Neurocomputing 198 (2016) 80–90

Contents lists available at ScienceDirect

Neurocomputing

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

Guaranteed cost neural tracking control for a class of uncertain nonlinear systems using adaptive dynamic programming ☆

Xiong Yang [a], Derong Liu [b,*], Qinglai Wei [a], Ding Wang [a]

[a] The State Key Laboratory of Management and Control for Complex Systems Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China
[b] School of Automation and Electrical Engineering, University of Science and Technology Beijing, Beijing 100083, China

ARTICLE INFO

Article history:
Received 28 March 2015
Received in revised form
17 July 2015
Accepted 22 August 2015
Available online 10 March 2016

Keywords:
Adaptive dynamic programming
Guaranteed cost control
Hamilton–Jacobi–Bellman equation
Neural network
Nonlinear system
Reinforcement learning

ABSTRACT

This paper presents an adaptive dynamic programming-based guaranteed cost neural tracking control algorithm for a class of continuous-time matched uncertain nonlinear systems. By introducing an augmented system and employing a modified cost function with a discount factor, the guaranteed cost tracking control problem is transformed into an optimal tracking control problem. Unlike existing optimal tracking control algorithms often requiring the control matrix to be invertible, the developed control algorithm relaxes this restrictive condition under the assumption that the system is controllable. A single critic neural network (NN) is constructed to approximate the solution of the modified Hamilton–Jacobi–Bellman equation corresponding to the nominal augmented error dynamics. Utilizing the newly developed critic NN, the optimal tracking control can be derived without policy iteration. All signals in the closed-loop system are proved to be uniformly ultimately bounded via Lyapunov's direct method. In addition, the developed control scheme is verified to guarantee that the tracking errors converge to an adjustable neighborhood of the origin. Two numerical examples are provided to illustrate the effectiveness and applicability of the developed approach.

1. Introduction

During the past several decades, controller designs for uncertain nonlinear systems have become a significant and active research area in the control community. Many remarkable results in this field have been obtained owing to the advance in nonlinear control theory, especially Lyapunov's direct methods, feedback linearization techniques, and neural network (NN) control [1–3]. Nevertheless, most of the existing literature only focuses on designing stable controllers for uncertain nonlinear systems. In fact, stability is only a bare minimum requirement in a system design. It is often desirable to design a controller which not only keeps the closed-loop system stable but also guarantees an adequate level of performance. One approach to address this problem is the so-called guaranteed cost control method, which provides an upper bound on a prescribed cost [4,5].

The guaranteed cost control problems of nonlinear systems with uncertainties are often converted to nonlinear optimal control problems [6,7]. A major challenge of getting solutions for nonlinear optimal control problems is that it usually falls to solve the Hamilton–Jacobi–Bellman (HJB) equations. The HJB equation is actually a nonlinear partial differential equation (PDE), which is intractable to solve by analytical approaches. Though dynamic programming provides a way to tackle the problem [8], it is implemented backward in time which makes the computation untenable to run with increase in the dimensionality of nonlinear systems. To overcome the difficulty of applying dynamic programming, Werbos introduced adaptive dynamic programming (ADP) methods [9,10]. A remarkable feature of the ADP approach is that it employs NNs to get approximate solutions of the HJB equation forward in time. The general structure utilized to implement ADP algorithms is the actor-critic architecture, where two kinds of NNs referred to as critic NNs and actor NNs are employed to approximate the optimal cost function and the optimal control, respectively. There are several synonyms used for ADP, including "adaptive dynamic programming" [11–18], "approximate dynamic programming" [19], "adaptive critic designs" [20], "neural dynamic programming" [21], and "reinforcement learning (RL)" [22–26].

In the past several years, applications of ADP approaches to optimal tracking control have been extensively studied in the literature [27–32]. It is well-known that the optimal tracking

☆This work was supported in part by the National Natural Science Foundation of China under Grants 61233001, 61273140, 61304086, and 61374105, in part by Beijing Natural Science Foundation under Grant 4132078, and in part by the Early Career Development Award of the State Key Laboratory of Management and Control for Complex Systems (SKLMCCS).
* Corresponding author. Tel.: +86 10 82544761; fax: +86 10 82544799.
E-mail addresses: xiong.yang@ia.ac.cn (X. Yang), derong@ustb.edu.cn (D. Liu), qinglai.wei@ia.ac.cn (Q. Wei), ding.wang@ia.ac.cn (D. Wang).

http://dx.doi.org/10.1016/j.neucom.2015.08.119
0925-2312/

controller for nonlinear systems is made of two controllers: the feedforward controller and the feedback controller [33]. All the above-mentioned literature uses feedback linearization techniques to derive the feedforward controller *a priori* and only utilizes ADP methods to get the optimal feedback controller. However, for purpose of utilizing feedback linearization techniques, the control matrix is often required to be invertible. This requirement is restrictive and hard to satisfy in engineering applications. To overcome the defect, in [34], Heydari and Balakrishnan proposed a single network adaptive critic architecture to obtain the optimal tracking control for continuous-time (CT) nonlinear systems. Based on the architecture, the control matrix was no longer required to be invertible. In [35], Modares and Lewis introduced a discounted value function for the CT linear quadratic tracking problem. By using an online integral RL algorithm, the optimal tracking control was obtained without requiring the control matrix to be invertible. After that, Modares and Lewis applied the proposed algorithm to derive the optimal tracking control for CT constrained-input nonlinear systems with a discounted value function [36]. More recently, in [37], Kiumarsi and Lewis extended the work of [36] and employed a similar approach as in [36] to obtain the optimal tracking control for discrete-time constrained-input nonlinear systems with a discounted value function.

Though the aforementioned literature provides important and valuable insights into giving the optimal tracking control for nonlinear systems, most of them developed control algorithms based on the policy iteration method, which requires an initial admissible control. To the best of our knowledge, there is no general approach proposed to derive such a control. From mathematical perspectives, the initial admissible control is actually a suboptimal control. The suboptimal control is difficult to obtain, for it is often impossible to get analytical solutions of nonlinear PDEs. Therefore, how to obtain the initial admissible control is a challenging issue. In addition, for using the policy iteration algorithm, it is necessary to guarantee the newly derived control to be admissible at each iteration. As far as we know, except for the work of [38], there is no literature to show that the newly obtained control for CT nonlinear systems with a discounted cost function is admissible at each iteration. Unfortunately, if the discount term $e^{-\alpha(\tau-t)}$ ($\alpha>0, \tau>t$) is taken into consideration in the cost function, then the admissible control might not be obtained at each iteration by using the algorithm developed in [38] (a detailed explanation is provided in Section 3). In this sense, the policy iteration algorithm cannot be used.

In this paper, a novel ADP-based guaranteed cost neural tracking control algorithm will be developed for a class of CT matched uncertain nonlinear systems. By introducing an augmented system and employing a modified cost function with a discount factor, the guaranteed cost tracking control problem is converted to an optimal tracking control problem. Unlike existing optimal tracking control algorithms generally requiring the control matrix to be invertible, the developed control algorithm relaxes this restrictive condition under the assumption that the system is controllable. A single critic NN is constructed to approximate the solution of the modified HJB equation corresponding to the nominal augmented error dynamics. Based on the newly developed critic NN, the optimal tracking control can be obtained without policy iteration. All signals in the closed-loop system are proved to be uniformly ultimately bounded (UUB) via Lyapunov's direct method. In addition, the developed control scheme is verified to guarantee that the tracking errors converge to an adjustable neighborhood of the origin.

It should be mentioned that, in some fields, such as aerocraft industry, it is often necessary to develop a controller to deliver a desired tracking performance for nonlinear systems, while guaranteeing an adequate level of cost [39]. Due to the uncertainties often involved in nonlinear systems, it gives rise to lots of challenges for directly designing such a controller. Moreover, the control matrix is often singular, which brings about more difficulties to obtain the tracking control. In this paper, the present control algorithm shall provide an effective way for handling the above-mentioned problems. In other words, the developed control algorithm in this paper can provide novel perspectives to successfully solve guaranteed cost tracking control problems from industry.

The rest of the paper is organized as follows. In Section 2, the problem statement and preliminaries are presented. In Section 3, we show that the guaranteed cost tracking control of uncertain nonlinear systems can be derived by solving an HJB equation. In Section 4, we develop an ADP-based control scheme to give the approximate solution of the HJB equation. Meanwhile, the stability analysis is provided. In Section 5, the theoretical results are applied to two simulation examples. Finally, in Section 6, we give several concluding remarks.

Notations: $\mathbb{R}$ represents the set of all real numbers. $\mathbb{R}^m$ denotes the Euclidean space of all real m-vectors. $\mathbb{R}^{n\times m}$ denotes the space of all $n\times m$ real matrices. I_n represents the $n\times n$ identity matrix. T is the transposition symbol. Ω ($\tilde{\Omega}$) is a compact set of $\mathbb{R}^n$ ($\mathbb{R}^{2n}$), $C^m(\Omega)$ ($C^m(\tilde{\Omega})$) represents the class of functions having continuous mth derivative on Ω ($\tilde{\Omega}$). $\xi=[\xi_1,\ldots,\xi_m]^{\mathsf{T}}\in\mathbb{R}^m$, $\|\xi\|=\left(\sum_{i=1}^m|\xi_i|^2\right)^{1/2}$ denotes the Euclidean norm of ξ. When $A\in\mathbb{R}^{m\times m}$, $\|A\|=\left(\lambda_{\max}\left(A^{\mathsf{T}}A\right)\right)^{1/2}$ denotes the 2-norm of A, where $\lambda_{\max}(A^{\mathsf{T}}A)$ represents the maximum eigenvalue of $A^{\mathsf{T}}A$.

2. Problem statement and preliminaries

2.1. Guaranteed cost tracking control problem

Consider the uncertain CT nonlinear system described by

$$\dot{x}(t)=f(x(t))+g(x(t))u(t)+\Delta f(x(t)), \tag{1}$$

where $x(t)\in\mathbb{R}^n$ is the state, $u(t)\in\mathbb{R}^m$ is the control input, $f(x(t))\in\mathbb{R}^n$ and $g(x(t))\in\mathbb{R}^{n\times m}$ are known functions with $f(0)=0$, and $\Delta f(x(t))\in\mathbb{R}^n$ is an unknown perturbation. For convenience of later analysis, we provide the following assumptions.

Assumption 1. $f(x)+g(x)u$ is Lipschitz continuous on a compact set $\Omega\subset\mathbb{R}^n$ containing the origin and system (1) is controllable. There exists a known constant $g_M>0$ such that $0<\|g(x)\|\le g_M$ for every $x\in\mathbb{R}^n$.

Assumption 2. The perturbation term $\Delta f(x)$ satisfies the matching condition [40]. It implies that $\Delta f(x)=g(x)d(x)$, where $d(x)\in\mathbb{R}^m$ is an unknown function bounded by a known function $d_M(x)$, i.e., $\|d(x)\|\le d_M(x)$. In addition, $d(0)=0$ and $d_M(0)=0$.

Assumption 3. Let the desired trajectory of system (1) be $x_d(t)$, which is a bounded function. Meanwhile, $x_d(t)$ is produced by the command generator model

$$\dot{x}_d(t)=\eta(x_d(t)),$$

where $\eta:\mathbb{R}^n\to\mathbb{R}^n$ is a Lipschitz continuous function with $\eta(0)=0$.

Define the tracking error as

$$e_{\mathrm{err}}(t)=x(t)-x_d(t).$$

Then we obtain the tracking error dynamics as

$$\begin{aligned}\dot{e}_{\mathrm{err}}(t)=\;& f(x_d(t)+e_{\mathrm{err}}(t))+g(x_d(t)+e_{\mathrm{err}}(t))u(t)-\eta(x_d(t))\\ &+\Delta f(x_d(t)+e_{\mathrm{err}}(t)).\end{aligned} \tag{2}$$