

외란에 강인한 심층 강화학습 기반 로봇 주행

Robust Robot Navigation against External Disturbance using Deep Reinforcement Learning

유형열¹ · 윤민성¹ · 박대형² · 윤성의[†]

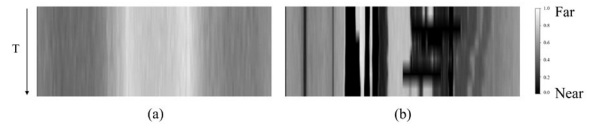
Hyeongyeol Ryu¹, Minsung Yoon¹, Daehyung Park², Sung-Eui Yoon[†]

Abstract: With recent advances of deep reinforcement learning (DRL) in complex robotic navigation, mobile robots have worked in diverse places. Nonetheless, as the robots operate in a wider set of environments including cluttered or messed up environments (e.g., cleanup robots on a dusty construction site), they can get exposed to external disturbances such as dust or stains. These could hinder the normal operation or even cause catastrophic behaviors to the robots. To deal with this problem, we propose a navigation method robust to disturbances using deep reinforcement learning that efficiently follows the path to a goal and secures safety by sensing the areas where some parts of the sensor cannot observe well due to the disturbances. Our approach utilizes *Confidence map*, which constructs the robot's local regions requiring more observation to avoid collision and identifies where to sense next via entropy. We empirically demonstrate the influences of the disturbances on the sensor and compare the existing DRL method with our proposed method under the disturbances. The result shows that our method achieves higher performances over the existing method in proposed situations.

Keywords: Deep reinforcement learning, External disturbance, Robot navigation

1. 서론

자율주행은 모바일 로봇의 핵심 기술 중 하나로, 충돌을 회피하며 목적지까지 도달하는 기술이다. 특히 모바일 로봇 주행은 평면상의 기하학적인 구조를 바탕으로 수행된다는 점에서 2D-LiDAR 센서가 많이 사용되는 추세다[1]. 최근, 복잡한 상황에서 유연하고 강인한 행동을 학습할 수 있는 특성으로 인해 심층 강화학습이 널리 활용되고 있다[2]. 하지만 로봇의 활동 범위가 점차 넓어지고 다양해짐에 따라 로봇의 센서 역시 더 많은 외란에 노출될 우려가 있다. 일례로, 먼지 혹은 장기간 사용으로 인한 얼룩이 센서의 정상적인 작동을 방해할 경우, 로봇이 예기치 않은 행동을 할 수 있다. 본 논문에서는 이러한 외란이 2D-LiDAR 센서 및 주행에 미치는 영향을 실험적으로 관측하고, 외란에 의해 올바르게



[Fig. 1] Visualization of the sensor observation during the time T. (a) Normal sensor data and (b) disturbed sensor data. The color shows the distance between the sensor and obstacles. X-axis indicates the field of view of the sensor: 240° in our experiment.

관측되지 못해 충돌 위험이 있는 영역을 온전한 센서로 감지해 회피할 수 있도록 하는 *Confidence map*을 제안한다. 나아가, 이를 강화학습에 적용해 효율적이면서도 안전한 로봇 주행 방식을 학습하고자 한다.

2. 본론

2.1 외란이 2D-LiDAR 센서에 미치는 영향

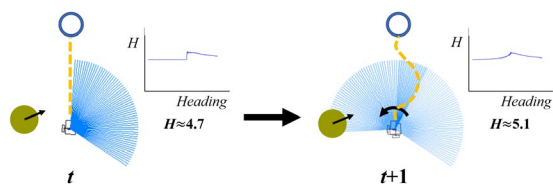
현실에서 발생 가능한 문제 상황(예, 센서에 오수가 묻는 상황)을 실제 센서에 재현하고 시간에 따른 관측 데이터를 시각화하였다. 실험 결과, 외란이 발생할 경우 센서가 주위 환경의 기하학적인 구조를 제대로 반영하지 못했고 [Fig. 1], 센서의 특정 영역이 외란에

※ This work was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (IITP-2015-0-00199).

1. Master Candidate, KAIST, Daejeon, Korea (hy.ryu@kaist.ac.kr, ymsalstjd@kaist.ac.kr)

2. Assistant Professor, School of Computing, KAIST, Daejeon, Korea (daehyung@kaist.ac.kr)

† Professor, Corresponding author: School of Computing, KAIST, Daejeon, Korea (sungeul@kaist.edu)



[Fig. 2] Modified trajectory to a goal after observing the dynamic obstacle: some parts of the sensor data (left) are unavailable due to the disturbances. *Confidence score* to get higher entropy (H) leads the robot to sense the areas requiring more observation, contributing to the collision avoidance.

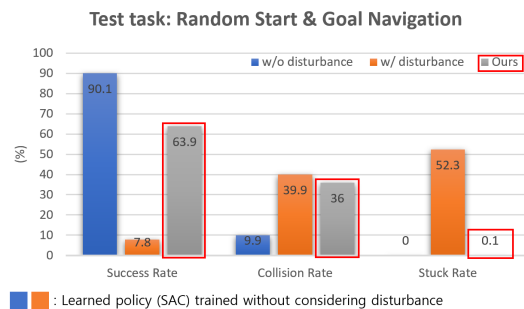
의해 지속해서 가려지는 경우, 관측된 데이터 값은 센서의 noise를 포함한 규격상의 최소 거리보다 작았다.

2.2 제안 방법

2.1 실험을 통해 지속해서 가려진 센서 영역의 값은 센서의 noise를 포함한 규격상의 최소 거리보다 작음을 확인하였다. 이 값을 임계값으로 사용하면 주행에 필요한 일부 온전한 센서 데이터를 확보할 수 있다. 이때, 충돌회피를 위해 온전한 센서로 감지 영역을 단순히 최대로 확보하는 것(로봇이 끊임없이 회전하는 것)은 매우 비효율적일 뿐만 아니라 안정성 문제도 초래할 수 있다. 따라서 주변 장애물을 적절히 감지하며 효율적으로 목적지까지 도달하는 행동이 필요하다. 본 논문에서는 강화학습을 통해 안전성과 효율성을 모두 만족하는 행동을 학습할 수 있는 *Confidence map*을 제안한다. *Confidence map* c 는 온전한 센서를 통해 로봇의 지역 정보를 T 시간 동안 기록하여 2차원의 circular 맵을 구성한다 (식 (1) 참고). γ 는 최근 관측된 영역에 더 많은 가중치를 부여하는 상수). 같은 방식으로 거리에 따른 가중치를 계산하면 로봇이 바라보는 방향에 따른 1차원의 *Confidence score*를 구할 수 있다. 이때 *Confidence score*

$$c_{x,y}^t = \frac{1}{T} \sum_{k=1}^T \gamma^{k-1} * c_{x,y}^{t-k} \quad (1)$$

그래프를 엔트로피 관점에서 해석할 경우, 엔트로피 값이 커짐에 따라 관측한 센싱 영역은 더 커지게 된다 ([Fig. 2] 참고). 결과적으로 T 시간 동안 최대의 센싱 영역을 확보할 수 있도록 엔트로피를 최대화할 경우 주변 정보를 더 잘 파악하여 충돌을 회피할 가능성을 높일 수 있다. 학습을 위한 목적함수는 Soft Actor-critic (SAC)을 토대로 구성하였고, 동시에 *Confidence score*의 엔트로피 값을 최대화할 수 있도록 하는 항을 추가하였다. 최종적으로, 강화학습을 통해 제안한 방식을 학습하여 효율적이면서도 안전한 주행을 수행할 수 있도록 한다.



[Fig. 3] Results of the experiments. While the performances without considering disturbances significantly decrease when disturbances occur, our proposed method achieves reasonable results even in that case. Results w/o disturbance (blue) can be considered as upper bounds that ours can achieve. Note that stuck rate means the robot cannot reach the goal within the time limit.

2.3 외란 상황에서의 주행 실험 및 결과

실험을 위해 무작위로 출발, 목적지를 구성한 후, 기존 DRL 방식은 외란이 없는 환경에서 SAC로 학습하였다. 이후, 테스트 시 외란을 발생시켜 성능의 변화를 측정하였다. 1,000회 시행 결과 ([Fig. 3]), 외란이 고려되지 않은 로봇은 외란이 발생했을 때 성능이 큰 폭으로 하락하였으나, 제안한 방법은 외란이 발생하였을 경우 온전한 센서로 센싱 영역을 넓혀 장애물과의 충돌을 회피하며 ([Fig. 2] 참고) 목적지까지 도달하는 모습을 보였다.

3. 결론 및 향후 연구

본 논문에서는 외란이 센서 및 로봇 주행에 미치는 영향을 확인하고, 이를 효율적이고 안전하게 해결하는 *confidence map* 방식을 제안하였다. 제안한 방식은 외란이 발생한 상황에서도 외란을 고려하지 않은 기존 방식 대비 더 우수한 성능을 보였다. 향후 연구를 통해 성능을 개선하여 추후 실제 로봇 주행에서도 외란에 강인한 행동을 수행하도록 할 것이다.

4. 참고문헌

- [1] J. Lundell, F. Verdoja and V. Kyrki, "Hallucinating Robots: Inferring Obstacle Distances from Partial Laser Measurements," *IEEE/RSJ International Conference on Intelligent Robots and Systems*, Madrid, Spain, pp. 4781-4787, 2018.
- [2] F. Leiva and J. Ruiz-del-Solar, "Robust RL-Based Map-Less Local Planning: Using 2D Point Clouds as Observations," *IEEE Robotics and Automation Letters*, vol. 5, no. 4, pp. 5787-5794, Oct. 2020.