ĐẠI HỌC THÁI NGUYÊN
TRƯỜNG ĐẠI HỌC KỸ THUẬT CÔNG NGHIỆP

BÁO CÁO TỔNG KẾT
ĐỀ TÀI KHOA HỌC VÀ CÔNG NGHỆ CẤP TRƯỜNG

ĐIỀU KHIỂN HỆ ROBOTIC CÓ ĐÁNH GIÁ ĐẾN MIỀN HẤP DẪN

Mã số: T2019-B11

Xác nhận của tổ chức chủ trì	Chủ nhiệm đề tài
KT. HIỆU TRƯỞNG	*(ký, họ tên)*
PHÓ HIỆU TRƯỞNG	
BỘ GIÁO DỤC VÀ ĐÀO TẠO ★ TRƯỜNG ĐẠI HỌC KỸ THUẬT CÔNG NGHIỆP ★ ĐẠI HỌC THÁI NGUYÊN	
PGS.TS. Vũ Ngọc Pi	Trần Thị Hải Yến

Thái Nguyên, 07/2021

NHỮNG NGƯỜI THAM GIA THỰC HIỆN ĐỀ TÀI

ThS. Nguyễn Vĩnh Thụy - Khoa Điện – Trường ĐHKT Công nghiệp.

ThS. Đỗ Thị Phương Thảo - Khoa Điện – Trường ĐHKT Công nghiệp.

ThS. Nguyễn Thị Chinh- Khoa Điện – Trường ĐHKT Công nghiệp.

ThS. Lâm Hùng Sơn- Khoa Điện – Trường ĐHKT Công nghiệp.

ThS. Nguyễn Hồng Quang - Khoa Điện – Trường ĐHKT Công nghiệp.

ThS. Dương Hòa An - Khoa Điện – Trường ĐHKT Công nghiệp.

ThS. Vũ Xuân Tùng - Khoa Điện – Trường ĐHKT Công nghiệp.

ThS. Trần Thị Thanh Thảo - Khoa Điện – Trường ĐHKT Công nghiệp.

ThS. Dương Quỳnh Nga - Khoa Điện – Trường ĐHKT Công nghiệp.

ThS. Nguyễn Văn Huỳnh - Khoa Điện – Trường ĐHKT Công nghiệp.

TABLE OF CONTENTS

TRƯỜNG ĐẠI HỌC
KỸ THUẬT CÔNG NGHIỆP

Đơn vị: Khoa Điện

THÔNG TIN KẾT QUẢ NGHIÊN CỨU

1. Thông tin chung:

- Tên đề tài: Điều khiển hệ Robotic có đánh giá đến miền hấp dẫn
- Mã số: T2019-B11
- Chủ nhiệm đề tài: Trần Thị Hải Yến.
- Cơ quan chủ trì: Trường Đại học Kỹ thuật Công nghiệp.
- Thời gian thực hiện: 07/2019 – 07/2021

2. Mục tiêu:

Hệ thống chuyển động robot là một hệ thống có tính phi tuyến mạnh và ràng buộc cao, các tham số động lực học như mô men quán tính, khối lượng tải thường biến đổi và không được xác định chính xác. Việc điều khiển Robot bám chính xác quỹ đạo đặt với các điều kiện nhiễu tác động bên ngoài là không biết trước cũng như kháng được nhiễu nội của hệ thống sinh ra là điều luôn luôn được quan tâm của các nhà nghiên cứu. Việc giải quyết bài toán điều khiển hệ Robotic có đánh giá đến miền hấp dẫn là nội dung cần được giải quyết khi đề cập đến nâng cao chất lượng điều khiển.

3. Kết quả nghiên cứu:

Tác giả thực hiện mô hình hóa đối tượng điều khiển, đề xuất các thuật toán điều khiển cho đối tượng điều khiển, mô phỏng kiểm chứng kết quả. Các kết quả thu được công bố bằng các bài báo quốc tế .

4. Sản phẩm:

- Sản phẩm đào tạo:
- Sản phẩm khoa học: 02 bài báo ISI/Scopus Q2.

- Sản phẩm ứng dụng:

5. Hiệu quả:

Kết quả nghiên cứu của nhóm tác giả được công bố trên các tạp chí khoa học có uy tín nằm trong danh mục ISI/Scopus.

6. Khả năng áp dụng và phương thức chuyển giao kết quả nghiên cứu:

Cung cấp tài liệu chuyên ngành tham khảo cho sinh viên, học viên cao học, nghiên cứu sinh ngành Kỹ thuật điều khiển và Tự động hóa. Các kết quả của đề tài có thể sử dụng để hướng dẫn đề tài luận văn cao học cho học viên ngành Kỹ thuật điều khiển và Tự động hóa.

Ngày 15 tháng 7 năm 2021

Cơ quan chủ trì	Chủ nhiệm đề tài
KT.HIỆU TRƯỞNG	
PHÓ HIỆU TRƯỞNG	
(BỘ GIÁO DỤC VÀ ĐÀO TẠO – TRƯỜNG ĐẠI HỌC KỸ THUẬT CÔNG NGHIỆP – ĐẠI HỌC THÁI NGUYÊN)	
PGS.TS. Vũ Ngọc Pi	**Trần Thị Hải Yến**

THAI NGUYEN UNIVERSITY
OF TECHNOLOGY
Faculty Electrical Engineering

INFORMATION ON RESEARCH RESULTS

1. General information:

- Project title: Robotic Control System considering the Attraction Domains.
- Code number: T2019-B011
- Coordinator: Tran Thi Hai Yen
- Implementing Institution: Thai Nguyen University of Technology.
- Duration: From 07/2019 – to 07/2021.

2. Objectives:

The robot motion system is a system with solid nonlinearity and high constraints. Dynamic parameters such as a moment of inertia, load mass are often variable and not precisely determined. Controlling the Robot to follow the exact trajectory set with the conditions of unknown external disturbances and resist the internal disturbances of the generated system is always the concern of the researchers. Solving the problem of controlling the Robotic system with an assessment of the attraction domain is the content that needs to be solved when it comes to improving the control quality.

3. Research results:

The author performs modeling of control objects, proposes control algorithms for control objects, simulation verifies the results. The results are published in international journal articles.

4. Products:

- Training products:
- Scientific products: 02 ISI/Scopus articles (Q2).
- Application products:

5. Effects:

Research results of the authors group are published in prestigious scientific journals in the ISI/Scopus list.

6. Transfer alternatives of reserach results andapplic ability:

Provide specialized reference materials for students, graduate students, PhD students in Automation Engineering. The research results can be used to guide the master thesis for graduate students in Automation Engineering.

July 15, 2021

CHAPTER 1

OVERVIEW OF RESEARCH ON ADAPTIVE DYNAMIC PROGRAMMING

Dynamic systems are universal in nature. Stability analysis of dynamic systems has been a hot research topic for a long time, and a series of methods have been put forward. However, researchers in control theory field not only devote to guarantee the stability of the control system, but also to acquire the optimal solution. When it came to the 50 s and 60 s, because of the development of the space technology and digital computer, dynamic system optimization theory had been rapidly developed, and an important subject branch, named optimal control, emerged. It can be more and more extensively applicable in the space technology, system engineering, economic management and decision, population control, multistage process equipment optimization, and many other areas. In 1957, Bellman presented an effective tool the dynamic programming (DP) method, which can be used for solving the optimal control problem. The Bellman principle of optimality is the key of above method, which is described as: An optimal policy has the property that whatever the initial state and initial decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision. This principle can be summed up in a basic recurrence formula. When solving the multistage decision problem, we should reverse recurrence. Therefore, this principle can be applicable extensively, such as discrete systems, continuous systems, linear systems, nonlinear systems, deterministic systems, stochastic systems, and etc.

Next the DP principle is introduced in two cases respectively: discrete systems and continuous systems. First discrete nonlinear systems are considered. Suppose that the system dynamic equation can be described as

$$x(k+1) = F(x(k), u(k), k), k = 0,1,\cdots, \tag{1}$$

where $x \in \mathbf{R}^n$ represents the state vector of the system and $u \in \mathbf{R}^m$ is the control input vector. The corresponding cost function (or performance index function) has the form as

$$J(x(i),i)=\sum_{k=i}^{\infty}\gamma^{k-i}l(x(k),u(k),k) \tag{2}$$

Where $x(k) = x_k$ is given, $l(x(k), u(k), k)$ is called the utility function, γ is the discount factor with $0<\gamma\leq 1$. The objective of dynamic programming problem is to find a control sequence $u(k)$, $k = i, i+1, \cdots$, so that the cost function in (2) is minimized. According to Bellman principle, the minimum cost of any state from time k consists of two parts. One is the minimum cost at time k, and the other is the accumulated sum of the minimum cost from time k + 1 to infinity, that is

$$u^*(k)=\arg\min_{u(k)}\{l(x(k),u(k))+\gamma J^*(x(k+1))\} \tag{3}$$

At the same time, the control policy u(k) at time k achieves the minimum, i.e.,

$$u^*(k)=\arg\min_{u(k)}\{l(x(k),u(k))+\gamma J^*(x(k+1))\} \tag{4}$$

Now we consider about the optimal control problem of nonlinear continuous-time (time-varying) dynamic (deterministic) systems, which can be described by

$$\dot{x}(t)=F(x(t),u(t),t), t\geq t_0 \tag{5}$$

where F(x, u, t) is any continuous function. The objective is to choose the admissible control u(t) such that the cost function (or performance index function) achieves the minimum.

$$J(x(t),t)=\int_t^{\infty}l(x(\tau),u(\tau))\mathrm{d}\tau \tag{6}$$

In general, we can transform the continuous-time problem into a discrete-time problem by using the discretization method, and then use the discrete dynamic programming method to find the optimal control solution. When the discretization time interval tends to be zero, both of them will tend to be consistent. With the application of the Bellman optimality principle, we can get the continuous form of DP as

$$-\frac{\partial J^*}{\partial t} = \min_{u \in U}\left\{ l(x(t),u(t),t) + \left(\frac{\partial J^*}{\partial x(t)}\right)^{\mathrm{T}} F(x(t),u(t),t)\right\}$$
$$= l\left(x(t),u^*(t),t\right) + \left(\frac{\partial J^*}{\partial x(t)}\right)^{\mathrm{T}} F\left(x(t),u^*(t),t\right) \quad (7)$$

From above equation, we can see that $J^*(x(t),t)$ is a first order nonlinear partial differential equation with independent variable x(t), t. In mathematics, we call it as Hamilton-Jacobi-Bellman (HJB) equation.

If the system is linear and the cost function has the quadratic form with respect to the state and control input, the optimal control can be expressed as a linear feedback of the states, where the gains are obtained by solving a standard Riccati equation. However if the system is nonlinear and the cost function does not have the quadratic form with respect to the state and control input, we have to solve HJB equation to achieve the optimal control. However, it is very difficult to solve HJB equation. In addition, DP method has obvious weaknesses. With the dimension of x and u increasing, it is often computationally untenable to run true dynamic programming due to the backward numerical process required for its solution, i.e., as a result of the well known "curse of dimensionality" [1−2]. In order to overcome these weaknesses, Werbos first propose the framework of adaptive dynamic programming (ADP) [3], in which the idea is to use an approximate structure of function (such as neural network, fuzzy model, polynomial, etc.) to estimate the cost function and to solve DP problem forward-in-time.

In recent years, ADP scheme has received a widespread attention. A series of synonyms arose, for example, adaptive evaluation design [4−7], heuristic dynamic programming [8−9], neuron dynamic programming[10] adaptive dynamic programming [12] and reinforcement learning[13], etc. In 2006, National Science Foundation organized "2006 NSF Workshop and Outreach Tutorials on Approximate Dynamic Programming" seminar, where it was suggested that the kind of method is called "Adaptive/Approximate Dynamic Programming". Bertsekas et al. summarized the neuron dynamic programming [10–11]. They introduced the dynamic programming, the structure of neural network and the training algorithm, and presented many effective methods for application of neuron dynamic programming. Si et al. discussed the development of ADP scheme among inter-disciplines [14], and specially introduced the connection between DP/ADP scheme and artificial intelligence, approximation theory, control theory, operational research and statistics. In [15], Powell showed how to use ADP scheme to solve deterministic or stochastic optimization problem, and pointed out the development direction of ADP scheme. Balakrishnan et al. summarized the methods of designing feedback controller for dynamic systems by using ADP before, with the consideration of two cases i.e. for model-based systems and for model-free systems in 16. Reference [17] discussed ADP scheme from the view of whether requiring initial stability or not. Based on the research achievements of our group and the previous studies, this paper summarizes the latest development of ADP scheme.

1.1. Development of ADP Structure

In order to execute ADP scheme, Werbos proposed two basic structures: heuristic dynamic programming (HDP) and dual heuristic programming (DHP). The structures are shown in Fig. 1 and Fig. 2, respectively [3].

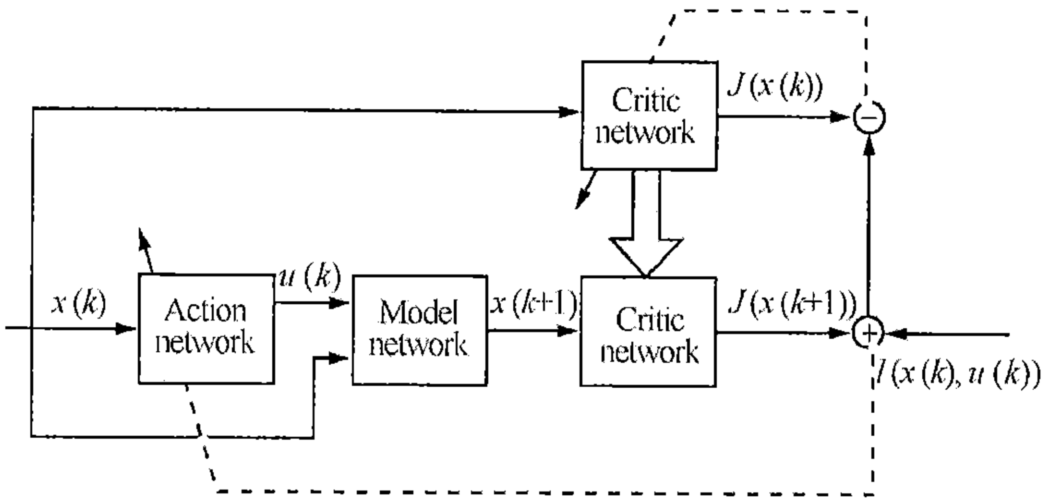


Fig. 1 The HDP structure

HDP is the most basic and widely used structure of ADP. The purpose is to estimate the system cost function. Generally three networks: critic network, action network and model network are adopted. The output of the critic network is used to estimate the cost function J(x(k)). The action network is used to map the relationship between state variable and the control input. The model network is used to estimate the system states for the next time. But the DHP is a method for estimating the gradient of the cost function. The definition of action network and model network is the same as the HDP, and output of the critic network is the gradient of the cost function, $\partial J(x(k))/\partial x(k)$.

Werbos further gave two other versions called "actiondependent critics", namely, ADHDP and ADDHP. The main difference from DHP and HDP is that the input of critic network is not only dependent on the system states, but including the control action. On the basis of that, Prokhorov and Wunsch presented two new structures: globalized dual heuristic programming (GDHP) and action dependent GDHP (ADGDHP) [18], whose characteristic is that the critic network can estimate not only the cost function itself but also the gradient of the cost function. All of the above ADP structures can be used to solve the optimal control policy, but the computation burden and computation precision are different. HDP is simple relatively. The computation speed is fast, but the computation precision is low. As for GDHP, the computation precision is high, but the calculation process need more time. The specific comparison is discussed in detail in [18].

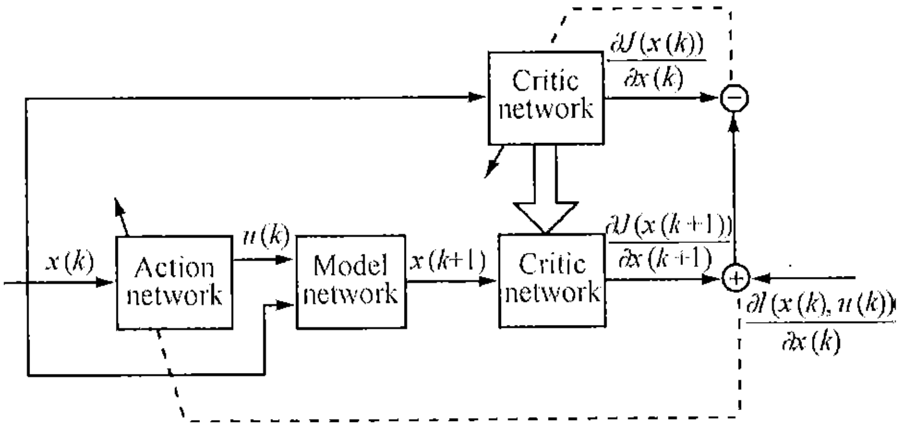


Fig. 2 The DHP structure

With the further development of ADP scheme, the two network structures with both critic network and action network are not needed anymore. Padhi et al. presented single network adaptive critic (SNAC) method. The structure is shown in Fig. 3. The action network is left and only critic network is kept in SNAC scheme. The output of critic network is the gradient of cost function, defined as costate vector $\lambda(k) = \partial J(x(k))/\partial x(k)$. This single network structure can reduce the computation burden and eliminate the approximation error of the action network. But the precondition of using SNAC scheme is that the optimal control policy can be explicitly expressed through the state vector and costate vector. Therefore, this method can be only used to solve the optimal control problem about linear systems with general quadratic form cost function or affine nonlinear systems. Because ADP scheme for continuous systems is developed on the basis of that for the discrete systems, the structure for continuous systems is roughly the same as that for discrete systems. The difference is only that the iteration of variables is running in the continuous space.

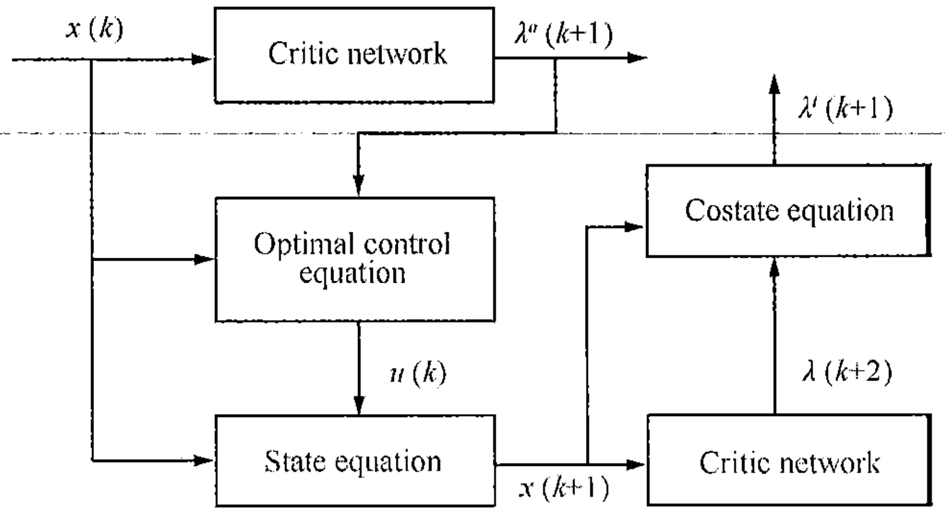


Fig. 3 The SNAC structure

With the further development of ADP scheme, the two network structures with both critic network and action network are not needed anymore. Padhi et al. presented single network adaptive critic (SNAC) method. The structure is shown in Fig. 3. The action network is left and only critic network is kept in SNAC scheme. The output of critic network is the gradient of cost function, defined as costate vector $\lambda(k) = \partial J(x(k))/\partial x(k)$.

This single network structure can reduce the computation burden and eliminate the approximation error of the action network. But the precondition of using SNAC scheme is that the optimal control policy can be explicitly expressed through the state vector and costate vector. Therefore, this method can be only used to solve the optimal control problem about linear systems with general quadratic form cost function or affine nonlinear systems. Because ADP scheme for continuous systems is developed on the basis of that for the discrete systems, the structure for continuous systems is roughly the same as that for discrete systems. The difference is only that the iteration of variables is running in the continuous space.

1.2. Algorithms of ADP

The algorithm of ADP experiences a process from offline iteration to online update realization. The theoretical research is mainly concerned with the stability analysis and the proof of convergence.

1.2.1. Offline iterative algorithm

Murray et al. first proposed an iterative ADP algorithm for continuous-time nonlinear systems in 2002 [12]. Consider a continuous-time differential equation

$$\dot{x} = f(x) + g(x)u, x(t_0) = x_0 \tag{8}$$

The corresponding cost function has been defined as (6), where $l(x,u) = Q(x) + u^{\mathrm{T}} R(x) u$. In this case, the optimal control can be expressed as

$$u^*(x) = -\frac{1}{2} R^{-1}(x) g^{\mathrm{T}}(x) \left(\frac{\partial J^*(x)}{\partial x} \right)^{\mathrm{T}} \tag{9}$$

$J^*(x)$ needs to be solved from the HJB equation (7). But we know the partial differential equation is difficult to find out an analytical solution so the following iteration algorithm is proposed. At first, an initial stable control policy is given, and then an iteration process is running between the following two formulas,

$$J_i(x_0) = \int_{t_0}^{+\infty} l(x_{i-1}, u_{i-1}) \mathrm{d}t \tag{10}$$

$$u_i(x) = -\frac{1}{2}R^{-1}(x)g^{T}(x)\left(\frac{\partial J_i(x)}{\partial x}\right)^{T} \tag{11}$$

Murray et al. gave the convergence analysis of the iterative ADP scheme and the stability proof of the system in [12]. This was the first time to prove that the iteration algorithm from an initial stable control policy can guarantee the stability of system and the convergence of iterative performance index mathematically, which is a great break through in ADP theory. Then Abu-Khalaf and Lewis studied about the optimal control problem of continuous nonlinear systems with saturation constraints [20], and proposed an iterative ADP algorithm based on the generalized HJB equation. An approximate optimal saturation controller is obtained as a result, and the convergence of the algorithm is proved strictly. Comparing the iteration ADP algorithm with [12], the policy iteration algorithm was used in [20], in which the policy equation is updated after each iteration.

However, value iteration algorithm was adopted in [12], in which the value function is updated after each iteration. For nonlinear discrete-time systems, Lewis and Zhang proposed an ADP iteration algorithm [21–23], which does not require an initially stable policy. Consider the following discrete-time system

$$x(k+1)=f(x(k))+g(x(k))\,u(k) \tag{12}$$

The corresponding cost function is shown as (2), where
$\gamma = 1, l(x,u) = x^{T}(k)Qx(k) + u^{T}(k)Ru(k)$ Q and R are positive definite matrices. The control objective is to find the optimal control policy so as to make the cost function minimum. This iteration algorithm starts from the initial value function $V_0(\cdot) = 0$, and the iteration process is running between the control policy and value function,

$$u_i(x(k)) = -\frac{1}{2}R^{-1}g^{T}(x(k))\left(\frac{\partial V_i(x(k))}{\partial x(k)}\right)^{T} \tag{13}$$

$$V_{i+1}(x(k)) = x(k)^{T}Qx(k) + u_i^{T}(x(k))Ru_i(x(k)) + V_i(x(k+1)) \tag{14}$$

Where:

$$x(k+1)=f(x(k))+g(x(k))\,u\{i\}(x(k))$$

Zhang et al. first proved that the iterative control policy converges to the optimal one and the value function sequence converges to the optimal cost function in theory. That is to say, all admissible control strategies achieve the minimum among all the cost functions. At the same time, they also proved that the optimal cost function can satisfy HJB equation, i.e. when $i \to \infty$,we have $V_\infty(x(k)) = J^*(x(k))$ and $u_\infty(x(k)) = u^*(x(k))$. Reference [22] employed HDP iteration algorithm to solve the optimal tracking control problem for a class of discrete systems. Due to the problem that the tracking dynamics may lead to the existing cost function tending to infinity, a new cost function needs to be defined. Using a system transformation, the optimal tracking problem is transformed into the optimal regulator problem and then the iterative HDP algorithm is introduced to obtain the approximate optimal tracking controller.

On the basis of literature above mentioned, [24–26] presented an approximate optimal control scheme via the ADP-based method for nonlinear time delay systems. Consider the discrete time delay system as

$$\begin{aligned} x(k+1) &= f\left(x\left(k-\sigma_0\right),\cdots,x\left(k-\sigma_m\right)\right)+ \\ &\quad g\left(x\left(k-\sigma_0\right),\cdots,x\left(k-\sigma_m\right)\right)u(k) \\ x(k) &= \lambda(k), -\sigma_m \le k \le 0 \end{aligned} \tag{15}$$

Where $\lambda(k)$ is the initial state, $\sigma_i, i = 0,1,\cdots,m$ are the time delays and satisfy $0 = \sigma_0 < \sigma_1 < \cdots < \sigma_m$, which are nonnegative integers. The corresponding cost function shows as (2). The control purpose is to find the optimal control policy to make the cost function minimum. First of all, an initial value function V $(\cdot) = 0$ is given. For any given initial state $\lambda(k)$ and initial control $\beta(k)$, one can begin to find the optimal control from i = 0. Then the HDP iteration algorithm is running among the control policy, value function, and system states.

$$u_i(k) = \arg\inf_{v(k)} \left\{x^{\mathrm{T}}(k)Qx(k) + u^{\mathrm{T}}(k)Ru(k) + V_i(x(k+1))\right\} \tag{16}$$

$$V_{i+1}\left(x_i(k)\right) = x_i(k)^{\mathrm{T}} Q x_i(k) + u_i(k)^{\mathrm{T}} R u_i(k) + V_i\left(x_{i-1}(k+1)\right) \tag{17}$$

$$x_i(t+1)=\begin{cases} f\left(x_{i+1}(t-\sigma_0),\cdots,x_{i+1}(t-\sigma_m)\right)+ \\ g\left(x_{i+1}(t-\sigma_0),\cdots,x_{i+1}(t-\sigma_m)\right)\times \\ \quad u_{i+1}(t),t\geq k \\ f\left(x_i(t-\sigma_0),\cdots,x_i(t-\sigma_m)\right)+ \\ g\left(x_i(t-\sigma_0),\cdots,x_i(t-\sigma_m)\right)\times \\ \quad u_i(t),0\leq t<k \end{cases} \tag{18}$$

$$x_i(t)=\lambda(t),-\sigma_m\leq t\leq 0$$

The convergence of the iteration algorithm is also proved. When $i\to\infty$ it can be seen that $u_\infty(k)=u^*(k)$, $V_\infty(x(k))=J^*(x(k))$ and $x_\infty(k)=x^*(k)$.

ADP scheme is also used to solve differential game problem forward-in-time [27–32]. During the system optimization design, we usually require the control variables to make the performance index minimum on the one hand. On the other hand, when the affect of the disturbance is adequately large, we do require the disturbance variables to make the performance index maximum simultaneously. Or we can say that one side chases, and another side escapes, which leads to the bilateral optimization problem of dynamic systems, namely, differential game problem.

Next, we utilize the two-person-zero-sum game as an example to describe how to use iterative ADP scheme to solve differential game problem. Consider the following system.

$$\dot{x}=f(x)+g(x)u+k(x)w \tag{19}$$

with the cost function as

$$J(x,u,w,t)=\int_t^\infty l(x(\tau),u(\tau),w(\tau))\mathrm{d}\tau \tag{20}$$

The objective of controller u is to make the system cost function (20) minimum, while the controller w is to achieve maximum of the cost function. So we need to define the upper bound function and the lower bound function respectively as

$$\overline{V}(x):=\inf_{u\in U[t,\infty)}\sup_{w\in W[t,\infty)}J(x,u,w) \tag{21}$$

$$\underline{V}(x) := \sup_{w \in W[t,\infty)} \inf_{u \in U[t,\infty)} J(x,u,w) \tag{22}$$

The corresponding control is defined as $(\overline{u},\overline{w})$ and $(\underline{u},\underline{w})$ respectively. So we have $\overline{V}(x)=J(x,\overline{u},\overline{w})$, $\underline{V}(x)=J(x,\underline{u},\underline{w})$. If $\overline{V}(x)$ and $\underline{V}(x)$ exist, and $\overline{V}(x)=\underline{V}(x)=V^{*}(x)$ is satisfied, it can be concluded that the optimal it can be concluded that the optimal control pair (u^{*},w^{*}) of the two-person-zero-sum game exists, namely, the saddle point exists. Assuming that the saddle point exists, Lewis et al. studied the two-personzero-sum differential game problem of discrete linear systems and continuous affine nonlinear systems by using iteration ADP scheme and combining with H_∞ control [27−29]. This iteration scheme was divided into inner loop iteration and outer loop iteration. Firstly given a stable control u_j, we implement the inner loop iteration to update w^i. After convergence of w^i, u_{j+1} can be updated in the outer loop. Then the algorithm goes to the inner loop iteration again, until the value function converges to the optimum, u_j converges to u , and w converges to w . In [30], the non-affine nonlinear two-person-zero-sum differential game problem in finite time domain is discussed, in which the non-affine nonlinear game problem is decomposed into a series of linear game problems. It is worth noting that the above researches are based on the assumption of the existence of saddle point. However, in practice, the saddle point may not exist for some nonlinear two-person-zero-sum game problems, namely $\overline{V}(x) \neq \underline{V}(x)$. So we have to obtain the hybrid optimal solution. Reference [31] first concerned how to solve the hybrid optimal solution $V^{o}(x), \underline{V}(x) \leq V^{o}(x) \leq \overline{V}(x)$ x by using iteration ADP scheme, when the saddle point does not exist. Obviously, this method can also be applicable to the case of the existence of saddle point. In the practical application, the systems are usually needed to achieve a certain performance index in a finite time, such as the implementation of a stabilization problem or a tracking problem. But the existing results based on ADP scheme mostly concerned the approximate optimal control problem in infinite time domain. An ADP-based optimal control scheme is proposed in the finite time domain in [33−34] in order to deal with above problem. The optimal control policy can be obtained by an iteration method, which makes the cost function of the system close to the optimal value infinitely in a bound ε and the number of optimal control steps are also obtained. The optimal control problem in finite time domain opens up a new field for the study of ADP scheme which has yet to be further studied. For example, the stabilization or tracking problem of continuous systems and time delay systems needs to be solved in finite time domain.

1.2.2. Online adaptive algorithm

In recent years, some new ADP algorithms have been presented by researchers. These algorithms no longer use offline iteration, but take online adaptive way to achieve the optimal control solution [35−38], which overcomes the shortcoming that once the system's parameters change, the iteration algorithm needs to be calculated offline again. Vamvoudakis and Lewis proposed an online adaptive update algorithm based on policy iteration to solve the optimal control problem of continuous nonlinear systems, and proved the stability of this online adaptive algorithm in theory in [35−36]. The kind of online adaptive algorithms also was applied to discrete systems. Using the online adaptive algorithm, [37−38] studied the optimal stabilization problem and the optimal tracking problem of a class of discrete affine nonlinear systems.

Now we introduce the basic principle of the online adaptive algorithm for continuous systems. The basic ideas for discrete systems are similar to that for continuous systems. Considering the space limitation, the detailed description is omitted here.

Consider a continuous affine nonlinear system described as (8), with the cost function as (6). The corresponding Hamiltonian function is

$$H\left(x,u,J_x\right)=l(x,u)+J_x^{\mathrm{T}}(f(x)+g(x)u) \qquad (23)$$

where J_x is the partial derivative of the cost function J(x) with respect to x. When the control policy and cost function get the optimal value, the HJB equation is satisfied, that is $H\left(x,u^*,J_x^*\right)=0$. Generally we construct the critic network and action network by neural networks, respectively, shown as

$$J(x)=W_c^{\mathrm{T}}\phi_c(x)+\varepsilon_c \qquad (24)$$

$$u(x) = W_a^T \phi_a(x) + \varepsilon_a \tag{25}$$

where W_c and W_a are the neural network target weights, $\varphi_c(\cdot)$ and $\varphi_a(\cdot)$ are the activation functions, ε_c and ε_a are the bounded approximation errors of neural network.

The actual output of the critic network is expressed as $\hat{V}(x) = \hat{W}_c^T \phi_c(x)$. Substituting it into Hamiltonian function (23), we can get $H\left(x, u, \hat{W}_c\right) = e_c$. The goal of critic network is to make $e_c = 0$, so as to satisfy HJB equation, which realizes that the output of critic network approximates the optimal value of the cost function.

The actual output of the action network is expressed as $\hat{V}(x) = \hat{W}_c^T \phi_c(x)$. The goal of action network is to realize that the output approaches the approximate optimal control policy, which is decided by the output of critic network, where $\nabla \phi_c(x)$ is the partial derivative of $\phi_c(x)$ with respect to x. The difference between the outputs of these two networks is defined as e_a. Then the target of action network can be expressed as $e_a = 0$. Based on the above ideas, we need to design weights updating rules of critic network and action network and make the weights of two networks to update together.

The weights of neural network are adjusted by using the online adaption algorithm. Over time, the weights of neural network converge finally, which realize that the output of critic network gradually approximates the optimal cost and the output of action network approximates the optimal control policy. We can prove the convergence of the weights and the stability of the system based on the online adaptive algorithm by Lyapunov theorem.

In order to relax the requirements that the system model is completely or partially known, Dierks et al. proposed an online system identification scheme for discrete affine nonlinear systems in [39]. In this scheme, the system dynamics were reconstructed by the identification structure of neural network based on the uniform approximation of neural network, as shown in Fig. 4.

Then we can use the ADP-based method to search the solution of optimal control policy. Zhang et al. proposed a data-driven robust approximate optimal tracking policy for non-affine nonlinear systems in [40]. The data-driven model with the available data is used to reconstruct system dynamics. Based on the establishment of data-driven model, the approximate optimal tracking policy can be solved by the online adaptive algorithm. And a novel robust term is designed first to ensure that the tracking error asymptotically converges to zero.

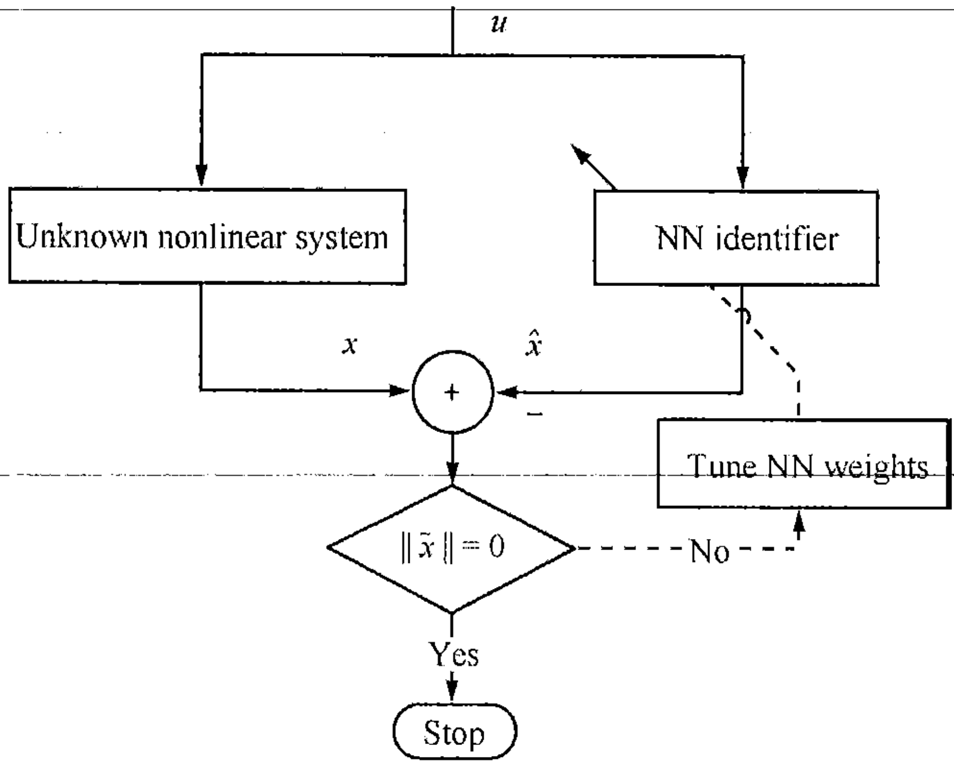


Fig. 4 NN identification structure

In the same way, to solve differential game problem with ADP-based method has come to online learning direction. Reference [42] presented an online adaptive control scheme based on policy iteration for the multi-personnon-zero-sum differential game problem. The cooperations and competitions coexist for each controller in the multi-player-non-zero-sum games, which leads to the intercoupling Hamilton-Jacobi (HJ) equation. The scheme proposed by Vamvoudakis et al. can approximate optimal strategies and Nash equilibrium point in real time based on the online adaptive algorithm. For each controller, there is a corresponding critic network and action network. These networks update synchronously, and also guarantee the stability of the overall closed-loop system.

It is worth noting that the critic network and action network are both used in above methods. Besides, it requires a given initial stable control in general in order to guarantee the stability when the system is operating online. To relax these two conditions, [41] proposed online control scheme with a single network, which is used to deal with the optimal control problem of continuous affine nonlinear systems.In the scheme, only critic network is adopted to approximate the cost function of the system and the control network is omitted. The estimated value of the optimal control policy can be obtained directly through the calculation of the equation $\hat{u} = -R^{-1} g^{T}(x) \nabla \phi_c^{T}(x) W_c / 2$ and the output of critic network. A new parameter training algorithm was presented in [41], which can remove the requirements of the initial stability. The boundedness of system states can be guaranteed in the process of online learning in the circumstance that the initial control is not an admissible control.

1.3. Application of ADP

Comparing with existing other optimal control methods, ADP scheme has its own unique algorithm and structure, which overcomes the shortcoming that the classical variational theory cannot deal with optimal control problem with the constraint condition of control variables in closed set. Same as maximum principle, ADP scheme is not only suitable for dealing with optimal control problem with the condition of open set constraints, but also of closed set constraints. In fact, maximum principle just offers the necessary condition of optimal control problem, while DP and ADP scheme offers the sufficient conditions of optimal control problem. However, it is very difficult to apply DP scheme directly because of the difficulty to solve HJB equation and "curse of dimensionality" problem. Therefore, as the approximate solution of DP scheme, ADP scheme overcomes DP scheme limitations, so it is more suitable for application to the system with strong coupling, strong nonlinearity, and high complexity.

Power system belongs to a kind of highly complex multivariate nonlinear systems, which is difficult to control. Though the dynamic characteristics change obviously with the load changing in power system, we have to ensure the stability of the power system in the time-varying operation condition (or in case of different faults). Specially with the development of smart grid, the traditional linearization method cannot fully satisfy the new demands. It is urgent to develop smart node, which can realize global optimum and coordinate control (including breaker, recloser device, transformer substation, etc). The optimization control based on ADP scheme offers the theoretical foundation, which has been applied successfully in recent years. For example, [43] applied HDP scheme to the real-time control of a single turbo generator power system, which overcomes the shortcomings that the performance of operating power system cannot be guaranteed by using the lead-lag compensator based on traditional phase complement theory in the frequency domain. Reference [44] applied ADP scheme to the control of synchronous generator, which replaces the traditional automatic voltage regulator. Reference [45] applied HDP method into the excitation control of generators in the multi-machine power system. Reference [46] applied DHDP method into static reactive power compensator to realize additional damping control. Reference [47] proposed the DHDP structure and algorithm based on iterative PID neural network. This structure can utilize existed PID parameters to serve for the selection of initial values. The simulation of additional damping control of the static var compensator in the 4 machines 2 regions system showed that the proposed algorithm can restrain the low frequency oscillation of inter-power-grid.

The research of intelligent traffic system is a hot topic in the field of optimal control as well. The traffic signal control system of intelligent traffic system is a complex large scale nonlinear system, which not only includes the block traffic system consisting of intersection signal light regulation, but also the urban traffic network coupled with on/off-ramp of city express way and block traffic network. In recent years, ADP scheme has been preliminarily applied to the optimization control of a single intersection traffic signal and on-ramp signal of the express way [48-53]. At present, the advanced urban traffic signal control and management system adopts the hierarchical distributed control scheme.

The upstream or downstream traffic states are acquired by the communication among the multi-agents, which can be utilized to construct their own performance index function. Thus, the mutual coordination and overall performance optimization are realized in the process of learning and optimizing of above performance index function[54]. Besides, ADP-based optimization control has been applied successfully in the field of avigation systems [55] aircrafts [56–57], communication systems[58], etc.

1.4. Conclusion

The optimal control of nonlinear systems has been one of the hot and difficult topics in control field. ADP scheme combines with neural network, adaptive evaluation design, enhancement learning, and classical dynamic programming theory as a new method of approximate solution in optimal control problem. ADP scheme can overcome the "curse of dimensionality" of DP scheme and obtain the approximate optimal closed-loop feedback control law. So it is considered as an effective method of solving optimal control ofnonlinear systems and attracts a lot of researchers' attention. Therefore, further research of ADP theory and its algorithm has important theoretical significance and practical value for solving the optimal control problem of nonlinear systems. The study of ADP scheme is still in the rise period. We hope that the readers would have a preliminary understanding about ADP scheme through this paper, and ADP scheme would be applied to solve various optimization problems in science and engineering fields.

CHAPTER 2

ADAPTIVE DYNAMIC PROGRAMING BASED OPTIMAL CONTROL FOR A ROBOT MANIPULATOR

In recent years, the control methodology for robotic systems has been widely developed not only in practical applications [59,60], but also in theoretical analysis [61-64]. The main challenges of the control design have been considered, such as robust adaptive control problem, motion/force control, input saturation and full state constraints [65,66] and the path planning problem [67]. Several control techniques have been employed for manipulators to tackle the issue of input saturation by adding more terms into the designed control input considering the absence of input Constraint [62], [63], [67-71]. In [62], authors proposed a new reference of control system due to the input saturation. The additional term world be computed based on the derivative of previous Lyapunov candidate function along the state trajectory under the control input saturation [62]. Furthermore, authors in [63] give a new approach to address the input constraints as well as combining with handling the disturbances. The proposed sliding surface was employed the Sat function of joint variables. In order to realize the disadvantage of state constraints in manipulator, the authors in [65,66] proposed the framework of Barrier Lyapunov function and Moore-Penrose inverse, Fuzzy-Neural Network technique. The equivalent sliding mode control algorithm was designed then the boundedness of control input was estimated. The advantage of this approach is that input boundedness absolutely adjusted by selecting several parameters. The work in [68-71] presents a technique to implement the input constraint using a modified Lyapunov Candidate function. Because of the actuator saturation, the Lyapunov function would be added more the quadratic term from the difference between the control input from controller and the real signal applied to object. The control design was obtained after considering the Lyapunov function derivative along the system trajectory. However, these aforementioned traditional nonlinear techniques have several drawbacks, such as difficulties in finding equivalent Lyapunov function, dynamic of additional terms. Optimization

Technique using GA (genetic algorithm), PSO (particle swarm optimization) were adressed to solve the papth planning problem [67]. The MPC (model predictive control) solution, which is the special case of optimal control design, has been investigated for linear motor not only online min-max technique in [72,73] but also offline algorithm in [74]. In order to consider for robot manipulators. Optimal control algorithm obtains the control design that can tackle the input, state constraint based on considering the optimization problem in presence of constraint. An asymptotic optimal control design was presented in [61] by solving directly the Riccati equation in linear systems. However, it is difficult to find the explicit solution of Riccati equation as well as partial differential HJB (Hamilton-Jacobi-Bellman) equation in general case. The approximate/adaptive dynamic programming (ADP) has been paid much attention for optimal control problem in recent years because it is necessary to solve not only Riccati equation for linear systems but also HJB equation for nonlinear systems. Thanks to Kronecker product technique, authors in [75] proposed the online solution for linear systems without the knowledge of system matrix based on the least-squares solution from acquisition of a sufficient number of data points. In [76], Zong-Ping Jiang et al. extend the above online solution to obtain the completely unknown dynamics by means that does not depend on either matrix A or matrix B of linear systems. The fact that Riccati equation was considered in more detail in the computation problem as well as data acquisition. Moreover, the exploration noise on the time interval was mentioned in proposed algorithm [76]. Instead of the approach of employing Kronecker product for the case of linear systems, the neural network approximation was mentioned for cost function to implement online adaptive algorithm on the Actor/Critic structure for continuous time nonlinear systems [77]. However, the proposed algorithm required the knowledge of input-to-state dynamics to update the control policy as well as persistent condition was not considered [77]. The weight parameters in neural network were tuned to minimize the objective in the least-squares sense [77]. The theoretical analysis about convergence of cost function and control input in adaptive/approximate dynamic programming (ADP) was the extension of the work in [78]. Thanks to the theoretical analysis about the neural network approximation, authors in [79] presented the novel online ADP algorithm which enables to tune simultaneously both actor and critic neural networks. The weights training problem of critic neural network (NN) was implemented by modified Levenberg-Marquardt algorithm to minimize the square residual error. Moreover, the tuning of weights in actor and critic NN depend on each other to obtain the weights convergence. It is worth noting that the persistence of excitation (PE) condition need to be satisfied and Lyapunov stability theory was employed to analysis the convergence problem [79]. Extension of the work in [79], based on the analysis of approximate Bellman error, the proposed algorithm in [80] enables to online simultaneously implement without the knowledge of drift term. In [81], the identifier along with adaptation law can be described using a Neural Network to approximate the dynamic uncertainties of nonlinear model. An extension using special cost function has been proposed in [82,83] to enable handling of input constraint. The framework of ADP technique and classical sliding mode control was presented to design the optimal control for an inverted pendulum [84]. However, the effectiveness of ADP has been still not considered for a robot manipulator in aforementioned researches. This work proposed the control algorithm combining exact linearization, Robust Integral of the Sign of the Error (RISE [61]) and ADP technique for manipulators in absence of holonomic constraint. This ADP technique was implemented using simultaneous tuning method to satisfy the weight convergence and stability.

2.1. Dynamic Model of a Robot Manipulator and Control Objective

Consider the following robot manipulator without constraint:

$$M(q)\ddot{q}+C(q,\dot{q})\dot{q}+G(q)+F(\dot{q})+\tau_d(t)=\tau \qquad (1)$$

Several appropriate assumptions [61] will be considered to develop the control design in next chapters.

Assumption 1. The inertia matrix M(q) is symmetric, positive definite, and guarantees the inequality $\forall \xi(t) \in \mathbb{R}^n$ as follows:

$$m_1\|\xi\|^2 \le \xi^T M(q)\xi \le \bar{m}(q)\|\xi\|^2, \tag{2}$$

where $m_1 \in \mathbb{R}$, $\bar{m}(q) \in \mathbb{R}$, $\|\cdot\|$ is a known positive constant, a known positive function, and the standard Euclidean norm, respectively.

Assumption 2. The relationship between an inertia matrix M(q) and the Coriolis matrix $C(q,\dot{q})$ can be represented as follows:

$$\xi^T(\dot{M}(q) - 2C(q,\dot{q}))\xi = 0 \ \ \forall \xi \in R^n. \tag{3}$$

It should be noticed that this manipulator is considered in the absence of holonomic constraint force. The control objective is to find the control algorithm being the framework of exact linearization, RISE and ADP technique enabling the position tracking control in manipulators control system (Fig. 1). ADP algorithm will be employed to implement optimal control design as desribed in next chapter.

2.2. Adaptive Dynamic Programming Approach for a Robot Manipulator

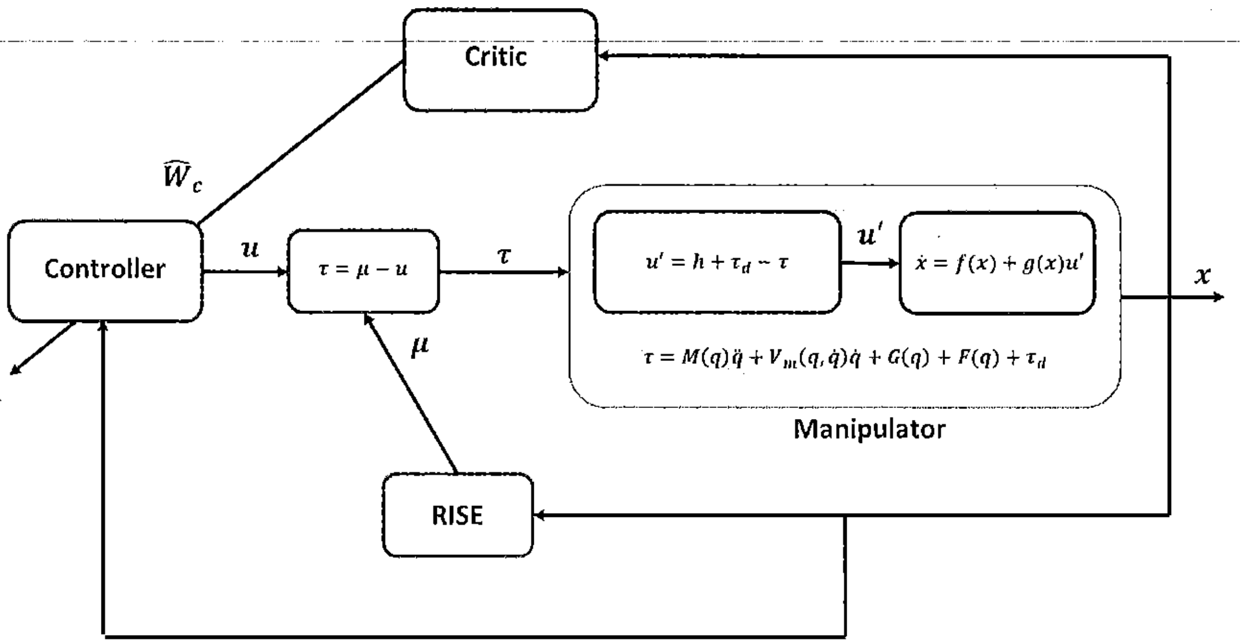


Fig. 1. Control Structure

2.2.1 ADP Algorithm

In [61], by using the control input (4) for manipulator (1) with nonlinear function (5) obtaining from (6), (7), (8), we lead to the nonlinear model (9):

$$u = -\tau + h + \tau_d \tag{4}$$

$$h = M(\alpha_1 \dot{e}_1) + C(\alpha_1 e_1) + G(q) + F(\dot{q}) \tag{5}$$

$$e_1 = q_d - q \tag{6}$$

$$e_2 = \dot{e}_1 + \alpha_1 e_1 \tag{7}$$

$$r = \dot{e}_2 + \alpha_2 e_2 \tag{8}$$

$$\dot{x} = f(x) + g(x)u \tag{9}$$

where

$$x = \begin{bmatrix} e_1 \\ e_2 \end{bmatrix}, f(x) = \begin{bmatrix} -\alpha_1 & I_{n\times n} \\ 0_{n\times n} & -M^{-1}C \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \end{bmatrix} \text{ and } g(x) = \begin{bmatrix} 0_{n\times n} \\ -M^{-1} \end{bmatrix}.$$

Now, the control object is to design a control law u to guarantee not only stabilization (9) but also minimizing the quadratic cost function with infinite horizon as follows:

$$V(x_0) = \int_0^\infty r(x,u)\,dt \tag{10}$$

$$r(x,u) = Q(x) + u^T R u \tag{11}$$

In which, $Q(x)$ and R is positive definite function of x, symmetric definite positive matrix, respectively.

This work presents a solution for approximate approach called adaptive dynamic programming (ADP) for optimal control design. In [79,80], consider the following affine system.

$$\dot{x} = f(x) + g(x)u \tag{12}$$

where $x \in \chi \subseteq R^n$, $u \in U \subseteq R^m$. $f(x)$ and $g(x)$ satisfy Lipschitz condition and $f(0) = 0$.

The cost function is defined as (10). The next definition was given in [75,76] to show that the optimal control solution will be considered in the set of admissible control.

Definition 1: A control policy $\mu(x)$ is defined as admissible policy if $\mu(x)$ stabilize system (12) and the equivalent value function $V^{\mu}(x)$ is finite. $\Psi(\chi)$ is denoted set of admissible control policy.

For any admissible policy $\mu(x)$, the nonlinear Lyapunov Equation (NLE) can be formulated

$$r\left(x,\mu(x)\right)+\left(\partial V/\partial x\right)^{T}\left(f\left(x\right)+g\left(x\right)\mu\left(x\right)\right)=0 \tag{13}$$

Defining Hamilton function and optimal cost function as follows:

$$H\left(x,\mu,V_x\right)=r\left(x,\mu\right)+\left(V_x^{\mu}\right)^{T}\left(f\left(x\right)+g\left(x\right)\mu\right) \tag{14}$$

$$V^{*}(x)=\min_{\mu\in\Psi(\chi)}\left(\int_{t}^{\infty}r(x,\mu)\right)$$

We lead to the following HJB equation:

$$0=\min_{\mu\in\Psi(\chi)}H(x,\mu,V_x^{*})=H(x,\mu^{*},V_x^{*}) \tag{15}$$

It can be noticed that, μ^{*} is optimal policy corresponding with the optimal cost function and $H(x,\mu,V_x^{\mu})=0$ with any admissible policy is NLE.

Now, the optimal control policy can be obtained by taking the derivative of Hamilton problem with respect to policy μ

$$\mu^{*}=-\frac{1}{2}\left(R^{-1}g^{T}V_x^{*}\right) \tag{16}$$

This work present Policy Iteration (PI) algorithm for a robot manipulator including 2 steps as follows:

Initiate admissible control policy $\mu^{0}(x)$

Repeat

Step 1: Policy Evaluation

Solve NLE for $V^{i}(x)$ corresponding given control policy μ^{i}

$$r\left(x,\mu^i(x)\right)+\left(V_x^i\right)^T\left(f(x)+g(x)\mu^i(x)\right)=0 \quad (17)$$

Step 2: Policy improvement

Update new policy according to

$$\mu^{i+1}=-\frac{1}{2}\left(R^{-1}g^TV_x^i\right) \quad (18)$$

Until $n=n_{max}$ or $\left|V^{i+1}-V^i\right|\le\varepsilon_v$.

Where n_{max} is a number of limited iteration and ε_v is an arbitrary given small positive number.

This algorithm is considered in [79] that prove each policy control μ^i is admissible control. The cost function V^i was reduced at each step until converge to optimal policy and μ^i converge toward optimal policy as well.

However, the nonlinear Lyapunov equation (17) is hard to solve directly. Therefore, in recent years, finding an indirectly way to solve this equation has been concerned by many researches [78-83]. In the next steps, two neural networks called Actor-Critic (AC) are trained simultaneously to solve approximately the HJB equation.

The cost function and its associated policy can be represented by using a neural network (NN) as follows

$$\begin{cases} V^*=W^T\phi(x)+\varepsilon_v \\ u^*=-\frac{1}{2}R^{-1}g^T\left(\nabla\phi(x)\right)^TW+\varepsilon_a \end{cases} \quad (19)$$

Where, $\phi(x)$ is corresponding function of NN that usually being selected as polynomial, Gausses, sigmoid function and so on. ∇ is denoted $\partial/\partial x$.

Approximated optimal cost function and optimal policy are presented:

$$\begin{cases} \hat{V}=\hat{W}_c^T\phi(x) \\ \hat{u}=-\frac{1}{2}R^{-1}g^T\left(\nabla\phi(x)\right)^T\hat{W}_a \end{cases} \quad (20)$$

Note that, to approximate HJB solution, we need to find only term $\hat{W}_c$. However, to stabilize closed-loop system, both $\hat{W}_a$, $\hat{W}_c$ are employed, which leads to the flexibility that can help handling the stability of system in learning process.

By replacing the optimal policy and the optimal cost function and by Actor-Critic networks in HJB equation (17), HJB error can be obtained.

$$Q(x)+\hat{u}^T R\hat{u}+\hat{W}_c^T\nabla\phi\left(f(x)+g(x)\hat{u}\right)=\varepsilon_{hjb} \tag{21}$$

$$Q(x)+\frac{1}{4}\hat{W}_a^T\nabla\phi^T G\nabla\phi\hat{W}_a+\hat{W}_c^T\nabla\phi\left(f(x)-\frac{1}{2}gR^{\text{l}}g^T\nabla\phi\hat{W}_a\right)=\varepsilon_{hjb} \tag{22}$$

Where $G=g^T R^{-1}g$.

The tuning law for $\hat{W}_c$ is described as follows

$$\dot{\hat{W}}_c=-\eta c\Gamma\frac{\omega}{1+\upsilon\omega^T\Gamma\omega}\varepsilon_{hjb} \tag{23}$$

$$\dot{\Gamma}=-\eta_c\Gamma\frac{\omega\omega^T}{1+\upsilon\omega\Gamma\omega^T}\Gamma \tag{24}$$

$\Gamma(t_r^+)=\Gamma(0)=\varphi_0 I$. Where t_r^+ is resetting time. To avoid slow convergence on $\hat{W}_c$, the matrix Γ is considered with default matrix $\Gamma(0)$ when minimum eigenvalue of Γ reach a given small positive number. $\omega(x)=\nabla\phi^T\left(f(x)+g(x)u\right)$ and $1+\upsilon\omega^T\Gamma\omega$ is normalization factor.

To make sure the convergence of $\hat{W}_c$ with update law (24), $\omega(x)$ must satisfy the Persistence Excitation (PE) condition [79].

$$\mu_1 I\ge\int_{t_0}^{t_0+T}\psi(\tau)\psi(\tau)^T d\tau\ge\mu_2 I \tag{25}$$

for several positive numbers μ_1, μ_2, T.

Where $\psi(\tau)=\dfrac{\omega(t)}{\sqrt{1+\upsilon\omega^T\Gamma\omega}}$.

On the other hands, (22) is nonlinear equation of $\hat{W}_a$. Therefore, the tuning law for $\hat{W}_a$ is formulated based on GD algorithm to minimize the cost $\left(\varepsilon_{hjb}(t)\right)^2$.

$$\dot{\hat{W}}_a = proj\left\{-\eta_{a1}\frac{1}{\sqrt{1+\omega^T\omega}}\nabla\phi G\nabla\phi^T\left(\hat{W}_a-\hat{W}_c\right)\varepsilon_{HJB}-\eta_{a2}\left(\hat{W}_a-\hat{W}_c\right)\right\} \quad (26)$$

Where $proj\{\bullet\}$ is a projection operator [22] that ensure the boundedness of updatation law.

Note that, these parameters of both two NN's update law η_c, η_{a1}, η_{a2} must be selected to satisfy some conditions [22] to ensure stability of closed-loop system. One can also find the complete proof of convergence of parameters and stability of system in [80].

2.2.2 RISE Feedback Control Design

In [3], the control term μ(t) is designed based on the RISE framework as follows:

$$\mu(t) \triangleq (k_s+1)e_2(t)-(k_s+1)e_2(0)+\upsilon(t) \quad (27)$$

Where $\upsilon(t)\in\mathbb{R}^n$ is described as:

$$\dot{\upsilon}=(k_s+1)\alpha_2 e_2+\beta_1 sgn(e_2) \quad (28)$$

$k_s\in\mathbb{R}$ is positive constant control gain, and $\beta_1\in\mathbb{R}$ can be selected being a positive control gain selected according to the following sufficient condition

$$\beta_1 > \zeta_1+\frac{1}{\alpha_2}\zeta_2 \quad (29)$$

Remark 1: It is different from the work in [3], in our work the ADP algorithm is presented to find the intermediate optimal control input in the absence of dynamic uncertainty. Furthermore, ADP technique was considered in [78-84] was still not to apply for a robotic manipulator.

Remark 2: In compare with the work of Dixon [61] that design optimal control solving Riccati equation, this work requires partial knowledge of manipulator's dynamic including matrices M, C. However, using the ADP approach, the optimal control problem is addressed in general case for any given cost function as (10) without constraint.

2.3. Offline Simulation Results

Consider the offline simulation of a two-link manipulator control system using ADP technique and RISE algorithm.

The general dynamic of two-link manipulator is represented by (1) with

$$M = \begin{bmatrix} 5 + 2\cos(q_2) & 1 + \cos(q_2) \\ 1 + \cos(q_2) & 1 \end{bmatrix},$$

$$C = \begin{bmatrix} -\dot{q}_2 \sin(q_2) & -(\dot{q}_1 + \dot{q}_2)\sin(q_2) \\ \dot{q}_1 \sin(q_2) & 0 \end{bmatrix}$$

$$G = 9.8\begin{bmatrix} 1.2\cos(q_1) + \cos(q_1 + q_2) \\ \cos(q_1 + q_2) \end{bmatrix},$$

$$F = -0.1 sign(\dot{q}), \tau_d = \begin{bmatrix} 0.1\sin(t) \\ 0.1\cos(t) \end{bmatrix}.$$

Value function is (10) with the term: $Q(x) = x^T Q_0 x$.

$$Q_0 = \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix}, Q_{11} = \begin{bmatrix} 40 & 2 \\ 2 & 40 \end{bmatrix},$$

$$Q_{12} = Q_{21} = \begin{bmatrix} -4 & 4 \\ 4 & -6 \end{bmatrix}, Q_{22} = \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix},$$

$$R = \begin{bmatrix} 0.25 & 0 \\ 0 & 0.25 \end{bmatrix}, \alpha = \begin{bmatrix} 15.6 & 10.6 \\ 10.6 & 10.4 \end{bmatrix}$$

Without loss of generality, the set-point is selected as $q_d = \begin{bmatrix} 0 & 0 \end{bmatrix}^T$, initial state is $q_0 = \begin{bmatrix} 0.1598 & 0.2257 \end{bmatrix}^T$.

The optimal value function which is solved directly in [61] is

$$V^* = x^T \begin{bmatrix} -Q_{12} & 0_{n \times n} \\ 0_{n \times n} & M \end{bmatrix} x = 2x_1^2 - 4x_2^2 + 3x_1 x_2$$

$$+2.5x_3^2 + x_3^2 \cos(x_2) + x_4^2 + x_3 x_4 + 0.5 x_3 x_4 \cos(x_2)$$

The updatation law of $\dot{\hat{W}}_c$ and $\dot{\hat{W}}_a$ are represented in (23) and (26) with

$\eta_c = 800,\ \upsilon = 1,\ \Gamma(0) = 100,\ \varepsilon_T = 0.001,\ \eta_{a1} = 0.01,\ \eta_{a2} = 1$.

NN activation function is selected as

$\phi(x) = \begin{bmatrix} x_1^2 & x_2^2 & x_1x_2 & x_3^2 & x_3^2\cos(x_2) & x_4^2 & x_3x_4 & x_3x_4\cos(x_2) \end{bmatrix}^T$.

The optimal parameter $W = \begin{bmatrix} 2 & -4 & 3 & 2.5 & 1 & 1 & 1 & 0.5 \end{bmatrix}$ that is obtained by solving directly HJB equation in [3]. Fig (1) and (2) show the convergence of $\hat{W}_c$, $\hat{W}_a$. The value of $\hat{W}_c$ after 110s is $\begin{bmatrix} 2 & -4 & 3 & 2.5 & 1 & 1 & 1 & 0.5 \end{bmatrix}$. To satisfy PE condition as in (25), a probing signal is added in system input. Moreover, system's error evolution is shown in Fig (3) determining the stability of control system.

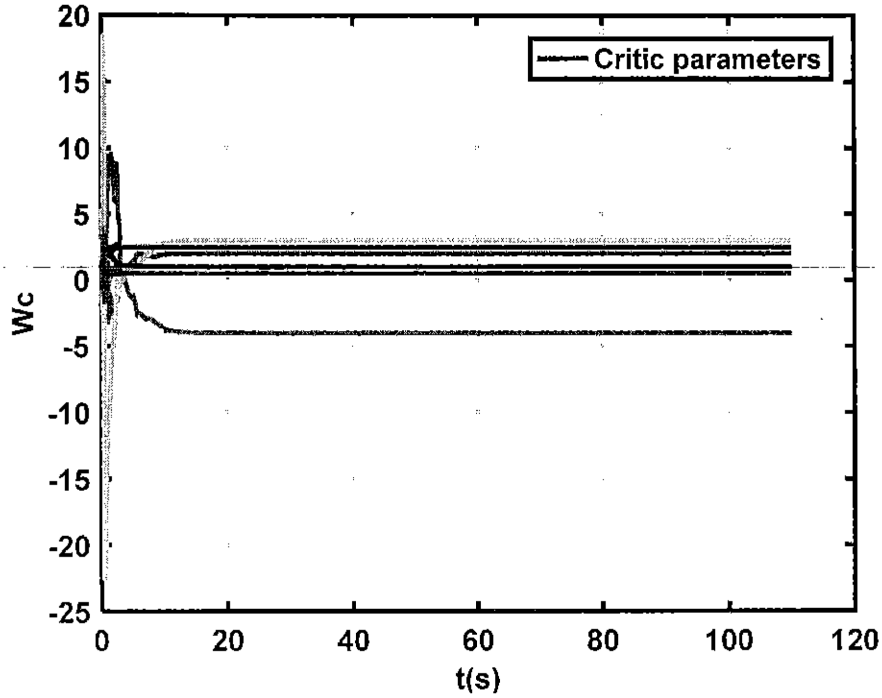


Fig 2. Convergence of Critic's parameters

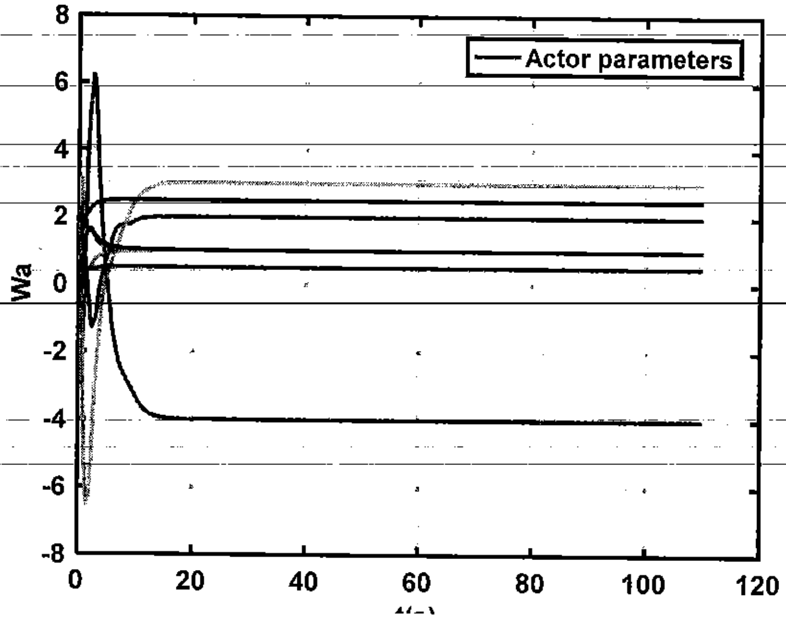


Fig 3. Convergence of Actor's parameters

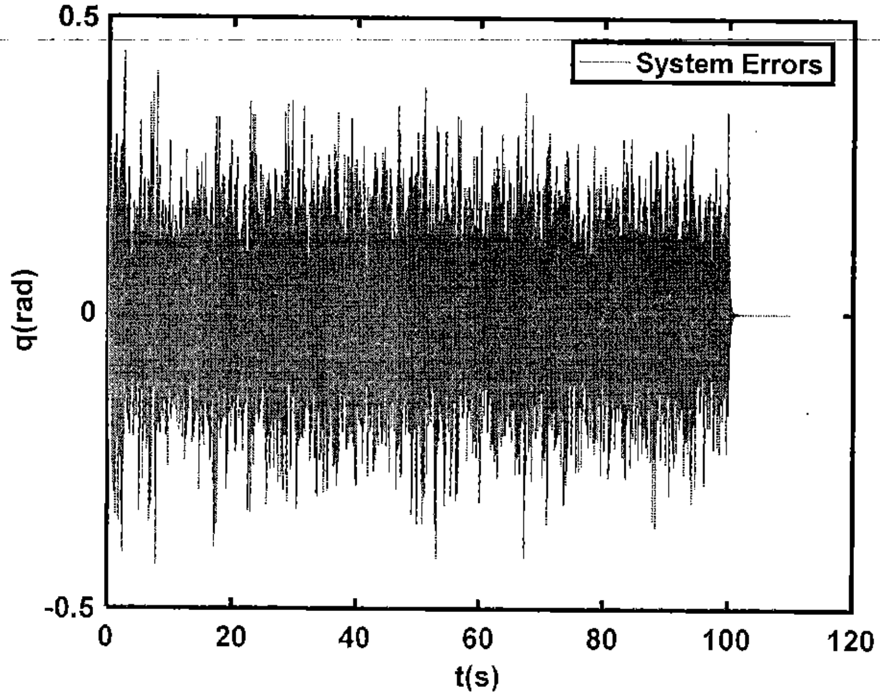


Figure 4. State's evolution

CHAPTER 3

ADAPTIVE DYNAMIC PROGRAMMING ALGORITHM FOR UNCERTAIN NONLINEAR SWITCHED SYSTEMS

It is worth noting that many systems in industry can be described by switched system such as DC/DC converter [85-86], H-bridge inverter [87], multilevel inverter [88], photovoltaic inverter [89] . Although many different approaches for switched systems have been proposed, e.g., switching-delay tolerant control [90,91], classical nonlinear control [92-96], the optimization approaches with the advantage of mentioning the input/state constraint has not been mentioned much. The approaches of fuzzy and neural network as well as ANN, particle swarm optimization (PSO) technique were investigated in several different systems such as photovoltaic inverter, transmission line,... [97-101]. Adaptive dynamic programming has been considered in many situations, such as nonlinear continuous time systems [102], actuator saturation [103], linear systems [104-106], output constraint [107]. In the case of nonlinear systems, the algorithm should be implemented based on Neural Networks (NNs). However, Kronecker product was employed in linear systems. Furthermore, the data driven technique should to be mentioned to compute the actor/critic precisely. It should be noted that the robotic systems has been controlled by ADP algorithm [108-109]. Our work proposed the solution of adaptive dynamic programming in nonlinear perturbed switching systems based on the neural networks. The consideration of the Halminton function enables us obtaining the learning technique of these neural networks. The UUB stability of closed system is analyzed and simulation results illustrate the high performance of the given controller.

3.1. Problem Statement

Consider the following uncertain continuous time nonlinear switched systems of the form:

$$\dot{x} = f_i(x) + g_i(x)\left(u + \Delta(x,t)\right) \qquad (1)$$

Where $x(t) \in \Omega_x \in \mathbb{R}^n$ denotes the vector of state variables, $u(t) \in \Omega_u \in \mathbb{R}^m$ is the vector of control inputs. The function $\sigma : [0, +\infty) \mapsto \Omega = \{1, 2, ..., l\}$ is a signal of switching processing, which is a piecewise continuous function with respect to time, and l is the subsystems number $f_i(x)$ are unknown smooth vector functions with $f_i(0) = 0$. $g_i(x)$ are known smooth vector functions such that $G_{\min} \leqslant \|g_i(x)\| \leqslant G_{\max}$.

The switching index $\sigma(t)$ is unknown.

Assumption 1: $\Delta(x,t)$ is bounded by a known function $\rho(x)$ as $\|\Delta(x,t)\| \leqslant \rho(x)$

Consider the performance index for the uncertain switched system (1):

$$J(x(t), u(t)) = \int_t^{\infty} r(x(\tau), u(\tau)) d\tau \tag{2}$$

Where $r(x(\tau), u(\tau)) = x^T Q x + u^T R u$ and $Q = Q^T > 0; R = R^T > 0$

The control objective is to design the state feedback controller and give the upper bound function to guarantee that the closed systems under this controller is robustly stable. Additionally, the performance index (2) is bounded as $J(x,u) \leq K(x,u) \leq M$

Definition: The function K(u) can be known as the guaranteed cost function. Therefore, the control law u^* with $u^* = \arg\min_{u \in \Omega_u} K(x,u)$ is known as the optimal guaranteed cost control law.

3.2. Control Design

The obtained nominal system after eliminating the disturbance in the system (3) is described by:

$$\dot{x} = f_i(x) + g_i(x) u \tag{3}$$

The performance index of system (3) is defined as:

$$J_1(x(t), u(t)) = \int_t^{\infty} \left[r(x(\tau), u(\tau)) + \lambda (\rho(x))^2 \right] d\tau \tag{4}$$

We prove that $J_1(x(t), u(t))$ with $\lambda \geqslant \|R\|$ is the one of guaranteed cost function of system (1). Define: $V^*(t) = \min_{u \in \Omega_u} J_1(x(t), u(t))$, we have:

$$V^*(t)=\min_{u\in\Omega_u}\int_t^{\infty}\left[r(x(\tau),u(\tau))+\lambda(\rho(x))^2\right]d\tau \quad (5)$$

$$\begin{aligned}V^*(t)&=\min_{u\in\Omega_u}\int_t^{t+\Delta t}\left[r(x(\tau),u(\tau))+\lambda(\rho(x))^2\right]d\tau\\&+\min_{u\in\Omega_u}\int_{t+\Delta t}^{\infty}\left[r(x(\tau),u(\tau))+\lambda(\rho(x))^2\right]d\tau\end{aligned} \quad (6)$$

Consider the Halminton function obtaining the nominal system and performance:

$$H(x,u,V^*)=r(x(t),u(t))+\lambda\rho^2(x)+(\nabla V^*)^T(f_i(x)+g_i(x)u) \quad (7)$$

We consider that:

$$H(x,u^*,V^*)=\min_{u\in\Omega_u}H(x,u,V^*)=0 \quad (8)$$

$$\left.\frac{\partial H(x,u,V^*)}{\partial u}\right|_{u=u^*}=0\Rightarrow u^*=-\frac{1}{2}R^{-1}(g_i(x))^T\nabla V^* \quad (9)$$

We continue to utilize this control law (9) for nonlinear switched system (1) and obtain the following result:

Theorem 1: The system (1) under the state feedback control law $u(x)=-\frac{1}{2}R^{-1}(g_i(x))^T\nabla V$ is stable with the associated Lyapunov function candidate: $V(t)=\int_t^{\infty}\left[r(x(\tau),u(\tau))+\lambda(\rho(x))^2\right]d\tau$ where $\lambda\geqslant\|R\|$

Proof: The derivative of V is given by the following formula:

$$\dot{V}(t)=(\nabla V)^T(f_i(x)+g_i(x)(u+\Delta(x,t)))$$

By using $u(x)=-\frac{1}{2}R^{-1}(g_i(x))^T\nabla V^*$

we infer:

$$\begin{aligned}\dot{V}(t)&=-r(x(t),u(t))-\lambda\rho^2(x)+(g_i(x))^T\nabla V^*\Delta(x,t)\\&=-x^TQx-u^TRu-\lambda\rho^2(x)-2u^TR\Delta(x,t)\end{aligned} \quad (10)$$

$$=-x^TQx-\lambda\rho^2(x)-(u^T+\Delta(x,t)^T)R(u+\Delta(x,t))+\Delta(x,t)^TR\Delta(x,t) \quad (11)$$

$$= -x^T Qx - \left(\lambda\rho^2(x) - \Delta(x,t)^T R\Delta(x,t)\right) - \left(u + \Delta(x,t)\right)^T R\left(u + \Delta(x,t)\right) \quad (12)$$

From assumption 1 and $\lambda \geqslant \|R\|$

we have:

$$\dot{V}(t) \leqslant -x^T Qx \quad (13)$$

Therefore, the system (1) is robustly stable. It is impossible to solve the analytic solution of HJB nonlinear equation (13). Hence, the optimal performance for system (3) can be described based on a neural network as follows:

$$V^* = w^T \sigma(x) + \varepsilon(x) \quad (14)$$

Where $\sigma(x): R^n \to R^N; \sigma(0) = 0$ is the vector of NN activation function, N is the number of neurons in the hidden layer, and $\varepsilon(x)$ is the NN approximation error, $w \in R^N$ is the NN constant weight vector. $\sigma(x)$ can be selected such that when $N \to \infty$, we have: $\varepsilon(x) \to 0$ and $\nabla\varepsilon(x) \to 0$, so for the fixed N, we can assume that:

Assumption 2:

$\|\varepsilon(x)\| \leqslant \varepsilon_{max}; \|\nabla\varepsilon(x)\| \leqslant \nabla\varepsilon_{max}; \nabla\sigma_{min} \leqslant \|\nabla\sigma(x)\| \leqslant \nabla\sigma_{max}; \|w\| \leqslant w_{max}$

Combining two formulas (13) and (14) we infer:

$$H(x,u^*,V^*) = r(x(t),u^*(t)) + \lambda\rho^2(x) + (\nabla V^*)^T \left(f_i(x) + g_i(x)u^*\right) = 0 \quad (15)$$

$$= x^T Qx + \lambda\rho^2(x) + (\nabla V^*)^T f_i(x) - \frac{1}{4}(\nabla V^*)^T g_i(x) R^{-1} g_i(x)^T (\nabla V^*) = 0 \quad (16)$$

$$\nabla V^* = \left(\nabla\sigma(x)\right)^T w + \nabla\varepsilon(x) \quad (17)$$

Obtain the NN based HJB equation as follows:

$$e_{NN} = x^T Qx + \lambda\rho^2(x) + \left(\nabla\sigma(x)^T w\right)^T f_i(x) - \frac{1}{4}\left(\nabla\sigma(x)^T w\right)^T g_i(x) R^{-1} g_i(x)^T \left(\nabla\sigma(x)^T w\right) \quad (18)$$

$$e_{NN} = x^T Qx + \lambda\rho^2(x) + w^T \nabla\sigma(x) f_i(x) \\ -\frac{1}{4} w^T \nabla\sigma(x) g_i(x) R^{-1} g_i(x)^T \nabla\sigma(x)^T w \tag{19}$$

The residual error formed by the function approximation error:

$$e_{NN} = -\nabla\varepsilon(x)^T \left(f_i(x) + g_i(x)u^*\right) \\ +\frac{1}{4}\nabla\varepsilon(x)^T g_i(x) R^{-1} g_i(x)^T \nabla\varepsilon(x) \tag{20}$$

It follows that e_{NN} converges uniformly to zero as $N \to \infty$. For each number N, e_{NN} is bounded on a region as $e_{NN} \leqslant e_{max}$. Under the structure of ADP-based optimal controller, a critic neural network is given the following estimated weight vector $\hat{w}$:

$$\hat{V} = \hat{w}^T \sigma(x) = \sigma(x)^T \hat{w}; \hat{u} = -\frac{1}{2} R^{-1} \left(g_i(x)\right)^T \nabla\hat{V} \tag{21}$$

The approximate error of the critic part:

$$e_{HJB} = r\left(x(t), \hat{u}(t)\right) + \lambda\rho^2(x) + \left(\nabla\sigma(x)^T \hat{w}\right)^T \left(f_i(x) + g_i(x)\hat{u}\right) \tag{22}$$

$$e_{HJB} = x^T Qx + \lambda\rho^2(x) + \hat{w}^T \nabla\sigma(x) f_i(x) \\ -\frac{1}{4}\hat{w}^T \nabla\sigma(x) g_i(x) R^{-1} g_i(x)^T \nabla\sigma(x)^T \hat{w} \tag{23}$$

The weight vector is trained based on a steepest descent method:

$$\frac{d}{dt}\hat{w} = -\alpha \frac{\partial E}{\partial \hat{w}} \tag{24}$$

with $E = \frac{1}{2} e_{HJB}^T e_{HJB}$

Remark 1: The weight $\hat{w}$ is trained to minimize the network error part $E = \frac{1}{2} e_{HJB}^T e_{HJB}$.

The fact is that:

$$\dot{E} = \frac{\partial E}{\partial t} = \frac{\partial E}{\partial \hat{w}} \cdot \frac{\partial \hat{w}}{\partial t} = \frac{\partial E}{\partial \hat{w}} \cdot \left(\frac{d}{dt} \hat{w} \right) = -\alpha \left(\frac{\partial E}{\partial \hat{w}} \right)^2 \quad (25)$$

Theorem 2: Consider the feedback controller in (21) and the weight vector of the critic part is updated by (25), the weight estimate error $\tilde{w} = w - \hat{w}$ and the closed-loop system's state vector $x(t)$ are uniform ultimate bounded.

Proof: Define: $\tilde{w} = w - \hat{w} \Rightarrow \dot{\tilde{w}} = -\dot{\hat{w}}$ Consider the Lyapunov function:

$V(t) = V_1(t) + V_2(t)$,

where: $V_1(t) = \frac{1}{2\alpha} \tilde{w}(t)^T \tilde{w}(t), V_2(t) = V^*$

For deriving the term $V_1(t)$, we obtain that:

$$\begin{aligned} \frac{d}{dt} V_1(t) &= \frac{1}{\alpha} \tilde{w}(t)^T \left(\frac{d}{dt} w(t) \right) \\ &= -\frac{1}{\alpha} \tilde{w}(t)^T \left(\frac{d}{dt} \hat{w}(t) \right) = \tilde{w}(t)^T \frac{\partial E}{\partial \hat{w}} \end{aligned}$$

$$\dot{V}_1 = \tilde{w}^T e_{HJB} \nabla \sigma(x) \left(f_i(x) + g_i(x) \hat{u} \right)$$

From (14), (25) we have:

$$\hat{u} - u^* = -\frac{1}{2} R^{-1} \left(g_i(x) \right)^T \left(\nabla \hat{V} - \nabla V^* \right)$$

$$\begin{aligned} &= -\frac{1}{2} R^{-1} \left(g_i(x) \right)^T \left(\nabla \sigma(x)^T \hat{w} - \left(\nabla \sigma(x) \right)^T w - \nabla \varepsilon(x) \right) \\ &= \frac{1}{2} R^{-1} \left(g_i(x) \right)^T \left(\left(\nabla \sigma(x) \right)^T \tilde{w} + \nabla \varepsilon(x) \right) \end{aligned}$$

Furthermore, we have:

$$\nabla \sigma(x) \left(f_i(x) + g_i(x) \hat{u} \right) = \nabla \sigma(x) \left(f_i(x) + g_i(x) u^* \right) + \nabla \sigma(x) g_i(x) \left(\hat{u} - u^* \right)$$

$$\nabla\sigma(x)\left(f_i(x)+g_i(x)\hat{u}\right) = \nabla\sigma(x)\begin{pmatrix} f_i(x)+g_i(x)u^* \\ +\frac{1}{2}g_i(x)R^{-1}g_i(x)^T\left(\nabla\sigma(x)^T\tilde{w}+\nabla\varepsilon(x)\right)\end{pmatrix} \tag{26}$$

we obtain:

$$\begin{aligned} &e_{HJB}-e_{NN} \\ &=-\tilde{w}^T\nabla\sigma(x)f_i(x)+\frac{1}{2}\tilde{w}^T\nabla\sigma(x)g_i(x)R^{-1}g_i(x)^T\nabla\sigma(x)^T w \\ &-1/4\tilde{w}^T\nabla\sigma(x)g_i(x)R^{-1}g_i(x)^T\nabla\sigma(x)^T\tilde{w}\end{aligned} \tag{27}$$

Because

$$u^*=-\frac{1}{2}R^{-1}\left(g_i(x)\right)^T\left(\left(\nabla\sigma(x)\right)^T w+\nabla\varepsilon(x)\right),$$

we have:

$$\begin{aligned} &e_{HJB}-e_{NN}= \\ &-\tilde{w}^T\nabla\sigma(x)f_i(x)+\tilde{w}^T\nabla\sigma(x)g_i(x)\left(-u^*-\frac{1}{2}R^{-1}g_i(x)^T\nabla\varepsilon(x)\right) \\ &-1/4\tilde{w}^T\nabla\sigma(x)g_i(x)R^{-1}g_i(x)^T\nabla\sigma(x)^T\tilde{w}\end{aligned} \tag{28}$$

Assumption 3: $\left\|f_i(x)+g_i(x)u^*\right\|\leqslant\mu_{\max}$

Define:

$$\mu_i=f_i(x)+g_i(x)u^*;G_i=g_i(x)R^{-1}g_i(x)^T;\nabla\sigma=\nabla\sigma(x);\nabla\varepsilon=\nabla\varepsilon(x)$$

we obtain:

$$\begin{aligned} &\dot{V}_1(t)= \\ &-\tilde{w}^T\begin{pmatrix}-e_{NN}+\tilde{w}^T\nabla\sigma\mu_i+\frac{1}{2}\tilde{w}^T\nabla\sigma G_i\nabla\varepsilon \\ +\frac{1}{4}\tilde{w}^T\nabla\sigma G_i\nabla\sigma^T\tilde{w}\end{pmatrix}\nabla\sigma(x)\left(\mu_i+\frac{1}{2}G_i\left(\nabla\sigma^T\tilde{w}+\nabla\varepsilon\right)\right)\end{aligned} \tag{29}$$

Define:

$$A=\tilde{w}^{T}\nabla\sigma G_{i}\nabla\sigma^{T}\tilde{w};B=\frac{3}{4}\left(\tilde{w}^{T}\nabla\sigma\mu_{i}\right)+\frac{1}{4}\left(\tilde{w}^{T}\nabla\sigma G_{i}\nabla\varepsilon\right)+\frac{1}{2}e_{NN};$$

$$C=\tilde{w}^{T}\nabla\sigma\mu_{i};D=e_{NN}-\frac{1}{2}\left(\tilde{w}^{T}\nabla\sigma G_{i}\nabla\varepsilon\right)$$

It is obvious that:

$$\begin{aligned}&\dot{V}_{1}(t)\\&=-\frac{1}{8}A^{2}-BA-C^{2}-DC=-\frac{1}{8}(A+4B)^{2}+2B^{2}-\left(C+\frac{D}{2}\right)^{2}+\frac{D^{2}}{4}\\&\leqslant-\frac{1}{8}\left[(A+4B)^{2}-\left(16B^{2}+2D^{2}\right)\right]\end{aligned}\tag{30}$$

We have:

$$\begin{aligned}&A+4B\geqslant\\&\|\tilde{w}\|^{2}\left(G_{\min}\right)^{2}\lambda_{\min}\left(R^{-1}\right)\left(\nabla\sigma_{\min}\right)^{2}\\&-\|\tilde{w}\|\left(3\nabla\sigma_{\max}\mu_{\max}+\nabla\sigma_{\max}\left(G_{\max}\right)^{2}\lambda_{\max}\left(R^{-1}\right)\nabla\varepsilon_{\max}\right)-2e_{\max}\end{aligned}\tag{31}$$

And

$$\begin{aligned}&16B^{2}+2D^{2}\leqslant\left(\|\tilde{w}\|\left(3\nabla\sigma_{\max}\mu_{\max}+\nabla\sigma_{\max}\left(G_{\max}\right)^{2}\lambda_{\max}\left(R^{-1}\right)\nabla\varepsilon_{\max}\right)+2e_{\max}\right)^{2}\\&+2\left(e_{\max}+\frac{1}{2}\|\tilde{w}\|\nabla\sigma_{\max}\nabla\sigma_{\max}\left(G_{\max}\right)^{2}\lambda_{\max}\left(R^{-1}\right)\nabla\varepsilon_{\max}\right)^{2}\end{aligned}\tag{32}$$

According to (29) and (30) the inequality

$$(A+4B)^{2}-\left(16B^{2}+2D^{2}\right)\geqslant\pi_{1}$$

with $\pi_{1}>0$ leads to other inequalities which having the polynomial quadratic form with the variable $\|\tilde{w}\|$ and the highest order coefficient

$$\left(\left(G_{\min}\right)^{2}\lambda_{\min}\left(R^{-1}\right)\left(\nabla\sigma_{\min}\right)^{2}\right)^{2}>0$$

so we can find the positive number ϑ_1 such that: $\forall \|\tilde{w}\| > \vartheta_1$

we have $(A+4B)^2 - (16B^2 + 2D^2) \geqslant \pi_1$,

we obtain: $\dot{V}_1(t) \leqslant -\pi_1$.

For the term $V_2(t)$, from (00) we have:

$$\begin{aligned}\dot{V}_2 &= (\nabla V^*)^T (f_i + g_i(\hat{u} + \Delta)) \\ &= -(x^T Q x + \lambda \rho^2(x)) - \frac{1}{4}(\nabla V^*)^T g_i R^{-1} g_i^T (\nabla V^*) \\ &+ \frac{1}{2}(\nabla V^*)^T g_i R^{-1} g_i^T (\nabla \sigma(x)^T \tilde{w} + \nabla \varepsilon(x)) + (\nabla V^*)^T g_i \Delta\end{aligned} \tag{33}$$

Assume that $\rho(x) = \varpi \|x\|$. From (30) we have:

$$\dot{V}_2 \leqslant -(\lambda_{\min}(Q) + \lambda \varpi)\|x\|^2 + \theta^2 \tag{34}$$

With

$$\begin{aligned}\theta^2 &= -\frac{1}{4}(\nabla V^*)^T g_i R^{-1} g_i^T (\nabla V^*) \\ &+ \frac{1}{2}(\nabla V^*)^T g_i R^{-1} g_i^T (\nabla \sigma(x)^T \tilde{w} + \nabla \varepsilon(x)) \\ &+ (\nabla V^*)^T g_i \Delta\end{aligned}$$

Based on the two above assumptions, we have:

$$\begin{aligned}\theta^2 &\leqslant \frac{1}{4}(w_{\max} \nabla \sigma_{\max} + \nabla \varepsilon_{\max})^2 g_{\max^2} \lambda_{\max}(R^{-1}) \\ &+ \frac{1}{2}(\vartheta \nabla \sigma_{\max} + \nabla \varepsilon_{\max})^2 g_{\max^2} \lambda_{\max}(R^{-1}) \\ &+ (w_{\max} \nabla \sigma_{\max} + \nabla \varepsilon_{\max}) g_{\max} \varpi \|x\|\end{aligned}$$

It is obvious that $(\lambda_{\min}(Q) + \lambda \varpi)\|x\|^2 - \theta^2 \geqslant \pi_2$ with $\pi_2 > 0$ also leads to other inequalities which having the polynomial quadratic form with the variable $\|x\|$ and the highest order coefficient $\left(\lambda_{\min}(Q)+\lambda\varpi\right)>0$ so we can find the positive number ϑ_2 such that: $\forall\|x\|>\vartheta_2$ we have $\left(\lambda_{\min}(Q)+\lambda\varpi\right)\|x\|^2-\theta^2\geqslant\pi_2$, we obtain:

$\dot{V}_2(t)\leqslant-\pi_2$

Remark 2: The numbers $\vartheta_1;\vartheta_2$ can be changed by renovating the neural network of the optimal performance index. Moreover, for any switching index, after $\dfrac{V(0)}{\min(\pi_1;\pi_2)}$ the variable $\|x\|$ and $\|\tilde{w}\|$ always are in the specified domains. The ADP control law $\hat{u}$ is proposed in (21), which tends to the neighborhood of the optimal controller u^*.

Proof: From (34) we have:

$$\|\hat{u}-u^*\|=\frac{1}{2}\left\|R^{-1}(g_i(x))^T\left((\nabla\sigma(x))^T\tilde{w}+\nabla\varepsilon(x)\right)\right\|$$
$$\leqslant\frac{1}{2}\lambda_{\max}(R^{-1}).G_{\max}.(\nabla\sigma_{\max}.\upsilon_1+\nabla\varepsilon_{\max})=\vartheta_3$$

Thus the proof is completed.

3.3. Simulation Results

In this section, we verify the effectiveness and performance of the proposed controller:

Let $N=2$ and the subsystems of the switched system are :

$$\begin{cases}\dot{x}_1=-x_1^3-2x_2+\left(u+\Delta_1(x,t)\right)\\ \dot{x}_2=x_1+0.5\cos(x_1^2)\sin(x_2^3)-\left(u+\Delta_1(x,t)\right)\end{cases}\qquad(35)$$

$$\begin{cases}\dot{x}_1=-x_1^5\sin(x_2)+\left(u+\Delta_2(x,t)\right)\\ \dot{x}_2=\dfrac{1}{2}x_1-\cos(x_1)\cos(x_2^3)-\left(u+\Delta_2(x,t)\right)\end{cases}\qquad(36)$$

The initial state vectors can be chosen : $x(0)=\begin{bmatrix}5 & -5\end{bmatrix}^T$

Choosing that the parameter matrices: $R=\begin{bmatrix}2&0\\0&2\end{bmatrix};Q=\begin{bmatrix}1&0\\0&3\end{bmatrix};\alpha=0.1;\lambda=5$

The simulation results shown in Fig.3.1, Fig.3.2 validate the effectiveness of proposed algorithm:

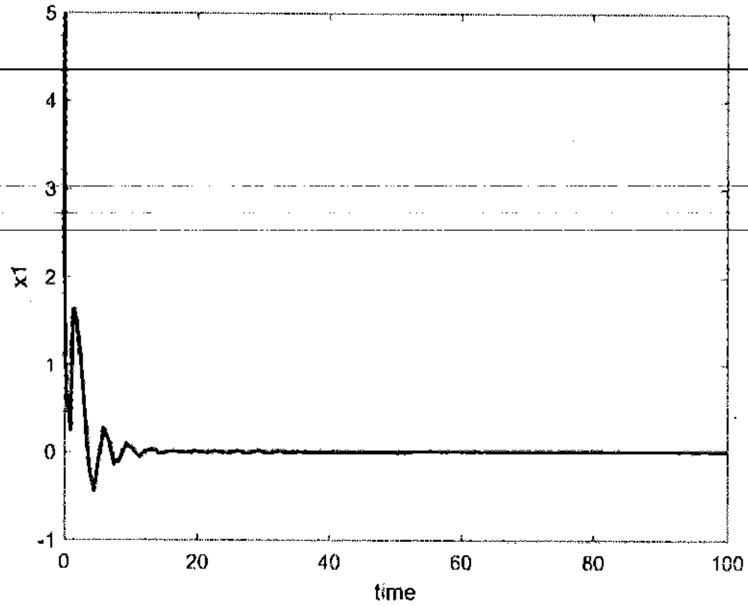


Figure 3.1. The response of x_1

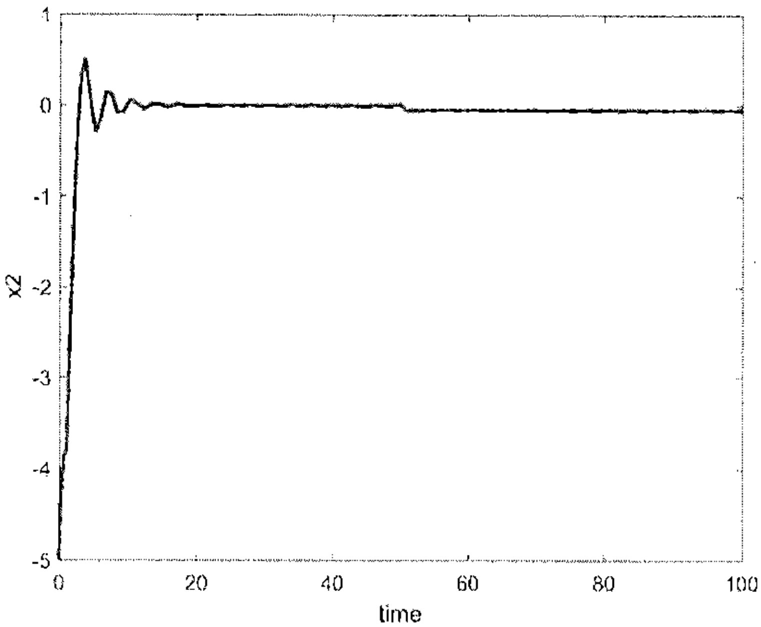


Figure3.2. The response of x_2

CONCLUSION AND DEVELOPMENT DIRECTION OF THE PROJECT

1. Conclusion

This Project mentioned the problem of optimal control design for a manipulator in combination with RISE and exact linearization. With the ADP technique, the solution of HJB equation was found by iteration algorithm to obtain the controller satisfying not only the convergence of weight but also the position tracking. Offline simulations were implemented to validate the performance and effectiveness of the optimal control for manipulators.

We consider previously for nominal systems by eliminating the disturbance, then using classical nonlinear control technique. The neural networks have been designed to approximate the actor and critic part of iterative algorithm. It is possible to develop the learning algorithm with simultaneous tuning. Finally, UUB stability problem of the closed-loop system are guaranteed under this solution.

2. Development Direction of the Project

The research on ADP scheme, as a newly emerging approximate optimal algorithm, is just a beginning. The following part is a brief introduction about the research focus and shortcomings of the existing ADP, by which we hope to show the development tendency about ADP scheme for readers.

1) The proposal of new-type ADP algorithms. At present, ADP scheme is still in the developing stage. Each existing algorithm has its shortcoming. So the new algorithm should be proposed aiming at these shortcomings.

2) The research of ADP scheme in finite time. In the practical application, one usually needs the system to achieve a certain performance index in finite time. So the issue of exploring the optimal control problem in finite time is still a difficult problem.

3) Output feedback of ADP scheme. Until now, most results of ADP scheme focus on the state feedback aspect, whereas the results based on output feedback are limited and still in the infancy period.

4) Improvement of online adaptive algorithm. An iter ation algorithm requires a longer offline calculation time. Once the system changes, we need offline calculation again. Through designing update rate of weights, the online operation using adaptive algorithm is an inevitable trend for ADP scheme. ADP scheme to be applied to approximate optimal.

5) control of large scale time delay uncertain systems (even varying time delay uncertain systems). The control problem of systems with time delay belongs to infinite dimensional system control. The ADP-based optimal control of infinite dimensional systems needs to be further researched.

REFERENCES

[1] Bellman R E. *Dynamic Programming*. Princeton: Princeton University Press, 1957

[2] Dreyfus S E, Law A M. *The Art and Theory of Dynamic Programming*. New York: Academic Press, 1977

[3] White D A, Sofge D A. *Handbook of Intelligent Control: Neural, Fuzzy, and Adaptive Approaches*. New York: Van Nostrand Reinhold, 1992

[4] Werbos P J. Advanced forecasting methods for global crisis warning and models of intelligence. *General Systems Yearbook*, 1977, **22**: 25−38

[5] Werbos P J. *A Menu of Designs for Reinforcement Learning over Time*. Cambridge, MA: MIT Press, 1990. 67−95

[6] Widrow B, Gupta N, Maitra S. Punish/reward: learning with a critic in adaptive threshold systems. *IEEE Transactions on Systems, Man, and Cybernetics*, 1973, **3**(5): 455−465

[7] Chen Zong-Hai, Wen Feng, Wang Zhi-Ling. Neural network control of nonlinear systems based on adaptive critic. *Control and Decision*, 2007, **22**(7): 765−768, 773 (in Chinese)

[8] Lendaris G G, Paintz C. Training strategies for critic and action neural networks in dual heuristic programming method. In: Proceedings of the 1997 IEEE International Conference on Neural Networks. Houston, USA: IEEE, 1997. 712−717

[9] Werbos P J. Consistency of HDP applied to a simple reinforcement learning problem. *Neural Networks*, 1990, **3**(2): 179−189

[10] Bertsekas D P, Tsitsiklis J N. *Neuro-Dynamic Programming*. Belmont: Athena Scientific, 1996

[11] Bertsekas D P. Dynamic programming and optimal control. *Approximate Dynamic Programming (Fourth edition) II*. Belmont: Athena Scientific, 2012

[12] Murray J J, Cox C J, Lendaris G G, Saeks R. Adaptive dynamic programming. *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and reviews*, 2002, **32**(2): 140−153

[13] Sutton R S, Barto A G. *Reinforcement Learning: An Introduction*. Cambridge, MA: The MIT Press, 1998

[14] Si J, Barto A G, Powell W B, Wunsch D. *Handbook of Learning and Approximate Dynamic Programming*. Hoboken: Wiley-IEEE Press, 2004

[15] Powell W B. *Approximate Dynamic Programming: Solving the Curses of Dimensionality*. Princeton: Wiley, 2007

[16] Balakrishnan S N, Ding J, Lewis F L. Issues on stability of ADP feedback controllers for dynamical systems. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 2008, **38**(4): 913−917

[17] Wang F Y, Zhang H G, Liu D R. Adaptive dynamic programming: an introduction. *IEEE Computational Intelligence Magazine*, 2009, **4**(2): 39−47

[18] Prokhorov D V, Wunsch D C II. Adaptive critic designs. *IEEE Transactions on Neural Networks*, 1997, **8**(5): 997−1007

[19] Padhi R, Unnikrishnan N, Wang X H, Balakrishnan S N. A single network adaptive critic (SNAC) architecture for optimal control synthesis for a class of nonlinear systems. *Neural Networks*, 2006, **19**(10): 1648−1660
[20] Abu-Khalaf M, Lewis F L. Nearly optimal control laws for nonlinear systems with saturating actuators using a neural network HJB approach. *Automatica*, 2005, **41**(5): 779−791
[21] Al-Tamimi A, Lewis F L, Abu-Khalaf M. Discrete-time nonlinear HJB solution using approximate dynamic programming: convergence proof. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 2008, **38**(4): 943−949
[22] Zhang H G,Wei Q L, Luo Y H. A novel infinite-time optimal tracking control scheme for a class of discrete-time nonlinear systems via the greedy HDP iteration algorithm. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 2008, **38**(4): 937−942
[23] Zhang H G, Luo Y H, Liu D R. Neural-network-based nearoptimal control for a class of discrete-time affine nonlinear systems with control constraints. *IEEE Transactions on Neural Networks*, 2009, **20**(9): 1490−1503
[24] Wei Q L, Zhang H G, Liu D R, Zhao Y. An optimal control scheme for a class of discrete-time nonlinear systems with time delays using adaptive dynamic rogramming. *Acta Automatica Sinica*, 2010, **36**(1): 121−129
[25] Song R Z, Zhang H G, Luo Y H, Wei Q L. Optimal control laws for time-delay systems with saturating actuators based on heuristic dynamic programming. *Neurocomputing*, 2010, **73**(16−18): 3020−3027
[26] Zhang H G, Song R Z, Wei Q L, Zhang T Y. Optimal tracking control for a class of nonlinear discrete-time systems with time delays based on heuristic dynamic programming. *IEEE Transaction on Neural Networks*, 2011, **22**(12): 1851−1862
[27] Al-Tamimi A, Abu-Khalaf M, Lewis F L. Adaptive critic designs for discrete-time zero-sum games with application to $H\infty$ control. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 2007, **37**(1): 240−247
[28] Abu-Khalaf M, Lewis F L, Huang J. Policy iterations on the Hamilton-Jacobi-Isaacs equation for $H\infty$ state feedback control with input saturation. *IEEE Transactions on Automatic Control*, 2006, **51**(12): 1989−1995
[29] Abu-Khalaf M, Lewis F L, Huang J. Neurodynamic programming and zero-sum games for constrained control systems. *IEEE Transactions on Neural Networks*, 2008, **19**(7): 1243−1252
[30] Zhang X, Zhang H G, Wang X Y, Luo Y H. A new iteration approach to solve a class of finite-horizon continuous-time nonaffine nonlinear zero-sum game. *International Journal of Innovative Computing, Information and Control*, 2011,**7**(2): 597−608
[31] Zhang H G, Wei Q L, Liu D R. An iterative adaptive dynamic programming method for solving a class of nonlinear zero-sum differential games. *Automatica*, 2011, **47**(1): 207− 214
[32] Wei Q L, Zhang H G, Cui L L. Data-based optimal control for discrete-time zero-sum games of 2-D systems using adaptive critic designs. *Acta Automatica Sinica*, 2009, **35**(6): 682−692

[33] Wang F Y, Jin N, Liu D R, Wei Q L. Adaptive dynamic programming for finite-horizon optimal control of discrete-time nonlinear systems with ε-error bound. *IEEE Transactions on Neural Networks*, 2011, **22**(1): 24−36
[34] Lin Xiao-Feng, Zhang Heng, Song Shao-Jian, Song Chun-Ning. Adaptive dynamic programming with ε-error bound for nonlinear discrete-time systems. *Control and Decision*, 2011, **26**(10): 1586−1590, 1595 (in Chinese)
[35] Vamvoudakis K G, Vrabie D, Lewis F L. Online policy iteration based algorithms to solve the continuous-time infinite horizon optimal control problem. In: Proceedings of the 2009 IEEE Symposium on Adaptive Dynamic Programming and Reinforcement Learning. Nashville, USA: IEEE, 2009. 36−41
[36] Vamvoudakis K G, Lewis F L. Online actor-critic algorithm to solve the continuous-time infinite horizon optimal control problem. *Automatica*, 2010, **46**(5): 878−888
[37] Dierks T, Jagannthan S. Optimal control of affine nonlinear discrete-time systems. In: Proceedings of the 17th Mediterranean Conference on Control and Automation. Thessaloniki, Greece: IEEE, 2009. 1390−1395
[38] Dierks T, Jagannathan S. Optimal tracking control of affine nonlinear discrete-time systems with unknown internal dynamics. In: Proceedings of the 48th IEEE Conference on Decision and Control and Conference on Chinese Control. Shanghai, China: IEEE, 2009. 6750−6755
[39] Dierks T, Thumati B T, Jagannathan S. Optimal control of unknown affine nonlinear discrete-time systems using offline-trained neural networks with proof of convergence. *Neural Networks*, 2009, **22**(5−6): 851−860
[40] Zhang H G, Cui L L, Zhang X, Luo Y H. Data-driven robust approximate optimal tracking control for unknown general nonlinear systems using adaptive dynamic programming method. *IEEE Transactions on Neural Networks*, 2011, **22**(12): 2226−2236
[41] Vamvoudakis K G, Lewis F L. Multi-player non-zerosum games: online adaptive learning solution of coupled Hamilton-Jacobi equations. *Automatica*, 2011, **47**(8): 1556−1569
[42] Dierks T, Jagannathan S. Optimal control of affine nonlinear continuous-time systems. In: Proceedings of the 2010 American Control Conference (ACC). Baltimore, USA: IEEE, 2010. 1568−1573
[43] Liu W X, Venayagamoorthy G K, Wunsch D C II. A heuristic-dynamic-programming-based power system stabilizer for a turbogenerator in a single-machine power system. *IEEE Transactions on Industry Applications*, 2005, **41**(5):1377−1385
[44] Park J W, Harley R G, Venayagamoorthy G K. Adaptivecritic- based optimal neurocontrol for synchronous generators in a power system using MLP/RBF neural networks. *IEEE Transactions on Industry Applications*, 2003, **39**(5):1529−1540
[45] Venayagamoorthy G K, Harley R G, Wunsch D C. Dual heuristic programming excitation neurocontrol for generators in a multimachine power system. *IEEE Transactions on Industry Applications*, 2003, **39**(2): 382−394

[46] Lu C, Si J, Xie X R. Direct heuristic dynamic programming for damping oscillations in a large power system. *IEEE Transactions on Systems, Man, and Cybernetics, Part B:Cybernetics*, 2008, **38**(4): 1008−1013
[47] Sun Jian, Liu Feng, Si J, Guo Wen-Tao, Mei Sheng-Wei. An improved approximate dynamic programming and its application in SVC control. *Electric Machines and Control*, 2011, **15**(5): 95−102 (in Chinese)
[48] Bazzan A L C. A distributed approach for coordination of traffic signal agents. *Autonomous Agents and Multi-Agent Systems*, 2005, **10**(1): 131−164
[49] Zhao Dong-Bin, Liu De-Rong, Yi Jian-Qiang. An overview on the adaptive dynamic programming based urban city traffic signal optimal control. *Acta Automatica Sinica*, 2009,**35**(6): 677−681 (in Chinese)
[50] Ray S, Venayagamoorthy G K, Chaudhuri B, Majumder R. Comparison of adaptive critic-based and classical wide-area controllers for power systems. *IEEE Transactions Systems, Man, and Cybernetics, Part B: Cybernetics*, 2008, **38**(4):1002−1007
[51] Li T, Zhao D B, Yi J Q. Heuristic dynamic programming strategy with eligibility traces. In: Proceedings of the 2008 American Control Conference. Seattle, USA: IEEE, 2008.4535−4540
[52] Bai X R, Zhao D B, Yi J Q, Xu J. Coordinated control of multiple ramp metering based on DHP(λ) controller. In: Proceedings of the 11th IEEE International Conference on Intelligent Transportation Systems. Beijing, China: IEEE,2008. 351−356
[53] Cai C. An approximate dynamic programming strategy for responsive traffic signal control. In: Proceedings of the 2007 IEEE International Symposium on Approximate Dynamic Programming and Reinforcement Learning. Honolulu,USA: IEEE, 2007. 303−310
[54] Li T, Zhao D B, Yi J Q. Adaptive dynamic programming for multi-intersections traffic signal intelligent control. In: Proceedings of the 11th IEEE International Conference on Intelligent Transportation Systems. Beijing, China: IEEE,2008. 286−291
[55] Bertsekas D P, Homer M L, Logan D A, Patek S D, Sandell N R. Missile defense and interceptor allocation by neuro-dynamic programming. *IEEE Transactions on Systems, Man, and Cybernetics, Part A: Systems and Humans*, 2000, **30**(1): 42−51
[56] Ferrari S, Stengel R F. Online adaptive critic flight control. *Journal of Guidance, Control, and Dynamics*, 2004, **27**(5):777−786
[57] Liu D R, Javaherian H, Kovalenko O, Huang T. Adaptive critic learning techniques for engine torque and air-fuel ratio control. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 2008, **38**(4): 988−993
[58] Liu D R, Zhang Y, Zhang H G. A self-learning call admission control scheme for CDMA cellular networks. *IEEE Transactions on Neural Networks*, 2005, **16**(5): 1219−1228
[59] Mohammed A. A. Al-Mekhlafi, Herman Wahid, Azian Abd Aziz, "*Adaptive Neuro-Fuzzy Control Approach for a Single Inverted Pendulum System*", International Journal of Electrical and Computer Engineering (IJECE), (ISSN 2088-8708), Vol. 8, No. 5, October 2018, pp. 3657~3665.

[60] Dwi Prihanto, Irawan Dwi Wahyono, Suwasono and Andrew Nafalski. "*Virtual Laboratory for Line Follower Robot Competition*", International Journal of Electrical and Computer Engineering (IJECE), (ISSN 2088-8708), Vol. 7, No. 4, August 2017, pp. 2253~2260.

[61] Keith Dupree, Parag M. Patre, Zachary D. Wilcox, Warren E. Dixon, "*Asymptotic optimal control of uncertain nonlinear Euler–Lagrange systems*", Automatica 47, pp. 99–107, 2011.

[62] Xin Hu, Xinjiang Wei, Huifeng Zhang, Jian Han, Xiuhua Liu, "*Robust adaptive tracking control for a class of mechanical systems with unknown disturbances under actuator saturation*", Int. J. Robust & Nonlinear Control, Vol. 29, Issue. 6, pp. 1893-1908, 2019.

[63] Yong Guo, Bing Huang, Ai-jun Li, Chang-qing Wang, "*Integral sliding mode control for Euler-Lagrange systems with input saturation*" Int. J. Robust & Nonlinear Control, Vol. 29, Issue. 4, pp. 1088-1100, 2018.

[64] Changjiang Xi, Jiuxiang Dong, "*Adaptive reliable guaranteed performance control of uncertain nonlinear systems by using exponent-dependent barrier Lyapunov function*", Int. J. Robust & Nonlinear Control, Vol. 29, Issue. 4, pp. 1051-1062, 2019.

[65] Wei He, Yuhao Chen, Zhao Yin, "*Adaptive Neural Network Control of an Uncertain Robot With Full-State Constraints*", IEEE Transactions on Cybernetic, Vol. 46, Issue. 3, pp. 620-629, 2016.

[66] Wei He, Yiting Dong, "*Adaptive Fuzzy Neural Network Control for a Constrained Robot Using Impedance Learning*" , IEEE Transactions on Neural Networks and Learning Systems, Vol. 29, Issue. 6, pp. 1174-1186, 2018.

[67] Panigrahi, Pratap Kumar et al., "*Comparison of GSA, SA and PSO Based Intelligent Controllers for Path Planning of Mobile Robot in Unknown Environment*", (2015).

[68] Wei He, Yiting Dong, Yiting Dong, Changyin Sun "*Adaptive Neural Impedance Control of a Robotic Manipulator With Input Saturation*", IEEE Transactions on Systems, Man and Cybernetics: Systems, Vol. 46, Issue. 3, pp. 334-344, 2016.

[69] Ziting Chen, Zhijun Li, Philip Chen "*Adaptive Neural Control of Uncertain MIMO Nonlinear Systems With State and Input Constraints*", IEEE Transactions on Neural Networks and Learning Systems, Vol. 28, Issue. 6, pp. 1318-1330, 2017.

[70] Guanyu Lai, Zhi Liu, Yun Zhang, Chun Lung Philip Chen, Shengli Xie,"*Asymmetric Actuator Backlash Compensation in Quantized Adaptive Control of Uncertain Networked Nonlinear Systems*", IEEE Transactions on Neural Networks and Learning Systems, Vol. 28, Issue. 2, pp. 294 -307, 2017.

[71] Tarek Madani, Boubaker Daachi, and Karim Djouani, "*Modular Controller Design Based Fast Terminal Sliding Mode for Articulated Exoskeleton Systems*", IEEE Transactions on Control Systems Technology, Vol. 25, Issue. 3, pp. 1133 - 1140, 2016.

[72] Quang N.H., et al., "*Min Max Model Predictive Control for Polysolenoid Linear Motor*", International Journal of Power Electronics and Drive System (IJPEDS), (ISSN 2088-8694), Vol. 9, No. 4, December 2018, pp. 1666~1675.

[73] Quang N.H., et al., "*On tracking control problem for polysolenoid motor model predictive approach*", International Journal of Electrical and Computer Engineering (IJECE), (ISSN 2088-8708), Vol. 10, No. 1, February 2020, pp. 849~855.

[74] Quang N.H., et al., " *Multi parametric model predictive control based on laguerre model for permanent magnet linear synchronous motors*", International Journal of Electrical and Computer Engineering (IJECE), (ISSN 2088-8708), Vol. 9, No. 2, April 2019, pp. 1067~1077.

[75] Vrabie, D., Pastravanu, O., Abu-Khalaf, M., & Lewis, F. L., "*Adaptive optimal control for continuous-time linear systems based on policy iteration*", Automatica, 45 (2), pp. 477–484, 2009.

[76] Yu Jiang, Zhong-Ping Jiang, "*Computational adaptive optimal control for continuous-time linear systems with completely unknown dynamics*", Automatica, 48, pp. 2699 – 2704, 2012.

[77] Vrabie, D., & Lewis, F. L., "*Neural network approach to continuous-time direct adaptive optimal control for partially unknown nonlinear systems*", Neural Networks, 22 (3), pp. 237–246, 2009.

[78] Murad Abu-Khalaf, Frank L.Lewis, "*Nearly optimal control laws for nonlinear systems with saturating actuators using a neural network HJB approach*" , Automatica, Vol. 49, Issue. 1, pp. 779-791, 2005.

[79] Vamvoudakis, K. G., Lewis, F. L., "*Online actor-critic algorithm to solve the continuous-time infinite horizon optimal control problem*", Automatica, 46(5), pp. 878–888, 2010.

[80] Kyriakos G. Vamvoudakis1, Draguna Vrabie, Frank L. Lewis, "*Online adaptive algorithm for optimal control with integral reinforcement learning*", Int. J. Robust & Nonlinear Control, Vol. 24, Issue. 17, pp. 2686-2710, 2014.

[81] S. Bhasin, R. Kamalapurkar, M. Johnson, K.G. Vamvoudakis, F.L. Lewis, W.E. Dixon, "*A novel actor–critic–identifier architecture for approximate optimal control of uncertain nonlinear systems*" Automatica, Vol. 49, Issue. 1, pp. 82-92, 2013.

[82] Hamidreza Modares, Frank L. Lewis, Mohammad-Bagher Naghibi-Sistani, "*Adaptive Optimal Control of Unknown Constrained-Input Systems Using Policy Iteration and Neural Networks*", IEEE Transactions on Neural Networks and Learning Systems, Vol. 24, Issue. 10, pp. 1513 - 1525, 2013.

[83] Hamidreza Modares, Frank L. Lewis, Mohammad-Bagher Naghibi-Sistani, "*Integral reinforcement learning and experience replay for adaptive optimal control of partially-unknown constrained-input continuous-time systems*" , Automatica, Vol. 50, Issue. 1, pp. 193 - 202, 2014.

[84] Nam D.P, et al., "*Adaptive Dynamic Programming based Integral Sliding Mode Control Law for Continuous-Time Systems: A Design for Inverted Pendulum Systems*", International Journal of Mechanical Engineering and Robotics Research, Vol. 8, No. 2, pp. 279-283, March 2019.

[85] Vu, Tran Anh and Nam, Dao Phuong and Huong, Pham Thi Viet, "Analysis and control design of transformerless high gain, high efficient buck-boost DC-DC converters," 2016 IEEE International Conference on Sustainable Energy Technologies (ICSET), pp. 72-77, 2016.

[86] Nam, Dao Phuong and Thang, Bui Minh and Thanh, Nguyen Truong, "Adaptive Tracking Control for a Boost DC--DC Converter: A Switched Systems Approach," in 2018 4th International Conference on Green Technology and Sustainable Development (GTSD), pp. 702-705, 2018.

[87] Thanh, Nguyen Truong and Sam, Pham Ngoc and Nam, Dao Phuong, "An Adaptive Backstepping Control for Switched Systems in presence of Control Input Constraint," 2019 International Conference on System Science and Engineering (ICSSE), pp. 196-200, 2019.

[88] Panigrahi, Swetapadma and Thakur, Amarnath, "Modeling and simulation of three phases cascaded H-bridge grid-tied PV inverter," Bulletin of Electrical Engineering and Informatics,vol. 8, pp. 1-9, 2019.

[89] Devarajan, N and Reena, A, "Reduction of switches and DC sources in Cascaded Multilevel Inverter," Bulletin of Electrical Engineering and Informatics, vol. 4, pp. 186-195, 2015.

[90] Venkatesan, M and Rajeshwari, R and Deverajan, N and Kaliyamoorthy, M, "Comparative study of three phase grid connected photovoltaic inverter using pi and fuzzy logic controller with switching losses calculation," International Journal of Power Electronics and Drive Systems, vol. 7, pp. 543-550, 2016.

[91] Zhang, Lixian and Xiang, Weiming, "Mode-identifying time estimation and switching-delay tolerant control for switched systems: An elementary time unit approach," Automatica, vol. 64, pp. 174-181, 2016.

[92] Yuan, Shuai and Zhang, Lixian and De Schutter, Bart and Baldi, Simone, "A novel Lyapunov function for a non-weighted L2 gain of asynchronously switched linear systems," Automatica, vol. 87, pp. 310-317, 2018.

[93] Xiang, Weiming and Lam, James and Li, Panshuo, "On stability and $H\infty$ control of switched systems with random switching signals," Automatica, vol. 95, pp. 419-425, 2018.

[94] Lin, Jinxing and Zhao, Xudong and Xiao, Min and Shen, Jingjin, "Stabilization of discrete-time switched singular systems with state, output and switching delays," Journal of the Franklin Institute, vol. 356, pp. 2060-2089, 2019.

[95] Briat, Corentin, "Convex conditions for robust stabilization of uncertain switched systems with guaranteed minimum and mode-dependent dwell-time," Systems \& Control Letters, vol. 78, pp. 63-72, 2015.

[96] Lian, Jie and Li, Can, "Event-triggered control for a class of switched uncertain nonlinear systems," Systems \& Control Letters, vol. 135, pp. 1-5, 2020.

[97] Anyaka, Boniface O and Manirakiza, J Felix and Chike, Kenneth C and Okoro, Prince A, "Optimal unit commitment of a power plant using particle swarm optimization approach," International Journal of Electrical and Computer Engineering (IJECE), vol. 10, pp. 1135-1141, 2020.
[98] Devi, Palakaluri Srividya and Santhi, R Vijaya, "Introducing LQR-fuzzy for a dynamic multi area LFC-DR model," International Journal of Electrical \& Computer Engineering, vol. 9, pp. 861-874, 2019.
[99] Omar, Othman AM and Badra, Niveen M and Attia, Mahmoud A, "Enhancement of on-grid pv system under irradiance and temperature variations using new optimized adaptive controller," International Journal of Electrical and Computer Engineering, vol. 8, pp. 2650-2660, 2018.
[100] Sharma, Purva and Saini, Deepak and Saxena, Akash, "Fault detection and classification in transmission line using wavelet transform and ANN," Bulletin of Electrical Engineering and Informatics, vol. 5, pp. 284-295, 2016.
[101] Ilamathi, P and Selladurai, V and Balamurugan, K, "Predictive modelling and optimization of nitrogen oxides emission in coal power plant using Artificial Neural Network and Simulated Annealing," IAES International Journal of Artificial Intelligence, vol. 1, pp. 11-18, 2012.
[102] Vamvoudakis, Kyriakos G and Vrabie, Draguna and Lewis, Frank L, "Online adaptive algorithm for optimal control with integral reinforcement learning," International Journal of Robust and Nonlinear Control, vol. 24, pp. 2686-2710, 2014.
[103] Bai, Weiwei and Zhou, Qi and Li, Tieshan and Li, Hongyi, "Adaptive reinforcement learning neural network control for uncertain nonlinear system with input saturation," IEEE transactions on cybernetics, 2020.
[104] Chen, Ci and Modares, Hamidreza and Xie, Kan and Lewis, Frank L and Wan, Yan and Xie, Shengli, "Reinforcement learning-based adaptive optimal exponential tracking control of linear systems with unknown dynamics," in IEEE Transactions on Automatic Control, vol. 64, pp. 4423-4438, 2019.
[105] Vamvoudakis, Kyriakos G and Ferraz, Henrique, "Model-free event-triggered control algorithm for continuous-time linear systems with optimal performance," in Automatica, vol. 87, pp. 412-420, 2018.
[106] Gao, Weinan and Jiang, Yu and Jiang, Zhong-Ping and Chai, Tianyou, "Output-feedback adaptive optimal control of interconnected systems based on robust adaptive dynamic programming," Automatica, vol. 72, pp. 37-45, 2016.
[107] Zhang, Tianping and Xu, Haoxiang, "Adaptive optimal dynamic surface control of strict-feedback nonlinear systems with output constraints," International Journal of Robust and Nonlinear Control, 2020.
[108] Wang, Ding and Mu, Chaoxu, "Adaptive-critic-based robust trajectory tracking of uncertain dynamics and its application to a spring--mass--damper system," { IEEE Transactions on Industrial Electronics, vol. 65, pp. 654-663, 2017.
[109] Wen, Guoxing and Ge, Shuzhi Sam and Chen, CL Philip and Tu, Fangwen and Wang, Shengnan, "Adaptive tracking control of surface vessel using optimized backstepping technique," IEEE transactions on cybernetics, vol. 49, pp. 3420-3431, 2018.

THUYẾT MINH ĐỀ TÀI
KHOA HỌC VÀ CÔNG NGHỆ CẤP TRƯỜNG NĂM 2019

1. TÊN ĐỀ TÀI: Điều khiển hệ Robotic có đánh giá đến miền hấp dẫn.	**2. MÃ SỐ: T2019-B11**
3. LĨNH VỰC NGHIÊN CỨU Khoa học Tự nhiên ☐ Khoa học KT và CN ☒ Khoa học Y, dược ☐ Khoa học Nông nghiệp ☐ Khoa học Xã hội ☐ Khoa học Nhân văn ☐	**4. LOẠI HÌNH NGHIÊN CỨU** Cơ bản ☒ Ứng dụng ☒ Triển khai ☐

5. THỜI GIAN THỰC HIỆN: 12 tháng

Từ tháng 7 năm 2019 đến tháng 7 năm 2020

6. CHỦ NHIỆM ĐỀ TÀI

Họ và tên: Trần Thị Hải Yến
Học vị: Thạc sỹ
Chức danh khoa học:
Năm sinh: 1986
Địa chỉ cơ quan:
Điện thoại di động: 0978610626
Điện thoại cơ quan:
Fax:
E-mail: haiyentran86@gmail.com

7. NHỮNG THÀNH VIÊN THAM GIA NGHIÊN CỨU ĐỀ TÀI

TT	Họ và tên	Đơn vị công tác và lĩnh vực chuyên môn	Nội dung nghiên cứu cụ thể được giao	Chữ ký
1	ThS. Nguyễn Vĩnh Thụy	Khoa Điện	Cấu trúc điều khiển mạch vòng trong cho hệ Tracking Trailer; **Mô phỏng hệ trong điều kiện lý tưởng** : Cấu trúc điều khiển nhiều mạch vòng trong cho hệ Tracking Trailer với điều kiện lý tưởng	[signature]
2	ThS. Đỗ Thị Phương Thảo	Khoa Điện	Cấu trúc điều khiển mạch vòng trong cho hệ xe tự hành; **Mô phỏng hệ trong điều kiện lý tưởng** : Cấu trúc điều khiển nhiều mạch vòng trong cho hệ xe tự hành với điều kiện lý tưởng	[signature]
3	ThS. Nguyễn Thị Chinh	Khoa Điện	Cấu trúc điều khiển mạch vòng trong cho hệ cầu treo; **Mô phỏng hệ trong điều kiện lý tưởng** : Cấu trúc điều khiển hai mạch vòng trong cho hệ cầu treo với điều kiện lý tưởng	[signature]

4	ThS. Lâm Hùng Sơn	Khoa Điện	Cấu trúc điều khiển nhiều mạch vòng trong cho hệ Tracking Trailer; **Mô phỏng hệ trong điều kiện lý tưởng** : Cấu trúc điều khiển nhiều mạch vòng trong cho hệ cầu treo với điều kiện lý tưởng	
5	ThS. Nguyễn Hồng Quang	Khoa Điện	Cấu trúc điều khiển nhiều mạch vòng trong cho hệ xe tự hành; **Mô phỏng hệ trong điều kiện có nhiễu ngoại lực và nhiễu đo:** Cấu trúc điều khiển mạch vòng trong cho hệ Tracking Trailer với điều kiện có ngoại lực và nhiễu đo	
6	ThS. Dương Hòa An	Khoa Điện	Cấu trúc điều khiển hai mạch vòng trong cho hệ cầu treo; **Mô phỏng hệ trong điều kiện có nhiễu ngoại lực và nhiễu đo:** Cấu trúc điều khiển mạch vòng trong cho hệ xe tự hành với điều kiện có ngoại lực và nhiễu đo	
7	ThS. Vũ Xuân Tùng	Khoa Điện	Cấu trúc điều khiển nhiều mạch vòng trong cho hệ cầu treo; **Mô phỏng hệ trong điều kiện có nhiễu ngoại lực và nhiễu đo:** Cấu trúc điều khiển mạch vòng trong cho hệ cầu treo với điều kiện có ngoại lực và nhiễu đo	
8	ThS. Trần Thị Thanh Thảo	Khoa Điện	**Mô phỏng hệ trong điều kiện lý tưởng** : Cấu trúc điều khiển mạch vòng trong cho hệ Tracking Trailer với điều kiện lý tưởng; **Mô phỏng hệ trong điều kiện có nhiễu ngoại lực và nhiễu đo:** Cấu trúc điều khiển nhiều mạch vòng trong cho hệ Tracking Trailer với điều kiện có ngoại lực và nhiễu đo	
9	ThS. Dương Quỳnh Nga	Khoa Điện	**Mô phỏng hệ trong điều kiện lý tưởng** : Cấu trúc điều khiển mạch vòng trong cho hệ xe tự hành với điều kiện lý tưởng; **Mô phỏng hệ trong điều kiện có nhiễu ngoại lực và nhiễu đo:** Cấu trúc điều khiển nhiều mạch vòng trong cho hệ xe tự hành với điều kiện có ngoại lực và nhiễu đo	

10	ThS. Nguyễn Văn Huỳnh	Khoa Điện	**Mô phỏng hệ trong điều kiện lý tưởng** : Cấu trúc điều khiển mạch vòng trong cho hệ cầu treo với điều kiện lý tưởng; **Mô phỏng hệ trong điều kiện có nhiễu ngoại lực và nhiễu đo:** Cấu trúc điều khiển hai mạch vòng trong cho hệ cầu treo với điều kiện có ngoại lực và nhiễu đo	

8. ĐƠN VỊ PHỐI HỢP CHÍNH

Tên đơn vị trong và ngoài nước	Nội dung phối hợp nghiên cứu	Họ và tên người đại diện đơn vị

9. TỔNG QUAN TÌNH HÌNH NGHIÊN CỨU THUỘC LĨNH VỰC CỦA ĐỀ TÀI Ở TRONG VÀ NGOÀI NƯỚC

9.1. Tổng quan tình hình nghiên cứu thuộc lĩnh vực của đề tài:

Với sự phát triển của khoa học và công nghệ ngày nay, các phương pháp phân tích và tổng hợp hệ thống trên cơ sở lý thuyết điều khiển phi tuyến ngày càng đưa con người đến gần hơn trong các ứng dụng thực tế. Cùng với sự phát triển và phổ biến của vi xử lý, vi điều khiển... việc thực hiện các thuật toán trong điều khiển phi tuyến ngày càng được thực hiện dễ dàng và hiệu quả hơn so với trước. Bằng việc kế thừa các kết quả của điều khiển tuyến tính cũng như đưa ra khái niệm hàm điều khiển Lyapunov (CLF) đã giúp cho việc giải quyết bài toán thiết kế các bộ điều khiển ổn định được chặt chẽ hơn, rõ ràng hơn, nhất là khi đối tượng có mô hình bất định (*uncertainties*) hoặc bị ảnh hưởng của nhiễu.

Có nhiều tư tưởng điều khiển cho lớp các hệ thống mang những đặc điểm trên mà trong đó việc thiết kế bộ điều khiển trên tư tưởng Tuyến tính hóa chính xác kết hợp với các khâu chỉnh định thích nghi và bền vững là một trong những phương pháp mang lại hiệu quả cao, điều đó đã chứng minh không những trong các nghiên cứu lý thuyết mà còn trên những kết quả thực tế đạt được. Vấn đề điều khiển cho hệ robotic là một chủ đề rất thú vị, mang lại ứng dụng lớn trong thực tiễn công nghiệp cũng như các ứng dụng dân sự. Các kết quả nghiên cứu về vấn đề này của các nhà khoa học ở nhiều lĩnh vực khác nhau như Cơ khí, Điều khiển, Điện, Lập trình...

9.2. Danh mục các công trình đã công bố thuộc lĩnh vực của đề tài của chủ nhiệm và những thành viên tham gia nghiên cứu (*họ và tên tác giả; bài báo; ấn phẩm; các yếu tố về xuất bản*)

a) Của chủ nhiệm đề tài:

1. **Apply digital signal processor tms320 to linear motor control**, *hội nghị khoa học, sáng tạo trẻ Đại học Thái nguyên lần thứ 3- 2016*
2. **Modelling and Fuzzy Control of Brushless DC Motor**, *International Journal of Electrical Electronics & Computer Science Engineering Volume 5, Issue 3 (June, 2018) | E-ISSN : 2348-2273 | P-ISSN : 2454-1222*

b) Của các thành viên tham gia nghiên cứu

1. **Multi parametric model predictive control based on laguerre model for permanent magnet linear synchronous motors** *International Journal of Electrical and Computer Engineering (IJECE) (IJECE, ISSN: 2088-8708), Vol 9, No. 2, April 2019* 2019-04 | journal-article; **Source:** Hong Quang Nguyen
2. **Vibration Suppression Control of a Flexible Gantry Crane System with Varying Rope Length** *Journal of Control Science and Engineering* 2019-02-11 | journal-article DOI: 10.1155/2019/9640814 **Source:** Crossref
3. **Modelling Polysolenoid Permanent Stimulation Linear Motors for Real Time Simulation Problem***International Journal of Electrical Electronics & Computer Science Engineering. E-ISSN : 2348-2273 | P-ISSN : 2454-1222* 2019-02 | journal-article **Source:** Hong Quang Nguyen
4. **An Adaptive Backstepping Trajectory Tracking Control of a Tractor Trailer Wheeled Mobile Robot***International Journal of Control, Automation and Systems* 2019-01 | journal-article DOI: 10.1007/s12555-017-0711-0 *Part of*ISSN: 1598-6446 **Source:** Crossref Metadata Search
5. **Tube based robust model predictive control for disturbed nonlinear systems via solving linear matrix inequalities** *ACM International Conference Proceeding Series* 2019 | conference-paper **Source:** Hong Quang Nguyen

6. **A New Approach of a Tube Based Output Feedback Model Predictive Control: Control Design for 2D Overhead Crane** *Advances in Engineering Research and Application* 2018-11 | other DOI: 10.1007/978-3-030-04792-4_9 *Part of* ISBN: 9783030047917 *Part of* ISSN: 2367-3370 **Source:** Crossref Metadata Search
7. **Adaptive Control to Load Disturbance for Brushless DC Motor Operates at Low peed** *Advances in Engineering Research and Application* 2018-11 | other DOI: 10.1007/978-3-030-04792-4_19 *Part of* ISBN: 9783030047917 *Part of* ISSN: 2367-3370 **Source:** Crossref Metadata Search
8. **A Laguerre Model-Based Model Predictive Control Law for Permanent Magnet Linear Synchronous Motor***Information Systems Design and Intelligent Applications* 2018 | other DOI: 10.1007/978-981-10-7512-4_31 *Part of* ISBN: 9789811075117 *Part of* ISSN: 2194-5357 **Source:** Crossref Metadata Search
9. **Min Max Model Predictive Control for Polysolenoid Linear Motor***International Journal of Power Electronics and Drive Systems (IJPEDS)* 2018 | journal-article DOI: 10.11591/ijpeds.v9.i4.pp1666-1675 *Part of* ISSN: 2088-8694 **Source:** Crossref Metadata Search
10. **Study on Controlling Brushless DC Motor in Current Control Loop Using DC-Link** *American Journal of Engineering Research (AJER). e-ISSN: 2320-0847 p-ISSN :*
Current *2320-0936* 2018 | journal-article **Source:** Hong Quang Nguyen
11. **Studying the Effect of AC Source's Frequencies to Micro Electromechanical System** *International Journal of Electrical Electronics & Computer Science Engineering. E-ISSN : 2348-2273 | P-ISSN : 2454-1222* 2018 | journal-article **Source:** Hong Quang Nguyen
12. **Multi Parametric Programming based Model Predictive Control for tracking Control of Polysolenoid Linear Motor** *Special issue on Measurement, Control and Automation.ISSN: 1859-0551* 2017-08 | journal-article **Source:** Hong Quang Nguyen
13. **Multi parametric programming and exact linearization based model predictive control of a permanent magnet linear synchronous motor** *2017 International Conference on System Science and Engineering (ICSSE)* 2017-07 | conference-paper DOI: 10.1109/icsse.2017.8030975 *Part of* ISBN: 9781538634226 **Source:** Crossref Metadata Search
14. **Design an Exact Linearization Controller for Permanent Stimulation Synchronous Linear Motor Polysolenoid** *International Journal of Electrical and Electronics Engineering* 2017-01 | journal-article DOI: 10.14445/23488379/ijeee-v4i1p102 *Part of* ISSN: 2348-8379 **Source:** Crossref Metadata Search
15. **Modeling of the Polysolenoid Linear Motor and its Control Problems** *The 4th Vietnam International Conference and Exhibition on Control and Automation (VCCA 2017). Ho Chi Minh 2017. ISBN: 978-604-73-5569-3* 2017 | conference-paper **Source:** Hong Quang Nguyen
16. **Flatness Based Control Structure for Polysolenoid Permanent Stimulation Linear Motors** *International Journal of Electrical and Electronics Engineering* 2016 | journal-article DOI: 10.14445/23488379/ijeee-v3i12p110 *Part of* ISSN: 2348-8379 **Source:** Crossref Metadata Search
17. **Cascade Motion/Force Control Strategy of nonholonomic Wheeled Mobile Robotic Systems** *ACM International Conference Proceeding Series* conference-paper. **Source:** Hong Quang Nguyen

10. TÍNH CẤP THIẾT CỦA ĐỀ TÀI

Hệ thống chuyển động robot là một hệ thống có tính phi tuyến mạnh và ràng buộc cao, các tham số động lực học như mô men quán tính, khối lượng tải thường biến đổi và không được xác định chính xác. Việc điều khiển Robot bám chính xác quỹ đạo đặt với các điều kiện nhiễu tác động bên ngoài là không biết trước cũng như kháng được nhiễu nội của hệ thống sinh ra là điều luôn luôn được quan tâm của các nhà nghiên cứu. Một vấn đề quan trọng trong điều khiển quỹ đạo robot cần được xét đến là điểm cuối của quỹ đạo mong muốn. Khi quỹ đạo tiến về điểm gốc đặt trước theo lý thuyết ổn định ISS là ổn định tiệm cận, điều này khá đẹp đẽ về mặt toán học nhưng để thực hiện được điều đó quỹ đạo cần vượt qua miền hấp dẫn chính là miền lân cận quanh điểm ổn định tiệm cận, khi tiến vào vùng này sẽ có sự nhiễu loạn tác động mạnh mẽ lên hệ thống. Việc giải quyết bài toán điều khiển hệ Robotic có đánh giá đến miền hấp dẫn là nội dung cần được giải quyết khi đề cập đến nâng cao chất lượng điều khiển.

11. MỤC TIÊU ĐỀ TÀI

Để vượt qua được miền hấp dẫn này các nhóm thuật toán điều khiển được đề xuất phải thỏa mãn được nguyên lý tách đồng thời dự báo được miền hấp dẫn của hệ Robotic. Một trong những hướng đi tiềm năng được đề xuất là bộ điều khiển hai thành phần trong đó một thành phần có tác dụng thực hiện quỹ đạo chuyển động trung tâm và thành phần còn lại đảm bảo đối phó được với sự nhiễu loạn tại miền hấp dẫn. Mục đích cần phải đạt được là phải xử lý được sự sai khác về trạng thái của hệ thống thực và hệ thống danh nghĩa. Các tác giả tập trung cụ thể hóa những nghiên cứu của nhóm nghiên cứu thông qua các bài báo quốc tế được phát triển từ đề tài.

12. ĐỐI TƯỢNG, PHẠM VI NGHIÊN CỨU

12.1. Đối tượng nghiên cứu: Điều khiển hệ Robotic có đánh giá đến miền hấp dẫn.

12.2. Phạm vi nghiên cứu : Xuất phát từ cách tiếp cận kinh điển, đầu tiên, mô hình toán học của hệ thống sẽ được thiết lập kèm theo những yêu cầu điều khiển và các ràng buộc cụ thể. Dựa trên cơ sở mô hình toán học và mục tiêu điều khiển, bộ điều khiển phù hợp với đối tượng sẽ được xây dựng và tính ổn định của hệ kín được chứng minh một các chi tiết. Tiếp theo, quá trình mô phỏng hệ kín được tiến hành đi kèm theo những đánh giá định tính và định lượng về động học hệ thống.

13. CÁCH TIẾP CẬN, PHƯƠNG PHÁP NGHIÊN CỨU

13.1. Cách tiếp cận: nghiên cứu lý thuyết và thực nghiệm.

13.2. Phương pháp nghiên cứu: Nghiên cứu trong tài liệu (từ các sách, bài báo, tạp chí khoa học...). Thừa kế, tham khảo các kết quả nghiên cứu gần và có liên quan. Tham gia các diễn đàn, hội thảo khoa học trong và ngoài nước, trao đổi các ý tưởng và kết quả nghiên cứu với các nhà khoa học trong và ngoài nước. Thử nghiệm trên mô hình hoá bằng máy tính và trên thiết bị thực tế.

14. NỘI DUNG NGHIÊN CỨU VÀ TIẾN ĐỘ THỰC HIỆN

14.1. Nội dung nghiên cứu:

- Mô hình hóa đối tượng điều khiển, đề xuất các thuật toán điều khiển cho đối tượng điều khiển. Mô phỏng kiểm chứng kết quả. Xây dựng báo cáo tổng kết, viết báo cáo khoa học, viết báo.
- Báo cáo các bài báo bằng tiếng Anh (là sản phẩm của đề tài sau khi được chấp nhận đăng) tại hội thảo đơn vị

14.2. Tiến độ thực hiện

STT	Các nội dung, công việc thực hiện	Sản phẩm	Thời gian (bắt đầu-kết thúc)	Người thực hiện
1	Xây dựng thuyết minh đề tài	Báo cáo	7/2019	Trần Thị Hải Yến

2	Mô hình hóa một số phần tử trong hệ robotic	Báo cáo	8/2019	Trần Thị Hải Yến
3	Cấu trúc điều khiển mạch vòng trong cho hệ Tracking Trailer	Báo cáo	8/2019-10/2019	Nguyễn Vĩnh Thụy
4	Cấu trúc điều khiển mạch vòng trong cho hệ xe tự hành	Báo cáo	8/2019-10/2019	Đỗ Thị Phương Thảo
5	Cấu trúc điều khiển mạch vòng trong cho hệ cầu treo	Báo cáo	8/2019-10/2019	Nguyễn Thị Chinh
6	Cấu trúc điều khiển nhiều mạch vòng trong cho hệ Tracking Trailer	Báo cáo	8/2019-10/2019	Lâm Hùng Sơn
7	Cấu trúc điều khiển nhiều mạch vòng trong cho hệ xe tự hành	Báo cáo	8/2019-10/2019	Nguyễn Hồng Quang
8	Cấu trúc điều khiển hai mạch vòng trong cho hệ cầu treo	Báo cáo	8/2019-10/2019	Dương Hòa An
9	Cấu trúc điều khiển nhiều mạch vòng trong cho hệ cầu treo	Báo cáo	8/2019-10/2019	Vũ Xuân Tùng
10	**Mô phỏng hệ trong điều kiện lý tưởng** : Cấu trúc điều khiển mạch vòng trong cho hệ Tracking Trailer với điều kiện lý tưởng	Báo cáo	8/2019-10/2019	Trần Thị Thanh Thảo
11	**Mô phỏng hệ trong điều kiện lý tưởng** : Cấu trúc điều khiển mạch vòng trong cho hệ xe tự hành với điều kiện lý tưởng	Báo cáo	8/2019-10/2019	Dương Quỳnh Nga
12	**Mô phỏng hệ trong điều kiện lý tưởng** : Cấu trúc điều khiển mạch vòng trong cho hệ cầu treo với điều kiện lý tưởng	Báo cáo	8/2019-10/2019	Nguyễn Văn Huỳnh
13	**Mô phỏng hệ trong điều kiện lý tưởng** : Cấu trúc điều khiển nhiều mạch vòng trong cho hệ Tracking Trailer với điều kiện lý tưởng	Báo cáo	08/2019-10/2019	Nguyễn Vĩnh Thụy
14	**Mô phỏng hệ trong điều kiện lý tưởng** : Cấu trúc điều khiển nhiều mạch vòng trong cho hệ xe tự hành với điều kiện lý tưởng	Báo cáo	08/2019-10/2019	Đỗ Thị Phương Thảo
15	**Mô phỏng hệ trong điều kiện lý tưởng** : Cấu trúc điều khiển hai mạch vòng trong cho hệ cầu treo với điều kiện lý tưởng	Báo cáo	08/2019-10/2019	Nguyễn Thị Chinh

16	**Mô phỏng hệ trong điều kiện lý tưởng** : Cấu trúc điều khiển nhiều mạch vòng trong cho hệ cầu treo với điều kiện lý tưởng	Báo cáo	08/2019-10/2019	Lâm Hùng Sơn
17	**Mô phỏng hệ trong điều kiện có nhiễu ngoại lực và nhiễu đo:** Cấu trúc điều khiển mạch vòng trong cho hệ Tracking Trailer với điều kiện có ngoại lực và nhiễu đo	Báo cáo	08/2019-10/2019	Nguyễn Hồng Quang
18	**Mô phỏng hệ trong điều kiện có nhiễu ngoại lực và nhiễu đo:** Cấu trúc điều khiển mạch vòng trong cho hệ xe tự hành với điều kiện có ngoại lực và nhiễu đo	Báo cáo	08/2019-10/2019	Dương Hòa An
19	**Mô phỏng hệ trong điều kiện có nhiễu ngoại lực và nhiễu đo:** Cấu trúc điều khiển mạch vòng trong cho hệ cầu treo với điều kiện có ngoại lực và nhiễu đo	Báo cáo	08/2019-10/2019	Vũ Xuân Tùng
20	**Mô phỏng hệ trong điều kiện có nhiễu ngoại lực và nhiễu đo:** Cấu trúc điều khiển nhiều mạch vòng trong cho hệ Tracking Trailer với điều kiện có ngoại lực và nhiễu đo	Báo cáo	08/2019-10/2019	Trần Thị Thanh Thảo
21	**Mô phỏng hệ trong điều kiện có nhiễu ngoại lực và nhiễu đo:** Cấu trúc điều khiển nhiều mạch vòng trong cho hệ xe tự hành với điều kiện có ngoại lực và nhiễu đo	Báo cáo	08/2019-10/2019	Dương Quỳnh Nga
22	**Mô phỏng hệ trong điều kiện có nhiễu ngoại lực và nhiễu đo:** Cấu trúc điều khiển hai mạch vòng trong cho hệ cầu treo với điều kiện có ngoại lực và nhiễu đo	Báo cáo	08/2019-10/2019	Nguyễn Văn Huỳnh
23	Viết bài báo quốc tế ISI/SCOPUS	Bài Báo	7/2019-02/2019	Trần Thị Hải Yến
24	Báo cáo bài báo	Báo cáo	3/2019-5/2020	Trần Thị Hải Yến
25	Viết báo cáo các kết quả nghiên cứu	Báo cáo	6/2020-7/2020	Trần Thị Hải Yến

15. SẢN PHẨM

STT	Tên sản phẩm	Số lượng	Yêu cầu chất lượng sản phẩm *(mô tả chi tiết chất lượng sản phẩm đạt được như nội dung, hình thức, các chỉ tiêu, thông số kỹ thuật,...)*
I	Sản phẩm khoa học (Các công trình khoa học sẽ được công bố: sách, bài báo khoa học, ..)		
	Bài báo quốc tế	02	ISI/Scopus, loại Q2
II	Sản phẩm đào tạo (cử nhân, thạc sĩ, tiến sĩ,...): 0		
III	Sản phẩm ứng dụng: 0		

16. PHƯƠNG THỨC CHUYỂN GIAO KẾT QUẢ NGHIÊN CỨU VÀ ĐỊA CHỈ ỨNG DỤNG

16.1. Phương thức chuyển giao:

16.2. Địa chỉ ứng dụng

17. TÁC ĐỘNG VÀ LỢI ÍCH MANG LẠI CỦA KẾT QUẢ NGHIÊN CỨU

17.1. Đối với lĩnh vực giáo dục và đào tạo

17.2. Đối với lĩnh vực khoa học và công nghệ có liên quan

17.3. Đối với phát triển kinh tế-xã hội

17.4. Đối với tổ chức chủ trì và các cơ sở ứng dụng kết quả nghiên cứu

17. KINH PHÍ THỰC HIỆN ĐỀ TÀI

Tổng kinh phí: 120.000.000Đ

Bằng chữ: Một trăm hai mươi triệu đồng chẵn

(Dự toán chi tiết các mục chi đính kèm có xác nhận của các đơn vị liên quan.)

Ngày 10 tháng 7 năm 2019

CHỦ NHIỆM ĐỀ TÀI | **XÁC NHẬN CỦA ĐƠN VỊ**

Trần Thị Hải Yến

KT. HIỆU TRƯỞNG
PHÓ HIỆU TRƯỞNG

BỘ GIÁO DỤC VÀ ĐÀO TẠO ★ TRƯỜNG ĐẠI HỌC KỸ THUẬT CÔNG NGHIỆP ★ ĐẠI HỌC THÁI NGUYÊN

PGS.TS. Vũ Ngọc Pi

TRƯỞNG PHÒNG KHCN&HTQT

PGS.TS. Phạm Thành Long

DỰ TOÁN KINH PHÍ ĐỀ TÀI KH&CN CẤP TRƯỜNG NĂM 2019

Tên đề tài: Điều khiển hệ Robotic có đánh giá đến miền hấp dẫn; Mã số: T2019-B11

Chủ nhiệm đề tài: ThS. Trần Thị Hải Yến

Thành viên chính: Nguyễn Vĩnh Thụy, Đỗ Thị Phương Thảo, Nguyễn Thị Chinh, Lâm Hùng Sơn, Nguyễn Hồng Quang, Dương Hòa An, Vũ Xuân Tùng, Trần Thị Thanh Thảo, Dương Quỳnh Nga, Nguyễn Văn Huỳnh

ĐVT: VNĐ

STT	Nội dung	Dự toán			
I	**Mục chi tiền công lao động tham gia trực tiếp (1)**	**Người thực hiện**	**Số ngày công**	**Hệ số tiền công theo ngày (2)**	**Thành tiền**
1.1	Xây dựng thuyết minh đề tài	Trần Thị Hải Yến	4	0,45	2.502.000
1.2	Mô hình hóa một số phần tử trong hệ robotic	Trần Thị Hải Yến	6	0,45	3.753.000
1.3	Cấu trúc điều khiển mạch vòng trong cho hệ Tracking Trailer	Nguyễn Vĩnh Thụy	12	0,3	5.004.000
1.4	Cấu trúc điều khiển mạch vòng trong cho hệ xe tự hành	Đỗ Thị Phương Thảo	12	0,3	5.004.000
1.5	Cấu trúc điều khiển mạch vòng trong cho hệ cầu treo	Nguyễn Thị Chinh	12	0,3	5.004.000
1.6	Cấu trúc điều khiển nhiều mạch vòng trong cho hệ Tracking Trailer	Lâm Hùng Sơn	12	0,3	5.004.000
1.7	Cấu trúc điều khiển nhiều mạch vòng trong cho hệ xe tự hành	Nguyễn Hồng Quang	8	0,3	3.336.000
1.8	Cấu trúc điều khiển hai mạch vòng trong cho hệ cầu treo	Dương Hòa An	12	0,3	5.004.000
1.9	Cấu trúc điều khiển nhiều mạch vòng trong cho hệ cầu treo	Vũ Xuân Tùng	12	0,3	5.004.000
1.10	Mô phỏng hệ trong điều kiện lý tưởng : Cấu trúc điều khiển mạch vòng trong cho hệ Tracking Trailer với điều kiện lý tưởng	Trần Thị Thanh Thảo	12	0,3	5.004.000
1.11	Mô phỏng hệ trong điều kiện lý tưởng : Cấu trúc điều khiển mạch vòng trong cho hệ xe tự hành với điều kiện lý tưởng	Dương Quỳnh Nga	12	0,3	5.004.000
1.12	Mô phỏng hệ trong điều kiện lý tưởng : Cấu trúc điều khiển mạch vòng trong cho hệ cầu treo với điều kiện lý tưởng	Nguyễn Văn Huỳnh	12	0,3	5.004.000

1.13	Mô phỏng hệ trong điều kiện lý tưởng: Cấu trúc điều khiển nhiều mạch vòng trong cho hệ Tracking Trailer với điều kiện lý tưởng	Nguyễn Vĩnh Thụy	12	0,3	5.004.000
1.14	Mô phỏng hệ trong điều kiện lý tưởng : Cấu trúc điều khiển nhiều mạch vòng trong cho hệ xe tự hành với điều kiện lý tưởng	Đỗ Thị Phương Thảo	12	0,3	5.004.000
1.15	Mô phỏng hệ trong điều kiện lý tưởng : Cấu trúc điều khiển hai mạch vòng trong cho hệ cầu treo với điều kiện lý tưởng	Nguyễn Thị Chinh	12	0,3	5.004.000
1.16	Mô phỏng hệ trong điều kiện lý tưởng : Cấu trúc điều khiển nhiều mạch vòng trong cho hệ cầu treo với điều kiện lý tưởng	Lâm Hùng Sơn	12	0,3	5.004.000
1.17	Mô phỏng hệ trong điều kiện có nhiễu ngoại lực và nhiễu đo: Cấu trúc điều khiển mạch vòng trong cho hệ Tracking Trailer với điều kiện có ngoại lực và nhiễu đo	Nguyễn Hồng Quang	8	0,3	3.336.000
1.18	Mô phỏng hệ trong điều kiện có nhiễu ngoại lực và nhiễu đo: Cấu trúc điều khiển mạch vòng trong cho hệ xe tự hành với điều kiện có ngoại lực và nhiễu đo	Dương Hòa An	12	0,3	5.004.000
1.19	Mô phỏng hệ trong điều kiện có nhiễu ngoại lực và nhiễu đo: Cấu trúc điều khiển mạch vòng trong cho hệ cầu treo với điều kiện có ngoại lực và nhiễu đo	Vũ Xuân Tùng	12	0,3	5.004.000
1.20	Mô phỏng hệ trong điều kiện có nhiễu ngoại lực và nhiễu đo: Cấu trúc điều khiển nhiều mạch vòng trong cho hệ Tracking Trailer với điều kiện có ngoại lực và nhiễu đo	Trần Thị Thanh Thảo	12	0,3	5.004.000
1.21	Mô phỏng hệ trong điều kiện có nhiễu ngoại lực và nhiễu đo: Cấu trúc điều khiển nhiều mạch vòng trong cho hệ xe tự hành với điều kiện có ngoại lực và nhiễu đo	Dương Quỳnh Nga	12	0,3	5.004.000
1.22	Mô phỏng hệ trong điều kiện có nhiễu ngoại lực và nhiễu đo: Cấu trúc điều khiển hai mạch vòng trong cho hệ cầu treo với điều kiện có ngoại lực và nhiễu đo	Nguyễn Văn Huỳnh	9	0,3	3.753.000

1.32	Viết bài báo quốc tế ISI/SCOPUS	Trần Thị Hải Yến	6	0,45	3.753.000
		Nguyễn Hồng Quang	4	0,3	1.668.000
1.33	Báo cáo bài báo	Trần Thị Hải Yến	5	0,45	3.127.500
1.36	Xây dựng báo cáo khoa học tổng kết đề tài	Trần Thị Hải Yến	4	0,45	2.502.000
	Tổng 1		258		**112.798.500**
2	***Chi mua nguyên vật liệu, dịch vụ chuyên môn kỹ thuật***	***Đơn vị tính***	***Đơn giá***	***Số lượng***	***Thành tiền***
					0
	Tổng 2				0
3	***Chi phí khác***				**0**
3.1	Hội thảo nội bộ của đề tài (3)				**700.000**
3.2	Đánh giá, nghiệm thu (4)				*150.000*
	Chủ tịch		*150.000*	*1*	*400.000*
	Ủy viên hội đồng		*100.000*	*4*	*150.000*
	Ủy viên phản biện		*150.000*	*1*	**300.000**
3.3	Thư ký hành chính (5)				**201.500**
3.4	Văn phòng phẩm, in ấn (6)				
3.5	Quản lý chung của cơ quan chủ trì (5% tổng kinh phí đề tài)				**6.000.000**
	Tổng 3				**7.201.500**
	Tổng 1+2+3				**120.000.000**

Cơ quan chủ trì
KT. HIỆU TRƯỞNG
PHÓ HIỆU TRƯỞNG

BỘ GIÁO DỤC VÀ ĐÀO TẠO – TRƯỜNG ĐẠI HỌC KỸ THUẬT CÔNG NGHIỆP – ĐẠI HỌC THÁI NGUYÊN

PGS.TS. Vũ Ngọc Pi

TRƯỞNG PHÒNG KHCN&HTQT

PGS.TS. Phạm Thành Long

CHỦ NHIỆM ĐỀ TÀI

ThS. Trần Hải Yến

TRƯỞNG PHÒNG KH-TC

TS. Đoàn Quang Thiệu

Home Journal Rankings Country Rankings Viz Tools Help About Us

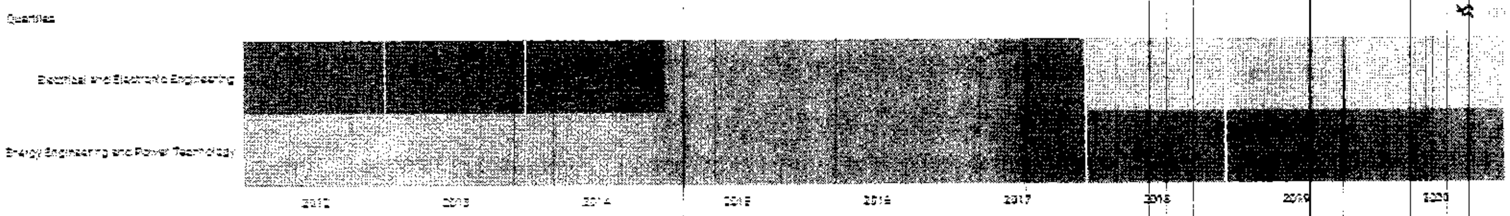
International Journal of Power Electronics and Drive Systems

COUNTRY	SUBJECT AREA AND CATEGORY	PUBLISHER	H-INDEX
Indonesia Universities and research institutions in Indonesia	Energy — Energy Engineering and Power Technology Engineering — Electrical and Electronic Engineering	Institute of Advanced Engineering and Science (IAES)	21

PUBLICATION TYPE	ISSN	COVERAGE	INFORMATION
Journals	20888694	2011-2020	Homepage How to publish in this journal auzani@utem.edu.my

SCOPE

International Journal of Power Electronics and Drive Systems (IJPEDS) is the official publication of the Institute of Advanced Engineering and Science (IAES). The journal is open to submission from scholars and experts in the wide areas of power electronics and electrical drive systems from the global world. The scope of the journal includes all issues in the field of Power Electronics and drive systems. Included are techniques for advanced power semiconductor devices, control in power electronics, low and high power converters (inverters, converters, controlled and uncontrolled rectifiers), Control algorithms and techniques applied to power electronics, electromagnetic and thermal performance of electronic power converters and inverters, power quality and utility applications, renewable energy, electric machines, modelling, simulation, analysis, design and implementations of the application of power circuit components (power semiconductors, inductors, high frequency transformers, capacitors), EMI/EMC considerations, power devices and components, sensors, integration and packaging, applications in motor drives, wind energy systems, solar, battery chargers, UPS and hybrid systems and other applications.



FIND SIMILAR JOURNALS

IEEE Transactions on Transportation Electrification

45%

International Journal of Power Electronics

44%

EPE Journal (European Power Electronics and Drives

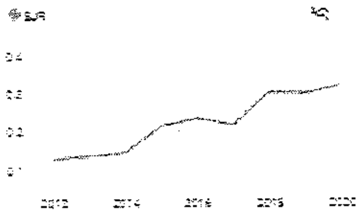
41%

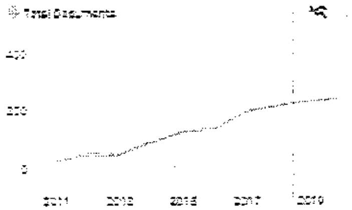
International Journal of Power and Energy Systems

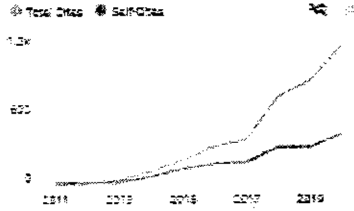
41%

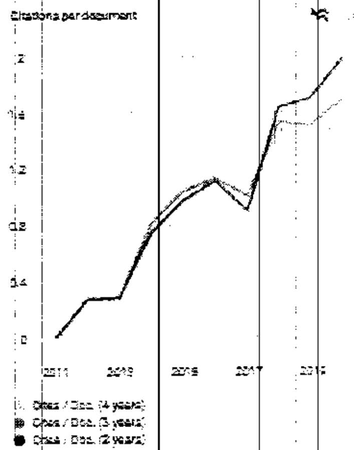
Journal of Green Engineering

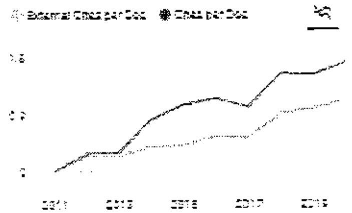
41%

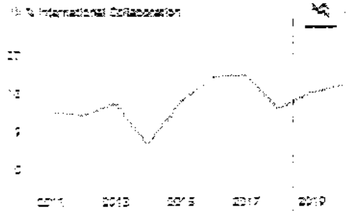


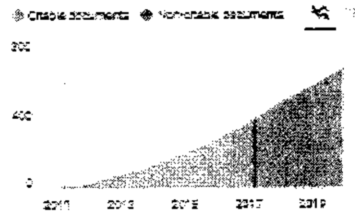


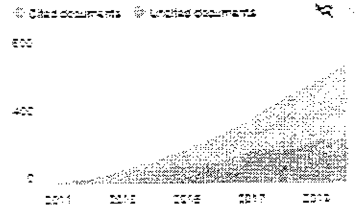


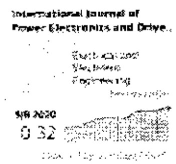












Show this widget in your own website

Just copy the code below and paste within your html code:

G SCImago Graphica

Explore, visually communicate and make sense of data with our new free tool.

Get it

International Journal of Power Electronics and Drive System (IJPEDS)
Vol. 11, No. 3, September 2020, pp. 1123~1131
ISSN: 2088-8694, DOI: 10.11591/ijpeds.v11.i3.pp1123-1131

Adaptive dynamic programing based optimal control for a robot manipulator

Dao Phuong Nam[1], Nguyen Hong Quang[2], Tran Phuong Nam[3], Tran Thi Hai Yen[4]
[1,3] Hanoi University of Science and Technology, Vietnam
[2,4] Thai Nguyen University of Technology, Vietnam

Article Info

Article history:

Received Sep 2, 2019
Revised Nov 9, 2019
Accepted Feb 4, 2020

Keywords:

Adaptive dynamic programming (ADP)
Input constraint
Neural network
Robot manipulator
Robust integral of the sign of the error (RISE)

ABSTRACT

In this paper, the optimal control problem of a nonlinear robot manipulator in absence of holonomic constraint force based on the point of view of adaptive dynamic programming (ADP) is presented. To begin with, the manipulator was intervened by exact linearization. Then the framework of ADP and Robust Integral of the Sign of the Error (RISE) was developed. The ADP algorithm employs Neural Network technique to tune simultaneously the actor-critic network to approximate the control policy and the cost function, respectively. The convergence of weight as well as position tracking control problem was considered by theoretical analysis. Finally, the numerical example is considered to illustrate the effectiveness of proposed control design.

Corresponding Author:

Nguyen Hong Quang,
Thai Nguyen University of Technology, Vietnam,
666, Street 3/2, Tich Luong Ward, Thai Nguyen City - Thai Nguyen Province, Vietnam.
Email: quang.nguyenhong@tnut.edu.vn

1. INTRODUCTION

In recent years, the control methodology for robotic systems has been widely developed not only in practical applications [1, 2], but also in theoretical analysis [3-6]. The main challenges of the control design have been considered, such as robust adaptive control problem, motion/force control, input saturation and full state constraints [7, 8] and the path planning problem [9]. Several control techniques have been employed for manipulators to tackle the issue of input saturation by adding more terms into the designed control input considering the absence of input Constraint [4, 5, 10-13]. In [4], authors proposed a new reference of control system due to the input saturation. The additional term world be computed based on the derivative of previous Lyapunov candidate function along the state trajectory under the control input saturation [4].

Furthermore, authors in [5] give a new approach to address the input constraints as well as combining with handling the disturbances. The proposed sliding surface was employed the Sat function of joint variables. In order to realize the disadvantage of state constraints in manipulator, the authors in [7, 8] proposed the framework of Barrier Lyapunov function and Moore-Penrose inverse, Fuzzy-Neural Network technique. The equivalent sliding mode control algorithm was designed then the boundedness of control input was estimated. The advantage of this approach is that input boundedness absolutely adjusted by selecting several parameters.

The work in [10-13] presents a technique to implement the input constraint using a modified Lyapunov Candidate function. Because of the actuator saturation, the Lyapunov function would be added more the quadratic term from the difference between the control input from controller and the real signal applied to object. The control design was obtained after considering the Lyapunov function derivative along the system trajectory. However, these aforementioned traditional nonlinear techniques have several drawbacks, such as difficulties in finding equivalent Lyapunov function, dynamic of additional terms [7, 8, 10-13]. Optimization Technique using GA (genetic algorithm), PSO (particle swarm optimization) were adressed to solve the papth planning problem [9]. The MPC (model predictive control) solution, which is the special case of optimal control design, has been investigated for linear motor not only online min-max technique in [14, 15] but also offline algorithm in [16]. In order to consider for robot manipulators. Optimal control algorithm obtains the control design that can tackle the input, state constraint based on considering the optimization problem in presence of constraint. An asymptotic optimal control design was presented in [3] by solving directly the Riccati equation in linear systems. However, it is difficult to find the explicit solution of Riccati equation as well as partial differential HJB (Hamilton-Jacobi-Bellman) equation in general case. The approximate/adaptive dynamic programming (ADP) has been paid much attention for optimal control problem in recent years because it is necessary to solve not only Riccati equation for linear systems but also HJB equation for nonlinear systems. Thanks to Kronecker product technique, authors in [17] proposed the online solution for linear systems without the knowledge of system matrix based on the least-squares solution from acquisition of a sufficient number of data points. In [18], Zong-Ping Jiang et al. extend the above online solution to obtain the completely unknown dynamics by means that does not depend on either matrix A or matrix B of linear systems. The fact that Riccati equation was considered in more detail in the computation problem as well as data acquisition. Moreover, the exploration noise on the time interval was mentioned in proposed algorithm [18]. Instead of the approach of employing Kronecker product for the case of linear systems, the neural network approximation was mentioned for cost function to implement online adaptive algorithm on the Actor/Critic structure for continuous time nonlinear systems [19].

However, the proposed algorithm required the knowledge of input-to-state dynamics to update the control policy as well as persistent condition was not considered [19]. The weight parameters in neural network were tuned to minimize the objective in the least-squares sense [19]. The theoretical analysis about convergence of cost function and control input in adaptive/approximate dynamic programming (ADP) was the extension of the work in [20]. Thanks to the theoretical analysis about the neural network approximation, authors in [21] presented the novel online ADP algorithm which enables to tune simultaneously both actor and critic neural networks. The weights training problem of critic neural network (NN) was implemented by modified Levenberg-Marquardt algorithm to minimize the square residual error. Moreover, the tuning of weights in actor and critic NN depend on each other to obtain the weights convergence. It is worth noting that the persistence of excitation (PE) condition need to be satisfied and Lyapunov stability theory was employed to analysis the convergence problem [21]. Extension of the work in [21], based on the analysis of approximate Bellman error, the proposed algorithm in [22] enables to online simultaneously implement without the knowledge of drift term. In [23], the identifier along with adaptation law can be described using a Neural Network to approximate the dynamic uncertainties of nonlinear model. An extension using special cost function has been proposed in [24, 25] to enable handling of input constraint. The framework of ADP technique and classical sliding mode control was presented to design the optimal control for an inverted pendulum [26]. However, the effectiveness of ADP has been still not considered for a robot manipulator in aforementioned researches. This work proposed the control algorithm combining exact linearization, Robust Integral of the Sign of the Error (RISE [3]) and ADP technique for manipulators in absence of holonomic constraint. This ADP technique was implemented using simultaneous tuning method to satisfy the weight convergence and stability.

2. DYNAMIC MODEL OF A ROBOT MANIPULATOR AND CONTROL OBJECTIVE

Consider the following robot manipulator without constraint:

$$M(q)\ddot{q}+C(q,\dot{q})\dot{q}+G(q)+F(\dot{q})+\tau_d(t)=\tau \quad (1)$$

Several appropriate assumptions [3] will be considered to develop the control design in next chapters.

Assumption 1. The inertia matrix M(q) is symmetric, positive definite, and guarantees the inequality $\forall\xi(t)\in\mathbb{R}^n$ as follows:

$$m_1\|\xi\|^2\le\xi^T M(q)\xi\le\bar{m}(q)\|\xi\|^2, \quad (2)$$

where $m_1 \in \mathbb{R}$, $\bar{m}(q) \in \mathbb{R}$, $\|\cdot\|$ is a known positive constant, a known positive function, and the standard Euclidean norm, respectively.

Assumption 2. The relationship between an inertia matrix M(q) and the Coriolis matrix $C(q,\dot{q})$ can be represented as follows:

$$\xi^T(\dot{M}(q)-2C(q,\dot{q}))\xi = 0 \ \ \forall \xi \in R^n. \tag{3}$$

It should be noticed that this manipulator is considered in the absence of holonomic constraint force. The control objective is to find the control algorithm being the framework of exact linearization, RISE and ADP technique enabling the position tracking control in manipulators control system as shown in Figure 1. ADP algorithm will be employed to implement optimal control design as desribed in next chapter.

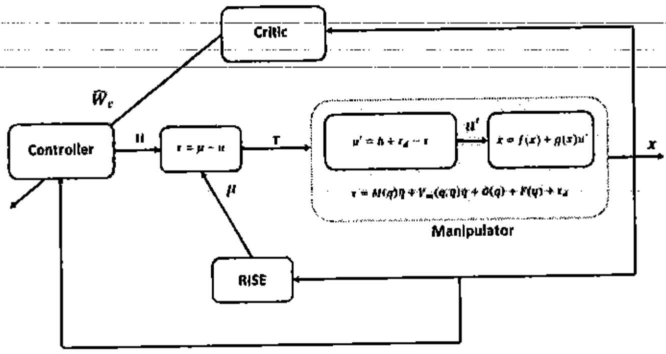


Figure 1. Control structure

3. ADAPTIVE DYNAMIC PROGRAMMING APPROACH FOR A ROBOT MANIPULATOR

3.1. ADP algorithm

In [3], by using the control input (4) for manipulator (1) with nonlinear function (5) obtaining from (6)-(8), we lead to the nonlinear model (9):

$$u = -\tau + h + \tau_d \tag{4}$$

$$h = M(\alpha_1 \dot{e}_1) + C(\alpha_1 e_1) + G(q) + F(\dot{q}) \tag{5}$$

$$e_1 = q_d - q \tag{6}$$

$$e_2 = \dot{e}_1 + \alpha_1 e_1 \tag{7}$$

$$r = \dot{e}_2 + \alpha_2 e_2 \tag{8}$$

$$\dot{x} = f(x) + g(x)u \tag{9}$$

where $x = \begin{bmatrix} e_1 \\ e_2 \end{bmatrix}$, $f(x) = \begin{bmatrix} -\alpha_1 & I_{n\times n} \\ 0_{n\times n} & -M^{-1}C \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \end{bmatrix}$ and $g(x) = \begin{bmatrix} 0_{n\times n} \\ -M^{-1} \end{bmatrix}$

Now, the control object is to design a control law u to guarantee not only stabilization (9) but also minimizing the quadratic cost function with infinite horizon as follows:

$$V(x_0) = \int_0^\infty r(x,u)\,dt \tag{10}$$

$$r(x,u) = Q(x) + u^T R u \tag{11}$$

In which, $Q(x)$ and R is positive definite function of x, symmetric definite positive matrix, respectively.

This work presents a solution for approximate approach called adaptive dynamic programming (ADP) for optimal control design. In [21, 22], consider the following affine system.

$$\dot{x} = f(x) + g(x)u \tag{12}$$

where $x \in \chi \subseteq R^n$, $u \in U \subseteq R^m$. $f(x)$ and $g(x)$ satisfy Lipschitz condition and $f(0) = 0$.

The cost function is defined as (10). The next definition was given in [17, 18] to show that the optimal control solution will be considered in the set of admissible control.

Definition 1: A control policy $\mu(x)$ is defined as admissible policy if $\mu(x)$ stabilize system (12) and the equivalent value function $V^{\mu}(x)$ is finite. $\Psi(\chi)$ is denoted set of admissible control policy.

For any admissible policy $\mu(x)$, the nonlinear Lyapunov Equation (NLE) can be formulated

$$r(x,\mu(x)) + \left(\partial V/\partial x\right)^T \left(f(x) + g(x)\mu(x)\right) = 0 \tag{13}$$

Defining Hamilton function and optimal cost function as follows:

$$H(x,\mu,V_x) = r(x,\mu) + \left(V_x^{\mu}\right)^T \left(f(x) + g(x)\mu\right) \tag{14}$$

$$V^*(x) = \min_{\mu\in\Psi(\chi)} \left(\int_t^{\infty} r(x,\mu) \right)$$

We lead to the following HJB equation:

$$0 = \min_{\mu\in\Psi(\chi)} H(x,\mu,V_x^*) = H(x,\mu^*,V_x^*) \tag{15}$$

It can be noticed that, μ^* is optimal policy corresponding with the optimal cost function and $H(x,\mu,V_x^{\mu}) = 0$ with any admissible policy is NLE.

Now, the optimal control policy can be obtained by taking the derivative of Hamilton problem with respect to policy μ,

$$\mu^* = -\frac{1}{2}\left(R^{-1} g^T V_x^*\right) \tag{16}$$

This work present Policy Iteration (PI) algorithm for a robot manipulator including 2 steps as follows:

Initiate admissible control policy $\mu^0(x)$,

Repeat

Step 1: Policy Evaluation

Solve NLE for $V^i(x)$ corresponding given control policy μ^i,

$$r\left(x,\mu^i(x)\right) + \left(V_x^i\right)^T \left(f(x) + g(x)\mu^i(x)\right) = 0 \tag{17}$$

Step 2: Policy improvement

Update new policy according to,

$$\mu^{i+1} = -\frac{1}{2}\left(R^{-1}g^T V_x^i\right) \quad (18)$$

Until $n = n_{\max}$ or $\left|V^{i+1} - V^i\right| \le \varepsilon_v$.

Where $n_{\max}$ is a number of limited iteration and ε_v is an arbitrary given small positive number.

This algorithm is considered in [21] that prove each policy control μ^i is admissible control. The cost function V^i was reduced at each step until converge to optimal policy and μ^i converge toward optimal policy as well.

However, the nonlinear Lyapunov (17) is hard to solve directly. Therefore, in recent years, finding an indirectly way to solve this equation has been concerned by many researches [20-25]. In the next steps, two neural networks called Actor-Critic (AC) are trained simultaneously to solve approximately the HJB equation.

The cost function and its associated policy can be represented by using a neural network (NN) as follows,

$$\begin{cases} V^* = W^T\phi(x) + \varepsilon_v \\ u^* = -\frac{1}{2}R^{-1}g^T\left(\nabla\phi(x)\right)^T W + \varepsilon_a \end{cases} \quad (19)$$

Where, $\phi(x)$ is corresponding function of NN that usually being selected as polynomial, Gausses, sigmoid function and so on. ∇ is denoted $\partial/\partial x$.

Approximated optimal cost function and optimal policy are presented:

$$\begin{cases} \hat{V} = \hat{W}_c^T\phi(x) \\ \hat{u} = -\frac{1}{2}R^{-1}g^T\left(\nabla\phi(x)\right)^T \hat{W}_a \end{cases} \quad (20)$$

Note that, to approximate HJB solution, we need to find only term $\hat{W}_c$. However, to stabilize closed-loop system, both $\hat{W}_a$, $\hat{W}_c$ are employed, which leads to the flexibility that can help handling the stability of system in learning process.

By replacing the optimal policy and the optimal cost function and by Actor-Critic networks in HJB (17), HJB error can be obtained.

$$Q(x) + \hat{u}^T R\hat{u} + \hat{W}_c^T\nabla\phi\left(f(x) + g(x)\hat{u}\right) = \varepsilon_{hjb} \quad (21)$$

$$Q(x) + \frac{1}{4}\hat{W}_a^T\nabla\phi^T G\nabla\phi\hat{W}_a + \hat{W}_c^T\nabla\phi\left(f(x) - \frac{1}{2}gR^{1}g^T\nabla\phi\hat{W}_a\right) = \varepsilon_{hjb} \quad (22)$$

Where $G = g^T R^{-1} g$.

The tuning law for $\hat{W}_c$ is described as follows,

$$\dot{\hat{W}}_c = -\eta c\Gamma\frac{\omega}{1 + \upsilon\omega^T\Gamma\omega}\varepsilon_{hjb} \quad (23)$$

$$\dot{\Gamma} = -\eta_c\Gamma\frac{\omega\omega^T}{1 + \upsilon\omega\Gamma\omega^T}\Gamma \quad (24)$$

$\Gamma(t_r^+)=\Gamma(0)=\varphi_0 I$. Where t_r^+ is resetting time. To avoid slow convergence on $\hat{W}_c$, the matrix Γ is considered with default matrix $\Gamma(0)$ when minimum eigenvalue of Γ reach a given small positive number. $\omega(x)=\nabla\phi^T\left(f(x)+g(x)u\right)$ and $1+\upsilon\omega^T\Gamma\omega$ is normalization factor.

To make sure the convergence of $\hat{W}_c$ with update law (24), $\omega(x)$ must satisfy the Persistence Excitation (PE) condition [21].

$$\mu_1 I \geq \int_{t_0}^{t_0+T} \psi(\tau)\psi(\tau)^T d\tau \geq \mu_2 I \tag{25}$$

for several positive numbers μ_1, μ_2, T.

Where $\psi(\tau)=\dfrac{\omega(t)}{\sqrt{1+\upsilon\omega^T\Gamma\omega}}$.

On the other hands, (22) is nonlinear equation of $\hat{W}_a$. Therefore, the tuning law for $\hat{W}_a$ is formulated based on GD algorithm to minimize the cost $\left(\varepsilon_{hjb}(t)\right)^2$.

$$\dot{\hat{W}}_a = proj\left\{-\eta_{a1}\frac{1}{\sqrt{1+\omega^T\omega}}\nabla\phi \mathrm{G}\nabla\phi^T\left(\hat{W}_a-\hat{W}_c\right)\varepsilon_{HJB}-\eta_{a2}\left(\hat{W}_a-\hat{W}_c\right)\right\} \tag{26}$$

Where $proj\{\cdot\}$ is a projection operator [22] that ensure the boundedness of updatation law.

Note that, these parameters of both two NN's update law η_c, η_{a1}, η_{a2} must be selected to satisfy some conditions [22] to ensure stability of closed-loop system. One can also find the complete proof of convergence of parameters and stability of system in [22].

3.2. RISE feedback control design

In [3], the control term μ(t) is designed based on the RISE framework as follows:

$$\mu(t)\triangleq(k_s+1)e_2(t)-(k_s+1)e_2(0)+\upsilon(t) \tag{27}$$

Where $\upsilon(t)\in\mathbb{R}^n$ is described as:

$$\dot{\upsilon}=(k_s+1)\alpha_2 e_2+\beta_1 sgn(e_2) \tag{28}$$

$k_s\in\mathbb{R}$ is positive constant control gain, and $\beta_1\in\mathbb{R}$ can be selected being a positive control gain selected according to the following sufficient condition,

$$\beta_1>\zeta_1+\frac{1}{\alpha_2}\zeta_2 \tag{29}$$

Remark 1: It is different from the work in [3], in our work the ADP algorithm is presented to find the intermediate optimal control input in the absence of dynamic uncertainty. Furthermore, ADP technique was considered in [20-26] was still not to apply for a robotic manipulator.

Remark 2: In compare with the work of Dixon [3] that design optimal control solving Riccati equation, this work requires partial knowledge of manipulator's dynamic including matrices M, C. However, using the ADP approach, the optimal control problem is addressed in general case for any given cost function as (10) without constraint.

4. OFFLINE SIMULATION RESULTS

Consider the offline simulation of a two-link manipulator control system using ADP technique and RISE algorithm.

The general dynamic of two-link manipulator is represented by (1) with

$$M=\begin{bmatrix}5+2\cos(q_2) & 1+\cos(q_2)\\ 1+\cos(q_2) & 1\end{bmatrix},\quad C=\begin{bmatrix}-\dot{q}_2\sin(q_2) & -(\dot{q}_1+\dot{q}_2)\sin(q_2)\\ \dot{q}_1\sin(q_2) & 0\end{bmatrix}$$

$$G=9.8\begin{bmatrix}1.2\cos(q_1)+\cos(q_1+q_2)\\ \cos(q_1+q_2)\end{bmatrix},\quad F=-0.1sign(\dot{q}),\quad \tau_d=\begin{bmatrix}0.1\sin(t)\\ 0.1\cos(t)\end{bmatrix}.$$

Value function is (10) with the term: $Q(x)=x^TQ_0x$.

$$Q_0=\begin{bmatrix}Q_{11} & Q_{12}\\ Q_{21} & Q_{22}\end{bmatrix},\quad Q_{11}=\begin{bmatrix}40 & 2\\ 2 & 40\end{bmatrix},\quad Q_{12}=Q_{21}=\begin{bmatrix}-4 & 4\\ 4 & -6\end{bmatrix},\quad Q_{22}=\begin{bmatrix}4 & 0\\ 0 & 4\end{bmatrix},\quad R=\begin{bmatrix}0.25 & 0\\ 0 & 0.25\end{bmatrix},$$

$$\alpha=\begin{bmatrix}15.6 & 10.6\\ 10.6 & 10.4\end{bmatrix}$$

Without loss of generality, the set-point is selected as $q_d=[0\quad 0]^T$, initial state is $q_0=[0.1598\quad 0.2257]^T$.

The optimal value function which is solved directly in [3] is

$$V^*=x^T\begin{bmatrix}-Q_{12} & 0_{n\times n}\\ 0_{n\times n} & M\end{bmatrix}x=2x_1^2-4x_2^2+3x_1x_2+2.5x_3^2+x_3^2\cos(x_2)+x_4^2+x_3x_4+0.5x_3x_4\cos(x_2)$$

The updatation law of $\dot{\hat{W}}_c$ and $\dot{\hat{W}}_a$ are represented in (23) and (26) with,

$$\eta_c=800,\ \upsilon=1,\ \Gamma(0)=100,\ \varepsilon_T=0.001,\ \eta_{a1}=0.01,\ \eta_{a2}=1.$$

NN activation function is selected as,

$$\phi(x)=\begin{bmatrix}x_1^2 & x_2^2 & x_1x_2 & x_3^2 & x_3^2\cos(x_2) & x_4^2 & x_3x_4 & x_3x_4\cos(x_2)\end{bmatrix}^T.$$

The optimal parameter $W=[2\quad -4\quad 3\quad 2.5\quad 1\quad 1\quad 1\quad 0.5]$ that is obtained by solving directly HJB as shown in [3]. Figures (1) and (2) show the convergence of $\hat{W}_c$, $\hat{W}_a$. The value of $\hat{W}_c$ after 110s is $[2\quad -4\quad 3\quad 2.5\quad 1\quad 1\quad 1\quad 0.5]$. To satisfy PE condition as in (25), a probing signal is added in system input. Moreover, system's error evolution is shown in Figure (3) determining the stability of control system and state's evolution as shown in Figure 4.

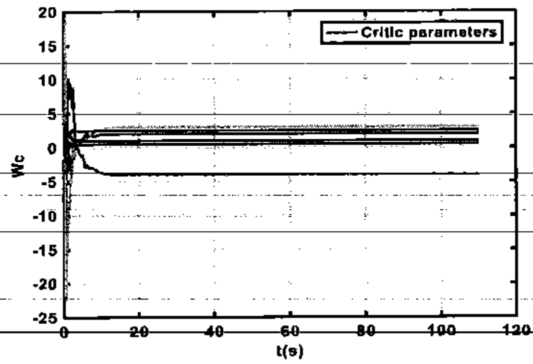


Figure 2. Convergence of critic's parameters

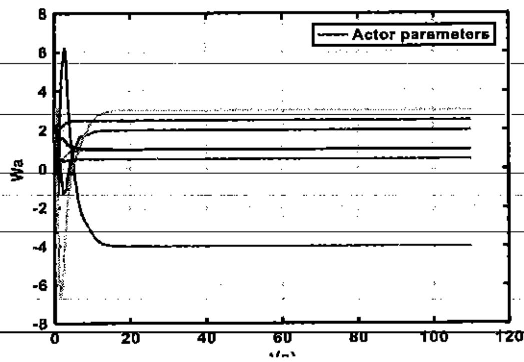


Figure 3. Convergence of actor's parameters

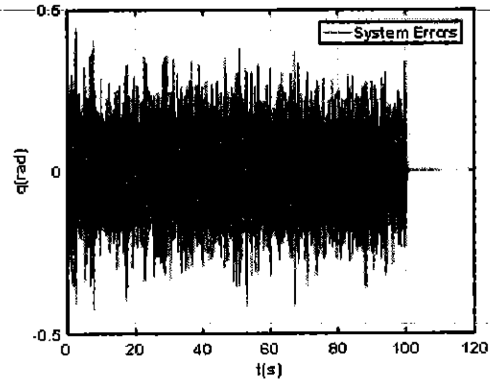


Figure 4. State's evolution

5. CONCLUSION

This paper mentioned the problem of optimal control design for a manipulator in combination with RISE and exact linearization. With the ADP technique, the solution of HJB equation was found by iteration algorithm to obtain the controller satisfying not only the convergence of weight but also the position tracking. Offline simulations were implemented to validate the performance and effectiveness of the optimal control for manipulators.

ACKNOWLEDGEMENTS

This research was supported by Research Foundation funded by Thai Nguyen University of Technology.

REFERENCES

[1] Mohammed A. A. Al-Mekhlafi, Herman Wahid, Azian Abd Aziz, "Adaptive Neuro-Fuzzy Control Approach for a Single Inverted Pendulum System", *International Journal of Electrical and Computer Engineering (IJECE)*, Vol. 8, No. 5, pp. 3657-3665, 2018.

[2] Dwi Prihanto, Irawan Dwi Wahyono, Suwasono and Andrew Nafalski. "Virtual Laboratory for Line Follower Robot Competition", *International Journal of Electrical and Computer Engineering (IJECE)*, Vol. 7, No. 4, pp. 2253-2260, 2017.

[3] Keith Dupree, Parag M. Patre, Zachary D. Wilcox, Warren E. Dixon, "Asymptotic optimal control of uncertain nonlinear Euler–Lagrange systems", *Automatica*, Vol. 47, pp. 99-107, 2011.

[4] Xin Hu, Xinjiang Wei, Huifeng Zhang, Jian Han, Xiuhua Liu, "Robust adaptive tracking control for a class of mechanical systems with unknown disturbances under actuator saturation", *Int. J. Robust & Nonlinear Control*, Vol. 29, Issue. 6, pp. 1893-1908, 2019.

[5] Yong Guo, Bing Huang, Ai-jun Li, Chang-qing Wang, "Integral sliding mode control for Euler-Lagrange systems with input saturation" *Int. J. Robust & Nonlinear Control*, vol. 29, no. 4, pp. 1088-1100, 2018.

[6] Changjiang Xi, Jiuxiang Dong, "Adaptive reliable guaranteed performance control of uncertain nonlinear systems by using exponent-dependent barrier Lyapunov function", *Int. J. Robust & Nonlinear Control*, vol. 29, no. 4, pp. 1051-1062, 2019.

[7] Wei He, Yuhao Chen, Zhao Yin, "Adaptive Neural Network Control of an Uncertain Robot With Full-State Constraints", *IEEE Transactions on Cybernetic*, Vol. 46, No. 3, pp. 620-629, 2016.

[8] Wei He, Yiting Dong, "Adaptive Fuzzy Neural Network Control for a Constrained Robot Using Impedance Learning", *IEEE Transactions on Neural Networks and Learning Systems*, vol. 29, no. 6, pp. 1174-1186, 2018.

[9] Panigrahi, Pratap Kumar et al., "Comparison of GSA, SA and PSO Based Intelligent Controllers for Path Planning of Mobile Robot in Unknown Environment", 2015.

[10] Wei He, Yiting Dong, Yiting Dong, Changyin Sun "Adaptive Neural Impedance Control of a Robotic Manipulator With Input Saturation", *IEEE Transactions on Systems, Man and Cybernetics: Systems*, vol. 46, no. 3, pp. 334-344, 2016.

[11] Ziting Chen, Zhijun Li, Philip Chen "Adaptive Neural Control of Uncertain MIMO Nonlinear Systems With State and Input Constraints", *IEEE Transactions on Neural Networks and Learning Systems*, vol. 28, no. 6, pp. 1318-1330, 2017.

[12] Guanyu Lai, Zhi Liu, Yun Zhang, Chun Lung Philip Chen, Shengli Xie, "Asymmetric Actuator Backlash Compensation in Quantized Adaptive Control of Uncertain Networked Nonlinear Systems", *IEEE Transactions on Neural Networks and Learning Systems*, vol. 28, no. 2, pp. 294-307, 2017.

[13] Tarek Madani, Boubaker Daachi, and Karim Djouani, "Modular Controller Design Based Fast Terminal Sliding Mode for Articulated Exoskeleton Systems", *IEEE Transactions on Control Systems Technology*, vol. 25, no. 3, pp. 1133-1140, 2016.

[14] Quang N.H., et al., "Min Max Model Predictive Control for Polysolenoid Linear Motor", *International Journal of Power Electronics and Drive System (IJPEDS)*, Vol. 9, No. 4, pp. 1666-1675, 2018.

[15] Quang N.H., et al., "On tracking control problem for polysolenoid motor model predictive approach", *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 10, no. 1, pp. 849-855, 2020.

[16] Quang N.H., et al.," Multi parametric model predictive control based on laguerre model for permanent magnet linear synchronous motors", *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 9, no. 2, pp. 1067-1077, 2019.

[17] Vrabie, D., Pastravanu, O., Abu-Khalaf, M., & Lewis, F. L., "Adaptive optimal control for continuous-time linear systems based on policy iteration", *Automatica*, vol. 45, no. 2, pp. 477-484, 2009.

[18] Yu Jiang, Zhong-Ping Jiang, "Computational adaptive optimal control for continuous-time linear systems with completely unknown dynamics", *Automatica*, vol. 48, pp. 2699-2704, 2012.

[19] Vrabie, D., & Lewis, F. L., "Neural network approach to continuous-time direct adaptive optimal control for partially unknown nonlinear systems", *Neural Networks*, vol. 22, no. 3, pp. 237-246, 2009.

[20] Murad Abu-Khalaf, Frank L.Lewis, "Nearly optimal control laws for nonlinear systems with saturating actuators using a neural network HJB approach", *Automatica*, vol. 49, no. 1, pp. 779-791, 2005.

[21] Vamvoudakis, K. G., Lewis, F. L., "Online actor-critic algorithm to solve the continuous-time infinite horizon optimal control problem", *Automatica*, vol. 46, no. 5, pp. 878-888, 2010.

[22] Kyriakos G. Vamvoudakis1, Draguna Vrabie, Frank L. Lewis, "Online adaptive algorithm for optimal control with integral reinforcement learning", *Int. J. Robust & Nonlinear Control*, vol. 24, no. 17, pp. 2686-2710, 2014.

[23] S. Bhasin, R. Kamalapurkar, M. Johnson, K.G. Vamvoudakis, F.L. Lewis, W.E. Dixon, "A novel actor–critic–identifier architecture for approximate optimal control of uncertain nonlinear systems" *Automatica*, vol. 49, no. 1, pp. 82-92, 2013.

[24] Hamidreza Modares, Frank L. Lewis, Mohammad-Bagher Naghibi-Sistani, "Adaptive Optimal Control of Unknown Constrained-Input Systems Using Policy Iteration and Neural Networks", *IEEE Transactions on Neural Networks and Learning Systems*, vol. 24, no. 10, pp. 1513-1525, 2013.

[25] Hamidreza Modares, Frank L. Lewis, Mohammad-Bagher Naghibi-Sistani, "Integral reinforcement learning and experience replay for adaptive optimal control of partially-unknown constrained-input continuous-time systems", *Automatica*, vol. 50, no. 1, pp. 193-202, 2014.

[26] Nam D.P, et al., "Adaptive Dynamic Programming based Integral Sliding Mode Control Law for Continuous-Time Systems: A Design for Inverted Pendulum Systems", *International Journal of Mechanical Engineering and Robotics Research*, Vol. 8, No. 2, pp. 279-283, March 2019.

International Journal of Power Electronics and Drive System (IJPEDS)
Vol. 12, No. 1, March 2021, pp. 551~557
ISSN: 2088-8694, DOI: 10.11591/ijpeds.v12.i1.pp551-557

Adaptive dynamic programming algorithm for uncertain nonlinear switched systems

Dao Phuong Nam[1], Nguyen Hong Quang[2], Nguyen Nhat Tung[3], Tran Thi Hai Yen[4]
[1]School of Electrical Engineering, Hanoi University of Science and Technology, Bách Khoa, Hai Bà Trung, Hà Noi, Vietnam
[2,4]Thai Nguyen University of Technology, So 666 D. 3/2, P, Thành pho Thái Nguyên, Thái Nguyên, Vietnam
[3]Electric Power University, 235 Hoàng Quoc Viet, Co Nhue, Tu Liêm, Hà Noi 129823, Vietnam

Article Info

Article history:

Received Feb 2, 2020
Revised Dec 15, 2020
Accepted Jan 10, 2021

Keywords:

Adaptive dynamic programming
HJB equation
Lyapunov
Neural networksstability
Nonlinear switched systems

ABSTRACT

This paper studies an approximate dynamic programming (ADP) strategy of a group of nonlinear switched systems, where the external disturbances are considered. The neural network (NN) technique is regarded to estimate the unknown part of actor as well as critic to deal with the corresponding nominal system. The training technique is simultaneously carried out based on the solution of minimizing the square error Hamilton function. The closed system's tracking error is analyzed to converge to an attraction region of origin point with the uniformly ultimately bounded (UUB) description. The simulation results are implemented to determine the effectiveness of the ADP based controller.

Corresponding Author:

Nguyen Hong Quang
Thai Nguyen University of Technology
So 666 D. 3/2, P, Thành pho Thái Nguyên, Thái Nguyên, Vietnam
Email: quang.nguyenhong@tnut.edu.vn

1. INTRODUCTION

It is worth noting that many systems in industry can be described by switched system such as DC-DC converter [1]-[3], H-bridge inverter [4], multilevel inverter [5], photovoltaic inverter [6]. Although many different approaches for switched systems have been proposed, e.g., switching-delay tolerant control [7], classical nonlinear control [8]-[12], the optimization approaches with the advantage of mentioning the input/state constraint has not been mentioned much. The approaches of fuzzy and neural network as well as ANN, particle swarm optimization (PSO) technique were investigated in several different systems such as photovoltaic inverter, transmission line. [13]-[17].

Adaptive dynamic programming has been considered in many situations, such as nonlinear continuous time systems [18], actuator saturation [19], linear systems [20]-[22], output constraint [23]. In the case of nonlinear systems, the algorithm should be implemented based on Neural Networks (NNs). However, Kronecker product was employed in linear systems. Furthermore, the data driven technique should to be mentioned to compute the actor/critic precisely. It should be noted that the robotic systems has been controlled by ADP algorithm [24]-[25].

Our work proposed the solution of adaptive dynamic programming in nonlinear perturbed switching systems based on the neural networks. The consideration of the Halminton function enables us obtaining the learning technique of these neural networks. The UUB stability of closed system is analyzed and simulation results illustrate the high effectiveness of given controller.

2. PROBLEM STATEMENTS

Consider the following uncertain nonlinear continuous time switched systems of the form:

$$\frac{d}{dt}\xi(t) = f_i\left(\xi(t)\right) + g_i\left(\xi(t)\right)\left(u + \Delta\left(\xi, t\right)\right) \tag{1}$$

where $\xi(t) \in \Omega_x \in \mathbb{R}^n$ denotes the state variables and $u(t) \in \Omega_u \in \mathbb{R}^m$ describes the control variables. The function $\beta : [0, +\infty) \mapsto \Omega = \{1, 2, ..., l\}$ is a information of switching processing, which is known as a function with many continuous piecewise depending on time, and l is the subsystems number. $f_i(\xi)$ are uncertain smooth vector functions with $f_i(0) = 0$. $g_i(\xi)$ are mentioned as smooth vector functions with the property $G_{\min} \leqslant \|g_i(\xi)\| \leqslant G_{\max}$. The switching index $\beta(t)$ is unknown.

Assumption 1: $\Delta(\xi, t)$ is bounded by a certain function $\varrho(\xi)$ as $\|\Delta(\xi, t)\| \leqslant \varrho(\xi)$

Consider the cost function connected with the uncertain switched system (1):

$$J(\xi, u) = \int_t^\infty r\left(\xi(\tau), u(\tau)\right) d\tau \tag{2}$$

where $r(\xi, u) = \xi^T Q\xi + u^T Ru$ and $Q = Q^T > 0$; $R = R^T > 0$.

The main purpose is to achieve the state feedback control design and give the upper bound term to guarantee the closed systems under this controller is robustly stable. Additionally, the performance index (2) is bounded as $J \leq K(\xi, u) \leq M$.

Definition: The term $K(u)$ is given by the appropriate performance index. As a result, the control input $u^* = \arg\min_{u\in\Omega_u} K(\xi, u)$ is mentioned as the optimal appropriate performance index method.

3. CONTROL DESIGN

The obtained nominal system after eliminating the disturbance in switched system (3) is described by:

$$\frac{d}{dt}\xi = f_i(\xi) + g_i(\xi)u \tag{3}$$

The performance index of system (3) is modified as (4)

$$Q_1(\xi, u) = \int_t^\infty \left[r(\xi, u) + \gamma\left(\rho(\xi)\right)^2\right] d\tau \tag{4}$$

We prove that $Q_1(\xi, u)$ with $\gamma \geqslant \|R\|$ is the one of appropriate performance indexes of dynamical system (1). Define: $V^*(t) = \min_{u\in\Omega_u} Q_1(\xi, u)$, we have (5)

$$V^*(t) = \min_{u\in\Omega_u} \int_t^\infty \left\{r(\xi, u) + \gamma\rho^2(\xi)\right\} d\lambda \tag{5}$$

based on nominal system and cost function (4), it leads to Halminton function as (6)

$$H(\xi, u, V^*) = r(\xi, u) + \gamma\rho^2(\xi) + \left(\frac{\partial V^*}{\partial \xi}\right)^T \left(f_i(\xi) + g_i(\xi)u\right) \tag{6}$$

by using optimality principle, the optimal control input can be obtained as (7).

$$u^*(\xi) = -\frac{1}{2}R^{-1}\left(g_i(\xi)\right)^T \frac{\partial V^*}{\partial \xi} \tag{7}$$

We continue to utilize this control law (7) for nonlinear continuous SW system (1) and obtain that:

Theorem 1: The system (1) under the controller $u^*(\xi) = -\frac{1}{2}R^{-1}\left(g_i(\xi)\right)^T \frac{\partial V^*}{\partial \xi}$ is stable with the associated Lyapunov function candidate:

$$V(t) = \int_t^\infty \{r(\xi,u) + \gamma \varrho^2(\xi)\}\, d\lambda \tag{8}$$

where $\gamma \geqslant \|R\|$.

Proof: Taking the derivative of V under the control input $u(\xi) = -\frac{1}{2}R^{-1}(g_i(\xi))^T \nabla V^*$, we imply that (9):

$$\frac{d}{dt}V = -\xi^T Q\xi - \left(\gamma\varrho^2(\xi) - \Delta(\xi,t)^T R\Delta(\xi,t)\right) - (u + \Delta(\xi,t))^T R(u + \Delta(\xi,t)) \tag{9}$$

It is able to conclude that (10):

$$\dot{V}(t) \leqslant -\xi^T Q\xi \tag{10}$$

Therefore, the system (1) is robustly stable. However, it is impossible to solve directly HJB equation. Hence, the optimal performance index V^* for system (3) can be described based on a NN as (11)

$$V^* = w^T\sigma(\xi) + \varepsilon(\xi) \tag{11}$$

where $\sigma(x) : R^n \to R^N; \sigma(0) = 0$, $w \in R^N$ is the NN constant weight vector. $\sigma(x)$ can be found to guarantee that when $N \to \infty$, we have: $\varepsilon(\xi) \to 0$ and $\nabla\varepsilon(\xi) \to 0$, so for fixed N, we can assume that:

Assumption 2: $\|\varepsilon(\xi)\| \leqslant \varepsilon_{\max}; \|\nabla\varepsilon(\xi)\| \leqslant \nabla\varepsilon_{\max}; \nabla\sigma_{\min} \leqslant \|\nabla\sigma(\xi)\| \leqslant \nabla\sigma_{\max}; \|w\| \leqslant w_{\max}$.

Combining two formulas (10) and (11) we imply (12)

$$H(\xi, u^*, V^*) = \xi^T Q\xi + \lambda\varrho^2(\xi) + (\nabla V^*)^T f_i(\xi) - \frac{1}{4}(\nabla V^*)^T g_i(\xi) R^{-1} g_i(\xi)^T (\nabla V^*) = 0 \tag{12}$$

Formula (19) leads to (13).

$$\nabla V^* = (\nabla\sigma(\xi))^T w + \nabla\varepsilon(\xi) \tag{13}$$

Obtain the description as (14).

$$e_{NN} = -\nabla\varepsilon(\xi)^T (f_i(\xi) + g_i(\xi)u^*) + \frac{1}{4}\nabla\varepsilon(\xi)^T g_i(\xi) R^{-1} g_i(\xi)^T \nabla\varepsilon(\xi) \tag{14}$$

It follows that e_{NN} converges uniformly to zero as $N \to \infty$. For each number N, e_{NN} is bounded on a region as $e_{NN} \leqslant e_{\max}$. Under the structure of ADP-based controller, a critic NN is computed as (15).

$$\hat{V} = \hat{w}^T\sigma(\xi) = \sigma(\xi)^T\hat{w}; \hat{u} = -\frac{1}{2}R^{-1}(g_i(\xi))^T \nabla\hat{V} \tag{15}$$

It is able to achieve that:

$$e_{HJB} = \xi^T Q\xi + \lambda\varrho^2(\xi) + \hat{w}^T\nabla\sigma(\xi) f_i(\xi) - \frac{1}{4}\hat{w}^T\nabla\sigma(\xi) g_i(\xi) R^{-1} g_i(\xi)^T \nabla\sigma(\xi)^T\hat{w} \tag{16}$$

The training law is handled based on a steepest descent method:

$$\frac{d}{dt}\widehat{w} = -\alpha\frac{\partial E}{\partial\widehat{w}} \tag{17}$$

with $E = \frac{1}{2}e_{HJB}^T e_{HJB}$.

Remark 1: The weight $\widehat{w}$ is trained to minimize the network error part $G = \frac{1}{2}e_{HJB}^T e_{HJB}$. This result is obtained from (18).

$$\frac{\partial G}{\partial t} = -\alpha\left(\frac{\partial G}{\partial\widehat{w}}\right)^2 \tag{18}$$

 ISSN: 2088-8694

Theorem 2: Consider the feedback controller in (15) and the critic weight is updated by (18), the weight estimate error $\tilde{w} = w - \hat{w}$ and the closed system's state vector $x(t)$ are uniformly ultimately bounded (UUB).

Proof: Let's choose the Lyapunov function:

$$V(t) = V_1(t) + V_2(t), \text{ where: } V_1(t) = \frac{1}{2\alpha}\tilde{w}(t)^T \tilde{w}(t), V_2(t) = V^* \tag{19}$$

Using the **Assumption 3:** $\|f_i(\xi) + g_i(\xi)u^*\| \leqslant \rho_{\max}$ and the definition: $\rho_i = f_i(\xi) + g_i(\xi)u^*; G_i = g_i(\xi)R^{-1}g_i(\xi)^T; \nabla\sigma = \nabla\sigma(\xi); \nabla\varepsilon = \nabla\varepsilon(\xi)$. Taking the derivative of $V_1(t)$, we imply that:

$$\dot{V}_1(t) = -\tilde{w}^T\left(-e_{NN} + \tilde{w}^T\nabla\sigma\mu_i + \frac{1}{2}\tilde{w}^T\nabla\sigma G_i\nabla\varepsilon + \frac{1}{4}\tilde{w}^T\nabla\sigma G_i\nabla\sigma^T\tilde{w}\right)$$

$$\nabla\sigma(x)\left(\mu_i + \frac{1}{2}G_i\left(\nabla\sigma^T\tilde{w} + \nabla\varepsilon\right)\right) \tag{20}$$

It leads to the estimation: $\dot{V}_1(t) \leqslant -\pi_1$. For the term $V_2(t)$, from (20) we have (21).

$$\dot{V}_2 = (\nabla V^*)^T(f_i + g_i(\hat{u} + \Delta)) = -\left(\xi^T Q\xi + \lambda\rho^2(\xi)\right) - \frac{1}{4}(\nabla V^*)^T$$

$$g_i R^{-1} g_i^T(\nabla V^*) + \frac{1}{2}(\nabla V^*)^T g_i R^{-1} g_i^T\left(\nabla\sigma(\xi)^T\tilde{w} + \nabla\varepsilon(\xi)\right) + (\nabla V^*)^T g_i\Delta \tag{21}$$

Assume that $\rho(\xi) = \varpi\|\xi\|$. From (40) we have (22).

$$\dot{V}_2 \leqslant -(\lambda_{\min}(Q) + \lambda\varpi)\|\xi\|^2 + \theta^2 \tag{22}$$

with $\theta^2 = -\frac{1}{4}(\nabla V^*)^T g_i R^{-1} g_i^T(\nabla V^*) + \frac{1}{2}(\nabla V^*)^T g_i R^{-1} g_i^T\left(\nabla\sigma(x)^T\tilde{w} + \nabla\varepsilon(x)\right) + (\nabla V^*)^T g_i\Delta$.

Based on the two above assumptions, we have (23).

$$\theta^2 \leqslant \frac{1}{4}(w_{\max}\nabla\sigma_{\max} + \nabla\varepsilon_{\max})^2 g_{\max}^2\lambda_{\max}\left(R^{-1}\right) + \frac{1}{2}(\vartheta\nabla\sigma_{\max} + \nabla\varepsilon_{\max})^2 g_{\max}^2\lambda_{\max}\left(R^{-1}\right)$$

$$+ (w_{\max}\nabla\sigma_{\max} + \nabla\varepsilon_{\max})g_{\max}\varpi\|x\| \tag{23}$$

It is obvious that $(\lambda_{\min}(Q) + \lambda\varpi)\|x\|^2 - \theta^2 \geqslant \pi_2$ with $\pi_2 > 0$ and we obtain (24).

$$\dot{V}_2(t) \leqslant -\pi_2 \tag{24}$$

Remark 2: The coefficients $\vartheta_1; \vartheta_2$ can be chosen by renovating the NN of the optimal performance index. Moreover, for arbitrary switching index, after $\frac{V(0)}{\min(\pi_1;\pi_2)}$ the variable $\|\xi\|$ and $\|\tilde{w}\|$ tend to the accurate domains. The ADP controller $\hat{u}$ is proposed in (15), which tends to the neighborhood of u^*.

Proof: The deviation of control input is estimated as (25).

$$\|\hat{u} - u^*\| = \frac{1}{2}\left\|R^{-1}(g_i(\xi))^T\left((\nabla\sigma(\xi))^T\tilde{w} + \nabla\varepsilon(\xi)\right)\right\|$$

$$\leqslant \frac{1}{2}\lambda_{\max}\left(R^{-1}\right).G_{\max}.(\nabla\sigma_{\max}.\upsilon_1 + \nabla\varepsilon_{\max}) = \vartheta_3 \tag{25}$$

Thus the proof is completed.

4. SIMULATION RESULTS

In this section, we consider the simulations to validate the performance of the established control scheme. Let $N = 2$ and the subsystems of the switched system are (26) and (27).

$$\begin{cases} \dot{x}_1 = -x_1^3 - 2x_2 + (u + \Delta_1(x,t)) \\ \dot{x}_2 = x_1 + 0.5\cos(x_1^2)\sin(x_2^3) - (u + \Delta_1(x,t)) \end{cases} \tag{26}$$

$$\begin{cases} \dot{x}_1 = -x_1^5 \sin(x_2) + (u + \Delta_2(x,t)) \\ \dot{x}_2 = \frac{1}{2}x_1 - \cos(x_1)\cos(x_2^3) - (u + \Delta_2(x,t)) \end{cases} \tag{27}$$

The initial state vectors can be chosen as (28).

$$x(0) = \begin{bmatrix} 5 & -5 \end{bmatrix}^T \tag{28}$$

Choosing that the parameter matrices: $R = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}; Q = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}; \alpha = 0.1; \lambda = 5.$

The simulation results shown in Figure 1 and Figure 2 validate the effectiveness of proposed algorithm.

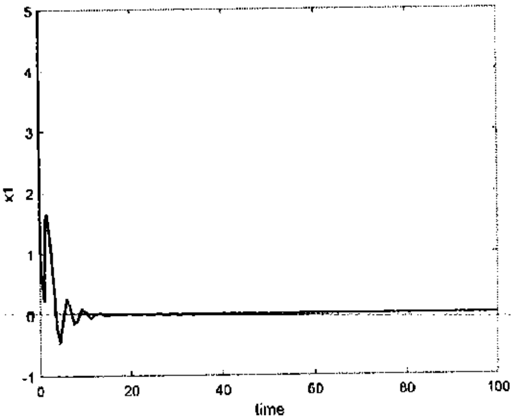


Figure 1. The response of x_2

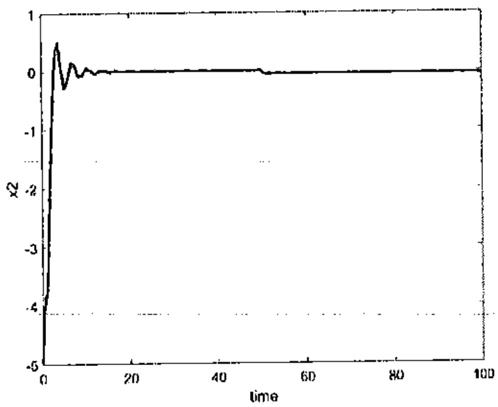


Figure 2. The response of x_2

5. CONCLUSION

This paper has investigated the ADP problem of switched nonlinear systems under the external disturbance. We consider previously for nominal system by eliminating the disturbance, then using classical nonlinear control technique. The neural networks have been designed to estimate the actor and critic NN of iteration. It is possible to develop the learning algorithm with simultaneous tuning. Finally, UUB description of the closed system is guaranteed under this work.

ACKNOWLEDGEMENT

This research was supported by research foundation funded by Thai Nguyen University of Technology.

REFERENCES

[1] Vu, Tran Anh and Nam, Dao Phuong and Huong, Pham Thi Viet, "Analysis and control design of transformerless high gain, high efficient buck-boost DC-DC converters," in *2016 IEEE International Conference on Sustainable Energy Technologies (ICSET)*, Hanoi, 2016, pp. 72-77, doi: 10.1109/ICSET.2016.7811759.

[2] Nam, Dao Phuong and Thang, Bui Minh and Thanh, Nguyen Truong, "Adaptive Tracking Control for a Boost DC–DC Converter: A Switched Systems Approach," in *2018 4th International Conference on Green Technology and Sustainable Development (GTSD)*, Ho Chi Minh City, 2018, pp. 702-705, doi: 10.1109/GTSD.2018.8595580.

[3] Thanh, Nguyen Truong and Sam, Pham Ngoc and Nam, Dao Phuong, "An Adaptive Backstepping Control for Switched Systems in presence of Control Input Constraint," in *2019 International Conference on System Science and Engineering (ICSSE)*, Dong Hoi, Vietnam, 2019, pp. 196-200, doi: 10.1109/ICSSE.2019.8823125.

[4] Panigrahi, Swetapadma and Thakur, Amarnath, "Modeling and simulation of three phases cascaded H-bridge grid-tied PV inverter," *Bulletin of Electrical Engineering and Informatics (BEEI)*, vol. 8, no. 1, pp. 1-9, 2019, doi: 10.11591/eei.v8i1.1225.

[5] Devarajan, N and Reena, A, "Reduction of switches and DC sources in Cascaded Multilevel Inverter," *Bulletin of Electrical Engineering and Informatics (BEEI)*, vol. 4, no. 3, pp. 186-195, 2015, doi: 10.11591/eei.v4i3.320.

[6] Venkatesan, M and Rajeshwari, R and Deverajan, N and Kaliyamoorthy, M, "Comparative study of three phase grid connected photovoltaic inverter using pi and fuzzy logic controller with switching losses calculation," *International Journal of Power Electronics and Drive Systems (IJPEDS)*, vol. 7, no. 2, pp. 543-550, 2016.

[7] Zhang, Lixian and Xiang, Weiming, "Mode-identifying time estimation and switching-delay tolerant control for switched systems: An elementary time unit approach," *Automatica*, vol. 64, pp. 174-181, 2016, doi: 10.1016/j.automatica.2015.11.010.

[8] Yuan, Shuai and Zhang, Lixian and De Schutter, Bart and Baldi, Simone, "A novel Lyapunov function for a non-weighted L2 gain of asynchronously switched linear systems," *Automatica*, vol. 87, pp. 310-317, 2018, doi: 10.1016/j.automatica.2017.10.018.

[9] Xiang, Weiming and Lam, James and Li, Panshuo, "On stability and H control of switched systems with random switching signals," *Automatica*, vol. 95, pp. 419-425, 2018, doi: 10.1016/j.automatica.2018.06.001.

[10] Lin, Jinxing and Zhao, Xudong and Xiao, Min and Shen, Jingjin, "Stabilization of discrete-time switched singular systems with state, output and switching delays," *Journal of the Franklin Institute*, vol. 356, pp. 2060-2089, 2019, doi: 10.1016/j.jfranklin.2018.11.034.

[11] Briat, Corentin, "Convex conditions for robust stabilization of uncertain switched systems with guaranteed minimum and mode-dependent dwell-time," *Systems & Control Letters*, vol. 78, pp. 63-72, 2015, doi: 10.1016/j.sysconle.2015.01.012.

[12] Lian, Jie and Li, Can, "Event-triggered control for a class of switched uncertain nonlinear systems," *Systems & Control Letters*, vol. 135, pp. 1-5, 2020, doi: 10.1016/j.sysconle.2019.104592.

[13] Anyaka, Boniface O and Manirakiza, J Felix and Chike, Kenneth C and Okoro, Prince A, "Optimal unit commitment of a power plant using particle swarm optimization approach," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 10, no.2, pp. 1135-1141, 2020, doi: 10.11591/ijece.v10i2.pp1135-1141.

[14] Devi, Palakaluri Srividya and Santhi, R Vijaya, "Introducing LQR-fuzzy for a dynamic multi area LFC-DR model," *International Journal of Electrical & Computer Engineering*, vol. 9, no. 2, pp. 861-874, 2019, doi: 10.11591/ijece.v9i2.pp861-874.

[15] Omar, Othman AM and Badra, Niveen M and Attia, Mahmoud A, "Enhancement of on-grid pv system under irradiance and temperature variations using new optimized adaptive controller," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 8, no. 5, pp. 2650-2660, 2018, doi: 10.11591/ijece.v8i5.2650-2660.

[16] Sharma, Purva and Saini, Deepak and Saxena, Akash, "Fault detection and classification in transmission line using wavelet transform and ANN," *Bulletin of Electrical Engineering and Informatics (BEEI)*, vol. 5, no. 3, pp. 284-295, 2016.

[17] Ilamathi, P and Selladurai, V and Balamurugan, K, "Predictive modelling and optimization of nitrogen oxides emission in coal power plant using Artificial Neural Network and Simulated Annealing," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 1, no. 1, pp. 11-18, 2012.

[18] Vamvoudakis, Kyriakos G and Vrabie, Draguna and Lewis, Frank L, "Online adaptive algorithm for optimal control with integral reinforcement learning," *International Journal of Robust and Nonlinear Control*, vol. 24, no. 17, pp. 2686-2710, 2013, doi: 10.1002/rnc.3018.

[19] Bai, Weiwei and Zhou, Qi and Li, Tieshan and Li, Hongyi, "Adaptive reinforcement learning neural network control for uncertain nonlinear system with input saturation," *IEEE transactions on cybernetics*, vol. 50, no. 8, pp. 3433-3443, Aug. 2020, doi: 10.1109/TCYB.2019.2921057.

[20] Chen, Ci and Modares, Hamidreza and Xie, Kan and Lewis, Frank L and Wan, Yan and Xie, Shengli, "Reinforcement learning-based adaptive optimal exponential tracking control of linear systems with unknown dynamics," in *IEEE Transactions on Automatic Control*, vol. 64, no. 11, pp. 4423-4438, Nov. 2019, doi: 10.1109/TAC.2019.2905215.

[21] Vamvoudakis, Kyriakos G and Ferraz, Henrique, "Model-free event-triggered control algorithm for continuous-time linear systems with optimal performance," in *Automatica*, vol. 87, pp. 412-420, 2018, doi: 10.1016/j.automatica.2017.03.013.

[22] Gao, Weinan and Jiang, Yu and Jiang, Zhong-Ping and Chai, Tianyou, "Output-feedback adaptive optimal control of interconnected systems based on robust adaptive dynamic programming," *Automatica*, vol. 72, pp. 37-45, 2016, doi: 10.1016/j.automatica.2016.05.008.

[23] Zhang, Tianping and Xu, Haoxiang, "Adaptive optimal dynamic surface control of strict-feedback nonlinear systems with output constraints," *International Journal of Robust and Nonlinear Control*, vol. 30, no. 5, pp. 2059–2078, 2020,

doi: 10.1002/rnc.4864.

[24] Wang, Ding and Mu, Chaoxu, "Adaptive-critic-based robust trajectory tracking of uncertain dynamics and its application to a spring–mass–damper system," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 1, pp. 654-663, Jan. 2018, doi: 10.1109/TIE.2017.2722424.

[25] Wen, Guoxing and Ge, Shuzhi Sam and Chen, CL Philip and Tu, Fangwen and Wang, Shengnan, "Adaptive tracking control of surface vessel using optimized backstepping technique," *IEEE transactions on cybernetics*, vol. 49, no. 9, pp. 3420-3431, Sept. 2019, doi: 10.1109/TCYB.2018.2844177.