Adaptive Reinforcement Learning Motion/Force Control of Multiple Uncertain Manipulators

Phuong Nam Dao(✉), Dinh Duong Pham, Xuan Khai Nguyen, and Tat Chung Nguyen

School of Electrical Engineering,
Hanoi University of Science and Technology, Hanoi, Vietnam
nam.daophuong@hust.edu.vn

Abstract. Cooperating Mobile Manipulators (CMMs) have been successfully employed as a powerful equipment in industrial automation. The classical nonlinear controllers have been implemented in presence of dynamic uncertainties, external disturbances. However, they are not appropriate in the situation of actuator saturation, input constraint as well as optimality requirements. This paper presents the adaptive reinforcement learning (ARL) based optimal control for CMMs. This approach enables us to overcome the disadvantage of solving Hamilton Jacobi Bellman (HJB) equation to obtain the optimal controller after obtaining the modified motion dynamic model. The simulation results illustrate the performance of the proposed control algorithm.

Keywords: Cooperating mobile manipulators · Adaptive reinforcement learning · Robust adaptive control

1 Introduction

Last decades have witnessed nonlinear control as a promising and powerful technique in designing control scheme for robotic systems. In [1–3], the Lyapunov stability theory are utilized for considering the selection of appropriate Lyapunov function candidate and backstepping technique in Bilateral Teleoperation systems and mobile robots. For Cooperating Mobile Manipulator (CMM) Systems, the control objective was classified into several situations depending on the control design requirement [4–12]. Firstly, multiple mobile robot manipulators in cooperation carrying a common object in presence of dynamic uncertainties, disturbances [11]. Secondly, one of them tightly holds object by end-effector and the end effector of remaining mobile manipulator follows a trajectory on the surface of the object [7]. Due to these control objectives, one realized the methods including centralized or decentralized, cooperation control methods with the tracking requirement of object as well as each mobile manipulators. The scheme in [10] implemented the centralized coordination control to ensure the tracking of each mobile manipulators using approximated Neural Networks by considering the actuators. However, almost existing control strategies for CMMs are implemented by conventional nonlinear control technique. As a result, it is difficult to

D.-T. Tran et al. (Eds.): ICISN 2021, LNNS 243, pp. 265–271, 2021.
https://doi.org/10.1007/978-981-16-2094-2_33

handle the situation of actuator saturations, full-state constraint, etc. This paper considers the novel direction of employing the optimal control for CMMs with the adaptive reinforcement learning (ARL) technique. As we have known that it is impossible to analytically solve Hamilton-Jacobi-Bellman (HJB) equation being an intermediate step of finding the optimal controller. To overcome this challenges, the approach using ARL technique is able to solve online the HJB equation to obtain the optimal scheme.

2 Preliminaries and Problem Statement

We investigate a n two-link manipulators, as shown in Fig. 1, which includes n robots interacted by each pair of manipulators, the unknown environment with the original Cartesian coordinate system $O(X, Y, Z)$ and the corresponding coordinate E_i of end-effector of each robot. The interaction between each pair of manipulators leads to constraint condition as described in the following equation:

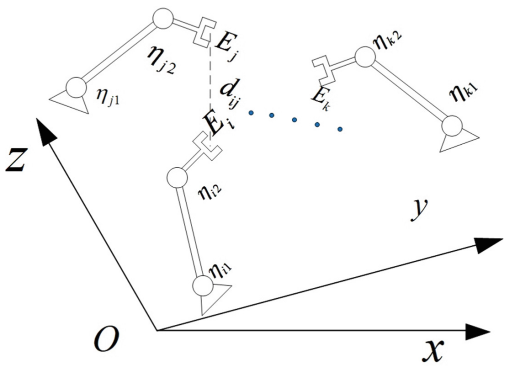

Fig. 1. The cooperating manipulator systems

$$\Xi_{ij} = \left|\mathrm{OE}_i - \mathrm{OE}_j\right| - d_{ij} = 0 (\forall 1 \le i < j \le n) \tag{1}$$

where OE_i is a vector with the first point being the original point and the end-effector E_i under the Cartesian coordinate system, d_{ij} is the distance between end-effectors of the i th robot manipulator and the j th robot manipulator. According to (1), we can be obtain the constraint vector as follows:

$$\Xi(\eta) = \left[\ldots, \quad \Xi_{ij}\left(\eta_i, \eta_j\right), \quad \ldots \right]^T = 0 \tag{2}$$

where $\Xi(\eta) \in R^{m\times 1}$, η is the joint vector of n two-link manipulators systems.

Assumption 1. The number of joints is more than the number of constraint conditions, i.e., $2n > m$.

In [8], the dynamic equation of n two-link robot manipulators with constraint force by interaction between each couples of manipulators can be described as:

$$M(\eta)\ddot{\eta} + C(\eta, \dot{\eta})\dot{\eta} + F(\eta, \dot{\eta}, t) = \tau - J^T(\eta)\lambda \tag{3}$$

where $J(\eta) = \frac{\partial \Xi}{\partial \eta} \in R^{m \times 2\mathrm{n}}$, $\lambda \in R^{m \times 1}$ is the vector of Lagrange multiplier.

The control objective is to obtain not only the tracking of joints to desired trajectories but also guaranteeing the constant distance between end-effectors of each pair of manipulators. Moreover, the Motion/Force control objective is also mentioned in the Largarangian coefficient (3). It is worth emphasizing that the consideration of ARL technique in control design is a significant difference between this work and [11].

3 Motion/Force Control for Multiple Uncertain Manipulators

In this section, we focus on the motion/force control scheme for multiple uncertain manipulators as described in Fig. 2. It will be seen that the robust control algorithm is the starting point of ARL based control design. Furthermore, the stability will be mentioned in these controllers.

3.1 Robust Controller for Multiple Uncertain Manipulators

In this section, the motion/force control design can be established from motion dynamic model and constraint force model. In order to obtain the motion dynamic model, the constraint force $J^T(\eta)\lambda$ in (3) needs to be eliminated by multiplying the matrix $Q = I - A^T(AA^T)^{-1}A$ on both sides of (3) to achive the following motion dynamic equation:

$$QM(\eta)\ddot{\eta} + QC(\eta, \dot{\eta}) + QF(\eta, \dot{\eta}, t) = \tau_{||} \tag{4}$$

It sholud be noted that due to the multiple manipulators system is considered in the case of $2n > m$ (assumption 1), it leads to that the completed motion dynamic model (5) is the differential equation of independent joint $\eta_{ind} \in R^{(2n-m) \times 1}$ with the detailed transformation being seen in [11]:

$$K_1\overline{M}\ddot{\eta}_{ind} + K_1\overline{C}(\eta, \dot{\eta})\dot{\eta}_{ind} + K_1\overline{F} = \tau_{||ind} \tag{5}$$

Similar to the work in [11], the constraint force can be obtained by the following equation:

$$F_c = \tau_{\perp} - h_{\perp} - F_{\perp} - (I - Q)MM_c^{-1}\left(\tau_{||} - Q(C\dot{\eta} + F(\eta, \dot{\eta}, t)) + M(\eta)S\dot{\eta}\right) \tag{6}$$

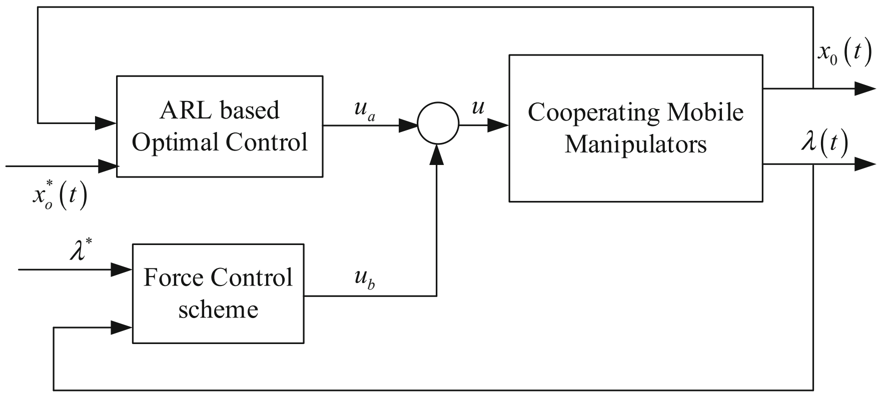

Fig. 2. The motion/force controller of cooperating mobile manipulators.

Based on the certainty equivalence principle, appropriate assumptions as described in [8], the robust adaptive controller can be designed as follows:

$$\begin{aligned} \tau_{||ind} &= Y\hat{\theta} - e_1 - \overline{\overline{M}}e_2 - k_2 e_2 - k_3 \|e_1\|^2 e_2 \\ \tau_{||idep} &= f\left(\tau_{||ind}, \dot{\eta}_{ind}\right) \end{aligned} \tag{7}$$

3.2 Adaptive Reinforcement Learning for Multiple Uncertain Manipulators

In this section, we further extend the results in the previous section to ARL based control design with the consideration of motion dynamic model (5) and the appropriate tracking error model. It can be seen that the tracking error model is a time-varying system. As a result, it should be transformed into the autonomous systems by using the additional terms as:

$$\begin{gathered} \tau_d = K_1\overline{M}\ddot{\eta}_{ind}^{ref} + K_1\overline{C}\dot{\eta}_{ind}^{ref} + K_1\overline{F}; \\ X = \begin{pmatrix} z_{dq} \\ z_{qind} \\ q_{ind}^d \end{pmatrix}; u = \tau - \tau_d; l(x) = K_1(x)\overline{C(x)}.x; h_1 = \dot{\eta}_{ind}^d; h_2 = \ddot{\eta}_{ind}^d \end{gathered} \tag{8}$$

According to (5) and (8), the affine system can be given as:

$$\dot{X} = \begin{bmatrix} -(K_1\overline{M})^{-1}.l(z_{dq} + \dot{\eta}_{ind}^{ref}) + (K_1\overline{M})^{-1}.l(\dot{\eta}_{ind}^{ref}) \\ (z_{qind} + \eta_{ind}^d)z_{dq} - \beta z_{qind} \\ h_1 \end{bmatrix} + \begin{bmatrix} (K_1\overline{M})^{-1} \\ 0 \\ 0 \end{bmatrix} u \tag{9}$$

Therefore, the ARL based controller can be proposed with performance index $V = \int_{0}^{\infty} (X^T Q_T X + u^T Ru) ds$ as follows:

$$\hat{V} = \hat{W}_C^T \phi(X) \tag{10}$$

$$\hat{u}(X) = -\frac{1}{2} R^{-1} H^T(X) \left(\frac{\partial \phi}{\partial X} \right)^T \hat{W}_a \tag{11}$$

$$\dot{\hat{W}}_C = -\eta_C \Gamma \frac{\gamma}{1 + v\gamma^T \Gamma \gamma} \sigma_{hjb} \tag{12}$$

$$\dot{\hat{W}}_a = -\frac{\eta_{a1}}{\sqrt{1 + \gamma^T \gamma}} . \frac{\partial \phi}{\partial X} .H(X).R^{-1} H^T(X). \left(\frac{\partial \phi}{\partial X} \right)^T .(\hat{W}_a - \hat{W}_c)\sigma_{hjb} - \eta_{a2}(\hat{W}_a - \hat{W}_c) \tag{13}$$

4 Simulation Results

In this section, the system contains three two-link manipulators is considered to simulate with the purpose of validating the proposed controller. The parameters of three manipulators used in this simulation are as follows: D_{12} = 7 m, D_{13} = 3.5 m, $l_1 = l_2$ = 1 m, $m_1 = m_2$ = 1 kg, $r_1 = r_2$ = 0.9 m, $l_3 = l_4$ = 1.5 m, $m_3 = m_4$ = 2 kg, $r_3 = r_4$ = 1.2 m, $l_5 = l_6$ = 1.2 m, $m_5 = m_6$ = 1.5 kg, $r_5 = r_6$ = 1 m. Matrix M_1, M_2, and M_3 of three manipulators can be rewritten as

$$M_1 = \left[\widehat{\theta}_1 + 2\widehat{\theta}_2 \cos\eta_2 \widehat{\theta}_3 + \widehat{\theta}_2 \cos\eta_2; \widehat{\theta}_3 + \widehat{\theta}_2 \cos\eta_2 \widehat{\theta}_3 \right]$$
$$M_2 = \left[\widehat{\theta}_4 + 2\widehat{\theta}_5 \cos\eta_4 \widehat{\theta}_6 + \widehat{\theta}_5 \cos\eta_4; \widehat{\theta}_6 + \widehat{\theta}_5 \cos\eta_5 \widehat{\theta}_6 \right]$$
$$M_3 = \left[\widehat{\theta}_7 + 2\widehat{\theta}_8 \cos\eta_6 \widehat{\theta}_9 + \widehat{\theta}_8 \cos\eta_6; \widehat{\theta}_9 + \widehat{\theta}_8 \cos\eta_6 \widehat{\theta}_9 \right]$$

The desired trajectories for four independent joints are previously established as $\eta_{1d} = \frac{\pi}{12}; \eta_{2d} = 1,91\pi + 0,2\sin(t); \eta_{3d} = 0,51\pi; \eta_{5d} = 0,191\pi$. The tracking effectiveness of join variables responses in three manipulators system and adaptation Law of uncertain parameters in (7) are shown in Fig. 3.

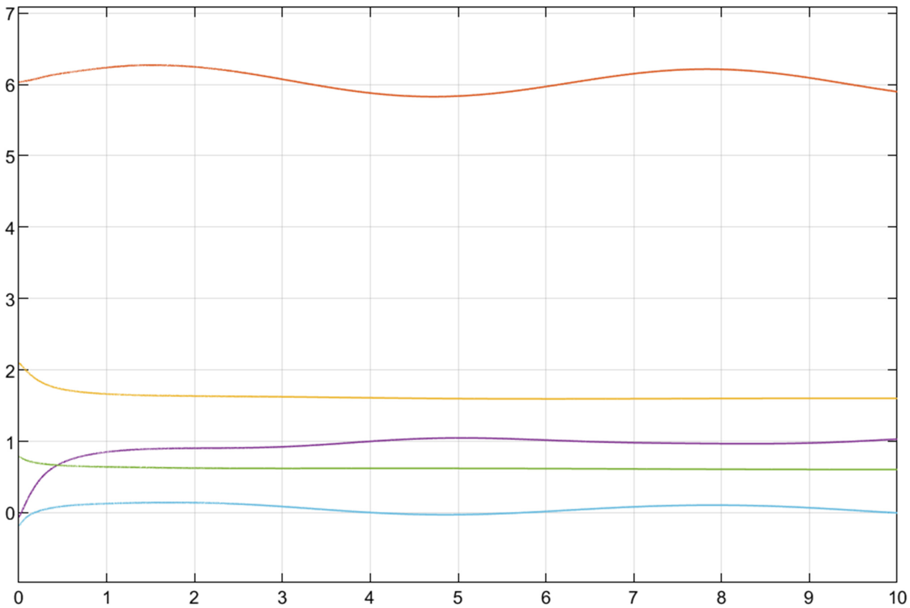

Fig. 3. The response of trajectories of η_i

5 Conclusions

In this paper, ARL based motion/force control have been proposed for solving the optimal control problem of multiple uncertain manipulators. The motion/force control structure is given to develop the ARL technique based on motion dynamic model after the separation technique of cooperating manipulators was implemented. However, the control design requires the ARL based solution for time-varying systems. For this reason, the additional terms in (8) are proposed to handle this disadvantage. Finally, the online Actor/Critic ARL is developed with Neural Network and optimality principle for modified motion dynamic model.

References

1. Liu, Y., Dao, P.N., Zhao, K.Y.: On robust control of nonlinear teleoperators under dynamic uncertainties with variable time delays and without relative velocity. IEEE Trans. Ind. Inform. **16**(2), 1272–1280 (2020). https://doi.org/10.1109/TII.2019.2950617
2. Nguyen, T.B., et al.: An adaptive backstepping trajectory tracking control of a tractor trailer wheeled mobile robot. International J. Control Autom. Syst. **17**(2), 465–473 (2019)
3. Van Tinh, N., et al.: A Gaussian wavelet network-based robust adaptive tracking controller for a wheeled mobile robot with unknown wheel slips. Int. J. Control (2019). https://doi.org/10.1080/00207179.2018.1458156
4. Baek, J., Kwon, W., Kim, B., Han, S.: A widely adaptive time delayed control and its application to robot manipulators. IEEE Trans. Ind. Electron. **66**, 5332–5342 (2018)

5. Bian, T., Jiang, Y., Jiang, Z.P.: Adaptive dynamic programming and optimal control of nonlinear nonaffine systems. Automatica **50**, 2624–2632 (2014)
6. Cao, F., Liu, J.: Three-dimensional modeling and input saturation control for a two-link flexible manipulator based on infinite dimensional model. J. Franklin Inst. **357**, 1026–1042 (2020)
7. Chen, C., Liu, Z., Zhang, Y., Xie, S.: Coordinated motion/force control of multiarm robot with unknown sensor nonlinearity and manipulated object's uncertainty. IEEE Trans. Syst. Man Cybern.: Syst. **47**, 1123–1134 (2016)
8. Deng, M., Li, Z., Kang, Y., Chen, C.P., Chu, X.: A learning based hierarchical control scheme for an exoskeleton robot in human–robot cooperative manipulation. IEEE Trans. Cybern. **50**, 112–125 (2018)
9. Dohmann, P.B.G., Hirche, S.: Distributed control for cooperative manipulation with event-triggered communication. IEEE Trans. Robot. **36**(4), 1038–1052 (2020)
10. Dong, Y., Ren, B.: UDE-based variable impedance control of uncertain robot systems. IEEE Trans. Syst. Man Cybern.: Syst. **49**(12), 2487–2498 (2017)
11. Dou, H., Wang, S.: Robust adaptive motion/force control for motion synchronization of multiple uncertain two-link manipulators. Mech. Mach. Theory **67**, 77–93 (2013)
12. Guo, X., Yan, W., CUI, R.: Integral reinforcement learning based adaptive NN control for continuous-time nonlinear MiMo systems with unknown control directions. IEEE Trans. Syst. Man Cybern.: Syst. **50**(11), 4068–4077 (2019)