自动化学报
ACTA AUTOMATICA SINICA

基于动态性能均衡的无人机集群元学习协同避障控制

苗金钊 刘金良 田恩刚 彭晨 黄廷文

Dynamic Performance Balancing-based Meta-learning Cooperative Obstacle Avoidance Control for Unmanned Aerial Vehicle Swarms

MIAO Jin-Zhao, LIU Jin-Liang, TIAN En-Gang, PENG Chen, HUANG Ting-Wen

在线阅读 View online: https://doi.org/10.16383/j.aas.c250695

您可能感兴趣的其他文章

性能约束下的四旋翼无人机协同吊挂系统分布式避碰跟踪控制

Distributed Collision Avoidance Tracking Control for Quadrotor Cooperative Suspension System Under Performance Constraints

自动化学报. 2024, 50(12): 2392−2406 https://doi.org/10.16383/j.aas.c240349

性能函数引导的无人机集群深度强化学习控制方法

Performance Function-guided Deep Reinforcement Learning Control for UAV Swarm

自动化学报. 2025, 51(5): 905−916 https://doi.org/10.16383/j.aas.c240519

基于深度强化学习的无人机虚拟管道视觉避障

Virtual Tube Visual Obstacle Avoidance for UAV Based on Deep Reinforcement Learning

自动化学报. 2024, 50(11): 2245−2258 https://doi.org/10.16383/j.aas.c230728

基于元学习的双目深度估计在线适应算法

Online Adaptation Through Meta-learning for Stereo Depth Estimation

自动化学报. 2023, 49(7): 1446−1455 https://doi.org/10.16383/j.aas.c200286

基于鹰群环伺行为的无人机集群关联跟踪控制

Association Tracking Control of UAV Swarms Based on Harris Hawks´ Surrounding Behavior

自动化学报. 2026, 52(5): 980−991 https://doi.org/10.16383/j.aas.c250518

基于社会化协同的无人机集群目标合围控制

Target Enclosing Control of Unmanned Aerial Vehicle Swarm Based on Socialized Collaboration

自动化学报. 2026, 52(2): 284−295 https://doi.org/10.16383/j.aas.c250415

基于动态性能均衡的无人机集群元学习协同避障控制

苗金钊 [1] 刘金良 [1] 田恩刚 [2] 彭 晨 [3] 黄廷文 [4]

摘 要 针对无人机集群在突发环境变化下易出现性能退化的问题, 提出一种基于动态性能均衡的元学习控制方法, 以实现集群系统高效、可靠的在线自适应协同控制. 首先, 基于无人机不确定动力学模型, 结合轨迹跟踪、集群协同与动态障碍物规避等多重约束, 构建一类能够有效抵御系统不确定性与外部扰动的自适应控制策略. 其次, 为解决瞬态响应与稳态精度之间的性能冲突, 并克服传统优化算法计算耗时长的局限, 构建元学习控制架构. 该方法通过离线训练习得具备强泛化能力的初始化参数, 使系统在线仅需利用少量历史数据和单步梯度更新, 即可实现毫秒级自适应并快速收敛至满足动态性能均衡要求的最优增益. 实验结果验证了该方法在动态障碍物及剧烈环境变化场景下的显著适应能力, 实现了兼具鲁棒性、安全性与高性能的协同控制.

关键词 无人机集群; 协同避障; 元学习; 自适应控制

引用格式 苗金钊, 刘金良, 田恩刚, 彭晨, 黄廷文. 基于动态性能均衡的无人机集群元学习协同避障控制. 自动化学报, 2026, **52**(10): 2261−2273

DOI 10.16383/j.aas.c250695 **CSTR** 32138.14.j.aas.c250695

Dynamic Performance Balancing-based Meta-learning Cooperative Obstacle Avoidance Control for Unmanned Aerial Vehicle Swarms

MIAO Jin-Zhao[1] LIU Jin-Liang[1] TIAN En-Gang[2] PENG Chen[3] HUANG Ting-Wen[4]

Abstract To address the issue of performance degradation in unmanned aerial vehicle (UAV) swarms under abrupt environmental changes, a meta-learning control method based on dynamic performance balancing is proposed to achieve efficient, reliable, and online adaptive cooperative control for swarm systems. First, based on the uncertain dynamic model of UAV and incorporating multiple constraints such as trajectory tracking, swarm cooperation, and dynamic obstacle avoidance, an adaptive control strategy is constructed to effectively counteract system uncertainties and external disturbances. Second, to resolve the performance conflict between transient response and steady-state accuracy, and to overcome the limitations of high computational cost in traditional optimization algorithms, a meta-learning control architecture is constructed. By learning initialization parameters with strong generalization capabilities through offline training, this method enables the system to achieve millisecond-level adaptation and rapidly converge to optimal gains that satisfy dynamic performance balancing requirements online, utilizing only a small amount of historical data and a single-step gradient update. Experimental results demonstrate the significant adaptability of this method in scenarios involving dynamic obstacles and drastic environmental changes, realizing cooperative control that balances robustness, safety, and high performance.

Keywords unmanned aerial vehicle swarms; cooperative obstacle avoidance; meta-learning; adaptive control

Citation Miao Jin-Zhao, Liu Jin-Liang, Tian En-Gang, Peng Chen, Huang Ting-Wen. Dynamic performance balancing-based meta-learning cooperative obstacle avoidance control for unmanned aerial vehicle swarms. *Acta Automatica Sinica*, 2026, **52**(10): 2261−2273

随着人工智能与机器人技术的快速发展, 无人机集群凭借其高机动性、强协作能力与天然冗余特性, 在协同运输、环境监测、搜索救援及区域覆盖等任务中展现出广阔的应用前景[1]. 要确保无人机集群在复杂、动态乃至未知环境中安全且高效地完成

收稿日期 2025-12-01 录用日期 2026-03-04
Manuscript received December 1, 2025; accepted March 4, 2026
国家自然科学基金 (62373252), 南京信息工程大学引进人才启动基金 (2024r063) 资助
Supported by National Natural Science Foundation of China (62373252) and Startup Foundation for Introducing Talent of Nanjing University of Information Science and Technology (2024r063)
本文责任编委 付俊杰
Recommended by Associate Editor FU Jun-Jie

1. 南京信息工程大学计算机学院 南京 210044 2. 上海理工大学光电信息与计算机工程学院 上海 200093 3. 上海大学机电工程与自动化学院 上海 200444 4. 深圳理工大学计算机科学与控制工程学院 深圳 518107
1. School of Computer Science, Nanjing University of Information Science and Technology, Nanjing 210044 2. School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai 200093 3. School of Mechatronic Engineering and Automation, Shanghai University, Shanghai 200444 4. Faculty of Computer Science and Control Engineering, Shenzhen University of Advanced Technology, Shenzhen 518107

任务, 设计先进的集群控制算法至关重要[2]. 然而, 实际应用中的集群系统不仅存在复杂的耦合非线性动力学和时变外界扰动, 还必须同时满足集群内部的防碰撞约束与外部环境中的安全避障要求, 这些因素使协同控制问题面临极大的挑战[3]. 因此, 开发一种兼具安全性、鲁棒性与自适应能力的集群控制方法, 已成为无人机在复杂环境下实现自主协同作业的核心科学问题. 为应对上述的挑战, 国内外研究者已探索并应用多种协同控制策略, 例如微分博弈方法、人工势场法以及领航–跟随架构等[4–6]. 这些方法在提升无人机集群的鲁棒性与协同性能方面取得了显著进展, 并推动了集群自主控制技术的持续发展.

为应对无人机集群在任务执行过程中可能产生的内部碰撞风险, 以及在复杂动态环境中对外部障碍物进行安全规避的需求, 大量研究致力于发展高效可靠的安全控制策略[7–9]. 其中, 一类基于人工势场或控制障碍函数的典型方法, 通过在控制律中嵌入势能引导项或可行性约束, 以对无人机运动施加安全限制, 从而在保持协同编队的同时实现有效避障与碰撞防护[10]. 例如, 文献 [11] 利用多智能体强化学习构建与障碍物相关的势场函数, 使智能体能够在未知环境中自主学习协同避障策略. 文献 [12] 通过引入动态势场与代价函数进行在线路径规划, 使无人系统在动态复杂环境下具备高效的障碍规避能力. 文献 [13] 将安全约束函数与势场式避碰机制相结合, 在通信拓扑变化及外界扰动存在的情况下仍能保持无人机间的安全间距与编队稳定性. 此外, 文献 [14] 将预设性能控制与障碍函数融合, 通过构造可行性安全约束实现编队任务过程中的无碰撞运行. 尽管上述控制方法在确保系统安全性方面取得了显著进展, 但在实际应用中仍面临若干关键挑战. 首先, 传统控制策略难以在瞬态响应与稳态性能之间实现有效权衡. 虽然可以通过调整控制增益改善某一性能指标, 但由于缺乏系统化的多目标协调机制, 在复杂任务环境下往往难以同时兼顾多项性能目标. 其次, 现有方法普遍缺乏对集群内部多智能体协同行为的整体性考量, 容易导致个体最优与集群协同之间出现冲突, 难以系统地处理瞬态性能、稳态性能与安全约束之间的耦合关系.

随着无人机集群在环境监测、智能交通与空地协同等任务中的广泛部署, 其在复杂动态环境下的协同控制与性能优化问题愈发受到关注[15]. 传统控制策略通常依赖单一性能指标进行关键参数设计, 或基于经验确定控制增益, 但这类方法难以在瞬态性能、稳态精度、安全性与能效等多个性能指标之间实现系统化权衡[16–18]. 在实际应用中, 这些性能指标往往相互耦合且具有内在冲突, 例如提升响应速度可能导致稳态误差增大, 而强化安全约束又可能引起能耗上升, 因此单目标优化方案难以满足复杂任务场景的需求[19]. 为解决上述矛盾, 近年来多目标优化方法被广泛引入无人机集群控制, 通过构造可表征性能冲突关系的评价模型, 为多性能指标之间的协同折中提供了系统化手段[20–22]. 例如, 文献 [22] 面向多无人机数据采集任务, 设计同时优化任务完成时间与能耗的多目标优化方法, 实现了时效性与能效之间的有效平衡. 文献 [23] 针对集群感知任务中的资源竞争问题, 利用多目标优化理论协调了参与者容忍度、预算约束与覆盖率之间的冲突关系. 文献 [24] 在不完全干扰信息条件下构建同时约束集群自干扰与外部窃听吞吐量的多目标优化模型, 以应对动态通信环境中的安全性与服务质量权衡. 进一步地, 文献 [25] 针对无人机赋能的智能交通场景, 通过改进进化算法求解帕累托 (Pareto) 解集, 实现在降低能耗的同时保持多无人机任务执行的时间同步性. 尽管现有的多目标优化框架为控制增益的最优设计奠定了坚实基础, 但在实际应用中仍面临严峻挑战. 帕累托最优解的搜索通常涉及高维非线性规划, 计算复杂度极高, 导致传统迭代优化算法的耗时远超无人机底层控制回路的周期. 这种计算时延使得优化算法难以在线实时运行, 限制了集群在面对动态障碍物或环境突变时的即时适应能力. 针对这一难题, 元学习因其能够从多任务经验中提炼跨任务的结构性知识, 并在面对新环境时实现快速适应, 成为突破实时性瓶颈的潜在手段[26–28]. 元学习能够通过离线训练习得具备强泛化能力的初始化策略, 使系统仅需极少的在线更新步骤即可逼近最优解, 从而有望在保证控制高频响应的同时实现性能的在线优化[28]. 例如, 文献 [29] 提出的元学习框架能够在跨任务中提取优化模式, 并在在线阶段快速调度多目标进化算子, 从而显著提升动态场景下的适应性. 进一步地, 文献 [30] 通过分解式多目标元学习策略对子问题的优化经验进行建模, 使多目标优化算法在面对新的环境条件或目标偏好时能够实现快速泛化. 此外, 文献 [31] 提出一种基于元学习的深度强化学习方法, 通过快速适应不同目标偏好以提升多目标优化中最优解的效率与质量.

针对上述挑战, 本文面向无人机集群在突发环境变化下易出现性能退化的难题, 提出一种基于动态性能均衡的元学习控制框架. 该框架在严格保障轨迹跟踪、集群协同与安全避障等多重约束的前提下, 实现了控制性能的动态均衡与在线快速自适应,

从而显著增强了无人机集群在复杂动态环境中的协同控制能力. 本文的主要贡献概括如下:

1) 构建抗扰动的自适应安全协同控制策略. 针对存在模型不确定性与外部扰动的无人机集群系统, 统筹考虑轨迹跟踪、集群协同及动态避障约束, 设计一类融合非线性阻尼补偿与参数在线估计的自适应控制律, 理论分析严格证明了闭环系统误差的一致最终有界性.

2) 建立关键增益的性能均衡优化设计框架. 针对控制器瞬态响应速度与稳态控制精度间的内在制约关系, 构建多目标优化模型, 将关键增益的整定形式化为可求解的数学问题, 并通过量化性能指标间的耦合特性使无人机集群系统能够依据性能偏好实现最优性能均衡.

3) 提出双速率元学习实时控制架构. 为突破传统优化算法计算耗时长、难以满足高频控制实时性要求的瓶颈, 构建一种双速率元学习架构. 该方法通过离线训练习得具备强泛化能力的初始化参数, 使系统在线仅需利用单步梯度更新即可实现参数的快速自适应, 在保证极低计算代价的同时实现了对动态性能均衡解的实时逼近.

本文结构安排如下: 第 1 节给出系统模型与控制目标; 第 2 节设计自适应安全控制策略; 第 3 节建立关键增益的性能均衡优化框架; 第 4 节阐述双速率元学习架构及在线更新机制; 第 5 节进行仿真实验验证; 第 6 节总结全文.

1　问题提出

1.1　无人机动力学模型

本文考虑由 N 个无人机组成的集群系统, 第 $i \in \{1,\ 2,\ \cdots,\ N\}$ 个无人机的动力学模型描述如下:

$$\begin{cases} \dot{\boldsymbol{p}}_i = \boldsymbol{v}_i \\ M_i \dot{\boldsymbol{v}}_i = -M_i \boldsymbol{g} + \boldsymbol{u}_i + \boldsymbol{\omega}_i \end{cases} \tag{1}$$

其中, $\boldsymbol{p}_i = [x_i,\ y_i,\ z_i]^{\mathrm{T}} \in \mathbf{R}^3$ 与 $\boldsymbol{v}_i \in \mathbf{R}^3$ 分别表示无人机 i 在惯性坐标系下的位置与速度向量; $\boldsymbol{u}_i \in \mathbf{R}^3$ 为控制输入向量; $M_i = m_i + \Delta M_i$ 为无人机实际质量, 由标称质量 $m_i > 0$ 与质量偏差 ΔM_i 构成; $\boldsymbol{g} = [0,\ 0,\ \bar{g}]^{\mathrm{T}} \in \mathbf{R}^3$ 为重力加速度向量, $\bar{g}$ 为重力加速度的大小; $\boldsymbol{\omega}_i \in \mathbf{R}^3$ 表示外部环境扰动.

1.2　控制目标

为使本文研究的无人机集群系统在面临模型不确定性和外部扰动时, 仍能在复杂动态环境中实现稳定、高效的协同运行, 所设计的控制器需同时满足以下轨迹跟踪性能指标与安全约束:

1) 协同轨迹跟踪: 集群系统需要协同跟踪给定的参考轨迹 $\boldsymbol{p}_0$, 其满足 $\dot{\boldsymbol{p}}_0 = \boldsymbol{v}_0$, $\boldsymbol{v}_0$ 为参考轨迹的速度向量. 控制器需驱动无人机收敛并保持在参考轨迹 $\boldsymbol{p}_0$ 的预设有界邻域内, 即

$$\lim_{t\to\infty} \|\boldsymbol{p}_i - \boldsymbol{p}_0\| \le \gamma_i \tag{2}$$

其中, $\gamma_i > 0$ 为预设的误差容限参数.

2) 内部避碰约束: 为确保集群内部安全, 任意两架无人机 i 与 j 之间的欧氏距离需要满足:

$$\|\boldsymbol{p}_i - \boldsymbol{p}_j\| > r_{ij},\ i \ne j \tag{3}$$

其中, $r_{ij} = r_i + r_j$ 为避碰安全距离, $r_i > 0$ 为无人机 i 的安全半径.

3) 外部避障约束: 对于每架无人机 i, 当其感知范围 $\mathcal{O}_i$ 内探测到障碍物时, 无人机需与障碍物中心 $\boldsymbol{\varpi}_l$ 始终保持安全距离, 即

$$\|\boldsymbol{p}_i - \boldsymbol{\varpi}_l\| > d_{il},\ l \in \mathcal{O}_i \tag{4}$$

其中, $d_{il} = r_i + R_l + r_0$ 为避障安全距离, $R_l > 0$ 为障碍物 l 的半径, $r_0 > 0$ 为预设的避障缓冲距离.

在进行控制方法设计之前, 首先引入以下假设:

假设 1. 针对系统中的集总不确定项 $\boldsymbol{\Xi}_i(\boldsymbol{p}_i,\ \boldsymbol{v}_i)$, 假设存在一个未知但有界的正常数 ρ_i, 使得对于任意可达的系统状态 $\boldsymbol{p}_i$ 与 $\boldsymbol{v}_i$, 如下不等式恒成立:

$$\|\boldsymbol{\Xi}_i(\boldsymbol{p}_i,\ \boldsymbol{v}_i)\| \le \rho_i m_i \Phi_i(\boldsymbol{p}_i,\ \boldsymbol{v}_i) \tag{5}$$

其中, $\Phi_i(\cdot)$ 为关于系统状态的已知非负回归函数.

2　控制方法设计

为实现复杂动态环境下集群的协同避障控制, 基于无人机动力学模型 (1) 及控制目标, 提出一种基于动态性能均衡的元学习控制框架, 如图 1 所示.

2.1　运动学制导层

在运动学制导层, 本文利用人工势场法的叠加原理, 将参考轨迹引力、集群内机间避碰斥力以及外部障碍物斥力进行向量合成, 得到第 i 架无人机综合速度指令向量 $\boldsymbol{\delta}_i$.

1) 参考轨迹跟踪项 $\boldsymbol{\delta}_{i1}$: 用于驱动无人机跟踪预设的参考轨迹 $\boldsymbol{p}_0$, 其定义为:

$$\boldsymbol{\delta}_{i1} = \boldsymbol{p}_i - \boldsymbol{p}_0 \tag{6}$$

2) 集群协同与避碰项 $\boldsymbol{\delta}_{i2}$: 用于平衡集群的聚集协同与防碰撞约束, 确保无人机在保持紧凑编队构型的同时避免内部物理冲突, 其构造如下:

$$\boldsymbol{\delta}_{i2} = \sum_{j=1,\ j\ne i}^{N} (\boldsymbol{p}_i - \boldsymbol{p}_j) \left(a_{ij} - \frac{b_{ij}}{\|\boldsymbol{p}_i - \boldsymbol{p}_j\|^q - r_{ij}^q} \right) \tag{7}$$

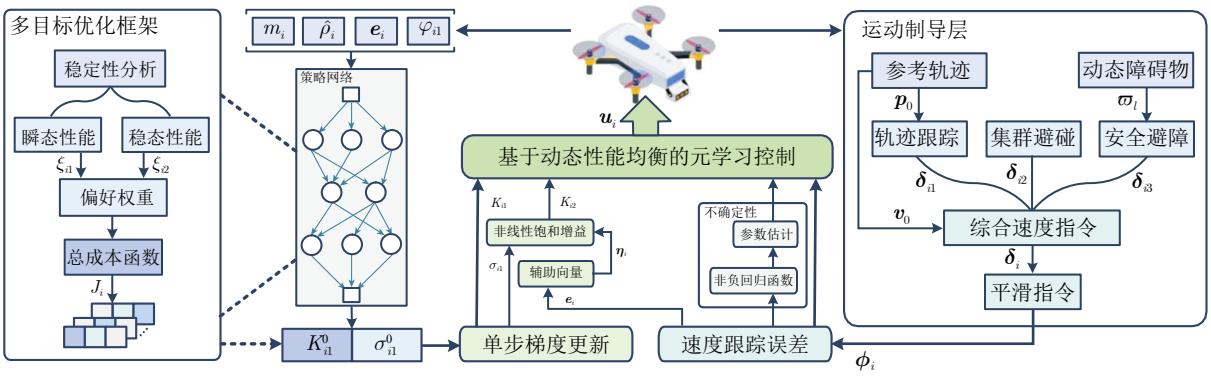


图 1 基于动态性能均衡的元学习控制框架

Fig.1 Dynamic performance balancing-based meta-learning control framework

其中, $a_{ij}>0$ 与 $b_{ij}>0$ 分别为势场吸引系数与排斥系数; $q\geq 1$ 为势场调节因子.

3) 动态障碍物规避项 $\boldsymbol{\delta}_{i3}$: 传统的静态人工势场法仅依赖位置信息建立斥力场, 难以感知动态障碍物的相对运动趋势, 易导致避障响应滞后甚至引发碰撞. 为此, 本文提出一种基于速度感知预测的分级避障策略. 设第 l 个障碍物的速度为 $\boldsymbol{v}_l$. 定义无人机 i 与障碍物 l 之间的相对位置向量为 $\boldsymbol{p}_{il}=\boldsymbol{p}_i-\boldsymbol{\varpi}_l$, 相对速度向量为 $\boldsymbol{v}_{il}=\boldsymbol{v}_i-\boldsymbol{v}_l$.

首先, 计算相对速度在相对位置方向上的投影标量 $\psi_{il}=\boldsymbol{p}_{il}^{\mathrm{T}}\boldsymbol{v}_{il}$. 若 $\psi_{il}<0$, 表明无人机与障碍物正在相互逼近. 此时, 引入基于相对速度幅值的预测修正项, 构造有效预测避障距离 ζ_{il} 为:

$$\zeta_{il}=\|\boldsymbol{p}_{il}\|-c_{i1}\mathbb{I}(\psi_{il}<0)\|\boldsymbol{v}_{il}\| \tag{8}$$

其中, c_{i1} 为预测时域系数; $\mathbb{I}(\cdot)$ 为指示函数, 当条件满足时取值为 1, 否则为 0. 根据有效距离 ζ_{il} 与避障安全距离 d_{il} 的关系, 设计分级避障斥力 $\boldsymbol{f}_{il}$ 如下:

$$\boldsymbol{f}_{il}=\begin{cases}F_i\dfrac{\boldsymbol{p}_{il}}{\|\boldsymbol{p}_{il}\|}, & \zeta_{il}\leq\tilde{d}_{il}\\ \mathrm{sat}_{\bar{F}}\left(\dfrac{c_{i2}+c_{i3}\|\boldsymbol{v}_{il}\|}{\zeta_{il}^2-d_{il}^2}\boldsymbol{p}_{il}\right), & \tilde{d}_{il}<\zeta_{il}<\bar{d}_{il}\\ 0, & \text{其他}\end{cases} \tag{9}$$

其中, $\tilde{d}_{il}=d_{il}+r_0$ 为紧急避障边界; $\bar{d}_{il}=d_{il}+d_{i0}$ 为探测与预警边界, $d_{i0}>0$ 为障碍物探测范围; $F_i>0$ 为紧急避障推力幅值; $\mathrm{sat}_{\bar{F}}(\boldsymbol{x})$ 表示饱和幅值为 $\bar{F}>0$ 的向量饱和函数; c_{i2}、$c_{i3}>0$ 为速度调节权重系数. 最终, 动态障碍物规避项 $\boldsymbol{\delta}_{i3}$ 构造为所有感知范围内障碍物斥力的叠加:

$$\boldsymbol{\delta}_{i3}=-\sum_{l\in\mathcal{O}_i}\boldsymbol{f}_{il}(\zeta_{il},\,\boldsymbol{v}_{il}) \tag{10}$$

将上述三类引导项进行线性加权组合, 可得到制导层输出的综合速度指令 $\boldsymbol{\delta}_i$:

$$\boldsymbol{\delta}_i=\boldsymbol{v}_0-(k_{i1}\boldsymbol{\delta}_{i1}+k_{i2}\boldsymbol{\delta}_{i2}+k_{i3}\boldsymbol{\delta}_{i3}) \tag{11}$$

其中, k_{i1}、k_{i2}、$k_{i3}>0$ 为各引导项的制导增益. 考虑到人工势场法在障碍物边界处可能产生突变的速度指令, 本文引入低通滤波器对 $\boldsymbol{\delta}_i$ 进行平滑处理, 得到最终的期望速度 $\boldsymbol{\phi}_i$:

$$\dot{\boldsymbol{\phi}}_i=k_{i4}(\boldsymbol{\delta}_i-\boldsymbol{\phi}_i) \tag{12}$$

其中, $k_{i4}>0$ 为滤波器带宽参数.

2.2 动力学跟踪层

动力学跟踪层旨在设计控制输入 $\boldsymbol{u}_i$, 确保无人机的实际速度 $\boldsymbol{v}_i$ 能够快速、精确地跟踪由运动学制导层生成的期望速度指令 $\boldsymbol{\phi}_i$. 为此, 定义速度跟踪误差为 $\boldsymbol{e}_i=\boldsymbol{v}_i-\boldsymbol{\phi}_i$.

为应对无人机集群系统中存在的不确定性, 本文基于假设 1 构建了用于对 ρ_i 进行在线估计的自适应更新律. 令 $\hat{\rho}_i$ 为 ρ_i 的估计值, 定义参数估计误差为 $\tilde{\rho}_i=\hat{\rho}_i-\rho_i$, 构造自适应更新律如下:

$$\dot{\hat{\rho}}_i=K_{i3}\left(\mu_i(\boldsymbol{e}_i)\|\boldsymbol{e}_i\|\Phi_i-K_{i1}\kappa_i\hat{\rho}_i\right) \tag{13}$$

其中, K_{i1}、K_{i3}、$\kappa_i>0$ 为待设计的增益参数; $\mu_i(\boldsymbol{e}_i)$ 为更新速率调节参数. 为抑制小误差区间内测量噪声对参数更新过程的影响, $\mu_i(\boldsymbol{e}_i)$ 设计如下:

$$\mu_i(\boldsymbol{e}_i)=\begin{cases}1, & \|\boldsymbol{e}_i\|\|\Phi_i\|\geq\epsilon_i\\ \dfrac{\|\boldsymbol{e}_i\|\|\Phi_i\|}{\epsilon_i}, & \|\boldsymbol{e}_i\|\|\Phi_i\|<\epsilon_i\end{cases} \tag{14}$$

其中, $\epsilon_i>0$ 为阈值参数. 同时, 为充分捕捉状态演化过程中的不确定性特征, 选取回归函数 Φ_i 为:

$$\Phi_i=\|\boldsymbol{e}_i\|+\|\dot{\boldsymbol{\phi}}_i\|+\|m_i\boldsymbol{g}\|+1 \tag{15}$$

此外, 为避免控制输入抖振问题, 设计了一种非线性饱和增益函数 $K_{i2}(\boldsymbol{\eta}_i)$. 定义辅助向量 $\boldsymbol{\eta}_i=$

$\hat{\rho}_i\Phi_i\boldsymbol{e}_i$, 构造 $K_{i2}(\boldsymbol{\eta}_i)$ 如下:

$$K_{i2}(\boldsymbol{\eta}_i)=\sigma_{i1}\frac{\sigma_{i2}\|\boldsymbol{\eta}_i\|^2}{\sigma_{i2}\|\boldsymbol{\eta}_i\|^2+\sigma_{i3}} \tag{16}$$

其中, σ_{i1}、σ_{i2}、$\sigma_{i3}>0$ 为待设计的增益参数.

综合上述设计, 本文提出的动力学控制输入为:

$$\boldsymbol{u}_i=m_i\big(\dot{\boldsymbol{\phi}}_i-K_{i1}\boldsymbol{e}_i-K_{i2}(\boldsymbol{\eta}_i)\hat{\rho}_i\Phi_i\boldsymbol{\eta}_i+\boldsymbol{g}\big) \tag{17}$$

2.3　稳定性分析

定理 1. 对于由 (1) 描述的无人机不确定动力学系统, 在假设 1 成立的前提下, 应用控制输入 (17) 及自适应更新律 (13), 则闭环系统的速度跟踪误差 $\boldsymbol{e}_i$ 和估计误差 $\tilde{\rho}_i$ 均是一致最终有界的.

证明. 定义如下的李雅普诺夫候选函数 V_i:

$$V_i=V_{i1}+V_{i2},\ \ V_{i1}=\tfrac{1}{2}\boldsymbol{e}_i^{\mathrm{T}}\boldsymbol{e}_i,\ \ V_{i2}=\tfrac{1}{2K_{i3}}\tilde{\rho}_i^2 \tag{18}$$

对 V_{i1} 求导可得:

$$\dot{V}_{i1}=\boldsymbol{e}_i^{\mathrm{T}}\dot{\boldsymbol{e}}_i=\boldsymbol{e}_i^{\mathrm{T}}\big(\dot{\boldsymbol{v}}_i-\dot{\boldsymbol{\phi}}_i\big) \tag{19}$$

将动力学模型 (1) 重写为:

$$m_i\dot{\boldsymbol{v}}_i=-m_i\boldsymbol{g}+\boldsymbol{u}_i+\boldsymbol{\Xi}_i \tag{20}$$

其中, $\boldsymbol{\Xi}_i=-\Delta M_i\dot{\boldsymbol{v}}_i-\Delta M_i\boldsymbol{g}+\boldsymbol{\omega}_i$ 为集总不确定项. 将所设计的控制输入 (17) 代入 (20), 可得:

$$m_i\dot{\boldsymbol{v}}_i=m_i\big(\dot{\boldsymbol{\phi}}_i-K_{i1}\boldsymbol{e}_i-K_{i2}(\boldsymbol{\eta}_i)\hat{\rho}_i\Phi_i\boldsymbol{\eta}_i\big)+\boldsymbol{\Xi}_i \tag{21}$$

因此, 速度跟踪误差 $\boldsymbol{e}_i$ 的导数为

$$\dot{\boldsymbol{e}}_i=-K_{i1}\boldsymbol{e}_i-K_{i2}(\boldsymbol{\eta}_i)\hat{\rho}_i\Phi_i\boldsymbol{\eta}_i+\frac{\boldsymbol{\Xi}_i}{m_i} \tag{22}$$

将 (22) 代入 (19), 得到:

$$\dot{V}_{i1}=-K_{i1}\|\boldsymbol{e}_i\|^2-K_{i2}(\boldsymbol{\eta}_i)\hat{\rho}_i\Phi_i\boldsymbol{e}_i^{\mathrm{T}}\boldsymbol{\eta}_i+\frac{\boldsymbol{e}_i^{\mathrm{T}}\boldsymbol{\Xi}_i}{m_i} \tag{23}$$

根据假设 1, 并利用柯西–施瓦茨不等式可得:

$$\dot{V}_{i1}\le-K_{i1}\|\boldsymbol{e}_i\|^2-K_{i2}(\boldsymbol{\eta}_i)\hat{\rho}_i\Phi_i\boldsymbol{e}_i^{\mathrm{T}}\boldsymbol{\eta}_i+\|\boldsymbol{e}_i\|\rho_i\Phi_i \tag{24}$$

对 V_{i2} 求导, 并代入自适应更新律 (13), 有:

$$\dot{V}_{i2}=\tilde{\rho}_i\big(\mu_i(\boldsymbol{e}_i)\Phi_i\|\boldsymbol{e}_i\|-K_{i1}\kappa_i\hat{\rho}_i\big) \tag{25}$$

根据 $\dot{V}_i=\dot{V}_{i1}+\dot{V}_{i2}$, 将 (24) 和 (25) 相加可得:

$$\begin{aligned}\dot{V}_i\le&-K_{i1}\|\boldsymbol{e}_i\|^2-K_{i2}(\boldsymbol{\eta}_i)\hat{\rho}_i\Phi_i\boldsymbol{e}_i^{\mathrm{T}}\boldsymbol{\eta}_i+\|\boldsymbol{e}_i\|\rho_i\Phi_i+\\&\tilde{\rho}_i\mu_i(\boldsymbol{e}_i)\Phi_i\|\boldsymbol{e}_i\|-K_{i1}\kappa_i\tilde{\rho}_i\hat{\rho}_i\end{aligned} \tag{26}$$

利用 $\rho_i=\hat{\rho}_i-\tilde{\rho}_i$ 代替 (26) 中的 ρ_i 可得:

$$\begin{aligned}\dot{V}_i\le&-K_{i1}\|\boldsymbol{e}_i\|^2-K_{i2}(\boldsymbol{\eta}_i)\hat{\rho}_i\Phi_i\boldsymbol{e}_i^{\mathrm{T}}\boldsymbol{\eta}_i+\hat{\rho}_i\Phi_i\|\boldsymbol{e}_i\|+\\&\big(\mu_i(\boldsymbol{e}_i)-1\big)\tilde{\rho}_i\Phi_i\|\boldsymbol{e}_i\|-K_{i1}\kappa_i\tilde{\rho}_i\hat{\rho}_i\end{aligned} \tag{27}$$

根据辅助向量 $\boldsymbol{\eta}_i=\hat{\rho}_i\Phi_i\boldsymbol{e}_i$ 的定义, 有 $\boldsymbol{e}_i^{\mathrm{T}}\boldsymbol{\eta}_i=\hat{\rho}_i\Phi_i\|\boldsymbol{e}_i\|^2$, 将其代入 (27) 中:

$$\begin{aligned}\dot{V}_i\le&-K_{i1}\|\boldsymbol{e}_i\|^2-K_{i2}(\boldsymbol{\eta}_i)\big(\hat{\rho}_i\Phi_i\|\boldsymbol{e}_i\|\big)^2+\\&\hat{\rho}_i\Phi_i\|\boldsymbol{e}_i\|+\big(\mu_i(\boldsymbol{e}_i)-1\big)\tilde{\rho}_i\Phi_i\|\boldsymbol{e}_i\|-\\&K_{i1}\kappa_i\tilde{\rho}_i\hat{\rho}_i\end{aligned} \tag{28}$$

为简化不等式 (28), 定义 $A_i=\|\boldsymbol{\eta}_i\|$, 推导可得 $A_i=\|\hat{\rho}_i\Phi_i\boldsymbol{e}_i\|=\hat{\rho}_i\Phi_i\|\boldsymbol{e}_i\|$. 进一步, 结合 $K_{i2}(\boldsymbol{\eta}_i)$ 的定义 (16), 可得:

$$\begin{aligned}&-K_{i2}(\boldsymbol{\eta}_i)\big(\hat{\rho}_i\Phi_i\|\boldsymbol{e}_i\|\big)^2+\hat{\rho}_i\Phi_i\|\boldsymbol{e}_i\|=\\&A_i-\left(\sigma_{i1}\frac{\sigma_{i2}\|\boldsymbol{\eta}_i\|^2}{\sigma_{i2}\|\boldsymbol{\eta}_i\|^2+\sigma_{i3}}\right)A_i^2=\\&A_i-\sigma_{i1}A_i^2+\frac{\sigma_{i1}\sigma_{i3}A_i^2}{\sigma_{i2}A_i^2+\sigma_{i3}}\end{aligned} \tag{29}$$

由于 $\sigma_{i1}\sigma_{i3}A_i^2/(\sigma_{i2}A_i^2+\sigma_{i3})\le\sigma_{i1}\sigma_{i3}/\sigma_{i2}$, 且当 $A_i=1/(2\sigma_{i1})$ 时, $A_i-\sigma_{i1}A_i^2$ 的最大值为 $1/(4\sigma_{i1})$, 因此 (29) 具有上界:

$$-K_{i2}(\boldsymbol{\eta}_i)\big(\hat{\rho}_i\Phi_i\|\boldsymbol{e}_i\|\big)^2+\hat{\rho}_i\Phi_i\|\boldsymbol{e}_i\|\le\frac{1}{4\sigma_{i1}}+\frac{\sigma_{i1}\sigma_{i3}}{\sigma_{i2}} \tag{30}$$

此外, 使用 $\hat{\rho}_i=\tilde{\rho}_i+\rho_i$ 和杨氏不等式, 可得:

$$-K_{i1}\kappa_i\tilde{\rho}_i\hat{\rho}_i\le-\frac{K_{i1}\kappa_i}{2}\tilde{\rho}_i^2+\frac{K_{i1}\kappa_i}{2}\rho_i^2 \tag{31}$$

将 (30) 和 (31) 代入 (28), 有:

$$\begin{aligned}\dot{V}_i\le&-K_{i1}\|\boldsymbol{e}_i\|^2-\frac{K_{i1}\kappa_i}{2}\tilde{\rho}_i^2+\Theta_i+\\&\big(\mu_i(\boldsymbol{e}_i)-1\big)\tilde{\rho}_i\Phi_i\|\boldsymbol{e}_i\|\end{aligned} \tag{32}$$

其中, $\Theta_i=1/(4\sigma_{i1})+\sigma_{i1}\sigma_{i3}/\sigma_{i2}+K_{i1}\kappa_i\rho_i^2/2$.

接下来, 针对式 (32) 中的 $\big(\mu_i(\boldsymbol{e}_i)-1\big)\tilde{\rho}_i\Phi_i\|\boldsymbol{e}_i\|$ 这一项, 分两种情况讨论:

情况 1. $\|\boldsymbol{e}_i\|\|\Phi_i\|\ge\epsilon_i$. 根据 (14) 可得 $\mu_i(\boldsymbol{e}_i)=1$, 因此式 (32) 可简化为:

$$\dot{V}_i\le-K_{i1}\|\boldsymbol{e}_i\|^2-\frac{K_{i1}\kappa_i}{2}\tilde{\rho}_i^2+\Theta_i \tag{33}$$

情况 2. $\|\boldsymbol{e}_i\|\|\Phi_i\|<\epsilon_i$. 在此情况下, $\mu_i(\boldsymbol{e}_i)=\|\boldsymbol{e}_i\|\|\Phi_i\|/\epsilon_i$, 且 $0\le|\mu_i(\boldsymbol{e}_i)-1|\le1$, 因此,

$$\big(\mu_i(\boldsymbol{e}_i)-1\big)\tilde{\rho}_i\Phi_i\|\boldsymbol{e}_i\|\le|\tilde{\rho}_i|\epsilon_i \tag{34}$$

进一步应用杨氏不等式, 得到

$$|\tilde{\rho}_i|\epsilon_i\le\frac{K_{i1}\kappa_i}{4}\tilde{\rho}_i^2+\frac{1}{K_{i1}\kappa_i}\epsilon_i^2 \tag{35}$$

将 (35) 代入 (32), 可得:

$$\dot{V}_i\le-K_{i1}\|\boldsymbol{e}_i\|^2-\frac{K_{i1}\kappa_i}{4}\tilde{\rho}_i^2+\frac{1}{K_{i1}\kappa_i}\epsilon_i^2+\Theta_i \tag{36}$$

综合 (33) 和 (37) 两种情况, $\dot{V}_i$ 总是满足以下形式的不等式:

$$\dot{V}_i \le -\lambda_{\min} V_i + \Omega_i \tag{37}$$

其中, $\lambda_{\min} = \min\left(2K_{i1},\ \frac{1}{2}\kappa_i K_{i1} K_{i3}\right)$, $\Omega_i = \Theta_i + \epsilon_i^2/(K_{i1}\kappa_i)$.

依据李雅普诺夫稳定性扩展定理, 式 (37) 表明 $\dot{V}_i$ 在集合 $\mathcal{B}_{V_i} = \{V_i \mid V_i \le \Omega_i/\lambda_{\min}\}$ 之外严格负定. 这意味着 $V_i(t)$ 最终将收敛至该有界区域, 即满足 $\lim_{t\to\infty} V_i(t) \le \Omega_i/\lambda_{\min}$. 鉴于 V_i 的定义 (18) 以及其各项的非负性, 可推导出速度跟踪误差与估计误差的渐近上界分别为 $\lim_{t\to\infty}\|\boldsymbol{e}_i(t)\| \le \sqrt{2\Omega_i/\lambda_{\min}}$ 与 $\lim_{t\to\infty}|\tilde{\rho}_i(t)| \le \sqrt{2K_{i3}\Omega_i/\lambda_{\min}}$. 上述结果表明误差 $\boldsymbol{e}_i$ 和 $\tilde{\rho}_i$ 均收敛于原点附近的紧集, 从而得证闭环系统是一致最终有界的. □

3 性能均衡优化

上述控制策略的有效性在很大程度上取决于关键增益 K_{i1} 和 σ_{i1} 的选取, 然而, 增益在提升瞬态响应速度与降低稳态误差之间存在固有的性能冲突. 鉴于此, 本节构建了一个多目标优化框架, 旨在定量分析这种制约关系, 并求解出能够满足性能均衡要求的最优增益配置.

3.1 性能指标定义

基于前述稳定性分析, 李雅普诺夫函数 $V_i(t)$ 在任意初始时刻 t_s 后的轨迹上界可表示为:

$$V_i(t) \le \left(V_i(t_s) - \frac{\Omega_i}{\lambda_{\min}}\right)\mathrm{e}^{-\lambda_{\min}(t-t_s)} + \frac{\Omega_i}{\lambda_{\min}} \tag{38}$$

为量化控制性能, 基于 (38) 定义如下性能指标:

1) 瞬态性能指标 ξ_{i1}: 该指标旨在衡量系统的累积瞬态响应误差, 定义为瞬态误差上界的能量积分.

$$\xi_{i1} := \int_{t_s}^{\infty}\left[\left(V_i(t_s) - \frac{\Omega_i}{\lambda_{\min}}\right)\mathrm{e}^{-\lambda_{\min}(t_\tau - t_s)}\right]^2 \mathrm{d}t_\tau = \frac{1}{2\lambda_{\min}}\left(V_i(t_s) - \frac{\Omega_i}{\lambda_{\min}}\right)^2 \tag{39}$$

2) 稳态性能指标 ξ_{i2}: 该指标用于衡量系统收敛后的稳态误差水平, 定义为最终上界的平方.

$$\xi_{i2} := \left(\frac{\Omega_i}{\lambda_{\min}}\right)^2 \tag{40}$$

本文的核心优化目标是同时最小化上述两个相互制约的性能指标 ξ_{i1} 和 ξ_{i2}. 分析表明, 这两项指标均是关键增益 K_{i1} 与 σ_{i1} 的复杂非线性函数, 调节增益以改善瞬态响应往往导致稳态性能恶化, 反之亦然. 因此, 本文构建了如下多目标优化框架以系统解决这一内在的性能冲突.

3.2 多目标优化框架

在构建多目标优化框架时, 本文选取对闭环系统性能具有决定性影响的两个关键增益 $K_{i1} \in (0,\ \infty)$ 与 $\sigma_{i1} \in (0,\ \infty)$ 作为决策变量, 其余参数 K_{i3}、κ_i、σ_{i2}、σ_{i3}、ϵ_i 均依据经验预设为正常数. 该多目标优化问题旨在上述两个相互制约的性能指标之间寻求最佳均衡, 进而求解出最优的关键增益.

基于性能指标 (39) 和 (40), 分别构建瞬态与稳态成本函数. 首先, 瞬态成本函数 J_{i1} 旨在最小化瞬态性能指标 ξ_{i1}, 并引入惩罚权重 β_{i1} 以约束增益 K_{i1} 的幅值. 为简化分析, 设定 $\lambda_{\min} = 2K_{i1}$, 即要求参数满足 $\kappa_i K_{i3} \ge 4$, 从而 J_{i1} 可表示为:

$$J_{i1}(K_{i1},\ \sigma_{i1}) = \frac{1}{4K_{i1}}\left(V_i(t_s) - \frac{\Omega_i}{2K_{i1}}\right)^2 + \beta_{i1}K_{i1}^2 \tag{41}$$

类似地, 稳态成本函数 J_{i2} 用于优化稳态性能指标 ξ_{i2}, 在引入相应的惩罚权重 β_{i2} 后可得:

$$J_{i2}(K_{i1},\ \sigma_{i1}) = \frac{\Omega_i^2}{4K_{i1}^2} + \beta_{i2}\sigma_{i1}^2 \tag{42}$$

依据多目标优化理论, 总成本函数 $J_i(K_{i1},\ \sigma_{i1})$ 可定义为上述两个子性能成本函数的加权组合:

$$J_i(K_{i1},\ \sigma_{i1}) = \varphi_{i1}J_{i1} + \varphi_{i2}J_{i2} \tag{43}$$

其中, 权重系数满足 φ_{i1}、$\varphi_{i2} > 0$ 且 $\varphi_{i1} + \varphi_{i2} = 1$, 代表瞬态响应与稳态精度之间的偏好权衡. 将 (41) 和 (42) 代入 (43), 展开可得:

$$J_i(K_{i1},\ \sigma_{i1}) = \frac{\varphi_{i1}V_i^2(t_s)}{4K_{i1}} - \frac{\varphi_{i1}V_i(t_s)}{4K_{i1}^2}\Omega_i + \left(\frac{\varphi_{i1}}{16K_{i1}^3} + \frac{\varphi_{i2}}{4K_{i1}^2}\right)\Omega_i^2 + \varphi_{i1}\beta_{i1}K_{i1}^2 + \varphi_{i2}\beta_{i2}\sigma_{i1}^2 \tag{44}$$

基于 (44), 计算总成本函数 J_i 关于 K_{i1} 的偏导数:

$$\frac{\partial J_i}{\partial K_{i1}} = -\frac{\varphi_{i1}V_i^2(t_s)}{4K_{i1}^2} + \frac{\varphi_{i1}V_i(t_s)}{2K_{i1}^3}\Omega_i - \left(\frac{3\varphi_{i1}}{16K_{i1}^4} + \frac{\varphi_{i2}}{2K_{i1}^3}\right)\Omega_i^2 + 2\varphi_{i1}\beta_{i1}K_{i1} \tag{45}$$

其次, 计算总成本函数 J_i 关于 σ_{i1} 的偏导数, 可得:

$$\frac{\partial J_i}{\partial \sigma_{i1}} = \left[-\frac{\varphi_{i1}V_i(t_s)}{4K_{i1}^2} + \left(\frac{\varphi_{i1}}{8K_{i1}^3} + \frac{\varphi_{i2}}{2K_{i1}^2}\right)\Omega_i\right] \times \left(\frac{\sigma_{i3}}{\sigma_{i2}} - \frac{1}{4\sigma_{i1}^2}\right) + 2\varphi_{i2}\beta_{i2}\sigma_{i1} \tag{46}$$

最终, 满足性能均衡要求的最优增益 $(K_{i1}^*,\ \sigma_{i1}^*)$ 可通过求解如下优化问题获得:

$$(K_{i1}^*,\ \sigma_{i1}^*) = \arg \min_{K_{i1}>0,\ \sigma_{i1}>0} J_i(K_{i1},\ \sigma_{i1}) \tag{47}$$

4　元学习控制方法

尽管第 3 节构建的多目标优化框架为最优控制增益的求解奠定了坚实的数学基础, 然而, 无论是基于梯度的数值优化方法还是基于种群的智能优化算法, 通常均需经历大量的迭代计算方能收敛. 这种显著的计算时延使得传统优化方法难以满足无人机高频控制回路对实时性的严苛要求. 为此, 本文提出一种基于动态性能均衡的元学习控制方法, 旨在赋予无人机集群在复杂动态环境下快速适应的能力.

4.1　元学习问题构建

定义以 θ 为参数的神经网络策略 π_θ. 该网络以当前无人机的状态特征向量 $\boldsymbol{P}_i = [m_i,\ \hat{\rho}_i,\ \|\boldsymbol{e}_i\|,\ \varphi_{i1}]^{\mathrm{T}}$ 为输入, 将其映射为关键控制增益的初始化值 $[K_{i1}^0,\ \sigma_{i1}^0]^{\mathrm{T}}$. 为严格满足定理 1 所述的李雅普诺夫稳定性条件, 并避免成本函数计算中的数值奇异性, 本文在网络输出层引入 Softplus 激活函数并施加安全阈值约束, 具体表示为:

$$[K_{i1}^0,\ \sigma_{i1}^0]^{\mathrm{T}} = \mathrm{Softplus}(\pi_\theta(\boldsymbol{P}_i)) + [\bar{K}_{i1},\ \bar{\sigma}_{i1}]^{\mathrm{T}} \tag{48}$$

其中, $\bar{K}_{i1}$、$\bar{\sigma}_{i1} > 0$ 为预设的增益安全下界.

对于任意给定的任务 $\mathcal{S}_j$, 该策略的优化目标是使网络输出的初始增益 $[K_{i1}^0,\ \sigma_{i1}^0]^{\mathrm{T}}$ 能够逼近最优性能均衡解所在的邻域. 在此基础上, 仅需基于总成本函数的解析梯度 (45) 与 (46) 执行单步更新, 即可快速收敛至当前任务下的近似最优增益.

4.2　元学习模型训练

元学习模型的离线训练旨在通过双层优化机制, 使策略网络 π_θ 习得一组具有强泛化能力的初始化参数. 在训练中, 一个任务 $\mathcal{S}_j$ 被定义为一组特定的环境参数与优化偏好的组合, 即 $\mathcal{S}_j = \{m_j,\ \rho_j,\ \varphi_{j1}\}$. 此外, 定义支持集 $\mathcal{Q}_j^1$ 与查询集 $\mathcal{Q}_j^2$ 分别为在任务 $\mathcal{S}_j$ 环境下采集长度为 H 的系统状态序列.

1) 内循环: 在每一轮训练迭代中, 从任务分布 $p(\mathcal{S})$ 中采样一批任务 $\mathcal{S}_j$ 及对应的支持集 $\mathcal{Q}_j^1$. 首先, 策略网络 π_θ 以支持集 $\mathcal{Q}_j^1$ 轨迹中最后时刻的状态特征 $\boldsymbol{P}_j(H)$ 为输入, 通过前向传播获取初始增益 $[K_j^0,\ \sigma_j^0]^{\mathrm{T}}$; 随后, 基于完整支持集计算成本函数关于增益的梯度, 并执行单步梯度下降更新, 得到该任务的适应性增益 $[K_j',\ \sigma_j']^{\mathrm{T}}$:

$$\begin{bmatrix} K_j' \\ \sigma_j' \end{bmatrix} = \begin{bmatrix} K_j^0 \\ \sigma_j^0 \end{bmatrix} - \alpha_1 \nabla_{[K,\ \sigma]} \mathcal{J}_{\mathcal{S}_j}\left(\mathcal{Q}_j^1;\ K_j^0,\ \sigma_j^0\right) \tag{49}$$

其中, α_1 为内循环学习率; $\mathcal{J}_{\mathcal{S}_j}(\cdot)$ 为定义在任务 $\mathcal{S}_j$ 上的总成本函数.

2) 外循环: 旨在评估适应性增益 $\left[K_j',\ \sigma_j'\right]^{\mathrm{T}}$ 在新数据查询集 $\mathcal{Q}_j^2$ 上的泛化性能. 训练的最终目标是寻找最优的全局网络参数 θ^*, 使得生成的初始增益 $[K_j^0,\ \sigma_j^0]^{\mathrm{T}}$ 在经过内循环的一步微调后, 能够在查询集 $\mathcal{Q}_j^2$ 上实现最小的期望综合加权成本:

$$\min_\theta \mathrm{E}_{\mathcal{S}_j \sim p(\mathcal{S})}\left[\mathcal{J}_{\mathcal{S}_j}\left(\mathcal{Q}_j^2;\ K_j',\ \sigma_j'\right)\right] \tag{50}$$

尽管内循环仅更新控制增益而不更新网络参数 θ, 但外循环中计算的梯度会通过链式法则反向传播至 θ. 这一机制驱动策略网络 π_θ 学习到一种易于适应的参数流形, 从而显著提升在线阶段的收敛速度. 元学习模型离线训练流程如算法 1 所示.

算法 1. 元学习模型离线训练流程

1. **初始化:** 内循环学习率 α_1, 外循环学习率 α_2, 任务批次大小 B, 最大迭代轮数 $S_{\max}$;
2. 随机初始化元参数 θ, 初始化迭代计数器 $s \leftarrow 1$;
3. **while** $s \le S_{\max}$ **do**
4. 　从任务分布 $p(\mathcal{S})$ 中采样一批任务 $\{\mathcal{S}_j\}_{j=1}^B$;
5. 　**for** 每个任务 $\mathcal{S}_j$ $(j = 1$ to $B)$ **do**
6. 　　**内循环 (增益适应):**
7. 　　从 $\mathcal{S}_j$ 中采集支持集 $\mathcal{Q}_j^1$ 并提取特征向量 $\boldsymbol{P}_j$;
8. 　　前向传播生成初始增益: $[K_j^0,\ \sigma_j^0]^{\mathrm{T}} = \pi_\theta(\boldsymbol{P}_j)$;
9. 　　计算 $\mathcal{J}_{\mathcal{S}_j}$ 关于增益的梯度: $\nabla_{[K,\ \sigma]}\mathcal{J}_{\mathcal{S}_j}(\mathcal{Q}_j^1;\ K_j^0,\ \sigma_j^0)$;
10. 　　更新任务适应性增益:
11. 　　$[K_j',\ \sigma_j']^{\mathrm{T}} \leftarrow [K_j^0,\ \sigma_j^0]^{\mathrm{T}} - \alpha_1 \nabla_{[K,\ \sigma]}\mathcal{J}_{\mathcal{S}_j}$;
12. 　　**外循环 (损失评估):**
13. 　　从 $\mathcal{S}_j$ 中采集查询集数据 $\mathcal{Q}_j^2$;
14. 　　计算基于适应后增益的损失 $\mathcal{J}_{\mathcal{S}_j}(\mathcal{Q}_j^2;\ K_j',\ \sigma_j')$;
15. 　**end for**
16. 　计算关于 θ 的元梯度: $\tilde{\theta} = \nabla_\theta \sum_{j=1}^B \mathcal{J}_{\mathcal{S}_j}(\mathcal{Q}_j^2;\ K_j',\ \sigma_j')$;
17. 　更新全局网络参数: $\theta \leftarrow \theta - \alpha_2 \tilde{\theta}$;
18. 　$s \leftarrow s + 1$;
19. **end while**
20. **return** 最优网络参数 θ^*.

4.3　元学习增益更新

在线运行时, 采用双速率时间尺度架构: 底层动力学控制律以高频控制周期实时运行, 而元学习

增益则以较低的频率更新. 需要特别说明的是, 尽管增益更新仅低频率周期触发, 但得益于策略网络的推理效率与解析梯度的低维特性, 单次增益更新的计算耗时严格小于一个底层控制周期. 这种设计确保了元学习算法能够在单一底层控制周期内完成, 从而在不阻塞实时控制的前提下实现性能的在线优化. 该过程包含以下核心步骤:

1) 动态偏好调整: 在每个增益更新周期, 系统根据无人机当前的感知状态动态调整性能偏好权重 φ_{i1}. 当检测到障碍物逼近或存在剧烈环境扰动时, 增大瞬态性能权重 φ_{i1}, 以允许更大的控制能量用于快速响应; 反之, 在平稳飞行区域则降低 φ_{i1}, 以优先保证稳态精度与能量效率.

2) 初始增益生成: 基于调整后的性能偏好权重以及当前系统状态构建特征向量 $\boldsymbol{P}_i$, 并通过策略网络 π_θ 的前向推理获取初始增益 $[K_{i1}^0,\ \sigma_{i1}^0]^{\mathrm{T}}$.

3) 增益在线微调: 利用长度为 H 的滑动窗口支持集 $\tilde{\mathcal{Q}}_i$, 对初始增益执行基于解析梯度的单步更新, 从而得到适应当前环境的近似最优增益 $[\hat{K}_{i1}^*,\ \hat{\sigma}_{i1}^*]^{\mathrm{T}}$:

$$\begin{bmatrix}\hat{K}_{i1}^*\\ \hat{\sigma}_{i1}^*\end{bmatrix} \leftarrow \begin{bmatrix}K_{i1}^0\\ \sigma_{i1}^0\end{bmatrix} - \frac{\alpha_1}{H}\sum_{h=1}^{H}\begin{bmatrix}\nabla_{K_{i1}} J_i(h)\\ \nabla_{\sigma_{i1}} J_i(h)\end{bmatrix}\Bigg|_{(K_{i1}^0,\ \sigma_{i1}^0)} \tag{51}$$

其中, $\nabla_{K_{i1}} J_i(h)$ 与 $\nabla_{\sigma_{i1}} J_i(h)$ 分别代表利用支持集 $\tilde{\mathcal{Q}}_i$ 中第 h 个时间步的系统状态数据, 依据式 (45) 和 (46) 计算得到的解析梯度值.

相比于依赖在线数值迭代的传统优化算法, 本文提出的架构通过将繁重的优化计算转移至离线元学习阶段, 使得在线更新仅涉及低维标量的解析梯度计算, 显著降低了计算负荷, 从而确保了复杂动态环境下无人机集群的实时控制性能优化.

5 仿真实验

5.1 实验设置与参数

为验证本文所提出的元学习协同避障控制方法的有效性, 本节在仿真环境中构建了由 $N = 6$ 架无人机组成的集群系统. 实验场景包含障碍物与外部环境扰动, 以全面评估所提出控制框架在复杂动态环境中的轨迹跟踪性能和安全避障效果. 实验场景的时间步长设定为 0.001 s, 总仿真时长 $T_m = 200$ s. 所有仿真实验均在配备 Intel Core i5-11400 CPU @ 2.60 GHz 处理器与 16 GB 内存的个人计算机上进行, 算法实现与仿真平台基于 MATLAB R2025a.

无人机的标称质量设定为 $m_i = 1.0$ kg, 并引入 $\Delta M_i = 0.1$ kg 的质量偏差以模拟模型不确定性, 重力加速度取 $\bar{g} = 9.8\ \mathrm{m/s^2}$. 为了验证系统的抗干扰能力, 施加了由背景噪声和突发气流组成的复合外部环境扰动 $\boldsymbol{\omega}_i = \boldsymbol{\omega}_{i1} + \boldsymbol{\omega}_{i2} + \boldsymbol{\omega}_{i3}$. 其中, $\boldsymbol{\omega}_{i1}(t) = [0.05\sin(0.5t) + 0.02\sin(5t),\quad 0.05\cos(0.5t) + 0.02 \times \cos(5t),\ 0.02\sin(t)]^{\mathrm{T}}$ 为全程存在的背景扰动; 在此基础上, 分别在 $t \in [60,\ 90]$ s 时段施加持续侧风扰动 $\boldsymbol{\omega}_{i2}(t) = [0.1\sin(0.5t),\ 0.1\cos(0.5t),\ 0.05]^{\mathrm{T}}$, 以及在 $t \in [130,\ 160]$ s 时段施加混合强风扰动 $\boldsymbol{\omega}_{i3}(t) = [0.2 + 0.1\sin(0.5t),\ 0.2 + 0.1\cos(0.5t),\ 0.1]^{\mathrm{T}}$.

元学习模型离线训练阶段采用了基于任务分布的采样策略, 构建了涵盖多样化性能偏好、质量偏差及环境不确定性的任务空间 $p(\mathcal{S})$. 其中, 无人机质量偏差 ΔM_i 在 $[-0.2,\ 0.2]$ kg 范围内均匀采样, 系统不确定性上界 ρ_i 在 $[0.1,\ 2.0]$ 区间内随机生成. 针对性能偏好参数 φ_{i1}, 本文设计了双模态混合采样机制: 以 50% 的概率在 $[0.65,\ 0.80]$ 区间内采样以模拟高动态响应需求的紧急避障工况, 或在区间 $[0.05,\ 0.60]$ 内采样以模拟平稳巡航工况. 训练过程基于 MATLAB 深度学习工具箱具体实现, 共执行 $S_{\max} = 1\,000$ 轮次迭代, 任务批次大小设为 $B = 16$. 对于从 $p(\mathcal{S})$ 中采样的每个任务, 通过随机初始化误差状态并执行短时动力学仿真, 收集包含 $H = 10$ 个连续时间步的支持集与查询集轨迹数据. 策略网络 π_θ 构建为包含两个隐藏层的全连接网络, 每层配置 64 个神经元. 元参数 θ 的更新采用 Adam 优化器, 内循环与外循环学习率分别设定为 $\alpha_1 = 0.005\,0$ 和 $\alpha_2 = 0.000\,5$, 惩罚权重配置为 $\beta_{i1} = 0.001\,7$ 及 $\beta_{i2} = 0.010\,0$. 在在线验证阶段, 系统维护一个长度为 $H = 10$ 的滑动窗口作为实时支持集 $\tilde{\mathcal{Q}}_i$. 为了动态平衡安全性与控制性能, 实验设定了自适应性能偏好权重调整策略: 当检测到无人机与最近障碍物的距离小于 8.0 m, 或处于强环境扰动时段 $t \in [130,\ 160]$ s 时, 系统自动将瞬态性能偏好权重 φ_{i1} 切换至 0.7 以增强快速响应能力; 在其他常规飞行阶段, 该权重保持在 $\varphi_{i1} = 0.5$ 以兼顾稳态控制精度.

人工势场相关的参数设定为轨迹跟踪增益 $k_{i1} = 0.10$、集群协同增益 $k_{i2} = 0.08$、避障排斥增益 $k_{i3} = 0.04$、吸引系数 $a_{ij} = 0.2$、排斥系数 $b_{ij} = 1.8$ 及指数因子 $q = 1.0$. 障碍物半径设为 $R_l = 3.0$ m, 无人机安全半径 $r_i = 0.8$ m. 指令滤波器的带宽参数设定为 $k_{i4} = 3.0$. 自适应增益函数的相关参数设定为 $\sigma_{i2} = 1.0$ 与 $\sigma_{i3} = 0.1$, 辅助控制增益设为 $K_{i3} = 2.1$, 自适应更新律的增益与阈值分别设为 $\kappa_i = 2.0$ 和 $\epsilon_i = 0.01$. 动态障碍物规避策略的相关参数为预测时域系数 $c_{i1} = 0.6$、避障安全余量 $r_0 = 0.3$ m、紧急避障推力 $F_i = 50.0$ N; 障碍物探测范围设定为

$d_{i0} = 4.0$ m, 速度调节系数分别为 $c_{i2} = 2.0$ 与 $c_{i3} = 1.5$, 幅值上限设定为 $\bar{F} = 20.0$ N.

5.2　实验结果及分析

为全面评估所提方法在多样化任务环境下的适应性与鲁棒性, 本文构建了四个包含不同动态障碍物分布及参考轨迹的实验场景, 相应的实验结果如图 2 ~ 图 7 所示. 在此基础上, 为了对算法性能进行更深入的量化评估与对比分析, 图 8 ~ 图 10 以及表 1 详细展示了在实验场景 1 下的具体实验数据.

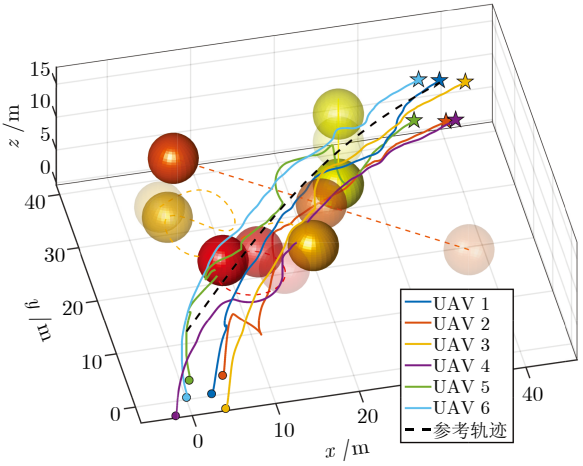


图 2　场景 1 下无人机集群的飞行轨迹

Fig. 2　Flight trajectories of the UAV swarm in Scenario 1

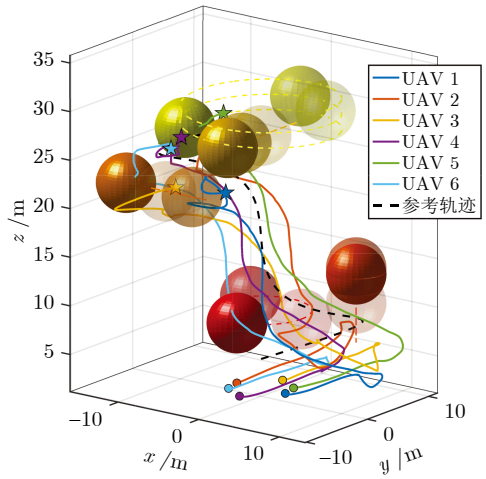


图 3　场景 2 下无人机集群的飞行轨迹

Fig. 3　Flight trajectories of the UAV swarm in Scenario 2

图 2 ~ 图 5 直观展示了无人机集群在包含障碍物的三维空间内的飞行轨迹演化过程. 观察轨迹曲线可知, 在任务全周期内, 无人机集群整体保持对参考轨迹 $\boldsymbol{p}_0$ 的稳定跟踪. 当检测到前方存在障碍物威胁时, 各无人机能够基于安全控制律自发生成平滑且有效的规避机动, 在严格保持与障碍物及邻机安全间隔的同时, 未发生任何碰撞现象. 上述结果充分验证了本文所提方法在处理非线性动力学、强耦合约束及环境不确定性方面的有效性, 成功实现了复杂动态场景下的高性能协同跟踪与可靠安全避障.

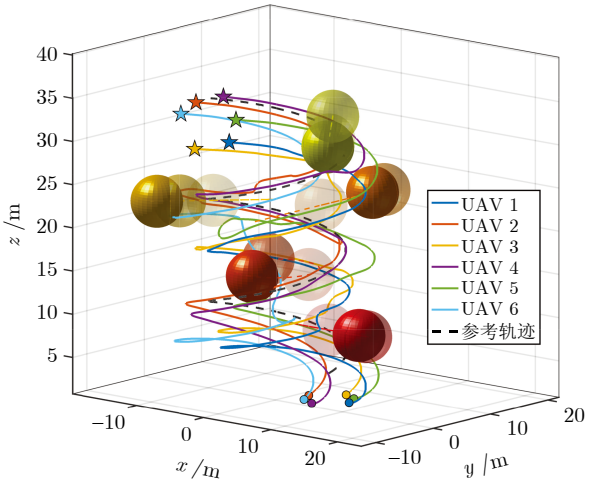


图 4　场景 3 下无人机集群的飞行轨迹

Fig. 4　Flight trajectories of the UAV swarm in Scenario 3

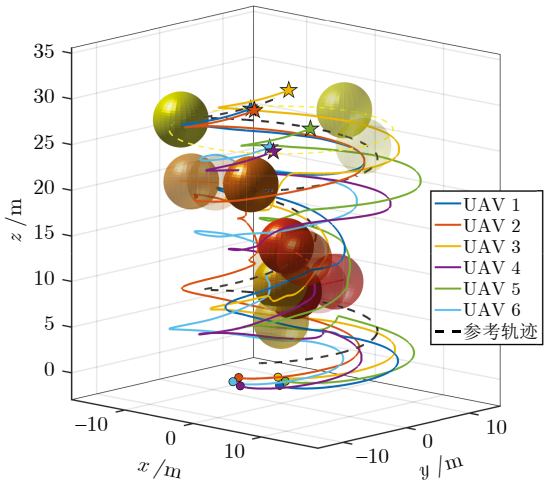


图 5　场景 4 下无人机集群的飞行轨迹

Fig. 5　Flight trajectories of the UAV swarm in Scenario 4

定义 $d_{i,\,\min} = \min_l \|\boldsymbol{p}_i - \boldsymbol{\varpi}_l\|$ 为无人机 i 与最近障碍物中心的欧氏距离. 图 6 展示了在四种不同动态场景下该最小距离随时间的变化曲线, 红色虚线代表了由障碍物半径 R_l 与无人机半径 r_i 构成的物理碰撞边界 3.8 m. 观察曲线可知, 尽管在规避过程中无人机与障碍物之间的最小距离 $d_{i,\,\min}$ 迅速逼

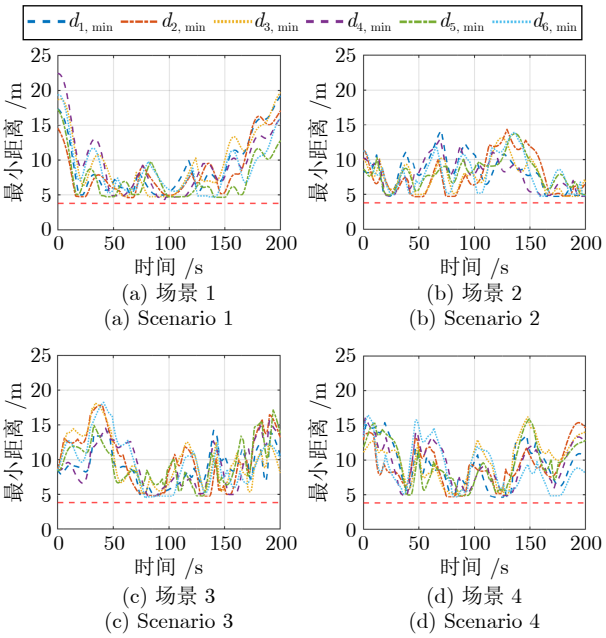


图 6 四种不同场景下无人机与障碍物的最小距离

Fig. 6 Minimum distance between UAVs and obstacles in four different scenarios

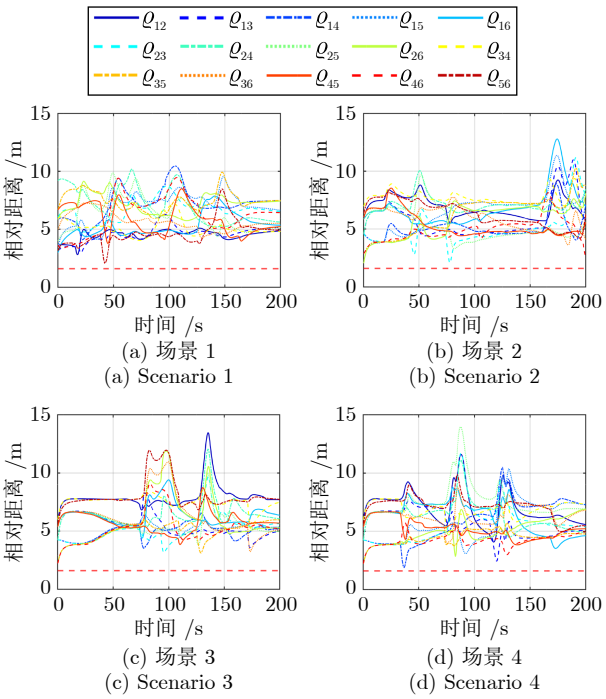


图 7 四种不同场景下无人机之间的相对距离

Fig. 7 Relative distances between UAVs in four different scenarios

近安全边界, 但始终未跌破该物理阈值. 这一现象验证了元学习增益更新策略的有效性, 通过生成具有更快响应速度的控制增益, 确保动力学底层能够准确跟踪避障指令, 从而在保证极高机动性的同时

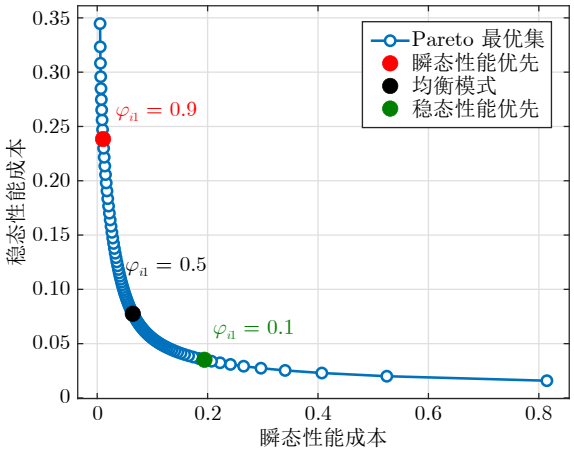


图 8 性能均衡优化前沿与不同偏好模式

Fig. 8 Performance balancing optimization front and different preference modes

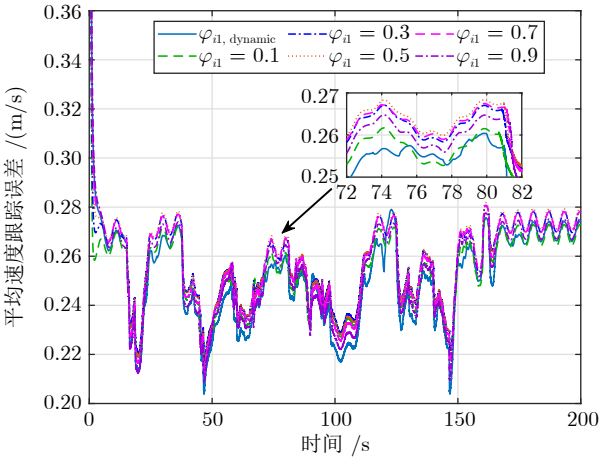


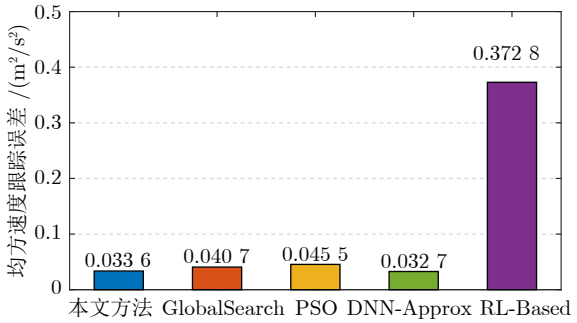
图 9 不同性能偏好权重下的平均速度跟踪误差 E

Fig. 9 Average velocity tracking errors E under different performance preference weights

实现了可靠的安全防护.

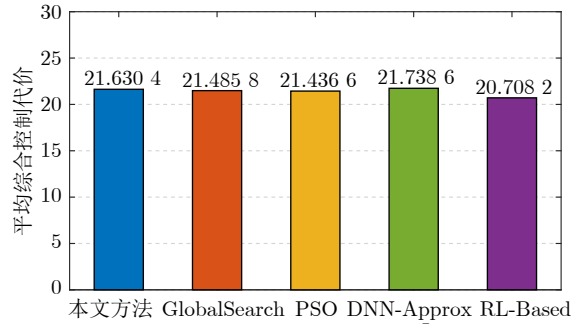
定义 $\varrho_{ij} = \|\boldsymbol{p}_i - \boldsymbol{p}_j\|$ 为集群内部无人机 i 与 j 之间的相对欧氏距离. 图 7 详细记录了在四种测试场景下相对距离随飞行时间的动态演化过程, 图 7 中红色虚线标示了由无人机物理半径决定的最小安全碰撞阈值 $r_{ij} = 1.6$ m. 尽管机间距离因轨迹调整出现了不同程度的收敛波动, 但所有无人机的相对距离始终严格维持在安全阈值之上. 这一结果有力证明了所设计的集群协同与避碰控制项 $\boldsymbol{\delta}_{i2}$ 能够在多重约束耦合及外部干扰存在的复杂工况下, 有效调节无人机间的相对运动状态, 在保障编队拓扑弹性的同时实现了零碰撞的安全协同飞行.

为综合评估所提方法在计算效率与寻优精度之间的协同权衡能力, 本文选取所有对比算法在同一时刻求解出的最小综合加权成本作为理论最优基



(a) 均方速度跟踪误差 $\bar{E}$

(a) Mean square velocity tracking errors $\bar{E}$



(b) 平均综合控制代价 $\bar{J}$

(b) Average composite control cost $\bar{J}$

图 10　不同控制策略下的跟踪误差与控制代价对比

Fig. 10　Comparison of tracking errors and control cost under different control strategies

准, 并以此计算各算法的增益偏差与成本偏差. 表 1 详细列出了各算法的性能指标. 分析数据可知, 尽管 GlobalSearch 算法能够以接近零的偏差逼近理论最优解, 但其 33.505 1 ms 的平均在线耗时远超无人机典型控制周期 5 ~ 10 ms 的时间容限, 无法满足实时性要求; Interior-Point 算法虽然精度较高, 但其 4.907 5 ms 的平均在线耗时已占据了控制周期的主要部分, 缺乏足够的计算安全余量. 在智能优化算法方面, GA 和 PSO 算法难以同时兼顾高频实时控制与高精度寻优的要求, 在有限的计算时间内产生了较大的成本偏差. 虽然 TD3 和 DNN-Approx 等深度学习方法实现了 0.166 9 ms 和 0.044 6 ms 的极致推理速度, 但其成本偏差高达 80.483 9% 和 74.015 7%, 揭示了纯数据驱动方法在处理复杂动态环境约束时存在的泛化瓶颈. 相比之下, 本文提出的元学习增益更新策略成功将平均在线耗时压缩至 1.674 1 ms, 其成本偏差仅为 0.634 1%, 有力证实了所提方法能够以极低的计算代价换取高度可靠的实时响应能力与寻优精度, 实现了实时性与精确性的最佳平衡. 此外, 尽管本文方法需要一定的离线训练时间成本, 但作为一次性计算投入, 这种策略以可接受的离线代价换取了在线阶段毫秒级的极速响应与高精度控制, 完全符合无人机高频控制的工程需求.

图 8 描绘了系统瞬态与稳态性能成本之间的性能均衡优化前沿. 该曲线是通过对性能偏好权重 φ_{i1} 在区间 [0.01, 0.99] 内进行高密度遍历求解生成的. 针对不同性能偏好权重计算出的最优控制增益构成了图 8 中所示的连续凸前沿, 清晰刻画了系统在不同偏好权重下的性能边界. 该曲线呈现出典型的光滑凸状特性, 揭示了改善瞬态性能必然伴随稳态代价增加的制约规律, 直观证实了所构建的多目标优化框架能够有效覆盖完整的性能均衡空间.

为量化无人机集群在动态环境下的协同跟踪精度, 定义平均速度跟踪误差为 $E = (1/N)\sum_{i=1}^{N}\|\boldsymbol{v}_i - \boldsymbol{\phi}_i\|$. 图 9 展示了该误差指标在全飞行周期内的动态演化过程. 观察曲线可知, 固定性能偏好权重 φ_{i1} 体现了显著的性能局限性, 较高的偏好权重导致了较大的平均跟踪误差, 表明过分侧重瞬态响应反而可能引入不必要的控制振荡. 虽然较低的权重可以将稳态误差维持在相对较低的水平, 但相比之下, 本文提出的采用动态性能偏好权重的策略 $\varphi_{i1,\,\mathrm{dynamic}}$ 展现了更优的综合性能, 成功融合了不同性能偏好模式的优势, 在全飞行周期内实现了跟踪精度与抗扰能力的综合优化.

为了综合评估各类算法在长时飞行任务中的综合效能, 本文引入量化指标均方速度跟踪误差 $\bar{E}$, 其计算方法为:

表 1　不同算法的详细性能指标对比

Table 1　Detailed performance metric comparison of different algorithms

算法类别	具体算法	离线训练 (h)	平均在线耗时 (ms)	增益偏差 K_{i1}^{*} (%)	增益偏差 σ_{i1}^{*} (%)	成本偏差 J_i (%)
数值优化	Interior-Point	—	4.907 5	0.023 9	0.080 3	0.000 2
	GlobalSearch	—	33.505 1	0.000 6	0.000 4	0.000 1
智能优化	GA	—	5.916 8	9.192 3	11.283 7	4.972 5
	PSO	—	2.199 8	6.202 7	7.981 1	1.436 6
深度学习	TD3	5.47	0.166 9	12.685 8	59.916 0	80.483 9
	DNN-Approx	0.08	0.044 6	54.577 0	1.593 0	74.015 7
元学习	本文方法	5.13	1.674 1	5.662 1	4.499 7	0.634 1

$$\bar{E} = \frac{1}{T_m}\int_0^{T_m}\frac{1}{N}\sum_{i=1}^{N}\|\boldsymbol{v}_i(t)-\boldsymbol{\phi}_i(t)\|^2\mathrm{d}t \tag{52}$$

以及平均综合控制代价 $\bar{J}$, 其计算方法定义为:

$$\bar{J} = \frac{1}{T_m}\int_0^{T_m}\left(\frac{1}{N}\sum_{i=1}^{N}\left(\|\boldsymbol{\delta}_{i1}(t)\|^2+\varkappa\|\boldsymbol{u}_i(t)\|^2\right)\right)\mathrm{d}t \tag{53}$$

其中, 权重系数取 $\varkappa = 0.01$, 用于调节误差收敛与控制能量开销之间的相对重要性.

图 10 对比了五种不同控制策略下的跟踪误差与控制代价. 分析图 10(a) 可知, 本文方法、GlobalSearch、PSO 以及 DNN-Approx 在均方速度跟踪误差 $\bar{E}$ 上均具备同等水平的高精度跟踪特性, 且显著优于基于强化学习的控制策略 RL-Based[32]. 实验结果表明, 本文方法成功达到了与 GlobalSearch 等复杂迭代优化算法相当的跟踪精度. 进一步观察图 10(b) 的综合控制代价统计, 部分策略较高的综合代价表明其跟踪精度的提升是以牺牲控制能效为代价的, 验证了元学习策略能够实现系统控制精度与能效水平的协同优化. 结合表 1 的计算效率数据, 鉴于传统优化算法因计算耗时过高而难以应用于高频实时控制回路, 本文提出的元学习增益更新策略凭借毫秒级的在线推理速度, 实现了对数值最优解的精准复现, 从而有力证实了其是统筹实时性、精确性与能效性的最佳协同控制方案.

6 结束语

本文的研究工作致力于解决无人机集群在非结构化动态环境下的自适应协同控制难题, 主要贡献在于建立了一种基于动态性能均衡的元学习控制框架. 通过理论推导, 证实了所提出的自适应控制策略能够有效补偿模型不确定性, 确保系统状态始终收敛于李雅普诺夫稳定域内. 针对瞬态响应与稳态精度之间的性能冲突, 本研究构建了关键增益的性能均衡优化机制, 并进一步通过元学习策略成功突破了在线计算瓶颈, 将复杂的参数寻优过程转化为离线经验学习与在线快速推理. 实验结果表明, 在毫秒级单步更新机制的驱动下, 该控制架构不仅具备处理动态障碍物与剧烈扰动的强鲁棒性, 更在计算效率与控制品质的动态均衡上展现出显著优越性, 为高频实时系统的智能化设计提供了新的理论范式.

参考文献

1 Li Zhong-Kui, Wang Jun-Jie, Zhang Yun-Yi, Zhang Shuo, Guo Meng, Sun Zhi-Yong. Formal logic-based cooperative task planning for multi-robot systems: Survey of recent advances and future directions. *Acta Automatica Sinica*, 2025, **51**(10): 2211−2231
(李忠奎, 王俊杰, 张云奕, 张硕, 国萌, 孙志勇. 集群协同任务规划的形式逻辑方法: 综述与展望. 自动化学报, 2025, **51**(10): 2211−2231)

2 Xu T, Yi X J, Wen G H. Distributed fuzzy formation control of multi-UAV systems with directed communication networks. *IEEE Transactions on Fuzzy Systems*, 2025, **33**(10): 3582−3594

3 Wen Guang-Hui, Yu Xing-Huo, Huang Ting-Wen, Zhou Yan. Fixed-time bipartite formation tracking control for multi-USV systems with uncertain model parameters. *Acta Automatica Sinica*, 2025, **51**(3): 669−677
(温广辉, 余星火, 黄廷文, 周艳. 模型参数不确定下多无人艇系统固定时间二分编队跟踪控制. 自动化学报, 2025, **51**(3): 669−677)

4 Cheng M, Liu H, Wen G H, Lewis F L. Multiquadrotor differential graphical containment game with noncooperative active leaders subject to external disturbances. *IEEE Transactions on Automatic Control*, 2026, **71**(1): 490−495

5 Ren Z M, Liu H, Wen G H, Lü J H. Event-triggered data-driven security formation control for quadrotors under denial-of-service attacks and communication faults. *IEEE Transactions on Cybernetics*, 2025, **55**(11): 5226−5236

6 Xu B Y, Dai Y F, Suleman A, Shi Y. Distributed fault-tolerant control of multi-UAV formation for dynamic leader tracking: A Lyapunov-based MPC framework. *Automatica*, 2025, **175**: Article No. 112179

7 Yang H J, Sun S P, Xia Y Q, Li P. Digital twin-based obstacle avoidance for unmanned aerial vehicles using feedforward-feedback control. *IEEE Transactions on Vehicular Technology*, 2025, **74**(6): 8721−8733

8 Du Z X, Zhang H, Wang Z P, Yan H C. Model predictive formation tracking-containment control for multi-UAVs with obstacle avoidance. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2024, **54**(6): 3404−3414

9 Liu B W, Wang Y X, Mofid O, Mobayen S, Khooban M H. Barrier function-based backstepping fractional-order sliding mode control for quad-rotor unmanned aerial vehicle under external disturbances. *IEEE Transactions on Aerospace and Electronic Systems*, 2024, **60**(1): 716−728

10 Pang W, Zhu D Q, Sun C Y. Multi-AUV formation reconfiguration obstacle avoidance algorithm based on affine transformation and improved artificial potential field under ocean currents disturbance. *IEEE Transactions on Automation Science and Engineering*, 2024, **21**(2): 1469−1487

11 Zhang X Y, Zong H, Wu W H. Cooperative obstacle avoidance of unmanned system swarm via reinforcement learning under unknown environments. *IEEE Transactions on Instrumentation and Measurement*, 2025, **74**: Article No. 7500615

12 Wang D, Chen H M, Lao S H, Drew S. Efficient path planning and dynamic obstacle avoidance in edge for safe navigation of USV. *IEEE Internet of Things Journal*, 2024, **11**(6): 10084−10094

13 Yu Y J, Chen C, Guo J, Chadli M, Xiang Z R. Adaptive formation control for unmanned aerial vehicles with collision avoidance and switching communication network. *IEEE Transactions on Fuzzy Systems*, 2024, **32**(3): 1435−1445

14 Fang Hao, Zhao Xin-Yue, Chen Jie. Autonomous control of unmanned aerial vehicle swarms: Prescribed performance driven safety formation control. *Acta Automatica Sinica*, 2025, **51**(5): 931−941
(方浩, 赵欣悦, 陈杰. 无人飞行器集群自主控制: 预设性能驱动的安全编队控制. 自动化学报, 2025, **51**(5): 931−941)

15 Wang Yao-Nan, Hua He-An, Zhang Hui, Zhong Hang, Fan Ye-Xin, Liang Hong-Tao, et al. Performance function-guided deep reinforcement learning control for UAV swarm. *Acta Automatica Sinica*, 2025, **51**(5): 905−916
(王耀南, 华和安, 张辉, 钟杭, 樊叶心, 梁鸿涛, 等. 性能函数引导的无人机集群深度强化学习控制方法. 自动化学报, 2025, **51**(5): 905−916)

16 Zhao X M, Cui Z R, Dong F F, Chen Y H, Huang J. Fuzzy game-theoretic control design for unmanned ground swarm systems: An integrated method. *IEEE Transactions on Fuzzy Systems*, 2026, **34**(2): 368−381

17 Liu J L, Liu Z H, Li Y, Tian E G, Peng C. Privacy-protected formation control for unmanned surface vehicles: A neural predictor-based noncooperative game framework. *IEEE Transactions on Vehicular Technology*, 2026, **75**(7): 12137−12147

18 Liu J L, Dong Y H, Zha L J, Xie X P, Tian E G. Reinforce-

ment learning-based tracking control for networked control systems with DoS attacks. *IEEE Transactions on Information Forensics and Security*, 2024, **19**: 4188−4197

19 Liu J L, Zhang N, Zha L J, Xie X P, Tian E G. Reinforcement learning-based decentralized control for networked interconnected systems with communication and control constraints. *IEEE Transactions on Automation Science and Engineering*, 2024, **21**(3): 4674−4685

20 Hou P, Huang Y, Zhu H B, Lu Z H, Huang S C, Yang Y, et al. Distributed DRL-based intelligent over-the-air computation in unmanned aerial vehicle swarm-assisted intelligent transportation system. *IEEE Internet of Things Journal*, 2024, **11**(21): 34382−34397

21 Zhou Z H, Tang J, Feng W M, Zhao N, Yang Z T, Wong K K. Optimized routing protocol through exploitation of trajectory knowledge for UAV swarms. *IEEE Transactions on Vehicular Technology*, 2024, **73**(10): 15499−15512

22 Jia R H, Fu Q Y, Zheng Z L, Zhang G L, Li M L. Energy and time trade-off optimization for multi-UAV enabled data collection of IoT devices. *IEEE/ACM Transactions on Networking*, 2024, **32**(6): 5172−5187

23 Liu X B, Wang Y, Gao H, Ngai E C, Zhang B, Wang C H, et al. A coverage-aware task allocation method for UAV-assisted mobile crowd sensing. *IEEE Transactions on Vehicular Technology*, 2024, **73**(7): 10642−10654

24 Li K X, Liu J J, Gu X, Yang Y D, Chang C, Chen H P, et al. Dynamic decision-making of UAV swarm based on constrained multi-objective optimization under incomplete interference information. *Chinese Journal of Aeronautics*, 2026, **39**(7): Article No. 103846

25 Peng C D, Wu Z X, Huang X M, Wu Y, Kang J W, Huang Q, et al. Joint energy and completion time difference minimization for UAV-enabled intelligent transportation systems: A constrained multi-objective optimization approach. *IEEE Transactions on Intelligent Transportation Systems*, 2024, **25**(10): 14040−14053

26 Li C L, Deng C Y, Zhang Y, Wan S H. Federated meta-learning based computation offloading approach with energy-delay tradeoffs in UAV-assisted VEC. *IEEE Transactions on Mobile Computing*, 2025, **24**(10): 10978−10991

27 Zhang Y, Wen L, Huang Y H, Sun S M, Bing Z S, He W, et al. Meta-learning-based safety-critical control in multi-obstacles environments. *IEEE Transactions on Automation Science and Engineering*, 2025, **22**: 15299−15313

28 Wei M X, Zheng L X, Wu Y, Mei R D, Cheng H. Meta-learning enhanced model predictive contouring control for agile and precise quadrotor flight. *IEEE Transactions on Robotics*, 2025, **41**: 3590−3608

29 Ye F Y, Lin B J, Yue Z X, Zhang Y, Tsang I W. Multi-objective meta-learning. *Artificial Intelligence*, 2024, **335**: Article No. 104184

30 Li S X, Pang Y S, Huang Z R, Chu X H. An offline-online learning framework combining meta-learning and reinforcement learning for evolutionary multi-objective optimization. *Swarm and Evolutionary Computation*, 2025, **97**: Article No. 102037

31 Zhang Z Z, Wu Z Y, Zhang H, Wang J H. Meta-learning-based deep reinforcement learning for multiobjective optimization problems. *IEEE Transactions on Neural Networks and Learning Systems*, 2023, **34**(10): 7978−7991

32 Ma Q, Jin P, Lewis F L. Guaranteed cost attitude tracking control for uncertain quadrotor unmanned aerial vehicle under safety constraints. *IEEE/CAA Journal of Automatica Sinica*, 2024, **11**(6): 1447−1457



苗金钊 南京信息工程大学计算机学院博士研究生. 主要研究方向为无人系统, 强化学习和神经网络.
E-mail: 202511200006@nuist.edu.cn
(**MIAO Jin-Zhao** Ph.D. candidate at the School of Computer Science, Nanjing University of Information Science and Technology. His research interests include unmanned systems, reinforcement learning, and neural networks.)



刘金良 南京信息工程大学计算机学院教授. 主要研究方向为信息物理系统, 隐私保护和博弈论. 本文通信作者.
E-mail: liujinliang@vip.163.com
(**LIU Jin-Liang** Professor at the School of Computer Science, Nanjing University of Information Science and Technology. His research interests include cyber-physical systems, privacy preservation, and game theory. Corresponding author of this paper.)



田恩刚 上海理工大学光电信息与计算机工程学院教授. 主要研究方向为网络化控制系统, 网络攻击和非线性随机控制.
E-mail: tianengang@163.com
(**TIAN En-Gang** Professor at the School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology. His research interests include networked control systems, cyber attack, and nonlinear stochastic control.)



彭 晨 上海大学机电工程与自动化学院教授. 主要研究方向为多智能体系统, 电力系统和互联系统.
E-mail: c.peng@shu.edu.cn
(**PENG Chen** Professor at the School of Mechatronic Engineering and Automation, Shanghai University. His research interests include multi-agent systems, power systems, and interconnected systems.)



黄廷文 深圳理工大学计算机科学与控制工程学院教授. 主要研究方向为机器学习, 隐私保护和神经网络.
E-mail: huangtw2024@163.com
(**HUANG Ting-Wen** Professor at the Faculty of Computer Science and Control Engineering, Shenzhen University of Advanced Technology. His research interests include machine learning, privacy preservation, and neural networks.)