vol. 77, no. 2, pp. 391-409, 2015, DOI: 10.1007/s10846-013-9901-z.
- [8] T. Kollar and N. Roy, "Trajectory Optimization using Reinforcement Learning for Map Exploration," *The International Journal of Robotics Research*, vol. 27, no. 2, pp. 175-196, 2008.
- [9] L. Murphy and P. Newman, "Risky Planning on Probabilistic Costmaps for Path Planning in Outdoor Environments," *IEEE Transactions on Robotics*, vol. 29, no. 2, pp. 445-457, 2012, DOI: 10.1109/TRO.2012.2227216.
- [10] L. Nardi and C. Stachniss, "Actively Improving Robot Navigation On Different Terrains Using Gaussian Process Mixture Models," *2019 International Conference on Robotics and Automation (ICRA)*, Montreal, QC, Canada, 2019, DOI: 10.1109/ICRA.2019.8794079.
- [11] T. Shan and B. Englot, "LeGO-LOAM: Lightweight and Ground-Optimized Lidar Odometry and Mapping on Variable Terrain," *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Madrid, Spain, 2018, DOI: 10.1109/IROS.2018.8594299.
- [12] A. Dosovitskiy, G. Ros, F. Codevilla, A. Lopez, and V. Koltun, "CARLA: An Open Urban Driving Simulator," *1st Annual Conference on Robot Learning*, PMLR, vol. 78, pp. 1-16, 2017, [Online], <http://proceedings.mlr.press/v78/dosovitskiy17a.html>.



조 영 훈

2018 KAIST 건설 및 환경공학과(학사)
 2020 KAIST 건설 및 환경공학과(석사)
 2020~현재 KAIST 건설및 환경공학과 박사과정

관심분야: SLAM, 경로 계획, 강화 학습



김 아 영

2005 서울대학교 기계항공공학부(공학사)
 2007 서울대학교 기계항공공학전공(공학석사)
 2012 미시간대학교 기계공학전공(공학박사)
 2014~현재 한국과학기술원 건설및환경공학과 부교수

관심분야: 영상기반 SLAM