

강화학습 기반의 지역 경로 탐색 및 장애물 회피 시스템

이세훈*, 염대훈^o, 김풍일**

^o인하공업전문대학 컴퓨터시스템과

**한국콘베어공업(주) 기술연구소

e-mail : seihoon@inhac.ac.kr*, ydh950901@naver.com^o, pikim0916@naver.com**

Local Path Planning and Obstacle Avoidance System based on Reinforcement Learning

Se-Hoon Lee*, Dae-Hoon Yeom^o, Pung-Il Kim**

^oDept. of Computer Systems & Engineering, Inha Technical College

**Research institute of Technology, Korea Conveyor Ind.Co.,LTD.

● 요약 ●

WCS에서 AGV의 스케줄링과 동적, 정적 장애물 인식 및 충돌 회피문제는 오래전부터 다뤄져 온 중요한 문제이다. 본 논문에서는 위의 문제를 해결하기 위해 Lidar 센서를 중심으로 다양한 데이터를 기반으로 한 강화학습 시스템을 제안한다. 제안하는 시스템은 기존의 명시적인 알고리즘에 비해 다양하고 유동적인 환경에서 경로 계획과 동적·정적 장애물을 인식하고 안정적으로 회피하는 것을 확인하였으며 산업 현장에 도입 가능성을 확인하였다. 또한 강화학습의 적용 범위, 적용 방안과 한계에 대해서 시사한다.

키워드: 강화학습(Reinforcement Learning), 장애물 회피(Obstacle avoidance), 경로 계획법(Path Planning)

I. Introduction

산업 혹은 물류 현장을 실시간으로 제어 및 운영하는 솔루션인 WCS(Warehouse Control System)에서 AGV(Automated Guided Vehicle)는 매우 중요한 요소이다. 이와 같은 AGV를 운용함에 있어 스케줄링과 동적, 정적 장애물 인식 및 충돌 회피문제는 오래전부터 활발한 연구가 이루어져 온 만큼 중요한 문제이다. 기존의 A*, Dijkstra, DWA, Potential Field와 같은 AGV 주행 관련 알고리즘과 기술들이 있지만 개별적으로 구현되어 있고 각각의 기술들이 유기적으로 연관되어 있지 않은 경우에 효율성을 저하시킬 수 있는 만큼 계층화 및 구조화 또한 중요한 문제이다. 따라서 본 논문에서는 강화학습을 통해 위와 같은 문제들의 해결방안을 제시 및 가능성을 확인한다.

다. AGV의 가중치 계산은 목표물과의 거리, Map weight를 더한 숫자를 이용하였다. 현재 AGV 위치 주변 8곳의 가중치를 학습의 보상으로 사용한다. state는 AGV의 360도 Lidar 데이터, 목표물까지의 거리, 각도를 input data로 학습을 진행하게 된다.

II. Design of the System

1. System Architecture

Fig 1은 강화학습 기반의 장애물 회피 시스템의 구조도이다. AGV는 Gazebo 시뮬레이터에서 동작하고, 강화학습의 State는 AGV의 Lidar 데이터를 이용하였다. 학습에서의 보상은 장애물의 가중치가 추가된 목표물과의 거리로 step마다 업데이트한다. Mapping을 통해 획득한 Map data를 장애물, 비장애물 지역으로 가중치를 주어 저장한

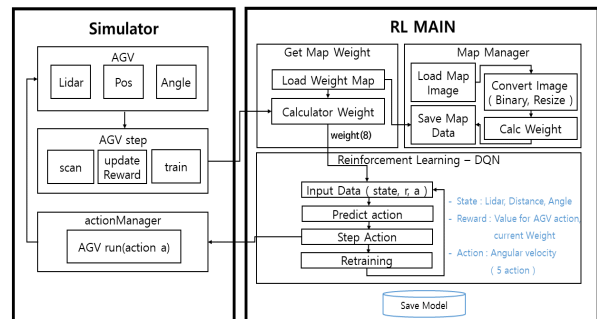


Fig. 1. System Architecture

Fig 2는 강화학습 장애물 회피 시스템의 흐름도이다. 학습 모델과 맵을 관리할 MODEL 부와 학습모델과 주행모델이 포함 된 TRAIN MODULE 부와 가상환경 부로 나누어져 있다. 학습이 시작되면 맵을 로드하게 된다. 가상환경에서의 상태(state, reward, action)을

불러와 학습을 진행하게 된다. 가중치를 얻기 위해 map을 숫자처리, resize를 진행하고 그를 csv파일로 저장한다. 가중치는 현재 AGV와 목표물까지의 거리, 주변 장애물의 가중치를 더하여 계산한다. 스텝마다 AGV에서 획득한 Lidar값을 input으로 하여 학습을 진행하고, AGV가 action을 수행하며 계산된 weight를 reward로 하여 학습을 반복 진행하게 된다.

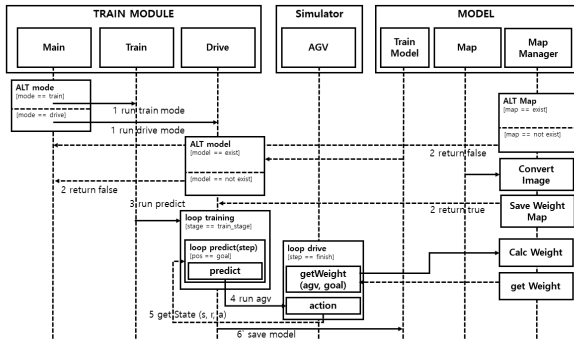


Fig. 2. Sequence Diagram

2. Experiments

주변 장애물의 가중치를 포함한 weight를 획득하는 방법은 다음과 같다.

1. 시뮬레이터상의 좌표와 csv파일의 좌표 매핑
 $((2.5 - x) * 10) * 1.3$
2. heuristic 값 계산
 $(\text{현재}x - \text{목적지}x) + (\text{현재}y - \text{목적지}y) + \text{weight}$
 * weight : 이동 가능 구역 0 / 장애물 구역 20
3. 현재 좌표를 제외한 근접 좌표(8개)의 heuristic 값 반환

본 시스템의 시나리오는 다음과 같다.

- agent : Single AGV
- goal : 10 nodes (Random)
- action : 5 actions (Angular Velocity)
- reward : Collision detection (reward = - 500)
 Visit Goal (reward = 1000)
 Step (reward = - Average(weight) / 10)

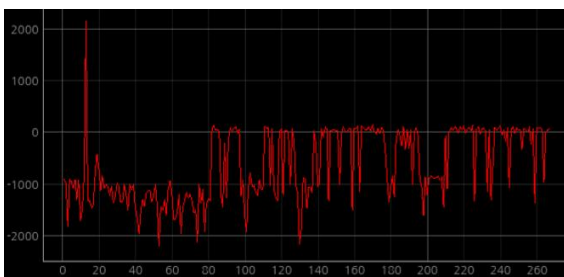


Fig. 3. Experiments Result

실험 결과 에피소드가 거듭 될수록 장애물 회피 성능의 향상을 확인 할 수 있었다. 학습 초기에는 정적 장애물에 많이 부딪히며 에피소드가 빨리 끝나는 경향을 보였다. 점차적으로 학습을 진행할수록 동적 장애물 근처로 접근하지 않으려는 모습을 보이고 목적지 근처로 가는 것을 확인할 수 있었다. 또한 충돌로 에피소드가 종료되는 횟수가 감소하고 오히려 목적지 도착으로 에피소드가 종료되는 경향을 보였다.

실험 결과 목적지로 도착하며 획득하는 보상의 증가를 확인 하였고 이는 곧 제안하는 시스템의 성능을 의미한다.

III. Conclusions

본 논문에서는 동적/정적 장애물을 인식하고 회피하기 위한 강화학습 기반의 시스템을 제안하였다. 현재 AGV 위치에서의 장애물까지의 Lidar 데이터를 state로 사용하였고 충돌, 도착시에 각각의 보상과 step마다 계산한 heuristic 값을 보상으로 부여하였다. 장애물을 포함한 보상 부여를 통해 좀 더 장애물 회피에 효율적이었고, 실험 결과 보상의 증가 또한 확인할 수 있었다. 이를 통해 강화학습 기반의 유동적인 장애물을 인식, 회피가 가능할 것으로 사료한다.

REFERENCES

- [1] Alex Kendall Jeffrey Hawke David Janz Przemyslaw Mazur Daniele Reda John-Mark Allen Vinh-Dieu Lam Alex Bewley Amar Shah, Learning to Drive in a Day
- [2] Ki Chang Choi, Moon Jong Kim, Ji Hoon Yang, Byong Hwa Oh. (2014). Artificial Intelligence : Local Path Generation Method for Unmanned Autonomous Vehicles Using Reinforcement Learning.