Neurocomputing 198 (2016) 91–99

Contents lists available at ScienceDirect

Neurocomputing

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

Neural network-based online H_∞ control for discrete-time affine nonlinear system using adaptive dynamic programming

Chunbin Qin [a,b], Huaguang Zhang [c,d,*], Yingchun Wang [c], Yanhong Luo [c]

[a] The College of Computer and Information Engineering, Henan University, Kaifeng, Henan 475004, China
[b] The College of Environment and Planning, Henan University, Kaifeng, Henan 475004, China
[c] The College of Information Science and Engineering, Northeastern University, Shenyang, Liaoning 110004, China
[d] The State Key Laboratory of Synthetical Automation for Process Industries (Northeastern University), Shenyang 110819, China

ARTICLE INFO

Article history:
Received 31 March 2015
Received in revised form
10 August 2015
Accepted 31 August 2015
Available online 8 March 2016

Keywords:
H_∞ control
Adaptive dynamic programming
Neural networks
Nonlinear discrete-time system
Two-person zero-sum game

ABSTRACT

In this paper, the problem of H_∞ control design for affine nonlinear discrete-time systems is addressed by using adaptive dynamic programming (ADP). First, the nonlinear H_∞ control problem is transformed into solving the two-player zero-sum differential game problem of the nonlinear system. Then, the critic, action and disturbance networks are designed by using neural networks to solve online the Hamilton–Jacobi–Isaacs (HJI) equation associating with the two-player zero-sum differential game. When novel weight update laws for the critic, action and disturbance networks are tuned online by using data generated in real-time along the system trajectories, it is shown that the system states, all neural networks weight estimation errors are uniformly ultimately bounded by using Lyapunov techniques. Further, it is shown that the output of the action network approaches the optimal control input with small bounded error and the output of the disturbance network approaches the worst disturbance with small bounded error. At last, simulation results are presented to demonstrate the effectiveness of the new ADP-based method.

1. Introduction

It is well known that the control performance for practical systems is often affected by the presence of unknown disturbances such as measurement noise, input disturbances and other exogenous signals, which invariably occur in most applications because of plant interactions with the environment. H_∞ control is one of the most powerful control methods for attenuating the effect of disturbances in dynamical systems [1]. The formulation of the H_∞ Control for dynamical systems was studied in the frame work of Hamilton–Jacobi equations by van der Schaft [2] and Isidori and Astolfi [3]. It is worth noting that conditions for the existence of smooth solutions of the Hamilton–Jacobi equation were studied through invariant manifolds of Hamiltonian vector fields and the relation with the Hamiltonian matrices of the corresponding Riccati equation for the linearized problem [2]. Some of these conditions were relaxed into critical and noncritical cases by Isidori and Astolfi [3]. Later, Basar and Bernhard in [4] stated that the H_∞ control problem could be posed as the zero-sum two-person differential game, in which the input controller is a minimizing player and the unknown disturbance is a maximizing player. Although the formulation of the nonlinear H_∞ control theory has been well developed, the main bottleneck for its practical application is the need to solve the Hamilton–Jacobi–Isaacs (HJI) equation, which is difficult or impossible to solve and may not have global analytic solutions [5]. Therefore, solving the HJI equation remains a challenge.

Over the past decades, some methods have been proposed to solve the HJI equation [6–8]. The smooth solution of the HJI equation has been determined directly by solving for the coefficients of the Taylor series expansion of the value function in a very efficient manner, as it has been presented in [6]. Beard and McLain [7] proposed an iterative-based policy to successively solve the HJI equation by breaking the nonlinear differential equation to a sequence of linear differential equations. On the basis of the work [6] and [7], a similar iterative-based policy was proposed in [8] to the HJI equation for nonlinear systems with input constraints.

In recent years, adaptive dynamic programming (ADP) [9–13] has appeared to be promising methodologies for solving H_∞ control problems [15–23]. Adaptive dynamic programming is a kind of machine learning method for learning the feedback control laws online in real time based on system performance without necessarily knowing the system dynamics, which overcomes the curse of dimensionality [14] of dynamic programming. Al-Tamimi et al. in [15]

* Corresponding author at: The College of Information Science and Engineering, Northeastern University, Shenyang, Liaoning 110004, China.
E-mail address: zhanghuaguangieee@gmail.com (H. Zhang).

http://dx.doi.org/10.1016/j.neucom.2015.08.120
0925-2312/

derived adaptive critic designs corresponding to heuristic dynamic programming and dual heuristic dynamic programming to solve online the H_∞ control problem of the linear discrete-time system in a forward-in-time manner. Based on this work, authors in [16] proposed an iterative adaptive critic design algorithm to find the optimal controller of a class of discrete-time two-person zero-sum games for Roesser types 2-D systems. Further, a novel data-based adaptive critic design was proposed by using output feedback of unknown discrete-time zero-sum games [17]. Besides, optimal strategies based Q-learning were proposed for the H_∞ optimal control problem without knowing the system dynamical matrices in [18] and [19]. For the nonlinear case, Mehraeen et al. [20,21] developed an off-line iterative approach to solve the HJI equation by using a successive approximation approach. Liu et al. in [22] proposed value iteration methods corresponding to heuristic dynamic programming and dual heuristic dynamic programming to solve the HJI equation for constrained input systems. Later, Liu et al. [23] proposed an iterative adaptive dynamic programming algorithm to solve the zero-sum game problems for affine nonlinear discrete-time systems. Nevertheless, a common feature of the above ADP-based results for solving the H_∞ control problem is that sequential iterative approaches are utilized to solve the HJI equation, which contain more than one iteration loop, i.e., the value function and the control and disturbance policies are asynchronously updated. However, such a procedure may lead to redundant iterations, and result in low efficiency [24], which motivates us to carry out the work of this paper.

In this paper, a new ADP-based method is proposed to solve online the H_∞ control problem of the nonlinear system, in which three online parametric structures are designed by using three neural networks for solving online the Hamilton–Jacobi–Isaacs equation appearing in the H_∞ control problem of the nonlinear system. The main contributions of this paper have two folds. First, we present a new ADP-based method in which the weights of three online parametric structures are tuned simultaneously along the system trajectories to converge to the solution of the HJI equation, which is different from the sequential algorithms in [15–23]. Second, while explicitly considering the neural network approximation errors in contrast to the works [20,22], Lyapunov theory is utilized to demonstrate that the system states and the weight estimation errors of three online parametric structures are uniformly ultimately bounded. Besides, it is shown that the pair of the approximated control signal and the disturbance input signal converges to the approximate Nash equilibrium solution of the two-player zero-sum differential game.

The remainder of this paper is organized as follows. In Section 2, the problem statement is shown. In Section 3, we present a new ADP-based method for solving HJI equation of nonlinear discrete-time systems and the rigorous proof of convergence is given. Section 4 presents an example to demonstrate the effectiveness of the proposed method. Finally, conclusions are drawn in Section 5.

2. Problem formulation

In this paper, we consider the following affine nonlinear discrete-time system in the presence of the disturbance $d(k)$:

$$x_{k+1}=f(x_k)+g(x_k)u(k)+d(k) \tag{1}$$

$$z(k)=[Cx_k \;\; Du(k)]^T \tag{2}$$

where $x_k\in R^n$ is the system state, $u(k)\in R^m$ is the system control input, $d(k)\in R^n$ is the disturbance signal with $d(k)\in L_2[0,\infty]$, $z(k)$ is the system fictitious output. Assume that $\|g(x_k)\|_F\le g_M$ [25], where $\|\cdot\|_F$ denotes the Frobenius norm.

The H_∞ control for the nonlinear discrete-time system (1) and (2) is to find a state feedback control

$$u(k)=u(x_k) \tag{3}$$

such that the closed-loop system (1) and (2) with (3) is asymptotically stable, and has L_2 gain less than or equal to γ, i.e.

$$\sum_{k=0}^{\infty} z^T(k)z(k)\le\gamma^2\sum_{k=0}^{\infty} d^T(k)d(k) \tag{4}$$

for all $d(k)\in L_2[0,\infty]$, where $\gamma>0$ is some prescribed level of disturbance attenuation. Note that throughout this paper we shall assume that γ is fixed and $\gamma\ge\gamma^*$, where γ^* is the minimum γ for which Eq. (4) can hold.

According to [4], it is well known that the H_∞ control problem can be posed as a zero-sum two-player differential game, in which the system control input $u(k)$ is regarded as a minimizing player and the disturbance $d(k)$ is regarded as a maximizing one. Correspondingly, we can define the following infinite horizon quadratic cost function for the zero-sum two-player differential game,

$$J(x(0),u,d)=\sum_{k=0}^{\infty}U(x_k,u_k,d_k) \tag{5}$$

where $U(x_k,u_k,d_k)=x_k^TQx_k+u_k^TRu_k-\gamma^2d_k^Td_k$, $Q=C^TC$, $R=D^TD$, $x_k=x(k)$, $u_k=u(k)$, $d_k=d(k)$.

For the given system control input u_k and the bounded disturbance d_k, we can define the corresponding value function as

$$V(x_k,u_k,d_k)=\sum_{i=k}^{\infty}U(x_i,u_i,d_i). \tag{6}$$

Correspondingly, the Hamilton function can be defined as

$$H(x_k,u_k,d_k)=V(x_{k+1})-V(x_k)+U(x_k,u_k,d_k), \tag{7}$$

where $x_{k+1}=f(x_k)+g(x_k)u_k+d_k$.

Therefore, for the zero-sum two-player differential game of the nonlinear discrete-time system (1) and (2), our aim is to find a state feedback saddle point (u^*, d^*) such that

$$V(u^*,d^*)=\min_u\max_d V(u,d), \tag{8}$$

that means

$$V(u^*,d)\le V(u^*,d^*)\le V(u,d^*), \tag{9}$$

where $u^*=\mu(x_k)$ and $d^*=\eta(x_k)$, $\mu(\cdot)$ and $\eta(\cdot)$ are smooth functions.

According to Bellman's optimality principle, we can obtain that the optimal value function $V^*(x_k)$ satisfies the following discrete-time HJI equation:

$$V^*(x_k)=\min_u\max_d\{U(x_k,u_k,d_k)+V^*(x_{k+1})\}. \tag{10}$$

At the same time, we can obtain the saddle point (u^*,d^*) of the zero-sum two-player differential game as follows:

$$u^*(x_k)=-\frac{1}{2}R^{-1}g(x_k)^T\frac{\partial V^*(x_{k+1})}{\partial x_{k+1}}, \tag{11}$$

and

$$d^*(x_k)=\frac{1}{2\gamma^2}\frac{\partial V^*(x_{k+1})}{\partial x_{k+1}}. \tag{12}$$

Inserting (11) and (12) into (10), the discrete-time HJI equation can be rewritten as

$$0=V^*(x_{k+1})-V^*(x_k)+\frac{1}{4}\frac{\partial V^{*T}(x_{k+1})}{\partial x_{k+1}}g(x_k)R^{-1}g^T(x_k)\frac{\partial V^*(x_{k+1})}{\partial x_{k+1}} -\frac{1}{4\gamma^2}\frac{\partial V^{*T}(x_{k+1})}{\partial x_{k+1}}\times\frac{\partial V^*(x_{k+1})}{\partial x_{k+1}}+x_k^TQx_k, \tag{13}$$

where $x_{k+1}=f(x_k)+g(x_k)u^*(x_k)+d^*(x_k)$, $V^*(0)=0$.