

Maximum Likelihood Estimation 기반 주행 특성 추론 연구

이동수, 권민혜*

송실대학교

movementwater@soongsil.ac.kr, minhae@ssu.ac.kr*

Maximum Likelihood Estimation Based Driving Characteristic Inference

Dongsu Lee, Minhae Kwon*

Soongsil Univeristy

요 약

자율주행시스템 기술의 빠른 성장과 함께 자율주행차량의 상용화가 논의되고 있으며, 이는 자연스럽게 자율주행차량과 인간 주행 차량이 도로를 공유하는 도로 환경인 mixed-autonomous traffic을 조성할 것이다. 이러한 혼합 도로 환경에서 특정 주행 차량의 주행 특성을 파악하는 것은 주행 행동 예측으로 이어질 수 있으며, 이는 인간 주행자 및 자율주행차량의 적응적 의사결정에 도움이 될 수 있다. 본 연구에서는 목표 주행 차량의 관찰을 통해 경로 데이터를 기반으로 역 강화학습(Inverse Reinforcement Learning) 및 Maximum Likelihood Estimation (MLE) 방법을 통해 주행 특성을 추론한다. 결과적으로 선행 학습 시킨 주행 차량을 관측하여 주행 특성 계수를 추론하는데 성공하였다.

1. 서 론

자율주행 시스템 기술의 성장과 함께 자율주행차량의 상용화에 대한 논의가 계속해서 이루어지고 있다. 자율주행차량의 상용화는 자연스럽게 자율주행차량과 비 자율주행차량이 혼재되어 주행하는 도로 환경인 mixed-autonomous traffic을 조성하게 될 것이다[1]. 이와 같은 도로 환경에서 자율주행차량이 다른 주행 차량의 행동 패턴을 유추할 수 있다면 적응적인 의사결정에 중요한 정보를 제공할 수 있다. 따라서, 본 연구에서는 개별적인 목표 차량의 주행 특성을 추론하기 위한 전체 프레임워크를 제안하고자 한다. 구체적으로 연속적이며 비선형적인 동적 공간 및 혼재된 행동 공간에서 부분적인 관측 정보를 기반으로 동작하는 개체의 내부 모델(internal model) 추론을 위해 inverse rational control [2] 및 MLE 기반의 프레임워크를 제안한다. 이를 위해 자율 주행 차량의 학습 문제를 정의하는 Partially Observable Markov Decision Process(POMDP), 그리고 MLE 기반의 주행 차량의 특성 추론 방법을 확인한다. 이후 시뮬레이션을 통해 제안된 방법의 성능을 확인한다.

II. 자율주행차량의 주행 특성 추론 방법

본 장에서는 특성을 고려한 개체의 학습 문제를 정의하는 POMDP 모델 확인 후, MLE를 기반으로 개체의 특성 추론 방법을 제안한다.

II.1. Partially Observable Markov Decision Process 및 문제 정의

관측을 위한 차량의 사전 학습 문제는 POMDP 모델로 정의할 수 있다. POMDP 모델은 튜플 $M = \langle S, O, A, R, T, \Omega, \gamma \rangle$ 로 정의하며, 연속적인 상태 공간 $s_t \in S$, 연속적인 관측 공간 $o_t \in O$, 혼재된 (hybrid) 행동 공간 $a_t = [a_t^{acc}, a_t^{lc}] \in A$, 보상함수 R , 상태 전이 확률(state transition probabilities) $T(s_{t+1}|s_t, a_t)$, 관측 전이 확률 $\Omega(o_t|s_t)$, 그리고 $\gamma = [0, 1]$ 는 감가율(discounted factor)을 의미한다. 구체적으로 보상 r_t 는 특성 계수(character coefficient) c_n 과 보상항 R_n 의 선형 결합으로

$$R(s_t, a_t, s_{t+1}; c) = \sum_{n=1}^N c_n R_n \text{와 같이 구성된다.}$$

POMDP 모델은 개체가 환경과의 상호작용을 통해 최대 누적 보상을 획득할 수 있는 최적의 정책을 찾는 것을 목표로 한다. 본 연구에서는 복잡도가 높은 연속적인 공간 및 혼재된 행동 공간의 조합에 대한 Q 함수 및 정책 π 을 고려하기 때문에 신경망 학습을 고려한다.

II.2. Maximum Likelihood Estimation 기반 주행 특성 추론 방법

본 연구에서 주행 차량의 주행 특성 파악을 위한 관찰자는 목표 주행 차량 e_i 의 주행 정보를 관측할 수 있다. 구체적으로 관찰자는 t timestep에서 주행 차량의 관측 및 행동 $(o_i, a_i)_t$ 을 관측할 수 있다. 관찰자는 주행 차량의 실제 주행 특성 c^* 를 추론하기 위해 관측 데이터 $(o_i, a_i)_{1:T}$ 의 log-likelihood를 최대화 하는 값을 찾는다. 이때 전체 관측 데이터를 이용하여 MLE 기반한 주행 특성 추론 방식은 다음과 같다.

Theorem 1. 목표 개체 e_i 의 특성 추론 문제

$\hat{c}_i = \arg \max_c \ln P(o_{i,1:T}, a_{i,1:T} | c)$ 는 관찰자에 의해 관측된 행동 o_i 및 추론된 특성 \hat{c}_i 을 기반으로 예측된 행동 $a_i = [a_{i,t}^{acc}, a_{i,t}^{lc}] = \pi(o_{i,t}; \hat{c}_i)$ 과 실제 행동 $a_i^* = [a_{i,t}^{*lc}, a_{i,t}^{*acc}]$ 사이의 Likelihood를 최적화하는

$$\hat{c}_i = \arg \max_c \sum_{t=1}^T \ln \pi(a_{i,t}^{acc} | o_{i,t}; c) + \ln \pi(a_{i,t}^{lc} | o_{i,t}; c) \text{ 문제와 동일하다.}$$

proof.

$$\begin{aligned} \hat{c}_i &= \arg \max_c \ln P(o_{i,1:T}, a_{i,1:T} | c) \\ &= \arg \max_c \ln \int P(s_{1:T}, o_{i,1:T}, a_{i,1:T} | c) ds_{1:T} \quad (1) \end{aligned}$$

수식 (1)은 Markov property에 따라 다음과 같이 전개할 수 있다.

* 교신저자(corresponding author)

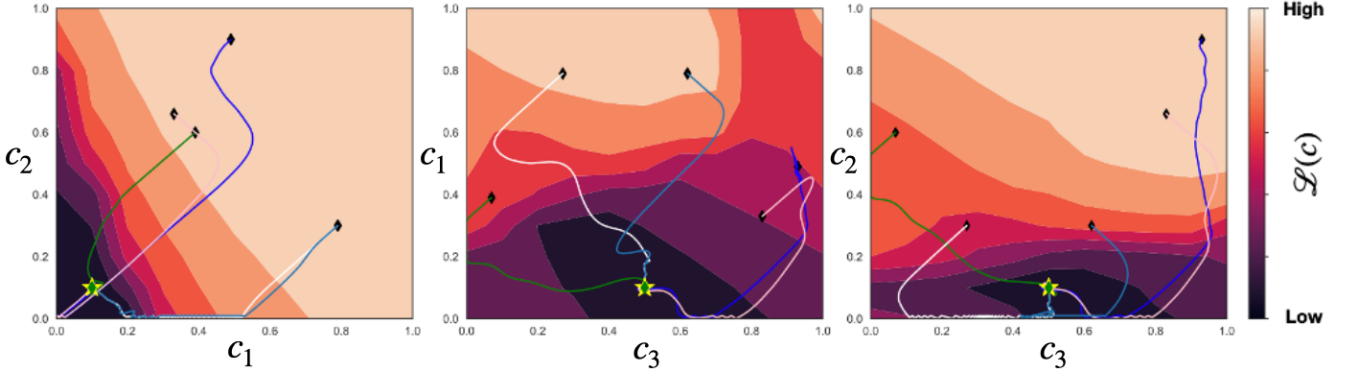


Figure 1. 주행 특성 계수 추론에 대한 손실 등고선 그래프. 검은색 다이아몬드: 초기 추론 특성 계수, 초록색 다이아몬드: 최종 추론 특성 계수, 노란색 별: 실제 주행 특성 계수, 각 실선: 추론 기간에 따른 추론된 특성 계수 변화.

$$\begin{aligned} \hat{c}_i &= \arg \max_c \int P(s_{1:T} | o_{i,1:T}, a_{i,1:T}) [\ln P(s_1) + \sum_{t=1}^T \ln \Omega(o_{i,t} | s_t) \\ &\quad + \sum_{t=1}^T \ln \pi(a_{i,t} | o_{i,t}; c) + \sum_{t=1}^T \ln T(s_{t+1} | s_t, a_t)] ds_{1:T} \quad (2) \\ &= \arg \max_c \sum_{t=1}^T \ln \pi(a_{i,t} | o_{i,t}; c) \times \int P(s_{1:T} | o_{i,1:T}, a_{i,1:T}) ds_{1:T} \\ &= \arg \max_c \sum_{t=1}^T \ln \pi(a_{i,t} | o_{i,t}; c) \quad (3) \end{aligned}$$

수식 (2)에서 $\Omega(o_{i,t} | s_t)$ 와 $T(s_{t+1} | s_t, a_t)$ 는 c 와 관계되어 있지 않기 때문에 무시할 수 있으며, 마찬가지로 $\int P(s_{1:T} | o_{i,1:T}, a_{i,1:T}) ds_{1:T}$ 역시 생략할 수 있다. 본 연구에서는 혼재된 행동 공간에 대해 고려하기 때문에 최종적으로 식 (3)은 다음과 같이 정의할 수 있다.

$$\hat{c}_i = \arg \max_c \sum_{t=1}^T \ln \pi(a_{i,t}^{acc} | o_{i,t}; c) + \ln \pi(a_{i,t}^c | o_{i,t}; c) \quad (4)$$

□

\hat{c}_i 의 최적화는 주행 차량의 주행 특성 계수에 대한 경사 상승법(gradient ascent)를 통해 이루어진다. 주행 특성 계수에 대한 경사 상승법은 $\hat{c}_i \leftarrow \hat{c}_i + \alpha \nabla_c L$ 로 정의되며, 여기서 α 는 학습률(learning rate)을 의미하며, L 은 손실함수로 $L(c) \leftarrow L(c) + (4)$ 로 정의된다.

III. 주행 특성 추론을 위한 모의 실험

본 연구에서 제안한 방법의 검증을 위해 우리는 차량 시뮬레이터인 SUMO를 기반으로 구성된 FLOW[3] 프레임워크를 이용한다. 구체적으로 우리는 다수의 차량이 다차선의 원형도로에서 주행하는 환경을 고려한다. 자율주행차량은 심층강화학습 알고리즘인 TD3[4]를 기반으로 선행 학습되었으며, 주행 특성 계수의 범위는 $c_i \in [0, 1]$ 로 설정하였다. 주행 특성 계수의 설정에 따라 자율주행차량은 공격적(목표 속도의 도달을 위해 다수의 차선 변경을 수행)하거나 방어적(차선 변경을 최소화하고 안전거리를 유지하며 주행)인 주행 특성을 갖고 주행할 수 있다. 주행특성의 추론 과정에서 손실 함수의 역전파 수행을 위해 Adam optimizer를 사용하며, 학습률 $\alpha = 1.0 \times 10^{-3}$ 로 설정하였다.

III.1. 주행 특성 추론

Figure 1은 자율주행차량의 주행 3가지 특성 계수 조합에 따른 손실 값의 등고선을 나타낸 그래프이다. 본 실험을 위해 선행 학습된 자율주행차량의 주행 특성 계수는 $c = [0.1, 0.1, 0.5]$ 로 설정하였다. 해당 그림에서 검은색 다이아몬드는 추론 초기에 무작위로 초기화된 추론 계수 값을 의미하며, 노란색 별은 선행 학습된 차량의 실제 주행 특성 계수를 의미한다. 총 5번의 실험을 수행하였으며 각각의 선분은 무작위로 선택된 초기 특성에서 추를 위한 estimation 과정의 반복을 통해 추론 계수의 변화를 나타낸다. 결과적으로 5번의 실험 모두 실제 주행 특성 계수와 매우 가까운 주행 특성 계수를 추론하는데 성공하였다.

IV. 결론

본 논문에서는 주행 차량의 주행 특성 계수를 추론하기 위한 MLE 및 역강화학습을 기반으로하는 전체 프레임워크를 제안하였다. 제안된 프레임워크의 검증을 위해 강화학습을 기반으로 특성 계수를 고려하여 자율주행차량을 선행 학습시켰다. 이어서 자율주행차량의 행동 및 관측 쌍을 수집하여 이를 설계된 손실함수를 최적화한다. 결과적으로 특성 계수의 초기 값에 관계없이 실제 주행 특성계수를 추론하는 것을 확인하였다. 제안된 추론 방식으로 추론된 특성은 개체의 행동 패턴 및 특정 상태에서의 행동을 예측할 수 있다. 이는 다중 개체 환경에서의 적응적 의사결정을 위한 다양한 연구에 기여할 수 있다.

사 사

이 논문은 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 (No. 2021-0-00739, 분산/협력 AI 기반 5G+ 네트워크 데이터 분석 기능 및 제어 기술 개발) 및 한국연구재단의 지원을 받아 수행된 연구임(No. NRF-2020R1F1A1069182).

참 고 문 헌

- [1] Lee, Dongsu, and Minhae Kwon, "Stability Analysis in Mixed-Autonomous Traffic With Deep Reinforcement Learning," IEEE Transactions on Vehicular Technology, 2022.
- [2] M. Kwon, et al., "Inverse rational control with partially observable continuous nonlinear dynamics," *NeurIPS*, 2020.
- [3] C. Wu, et al., "Flow: Architecture and Benchmarking for Reinforcement Learning inTraffic Control," arXiv preprint arXiv:1710.05465, 2017.
- [4] S. Fujimoto, et al., "Addressing Function Approximation Error in Actor-Critic Methods," ICML, 2018.