5U-05

Path Finding of Autonomous Mobile Robot by Reinforcement Learning

Hafiyanda Razan† Shohei Kato†‡

Department of Computer Science and Engineering, Graduate School of Engineering, Nagoya Institute of Technology†

Frontier Research Institute for Information Science, Nagoya Institute of Technology‡

1 Introduction

In recent years, automated warehouse establishment is increasing rapidly [1]. In the other side, conventional warehouses have not distinguished yet. There is a large number of employees who is needed to maintain and run a warehouse. It is not easy to cut the cost of running a warehouse. One of the way is cutting the number of employee needed. Because of that, the number of tasks which have to done by a person is getting bigger. This research is going to cut some tasks by supporting an employee at warehouse environment in a special case.

There is a job which always need a work cart nearby. Therefore, an employee has to pull the cart around the warehouse. If the cart autonomously follow the employee, this will simplify the task and reduce errors of each task. We are developing an autonomous mobile robot as a work cart for supporting in the conventional warehouse environment. This research will be applied to a 4WD omni wheel mobile robot as shown in Fig. 1 which will be used as a human support in warehouse environment. As an autonomous robot, the robot uses two *LiDAR* sensors as inputs to recognize the condition of the environment and four motor speeds as outputs to move the robot.

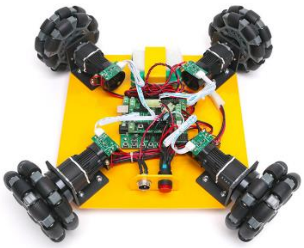

Fig. 1: 4WD Omni Wheel Mobile Robot

2 Objective

The mobile robot have to avoid any obstacles while finding the shortest path to a goal point which is an employee in this case. Conventional warehouse environment still has many employees inside that move freely and stocks which stay still. They are dynamic obstacles and static obstacles respectively. There are some times that the robot lose the employee who must have been followed. That time, the robot has to know where to go to catch up with that employee as fast as it can while it has to overcome the obstacles before reaching the employee.

3 Learning Algorithm

The conventional warehouse environment some changing physical factors as obstacles as mentioned before which change over time. In this situation a learning method which is able to decide the best move at the moment in the condition that it has not learned yet based on the learned experiences is needed. Reinforcement learning is a learning method which can achieve that specification. One of the basics of reinforcement learning is *Q-learning* [2].

$$Q(s_t, a_t) \leftarrow (1-\alpha)Q(s_t, a_t) + \alpha(r_t + \gamma \max_a Q(s_{t+1}, a))$$

4 Experiment

4.1 Environment

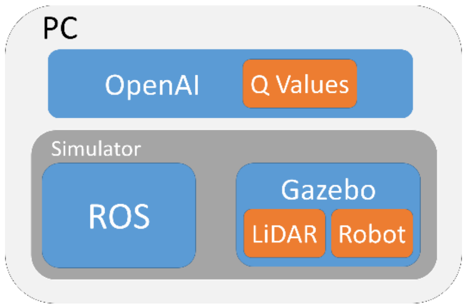

Fig. 2: Experiment Environment

Fig. 2 shows the environment of the experiment. The experiment uses Robot Operating System (ROS) as control system for the robot model. Gazebo, as a 3D simulator, simulates a 45 [cm] diameter of mobile robot with a 270 degree *LiDAR* sensor. *OpenAI* supports the learning calculation and produces Q values to be applied in the real wheeled robot.

4.2 Simulation Conditions

4.2.1 Maps

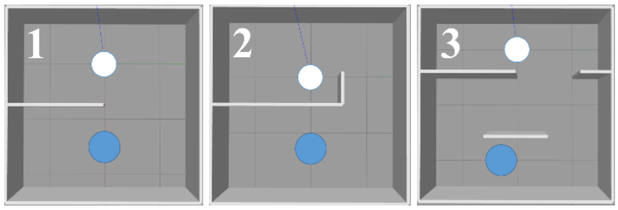

Fig. 3: Simulation Maps

Fig. 3 shows three maps which are used in the experiment. The white and the blue circles in each map shows the start and the goal points, respectively, of the simulation. Each of the maps is 3 [m] squared which the goal points have 25 [cm] radius.

The learning parameters α and γ in all of the conditions are 0.2 and 0.8 respectively.

4.2.2 States

The states for the *Q-learning* are nine of the average range data of the *LiDAR* sensor in meter. The data for the states are set to have an increment of 0.25 [m] with range from 0 to 4.25 [m].

$$Data = \{0.0, 0.25, 0.5, \ldots, 3.75, 4.0, 4.25\}$$

4.2.3 Actions

The robot has four actions: front, back, right, and left. The speed is set to 50 [cm/s]. In one action the robot moves for 0.5 ± 0.01 [s] of simulator time.

4.2.4 Rewards

This experiment has four conditions of rewards in the learning. The first one, if the robot model hit the wall, it has -(10^6) reward points. The second one, if the robot model is in the same distance as before or getting further from the goal point, it has -10 reward points. The third one, if the robot model is getting closer to the goal point, it has -1 reward points. The last one, if the robot model reaches the goal, it has 0 reward points.

5 Results

Table 1: Experiments results of three times trials.

Map	Total Learning Episodes	Simulation speed (x times)	Total succeed episodes	Average succeed actions	Fewest actions
1	1000	50	590	16	9
2			15.67	82	42.67
3			5.33	41.33	34

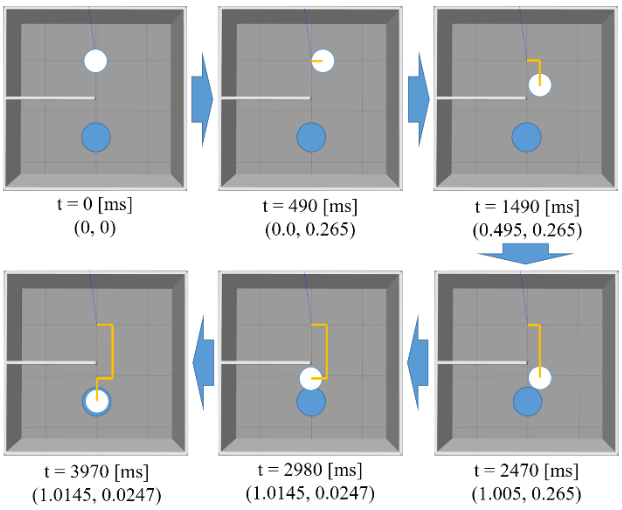

Fig. 4: One of the fewest actions in map 1.

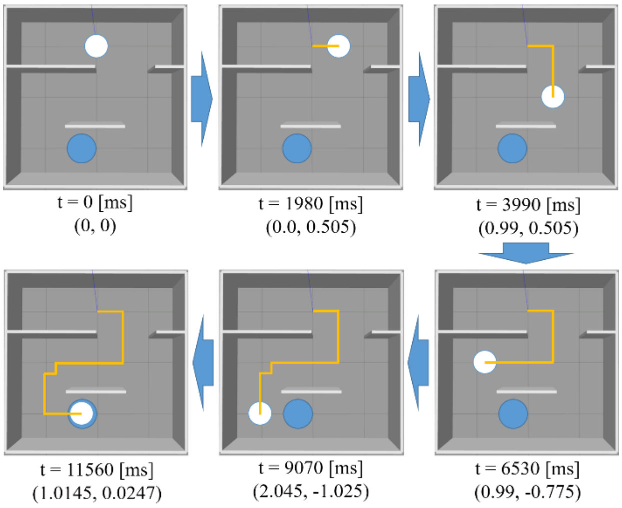

Fig. 5: One of the fewest actions in map 3.

Table. 1 shows the results of the average of three times trials of the experiment in each map and the average of total succeed episodes in each experiment per condition.

Fig. 4 and Fig. 5 show one of the fewest action over all of the experiment in map 1 and map 3 respectively. From these results we know that just by the *Q-learning*, which is one of the basic of the reinforcement learning, a mobile robot is able to find a way to goal while avoiding static obstacles.

6 Future Works

Currently, the movement of the robot model are just four movements which the real robot, omni wheel mobile robot, is able to move freely by controlling the x vector and y vector of its velocity, and its angular velocity. We plan to apply that movement control method in the learning actions by applying discrete speed vector or changing the actions to continuous actions. Then, we plan to apply the other algorithms, such as *deep q-learning*, *double q-learning*, *double deep q-learning*, *actor-critic* and compare the performance between them. From there, we plan to find the shortest path to the goal for all conditions with just static obstacles. And, to be more realistic, dynamic obstacles will be included in the simulation.

References

[1] M. Trottman and S. Zang, *White Paper: The Trend Towrds Warehouse Automation*. Reading, Westernacher Consulting AG, 2017. [E-book] Available: https://westernacher-consulting.com/wp-content/uploads/2017/11/Whitepaper_Trend_to_Automation_FINAL_s.pdf.

[2] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction, second edition*. Reading, MIT Press, Cambridge, MA, 2018. [E-book] Available: http://incompleteideas.net/book/RLbook2018.pdf.