

V2X 환경에서 교통 혼잡 방지를 위한 심층 강화학습 기반 자율주행

조영제*, 정현주*, 황경호^o

Autonomous Vehicle Based on Deep Reinforcement Learning to Prevent Traffic Congestion in V2X

Yeong-je Jo*, Hyeon-joo Jeong*, Gyung-Ho Hwang^o

요약

본 논문에서는 V2X 환경에서 자율주행 차량에 심층 강화학습 알고리즘인 PPO 알고리즘을 적용하고 신호등 제어를 통해 교통 혼잡을 방지하였으며 시뮬레이션을 통해 도로 속 차량들의 평균 속도를 비교 분석하였다.

Key Words : Autonomous vehicle, V2X, PPO, Deep Reinforcement Learning, Speed control

ABSTRACT

In this letter, traffic congestion was prevented using autonomous vehicles and traffic light control based on the PPO algorithm in the V2X environment. The average speed of vehicles was compared and analyzed through simulations.

I. 서론

본 논문에서는 매년 증가하는 교통혼잡비용 문제를

감소시키기 위해 교통 체증의 대표적인 원인인 차선이 점차 줄어드는 도로에서 교통 혼잡 방지를 위한 연구를 진행한다. 두 개 이상의 차선이 하나의 차선으로 합쳐지는 병합 차선 같은 경우 차들의 속도 감·가속, 차선변경 등과 같은 이유로 유령 체증, 병목현상과 같은 교통 체증의 원인이 된다. 본 논문에서는 주행 차량에 심층 강화학습 알고리즘인 PPO 알고리즘을 적용한 자율주행 기법을 제안하고 비자율 주행, 신호등 제어환경에서 시뮬레이션을 통하여 병목현상이 발생하는 도로 환경에서 주행 차량들의 주행 평균 속도를 비교 분석한다.

II. 관련 연구

2.1 V2X 통신 기술

V2X 기술은 ADAS(Advanced Driver Assistance System) 자율주행 자동차의 한계를 개선하기 위한 통신 기술로 현재 ADAS 기반의 자율주행 차량 같은 경우 차량 내, 외부에 카메라, 레이더 등의 센서를 탑재하여 센터를 통해 주변 상황을 감지하고 주변 차량, 보행자, 야생동물 등으로부터의 충돌을 방지하고 있다. 하지만 ADAS 기반의 자율주행 차량은 센서가 인식 가능한 범위 내에서의 충돌 방지, 주행 보조를 해주기 때문에 주변 현수막, 구조물 등으로 인한 시야 가림 및 폭우, 폭설, 안개등과 같은 날씨에서 센서가 주변 상황을 제대로 인식하지 못하여 잘못된 판단을 할 수 있는 위험이 있는 반면 V2X통신 기술은 V2V(Vehicle to Vehicle), V2I(Vehicle to Infrastructure), V2N(Vehicle to Nomadic Device), V2P(Vehicle to Pedestrian) 등 주변 통신기기들과의 통신을 통해 주변 상황을 판단하여 충돌 방지, 주행 보조를 하는 통신 기술이다.^[1]

2.2 강화학습 기반 자율주행

주행 차량에 강화학습 알고리즘을 적용한 강화학습 기반 자율주행 차량의 경우 센서 및 V2X 통신을 통해 얻게 된 정보를 바탕으로 최적의 행동을 선택하는 방식으로 단순 사고방지를 위한 주행뿐만 아니라 도로 속 교통 혼잡의 원인인 병목현상, Stop and go wave 등의 현상을 방지하여 교통 혼잡을 줄이기 위한 연구가 진행

※ 본 과제(결과물)는 2023년도 교육부의 재원으로 한국연구재단의 지원을 받아 수행된 지자체-대학 협력기반 지역혁신 사업의 결과입니다. (2021RIS-004)

• First Author : (ORCID:0000-0002-0663-0746) WIFIVE, yeongje3301@wifive.co.kr, 연구원, 정희원

◦ Corresponding Author : (ORCID:0000-0001-6795-8086) Dept. Computer Engineering, Hanbat National University, gabriel@hanbat.ac.kr, 정교수, 중신회원

* (ORCID:0009-0005-9611-1665) Dept. Computer Engineering, Hanbat National University, HyeonJJ@edu.hanbat.ac.kr, 학생(석사), 학생회원

논문번호 : 202307-138-B-LU, Received July 1, 2023; Revised July 7, 2023; Accepted July 7, 2023

중이다.^[2] 선행 연구에서는 신호등이 없는 도로의 단일 차선에서 같은 수의 차량을 대상으로 연구를 진행하여 차선변경으로 인한 유령 체증 현상, 신호등 제어 등을 고려하지 못하였다. 본 논문에서는 여러 차선이 존재하는 도로에서의 차량 병목현상을 감소하는 기법을 제안한다.

III. 심층 강화학습 기반 자율주행

본 논문에서는 자율주행 차량의 감·가속도 제어를 통해 차선이 점차 줄어드는 도로 환경 속에서 병목현상의 개선 정도를 시뮬레이션을 통해 확인한다. 자율주행 차량에는 심층 강화학습 알고리즘 중 하나인 PPO 알고리즘을 적용하여 차량의 감·가속도를 제어한다. PPO 알고리즘은 Actor Critic 방식의 대표적인 알고리즘으로 보상을 최대화할 수 있는 행동을 결정하는 Actor 네트워크와 상태 가치를 평가하는 Critic 네트워크로 구성되어 있다.^[3] 자율주행 차량은 time step마다 주변 차량의 정보를 관찰하고 관찰 정보를 바탕으로 최적의 행동을 취하여 병목현상을 감소한다.

V2X 환경에서 도로 속 자율주행 차량은 앞, 뒤 차량 간의 거리와 주행 속도 그리고 자신의 현재 주행 속도를 포함한 총 5개의 관찰 공간을 가지게 된다. 자율주행 차량의 행동 결정은 관찰한 정보를 가지고 병목현상을 방지할 수 있도록 도로 속 가능한 속도 범위 내에서 차량의 불필요한 속도 감·가속으로 인하여 발생하는 교통체증을 방지하기 위하여 자율주행 차량의 감·가속도를 결정하여 교통 체증을 최소화하는 차량의 주행 속도를 제어한다.

행동에 대한 보상으로는 식(3)과 같다. 본 논문에서는 ϵ_1 는 16, ϵ_2 는 34로 설정하였으며, $cost_1$ 은 자율주행 차량의 주행 속도가 주행 목표속도인 v_0 에 가까워질수록 보상을 높게 주어 도로 속 차들의 주행 속도를 높일 수 있도록 설정하였고, $cost_3$ 은 각 자율주행 차량들의 $cost_2$ 값을 합한 값으로 h 는 앞 차량과의 거리를 의미하고, T 는 앞 차량과의 허용 가능한 최소 시간으로 본 논문에서는 1초로 설정하였다. $cost_2$ 는 차량과의 간격이 짧아 사고 위험이 있을 때 보상을 낮추는 처벌 향으로 자율주행 차량이 안전하게 주행을 할 수 있도록 한다.

$$cost_1 = \frac{\max(\|v_0\| - \|v - v_0\|, 0)}{\|v_0\|} \tag{1}$$

$$cost_2 = \min\left(\frac{\max\left(\frac{h}{v}, 0\right) - T}{T}, 0\right) \tag{2}$$

$$\max(\epsilon_1 * cost_1 + \epsilon_2 * cost_3, 0) \tag{3}$$

IV. 시뮬레이션

심층 강화학습 기반 자율주행 차량으로 인하여 도로 속 병목현상이 개선되었는지를 확인하기 위하여 Flow 시뮬레이터를 이용하여 비 자율주행, 자율주행, 신호등 제어환경에서 시뮬레이션을 진행한다.^[4] 도로의 최대 주행 속도는 30m/s로 설정하였으며 차량의 가속, 감속의 최댓값은 $3m/s^2$ 으로 설정하였다. 심층 강화학습 알고리즘을 적용하지 않은 비 자율주행 차량은 Intelligent Driver Model(IDM) 방식을 사용하여 가속도를 결정하여 주행하도록 하였다.^[5]

그림 1은 시뮬레이션에서 진행한 신호등 제어를 나타내는 순서도로 N_t 는 해당 차선의 10초 동안의 평균 차량수를 의미하며 K 는 학습의 정도를 조절하는 역할을 한다. CT_t 는 Th_2 동안의 순환 시간을 의미하며 RT_t 는 현재 신호를 얼마 동안 유지했는지를 나타내게 된다. 이렇게 구해진 RT_t 값이 Th_3 값보다 낮으면 신호등을 파란불로 바꾸고 Th_3 값보다 크면 빨간불로 바뀌도록 설정하였다. 표 1은 시뮬레이션에서 사용한 파라미터 정보이다.

그림 2는 비 자율주행, 그림 3은 신호등 제어, 그림 4는 자율주행 시뮬레이션 진행을 나타낸 그림으로 도로 환경은 4차선에서 2차선, 1차선 순으로 점차 줄어가고

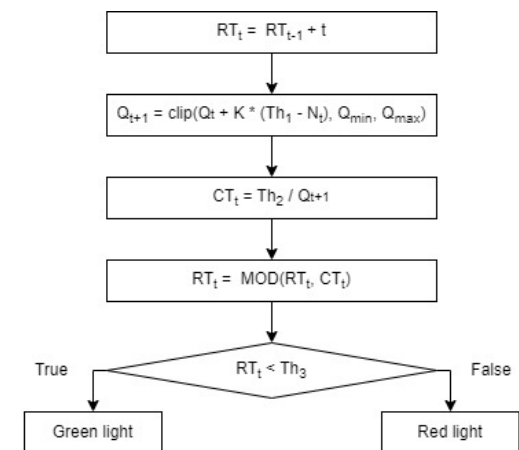


그림 1. 신호등 제어 순서도
Fig. 1. Traffic light control flowchart

표 1. 환경 변수
Table 1. Environment variables

Parameter	Value
ϵ_1	16us
ϵ_2	34us
K	20
Q_{min}	275
Q_{max}	1100
Th_1	35
Th_2	7200
Th_3	4

표 2. 시뮬레이션 결과
Table 2. Simulation Results

	non Autonomous vehicle	Autonomous vehicle	Traffic light control
Velocity mean (m/s)	9.86	11.423	10.125

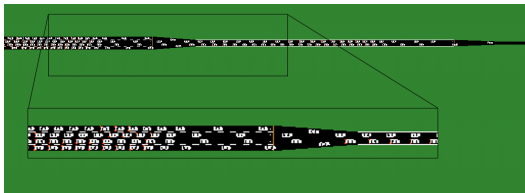


그림 2. 비 자율주행 시뮬레이션
Fig. 2. Non autonomous vehicle simulation

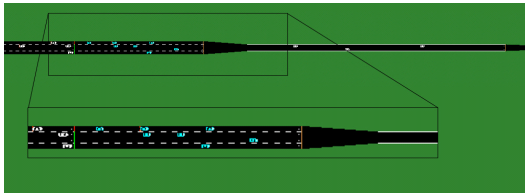


그림 3. 신호등 제어 시뮬레이션
Fig. 3. Traffic control simulation

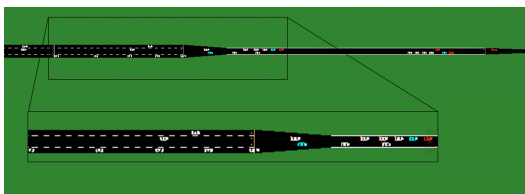


그림 4. 자율주행 시뮬레이션
Fig. 4. Autonomous vehicle simulation

비 자율주행 시뮬레이션 시에는 5초마다 비 자율주행 차량이 5대씩 진입하도록 설정하였고 자율주행 차량 시뮬레이션 시에는 5초마다 비 자율주행 차량이 4대, 자율주행 차량이 1대씩 진입하도록 설정하여 시뮬레이션을 진행하였다. 시뮬레이션 결과 총 5분간 차들의 평균 속도가 비 자율주행 시 9.86m/s, 자율주행 11.432m/s, 신호등 제어 10.125m/s로 자율주행 시 평균 속도가 비 자율주행 대비 15.94% 향상된 것을 알 수 있었다.

V. 결론

본 논문에서는 매년 증가하는 교통혼잡비용을 감소 시키기 위하여 병목현상이 자주 발생하는 도로 속에서 차량의 주행 속도를 향상하는 연구를 진행하였다. Flow 시뮬레이터를 사용해 시뮬레이션을 진행해본 결과 심층 강화학습 알고리즘인 PPO 알고리즘을 적용한 자율주행 차량이 신호등 제어 및 비 자율주행 차량의 주행 대비 평균 주행 속도가 향상된 것을 확인하였다.

References

- [1] B. C. Jo, "Trend of V2X communication technology for connected & automated vehicles," *AUTO J*, vol. 42, no. 12, pp. 23-27, Dec. 2020.
- [2] D. S. Lee and M. H. Kwon, "Combating stop-and-go wave problem at a ring road using deep reinforcement learning based autonomous vehicles," *J. KICS*, vol. 46, no. 10, pp. 1667-1682, Oct. 2021. (<https://doi.org/10.7840/kics.2021.46.10.1667>)
- [3] J. Schulman, F. Wolski, P. Dhariwal, et al., "Proximal policy optimization algorithms," *arXiv preprint arXiv:1707.06347*, 2017. (<https://doi.org/10.48550/arXiv.1707.06347>)
- [4] C. Wu, A. Kreidieh, K. Parvate, E. Vinitsky, and A. Bayen, "Flow: Architecture and benchmarking for reinforcement learning in traffic control," *arXiv preprint arXiv:1710.05465*, 2017.
- [5] M. Treiber, A. Hennecke, and D. Helbing, "Congested traffic states in empirical observations and microscopic simulations," *Phys. Rev. E*, vol. 62, pp.1805-1824, 2000. (<https://doi.org/10.1103/physreve.62.1805>)