

강화학습과 회전 반경 계인을 활용한 4륜 독립 조향 EV의 경로 추종 전략

김 태 윤 · 이 훈 기 · 김 경 호 · 황 성 호*

성균관대학교 기계공학과

Path-Following Strategies for 4-Wheel Independent Steering EVs Using PPO Reinforcement Learning and Turning Radius Gain

Taeyun Kim · Hunki Lee · Kyungho Kim · Sung-Ho Hwang*

Department of Mechanical Engineering, Sungkyunkwan University, Gyeonggi 16419, Korea

(Received 10 April 2023 / Revised 25 April 2023 / Accepted 3 May 2023)

Abstract : Recent developments in the automobile industry include the introduction of e-Corner modules, which incorporate driving, braking, steering, and suspension systems within each wheel of a vehicle. Accordingly, research on the control strategy of a four-wheel independent drive and steering EV is actively underway. This paper proposes a path-tracking controller by applying PPO reinforcement learning in four-wheel independent steering vehicles. An algorithm for switching the steering mode of a four-wheel independent steering vehicle was used with the turning radius gain variable. Then, a learning environment was developed and evaluated by using a vehicle dynamics simulator with repeated low and high curvature paths. Next, the learned agents compared performance with commonly used lateral controllers, namely, Pure Pursuit and Stanley. The proposed controller's lateral path offset error performance was less than 30 cm in the low-curvature scenario, while, in the high-curvature scenario, the paths following performance and stability were better than those of Pure Pursuit and Stanley.

Key words : Autonomous driving shuttle(자율주행 셔틀), 4-Wheel independent steering(4륜 독립 조향), Path tracking algorithm(경로 추종 알고리즘), Reinforcement learning(강화학습), Turning radius gain(회전 반경 계인)

Nomenclature

m	: mass, kg	κ_{ego}	: inverse of vehicle turning radius, 1/m
L	: wheel base, m	δ	: steering angle, deg
w	: tread, m	δ_o	: outer wheel steering angle, deg
h	: height, m	δ_i	: inner wheel steering angle, deg
R	: vehicle turning radius, m	δ_{pp}	: Pure Pursuit steering angle, deg
v_{target}	: target velocity, m/s	$\delta_{stanley}$: stanley steering angle, deg
v_x	: longitudinal velocity, m/s	e_y	: lateral path offset error, m
a_x	: longitudinal acceleration, m/s ²	$e_\varphi = \varphi$: heading offset error, rad
T	: motor torque, Nm	k_r	: turning radius gain
a	: look-ahead point angle, rad	k_{ld}	: look-ahead distance gain
l_d	: look-ahead distance, m	k	: lateral offset error gain, 1/s
κ	: road curvature, 1/m	k_{soft}	: stanley - softening factor, m/s
		MDP	: markov decision process

*Corresponding author, E-mail: hsh0818@skku.edu

*This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License(<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium provided the original work is properly cited.

1. 서론

최근 자동차 산업이 발전함에 따라 새로운 교통수단과 교통서비스가 등장하고 있다. 그중에서도 미래 모빌리티 사회에서 다양한 주행 서비스 제공이 가능한 핵심 운송 수단, 즉 MaaS(Mobility as a Service) 개념의 교통수단으로 자율주행 셔틀이 제시, 개발되고 있다.^{1,2)} 국내에서는 차세대융합기술연구원이 개발한 제로셔틀을 시작으로 한국교통대학교의 아이브, 스포링클라우드의 스포링카, 언맨드솔루션의 위드어스, 한국자동차연구원의 카모 등이 개발되었다. 최근에는 차량의 기동성을 극대화하기 위해 4륜 독립 구동·조향이 가능한 e-Corner 모듈을 사용한 자율셔틀 공용재시플랫폼이 개발되고 있다.

e-Corner 모듈은 차량의 각 바퀴 안에 구동, 제동, 조향, 현가 시스템을 통합해 놓은 모듈이다.^{3,4)} e-Corner 모듈에 대한 개발이 활발하게 진행됨과 동시에 4륜 독립 구동·조향 전기차의 제어 전략에 대한 연구가 활발히 진행되고 있다.⁵⁻¹⁰⁾ 많은 연구에서 전륜조향(Front wheel steering; FWS), 후륜조향(Rear wheel steering; RWS), 4륜조향(Four wheel steering; 4WS) 등 여러가지 조향 모드에 대한 연구를 진행하였지만, 다중 조향 모드의 동적 전환이 불가능하며 조향 모드 변경 시 차량 정지 후 모드 전환이라는 제약이 있다.¹¹⁾ 주행 중 조향 모드의 동적 전환에 대한 연구도 진행되었지만 전환 시 차량 회전 반경을 유지시켜야 한다는 문제가 있다.¹²⁾



Photo. 1 Unmanned solution's WITH:US & e-Corner module

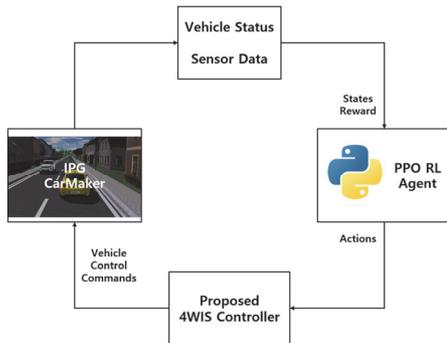


Fig. 1 Reinforcement learning framework

한편, 딥러닝 기술이 발전함에 따라 강화학습을 자율주행 차량의 제어에 활용하는 연구가 활발히 진행되고 있다. 강화학습은 주어진 환경에서 에이전트가 최대한의 보상을 위해 행동을 개선하는 방향으로 학습하는 알고리즘이다. 자율주행 자동차의 적응형 순항 제어, 차선 유지시스템 등과 같은 운전자 보조 시스템이나 복잡한 시나리오에서의 의사결정에 강화학습을 적용하는 연구가 진행되었다.¹³⁻¹⁶⁾

본 논문에서는 회전 반경 개인 변수를 도입한 4륜 독립 조향 차량의 조향 알고리즘을 제안하며 경로 추종 중 회전 반경 개인 변수의 변화를 통해 조향 모드를 변화시키고자 하였다. 도로 곡률 및 경로 추종 상태에 따른 회전 반경 개인 변화량과 조향 값의 변화량을 강화학습을 통해 학습하였다. 에이전트가 현재 상태에서 취할 수 있는 모든 행동에 대해 탐색을 수행하고, 해당 경험을 통해 학습해야 하므로 시뮬레이션 환경을 사용하였다. IPG CarMaker 소프트웨어를 활용하여 4륜 독립 구동·조향이 가능한 차량을 구성하였으며, Fig. 1과 같이 시뮬레이터와 Python의 강화학습 에이전트 간의 프레임워크를 구성하였다. 시뮬레이터와 에이전트 간의 통신으로는 MMF(Memory mapped file) 기법을 활용하였다. 저 곡률, 고 곡률 시나리오에서의 경로 추종 성능을 대표적인 횡제어기인 Pure pursuit, Stanley와 비교하였으며 제안하는 알고리즘을 활용한 4륜 독립 조향 차량의 FWS, 4WS 모드 변환에 대해 분석하였다.

2. 학습 환경 구성

2.1 차량 구성

CarMaker 시뮬레이터를 활용하여 4륜 독립 구동·조향 차량을 구성하였다. 차량 제원 및 e-Corner 모듈 제원은 국내에서 개발이 진행되고 있는 캐빈교체형 자율셔틀을 따랐으며 Table 1, Fig. 2와 같다.

Table 1 Simulation vehicle specifications

Vehicle specifications	
Overall length, width, height (m)	5.3, 2.1, 1.1
Wheelbase (m)	3.7
Overall mass (kg)	3,400
Motor maximum output (Nm)	630
Motor maximum speed (rpm)	450
Wheel steering speed (deg/s)	213
Wheel steering range (deg)	Toe-in 90 Toe-out 45

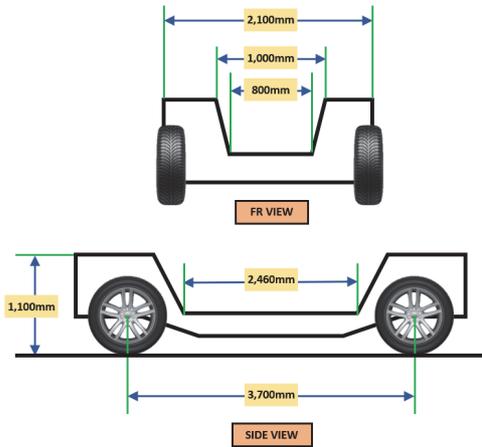


Fig. 2 Simulation vehicle schematics

2.2 조향 알고리즘 구성

2.2.1 애커만 조향(Ackermann Steering)

애커만 조향은 1817년 Georg Lankensperger가 발명한 기하학적인 조향 방식이다.¹⁷⁾ 전륜의 내부 바퀴와 외부 바퀴가 동일한 조향 값을 가지게 되면 서로 다른 회전 중심을 가지게 된다. 애커만 조향은 이를 개선하기 위해 Fig. 3과 같은 방법으로 전륜의 내부, 외부 바퀴의 조향각을 달리하여 4륜의 회전 중심을 일치시키는 방법이다.

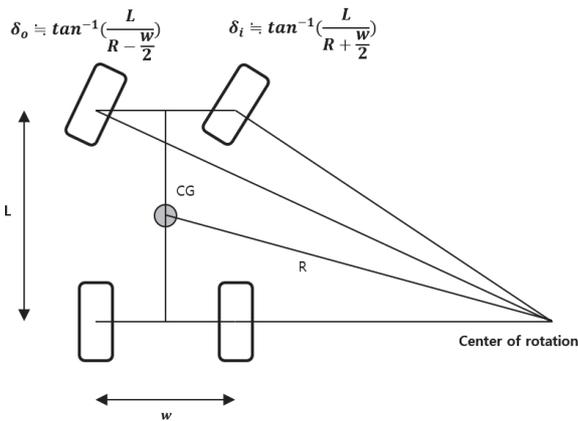


Fig. 3 Ackermann steering geometry

2.2.2 회전 반경 계인을 활용한 4륜 독립 조향 알고리즘

4륜 독립 조향 차량의 경우 Fig. 4와 같이 FWS, RWS 이외에도 후륜을 전륜과 같은 방향으로 조향하는 방법, 다른 방향으로 조향하는 방법 등을 활용할 수 있다. 하지만 주행 중 조향 방법을 전환하는 데에는 많은 어려움이 따르므로 본 논문에서는 회전 반경 계인(k_r)을 활용하여 조향 방법을 전환하는 방법을 제안한다. 제안하는 회전 반경 계인을 활용한 4륜 독립 조향 알고리즘은 Fig. 5와 같다.

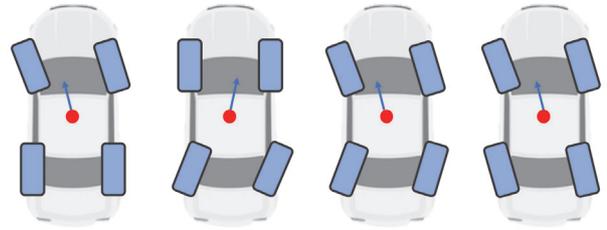


Fig. 4 Steering modes of 4WIS vehicle

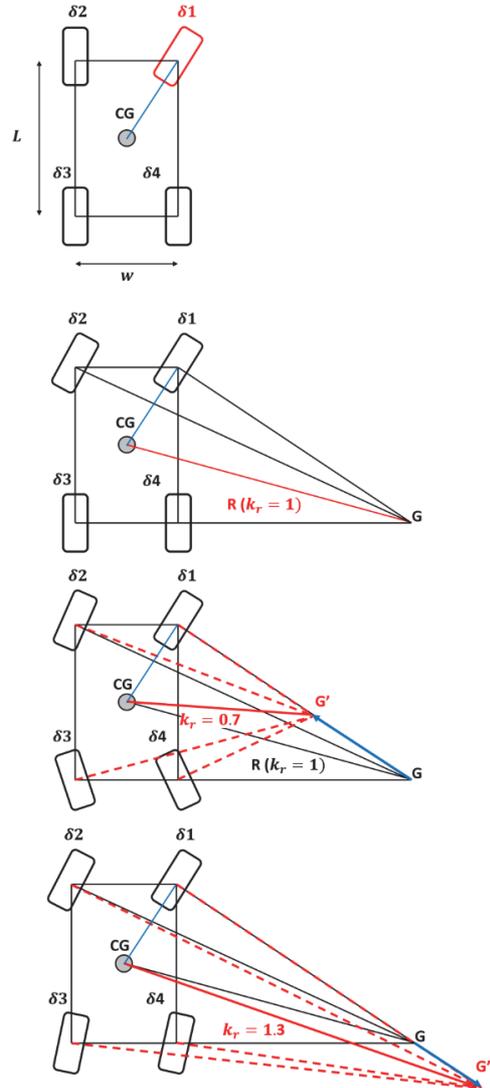


Fig. 5 Steering algorithms using turning radius gain

차량의 주 조향륜을 우선회 시 우측 전륜, 좌선회 시 좌측 전륜으로 선정한다. 주 조향륜의 조향 값인 δ_1 이 결정되었을 때, 애커만 조향 알고리즘을 통해 회전 중심 G 좌표를 선정, 무게중심까지의 거리인 회전반경 R을 계산한다. 이때의 회전 반경 계인 k_r 은 1이 된다.

회전 반경 계인을 변경하였을 때, 주 조향륜을 제외한 세 바퀴는 감소된 회전 반경 중심 G' 에 수직이 되도록 변화한다. 회전 반경 계인을 감소시킬 경우 Fig. 5의 세 번째 그림과 같이 후륜이 전륜과 역상으로 위치하며 회전 반경이 감소하게 된다. 또한 회전 반경 계인을 증가시킬 경우 Fig. 5의 네 번째 그림과 같이 후륜이 전륜과 동상으로 위치하며 회전 반경이 증가한다.

2.3 강화학습을 위한 MDP 정의

강화학습은 주어진 환경에서 누적 보상을 최대화하기 위해 행동의 주체인 에이전트가 시행착오를 통해 행동을 교정하는 학습 과정이다. 에이전트는 시물레이션 혹은 시스템 환경으로부터 관찰되는 복수의 상태를 바탕으로 행동을 결정한다. 설정된 보상함수를 바탕으로 행동에 대한 상태의 변화를 통해 보상을 받으며, 보상이 최대가 되도록 학습한다.

Markov 결정 과정(MDP)은 순차적으로 행동을 결정해야 하는 문제를 수학적으로 표현한 것으로, 상태의 집합 S , 행동의 집합 A , 상태 전이 확률 P , 보상 함수 R , 감쇠 인자 γ 로 구성된다. 강화학습은 MDP로 정의된 문제를 푸는 것으로, 경로 추종을 위한 4륜 독립 조향 알고리즘을 강화학습에 적용하기 위해 MDP를 다음과 같이 정의했다.

2.3.1 상태 공간 정의

상태 공간은 경로 추종을 위한 자차(Ego vehicle)의 목표 속도, 종 방향 속도, 종 방향 가속도, 경로 횡 오차, 경로 헤딩 오차, 도로 곡률, 도로 곡률 절댓값, 조향각, 회전 반경 계인, 자차 회전 반경의 역수로 구성하였으며 Table 2와 같다. 도로 곡률 절댓값은 하이퍼볼릭 탄젠트 활성화 함수를 사용한 네트워크에서 절댓값을 유도하기 힘들다고 판단하여 추가한 상댓값이다.

4륜 독립 조향이 가능한 차량의 특성을 고려하여 도로 곡률을 기반으로 상태를 입력하였으며 조향 값과 회전 반경 계인 값을 통해 자차 회전 반경의 역수를 계산하여 상태로 입력하였다.

2.3.2 행동 공간 정의

행동 공간은 Table 3과 같이 4륜에 작용하는 구동 토크, 회전 반경 계인 변화량, 조향 값 변화량으로 구성하였다. 구동 토크는 4륜에 동일하게 입력하였으며, 회전 반경 계인 변화량은 한 Step 당 최대 0.1, 조향 값 변화량은 최대 21.3 deg로 제한하였다. 또한 시물레이션 환경에서 검증 시 실차의 하드웨어 딜레이를 고려하기 위해 시상수 τ 를 0.2로 한 1차 전달함수를 적용하였다.

Table 2 State space definition

State	Meaning
S_1	Target velocity, v_{target}
S_2	Longitudinal velocity, v_x
S_3	Longitudinal acceleration, a_x
S_4	Lateral path offset error, e_y
S_5	Heading offset error, $e_\phi = \phi$
S_6	Road curvature, κ
S_7	Absolute road curvature, $ \kappa $
S_8	Steering angle, δ
S_9	Turning radius gain, k_r
S_{10}	Inverse of vehicle turning radius, κ_{ego}

Table 3 Action space definition

Action	Meaning
A_1	Motor torque, T
A_2	Turning radius gain change, Δk_r
A_3	Steering angle change, $\Delta \delta$

2.3.3 보상 함수 정의

보상 함수는 식 (1)과 같이 목표 속도 유지를 위한 항, 횡 오차에 대한 페널티, 조향 및 회전 반경 계인 변화에 대한 페널티, 승차감을 위한 가가속도 페널티로 구성하였다. 승차감을 위해 회전 반경 계인을 높게 선정하여 주행할 경우 전륜에 비해 후륜 중심이 경로에서 많이 벗어나는 문제를 방지하기 위해 횡 오차에 대한 페널티는 전륜의 횡 오차와 후륜의 횡 오차 모두 고려하도록 설계하였다.

$$R = r_v + r_{path_front} + r_{path_rear} + r_{steer} + r_{radius} + r_{jerk} \tag{1}$$

각 보상 함수가 현재 속도에 비례한 페널티를 받을 수 있도록 식 (2)와 같이 목표 속도에 대한 현재 속도 비율을 정의하여 곱해주었다. 또한 경로 횡 오차가 한 개 차로 폭에 해당하는 3.5 m를 넘어가거나 경로 헤딩오차가 경로에서 수직이 되는 각도인 $\frac{\pi}{2}$ 를 넘어가는 경우 이후의 경로 추종은 무의미하다고 판단하여 큰 페널티를 부과한 후 시나리오를 종료하였다. 세부적인 보상의 정의 및 계인은 Table 4와 같다.

Table 4 Reward function definition

Reward	Gain	Definition
r_v	-1.0	$\frac{ v_{target} - v_x }{v_{target}}$
r_{path_front}	-1.0	$\tilde{v} \cdot e_{y_front} $
r_{path_rear}	-0.5	$\tilde{v} \cdot e_{y_rear} $
r_{steer}	-1.0	$\tilde{v} \cdot \Delta\delta $
r_{radius}	-2.0	$\tilde{v} \cdot \Delta k_r $
r_{jerk}	-0.5	$\tilde{v} \cdot (j_y + j_x)$

Table 5 Hyperparameters of PPO reinforcement learning

Parameter	Meaning	Value
K_{epoch}	Num of epochs in 1 update	80
ϵ	Clipped surrogate objective's maximum degree of clipping	0.2
γ	Discount factor	0.99
lr_{actor}	Learning rate of actor	3e-4
lr_{critic}	Learning rate of critic	1e-3
t_{max}	Maximum training timesteps	5e6
f_{update}	Update frequency in timestep	3e3
σ_{init}	Initial standard deviation for action distribution	0.6
σ_{min}	Minimum standard deviation	0.1
$\Delta\sigma$	Standard deviation decaying rate in timestep	0.05
f_{σ}	Standard deviation decaying frequency	4e5

$$\tilde{v} = \min \left[\frac{v_x}{v_{target}}, 1 \right] \quad (2)$$

2.3.4 강화학습 네트워크 선정

에이전트의 행동 범위가 연속적일 경우 DQN(Depp-Q-Network)과 같은 가치 기반 방법론^{18,19)}이 아닌 REINFORCE와 같은 정책 기반 방법론²⁰⁻²²⁾을 사용해야 한다. 본 논문에서는 정책 기반 방법론 중 다양한 환경에

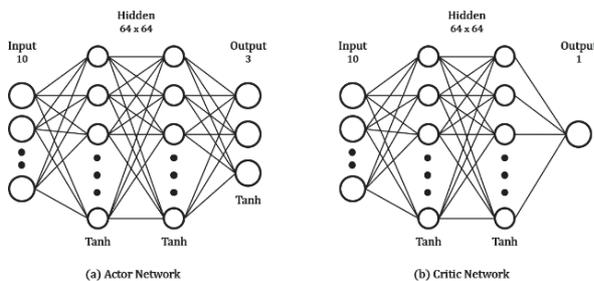


Fig. 6 Actor & Critic network structure

서 좋은 성능을 보여주며, 구현 방법이 간단한 PPO (Proximal policy optimization)²³⁾ 알고리즘을 사용하였다. PPO 알고리즘은 TRPO(Trust region policy optimization)²⁴⁾의 신뢰 영역 정책 최적화 방법을 개선한 Clipped surrogate 목적함수를 사용하고, 학습 데이터를 재사용하여 수렴 속도가 빠른 장점이 있다. 하지만 하이퍼 파라미터 설정이 잘못되면 학습이 불안정해지거나 느려지는 단점이 있다. Table 5는 본 논문에서 사용한 PPO 알고리즘의 하이퍼 파라미터이다.

2.3.5 네트워크 신경망 구성

액터와 크리틱의 신경망 구조는 Fig. 6과 같다. 신경망은 10개의 상태 정보를 입력으로 받아 64 × 64 크기의 은닉층을 거쳐 3개의 출력값을 계산한다. 각 은닉층은 출력값을 -1에서 1 사이로 정규화하는 하이퍼볼릭 탄젠트 활성화 함수를 사용하였으며, 액터의 최종 출력에는 하이퍼볼릭 탄젠트 활성화 함수를 한 번 더 사용하여 출력의 범위와 행동의 범위를 일치시켜 주었다.

3. 학습 시나리오 및 평가 시나리오

3.1 학습 시나리오

강화학습 시 학습 시나리오가 특정된다면 해당 시나리오에 과적합 되어 다른 시나리오에서의 성능을 보장할 수 없게 된다. 이를 방지하기 위해 직진, 좌선회, 우선회, U-turn 등의 경로가 다양하게 혼합될 수 있는 시나리오를 구성하였다. 각 경로의 회전 반경 범위와 회전 각도는 Table 6과 같으며, 랜덤하게 생성된 경로의 예시는 Fig. 7과 같다.

Table 6 Training scenario turn descriptions

Type	Turning radius	Turning degree
Straight	10 ~ 100 m (Length)	-
Left turn	15 ~ 100 m	50 ~ 90 deg
Right turn	15 ~ 100 m	50 ~ 90 deg
U-turn	5 ~ 15 m	180 deg

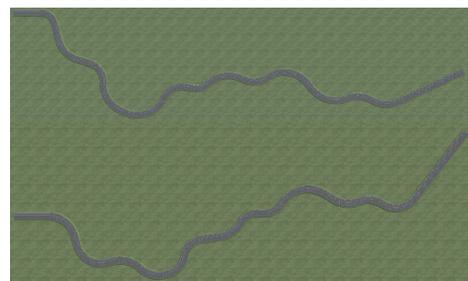


Fig. 7 Examples of randomly generated scenario

3.2 평가 시나리오

학습된 에이전트와 비교 제어를 평가할 때는 고정된 시나리오를 사용하였다. 경로 추종 성능 평가는 저 곡률 도로에 대한 경로 추종 성능과 U-turn을 포함한 고 곡률 경로 추종 성능으로 구분하여 진행하였다.

Table 7 Low curvature test scenario description

Type	Turning radius	Turning degree
Right	30 m	120 deg
Right	85 m	42 deg
Left	65 m	38 deg
Right	80 m	45 deg
Right	80 m	70 deg
Left	40 m	30 deg
Right	40 m	35 deg
Right	49 m	160 deg
Left	60 m	120 deg

저 곡률 시나리오는 국토교통부 국토교통과학기술진흥원에서 진행한 자율주행 자동차 안전성 평가기술 및 테스트베드 개발 보고서를 기반으로 Fig. 8과 같이 최소 회전 반지름 30 m, 최대 회전 반지름 85 m의 경로를 생성하였다. 각 선회 구간에 대한 회전 반지름은 Table 7과 같으며 각 선회 영역 사이에는 연속적인 곡률을 가지는 클로소이드 곡선 형상의 도로를 추가하였다.

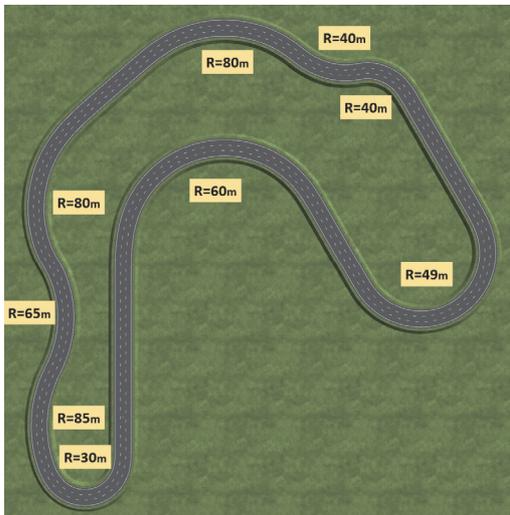


Fig. 8 Low curvature test scenario

고 곡률 시나리오는 Fig. 9와 같이 도심지에서 마주할 수 있는 우회전, U-turn, 좌회전 구간의 회전 반경으로 구성되었다. 특히 U-turn 구간의 회전 반경은 6 m로 “도로

의 구조·시설 기준에 관한 규칙” 제5조의 승용차의 최소 회전 반지름을 기준으로 하였다.



Fig. 9 High curvature test scenario

4. 주행 실험 및 결과

4.1 비교 제어기

학습된 에이전트의 경로 추종 성능을 비교하기 위해 기하학적 횡 제어 방법인 Pure pursuit 제어기와 Stanley 제어기를 사용하였다.²⁵⁾

4.1.1 Pure Pursuit Controller

Pure Pursuit 제어기는 차량의 현재 위치와 속도를 기반으로 경로를 따라 이동하는 데 필요한 조향각을 계산하는 기하학적 제어 방법이다.²⁶⁾ 제어 목표 지점인 예견점(Look-ahead point)을 선정할 후 예견점과 차량 헤딩 방향과의 각도 차이 α 를 sin 법칙을 통해 유도한다. 이후 아커만 조향각(Ackermann steering angle) 모델을 적용해 식 (3)과 같이 목표 조향각을 계산한다.

$$\delta_{pp} = \tan^{-1} \left(\frac{2L \sin \alpha}{l_d} \right) \quad (3)$$

$$l_d = k_{l_d} \cdot v \quad (k_{l_d} = 1 \cdot s) \quad (4)$$

기본적으로 Pure pursuit 제어기는 차량 속도에 따라 예견점을 결정하나,²⁷⁾ 주행 시 예견점이 멀 경우 경로의 안쪽을 파고드는 Cut-corner 현상과 미세한 경로 횡 오차를 잡지 못하고 유지하며 주행하는 Steady-state offset 현상이 발생할 수 있다. 본 논문에서는 예견점을 식 (4)와 같이 설정하였다.

4.1.2 Stanley Controller

Stanley 제어기는 전륜 기준의 경로 횡 오차, 경로 헤딩 오차를 이용하는 제어 방법이다. 전륜과 가장 가까운 경로상의 점을 기준으로 횡 오차 e_y , 헤딩 오차 e_ϕ 를 계산

한 후 식 (5)와 같이 목표 조향각을 계산한다.²⁸⁾

본 논문에서는 저속에서 경로 횡 오차에 대한 항이 너무 커지는 것을 방지하기 위한 k_{soft} 값을 1m/s로 설정하였으며, 경로 횡 오차에 대한 계인 k 는 $2.5 \frac{1}{s}$ 로 설정하였다.

$$\delta_{stanley} = e_{\varphi} + \tan^{-1} \left(\frac{k \cdot e_y}{k_{soft} + v_x} \right) \quad (5)$$

4.2 저 곡률 경로 추종 시나리오 결과

강화학습 에이전트의 저 곡률 시나리오 경로 추종 성능을 Pure pursuit, Stanley 제어기와 비교하였다. 목표 속도는 20 km/h로 하였으며 Pure pursuit과 Stanley 횡 제어기 평가에는 Intelligent driver model을 종 방향 제어로 사용하여 목표 속도를 추종하였다.

주행 성능 평가로는 경로 횡 오차, 경로 헤딩 오차를 사용하였으며, 차량 승차감 평가로는 횡 방향 가속도 (Lateral Jerk)[m/s³]를 사용하였다. 세 제어기의 저 곡률

Table 8 Low curvature test scenario results

	Max e_y	Mean e_y	Mean e_{φ}	Max Jerk	Mean Jerk
Pure pursuit	0.265	0.047	0.046	1.154	0.053
Stanley	0.193	0.044	0.045	3.972	0.055
Reinforcement learning	0.216	0.043	0.062	3.563	0.157

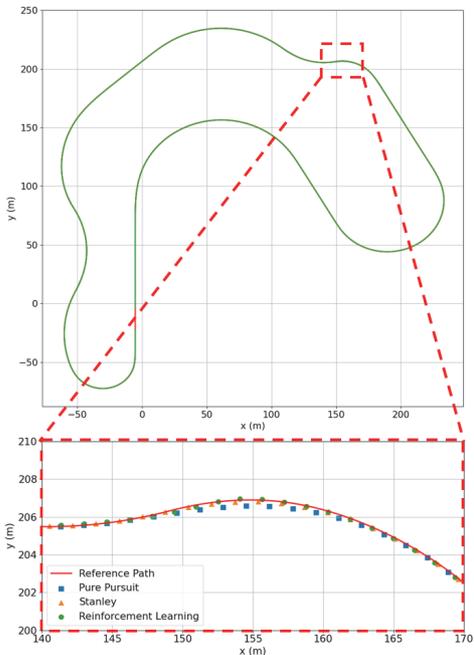


Fig. 10 Trajectory results in low curvature scenario

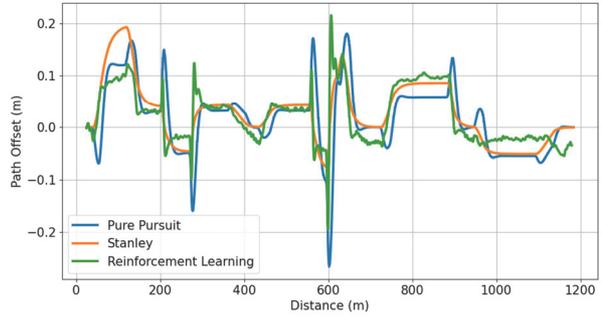


Fig. 11 Path offset results in low curvature scenario

시나리오 주행 결과는 Table 8과 같다. 경로 횡 오차는 Stanley, Reinforcement learning, Pure pursuit 순으로 좋은 결과를 보였으며 경로 헤딩 오차는 Stanley, Pure pursuit, Reinforcement learning 순으로 좋은 결과를 보였다. 횡 방향 평균 Jerk는 Pure pursuit, Stanley, Reinforcement learning 순으로 좋은 결과를 보였으나 최대 Jerk는 Stanley가 가장 높은 것으로 나타났다.

4.3 고 곡률 경로 추종 시나리오 결과

동일하게 학습된 에이전트를 고 곡률 경로 추종 시나리오에서 평가하였다. 고 곡률 경로 시나리오는 회전 반경 6 m가 포함된 U-turn 구간이 있어 Pure pursuit 제어기의 Cut-corner 현상과 Stanley 제어기의 헤딩 오차에 대한 불안정성이 나타날 수 있는 시나리오이다.

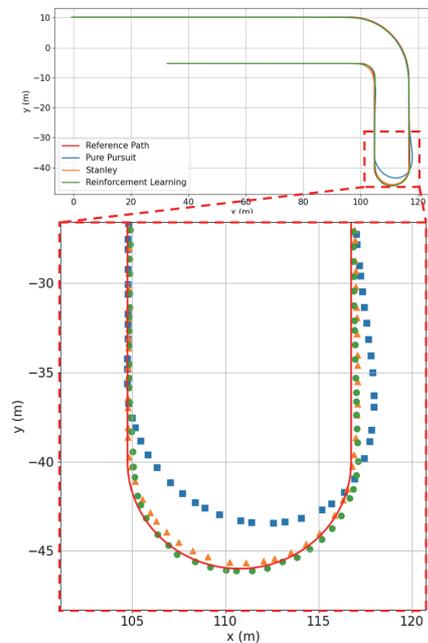


Fig. 12 Trajectory results in high curvature scenario

Table 9 High curvature test scenario results

	Max e_y	Mean e_y	Mean e_ϕ	Max Jerk	Mean Jerk
Pure pursuit	2.256	0.157	0.084	4.086	0.088
Stanley	0.805	0.078	0.076	3.612	0.076
Reinforcement learning	0.65	0.095	0.046	2.535	0.121

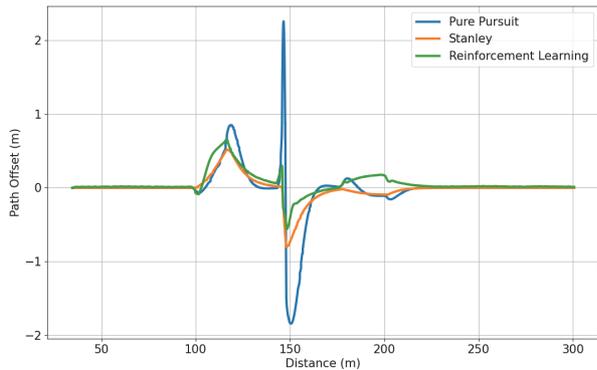


Fig. 13 Path offset results in high curvature scenario

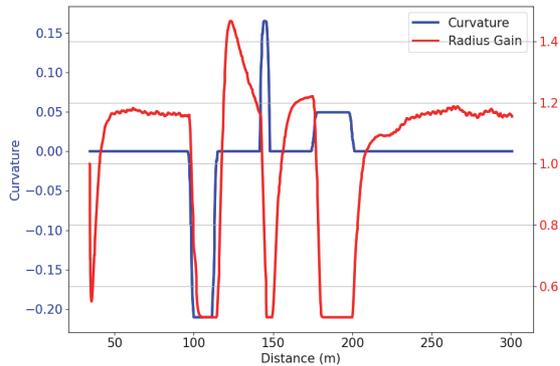


Fig. 14 Curvature and turning radius gain in high curvature scenario

고 곡률 시나리오의 주행 성능 평가로도 경로 횡 오차, 경로 헤딩 오차를 사용하였으며 승차감 평가로 횡 방향 Jerk를 사용하였다. 제어기의 고 곡률 시나리오 주행 결과는 Table 9와 같다. 경로 횡 오차는 Reinforcement learning, Stanley, Pure pursuit 순으로 좋은 결과를 보였으며 Pure pursuit의 경우 Cut-corner 현상이 심하게 나타나 경로에서 2.256 m 벗어난 결과를 보였다. 경로 헤딩 오차는 Reinforcement learning, Stanley, Pure pursuit 순으로 좋은 결과를 보였으며 횡 방향 최대 Jerk 역시 Reinforcement learning, Stanley, Pure pursuit 순으로 좋은 결과를 보였다.

Fig. 14는 거리에 따른 도로 곡률, 에이전트의 회전 반경 계인을 나타낸 것이다. 곡률이 낮을 때 회전 반경 계

인은 1.5가 되어 차량의 후륜을 전륜과 동상으로 주행하고, 곡률이 높을 때 회전 반경 계인을 낮추어 회전 반경을 낮추는 것을 확인할 수 있었다.

5. 결론

본 논문에서는 회전 반경 계인을 활용한 4륜 독립 조향 차량의 경로 추종 제어를 제안하였다. 강화학습을 통해 제어기를 학습시키고, 4륜 독립 조향 차량의 여러 가지 조향 모드를 회전 반경 계인이라는 변수를 통해 전환할 수 있었다.

제한한 회전 반경 계인을 활용한 제어기의 경로 추종 성능을 저 곡률 시나리오, 고 곡률 시나리오에서 확인하였다. 저 곡률 시나리오의 경우 Pure pursuit 제어기보다 경로 추종에서 좋은 성능을 보였지만, 차량의 안정성 부분에서 좋지 않은 결과를 보였다. Stanley 제어기에 비해서는 경로 추종에서 좋지 않은 결과를 보였으나 안정성에서 좋은 결과를 보였다. 고 곡률 시나리오의 경우 Pure pursuit 제어기와 Stanley 제어기보다 경로 추종 성능, 횡 방향 Jerk가 좋은 것으로 나타났다. 또한, 강화학습을 활용한 제어기가 도로 곡률을 기반으로 회전 반경 계인을 변화시키며 차량의 후륜을 전륜과 동상으로 하여 승차감을 개선시키고, 역상으로 하여 회전 반경을 감소시키는 것을 확인할 수 있었다.

추후 연구에서는 차선 변경 경로 추종, 보행자 회피 등 다양한 시나리오에서의 학습을 통해 경로 추종 상황이 외에도 범용적으로 사용할 수 있는 제어기를 개발할 계획이다. 또한, 본 논문에서 제안한 회전 반경 계인 알고리즘을 활용하여 강화학습이 아닌 동역학 수식을 기반으로 한 제어기를 설계할 예정이며, RWS 조향이 불가능한 점을 개선하기 위해 회전 반경 계인 변수를 재정의하여 다양한 조향 모드를 활용할 수 있도록 개선할 계획이다.

후 기

이 논문은 2023년도 정부(산업통상자원부)의 재원으로 한국산업기술진흥원의 지원(P0017120, 2023년 산업혁신인재성장지원사업)과 2023년도 산업통상자원부 및 산업기술평가관리원(KEIT) 연구비 지원에 의한 연구임(‘20014983’).

References

1) A. Bucchiarone, S. Battisti, A. Marconi, R. Maldacea and D. C. Ponce, “Autonomous Shuttle-as-a-Service (ASaaS): Challenges, Opportunities, and Social Implications,” IEEE Transactions on

- Intelligent Transportation Systems, Vol.22, No.6, pp.3790-3799, 2020.
- 2) C. Iclodean, N. Cordos and B. O. Varga, "Autonomous Shuttle Bus for Public Transportation: A Review," *Energies*, Vol.13, No.11, Paper No. 2917, 2020.
 - 3) J. Hag, *Wheel Corner Modules: Technology and Concept Analysis*, Ph. D. Dissertation, KTH Royal Institute of Technology, Sweden, 2011.
 - 4) C. Wilwert, N. Navet and F. Simonot-Lion, *Design of Automotive X-by-Wire Systems*, The Industrial Communication Technology Handbook, CRC Press, United States, 2005.
 - 5) S. H. Lee, U. K. Lee, S. K. Ha and C. S. Han, "Four-wheel Independent Steering (4WIS) System for Vehicle Handling Improvement by Active Rear Toe Control," *JMSE International Journal Series C Mechanical Systems, Machine Elements and Manufacturing*, Vol.42, No.4, pp.947-956, 1999.
 - 6) Y. Lei, G. Wen, Y. Fu, X. Li, B. Hou and X. Geng, "Trajectory-following of a 4WID-4WIS Vehicle Via Feedforward-backstepping Sliding-mode Control," *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, Vol.236, Nos.2-3, pp.322-333, 2022.
 - 7) H. Zheng and S. Yang, "A Trajectory Tracking Control Strategy of 4WIS/4WID Electric Vehicle with Adaptation of Driving Conditions," *Applied Sciences*, Vol.9, No.1, Paper No.168, 2019.
 - 8) J. H. Kim, H. S. Ham, H. M. Jin and C. W. Moon, "Lateral Behavior Simulation of Independent Four Wheel Steering Autonomous Shuttle without Steering Linked System," *KSAE Fall Conference Proceedings*, p.709, 2022.
 - 9) H. Chen, S. Chen, R. Zhou, X. Huang and S. Zhu, "Research on Four-wheel Independent Steering Intelligent Control Strategy Based on Minimum Load," *Concurrency and Computation: Practice and Experience*, Vol.33, No.9, e6145, 2021.
 - 10) P. Hang and X. Chen, "Towards Autonomous Driving: Review and Perspectives on Configuration and Control of Four-wheel Independent Drive/Steering Electric Vehicles," *Actuators*, Vol.10, No.8, p.184, 2021.
 - 11) T. L. Lam, H. Qian and Y. Xu, "Omnidirectional Steering Interface and Control for a Four-wheel Independent Steering Vehicle," *IEEE/ASME Transactions on Mechatronics*, Vol.15, No.3, pp.329-338, 2009.
 - 12) F. Xu, X. Liu, W. Chen and C. Zhou, "Dynamic Switch Control of Steering Modes for Four Wheel Independent Steering Rescue Vehicle," *IEEE Access*, Vol.7, pp.135595-135605, 2019.
 - 13) C. Desjardins and B. Chaib-Draa, "Cooperative Adaptive Cruise Control: A Reinforcement Learning Approach," *IEEE Transactions on Intelligent Transportation Systems*, Vol.12, No.4, pp.1248-1260, 2011.
 - 14) Y. Lin, J. McPhee and N. L. Azad, "Comparison of Deep Reinforcement Learning and Model Predictive Control for Adaptive Cruise Control," *IEEE Transactions on Intelligent Vehicles*, Vol.6, No.2, pp.221-231, 2020.
 - 15) H. K. Lee, T. Y. Kim, H. B. Kim and S. H. Hwang, "Reinforcement Learning Based Autonomous Emergency Steering Control in Virtual Environments," *Journal of Drive and Control*, Vol.19, No.4, pp.110-116, 2022.
 - 16) A. E. Sallab, M. Abdou, E. Perot and S. Yogamani, "End-to-end Deep Reinforcement Learning for Lane Keeping Assist," *arXiv preprint arXiv:1612.04340*, 2016.
 - 17) W. C. Mitchell, A. Staniforth and I. Scott, "Analysis of Ackermann Steering Geometry," *SAE 2006-01-3638*, 2006.
 - 18) V. Mnih, K. Kavukcuoglu, D. Silver, A. Graves, I. Antonoglou, D. Wierstra and M. Riedmiller, "Playing Atari with Deep Reinforcement Learning," *arXiv preprint arXiv:1312.5602*, 2013.
 - 19) H. V. Hasselt, A. Guez and D. Silver, "Deep Reinforcement Learning with Double Q-learning," *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol.30, No.1, 2016.
 - 20) T. P. Lillicrap, J. J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa, D. Silver and D. Wierstra, "Continuous Control with Deep Reinforcement Learning," *arXiv preprint arXiv:1509.02971*, 2015.
 - 21) S. Fujimoto, H. Hoof and D. Meger, "Addressing Function Approximation Error in Actor-critic Methods," *International Conference on Machine Learning*, pp.1587-1596, 2018.
 - 22) T. Haarnoja, A. Zhou, P. Abbeel and S. Levine, "Soft Actor-critic: Off-policy Maximum Entropy Deep Reinforcement Learning with A Stochastic Actor," *International Conference on Machine Learning*, pp.1861-1870, 2018.
 - 23) J. Schulman, F. Wolski, P. Dhariwal, A. Radford and O. Klimov, "Proximal Policy Optimization Algorithms," *arXiv preprint arXiv:1707.06347*, 2017.
 - 24) J. Schulman, S. Levine, P. Moritz, M. Jordan and P.

- Abbeel, "Trust Region Policy Optimization," International Conference on Machine Learning, pp.1889-1897, 2015.
- 25) H. G. Lee, T. Y. Kim, D. Y. Yu and S. H. Hwang, "Path-following Correction Control Algorithm using Vehicle State Errors," Transactions of KSAE, Vol.30, No.2, pp.123-131, 2022.
- 26) R. C. Coulter, Implementation of the Pure Pursuit Path Tracking Algorithm, Carnegie-Mellon UNIV Pittsburgh PA Robotics INST, 1992.
- 27) H. J. Joo and K. B. Lee, "Changes in Path Tracking Performance of Autonomous Vehicle according to Design Feedback Gain and Look-ahead Distance of Pure-pursuit Algorithm," Transactions of KSAE, Vol.29, No.9, pp.839-846, 2021.
- 28) G. M. Hoffmann, C. J. Tomlin, M. Montemerlo and S. Thrun, "Autonomous Automobile Trajectory Tracking for Off-road Driving: Controller Design, Experimental Validation and Racing," American Control Conference, pp.2296-2301, 2007.