基于深度强化学习的ACC控制算法设计

目录

[基于深度强化学习的ACC控制算法设计 1](#_Toc21464518)

[第1章. 绪论 2](#_Toc21464519)

[第2章. 基于深度强化学习的ACC控制算法设计 3](#_Toc21464520)

[第3章. 基于逆向深度强化学习的拟人化改进 3](#_Toc21464521)

[第4章. ACC控制算法仿真研究 3](#_Toc21464522)

[第5章. ACC控制算法实验研究 3](#_Toc21464523)

[第6章. 总结与展望 3](#_Toc21464524)

# 绪论

## 课题背景和意义

自动驾驶技术能够减少驾驶员操作失误，减少因此导致的事故，使汽车驾驶更加方便和安全[[1]](#endnote-1)。然而，考虑到完全实现自动驾驶之前我们可能面临长达数十年的过渡期，在完全实现自动驾驶之前自动驾驶车辆与人类驾驶的车辆将共用道路，因此自动驾驶的一个重要挑战是在安全驾驶的前提下模仿人类驾驶风格，实现拟人化驾驶[[2]](#endnote-2)[[3]](#endnote-3)[[4]](#endnote-4)。拟人化驾驶将（1）为乘客提供舒适的驾驶体验，提升乘客对自动驾驶汽车的信任，以及（2）使周围的人类驾驶员能够更好地理解和预测自动驾驶车辆的行为，以便他们能够自然地与之交互1[[5]](#endnote-5)。

在算法设计中拟人化驾驶可以分为两个部分：决策拟人化和操作拟人化。决策拟人化即自动驾驶算法可以在不同的外部环境下模仿人类选择对应的决策策略，在本文中决策拟人化的决策目标为自车和前车的相对距离变化量；操作拟人化是指自动驾驶算法为了完成决策目标而在不同外部环境下模仿人类执行合理的动作，在本文中算法执行的动作为自车加速度的选择。总体而言设计拟人化驾驶算法的主要任务是使其根据外部环境条件选择与人类决策结果最接近的动作。而在强化学习算法中智能体就是通过与外部环境交互的方法提升决策能力，因此本文使用强化学习算法作为拟人化ACC控制算法的决策算法。

2015年谷歌DeepMind团队提出了深度强化学习算法，其将强化学习算法与深度神经网络结合起来，可以设计出能够在复杂情况下智能决策的算法[[6]](#endnote-6)。与传统跟车模型相比，深度强化学习有诸多优点：（1）深度神经网络擅长通用函数逼近[[7]](#endnote-7)，可以通过训练快速得出与传统跟车模型性能相近的神经网络;（2）因为深度强化学习算法是从训练数据中学习决策机制，而不是通过对数据进行参数估计[[8]](#endnote-8)，因此强化学习可以实现更好的泛化性能;（3）通过不断学习已有的驾驶数据，深度强化学习可以使汽车根据外部具体环境评估采取各种动作的价值选择驾驶操作，达到拟人化的效果。

深度强化学习在游戏中表现良好[[9]](#endnote-9)[[10]](#endnote-10)[[11]](#endnote-11)是因为游戏设计者们已经为游戏设计了一套合理的奖励策略，但是在解决实际问题时奖励函数的设计较为困难，为此研究者们提出了不同的解决方案。虽然这些奖励函数在具体问题中均取得了较为理想的控制效果，但是奖励函数的设计存在三个问题：

1. 奖励函数的设计较为主观，函数类型和相关参数的选择缺乏理论依据；
2. 一个表现良好的奖励函数往往需要需要领域内经验丰富的专家来设计；
3. 奖励函数需要大量的手动调试工作才能实现较好的控制效果。

基于这一问题，为了完成拟人化的ACC控制算法设计，本文提出了一种基于分层强化学习的ACC控制算法，并基于真实跟驰驾驶数据和逆向强化学习算法设计了ACC控制算法的奖励函数，解决了拟人化ACC的算法结构设计和统用奖励函数设计连个问题。

吉林大学朱冰[[12]](#endnote-12)等人在使用深度强化学习方法解决车辆跟驰控制问题中将奖励函数定义为以下形式：

其中为跟踪性指标，

表示当前间距和期望间距之差，表示车距权重系数。

为舒适性指标，

表示k时刻的加速度，为冲击度权重系数。

为燃油经济性系数，

Q为控制变量的权重系数。

为安全性系数，

当车速超过预定车速或间距小于理想间距有时为-1，否则为0.

Zhu M，Wang X等人[[13]](#endnote-13)在研究拟人化纵向控制算法时将奖励函数定义为以下形式：

其中表示速度跟随误差，

表示驾驶平稳性，u表示加速度

虽然上述研究均实现了较为理想的控制效果，但是奖励函数的设计存在三个问题：

1. 奖励函数的设计较为主观，函数类型和相关参数的选择缺乏理论依据；
2. 一个表现良好的奖励函数往往需要需要领域内经验丰富的专家来设计；
3. 奖励函数需要大量的调试工作才能实现较好的控制效果。

针对这个问题本文借鉴Markus Wulfmeier等人[[14]](#endnote-14)提出的使用神经网络拟合奖励函数的思想，基于深度神经网络和真实跟驰驾驶数据为深度强化学习算法拟合了拟人化驾驶的奖励函数。本文使用的真实跟驰驾驶数据是中国智能网联汽车研究院在中国10个代表性的城市采集真实交通场景并后处理得到的。

## 深度强化学习研究现状

## 拟人化ACC控制算法研究现状

## 本文主要研究内容

# 基于深度强化学习的ACC控制算法设计

# 基于逆向深度强化学习的拟人化改进

# ACC控制算法仿真研究

# ACC控制算法实验研究

# 总结与展望

1. Wei J, Dolan J M, Litkouhi B. A learning-based autonomous driver: emulate human driver's intelligence in low-speed car following[C]//Unattended Ground, Sea, and Air Sensor Technologies and Applications XII. International Society for Optics and Photonics, 2010, 7693: 76930L. [↑](#endnote-ref-1)
2. Gu T, Dolan J M. Toward human-like motion planning in urban environments[C]//2014 IEEE Intelligent Vehicles Symposium Proceedings. IEEE, 2014: 350-355. [↑](#endnote-ref-2)
3. Lefevre S, Carvalho A, Borrelli F. Autonomous car following: A learning-based approach[C]//2015 IEEE Intelligent Vehicles Symposium (IV). IEEE, 2015: 920-926. [↑](#endnote-ref-3)
4. Kuderer M, Gulati S, Burgard W. Learning driving styles for autonomous vehicles from demonstration[C]//2015 IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2015: 2641-2646. [↑](#endnote-ref-4)
5. Wang X, Zhu M, Chen M, et al. Drivers’ rear end collision avoidance behaviors under different levels of situational urgency[J]. Transportation research part C: emerging technologies, 2016, 71: 419-433. [↑](#endnote-ref-5)
6. Hausknecht M, Stone P. Deep reinforcement learning in parameterized action space[J]. arXiv preprint arXiv:1511.04143, 2015. [↑](#endnote-ref-6)
7. Hornik K, Stinchcombe M, White H. Multilayer feedforward networks are universal approximators[J]. Neural networks, 1989, 2(5): 359-366. [↑](#endnote-ref-7)
8. Mnih V, Kavukcuoglu K, Silver D, et al. Human-level control through deep reinforcement learning[J]. Nature, 2015, 518(7540): 529. [↑](#endnote-ref-8)
9. Playing Atari with Deep Reinforcement Learning [↑](#endnote-ref-9)
10. Human-level control through deep reinforcement learning [↑](#endnote-ref-10)
11. On Reinforcement Learning for Full-length Game of StarCraft [↑](#endnote-ref-11)
12. 朱冰, 蒋渊德, 赵健, 等. 基于深度强化学习的车辆跟驰控制[J]. 中国公路学报, 32(6): 53-60. [↑](#endnote-ref-12)
13. Zhu M, Wang X, Wang Y. Human-like autonomous car-following model with deep reinforcement learning[J]. Transportation research part C: emerging technologies, 2018, 97: 348-368. [↑](#endnote-ref-13)
14. Deep Inverse Reinforcement Learning [↑](#endnote-ref-14)