

复杂环境下多无人机协同围捕的分层规划与控制

褚晶¹, 张妮¹, 黄勇², 刘雨濛¹

(1.西安邮电大学 自动化学院, 陕西 西安 710061; 2.西北工业大学 航天学院, 陕西 西安 710072)

摘要:在复杂环境下,多无人机协同围捕入侵目标时,常面临目标轨迹预测精度欠佳、任务分配过程中系统能耗过高及动态环境下避障能力不足等关键挑战,严重制约了围捕任务的效率与可靠性。针对上述问题,提出一种包含聚集阶段与引导阶段的两阶段多无人机协同围捕框架。在聚集阶段,为实现对入侵目标的提前拦截,设计插值与误差校正的多项式拟合(IEC-PF)方法,有效提升目标轨迹预测精度;针对围捕点分配问题,提出基于迭代凸优化的任务分配策略,在严格满足最终围捕编队约束条件的前提下,实现系统总能耗最小化。在引导阶段,无人机依据分配的围捕点构建最优包围编队,基于生成的无障碍凸区域,通过非线性优化算法求解最优全局路径及编队构型参数;引入基于应力矩阵的仿射编队控制策略,保障无人机在路径跟踪过程中编队的稳定与重构;融入基于深度强化学习(DRL)的局部避障机制,显著提升系统对动态障碍物的实时响应能力。仿真实验结果表明,所提框架能够显著增强系统在复杂环境下的避障性能,有效确保多无人机成功协同围捕入侵目标并将其引导至目标区域。

关键词:多智能体协同围捕;轨迹预测;任务分配;非线性优化;仿射编队跟踪控制;动态避障
中图分类号:V249 **文献标志码:**A **文章编号:**1000-2758(2026)02-0380-13

低空经济作为新兴经济形态,其快速发展推动了无人机在物流配送、城市巡检、应急救援等领域的广泛应用,但“乱飞”等违规问题频发,无人机未经报备闯入敏感低空空域,会干扰交通秩序、威胁人员财产安全^[1]。而传统围捕方法中人工巡查与识别效率低,难以应对复杂低空空域内的无人机活动,现有固定式围捕系统则存在部署成本高、灵活性不足的问题,难以适配动态防控需求,导致围捕响应滞后、漏捕率高。当前低空空域无人机安全管控与围捕能力亟待提升。

多智能体协同围捕作为应对动态目标拦截的有效手段,已成为无人系统领域的研究热点,并取得一系列进展。早期研究以仿生启发式算法为核心。Cao 等^[2]提出仿生协同搜索学习算法(biomimetic cooperative search learning algorithm, BCSLA),实现

动态目标捕获与避障强化;吴子沉等^[3]通过神经网络实现协同围捕,但仅适用于单任务场景。近年来,研究重心逐渐转向强化学习(reinforcement learning, RL)方法。杨梓豪等^[4]将基于选项的分层强化学习方法引入到多无人机目标协同围捕过程,但仅适用于小型无人机集群的简单任务;Zheng 等^[5]融合博弈论与 Q-learning 学习,优化围捕与目标追踪效果;孙懿豪等^[6]结合深度强化学习(deep reinforcement learning, DRL)与博弈论,提升未知环境下的捕获效率;针对无人水面艇(unmanned surface vehicles, USVs),Xia 等^[7]提出多智能体近端策略优化(multi-agent proximal policy optimization, MAPPO),优化围捕效率与安全性。

围捕任务通常涉及对智能体的路径规划及多智能体的控制。路径规划对于围捕任务的可靠运行至关重要,需综合考虑各类约束以生成安全无碰撞的路径。传统路径规划算法(Dijkstra、A*、APF)^[8-10]对动态障碍物响应滞后;环境分解方法(如梯形分解和三角形分解)^[11-12]虽能实现完整区域划分,但生成的单元较小,会降低规划效率。Voronoi 方法能

收稿日期:2025-10-12

基金项目:国家自然科学基金(61703336)与陕西省自然科学基金(2023-JC-QN-0727)资助

作者简介:褚晶(1989—),副教授

通信作者:黄勇(1976—),副教授 e-mail:huangyong@nwpu.edu.cn

有效定义静态环境的障碍边界^[13],但在实际应用中,用凸区域近似非凸自由空间往往更有效。Liu等^[14]提出的“安全飞行走廊”通过构建重叠凸集增强避障能力;Dai等^[15]采用双层优化方法迭代生成多边形,在最大化椭球体积的同时分离智能体与障碍物,但计算成本较高;Marcucci等^[16]采用矩形表示无碰撞通道以实现更平滑的路径设计;Chen等^[17]利用局部感知点云数据生成星形区域,但其非凸性增加了编队优化的复杂性。

为应对动态环境,DRL已被应用于基于视觉输入与传感器融合的智能体导航及障碍物规避^[18]。深度确定性策略梯度(deep deterministic policy gradient,DDPG)已被应用于三维城市环境中的无人机导航^[19],但仍存在探索效率低下和实时适应性差的问题^[20]。由于强化学习依赖试错机制,智能体在训练初期难以快速习得有效策略,导致对样本量要求较高^[21]。为解决此问题,Niu等^[22]提出基于人类示范的端到端强化学习方法,在DDPG框架中融入优先经验回放(prioritized experience replay,PER)机制,显著提升学习效率。该方法可实现实时导航避障,但在已知特定环境中应用时,可能因耗时长、探索路径冗长导致不必要的能量消耗。

在多智能体的编队控制方面,传统控制方法存在明显局限:领导者-跟随者方法^[23]易过度依赖领导者,易因领导者故障影响整体;基于行为方法^[24]难以定义整体行为,协调性不足;虚拟结构法^[25]对动态环境适应性差,难以应对突发障碍。近年来,编队控制通过融合一致性理论,引入了复拉普拉斯矩阵^[26]和应力矩阵^[27],提升目标编队的定义精度。Han等^[28]提出的分布式控制策略,借助复拉普拉斯矩阵定义编队形状,可确保智能体收敛至目标构型,但仅适用于二维环境;应力矩阵也是实现编队机动的有效工具,Zhao^[27]提出仿射编队控制律,能在机动过程中实现动态调整编队形状,但局限于简单线性运动学,难以应用于非线性运动学模型。

尽管上述研究已分别解决了围捕任务中各个子问题,但仍存在不足,例如目标预测精度低、任务分配效率低下、避障能力不足。为应对上述挑战,本文提出一种分层协作围捕框架,将围捕任务分解为聚集与引导2个控制阶段。每个阶段均针对特定子问题设计策略,从而实现动态环境下高效、灵活且安全的多无人机围捕。本文主要贡献包括:

1) 提出一种分层控制架构,将轨迹预测、能量

最优的任务分配、全局路径规划、局部动态避障、编队控制结合,完成对入侵目标的围捕;

2) 提出基于迭代凸优化的围捕点分配方法,求解能量最优的分配并构建满足约束的围捕编队,结合插值及误差校正的多项式拟合(interpolation and error correction polynomial fitting,IEC-PF)方法,精准预测入侵目标轨迹并实现拦截,降低无人机能耗;

3) 提出融合了基于非线性优化的全局路径规划与基于深度学习的局部动态避障的策略,通过非线性优化全局路径,得到具有最优编队构型参数的无碰撞路径,并由深度学习策略实现局部动态避障,进一步保障围捕的安全性与实时性。

1 问题描述

1.1 图论

考虑在空间 $\mathbf{R}^d$ 中由 n 个固定翼无人机组成的多智能体系统,各无人机间的拓扑关系由无向图 $\mathbf{G}=(\mathbf{V},\mathbf{E})$ 描述。图的节点集记为 $\mathbf{V}=\{1,2,\dots,n\}$,边集 $\mathbf{E}\subseteq\mathbf{V}\times\mathbf{V}$ 表示节点之间的边集合。边 $(i,j)\in\mathbf{E}$ 表示无人机 i 从无人机 j 获取信息。编队 $(\mathbf{G},\mathbf{p})$ 将多个固定翼无人机关联为无向图 $\mathbf{G}$,其中 $\mathbf{p}_i\in\mathbf{R}^d$ 表示无人机 i 的位置。无人机分为领导者和跟随者2类, $\mathbf{V}_l=\{1,2,\dots,n_l\}$ 被定义为领导者, $\mathbf{V}_f=\mathbf{V}\setminus\mathbf{V}_l$ 为跟随者,位置分别由 $\mathbf{p}_{n_l}=[\mathbf{p}_1^T,\mathbf{p}_2^T,\dots,\mathbf{p}_{n_l}^T]^T$ 和 $\mathbf{p}_f=[\mathbf{p}_{n_l+1}^T,\mathbf{p}_{n_l+2}^T,\dots,\mathbf{p}_n^T]^T$ 表示。

1.2 应力矩阵

应力矩阵 $\boldsymbol{\Omega}$ 可以同时表征编队的通信拓扑结构与几何约束条件。对于编队 $(\mathbf{G},\mathbf{p})$,应力定义为分配给该编队边集的一组标量参数,即对于编队中的任意边 $(i,j)\in\mathbf{E}$,均对应一个标量应力值 $\tilde{\omega}_{ij}$ 。当应力值满足 $\sum_{j\in\mathbf{N}_i}\tilde{\omega}_{ij}(\mathbf{p}_j-\mathbf{p}_i)=0$ 时,编队的平衡应力状态即得以建立。

平衡约束的矩阵形式为 $(\boldsymbol{\Omega}\otimes\mathbf{I}_d)\mathbf{p}=0$,其中 $\mathbf{I}_d$ 为单位矩阵, $\otimes$ 表示克罗内克积。 $\bar{\boldsymbol{\Omega}}=\boldsymbol{\Omega}\otimes\mathbf{I}_n$ 可根据领导者和跟随者被分为

$$\bar{\boldsymbol{\Omega}}=\begin{bmatrix}\bar{\boldsymbol{\Omega}}_{ll}&\bar{\boldsymbol{\Omega}}_{lf}\\\bar{\boldsymbol{\Omega}}_{fl}&\bar{\boldsymbol{\Omega}}_{ff}\end{bmatrix}\quad(1)$$

式中: $\bar{\boldsymbol{\Omega}}_{ll}\in\mathbf{R}^{(dN_l)\times(dN_l)}$, $\bar{\boldsymbol{\Omega}}_{lf}\in\mathbf{R}^{(dN_l)\times(dN_f)}$, $\bar{\boldsymbol{\Omega}}_{fl}\in\mathbf{R}^{(dN_f)\times(dN_l)}$, $\bar{\boldsymbol{\Omega}}_{ff}\in\mathbf{R}^{(dN_f)\times(dN_f)}$,如果仿射可定性条件成立^[27],则跟随者的位置可以由领导者的位置唯一

确定,即 $\mathbf{p}_f = -\bar{\mathbf{\Omega}}_{ff}^{-1} \bar{\mathbf{\Omega}}_{ff} \mathbf{p}_i$ 。

1.3 仿射编队

仿射变换是包含线性变换与平移操作的几何变换,可实现对目标对象的平移、缩放、剪切、旋转,以及由多个单一变换构成的复合变换(见图1)。

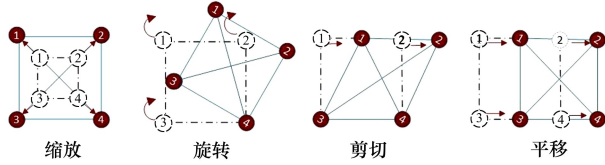


图1 仿射变换

标称编队被定义为 $(\mathbf{G}, \mathbf{r})$, $\mathbf{r} = [\mathbf{r}_1^T, \mathbf{r}_{i+1}^T, \dots, \mathbf{r}_n^T]^T$ 表示固定翼无人机编队的仿射变换基。时变仿射编队被定义为

$$\bar{\mathbf{p}}(t) = [\mathbf{I}_n \otimes \mathbf{A}(t)] \mathbf{r} + \mathbf{1}_n \otimes \mathbf{b}(t), \quad (2)$$

式中: $\mathbf{A}(t) \in \mathbf{R}^{d \times d}$ 是时变变换矩阵,包括选择、剪切、缩放; $\mathbf{b}(t)$ 表示平移; $\mathbf{1}_n$ 是单位向量。无人机 i 的期望位置由 $\bar{\mathbf{p}}_i(t) = \mathbf{A}(t) \mathbf{r}_i + \mathbf{b}(t)$ 得到。

因此,固定翼无人机的控制目标设定为跟踪其目标时变位置,即当 $t \rightarrow \infty$ 时,编队收敛于目标 $\mathbf{p}(t) \rightarrow \bar{\mathbf{p}}(t)$ 。

1.4 无障碍区域

假设在全局地图中存在一组静态障碍物, $\bar{\mathbf{O}}$ 定义膨胀障碍物集合,固定翼无人机在位置 $\mathbf{p}$ 所占据的区域为 $\mathbf{S}(\mathbf{p}) \subset \mathbf{R}^2$,则障碍物集合可以表示为

$$\bar{\mathbf{O}} = \{\mathbf{p} \in \mathbf{R}^2 | \mathbf{S}(\mathbf{p}) \cap \mathbf{O} \neq \emptyset\} \quad (3)$$

对于静态障碍物,自由空间可以表示为

$$\bar{\mathbf{F}} = \mathbf{R}^2 \setminus \bar{\mathbf{O}} \subset \mathbf{R}^2 \quad (4)$$

1.5 运动学模型

假设所有固定翼无人机均在同一高度飞行,则每架无人机可被简化为质点。基于此假设,固定翼无人机的运动方程可表述为^[29]

$$\begin{cases} \dot{x}_i = v_i \cos \theta_i \\ \dot{y}_i = v_i \sin \theta_i, & i = 1, 2, \dots, n \\ \dot{\theta}_i = \omega_i \end{cases} \quad (5)$$

式中: $\mathbf{p}_i = (x_i, y_i)^T$ 表示无人机 i 的位置; θ_i 为无人机 i 的航向角。 $\mathbf{X}_i = [x_i, y_i, \theta_i]^T$ 表示无人机 i 的状态变量,控制量为 $\mathbf{U}_i = [u_{i1}, u_{i2}]^T = [v_i, w_i]^T$,线速度 v_i 和角速度 w_i 分别受到(6)式约束。

$$\begin{cases} v_i \in [0, v_{\max}] \\ \omega_i \in [-\omega_{\max}, \omega_{\max}] \end{cases} \quad (6)$$

式中: $v_{\max}$ 分别为最大线速度; $\omega_{\max}$ 为最大角速度。

2 方法

为实现复杂环境下固定翼无人机的协同围捕任务,本文提出了一种包含聚集阶段与引导阶段的分层规划与控制框架。在聚集阶段,首先对入侵者轨迹进行预测,再结合地图环境信息与入侵者实时位置,完成围捕点的优化分配;在引导阶段,首先,基于无障碍凸区域,通过非线性优化算法求解全局无碰撞路径;同时,对于动态障碍物, DRL 算法利用领导者的激光雷达数据、目标点和速度信息,生成实时速度控制指令;最后,采用基于应力矩阵的仿射编队控制策略,进行目标跟踪与编队重构。该分层规划与控制框架的整体结构如图2所示。

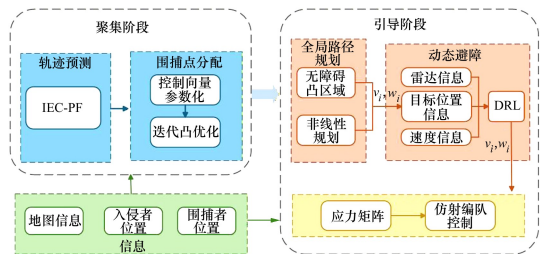


图2 多无人机协同围捕的分层规划与控制

2.1 聚集阶段

围捕无人机初始被随机部署于任务区域内,其首要任务是获取入侵者位置信息,基于该位置预测入侵者的短期运动轨迹,在入侵者周围生成目标围捕点,并完成围捕点的分配,使得所有无人机均能沿最短路径抵达各自分配的目标围捕点。

2.1.1 多项式拟合轨迹预测

在二维空间中,移动目标的运动轨迹是一条连续的曲线,而实际观测中仅能通过特定频率对其采样,得到离散轨迹点。根据采样定理可知,连续信号可以由固定时间间隔的离散采样点重构,因此,移动目标轨迹的预测问题可转化为对离散采样点的预测问题。

轨迹预测为基于有限数量的观测轨迹样本构建预测模型,再利用该模型估计目标随时间推移的未来位置。对于 n 个轨迹采样点 $P_1(x_1, y_1), P_1(x_2, y_2), \dots, P_n(x_n, y_n)$, 可以通过拟合一个 m 次多项式对轨迹进行近似表示,该多项式用于计算 x_i 对应的位置 $y_i (i = 1, 2, \dots, n)$, 表示为

$$y_i = a_0 + a_1x + a_2x^2 + \cdots + a_mx^m = \sum_{j=0}^m a_jx^j (m < n) \quad (7)$$

式中: m 表示多项式的阶数; $[a_0, a_1, \cdots, a_m]^T$ 为通过最小二乘法求解得到的多项式系数矩阵, 将系数矩阵代入(7)式中, 可以推导出曲线拟合的最优解, 利用该最优解, 即可实现对目标下一时刻轨迹点

$$P_{n+1} \text{ 的预测, 即 } P_{n+1} = \sum_{j=0}^m a_j x_{n+1}^j$$

对于复杂运动轨迹, 短期内目标的未来轨迹主要依赖近期运动状态, 与远期的历史轨迹无显著关联性。因此, 对入侵目标移动轨迹进行实时采样, 获取多项式拟合所需的历史轨迹点集, 假设目标短期沿拟合轨迹运动; 随着目标的移动, 采样数据实时更新并重新计算预测位置, 实现采样窗口动态滑动如图3所示, 目标移动时远期采样点被近期采样点逐步替换, 即实现采样窗口的动态滑动。

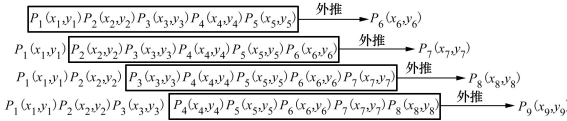


图3 动态滑动采样轨迹预测模型

采样频率决定的采样间隔会影响轨迹预测精度。相邻2次预测的误差呈现出一致的趋势, 该特性可用于制定预测值更新策略。由于预测误差随预测点数量增加而增大, 且同一预测位置误差大小相近, 本文进一步提出插值与误差校正多项式拟合(IEC-PF)方法, 通过对采样点插值获得更密集的点, 并利用预测值之间的误差更新预测值, 以提升目标轨迹预测的准确性。

该算法的输入是待拟合点集 P , 其由近期历史轨迹点构成。算法主要重复以下步骤, 输出最终预测点集合:

- 1) 初始化采样窗口大小 W 和预测次数 K ;
- 2) 对输入点集 P 进行插值, 生成更密集的点集 P_{new} ;
- 3) 执行 K 次拟合预测以生成 K 个预测点: 在第 i 次预测时, 选取 P_{new} 中最后的 W 个点 $\tilde{P}_i$ 进行曲线拟合, 计算得到新预测点 $P_{i, \text{temp}}$, 并将该点存入预测集合 P' 中, 该过程重复执行 K 次, 形成最终的初始预测点集合 P' ;
- 4) 实施误差修正以提升预测精度: 计算第 $i-1$

次预测的误差 $e_{i-1} = P'_{i-1} - P_{i-1, \text{new}}$, 其中 P'_{i-1} 表示预测值, $P_{i-1, \text{new}}$ 为对应时刻的实际观测值; 然后基于 e_{i-1} 进行单次误差修正, 更新预测值 $P'_{i, \text{new}}$, 并计算修正后的误差 $e_{i, \text{new}}$; 通过多轮误差修正, 最终得到目标预测结果 P^* 。

算法1 插值和误差修正多项式拟合

- 1 输入: 待拟合点集 P
- 2 输出: 预测点集 P^*
- 3 初始化滑动窗口大小 W 、预测点数 K
- 4 $P_{\text{new}} = \text{interp1}(P)$ // 获取更密集的点集
- 5 for $i = 1, 2, \cdots, K$
- 6 $\tilde{P}_i = P_{i, \text{new}}(P_{\text{end}} - W, P_{\text{end}})$ // 对于第 i 次预测, 选取最后 W 个
- 7 $P_{i, \text{temp}} = \text{curvefitPredict}(\tilde{P}_i)$ // 拟合预测点
- 8 $P' = \text{append}(P_{i, \text{temp}})$ // 保存每次预测结果
- 9 end for
- 10 // 误差修正
- 11 for $i = 2, 3, \cdots, K$
- 12 $e_{i-1} = P'_{i-1} - P_{i-1, \text{new}}$ // 第 $i-1$ 次的预测误差
- 13 $P'_{i, \text{new}} = P'_i - k_1 \times e_{i-1}$
- 14 $e_{i, \text{new}} = P'_{i, \text{new}} - P_{i, \text{new}}$
- 15 $P_i^* = P'_i - k_1 \times e_{i-1} - k_2 \times e_{i-1, \text{new}}$ // 最终预测值
- 16 end for

2.1.2 围捕点分配

在围捕任务中, 围捕点分配的核心目标是基于入侵者的预测位置, 为各无人机计算最优围捕点(见图4), 构建满足预设几何约束的围捕编队, 同时实现编队建立过程的能量消耗最小化。编队的几何约束可由一组参考参数确定, 例如圆形编队, 若已知圆心、半径及无人机数量, 参考参数可以视为无人机的圆周偏转角 δ , 相邻无人机的角度间隔由 $\phi = 2\pi/n$ 固定。因此, 围捕点的分配本质是通过优化参考参数, 结合无人机运动学约束、控制输入约束等, 使各无人机从初始位置运动至目标围捕点的系统总

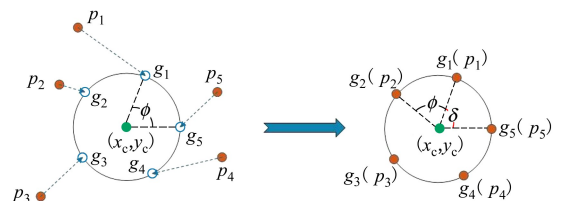


图4 围捕点分配

能耗最低。

为实现能耗最小化且满足相关约束,将围捕点分配与编队建立过程统一建模为最优控制问题。

以控制输入能耗总和最小为优化目标,目标函数定义为

$$J = \frac{1}{2} \int_{t_0}^{t_f} \sum_{i=1}^n \mathbf{U}_i^T \mathbf{U}_i dt \quad (8)$$

该最优控制问题受到的约束有:

1) 动力学约束:(5)式中的动力学约束可表述为

$$\dot{\mathbf{X}}_i = f_i(\mathbf{X}_i, \mathbf{U}_i), i = 1, 2, \dots, n \quad (9)$$

式中, $f_i \in \mathbf{R}^3$ 为(5)式等号右边项。

2) 控制输入约束:(6)式中的控制输入约束可以表示为

$$|U_{i,j}(t)| \leq \bar{U}_{ij}, i = 1, 2, \dots, n, j = 1, 2, \forall t \in [t_0, t_f] \quad (10)$$

式中: $U_{i,j}$ 是个无人机 i 的控制向量的第 j 个分量; $\bar{U}_{ij}$ 为第 j 个分量的最大控制量; t_0 和 t_f 分别为初始时间和终止时间。

3) 初始约束

$$\mathbf{X}_i(t_0) = \mathbf{X}_{i0}, i = 1, 2, \dots, n, \quad (11)$$

式中, $\mathbf{X}_{i0}$ 是个无人机 i 的初始状态。

4) 状态约束

$$|\theta_i(t)| \leq \bar{\theta}_i, i = 1, 2, \dots, n, \forall t \in [t_0, t_f] \quad (12)$$

式中, $\bar{\theta}_i$ 为无人机 i 的航向角上界,该约束用于避免无人机因过度转向导致失稳。

5) 最终编队队形约束:各无人机的目标围捕点需满足圆形编队的几何要求,即无人机最终状态 $\mathbf{X}_i(t_f)$ 由参考参数 $\mathbf{X}_{\text{ref}}$ 决定。

$$\mathbf{X}_i(t_f) = f(\mathbf{X}_{\text{ref}}), i = 1, 2, \dots, n \quad (13)$$

在圆形编队中,将偏转角 δ 作为 $\mathbf{X}_{\text{ref}}$,与最终状态 $\mathbf{X}_i(t_f)$ 的关系如(14)式所示。

$$\begin{cases} \mathbf{X}_i(t_f) = f(\mathbf{X}_{\text{ref}}) = \begin{bmatrix} x_c \\ y_c \end{bmatrix} + R \begin{bmatrix} \cos((i-1)\phi + \delta) \\ \sin((i-1)\phi + \delta) \end{bmatrix} \\ \phi = 2\pi/n \end{cases} \quad (14)$$

圆形编队的圆心表示为 $[x_c, y_c]$ 、半径为 R , ϕ 为 n 个无人机之间的角度间隔。

对于该最优控制问题,首先将其转换为一个非线性规划问题,采用控制向量参数化方法实现;其次,确定状态约束和最终编队约束相对于所有优化

变量的梯度;最后将能量最优控制问题近似为迭代凸规划(SCP)问题进行求解。

1) 控制向量参数化

根据标准的控制向量参数化方法将控制向量近似成一个分段常值函数。首先,将时间区间 $[t_0, t_f]$ 分成 N 个子区间,即

$$0 = \tau_0 < \tau_1 < \dots < \tau_N = t_f \quad (15)$$

对于无人机 i ,在每个区间内,控制向量 $\mathbf{U}_i$ 近似为常值 σ_i^k ,如(16)式所示。

$$\mathbf{U}_i(\tau) \approx \sigma_i^k, t \in [\tau_{k-1}, \tau_k), k = 1, 2, \dots, N \quad (16)$$

控制参数化后,能量最优控制问题可以被重新建立,将(16)式代入(9)式,动力学方程可以表示为

$$\dot{\mathbf{X}}_i(\tau) = f_i(\tau, \mathbf{X}_i, \sigma_i), \tau \in [0, t_f], i = 1, 2, \dots, n \quad (17)$$

对于无人机 i ,将所有满足控制约束(10)式的常值控制参数 σ_i 的集合记为 Ξ_i ,若给定 $\sigma_i \in \Xi_i$,可以利用(11)式中的初始条件求解(17)式,得到 $\mathbf{X}_i(\tau)$ 。该问题可归为初始值问题(IVP),求解方法主要分为单步法、多步法和配置法。常见的单步法有龙格-库塔(RK)方法,直接通过当前步状态计算下一步状态;多步法根据问题特性选择适配方法,如用于非刚性问题的 Adams-Bashforth 法、用于刚性问题的向后差分法;配置法通过分段多项式近似各子区间内的状态变量以实现求解。求解得到无人机飞行轨迹后,即可重新构建如(18)式所示的编队队形约束。

$$\mathbf{X}_i(t_f | \sigma_i) = \mathbf{X}_{\text{ref}}, i = 1, 2, \dots, n \quad (18)$$

式中, $\mathbf{X}_i(t_f | \sigma_i)$ 为由 N 个子区间对应的 σ_i 求得的无人机 i 的最终状态解。

基于上述分析,可以将围捕点的分配与编队建立问题表达为更简单的形式

$$\begin{aligned} \min J &\approx \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^2 \sum_{k=1}^N (\sigma_{ij}^k)^2 \\ \text{subject to } &|\sigma_{ij}^k| \leq \bar{U}_{ij}, \\ &i = 1, 2, \dots, n, j = 1, 2, k = 1, 2, \dots, N \\ &|\theta_i(t)| \leq \bar{\theta}_i, \end{aligned} \quad (19)$$

$$\mathbf{X}_i(t_f | \sigma_i) = \mathbf{X}_{\text{ref}}, i = 1, 2, \dots, n$$

该问题可以视为非线性优化问题,其优化变量为参考参数 $\mathbf{X}_{\text{ref}}$ 及常值控制参数 σ_i ,为便于简化后续推导,将优化变量记为 $\xi = [\sigma_1^T, \sigma_2^T, \dots, \sigma_n^T, \mathbf{X}_{\text{ref}}]^T$ 。对于非线性规划问题,可以使用迭代二次规

划方法(SQP)、内点法或者迭代凸规划(SCP)方法求解,因SCP方法简单高效且开源,本文采用SCP方法。

2) 迭代凸规划方法

采用灵敏度法计算状态约束(12)式与最终编队队形约束(13)式相对于优化变量 ξ 的梯度,通过联立求解动力学方程与灵敏度辅助方程,得到状态变量对 σ_i 的偏导数,为凸规划提供精确的线性化基础。

利用凸规划(SCP)将非线性规划问题近似为一系列凸优化问题,通过迭代逐步逼近最优解。在问题(19)中,仅需对状态约束与最终队形约束进行近似。这2类约束相对于优化变量的梯度可通过积分求得,因此采用一阶泰勒展开完成近似。对于第 k 次迭代,通过引入优化变量 ξ ,利用问题(20)对问题(19)进行凸逼近,(20)式为典型的凸优化问题。为确保初始猜测值任意选择时的全局收敛性,需在 ξ 上施加信任域,但目标函数中仅含优化变量中的控制参数,因此信任域仅需针对该部分进行参数设定,如(21)式所示。

$$\min J = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^2 \sum_{k=1}^N (\sigma_{ij}^k)^2 \quad (20)$$

$$\text{subject to } \mathbf{X}_{i,k} + \frac{\partial \mathbf{X}_i(t_f | \sigma_{i,k})}{\partial \sigma_i} (\sigma_i - \sigma_{i,k}) = \mathbf{X}_{\text{ref}}$$

$$|\sigma_{ij}^p| \leq \bar{U}_{ij}, i = 1, 2, \dots, n, j = 1, 2, p = 1, 2, \dots, N$$

$$\left| \frac{\partial \theta_i^p}{\partial \sigma_i} (\sigma_i - \sigma_{i,k}) \right| \leq \bar{\theta}_i,$$

$$i = 1, 2, \dots, n, p = 1, 2, \dots, N$$

$$|\sigma_{ij}^p - \sigma_{ij,k}^p| < \rho_k, i = 1, 2, \dots, n,$$

$$j = 1, 2, p = 1, 2, \dots, N \quad (21)$$

迭代过程为:

1) 初始化控制参数 σ_i^0 ;

2) 在第 k 次迭代中,基于当前迭代点 ξ_k ,对状态约束与最终编队约束进行一阶泰勒展开,将目标函数(8)式近似为二次型函数,形成凸优化子问题,通过开源CVX求解;

3) 若满足收敛条件 $\|\xi_{k+1} - \xi_k\| < \varepsilon$,迭代终止,得到最优偏转角,代入(14)式可得最优围捕点;若未收敛,则进入下一次迭代。

通过上述流程,可在满足所有约束的前提下,为各无人机分配能耗最优的围捕点,同时确保多机可快速构建预设的圆形围捕编队,为后续围捕任务的

执行奠定基础。

2.2 引导阶段

当围捕无人机成功抵达指定的围捕点并建立围捕编队后,进入引导阶段。这一阶段在无障碍凸区域内进行非线性全局路径规划,并使用深度强化学习方法实现局部避障,为了追踪轨迹并重构编队,采用了基于应力矩阵的仿射编队跟踪控制。

2.2.1 全局路径规划

为确保围捕编队在引导入侵者至安全区域的过程中,既满足无碰撞约束又能实现队形构型优化,本文对基于非线性优化的全局路径规划方法进行改进。首先在任务区域自由空间内生成离散采样点,并基于该采样点集确定对应的无障碍凸区域;再引入约束优化,既确保编队在每个采样点处的构型均在无障碍凸区域内,又避免编队内部发生碰撞;最后采用图搜索算法,对所有采样点进行路径搜索,最终确定安全且成本最低的全局最优路径。

1) 无障碍凸区域:无障碍凸区域是指在自由空间中不包含任何障碍物的凸多边形,如图5所示。无障碍凸区域可以通过一组线性约束方程描述,即

$$\mathbf{Q} = \{ \mathbf{x} \in \mathbf{R}^d | \mathbf{A}_q \mathbf{x} \leq \mathbf{b}_q, \mathbf{A}_q \in \mathbf{R}^{n \times d}, \mathbf{b}_q \in \mathbf{R}^n \} \quad (22)$$

式中: $\mathbf{x}$ 为采样点; $\mathbf{A}_q$ 表示超平面系数矩阵; $\mathbf{b}_q$ 为列向量。

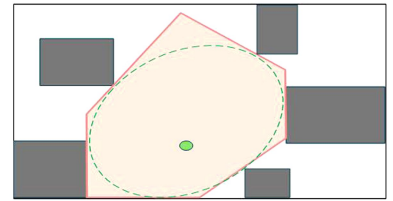


图5 无障碍凸区域

本文使用IRIS^[15]在任务环境中生成无障碍的凸区域。针对任意采样点,首先构建一组超平面将该采样点与周围障碍物分隔开,然后将该区域膨胀,生成该超平面界定的最大体积椭球体;重复该过程,寻找能将当前椭球体与障碍物进一步分隔的新超平面,最终形成最大的无障碍凸区域。

2) 编队构型的非线性优化:为了在无障碍区域内实现最优的编队构型,本文采用了非线性优化方法,目标为最小化构型参数的误差。

$$J(z) = \min(w_t \|\mathbf{t} - \mathbf{g}\|^2 + w_s \|\mathbf{s} - \bar{\mathbf{s}}\|^2 + c) \quad (23)$$

式中, w_t, w_s 表示设计权重; c 表示成本; z 表示编队

构型,它由大小 s 和平移 t 组成。编队中固定翼无人机的位置可以通过构型 z 获取。

$$v_j = t + sm_j, \forall j \in [1, n] \quad (24)$$

式中: m_j 为当前构型的编队顶点; v_j 是在编队重构之后的编队顶点,如图6所示。

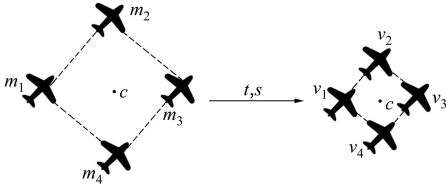


图6 编队重构

约束条件1 为了避免编队与障碍物发生碰撞,每个无人机的位置必须位于无障碍的凸区域 Q 内,即 $A_q[t + sm_j] \leq b_q$ 。

约束条件2 考虑到编队中无人机间的碰撞问题,所有无人机之间要保持一个安全相对距离 d ,本文设置 d 不小于无人机半径的2倍,此约束条件可表示为 $d \geq 2r_d$ 。

上述优化通过非线性求解器 MOSEK 求解,生成能够避免碰撞并实现编队重构的编队构型。

3) 全局路径规划: 全局路径规划生成编队的可行路径及中间构型。本文将无障碍凸区域与非线性优化方法相结合,该方法可实现从起始点 s 到目标点 g 的全局路径求解。

在全局路径规划中,图 $G = (V, E)$ 表示编队路径与构型之间的关系,每个节点 V 对应一个编队构型和位置,如果构型 z 完全位于一个凸区域 Q 内,则构型 z 被视为有效。因此,图中的每个节点都对应着一个可行的构型。边 E 将不同的节点连接起来,表示构型之间的可能转化,对于每一个凸区域 $Q \in \mathcal{Q}$,其中 $\mathcal{Q}$ 是现有凸区域列表,不断更新构型列表,包含所有编队构型完全位于该凸区域内的情况。节点列表 V 由初始构型 z_s 和最终构型 z_g 初始化,其质心分别为 s 和 g 。凸区域列表分别初始化为 Q_s 和 Q_g 。

基于采样方法,在工作区域内随机采样,每个采样点对应一个节点,然后在节点之间建立边,表示从起始点到目标点的可能路径。如果采样点位于障碍物或列表 $\mathcal{Q}$ 中的任何多边形内,则剔除,否则,执行以下步骤:

1) 随机生成采样点 p ,生成凸区域 Q_p ;

2) 对于与 Q_p 相交的每一个凸区域 $Q \in \mathcal{Q}$,计算使得编队完全位于这2个凸区域交集内的构型 z ,并最小化其质心到目标点 g 的距离。该构型通过上述非线性优化方法计算得出;

3) 当有效构型 z 被添加到节点列表中时,会为所有构型添加一条边 $\{z, z_i, Q_p\}$,并会为所有构型 $z_i \in \mathcal{L}_{Q_p}$ 添加另一条边。这样就能确保该结构能够通过相关的凸区域在2个构型之间进行移动。

随着采样的进行,动态更新图的节点和边,实现图结构的逐步扩展。当图构建完成后,采用 A^* 算法对全局路径进行搜索。对于连接2个构型 z_1 和 z_2 的每条边,其代价定义为 z_1 和 z_2 的质心之间的欧氏距离 $d(z_1, z_2)$ 。

2.2.2 局部动态避障

为应对全局路径跟踪过程中的突发动态障碍物,本文采用基于强化学习(DRL)的局部避障策略。该算法不需要精确的地图信息,能将传感器数据直接映射为速度指令,实时响应突发障碍物。本文采用领导者-跟随者结构,上述生成的速度指令将作为领导者的控制输入,以引导整个编队完成避障动作。

1) 观测状态与动作空间: 观测状态由向量 σ_t 表示,包含激光雷达数据 L_t 、速度信息 v_t 及目标相对位置 g_t 。 L_t 是最大距离归一化的数组,速度 v_t 用极坐标表示,动作空间 v_t 包含线速度 v_{lt} 与角速度 $v_{\omega t}$ 。

2) 奖励空间: 当无人机沿规划的全局路径行驶时,需应对环境中未知的动态障碍物。奖励函数用于驱动无人机飞向目标航点并规避未知障碍物,可以被设计为

$$\gamma_{\text{sum}} = \gamma_g + \gamma_c + \gamma_v \quad (25)$$

式中: γ_{sum} 表示总奖励, γ_g 、 γ_c 和 γ_v 分别表示目标距离奖励、避障奖励和速度奖励。目标距离奖励 γ_g 可由(26)式计算。

$$\gamma_c = \begin{cases} \gamma_a, & d_g < d_{\text{gmin}} \\ k_g \Delta d_g, & \text{其他} \end{cases} \quad (26)$$

当无人机与目标的距离 d_g 低于阈值 d_{gmin} 时,无人机获得到达奖励 γ_a ,否则,获得的奖励等于上一个时间步长内移动的距离。 k_g 为可调参数。避障奖励 γ_c 可定义为

$$\gamma_c = \begin{cases} \gamma_{co}, & d_o < d_{\text{omin}} \\ 0, & \text{其他} \end{cases} \quad (27)$$

式中, d_o 表示无人机与障碍物的距离,若 d_o 小于阈

值 $d_{\min}$, 则无人机避障获得负奖励。速度奖励 γ_v 定义为

$$\gamma_v = \gamma_{vl} + \gamma_{vw} \quad (28)$$

$$\gamma_{vw} = \begin{cases} -\gamma_w, & |v_w| > v_{w\max} \\ 0, & \text{其他} \end{cases} \quad (29)$$

$$\gamma_{vl} = \begin{cases} -\gamma_l, & v_l < v_{l\min} \\ 0, & \text{其他} \end{cases} \quad (30)$$

式中: $v_{w\max}$ 和 $v_{l\min}$ 分别定义为最大角速度和最小线速度。鼓励无人机以更快的线速度和更慢的角速度抵达目标位置, 因此, 若角速度 v_w 超过 $v_{w\max}$, 或线速度 v_l 低于 $v_{l\min}$, 无人机将分别获得惩罚奖励 $-\gamma_w$ 和 $-\gamma_l$ 。

3) 网络结构: 图7~8为动作网络与评价网络的结构。动作网络的输入为24维的激光雷达数据、2维的相对位置及2维的速度信息, 由3层全连接层处理后输出线速度与角速度。评价网络融合动作网络的输入与输出, 应用3层带ReLU激活函数的全连接层, 最终通过线性激活函数输出 Q 值。

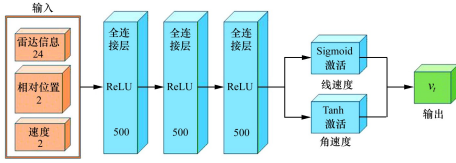


图7 动作网络的结构

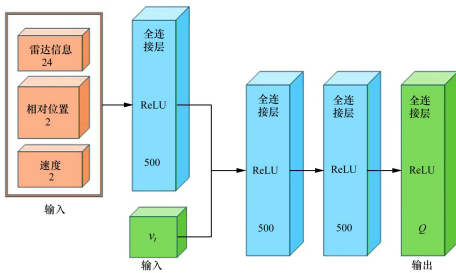


图8 评价网络的结构

4) 人类演示数据采集: 该方法使无人机能够直接从人类操作中学习。状态数据 σ_t 包含激光扫描信息 L_t 、相对位置信息 g_t 和速度信息 v_t 。速度指令 v_t 通过键盘输入, 奖励值由(25)式计算得出。经验数据集 $(\sigma_t, v_t, \sigma_{t+1}, \gamma_t)$ 存储于数据缓冲区中。

5) 优先经验回放: 为提升训练效率, 引入优先经验回放 (prioritized experience replay, PER) 算法, 并将其与 DDPG 算法融合。与传统均匀采样策略不同, PER 通过对经验池数据按重要性赋予优先

级, 优先采样对模型更新贡献更大的关键经验。优先经验回放算法的设计原理为

$$p_i = \rho_i^a / \sum_k \rho_k^a \quad (31)$$

式中: 采样概率 p_i 对应第 i 次转换; a 为分布系数; ρ_i 表示转换数据集优先级, 即

$$\rho_i = \delta_i^2 + \mu \|\nabla_a Q(s_i, \alpha_i | \theta^Q)\|^2 + \varphi + \varphi_D \quad (32)$$

式中: δ 表示时间误差; $\mu \|\nabla_a Q(s_i, \alpha_i | \theta^Q)\|^2$ 代表损失; μ 为贡献系数; φ 指每次转换的次要正采样概率; φ_D 表示额外采样概率, 用于提高运行数据的选取率。每次转换均被评估为

$$\lambda_i = (1/(N \cdot p_i))^b, \quad (33)$$

式中: N 为批次大小; b 是调整采样权重的常数; λ_i 为动作网络中的采样权重, 反映每个状态转移数据的重要性。

2.2.3 编队跟踪控制

为实现编队对全局路径的稳定跟踪与动态重构, 本文采用基于应力矩阵的仿射编队控制策略。在领导者-跟随者控制框架下, 该控制律能够使无人机追踪时变目标构型, 确保编队重构与维持。在该控制律的作用下, 无人机当前位置 $p_f(t)$ 可收敛至期望位置 $p_f^*(t)$, 确保(34)式的位置误差趋近于零。

$$\delta_{p_f}(t) = p_f(t) - p_f^*(t) = p_f(t) + \bar{\Omega}_{ff}^{-1} \bar{\Omega}_{ff} \tilde{p}_1(t) \quad (34)$$

1) 无饱和和约束: 对于无速度饱和约束的跟随无人机, 其控制律可设计为

$$\begin{aligned} v_i &= h_i^T \left[-\sum_{j \in N_i} \tilde{\omega}_{ij} (p_i - p_j) + k_i \bar{v}_1 + \omega_c h_i^\perp \right] \\ \omega_i &= (h_i^\perp)^T \left[-\sum_{j \in N_i} \tilde{\omega}_{ij} (p_i - p_j) + k_i \bar{v}_1 + \omega_c h_i^\perp \right] \end{aligned} \quad (35)$$

式中, $k_i \in \mathbf{R}^{2 \times 2n_1}$ 由矩阵 $K = -\bar{\Omega}_{ff}^{-1} \bar{\Omega}_{ff}$ 的第 $2i-1$ 行和第 $2i$ 行组成。 $\bar{v}_1$ 被定义为 $\bar{v}_1 = \bar{v}_1 h_1 \in \mathbf{R}^{2n}$, 其中 $\bar{v}_1 = \text{diag}(v_1) \otimes I_2 \in \mathbf{R}^{2n_1 \times 2n_1}$ 。设 $h_i = [\cos \theta_i, \sin \theta_i]^T$ 及 $h_i^\perp = [-\sin \theta_i, \cos \theta_i]^T$, 其中 h_i 表示无人机的方向向量, $h_i^\perp$ 与 h_i 正交。

$$H = \begin{bmatrix} h_1 & & & & \\ & h_2 & & & \\ & & \ddots & & \\ & & & h_{n_1} & \\ & & & & h_{n_1+1} \\ & & & & & h_{n_1+2} \\ & & & & & & \ddots \\ & & & & & & & h_n \end{bmatrix} =$$

$$\begin{bmatrix} H_1 \\ H_f \end{bmatrix} \in \mathbf{R}^{2n \times n} \quad (36)$$

因此,运动学模型(5)式可以被简化为

$$\begin{cases} \dot{p} = H v \\ \dot{h} = H^\perp \omega \end{cases} \quad (37)$$

因此,对于无人机 $i \in v_f$,当 $v_i > 0$ 时,跟随者的跟踪误差 δ_{p_i} 可收敛至零。

控制律(35)式的证明过程详见文献[29],该控制律可分为2个主要部分:第一部分为固定翼无人机构建位置共识协议,用于实现多机位置的协同收敛;第二部分则聚焦于协调领航机与跟飞无人机间的速度匹配,以保障编队运动的一致性。基于上述控制律,所提多固定翼无人机系统可实现对多种类型参考轨迹的稳定追踪,及任意时变仿射变换下的编队构型调整。

2) 饱和约束:实际应用中,固定翼无人机通常受速度等约束限制。因此,考虑(6)式的速度饱和约束,控制律可设计为

$$\begin{cases} v_i = h_i^T \left(- \sum_{j \in N_i} \tilde{\omega}_{ij} (p_i - p_j) + k_i \bar{v}_1 + \omega_c h_i^\perp \right) \\ \omega_i = h_i^{\perp T} \left(- \sum_{j \in N_i} \tilde{\omega}_{ij} (p_i - p_j) + k_i \bar{v}_1 + \omega_c h_i^\perp \right) \\ v_i \in [v_{\min}, v_{\max}] \\ \omega_i \in [-\omega_{\max}, \omega_{\max}] \end{cases} \quad (38)$$

3 仿真结果

为了验证多无人机协同围捕分层规划与控制框架的有效性,从目标轨迹预测、围捕点分配、路径规划及避障、编队跟踪控制等多方面进行了 MATLAB 仿真。任务区域设置为 $15 \text{ m} \times 15 \text{ m}$,其中随机分布蓝色静态障碍物和黄色动态障碍物(见图9a))。5个围捕无人机被随机分布在环境内,一旦检测到入侵者,则触发协同围捕分层规划与控制框架。

图9a)~9b)展示了多无人机协同围捕分层规划与控制框架的聚集阶段,该阶段包含入侵者的轨迹预测及围捕点分配,形成围捕编队。对于入侵者轨迹预测,采用插值与误差校正多项式拟合(IEC-PF)方法,设置采样窗口大小 $W=3$ 、预测次数 $K=5$ 、多项式阶数 $m=3$ 。基于迭代凸优化的围捕点分配中,对于状态约束,无人机航向角上界都设置为 90° ,围捕无人机的初始状态 X_{i0} 为随机位置,所有控制参数的信任域半径设为 0.02 ,限制迭代次数为 50 ,该最优控制问题使用 CVX 进行求解。

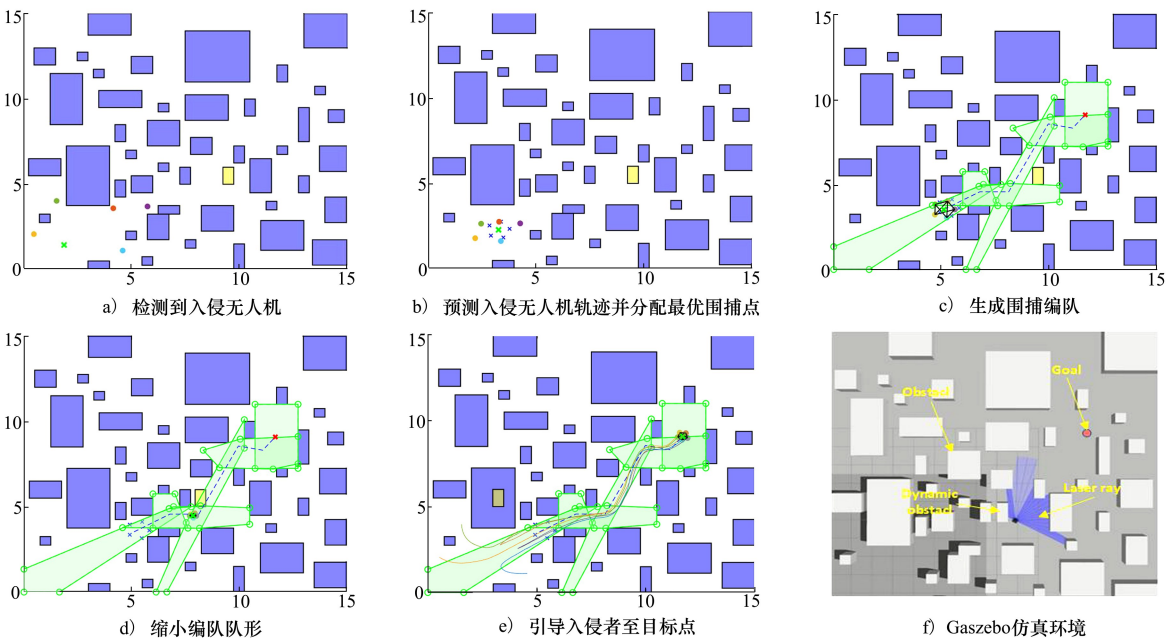


图9 无人机在动态障碍环境中的围捕仿真结果

如图 9c) 所示,当围捕者抵达分配的围捕点时,会自动形成围捕编队,系统进入引导阶段。标称编队 (G,r) 如图 10 所示,前 3 个无人机被指定为领导者,应力矩阵的权重分布于每条边上, $P(r)$ 为无人机在标称编队中的位置。以满足仿射可定位性条件,并计算标称编队的应力矩阵 $\Omega^{[27]}$ 。随后进入引导阶段,该阶段包含全局路径规划、编队跟踪控制及动态障碍物躲避。

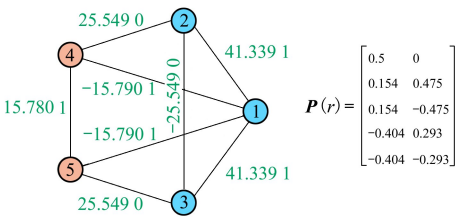


图 10 标称编队及应力值

在有密集障碍物的静态环境中,将基于非线性优化的全局路径规划方法与基于 Voronoi 的方法进行对比。非线性方法可以显著提升路径质量,2 种场景中路径长度分别缩短了 21.9% 和 11.9%,曲率降低了 76.9% 和 68.4%,提升了多无人机路径规划效率、路径平滑度,进而减少能量消耗。

使用具有速度约束的仿射编队控制跟踪全局无碰撞路径,结果表明,所有无人机均能在无障碍凸区域内实现跟踪路径(如图 9d)~9e)所示)。由非线性路径规划得到中间路径点的编队需缩放 0.3 倍,在图 12b)中用不同底色区分聚集和引导 2 个阶段,可以看出,各无人机的间距始终保持在安全阈值 $2r_d=0.18\text{ m}$ 以上,实现了无碰撞运行。当遇到未知动态障碍物时,采用基于深度强化学习(DRL)的局部避障方法,在 Gazebo 环境中完成训练,并在 TurtleBot3 机器人上进行测试(如图 9f)所示),使机器人在导航至红色目标的同时进行避障,训练参数详见表 1。遇到动态障碍物时,领导者通过 DRL 算法调整线速度与角速度,在跟踪全局路径的同时实现避障。图 9d)~9e)、图 11 也验证了系统的实时避障性能。

表 1 训练参数

参数	批次大小 N	贡献因子 μ	采样概率	网络速率	经验回放池大小
数值	512	0.2	0.2	0.000 1	1 000 000

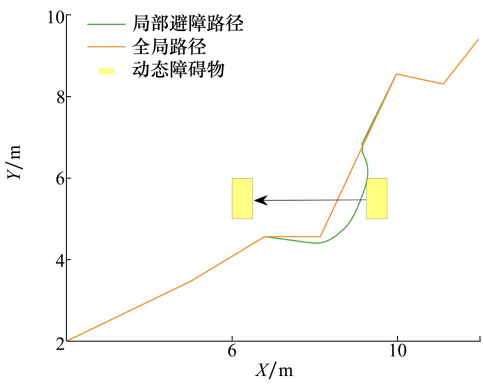


图 11 局部动态避障路径

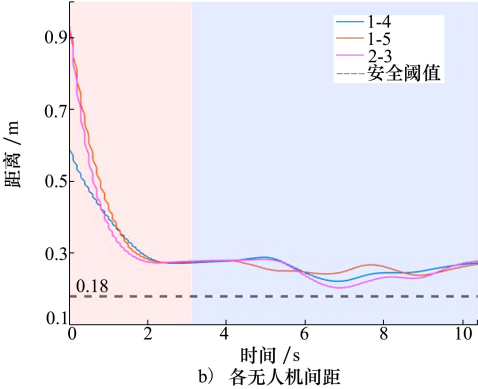
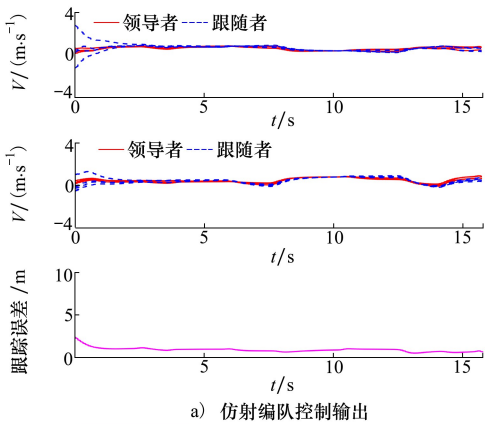


图 12 围捕过程相关输出量

总体而言,实验结果表明本文所提出的分层框架可有效、安全地完成围捕任务,在预测精度、任务分配、路径规划、编队跟踪及重构能力和避障性能方面均取得显著提升。

4 结 论

本文针对复杂环境下多无人机协同围捕面临的预测精度低、围捕任务分配能耗高、避撞能力不足等问题,提出包含聚集阶段与引导阶段的多无人机协

同围捕分层规划控制框架。在聚集阶段引入 IEC-PF 方法增强对入侵目标的短期轨迹预测,相较传统多项式拟合显著提升了预测精度,针对围捕点的分配提出基于迭代凸优化的分配方法,最小化系统能耗并构建满足约束的围捕编队。引导阶段采用基于非线性优化的全局路径规划,构建无障碍凸区域,生成的编队路径较 Voronoi 方法更平滑、更短,路径长

度与曲率分别降低 21.9% 和 76.9%;基于深度强化学习的局部避障算法显著提升动态障碍物实时响应能力;在固定翼无人机运动学约束下,仿射编队控制可确保路径追踪与编队重构。本文所提框架可显著提升系统在复杂环境中的避障能力,且能有效保障多无人机协同完成对入侵目标的围捕任务,并将其引导至指定目标区域。

参考文献:

- [1] 周萌, 李建宇, 王昶, 等. 多机器人协同围捕方法综述[J]. 自动化学报, 2024(12): 2325-2358.
ZHOU Meng, LI Jianyu, WANG Chang, et al. Multi-robot cooperative hunting: a survey[J]. Acta Automatica Sinica, 2024(12): 2325-2358. (in Chinese)
- [2] CAO Z, GU N, TAN M, et al. Multi-robot hunting in dynamic environments[J]. Intelligent Automation & Soft Computing, 2008, 14(1): 61-72.
- [3] 吴子沉, 胡斌. 基于态势认知的无人机集群围捕方法[J]. 北京航空航天大学学报, 2021, 47(2): 424-430.
WU Zichen, HU Bin. Swarm rounding up method of UAV based on situation cognition[J]. Journal of Beijing University of Aeronautics and Astronautics, 2021, 47(2): 424-430. (in Chinese)
- [4] 杨梓豪, 王庆领. 基于分层强化学习的多无人机协同围捕方法[J]. 海军航空大学学报, 2025, 40(4): 567-575.
YANG Zihao, WANG Qingling. Multi-UAV cooperative encirclement method based on hierarchical reinforcement learning[J]. Journal of Naval Aviation University, 2025, 40(4): 567-575. (in Chinese)
- [5] ZHENG Y, FAN W, HAN M. Research on multi-agent collaborative hunting algorithm based on game theory and Q-learning for a single escaper[J]. Journal of Intelligent & Fuzzy Systems, 2021, 40(1): 205-219.
- [6] 孙懿豪, 闫超, 相晓嘉, 等. 基于分层强化学习的多无人机协同围捕方法[J]. 控制理论与应用, 2025(1): 96-108.
SUN Yihao, YAN Chao, XIANG Xiaojia, et al. Multi-UAV collaborative pursuit method via hierarchical reinforcement learning[J]. Control Theory & Applications, 2025(1): 96-108. (in Chinese)
- [7] XIA J, LUO Y, LIU Z, et al. Cooperative multi-target hunting by unmanned surface vehicles based on multi-agent reinforcement learning[J]. Defence Technology, 2023, 29(11): 80-94.
- [8] XIA J W, LUO Y S, LIU Z K, et al. Cooperative multi-target hunting by unmanned surface vehicles based on multi-agent reinforcement learning[J]. Defence Technology, 2023, 29: 80-94.
- [9] BOROUJENI Z, GOEHRING D, ULBRICH F, et al. Flexible unit A-star trajectory planning for autonomous vehicles on structured road maps[C]//2017 IEEE International Conference on Vehicular Electronics and Safety, Vienna, 2017: 7-12.
- [10] 章翔峰, 田家全, 姜宏, 等. 改进人工势场法在机器人路径规划的应用研究[J]. 机械设计与制造, 2026(3): 332-336.
ZHANG Xiangfeng, TIAN Jiaquan, JIANG Hong, et al. Application research of improved artificial potential field method in robot path planning[J]. Machinery Design & Manufacture, 2026(3): 332-336. (in Chinese)
- [11] SEIDEL R. A simple and fast incremental randomized algorithm for computing trapezoidal decompositions and for triangulating polygons[J]. Computational Geometry, 1999, 1(1): 556-564.
- [12] AYANIAN N, KALLEM V, KUMAR V. Synthesis of feedback controllers for multiple aerial robots with geometric constraints [C]//2011 IEEE/RSJ International Conference on Intelligent Robots and Systems, 2011: 3126-3131.
- [13] 王钦禾, 尹永鑫, 戴丽, 等. 基于 Voronoi 图和蚁群算法的无人机航迹规划[J]. 导航定位与授时, 2021(2): 66-73.
WANG Qinhe, YIN Yongxin, DAI Li, et al. Trajectory planning of UAV based on Voronoi diagram and ant colony optimization algorithm[J]. Navigation Positioning & Timing, 2021(2): 66-73. (in Chinese)
- [14] LIU S, WATTERSON M, MOHTA K, et al. Planning dynamically feasible trajectories for quadrotors using safe flight corridors in 3-D complex environments[J]. IEEE Robotics & Automation Letters, 2017, 2(3): 1688-1695.
- [15] DAI H, AMICE A, WERNER P, et al. Certified polyhedral decompositions of collision-free configuration space[J]. International Journal of Robotics Research, 2024, 43(9): 1322-1341.

- [16] MARCUCCI T, NOBEL P, TEDRAKE R, et al. Fast path planning through large collections of safe boxes[J]. IEEE Trans on Robotics, 2024, 40: 3795-3811.
- [17] CHEN K, LIU H, LI Y, et al. Robot navigation in unknown and cluttered workspace with dynamical system modulation in star-shaped roadmap[C]//2025 IEEE International Conference on Robotics and Automation (ICRA 2025), Los Alamitos, 2025: 10140-10146.
- [18] 丁德锐, 张豪华, 孙然, 等. 基于双分支感知融合网络的 GRU-Attention-TD3 无地图自主导航避障方法[J/OL]. (2026-01-10) [2026-04-17]. <https://doi.org/10.14107/j.cnki.kzgc.20250423>.
DING Derui, ZHANG Haohua, SUN Ran, et al. Map-free autonomous navigation and obstacle avoidance method of GRU-Attention-TD3 based on dual-branch perception fusion network[J/OL]. (2026-01-10) [2026-04-17]. <https://doi.org/10.14107/j.cnki.kzgc.20250423>. (in Chinese)
- [19] BOUHAMED O, GHAZZAI H, BESBES H, et al. Autonomous UAV navigation: a DDPG-based deep reinforcement learning approach[C]//2020 IEEE International Symposium on Circuits and Systems, Seville, 2020: 1-5.
- [20] ZHAO W, QUERALTA J P, WESTERLUND T. Sim-to-real transfer in deep reinforcement learning for robotics: a survey[C]//2020 IEEE Symposium Series on Computational Intelligence, Canberra, 2020.
- [21] LI G, SHI L, CHEN Y, et al. Breaking the sample complexity barrier to regret-optimal model-free reinforcement learning[J]. Information and Inference: A Journal of the IMA, 2023, 12(2): 969-1043.
- [22] NIU H, JI Z, ARVIN F, et al. Accelerated sim-to-real deep reinforcement learning: learning collision avoidance from human player[C]//2021 IEEE/SICE International Symposium on System Integration, New York, 2021: 144-149.
- [23] SHAO J Y, XIE G M, YU J Z, et al. Leader-following formation control of multiple mobile robots[C]//2005 IEEE International Symposium on Mediterrean Conference on Control and Automation Intelligent Control, Limassol, 2005: 808-813.
- [24] BALCH T, ARKIN R C. Behavior-based formation control for multirobot teams[J]. IEEE Trans on Robotics & Automation, 1998, 14(6): 926-939.
- [25] 李正平, 鲜斌. 基于虚拟结构法的分布式多无人机鲁棒编队控制[J]. 控制理论与应用, 2020, 37(11): 2423-2431.
LI Zhengping, XIAN Bin. Robust distributed formation control of multiple unmanned aerial vehicles basedon virtual structure[J]. Control Theory & Applications, 2020, 37(11): 2423-2431. (in Chinese)
- [26] FANG X, XIE L. Distributed formation maneuver control using complex laplacian[J]. IEEE Trans on Automatic Control, 2024, 69(3): 1850-1857.
- [27] ZHAO S. Affine formation maneuver control of multiagent systems[J]. IEEE Trans on Automatic Control, 2018, 63(12): 4140-4155.
- [28] HAN Z, WANG L, LIN Z, et al. Formation control with size scaling via a complex laplacian-based approach[J]. IEEE Trans on Cybernetics, 2016, 46(10): 2348-2359.
- [29] LI H, CHEN H, WANG X. Affine formation tracking control of unmanned aerial vehicles[J]. Frontiers of Information Technology & Electronic Engineering, 2022, 23(6): 909-919.

Hierarchical planning and control for multi-UAV cooperative encirclement in complex environments

CHU Jing¹, ZHANG Ni¹, HUANG Yong², LIU Yumeng¹

(1.School of Automation, Xi'an University of Posts and Telecommunications, Xi'an 710061, China;
2.School of Astronautics, Northwestern Polytechnical University, Xi'an 710072, China)

Abstract: In complex environments, multi-UAV cooperative encirclement tasks often face critical challenges such as low intruder trajectory prediction accuracy, excessive system energy consumption during task allocation, and insufficient obstacle avoidance capability under dynamic conditions. These issues severely limit the efficiency and reliability of encirclement operations. To address these problems, a two-stage cooperative encirclement framework consisting of a gathering phase and a guidance phase is proposed in this paper. In the gathering phase, an interpolation and error-correction polynomial fitting (IEC-PF) method is designed to achieve early interception and significantly improve the accuracy of intruder trajectory prediction. For the task allocation of capture points, an iterative convex optimization-based allocation strategy is developed to minimize total system energy consumption while strictly satisfying the final capture formation constraints. In the guidance phase, UAVs construct an optimal encircling formation based on the assigned capture points. Within the generated obstacle-free convex regions, a nonlinear optimization method is used to compute optimal global paths and formation configuration parameters. An affine formation control strategy based on stress matrices is further introduced to ensure formation stability and reconfiguration during path tracking. Additionally, a deep reinforcement learning (DRL)-based local obstacle avoidance mechanism is integrated to enhance the system's real-time responsiveness to dynamic obstacles. Simulation results demonstrate that the proposed framework substantially improves obstacle avoidance performance in complex environments and enables reliable multi-UAV cooperative capture of intruding targets, as well as guiding them to designated areas.

Keywords: multi-agent cooperative hunting; trajectory prediction; task allocation; nonlinear optimization; affine formation tracking control; dynamic obstacle avoidance

引用格式: 褚晶, 张妮, 黄勇, 等. 复杂环境下多无人机协同围捕的分层规划与控制[J]. 西北工业大学学报, 2026, 44(2): 380-392.

CHU Jing, ZHANG Ni, HUANG Yong, et al. Hierarchical planning and control for multi-UAV cooperative encirclement in complex environments[J]. *Journal of Northwestern Polytechnical University*, 2026, 44(2): 380-392. (in Chinese)