

基于 Q-Learning 的双无人机覆盖路径规划

陈佳雨¹, 李文¹, 李泰融², 李志茹¹, 王子怡¹, 陈鹏云¹
(1 中北大学 太原 030000;
2 济南绿色城市发展投资集团有限公司 济南 250000)

摘要: 覆盖路径规划的目标是确保无人机能够实现目标区域的完全覆盖。在以往的研究中, 无人机的工作模式为分别负责每个子区域的覆盖任务, 而本研究中两架无人机在整个搜索区域中协同工作, 能够在提高覆盖效率的基础上更加灵活地实现覆盖任务。针对传统方法求解无人机覆盖路径规划易导致规划总代价高的问题, 本文提出一种基于 Q-Learning (Q 学习) 的双无人机覆盖路径规划算法。为了节省无人机完成覆盖搜索任务的时间, 采用基于网格的旋转区域划分算法最小化待搜索区域。通过建立无人机覆盖路径规划模型, 将路径规划转化为多目标函数优化问题, 并利用 Double-Q-Learning (双 Q 学习) 算法均衡算法全局搜索与局部开发能力, 以综合考虑距离代价和转折代价的总代价函数, 对路径规划迭代寻优。仿真实验结果表明: 所提算法规划的路径能够以更低的总代价在不同的目标区域中实现两架无人机对目标区域的完全覆盖。

关键词: 覆盖路径规划; 双无人机; Double-Q-Learning; 协同控制; 旋转区域; 多目标函数

中图分类号: V27

文献标志码: A

文章编号: 2095-1000(2025)04-0096-09

DOI: 10.12347/j.ycyk.20241022001

CSTR: 32406.14.ycyk.20241022001

引用格式: 陈佳雨, 李文, 李泰融, 等. 基于 Q-Learning 的双无人机覆盖路径规划[J]. 遥测遥控, 2025, 46(4): 96-104.

Q-Learning Based Dual Drone Coverage Path Planning

CHEN Jiayu¹, LI Wen¹, LI Tairong², LI Zhiru¹, WANG Ziyi¹, CHEN Pengyun¹

(1. North University of China, Taiyuan 030000, China;

2. Jinan Green City Development Investment Group Co., Ltd, Jinan 250000, China)

Abstract: The objective of coverage path planning is to ensure that Unmanned Aerial Vehicles (UAVs) achieve complete coverage of the target area. Previous studies assigned UAVs the task of covering each sub-area separately. However, this study proposes a new methodology in which two UAVs collaborate across the entire search area, achieving coverage tasks more flexibly while enhancing efficiency. This paper aims to address the high cost of traditional UAV coverage path planning by proposing a dual-UAV coverage path planning algorithm based on Q-Learning. To reduce the time taken for the process, a grid-based rotating area partitioning algorithm is used to minimize the search area. The path planning is transformed into a multi-objective function optimisation problem, and the Double-Q-Learning algorithm balances global search and local exploitation, iteratively optimising the path with a total cost function that considers distance and turning costs. The simulation results demonstrate that the proposed algorithm can achieve complete coverage of different target areas with a lower total cost.

Keywords: Coverage path planning; Dual unmanned aerial vehicles; Double-Q-Learning; Collaborative control; Rotating area; Multi-Objective function

Citation: CHEN Jiayu, LI Wen, LI Tairong, et al. Q-Learning Based Dual Drone Coverage Path Planning[J]. Journal of Telemetry, Tracking and Command, 2025, 46(4): 96-104.

0 引言

随着社会发展与需求增长, 各领域对大面积区域精准覆盖的需求日益迫切, 包括军事侦察、

搜索救援、地形测绘等^[1-3]。在地形测绘领域, 无论是对广袤山脉的地形绘制, 还是对辽阔海洋的海底地貌测绘, 或是对复杂城市地形的精准测量, 都对大面积区域精准覆盖提出了极高要求。面对

日益复杂的覆盖路径规划 (Coverage Path Planning, CPP)^[4-6]问题, 单台无人机在执行任务时逐渐显露出其局限性。以往许多依赖单台无人机执行的任务, 如今正越来越多地被多无人机协同所取代, 从而能更为高效地达成目标。

覆盖路径规划是指在特定区域中, 确保无人机能够完全遍历给定区域。现有的CPP解决方法主要分为经典方法和启发式方法。经典方法中, 来回扫描^[7,8]和螺旋扫描^[9,10]在规则区域或凸地形中覆盖效率较高, 但在非凸区域及动态环境中扫描方向难以灵活调整; 梯形分解^[11]和Boustrophedon (牛耕式) 分解^[12]均属于精确细胞分解方法, 与梯形分解相比Boustrophedon分解能够减少单元数量和覆盖路径长度, 但后者仅考虑障碍物顶点, 容易创建非凸的较大单元; 近似细胞分解技术将目标区域离散化为一组规则的细胞^[13], 能够适应不同的环境模型, 提高处理问题的效率; 简单随机游走算法 (RW)^[14]作为一种随机覆盖算法, 不需要精确的环境模型, 但难以保证高覆盖率和低重复率; 旋转树覆盖算法 (STC)^[15]可高效覆盖不同区域。将单无人机STC算法扩展到多无人机时, 虽然STC算法的重复率较低, 但匝数较高。启发式算法^[16-18]在处理CPP问题中, 通过利用经验法则和直观推断, 在合理的时间内为路径规划提供近似最优的解决方案。强化学习也是一种启发式算法^[19], 它根据特定的奖励机制来调整行为策略, 具有极强的自主性, 无需人工预先设定具体路径, 智能体便能自主探索, 但强化学习训练时间较长会影响实际应用效率。

Dong等人^[20]提出基于深度强化学习的无人机区域覆盖路径规划框架, 验证了无人机可在静态和动态环境下实现完全覆盖。Bayerlein等人^[21]提出了一种多智能体强化学习方法, 该方法能够适应采集任务中场景参数的变化。Zhang等人^[22]针对随机和动态环境中路径规划问题提出基于深度强化学习的方法, 验证了多架无人机能够在多障碍物环境中执行导航任务。Tan等人^[23]提出了一种基于Q学习的完全覆盖路径计划算法, 通过强化学习方法引入了全局环境信息, 优化了原始算法在障碍物附近的路径规划策略。Zhang等^[24]人提出新型奖励函数来改进基于Q-Learning的覆盖路径规划算法, 极大降低了覆盖路径的重复率。

目前, 许多专家认为强化学习可提升覆盖路

径规划性能, 但多无人机使用强化学习会增加计算量和时间成本。采用Q-Learning对两架无人机进行覆盖路径规划虽能完成覆盖任务, 但需要充分考虑无人机之间的协同问题, 并且由于时间和计算资源的限制, 不能期望每架无人机都能找到最优路线。除此之外, 在研究多机协同覆盖问题时, 多是对无人机进行任务划分, 每架无人机负责各自子区域的覆盖导致无人机之间的协同程度不高。

本文针对以上问题, 提出基于Q-Learning覆盖路径规划的三阶段方法。第一阶段提出基于网格的旋转区域划分方法, 网格覆盖区域大小取决于不同角度的分解方式。第二阶段使用Double-Q-Learning对双无人机进行协同覆盖路径规划。最后, 在不同目标区域下, 对Double-Q-Learning算法进行无人机覆盖路径规划的有效性验证, 并与同类算法进行性能对比。

1 覆盖路径规划模型

1.1 无人机运动模型

本文针对覆盖路径规划问题, 采用四旋翼无人机作为研究模型。在二维飞行网格中, 四旋翼无人机的飞行路径存在8种方向, 如图1所示。

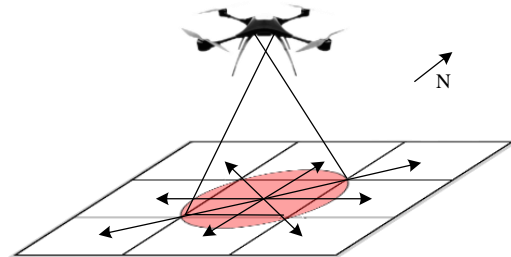


图1 无人机运动方向示意

Fig. 1 Schematic of drone movement direction

为保证多无人机之间的协同性, 就需要无人机在覆盖过程中进行变速飞行, 确保每一架无人机到达下一个路径点的时间是相同的, 以消除个体之间的时间差异, 初步达到协同的目的。无人机数量用 m 表示, 第 i 台无人机被标记为 $U_{UAV,i}$, 在 t 时刻的状态信息如式(1)所示。

$$U_{UAV,i}(t) = \{x(t), y(t), \theta(t) | \theta(t) \in [0, 2\pi]\} \quad (1)$$

式中, $x(t)$ 和 $y(t)$ 表示 $U_{UAV,i}$ 二维平面中 t 时刻的横纵坐标, $\theta(t)$ 表示 $U_{UAV,i}$ 在 t 时刻偏转角。偏转角 θ 的大小为无人机运动方向与水平轴正方向的顺时针

针方向夹角。

当无人机运动到其他位置时, 无人机的状态发生了改变, 则位置信息也要更新, 无人机的运动方程如式 (2) 所示。

$$\begin{cases} x(t+1)=x(t)+\Delta t \cdot V \cdot \cos \theta(t) \\ y(t+1)=y(t)+\Delta t \cdot V \cdot \sin \theta(t) \\ \theta(t+1)=\theta(t)+\Delta \theta \end{cases} \quad (2)$$

式中, Δt 表示单次移动的时间变化量; V 表示无人机的飞行速度, $\Delta \theta$ 表示偏转角的变化量。

1.2 覆盖区域建模

在网格化的目标区域中, 每个网格单元都携带着网格内区域的环境信息, 包括位置、大小、障碍、目标物及被搜索情况等。在本文中, 网格化目标区域主要作用是表示网格内的地区是否为任务区域以及是否已被搜索。

$$S(g_{x,y}, t) = \begin{cases} 0, & g_{x,y} \in Mn \\ 1, & g_{x,y} \in Ms \\ -1, & g_{x,y} \in Mu \end{cases} \quad (3)$$

式(3)中, $S(g_{x,y}, t)$ 表示栅格 $g_{x,y}$ 在 t 时刻的被搜索状态, Mn 表示非搜索区域, Ms 表示已搜索区域, Mu 表示未搜索区域。获得状态的信息后, 网格化的搜索区域可以用二维矩阵 M_{cov} 来表示, 如式 (4) 示。

$$M_{cov} = \begin{bmatrix} S(g_{1,1}, t) & \cdots & S(g_{1,n}, t) \\ \vdots & \ddots & \vdots \\ S(g_{m,1}, t) & \cdots & S(g_{m,n}, t) \end{bmatrix} \quad (4)$$

在搜索过程中, 无人机的飞行高度会影响视窗大小, 视窗是指无人机机载摄像机的有效监视范围。一般情况下, 无人机水平飞行时的视窗形状呈圆形, 视窗示意如图 2 所示。



图 2 无人机视窗示意

Fig. 2 Drone camera shooting range

1.3 约束条件和总代价目标

为了避免重复覆盖, 提高无人机的总体覆盖

效率, 本文中, 覆盖路径规划要求所有目标都被访问到, 每个区域只能访问一次, 并且每个目标只被一架无人机访问一次。多架无人机之间的协同访问约束如式 (5):

$$\begin{cases} P_{path_i} \cap P_{path_j} = \emptyset, \forall i \neq j, 1 \leq i, j \leq m \\ \bigcup_{i=1}^{i=m} P_{path_i} = \{T_j, 1 \leq j \leq n\} \end{cases} \quad (5)$$

式中, P_{path_i} 表示第 i 架无人机的路径; m 为无人机总数; T_j 表示路径点的集合。

覆盖路径规划要求所有无人机只能从当前位置向相邻区域转移, 不能跨区域移动, 式 (6) 为无人机的动作转移约束。

$$\sum_{i \in N_j} \sum_{j \in N^{Adj}(i)} \sum_{k \in N^{Adj}(j)} x_{ijk} = 1, h \in m \quad (6)$$

式中, $N^{Adj}(i)$ 为节点 i 的相邻节点集; $N^{Adj}(j)$ 为节点 j 的相邻节点集; m 为无人机数量; N 为搜索区域的位置节点集合; h 为当前状态的无人机编号; x_{ijk} 表示如果无人机 $U_{UAV,h}$ 以路径 $(i \rightarrow j \rightarrow k)$ 完成了覆盖任务, $x_{ijk} = 1$, 否则 $x_{ijk} = 0$ 。

一般来说, 无人机的飞行目的是在保证安全的情况下尽量减少时间、降低成本, 故而路径规划中的路径长度至关重要。飞行的总距离代价 F_{pc} 可表示为:

$$F_{pc} = \sum_{m=1}^2 \sum_{i=1}^{n-1} \|(x_{i+1}, y_{i+1}) - (x_i, y_i)\|_2 \quad (7)$$

式中, n 为无人机路径航迹点; (x_i, y_i) 表示无人机在第 i 个航迹点的坐标; (x_{i+1}, y_{i+1}) 表示无人机在第 $i+1$ 个航迹点的坐标; m 为无人机总数量。

对于移动机器人的应用, 拥有一条更具吸引力的路径直线 (转弯较少) 是有益的。因此, 式 (8) 为一段路径的平滑度和路径转折代价:

$$\begin{cases} R^s(o_i) = \frac{\angle o_{i-1} o_i o_{i+1}}{180^\circ} \\ F_{sc} = \sum_{m=1}^2 \sum_{i=2}^{n-1} \frac{1}{R^s(o_i)} \end{cases} \quad (8)$$

式中, $R^s(o_i) \in (0, 1]$ 表示路径的平滑程度; n 为无人机路径航迹点; m 为无人机总数量; $\angle o_{i-1} o_i o_{i+1}$ 为第 i 步路径规划的节点位置 o_i 与第 $i-1$ 步路径规划的节点位置 o_{i-1} 和第 $i+1$ 步路径规划的节点位置 o_{i+1} 构成的角度, 且 $\angle o_{i-1} o_i o_{i+1} \in (0, 180^\circ]$; F_{sc} 为路径转折代价; m 为无人机总数; n 为路径航迹点。

综合考虑无人机飞行的距离代价和转折代价, 则无人机二维平面内的覆盖路径规划问题的总目

标 z 可表示为:

$$\begin{cases} z = \min(F_{tc}) \\ F_{tc} = \omega_1 \times F_{pc} + \omega_2 \times F_{sc}, \sum_{i=1}^2 \omega_i = 1, \omega_i \geq 0 \end{cases} \quad (9)$$

式中, F_{tc} 为总代价, $\omega_i(i=1,2)$ 为权重系数。

2 基于网格的旋转区域划分

当搜索区域被分解为正方形网格单元时,

通过移动到正方形网格单元的相邻一个单元来生成覆盖路径。搜索区域的大小会影响无人机完成覆盖搜索任务的时间,通过旋转覆盖网格来最小化待搜索区域的面积。如图3所示,基于网格的分解方法根据分解的角度不同,导致要覆盖的搜索区域也不相同。对此,本文提出一种基于网格的旋转区域划分方法。

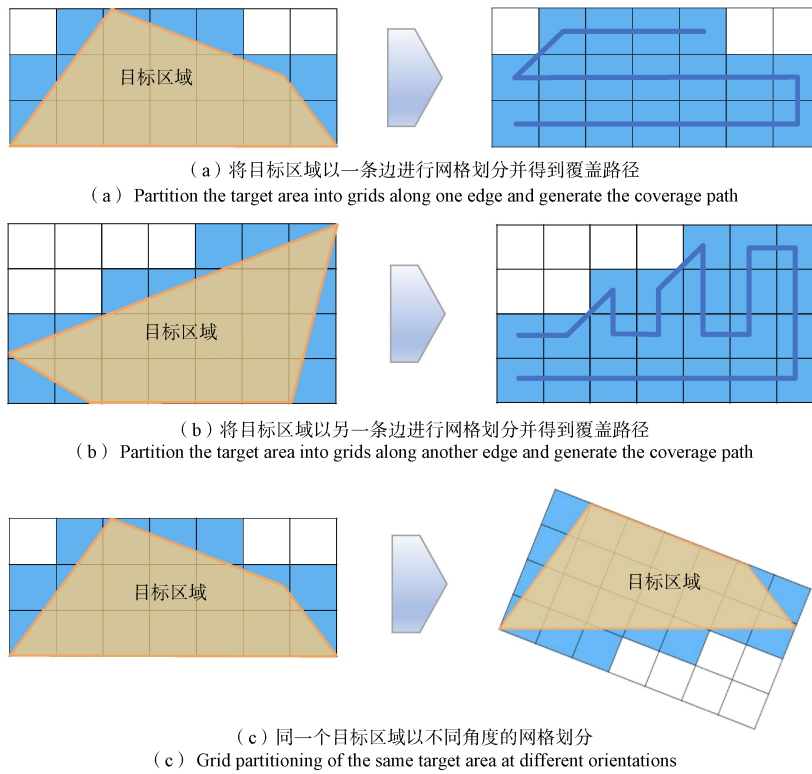


图3 基于网格的旋转区域划分

Fig. 3 Grid based region partition

在二维平面中,用笛卡尔坐标系 OXY 来表示。用矩阵来表示二维平面坐标的旋转:

$$\mathbf{R}(\theta) = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \quad (10)$$

式中, $\mathbf{R}(\theta)$ 为旋转矩阵; $\theta \in [0, \frac{\pi}{2}]$ 为旋转角度。

目标区域各顶点在坐标系 OXY 中的坐标向量为 $\vec{v} = \begin{bmatrix} x \\ y \end{bmatrix}$, 通过旋转坐标系得到新的顶点坐标向量为 $\vec{v}' = \begin{bmatrix} x' \\ y' \end{bmatrix}$, 则有 $\vec{v}' = \mathbf{R}^{-1}(\theta)\vec{v}$ 。

如图4所示,将目标区域划分为规则的网格单元。其中, $x' \in [x'_{min}, x'_{max}]$, $y' \in [y'_{min}, y'_{max}]$ 。将区域

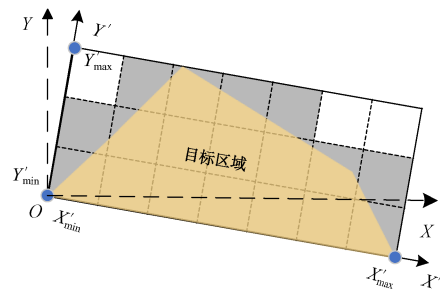


图4 区域分解示意

Fig. 4 Schematic of regional decomposition

空间划分成大小相等的矩形单元:

$$m = \left\lceil \frac{y'_{max} - y'_{min}}{a} \right\rceil \quad (11)$$

$$n = \left\lceil \frac{x'_{\max} - x'_{\min}}{a} \right\rceil \quad (12)$$

式中, m 表示 y 方向划分了 m 个单元; n 表示 x 方向划分了 n 个单元; a 表示相机正方形视窗的边长。

网格单元的边长如式 (13) 和式 (14) 所示, 分别为 Δx 和 Δy :

$$\Delta x = \frac{x'_{\max} - x'_{\min}}{n} \quad (13)$$

$$\Delta y = \frac{y'_{\max} - y'_{\min}}{m} \quad (14)$$

每个网格单元在空间中都有一个唯一的位置标识, 通常可以用其行列号来表示。位于第 i 行、第 j 列的网格单元, 其中心位置在空间中的坐标如式 (15) 所示:

$$S_{ij} = (x_{\min} + (j - 0.5)\Delta x, y_{\min} + (i - 0.5)\Delta y) \quad (15)$$

当前划分的网格中还存在部分未覆盖区域, 需对离散后的网格 N_{ij} 进行覆盖筛选:

$$N_{ij} = \begin{cases} 1, & \text{包含目标区域} \\ 0, & \text{其他} \end{cases} \quad (16)$$

最后, 计算待搜索区域的总面积 Z_k :

$$Z_k = \Delta x \times \Delta y \times \sum_i \sum_j N_{ij} \quad (17)$$

3 基于 Q-Learning 的双无人机覆盖路径规划

3.1 Q-Learning 的基本原理

Q-Learning 是从生物行为主义理论获得启发, 通过智能体 (Agent) 对环境 (Environment) 状态 (State) 的感知, 选择相应的动作 (Action), 记录从环境接受到的回报 (Reward), 如图 5 所示。按照一定规则多次探索后, 建立环境的输入状态与动作之间的关系, 使得智能体再次遇到相同状态时做出更优的动作选择。

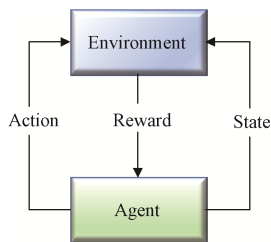


图 5 Q-Learning 模型

Fig. 5 Q-Learning model

Q-Learning 中的无人机可以表示为以下四

元组:

$$U_{UAV}(s, a, r, s') \quad (18)$$

式中, 无人机的当前状态为 s ; a 表示无人机的不同动作; r 为回报奖励; s' 为下一状态。

在学习过程开始时, 以 0 初始化矩阵 Q , 并在迭代学习过程中通过式 (19) 进行更新:

$$Q[s, a] = (1 - \alpha)Q[s, a] + \alpha \left(R(s, a) + \lambda \left(\max_{a'} Q[s', a'] \right) \right) \quad (19)$$

式中, α 为学习率, 取值范围 $0 < \alpha \leq 1$, 为简化算法, 通常取 $\alpha = 1$; λ 为回报的折扣因子, 取值范围 $0 < \lambda \leq 1$ 。 $R(s, a)$ 表示在状态 s 下执行动作 a 所带来的奖励。 $R(s, a)$ 定义为式 (20)。当无人机完成一次覆盖任务时可获得奖励, 即:

$$R(s, a) = \begin{cases} m, & \text{当完成一次覆盖任务} \\ 0, & \text{其他} \end{cases} \quad (20)$$

式中, m 为常数。

3.2 多机协同机制

本文中双无人机协同是全区域、全过程进行的, 实现了完全覆盖和无人机防撞。两架无人机共用一个记录区域访问信息的矩阵, 且都具有采集和更新矩阵数据的权限。为避免发生两架无人机同时被同一侦察价值较高的区域吸引, 引入依次规划的规则。每片区域只能被一架无人机访问一次, 给下一架无人机作为规划参考, 实现两架无人机之间的信息分享和协同覆盖搜索。

每架无人机规划路径时都要避免发生碰撞。由于无人机的动作选择, 无人机之间会发生 2 种情况的碰撞, 分别是节点处碰撞和路径交叉处碰撞。

由于 2 架无人机在同一时间选择了相同的路径点而产生碰撞, 其特点可以表达为式 (21):

$$\begin{cases} x_i(t+1) = x_j(t+1) \\ y_i(t+1) = y_j(t+1) \end{cases} \quad i, j = 1, 2, \dots, N \quad (21)$$

式中, $x_i(t+1)$, $y_i(t+1)$ 为第 i 架无人机在 $(t+1)$ 时刻的位置, N 为无人机总数量。

由于 2 条路径产生了交叉, 导致无人机碰撞。无人机为了达到协同效果, 在执行全覆盖路径任务时均采用可变速飞行, 在飞行过程中速率相同。一旦 2 架无人机路径产生交叉, 必然会发生碰撞, 其特点可以表达为式 (22):

$$\begin{cases} x_i(t)+x_i(t+1)=x_j(t)+x_j(t+1) \\ y_i(t)+y_i(t+1)=y_j(t)+y_j(t+1) \end{cases} \quad i,j=1,2,\dots,N \quad (22)$$

式中, $x_i(t)$, $y_i(t)$ 为第*i*架无人机在*t*时刻的位置;
 $x_i(t+1)$, $y_i(t+1)$ 为第*i*架无人机在*t+1*时刻的位置;
*N*为无人机总数量。

3.3 Double-Q-Learning 双无人机覆盖路径规划

使用Q学习进行双无人机覆盖路径规划时, 需要充分考虑无人机之间的协同问题, 并且不能期望每架无人机都能找到最短路线。Q学习虽然可以进行一定程度的探索, 但由于时间和计算资源的限制, 很难遍历所有可能的路径。为了提高双无人机覆盖路径规划的性能, 本文基于提出Double-Q-Learning 双无人机覆盖路径规划算法。

用Q-Learning进行双无人机覆盖路径规划是从

全局的角度出发, 对整个任务区域进行双无人机的任务分配。一次Q学习的侧重点在于考虑整体任务分配, 将总路径长度*D*作为其评价指标。

$$D = \sum_{m=1}^2 \sum_{i=1}^{n-1} \|(x_{i+1}, y_{i+1}) - (x_i, y_i)\|_2 \quad (23)$$

式中, *n*为无人机路径航迹点; (x_i, y_i) 表示无人机在第*i*个航迹点的坐标; (x_{i+1}, y_{i+1}) 表示无人机在第*i+1*个航迹点的坐标; *m*为无人机总数量。

二次Q-Learning从局部的角度出发, 可以对全局规划进行进一步优化。每架无人机在执行任务时, 可以根据自身所处的局部环境和实时状态, 进行局部的路径调整和决策, 能够快速收敛到最优的路径策略, 减少两架无人机在路径规划过程中的搜索时间和计算资源消耗, 提高任务执行效率, 具体流程如图6所示。

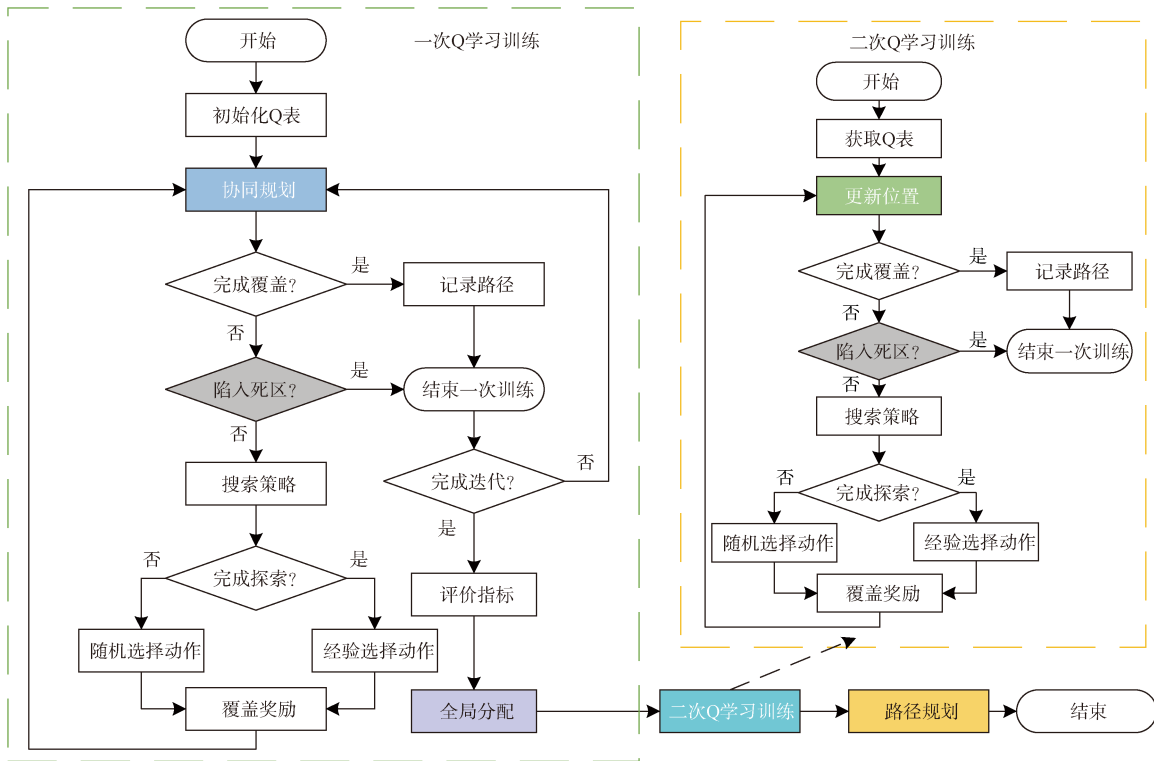


图6 Double-Q-Learning的覆盖路径规划流程图

Fig. 6 Flow chart of coverage path planning for Double-Q-Learning

4 仿真算例

为了验证模型的合理性和所提出的基于Q-Learning的双无人机覆盖路径规划算法的有效性, 对两架无人机进行了多组仿真实验。仿真算例的模型参数的设定如表1所示。

4.1 基于网格的旋转区域划分结果

随机生成多组初始目标区域, 通过基于网格的旋转区域划分可以得到不同大小的搜索区域。从表2可以看出: 通过旋转坐标系可以得到初始目标区域的不同大小的网格搜索范围。优先选取搜索面积最小的区域, 可以有效减少无人机覆盖路

表 1 模型参数

Table 1 Model parameters

状态参数	参数设定	数值
d	视窗直径	20 m
n	最大迭代次数	10 000
ω_1	距离代价系数	0.3
ω_2	转折代价系数	0.7
λ	折扣因子	0.8
α	学习率	1
m	回报奖励	10

表 2 基于网格的搜索区域划分对比

Table 2 Comparison of grid based search area partition

目标区域	最大搜索面积 (m ²)	最小搜索面积 (m ²)	差值(m ²)
a	20 937.928 3	17 606.894 3	3 331.034 0
b	23 260.527 9	18 311.479 4	4 949.048 5
c	30 130.862 7	14 810.085 1	15 320.777 6
d	26 061.568 1	23 218.488 0	2 843.080 1
e	27 535.317 0	22 028.253 6	5 507.063 4
f	28 903.369 8	20 930.026 4	7 973.343 4

径规划的时间和能耗, 提高覆盖路径质量。

4.2 双无人机覆盖路径规划结果

针对搜索区域, 采取 2 架无人机进行覆盖搜索任务。考虑到 2 架无人机之间距离太近会对其决策

产生影响, 为消除无人机之间的初始干涉, 让无人机均从相同的初始位置出发, 分别到达距离搜索区域最近的地点。图 7 是用 Double-Q-Learning 进行双无人机覆盖路径规划结果。

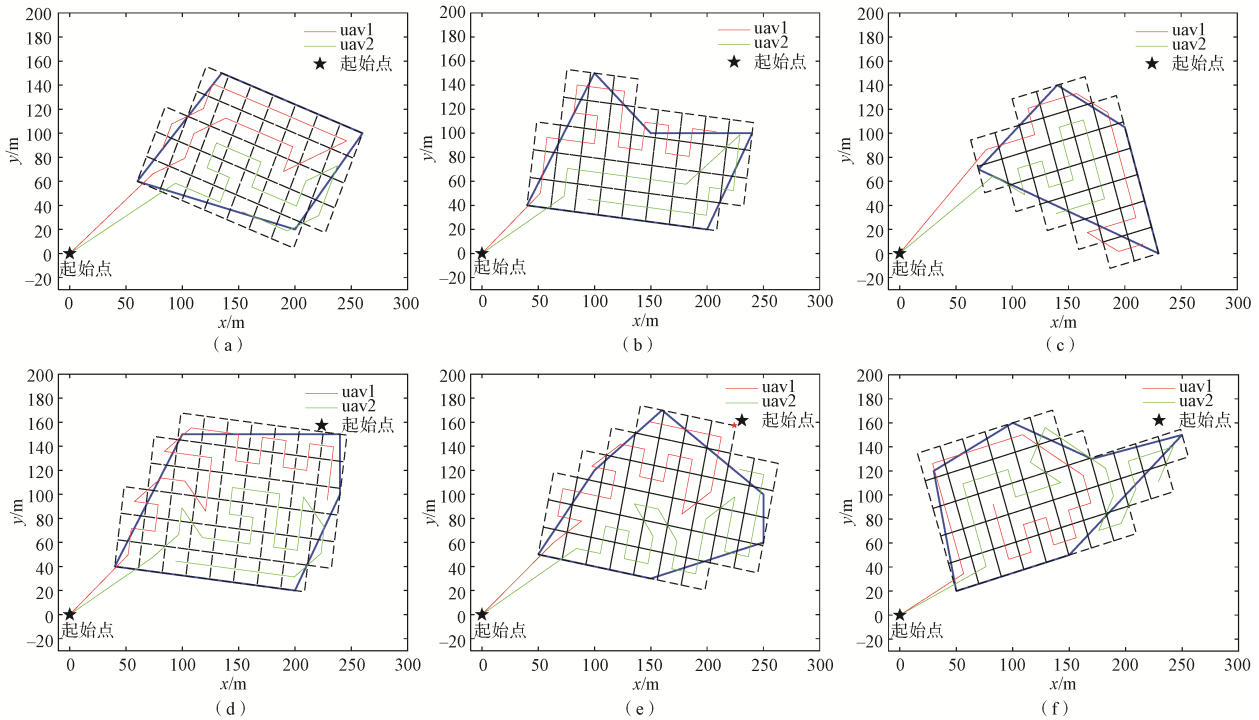


图 7 基于 Double-Q-Learning 的双无人机覆盖路径规划结果

Fig. 7 Results of dual drone coverage path planning based on Double-Q-Learning

4.3 算法对比

为对比不同算法在不同形状区域中的性能表现, 本文选取了具有典型形状特征的两类区域 (凸区域和凹区域), 作为对比算法的初始化环境模型。图 8 是三种算法分别在两种区域下覆盖路径规划结果。根据表 3 可得, 在凸区域中, 所提出的算法对比上述两种算法, 解决方案的总代价目标 Z 减少了 5.46% 和 14.31%, 单架最大代价 Z_{\max} 平均减少了 5.11% 和 21.28%。在凹区域中, 所提出的算

法对比上述两种算法, 解决方案的总代价目标 Z 减少了 5.28% 和 68.87%, 单架最大代价 Z_{\max} 对比上述两种算法减少了 7.2% 和 82.23%。

Q-Learning 算法虽然可以进行一定程度的探索, 但由于时间和计算资源的限制, 很难遍历所有可能的路径。BCD (分支切割与下潜) 算法通常依赖于预先定义的空间划分规则和固定的策略来生成路径, 很难根据环境的实时变化和任务执行过程中的反馈来灵活调整路径。采用 Double-Q-

Learning的路径规划解可以得到更小的总代价。原因在于该算法对全局分配结果加入了二次Q-Learning的路径重规划算法，既避免了Q-Learning

算法缺乏局部优化的问题，又解决了BCD算法过于依赖模型的情况，实现了路径寻优中全局分配与局部优化能力的均衡。

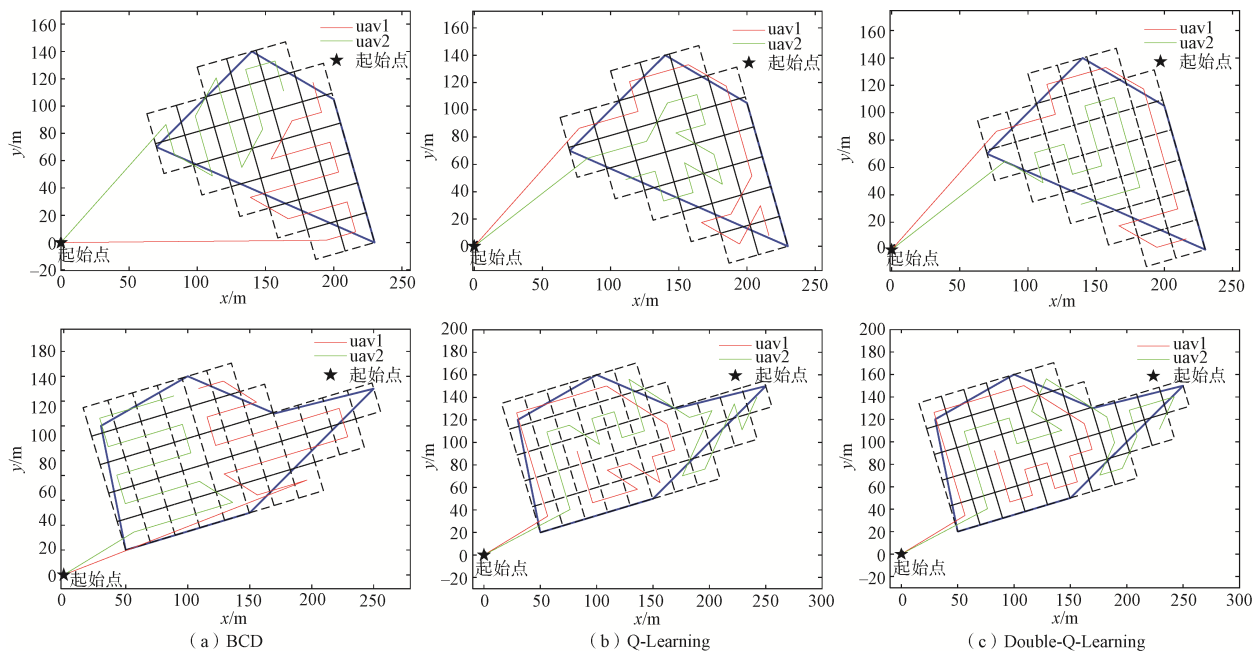


图8 三种算法下覆盖路径规划结果

Fig. 8 Coverage path planning results under three algorithms

表3 三种算法下覆盖路径规划的性能比较

Table 3 Performance comparison of coverage path planning under three algorithms

算法	区域	总代价	平均代价	单架最大代价
Double-Q-Learning	凸区域	289.183 5	144.591 8	149.994 9
	凹区域	474.103 1	237.052 0	242.558 4
Q-Learning	凸区域	305.871 3	152.935 7	158.079 5
	凹区域	500.552 6	250.276 3	261.462 3
BCD	凸区域	337.468 3	168.734 2	190.550 2
	凹区域	1 522.884 7	761.442 4	1 356.821 2

5 结束语

覆盖路径规划旨在让无人机实现对目标区域的完全覆盖。为降低覆盖路径的总代价函数，完成双无人机协同覆盖搜索，本文提出一种基于Q-Learning的双无人机覆盖路径规划算法。采用基于网格的旋转区域划分算法，通过减少待搜索区域的覆盖面积，节省无人机完成覆盖搜索任务的时间。引入Double-Q-Learning算法，将覆盖路径规划任务分为全局任务分配和局部路径重规划两部分。实现了路径寻优中全局分配与局部优化能力的均衡，提升了算法的寻优精度和寻优效率。通

通过对两架无人机进行多组仿真实验和现有的两种算法进行对比可以得出：Double-Q-Learning算法不仅能够适应不同大小和特点的搜索区域，还能以更低的总代价构建双无人机完成覆盖任务的最优路径。

本文为了降低计算复杂度，加快算法训练速度，对双无人机覆盖路径规划中的环境设置较为简单。然而，在实际应用中，两架及以上的多无人机在执行完全覆盖任务过程中涉及的环境可能性很多，需要对环境问题进行综合考虑。因此，下一步的工作是扩展算法。例如，可以使用该算法在多障碍场景和大规模场景等复杂环境条件下实现多无人机的覆盖路径规划。

参考文献

- [1] WAHARTE S, TRIGONI N. Supporting search and rescue operations with UAVs[C]//2010 International Conference on Emerging Security Technologies. New York: IEEE, 2010: 142-147.
- [2] MAZA I, CABALLERO F, CAPITÁN J, et al. Experimental results in multi-UAV coordination for disaster management and civil security applications[J]. Journal of Intelligent & Robotic Systems: with a special section on Unmanned Systems, 2011, 61(1-4): 563-585.
- [3] CESETTI A, FRONTONI E, MANCINI A, et al. A visual global positioning system for unmanned aerial vehicles used in photogrammetric applications[J]. Journal of Intelligent & Robotic Systems: With a special section on Unmanned Systems, 2011, 61(1-4): 157-168.
- [4] SKOROBOGATOV G, BARRADO C, SALAMÍ E. Multiple UAV systems: A survey[J]. Unmanned Systems, 2020, 8(2): 149-169.
- [5] TAN C S, MOHD-MOKHTAR R, ARSHAD M R. A comprehensive review of coverage path planning in robotics using classical and heuristic algorithms[J]. IEEE Access, 2021, 9: 119310-119342.
- [6] HUANG X, SUN M, ZHOU H, et al. A multi-robot coverage path planning algorithm for the environment with multiple land cover types[J]. IEEE Access, 2020, 8: 198101-198117.
- [7] HUANG W H. Optimal line-sweep-based decompositions for coverage algorithms[C]//Proceedings of the 2001 IEEE International Conference on Robotics and Automation (ICRA). New York: IEEE, 2001: 27-32.
- [8] SKOROBOGATOV G, BARRADO C, SALAMÍ E, et al. Flight planning in multi-unmanned aerial vehicle systems: Nonconvex polygon area decomposition and trajectory assignment[J]. International Journal of Advanced Robotic Systems, 2021, 18(1): 1729881421989551.
- [9] ACEVEDO J J, ARRUE B C, MAZA I, et al. Cooperative large area surveillance with a team of aerial mobile robots for long endurance missions[J]. Journal of Intelligent & Robotic Systems, 2013, 70: 329-345.
- [10] BALAMPANIS F, MAZA I, OLLERO A. Area decomposition partition and coverage with multiple remotely piloted aircraft systems operating in coastal regions[C]//2016 International Conference on Unmanned Aircraft Systems (ICUAS). New York: IEEE, 2016: 275-283.
- [11] OKSANEN T, VISALA A. Coverage path planning algorithms for agricultural field machines[J]. Journal of Field Robotics, 2009, 26: 651-668.
- [12] COOMBES M, CHEN W H, LIU C. Boustrophedon coverage path planning for UAV aerial surveys in wind [C]//Proceedings of the 2017 International Conference on Unmanned Aircraft Systems (ICUAS). New York: IEEE, 2017: 1563-1571.
- [13] CHO S W, PARK H J, LEE H, et al. Coverage path planning for multiple unmanned aerial vehicles in maritime search and rescue operations[J]. Computers & Industrial Engineering, 2021, 161: 107680.
- [14] BARTUMEUS F, DA LUZ M G E, VISWANATHAN G M, et al. Animal search strategies: A quantitative random-walk[J]. Ecology, 2005, 86(12): 3078-3087.
- [15] ARRIBAS E, MANCUSO V, CHOLVI V. Coverage optimization with a dynamic network of drone relays[J]. IEEE Transactions on Mobile Computing, 2020, 19(10): 2278-2298.
- [16] KOENIG S, LIU Y. Terrain coverage with ant robots: A simulation study[C]//Proceedings of the Fifth International Conference on Autonomous Agents. New York: ACM, 2001: 600-607.
- [17] SADEK M G, MOHAMED A E, EL-GARHY A M. Augmenting multiobjective genetic algorithm and dynamic programming for online coverage path planning [C]//Proceedings of the 13th International Conference on Computational Engineering Systems (ICCES). New York: IEEE, 2018: 475-480.
- [18] KENNEDY J. The particle swarm: Social adaptation of knowledge[C]//Proceedings of 1997 IEEE International Conference on Evolutionary Computation, April 13-16, 1997, Indianapolis, USA. New York: IEEE, 1997: 303-308.
- [19] JIA J, WANG W. Review of reinforcement learning research[C]//2020 35th Youth Academic Annual Conference of Chinese Association of Automation (YAC). New York: IEEE, 2020: 186-191.
- [20] DONG J X. Area coverage path planning of UAV based on deep reinforcement learning[J]. Industrial Control and Computing, 2021, 34(2): 80-82.
- [21] BAIERLEIN H, THEILE M, CACCAMO M, et al. Multi-UAV path planning for wireless data harvesting with deep reinforcement learning[J]. IEEE Open Journal of the Communications Society, 2021, 2: 1171-1187.
- [22] ZHANG S T, LI Y B, DONG Q H. Autonomous navigation of UAV in multi-obstacle environments based on a deep reinforcement learning approach[J]. Applied Soft Computing, 2022, 115: 108194.
- [23] TAN X Q, HAN L H, GONG H, et al. Biologically inspired complete coverage path planning algorithm based on Q-learning[J]. SENSORS, 2023, 23(10): 4647.
- [24] ZHANG M Y, CAI W Y, PANG L F. Predator-prey reward based Q-learning coverage path planning for mobile robot[J]. IEEE Access, 2023, 11: 29673-29683.

[作者简介]

- 陈佳雨 2001年生, 硕士研究生。
李文 2000年生, 本科。
李泰融 1998年生, 本科。
李志茹 2001年生, 硕士研究生。
王子怡 2001年生, 硕士研究生。
陈鹏云 1987年生, 博士, 副教授。

(本文编辑: 傅杰)