基于自适应动态规划的移动装弹机械臂轨迹控制

唐志国[1], 张富尧[2], 马　彦[1†]
(1. 吉林大学 通信工程学院, 吉林 长春 130022; 2. 吉林大学 机械与航天工程学院, 吉林 长春 130022)

摘要: 针对移动装弹机械臂系统非线性、强耦合、受多种不确定因素影响的问题, 本文基于自适应动态规划方法, 提出了仅包含评价网络结构的轨迹跟踪控制方法, 有效减小了系统跟踪误差. 首先, 考虑到系统非线性特性、变量间强耦合作用及重力因素的影响, 通过拉格朗日方程建立了移动装弹机械臂的动力学模型. 其次, 针对系统存在不确定性上界未知的问题, 建立单网络评价结构, 通过策略迭代算法, 求解哈密顿–雅可比–贝尔曼方程, 基于李雅普诺夫稳定性理论, 设计了自适应动态规划轨迹跟踪控制方法. 最后, 通过仿真实验将该控制方法与自适应滑模控制方法进行了对比, 进一步检验了所设计控制方法的有效性.

关键词: 移动机械臂; 自适应动态规划; 轨迹控制; 拉格朗日方程

引用格式: 唐志国, 张富尧, 马彦. 基于自适应动态规划的移动装弹机械臂轨迹控制. 控制理论与应用, 2021, 38(9): 1442 – 1451

DOI: 10.7641/CTA.2021.00702

Adaptive dynamic programming based trajectory tracking control of mobile missile-loading manipulator

TANG Zhi-guo[1], ZHANG Fu-yao[2], MA Yan[1†]
(1. College of Communication Engineering, Jilin University, Changchun Jilin 130022, China;
2. College of Mechanical and Aerospace Engineering, Jilin University, Changchun Jilin 130022 China)

Abstract: In this paper, the trajectory tracking control with critic-only structure is studied to reduce tracking error based on adaptive dynamic programming in the mobile missile-loading manipulator system, when there are nonlinear, strong coupling and many kinds of uncertainties. Firstly, a dynamic model using Lagrange equation is established for the mobile missile-loading manipulator while all of the above impacts, besides gravity of manipulator are simultaneously considered. Furthermore, in the case of unknown upper bound uncertainty in the system, an adaptive dynamic programming trajectory tracking controller is presented to improve the control precision. According to neural network algorithm, critic-only policy iteration algorithm for mobile missile-loading manipulator is proposed. Using the policy iteration method can solve the HJB equation and approximate the optimal control strategy. And then the asymptotic stability of closed-loop system is proved by Lyapunov stability theory. Finally, the effectiveness of the designed control method is further verified by simulation, compared with the adaptive sliding mode controller.

Key words: mobile missile-loading manipulator; adaptive dynamic programming; trajectory tracking control; Lagrange equation

Citation: TANG Zhiguo, ZHANG Fuyao, MA Yan. Adaptive dynamic programming based trajectory tracking control of mobile missile-loading manipulator. *Control Theory & Applications*, 2021, 38(9): 1442 – 1451

1　引言

随着现代战争信息化的快速发展, 战术打击速度不断提升, 战争进程急剧缩短, 因此现代战争除了对导弹自身参数提出严格要求外, 对导弹的补给, 即装填速度同样也提出了更高的要求[1]. 然而对于质量大、体积大的常规导弹, 以往通常是采用人工指挥吊车的方法完成导弹装载工作, 会产生准确性低、实时性差、费效比低等问题, 不利于现代化战争的进程[2–3]. 为了适应国防现代化发展的需要, 将机械臂安装在汽车类移动平台上, 实现了通过一个机械臂系统

收稿日期: 2020−10−14; 录用日期: 2021−04−23.
†通信作者. E-mail: mayan_maria@163.com; Tel.: +86 13944003569.
本文责任编委: 黄攀峰.
省校共建项目(SXGJSF2017–2–1–1), 吉林省科技发展计划项目(20170520060JH), 吉林省产业创新专项基金项目(2018C035–2)资助.
Supported by the Provincial-University Co-Construction Project (SXGJSF2017–2–1–1), the Jilin Science and Technology Development Project (20170520060JH) and the Industrial Innovation Special Funds of Jilin Province (2018C035–2).

一次性完成吊装导弹与运输转载两种任务的目的. 该类机械臂系统因其具有功能多、效率高、平稳性强、快速性好及作用范围大等优点, 受到越来越多国内外学者的密切关注[4–8].

考虑到大型移动机械臂是一个高度非线性、强耦合的时变系统, 而为了确保装载的平稳性和精确性, 装载机械臂的移动平台通常情况下会呈现出质量大、动力学响应速度慢的特性, 这些均增加了系统建模与控制的难度. 文献[9–11]在微分同胚和非线性输入变换基础上, 将移动机械臂的动力学模型降阶分解为4个低维子系统. 在不加装关节力/力矩传感器的情况下, 分别采用滑模和二阶滑模控制算法设计了鲁棒跟踪控制器. 虽然采用滑模控制可以克服系统的不确定性、提高系统的抗干扰性, 但输出力矩可能会产生抖振. 文献[12]利用模糊控制来改善滑模趋近律, 有效抑制了控制力矩的抖振. 文献[13]对已经整合了四轮移动平台和三自由度机械臂的系统分别设计出了滑模和非奇异端滑模控制器. 此外, 文献[14]在对移动机械臂的动力模型简化处理后, 在参考坐标系下, 研究了包含电机驱动动力学的滑模控制问题. 为确保系统在不确定性和外部扰动的作用下仍能正常工作, 文献[15]研究了非完整约束轮式移动机械臂滑模轨迹跟踪控制问题, 利用反步法保证了速度跟踪性能. 基于串级控制思想, 设计了考虑控制饱和的最小范数运动学控制器及动力学补偿器.

当大型移动机械臂系统受到模型参数不确定性、未建模动态及外界干扰等多种不确定性因素的影响时, 特别是当系统不确定性难以测量, 且上界未知的情况下, 传统的滑模控制方法难以应对. 自适应动态规划(adaptive dynamic programming, ADP)是一种综合了动态规划、神经网络和强化学习的近似智能控制方法, 可以有效处理动态规划中的"维数灾"难题, 对复杂非线性系统的控制具有重要意义[16]. 近年来该算法已在高超声速飞行器[17]、轮式机器人控制[18]、导弹制导律[19]、随机系统[20]等领域得到大量应用, 逐渐成为利用最优性原理控制高度非线性、强耦合时变系统的一种非常有用的工具.

因此, 本文针对移动装弹机械臂系统, 基于自适应动态规划方法, 提出了仅包含评价网络结构的轨迹跟踪控制方法, 在保证轨迹跟踪误差最终一致有界的前提下, 有效提高了系统的实时性. 首先, 建立了移动装弹机械臂的动力学模型, 并定义了总体不确定性函数. 然后, 利用单网络评价结构策略迭代计算的自适应动态规划算法, 求解哈密顿–雅可比–贝尔曼(Hamilton–Jacob–Behrman, HJB)方程, 完成近似最优控制方法设计. 最后, 仿真实验验证了本文所设计控制方法的有效性.

2　移动装弹机械臂系统描述

移动装弹机械臂由连杆、套筒、3个伸缩关节、3个旋转关节和两个回转关节组成, 共计8个自由度, 且其驱动力由液压系统提供, 与传统机械臂相比更为复杂. 移动装弹机械臂示意图如图1所示.

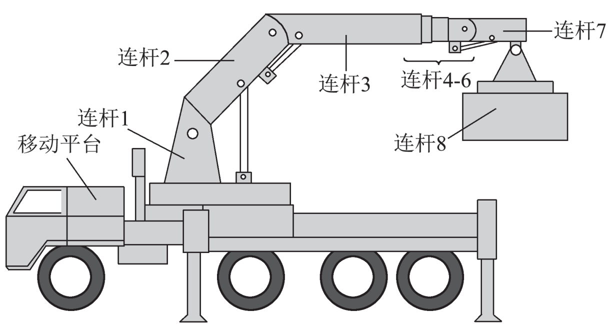


图 1　移动装弹机械臂示意图

Fig. 1　Schematic diagram of mobile missile-loading manipulator

机械臂的基座定义为自由度1, 调整末端姿态的关节定义为自由度8, 二者皆为回转关节, 可以通过控制回转的角度, 调节机械臂抓取导弹的空间位置指向; 机械臂的两个旋转关节分别定义为自由度2和3; 机械臂的伸缩关节具有三段式同步伸缩功能, 定义为自由度4, 5和6, 该结构可以让机械臂末端达到更远的位置; 此外, 机械臂还有一个腕关节, 定义为自由度7, 该关节的作用是在导弹质量过重的情况下减小末端重物产生的附加力矩, 从而保证负载的方向始终垂直向下.

移动装弹机械臂各关节坐标系建立情况如图2所示. 规定连杆i的轴线方向为Z_i轴, X_i轴在垂直于Z_i轴的方向上, Y_i轴是X_iOZ_i平面的法线方向且满足右手法则, 规定逆时针为正. 定义参考坐标系$\{0\}$, 且其固定于移动装弹机械臂的基座上, 当回转关节θ_1为0时, 坐标系$\{0\}$与坐标系$\{1\}$重合, 且Z_0轴与连杆1轴线重合.

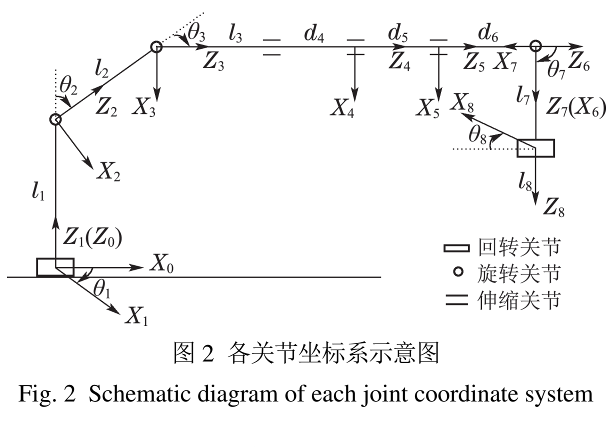


图 2　各关节坐标系示意图

Fig. 2　Schematic diagram of each joint coordinate system

3　移动装弹机械臂动力学模型

3.1　动力学模型建立

拉格朗日方程是基于功–能平衡法构建系统动力

学模型. 首先, 分别求取机械臂系统各连杆的动能和势能, 以连杆2和连杆4为例.

连杆2上任一点在参考坐标系下表示为

$$ {}^0\gamma = {}^0T_2{}^2\gamma = \begin{bmatrix} c_1 s_2\gamma \\ s_1 s_2 \gamma \\ c_2\gamma + l_1 \\ 1 \end{bmatrix}. \tag{1}$$

连杆2的动能和势能分别为

$$\begin{aligned} T_2 = & \int_{\text{link}_2} \frac{1}{2}{}^0\dot{\gamma}^{\mathrm{T}0}\dot{\gamma}dm + \frac{1}{2}{}^0\dot{\gamma}_{l_0}^{\mathrm{T}0}\dot{\gamma}_{l_0}m_3 = \\ & \frac{1}{6}\rho_2 l_2^3(s_2^2\dot{\theta}_1^2 + \dot{\theta}_2^2) + \frac{1}{2}m_3 l_2^2(s_2^2\dot{\theta}_1^2 + \dot{\theta}_2^2), \end{aligned} \tag{2}$$

$$\begin{aligned} V_2 = & \int_{\text{link}_2} \mathrm{g}(rc_2 + l_1)\rho_2 d\gamma + m_3\mathrm{g}(l_2c_2 + l_1) = \\ & \rho_2\mathrm{g}(l_1l_2 + \frac{1}{2}l_2^2c_2) + m_3\mathrm{g}(l_2c_2 + l_1). \end{aligned} \tag{3}$$

伸缩杆可视为连杆4, 其上面任一点在参考坐标系下表示为

$$ {}^0\gamma = {}^0T_4{}^4\gamma = \begin{bmatrix} c_1s_{23}(l_3 + d - \gamma) + c_1s_2l_2 \\ s_1s_{23}(l_3 + d - \gamma) + s_1s_2l_2 \\ c_{23}(l_3 + d - \gamma) + c_2l_2 + l_1 \\ 1 \end{bmatrix}. \tag{4}$$

由于伸缩关节坐标建立在伸缩杆末端, 在计算该部分的动能和势能时需考虑该关节及下一关节的动能和势能. 设三个伸缩杆中每段的最大伸缩长度均为ΔD, 当伸缩杆全部缩回时的长度为固定在连杆3中的长度D. 由于伸缩部分采用套筒式同步结构, 且为等速伸缩, 即当一个伸缩杆伸出Δx, 则三段伸缩杆总伸缩长度为$d = 3\Delta x$; 根据多级伸缩结构可知, 伸缩部分整体伸出长度为$D + \frac{2}{3}d$, 如图3所示.

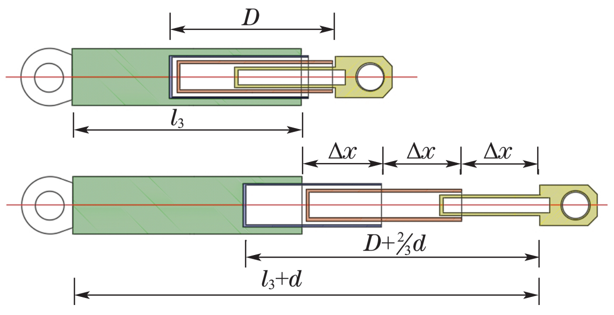


图 3 伸缩连杆示意图

Fig. 3 Schematic diagram of telescopic connecting rod

伸缩杆部分的动能和势能分别为

$$\begin{aligned} T_4 = & \frac{1}{2}M_4\{[(l_3+d)^2 - (l_3+d)(D+\frac{2}{3}d) + \frac{1}{3}(D+ \\ & \frac{2}{3}d)^2](s_{23}^2\dot{\theta}_1^2 + (\dot{\theta}_2+\dot{\theta}_3)^2) + l_2^2(s_2^2\dot{\theta}_1^2 + \dot{\theta}_2^2) + \\ & \dot{d}^2 + [2(l_3+d) - (D + \frac{2}{3}d)]l_2(s_2s_{23}\dot{\theta}_1^2 + \\ & C_3\dot{\theta}_2(\dot{\theta}_2 + \dot{\theta}_3) + 2l_2s_3\dot{\theta}_2\dot{d}\}, \end{aligned} \tag{5}$$

$$\begin{aligned} V_4 = & M_4\mathrm{g}[l_1 + l_2c_2 + c_{23}(l_3 + d) - \\ & \frac{1}{2}c_{23}(D + \frac{2}{3}d)] = \\ & M_4\mathrm{g}[l_1 + l_2c_2 + c_{23}(l_3 + \frac{2}{3}d - \frac{1}{2}D)]. \end{aligned} \tag{6}$$

伸缩杆后模拟手腕功能的连杆用连杆7描述, 矩形刚体夹具用连杆8描述, 以及连杆1和连杆3, 它们动能与势能的计算步骤、计算方法与连杆2类似. 将所求的所有动能与势能代入到拉格朗日方程

$$\frac{\mathrm{d}}{\mathrm{d}t}(\frac{\partial L}{\partial \dot{q}_i}) - \frac{\partial L}{\partial q_i} = Q_i,\ i = 1, 2, \cdots, n, \tag{7}$$

其中: $L = T - V$; T为系统总动能; V为系统总势能; q为广义坐标; Q为广义力; n为连杆个数.

推导、整理可得系统动力学方程为

$$M(q)\ddot{q} + C(q, \dot{q})\dot{q} + G(q) = u, \tag{8}$$

其中: $M(q)$为惯性矩阵; $C(q, \dot{q})$为离心力与哥氏力项; $G(q)$为重力项; $u = [u_1\ u_2\ u_3\ u_4\ u_7\ u_8]^{\mathrm{T}}$为广义力, 即机械臂的控制力矩, 由液压驱动系统提供; $q = [\theta_1\ \theta_2\ \theta_3\ d\ \theta_7\ \theta_8]^{\mathrm{T}}$为广义坐标, 包含回转角度、旋转角度及伸缩长度.

动力学方程中具体变量含义为: J_1和J_2分别为连杆1和关节2的转动惯量; ρ_1为连杆1的面密度; A_1为连杆1的微元面积; ρ_2, ρ_3, ρ_7和ρ_8分别为连杆2、连杆3、连杆7和矩形刚体夹具的线密度; m_2, m_3和m_8分别为关节2、关节3和末端回转关节的质量; M_4为伸缩杆的总质量; l_1, l_2, l_3, l_7和l_8分别为连杆1、连杆2、连杆3、连杆7和矩形刚体夹具的长度; θ_1和θ_8为回转关节的回转角度; θ_2, θ_3和θ_7为旋转关节的旋转角度; d为伸缩关节的伸缩长度; g为重力加速度. 此外, s_i 及 c_{ijk}等均为正弦和余弦的简写形式, 如: $c_{237} = \cos(\theta_2 + \theta_3 + \theta_7)$, $s_3 = \sin\theta_3$ 等.

3.2 不确定性及状态空间描述

移动装弹机械臂的动力学方程是一个非线性、强耦合、时变的复杂系统, 考虑到实际系统中会存在参数不确定性和外部干扰等影响, 将系统动力学方程整理为

$$\bar{M}(q)\ddot{q} + \bar{C}(q, \dot{q})\dot{q} + \bar{G}(q) + \psi = u, \tag{9}$$

$$\psi = \Delta M(q)\ddot{q} + \Delta C(q, \dot{q})\dot{q} + \Delta G(q) + d^*, \tag{10}$$

其中: ψ表示机械臂系统的总体不确定性; $\Delta M(q)$, $\Delta C(q, \dot{q})$和$\Delta G(q)$为系统参数中的不确定性; d^*为外部干扰项; $\bar{M}(q)$, $\bar{C}(q, \dot{q})$和$\bar{G}(q)$为系统模型的标称参数. 下文中会将$\bar{M}(q)$简写成$\bar{M}$, 其他类似.

假设 1 不确定性ψ具有未知上界ψ^*, 即有

$$\|\psi\| \leqslant \psi^*. \tag{11}$$

设$x=[x_1\ \ x_2]^{\mathrm{T}}=[q\ \ \dot{q}]^{\mathrm{T}}$, 则有

$$\begin{cases}\dot{x}_1=x_2,\\ \dot{x}_2=f(x)+g(x)u,\\ y=x_1,\end{cases}\tag{12}$$

其中: $f(x)=\bar{M}^{-1}(-\bar{C}x_2-\bar{G}-\psi)$, $g(x)=\bar{M}^{-1}$, 且$f(x)$和$g(x)$均为Lipschitz函数.

4 基于ADP的轨迹控制方法设计

4.1 控制器设计

在经典的ADP控制结构中, 既包含了执行网络, 又包含了评价网络, 最优反馈控制律将由二者共同求得. 而本文将执行网络舍弃, 仅保留评价网络, 构成单网络评价结构ADP控制, 其最优反馈控制仅依赖于评价网络输出的最优性能指标函数的梯度, 通过在线迭代求得[21–22]. 不但简化了训练过程, 而且消除了两网络间的近似误差. 移动装弹机械臂轨迹控制系统结构框图如图4所示, 具体设计如下.

假设 2 期望关节角度、角速度及角加速度皆具有上界, 显然, $f(x)$和$g(x)$亦有界.

令x_{d}表示系统的期望轨迹, x表示系统的实际轨迹, 则系统的轨迹跟踪误差为

$$e=x-x_{\mathrm{d}}.\tag{13}$$

定义性能指标如下:

$$J(e(\tau))=\int_0^{\infty}N(e(\tau),u(e(\tau)))\mathrm{d}\tau,\tag{14}$$

其中$N(e(\tau),u(e(\tau)))=e^{\mathrm{T}}Qe+u^{\mathrm{T}}Ru$为效应函数. 有$N(0,0)=0$, 对所有的$e$和$u$均有$N(e,u)>0$成立. 且$Q\in\mathbb{R}^{n\times n}$, $R\in\mathbb{R}^{m\times m}$为正定矩阵.

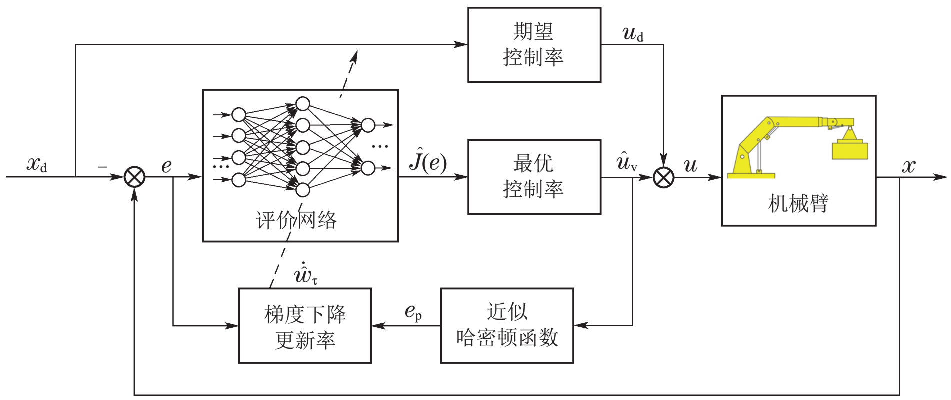


图 4 移动装弹机械臂系统控制结构框图

Fig. 4 Block diagram of mobile missile-loading manipulator control system

令u_{d}表示期望控制律, 则

$$u_{\mathrm{d}}=g^{+}(x_{\mathrm{d}})(\dot{x}_{\mathrm{d}}-f(x_{\mathrm{d}})).\tag{15}$$

因此

$$\dot{e}=f(x)-f(x_{\mathrm{d}})+g(x)u-g(x_{\mathrm{d}})u_{\mathrm{d}}.\tag{16}$$

系统控制律u包含期望控制律u_{d}和最优控制律u_{v}两部分, 即

$$u=u_{\mathrm{d}}+u_{\mathrm{v}}.\tag{17}$$

因此

$$\begin{aligned}\dot{e}=&f(x)-f(x_{\mathrm{d}})+g(x)u-g(x_{\mathrm{d}})u_{\mathrm{d}}=\\ &f_{\mathrm{e}}+[g(x)-g(x_{\mathrm{d}})]u+g(x_{\mathrm{d}})u_{\mathrm{v}}.\end{aligned}\tag{18}$$

最优控制律u_{v}可以保证系统的轨迹跟踪误差以最优方式收敛于系统的稳定状态.

式(14)可改写为

$$J(e(\tau))=\int_0^{\infty}N(e(\tau),u_{\mathrm{v}}(e(\tau)))\mathrm{d}\tau,\tag{19}$$

$$u_{\mathrm{v}}\in\varPhi(\varOmega),\tag{20}$$

其中: $N(e,u_{\mathrm{v}})=e^{\mathrm{T}}Qe+u_{\mathrm{v}}^{\mathrm{T}}Ru_{\mathrm{v}}$为效应函数, 且$N(0,0)=0$, 对所有的$e$和$u_{\mathrm{v}}$都有$N(e,u_{\mathrm{v}})>0$成立. $\varPhi(\varOmega)$为一组容许控制序列.

定理 1 针对移动装弹机械臂控制系统(18), 对$\forall e\in\varOmega$, 若存在一组容许控制$u(e)\in\varPhi(\varOmega)$, 且在$\varOmega$上连续, 并满足$u(e)=0$, 则$u(e)$即可保证机械臂系统在紧集$\varOmega\in\mathbb{R}^n$上收敛, 且性能指标函数$\varOmega\in\mathbb{R}^n$有限.

若性能指标(19)连续可微, 则其无穷小的形式可表示为

$$\begin{aligned}0=&N(e,u_{\mathrm{v}})+(\nabla J(e))^{\mathrm{T}}(f_{\mathrm{e}}+\\ &[g(x)-g(x_{\mathrm{d}})]u+g(x_{\mathrm{d}})u_{\mathrm{v}}),\end{aligned}\tag{21}$$

其中: $J(0)=0$, $N(0,0)=0$, $\nabla J(e)$为$J(e)$的关于e的偏导数, 即$\nabla J(e)=\dfrac{\partial J(e)}{\partial e}$.

定义哈密顿函数及最优性能指标分别为

$$\begin{aligned}H(e,u,\nabla J(e))=&N(e,u_{\mathrm{v}})+(\nabla J(e))^{\mathrm{T}}(f_{\mathrm{e}}+\\ &[g(x)-g(x_{\mathrm{d}})]u+g(x_{\mathrm{d}})u_{\mathrm{v}}),\end{aligned}\tag{22}$$

$$J^*(e)=\min_{u_{\rm v}}\int_0^\infty N(e(\tau),u_{\rm v}(e(\tau)))\mathrm{d}\tau, \tag{23}$$

显然, $J^*(e)$满足

$$0=\min_{u_{\rm v}}H(e,u_{\rm v},\nabla J^*(e)), \tag{24}$$

其中$\nabla J^*(e)=\dfrac{\partial J^*(e)}{\partial e}$.

若$J^*(e)$存在, 且连续可微, 则可以通过单网络评价结构策略迭代算法, 循环迭代求解出最优反馈控制律为

$$u_{\rm v}^*=-\frac{1}{2}R^{-1}g^{\rm T}(x)\nabla J^*(e). \tag{25}$$

综合式(21)和式(24), 有

$$-e^{\rm T}Qe-u_{\rm v}^{\rm T}Ru_{\rm v}=(\nabla J(e))^{\rm T}(f_{\rm e}+[g(x)-g(x_{\rm d})]u+g(x_{\rm d})u_{\rm v}). \tag{26}$$

单网络评价结构策略迭代流程如图5所示, 在算法中, 使用式(21)进行控制策略评价, 基于评价结果利用式(25)求取最优反馈控制律, 提升系统调节效果, 性能指标函数$J(e)$采用神经网络近似, 有

$$J(e)=w_\tau^{\rm T}\sigma_\tau(e)+\varepsilon_\tau, \tag{27}$$

其中: $w_\tau\in\mathbb{R}^l$为理想神经网络的权值; l为隐含层神经元个数; $\sigma_\tau(e)$为神经网络激活函数; ε_τ为评价网络的近似误差. 则$J(e)$的梯度可表示成

$$\nabla J(e)=(\nabla\sigma_\tau(e))^{\rm T}w_\tau+\nabla\varepsilon_\tau, \tag{28}$$

其中$\nabla\sigma_\tau(e)=\dfrac{\partial\sigma_\tau(e)}{\partial e}\in\mathbb{R}^{N\times n}$与$\nabla\varepsilon_\tau$分别用来表示$\sigma_\tau(e)$与$\varepsilon_\tau$的梯度.

将式(28)代入式(21), 可得哈密顿函数为

$$H(e,u_{\rm v},w_\tau)=N(e,u_{\rm v})+(w_\tau^{\rm T}\nabla\sigma_\tau(e))\dot e=-\nabla\varepsilon_\tau^{\rm T}\dot e\triangleq e_{\rm PH}, \tag{29}$$

其中$e_{\rm PH}$为逼近神经网络的残余误差.

定义$\hat w_\tau$和$\tilde w_\tau$分别为评价神经网络权值w_τ的估计值和估计误差. 因此, 评价网络的输出$\hat J(e)$及其梯度分别为

$$\hat J(e)=\hat w_\tau^{\rm T}\sigma_\tau(e), \tag{30}$$

$$\nabla\hat J(e)=(\nabla\sigma_\tau(e))^{\rm T}\hat w_\tau. \tag{31}$$

近似的哈密顿函数为

$$H(e,u_{\rm v},\hat w_\tau)=N(e,u_{\rm v})+(\hat w_\tau^{\rm T}\nabla\sigma_\tau(e))\dot e\triangleq e_{\rm P}. \tag{32}$$

神经网络训练过程需要最小化的性能准则[21]为

$$E_{\rm P}=\frac{1}{2}e_{\rm P}^{\rm T}e_{\rm P}. \tag{33}$$

权值则采用梯度下降法来更新, 有

$$\dot{\hat w}_\tau=-\alpha_\tau e_{\rm P}\eta, \tag{34}$$

其中$\eta=\nabla\sigma_\tau(e)\dot e$, $\alpha_\tau>0$为评价网络的学习率.

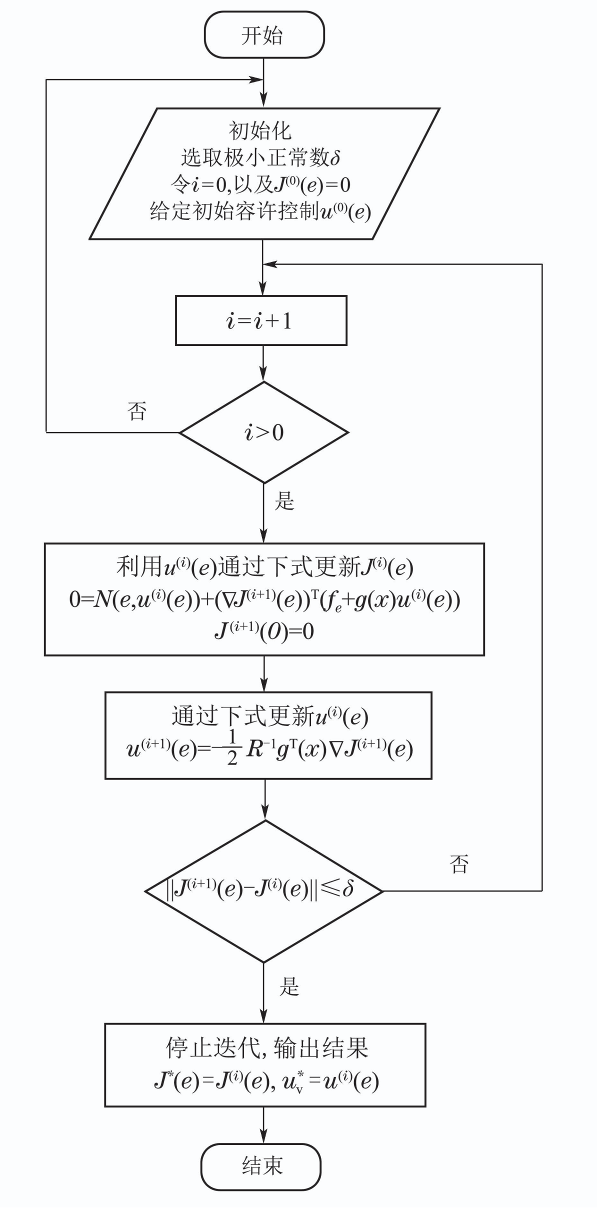


图 5 单网络评价结构策略迭代算法流程图

Fig. 5 Flowchart of critic-only policy iteration algorithm

由于

$$\tilde w_\tau=w_\tau-\hat w_\tau, \tag{35}$$

所以有

$$e_{\rm P}=e_{\rm PH}-w_\tau^{\rm T}\nabla\sigma_\tau(e)\dot e, \tag{36}$$

因此, 权值估计误差更新率为

$$\dot{\tilde w}_\tau=\alpha_\tau(e_{\rm PH}-\tilde w_\tau^{\rm T}\nabla\sigma_\tau(e)\dot e)\nabla\sigma_\tau(e)\dot e, \tag{37}$$

故, 理想最优反馈控制律及相应迭代控制律分别为

$$u_{\rm v}=-\frac{1}{2}R^{-1}g^{\rm T}(x)((\nabla\sigma_\tau(e))^{\rm T}w_\tau+\nabla\varepsilon_\tau), \tag{38}$$

$$\hat u_{\rm v}=-\frac{1}{2}R^{-1}g^{\rm T}(x)(\nabla\sigma_\tau(e))^{\rm T}\hat w_\tau. \tag{39}$$

假设 3　期望控制律u_{d}与η均有未知上界, 即

$$\|u_{\mathrm{d}}\| \leqslant w_{\mathrm{ud}}, \tag{40}$$

$$\|\eta\| \leqslant \eta_{\mathrm{M}}. \tag{41}$$

定理 2　在假设1–3的条件下, 若基于神经网络的方程的解存在, 考虑移动装弹机械臂系统状态空间模型(12)与评价网络权值更新率(37), 若系统轨迹跟踪最优控制律选为

$$\begin{aligned} u = {} & u_{\mathrm{d}} + \hat{u}_{\mathrm{v}} = \\ & g^{+}(x_{\mathrm{d}})(\dot{x}_{\mathrm{d}} - f(x_{\mathrm{d}})) - \\ & \frac{1}{2}R^{-1}g^{\mathrm{T}}(x)(\nabla\sigma_{\tau}(e))^{\mathrm{T}}\hat{w}_{\tau}, \end{aligned} \tag{42}$$

即可保证权值近似误差与系统轨迹跟踪误差均为最终一致有界.

4.2　稳定性分析

证　定义Lyapunov函数为

$$V = \frac{1}{2}e^{\mathrm{T}}e + J^{*}(e) + \frac{1}{2\alpha_{\tau}}\tilde{w}_{\tau}^{\mathrm{T}}\tilde{w}_{\tau}. \tag{43}$$

对时间求导, 得

$$\begin{aligned} \dot{V} = {} & e^{\mathrm{T}}\dot{e} + (\nabla J^{*}(e))^{\mathrm{T}}\dot{e} + \frac{1}{\alpha_{\tau}}\tilde{w}_{\tau}^{\mathrm{T}}\dot{\tilde{w}}_{\tau} = \\ & e^{\mathrm{T}}[f_{\mathrm{e}} + [g(x) - g(x_{\mathrm{d}})]u + g(x_{\mathrm{d}})u_{\mathrm{v}}] + \\ & (\nabla J^{*}(e))^{\mathrm{T}}\dot{e} + \frac{1}{\alpha_{\tau}}\tilde{w}_{\tau}^{\mathrm{T}}\dot{\tilde{w}}_{\tau} = \\ & e^{\mathrm{T}}[f_{\mathrm{e}} + [g(x) - g(x_{\mathrm{d}})]u + g(x_{\mathrm{d}})u_{\mathrm{v}}] - \\ & e^{\mathrm{T}}Qe - u_{\mathrm{v}}^{\mathrm{T}}Ru_{\mathrm{v}} + \frac{1}{\alpha_{\tau}}\tilde{w}_{\tau}^{\mathrm{T}}\dot{\tilde{w}}_{\tau}. \end{aligned} \tag{44}$$

因为$f(x)$是Lipschitz函数, 则一定存在$L_{\mathrm{f}} > 0$, 使得不等式$\|f_{\mathrm{e}}\| \leqslant L_{\mathrm{f}}\|e\|$成立, 根据假设1, 可知$g(x)$及$g(x_{\mathrm{d}})$均有界, 不妨设

$$\|g(x)\| \leqslant w_{\mathrm{g}},\ \|g(x_{\mathrm{d}})\| \leqslant w_{\mathrm{gd}}, \tag{45}$$

进而得

$$\|g(x) - g(x_{\mathrm{d}})\| \leqslant \Delta w_{\mathrm{g}}. \tag{46}$$

利用三角不等式, 有

$$\begin{aligned} \dot{V} \leqslant {} & L_{\mathrm{f}}\|e\|^{2} + \Delta w_{\mathrm{g}}\|u\|\|e\| + w_{\mathrm{gd}}\|u_{\mathrm{v}}\|\|e\| - \\ & e^{\mathrm{T}}Qe - u_{\mathrm{v}}^{\mathrm{T}}Ru_{\mathrm{v}} + \frac{1}{\alpha_{\tau}}\tilde{w}_{\tau}^{\mathrm{T}}\dot{\tilde{w}}_{\tau} \leqslant \\ & L_{\mathrm{f}}\|e\|^{2} + \Delta w_{\mathrm{g}}\|u_{\mathrm{d}} + u_{\mathrm{v}}\|\|e\| - e^{\mathrm{T}}Qe - \\ & u_{\mathrm{v}}^{\mathrm{T}}Ru_{\mathrm{v}} + w_{\mathrm{gd}}\|u_{\mathrm{v}}\|\|e\| + \tilde{w}_{\tau}^{\mathrm{T}}(e_{\mathrm{PH}} - \tilde{w}_{\tau}^{\mathrm{T}}\eta)\eta \leqslant \\ & L_{\mathrm{f}}\|e\|^{2} + \Delta w_{\mathrm{g}}\|u_{\mathrm{d}}\|\|e\| + \frac{1}{2}\Delta {w_{\mathrm{g}}}^{2}\|u_{\mathrm{v}}\|^{2} + \\ & \frac{1}{2}\|e\|^{2} + \frac{1}{2}w_{\mathrm{gd}}^{2}\|e\|^{2} + \frac{1}{2}\|u_{\mathrm{v}}\|^{2} - \|\tilde{w}_{\tau}^{\mathrm{T}}\eta\|^{2} - \\ & \lambda_{\min}(Q)\|e\|^{2} - \lambda_{\min}(R)\|u_{\mathrm{v}}\|^{2} + \tilde{w}_{\tau}^{\mathrm{T}}e_{\mathrm{PH}}\eta \leqslant \\ & -[(\lambda_{\min}(Q) - L_{\mathrm{f}} - \frac{1}{2}w_{\mathrm{gd}}^{2} - \frac{1}{2})\|e\| - \\ & \Delta w_{\mathrm{g}}\|u_{\mathrm{d}}\|]\|e\| - (\lambda_{\min}(R) - \frac{1}{2}\Delta {w_{\mathrm{g}}}^{2} - \\ & \frac{1}{2})\|u_{\mathrm{v}}\|^{2} + \frac{1}{2}e_{\mathrm{PH}}^{2} - \frac{1}{2}\|\tilde{w}_{\tau}^{\mathrm{T}}\eta\|^{2} \leqslant \\ & -[(\lambda_{\min}(Q) - L_{\mathrm{f}} - \frac{1}{2}w_{\mathrm{gd}}^{2} - \frac{1}{2})\|e\| - \\ & \Delta w_{\mathrm{g}}w_{\mathrm{ud}}]\|e\| - (\lambda_{\min}(R) - \frac{1}{2}\Delta w_{\mathrm{g}}^{2} - \\ & \frac{1}{2})\|u_{\mathrm{v}}\|^{2} + \frac{1}{2}e_{\mathrm{PH}}^{2} - \frac{1}{2}\|\tilde{w}_{\tau}^{\mathrm{T}}\eta\|^{2}. \end{aligned} \tag{47}$$

显然, 在集合

$$\varOmega_{1} = \{\tilde{w}_{\tau} : \|\tilde{w}_{\tau}\| \leqslant \frac{e_{\mathrm{PH}}}{\eta_{\mathrm{M}}}\} \tag{48}$$

与集合

$$\varOmega_{2} = \{e : \|e\| \leqslant \frac{\Delta w_{\mathrm{g}}w_{\mathrm{ud}}}{\lambda_{\min}(Q) - L_{\mathrm{f}} - \frac{1}{2}w_{\mathrm{gd}}^{2} - \frac{1}{2}}\} \tag{49}$$

之外, 且需满足

$$\lambda_{\min}(Q) \geqslant L_{\mathrm{f}} + \frac{1}{2}w_{\mathrm{gd}}^{2} + \frac{1}{2}, \tag{50}$$

$$\lambda_{\min}(R) \geqslant \frac{1}{2}\Delta w_{\mathrm{g}}^{2} + \frac{1}{2} \tag{51}$$

条件时, 有

$$\dot{V} \leqslant 0. \tag{52}$$

由此可见, 根据Lyapunov稳定性定理, 神经网络权值近似误差与移动装弹机械臂系统轨迹跟踪误差均最终一致有界.　　证毕.

5　仿真研究

为验证系统建模的准确性及所设计控制方法的有效性, 以军事工程中某型号移动装弹机械臂系统为研究对象, 利用MATLAB仿真平台进行控制效果验证. 机械臂系统参数如表1所示, 关节角期望轨迹如表2所示.

在自适应动态规划控制方法中, 评价网络采用12–10–1结构的神经网络, 其权重初值选为$\widehat{w}_{\tau} = [30\ \ 30\ \ 20\ \ 45\ \ 20\ \ 40\ \ 30\ \ 45\ \ 50\ \ 40]^{\mathrm{T}}$, 激活函数选为S型函数, 其余参数分别为: $Q_{i} = 5I_{4}$, $R_{i} = 2I_{2}$, $\alpha_{i} = 0.001$, 这里I_{4}和I_{2}分别为4阶和2阶的单位对角矩阵.

对比的自适应滑模控制 (adaptive sliding mode control, ASMC)方法为

$$u = \bar{M}\ddot{q}_{r} + \bar{C}\dot{q}_{r} + \bar{G} + \hat{\psi} + \varepsilon\,\mathrm{sgn}\,s, \tag{53}$$

$$\dot{q}_{r} = \dot{q}_{\mathrm{d}} + \lambda(q_{\mathrm{d}} - q), \tag{54}$$

$$\dot{\hat{\psi}} = \varGamma^{\mathrm{T}}s. \tag{55}$$

为解决抖振问题, 用饱和函数$\mathrm{sat}\,s$代替式(53)中符号函数$\mathrm{sgn}\,s$, 边界层Δ选为0.01, 其他控制参数选取

为

$$\varepsilon = \mathrm{diag}\{1200, 1200, 1200, 1200, 1200, 1200\},$$

$$\Gamma = \mathrm{diag}\{120, 160, 95, 180, 100, 100\},$$

$$\lambda = \mathrm{diag}\{25, 20, 10, 80, 15, 20\}.$$

表 1 移动装弹机械臂系统参数

Table 1 Parameters of mobile missile-loading manipulator system

参数	符号	单位	数值
连杆1转动惯量	J_1	$\mathrm{kg\cdot m^2}$	1.74
关节2转动惯量	J_2	$\mathrm{kg\cdot m^2}$	1.12
关节2质量	m_2	kg	30
关节3质量	m_3	kg	25
伸缩杆质量	M_4	kg	100
关节7质量	m_7	kg	15
夹具质量	m_8	kg	10
连杆1长度	l_1	m	1.74
连杆2长度	l_2	m	2.5
连杆3长度	l_3	m	3
连杆7长度	l_7	m	0.8
吊具长度	l_8	m	1.2
伸缩部分长度	D	m	1.5
连杆1面密度	ρ_1	$\mathrm{kg/m^2}$	42.825
连杆2线密度	ρ_2	kg/m	42.825
连杆3线密度	ρ_3	kg/m	46.8
连杆7线密度	ρ_7	kg/m	45.8
夹具线密度	ρ_8	kg/m	46

表 2 关节角期望轨迹

Table 2 Desired trajectories of joints

关节1	$\theta_1(t) = \pi/12 + 0.1t + 0.1t^2 + 0.1t^3$
关节2	$\theta_2(t) = \pi/6 + 0.2t + 0.1t^2 + 0.01t^3$
关节3	$\theta_3(t) = \pi/3 + 0.1t + 0.1t^2 + 0.01t^3$
伸缩杆	$d(t) = 0.3t + 0.1t^2 + 0.1t^3$
关节7	$\theta_7(t) = -0.1t - 0.2t^2 - 0.01t^3$
关节8	$\theta_8(t) = -\pi/8 + 0.1t + 0.1t^2 + 0.2t^3$

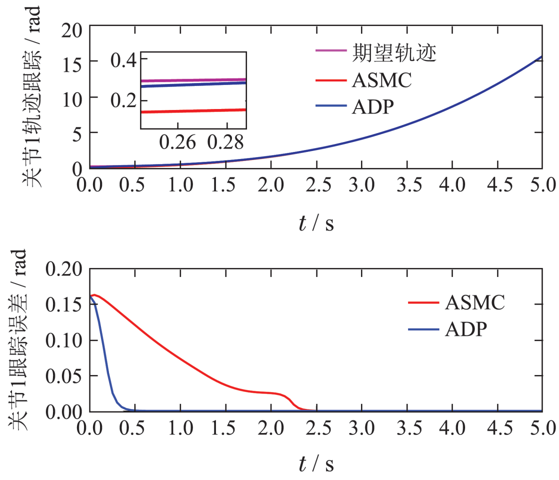
仿真时间为5 s, 仿真结果曲线如图6–8所示, 系统轨迹跟踪平均绝对误差如表3所示.

表 3 关节角轨迹跟踪平均绝对误差

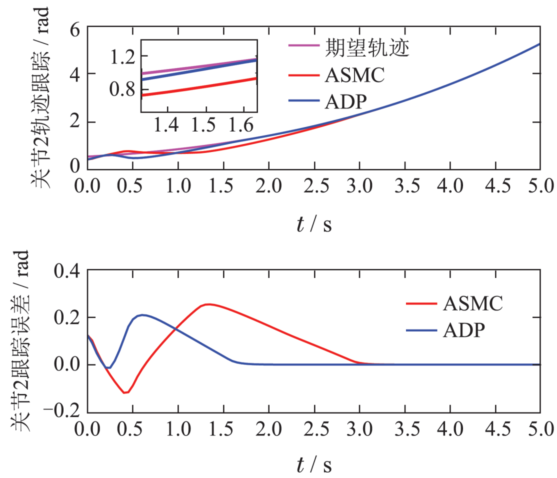
Table 3 Trajectory tracking mean absolute errors of joints

关节角	ADP 平均绝对误差	ASMC 平均绝对误差
θ_1	2.95×10^{-7}	3.33×10^{-7}
θ_2	1.34×10^{-5}	6.52×10^{-5}
θ_3	1.08×10^{-5}	1.40×10^{-5}
d	2.62×10^{-6}	8.14×10^{-6}
θ_7	3.78×10^{-6}	8.55×10^{-6}
θ_8	2.70×10^{-7}	4.45×10^{-7}

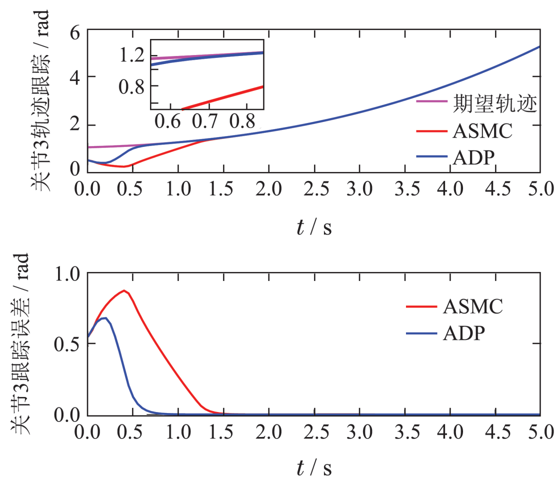
图6中每个窗口上图为轨迹跟踪实时曲线, 下图为轨迹跟踪误差曲线. 在ASMC作用下, 移动装弹机械臂的回转关节θ_1和旋转关节θ_2至少需要3 s才可跟踪上期望轨迹; 而在ADP作用下, 旋转关节θ_2和关节θ_3跟踪上期望轨迹最多只需2 s. 由于关节θ_1和关节θ_8均为回转关节, 关节d实为套筒伸缩部分, 在ADP作用下, 它们的跟踪时间与旋转关节相比更短, 仅0.5 s. 此外, 从表3中两种控制方法作用下轨迹跟踪平均绝对误差数据分析, ADP的控制精度也是高于ASMC.



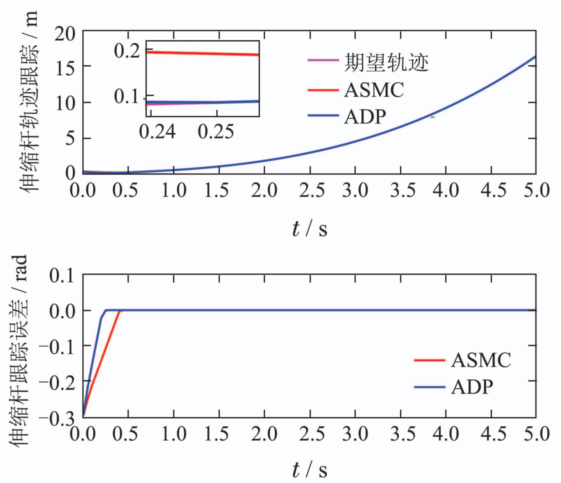
(a) 回转关节θ_1



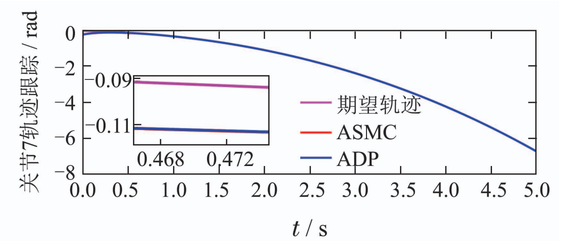
(b) 旋转关节θ_2



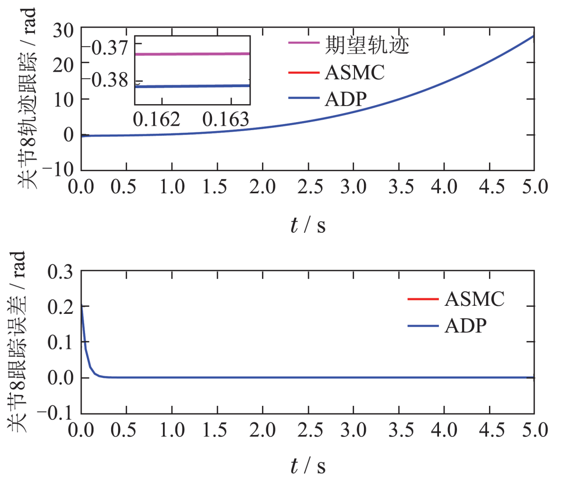
(c) 旋转关节θ_3



(d) 伸缩关节d



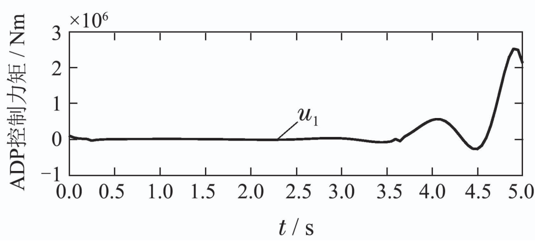
(e) 旋转关节θ_7

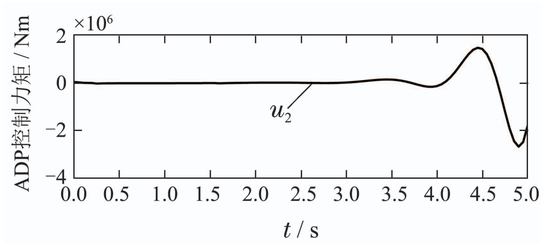


(f) 回转关节θ_8

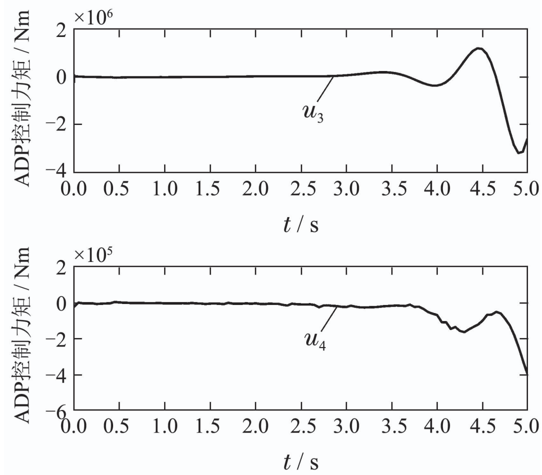
图 6 关节角轨迹跟踪曲线与跟踪误差

Fig. 6 Trajectory tracking curves and tracking errors of joints

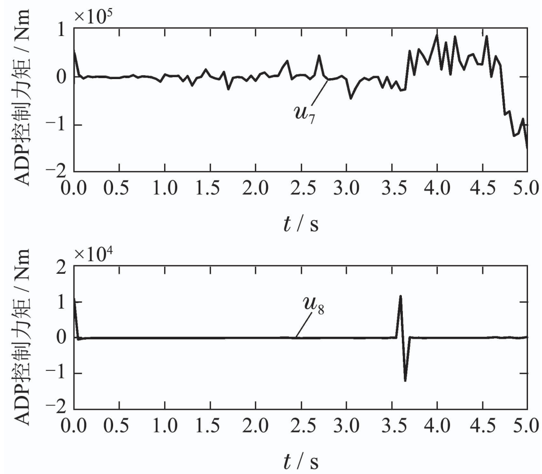




(a) 控制力矩u_1, u_2



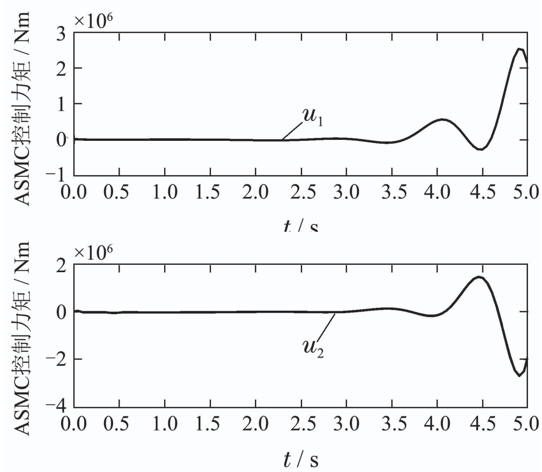
(b) 控制力矩u_3, u_4



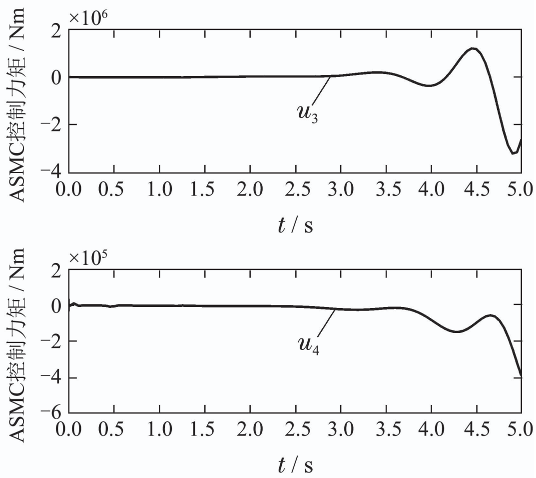
(c) 控制力矩u_7, u_8

图 7 ADP控制力矩

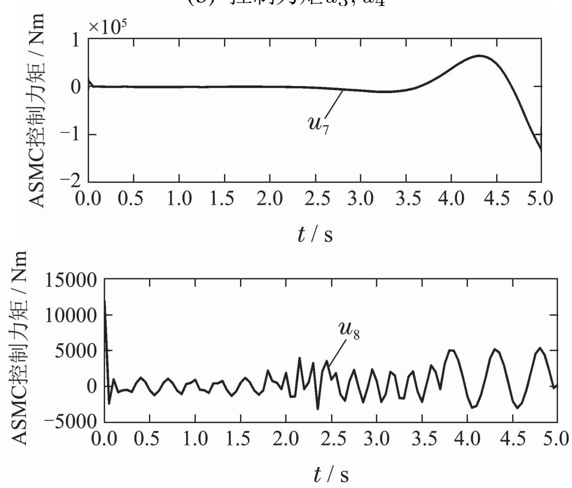
Fig. 7 Control torques with ADP



(a) 控制力矩u_1, u_2



(b) 控制力矩u_3, u_4



(c) 控制力矩u_7, u_8

图 8 ASMC控制力矩

Fig. 8 Control torques with ASMC

图7和图8分别为移动装弹机械臂在ADP与ASMC两种控制作用下控制力矩输出曲线, 其中不同控制作用下同一关节的输出力矩用同一颜色表征. 在ASMC控制过程中, 回转关节θ_8的控制力矩u_8因饱和函数代替了符号函数, 虽仍有波动趋势, 但已无抖振, 其他各关节控制力矩与ADP相类似, 均比较平滑.

综合比较ADP控制作用下的各关节输出力矩, 最大值为4×10^6 Nm, 即4 MNm, 有多种型号的液压缸可满足该输出力矩需求, 以合丰大吨位液压油缸为例[23], 部分型号如表4所示.

表 4 合丰大吨位液压油缸参数

Table 4 Parameters of Hefeng large tonnage hydraulic cylinder

型号	吨位/T	外径/mm	内径/mm	杆直径/mm	压力/MPa
QF400	400	402	280	200	63
QF500	500	420	320	250	63
QF630	630	505	360	280	63
QF800	800	550	400	320	63
QF1000	1000	625	450	360	63

液压驱动系统中液压缸输出力矩u与压强P、内径ϕ_D的近似关系为

$$u = P \times \frac{\pi \phi_D^2}{4} \times L \times \beta, \tag{56}$$

其中: 按经验值选取负荷率β为0.8; 连杆长度L取表1中最短连杆长度0.8 m; 选择QF630型号液压油缸, 做简单估算有

$$u \approx 4.1\ \mathrm{MNm}. \tag{57}$$

综上所述, 当系统存在未知上界不确定性和外部干扰时, 相对于ASMC, ADP的控制性能更好. 由于ADP的控制过程包含神经网络学习过程, 在控制初期部分关节的实际轨迹并没有跟踪上期望轨迹. 在ASMC和ADP作用下, 当实际轨迹均跟踪上期望轨迹后, ADP作用下的控制精度更高. 此外, 本文设计的控制方法在工程上有液压油缸可保障其实现.

6 结论

本文应用拉格朗日方程建立了移动装弹机械臂系统的动力学模型. 当系统存在外界干扰和参数不确定性时, 通过定义总体不确定项, 设计了自适应动态规划轨迹控制器. 仿真结果表明, 文中所设计的基于单网络评价结构的控制器可以很好地跟踪上系统的期望轨迹. 与自适应滑模控制器相比, 自适应动态规划控制在保证系统稳定性的前提下, 缩短了系统的响应时间, 减小了跟踪误差, 使系统控制力矩更加平滑, 提高了系统的控制精度, 达到了更好的调节效果.

参考文献:

[1] ZHANG Dongqing, JIANG Qi. Reviews and outlooks of global missiles evolution in 5 years version. *Tactical Missile Technology*, 2016, 1: 1 – 8.
(张冬青, 蒋琪. 世界导弹武器装备与技术5年发展回顾和展望. 飞航导弹, 2016, 1: 1 – 8.)

[2] WANG Guohui, MA Xiaofei, LI Xiangrong, et al. Research on combined fault diagnosis technology in fault diagnosis of autoloader. *Journal of Gun Launch & Control*, 2020, 159(3): 83 – 88.
(王国辉, 马啸飞, 李向荣, 等. 组合故障诊断技术在自动装弹机故障诊断中的研究. 火炮发射与控制学报, 2020, 159(3): 83 – 88.)

[3] MIAO Donghui, MA Zhangjian, LI Jiasheng, et al. Simulation analysis for hoisting stability of a missile trans-loading vehicle. *Air & Sport Defense*, 2018, 1(4): 58 – 64.
(缪东辉, 马张健, 李佳圣, 等. 导弹运输装填发射车吊装稳定性仿真研究. 空天防御, 2018, 1(4): 58 – 64.)

[4] WANG Liling, DONG Liyuan, MA Dong, et al. Active disturbance rejection tracking control of wheeled mobile robots under sliding and slipping conditions. *Control Theory & Applications*, 2020, 37(2): 431 – 438.
(王立玲, 董力元, 马东, 等. 滑动与打滑条件下的轮式移动机器人自抗扰跟踪控制. 控制理论与应用, 2020, 37(2): 431 – 438.)

[5] GERASIMOS R, KRISHNA B. Flexible-link robots. *Robotic Manipulators and Vehicles*, 2018, 152: 271 – 300.

[6] HEIDARI H, KORAYEM M H, HAGHPANAHI M. Optimal trajectory planning for increased stability of mobile flexible manipulators undergoing large deflection. *Proceedings of the Institution of*

Mechanical Engineers, Part B: Journal of Engineering Manufacture, 2017, 231(1): 85 – 95.

[7] DAVID G W, JOHN E, KEVIN K. Quasi-real-time confined environment path generation for mobile robotic manipulator arms. *Proceedings of The Institution of Mechanical Engineers Part I-Journal of Systems and Control Engineering*, 2018, 232(3): 270 – 284.

[8] LI Q H, MU Y Q, YOU Y, et al. A hierarchical motion planning for mobile manipulator. *IEEJ Transactions on Electrical and Electronic Engineering*, 2020, 15(9): 1390 – 1399.

[9] ZHANG Zhuo, ZHANG Cheng. Adaptive robust control of trajectory tracking at the end of robotic arm. *Robot Technology and Application*, 2018, 4: 31 – 35.
(张卓, 张程. 机械臂末端轨迹跟踪的自适应鲁棒控制. 机器人技术与应用, 2018, 4: 31 – 35.)

[10] LI Yuanchun, WANG Meng, SHENG Lihui, et al. Adaptive second order sliding mode control for hydraulic manipulator based on backstepping. *Journal of Jilin University (Engineering and Technology Edition)*, 2014, 45(1): 193 – 201.
(李元春, 王蒙, 盛立辉, 等. 液压机械臂基于反演的自适应二阶滑模控制. 吉林大学学报(工学版), 2014, 45(1): 193 – 201.)

[11] WU Y X, FENG Y, HU Y M. Output tracking control of mobile manipulators based on dynamical sliding-mode control. *Frontiers of Mechanical Engineering*, 2007, 2(1): 110 – 115.

[12] GAO Q J, LIU K, LIU L L. Study on control for mobile manipulators based on fuzzy sliding mode control. *Journal of System Simulation*, 2008, 20(16): 4323 – 4325.

[13] SZUSTER M, GIERLAK P. Approximate dynamic programming in tracking control of a robotic manipulator. *International Journal of Advanced Robotic Systems*, 2016, 16(13): 47 – 49.

[14] WANG C C. Design of a variable structure controller for mobile manipulators. *Second International Conference on Innovative Computing, Information and Control*. Kumamoto, Japan: IEEE, 2007, 252: 333 – 336.

[15] CHENG M B, TSAI C C. Hybrid robust tracking control for a mobile manipulator via sliding-mode neural network. *2005 IEEE International Conference on Mechatronics*. Taibei, China: IEEE, 2005: 437 – 542.

[16] HONG Chengwen, FU Yue. Nonlinear robust approximate optimal tracking control based on adaptive dynamic programming. *Control Theory & Applications*, 2016, 37(2): 225 – 232.
(洪成文, 富月. 基于自适应动态规划的非线性鲁棒近似最优跟踪控制. 控制理论与应用, 2018, 35(9): 1285 – 1292.)

[17] SUN Lei, FU Bin, WAN Shizheng, et al. Differential game guidance law based on adaptive dynamic programming for the interception of hypersonic target. *Fire Control & Command Control*, 2020, 48(6): 42 – 48, 72.
(孙磊, 付斌, 万世正, 等. 基于自适应动态规划的反高超武器微分对策制导律. 航空工程进展, 2020, 48(6): 42 – 48, 72.)

[18] MARCIN S. Dual-heuristic dynamic programming in the three-wheeled mobile transport robot control. *Artificial Intelligence and Soft Computing*, 2018, 10842: 763 – 776.

[19] SUN Jingliang, LIU Chunsheng. An overview on the adaptive dynamic programming based missile guidance law. *Acta Automatica Sinica*, 2017, 43(7): 1101 – 1113.
(孙景亮, 刘春生. 基于自适应动态规划的导弹制导律研究综述. 自动化学报, 2017, 43(7): 1101 – 1113.)

[20] GE Y Y, LIU X K, LI Y. Pareto optimal control of the mean-field stochastic systems by adaptive dynamic programming algorithm. *ISA Transactions*, 2020, 102: 81 – 90.

[21] LIU D R, WANG D, LI H L. Decentralized stabilization for a class of continuous-time nonlinear interconnected systems using online learning optimal control approach. *IEEE Transactions on Neural Networks and Learning Systems*, 2014, 25(2): 411 – 428.

[22] ZHAO B, LIU D R, LI Y C. Online fault compensation control based on policy iteration algorithm for a class of affine nonlinear systems with actuator failures. *IET Control Theory & Applications*, 2016, 10(15): 1816 – 1823.

[23] 百度爱米购. 德州合丰液压机械有限公司600吨液压油缸. https://b2b.baidu.com/land?id=11723655338946e6a6c34dad16e335 a110.

作者简介:

唐志国　博士, 讲师, 目前研究方向为跨尺度控制系统、智能机械与机器人, E-mail: tangzhiguo@jlu.edu.cn;

张富尧　硕士研究生, 目前研究方向为数控装备关键功能部件可靠性, E-mail: 127424018@qq.com;

马　彦　博士, 教授, 目前研究方向为控制系统鲁棒滤波、状态估计方法、车载动力系统的建模和应用, E-mail: mayan_maria@163.com.