

Adaptive dynamic programming based optimal control for a robot manipulator

Dao Phuong Nam¹, Nguyen Hong Quang², Tran Phuong Nam³, Tran Thi Hai Yen⁴

^{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.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



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 would 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 addressed to solve the path 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:

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

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:

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

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 described in next chapter.

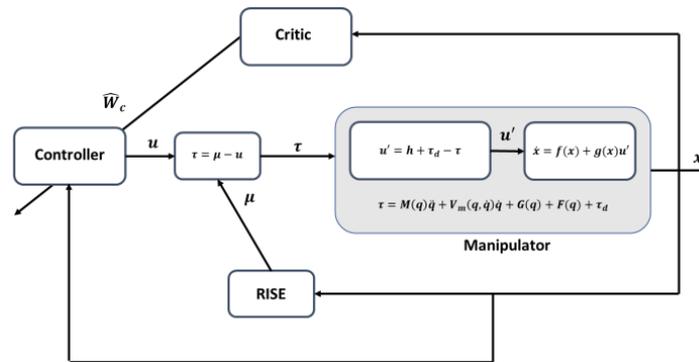


Figure 1. Control structure

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):

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

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

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

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

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

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

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.

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

where $x \in \mathcal{X} \subseteq \mathbb{R}^n$, $u \in U \subseteq \mathbb{R}^m$. $f(x)$ 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. $\Psi(\mathcal{X})$ is denoted set of admissible control policy.

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

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

Defining Hamilton function and optimal cost function as follows:

$$H(x, \mu, V_x) = r(x, \mu) + (V_x^\mu)^T (f(x) + g(x)\mu) \quad (14)$$

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

We lead to the following HJB equation:

$$0 = \min_{\mu \in \Psi(\mathcal{X})} H(x, \mu, V_x^*) = H(x, \mu^*, V_x^*) \quad (15)$$

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 ,

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

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 ,

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

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 $|V^{i+1} - V^i| \leq \varepsilon_v$.

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,

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

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

Approximated optimal cost function and optimal policy are presented:

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

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.

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

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

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

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

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

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

$\Gamma(t_r^+) = \Gamma(0) = \varphi_o I$. 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. $\omega(x) = \nabla \phi^T (f(x) + g(x)u)$ and $1 + \nu \omega^T \Gamma \omega$ 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].

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

for several positive numbers μ_1, μ_2, T .

Where
$$\psi(\tau) = \frac{\omega(t)}{\sqrt{1 + \nu \omega^T \Gamma \omega}}$$

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 $(\varepsilon_{hib}(t))^2$.

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

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 $\eta_c, \eta_{a1}, \eta_{a2}$ 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 $\mu(t)$ is designed based on the RISE framework as follows:

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

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

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

$k_s \in \mathbb{R}$ 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 [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

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

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

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

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

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

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

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 updation law of \hat{W}_c and \hat{W}_a are represented in (23) and (26) with,

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

NN activation function is selected as,

$$\phi(x) = [x_1^2 \quad x_2^2 \quad x_1x_2 \quad x_3^2 \quad x_3^2 \cos(x_2) \quad x_4^2 \quad x_3x_4 \quad x_3x_4 \cos(x_2)]^T.$$

The optimal parameter $W = [2 \ -4 \ 3 \ 2.5 \ 1 \ 1 \ 1 \ 0.5]$ 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 $[2 \ -4 \ 3 \ 2.5 \ 1 \ 1 \ 1 \ 0.5]$. 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.

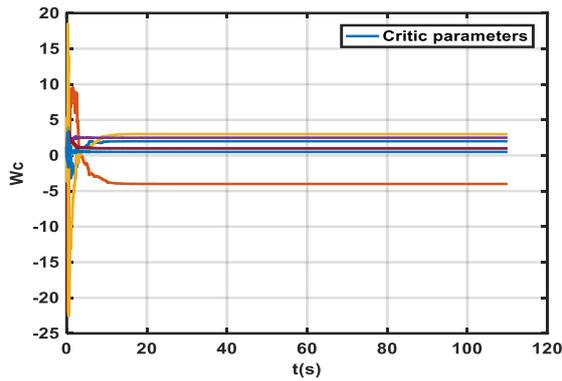


Figure 2. Convergence of critic's parameters

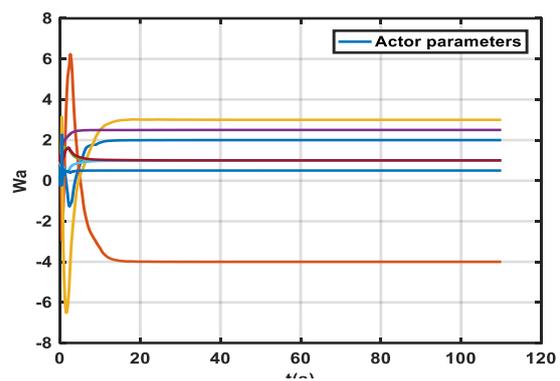


Figure 3. Convergence of actor's parameters

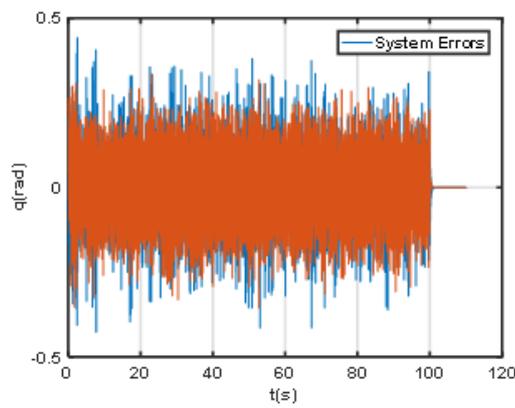


Figure 4. State's evolution

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. Vamvoudakis, 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.