

자율주행 차량의 고정밀 경로 추적을 위한 심층 강화학습 기반 가속도 및 조향각 공동 제어

양승건[†], 조은호[†], 김정윤[‡], 임승찬[†]

[†] 한경국립대학교, [‡] 한국과학기술원

{tmdrjs1543, jehwow, sclim}@hknu.ac.kr[†], jykim@kaist.ac.kr[‡]

Joint Acceleration and Steering Control for High-Precision Path Tracking in Autonomous Vehicles via Deep Reinforcement Learning

Seung Geon Yang[†], Eun Ho Cho[†], Jeongyun Kim,[‡] Seung-Chan Lim[†]

[†] Hankyong National University, [‡] Korea Advanced Institute of Science and Technology

요약

본 논문에서는 무인 이동체의 고정밀 자율주행을 위한 심층 강화학습 기반 경로 추적 알고리즘을 제안한다. 제안된 기법은 가속도와 조향각을 공동 액션으로 구성하고, 이를 속도 및 조향 제어 작업으로 통합하여 보다 효율적인 학습과 정책 최적화를 가능하게 한다. 시뮬레이션을 통해, 제안된 알고리즘이 기준 경로 추적 기법들에 비해 더 높은 추적 정확도와 주행 안정성을 달성함을 확인하였다.

I. 서론

차율 주행 차량의 경로 추적은 차량의 움직임을 제어하여 지정된 경로를 정확하고 안정적으로 따라 주행하도록 하는 핵심 기술로, 다양한 도로 환경에서 차율주행 차량이 신뢰성 높은 판단을 수행하기 위한 필수적인 기반이 된다 [1]. 그동안 기하학 기반 및 최적화 기반 기법들이 개발됐으나, 모델의 정확성과 매개변수 설정에 대한 민감성으로 인해 실제 적용에는 한계가 있었다. 또한, 강화학습 기반 경로 추적 기법도 제안되었지만, 대부분 조향 제어에 집중되어 가속도 제어를 포함하는 통합 제어에는 제약이 존재한다 [2].

본 연구에서는 가속도와 조향각을 동시에 제어하는 심층 Q-네트워크(deep Q-network, DQN) 기반 경로 추적 알고리즘을 제안한다. 제안된 기법은 가속 및 조향 제어 입력을 2차원 튜플 형태의 공동 액션으로 구성함으로써, 두 제어 변수를 통합적으로 처리할 수 있도록 설계되었다. 시뮬레이션을 통해, 제안된 알고리즘이 기준 기법들에 비해 더 높은 경로 추적 정확도와 주행 안정성을 달성함을 확인하였다.

II. 시스템 모델

본 논문에서는 무인 이동체가 정해진 기준 경로를 추적하는 시나리오를 고려한다. 차량의 축거를 L , 차량 중심에서 뒷바퀴 축까지의 거리를 L_r 이라고 하자. 시간 단계 k 에서 차량 중심의 위치는 $\mathbf{c}_k = [x_{c,k}, y_{c,k}]^T$ 로 정의되며, 이때 차량 중심의 진행 방향은 $\theta_{c,k}$ 로 나타낸다. 또한, 시간 단계 k 에서 앞바퀴 축의 위치는 $\mathbf{f}_k = [x_{f,k}, y_{f,k}]^T$ 로 정의되고, 진행 방향은 $\theta_{f,k}$ 로 나타낸다.

각 시간 단계 k 에서 기준 경로를 정확히 추적하기 위해, 차량의 가속도 \dot{v}_k 와 조향각 δ_k 는 경로 추적 알고리즘에 의해 제어된다. 주어진 \dot{v}_k 와 δ_k 에 따라, kinematic bicycle model을 따르는 차량의 속도 v_k , 슬립 각도 β_k , 각속도 ω_k 는 다음과 같이 각각 표현된다:

$$v_k = v_{k-1} + \dot{v}_k \Delta k, \quad (1a)$$

$$\beta_k = \tan^{-1}\left(\frac{L_r}{L} \tan(\delta_k)\right), \quad (1b)$$

$$\omega_k = \frac{v_k}{L} \tan(\delta_k) \cos(\beta_k). \quad (1c)$$

여기서 Δk 는 시간 간격을 나타내며, v_k 는 속도 범위 $[v_{\min}, v_{\max}]$ 내에 존재한다. 이후, 시간 단계 $k+1$ 에서 차량 중심 자세는 다음과 같이 갱신된다:

$$x_{c,k+1} = x_{c,k} + v_k \cos(\theta_{c,k} + \beta_k) \Delta k + x_{n,k}, \quad (2a)$$

$$y_{c,k+1} = y_{c,k} + v_k \sin(\theta_{c,k} + \beta_k) \Delta k + y_{n,k}, \quad (2b)$$

$$\theta_{c,k+1} = \theta_{c,k} + \omega_k \Delta k + \theta_{n,k}. \quad (2c)$$

여기서 $x_{n,k}$, $y_{n,k}$, $\theta_{n,k}$ 는 실제 주행 환경에서의 불확실성을 반영하기 위한 랜덤 잡음을 나타낸다.

시간 단계 k 에서, 차량 중심으로부터 기준 경로상 가장 가까운 점의 위치를 \mathbf{q}_k 라고 정의할 때, 차량 중심과 기준 경로 간의 횡방향 오차는 다음과 같이 정의된다:

$$d_k = \| \mathbf{c}_k - \mathbf{q}_k \| . \quad (3)$$

III. DQN 기반 가속도 및 조향각 공동 제어 알고리즘

본 연구에서는 DQN 기법을 활용하여, 차량이 자율 주행 환경과의 상호작용을 통해 가속도 \dot{v}_k 와 조향각 δ_k 를 공동 최적화하는 정책을 학습하는 방법을 제안한다. 이를 위하여 다음과 같이 마르코프 결정 과정(Markov decision process, MDP)으로 문제를 정의하였다.

- 상태:** 상태 s_k 는 5차원 튜플로 구성되며, 이는 앞바퀴 축 자세 $[x_{f,k}, y_{f,k}, \theta_{f,k}]$ 와 가장 가까운 기준 경로 상의 점과의 오차를 차량의 로컬 좌표계로 변환한 세 가지 오차 성분, 식 (1c)의 각속도, 그리고 식 (1a)의 속도를 최대 속도 v_{\max} 로 정규화한 값으로 이루어진다.

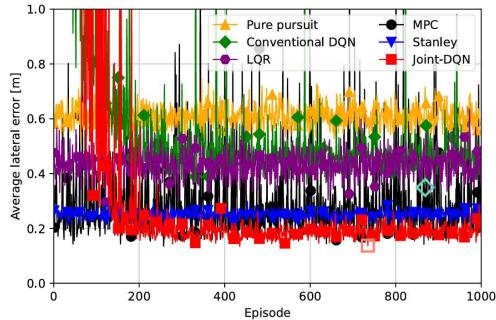


그림 1. 평균 횡방향 오차 성능 비교.

- 액션:** 액션 a_k 는 가속도 v_k , 조향각 δ_k 로 구성된 2차원 투플 (v_k, δ_k) 로 정의된다. 가속도 액션 공간은 $[v_{\min}, v_{\max}]$ 범위에서 균일한 간격으로 이산화되어 N_v 개의 값을 갖도록 설정된다. 또한, 조향각 액션 공간은 $[\delta_{\min}, \delta_{\max}]$ 범위에서 균일한 간격으로 N_δ 개의 값을 갖도록 정의된다.
- 보상:** 횡방향 오차를 최소화하면서도 준수한 속력을 유지하도록 다음과 같이 설계한다:

$$r_k = -d_{k+1} + \frac{v_{k+1}}{v_{\max}} \Delta k. \quad (4)$$

제안된 알고리즘은 학습은 ϵ -greedy 정책에 따라 현재 상태 s_k 에서 액션 a_k 를 선택하고, 환경으로부터 보상 r_k 를 획득한 후 다음 상태로 전이하는 과정을 반복한다. 수집된 경험은 리플레이 메모리에 저장되며, 미니배치 샘플을 이용해 타깃 Q-값 기반의 손실 함수를 계산하고, 이를 최소화하도록 Q-네트워크를 업데이트한다.

III. 시뮬레이션 결과

실제 도로 환경을 모사한 기준 경로에 대하여 제안된 경로 추적 알고리즘(Joint-DQN)의 성능을 시뮬레이션을 통해 평가하였다. 차량은식 (1)과 (2)의 모델에 따라 주행하며, 랜덤 잡음 $x_{n,k}$, $y_{n,k}$, $\theta_{n,k}$ 는 각각 분산 $10^{-2}, 10^{-2}, 10^{-3}$ 를 갖는 평균 0의 가우시안 확률 변수로 모델링되었다. 성능 평가 지표로는 다음과 같이 정의된 평균 횡방향 오차를 사용하였다:

$$\frac{1}{K} \sum_{k=1}^K d_k. \quad (5)$$

여기서 K 는 알고리즘의 경로 추정이 종료되는 시점의 시간 단계이다. 비교 대상 알고리즘으로는 PID 제어기를 통해 가속도를 제어하는 Pure pursuit, Stanley, Conventional DQN, 그리고 LQR, MPC를 포함하였다.

그림 1은 제안된 Joint-DQN의 학습 과정에 따른 평균 횡방향 오차의 변화를 보여준다. Pure pursuit은 약 0.6m, Conventional DQN은 약 0.48m, LQR은 약 0.43m의 평균 횡방향 오차를 기록하였으며, 이는 일반적인 도로 폭을 고려할 때 안전한 주행 성능을 확보하기에 미흡한 수준이다. MPC의 평균 횡방향 오차는 낮은 편이나, 랜덤 잡음에 매우 민감하게 반영하여 주행 안정성이 저하되는 경향을 보였다. Stanley는 약 0.25m의 오차로 비교적 우수한 추적 성능을 나타냈으나, 조향 제어만을 대상으로 하므로 가속도와 조향을 통합적으로 제어하는 고정밀 주행에는 한계가 있다. 반면, 제안된 Joint-DQN은 약 0.19m 수준으로 수렴하여, 기존 기법들에 비해 더욱 정밀한 경로 추적과 높은 제어 안정성을 동시에 달성하였다.

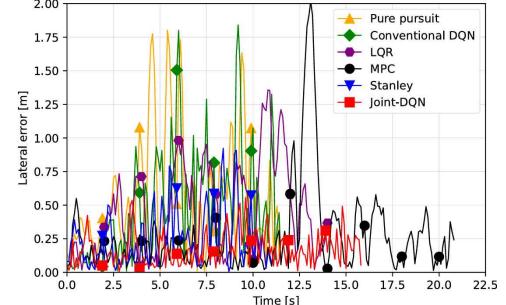


그림 2. 특정 에피소드에 대한 횡방향 오차 성능 비교.

그림 2는 각 경로 추적 알고리즘의 평균 횡방향 오차가 충분히 수렴한 에피소드에서, 주행 종료 시점까지의 횡방향 오차 변화를 비교한 결과를 보여준다. MPC는 최대 2m, Pure pursuit과 Conventional DQN은 최대 1.75m, LQR은 최대 1.25m 수준의 횡방향 오차를 기록하며 전반적으로 낮은 주행 정밀도와 불안정한 추적 성능을 보였다. Stanley는 약 12초 만에 경로를 완주하였지만, 최대 횡방향 오차가 0.9m에 달해 단일 차선 내에서 안정적인 주행을 유지하기 어려운 한계를 드러냈다. 반면, 제안된 Joint-DQN은 약 15.5초에 주행을 마치는 동시에, 최대 횡방향 오차가 0.5m로 가장 낮은 수준을 기록하며 뛰어난 경로 추적 안정성을 보였다. 종합적으로, Joint-DQN은 가속도와 조향각에 대한 제어 정책을 통합적으로 학습하는 공동 제어 구조를 바탕으로, 기존 기법들보다 우수한 경로 추적 정밀도와 주행 안정성을 달성하였으며, 고정밀 자율주행 경로 추적에 적합한 알고리즘임을 검증하였다.

IV. 결론

본 연구에서는 가속도와 조향각을 동시에 제어하는 DQN 기반 경로 추적 알고리즘을 제안하였다. 제안된 기법은 두 제어 변수를 2차원 투플 형태의 공동 액션으로 구성함으로써, 보다 정밀하고 안정적인 경로 추적을 가능하게 하였다. 시뮬레이션 결과, 제안된 알고리즘은 기존의 기하학 기반, 최적화 기반, 그리고 조향 제어 중심 DQN 기법들과 비교하여 평균 및 최대 횡방향 오차 측면에서 가장 우수한 성능을 보였다. 향후 연구에서는 다양한 도로 조건과 복잡한 교통 환경을 고려한 시나리오로 확장하여, 제안 기법의 일반화 성능과 실용 가능성성을 추가적으로 검증할 예정이다.

ACKNOWLEDGMENT

이 논문은 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원-대학ICT연구센터(ITRC)의 지원(IITP-2025-RS-2023-00259991, 50%)과 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(RS-2024-00394286, 50%).

참고 문헌

- J. M. Snider, "Automatic steering methods for autonomous automobile path tracking," Carnegie Mellon University, Pittsburgh, PA, Tech. Rep. CMU-RI-TR-09-08, Feb. 2009.
- S. G. Yang, E. H. Cho, J. Kim, and S.-C. Lim, "Deep reinforcement learning-based path-tracking for unmanned vehicle navigation enhancement," in Proc. Int. Conf. Electron., Inf. Commun. (ICEIC), Jan. 2024, pp. 374 - 377.