Neurocomputing 125 (2014) 57–67

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Full-range adaptive cruise control based on supervised adaptive dynamic programming ☆

Dongbin Zhao [a], Zhaohui Hu [a,b], Zhongpu Xia [a,*], Cesare Alippi [a,c], Yuanheng Zhu [a], Ding Wang [a]

[a] State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China
[b] Electric Power Research Institute of Guangdong Power Grid Corporation, Guangzhou 510080, China
[c] Dipartimento di Elettronica e Informazione, Politecnico, di Milano, 20133 Milano, Italy

ARTICLE INFO

Available online 26 February 2013

Keywords:
Adaptive dynamic programming
Supervised reinforcement learning
Neural networks
Adaptive cruise control
Stop and go

ABSTRACT

The paper proposes a supervised adaptive dynamic programming (SADP) algorithm for a full-range adaptive cruise control (ACC) system, which can be formulated as a dynamic programming problem with stochastic demands. The suggested ACC system has been designed to allow the host vehicle to drive both in highways and in Stop and Go (SG) urban scenarios. The ACC system can autonomously drive the host vehicle to a desired speed and/or a given distance from the target vehicle in both operational cases. Traditional adaptive dynamic programming (ADP) is a suitable tool to address the problem but training usually suffers from low convergence rates and hardly achieves an effective controller. A SADP algorithm which introduces the concept of inducing region is here introduced to overcome such training drawbacks. The SADP algorithm performs very well in all simulation scenarios and always better than more traditional controllers. The conclusion is that the proposed SADP algorithm is an effective control methodology able to effectively address the full-range ACC problem.

1. Introduction

Nowadays, driving safety and driver-assistance systems are of paramount importance: by implementing these techniques accidents reduce and driving safety significantly improves [1]. There are many applications derived from this concept, e.g., anti-lock braking systems (ABS), electronic braking systems (EBS), electronic brake-force distribution systems (EBD), traction control systems (TCS), and electronic stability program (ESP) [1].

1.1. Adaptive cruise control

Adaptive cruise control is surely another issue going in the direction of safe driving and, as such, of particular relevance. Nowadays, ACC is mounted in some luxury vehicles to increase both comfort and safety [2]. The system differentiates from the cruise control (CC) system mostly used in highway driving, which controls the throttle position to maintain the constant speed as set by the driver (eventually adjusted manually to adapt to environmental changes). However, the driver always has to brake when approaching the target vehicle proceeding at a lower speed. Differently, an ACC system equipped with a proximity radar [3] or sensors detecting the distance and the relative speed between the host vehicle and the one in front of it, proceeding in the same lane (target vehicle), can operate either on brake or the engine throttle valve to keep a safe distance.

As a consequence, the ACC does not only free the driver from frequent accelerations and decelerations but also reduces the stress of the driver as pointed out in [4]. Interestingly, [5] showed that if 25% vehicles driving in a highway were equipped with the ACC system, congestions could be avoided. The ACC problem could be solved by considering different techniques, e.g., a PID controller [12], a fuzzy controller as pointed out in [11], a sliding mode approach [9] or a neural network [18].

ACC systems suggested in the literature, and currently implemented in vehicles, work nicely at a vehicle speed over 40 km/h and in highways [1], but always fail at a lower speed hence requiring accelerations (action on the throttle) and decelerations (mostly breaking) to keep a safe clearance to the target vehicle in urban areas. In this case, the driving activity increases significantly, even more within an urban traffic with an obvious impact on fuel consumption and pollutant emissions. To address the problem the literature suggested solutions like stop and go, collision warning and collision avoidance [22]. When the ACC and

---

☆This work was supported partly by National Natural Science Foundation of China under Grant nos. 61273136, 61034002, and 60621001, Beijing Natural Science Foundation under Grant no. 4122083, and Visiting Professorship of Chinese Academy of Sciences.

* Corresponding author. Tel.: +86 13683277856; fax: +86 10 8261 9580.
E-mail addresses: dongbin.zhao@ia.ac.cn (D. Zhao), huzhaohui27@foxmail.com (Z. Hu), zhongpu.xia@gmail.com (Z. Xia), alippi@elet.polimi.it (C. Alippi), zyh7716155@163.com (Y. Zhu), ding.wang@ia.ac.cn (D. Wang).

0925-2312/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.neucom.2012.09.034

the SG solutions are considered together, we speak about a full-range ACC. A full-range ACC system with collision avoidance was proposed in [16]. There, driving situations were classified into three control modes based on the warning index and the time-to-collision: comfort, large deceleration and severe braking. Three controllers were proposed and combined to provide the ultimate control strategy. Moon and Yi [16] pointed out that how the full-range ACC problem was a nonlinear process requesting a nonlinear controller, for instance designed with reinforcement learning.

1.2. Reinforcement learning and adaptive dynamic programming

Reinforcement learning (RL) [21] is suited for the ACC problem, because it can grant quasi-optimal control performance through a trial and error mechanism in a changing environment. However, the convergence rate of RL might be a problem [23] also leading to some inefficiency. Most of the time, the agent (the software implementing the controller) will learn the optimal policy after a relatively long training, especially when the model is characterized by a large state space. This inefficiency can be fatal in some real time control systems.

Supervised reinforcement learning (SRL) can be introduced to mitigate the RL problem, by combining supervised learning (SL) and RL and, hence, taking advantage of both algorithms. Pioneering work has been done in Rosenstein and Barto's [7,19], where SRL was applied to solve the ship steering task and the manipulator control and the peg insertion task. All results clearly showed how SRL outperforms RL. In [17], a potential function was introduced to construct the shaping reward function; they proved that an optimal control policy could be gained. The results showed that such shaping method could be used also in dynamic models by dramatically shortening the learning time.

Our team applied the SRL control strategy to the ACC problem first in [14]. There, we showed that the speed and the distance control had enough accuracy and was robust with respect to the different drivers [14]. However, since the state and the action needed to be discretized, there are some drawbacks. Firstly, the discretization of the distance, speed, and acceleration, introduces some fluctuations in the continuous control problem. Secondly, the higher number of discretized states cause the larger state and the action spaces. As a consequence, there always exists a conflict between control accuracy and required training time.

For continuous reinforcement learning problem, ADP was proposed in [8,25] with neural networks mapping the relationships between states and actions and the relationships between states, actions and performance index. More in detail, the algorithm uses a single step computation of the neural network to approximate the performance index which will be obtained by iterating the dynamic programming algorithm. The method provides us with a feasible and effective way to address many optimal control problems; examples can be found in the cart-pole control [13,20], pendulum robot upswing control [26], urban intersection traffic signal control [15], freeway ramp metering [6,27], play Go-Moku [28], and so on. However, the learning inefficiency of RL is also inherited in ADP but can also be remedied with a supervisor to formulate SADP.

1.3. The idea

In this paper we propose a novel effective SADP algorithm able to deal with the full-range ACC problem. The considered framework is as follows:

(1) There are two neural networks in SADP, the Action and the Critic networks. The Action network is used to map the continuous state space to the control signal; the Critic network is used to evaluate the goodness of the action signals generated by the Action network and provides advice while training both networks. In this way we avoid the curse of dimensionality caused by the large dimension of the discrete state-action pairs.
(2) The supervisor can always provide information for RL, hence speeding up the learning process.

In this paper, the ACC problem is described as a Markov decision process. The main contributions are as follows:

(1) A simple single neural network controller is proposed and optimized to solve the full-range adaptive cruise control problem.
(2) An inducing region scheme is introduced as a supervisor, which is combined with ADP, provides an effective learning algorithm.
(3) An extensive experimental campaign is provided to show the effectiveness and robustness of the proposed algorithm.

The paper is organized as follows. Section 2 formalizes the full-range ACC problem. Section 3 proposes the SADP algorithm based on the inducing region concept and presents design details. Section 4 provides experimental results based on the typical driving scenarios. Section 5 summarizes the paper.

2. The adaptive cruise control

2.1. The ACC model

The ACC model is shown in Fig. 1 with the nomenclature give in Table 1.

During driving, the ACC system assists (or replaces) the driver to control the host vehicle. In other words, ACC will control the throttle and the brake to drive the vehicle safely despite the uncertainty scenarios it might encounter. More in detail, there are two controllers in the ACC system: the upper and the bottom ones. The upper controller generates the desired acceleration

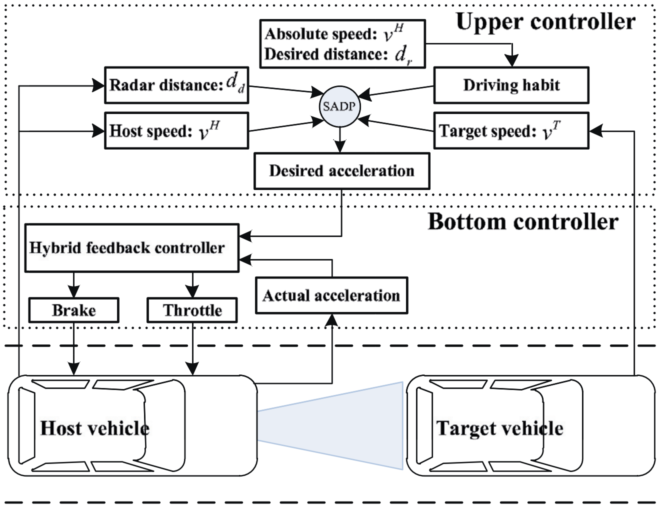

Fig. 1. The SADP framework for the full-range ACC. The radar detects the distance between the two vehicles and the target vehicle's speed. The host vehicle speed and the current acceleration come from the mounted sensors. The upper controller generates the desired acceleration signal by combining the relative speed and the relative distance information. The bottom controller maps the acceleration to the brake or the throttle control signals.