

一种无人机编队控制方法研究与仿真

陈泽坤*, 何杏宇

上海理工大学光电信息与计算机工程学院, 上海

收稿日期: 2024年4月20日; 录用日期: 2024年5月20日; 发布日期: 2024年5月27日

摘要

现有的无人机群编队控制研究难以在障碍物环境下兼顾无人机的避障一致性和避障灵活性。为此, 本文提出了一种基于强化学习的差异化无人机编队控制方法。该方法允许无人机群中的任一无人机根据其局部环境在编队聚集、编队保持和避障之间改变其编队控制策略, 并设计了一个强化学习模型为上述三种控制策略下的无人机计算最优偏移向量。仿真结果表明, 该方法可以有效地兼顾无人机群的避障灵活性和一致性, 从而提高其飞行效率并保持其网络连通性。

关键词

编队控制, 避障一致性, 避障灵活性, 强化学习

Research and Simulation of a UAV Formation Control Method

Zekun Chen*, Xingyu He

School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai

Received: Apr. 20th, 2024; accepted: May 20th, 2024; published: May 27th, 2024

Abstract

In existing formation control researches, it is hardly to reconcile both the obstacle avoidance consistency and obstacle avoidance flexibility of the UAV swarm. In view of this, this paper proposes a reinforcement learning based differentiated formation control method for UAVs. This method allows any UAV within the swarm to adapt its formation control strategy among formation aggregation, formation maintenance, and obstacle avoidance based on its local environment, and a rein-

*通讯作者。

forcement learning model is designed to compute the optimal offset vector for each UAV under the three aforementioned formation control strategies. Simulation results demonstrate that the method can effectively balance the maneuverability and consistency of obstacle avoidance for UAV swarms, thereby improving their flight efficiency and maintaining network connectivity.

Keywords

Formation Control, Obstacle Avoidance Consistency, Obstacle Avoidance Flexibility, Reinforcement Learning

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

无人机凭借其尺寸、速度和机动性等方面的优势,在民事领域和军事领域中已经得到广泛的应用。由多无人机组成的编队可以通过协同控制更高效、更彻底地完成单无人机难以完成的任务。例如,灾害救援和边境巡逻等[1] [2]。编队控制是无人机群协同执行任务研究中的一个重点,其关键点在于如何实现无人机编队保持与避障。

现有常见的编队控制方法有基于行为的方法、基于虚拟结构的方法以及主从编队控制方法等[3] [4]。基于行为的方法根据无人机本身特性和编队行为模式设计控制策略,通过模拟生物群体或物理系统的自组织规则来实现编队协同[5],文献[6]提出一种基于行为的编队控制方法,方法给出了编队保持和编队移动的运动矢量方程,并定义了不同行为的重要性权重。这种方法通过为每个要解决的问题添加行为,不仅可以解决运动问题,同时还能解决身份验证等其他问题;基于虚拟结构的方法将无人机编队视为一个称为虚拟结构的单一刚体系统,其中每个无人机的位置和姿态都是相对于连接到整个结构上的虚拟机身框架[7],文献[8]的作者提出了一种将虚拟结构与差分平坦反馈控制相结合的四旋翼无人机编队控制方法,以实现编队的敏捷机动;主从编队方法则指定一架无人机作为领导者,其它无人机作为追随者,领导者决定自身的飞行轨迹,追随者保持与领导者的相对距离来实现编队飞行[9] [10]。但这种方法存在过度依赖领导无人机的问題,为此,[11]的作者提出了一种动态领导者选择的编队控制方法,用于解决领导者出现故障导致编队失去其能效的问题。

从系统结构的角皮,上述方法大致可以分为两类,集中式编队控制(基于虚拟结构,主从式编队)和分布式编队控制(基于行为)。集中式编队控制方法有利于编队保持稳定的编队结构和通信拓扑,因此它们具有更好的避障一致性,但灵活性较差,无法有效地应对环境中的突发事件,例如遭遇密集障碍物[12]。相反,分布式编队控制方法由于其编队结构可变,可以极大地提高灵活性,但代价是编队的避障一致性较差,可能会出现编队通信拓扑被破坏,从而导致编队中的无人机无法实现有效地持续协同[13]。

通过上述分析可以发现,现有无人机编队控制方法中存在无法兼顾避障一致性和避障灵活性的两难局面。也就是说,固定队形的编队控制方法可以保持无人机之间稳定、连续的通信,但在障碍物场景下的避障灵活性有很大限制[14],而可变形无人机群控制方法通过允许无人机在避障时暂时放弃编队限制来提高避障灵活性[15],但它们在避障一致性的实现上仍然存在困难。由此可见,打破上述困境的关键是设计一种能够保持稳定通信连通性的松散编队控制方法。

鉴于上述问题,本文提出了一个同时兼顾无人机编队保持和灵活避障的编队控制方法。该方法允许

无人机机群中的任一无人机根据其局部环境在编队聚集、编队保持和避障之间改变其编队控制策略, 然后根据所选策略计算其当前时刻的移动子目标。编队聚集策略适用具有脱离编队整体风险的无人机, 编队保持策略用于已经与无人机编队整体保持一致的无人机, 而避障策略适用于遭遇障碍物的无人机, 它要求无人机共享其避障信息和避障行动, 从而让无人机选择一致的避障方向。此外, 该方法为无人机设计了一个通用的强化学习模型, 该模型能够根据无人机当前时刻自身状态以及该时刻所选择的控制策略, 输出无人机的最优偏移向量, 从而实现编队控制的效果。本文在不同的实验场景下进行了仿真实验, 实验结果表明与现有其它方法相比较, 本文所提出的方法能够有效地兼顾无人机的避障的灵活性和避障一致性, 从而实现更高的飞行效率和保持稳定的网络通信。

2. 编队控制模型设计

2.1. 无人机模型

在本文的无人机编队控制场景中, 假设每个无人机都配备了感知、计算、存储和通信设备, 无人机通过自身的传感器和无线通信网络获得自身和编队中其它无人机的状态信息。如图 1 所示, 我们将无人机的运动视为质点运动, 无人机的通信距离为 D_c , 处于 UAV_i 通信距离内的所有无人机定义为其邻居无人机。编队距离 D_f , 编队距离用于确保无人机形成相对稳定的编队。需要注意的是, D_f 总是小于 D_c , 从而确保无人机之间的连通性保持。碰撞距离 D_s , 无人机之间的距离不能小于 D_s 。处于 UAV_i 碰撞距离和编队距离之间的无人机为其“伙伴”无人机, 本文规定若某一架无人机有两个及以上伙伴, 则称其处于稳定状态。无人机的飞行目标是在尽快短的时间内到达预定义的全局目标位置 P_g 。

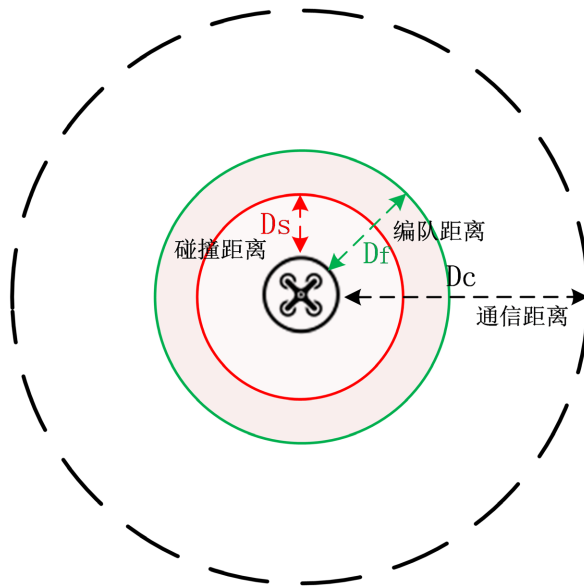


Figure 1. Communication, partnership and collision ranges of UAVs
图 1. 无人机的通信、伙伴关系和碰撞范围

定义 Δ_i 为 UAV_i 当前实际位置 P 和无人机目标位置 P_g 之间的理想的位置偏差, $\Delta_i = [\Delta_{ix}, \Delta_{iy}, \Delta_{iz}]^T \in R^3$, $i \in \Gamma$, 则无人机的控制目标表示为:

$$\lim_{t \rightarrow \infty} \sum_{i=1}^n \Delta_i = 0 \tag{1}$$

2.2. 差异化编队控制策略

差异化编队控制策略是指编队中的成员在同一时刻能够根据自身的状态执行不同的编队控制策略。本文提出的编队控制方法中分别定义了编队聚集、编队保持和避障三种不同的编队控制策略, 以及对应策略下的移动子目标计算方法。通过采用差异化编队控制策略, 一方面赋予无人机群中单架无人机足够的灵活性, 使其能够根据自身的环境状态选择性地选择合适的编队控制策略。另一方面, 它允许位置相近的无人机有效地实现编队聚集和协同避障。

编队聚集: 无人机通过执行聚集操作来达到稳定状态。执行聚集操作的无人机将通信范围内最近的邻居作为移动子目标, 直到与其成为伙伴。

编队保持: 若无人机已经处于稳定状态, 则执行编队保持操作, 与相邻个体的速度和方向保持一致向目的地移动, 此时, 将全局目标位置 P_g 作为无人机当前时刻的移动子目标。

避障: 若无人机通过传感器检测到当前的前进方向上有障碍物, 此时避障子目标的确定有两种情况:

1) 无人机没有收到来自其他无人机的避障信息。此时无人机会先将探测到的障碍物信息形成障碍物大致轮廓, 并根据与障碍物的相对位置, 将障碍轮廓投影在 XOY , XOZ 和 YOZ 平面上, 并计算障碍物投影在各个平面上的边缘端点之间的距离, 将距离最短的障碍物轮廓端点所在的平面作为最优避障平面。如图 2 所示, 无人机将障碍物端点距离更短的 $A'B'$ 所在的平面 XOY 作为避障平面。确定避障平面后, 选择距离全局目标位置更近的障碍物端点作为避障移动子目标, 从而绕过障碍物。同时广播自身的避障信息, 避障信息包括其无人机标识符、无人机位置信息、避障子目标。收到避障信息的无人机会将收到的避障信息加入自身的避障信息列表。

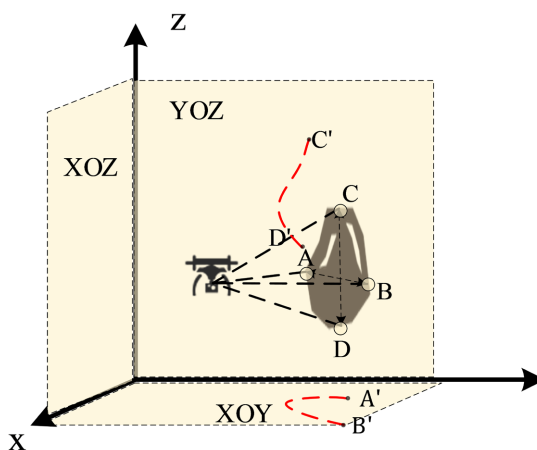


Figure 2. Obstacle avoidance plane selection

图 2. 避障平面选择

2) 若无人机已经收到了来自其它无人机的避障信息, 无人机将避障信息列表中距离自身最近的无人机的避障子目标作为自身当前时刻的移动子目标。通过该策略, 无人机能与自身相近的无人机保持避障方向的一致性。

为了协调无人机的策略选择, 本文定义了差异化编队控制策略选择方法。首先, 无人机根据传感器检测到的环境信息进行策略选择。如果某架无人机检测到当前前进方向存在障碍物, 执行避障策略, 绕过障碍物, 并广播避障信息, 确保收到信号的无人机在执行碰撞避免时与编队整体移动方向保持一致。如果无人机已经存在两个以上的“伙伴”, 无人机已处于稳定状态, 此刻以较大概率执行编队保持, 向

预定义的全局目标位置移动。为了避免无人机群分成多个稳定的小团体, 导致整个系统网络图不连通, 我们要让无人机以一定的概率 τ 在稳定状态执行编队聚集, 这个概率随时间增大而逐渐趋于零, 因为经过一段时间的飞行编队大概率已经处于整体联通的状态。若无人机通信范围内存在邻居, 且无人机还不处于稳定状态, 则执行编队聚集捕获伙伴, 以尽快形成稳定的状态。若无人机通信范围内没有邻居, 此时无人机因为意外情况的累积导致处于与系统整体断联, 只能执行编队保持向目的地移动, 这也是尽快与系统整体建立整体的唯一方法。

2.3. 强化学习模型

本文将无人机群向全局目标位置 P_g 移动的过程建模为一个强化学习顺序决策过程, 在这个过程中了, 为了兼顾编队在障碍物存在的环境下避障一致性和避障灵活性, 我们设计了一个通用的差异化编队控制策略强化学习模型, 如图 3 所示。

在该模型中, 无人机在 t 时刻的状态为 $S(t)$, 无人机根据 $S(t)$ 选择适合当前时刻的编队控制策略, 再根据所选择的策略计算得到 t 时刻自身的移动子目标 $P_{subgoal}(t)$, 然后将 $S(t)$ 和 $P_{subgoal}(t)$ 输入到经过良好训练的强化学习模型中, 得到移动偏移矢量 $\delta(t)$, 作用于无人机得到无人机下一时刻的状态 $S(t+1)$ 。通过迭代移动到全局目标位置 P_g 。

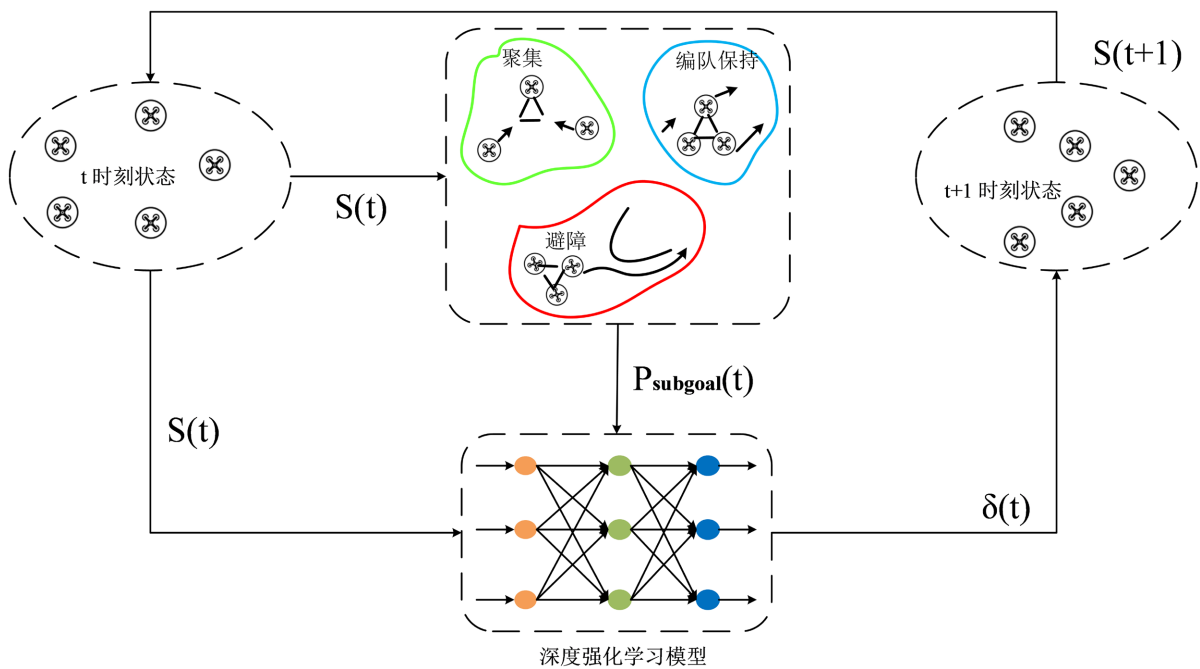


Figure 3. Formation control reinforcement learning model

图 3. 编队控制强化学习模型

在本文的编队控制强化学习模型中, 无人机的状态向量定义为:

$$s(t) = [P(t), d_{goal}(t), d_{subgoal}(t)] \tag{2}$$

其中 $P(t) = \{x(t), y(t), z(t)\}$, 表示当前无人机在 t 时刻的位置, d_{goal} , $d_{subgoal}$ 分别表示当前无人机距全局目标位置的距离和当前时刻子目标的距离。为了得到无人机不同情况下的最优策略 $\pi^* : (s(t), a(t))$, 本文考虑连续的动作空间 $a \in R^3$, 无人机的动作 $a(t)$ 为偏移矢量 $\delta(t)$:

$$P_i(t+1) = P_i(t) + \delta_i(t) \quad (3)$$

$$\pi_{\theta^\pi}(s_i(t), a_i(t)) = \delta_i(t) = [\delta_i^x(t), \delta_i^y(t), \delta_i^z(t)] \quad (4)$$

$$\|\delta_i(t)\| \leq \delta_{\max} \quad (5)$$

其中 $\delta_i^x(t), \delta_i^y(t), \delta_i^z(t)$ 分别表示无人机在 t 时刻在 x, y, z 方向上期望的位置增量, θ 表示策略网络的参数。此外, 为了确保智能体下一时刻的位置处于智能体自我规划的范围内, 我们给偏移矢量 δ 设定阈值 δ_{\max} 。为了鼓励无人机移动到子目标位置, 我们需要设计奖励函数来激励无人机向子目标位置移动, 同时惩罚不好的动作。奖励函数的设计如下:

$$r_i(s, a) = \begin{cases} r_{\text{general}} & \text{if } (p_i(t) = P_i^{\text{sub}}(t)) \\ r_{\text{punish}} & \text{if } (d_i^c(t) < ds) \\ r_{\text{const}}(d_i^{\text{sub}}(t-1) - d_i^{\text{sub}}(t)) & \text{otherwise} \end{cases} \quad (6)$$

其中 r_{general} 是 UAV i 成功移动到子目标位置的奖励常量, $P_i^{\text{sub}}(t)$ 是 UAV i 在 t 时刻的子目标位置, r_{punish} 是无人机与环境其它对象发生碰撞的惩罚常量, $d_i^c(t)$ 是 UAV i 在 t 时刻与其最近的无人机的距离。 r_{const} 是无人机向子目标位置移动的奖励常量, $d_i^{\text{sub}}(t)$ 是 UAV i 在 t 时刻距离子目标的距离。

本文使用深度确定性策略梯度(DDPG)算法来训练策略网络, DDPG 算法在训练连续动作空间强化学习策略网络参数的性能已经被证明[16]。无人机的行动决策模型由 Actor 网络和 Critic 网络组成, Actor 网络和 Critic 网络通过双向循环网络 BRNN 连接而成。模型的输入为无人机的状态信息, 输出无人机的偏移矢量。由于模型基于 BRNN 构建, 因此, 对于网络参数学习的思路是将网络展开成 n 个子网络计算反向梯度, 然后使用基于时间的反向传播算法更新网络参数。梯度在每个无人机个体的 Q 函数和策略函数中传播, 模型学习时, 各个无人机个体奖励值影响各个无人机的行动, 进而由此产生的梯度信息反向传播并更新模型参数。

本文定义无人机的目标函数为:

$$J(\theta) = E_{d_{\pi_\theta}(s)} \left[\sum_{i=1}^n r_i(s, a_\theta(s)) \right] \quad (7)$$

它表示奖励值的累加期望, 其中, $d_{\pi_\theta}(s)$ 是基于策略 π_θ 生成的马尔可夫链关于状态的静态分布。根据确定性策略梯度定理, 对 θ 求导的梯度可以表示为:

$$\nabla_\theta J(\theta) = E_{d_{\pi_\theta}(s)} \left[\sum_{i=1}^n \nabla_\theta a_\theta(s) \nabla_a Q_i^{a_\theta}(s, a_\theta(s)) \right] \quad (8)$$

由于本文的动作空间连续, 我们不直接计算每一个动作的概率, 而是学习概率分布的统计量。为了得到一个参数化的策略函数, 我们将策略定义为向量动作的正态概率密度, 其中均值和标准差由状态的参数化函数近似给出,

$$\pi_{\theta^\pi}(s_i) \doteq \frac{1}{\sigma(s, \theta) \sqrt{2\pi}} \exp \left(-\frac{(a - \mu(s, \theta))^2}{2\sigma(s, \theta)^2} \right) \quad (9)$$

其中 $\theta = [\theta_\mu, \theta_\sigma]^T$, θ_μ 用来近似均值, θ_σ 用来近似标准差。均值可以用一个线性函数来逼近, 标准差必须为正数, 因此 $\mu(s, \theta)$ 和 $\sigma(s, \theta)$ 可以近似计算如下:

$$\mu(s, \theta) \doteq \theta_{\mu}^T x_{\mu}(s) \quad (10)$$

$$\sigma(s, \theta) \doteq \exp(\theta_{\sigma}^T x_{\sigma}(s)) \quad (11)$$

其中, $x_{\mu}(s)$ 和 $x_{\sigma}(s)$ 是状态特征向量。基于公式(8), 采用随机梯度下降法优化 Actor 和 Critic 网络, 在交互学习的过程中, 通过试错获取的数据更新参数, 完成飞行策略的优化。

3. 仿真实验与结果分析

3.1. 对比方法和实验指标

在本节中, 将本文所提出的方法与 LF-DQN [17]、LF-DDPG [18]、DQN [19]和 DDPG [20]在不同障碍物分布场景下的网络连通性和飞行效率进行了比较, 以验证本文方法在避障灵活性和一致性的优势。在本文中, LF-DQN 和 LF-DDPG 是典型的基于强化学习的集中式编队控制方法, DQN 和 DDPG 是典型的基于强化学习的分布式编队控制方法。不同场景的设置如下:

场景 1: 障碍物大小为 5, 障碍物数量在 10 到 20 之间。两个相邻障碍物之间的距离在 20 到 40 之间, 大于无人机碰撞距离的两倍, 但小于无人机碰撞距离的三倍(当两个障碍物之间距离小于 30 时, 形成三角形的无人机群无法通过两个障碍之间的间隙), 这适合于测试无人机群的避障灵活性。

场景 2: 障碍物大小为 5, 障碍物数量在 10 到 20 之间。两个相邻障碍物之间的距离大于 60, 这足以让三角形编队的无人机群通过, 因此, 这适合于测试无人机群中的避障一致性。

为了验证本文方法有效性, 从无人机群飞行效率、网络连通性两项指标对本文方法和对比方法进行评价, 分析不同实验场景下对实验指标的影响。具体的, 无人机群飞行效率是指无人机群完成单次飞行任务所采取的动作次数。网络连通性是指无人机群在单次飞行过程中的连通性保持情况, 可通过深度优先遍历算法计算。仿真实验的相关超参数如表 1 所示。

Table 1. Parameter settings
表 1. 参数设置

参数	值
无人机规模	3
环境空间大小	200
碰撞距离 D_s	5
编队距离 D_f	10
通信距离 D_c	30
单步偏移量阈值 δ_{\max}	5
训练过程最大迭代次数	1000

3.2. 实验结果分析

3.2.1. 有效性验证

在模型的训练过程中, 环境空间中障碍物的数量、大小和位置是随机生成的。图 4 显示了所提出的模型在随机环境中无人机的飞行轨迹。可以看出, 无人机在飞行过程中都可以保持一定的队形, 并避免与障碍物发生碰撞。

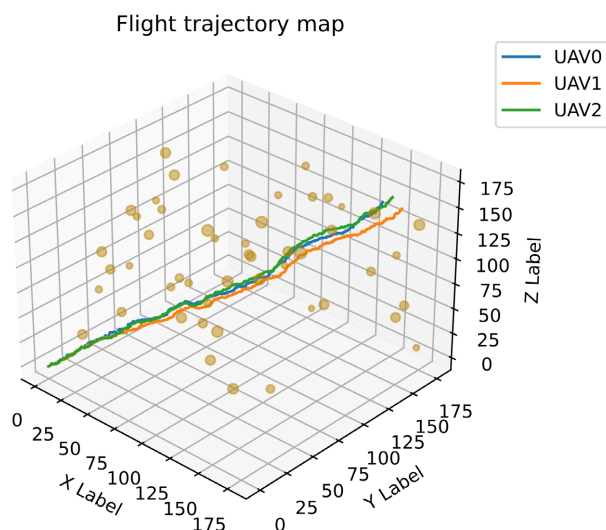
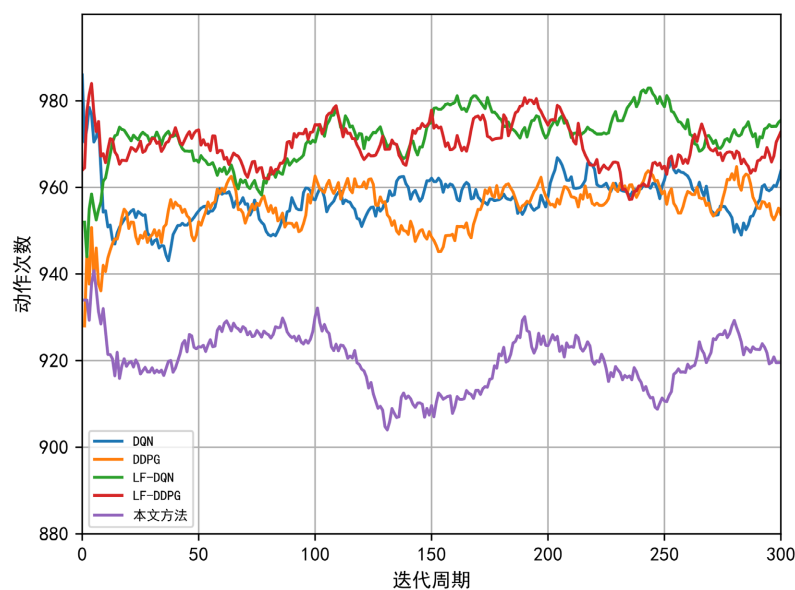


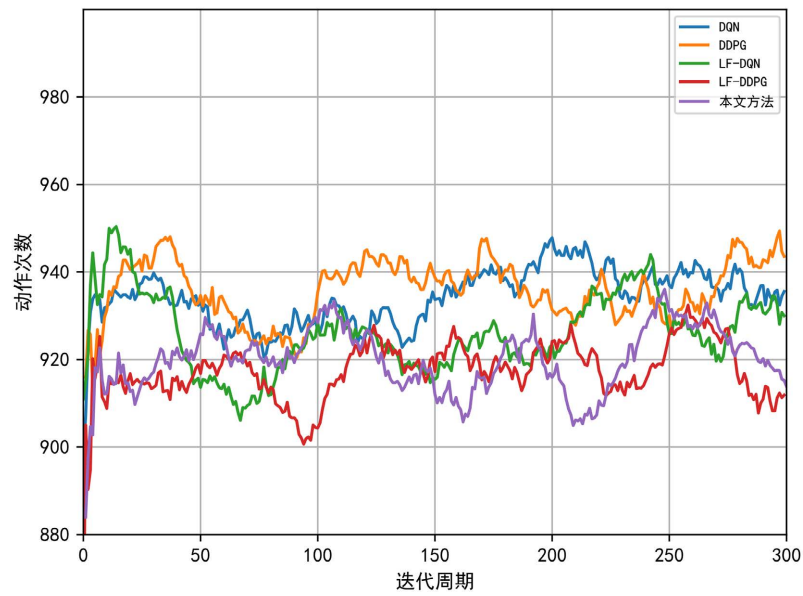
Figure 4. Flight trajectory map
图 4. 飞行轨迹图

3.2.2. 不同方法飞行效率比较

本文方法与 4 种对比方法的飞行效率比较如图 5 所示。具体而言, 在场景 1 中, 本文所提出的方法具有最高的飞行效率(所采取的平均动作次数为 919), DQN 和 DDPG 具有中等的飞行效率(DQN 和 DDPG 所采取的平均动作次数分别为 957 和 955), LF-DDPG 和 LF-DQN 表现出最低和第二低的飞行效率(LF-DQN 和 LF-DDPG 所采取的平均动作次数分别为 969 和 971)。可以解释为, 在场景 1 中, 集中控制方式由于其固定的编队控制, 避障灵活性较差, 集中控制方法中的无人机无法通过尺寸小于 30 的障碍物间隙, 并且需要花费更多的时间来搜索尺寸大于 30 的障碍物间隙。由于分布式编队控制不要求队形固定, 分布式控制方法具有良好的避障灵活性, 但其避障一致性较差, 从而影响其飞行效率。所提出的方法能够结合良好的避障灵活性和一致性, 从而使无人机以更高的效率移动到最终目标。



(a) 场景 1 中不同方法的飞行效率

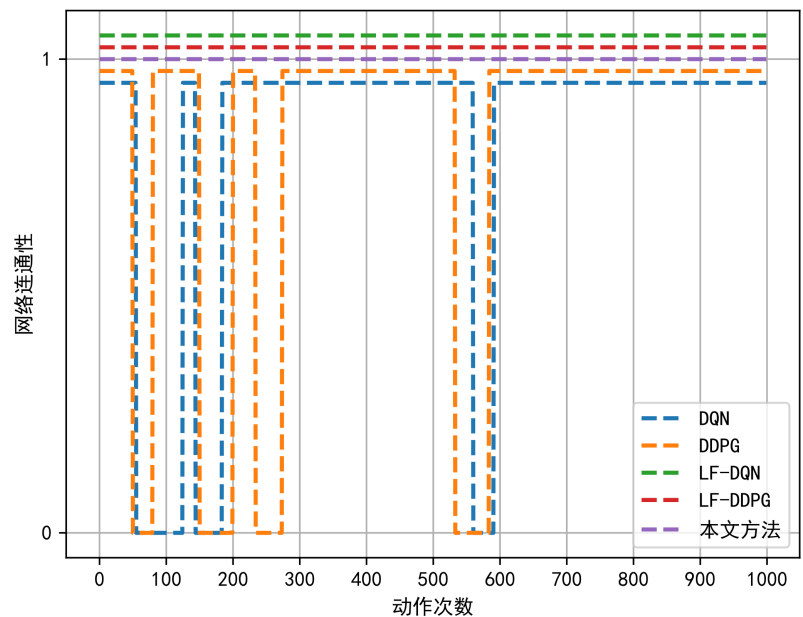


(b) 场景 2 中不同方法的飞行效率

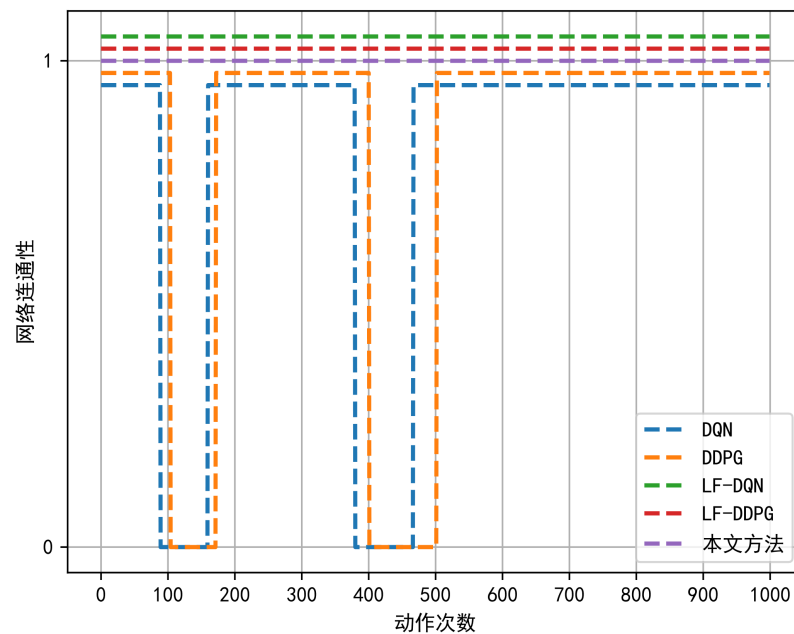
Figure 5. Flight efficiency comparison
图 5. 飞行效率比较

在场景 2 中, DQN 和 DDPG 表现出最低和第二低的飞行效率(DQN 和 DDPG 所采取的平均动作次数分别为 936 和 934), 本文所提出的方法的飞行效率与 LF-DDPG 和 LF-DDQN 非常接近(本文方法所采取的平均动作次数为 924, LF-DDPG 和 LF-DQN 分别为 917 和 926), 这是因为, 在场景 2 中, DQN 和 DDPG 中的无人机避障一致性较差, 其中一些无人机可能会选择不同并且更长的避障路线, 相反, 所提出的方法和集中编队控制方法中的无机可以在最佳避障路线选择上实现一致性。从而表现出更高的飞行效率。

3.2.3. 不同方法网络连通性比较



(a) 场景 1 中不同方法的网络连通性



(b) 场景 2 中不同方法的网络连通性

Figure 6. Network connectivity comparison

图 6. 网络连通性比较

图 6 比较了本文方法和 4 种对比方法在不同场景下编队的网络连通性, 如果连通性值为 1, 则无人机群处于完全联通状态; 否则机群不连通。

根据图 6 可以看出, 集中式编队方法由于其固定编队控制具有更好的避障一致性, 在两种场景中都能保持网络完全连通。分布式编队控制方法由于编队控制松散导致其避障一致性较差, 在两种场景中的任何一种情况下都无法保持完全的网络连通。具体来说, 网络连通性不完整的原因是无人机可能会选择不同的障碍物间隙通过。本文所提出的方法由于聚集策略的存在, 同样能够在两种场景中保持完全联通的状态。

4. 结论

本文提出了一种基于强化学习的无人机差异化编队控制方法。该方法通过差异化编队控制策略, 允许编队中的任一无人机根据其局部环境, 在编队聚集、编队保持和避障三种不同的控制策略之间灵活选择, 从而实现兼顾无人机避障一致性和避障灵活性的效果。在不同障碍物场景下的仿真实验结果表明, 相对于传统的分布式编队控制方法和集中式编队控制方法, 本文所提出的方法在编队飞行效率和维持编队网络连通性方面具有优势。具体而言, 本文通过对分布式编队控制方法的飞行一致性进行改善, 达到了保证编队网络连通性的同时并提升编队飞行效率的效果。

基金项目

国家自然科学基金资助项目(61602305, 61802257)、上海市自然科学基金资助项目(18ZR1426000, 19ZR1477600)。

参考文献

- [1] Zhang, H. and Hanzo, L. (2020) Federated Learning Assisted Multi-UAV Networks. *IEEE Transactions on Vehicular*

- Technology*, **69**, 14104-14109. <https://doi.org/10.1109/TVT.2020.3028011>
- [2] Wu, Y., Low, K.H. and Lv, C. (2020) Cooperative Path Planning for Heterogeneous Unmanned Vehicles in a Search-and-Track Mission Aiming at an Underwater Target. *IEEE Transactions on Vehicular Technology*, **69**, 6782-6787. <https://doi.org/10.1109/TVT.2020.2991983>
- [3] Wu, J., Luo, C., Luo, Y., et al. (2021) Distributed UAV Swarm Formation and Collision Avoidance Strategies Over Fixed and Switching Topologies. *IEEE Transactions on Cybernetics*, **52**, 10969-10979. <https://doi.org/10.1109/TCYB.2021.3132587>
- [4] Guo, J., Liu, Z., Song, Y., et al. (2023) Research on Multi-UAV Formation and Semi-Physical Simulation with Virtual Structure. *IEEE Access*, **11**, 126027-126039. <https://doi.org/10.1109/ACCESS.2023.3330149>
- [5] Nursyeha, M.A., Rivai, M. and Purwanto, D. (2020) LiDAR Equipped Robot Navigation on Behavior-Based Formation Control for Gas Leak Localization. 2020 *International Seminar on Intelligent Technology and Its Applications (ISITIA)*, Surabaya, 22-23 July 2020, 89-94. <https://doi.org/10.1109/ISITIA49792.2020.9163758>
- [6] Lee, G. and Chwa, D. (2018) Decentralized Behavior-Based Formation Control of Multiple Robots Considering Obstacle Avoidance. *Intelligent Service Robotics*, **11**, 127-138. <https://doi.org/10.1007/s11370-017-0240-y>
- [7] Chen, Q., Wang, Y. and Lu, Y. (2021) Formation Control for UAVs Based on the Virtual Structure Idea and Nonlinear Guidance Logic. 2021 *6th International Conference on Automation, Control and Robotics Engineering (CACRE)*, Dalian, 15-17 July 2021, 135-139. <https://doi.org/10.1109/CACRE52464.2021.9501340>
- [8] Zhou, D., Wang, Z. and Schwager, M. (2018) Agile Coordination and Assistive Collision Avoidance for Quadrotor Swarms Using Virtual Structures. *IEEE Transactions on Robotics*, **34**, 916-923. <https://doi.org/10.1109/TRO.2018.2857477>
- [9] Wen, G., Chen, C.L.P., Liu, Y.J., et al. (2016) Neural Network-Based Adaptive Leader-Following Consensus Control for a Class of Nonlinear Multiagent State-Delay Systems. *IEEE Transactions on Cybernetics*, **47**, 2151-2160. <https://doi.org/10.1109/TCYB.2016.2608499>
- [10] Xiao, H. and Chen, C.L.P. (2019) Leader-Follower Consensus Multi-Robot Formation Control Using Neurodynamic-Optimization-Based Nonlinear Model Predictive Control. *IEEE Access*, **7**, 43581-43590. <https://doi.org/10.1109/ACCESS.2019.2907960>
- [11] Le, J., Zhao, L., Wang, C., et al. (2020) Leader-Follower Optimal Selection Method for Distributed Control System in Active Distribution Networks. *CSEE Journal of Power and Energy Systems*, **10**, 314-323.
- [12] Lyu, Y., Hu, J., Chen, B.M., et al. (2019) Multivehicle Flocking with Collision Avoidance via Distributed Model Predictive Control. *IEEE Transactions on Cybernetics*, **51**, 2651-2662. <https://doi.org/10.1109/TCYB.2019.2944892>
- [13] Chen, Q., Jin, Y., Wang, T., et al. (2022) UAV Formation Control Under Communication Constraints Based on Distributed Model Predictive Control. *IEEE Access*, **10**, 126494-126507. <https://doi.org/10.1109/ACCESS.2022.3225434>
- [14] Luo, L., Wang, X., Ma, J., et al. (2021) GrpAvoid: Multigroup Collision-Avoidance Control and Optimization for UAV Swarm. *IEEE Transactions on Cybernetics*, **53**, 1776-1789. <https://doi.org/10.1109/TCYB.2021.3132044>
- [15] Zhu, X., Liang, Y. and Yan, M. (2019) A Flexible Collision Avoidance Strategy for the Formation of Multiple Unmanned Aerial Vehicles. *IEEE Access*, **7**, 140743-140754. <https://doi.org/10.1109/ACCESS.2019.2944160>
- [16] Liu, D., Dou, L., Zhang, R., et al. (2022) Multi Agent Reinforcement Learning-Based Coordinated Dynamic Task Allocation for Heterogenous UAVs. *IEEE Transactions on Vehicular Technology*, **72**, 4372-4383. <https://doi.org/10.1109/TVT.2022.3228198>
- [17] Yan, C., Xiang, X. and Wang, C. (2020) Fixed-Wing UAVs Flocking in Continuous Spaces: A Deep Reinforcement Learning Approach. *Robotics and Autonomous Systems*, **131**, Article 103594. <https://doi.org/10.1016/j.robot.2020.103594>
- [18] Raja, G., Essaky, S., Ganapathisubramaniyan, A., et al. (2022) Nexus of Deep Reinforcement Learning and Leader-Follower Approach for AIOT Enabled Aerial Networks. *IEEE Transactions on Industrial Informatics*, **19**, 9165-9172. <https://doi.org/10.1109/TII.2022.3226529>
- [19] Mousavifard, R., Alipour, K., Najafqolian, M.A., et al. (2022) Formation Control of Multi-Quadrotors Based on Deep Q-Learning. 2022 *10th RSI International Conference on Robotics and Mechatronics (ICRoM)*, Tehran, 22-24 November 2022, 172-177. <https://doi.org/10.1109/ICRoM57054.2022.10025251>
- [20] Wang, W., Wang, L., Wu, J., et al. (2022) Oracle-Guided Deep Reinforcement Learning for Large-Scale Multi-UAVs Flocking and Navigation. *IEEE Transactions on Vehicular Technology*, **71**, 10280-10292. <https://doi.org/10.1109/TVT.2022.3184043>