ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/261044713

Adaptive optimal control for the uncertain driving habit problem in adaptive cruise control system

Conference Paper · July 2013

DOI: 10.1109/ICVES.2013.6619622

CITATIONS
13

READS
206

2 authors, including:

Dongbin Zhao
Chinese Academy of Sciences

351 PUBLICATIONS **5,478** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Project StarCraft View project

All content following this page was uploaded by Dongbin Zhao on 09 December 2016.

The user has requested enhancement of the downloaded file.

Adaptive Optimal Control for the Uncertain Driving Habit Problem in Adaptive Cruise Control System

ZHAO Dongbin, XIA Zhongpu
The State Key Laboratory of Management and Control of Complex Systems,
Institute of Automation, Chinese Academy of Science,
Beijing 100190, CHINA
dongbin.zhao@ia.ac.cn, zhongpu.xia@gmail.com

***Abstract*—In this paper, a novel adaptive optimal control approach based on Q-function is proposed to address the problem as the driving habits change among drivers and over time in the adaptive cruise control system. The proposed approach, adopting the special structure Q-function of the linear discrete-time system, uses policy iteration method to derive the optimal control policy online. It repeats between policy evaluation where the polynomial neural network is employed to approximate the cost function of the system and policy improvement where the control policy is updated based on the converged neural network, until the optimal controller is achieved. Simulation is conducted and results show the effectiveness for uncertain driving habit problem in the adaptive cruise control system.**

Keywords—adaptive cruise control; driving habit; Q-function; adaptive optimal control; policy iteration

I. INTRODUCTION

Advanced driver assistance systems (ADAS) have been active topics of research and development since the 1990s owing to the potential for increasing vehicle safety and improving driving comfort [1], such as adaptive cruise control (ACC) [2, 3], stop & Go [4, 5], forward collision warning/avoidance [6]. As it is capable of assisting or replacing a driver to regulate the throttle or the brake for maintaining the vehicle from the one ahead in a safe distance and velocity, ACC is taken as the vital stage to the driverless technique and thus widely researched.

Lots of methodologies have been adopted to improve safety, enhance comfort, reduce driver's labour and save fuel within ACC systems. Proportion-Integration-Differentiation type control [7], sliding mode control [8] and linear quadratic optimal control [1, 9] performed well in improving the safety and comfort. In order to satisfy the requirements of human drivers, intelligent control laws are employed to construct ACC systems. Fuzzy control [10, 11] based on the skilled driver was applied to the ACC system to achieve a humanlike driving. Moon et. [12] categorized driving situation into safe, warning and dangerous modes using warning index and the time-to-collision, then the different control strategy was adopted depending on these modes and hence a same performance as manual-driving was obtained. Model predictive control was used to design the ACC system in [13-15] with the objectives of comfort, fuel-economy, safety and car-following.

This work is supported by National Natural Science Foundation of China (NSFC) under Grants No. 61273136, No. 61034002, and Beijing National Science Foundation under grant No. 4122083.

As driving habits change among drivers and over time in the ACC system, an intelligent ACC system should adapt to different driving habits. Otherwise a driver would intervene even in situations that ACC is able to manage, typically when it could nor meet the driver's expectations [16]. As so, an ACC system which takes the individual driving habit into consideration would contribute to its attraction and put into commercial production [16]. All the above mentioned ACC systems were designed according to the common driving habit model, so that it is difficult for them to be put into application. The online learning method is thought as a good solution to such problem and much work has been done.

A machine learning approach based on neural networks was adopted in [17, 18], where the networks were trained by current driver continuously to accommodate the driving habits as the changes among drivers and over time. In [19] the self-learning ACC system was proposed, where the parameters of the driving habit model were identified from the data in the manual operation phase, and the identified results were applied in the automatic control phase. This self-learning ACC system was tested in the real traffic environments and results showed that the acceptability was improved to some extent.

Reinforcement learning method, such as supervised adaptive dynamic programming [20] and supervised reinforcement learning [21], has been adopted to address such problem without prior knowledge of the system. However, the computation required by this method was too large to implement in real time. This paper adopts the adaptive optimal control approach based on special structure Q-function to address the problem in real time, which uses policy iteration to derive the optimal controller online by using the observed data rather than accessing to the model of the system. All required is to provide an initial stable controller, which can be acquired easily in the design phase. It is capable of searching the optimal control solution for systems whose parameters are uncertain or drift over time. Hence the proposed approach would be a perfect solution to the uncertain driving habit problem in ACC system.

The paper is organized as follows. The ACC model and adaptive optimal control based on Q-function are given in Section II. Section III implements the proposed approach in the ACC system. The performances of the presented approach including convergence, optimization ability and adaptability for different driving habits are verified in Section IV. The conclusion is given in Section V.

II. ACC MODEL AND OPTIMAL CONTROL

A. ACC Model

The ACC system structure is shown in Fig. 1. Here the vehicle mounted with the ACC system is named host vehicle and the one ahead in the same lane is called target vehicle. The ACC system is the part in the dot-dashed-box of Fig. 1. It detects the ambient information of the host vehicle, including the velocity and the acceleration of the host vehicle v_h and a_h respectively, the distance to the target vehicle d and the velocity of the target vehicle v_t. According to these data, the controller regulates the position of the throttle or brake in order to track the target vehicle in a safe distance and velocity.

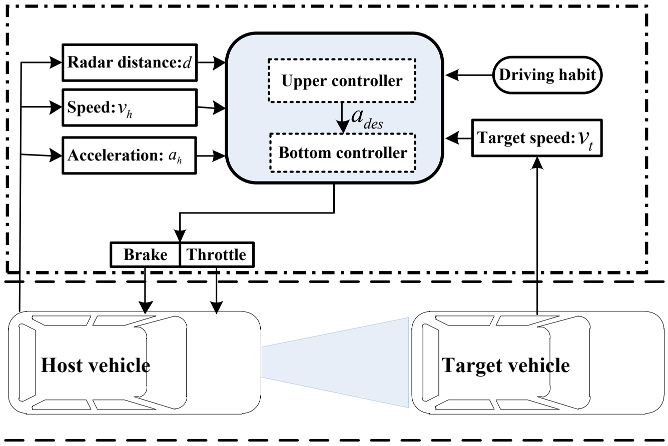

Fig. 1 The structure of the ACC system

The controller of the ACC system is always designed by the way of dividing it into two layers: the upper controller and the bottom controller [1, 22]. The upper controller generates the desired acceleration control signal according to the detected information, while the objective of the bottom controller is to transfer the desired acceleration signal a_{des} into the actual acceleration by coordinating the position of the throttle or the brake. In this paper, only the upper controller is considered while the bottom controller and the host vehicle are modeled together as a first-order system with a time constant τ_b as follows [23]

$$\tau_b \frac{da_h(t)}{dt} + a_h(t) = a_{des}(t) \quad . \tag{1}$$

As discussed above, driving habits change among drivers and over time: an intelligent ACC controller should learn the driving habits and react like the driver. Driving habits can be modeled as the function of the host velocity, the desired distance between the motionless host and target vehicle d_0 and the headway time τ_h [5, 22]

$$d_{des} = d_0 + v_h(t)\tau_h \quad . \tag{2}$$

It comes out that the headway time is long for conservative drivers, and short for sportive drivers.

According to the dynamics (1) and driving habit model (2), the ACC system can be modeled by the state space, given by

$$\begin{aligned}\dot{\mathbf{x}} &= \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u} + \mathbf{E}w \\ &= \begin{bmatrix} 0 & 1 & \tau_h \\ 0 & 0 & 1 \\ 0 & 0 & -1/\tau_b \end{bmatrix}\mathbf{x} + \begin{bmatrix} 0 \\ 0 \\ 1/\tau_b \end{bmatrix} a_{des} - \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} a_t \end{aligned} \tag{3}$$

where the states $\mathbf{x} = [d_{des} - d, v_h - v_t, a_h]^T$ are the error between the desired distance d_{des} and the detected distance d, the error between the host velocity v_h and target velocity v_t, and the host acceleration a_h, the input $\mathbf{u} = a_{des}$ is the desired acceleration, and the disturbance $w = a_t$ is the acceleration of the target vehicle.

As the controller is designed in the form of discrete time, here the continuous ACC system is converted into a discrete time system by zero-order-hold method with the sampling time T:

$$\mathbf{x}_{k+1} = \mathbf{G}\mathbf{x}_k + \mathbf{H}\mathbf{u}_k + \mathbf{L}w_k \quad . \tag{4}$$

B. Adaptive Optimal Control Based on Q-function

The system dynamics (4) converts the ACC tracking control problem into a regulator problem, the objective of the controller is to regulate the states of the system to the equilibrium states $\mathbf{x} = [0, 0, 0]^{\mathrm{T}}$. In order to regulate the system in an optimal manner, a sequential control action $\mathbf{u}^*(\mathbf{x}_k),\ k \in [0, \infty]$ is selected to minimize the infinite horizon cost function as follows

$$J(\mathbf{x}_k) = \sum_{j=0}^{\infty} (\mathbf{x}_{k+j}^{\mathrm{T}} \mathbf{Q} \mathbf{x}_{k+j} + \mathbf{u}_{k+j}^{\mathrm{T}} \mathbf{R} \mathbf{u}_{k+j}) \tag{5}$$

where $\mathbf{Q}, \mathbf{R}$ are the symmetric positive definite matrix, $\mathbf{x}_k$ is the state vector at the time step k after a sequence of actions. Based on the Bellman's principle of optimality, the optimal cost function $J^*(\mathbf{x}_k)$ is time invariant and satisfies the discrete time Bellman function [24]

$$J^*(\mathbf{x}_k) = \min_{\mathbf{u}_k}(r(\mathbf{x}_k, \mathbf{u}_k) + J^*(\mathbf{x}_{k+1})) \tag{6}$$

where $r(\mathbf{x}_k, \mathbf{u}_k) = \mathbf{x}_k^{\mathrm{T}} \mathbf{Q} \mathbf{x}_k + \mathbf{u}_k^{\mathrm{T}} \mathbf{R} \mathbf{u}_k$ is the reward for the system.

For linear discrete-time system, Q-function proposed by Watkins [25] is a cost function with respect to both state and action for a stable control policy **K**, which is also known as state-action function differing from the Bellman function also known as state function. The Q-function Bellman equation (6) is

$$Q(\mathbf{x}_k,\mathbf{u}_k)=r(\mathbf{x}_k,\mathbf{u}_k)+J(\mathbf{x}_{k+1})\,. \tag{7}$$

The value of $Q(\mathbf{x}_k,\mathbf{u}_k)$ represents the cost of taking any admissible action $\mathbf{u}_k$ at state $\mathbf{x}_k$. $J(\mathbf{x}_{k+1})$ is the cost at state $\mathbf{x}_{k+1}$.

By adopting the Q-function, more information is implied in the cost function than the Bellman function. It contains the cost for all actions with respect to each state, then the greedy policy can be used to improve the control policy as

$$\mathbf{u}_k=\mathbf{K}'\mathbf{x}_k=\arg\min_{\mathbf{u}_k}(r(\mathbf{x}_k,\mathbf{u}_k)+Q(\mathbf{x}_{k+1},\mathbf{K}\mathbf{x}_{k+1}))\ . \tag{8}$$

$\mathbf{K}'$ represents the control policy after improved here.

III. ADAPTIVE OPTIMAL CONTROL FOR ACC

The parameters τ_h , τ_b and d_0 change among drivers and over time, therefore the ACC system is a time varying linear system. Optimal control is available for the time-invariant linear system. However, the accurate model of the system must be known in advance. The class of technique called adaptive control is available for dealing with such problem with uncertain parameters. Unfortunately adaptive control is not optimal in a formal sense, only minimizing a cost function of the output error [26]. Adaptive optimal control based on Q-function would be a perfect solution to optimal control for time varying ACC systems as it can improve the control policy by interacting with the environment, then the optimal control policy can be yielded as long as all state-action space is visited. As the upper model of the ACC system can be described as a linear discrete-time system, the Q-function can be represented by an explicit form, which will contribute to the speed of converging and optimizing.

Here the neural network is employed to approximate the Q-function with

$$Q(\mathbf{x}_k,\mathbf{u}_k)=\mathbf{w}^{\mathrm{T}}\boldsymbol{\varphi}(\mathbf{x}_k,\mathbf{u}_k) \tag{9}$$

where the $\boldsymbol{\varphi}(\mathbf{x}_k,\mathbf{u}_k)$ is the bias function and $\mathbf{w}$ is the parameter vector. As the cost function of the linear quadratic regulator problem is the quadratic polynomial function with respect to the states and the actions [27], the neural network can approximate the Q-function well if the basis function is chosen as $\boldsymbol{\varphi}(\mathbf{x},\mathbf{u})=\left[x_1^2,x_1x_2,x_1x_3,x_1u,x_2^2,x_2x_3,x_2u,x_3^2,x_3u,u^2\right]^{\mathrm{T}}$ and $\mathbf{w}=\left[w_1,w_2,w_3,w_4,w_5,w_6,w_7,w_8,w_9,w_{10}\right]^{\mathrm{T}}$.

To find the greedy control policy, we take the partial deviation of Q-function with respect to $\mathbf{u}$, then $\mathbf{u}$ is derived as

$$\mathbf{u}=-\frac{1}{2}\frac{w_4x_1+w_7x_2+w_9x_3}{w_{10}}=\mathbf{K}\mathbf{x}\ . \tag{10}$$

And the control policy is

$$\mathbf{K}=-\frac{1}{2w_{10}}\begin{bmatrix}w_4 & w_7 & w_9\end{bmatrix}\ . \tag{11}$$

In order to achieve the optimal solution online, here the control policy **K** is yielded by adopting the policy iteration method which is implemented in two steps: policy evaluation and policy improvement. In the policy evaluation step, the cost function associated with the current stable controller iterates until **w** converges. In the policy improvement step, the policy with the minimal cost is determined based on (11). In this algorithm, all required is an initial stable control policy. The policy iteration method for linear discrete-time system is as the following:

- Policy evaluation:

$$\mathbf{w}_i^{\mathrm{T}}\boldsymbol{\varphi}(\mathbf{x}_k,\mathbf{u}_k)=r(\mathbf{x}_k,\mathbf{u}_k)+\mathbf{w}_i^{\mathrm{T}}\boldsymbol{\varphi}(\mathbf{x}_{k+1},\mathbf{u}_{k+1})\ . \tag{12}$$

- Policy improvement:

$$\mathbf{K}_{i+1}=-\frac{1}{2\mathbf{w}_i(10)}\begin{bmatrix}\mathbf{w}_i(4) & \mathbf{w}_i(7) & \mathbf{w}_i(9)\end{bmatrix}. \tag{13}$$

where the subscript i represents the i-th iteration for control policy, in order to differ from k which represents the k-th sampling step for the system. The algorithm alternates between policy evaluation and policy improvement steps until an update of the control policy will no longer improve the system performance.

It is seen that the policy iteration algorithm (12) (13) only utilizes the observed data such as state, action and reward from environment along the system trajectories, rather than the accurate model of the system. Hence the improvement of the control policy follows the change of the ACC system. So it will be a perfect solution to the different driving habits problem, where only an initial stable controller is required.

The implementation of an online ACC system controller in real time will be given below. The controller implemented with the algorithm is named online learning controller throughout the paper.

In order to find the parameters of the cost function, (12) is rewritten as

$$\mathbf{w}_i^{\mathrm{T}}(\boldsymbol{\varphi}(\mathbf{x}_k,\mathbf{u}_k)-\boldsymbol{\varphi}(\mathbf{x}_{k+1},\mathbf{u}_{k+1}))=r(\mathbf{x}_k,\mathbf{u}_k)\ . \tag{14}$$

So the parameters of the cost function can be derived by solving (14), if the states and the reward along the trajectory of the ACC system are obtained. In order to derive the value of

$\mathbf{w}_i$, making the error between the left side and right side of (14) in the least mean square senses, least-mean squares method is employed:

$$\mathbf{w}_i = (\mathbf{\Psi}\mathbf{\Psi}^{\mathrm{T}})^{-1}\mathbf{\Psi}\mathbf{Y} \tag{15}$$

where

$$\begin{aligned}\mathbf{\Psi} &= [\nabla\boldsymbol{\varphi}_{k-N}, \nabla\boldsymbol{\varphi}_{k-N+1}, \cdots, \nabla\boldsymbol{\varphi}_k] \\ \nabla\boldsymbol{\varphi}_k &= \boldsymbol{\varphi}(\mathbf{x}_k, \mathbf{u}_k) - \boldsymbol{\varphi}(\mathbf{x}_{k+1}, \mathbf{u}_{k+1}) \\ \mathbf{Y} &= [r(\mathbf{x}_{k-N}, \mathbf{u}_{k-N}), r(\mathbf{x}_{k-N+1}, \mathbf{u}_{k-N+1}), \cdots, r(\mathbf{x}_k, \mathbf{u}_k)]^{\mathrm{T}}\end{aligned}.$$

The least-mean square problem can be solved in real time if sufficient number of data points are collected $N \geq |\mathbf{w}|$ (symbol $|\cdot|$ is defined as the number of elements in a vector). Then the control policy updates according to (14).

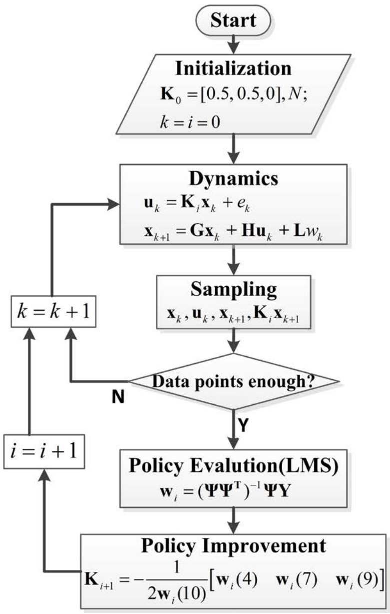

Fig. 2. Adaptive optimal control algorithm for ACC system

It is very important that the state-action space must be sufficiently explored for policy improvement. Otherwise, it may fall into the local optimal control policy or even cannot improve the control policy any more. Here noise from the normal distribution with zero mean is added to the action for the exploration.

$$\mathbf{u}_k = \mathbf{K}_i\mathbf{x}_k + e_k \tag{16}$$

where e_k acts as the exploration component of the action sign.

So far, the adaptive optimal control algorithm based on Q-function developed for the ACC system is summarized in the flowchart in Fig. 2.

IV. SIMULATION RESULT

The function of the ACC system is to adjust the acceleration of the host vehicle in order to keep a safe distance from the target vehicle, and adapt to different driving habits as well as possible. Simulation is conducted to show the effectiveness of the presented approach, including convergence, the ability of optimization and the adaptability for different driving habits.

In the simulation, the time constant for the bottom controller is selected as $\tau_b = 0.45s$ [23]. The parameters of the cost function are selected as $\mathbf{Q}$ = [0.8 0 0; 0 1 0; 0 0 0], $\mathbf{R}$ = 1. The initial stable control gain is selected as $\mathbf{K}_{\mathrm{init}}$ = [0.5, 0.5, 0]. The sampling time for the system is T = 0.05*s*. The parameters of driving habit are listed Table I:

Table I. DRIVING HABIT MODELS [5]

Human Driver	*Headway time* [*s*]	*Motionless Clearance* [*m*]
Driver 1	1.70	1.64
Driver 2	1.25	4.30
Driver 3	0.67	2.25

In order to show the performance of the adaptive optimal control algorithm for ACC systems, the simulation is separated into a learning phase where the acceleration of the target vehicle is generated randomly and a testing phase where the target acceleration is determined:

- **Learning phase:** the initial states of the host vehicle and the target one include the distance between the two vehicles d = 50*m*, the velocity of the host vehicle v_h = 72*km/h* , the velocity of the target vehicle v_t = 90*km/h*, the acceleration of the host a_h = 0m/s^2. The amplitude of the target acceleration in one period is generated by the uniform distribution on the interval [-1, 1]. Moreover, each period lasts for random steps from the integer uniform distribution [20, 40]. This period repeats in the learning phase which continues 800 steps. One of the experiment results through the learning phase is shown in Fig. 3.
- **Testing phase:** The initial states of the two vehicles are the same as the learning phase, but the acceleration of the target vehicle is determined for comparison. It is selected as 0 m/s^2 in the interval [0, 20] *s*, 0.5 m/s^2 in the interval (20, 25] *s*, 0 m/s^2 in the interval (25, 35] *s*, 0.5 m/s^2 in the interval (20, 25] *s*, 0 m/s^2 in the interval (25, 35] *s*. This whole phase lasts for 40*s*, that is 800 steps.

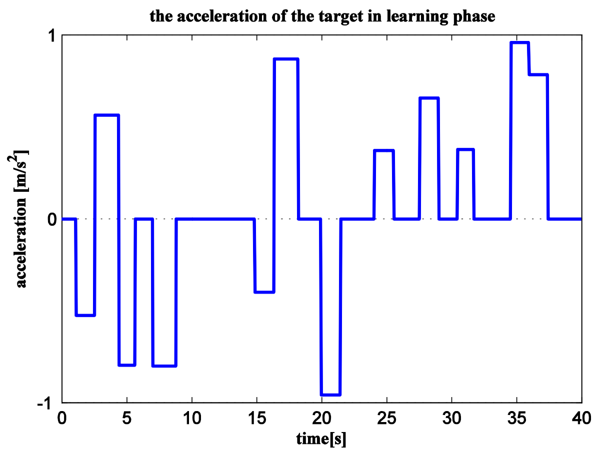

Fig. 3. The acceleration of the target vehicle generated randomly

In the learning phase, the online learning controller is improved by the policy iteration algorithm. In order to test the optimization ability of the algorithm, at the time 20s the driver changes from Driver 1 to Driver 3, and the time delay τ_b changes from 0.45s to 0.3s, too. Furthermore the LQR controller $\mathbf{K}_{\mathrm{lqr}} = [\ 0.8547,\ \ 1.0169,\ \ 0.7996]$ is obtained for Driver 1 based on the ACC model (4) without noises. The simulation results including the trajectories of the ACC system, the evolution of the control policy and the changes of the cost function over the learning phase are shown in Fig. 4, Fig. 5, Fig. 6, respectively.

From Fig. 4, it is shown that the host vehicle is able to track the target vehicle very well and keeps at a desired distance. It also can be seen that the desired distance curve changes abruptly at the time 20s as a result of the driving habit changing. However the host vehicle could track the target one successfully. At the same time, the control policy adapts to the new system, which is shown in Fig. 5.

The control policy improves as more and more observed data acquired and the parameters converge after the time 5s. The online controller search for optimal control policy again after the change of the driving habits and then converges at the time 25s. These are shown in Fig. 5. By the way, the marker in the figure represents the time of control policy updating.

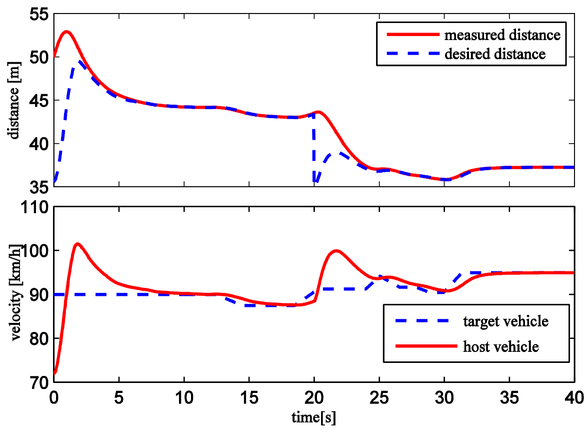

Fig. 4. The trajectories of the ACC system

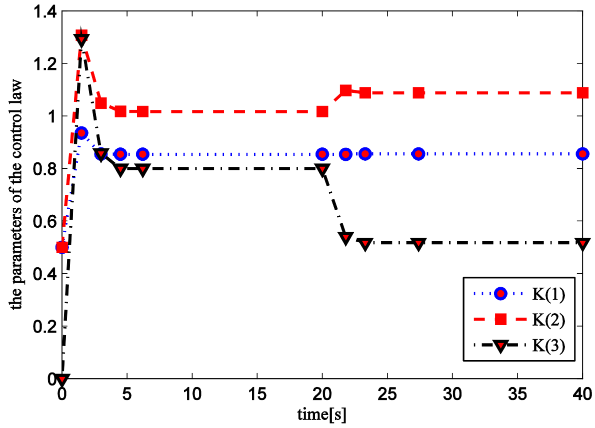

Fig. 5. The evolution of the online control policy

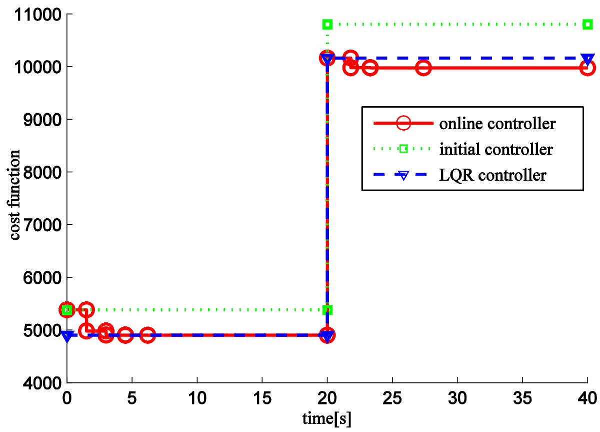

Fig. 6. The cost function for each control policy (cost function is calculated by accumulating the reward from the time 0s to 40s over the testing phase)

To evaluate the performance of the control law, the cost function (5) is employed. A comparison with the initial controller and LQR controller is considered where the cost function is calculated by accumulating the reward from the time 0s to 40s over the testing phase. The evolution of the cost is shown in Fig. 6. The markers in the figure stand for that the control policy is update and new cost function is calculated. The cost of online controller decreases as the process proceeds and converges to the cost of the optimal LQR controller before the changes, which means that the online controller improved to the optimal controller. However, the cost curves change abruptly as the ACC system changes at the time 20s. As the LQR controller and the initial controller are derived offline, which implies that they are fixed over the learning phase, the cost function of them does not change if no changes happen to the dynamics of the system. But the online controller improves in real time, it continues to explore the optimal control policy along the trajectories. So the online controller yields a lower cost than the LQR controller designed offline after the time 20s.

V. CONCLUSION

In this paper, we propose the Q-function based adaptive optimal control which adopts policy iteration to address the problem as the driving habits change among drivers and over time in the adaptive cruise control system. The proposed approach is model-free and can be implemented online, only an initial stable control policy is required in advance. It is applied to an ACC system with different driving habits in the simulation, and the results confirm the effectiveness of the presented approach, including convergence, the ability of optimization and the adaptability for different driving habits.

This paper has provided a solution to the upper controller of the ACC system. In further work we will build a full ACC system, which contains the upper controller, the bottom controller and the vehicle model. At the same time detected data with noise need be taken into consideration.

REFERENCES

[1] K. Yi, S. Lee, and Y.D. Kwon, “An investigation of intelligent cruise control laws for passenger vehicles,” Proc. Inst. Mech. Eng. Part D-J. Automob. Eng., 2001, pp. 159-169.

[2] P.A. Ioannou, and C.C. Chien, “Autonomous intelligent cruise control,” IEEE Trans. Veh. Technol., Vol. 42(4), pp. 657-672, 1993.

[3] L. Xiao, and F. Gao, “A comprehensive review of the development of adaptive cruise control systems,” Veh. Syst. Dyn., Vol. 48(10), pp. 1167-1192, 2010.

[4] M. Persson, F. Botling, E. Hesslow, and R. Johansson, “Stop and go controller for adaptive cruise control,” Proceedings of the IEEE International Conference on Control Applications, 1999, pp. 1692-1697.

[5] Y. Kyongsu, and M. Il-Ki, “A driver-adaptive stop-and-go Cruise control strategy,” IEEE International Conference on Networking, Sensing and Control, 2004, pp. 601-606.

[6] A. Vahidi, and A. Eskandarian, “Research advances in intelligent collision avoidance and adaptive cruise control,” IEEE Trans. Intell. Transp. Syst., Vol. 4(3), pp. 143-153, 2003.

[7] C.C. Chien, P. Ioannou, and M.C. Lai, “Entrainment and vehicle following controllers design for autonomous intelligent vehicles,” American Control Conference, 1994, pp. 6-10 vol.11.

[8] S.-B. Choi, and P. Devlin, “Throttle and brake combined control for intelligent vehicle highway systems,” Systems and issues in ITS, Vol. 12, pp. 53-60, 1995.

[9] C.Y. Liang, and H. Peng, “Optimal adaptive cruise control with guaranteed string stability,” Veh. Syst. Dyn., Vol. 32(4-5), pp. 313-330, 1999.

[10] J.E. Naranjo, C. Gonzalez, R. Garcia, and T. de Pedro, “ACC plus Stop&Go maneuvers with throttle and brake fuzzy control,” IEEE Trans. Intell. Transp. Syst., Vol. 7(2), pp. 213-225, 2006.

[11] J.E. Naranjo, C. Gonzalez, R. Garcia, and T. de Pedro, “Cooperative throttle and brake fuzzy control for ACC plus Stop&Go maneuvers,” IEEE Trans. Veh. Technol., Vol. 56(4), pp. 1623-1630, 2007.

[12] S. Moon, I. Moon, and K. Yi, “Design, tuning, and evaluation of a full-range adaptive cruise control system with collision avoidance,” Control Eng. Practice, Vol. 17(4), pp. 442-455, 2009.

[13] G.L. Naus, J. Ploeg, M.J.G. Molengraft, W.P.M.H. Heemels, and M. Steinbuch, A Model Predictive Control Approach to Design a Parameterized Adaptive Cruise Control, in Re, L., Allgöwer, F., Glielmo, L., Guardiola, C., and Kolmanovsky, I. (Eds.): ‘Automotive Model Predictive Control’ (Springer London, 2010), pp. 273-284.

[14] P. Shakouri, A. Ordys, and M.R. Askari, “Adaptive cruise control with stop&go function using the state-dependent nonlinear model predictive control approach,” ISA transactions, Vol. 51(5), pp. 622-631, 2012.

[15] L. Luo, H. Liu, P. Li, and H. Wang, “Model predictive control for adaptive cruise control with multi-objectives: comfort, fuel-economy, safety and car-following,” J. Zhejiang Univ. Sci. A, Vol. 11(3), pp. 191-201, 2010.

[16] P. Zheng, and M. McDonald, “Manual vs. adaptive cruise control – Can driver’s expectation be matched?,” Transportation Research Part C: Emerging Technologies, Vol. 13(5–6), pp. 421-431, 2005.

[17] F. Simonelli, G. Bifulco, V. De Martinis, and V. Punzo, “Human-like adaptive cruise control systems through a learning machine approach,” Applications of Soft Computing, Vol. 52, pp. 240-249, 2009.

[18] G.N. Bifulco, F. Simonelli, and R. Di Pace, “Experiments toward an human-like adaptive cruise control,” IEEE Intelligent Vehicles Symposium, 2008, pp. 919-924.

[19] J. Wang, L. Zhang, D. Zhang, and K. Li, “An adaptive longitudinal driving assistance system based on driver characteristics,” IEEE Transactions on Intelligent Transportation Systems, Vol. 14(1), pp. 1-12, 2013.

[20] D. Zhao, and Z. Hu, “Supervised adaptive dynamic programming based adaptive cruise control,” IEEE Symposium on Adaptive Dynamic Programming And Reinforcement Learning, 2011, pp. 318-323.

[21] D. Zhao, and Z. Hu, “Supervised reinforcement learning for adaptive cruise control,” Proceedings of the 4th International Symposiumon Computational Intelligence and Industrial Application, 2010, pp. 239-248.

[22] D. Zhao, Z. Hu, Z. Xia, C. Alippi, Y. Zhu, and D. Wang, “Full-range adaptive cruise control based on supervised adaptive dynamic programming,” Neurocomputing, http://dx.doi.org/10.1016/j.neucom.2012.09.034.

[23] D.Z. Zhang, K.Q. Li, and J.Q. Wang, “A curving ACC system with coordination control of longitudinal car-following and lateral stability,” Veh. Syst. Dyn., Vol. 50(7), pp. 1085-1102, 2012.

[24] T. Dierks, and S. Jagannathan, “Online optimal control of nonlinear discrete-time systems using approximate dynamic programming,” J. Control Theory Appl., Vol. 9(3), pp. 361-369, 2011.

[25] H. Watkins, “Learning from delayed rewards,” Cambridge University, Cambridge, England, Doctoral thesis, 1989.

[26] D. Vrabie, M. Abu-Khalaf, F.L. Lewis, and Y.Y. Wang, “Continuous-time ADP for linear systems with partially unknown dynamics,” IEEE International Symposium on Approximate Dynamic Programming and Reinforcement Learning, New York, 2007, pp. 247-253.

[27] S.J. Bradtke, B.E. Ydstie, and A.G. Barto, “Adaptive linear quadratic control using policy iteration,” American Control Conference, 1994, pp. 3475-3479.

[28] A. Al-Tamimi, F.L. Lewis, and M. Abu-Khalaf, “Model-free Q-learning designs for linear discrete-time zero-sum games with application to H-infinity control,” Automatica, Vol. 43(3), pp. 473-481, 2007.

View publication stats