含弹簧阻尼缓冲机构空间机器人捕获卫星操作的避撞柔顺强化学习控制

朱 安, 陈 力†

(福州大学 机械工程及自动化学院, 福建 福州 350116)

摘要: 针对空间机器人捕获卫星操作过程中, 关节因受冲击载荷而易造成冲击破坏的问题, 在关节电机与机械臂之间设计了一种弹簧阻尼缓冲机构. 缓冲机构不仅能利用弹簧实现捕获操作过程的柔顺化, 利用阻尼器实现碰撞能量的吸收及柔性振动的抑制; 还能通过合理设计与之配合的避撞柔顺策略使关节所受冲击力矩限定在安全范围内. 首先, 分别利用含耗散力Lagrange方程法与Newton-Euler法导出了碰撞前的空间机器人与被捕获卫星的分体系统动力学方程; 然后, 结合Newton第三定律、捕获点的速度约束、各分体的位置约束获得了捕获后的混合体系统动力学方程, 且基于动量守恒关系计算了碰撞冲击效应与冲击力; 最后, 提出了一种结合缓冲机构的避撞柔顺强化学习控制方案, 该方案通过实时与动态环境试错交互得到惩罚信号, 并利用惩罚信号对控制器进行优化, 实现对失稳混合体系统的镇定控制. 利用Lyapunov定理证明了系统的稳定性; 数值仿真验证了缓冲结构的抗冲击性能及所提策略的有效性.

关键词: 空间机器人; 弹簧阻尼缓冲机构; 避撞柔顺策略; 冲击效应; 强化学习

引用格式: 朱安, 陈力. 含弹簧阻尼缓冲机构空间机器人捕获卫星操作的避撞柔顺强化学习控制. 控制理论与应用, 2020, 37(8): 1727 – 1736

DOI:10.7641/CTA.2020.90839

Collision avoidance and compliance reinforcement learning control for space robot with spring-damper buffer device capturing satellite

ZHU An, CHEN Li†

(School of Mechanical Engineering and Automation, Fuzhou University, Fuzhou Fujian 350116, China)

Abstract: In order to protect joints from impact damage during the process of space robot capturing satellite, a spring-damper buffer device is designed between joint motor and manipulator. The device can not only use spring to achieve the compliance during the capture operation, use damper to absorb impact energy and suppress flexible vibration; but also limit the joint's impact torque to a safe range through reasonable and coordinated design the collision avoidance and compliance strategy. First of all, the dynamic models of space robot and satellite at collision time are derived by using Lagrange function based on dissipation theory and Newton-Euler function respectively. After that, combined with Newton's third law, velocity and position constraints of capture points constraints of bodies, the dynamic model of hybrid system after capture is obtained, the impact effect and impact force are calculated based on momentum conservation. Finally, a collision avoidance and compliance reinforcement learning control strategy with buffer device is proposed. The penalty signal is obtained by trial-and-error interaction with dynamic environment, be used to optimize the controller to stabilize the instability hybrid system. The stability of the system is proved by Lyapunov theorem, and the impact resistance of the device and the effectiveness of the proposed strategy are verified by numerical simulation.

Key words: space robot; spring-damper buffer device; collision avoidance and compliance strategy; impact effect; reinforcement learning

Citation: ZHU An, CHEN Li. Collision avoidance and compliance reinforcement learning control for space robot with spring-damper buffer device capturing satellite. *Control Theory & Applications*, 2020, 37(8): 1727 – 1736

收稿日期: 2019−10−10; 录用日期: 2020−04−23.

†通信作者. E-mail: chnle@fzu.edu.cn; Tel: +86 15345970061.

本文责任编委: 贾英民.

国家自然科学基金项目(11372073), 福建省工业机器人基础部件技术重大研发平台项目(2014H21010011), 福建省机器人基础部件与系统集成创新中心专项资金项目资助.

Support by the National Nature Science Foundation of China (11372073), the Fujian Provincial Industrial Robot Basic Components' Technology Research Platform (2014H21010011) and the Special Fund of Fujian Robot Basic Components and System Integration Innovation Center.

1 引言

随着人类在空间探索、卫星通信、气象观测、资源勘探、导航定位等领域研究的深入, 全球每年向太空发射的卫星数量正在不断地攀升. 这些卫星中难免会有一小部分未能到达预定轨道, 或者在轨道运行过程中发生失效, 若能对其进行回收与维修, 它们仍能继续使用. 另外, 一般情况下卫星达到使用寿命是因为其携带的燃料耗尽而不是出现故障, 如果可以进行燃料的补给, 这些卫星仍能继续工作. 而实现上述卫星回收与燃料补给任务的关键就是对卫星进行捕获操作. 目前, 使用空间机器人对卫星进行捕获操作具有广阔的应用前景. 已经受到了国内外学者的广泛关注[1–5].

一般情况下, 空间机器人在轨捕获卫星的过程可分为4个阶段: 1) 对被捕获卫星进行观测; 2) 空间机器人逐渐靠近被捕获卫星, 进行捕获阶段的准备; 3) 空间机器人末端抓手与被捕获卫星的捕获点接触、碰撞; 4) 捕获完成后, 对空间机器人与被捕获卫星形成的混合体系统进行镇定控制. 针对捕获第3阶段的动力学分析问题, 众多学者进行了深入的研究[6–9]. 程靖等[10]对空间机器人捕获卫星过程的动力学演化模拟进行了分析; Uyam等[11]对空间机器人与自由漂浮卫星的接触效应进行了实验的评估; Dimitrov等[12]详细建立了空间机器人捕获卫星过程的接触动力学模型, 并分析了捕获过程的动量交换问题; Yoshida等[13]基于动量守恒定律研究了空间机器人获卫星的碰撞动力学及运动学问题; 董楸煌等[14]利用假设模态法近似描述柔性杆的弹性变形, 然后利用动量冲量法分析了空间机械臂捕获卫星的碰撞动力学. 值得注意的是, 上述学者的主要关注点均在碰撞动力学上, 未考虑空间机器人关节的脆弱性, 从而忽略了对其的保护. 事实上, 若捕获操作中未对关节进行保护, 关节就有可能受到冲击破坏, 从而使捕获操作失败, 甚至造成空间机器人的损坏.

地面机器人中, 为了使机器人与外界环境发生碰撞时保护其关节不受冲击破坏, 在机器人关节电机与机械臂之间加入弹簧缓冲机构(spring buffer device, SBD)实现对冲击载荷的缓冲是一种行之有效的方法[15–19]. 但对于空间机器人, SDB的加入会增加关节的柔性, 由于空间机器人系统为无根树系统, 且其处于微重力、高真空的环境中, 关节柔性的存在会使机械臂在镇定控制过程中产生振动, 造成控制精度的下降, 严重时可能使系统失稳. 因此, 本文针对空间机器人, 尝试设计了一种弹簧阻尼缓冲机构(spring-damper buffer device, SDBD). 相较于SBD, SDBD 不仅能够实现对冲击载荷的快速缓冲卸载, 而且还能使柔性振动快速衰减, 实现对柔性振动的抑制.

针对捕获第4阶段的失稳混合体系统镇定控制问题, Liu等[20]采用阻抗控制对双臂空间机器人捕获卫星后的混合体系统进行了镇定控制方案; Huang等[21]针对空间机器人捕获卫星后质量特性与反作用轮结构发生变化的问题, 提出了一种改进的状态依赖Riccati方程最优控制器; Luo等[22]考虑了混合体系统的不可测状态、未知惯性特性和外部干扰, 提出了一种基于有限时间收敛的鲁棒无惯性预定性能控制方案. 然而, 上述控制方案均未考虑冲击效应, 冲击效应较大时空间机器人将处于严重失稳的状态, 此时一般的控制策略难以对其进行镇定控制. 强化学习能够通过智能系统从环境到行为映射的学习, 以使强化信号值达到最大. 强化学习不同于其他形式的机器学习方式, 其通过环境提供的强化信号来对机器产生动作的好坏作出一种评价(通常称为惩罚信号), 而不是告诉强化学习系统如何去产生正确的动作, 因此其能表现出极强的环境适应能力. Hu等[23]针对系统参数与死区未知的机械臂, 利用状态网络对机械臂的状态进行预测, 利用评价网络评估操作策略的性能指标, 提出了一种神经网络强化学习模型. Li等[24]针对末端执行器运动受限的机器人系统, 提出了一种自适应阻抗控制策略, 并利用积分强化学习方法解决了线性二次调节模型的运算问题. Li等[25]针对参数未知、具有共同期望轨迹的多机器人协调控制问题, 提出了一种自适应强化学习策略. 考虑到强化学习优越的环境适应能力, 本文针对严重失稳的混合体系统镇定控制问题, 提出一种强化学习自适应模糊控制方案. 考虑到混合体系统失稳严重, 若直接利用强化信号设计控制力矩, 很容易导致强化学习系统因信息源单一而使镇定控制失败; 因此, 本文采用一个代理器从固定增益控制器、自适应惩罚单元、模糊动作发生单元中收集信号, 并通过该信号设计控制力矩, 从而增强了学习系统的稳定性.

本文设计了一种SDBD以实现在空间机器人捕获卫星操作过程中保护其关节免受冲击破坏. 用含耗散力Lagrange方程法与Newton-Euler法导出了分体系统动力学方程; 利用Newton第3定律、捕获点的速度约束、各分体的位置约束计算了碰撞冲击效应、冲击力, 并结合动量守恒关系导出了混合体系统动力学方程; 提出了一种避撞柔顺强化学习控制方案实现对失稳混合体系统的镇定控制; 通过对捕获操作过程的仿真分析, 验证了所提策略的有效性.

2 SDBD模型结构及避撞策略

2.1 SDBD模型结构

SDBD模型结构示意图如图1所示, 其主要由旋转阻尼器与扭转弹簧组成. 为了更加真实的描述电机与机械臂处的阻力, 分别在其上添加了等效阻尼器. 图中k_{si}, $D_{ti}(i=1,2)$分别为扭转弹簧的刚度、旋转阻尼器的阻尼系数; D_{mi}, $D_{Li}(i=1,2)$分别为电机、机

械臂端等效阻尼器的阻尼系数.

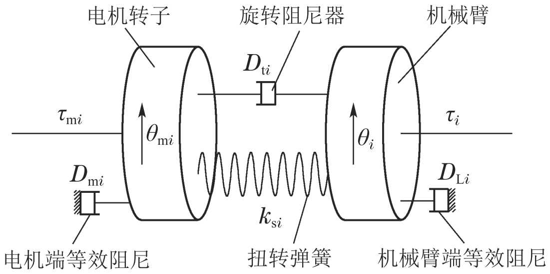

图 1 SDBD模型结构示意图

Fig. 1 Structural model of SDBD

2.2 避撞策略描述

在捕获的第3阶段, 空间机器人机械臂末端与被捕获卫星发生碰撞, 此时机械臂端将受到很大的冲击力矩. 在传导至电机转子的过程中, 该力矩会被弹簧和阻尼器快速卸载, 从而实现对关节的保护. 在捕获的第4阶段, 由于冲击效应的存在, 电机开启时会产生瞬时冲击力矩, 若该力矩超过关节所能承受的极限而未关停电机, 关节很可能会发生损坏. 因此, 需要根据关节所能承受的极限力矩值来设置一个关机力矩阈值, 当检测到冲击力矩超过所设阈值后电机关停, 此时SDBD中的弹簧将会提供弹力来减小关节所受冲击力矩, 阻尼器将会快速耗能抑制柔性振动. 然而, 在实际操作中, 如果只设定一个关机力矩阈值, 将导致电机频繁的开关机, 很容易造成电机的损坏. 由此, 本文所提的避撞策略同时设置了开、关机阈值. 当检测到关节所受冲击力矩超过关机力矩阈值时电机关停; 当SDBD将冲击力矩降低到开机阈值时电机再次开启.

3 空间机器人捕获卫星操作分析

3.1 分体系统动力学建模

配置SDBD的空间机器人系统与卫星系统如图2所示.

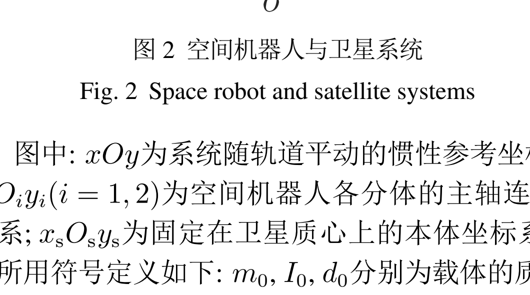

图 2 空间机器人与卫星系统

Fig. 2 Space robot and satellite systems

图中: xOy为系统随轨道平动的惯性参考坐标系; $x_iO_iy_i(i=1,2)$为空间机器人各分体的主轴连体坐标系; $x_{\mathrm{s}}O_{\mathrm{s}}y_{\mathrm{s}}$为固定在卫星质心上的本体坐标系. 文中所用符号定义如下: m_0, I_0, d_0分别为载体的质量、转动惯量、质心到第一个关节铰中心的距离; $m_{\mathrm{s}}, I_{\mathrm{s}}, d_{\mathrm{s}}$分别为卫星的质量、转动惯量、质心到末端把手的距离; $m_i, I_i, L_i(i=1,2)$分别为第i个机械臂的质量、转动惯量、长度; $I_{\mathrm{m}i}(i=1,2)$为第i个电机转子的转动惯量; $d_i(i=1,2)$为第i个关节铰中心到机械臂i质心的距离; $\theta_0, \theta_i, \theta_{\mathrm{s}}, \theta_{\mathrm{m}i}(i=1,2)$分别为载体姿态角、机械臂转角、卫星姿态角和电机转子转角.

由图2可导出载体质心O_0、机械臂i质心$O_{\mathrm{c}i}(i=1,2)$相对原点O的矢径为

$$\begin{cases}\boldsymbol{r}_0=(x_0\ \ y_0)^{\mathrm{T}},\\ \boldsymbol{r}_i=\boldsymbol{r}_0+\sum\limits_{j=0}^{i-1}L_j\boldsymbol{a}_j+d_i\boldsymbol{a}_i,\ i=1,2,\end{cases}\tag{1}$$

式中: x_0, y_0为载体质心坐标; $\boldsymbol{a}_j(j=1,2,3)$为$\boldsymbol{x}_j$轴的基矢量. 对式(1)求导可得空间机器人各分体质心的运动速度矢量, 基于此可得其动能表达式为

$$T_{\mathrm{r}}=\sum_{i=0}^{2}(\frac{1}{2}m_i\dot{\boldsymbol{r}}_i^2+\frac{1}{2}I_i\omega_i^2)+\sum_{j=1}^{2}\frac{1}{2}I_{\mathrm{m}j}\boldsymbol{\omega}_{\mathrm{m}j}^2,\tag{2}$$

式中: $\omega_i=\sum\limits_{j=0}^{i}\dot{\theta}_j,\ i=1,2,3, \omega_{\mathrm{m}j}=\dot{\theta}_{\mathrm{m}j},\ j=1,2.$ 若忽略太空的微重力影响, 空间机器人系统的势能来源于缓冲机构中的弹簧, 故其势能表达式为

$$U_{\mathrm{r}}=\frac{1}{2}\sum_{i=1}^{2}\left[k_{\mathrm{s}i}(\theta_{\mathrm{m}i}-\theta_i)^2\right].\tag{3}$$

由于SDBD的引入, 空间机器人系统存在非有势力, 因此需补充耗散函数:

$$\vartheta_{\mathrm{r}}=\sum_{i=1}^{2}(D_{\mathrm{m}i}\dot{\theta}_{\mathrm{m}i}^2+D_{\mathrm{t}i}(\dot{\theta}_{\mathrm{m}i}-\dot{\theta}_i)^2+D_{\mathrm{L}i}\dot{\theta}_i^2).\tag{4}$$

含耗散力的Lagrange方程为

$$\frac{\mathrm{d}}{\mathrm{d}t}(\frac{\partial L_{\mathrm{r}}}{\partial\dot{q}_{\mathrm{r}}})-\frac{\partial L_{\mathrm{r}}}{\partial q_{\mathrm{r}}}+\frac{\partial\vartheta_{\mathrm{r}}}{\partial\dot{q}_{\mathrm{r}}}=\boldsymbol{Q},\tag{5}$$

式中: $L_{\mathrm{r}}=T_{\mathrm{r}}-U_{\mathrm{r}}$为Lagrange函数, $\boldsymbol{Q}\in\mathbb{R}^{5\times1}$为系统广义力. 结合式(2)–(5)可得碰撞前的空间机器人系统动力学方程为

$$\begin{cases}\boldsymbol{M}_{\mathrm{r}}(\boldsymbol{q}_{\mathrm{r}})\ddot{\boldsymbol{q}}_{\mathrm{r}}+(\boldsymbol{H}_{\mathrm{r}}(\boldsymbol{q}_{\mathrm{r}},\dot{\boldsymbol{q}}_{\mathrm{r}})+\boldsymbol{D}_{\mathrm{L}})\dot{\boldsymbol{q}}_{\mathrm{r}}=\boldsymbol{\tau}_{\mathrm{r}}+\boldsymbol{J}_{\mathrm{r}}^{\mathrm{T}}\boldsymbol{F}_{\mathrm{p}},\\ \boldsymbol{I}_{\mathrm{m}}\ddot{\boldsymbol{q}}_{\mathrm{m}}+\boldsymbol{D}_{m\theta}\dot{\boldsymbol{q}}_{\mathrm{m}}+\boldsymbol{\tau}_{\theta}=\boldsymbol{\tau}_{\mathrm{m}},\\ \boldsymbol{K}_{\mathrm{s}}(\boldsymbol{q}_{\mathrm{m}}-\boldsymbol{q}_{\theta})+\boldsymbol{D}_{\mathrm{t}\theta}(\dot{\boldsymbol{q}}_{\mathrm{m}}-\dot{\boldsymbol{q}}_{\theta})=\boldsymbol{\tau}_{\theta},\end{cases}\tag{6}$$

式中: $\boldsymbol{q}_{\mathrm{r}}=[x_0\ \ y_0\ \ \theta_0\ \ \boldsymbol{q}_{\theta}^{\mathrm{T}}]^{\mathrm{T}}$为空间机器人系统的广义坐标, $\boldsymbol{q}_{\theta}=[\theta_1\ \ \theta_2]^{\mathrm{T}}, \boldsymbol{q}_{\mathrm{m}}=[\theta_{\mathrm{m}1}\ \ \theta_{\mathrm{m}2}]^{\mathrm{T}}$; $\boldsymbol{M}_{\mathrm{r}}(\boldsymbol{q}_{\mathrm{r}})\in\mathbb{R}^{5\times5}$为空间机器人系统对称、正定的惯量矩阵, $\boldsymbol{H}_{\mathrm{r}}(\boldsymbol{q}_{\mathrm{r}},\dot{\boldsymbol{q}}_{\mathrm{r}})$为包含科氏力、离心力的列向量; $\boldsymbol{D}_{\mathrm{L}}\in\mathbb{R}^{5\times5}$为增广的机械臂等效阻尼系数矩阵, $\boldsymbol{D}_{\mathrm{m}\theta}\in\mathbb{R}^{2\times2}$为电机等效阻尼系数矩阵, $\boldsymbol{D}_{\mathrm{t}\theta}\in\mathbb{R}^{2\times2}$为缓冲机构中阻尼器的阻尼系数矩阵; $\boldsymbol{\tau}_{\mathrm{r}}=[\boldsymbol{\tau}_{\mathrm{B}}^{\mathrm{T}}\ \ \tau_0\ \ \boldsymbol{\tau}_{\theta}^{\mathrm{T}}]^{\mathrm{T}}$, $\boldsymbol{\tau}_{\mathrm{B}}=[0\ \ 0]^{\mathrm{T}}$表示载体位置不受控, τ_0为载体控制力矩, $\boldsymbol{\tau}_{\theta}=[\tau_1\ \ \tau_2]^{\mathrm{T}}$为关节输入力矩列向量, $\boldsymbol{\tau}_{\mathrm{m}}=[\tau_{\mathrm{m}1}\ \ \tau_2]^{\mathrm{T}}$为电机输出力

矩列向量; $\boldsymbol{I}_{\rm m}\in\mathbb{R}^{2\times2}$为电机转子转动惯量矩阵; $\boldsymbol{K}_{\rm s}\in\mathbb{R}^{2\times2}$为缓冲机构中弹簧的刚度矩阵; $\boldsymbol{J}_{\rm r}\in\mathbb{R}^{3\times5}$为空间机器人末端抓手捕获点的运动雅克比矩阵; $\boldsymbol{F}_{\rm p}\in\mathbb{R}^{3\times1}$为捕获点所受作用力.

采用Newton-Euler法可获得碰撞前卫星的动力学方程为

$$\boldsymbol{M}_{\rm s}\ddot{\boldsymbol{q}}_{\rm s}=\boldsymbol{J}_{\rm s}^{\rm T}\boldsymbol{F}_{\rm p'},\tag{7}$$

式中: $\boldsymbol{M}_{\rm s}\in\mathbb{R}^{3\times3}$为卫星系统对称、正定的惯量矩阵; $\boldsymbol{q}_{\rm s}=[x_{\rm s}\ \ y_{\rm s}\ \ \theta_{\rm s}]^{\rm T}$为卫星系统广义坐标, $x_{\rm s}, y_{\rm s}$为卫星质心坐标; $\boldsymbol{J}_{\rm s}\in\mathbb{R}^{3\times3}$为卫星把手被捕获点的运动雅克比矩阵; $\boldsymbol{F}_{\rm p'}\in\mathbb{R}^{3\times1}$为被捕获点所受作用力, 由牛顿第三定律可知$\boldsymbol{F}+\boldsymbol{F}_{\rm p'}=\boldsymbol{0}_{3\times3}$.

3.2 碰撞冲击效应与碰撞力计算

空间机器人捕获卫星操作的过程中未受到外力, 因此整个系统满足动量守恒, 假设碰撞时间为Δt, 对式(6)第1式、式(7)在碰撞时间内积分得

$$\int_{t_0}^{t_0+\Delta t}[\boldsymbol{M}_{\rm r}\ddot{\boldsymbol{q}}_{\rm r}+(\boldsymbol{H}_{\rm r}+\boldsymbol{D}_{\rm L})\dot{\boldsymbol{q}}_{\rm r}]{\rm d}t=\int_{t_0}^{t_0+\Delta t}(\boldsymbol{\tau}_{\rm r}+\boldsymbol{J}_{\rm r}^{\rm T}\boldsymbol{F}_{\rm p}){\rm d}t,\tag{8}$$

$$\int_{t_0}^{t_0+\Delta t}\boldsymbol{M}_{\rm s}\ddot{\boldsymbol{q}}_{\rm s}{\rm d}t=\int_{t_0}^{t_0+\Delta t}\boldsymbol{J}_{\rm s}^{\rm T}\boldsymbol{F}_{\rm p'}{\rm d}t,\tag{9}$$

式中t_0为碰撞时刻. 由于碰撞时间Δt很短, 在这一时段可以认为系统的广义坐标未发生突变, 仅有广义速度和广义加速度发生突变. 为了保护关节电机, 在捕获碰撞阶段电机处于关机状态, 故式(8)–(9)可近似写为

$$\boldsymbol{M}_{\rm r}[\dot{\boldsymbol{q}}_{\rm r}(t_0+\Delta t)-\dot{\boldsymbol{q}}_{\rm r}(t_0)]=\boldsymbol{J}_{\rm r}^{\rm T}\boldsymbol{f}_{\rm p},\tag{10}$$

$$\boldsymbol{M}_{\rm s}[\dot{\boldsymbol{q}}_{\rm s}(t_0+\Delta t)-\dot{\boldsymbol{q}}_{\rm s}(t_0)]=\boldsymbol{J}_{\rm s}^{\rm T}\boldsymbol{f}_{\rm p'},\tag{11}$$

式中:

$$\boldsymbol{f}_{\rm p}=\int_{t_0}^{t_0+\Delta t}\boldsymbol{F}_{\rm p}{\rm d}t,\ \boldsymbol{f}_{\rm p'}=\int_{t_0}^{t_0+\Delta t}\boldsymbol{F}_{\rm p'}{\rm d}t$$

为碰撞冲量, 且有$\boldsymbol{f}+\boldsymbol{f}_{\rm p'}=\boldsymbol{0}_{3\times3}$.

碰撞结束后空间机器人末端抓手捕获点P、卫星把手被捕获点$p\prime$相对原点O的矢径为

$$\begin{cases}\boldsymbol{r}_{\rm p}=\boldsymbol{r}_0+L_0\boldsymbol{a}_0+L_1\boldsymbol{a}_1+L_2\boldsymbol{a}_2,\\ \boldsymbol{r}_{\rm p'}=\boldsymbol{r}_{\rm s}+d_{\rm s}\boldsymbol{a}_{\rm s},\end{cases}\tag{12}$$

式中: $\boldsymbol{r}_{\rm s}=[x_{\rm s}\ \ y_{\rm s}]^{\rm T}$为卫星质心相对原点$O$的矢径, $\boldsymbol{a}_{\rm s}$为$\boldsymbol{x}_{\rm s}$轴的基矢量.

对式(12)求导并进行增广, 可得碰撞后捕获点与被捕获点的运动速度矢量为

$$\begin{cases}\boldsymbol{S}_{\rm p}(t_0+\Delta t)=\boldsymbol{J}_{\rm r}\dot{\boldsymbol{q}}_{\rm r}(t_0+\Delta t),\\ \boldsymbol{S}_{\rm p'}(t_0+\Delta t)=\boldsymbol{J}_{\rm s}\dot{\boldsymbol{q}}_{\rm s}(t_0+\Delta t),\end{cases}\tag{13}$$

式中:

$$\boldsymbol{S}_{\rm p}(t_0+\Delta t)=[\dot{x}_{\rm p}\ \ \dot{y}_{\rm p}\ \ \omega_2]^{\rm T},$$

$$\boldsymbol{S}_{\rm p'}(t_0+\Delta t)=[\dot{x}_{\rm p'}\ \ \dot{y}_{\rm p'}\ \ \dot{\theta}_{\rm s}]^{\rm T}.$$

碰撞结束后空间机器人与卫星固连成一个整体, 此时满足速度约束

$$\boldsymbol{S}_{\rm p}(t_0+\Delta t)=\boldsymbol{S}_{\rm p'}(t_0+\Delta t),\tag{14}$$

结合式(10)–(14)可计算出碰撞冲击效应为

$$\dot{\boldsymbol{q}}_{\rm r}(t_0+\Delta t)=\boldsymbol{A}^{-1}[\boldsymbol{M}_{\rm r}\dot{\boldsymbol{q}}_{\rm r}(t_0)+\boldsymbol{B}\dot{\boldsymbol{q}}_{\rm s}(t_0)],\tag{15}$$

其中: $\boldsymbol{A}=\boldsymbol{M}_{\rm r}+\boldsymbol{B}\boldsymbol{J}_{\rm s}^{-1}\boldsymbol{J}_{\rm r}$, $\boldsymbol{B}=\boldsymbol{J}_{\rm r}^{\rm T}(\boldsymbol{J}_{\rm s}^{\rm T})^{-1}\boldsymbol{M}_{\rm s}$.

将式(15)代入式(10)可计算出碰撞冲量$\boldsymbol{f}_{\rm p}$为

$$\boldsymbol{f}_{\rm p}=(\boldsymbol{J}_{\rm r}^{\rm T})^{+}\boldsymbol{M}_{\rm r}[(\boldsymbol{A}^{-1}\boldsymbol{M}_{\rm r}-\boldsymbol{E}_{5\times1})\dot{\boldsymbol{q}}_{\rm r}(t_0)+\boldsymbol{B}\dot{\boldsymbol{q}}_{\rm s}(t_0)],\tag{16}$$

式中$(\boldsymbol{J}_{\rm r}^{\rm T})^{+}$为$\boldsymbol{J}_{\rm r}^{\rm T}$的伪逆. 由于碰撞时间很短, 可以将碰撞力$\boldsymbol{F}_{\rm p}$近似为

$$\boldsymbol{F}_{\rm p}=\frac{\boldsymbol{f}_{\rm p}}{\Delta t}.\tag{17}$$

3.3 混合体系统动力学建模

式(14)对时间求导, 并结合式(13)可得碰撞后卫星的加速度为

$$\ddot{\boldsymbol{q}}_{\rm s}(t_0+\Delta t)=\boldsymbol{J}_{\rm s}^{-1}(\dot{\boldsymbol{J}}_{\rm r}-\dot{\boldsymbol{J}}_{\rm s}\boldsymbol{J}_{\rm s}^{-1}\boldsymbol{J}_{\rm r})\dot{\boldsymbol{q}}_{\rm r}(t_0+\Delta t)+\boldsymbol{J}_{\rm s}^{-1}\boldsymbol{J}_{\rm r}\ddot{\boldsymbol{q}}_{\rm r}(t_0+\Delta t).\tag{18}$$

由于$\boldsymbol{F}+\boldsymbol{F}_{\rm p'}=\boldsymbol{0}_{3\times3}$, 结合式(6)–(7)可得混合系统动力学方程为

$$\begin{cases}\boldsymbol{M}_{\rm rs}(\boldsymbol{q}_{\rm r})\ddot{\boldsymbol{q}}_{\rm r}+(\boldsymbol{H}_{\rm rs}(\boldsymbol{q}_{\rm r},\dot{\boldsymbol{q}}_{\rm r})+\boldsymbol{D}_{\rm L})\dot{\boldsymbol{q}}_{\rm r}=\boldsymbol{\tau}_{\rm r},\\ \boldsymbol{I}_{\rm m}\ddot{\boldsymbol{q}}_{\rm m}+\boldsymbol{D}_{{\rm m}\theta}\dot{\boldsymbol{q}}_{\rm m}+\boldsymbol{\tau}_{\theta}=\boldsymbol{\tau}_{\rm m},\\ \boldsymbol{K}_{\rm s}(\boldsymbol{q}_{\rm m}-\boldsymbol{q}_{\theta})+\boldsymbol{D}_{{\rm t}\theta}(\dot{\boldsymbol{q}}_{\rm m}-\dot{\boldsymbol{q}}_{\theta})=\boldsymbol{\tau}_{\theta},\end{cases}\tag{19}$$

式中:

$$\boldsymbol{H}_{\rm rs}=\boldsymbol{H}_{\rm r}+\boldsymbol{B}\boldsymbol{J}_{\rm s}^{-1}(\dot{\boldsymbol{J}}_{\rm r}-\dot{\boldsymbol{J}}_{\rm s}\boldsymbol{J}_{\rm s}^{-1}\boldsymbol{J}_{\rm r}),$$

$$\boldsymbol{M}_{\rm rs}=\boldsymbol{M}_{\rm r}+\boldsymbol{B}\boldsymbol{J}_{\rm s}^{\rm T}\boldsymbol{J}_{\rm r}.$$

在实际的应用中, 为了延长空间机器人的使用寿命, 其载体位置一般处于无控的状态; 而为了保证与地面的通讯, 需要将其天线指定向特定区域, 因此载体姿态需受控. 由于$\boldsymbol{H}_{\rm rs}$矩阵的前两列元素均为零, 故对式(19)第1式进行如下分块:

$$\begin{bmatrix}\boldsymbol{M}_{\rm rs11}&\boldsymbol{M}_{\rm rs12}\\ \boldsymbol{M}_{\rm rs21}&\boldsymbol{M}_{\rm rs22}\end{bmatrix}\begin{bmatrix}\ddot{\boldsymbol{X}}\\ \ddot{\boldsymbol{q}}_{{\rm rs}\theta}\end{bmatrix}+\begin{bmatrix}\boldsymbol{H}_{\rm rs11}+\boldsymbol{D}_{\rm L11}&\boldsymbol{H}_{\rm rs12}+\boldsymbol{D}_{\rm L12}\\ \boldsymbol{H}_{\rm rs21}+\boldsymbol{D}_{\rm L21}&\boldsymbol{H}_{\rm rs22}+\boldsymbol{D}_{\rm L22}\end{bmatrix}\begin{bmatrix}\dot{\boldsymbol{X}}\\ \dot{\boldsymbol{q}}_{{\rm rs}\theta}\end{bmatrix}=\begin{bmatrix}\boldsymbol{\tau}_{\rm B}\\ \boldsymbol{\tau}_{{\rm rs}\theta}\end{bmatrix},\tag{20}$$

式中:

$$\boldsymbol{M}_{\rm rs11},\boldsymbol{H}_{\rm rs11},\boldsymbol{D}_{\rm rs11}\in\mathbb{R}^{2\times2},$$

$$\boldsymbol{M}_{\rm rs12},\boldsymbol{H}_{\rm rs12},\boldsymbol{D}_{\rm rs12}\in\mathbb{R}^{2\times3},$$

$$\boldsymbol{M}_{\rm rs21},\boldsymbol{H}_{\rm rs21},\boldsymbol{D}_{\rm rs21}\in\mathbb{R}^{3\times2},$$

$$\boldsymbol{M}_{\mathrm{rs22}}, \boldsymbol{H}_{\mathrm{rs22}}, \boldsymbol{D}_{\mathrm{rs22}} \in \mathbb{R}^{3\times3},$$
$$\boldsymbol{X} = [x_0\ \ y_0]^{\mathrm{T}},\ \ddot{\boldsymbol{q}}_{\mathrm{rs}\theta} = [\theta_0\ \ \theta_1\ \ \theta_2]^{\mathrm{T}}.$$

由式(20)前两行可解出$\ddot{x}_0, \ddot{y}_0$, 然后回代入式(20)的后3行, 可得完全能控形式的混合体系统动力学方程为

$$\begin{cases}\boldsymbol{M}_{\mathrm{rs}\theta}\ddot{\boldsymbol{q}}_{\mathrm{rs}\theta} + (\boldsymbol{H}_{\mathrm{rs}\theta} + \boldsymbol{D}_{\mathrm{Lrs}\theta})\dot{\boldsymbol{q}}_{\mathrm{rs}\theta} = \boldsymbol{\tau}_{\mathrm{rs}\theta},\\ \boldsymbol{I}_{\mathrm{m}}\ddot{\boldsymbol{q}}_{\mathrm{m}} + \boldsymbol{D}_{\mathrm{m}\theta}\dot{\boldsymbol{q}}_{\mathrm{m}} + \boldsymbol{\tau}_{\theta} = \boldsymbol{\tau}_{\mathrm{m}},\\ \boldsymbol{K}_{\mathrm{s}}(\boldsymbol{q}_{\mathrm{m}} - \boldsymbol{q}_{\theta}) + \boldsymbol{D}_{\mathrm{t}\theta}(\dot{\boldsymbol{q}}_{\mathrm{m}} - \dot{\boldsymbol{q}}_{\theta}) = \boldsymbol{\tau}_{\theta},\end{cases} \tag{21}$$

式中:

$$\boldsymbol{M}_{\mathrm{rs}\theta} = \boldsymbol{M}_{\mathrm{rs22}} - \boldsymbol{M}_{\mathrm{rs21}}\boldsymbol{M}_{\mathrm{rs11}}^{-1}\boldsymbol{M}_{\mathrm{rs12}},$$
$$\boldsymbol{H}_{\mathrm{rs}\theta} = \boldsymbol{H}_{\mathrm{rs22}} - \boldsymbol{M}_{\mathrm{rs21}}\boldsymbol{M}_{\mathrm{rs11}}^{-1}\boldsymbol{H}_{\mathrm{rs12}},\ \boldsymbol{D}_{\mathrm{Lrs}\theta} = \boldsymbol{D}_{\mathrm{L22}}.$$

4　控制器设计

由于本文考虑捕获的卫星具有高速、旋转特性, 捕获完成后的空间机器人与卫星形成的混合体系统将严重失稳, 若直接利用强化信号设计控制力矩, 很容易导致强化学习系统因信息源单一而使镇定控制失败. 因此, 为了提高强化学习系统的稳定性, 所提策略通过一个代理器从固定增益控制器、自适应惩罚单元、模糊动作发生单元中收集信号, 且利用该信号设计控制力矩, 从而实现失稳混合体系统的镇定控制.

由于空间机器人系统参数未知, 如液体燃料的消耗导致载体质量变化, 机械臂向阳面与背阳面温差导致机械臂质心偏移等, 为了保证系统的稳定和跟踪性能, 本文采用了关联搜索单元(associate search element, ASE)对应的模糊控制器, 对自适应惩罚单元(adaptive penalty element, APE)进行了改进, 使模糊控制器的权值可根据惩罚信号进行调整.

强化学习系统结构如图3所示, 其中APE调整的模糊控制器不直接生成控制信号, 而是用于逼近空间机器人系统的不确定项; 性能评测单元通过测量系统状态生成误差评测信号$\boldsymbol{S}$; APE通过对该信号的采集, 生成惩罚信号$\boldsymbol{f}$对模糊系统进行优化.

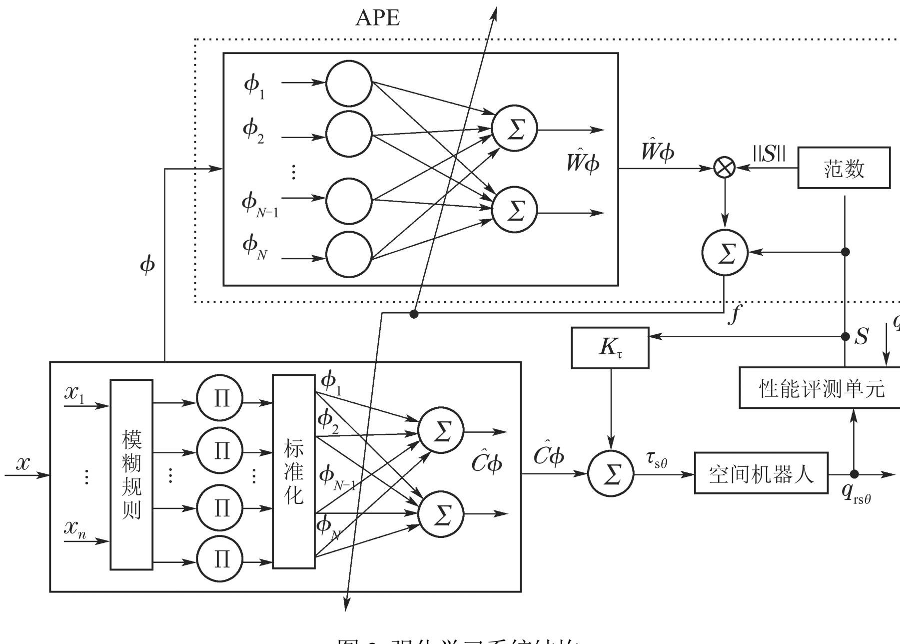

图 3　强化学习系统结构

Fig. 3　The diagram of reinforcement learning system

定义轨迹跟踪误差$\boldsymbol{e}$与误差评测信号$\boldsymbol{S}$为

$$\begin{cases}\boldsymbol{e} = \boldsymbol{q}_{\mathrm{d}} - \boldsymbol{q}_{\mathrm{rs}\theta},\\ \dot{\boldsymbol{e}} = \dot{\boldsymbol{q}}_{\mathrm{d}} - \dot{\boldsymbol{q}}_{\mathrm{rs}\theta},\end{cases} \tag{22}$$

$$\boldsymbol{S} = \dot{\boldsymbol{e}} + \boldsymbol{\Lambda}\boldsymbol{e}, \tag{23}$$

式中: $\boldsymbol{q}_{\mathrm{d}}, \dot{\boldsymbol{q}}_{\mathrm{d}} \in \mathbb{R}^{3\times1}$分别为载体与机械臂位置、速度的期望值, $\boldsymbol{\Lambda} \in \mathbb{R}^{3\times3}$为对称正定矩阵.

结合式(21)–(23)可得

$$\boldsymbol{M}_{\mathrm{rs}\theta}\dot{\boldsymbol{S}} = -(\boldsymbol{H}_{\mathrm{rs}\theta} + \boldsymbol{D}_{\mathrm{Lrs}\theta})\boldsymbol{S} + \boldsymbol{\rho} - \boldsymbol{\tau}_{\mathrm{rs}\theta}, \tag{24}$$

式中

$$\boldsymbol{\rho} = \boldsymbol{M}_{\mathrm{rs}\theta}(\ddot{\boldsymbol{q}}_{\mathrm{d}} + \boldsymbol{\Lambda}\dot{\boldsymbol{e}}) + (\boldsymbol{H}_{\mathrm{rs}\theta} + \boldsymbol{D}_{\mathrm{Lrs}\theta})(\dot{\boldsymbol{q}}_{\mathrm{d}} + \boldsymbol{\Lambda}\boldsymbol{e})$$

为系统的不确定项. 为了尽量消除系统不确定项对控制精度的影响, 因此采用ASE对应的模糊控制器对不确定项进行估计.

以$\boldsymbol{x} = [\ddot{\boldsymbol{q}}_{\mathrm{d}}\ \ \dot{\boldsymbol{e}}^{\mathrm{T}}\ \ \boldsymbol{e}^{\mathrm{T}}]^{\mathrm{T}}$为模糊控制器输入量, 通过IF–Then模糊规则对$\boldsymbol{\rho}$进行估计. 第n条模糊规则为

$$\boldsymbol{R}^n: \mathrm{IF}\ \ \boldsymbol{x}_1\ \ \mathrm{is}\ \ \Theta_1^n\ \ \mathrm{and}\ \ \boldsymbol{x}_2\ \ \mathrm{is}\ \ \Theta_2^n\ \ \mathrm{and}\cdots$$
$$\mathrm{Then}\ \ y^n\ \ \mathrm{is}\ \ \Psi^n,\ n = 1, 2, \cdots, N,$$

其中: Θ_i^n, Ψ^n是以$\mu_{\Theta_i^n}(x_i)$, $\mu_{\Psi^n}(y)$为隶属函数的模糊集合, N为模糊规则数量. 采用单值模糊控制器、

乘积推理机、面积中心解模糊控制器设计模糊系统. 模糊系统输出为

$$y(\boldsymbol{x})=\frac{\sum\limits_{n=1}^{N}\chi^n(\prod\limits_{i=1}^{9}\mu_{\Theta_i^n}(x_i))}{\sum\limits_{n=1}^{N}(\prod\limits_{i=1}^{9}\mu_{\Theta_i^n}(x_i))}, \tag{25}$$

式中χ^n为各个隶属函数为1时的值. $y(\boldsymbol{x})$可进一步简写为

$$y(\boldsymbol{x})=\boldsymbol{C}^{\mathrm{T}}\boldsymbol{\phi}(\boldsymbol{x}), \tag{26}$$

式中: $\boldsymbol{C}\in\mathbb{R}^{N\times m}$为理想权值矩阵;

$$\boldsymbol{\phi}(\boldsymbol{x})=[\boldsymbol{\phi}_1\ \ \boldsymbol{\phi}_2\ \ \cdots\ \ \boldsymbol{\phi}_N]^{\mathrm{T}}$$

为回归向量, 其定义为

$$\boldsymbol{\phi}_n=\frac{\prod\limits_{i=1}^{9}\mu_{\Theta_i^n}(x_i)}{\sum\limits_{n=1}^{N}(\prod\limits_{i=1}^{9}\mu_{\Theta_i^n}(x_i))}. \tag{27}$$

基于以上描述, $\boldsymbol{\rho}$可以被逼近为

$$\boldsymbol{\rho}=\boldsymbol{C}^{\mathrm{T}}\boldsymbol{\phi}(\boldsymbol{x})+\boldsymbol{\delta}(\boldsymbol{x}), \tag{28}$$

假设不确定项被模糊控制器逼近的估计值为

$$\hat{\boldsymbol{\rho}}=\hat{\boldsymbol{C}}^{\mathrm{T}}\boldsymbol{\phi}(\boldsymbol{x}), \tag{29}$$

式中$\hat{\boldsymbol{C}}\in\mathbb{R}^{N\times n}$为权值矩阵的估计值. 基于此, 设计如下控制信号:

$$\boldsymbol{\tau}_{\mathrm{rs}\theta}=\boldsymbol{K}_{\tau}\boldsymbol{S}+\hat{\boldsymbol{\rho}}, \tag{30}$$

式中$\boldsymbol{K}_{\tau}\in\mathbb{R}^{3\times3}$为正定对角矩阵.

将式(30)代入式(24)得

$$\boldsymbol{M}_{\mathrm{rs}\theta}\dot{\boldsymbol{S}}=-(\boldsymbol{K}_{\tau}+\boldsymbol{H}_{\mathrm{rs}\theta}+\boldsymbol{D}_{\mathrm{Lrs}\theta})\boldsymbol{S}+\tilde{\boldsymbol{\rho}}, \tag{31}$$

式中: $\tilde{\boldsymbol{\rho}}=\boldsymbol{\rho}-\hat{\boldsymbol{\rho}}=\tilde{\boldsymbol{C}}^{\mathrm{T}}\boldsymbol{\phi}(\boldsymbol{x})+\boldsymbol{\delta}(\boldsymbol{x}),\ \tilde{\boldsymbol{C}}=\boldsymbol{C}-\hat{\boldsymbol{C}}$.

通过APE可设计如下惩罚信号:

$$\boldsymbol{f}=\boldsymbol{S}+\|\boldsymbol{S}\|\hat{\boldsymbol{W}}^{\mathrm{T}}\boldsymbol{\delta}(\boldsymbol{x}), \tag{32}$$

式中$\hat{\boldsymbol{W}}\in\mathbb{R}^{N\times m}$为RBF神经网络权值矩阵的估计值. 因此, APE与ASE的自适应率设计为如下形式:

$$\begin{cases}\dot{\hat{\boldsymbol{W}}}=-\boldsymbol{K}_{\mathrm{W}}\|\boldsymbol{S}\|\boldsymbol{\phi}(\hat{\boldsymbol{C}}^{\mathrm{T}}\boldsymbol{\phi})^{\mathrm{T}}-\eta\boldsymbol{K}_{\mathrm{W}}\|\boldsymbol{S}\|\hat{\boldsymbol{W}},\\ \dot{\hat{\boldsymbol{C}}}=\boldsymbol{K}_{\mathrm{C}}\boldsymbol{\phi}\boldsymbol{f}^{\mathrm{T}}-\eta\boldsymbol{K}_{\mathrm{C}}\|\boldsymbol{S}\|\hat{\boldsymbol{C}},\end{cases} \tag{33}$$

式中: $\boldsymbol{K}_{\mathrm{W}},\boldsymbol{K}_{\mathrm{C}}\in\mathbb{R}^{3\times3}$为正定的对角矩阵, η为正数.

定理 1 $\dot{\boldsymbol{M}}_{\mathrm{rs}\theta}$与$\boldsymbol{H}_{\mathrm{rs}\theta}$满足斜对称性, 即

$$\frac{1}{2}\boldsymbol{z}\dot{\boldsymbol{M}}_{\mathrm{rs}\theta}\boldsymbol{z}^{\mathrm{T}}-\boldsymbol{z}\boldsymbol{H}_{\mathrm{rs}\theta}\boldsymbol{z}^{\mathrm{T}}=0,\ \forall\boldsymbol{z}\in\mathbb{R}^{3\times1}.$$

假设 1 RBF神经网络与模糊控制器的理想权值均有界, 即$\|\boldsymbol{W}\|\leqslant W_{\mathrm{m}},\ \|\boldsymbol{C}\|\leqslant C_{\mathrm{m}}$.

假设 2 模糊控制器逼近误差有界, 即$\|\boldsymbol{\delta}\|\leqslant\delta_{\mathrm{m}}$.

定理 2 若假设1–2均成立, 且采用式(23)所示的误差测评信号, 式(30)所示的控制信号, 式(32)所示的惩罚信号, 式(33)所示的自适应率, 则可保证闭环系统中的所有信号均有界.

证 定义如下形式的Lyapunov函数:

$$V=\frac{1}{2}\boldsymbol{S}^{\mathrm{T}}\boldsymbol{M}_{\mathrm{rs}\theta}\boldsymbol{S}+\frac{1}{2}\mathrm{tr}(\tilde{\boldsymbol{C}}^{\mathrm{T}}\boldsymbol{K}_{\mathrm{C}}^{-1}\tilde{\boldsymbol{C}})+\frac{1}{2}\mathrm{tr}(\tilde{\boldsymbol{W}}^{\mathrm{T}}\boldsymbol{K}_{\mathrm{W}}^{-1}\tilde{\boldsymbol{W}}), \tag{34}$$

式中: $\boldsymbol{M}_{\mathrm{rs}\theta}$为混合体系统正定惯量矩阵, $\tilde{\boldsymbol{W}}=\boldsymbol{W}-\hat{\boldsymbol{W}}$, 因此$V$是正定函数. 对式(34)求导得

$$\dot{V}=\boldsymbol{S}^{\mathrm{T}}\boldsymbol{M}_{\mathrm{rs}\theta}\dot{\boldsymbol{S}}+\frac{1}{2}\boldsymbol{S}^{\mathrm{T}}\dot{\boldsymbol{M}}_{\mathrm{rs}\theta}\boldsymbol{S}+\mathrm{tr}(\tilde{\boldsymbol{C}}^{\mathrm{T}}\boldsymbol{K}_{\mathrm{C}}^{-1}\dot{\tilde{\boldsymbol{C}}})+\mathrm{tr}(\tilde{\boldsymbol{W}}^{\mathrm{T}}\boldsymbol{K}_{\mathrm{W}}^{-1}\dot{\tilde{\boldsymbol{W}}}). \tag{35}$$

将式(31)(33)代入式(35)可得

$$\begin{aligned}\dot{V}=&-\boldsymbol{S}^{\mathrm{T}}(\boldsymbol{K}_{\tau}+\boldsymbol{D}_{\mathrm{Lrs}\theta})\boldsymbol{S}+\boldsymbol{S}^{\mathrm{T}}\tilde{\boldsymbol{\rho}}-\\&\boldsymbol{S}^{\mathrm{T}}\boldsymbol{H}_{\mathrm{rs}\theta}\boldsymbol{S}+\frac{1}{2}\boldsymbol{S}^{\mathrm{T}}\dot{\boldsymbol{M}}_{\mathrm{rs}\theta}\boldsymbol{S}+\\&\mathrm{tr}\big(\tilde{\boldsymbol{C}}^{\mathrm{T}}\boldsymbol{K}_{\mathrm{C}}^{-1}(-\boldsymbol{K}_{\mathrm{C}}\boldsymbol{\phi}\boldsymbol{f}^{\mathrm{T}}+\eta\boldsymbol{K}_{\mathrm{C}}\|\boldsymbol{S}\|\hat{\boldsymbol{C}})\big)+\\&\mathrm{tr}\big(\tilde{\boldsymbol{W}}^{\mathrm{T}}\boldsymbol{K}_{\mathrm{W}}^{-1}(\boldsymbol{K}_{\mathrm{W}}\|\boldsymbol{S}\|\boldsymbol{\phi}(\hat{\boldsymbol{C}}^{\mathrm{T}}\boldsymbol{\phi})^{\mathrm{T}}+\\&\eta\boldsymbol{K}_{\mathrm{W}}\|\boldsymbol{S}\|\hat{\boldsymbol{W}})\big).\end{aligned} \tag{36}$$

结合定理1, 式(36)可化简为

$$\begin{aligned}\dot{V}=&-\boldsymbol{S}^{\mathrm{T}}(\boldsymbol{K}_{\tau}+\boldsymbol{D}_{\mathrm{Lrs}\theta})\boldsymbol{S}+\boldsymbol{S}^{\mathrm{T}}\tilde{\boldsymbol{\rho}}+\\&\mathrm{tr}\big(-\tilde{\boldsymbol{C}}^{\mathrm{T}}\boldsymbol{\phi}\boldsymbol{f}^{\mathrm{T}}+\eta\tilde{\boldsymbol{C}}^{\mathrm{T}}\|\boldsymbol{S}\|\hat{\boldsymbol{C}}\big)+\\&\mathrm{tr}\big(\tilde{\boldsymbol{W}}^{\mathrm{T}}\|\boldsymbol{S}\|\boldsymbol{\phi}(\hat{\boldsymbol{C}}^{\mathrm{T}}\boldsymbol{\phi})^{\mathrm{T}}+\eta\tilde{\boldsymbol{W}}^{\mathrm{T}}\|\boldsymbol{S}\|\hat{\boldsymbol{W}}\big)=\\&-\boldsymbol{S}^{\mathrm{T}}(\boldsymbol{K}_{\tau}+\boldsymbol{D}_{\mathrm{Lrs}\theta})\boldsymbol{S}+\boldsymbol{S}^{\mathrm{T}}\tilde{\boldsymbol{\rho}}+\\&\mathrm{tr}\big(-\tilde{\boldsymbol{C}}^{\mathrm{T}}\boldsymbol{\phi}\boldsymbol{f}^{\mathrm{T}}+\eta\tilde{\boldsymbol{C}}^{\mathrm{T}}\|\boldsymbol{S}\|\hat{\boldsymbol{C}}+\\&\tilde{\boldsymbol{W}}^{\mathrm{T}}\|\boldsymbol{S}\|\boldsymbol{\phi}(\hat{\boldsymbol{C}}^{\mathrm{T}}\boldsymbol{\phi})^{\mathrm{T}}+\eta\tilde{\boldsymbol{W}}^{\mathrm{T}}\|\boldsymbol{S}\|\hat{\boldsymbol{W}}\big).\end{aligned} \tag{37}$$

令$\boldsymbol{K}_{\tau\mathrm{D}}=\boldsymbol{K}_{\tau}+\boldsymbol{D}_{\mathrm{Lrs}\theta}$, 且将$\tilde{\boldsymbol{\rho}}=\tilde{\boldsymbol{C}}^{\mathrm{T}}\boldsymbol{\phi}(\boldsymbol{x})+\boldsymbol{\delta}(\boldsymbol{x})$与式(32)代入式(37)可得

$$\begin{aligned}\dot{V}=&-\boldsymbol{S}^{\mathrm{T}}\boldsymbol{K}_{\tau\mathrm{D}}\boldsymbol{S}+\boldsymbol{S}^{\mathrm{T}}\boldsymbol{\delta}+\\&\|\boldsymbol{S}\|\mathrm{tr}\big(-\tilde{\boldsymbol{C}}^{\mathrm{T}}\boldsymbol{\phi}(\hat{\boldsymbol{W}}^{\mathrm{T}}\boldsymbol{\phi})^{\mathrm{T}}+\\&\eta\tilde{\boldsymbol{C}}^{\mathrm{T}}\hat{\boldsymbol{C}}+\tilde{\boldsymbol{W}}^{\mathrm{T}}\boldsymbol{\phi}(\hat{\boldsymbol{C}}^{\mathrm{T}}\boldsymbol{\phi})^{\mathrm{T}}+\eta\tilde{\boldsymbol{W}}^{\mathrm{T}}\hat{\boldsymbol{W}}\big).\end{aligned} \tag{38}$$

由于

$$\begin{aligned}&\|\boldsymbol{\delta}(\boldsymbol{x})\|\leqslant\delta_{\mathrm{m}},\ \tilde{\boldsymbol{C}}^{\mathrm{T}}\hat{\boldsymbol{C}}\leqslant\|\tilde{\boldsymbol{C}}\|C_{\mathrm{m}}-\|\tilde{\boldsymbol{C}}\|^2,\\&\tilde{\boldsymbol{W}}^{\mathrm{T}}\hat{\boldsymbol{W}}\leqslant\|\tilde{\boldsymbol{W}}\|W_{\mathrm{m}}-\|\tilde{\boldsymbol{W}}\|^2,\\&-\tilde{\boldsymbol{C}}^{\mathrm{T}}\boldsymbol{\phi}(\hat{\boldsymbol{W}}^{\mathrm{T}}\boldsymbol{\phi})^{\mathrm{T}}\leqslant\|\tilde{\boldsymbol{C}}\|\|\boldsymbol{\phi}\|^2W_{\mathrm{m}},\\&\tilde{\boldsymbol{W}}^{\mathrm{T}}\boldsymbol{\phi}(\hat{\boldsymbol{C}}^{\mathrm{T}}\boldsymbol{\phi})^{\mathrm{T}}\leqslant\|\tilde{\boldsymbol{W}}\|\|\boldsymbol{\phi}\|^2C_{\mathrm{m}}.\end{aligned}$$

因此得

$$\dot V\leqslant -\boldsymbol{S}^{\mathrm T}\boldsymbol{K}_{\tau\mathrm D}\boldsymbol{S}+\boldsymbol{S}^{\mathrm T}\boldsymbol{\delta}_{\mathrm m}+ \|\boldsymbol{S}\|\mathrm{tr}\big(\|\tilde{\boldsymbol{C}}\|W_{\mathrm m}\|\boldsymbol{\phi}\|^2+\eta\|\tilde{\boldsymbol{C}}\|C_{\mathrm m}+ \|\tilde{\boldsymbol{W}}\|\|\boldsymbol{\phi}\|^2C_{\mathrm m}+\eta\|\tilde{\boldsymbol{W}}\|W_{\mathrm m}- \eta\|\tilde{\boldsymbol{C}}\|^2-\eta\|\tilde{\boldsymbol{W}}\|^2\big)= -\boldsymbol{S}^{\mathrm T}\boldsymbol{K}_{\tau\mathrm D}\boldsymbol{S}+\boldsymbol{S}^{\mathrm T}\boldsymbol{\delta}_{\mathrm m}+ \|\boldsymbol{S}\|\big[(W_{\mathrm m}\|\boldsymbol{\phi}\|^2+\eta C_{\mathrm m})\|\tilde{\boldsymbol{C}}\|+(\|C_{\mathrm m}\|\boldsymbol{\phi}\|^2+ \eta W_{\mathrm m})\|\tilde{\boldsymbol{W}}\|-\eta\|\tilde{\boldsymbol{C}}\|^2-\eta\|\tilde{\boldsymbol{W}}\|^2\big]. \tag{39}$$

由模糊函数的定义可知$\|\boldsymbol{\phi}\|^2\leqslant N$, 故式(39)可化简为

$$\dot V\leqslant -\boldsymbol{S}^{\mathrm T}\boldsymbol{K}_{\tau\mathrm D}\boldsymbol{S}+\boldsymbol{S}^{\mathrm T}\boldsymbol{\delta}_{\mathrm m}+ \|\boldsymbol{S}\|\big[(W_{\mathrm m}N+\eta C_{\mathrm m})\|\tilde{\boldsymbol{C}}\|+(\|C_{\mathrm m}N+ \eta W_{\mathrm m})\|\tilde{\boldsymbol{W}}\|-\eta\|\tilde{\boldsymbol{C}}\|^2-\eta\|\tilde{\boldsymbol{W}}\|^2\big]\leqslant -K_{\min}\|\boldsymbol{S}\|^2-\eta\|\boldsymbol{S}\|\big[(\|\tilde{\boldsymbol{C}}\|-\lambda_{\mathrm C})^2+ (\|\tilde{\boldsymbol{W}}\|-\lambda_{\mathrm W})^2-(\lambda_{\mathrm C}^2+\lambda_{\mathrm W}^2+\lambda_\delta)\big], \tag{40}$$

式中: $K_{\min}$为$\boldsymbol{K}_{\tau\mathrm D}$的最小元素,

$$\lambda_{\mathrm C}=\frac{N\boldsymbol{W}_{\mathrm m}+N\boldsymbol{C}_{\mathrm m}}{2\eta},$$
$$\lambda_{\mathrm W}=\frac{N\boldsymbol{C}_{\mathrm m}+N\boldsymbol{W}_{\mathrm m}}{2\eta},\ \lambda_\delta=\frac{\delta_{\mathrm m}}{\eta}.$$

若要满足$\dot V\leqslant 0$, 则只需满足以下任意一个条件:

$$\|\boldsymbol{S}\|>\frac{\eta(\lambda_{\mathrm C}^2+\lambda_{\mathrm W}^2+\lambda_\delta)}{K_{\min}}=\delta_{\mathrm s}, \tag{41}$$

$$\|\tilde{\boldsymbol{C}}\|>\lambda_{\mathrm C}+\sqrt{\lambda_{\mathrm C}^2+\lambda_{\mathrm W}^2+\lambda_\delta}=\delta_{\mathrm C}, \tag{42}$$

$$\|\tilde{\boldsymbol{W}}\|>\lambda_{\mathrm W}+\sqrt{\lambda_{\mathrm C}^2+\lambda_{\mathrm W}^2+\lambda_\delta}=\delta_{\mathrm W}. \tag{43}$$

因此, 由式(41)–(43), 可以定义条件

$$\boldsymbol{\Omega}=\{(\|\boldsymbol{S}\|,\|\tilde{\boldsymbol{C}}\|,\|\tilde{\boldsymbol{W}}\|)|\|\boldsymbol{S}\|>\delta_{\mathrm s}, \|\tilde{\boldsymbol{C}}\|>\delta_{\mathrm C},\ \|\tilde{\boldsymbol{W}}\|>\delta_{\mathrm W}\}.$$

通过Lyapuno稳定性理论保证V收敛, 进一步由式(34)可得$\boldsymbol{S}, \tilde{\boldsymbol{C}}, \tilde{\boldsymbol{W}}$均收敛, 再由式(22)–(23)可得$\boldsymbol{q}_{\mathrm{rs}\theta}$, $\dot{\boldsymbol{q}}_{\mathrm{rs}\theta}$均收敛.

5　仿真模拟分析

5.1　碰撞过程弹簧阻尼缓冲机构抗冲击性能模拟

采用图2所示的空间机器人与卫星系统进行仿真试验. 空间机器人参数如下:

$$m_0=100\ \mathrm{kg},\ m_1=10\ \mathrm{kg},\ m_2=10\ \mathrm{kg},\ L_0=1\ \mathrm{m},$$
$$L_1=2\ \mathrm{m},\ L_2=2\ \mathrm{m},\ d_1=1\ \mathrm{m},\ d_2=1\ \mathrm{m},$$
$$I_0=64\ \mathrm{kg\cdot m^2},\ I_1=3.5\ \mathrm{kg\cdot m^2},\ I_2=3.5\ \mathrm{kg\cdot m^2},$$
$$I_{\mathrm{m1}}=0.05\ \mathrm{kg\cdot m^2},\ I_{\mathrm{m2}}=0.05\ \mathrm{kg\cdot m^2},$$
$$K_{\mathrm{s1}}=2865\ \mathrm{N/rad},\ K_{\mathrm{s2}}=2865\ \mathrm{N/rad},$$
$$D_{\mathrm{m1}}=28.65\ \mathrm{Ns/rad},\ D_{\mathrm{m2}}=28.65\ \mathrm{Ns/rad},$$
$$D_{\mathrm{t1}}=1146\ \mathrm{Ns/rad},\ D_{\mathrm{t2}}=1146\ \mathrm{Ns/rad},$$
$$D_{\mathrm{L1}}=28.65\ \mathrm{Ns/rad},\ D_{L2}=28.65\ \mathrm{Ns/rad}.$$

卫星参数如下:

$$m_{\mathrm s}=50\ \mathrm{kg},\ d_{\mathrm s}=0.5\ \mathrm{m},\ I_{\mathrm s}=8.5\ \mathrm{kg\cdot m^2}.$$

假设空间机器人初始位置为

$$\boldsymbol{q}_{\mathrm r}=[0\,\mathrm m\ \ 0\,\mathrm m\ \ 100^\circ\ \ 30^\circ\ \ 60^\circ]^{\mathrm T},$$

初始速度为

$$\dot{\boldsymbol{q}}_{\mathrm r}=[0\,\mathrm{m/s}\ \ 0\,\mathrm{m/s}\ \ 0\,(^\circ)/\mathrm s\ \ 0\,(^\circ)/\mathrm s\ \ 0\,(^\circ)/\mathrm s]^{\mathrm T}.$$

为了验证SDBD在空间机器人捕获卫星操撞击过程中的抗冲击性能, 在多组卫星速度下, 对关节所受冲击力矩进行力学模拟. 结果如表1所示.

表 1　不同卫星速度下SDBD抗冲击性能对比

Table 1　Comparison of impact resistance of SDBD at different satellite velocities

卫星速度/ $(\mathrm{m\cdot s^{-1}, m\cdot s^{-1}}, (^\circ)/\mathrm s)$	含SDBD最大冲击力矩/Nm	未含SDBD最大冲击力矩/Nm	降低百分比/%
$[0.05\ \ 0.05\ \ 8.6]^{\mathrm T}$	54.37	91.51	40.59
$[0.05\ \ 0\ \ 0]^{\mathrm T}$	18.24	35.43	48.52
$[0\ \ 0.05\ \ 0]^{\mathrm T}$	2.35	4.90	52.04
$[0\ \ 0\ \ 8.6]^{\mathrm T}$	31.80	55.35	42.55

从表1可以看出, 在碰撞过程中, 对于给定的不同卫星速度, SDBD均能显著的降低关节所受冲击力矩, 且最大可以降低52.04%, 因此可以认为其能在碰撞过程对关节起到较好的保护作用.

5.2　镇定控制过程柔顺策略性能模拟

为了实现捕获后严重失稳的混合体系统的镇定控制, 采用式(30)所示的控制力矩对其进行控制.首先, 通过性能评测单元测量系统状态继而生成误差评测信号$\boldsymbol{S}$; 然后, 利用APE对该信号的采集生成惩罚信号$\boldsymbol{f}$; 最后, 通过惩罚信号对系统状态进行优化, 从而不断地使模糊控制器达到最优逼近状态, 即实现控制的最优化.通过以上方式设计的强化学习系统, 可以有效的避免因信号源单一而造成的系统不稳定问题, 从而实现对失稳混合体系统的镇定控制.

控制参数:

$$\eta=1,\ \boldsymbol{K}_{\mathrm W}=\mathrm{diag}\{50,50,50\},$$
$$\boldsymbol{\varLambda}=\mathrm{diag}\{0.5,0.5,0.5\},\ \mathrm{K_C}=\mathrm{diag}\{50,50,50\},$$
$$\boldsymbol{K}_\tau=\mathrm{diag}\{2000,2000,2000\}.$$

空间机器人初始位置与速度同第5.1节, 卫星速度为

$\dot{\boldsymbol{q}}_{\rm s}(0) = [0.05\ \ 0.05\ \ 8.6(^\circ)/{\rm s}]^{\rm T}$, 混合体系统期望状态为$\boldsymbol{q}_{\rm d} = [100^\circ\ \ 30^\circ\ \ 60^\circ]^{\rm T}$. 为了尽可能的保护关节, 需要让SDBD将碰撞产生的冲击力矩卸载后电机才能开机, 经计算分析可知1 s内冲击力矩可卸载, 因此电机在发生碰撞1 s后开机. 结合式(15)(21)可计算出电机开机时混合体系统的位置为$\boldsymbol{q}_{{\rm rs}\theta} = [91.81^\circ\ \ 17.16^\circ\ \ 34.65^\circ]^{\rm T}$, 仿真时间为20 s.

为了突出镇定控制过程中SDBD的性能, 本文与目前较常用的一种SBD–串联弹性执行器(series elastic actuator, SEA)进行对比. 假设在电机负载情况下, 关节能承受的冲击力矩为60 Nm; 第1组仿真将关机力矩阈值设置为$F_{\rm C} = 40$ Nm, 开机力矩阈值设置为$F_{\rm O} = 5$ Nm; 考虑到随着空间机器人使用年数的增加, 关节所能承受冲击力矩将会下降, 因此第2组仿真将关机力矩阈值设置为$F_{\rm C} = 30$ Nm, 开机力矩阈值设置为$F_{\rm O} = 5$ Nm.

通过图4–5的对比可以发现, 在混合体系统实际状态与期望状态存在误差时, 自适应惩罚元件会不断地发出惩罚信号对模糊控制器进行优化, 最终实现控制器的最优化.

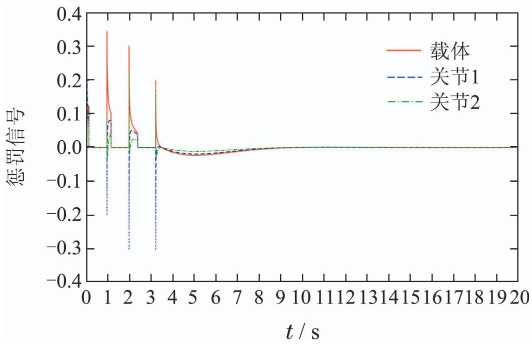

图 4 配置SDBD惩罚信号(第1组)

Fig. 4 Penalty signal with SDBD (1st group)

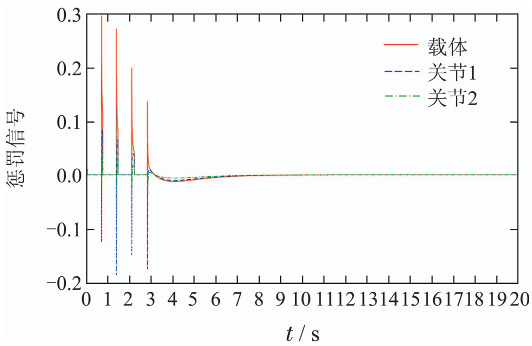

图 5 配置SEA惩罚信号(第1组)

Fig. 5 Penalty signal with SEA (1st group)

通过图6–7和图8–9的对比可看出, 由于SDBD中的阻尼器可以快速消耗冲击能量, 其系统表现为耗散系统, 因此配置SDBD比配置SEA的系统更加稳定(电机开、关机次数越多说明系统越不稳定), 且可以预测, 随着设置的关机力矩阈值继续减小, 配置SEA的系统将会失稳.

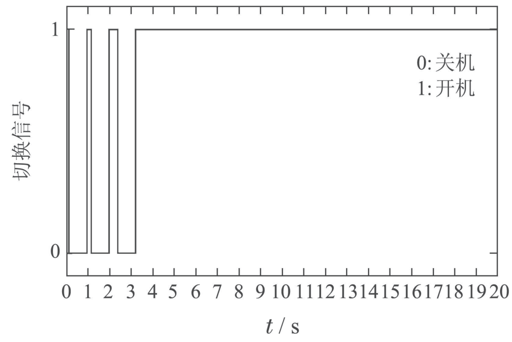

图 6 配置SDBD电机切换信号(第1组)

Fig. 6 Switch signal of motor with SDBD (1st group)

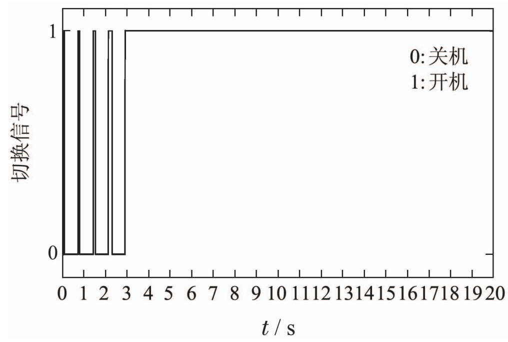

图 7 配置SEA电机切换信号(第1组)

Fig. 7 Switch signal of motor with SEA (1st group)

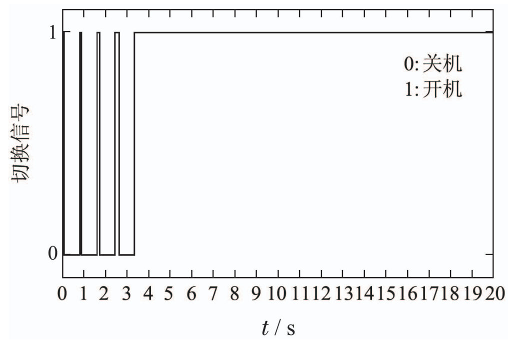

图 8 配置SDBD电机切换信号(第2组)

Fig. 8 Switch signal of motor with SDBD (2nd group)

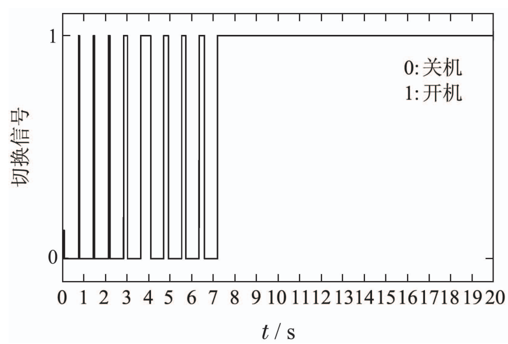

图 9 配置SEA电机切换信号(第2组)

Fig. 9 Switch signal of motor with SEA (2nd group)

通过图10–11和图12–13的对比可发现配置SDBD比配置SEA能更好的将冲击力矩限制在安全范围内, 即不会出现到达关机阈值后冲击力矩继续上升的现象, 考虑到随着使用时间的增加, 关节的抗冲击性能将会下降, 因此配置SDBD的空间机器人将比配置SEA的空间机器人具有更长的使用寿命.

通过图14–16可以看出, 配置SDBD与配置SEA均能使系统达到期望状态, 但配置SDBD相较于配置SEA具有更小的超调量, 也反映出配置SDBD具有更好的稳定性.

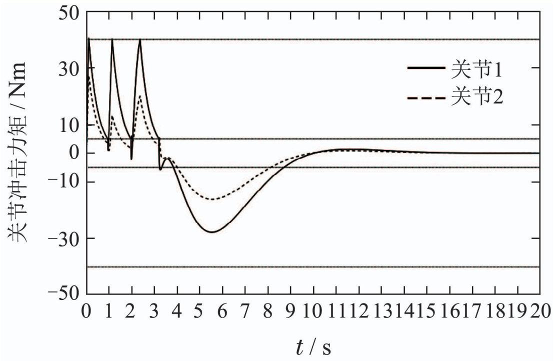

图 10 配置SDBD关节冲击力矩(第1组)

Fig. 10 Joint impact torque with SDBD (1st group)

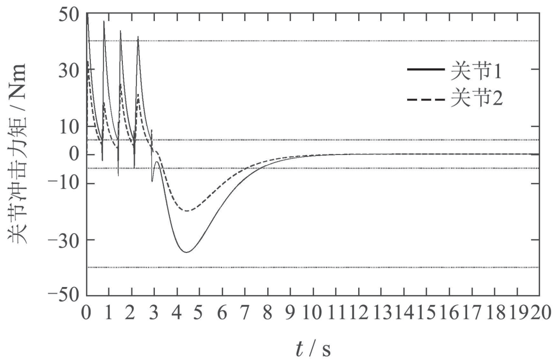

图 11 配置SEA关节冲击力矩(第1组)

Fig. 11 Joint impact torque with SEA (1st group)

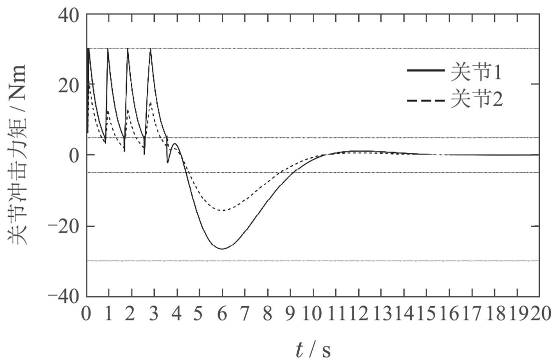

图 12 配置SDBD关节冲击力矩(第2组)

Fig. 12 Joint impact torque with SDBD (2nd group)

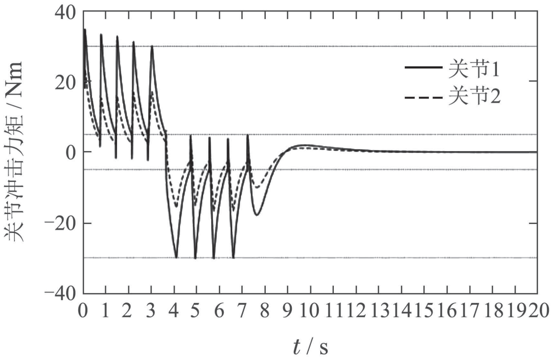

图 13 配置SEA关节冲击力矩(第2组)

Fig. 13 Joint impact torque with SEA (2nd group)

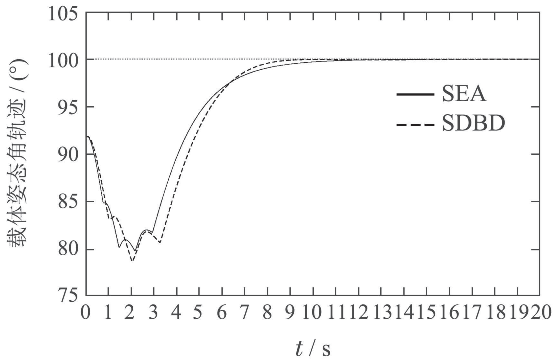

图 14 载体姿态角轨迹(第1组)

Fig. 14 Trajectory of attitude angle (1st group)

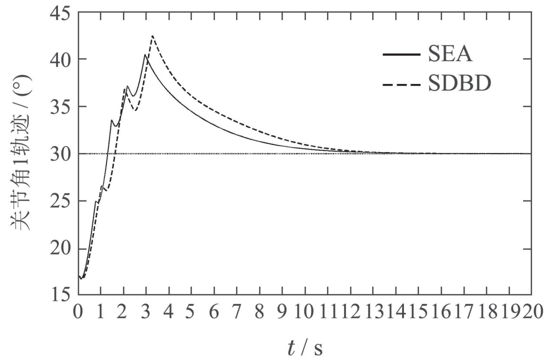

图 15 关节角1轨迹(第1组)

Fig. 15 Trajectory of joint angle (1st group)

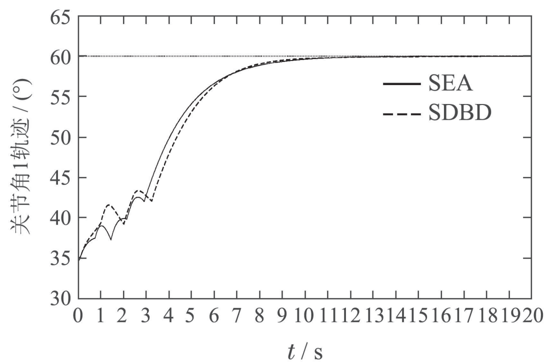

图 16 关节角2轨迹(第1组)

Fig. 16 Trajectory of joint angle 2 (1st group)

6 结论

本文为了在空间机器人捕获卫星操作过程中保护其关节不受冲击破坏, 在关节电机与机械臂之间设计了一种SDBD, 并给出了一种配合该装置的避撞柔顺策略; 计算了捕获过程的碰撞冲击效应、冲击力, 导出了失稳的混合体系统动力学方程, 并针对该系统的镇定控制提出了一种强化学习自适应模糊控制方案. 通过分析可以得出以下结论:

1) 捕获操作的碰撞过程中, 关节处会产生很大的冲击力矩, 在电机与机械臂之间添加SDBD可以实现冲击载荷的快速卸载;

2) 捕获操作的镇定控制过程中, 配合SDBD所设计的避撞柔顺策略可以将关节所受冲击力矩限制在安全范围内, 且配置SDBD的性能比配置SEA的性能更加优越, 系统更稳定.

参考文献:

[1] WANG M, LUO J, YUAN J, et al. Detumbling strategy and coordination control of kinematically redundant space robot after capturing a tumbling target. *Nonlinear Dynamics*, 2018, 92(3): 1023 – 1043.

[2] HUANG Yi, JIA Yingmin. Robust relative position and attitude control for non-cooperative fly-around mission. *Control Theory & Applications*, 2018, 35(10): 1405 – 1414.
(黄艺, 贾英民. 非合作目标绕飞任务的航天器鲁棒姿轨耦合控制. 控制理论与应用, 2018, 35(10): 1405 – 1414.)

[3] HUANG Xiuwei, DUAN Guangren. Attitude control and structure robust control allocation for combined spacecraft. *Control Theory & Applications*, 2018, 35(10): 1447 – 1457.
(黄秀韦, 段广仁. 组合航天器的姿态控制与结构鲁棒控制分配. 控制理论与应用, 2018, 35(10): 1447 – 1457.)

[4] YAN L, XU W, HU Z, et al. Virtual-base modeling and coordinated control of a dual-arm space robot for target capturing and manipulation. *Multibody System Dynamics*, 2019, 45(4): 431 – 455.

[5] GASBARRI P, PISCULLI A. Dynamic control interactions between flexible orbiting space-robot during grasping, docking and post-docking manoeuvres. *Acta Astronautica*, 2015, 110: 225 – 238.

[6] LIANG Jie, CHEN Li. Dynamics modeling for free-floating space-based robot during satellite capture and RBF neural network control for compound body stable movement. *Acta Aeronautica ET Astronautica Sinica*, 2013, 34(4): 970 – 978.
(梁捷, 陈力. 漂浮基空间机器人捕获卫星过程动力学模拟及捕获后混合体运动的RBF神经网络控制. 航空学报, 2013, 34(4): 970 – 978.)

[7] LIU X, LI H, CHEN Y, et al. Dynamics and control of capture of a floating rigid body by a spacecraft robotic arm. *Multibody System Dynamics*, 2015, 33(3): 315 – 332.

[8] GUO Wenhao, WANG Tianshu. Pre-impact configuration optimization for a space robot capturing target satellite. *Journal of Astronautics*, 2015, 36(4): 390 – 396.
(郭闻昊, 王天舒. 空间机器人抓捕目标星碰撞前构型优化. 宇航学报, 2015, 36(4): 390 – 396.)

[9] AGHILI F. A prediction and motion-planning scheme for visually guided robotic capturing of free-floating tumbling objects with uncertain dynamics. *IEEE Transactions on Robotics*, 2012, 28(3): 634 – 649.

[10] CHENG Jing, CHEN Li. Mechanical analysis and calm control of dual-arm space robot for capturing a satellite. *Chinese Journal of Theoretical and Applied Mechanics*, 2016, 48(4): 832 – 842.
(程靖, 陈力. 空间机器人双臂捕获卫星力学分析及镇定控制. 力学学报, 2016, 48(4): 832 – 842.)

[11] UYAMA N, HIRANO D, NAKANISHI H, et al. Impedance-based contact control of a free-flying space robot with respect to coefficient of restitution. *2011 IEEE/SICE International Symposium on System Integration*. Kyoto, Japan: IEEE, 2012: 1196 – 1201.

[12] DIMITROV D, YOSHIDA K. Utilization of the bias momentum approach for capturing a tumbling satellite. *2004 IEEE/RSJ International Conference on Intelligent Robots and Systems*. Sendai, Japan: IEEE, 2004: 3333 – 3338.

[13] YOSHIDA K, NAKANISHI H, UENO H, et al. Dynamics control and impedance matching for robotic capture of a non-cooperative satellite. *Advanced Robotics*, 2004, 2(2): 175 – 198.

[14] DONG Qiuhuang, CHEN Li. Composite control of robust stabilization and adaptive vibration suppression of flexible space manipulator capturing a satellite. *Robot*, 2014, 36(3): 342 – 348.
(董楸煌, 陈力. 柔性空间机械臂捕获卫星过程的鲁棒镇定与自适应抑振复合控制. 机器人, 2014, 36(3): 342 – 348.)

[15] SARIYILDIZ E, CHEN G, YU H Y. An acceleration-based robust motion controller design for a novel series elastic actuator. *IEEE Transactions on Industrial Electronics*, 2016, 63(3): 1900 – 1910.

[16] LI X, PAN Y, CHEN G, et al. Continuous tracking control for a compliant actuator with two-stage stiffness. *IEEE Transactions on Automation Science and Engineering*, 2018, 15(1): 57 – 66.

[17] KEPPLER M, LAKATOS D, OTT C, et al. Elastic structure preserving (EPS) control for compliantly actuated robots. *IEEE Transactions on Robotics*, 2018, 34(2): 317 – 335.

[18] WANG M, SUN L, YIN W, et al. Nonlinear disturbance observer based torque control for series elastic actuator. *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems*. Daejeon, Korea: IEEE, 2016: 286 – 291.

[19] ZHANG Xiuli, GU Xiaoxu, ZHAO Hongfu, et al. Design of a compliant robotic arm based on series elastic actuator. *Robot*, 2016, 38(4): 385 – 394.
(张秀丽, 谷小旭, 赵洪福, 等. 一种基于串联弹性驱动器的柔顺机械臂设计. 机器人, 2016, 38(4): 385 – 394.)

[20] LIU S, WU L, LU Z. Impact dynamics and control of a flexible dual-arm space robot capturing an object. *Applied Mathematics and Computation*, 2007, 185(2): 1149 – 1159.

[21] HUANG P, WANG M, MENG Z, et al. Attitude takeover control for post-capture of target spacecraft using space robot. *Aerospace Science and Technology*, 2016, 51: 171 – 180.

[22] LUO J, WEI C, DAI H, et al. Robust inertia-free attitude takeover control of postcapture combined spacecraft with guaranteed prescribed performance. *ISA Transactions*, 2018, 74: 28 – 44.

[23] HU Y, SI B. A reinforcement learning neural network for robotic manipulator control. *Neural Computation*, 2018, 30(7): 1983 – 2004.

[24] LI Z, LIU J, HUANG Z, et al. Adaptive impedance control of human-robot cooperation using reinforcement learning. *IEEE Transactions on Industrial Electronics*, 2017, 64(10): 8013 – 8022.

[25] LI Y, CHEN L, TEE K P, et al. Reinforcement learning control for coordinated manipulation of multi-robots. *Neurocomputing*, 2015, 170: 168 – 175.

作者简介:

朱 安 博士研究生, 目前研究方向为空间机器人系统动力学与控制, E-mail: zhu_an24@sina.com;

陈 力 教授, 博士生导师, 目前研究方向为空间机器人系统动力学与控制, E-mail: chnle@fzu.edu.cn.