Active learning 기반 운전자 행동 모방 학습 기법 연구

황카이스 [1], 문명운 [1], 박지선 [1], 성연식 [1], 조경은 [1]*
[1] 동국대학교 멀티미디어공학과
*e-mail : cke@dongguk.edu (교신저자)

A Study on a Driving Behavior Imitation Learning Method Based on Active Learning

Kaisi Huang[1], Mingyun Wen[1], Jisun Park[1], Yunsick Sung[1], Kyungeun Cho[1]*
[1]Department of Multimedia Engineering, Dongguk University-Seoul

요 약

Simulated driving behavior is an important aspect of realistic simulation systems. To simulate natural driving behavior, this paper proposes an imitation learning method based on active learning that combines demonstration and experience. Driving demonstrations are collected from human drivers in a driving simulator. A driving behavior policy is learned from these demonstrations. The driving demonstration dataset is augmented with new demonstrations that the original demonstrations did not contain, in the form of behaviors from another driving behavior policy learned from experience. The final driving behavior policy is learned from an augmented demonstration dataset.

1. Introduction

Transportation simulation systems are utilized for training novice drivers and predicting traffic problems. Therefore, simulation of driving behavior is essential for developing high-quality, realistic transportation simulation systems, which have the potential to advance research in automotive safety.

The act of driving a vehicle is complex and varied, and requires relatively high levels of skill and experience. One traditional approach for implementing humanlike driver behavior in simulations is the rule-based approach. However, rule-based approaches often result in smooth and collision-free driving, which is not realistic compared to actual human driving behaviors. Therefore, we propose a driving behavior imitation learning method based on active learning. The proposed method can learn driving behavior policy automatically.

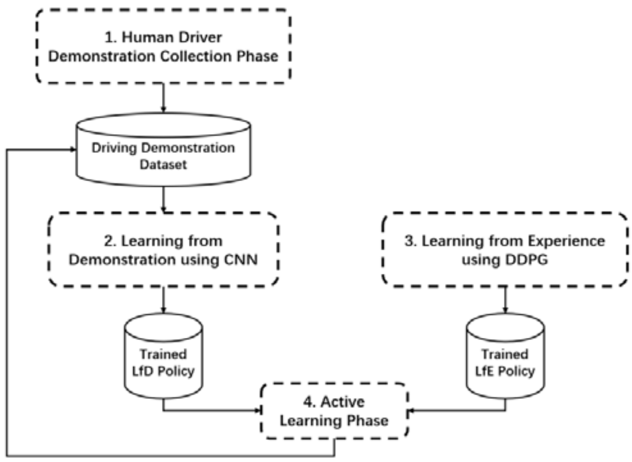
2. Related work

For humanlike behaviors, Imitation Learning (IL) based on human demonstrations can be utilized to learn a policy that behaves similar to experts. Behavioral Cloning (BC) is the simplest approach to imitation learning in which the goal is to learn the relationship between states and demonstration behaviors as supervised learning [1]. However, this approach requires a large amount of training data to guarantee that most of the states that may occur are covered. Another approach is Inverse Reinforcement Learning (IRL), which is the most successful approach to imitation learning [2,3]. This approach assumes that the demonstration behaviors the learners attempt to imitate are generated by experts and attempts to find a reward function that can explain the expert behaviors. However, it can take a significant amount of time to find the reward function.

To solve the aforementioned problems, we propose a method for learning a driving behavior policy that combines Learning from Demonstration (LfD) and Learning from Experience (LfE) via Active Learning (AL).

3. Enhanced imitation learning approach

A diagram of the proposed method is illustrated in Fig. 1. The driving demonstration is obtained from human drivers in a vehicle driving simulator in which it is assumed that the human drivers are experts. Each driving demonstration data point consists of two pieces of data: a top-view image and a behavior. Each behavior consists of information on the steering and acceleration of the vehicle agent. The LfD policy is learned by Convolutional Neural Networks (CNN).



(Figure 1) The flow of the proposed method

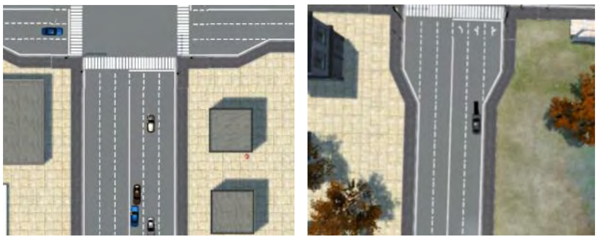
The LfE policy serves as another expert with which to generate driving demonstrations in the AL phase. The LfE policy is learned from experience by a Deep Deterministic Policy Gradient (DDPG). To train the LfE policy, a reward function based on basic driving rules is provided to evaluate the behavior taken by the LfE policy.

In the AL phase, a vehicle agent is controlled by the LfD policy in a virtual environment simulator. The LfD policy provides a probability distribution of possible behaviors by considering the current state. If the highest probability value is lower than a probability threshold, the vehicle agent executes the behavior prescribed by the LfE policy. The current state and behavior are then added to the driving demonstration dataset. The LfD policy is further trained using the augmented driving demonstration.

4. Experiments

For implementation, we would use a virtual environment simulator (Figure 2) that run in Unity3D for validating the proposed method. The neural network was implemented in Python. A vehicle agent in the virtual environment simulator was controlled by driving policy in Python.

To evaluate the proposed method, the results of driven paths by the proposed method would be compared to the result of driven paths by human drivers.



(Figure 2) Virtual environment simulator

5. Conclusion

This paper proposed an imitation learning method to simulate driving behaviors by combining both LfD and LfE via AL. During the LfD phase, CNN is utilized to learn the policy of driving behaviors from the driving demonstration. During the LfE phase, another driving policy is learned by DDPG. During the AL phase, a confidence value is utilized to compute a driving behavior predicted by the LfD policy for a given state. If the confidence value is lower than the threshold, the predicted behavior is provided by the LfE policy, and the state and predicted behavior are added to the demonstration dataset. The LfD policy continues to learn from the augmented driving demonstration dataset.

Acknowledgements

This research was supported by a grant from Defense Acquisition Program Administration and Agency for Defense Development, under contract #UE171095RD.

6. Reference

[1] Bojarski M., Del Testa D., Dworakowski D., Firner B., Flepp B., Goyal P., ..., Zhang X. "End to end learning for self-driving cars." arXiv preprint arXiv:1604.07316, 2016.
[2] Ziebart B. D., Maas A. L., Bagnell J. A., Dey A. K.. "Maximum entropy inverse reinforcement learning." In Aaai, Vol.8, 2008, pp.1433-1438.
[3] Ng, A. Y., Russell S.. "Algorithms for inverse reinforcement learning." In Icml, Vol.1, 2000, pp.2.
[4] Ho J., Gupta J., Ermon S.. "Model-free imitation learning with policy optimization." In International Conference on Machine Learning, 2016, pp.2760-2769.