机械手多任务均衡策略

朱 科[1,2],邢志明[1],康翔宇[1]

(1. 上海理工大学 光电信息与计算机工程学院,上海 200093;2. 江苏镭神智造科技有限公司,江苏 无锡 214142)

摘 要:并联机械手可实现高精度抓取和搬运动态物体,目前在智能工厂广泛应用。在物料输送过程中,单一机械手很难保证物料的无漏性,因此如何实现多机械手协同作业是一个亟待解决的问题。本文基于深度强化学习(DQN)算法提出最小化成本目标函数,并结合行动者-批评者算法对本研究算法进行改进,同时提出均衡优势函数以最大化优化均衡执行策略。实验显示,本研究算法的收敛性能提高17.1%。在既定实验条件下,与基线算法相比,本算法总成本降低约31.6%,显著提高了机械手在多任务场景下的整体执行效率。

关键词:机械手;深度强化学习;任务均衡;多任务

中图分类号:TP241;TP29 **文献标志码**:A **文章编号**:1671-5497(2025)08-2782-09

DOI:10.13229/j.cnki.jdxbgxb.20231309

Multi task balancing strategy for robot manipulators

ZHU Ke[1,2],XING Zhi-ming[1],KANG Xiang-yu[1]

(1. *School of Optical-Electrical and Computer Engineering*, *University of Shanghai for Science and Technology*, *Shanghai* 200093, *China*; 2. *Jiangsu Raysmart Intelligent Manufacturing Technology Co.*, *Ltd.*, *Wuxi* 214142, *China*)

Abstract: Continuing the research on multi-task balancing strategy for robot manipulators, the use of parallel robot manipulators in automated assembly lines allows for precise grasping and handling of dynamic objects on conveyors. However, the challenge lies in ensuring leakage-free handling of materials, which cannot be achieved by a single robot manipulator. This highlights the importance of multi-robot manipulator collaboration. In this study, a novel approach based on the deep Q-learning (DQN) algorithm is proposed to address the task balancing problem for robot manipulators. The objective is to minimize the cost function associated with the manipulation tasks. To improve the performance of the DQN algorithm, the study incorporates the actor-critic algorithm and introduces the concept of balancing advantage function. This allows for the optimization of the execution strategy, ensuring a balanced allocation of tasks among the robot manipulators. Experimental results validate the effectiveness of the proposed algorithm. It demonstrates a 17.1% improvement in convergence performance compared to the traditional DQN algorithm. Additionally, under predefined experimental conditions, the proposed algorithm achieves approximately a 31.6% reduction in overall cost when compared with the baseline algorithm. This

收稿日期:2023-11-28.

基金项目:国家重点研究开发计划项目(2017YFB0503102).

作者简介:朱科(1984-),男,高级工程师,博士.研究方向:新能源汽车充电系统研发,智能装备研发.
E-mail:zhuke12@126.com

significant reduction in cost enhances the overall execution efficiency of robot manipulators in multi-task scenarios.

Key words: robot manipulator; deep Q-learning algorithm; balanced allocation of task; multi-task

0　引　言

随着工业4.0的到来和智能制造理念的不断深化,现代工业生产的工作任务变得更加复杂,环境变得更加恶劣。传统单机械手虽已广泛应用,但在物料协同与处理方面存在明显缺陷。相比之下,多机械手协同工作能充分利用空间、信息和能耗等方面的优势[1],很好地解决单机械手场景存在的问题。Parikh等[2]提出了一种新型多机器人抓取任务分配问题的分类方法,从机器人、抓取任务和分配时间3个角度划分为ST-SR-IA等8种不同类型;文献[3]使用位置预测对目标物体进入空闲机械手的工作空间进行调度;文献[4]提出了一种基于非合作博弈的方法,即每个作业位在其作业范围内,对距其最近的目标物体和相邻最不可能被捡起的目标物体之间进行权衡选择,最终完成抓取任务分配;文献[5]利用改进粒子群算法对模糊控制规则权重及量化因子进行实时调整,设计了一个两输入三输出的自适应模糊PID抓取力控制器;文献[6]针对轮胎产业链条中机械手抓取姿态缺陷导致局部抓取力度不均匀、部分位置受力过大等问题,提出卷积神经网络模型,有效改善了传统机械手抓取姿态规划方法的不足;文献[7]利用位置预测,在目标进入空闲机械手工作空间时对其进行分配;文献[8]针对分拣和投递系统的快件堆放问题,提出并设计了一种基于视觉的机械手快递包裹分拣与放置系统,采用YOLO算法和PCL法向量确定包裹的取件点位置和法向量,规划并控制机械手移动到目标位置吸包;文献[9]提出了一种以非合作博弈为基础的方法,即每个操作员在其工作区间内,通过获得最近的目标和拾起邻居最不可能拾到的目标之间的权衡进行选择,以完成抓取任务分配;文献[10]针对六连杆采摘机械手运动规划的实时性问题,基于传统的人工势场(Artifical potential field, APF)优化算法,采用初等变换序列方法对机械手进行运动学建模,提出改进的人工势场(Impoved artificial potential field, IAPF)优化算法,以提升采摘机械手在避障运动规划中的实时性和准确性。针对现有抓取任务分配策略中存在的多个机械手负荷不一致、抓取任务分配不均以及缺乏整体计划的问题,本研究提出了一种基于改进深度强化学习的平衡抓取任务分配策略。该策略旨在保证抓取任务完整性的前提下,将抓取任务平均且合理地分配给多个机械手,以确保它们在工作中负荷相等,增强系统的稳定性。

1　机械手多抓取任务协同模型设计

本文以并联机械手抓取系统为研究对象,其实物结构如图1所示。在本文中,由于不考虑机械手的运动学分析,因此未对其物理结构进行建模,而是聚集于抓取任务分配计算模型的分析[11]。



图1　并联机械手抓取实物结构图

Fig. 1　Physical structure diagram of parallel manipulator grasping

1.1　机械手多任务协同操作基本流程

①在并联机械手抓取动作产生之前,需对组件进行识别、检测、判断,此过程需要检测组件的机械性能及电化学性能,以确定半导体参数,保障对组件的有效接触识别。②当组件进料出现在基底上,机械手通过对组件的长度、粗细等物理特性及电学特性进行感应识别,判断其各方面是否满足搭建合格器件的要求,决定产生抓取动作与否,使得组件与基底分离。③当机械手探针成功拾取组件后,需将其搬运到目标位置。若某些器件连接的位置存在一定角度或位置偏差,则在搬运过程中,需要通过2个机械臂配合调整抓手的相对位置实现碳纳米管不同程度的弯曲,将碳纳米管弯折形成一定的角度后放在相应位置。④对于合

格的组件,通过副操作器位姿调整后,基于图像视觉反馈将其对准被放置到预期位置,在切割模式下通过电流熔断实现切割组件使其与操作器分离。⑤主操作器与组件脱离后,复位并继续执行新的拾取任务。

1.1.1 复杂任务的简化

完整的机械手操作过程中,不同的任务又可以分解为多个子任务,包含定位(a)、测量(b)、接触(c)、分离(d),如表1所示。通过几个子任务的不同组合,可以完成复杂操作任务。

表1 不同任务流程和子任务组合

Table 1 Different task processes and subtask combinations

任务	定位识别	测量固定	拾取	搬运	释放
	1.探针复位位置 2.与半导体测试仪连接定位检测 3.移动至组件处	1.探针位置锁定固定	1.探针抓取 2.组件分离基底	1.目标位置确定 2.搬运目标位置	1.组件放置 2.组件着地
子任务组合	a-b	b	b-c	a-b-c	c-d

针对不同的操作任务,只需进行子任务的组合,将各子任务分配给对应的机械手,即可实现多机械手协调配合,完成组装过程中的各类任务。

1.1.2 不同子任务的操作策略设计

(1) 定位识别

在确定目标位置后,机械手操作探针首先以较快速度接近目标位置,当探针与目标位置距离小于10 m时,调整探针移动速度使其慢速接近目标位置,直到与目标位置的距离为零时停止运动。这种先快后慢的接近策略缩短了探针的定位时间,提高了效率,并且有利于避免探针在目标位置附近因速度过快导致与样品或与其他探针碰撞,导致操作失效,目标定位识别流程如图2所示。

(2) 接触抓取

由于操作过程在扫描电镜中进行,电子束从上方进入样品室扫描样品,如果采用从上方接近组件的方式,将无法观察到探针与组件接触的情况,因此采用由下往上的方式逐步接近组件。

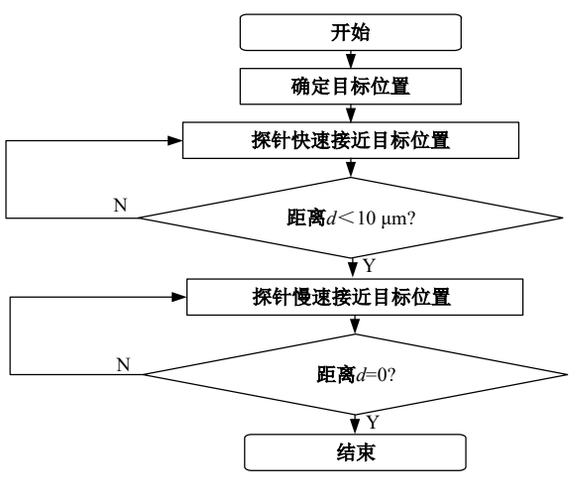

图2 目标定位识别流程

Fig. 2 Target positioning and recognition process

(3) 分离释放

在整个机械手操作过程中,分离的情况有2种:①拾取组件时使组件与基底分离;②将组件放置在目标位置后,使组件与探针分离。

(4) 测量固定

测量固定动作主要发生在抓取之前和抓取之后的目标放置阶段:一方面,完成一个完整抓取动作之后,需要复位、定位识别和固定测量,以便下一次准确抓取;另一方面,在将组件搬运到目标位置的过程中,需要对组件进行机械性能以及电学性能的检测,以便确保组件性能合格,从而节省整个组件组装的时间。

1.2 机械手多任务协同操作模型设计原理解析

1.2.1 计算模型设计

基于并联机械手的抓取场景,假设存在3排组并联机械手,其组件进料呈递进关系。设i为第i个抓取任务,$i\in\{1,2,\cdots,N\}$,j为第1排组第j个机械手节点,$j\in\{1,2,\cdots,M\}$。这些机械手节点可以对进料进行抓取和放置。在理想情况下,假定每个进料的抓取任务都是独立的作业单元,忽略抓取任务的关联性。当存在并发抓取任务请求时,需要优先确认抓取任务的延时上限和机械手执行完整套动作的预估时间,若机械手无法满足抓取任务响应需求,则发起负载转移,其抓取任务信息转移至其他机械手。为考虑进料的多样性,定义D_j为抓取任务信息的大小,生成如下抓取任务信息5元组:

$$F_{ij}=\left\{x_{ij},\lambda_{ij},\beta_{ij},D_i,T_i^{tot}\right\} \tag{1}$$

式中:x_{ij}为本地或负载执行的判断标识;λ_{ij}、β_{ij}分别为机械手j分配给任务i的作业时间和能耗比例;T_i^{wl}为最大延迟的上限。

机械手通过计算抓取任务的决策与执行,同时维护F_{ij}并同步分发到其他同一排组下的其他机械手节点,从而得出总抓取任务集合F:

$$F=\left\{F_{ij}|i\in\{1,2,\cdots,N\},j\in\{1,2,\cdots,M\}\right\} \quad (2)$$

通过总抓取任务集合F的计算模型,可以生成涵盖场景中所有机械手的计算模型。利用模型可以计算所有待抓取任务的负载执行成本,计算过程涉及任务执行所需的时间、能耗、材料和人力资源等因素。综合考虑这些因素,可以计算出每个待抓取任务的负载执行成本。这些成本数据对评估系统中机械手的资源利用率和执行效率至关重要。同时,通过比较不同的任务分配策略,可找到最佳的方案以最小化待抓取任务的负载执行成本,这一过程有助于提高机械手抓取任务的效率和经济性。

1.2.2　本地执行模型

当已分配的机械手能够满足任务的延迟限制时,任务将在本地执行。抓取任务的实际执行延迟可以表示为理论执行延迟加上运动学性能方面的补偿[12],即:

$$T'_{ij}=\left(1-x_{ij}\right)\frac{D_i}{\beta'_{ij}f'_i} \quad (3)$$

式中:β为机械手机械性能占比;f_j'为实际机械手的执行性能。能耗则定义为:

$$E'_{ij}=p'_iT'_{ij} \quad (4)$$

式中:p'_i为本地执行的功率。因此,结合上述可以得出能量消耗的总成本为:

$$C^{\mathrm{l}}_{ij}=\alpha T^{\mathrm{l}}_{ij}+(1-\alpha)E'_{ij} \quad (5)$$

式中:α为时间权重,$\alpha\in[0,1]$;$1-\alpha$为能耗权重,因为不同待抓取的需求任务权重差异,所以α可以根据抓取任务具体情况进行择优设定。

1.2.3　负载执行模型

当已有机械手无法提供所需的执行资源时,就会触发执行负载任务,整个任务过程可以分为5个阶段。①当需要执行负载任务i的抓取时,将向其他排组发送数据负载请求;②执行机械手节点基于接收的请求,做出最佳的负载决策,并发送给任务i;③i将待处理的数据上传到指定的机械手j,并通过传送带将相应的待抓取物分配给j;④机械手节点j为处理任务i分配执行资源;⑤机械手j向任务i返回执行抓取任务的结果[13]。当不同任务给定负载时,由于上行链路速率存在差异,因此假设抓取任务i在向机械手j负载时的上行链路速率为:

$$V^{\mathrm{up}}_{ij}=\lambda_{ij}B_j\log\left(1+\frac{P^{\mathrm{mp}}_i|H_i|}{N_0\Gamma\left(g_{\mathrm{up}}\right)d\left(i,j\right)^t}\right) \quad (6)$$

式中:λ_{ij}为机械手j分配给任务i的作业时间;B_j为当前节点j可达到的传送带承载能力;P^{mp}_i为任务i所需的执行功率;H_i为传送带中抓取任务i的能耗增益;N_0为传送带的能耗功率;g_{up}为一个目标位错率;$d\left(i,j\right)^t$为执行机械手i与j的距离;$\Gamma\left(g_{\mathrm{up}}\right)$为一个传送通道的损失指标。因此,传输延迟$T^{\mathrm{h}}_i$可以定义为:

$$T'^{j}_{ij}=x_{ij}\frac{D_i}{V^{up}_{ij}} \quad (7)$$

根据式(7),传输能耗E^{in}_{ij}可以定义为:

$$E^{\mathrm{hu}}_{ij}=p^{\mathrm{ap}}_iT^{\mathrm{hu}}_{ij} \quad (8)$$

式中:p^i_i为上行的传输功率。执行时间T^{e}_{ij}定义为:

$$T^{\mathrm{e}}_{ij}=x_{ij}\left(\frac{D_i}{\beta_{ij}f^{\mathrm{e}}_j}\right) \quad (9)$$

如果负载到机械手完成抓取后,执行节点会将结果反馈至控制中心,那么传输延迟可以定义为从执行节点发送结果到控制中心接收结果的时间间隔。传输延迟考虑了网络传输的时延、数据处理的时延以及其他影响因素,反映了控制中心获取执行结果的延迟情况,可以用来评估执行效率和系统响应时间。传输延迟定义为:

$$T^{ij}_{ij}=x_{ij}\frac{D^o_{ij}}{V^{ij}_{ij}} \quad (10)$$

D^o_{ij}在式(10)的基础上,结合结果信息的大小,可以得出反馈接收的整体系统能耗,如式(11)所示。该能耗可以用来评估系统在接收反馈结果时所消耗的能量[14]:

$$E^{\mathrm{td}}_{ij}=p^{\mathrm{do}}_iT^{\mathrm{id}}_{ij} \quad (11)$$

式中:p^{do}_i为整体执行机构的功率。

综上,抓取任务i的负载执行延迟为:

$$T'^{\mathrm{awl}}_{ij}=T'^{\mathrm{a}}_{ij}+T^{\mathrm{e}}_{ij}+T'^{j}_{ij} \quad (12)$$

持续能耗为:

$$E^*_{ij}=p^*_iT^{\mathrm{swal}}_{ij} \quad (13)$$

式中:p^*_i为持续等待过程中整体执行设备的能耗功率。

最终总能耗为:

$$E^{\mathrm{msw}}_{ij}=E'^{i}_{ij}+E'^{j}_{ij}+E^{\mathrm{v}}_{ij} \quad (14)$$

通过式(13)(14)计算总成本为:

$$C^{\mathrm{m}\infty\varpi}_{ij}=\alpha T^{\mathrm{mowi}}_{ij}+(1-\alpha)E^{\mathrm{posel}}_{ij} \quad (15)$$

1.3 机械手多任务协同设计不足及优化思路

待抓取目标的抓取动作时长和整体设备的能耗,在计算模型中被定义为双重优化目标。为综合考虑这2个指标,定义它们的加权和为:

$$\min_{\lambda_j,\beta_j,-,ij} C=\min_{\lambda_j,\beta_j,\tau_j}\sum_{j=0}^{M}\sum_{i=1}^{N}Z_{ij}\left[\left(1-x_{ij}\right)C_{ij}^{1}+x_{ij}C_{ij}^{\text{total}}\right] \tag{16}$$

在上述优化问题中,整个作业完成时间与能量消耗的加权和即为目标函数,如式(16)所示,用总费用C表示。同时包含约束:①无论是本地机械手执行引起的延迟,还是负载机械手引起的延时,都不能超过任务的最大可承受延迟;②单个抓取任务分配的能耗占比≤1,即向机械手负载的全部抓取任务所占用的资源总量不超过机械手的最大功能;③$z_{ij}=0$和$z_{ij}=1$分别代表不选择和选择既定机械手节点j执行任务。

2 多任务协同负载策略研究

该优化问题可以归类为混合整数规划问题,因此本研究提出基于DRL-AC3O的异步机械手多任务均衡算法。DRL-AC3O算法采用非同步多线程技术,支持并联机械手系统中不同排组间的机械手执行节点与周围环境交互学习,并向控制中心单元发送学习的梯度参数;执行节点从控制中心单元定期接受新的参数,可以更好地引导目前的机械手执行单元和后续学习。DRL-AC3O是动态智能强化学习模型,它通过持续评估机械手执行单元的环境影响因素,并和控制中心进行交互,最终输出最优抓取任务均衡执行策略。

假设在每个时间步骤中,整体环境的状态为s_t,执行动作为α,接收奖励为r_{r+1},执行策略为π。基于这些因素,通过累积奖励的最大化可获得最优的抓取任务执行策略。根据这种设定,算法的状态空间可以表示为时间步骤t时抓取任务执行策略的集合[15]。

$$S_t=\left(C_{ij}(t)\right) \tag{17}$$

动作空间定义为:

$$A_i=\left[z_{ij}(t),\lambda_{ij}(t),\beta(t)\right] \tag{18}$$

式中:$z_{ij}(t)$在t内选择进行分配的任务和机械手执行单元;节点在执行动作α后,在状态s条件下获得的奖励为$R(s,a)$。

研究最优问题是使所有任务的作业时间和能源消耗的加权和总费用最少,而解的目的是得到最大的奖励函数值,所以奖励值与总费用成反关系,设v为特定环境确定的奖励值[16],则奖励函数为:

$$r_t^{ij}=\begin{cases}r_{t-1}^{ij}+v, C_{ij}(t)<C_{ij}(t-1)\\ r_{t-1}^{ij}, C_{ij}(t)=C_{ij}(t-1)\\ r_{t-1}^{ij}-v, C_{ij}(t)>C_{ij}(t-1)\end{cases} \tag{19}$$

2.1 整体流程分析

本文算法的整体流程如图3所示。当机械手执行单元接收控制中心的所有抓取任务请求时,通过2个深度神经网络对执行单元中智能体的策略和价值函数进行近似估计。其中,策略网络用于估计一组动作的输出概率,价值网络用于评估状态价值。

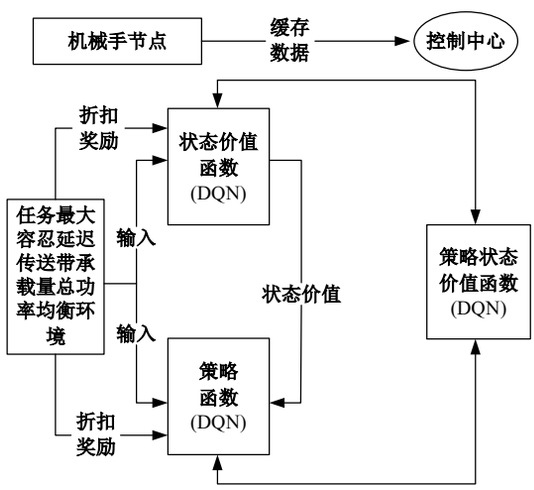

图3 DRL-AC3O执行框图

Fig. 3 DRL-AC3O execution block diagram

在节点上,智能体的运行采用最优分配策略。机械手执行单元将神经网络的参数梯度和缓存信息发送到控制中心,由控制中心学习后将最优参数反馈给各机械手执行单元,以指导智能体运行。这种机制使DRL-AC3O可以为抓取任务提供长期均衡优化服务。

2.2 分配策略函数

深度神经网络是一种基于策略函数的网络结构,它通过对输入数据的深度特征建模,为智能体的决策和控制奠定坚实的基础,能够实现对复杂系统的感知和决策。在这个基础上,将机械手执行单元智能体的不同观察值引入移动策略的开发中,并给出以下策略函数:

$$\pi(s_t,a_t)\approx\pi(a_t|s_t;\theta_t) \tag{20}$$

该方法通过不断更新权重实现优化。由于策

略函数的目标是最大化奖励,因此可以使用梯度上升算法计算奖励期望。θ网络的权重参数在控制网络中不断地被更新。策略梯度的更新公式为:

$$\nabla_\theta E[R_t]=\nabla_\theta \log\pi(a_t|s_t;\theta_t)R_t \tag{21}$$

式中:$\pi(a_t|s_t;\theta_t)$为s_t状态下选择动作a_t的概率;$\nabla_\theta \log\pi(a_t|s_t;\theta_t)R_t$为对$\nabla_\theta E[R_t]$的无偏性估计。

假定每一个动作的奖励值R_r都为正(即所有的梯度值>0),那么在梯度递增算法的作用下,每一次动作发生的概率都会逐渐增加,但可能降低学习速度,并导致梯度变化的增加。因此,将用于减小梯度变化的标准化运算引入该部分中,将式(21)修正为:

$$\nabla_\theta E[R_t]=\nabla_\theta \log\pi(a_t|s_t;\theta_t)\big(R_t-b_t(st)\big) \tag{22}$$

其中,策略函数是用R_t减去基准函数$b_t(st)$来减少估计方差,同时保持其无偏性。对于奖励值R_t的预期估算,通过计算基线b的梯度来更新参数θ,则在总体奖励超出基线操作的情况下,增加了其动作概率,并减小了梯度变化。

2.3　目标评估函数

在算法流程中,以价值函数为基础的深度神经网络,将任务分配至机械手执行单元的观察,观察数据包含了机械手执行单元的总功率和最大允许时延,并定义了行为值函数Q:

$$Q(s_t,a_t)\approx Q(s_t,a_t;w_t) \tag{23}$$

式中:w为深度神经网络参数。

基于这些参数,利用深度神经网络构建损失函数。损失函数衡量了本研究模型在任务中的预测与实际结果的差异。最小化损失函数可以优化网络权重,使其能更准确地预测任务的结果。

$$L(w_t)=E\left[\left(\text{Target } Q-Q(s_t,a_t;w_t^-)\right)^2\right] \tag{24}$$

在该部分中,利用多步Q-learning算法的对象行为值进行计算。该算法的优势在于,当前奖励r可以直接影响之前的状态-动作对,并且能够更好地模拟历史分配过程,从而显著改善算法的学习效果。多步Q-learning算法中的多步骤意味着计算包含接下来的多个步骤。因此,本文定义了目标值(Target)的计算方式,即:

$$\text{Target } Q=r_t+\gamma r_{t+1}+\cdots+\gamma^{n-1}r_{t+n-1}+\max_a\gamma Q(s_{t+n},a) \tag{25}$$

在DRL-AC3O算法中,基于策略函数的DRL-AC3O算法与以价值函数为基础的DRL-AC3O算法具有相似之处,只是在同一计算模型下输出的数据流不同。

2.4　均衡优势函数

在DRL-AC3O算法中,策略函数$\pi(a_t|s_t;\theta_t,b_t(st))$可看作一个评论者。优势函数可以抽象成行动者-评论者算法的损失函数,并结合竞争网络的特点,通过对均衡场景的调整和优化,使其能够更好地估算行为价值。在策略梯度更新过程中(式(23)),该更新规则采用了折现R_t来告知机械手执行单元哪些是正确的,而哪些是错误的。随后,网络进行更新以确定分配决策的质量和价值函数的最佳值。优势函数的表达式为:

$$A(s_t,a_t)=Q(s_t,a_t)-V(s_t) \tag{26}$$

式中:状态$V(s_t)$为对整体概率的期望,给定一组t条件下的所有动作值函数;动作$Q(s_t,a_t)$为每个执行动作的价值。

式(26)中$Q(s_t,a_t)-V(s_t)$可以评估基于平均值的动作值的大小。然而,由于在目标函数中无法对Q进行定性评估,因此将其作为一种估值,最终函数的表达式为:

$$A(s_t,a_t)=R(s_t,a_t)-V(s_t) \tag{27}$$

使用优势函数进行估算的优点在于,在折扣奖励的作用下,能够更精准地评估分配决策的质量。

3　实验分析

3.1　实验基础条件

根据抓取运动时间特性,不同速度和不同距离会产生不同的能耗。若机械手抓取操作时抓取装置末端的运动速度不当,会导致组件不同程度的损伤。为保障安全,减小组件助力机械手抓取运行时产生的振动和冲击,机械手夹紧力与离心力平衡十分关键。同时,为提高准确识别和定位,减少距离和速度能耗,需要设置目标物体在组件助力机械手抓取坐标系的方位角为θ,组件助力机械手末端执行器的方位角为β,保证组件助力机械手抓取时始终指向目标物体,并逐步移至目标物体的位置,通过探针复位、识别、定位、测量、抓取动作,完成x_{ij}任务。事实上,x_i到x_n不完全一致,存在距离差,由此会产生不同的时间差和功耗。当i和n相同时,能耗差最少;每差一个位置,

可参照总能耗公式式(14)计算,即可得出每多出一个距离,能耗增加0.001 J,具体参数配置如表2所示。

表2 实验配置

Table 2 Experimental configuration

机械手排组	传送带承载量/g	传送速率/(mm·s^{-1})	能耗/j
1	100	150	0.002
2	150	100	0.003
3	200	200	0.001

仿真实验通过一系列模拟来评价本文算法的性能,并将其与目前3种常见的混个表格方法和深度Q网络进行对比。具体来说,假定20个待抓取且每一个任务抓取执行功耗在0.01~0.03 J,并且随机产生各既定节点到负载节点的距离。此外,假定机械手的性能参数f_j'为30,资源消耗p_i'为0.02 J,为便于计算,每小时的传送带工作资源消耗均为0.001 J。

3.2 实验结果分析

由表2可知,能耗差异确实存在。结合反馈接收的整体系统能耗式(11)、总能耗式(14)进行计算,其结果如图4所示的DRL-AC3O损失函数,呈现一定的收敛性。这主要是因为初始阶段的行为会对奖励值产生较大的影响,在迭代次数从0至1 000的过程中,y轴损失函数值从350迅速降低,然后逐渐接近最佳值0。由于实际能耗不可能为0,损失函数无限趋近于0,此时神经网络参数收敛至最佳。

图5为一个基于机械手节点的信息反馈,呈现一定的周期性。从一个抓取动作完成到下一个抓取动作完成开始,由于深度神经网络具有收敛特性,每次新的机械手执行参数梯度被接收后,在

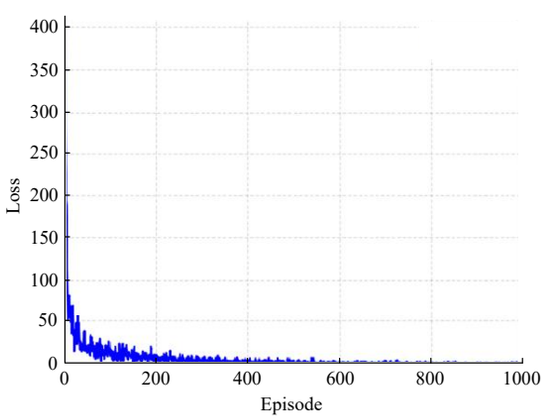

图4 损失函数收敛特性

Fig. 4 Convergence characteristics of loss function

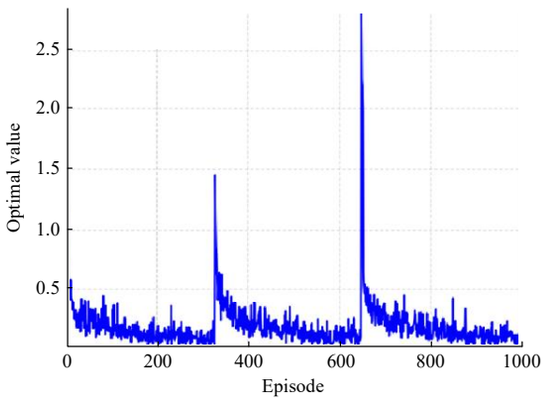

图5 优势函数收敛特性

Fig. 5 Convergence characteristics of optimal functions

约300次迭代后优势函数趋于稳定,且结果显示以能量消耗最少原则和能耗收敛规律,机械手出现最大抓取功耗和最小功耗的概率增加。从而证明了每次从机械手节点处收到的参数梯度反馈后,神经网络模型能在一定范围内实现对均衡策略的动态优化响应,如图5所示,0~300、300~600、600~1 000平均在350次左右。该算法在迭代次数内得到了最优函数的收敛,得到了最优网络的最优。

图6为5种算法的总成本对比,包括整体负载、局部计算、贪婪算法、DQN算法以及本研究提出的DRL-AC3O算法。由图6可知:①随着任务量的增大,各种方法的总成本均呈上升趋势;②初始阶段的工作时间较短,局部计算、贪婪算法、DQN算法、DRL-AC3O算法的总成本相同,原因是任务量未超过本地计算能力范围;③综合各种算法,DRL-AC3O算法是最接近贪婪算法所需的总成本。由于贪婪算法在执行速率和资源消耗上要远远高于DRL-AC3O算法,尤其是在复杂的

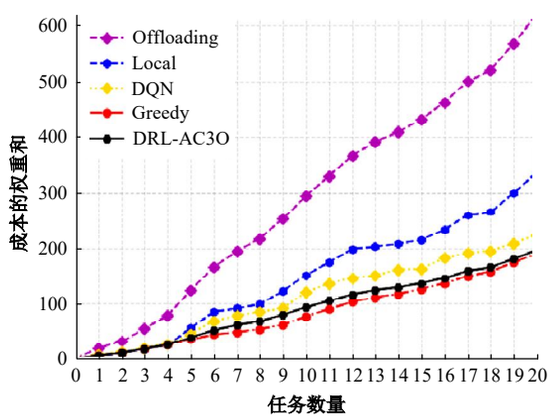

图6 总成本比较

Fig. 6 Total cost comparison

机械手工作环境下,本文提出的DRL-AC3O算法性能优势明显。

4　结束语

随着人工智能技术的深入研究和广泛应用,以及工业4.0和中国制造2025实现机械全面自动化之际,机械手应用普及水平逐步提高。随之而来的双碳目标和节能减排应用要求,机械手多任务协同能耗、传统负载精准性等成为新的时代研究主题。为此,在充分搜集国内外关于机械手多任务协同的研究资料的前提下,聚焦探讨并联机械手智能化、自动化实现的机制和原理,对能耗、传统带承载量和负载决策等综合考虑进行重点研究。

研究发现,在机械手多任务协同过程中,不同任务目标的损耗函数(位置、距离、速度差异)都会导致能耗变化,且能耗的规律呈现一定的收敛性:每多一个位置距离,则会多0.001 J的能耗。同时,由于每个机械手一次完整的定位识别、测量固定、抓取搬运、释放复位过程具有周期性,基于机械手节点的信息反馈呈现一定的周期性本文在行动者－批评者算法的基础上,给出了一种基于深度强化学习的异步机械手负载协作算法。结合机械手多任务协同的过程和模型设计原理解析,以及能耗存在的客观性,面对整体负载、局部计算、贪婪算法、DQN算法以及本文提出的DRL-AC3O算法等多种策略,将竞争神经网络的竞争特性与传统的损失函数相比较,构造均衡优势函数,以机械手多任务均衡策略对分配策略进行适应性调整,不仅能耗水平明显降低,而且在减少样本相关性的同时,能够利用奖励值更精准地估算行为价值,更有效地确保决策的动态性和多样性。可以预见,机械手神经网络模型的深度学习能力和DRL算法(如DQN)的均衡策略动态响应能力,具有最优能耗应用前景和实用潜力。

参考文献:

[1] 张映锋, 张党, 任杉. 智能制造及其关键技术研究现状与趋势综述[J]. 机械科学与技术, 2019, 38(3): 329-338.
Zhang Ying-feng, Zhang Dang, Ren Shan. Survey on current research and future trends of smart manufacturing and its key technologies[J]. Mechanical Science and Technology for Aerospace Engineering, 2019, 38(3): 329-338.

[2] Parikh P J, Lam S S Y. A hybrid strategy to solve the forward kinematics problem in parallel manipulators[J]. IEEE Transactions on Robotics, 2005, 21(1): 18-25.

[3] 国家制造强国建设战略咨询委员会. 中国制造2025蓝皮书[M]. 北京: 电子工业出版社, 2016.
National Manufacturing Power Construction Strategy Advisory Committee. China manufacturing 2025 bluebook[M]. Beijing: Electronic Industry Press, 2016.

[4] Chen X Y, Ren G Q. Key technologies and development trends of intelligent manufacturing and robot application[J]. IOP Conference Series: Earth and Environmental Science, 2020, 461(1): 012049.

[5] 管声启, 张理博, 刘通, 等. 基于改进粒子群算法的机械手抓取力自整定模糊PID控制[J]. 西安工程大学学报, 2024, 25(2):1-8.
Guan Sheng-qi, Zhang Li-bo, Liu Tong, et al. Self-tuning fuzzy PID control of manipulator grasping force based on improved particle swarm optimization algorithm[J]. Journal of Xi′an Polytechnic University, 2024, 25(2):1-8.

[6] 郭美君. 机械手控制系统的研究与应用[J]. 电动工具, 2023(6): 24-26.
Guo Mei-jun. Research and application of manipulator control system[J]. Electric Tools, 2023(6):24-26.

[7] 卢志琦. 基于卷积神经网络的煤矿支架搬运车轮胎助力机械手抓取姿态规划方法[J]. 煤矿机械, 2024, 45(2): 179-182.
Lu Zhi-qi. Grasping attitude planning method of tire assisted manipulator of support truck in coal mine based on convolutional neural network[J]. Coal Mine Machinery, 2024, 45(2): 179-182.

[8] 蒋智蓓, 张松亚, 邓佳成, 等. 基于计算机视觉的机械手包裹分拣与放置系统[J]. 机电工程技术, 2024, 53(1): 212-215.
Jiang Zhi-bei, Zhang Song-ya, Deng Jia-cheng, et al. A robot parcel sorting and placement system based on computer vision[J]. Mechanical & Electrical Engineering Technology, 2024, 53(1): 212-215.

[9] Chen X Y, Ren G Q. Key technologies and development trends of intelligent manufacturing and robot application[J]. IOP Conference Series: Earth and Environmental Science, 2020, 461(1): 012049.

[10] 姚国林, 王合闯. 改进的APF算法在采摘机械手运动规划中的应用[J]. 机床与液压, 2023, 51(23): 65-72.
Yao Guo-lin, Wang He-chuang. Application of improved APF algorithm in motion planning of picking

manipulator[J]. Machine Tool & Hydraulics, 2023, 51(23): 65-72.

[11] 张明文, 于振中. 工业机器人原理及应用:DELTA并联机器人[M]. 哈尔滨: 哈尔滨工业大学出版社, 2018.
Zhang Ming-wen, Yu Zhen-zhong. Principles and Applications of Industrial Robots: DELTA Parallel Robots[M]. Harbin: Harbin Institute of Technology Press, 2018.

[12] 张英坤. Delta并联机器人的研究进展[J]. 机床与液压, 2016, 44(21): 16-20.
Zhang Ying-kun. Research progress and status of Delta parallel robot[J]. Machine Tool & Hydraulics, 2016, 44(21): 16-20.

[13] 许礼进, 刘有余, 刘阳. DELTA机器人逆运动学的分析与控制[J]. 制造业自动化, 2014, 36(17): 66-69.
Xu Li-jin, Liu You-yu, Liu Yang. Inverse-kinematic analysis and control of a DELTA robot[J]. Manufacturing Automation, 2014, 36(17): 66-69.

[14] 张颖, 赵建国, 沈鑫, 等. Delta并联机器人的运动学分析及虚拟样机仿真[J]. 计量与测试技术, 2019, 46(11): 20-24.
Zhang Ying, Zhao Jian-guo, Shen Xin, et al. The kinematic analysis and virtual prototyping simulation of a Delta parallel manipulator[J]. Metrology & Measurement Technique, 2019, 46(11): 20-24.

[15] 车林仙. 6-CPS正交并联机器人位置正解分析[J]. 农业机械学报, 2009, 40(11): 212-218.
Che Lin-xian. Forward positional analysis of 6-CPS orthogonal parallel manipulators[J]. Transactions of the Chinese Society of Agricultural Machinery, 2009, 40(11): 212-218.

[16] 罗玉坤, 周健, 李立君, 等. 基于旋量理论的4-R(SS)^{2}并联机器人正运动学分析[J]. 机床与液压, 2018, 46(3): 33-37.
Luo Yu-kun, Zhou Jian, Li Li-jun, et al. Direct kinematic analysis of 4-R(SS)^2 parallel robot based on screw theory[J]. Machine Tool & Hydraulics, 2018, 46(3): 33-37.