

# 基于 MATD3 的无人机智能协同路径规划方法

杨闰麟<sup>1</sup>, 陈才轶<sup>1</sup>, 郭正玉<sup>2</sup>, 罗德林<sup>1,3\*</sup>

(1. 厦门大学航空航天学院, 福建 厦门 361102; 2. 中国空空导弹研究院, 河南 洛阳 471000;

3. 空基信息感知与融合全国重点实验室, 河南 洛阳 471000)

**摘要:** [目的] 多无人机协同路径规划是多无人机应用的关键技术之一, 本文旨在研究三维动态环境下的多无人机智能协同路径规划和优化方法。[方法] 本文提出了一种基于搜索的扰动流体动态系统(SIFDS)与速度控制策略结合的路径规划算法。该方法通过 SIFDS 方法将多障碍物环境简化为单障碍物环境, 并生成由反应系数构成的无人机动作空间; 再结合速度控制策略实现无人机间的时间协同; 最后利用多智能体双延迟深度确定性策略梯度(multi-agent twin delayed deep deterministic policy gradient, MATD3)算法, 训练得到的动作网络在线生成最优的反应系数, 以此提高多无人机路径规划的质量与效率。[结果] 本文所提出的方法能够在三维动态环境下使多无人机规划出安全且互不干扰的路径; 引入 MATD3 后, 无人机路径更短且平滑; 速度控制策略能够控制各无人机同时到达的时间误差在测量精度 0.5 s 以下。[结论] 本文所提出的方法, 能够使多无人机在规划路径的同时满足时间协同、空间协同的约束, 实现了多无人机协同路径规划, 且具有较强的稳定性, 为多无人机系统在复杂任务中的实际应用提供了一定的技术支持和理论参考。

**关键词:** 无人机; 协同路径规划; 扰动流体动态系统; 强化学习

中图分类号: TP 3

文献标志码: A

文章编号: 0438-0479(2025)02-0241-12

## Intelligent collaborative path planning method for unmanned aerial vehicles based on MATD3

YANG Runlin<sup>1</sup>, CHEN Caiyi<sup>1</sup>, GUO Zhengyu<sup>2</sup>, LUO Delin<sup>1,3\*</sup>

(1. School of Aerospace Engineering, Xiamen University, Xiamen 361102, China; 2. China Airborne Missile Academy, Luoyang

471000, China; 3. National Key Laboratory of Air-based Information Perception and Fusion, Luoyang 471000, China)

**Abstract:** [Objective] Coordinated path planning for multiple unmanned aerial vehicles (multi-UAVs) represents one of the pivotal technologies in multi-UAV applications. In this study, we aim to investigate intelligent coordinated path planning and optimization methods for multi-UAVs in three-dimensional dynamic environments. [Methods] Herein we propose a path planning algorithm that combines search-based interfered fluid dynamic system (SIFDS) with a velocity control strategy. This method utilizes SIFDS to simplify a multi-obstacle environment into a single obstacle environment and generates a UAV action space composed of reaction coefficients. Also, the velocity control strategy is employed to achieve temporal coordination among UAVs. Combined with the multi-agent twin delayed deep deterministic policy gradient (MATD3) algorithm, the trained action network generates the optimal reaction coefficient online, thus improving the quality and the efficiency of multi-UAV path planning. [Results] The proposed method enables multi-UAVs to plan safe and non-interfering paths in a three-dimensional dynamic environment. After introducing MATD3, the path of the drone is shorter and smoother. Furthermore, comparative experiments illustrate that the integrated velocity control strategy facilitates the simultaneous arrival of all UAVs at their destinations. The arrival time error is below 0.5 seconds of measurement

收稿日期: 2024-09-09 录用日期: 2025-01-26

基金项目: 空基信息感知与融合全国重点实验室与航空科学基金联合资助项目(20220001068001)

\* 通信作者: luodelin1204@xmu.edu.cn

引文格式: 杨闰麟, 陈才轶, 郭正玉, 等. 基于 MATD3 的无人机智能协同路径规划方法[J]. 厦门大学学报(自然科学版), 2025, 64(2): 241-252.

Citation: YANG R L, CHEN C Y, GUO Z Y, et al. Intelligent collaborative path planning method for unmanned aerial vehicles based on MATD3[J]. J Xiamen Univ Nat Sci, 2025, 64(2): 241-252. (in Chinese)



**accuracy.** [Conclusions] The proposed method allows multi-UAVs to meet both temporal and spatial collaboration constraints while planning their paths, so that effective collaborative path planning is achieved. Hopefully, our study can provide technical support and theoretical guidance for the practical deployment of multi-UAV systems in complex tasks.

**Keywords:** UAV; collaborative path planning; interfered fluid dynamic system; reinforcement learning

随着无人机技术日新月异的发展,无人机无论是在军事领域<sup>[1-5]</sup>还是在民用领域<sup>[6-8]</sup>的应用都日益广泛.然而,随着任务复杂度的不断提升,单一无人机的作业能力显得愈发局限.多无人机协同作业不仅能够提高任务执行的效率,还能增强系统的鲁棒性和容错能力.因此,多无人机协同作业成为了一种重要的趋势.

在多无人机协同作业中,协同路径规划是关键环节之一.协同路径规划不仅需要解决无人机自身避障的问题,还要解决多个无人机之间的路径冲突问题,以确保整个无人机集群能够协同一致、高效安全地完成既定任务.传统的路径规划方法通常基于启发式搜索算法(如 A\* 算法<sup>[9]</sup>、D\* 算法<sup>[10]</sup>)或群智能优化算法(如遗传算法<sup>[11]</sup>、粒子群算法<sup>[12]</sup>、灰狼算法<sup>[13]</sup>),在应对多无人机协同规划问题时,存在计算复杂度高、实时响应能力及动态环境适应性较差等方面的不足.

随着深度强化学习(DRL)技术的不断发展,该领域的研究者纷纷探索其在无人机路径规划方面的应用潜力.DRL通过模拟智能体与环境之间频繁且深入的交互学习,逐步优化并习得一种高效的路径规划策略;这使得无人机在应对复杂且动态变化的环境条件时,能够更加灵活且精准地执行路径规划任务,显著提升其适应性和效率.Zhang等<sup>[14]</sup>将强化学习与扰动流体动态系统(interfered fluid dynamic system, IFDS)<sup>[15]</sup>相结合,提出自适应 IFDS 算法,验证其在路径安全性上具有优越性.吴健发等<sup>[16]</sup>考虑无人机运动学模型,提出受约束的 IFDS(C-IFDS)算法,其与深度确定性策略梯度(deep deterministic policy gradient, DDPG)融合的算法有实时性好、路径质量高等优点.Zhang等<sup>[17]</sup>引入双倍体验重放缓冲区,对 DDPG 算法进行改进,极大地提高了训练效率.牟文心等<sup>[18]</sup>提出了基于孪生 DDPG(twin delayed DDPG, TD3)改进的 PP-CMNTD3 算法.该方法能够提高 DRL 在无人机航迹规划中的训练效率,有效应对复杂场景,并实现能耗与快速抵达目的地之间的平衡.然而,上述文献仅处理单无人机路径规划问题,并未考虑多无人机协同问题.王伟伦等<sup>[19]</sup>提出了一种基于多智能体软动作评价(multi-agent soft actor critic, MASAC)网络的无人机集群协同搜索路径规划方法,能够更高效地完成无人机集群探索任务.聂铭涛等<sup>[20]</sup>提出了一种基于多

智能体 DDPG(multi-agent DDPG, MADDPG)<sup>[21]</sup>的无人机协同路径规划算法,能够较好地完成多无人机协同突防以及汇合任务.杜江涛等<sup>[22]</sup>提出了基于 MADDPG 的路径规划策略.该方法能够解决不完全信息下无人机集群路径规划问题,实现快速有效的路径规划.当前,针对三维动态环境下的多无人机智能路径规划问题,相关的探索与研究尚显不足.

针对多无人机协同路径规划,本文将基于搜索的 C-IFDS 算法用于无人机路径规划,再利用多智能体 TD3(multi-agent TD3, MATD3)<sup>[23]</sup>算法将其与速度控制策略结合,求解动作反应系数,进而解决无人机空间与时间协同问题.本文的主要贡献包括:1) 提出一种基于搜索的 IFDS(search-based IFDS, SIFDS)算法,减少了多障碍环境下无人机所需优化参数数量.2) 在 SIFDS 路径规划算法的基础上,引入 MATD3 多智能体深度强化学习算法.这一融合策略实现了对生成的路径参数的实时优化,并显著提升了无人机路径的质量与效率.3) 在 MATD3 算法中使用参数整流线性(PReLU)作为动价(actor)网络与评价(critic)网络的激活函数,解决了死亡神经元的问题,提高了模型的收敛速度与训练效率.

本文的章节组织安排如下:第一节描述协同路径规划问题以及环境建模,第二节介绍 SIFDS 算法和时间协同规划方法,第三节描述 SIFDS 算法如何与 DRL 相结合,第四节进行实验与分析;第五节给出结论.

## 1 问题描述与建模

### 1.1 多无人机协同路径规划问题描述

考虑到无人机在不同环境下执行不同任务存在较大差异,对本文研究的任务场景做出以下说明:

1) 在任务区域中,存在若干静态障碍物和动态障碍物, $N$ 架己方无人机以及各无人机对应目标点,己方无人机以避开障碍物并同时到达各自目标点为任务目标.

2) 无人机在进行路径规划时,需要满足无人机自身的运动学约束,如最大最小速度、偏航角速率等;还需满足无人机机间最小距离和无人机与障碍物最小

距离的约束,保证安全性。

3) 每架无人机可以感知自身周围局部环境的障碍物信息。

4) 若最早和最晚到达终点区域的无人机时间差小于给定阈值,且飞行过程中各无人机未发生碰撞,则认为此次任务成功完成。

综上,无人机协同路径规划问题描述如下:在满足无人机飞行速度、偏航角速率、避障等约束条件的同时,利用各无人机的自身观测和共享信息,在相同时间或允许误差范围内达到各自目标点。

## 1.2 无人机运动学模型以及约束

无人机在三维坐标系下的运动学模型如下<sup>[16]</sup>:

$$\begin{cases} \dot{x} = V \cos \gamma \cos \chi, \\ \dot{y} = V \cos \gamma \sin \chi, \\ \dot{z} = V \sin \gamma, \\ \dot{V} = (n_x - \sin \gamma) g, \\ \dot{\gamma} = \frac{g(n_z - \cos \gamma)}{V}, \\ \dot{\chi} = \frac{gn_y}{V \cos \gamma}, \end{cases} \quad (1)$$

其中, $V$  为飞行速度, $\gamma$  和  $\chi$  分别为俯仰角和偏航角, $g$  为重力常数,作为控制输入的  $n_x, n_y, n_z$  表示沿航迹坐标系  $x, y, z$  轴的过载。此外,无人机还须满足如下约束条件:

$$\begin{cases} n_i \in [n_{i\min}, n_{i\max}], i = x, y, z, \\ \dot{\gamma} \in [-\dot{\gamma}_{\max}, \dot{\gamma}_{\max}], \dot{\chi} \in [-\dot{\chi}_{\max}, \dot{\chi}_{\max}], \\ \gamma \in [\gamma_{\min}, \gamma_{\max}], \\ V \in [V_{\min}, V_{\max}]. \end{cases} \quad (2)$$

## 1.3 障碍环境建模以及避碰约束

本文采用标准凸面体包络对地形或威胁进行等效。障碍物类型包括静态球体、动态球体与静态圆柱。建立如下障碍物等效标准凸面体包络方程:

$$\Gamma(\mathbf{P}) = \left( \frac{x - x_0}{a + R_A} \right)^{2p} + \left( \frac{y - y_0}{b + R_A} \right)^{2q} + \left( \frac{z - z_0}{c + R_A} \right)^{2r}, \quad (3)$$

式中: $\mathbf{P} = [x, y, z]^T$  表示无人机的三维位置; $a, b, c > 0$  和  $p, q, r > 0$  分别决定了障碍物的覆盖范围与形状,当  $p = q = r = 1$  且  $a = b = c$  时,障碍物为圆球;当  $p = q = 1, r > 1$  且  $a = b$  时,障碍物为圆柱; $\mathbf{P}_0 = [x_0, y_0, z_0]^T$  表示障碍物几何中心;无人机安全半径为  $R_A$ ; $\Gamma(\mathbf{P}) > 1, \Gamma(\mathbf{P}) = 1$  和  $\Gamma(\mathbf{P}) < 1$  分别表示无人机位于障碍物等效包络的外部、表面和内部。 $B$  为所有障碍物的内部及表面,定义如下:

$$B = \bigcup_{k=1}^K B_k, B_k = \{\mathbf{P} \mid \Gamma_k(\mathbf{P}) \leq 1\}, \quad (4)$$

式中, $K$  表示障碍物的个数,各无人机在飞行过程中不能与障碍物发生碰撞,需满足如下约束:

$$\mathbf{P}_i \notin B, i = 1, 2, \dots, N, \quad (5)$$

式中, $\mathbf{P}_i$  表示第  $i$  架无人机的三维位置, $N$  为无人机数量。

## 1.4 无人机协同约束

本文还考虑了无人机空间协同约束和时间协同约束。空间协同约束强调在无人机进行路径规划时,各无人机之间必须维持足够的安全间隔,以确保飞行过程中的互不干扰与安全性。需满足如下约束:

$$d_{ij} < d_{\text{safe}}, \quad (6)$$

式中, $d_{ij}$  表示无人机  $i$  与无人机  $j$  之间的距离, $d_{\text{safe}}$  表示无人机机间安全距离。

而时间协同约束则要求各无人机抵达它们各自的目标位置的时间差小于给定值。需满足如下约束:

$$t_{\text{last}} - t_{\text{first}} \leq C, \quad (7)$$

式中, $t_{\text{first}}$  为最早到达的无人机所花时间, $t_{\text{last}}$  为最晚到达的无人机所花时间, $C$  为允许最大差值。

# 2 多无人机协同规划

## 2.1 基于 SIFDS 的路径规划

Wang 等<sup>[15]</sup>根据自然界的流水避石原理提出了 IFDS 算法。IFDS 算法能够平滑地避开障碍物到达终点,但是会受到局部最优问题影响。因此,姚鹏等<sup>[24]</sup>提出了改进 IFDS(IIFDS)算法,在扰动矩阵中引入切向矩阵,解决了流线分布空间有限和局部最优的问题。吴健发等<sup>[16]</sup>在 IIFDS 算法的基础上引入无人机运动学模型以及运动约束,提出了 C-IFDS 算法,提高了规划路径的可跟踪性。为了减少路径规划时的计算量以及方便后续路径的优化,本文提出 SIFDS 路径规划方法,原理如下。

无障碍情况下,初始流速  $\mathbf{u}(\mathbf{P})$  定义如下:

$$\mathbf{u}(\mathbf{P}) = \left[ \frac{V(x_d - x)}{\|\mathbf{P} - \mathbf{P}_d\|} \quad \frac{V(y_d - y)}{\|\mathbf{P} - \mathbf{P}_d\|} \quad \frac{V(z_d - z)}{\|\mathbf{P} - \mathbf{P}_d\|} \right]^T, \quad (8)$$

式中, $\mathbf{P}$  和  $\mathbf{P}_d = [x_d, y_d, z_d]^T$  分别为无人机当前位置和目标点位置。

当环境中存在  $K$  个障碍物时,障碍物对  $\mathbf{u}(\mathbf{P})$  的干扰影响为总和扰动矩阵  $\overline{\mathbf{M}}(\mathbf{P})$ , 其计算量受到障碍物数量  $K$  的影响。在本文中通过实时搜索障碍物的方式,选取当前对无人机威胁最大的障碍物作为参考障

障碍物,选取方式为:无人机对搜索半径  $R$  内搜索到的障碍物进行筛选.具体地,先判断在无人机飞行方向上是否有障碍物阻挡:若有,则阻挡障碍物为参考障碍物;若无,则选择搜索半径内与无人机最近的障碍物作为参考障碍物.如图 1 所示.

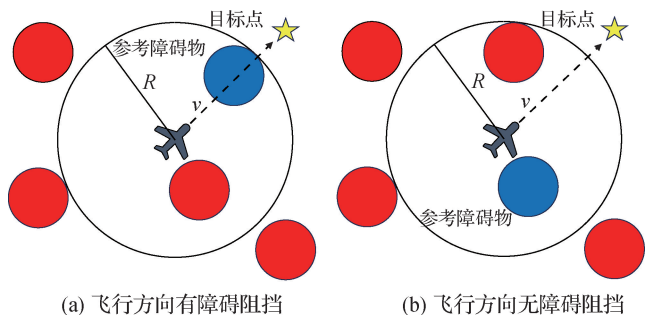


图 1 参考障碍物的选择

Fig. 1 Selection of reference obstacle

参考障碍物会根据无人机的当前位置进行更新,通过这种方式可以筛除对无人机飞行威胁较小的障碍物,专注于对无人机威胁较大的障碍物进行避障,将多障碍环境简化为单障碍环境.

考虑参考障碍物对  $u(\mathbf{P})$  的影响,总扰动矩阵  $\bar{\mathbf{M}}(\mathbf{P})$  可表示为:

$$\bar{\mathbf{M}}(\mathbf{P}) = \mathbf{M}(\mathbf{P}) = \mathbf{I} - \frac{\mathbf{n}(\mathbf{P})(\mathbf{n}(\mathbf{P}))^T}{|\Gamma(\mathbf{P})| \frac{1}{\rho} (\mathbf{n}(\mathbf{P}))^T \mathbf{n}(\mathbf{P})} + \frac{\mathbf{t}(\mathbf{P})(\mathbf{n}(\mathbf{P}))^T}{|\Gamma(\mathbf{P})| \frac{1}{\sigma} \|\mathbf{t}(\mathbf{P})\| \|\mathbf{n}(\mathbf{P})\|}, \quad (9)$$

式中:  $\mathbf{M}(\mathbf{P})$  为参考障碍物的扰动矩阵;  $\Gamma(\mathbf{P})$  表示由式(3)定义的障碍物包络方程;  $\rho$  为排斥反应系数;  $\sigma$  为切向反应系数;  $\mathbf{I}$  为三阶的单位吸引矩阵;  $\mathbf{n}(\mathbf{P})$  为径向向量<sup>[24]</sup>;  $\mathbf{t}(\mathbf{P})$  为惯性系  $Oxyz$  下的切向矩阵,推导过程如下.

在与  $\mathbf{n}(\mathbf{P})$  垂直的切平面  $S$  上定义两个相互垂直的切向量  $\mathbf{t}_1(\mathbf{P})$  和  $\mathbf{t}_2(\mathbf{P})$ :

$$\mathbf{t}_1(\mathbf{P}) = \begin{bmatrix} \frac{\partial \Gamma(\mathbf{P})}{\partial y} & -\frac{\partial \Gamma(\mathbf{P})}{\partial x} & 0 \end{bmatrix}, \quad (10)$$

$$\mathbf{t}_2(\mathbf{P}) = \begin{bmatrix} \frac{\partial \Gamma(\mathbf{P})}{\partial x} & \frac{\partial \Gamma(\mathbf{P})}{\partial z} \\ \frac{\partial \Gamma(\mathbf{P})}{\partial y} & \frac{\partial \Gamma(\mathbf{P})}{\partial z} \\ -\left(\frac{\partial \Gamma(\mathbf{P})}{\partial x}\right)^2 & -\left(\frac{\partial \Gamma(\mathbf{P})}{\partial y}\right)^2 \end{bmatrix}. \quad (11)$$

以  $\mathbf{t}_1(\mathbf{P})$ 、 $\mathbf{t}_2(\mathbf{P})$ 、 $\mathbf{n}(\mathbf{P})$  为  $x'$ 、 $y'$ 、 $z'$  三轴建立坐标系  $O'-x'y'z'$ , 则切平面  $S$  内任意单位切向量在  $O'-x'y'z'$  表示为:

$$\mathbf{t}'(\mathbf{P}) = [\cos \theta \quad \sin \theta \quad 0]^T, \quad (12)$$

式中,  $\theta \in [-\pi, \pi]$  为任意切向量与  $x'$  轴的夹角.

通过坐标旋转矩阵  $\mathbf{R}$  可将  $O'-x'y'z'$  下的  $\mathbf{t}'(\mathbf{P})$  转换为  $Oxyz$  下的  $\mathbf{t}(\mathbf{P})$ :

$$\mathbf{t}(\mathbf{P}) = \mathbf{R}(\mathbf{P})\mathbf{t}'(\mathbf{P}). \quad (13)$$

考虑动态威胁的影响,构建相对初始流场,先通过  $\bar{\mathbf{M}}(\mathbf{P})$  修正相对初始流场,求得相对流场扰动流速,最后得到实际扰动流速  $\bar{\mathbf{u}}(\mathbf{P})$ :

$$\bar{\mathbf{u}}(\mathbf{P}) = \bar{\mathbf{M}}(\mathbf{P})(\mathbf{u}(\mathbf{P}) - \mathbf{v}(\mathbf{P})) + \mathbf{v}(\mathbf{P}), \quad (14)$$

式中,  $\mathbf{v}(\mathbf{P})$  为参考障碍物的速度矢量.

最后,对  $\bar{\mathbf{u}}(\mathbf{P})$  积分得到下一个规划航点.

$$\mathbf{P}(t+1) = \mathbf{P}(t) + \bar{\mathbf{u}}(\mathbf{P})\Delta T. \quad (15)$$

考虑到无人机的运动学模型及其约束,实时规划时需要确保无人机规划路径满足式(2)的约束.如果下一个规划航点满足约束,可以直接使用;若不满足约束,则应当对该航路点进行修正,使其满足无人机运动学模型约束,进一步提高规划路径的平滑度与可跟踪性.

在计算扰动矩阵时,参考障碍物会对应一组反应系数,包括  $\rho$ 、 $\sigma$ 、切向方向系数  $\theta$ ;  $\rho$  和  $\sigma$  决定了规划路径的形状,值越大,规避障碍的反应越快;  $\theta$  决定了飞行方向,不同的系数组合对规划路径的影响如图 2 所示.

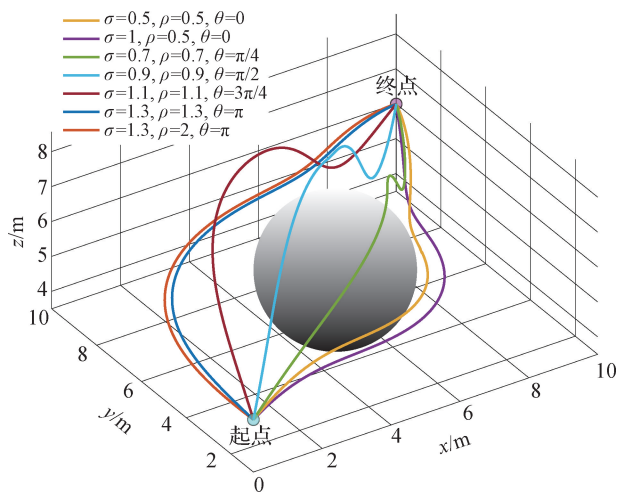


图 2 不同系数组合对路径的影响

Fig. 2 Effects of different coefficient combinations on the path

通过调整组合上述 3 个系数可得到一系列不同形状且分布于整个规划空间的流线.后续可对各个时刻的反应系数进行调整进而优化对应无人机的路径.

## 2.2 基于速度控制策略的时间协同

假设有  $N$  架无人机,对于无人机的速度控制,有:

$$\dot{v}_i = \alpha_{v,i}(v_i^c - v_i), \quad (16)$$

式中,  $v_i^c$  为第  $i$  架无人机速度指令,  $v_i$  为第  $i$  架无人机

当前飞行速度,  $\alpha_{v,i}$  为控制第  $i$  架无人机飞行速度指令调整快慢正常数。

假设所有无人机在  $t=0$  时刻出发. 此后, 各无人机开始沿规划路径飞行. 令  $L_i$  为  $t$  时刻第  $i$  架无人机剩余路径长度, 其值可根据无人机当前位置实时估计得到.  $v_i$  为  $t$  时刻第  $i$  架无人机的飞行速度, 可由无人机模块实时给出. 得到  $L_i$  和  $v_i$  以后, 第  $i$  架无人机能估算出  $t$  时刻的期望到达时刻:

$$\zeta_i = t + \tau_i = t + \frac{L_i}{v_i}, \quad (17)$$

式中,  $\tau_i$  为第  $i$  架无人机在  $t$  时刻剩余路径的期望到达时间. 基于一致性算法, 多机同时到达的控制目标是使所有无人机在  $t$  时刻的期望到达时刻  $\zeta$  趋于一致, 即:

$$\zeta_i \rightarrow \zeta_j, \forall i, j \in \{1, 2, \dots, N\}. \quad (18)$$

经过推导后, 得到如下分散化控制策略<sup>[25]</sup>:

$$\begin{cases} v_i^c = v_i - \frac{v_i u_i}{(\alpha_{v,i} \tau_i)}, \\ u_i = - \sum_{j=1}^n a_{ij} (\tau_i - \tau_j). \end{cases} \quad (19)$$

其中  $a_{ij}$  表示无人机  $i$  与无人机  $j$  之间的通信强度。

将上述速度控制策略与 SIFDS 方法相结合, 提出一种 SIFDS 协同路径规划方法, 令各无人机既能够规划出安全的路径, 又能减小各无人机到达各自的目标点的时间差, 使得多无人机能在一定时间差内到达各自目标点, 实现多无人机时间协同规划。

### 3 基于 MATD3 的动作空间参数优化

#### 3.1 马尔科夫决策链模型

本文无人机关注于完全合作的多智能体策略学习, 每架无人机在进行路径规划的同时, 需要控制与障碍物和其他无人机的距离, 从而避免碰撞, 并最终同时到达目标区域. 整个任务过程可以使用部分可观测的马尔科夫决策链 (Dec-POMDP) 进行建模<sup>[26]</sup>. 该马尔科夫决策模型可以用一个元组  $G$  来表示,  $G = (\mathbf{S}, \mathbf{O}, \mathbf{A}, P, R, N, \gamma)$ , 具体介绍如下:

- $\mathbf{S}$ : 无人机的状态空间,  $\mathbf{S} = [s_1, s_2, \dots, s_N]$ , 代表无人机当前观测和动作空间的合集;

- $\mathbf{O}$ : 无人机的观测空间,  $\mathbf{O} = [o_1, o_2, \dots, o_N]$ , 每架无人机的观测空间为  $o_i$ ,  $o_i = [x_i, y_i, v_i, \alpha_i]$ , 其中  $x_i$  为无人机当前位置到终点的向量,  $y_i$  为无人机当前位置到障碍物表面最短距离的向量,  $v_i$  为第  $i$  架无人机观测到的参考障碍物的速度,  $\alpha_i$  为无人机到终点和到障碍物中心的夹角;

- $\mathbf{A}$ : 联合动作空间,  $\mathbf{A} = a_1 \times a_2 \times \dots \times a_N$ , 其中每架无人机的动作空间为  $a_i$ ,  $a_i = [\rho_i, \sigma_i, \theta_i]$ ; 其中,  $\rho_i \in [0.1, 3], \sigma_i \in [0.1, 3], \theta_i \in [0, \pi]$ ;

- $P$ : 状态转移函数  $P_i(s_{t+1} | s_t) \in [0, 1]$ , 表示在当前状态  $s_t$  下, 选择动作  $a_t$ , 得到下一个状态  $s_{t+1}$  的概率.

- $r$ : 无人机获得的奖励, 表示无人机  $i$  在状态  $s_t$  下根据策略  $\pi_i$  执行动作  $a_i$  后获得的奖励, 第  $i$  架无人机的奖励函数为  $r_i$ ,  $\gamma_i$  为关于  $\mathbf{S}_t, a_{1t}, a_{2t}, \dots, a_{Nt}$  的函数, 具体参考 3.4 节;

- $\gamma$ : 折扣因子,  $\gamma \in [0, 1]$ , 用于权衡长期奖励和当前奖励的相对重要程度.

每架无人机  $i$  的训练目标是学习一个从  $o_i$  映射到动作空间分布的策略  $\pi_i: o_i \rightarrow P(a_i)$ , 该策略能够最大化无人机  $i$  的期望奖励:

$$J_i(\pi_i) = \mathbb{E}_{a_1 \sim \pi_1, \dots, a_N \sim \pi_N, s \sim P_i} \left[ \sum_t \gamma^t r_{it} \right]. \quad (20)$$

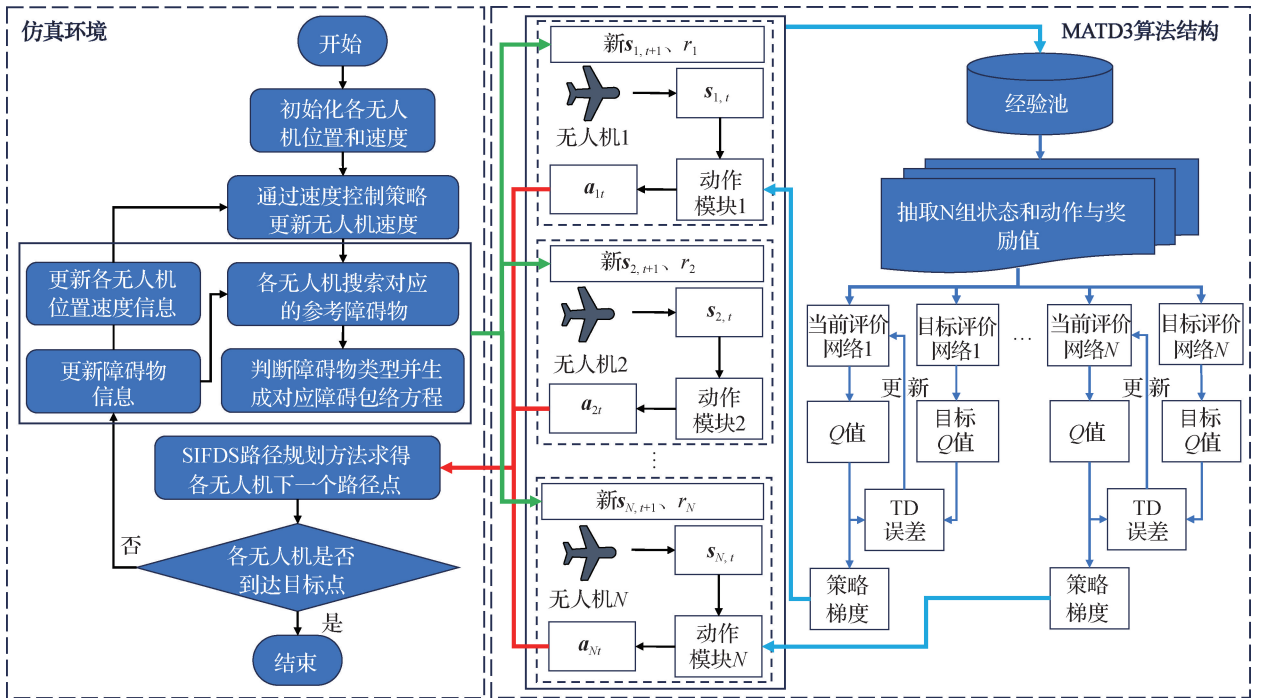
#### 3.2 基于 MATD3 的参数优化设计

DRL 拥有处理高维连续状态和动作空间的能力, 将 DRL 引入到本文提出的 SIFDS 路径规划算法中, 旨在通过在线优化 SIFDS 中的相应参数, 进一步提升实时规划航迹的质量. 具体流程如下: 将 SIFDS 算法作为路径规划的基础方法, 再经由 MATD3<sup>[23]</sup> 算法训练得到的神经网络, 在线优化各无人机对应的  $\rho_i, \sigma_i, \theta_i$ . 本文将 MATD3 算法与 SIFDS 路径规划算法相结合, 提出了 MATD3-SIFDS 算法, 其结构图如图 3 所示.

MATD3 算法是 MADDPG 算法的改进版本, 其算法结构与 MADDPG<sup>[21]</sup> 相似, 并同样可运用在连续状态/动作空间环境下. 在 MATD3 算法中, 每个智能体包括动作模块和评价模块, 每个模块又由当前网络和目标网络两个神经网络组成. 其中两个评价网络的参数随机给定, 两个动作网络的参数则通过深度学习对传统方法进行输入输出拟合而给定, 在这种方式下, 动作网络可以获得一个初始可行解, 随后通过梯度下降对网络进行训练, 从而提高算法的训练效率和收敛性.

具体在本文中, 首先让每架无人机维护独立的动作网络和双评价网络, 同时初始化对应的目标网络. 接着, 所有智能体与环境交互, 将收集的  $(s_t, a_t, r_t, s_{t+1})$  数据存入共享的经验回放池. 再采用双重 Q 学习机制, 通过两个评价网络分别计算最小值. 最后, 交替更新动作和评价网络, 最大化预期 Q 值, 输出当前环境下的最优动作. 下面给出 MATD3 的伪代码.

初始化每架无人机的动作在线网络  $\pi_\theta$  和对应的目标网络  $\pi_{\theta^-}$ , 两个评价在线网络  $Q_{\phi_1}, Q_{\phi_2}$  和对应的目标网络  $Q_{\phi_1^-}, Q_{\phi_2^-}$ ;



TD 误差为基于时序差分学习(temporal difference learning)的误差项,用于衡量当前 Q 值函数的估计值与目标值之间的差距。

图 3 MATD3-SIFDS 算法结构

Fig. 3 Algorithm structure of MATD3-SIFDS

for Episode  $i=1$  to  $M$  do

for step  $t=0$  to  $T-1$  do

根据当前策略  $\pi_\theta$  在状态  $s_t$ , 由公式  $a_t = \pi_\theta(s_t) + \text{clip}(\epsilon, -c, c)$  选择确定的动作  $a_t$  并执行, 得到  $r_t, s_{t+1}$ , 其中, clip 为裁剪函数,  $\epsilon$  为噪声,  $-c$  与  $c$  为引入噪声的上下限;

执行动作  $a_t$ , 得到  $r_{t+1}, s_{t+1}$ , 将采样的经验  $(s_t, a_t, r_t, s_{t+1})$  存储在经验回放池  $D$ ;

if 达到足够的训练样本 then

从经验池中回放 mini-batch;

根据采样的经验、当前策略  $\pi_\theta$  和噪声, 根据下式计算平滑的目标值;

$$a_{\theta-} = \pi_{\theta-}(s') + \text{clip}(\epsilon, -c, c)$$

$$y_t = r + \gamma \min_{i=1,2} Q_{\theta_i}(s', \text{clip}(a_{\theta-}, a_{\text{low}}, a_{\text{high}}))$$

根据下式计算两个 Q 网络的损失, 并进行后向梯度更新;

$$\theta = \text{argmax} \mathbb{E}_{s \sim D} [Q_{\theta_1}(s, \pi_\theta(s))]$$

if 达到动作的更新频率 then

根据  $\mathbb{E}_{s \sim D} [Q_{\theta_1}(s, \pi_\theta(s))]$  的确定性策略梯度更新在线策略网络的参数  $\theta$

更新目标网络参数  $\theta^-, \phi_1^-, \phi_2^-$

end if

end if

end for

end for

### 3.3 网络结构

上一节介绍了 MATD3<sup>[23]</sup>算法的基本流程, 其中的动作网络和评价网络结构设计如图 4 所示。

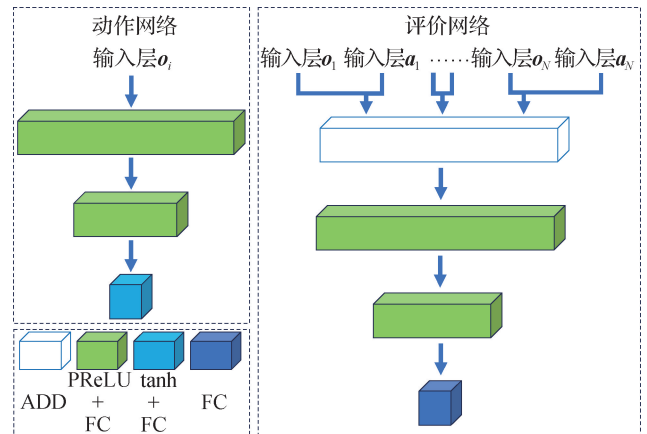


图 4 评价网络和动作网络的结构

Fig. 4 Structures of critic network and actor network

动作网络的输入为每个智能体的观测空间, 输出为该智能体选择的唯一动作, 由输入层、全连接层(FC)+参数整流线性(PReLU)层、FC+双曲正切(tanh)层构成; 评价网络则将所有智能体的观测-动作作为联合输入, 并输出状态动作值函数, 即 Q 值, 用

来指导动作网络进行更新,由输入层、添加层(ADD)、FC+PReLU 和 FC 层构成。

文献[23]中 MATD3 采用的激活函数为 ReLU,但采用 ReLU 作为激活函数时,若神经元的输入落入负值区间,其梯度会骤然变为 0[图 5(a)],它会导致该神经元在后续的训练迭代中无法再被有效激活或更新权重,这一现象称为“死亡神经元”。相比之下,PReLU 通过引入一个可自适应学习的参数  $\alpha$ ,当输入为负时,仍能保持输出非零的梯度[图 5(b)],有效解决了死亡神经元的问题。这一特性使得 PReLU 即使在处理负输入时也能持续学习,从而增强了其捕捉数据复杂特征的能力。因此,本文选用 PReLU 作为激活函数,以便更灵活地适应数据,减少波动和过拟合,使模型具有更强的鲁棒性与泛化能力。

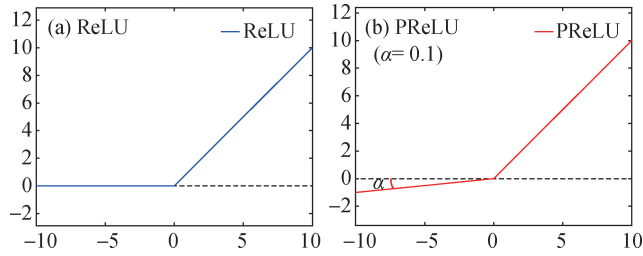


图 5 ReLU 与 PReLU

Fig. 5 ReLU and PReLU

### 3.4 奖励函数设计

在进行无人机协同路径规划时,需确保每架无人机都能有效规避环境中的各类潜在威胁,安全抵达预定目标点的同时,还需兼顾整个无人机群体到达目标点的时间相同,并严格实施无人机间的防撞策略。因此,既要考虑无人机自身的奖励,又要考虑各无人机的联合奖励。

自身奖励主要负责引导无人机安全平滑且高效地向目标点飞行,分为避障奖励、航迹角速率奖励、路径长度奖励、最佳飞行角奖励。联合奖励主要负责激励各无人机在保证各自安全的情况下协同工作,实现整体效能的最优化,分为机间避障奖励与到达时间奖励。

避障奖励  $r_{col}$  如下:

$$r_{col} = \begin{cases} -10, d \leq 0, \\ -3, 0 < d \leq 10, \\ -2, 10 < d \leq 20, \\ -1, 20 < d \leq 30, \\ 0, d > 30, \end{cases} \quad (21)$$

式中, $d$  表示无人机规划的下一个路径点到障碍物等效表面的最小距离, $d$  为负数代表规划的下一个路径点在障碍物内部。

航迹角速率奖励  $r_{ang}$  如下:

$$r_{ang} = -\frac{|\dot{\gamma}|}{\dot{\gamma}_{max}} - \frac{|\dot{\chi}|}{\dot{\chi}_{max}}, \quad (22)$$

式中, $\dot{\gamma}$  和  $\dot{\chi}$  表示路径规划得到下一个路径点而引起航迹角和爬升角变化时需要的角速率, $\dot{\gamma}_{max}$  和  $\dot{\chi}_{max}$  分别表示无人机偏航角速率与爬升角速率的最大值。

路径长度奖励  $r_{len}$  如下:

$$r_{len} = \begin{cases} (\| \mathbf{P}_d - \mathbf{P}(t) \| - \| \mathbf{P}_d - \mathbf{P}(t+1) \|) - \| \mathbf{P}(t) - \mathbf{P}(t+1) \|, \\ \| \mathbf{P}_d - \mathbf{P}(t+1) \| > R_{thr}, \\ (\| \mathbf{P}_d - \mathbf{P}(t) \| - \| \mathbf{P}_d - \mathbf{P}(t+1) \|) - \| \mathbf{P}(t) - \mathbf{P}(t+1) \| + r_a, \\ \| \mathbf{P}_d - \mathbf{P}(t+1) \| \leq R_{thr}, \end{cases} \quad (23)$$

式中: $\mathbf{P}_d$  表示目标点位置, $\mathbf{P}(t)$ 、 $\mathbf{P}(t+1)$  分别表示无人机当前点与下一个规划路径点的位置; $R_{thr}$  表示目标点误差范围半径,取 30; $r_a$  表示无人机到达目标点的奖励,取 10。

当无人机向目标点飞行存在障碍物阻挡时,引导无人机向正确的方向飞行可以减少最终规划的路径长度,在此引入飞行角度奖励<sup>[27]</sup>。

无人机观测到前方的障碍物,并且测得障碍物边缘切线  $L_1$  和  $L_2$  的方向角,已知目标点的位置,可以计算出  $L_1$  和  $L_2$  到飞机与目标点连线的夹角  $\theta_1$  和  $\theta_2$ ,以图 6 为例, $\theta_2$  大于  $\theta_1$ ,最佳的飞行方向为  $L_1$  切线上方,反之则为  $L_2$  切线下方,最佳飞行角  $\theta_{best}$  定义如下:

$$\theta_{best} = \begin{cases} \theta_{L_1} + \theta_e, \theta_2 > \theta_1, \\ \theta_{L_2} - \theta_e, \theta_2 \leq \theta_1, \end{cases} \quad (24)$$

式中: $\theta_{L_1}$  和  $\theta_{L_2}$  分别表示无人机飞行方向与切线  $L_1$  和  $L_2$  的夹角; $\theta_e$  为偏离切线方向的角度差,本文取  $2^\circ$ 。

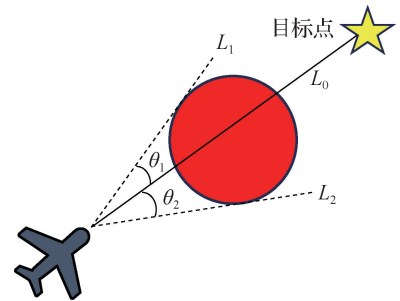


图 6 无人机、障碍物、目标点俯视示意图

Fig. 6 Top view diagram of UAV, obstacle, and destination

飞行角度奖励  $r_{best}$  如下:

$$r_{best} = \begin{cases} 1, |\theta_{real} - \theta_{best}| \leq 5, \\ 0, 5 < |\theta_{real} - \theta_{best}| \leq 15, \\ -1, |\theta_{real} - \theta_{best}| > 15. \end{cases} \quad (25)$$

机间避碰奖励  $r_{in}$  如下:

$$r_{in} = \sum_{j=1, j \neq i}^N p_{ij}, p_{ij} = \begin{cases} 0, & d_{ij} \geq d_{safe}, \\ -10, & d_{ij} < d_{safe}, \end{cases} \quad (26)$$

式中,  $d_{ij}$  表示无人机  $i$  与无人机  $j$  之间的距离,  $d_{safe}$  表示无人机机间安全距离, 本文取 30.

到达时间奖励  $r_{arr}$  如下:

$$r_{arr} = 5 - \sum_{j, j \neq i}^N |t_i - t_j|, \quad (27)$$

式中,  $t_i$  与  $t_j$  分别表示无人机  $i$  与无人机  $j$  到达目标点的时间.

第  $i$  架无人机的奖励函数  $r_i$  如下:

$$r_i = \omega_{col} r_{col} + \omega_{ang} r_{ang} + \omega_{len} r_{len} + \omega_{best} r_{best} + \omega_{in} r_{in} + \omega_{arr} r_{arr}, \quad (28)$$

式中,  $\omega_{col}, \omega_{ang}, \omega_{len}, \omega_{best}, \omega_{in}, \omega_{arr}$  为相应奖励的权重.

### 4 实验与分析

#### 4.1 实验环境设置与算法对比

本文所设计实验的硬件环境为: Intel CoreI i7-9750H CPU@2.60 GHz CPU, 16 GB RAM, NVIDIA Geforce RTX3090 graphics.

实验的软件环境为: Windows 11 64 位操作系统, 程序开发软件为 VS Code, Python3.8 以及 Pytorch 2.1.0 的深度学习框架.

实验对模型的超参数设置如表 1 所示.

表 1 超参数的设置值

Tab. 1 Hyperparameter setting values

超参数	数值
训练回合数	2 500
每回合训练最大步数	500
折扣率	0.95
批大小	128
隐藏层 1 神经元个数	256
隐藏层 2 神经元个数	128
软更新系数	0.01
梯度裁剪值	10
动作网络的学习率	0.000 1
评价网络的学习率	0.000 5
策略噪声	0.2
策略噪声裁剪值	0.5
策略更新频率	2

为了验证激活函数优化模型性能提升的效果, 将优化后的 MATD3 算法、原始的 MATD3 算法以及 MADDPG 算法分别与 SIFDS 算法相结合, 对多无人机进行训练, 并对训练结果进行对比分析. 在相同的测试条件下, 3 种算法的奖励值曲线如图 7 所示.

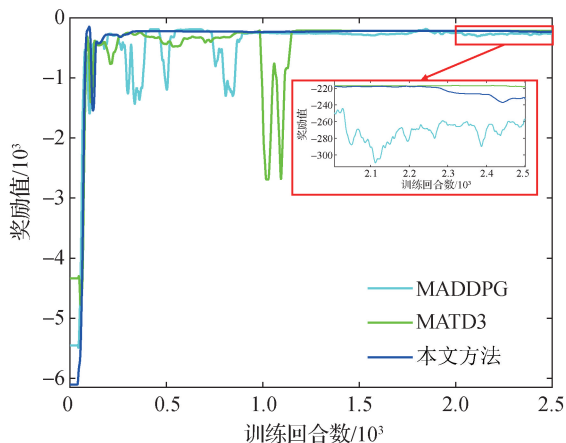


图 7 3 种算法奖励值曲线

Fig. 7 Reward value curves for 3 algorithms

由图 7 可知, 3 种算法最终都达到了收敛, MADDPG 算法最终收敛的奖励值略低于另外两种方法; MATD3 算法收敛速度在三者中较慢; 本文使用的方法收敛速度快且在收敛后稳定性较强(图 7 中插图), 使用该方法有利于无人机规划出高质量的路径.

为了验证本文引入 DRL 方法能够有效提高无人机规划的路径质量, 将 SIFDS 算法和 MATD3 + SIFDS 算法分别进行仿真实验, 并进行对比. 仅使用 SIFDS 算法时, 四架无人机的参数均设置如下: 排斥反应系数  $\rho=0.5$ , 切向反应系数  $\sigma=0.5$ , 切向方向系数  $\theta=0.1$ . 最终生成的三维路径如图 8 所示.

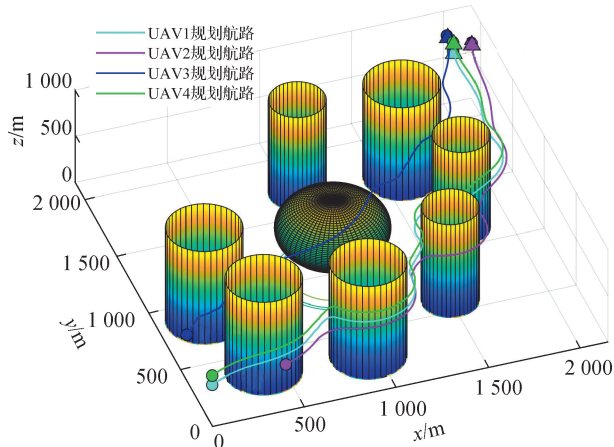


图 8 SIFDS 算法的仿真结果

Fig. 8 Simulation results of SIFDS algorithm

仅使用 SIFDS 算法生成的四条路径长度分别为 3 631.07, 3 223.07, 2 951.09 和 3 652.43 m. 四条路径对应俯仰角和偏航角变化的总和分别为 19.78, 14.41, 25.01 和 21.08 rad.

MATD3+SIFDS 算法生成的三维路径如图 9 所示.

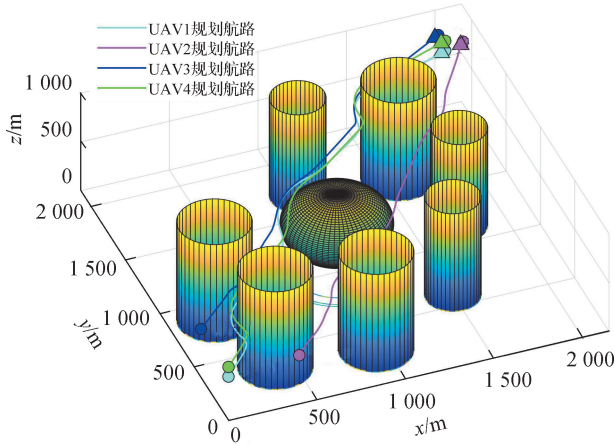


图 9 MATD3+SIFDS 算法的仿真结果

Fig. 9 Simulation results of MATD3+SIFDS algorithm

MATD3+SIFDS 算法生成的四条路径长度分别为 3 053.16, 2 654.44, 2 710.07 和 3 036.27 m. 四条路径对应俯仰角和偏航角变化的总和分别为 8.49, 6.11, 8.53 和 8.47 rad.

由仿真结果可知,引入 MATD3 后,无人机路径长度显著缩短且更平滑. 综上,引入 MATD3 能够有效提升无人机路径规划质量.

### 4.2 三维动态环境下路径规划仿真实验

本文设置了 4 种仿真环境,无人机数量  $N$  为 4. 在仿真环境 1 和 2 中,各无人机从大致相同的位置出发,初始飞行方向相同,最终到达各自目标点,具体设置如下:各无人机的初始位置分别为  $(0, 0, 450)$ 、 $(400, 0, 500)$ 、 $(0, 400, 500)$  和  $(0, 0, 550)$ , 目标点分别为  $(2 000, 2 000, 750)$ 、 $(2 100, 2 000, 800)$ 、 $(2 000, 2 100, 800)$  和  $(2 000, 2 000, 850)$ , 单位为 m; 各无人机的初始速度方向均为  $\gamma(0) = 0^\circ$  和  $\chi(0) = 45^\circ$ , 初始速度幅值均为 30 m/s, 步长为 0.5 s; 无人机运动约束为:  $\dot{\gamma}_{\max} = 40(^{\circ})/s$ 、 $\dot{\chi}_{\max} = 30(^{\circ})/s$ 、 $V \in [20, 60]$  m/s、 $\gamma \in [-75^\circ, 75^\circ]$ 、 $n_x \in [-0.5, 2]$ 、 $n_y \in [-2, 2]$ 、 $n_z \in [-1, 3]$ ; 在环境 1 中存在一个静态球体和多个静态圆柱体障碍(图 8). 在环境 2 中存在多个静态圆柱体障碍,还有两个等效安全半径为 100 m 的动态威胁(图 10),其运动模式如下.

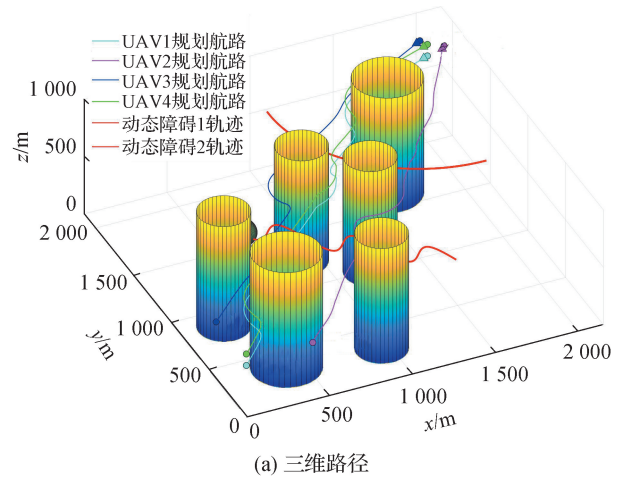
动态威胁 I 作蛇形运动:

$$\begin{cases} x(t) = 1\,400 - 10t, \\ y(t) = 300 + \frac{1\,000}{3\pi} \sin(0.06\pi t) + 8t, \\ z(t) = 600. \end{cases} \quad (29)$$

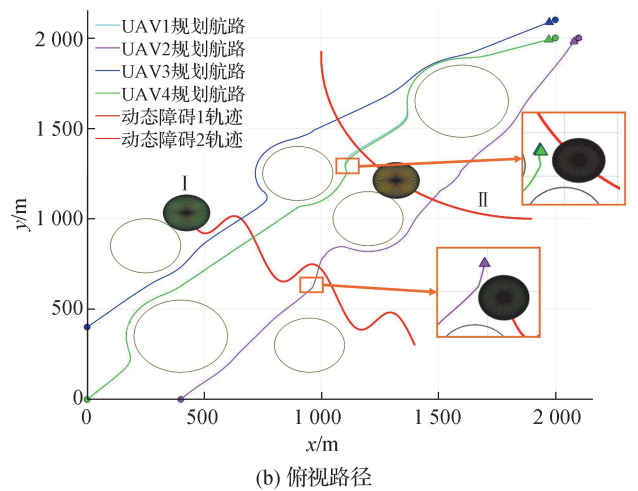
动态威胁 II 作圆弧往返运动:

$$\begin{cases} \Delta t = t \bmod 40, 40k \leq t < 40(k+1), \\ \quad k \text{ 为偶数,} \\ \Delta t = 40 - t \bmod 40, 40k \leq t < 40(k+1), \\ \quad k \text{ 为奇数,} \\ x(t) = 1\,900 - 900\sin(0.04\Delta t), \\ y(t) = 1\,900 - 900\cos(0.04\Delta t), \\ z(t) = 700. \end{cases} \quad (30)$$

环境 2 在线规划的三维航路如图 10(a)所示,其俯视图如图 10(b)所示.



(a) 三维路径



(b) 俯视路径

图 10 环境 2 下的路径

Fig. 10 The path in environment 2

飞行过程中,各无人机与参考障碍物表面最近距离的最小值分别为 29.5, 31.5, 25.9, 29.7 m.

仿真环境 3、4 下,各无人机从不同位置出发,初

始飞行方向不同,最终到达位于地图中心的各个目标点,具体设置如下:各无人机的初始位置分别为(0,0,450)、(2 000,0,500)、(0,2 000,500)和(2 000,2 000,550),目标点分别为(900,900,800)、(1 000,900,800)、(900,1 000,800)和(1 000,1 000,850),单位为m;各无人机的初始速度方向分别为 $\gamma_1(0) = 0^\circ$ 和 $\chi_1(0) = 45^\circ$ 、 $\gamma_2(0) = 0^\circ$ 和 $\chi_2(0) = 135^\circ$ 、 $\gamma_3(0) = 0^\circ$ 和 $\chi_3(0) = -135^\circ$ 、 $\gamma_4(0) = 0^\circ$ 和 $\chi_4(0) = -45^\circ$ 初始速度幅值均为30 m/s;无人机运动约束同仿真环境1和2,在环境3中存在多个圆柱体障碍.在环境4中存在多个静态圆柱体障碍,还有两个等效安全半径为150 m的动态威胁,其运动模式如下.

动态威胁 I 作直线运动:

$$\begin{cases} x(t) = 1\,000 - 10t, \\ y(t) = 300 + 10t, \\ z(t) = 600. \end{cases} \quad (31)$$

动态威胁 II 作圆周运动:

$$\begin{cases} x(t) = 1\,900 - 900\sin(0.04t), \\ y(t) = 1\,900 - 900\cos(0.04t), \\ z(t) = 700. \end{cases} \quad (32)$$

环境4在线规划的三维航路如图11(a)所示,其俯视图如图11(b)所示.

飞行过程中,各无人机与参考障碍物表面最近距离的最小值分别为31.8,26.1,31.6,44.0 m.

由上述仿真结果可知,在飞行过程中,各无人机对于静态/动态威胁能够及时做出避障反应,并规划出较为平滑的路径.综上所述,本文方法能够在复杂三维动态环境下完成多无人机路径规划.

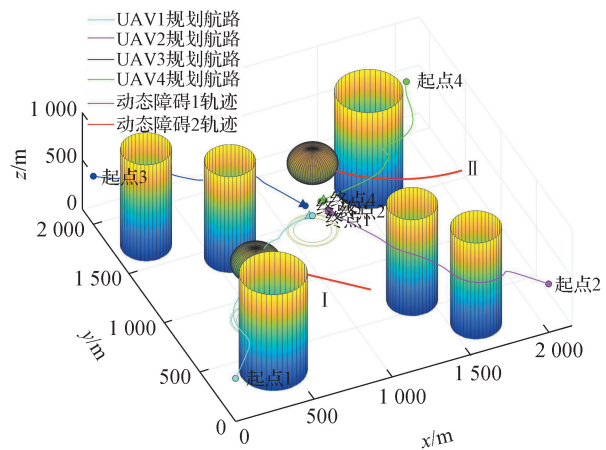
### 4.3 空间协同与时间协同仿真验证

由于本文是多无人机协同路径规划,在各无人机规划路径的同时,还应当满足前文所提到的空间协同与时间协同的约束.

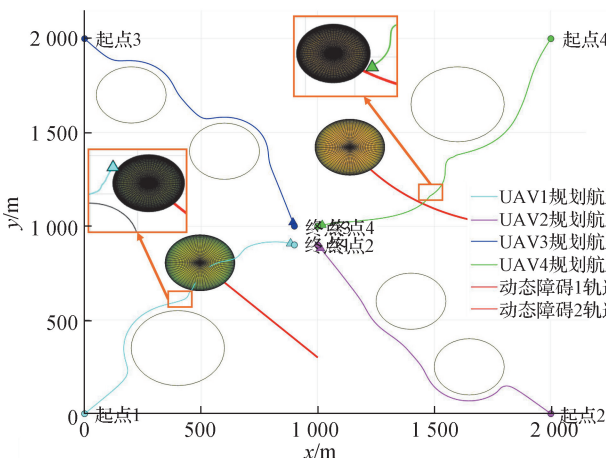
空间协同约束,即无人机在规划各自的航路时要保证各无人机的机间要保持一定的安全距离.四个环境下各无人机机间距离最小值如表2所示.

由表2可知,四个环境下各时刻无人机机间距离最小值为94.4 m,远大于无人机机间安全距离30 m.因此,在路径规划的同时,各无人机机间保持了一定的安全距离,不会发生机间碰撞,满足了无人机协同路径规划的空间协同约束.

时间协同约束,即无人机在规划各自的航路时要保证各无人机同时出发并在基本相同的时间到达各自的目标点.在所构建的4个环境下,分别测试使用2.3节提到的速度控制策略与不使用该策略在四个环



(a) 三维路径



(b) 俯视路径

图11 环境4下的路径

Fig. 11 The path in environment 4

境下的无人机到达情况,不使用该策略的无人机速度幅值恒定值30 m/s.

表2 各个环境下各无人机机间距离最小值

Tab. 2 Minimum distance between UAVs in each environments

无人机	最小距离/m			
	环境1	环境2	环境3	环境4
UAV1-UAV2	117.6	118.5	126.0	130.5
UAV1-UAV3	105.2	107.0	116.9	112.6
UAV1-UAV4	94.4	99.5	159.7	167.8
UAV2-UAV3	149.2	150.7	174.8	181.3
UAV2-UAV4	117.3	118.5	113.7	124.3
UAV3-UAV4	103.5	107.4	117.1	126.4

使用速度控制策略与速度恒定时的无人机到达时间分别如表3所示.不使用速度控制策略的情况

下,各无人机由于规划出的路径长度不同,到达目标点的时间不同,本文所使用的速度控制策略能够有效地减少各无人机到达目标点的时间差,满足了无人机协同路径规划的时间协同约束。

表 3 使用速度控制策略和恒速的各个环境下各无人机到达时间

Tab. 3 Arrival time of UAVs in each environments using speed control strategy and constant speed

速度	环境编号	到达时间/s			
		UAV1	UAV2	UAV3	UAV4
控制策略	1	96.0	96.0	96.0	96.0
	2	97.5	97.5	97.5	97.5
	3	54.5	54.5	54.5	54.5
	4	51.5	51.5	51.5	51.5
恒定	1	105.5	95.0	96.5	105.0
	2	110.5	88.5	90.0	106.5
	3	53.5	53.0	53.0	61.5
	4	51.5	53.0	52.0	58.0

综上所述,本文所提出的方法能够在三维动态环境下完成多无人机协同路径规划。

## 5 结 论

针对多无人机协同路径规划问题,本文提出了基于 SIFDS 与 MATD3 相结合的设计方案,该方案的特点是通过搜索障碍物的方式将多障碍环境简化为单障碍环境,减少了所需优化的反应系数,再结合一种速度控制策略,使得多无人机能满足时间协同约束,通过设计合理的观测与奖励函数,利用 MATD3 优化各无人机规划的路径,多无人机同时出发,各自规划出平滑、高效、安全的路径,最终同时到达各自的目标点.在所设置的 4 个三维复杂环境下进行仿真实验,验证了该方案的有效性。

本文的研究为多无人机系统在复杂任务中的实际应用提供了有力的技术支持和理论参考,展示了基于深度强化学习的多智能体协同决策在提升系统整体性能方面的巨大潜力.未来工作可进一步探索更复杂的场景和更多样的任务需求,使用更优的多智能体强化学习算法(如多智能体近端策略优化(MAPPO)<sup>[28]</sup>等),以应对更大规模无人机编队的协同规划问题,推动多无人机协同作业技术向更高水平发展。

## 参考文献:

- [1] 刘森,姜雪松,张宇晨,等. 林火搜救多无人机协同任务分配方法[J]. 中国新技术新产品,2024(6):133-136.
- [2] 柳汀,周国鑫,徐扬,等. 融合信息图的优化哈里斯鹰多无人机动态目标搜索[J]. 航空学报,2024,45(s1):1-13.
- [3] LEE K S, OVINIS M, NAGARAJAN T, et al. Autonomous patrol and surveillance system using unmanned aerial vehicles[C]//International Conference on Environment and Electrical Engineering. Piscataway: IEEE,2015:1291-1297.
- [4] XIE P Y, MA B, WANG B B, et al. 3D path planning of the solar powered UAV in the urban-mountainous environment with multi-objective and multi-constraint based on the enhanced sparrow search algorithm incorporating the Levy flight strategy [J]. Guidance, Navigation and Control, 2024,4(1):2450005.
- [5] ZHU K, HAN B, ZHANG T. Multi-UAV distributed collaborative coverage for target search using heuristic strategy [J]. Guidance, Navigation and Control, 2021, 1(1):2150002.
- [6] 陈博帆,潘可达,陈静川,等. 面向电力巡检场景的多无人机任务分配与路径规划方法[J/OL]. 南方电网技术,1-9. [2024-09-01]. <http://kns.cnki.net/kcms/detail/44.1643.TK.20240801.0934.002.html>.
- [7] 施晓航,徐勇勤. 旋翼无人机的多机协同自主探索决策技术综述[J]. 兵器装备工程学报,2023,44(10):182-190.
- [8] CHAO H Y, BAUMANN M, JENSEN A, et al. Band-reconfigurable multi-UAV-based cooperative remote sensing for real-time water management and distributed irrigation control[J]. IFAC Proceedings Volumes, 2008, 41(2):11744-11749.
- [9] VACHTSEVANOS G, KIM W, AL-HASAN S, et al. Autonomous vehicles: from flight control to mission planning using fuzzy logic techniques[C]// International Conference on Digital Signal Processing. Piscataway: IEEE, 1997:977-981.
- [10] STENTZ A. Optimal and efficient path planning for partially-known environments [C] // International Conference on Robotics and Automation. Piscataway: IEEE,1994:3310-3317.
- [11] PEHLIVANOGLU Y V, PEHLIVANOGLU P. An enhanced genetic algorithm for path planning of autonomous UAV in target coverage problems [J]. Applied Soft Computing,2021,112:107796.
- [12] 许诺,朱黔,谢晓阳,等. 基于改进混合粒子群优化算法的多无人机协同围捕方法研究[J]. 电光与控制,2024, 31(9):1-5.

- [13] DEWANGAN R K, SHUKLA A, GODFREY W W. Three dimensional path planning using grey wolf optimizer for UAVs[J]. *Applied Intelligence*, 2019, 49(6):2201-2217.
- [14] ZHANG Y F, WANG H L. Adaptive interfered fluid dynamic system algorithm based on deep reinforcement learning framework[C]// *International Conference on Autonomous Unmanned Systems*. Singapore: Springer, 2022:1388-1397.
- [15] WANG H L, LYU W T, YAO P, et al. Three-dimensional path planning for unmanned aerial vehicle based on interfered fluid dynamical system[J]. *Chinese Journal of Aeronautics*, 2015, 28(1):229-239.
- [16] 吴健发,王宏伦,王延祥,等. 无人机反应式扰动流体路径规划[J]. *自动化学报*, 2023, 49(2):272-287.
- [17] ZHANG Y T, ZHANG Y M, YU Z Q. Path following control for UAV using deep reinforcement learning approach[J]. *Guidance, Navigation and Control*, 2021, 1(1):2150005.
- [18] 牟文心,时宏伟. 基于改进 TD3 算法的无人机轨迹规划[J]. *计算机系统应用*, 2024, 33(12):197-209.
- [19] 王伟伦,尤明,孙磊,等. 未知环境下无人机集群智能协同探索路径规划[J]. *工程科学学报*, 2024, 46(7):1197-1206.
- [20] 聂铭涛,苏玉展,陈改改,等. 基于确定性策略搜索的无人机协同路径规划算法[J]. *飞行力学*, 2024, 42(4):48-55.
- [21] LOWE R, WU Y, TAMAR A, et al. Multi-agent actor-critic for mixed cooperative-competitive environments [EB/OL]. [2024-09-01]. <https://arxiv.org/pdf/1706.02275>.
- [22] 杜江涛,于家明,齐辉. 无人机集群不完全信息路径规划方法[J]. *哈尔滨工程大学学报*, 2024, 45(11):2210-2217.
- [23] ACKERMANN J, GABLER V, OSA T, et al. Reducing overestimation bias in multi-agent domains using double centralized critics [EB/OL]. [2024-09-01]. <https://arxiv.org/pdf/1910.01465>.
- [24] 姚鹏,王宏伦. 基于改进流体扰动算法与灰狼优化的无人机三维航路规划[J]. *控制与决策*, 2016, 31(4):701-708.
- [25] 袁利平,陈宗基,周锐,等. 多无人机同时到达的分散化控制方法[J]. *航空学报*, 2010, 31(4):797-805.
- [26] QIE H, SHI D X, SHEN T L, et al. Joint optimization of multi-UAV target assignment and path planning based on multi-agent reinforcement learning[J]. *IEEE Access*, 2019, 7:146264-146272.
- [27] 唐恒,孙伟,吕磊,等. 融合动态奖励策略的无人机编队路径规划方法[J]. *系统工程与电子技术*, 2024, 46(10):3506-3518.
- [28] YU C, VELU A, VINITSKY E, et al. The surprising effectiveness of MAPPO in cooperative, multi-agent games[EB/OL]. [2024-09-01]. <https://arxiv.org/pdf/2103.01955v1>.

(责任编辑:任滢滢)