

基于改进灰狼算法的冗余机械臂最优轨迹规划

崔靖凯,王凯,范正奇,朱明超,徐振邦,刘伟东

引用本文:

崔靖凯, 王凯, 范正奇, 等. 基于改进灰狼算法的冗余机械臂最优轨迹规划[J]. 控制与决策, 2025, 40(5): 1457-1466.

在线阅读 View online: https://doi.org/10.13195/j.kzyjc.2024.1072

您可能感兴趣的其他文章

Articles you may be interested in

面向全局搜索的自适应领导者樽海鞘群算法

Global search-oriented adaptive leader salp swarm algorithm 控制与决策. 2021, 36(9): 2152-2160 https://doi.org/10.13195/j.kzyjc.2020.0090

基于地标特征和元学习方法推荐最适用优化算法

Recommending best suitable metaheuristic based on landmarking feature and meta-learning approach 控制与决策. 2021, 36(5): 1223-1231 https://doi.org/10.13195/j.kzyjc.2019.0993

基于Frenet坐标系的自动驾驶轨迹规划与优化算法

Trajectory planning and optimization algorithm for automated driving based on Frenet coordinate system 控制与决策. 2021, 36(4): 815–824 https://doi.org/10.13195/j.kzyjc.2019.0748

基于改进蚁群算法的水面无人艇路径规划

Path planning for unmanned surface vehicle based on improved ant colony algorithm 控制与决策. 2021, 36(4): 847-856 https://doi.org/10.13195/j.kzyjc.2019.0839

基于动态资源权重的多技能项目调度启发式算法

Dynamic resource priority-based heuristics for multi-skill resource constrained project scheduling problem 控制与决策. 2021, 36(10): 2553-2561 https://doi.org/10.13195/j.kzyjc.2020.0070

基于改进灰狼算法的冗余机械臂最优轨迹规划

崔靖凯1,王 凯1,范正奇1,朱明超2,徐振邦2,刘伟东11

(1. 中国汽车技术研究中心有限公司中央研究院,天津 300300;2. 中国科学院长春光学精密机械与物理研究所,长春 130033)

摘 要:针对冗余机械臂时间-冲击最优轨迹规划问题,提出一种基于改进灰狼算法的最优轨迹规划器.首先,为 了克服灰狼算法 (GWO) 开发与探索不平衡的局限性,提出基于强化学习的灰狼算法 (QLGWO) 及其多目标版 本 (MOQLGWO): QLGWO 使用 Q学习指导灰狼个体基于经验和奖励选择探索或开发动作,以实现算法局部与 全局搜索的自主平衡; MOQLGWO 引入存档和领导选择机制,在搜索衡量多种优化目标的帕累托最优解的同时, 引导搜索方向朝未被探索的区域拓展,以逼近全局最优. 然后,使用两段五阶多项式来构造机械臂的运动轨迹, 需 要搜索的解由运行时间以及中间点的关节位置、速度、加速度组成.最后,在 12 个基准函数上,将 QLGWO 与 GWO 以及其他 4 种先进的元启发式算法进行对比,并使用 MOQLGWO 求解 9 自由度冗余机械臂的时间-冲击 最优轨迹规划问题. 仿真和实验结果表明:所提出 QLGWO 可有效提高 GWO 的性能;最优轨迹规划器能够在满 足关节约束的前提下获得安全、光滑的时间-冲击最优轨迹,其运行时间小于 14 s,冲击处于 -0.25 rad/s³ ~ 0.15 rad/s³之间.

关键词: 冗余机械臂; 轨迹规划; 多目标优化; 元启发式; 灰狼算法; 强化学习

中图分类号: TP241 文献标志码: A

DOI: 10.13195/j.kzyjc.2024.1072

引用格式: 崔靖凯, 王凯, 范正奇, 等. 基于改进灰狼算法的冗余机械臂最优轨迹规划 [J]. 控制与决策, 2025, 40(5): 1457-1466.

Optimal trajectory planning of robotic manipulator based on improved grey wolf optimizer

CUI Jing-kai¹, WANG Kai¹, FAN Zheng-qi¹, ZHU Ming-chao², XU Zhen-bang², LIU Wei-dong^{1†}

(1. Central Research Institute, China Automotive Technology and Research Center Co., Ltd., Tianjin 300300, China; 2. Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun 130033, China)

Abstract: This paper introduces an optimal trajectory planner for the time-impact problem of redundant robotic manipulators, utilizing an enhanced grey wolf optimizer. Firstly, to address the limitations of the grey wolf optimizer (GWO) due to imbalanced exploitation and exploration, this paper introduces a reinforcement learning-based GWO (QLGWO) algorithm and its multi-objective version(MOQLGWO). In the QLGWO, Q-learning guides search agents to choose exploration or exploitation actions based on experience and rewards, which is beneficial for achieving autonomous balance between local and global search. In the MOQLGWO, the archive and leadership selection mechanisms are used to search for Pareto optimal solutions that consider multiple optimization objectives. In addition, these mechanisms guide the search direction towards unexplored regions, enabling the algorithm to approach the global optimum. Then, two five-order polynomial functions are used to construct motion trajectories. The solution that needs to be searched consists of travelling time and joint positions, velocities, and accelerations at intermediate points. Finally, on 12 benchmark functions, the QLGWO is compared with the GWO and four other state-of-the-art metaheuristic algorithms. Furthermore, the MOQLGWO is employed to solve the time-impact optimal trajectory planning problem for a nine-degree-of-freedom redundant robotic manipulator. Simulation and experimental results demonstrate that the proposed QLGWO significantly enhances the performance of the GWO; the optimal trajectory planner can achieve a safe and smooth time-impact optimal trajectory, while adhering to joint constraints. The travel time of the searched trajectory is less than 14 seconds, and the impact is between -0.25 rad/s^3 and 0.15 rad/s^3 .

Keywords: redundant robotic manipulator; trajectory planning; multi-objective optimization; meta-heuristic; grey wolf optimizer; reinforcement learning

收稿日期: 2024-09-05; 录用日期: 2024-11-28.

基金项目: 国家自然科学基金项目 (62173047).

责任编委: 訾斌.

[†]通信作者. E-mail: liuweidong@catarc.ac.cn.

0 引 言

冗余机械臂是自由度数大于完成任务所需最小 自由度数的机械臂. 冗余的自由度提升了机械臂的 灵活性和适应性, 但是也增加了结构的复杂性, 使得 机械臂的轨迹规划面临诸多挑战. 如运动学参数和 约束条件的增加提高了轨迹规划的计算复杂度; 同 一末端位置对应无限的关节组合, 要求规划的轨迹 具有更高的效率和安全性, 即冗余机械臂的轨迹规 划被纳入优化问题的范畴.

最优轨迹规划旨在寻找最优的关节轨迹,使得 机械臂从起始位置安全、平稳、准确地移动到终止位 置,以实现相应的生产任务.优化的目标可以是最小 化运行时间^[1]、最小化能量^[2]、最小化冲击^[3](加速度 的变化率)等.20世纪80年代以来,如概率路线图^[4]、 快速探索随机树^[5]等传统轨迹规划方法得到了广泛 应用.然而,对于冗余机械臂,传统方法存在全局收 敛性弱、计算复杂度高、鲁棒性差等局限性^[6].具体 而言,传统方法易陷入局部最优停滞,寻优精度低、 速度慢,依赖于机械臂的数学模型,无法设计通用的 轨迹规划器.此外,这些方法生成的轨迹长且粗糙, 使得机械臂难以跟踪.

为了克服上述缺点,基于遗传算法^[7]、粒子群算 法^[8]、灰狼算法 (GWO)^[9] 等元启发式算法的最优轨 迹规划方法应运而生.这些方法基于种群的进化或 迭代来搜索最优解,高度随机性赋予其强大的探索 能力,能够突破局部最优停滞;此外,它们生成的光 滑轨迹能够被机械臂轻易地跟踪和执行.近几年来, 诸多文献报告了元启发式方法在冗余机械臂轨迹规 划领域成功应用的案例. Zhang 等^[10]提出了一种改 进的粒子群算法,将其应用于7自由度冗余机械臂 的轨迹规划问题,通过单个适应度函数最小化机械 臂基座的扰动,然而,该方法无法应对多目标轨迹优 化的需求; Li 等^[11] 将粒子群算法与鲸鱼算法混合, 以优化7自由度冗余机械臂的运动轨迹.适应度函 数由基座的扰动、运动时间和可操作度组成,然而, 算法的混合将增加计算量,降低优化效率,Liu等^[12] 基于遗传算法和样条插值设计了轨迹规划器,以最 小化时间和冲击为目标,搜索了7自由度冗余机械 臂的最优轨迹.

本文针对一种9自由度冗余机械臂,设计基于 改进灰狼算法的最优轨迹规划器,以搜索关节空间 中点到点的时间-冲击最优轨迹,在保证机械臂运动 效率的同时,尽可能地减小关节抖动以提高安全性. 本文的主要内容如下:1)提出基于强化学习(Q学习) 的灰狼算法 (QLGWO), 以实现探索与开发的自主平衡; 2) 引入存档和领导选择机制, 提出 QLGWO 的多目标版本 (MOQLGWO), 以应对多目标优化的需求; 3) 设计元启发式最优轨迹规划器, 使用分段插值法构建机械臂的运动方程, 并引入惩罚函数来保证解的可行性. 最后, 在 12 个基准函数上, 将所提出QLGWO 与 GWO 以及其他 4 种先进的元启发式算法进行对比测试来验证 QLGWO 的优越性; 在 9 自由度冗余机械臂上, 对所提出 MOQLGWO 和轨迹规划器进行仿真和实验, 验证所提出方法的有效性.

1 基于强化学习的灰狼算法

1.1 灰狼算法概述

GWO 是由 Mirjalili 等^[9]于 2014 年提出的一种 元启发式优化算法. 在 GWO 的每次迭代中, 种群中 的最优解、次优解和第 3 优的解被分别设置为领导 者 α 、 β 和 δ . 底层灰狼 ω 的位置将根据 α 、 β 和 δ 按照 下式进行更新:

 $X(t+1) = (X_1 + X_2 + X_3)/3;$ (1)

$$X_1 = X_{\alpha} - A_1 \cdot |C_1 \cdot X_{\alpha} - X|;$$
 (2)

$$X_{2} = X_{\beta} - A_{2} \cdot |C_{2} \cdot X_{\beta} - X|; \qquad (3)$$

$$X_3 = X_\delta - A_3 \cdot |C_3 \cdot X_\delta - X|; \tag{4}$$

$$A_i = 2a \cdot r_{1i} - a, \ i = 1, 2, 3; \tag{5}$$

$$a = 2 - 2t / \text{MaxIter}; \tag{6}$$

$$C_i = 2 \cdot r_{2i}, \ i = 1, 2, 3.$$
 (7)

其中: t为迭代次数, X(t+1)为灰狼个体在第t次迭 代后的新位置, X_1 、 X_2 和 X_3 分别为由 α 、 β 和 δ 引起 的位移, X_{α} 、 X_{β} 和 X_{δ} 分别为 α 、 β 和 δ 的位置, a、 A_i 和 $C_i(i=1,2,3)$ 为算法参数, r_{1i} 和 $r_{2i}(i=1,2,3)$ 为位于区间[0,1]的随机数, MaxIter 为最大迭代次 数.

1.2 Q学习概述

Q学习 (QL)^[13] 是一种不基于模型但是基于价值的强化学习算法,在每次迭代中,QL 将根据如下所示的贝尔曼 (Bellman) 方程更新每个个体的Q表:

$$Q_{t+1}(s_t, a_t) = Q_t(s_t, a_t) + \lambda [R_t(s_t, a_t) + \gamma \max Q_t(s_{t+1}; \cdot) - Q_t(s_t, a_t)].$$
(8)

其中: $Q_t(s_t, a_t)$ 为第t次迭代时, 智能体在状态 s_t 下 执行动作 a_t 的价值; $Q_{t+1}(s_t, a_t)$ 为第t + 1次迭代时, 上述价值的更新值; $R_t(s_t, a_t)$ 为第t次迭代时, 智能 体在状态 s_t 下执行动作 a_t 所获得的奖励或惩罚; max $Q_t(s_{t+1}, :)$ 为第t次迭代时, 预估的下一个状态 s_{t+1} 对应的所有动作的最大价值; λ 和 γ 分别为学习

1.3 改进算法设计

GWO 的搜索过程分为开发和探索两个阶段,且 搜索阶段的切换是由参数A控制的. 然而,这种切换 是静态的,即遵循预先设计的规律. 在面对不同的优 化问题时,上述规律难以保证算法开发与探索的最 优平衡.

为了克服上述局限,本文将 QL 引入 GWO,基 于经验和奖励自主决策开发或探索动作,以提升 算法性能.若将 GWO 的搜索阶段视为与 QL 交互的 环境,则该环境具有两个状态:开发状态和探索状态. 在 QLGWO 中,环境的状态*s*_i按照下述规则定义, *A*_i按照式 (5) 计算, |·|为绝对值:

$$1) \stackrel{\sum_{i=1}^{3} |A_i|}{3} > 1, \quad ||s_t = \frac{1}{2} |S_t| = \frac{1}{2} |A_i|$$

$$2) \stackrel{\sum_{i=1}^{3} |A_i|}{3} \leq 1, \quad ||s_t| = \frac{1}{2} |S_t| = \frac{1}{2}$$

QLGWO 的个体在每个状态下可能执行两种动作:开发动作和探索动作.其中:开发动作旨在使得智能体向最优解收敛,以实现局部搜索;探索动作旨在使得智能体分散到新的搜索空间,以实现全局搜索.在 QLGWO 中,若智能体执行开发动作,则灰狼个体的新位置*X*(*t* + 1)由下式更新:

 $X(t+1) = (X_1 + X_2 + X_3)/3.$ (9) 式 (9) 与 (1) 所述 GWO 的种群更新方程是一致的, 其各变量的含义可见第 1.1 节的描述. 现有的研究表 明, 该种群更新方式具有卓越的收敛性.

相反地,若智能体执行探索动作,则灰狼个体的 新位置*X*(*t* + 1)由下式更新:

$$X(t+1) = X_{R_1} - X_{R_2}.$$
 (10)

其中: t为当前迭代次数; X_{R_1} 和 X_{R_2} 为从种群中随机选择的两个个体的位置, 且 $X_{R_1} \neq X_{R_2}$.式 (10) 模拟了自然界中部分灰狼因食物匮乏而被迫迁移的现象. 这种随机分散策略能够使得部分灰狼个体分散到新的搜索空间, 这为算法突破局部最优、搜索更好的解提供了机会.

综上所述, QLGWO 的每个智能体具有以下 4 种可以选择的决策: 1) 在探索状态执行探索动作; 2) 在探索状态执行开发动作; 3) 在开发状态执行探 索动作; 4) 在开发状态执行开发动作. 每种决策均会 受到环境反馈的奖励或惩罚, 奖惩由执行动作后获 得的新个体*X*(*t*+1)和执行动作前所持有的旧个体 *X*(*t*)的适应度*f*来决定.对于最小化优化问题(适应 度值越小,解越好),奖惩规则设置如下所示:

1) 若 $f_{X(t+1)} < f_{X(t)}$, 则奖励值R = 1;

2) 否则, 惩罚值R = -1.

在 QLGWO 的贝尔曼方程中, 折扣因子被设置 为一个较高的值, 即 γ =0.8, 使得智能体对未来奖励 和长期收益足够重视. QLGWO 的学习率 λ 被设置为 线性递减的, 其计算方式如下所示:

$$\lambda(t) = 1 - t / \text{MaxIter.}$$
(11)

其中: *t*为当前迭代次数, MaxIter 为最大迭代次数. 随着迭代次数*t*的增加, 学习率λ从1到0线性递减. 这意味着在迭代的初期, 智能体的学习速度很快, 对 原本缺乏的经验依赖较少; 而在迭代的末期, 智能体 放缓学习速度, 更多地利用已有经验, 以保证算法的 稳定性和收敛性.

综上所述,所提出 QLGWO 的流程如图 1 所示. 在 QLGWO 中,灰狼个体通过与环境动态交互来获 得经验和奖励,以做出最优决策.当环境变化时,上 述决策会随之变化.这使得 QLGWO 具有更好的动



态性能,能够根据具体的需求在开发或探索中切换, 有利于寻找更好的解.

1.4 多目标算法设计

本节将存档和领导选择机制引入 QLGWO, 以 开发其多目标版本 (MOQLGWO).存档是一种存储 单元,用于保存和检索目前为止搜索的帕累托最优 解^[14].它利用网格机制实时评估搜索空间的拥挤程 度.在 MOQLGWO 中,设置网格膨胀系数*d*=0.1, 网格大小系数*n*=8,以保证网格的精细度.

存档具有最大成员数. 当存档已满时, 启动如 下所示的轮盘赌法, 在搜索空间较拥挤的网格删除 旧解:

$$p(i) = c^{\xi} / \sum_{i=1}^{I} c^{\xi},$$
 (12)

$$q(i) = \sum_{j=1}^{i} p(j).$$
 (13)

其中: p(i)为第i个网格的选择概率; q(i)为第i个网格的累积概率; c为记录网格中解的个数的数组; I为 含有解的网格数量; ξ 为一个常参数, 本文设置 $\xi=2$, 使得越拥挤的网格, 被选择的概率越高.

领导选择机制则将上述轮盘赌法的参数ξ设置 为-4,即在搜索空间中较不拥挤的区域选择 3 个帕 累托最优解,将其设置为α、β和δ.该机制有利于领 导者引导其他灰狼个体向搜索空间中探索程度较低 的区域移动,以增强算法的全局搜索能力.

综上所述,所提出 MOQLGWO 的流程如图 2



所示.

2 时间-冲击最优轨迹规划器

2.1 运动方程构造

针对机械臂关节空间中点到点的时间-冲击最 优轨迹规划问题,本节将机械臂每个关节的运动轨 迹分割为两段,每段均采用五阶多项式构造运动方 程,以获得连续的位置、速度和加速度曲线.

构造的机械臂单个关节的运动轨迹如图 3 所示. 图 3 中: 起始点和终止点的关节位置、速度、加速度 θ_i 、 $\dot{\theta}_i$ 、 $\ddot{\theta}_i$ 和 θ_f 、 $\dot{\theta}_f$ 、 $\ddot{\theta}_f$ 是已知的,它们与机械臂的起 始和终止构型有关;中间点的关节位置、速度、加速 度 θ_m 、 $\dot{\theta}_m$ 、 $\ddot{\theta}_m$ 和两段轨迹的运行时间 t_1 、 t_2 是未知 的,它们是轨迹规划器搜索的变量.



图3 单个关节的运动轨迹

两段运动轨迹遵循如下约束:1) 起始点和终止 点的关节位置是已知的常数,关节速度、加速度为0; 2) 两段轨迹在连接处(中间点)具有相同的位置、速 度和加速度值.

结合以上两条约束,可构造如下运动方程:

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & t_{1} & t_{1}^{2} & t_{1}^{3} & t_{1}^{4} & t_{1}^{5} \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 2t_{1} & 3t_{1}^{2} & 4t_{1}^{3} & 5t_{1}^{4} \\ 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 2 & 6t_{1} & 12t_{1}^{2} & 20t_{1}^{3} \end{bmatrix} \begin{bmatrix} a_{0} \\ a_{1} \\ a_{2} \\ a_{3} \\ a_{4} \\ a_{5} \end{bmatrix} = \begin{bmatrix} \theta_{i} & \theta_{m} & \dot{\theta}_{i} & \dot{\theta}_{m} & \ddot{\theta}_{i} & \ddot{\theta}_{m} \end{bmatrix}^{\mathrm{T}}, \qquad (14)$$
$$\begin{bmatrix} 1 & t_{1} & t_{1}^{2} & t_{1}^{3} & t_{1}^{4} & t_{1}^{5} \\ 1 & t_{s} & t_{s}^{2} & t_{s}^{3} & t_{s}^{4} & t_{s}^{5} \\ 0 & 1 & 2t_{1} & 3t_{1}^{2} & 4t_{1}^{3} & 5t_{1}^{4} \\ 0 & 1 & 2t_{s} & 3t_{s}^{2} & 4t_{s}^{3} & 5t_{s}^{4} \\ 0 & 0 & 2 & 6t_{1} & 12t_{1}^{2} & 20t_{1}^{3} \\ 0 & 0 & 2 & 6t_{s} & 12t_{s}^{2} & 20t_{s}^{3} \end{bmatrix} \begin{bmatrix} b_{0} \\ b_{1} \\ b_{2} \\ b_{3} \\ b_{4} \\ b_{5} \end{bmatrix} = \begin{bmatrix} \theta_{m} & \theta_{f} & \dot{\theta}_{m} & \dot{\theta}_{f} & \ddot{\theta}_{m} & \ddot{\theta}_{f} \end{bmatrix}^{\mathrm{T}}. \qquad (15)$$

其中: $t_s = t_1 + t_2$ 为两段轨迹的运行时间的和, $[a_0 \ a_1 \ a_2 \ a_3 \ a_4 \ a_5]^{T} 和 [b_0 \ b_1 \ b_2 \ b_3 \ b_4 \ b_5]^{T}$ 为决定轨迹形状的参数.

2.2 适应度函数设计

时间-冲击最优轨迹规划有两个优化目标,即最小化运行时间和最小化冲击.设计如下适应度函数

$$F(X) = \begin{cases} f_1 = t_1 + t_2, \\ f_2 = \max(|\ddot{\theta}_1|, |\ddot{\theta}_2|, \dots, |\ddot{\theta}_n|). \end{cases}$$
(16)

其中: *X*为该优化问题的解; f_1 为运行时间的适应度 函数 (单位: s); f_2 为冲击的适应度函数 (单位: rad/s³); t_1 和 t_2 分别为第 1 段和第 2 段轨迹的运行时间; *n*为 关节数量; $\tilde{\theta}_j(j=1,2,\ldots,n)$ 为关节*j*从起始点运动 至终止点的冲击, 该变量可由下式计算:

$$\ddot{\theta}_{i,m}(t) = 6a_3 + 24a_4t + 60a_5t^2, \tag{17}$$

$$\ddot{\theta}_{m,f}(t) = 6b_3 + 24b_4t + 60b_5t^2, \tag{18}$$

这里 $\ddot{\theta}_{i,m}(t)$ 和 $\ddot{\theta}_{m,f}(t)$ 分别为该关节从起始点运动 到中间点,以及从中间点运动到终止点的冲击曲线.

机械臂在运动过程中,应满足关节位置、速度和 加速度的约束.为了保证轨迹规划器搜索的最优解 均是可行的,本文采用文献 [15] 的约束处理方法来 修改违反约束的不可行解的适应度值.含惩罚函数 的适应度计算方法如下所示:

$$F(x) = \begin{cases} f(x), \ c_i(x) \ge 0, \ \forall i = 1, 2, \dots, I; \\ f_{\max} + \sum_{i=1}^{I} C_i(x), \ \text{otherwise}; \end{cases}$$
(19)

$$C_i(x) = \begin{cases} 0, \ c_i(x) \ge 0; \\ |c_i(x)|, \ \text{otherwise.} \end{cases}$$
(20)

其中: x为解, f_{max} 为种群中最差的个体的适应度值, $c_i(x)$ 为第i个约束. 若 $c_i(x) \ge 0$,则解x没有违反约束; 否则,解x违反约束,此时, $|c_i(x)|$ 越大,违反约束的 程度越大.由式 (19)和 (20),当解不违反任何约束 时,其适应度不变;否则,根据其违反约束的数量和 程度以及当前种群中最差的解修改其适应度,使得 不可行解失去竞争最优解的机会.

3 仿真和实验

3.1 基于基准函数的性能分析

为了验证 QLGWO 的优越性,本节基于如表 1 所示的 12 个基准函数^[9,16-17],对 QLGWO 进行寻优测 试,并将结果与 GWO、两种 GWO 的变体 (HGWO^[18] 和 IGWO^[19]),以及两种先进的元启发式算法 (APSO^[20] 和 SADE^[21])进行对比.上述算法的参数设置均与原 文献相同,如表 2 所示.

基准函数是检测元启发式算法性能的标准工具. 表1中: F₁ ~ F₆为单峰函数,只有一个全局最小值, 用于检测算法的开发能力; F₇ ~ F₁₂为多峰函数,具 有多个局部极小值,用于检测算法的探索能力.上述 函数的理论最小值均为0,维度n=30.

仿真测试是在一台配备英特尔 i7 处理器和 8 GB

表1 基准函数

函数	搜索范围
$F_1(x) = \sum_{i=1}^n x_i^2$	[-100, 100]
n n	

$$F_{2}(x) = \sum_{i=1} |x_{i}| + \prod_{i=1} |x_{i}|$$
[-10, 10]

$$F_3(x) = \sum_{i=1}^{3} \left(\sum_{j=1}^{3} x_j\right)^2$$
[-100, 100]

$$F_{5}(x) = \sum_{i=1}^{n} ix_{i}^{2}$$
[-10, 10]

$$F_{6}(x) = \sum_{i=1}^{n} ix_{i}^{4} + \text{random}[0, 1) \qquad [-1.28, 1.28]$$

$$F_7(x) = \sum_{i=1}^n [x_i^2 - 10\cos(2\pi x_i) + 10]$$
 [-5.12, 5.12]

$$F_8(x) = \frac{1}{4\,000} \sum_{i=1}^n x_i^2 - \prod_{i=1}^n \cos\left(\frac{x_i}{\sqrt{i}}\right) + 1 \qquad [-600, 600]$$

$$F_9(x) = \sum_{i=1}^{n} |x_i \sin(x_i) + 0.1x_i|$$
 [-10, 10]

$$F_{10}(x) = 0.1n - \left(0.1\sum_{i=1}^{n}\cos(5\pi x_i) - \sum_{i=1}^{n}x_i^2\right)$$
[-1, 1]

$$F_{11}(x) = \sum_{i=1}^{n} x_i^2 + \left(\sum_{i=1}^{n} 0.5ix_i\right)^2 + \left(\sum_{i=1}^{n} 0.5ix_i\right)^4 \qquad [-5, 10]$$

$$F_{12}(x) = -20\exp\left(-0.2\sqrt{\frac{1}{n}\sum_{i=1}^{n}x_{i}^{2}}\right) - \exp\left(\frac{1}{n}\sum_{i=1}^{n}\cos(2\pi x_{i})\right) + 20 + e$$
[-32, 32]

表2 算法的参数设置

算法	全称	参数设置
GWO	灰狼算法	$a \in [2,0]$
HGWO	混合灰狼算法	$a \in [2,0], W = 1, C_r \in [0,0.2]$
IGWO	改良灰狼算法	$a\in [2,0],r_{\rm 1,2,3}\in [0,1]$
APSO	自适应粒子群优化	$c_{\rm 1,2}=1.2,W_{\rm 1}=0.9,W_{\rm 2}=0.7$
SADE	自适应差分进化	$\tau_{\rm 1,2}=0.1,F_i=0.1,F_u=0.9$

内存的计算机上进行的,使用同一软件编写和运行 算法.为了保证比较的公平性,对于所有算法,种群 大小均被设置为 30,最大迭代次数均被设置为 500, 运行的停止条件是迭代次数达到最大值.对于每个 基准函数,所有算法独立运行 10 次.

各算法在各基准函数上的收敛曲线如图 4 和 图 5 所示,该曲线由 10 次优化结果的平均值绘制. 收敛曲线越逼近于 0,算法的寻优性能越好.显然, QLGWO 在 12 个基准函数上均获得了最好的结果. 在函数*F*₁₀上,IGWO 和 GWO 也获得了与 QLGWO 相同的结果,但是 QLGWO 的收敛速度更快.表 3 为 各算法 10 次运行结果的标准差 "std" 和最大值 "max"





(表示最差结果), 最佳结果用粗体表示. 显然, QLGWO 在所有函数上获得了最小的标准差. 即使在最差的 情况下, QLGWO 获得的优化结果在所有算法中也 是最好的. 在 *F*₁₀上, GWO 和 IGWO 与 QLGWO 并 列第 1. 这表明 QLGWO 中的强化学习机制有效地 增强了 GWO 的开发和探索能力, 并成功实现了两 者的自主平衡, 进而促使 QLGWO 具有更高的收敛 精度、更快的收敛速度和更稳定的优化性能.

除收敛精度,计算时间也是评价算法的重要指标.在相同的仿真设置下,本节对上述6种算法在12个基准函数的平均计算时间(单位:s)进行了测试和统计,结果如图6所示.图6中:GWO、HGWO、IGWO、APSO、SADE、QLGWO的计算时间分别为0.1376s、0.2328s、0.6091s、0.1931s、0.2305s、

実る	- 笛注 10 '	次法纪	护生田的	标准主	ミわまう	一估
125	- 开ム 10 /	人巴1	リッコンエリン	1小庄石	ニイド・取ノ	く回し

算法	-	F_1	F_2	F_3	F_4	F_5	F_6	F_7	F_8	F_9	F_{10}	F_{11}	F_{12}
QLGWO std max	std	6.7e-93	1.9e-51	0	0	0	4.4e-05	0	0	0	0	0	0
	max	2.1e-92	6.1e-51	0	0	0	1.3e-04	0	0	0	0	0	8.8e-16
GWO std max	std	6.8e-28	8.7e-17	8.5e-06	4.7e-07	7.4e-28	1.3e-03	4.0e+00	1.0e-02	5.0e-04	0	4.3e-07	1.7e-14
	max	2.3e-27	3.1e-16	2.1e-05	1.4e-06	2.2e-27	5.4e-03	1.2e+01	2.5e-02	1.4e-03	0	1.3e-06	1.1e-13
HGWO std max	std	5.2e-14	1.4e-09	2.6e-02	1.5e+00	2.4e-15	6.4e-03	2.9e+01	1.1e-02	3.3e+00	7.6e-01	2.3e-06	1.4e+00
	max	1.6e-13	4.2e-09	7.2e-02	5.7e+00	8.9e-15	2.6e-02	1.3e+02	2.7e-02	1.7e+01	2.9e+00	6.8e-06	2.9e+00
IGWO	std	2.4e-28	8.1e-18	9.5e-04	2.1e-05	1.7e-28	1.0e-03	4.4e+01	1.9e-02	2.6e-04	0	1.6e-07	1.0e-14
	max	7.8e-28	2.7e-17	2.4e-03	7.3e-05	5.7e-28	4.4e-03	1.5e+02	6.1e-02	8.8e-04	0	5.1e-07	7.9e-14
APSO st m	std	2.8e-02	6.6e-01	1.8e+02	1.0e+00	1.5e-01	9.6e-02	1.6e+01	2.2e+00	2.9e-01	1.6e-01	1.6e+01	4.0e-01
	max	9.2e-02	2.3e+00	6.9e+02	4.8e+00	5.4e-01	3.8e-01	7.8e+01	9.2e+00	9.5e-01	1.1e+00	6.3e+01	2.0e+00
SADE std ma	std	4.9e-02	7.0e-03	5.3e+03	1.8e+01	4.6e-02	2.8e-02	2.5e+01	1.9e-01	6.9e-01	1.4e-01	5.6e+01	2.1e+00
	max	1.2e-01	2.3e-02	1.5e+04	6.3e+01	1.4e-01	1.2e-01	1.3e+02	6.4e-01	2.2e+00	4.4e-01	2.2e+02	7.0e+00



0.1415 s. 所提出 QLGWO 因引入了强化学习机制, 计 算耗时比 GWO 高 2.83%. 考虑到 QLGWO 对 GWO 性能的显著提升, 这些时间消耗是可以接受的. 与其 他 4 种算法相比, QLGWO 在计算速度方面更具优 势.

3.2 时间-冲击最优轨迹规划

本节用于测试的 9 自由度冗余机械臂和关节坐标系^[22] 如图 7 所示,机械臂的关节约束如表 4 所示.



图7 9 自由度冗余机械臂

表4 机械臂的关节约束

		-	-	
关节	位置边界/(°)	速度边界/(°/s)	加速度边界/(°/s²)	
1	±180	±30	±1 000	
2	± 90	±30	$\pm 1 000$	
3	± 180	±30	$\pm 1 000$	
4	± 90	±30	$\pm 1 000$	
5	± 180	±30	$\pm 1 000$	
6	± 90	±30	$\pm 1 000$	
7	±180	±30	$\pm 1 000$	
8	± 90	±30	$\pm 1 000$	
9	±180	±30	± 1000	

在本节的仿真中,随机选择两个构型作为机械 臂轨迹规划的起始点和终止点.其中:起始点的关节 配置为 [0.1, 0.2, 0.5, -0.1, 0.2, 0.4, 0.1, -0.2, 0.1] (单 位: rad),终止点的关节配置为 [1.5, 1.3, 1.7, -1.5, 1.0, 1.2, 1.8, -1.8, 2.0] (单位: rad).

使用 MOQLGWO 在第 2 节提出的轨迹规划器 的框架中对上述轨迹规划问题进行求解. 算法的种 群大小、最大迭代次数和存档的最大成员数均被设 置为 100, 需要搜索的变量数为 29. MOQLGWO 的 收敛过程如图 8 所示. 显然, 随着迭代次数的增加, 灰狼个体成功收敛至由目标函数 *f*₁和 *f*₂构成的凹曲 线 (帕累托前沿).

基于图 8 的存档中较为折衷地衡量 $f_1 和 f_2$ 的解 A, 计算所得到关节运动轨迹的位置、速度、加速度 和冲击曲线如图 9 所示. 由图 9 可知, 各关节在满足 约束的前提下, 准确地完成了点到点的运动, 其位 置、速度和加速度曲线连续且光滑, 运行时间小于 14 s, 冲 击 处 于 $-0.25 \text{ rad/s}^3 \sim 0.15 \text{ rad/s}^3 \ge$ 间. 图 10 为机械臂执行上述运动轨迹所呈现的构型变 化, 以及由正运动学计算的机械臂末端的运动轨迹. 图 10 中, 机械臂的末端轨迹连续且光滑, 表明解 A能够保证机械臂末端运动的安全性和平稳性.



在存档的其他解处也可得到相同的结论,由于 篇幅限制不再赘述.在实际应用中,可根据需求在存 档中选择最优解,以图 8(d)所示的存档为例,若希望 获得更小的运行时间,则可以选择解B;若希望获得 更小的冲击,则可以选择解C.

对解A的运动轨迹在如图 11 所示的实验平台^[22] 上进行实验验证.对于每个关节,将规划的位置曲线 作为控制指令通过网线输入给驱动器,采样时间为 5 ms.实验中,机械臂运动过程的照片如图 12 所示.



图 12 中, 左上角的数字顺序表示运动的顺序. 图 13 为实验中由上位机测量的关节运动轨迹和轨迹跟踪 误差. 其运动轨迹与仿真结果的误差大致处于 [10⁻⁶, 10⁻³] (单位: rad) 之间, 表明机械臂能够准确地跟踪 最优轨迹. 综上, 轨迹规划器搜索的时间-冲击最优 轨迹能够使得机械臂安全、平稳、准确地完成指定的 点到点运动.





图11 轨迹规划的实验平台



图12 轨迹规划实验

4 结 论

为了解决冗余机械臂的时间-冲击最优轨迹规 划问题,本文提出了一种基于改进灰狼算法的最优 轨迹规划器.首先,将强化学习(Q学习)机制引入了 灰狼算法 (GWO),提出了一种改进灰狼算法 (QLGWO)及其多目标版本 (MOQLGWO).QLGWO 能够根据经验和奖励选取最优决策,自主平衡算法 的开发与探索;MOQLGWO中的存档和领导选择机 制能够搜索衡量多种目标函数的帕累托最优解.然 后,使用两段五阶多项式插值法构造机械臂各关节



图13 实验测量的关节运动轨迹和跟踪误差

的点到点运动轨迹,以最小化运行时间和冲击为目标,提出了元启发式最优轨迹规划器,旨在搜索连续、光滑的运动轨迹,并保证机械臂运动的高效性和 平稳性.最后,使用基准函数对QLGWO进行了性能 测试,并在9自由度冗余机械臂上对轨迹规划器进 行了仿真和实验验证.实验结果表明:QLGWO有效 地增强了GWO的性能,具有很强的优越性;轨迹规 划器成功地搜索到高效、平稳的运动轨迹,其运行时 间小于14 s,冲击处于 $-0.25 \text{ rad/s}^3 \sim 0.15 \text{ rad/s}^3$ 之间.

参考文献 (References)

 张韵悦, 孙志毅, 孙前来, 等. 基于强化学习的挖掘机 时间最优轨迹规划[J]. 控制与决策, 2024, 39(5): 1433-1440.

(Zhang Y Y, Sun Z Y, Sun Q L, et al. Time optimal trajectory planning of excavator based on deep reinforcement learning[J]. Control and Decision, 2024, 39(5): 1433-1440.)

- [2] Nonoyama K, Liu Z A, Fujiwara T, et al. Energyefficient robot configuration and motion planning using genetic algorithm and particle swarm optimization[J]. Energies, 2022, 15(6): 2074.
- [3] Liu J J, Wang H, Li X, et al. Robotic arm trajectory optimization based on multiverse algorithm[J]. Mathematical Biosciences and Engineering, 2023, 20(2): 2776-2792.
- [4] Zhao S Y, Zhang J Z, He C K, et al. Collision-free emergency planning and control methods for CAVs considering intentions of surrounding vehicles[J]. ISA

Transactions, 2023, 136: 535-547.

- [5] 陈丹, 谭钦, 徐哲壮. 基于采样点优化 RRT 算法的机 械臂路径规划[J]. 控制与决策, 2024, 39(8): 2597-2604.
 (Chen D, Tan Q, Xu Z Z. Robotic arm path planning based on sampling point optimization RRT algorithm[J]. Control and Decision, 2024, 39(8): 2597-2604.)
- [6] Zerrouki N, Goléa N, Benoudjit N. Particle swarm optimization of non uniform rational B-splines for robot manipulators path planning[J]. Periodica Polytechnica Electrical Engineering and Computer Science, 2017, 61(4): 337-349.
- [7] Zhang J W, Cheng L, Liu C, et al. Cost-aware scheduling systems for real-time workflows in cloud: An approach based on genetic algorithm and deep reinforcement learning[J]. Expert Systems with Applications, 2023, 234: 120972.
- [8] 于海波,朱秦娜,康丽,等.带偏向性轮盘赌的多算子协同粒子群优化算法[J]. 控制与决策, 2024, 39(4): 1167-1176.
 (Yu H B, Zhu Q N, Kang L, et al. A multi-operator collaborative particle swarm optimization algorithm with

biased roulette[J]. Control and Decision, 2024, 39(4): 1167-1176.)

- [9] Mirjalili S, Mirjalili S M, Lewis A. Grey wolf optimizer[J]. Advances in Engineering Software, 2014, 69: 46-61.
- [10] Zhang J X, Yang F C, Nie G H, et al. Optimal trajectory planning for minimizing base disturbance of a redundant space robot with IQPSO[J]. Journal of Electrical and Computer Engineering, 2022: 3398810.
- [11] Li R, Liu M, Teutsch J, et al. Constraint trajectory planning for redundant space robot[J]. Neural Computing and Applications, 2023, 35(34): 24243-24258.
- [12] Liu Y, Yi W M, Feng Z Q, et al. Design and motion planning of a 7-DOF assembly robot with heavy load in spacecraft module[J]. Robotics and Computer-Integrated Manufacturing, 2024, 86: 102645.
- [13] Watkins C J C H, Dayan P. Q-learning[J]. Machine Learning, 1992, 8(3): 279-292.
- [14] Mirjalili S, Saremi S, Mirjalili S M, et al. Multiobjective grey wolf optimizer: A novel algorithm for multi-criterion optimization[J]. Expert Systems with Applications, 2016, 47: 106-119.
- [15] Deb K. An efficient constraint handling method for genetic algorithms[J]. Computer Methods in Applied Mechanics and Engineering, 2000, 186(2/3/4): 311-338.
- [16] Gao W F, Liu S Y, Huang L L. A novel artificial bee colony algorithm based on modified search equation and orthogonal learning[J]. IEEE Transactions on

Cybernetics, 2013, 43(3): 1011-1024.

- [17] Long W, Jiao J J, Liang X M, et al. An explorationenhanced grey wolf optimizer to solve high-dimensional numerical optimization[J]. Engineering Applications of Artificial Intelligence, 2018, 68: 63-80.
- [18] Jayabarathi T, Raghunathan T, Adarsh B R, et al. Economic dispatch using hybrid grey wolf optimizer[J]. Energy, 2016, 111: 630-641.
- [19] Nadimi-Shahraki M H, Taghian S, Mirjalili S. An improved grey wolf optimizer for solving engineering problems[J]. Expert Systems with Applications, 2021, 166: 113917.
- [20] Deng H, Xie C. An improved particle swarm optimization algorithm for inverse kinematics solution of multi-DOF serial robotic manipulators[J]. Soft Computing, 2021, 25(21): 13695-13708.
- [21] Brest J, Greiner S, Boskovic B, et al. Self-adapting control parameters in differential evolution: A comparative study on numerical benchmark problems[J]. IEEE Transactions on Evolutionary Computation, 2006, 10(6): 646-657.
- [22] 崔靖凯,周宇飞,贺顺锋,等.基于改进灰狼算法的冗余机械臂轨迹跟踪与避障[J].光学精密工程,2023,31(24):3595-3605.
 (Cui J K, Zhou Y F, He S F, et al. Trajectory tracking and obstacle avoidance of a redundant robotic manipulator based on the improved grey wolf optimizer[J]. Optics and Precision Engineering, 2023, 31(24): 3595-3605.)

作者简介

崔靖凯 (1997-), 男, 工程师, 博士, 主要研究方向为元 启发式优化算法、机器人运动规划与控制, E-mail: ouccjk@ 163.com;

王凯 (1995-), 男, 工程师, 主要研究方向为机器人结构 设计与控制、仿生机械装备开发, E-mail: wangkai2021@ catarc.ac.cn;

范正奇 (1998-), 男, 助理工程师, 主要研究方向为仿生 机械装备开发, E-mail: fanzhengqi@catarc.ac.cn;

朱明超 (1980-), 男, 研究员, 博士, 主要研究方向为机器人运动规划与控制、深度学习与图像处理, E-mail: mingchaozhu@gmail.com;

徐振邦(1982-), 男, 研究员, 博士, 主要研究方向为空间智能机器人、空间遥感器设计、振动控制技术, E-mail: xuzhenbang@ciomp.ac.cn;

刘伟东(1990-),男,高级工程师,主要研究方向为元启 发式优化算法、仿生机械装备开发,E-mail: liuweidong@ catarc.ac.cn.