

基于 MDP 的无人机避撞航迹规划研究

阚 煌¹, 辛长范¹, 谭哲卿¹, 高 鑫¹, 史铭姝¹, 张 谦²

(1. 中北大学 机电工程学院, 太原 030051; 2. 滨州魏桥国科高等技术研究院, 山东 滨州 256606)

摘要: 无人机 (UAV) 进行避撞前提下的目标搜索航迹规划是指在复杂且众多的环境障碍约束中通过合理规划飞行路径, 以更快、更高效的形式找到目标; 研究了无障碍环境条件下有限位置马尔科夫移动的规律, 构建了相应的马尔科夫移动分布模型; 在借鉴搜索系统航迹规划的前沿研究成果之上, 结合马尔科夫决策过程理论 (MDP), 引入了负奖励机制对 Q-Learning 策略算法迭代; 类比“风险井”的可视化方式将障碍威胁区域对无人机的负奖励作用直观地呈现出来, 构建了复杂障碍约束环境下单无人机目标搜索航迹规划模型, 并进行仿真实验证明该算法可行, 对航迹规划算法的设计具有一定的参考意义。

关键词: 无人机; 航迹规划; 避撞; 静态目标搜索; 马尔科夫决策过程 (MDP); 风险井

Research on UAV Collision Avoidance Path Planning Based on MDP

KAN Huang¹, XIN Changfan¹, TAN Zheqing¹, GAO Xin¹, SHI Mingshan¹, ZHANG Qian²

(1. College of Mechatronic Engineering, North University of China, Taiyuan 030051, China;

2. Binzhou Institute of Technology, Weiqiao-UCAS Science and Technology Park, Binzhou 256606, China)

Abstract: Target search path planning of UAV on the premise of collision avoidance is to find the target in the faster and more efficient form by reasonable flight path planning in the complex and numerous environmental obstacles. Discussed the law of finite position Markov movement under barrier-free environment and constructed the corresponding Markov movement distribution model. Based on the cutting-edge research results of search system trajectory planning, combined with the MDP theory, the negative reward mechanism was innovatively introduced to iterate the Q-Learning strategy algorithm, and the single UAV target search model was constructed. By analogy with the “risk well” visualization method, the negative reward effect of the obstacle threat area on the UAV was intuitively presented, and the single UAV target search path planning model under complex obstacle constraint environment was constructed. and the simulation experiment showed that the algorithm is feasible, Which has certain reference significance for the design of the route planning algorithm.

Keywords: UAV; path planning; collision avoidance; static target search; MDP (Markov decision process); risk well

0 引言

无人机 (UAV, unmanned aerial vehicle) 是一种区别于有人驾驶的飞行器, 其依靠驱动电机高速转动作为动力来源, 通过机载的决策控制算法程序与地面基站的控制指令相互配合进行协同操控^[1]。近年来, 随着互联网科技、航空航天技术和无线通信等技术的迅猛式发展, 无人机领域的研究在国内外日趋火热。不仅在军事战争领域, 无人机技术得到了广泛应用; 而且在日常生活中, 其应用前景也日益扩大^[2]。可以预见未来行业将涌现更多人员致力于无人机技术开发工作当中^[3]。其中, 无人机的自主避撞飞行决策是研究内容的核心。同时, 作为航迹规划的重要一环, 避撞算法研究在实现自主飞行策略中发挥不可忽视的作用。它直接影响到无人机能否完成任务的成功率和效率。因此, 对进行避撞飞行的航迹规划算法进行深入研究, 对其实现自主决策飞行具有深远的意义。

对于无人机的开发研究, 国内外的研究方向基本一致。

集中探索无人机根据搜索如何进行复杂环境下进行避撞飞行的航迹规划。然而, 由于当前研究水平的限制, UAV 在独立执行搜索任务方面尚显薄弱, 其智能化程度往往未能达到人们的预期, 难以完全胜任各种类型任务要求。在实际应用中, 无人机大多基于预先存储的程序来执行目标搜索任务, 这导致其在应对突发障碍时的灵活性不足, 难以做到及时规避。而针对航迹规划过程中的避撞飞行策略, 优化制导算法作为一种常用且见效最快的抉择。2000 年, Szczerba 提出了一种稀疏 A* 算法, 该算法结合无人机的机动性能约束, 成功地缩减了搜索空间, 提高了算法效率, 但是算法只能进行平面环境下的搜索, 无法适应实际需求^[3]。文献 [4] 为了将 A* 算法扩展到三维空间, 在稀疏 A* 搜索算法的算法研究基础上提出了一种无人机主动改变步长的航迹规划策略。结合到战场环境中侦测的威胁信息来在航迹规划过程中主搜索步长在一定程度上缩短了搜索时间, 但是它只允许在二维平面空间进行航迹规划, 策略

收稿日期: 2023-05-24; 修回日期: 2024-05-26。

基金项目: 教育部产学研合作协同育人项目 (231106429103427)。

作者简介: 阚 煌 (1997-), 男, 硕士研究生。

引用格式: 阚 煌, 辛长范, 谭哲卿, 等. 基于 MDP 的无人机避撞航迹规划研究[J]. 计算机测量与控制, 2024, 32(6): 292-298.

算法无法适用于立体空间。文献 [5] 以现有 SAS 算法和目标分配算法为依据, 细致描述了航迹规划策略的算法原理, 解析了多机航迹规划算法的实现途径。但是在策略算法的设计过程中, 只简单讨论部分约束条件, 其在精确度方面却略显不足。文献 [6] 针对单机搜索目标航迹规划问题进行了探讨, 提出了一种新型搜索算法。无人机在每个搜索阶段, 基于当前探测区域信息预测目标未来可能轨迹, 并通过连续系列的决策过程, 最终实现通往目标的航迹规划。但是该算法没有考虑复杂障碍环境对航迹规划的干扰。

分析讨论了前人的搜索策略算法中存在诸多限制和考虑不够全面的情况, 作为本文算法的创新点和优化点。本文以静态目标搜索为基础, 基于有限状态马尔科夫概率分布对无人机移动区域搜索进行了仿真分析。同时对策略算法思路进行突破, 提出一种新型的避撞航迹规划策略。不仅实现了对实际复杂搜索环境进行障碍避撞完成搜索任务, 而且该算法具备较高的改进优化潜力, 以更适应多样化的目标搜索需求。

1 避撞航迹规划算法结构与原理

本文以静态目标搜索为基础, 基于有限状态马尔科夫概率型运动规律对无人机移动搜索过程进行了仿真分析。同时策略算法思路进行突破, 引入了负奖励机制, 并使用“风险井”机制对障碍物和约束条件进行近似类比建模。结合强化学习 Q-Learning 算法改进传统的策略迭代算法^[7], 提出一种全新的搜索避撞航迹规划算法。整个系统的结构如图 1 所示。

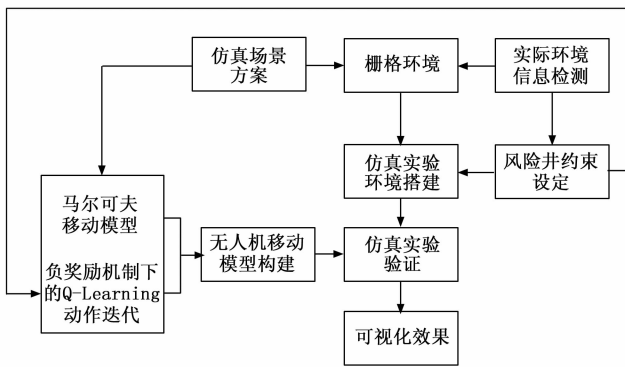


图 1 避撞航迹规划算法系统结构图

该算法不仅对无人机在目标搜索过程中无障碍区域自身运动规律进行深入研究; 同时也考虑到多样且复杂化障碍和约束对无人机搜索过程中的干扰, 进行避撞移动处理。运用了负奖励机制, 并使用“风险井”对障碍物和约束条件进行建模, 设法引入到无人机搜索运动的全过程中。将负奖励引入强化学习 Q-Learning 算法来改进传统的策略迭代算法, 避免了以前目标搜索航迹规划可能陷入局部最优解的局限性。该算法考虑到复杂环境条件下无障碍和有障碍约束条件下的目标搜索对无人机移动产生的影响, 分析了无人机自身的运动规律。考虑到诸多因素对目标搜索全

过程的作用, 该算法进行全面且深入地剖析, 最终可以满足多样化的目标搜索任务需求。

2 马尔科夫移动搜索模型分析与建立

2.1 马尔科夫链及转移概率矩阵

2.1.1 马尔科夫链

马尔科夫链 (Markov Chain) 是概率论和数理统计中的一类随机过程。经过近百年的不断发展, 马尔科夫链模型的相关理论在现代物流、经济学和医学研究等多个学科领域中得到了广泛应用^[8-10]。该特点表现为, 在确定当前状态之后, 以往的状态将不再对未来状态的预测产生影响。这些过程其实都可以通过给定数学公式来量化表达并且计算。

假设在 $t=0$ 时刻的随机变量 X_0 遵循概率分布 $P(X_0) = 0$, 我们称作初始状态分布, 在特定时间点 $t \geq 1$, X_t 与前时间点的 X_{t-1} 之间存在一种特定的条件分布 $P(X_t | X_{t-1})$ 。若 X_t 仅与 X_{t-1} 相关联, 而与之前所有时间点的随机变量 $\{X_0, X_1, X_2, \dots, X_{t-2}\}$ 无关, 这种特性被定义为马尔科夫性。此性质可用式 (1) 表示:

$$P(X_t | X_0, X_1, X_2, \dots, X_{t-1}) = P(X_t | X_{t-1}), t = 1, 2, \dots \quad (1)$$

把具备这样特性的数列命名马尔科夫链, 又称作马尔科夫过程 (Markov Process), 把条件概率分布 $P(X_t | X_{t-1})$ 称为马尔科夫链的转移概率, 转移概率直接决定了马尔科夫链的特性。

2.1.2 转移概率矩阵

转移概率矩阵是一个具有特殊性质的矩阵, 在该矩阵中, 所有元素的值都具备非负性, 同时满足每一行的元素之和恒为 1。这些元素代表了各种可能性转移概率, 所以得名为转移概率矩阵。

针对一个马尔科夫链 $\{X_t, t=1, 2, \dots\}$ 而言, 用 $P(m) = P_{ij}(m)$ 代表其 m 步转移矩阵中任意元素。特别地, 当 $m=1$ 时, $P(1) = P$, 代表其一步转移概率矩阵。

为了更直观地展示转移概率矩阵的求解过程, 如下进行例子展示。假定随机系统的状态空间有 4 个元素, 即 $S = \{1, 2, 3, 4\}$, 一直进行观测, 其表现为如下的连续状态:

4 3 2 1 4 3 1 1 2 3
1 3 1 1 2 3 3 4 1 1
2 1 2 2 3 4 2 3 2 1
4 3 2 4 1 3 2 4 3 1

依据此模型来估算出状态之间的转移概率 P_{ij} 。基于上述采集的数据样本, 可以清晰地呈现出各类转移的发生频次 n_{ij} , 并将其分类整理后列入表 1 中以供进一步分析。

表 1 状态 i 向状态 j 的转移次数

状态	1	2	3	4	n_i
1	3	4	2	1	10
2	3	1	4	2	10
3	4	4	1	2	11
4	2	1	4	0	7

n_i 是该系统从状态 i 移到其他状态的次数, n_{ij} 是由状态 i 到状态 j 转移次数, 则 P_{ij} 可以借由 $P_{ij} = n_{ij} / n_i$ 计算。计算结果如表 2 所示。

表 2 转移概率

转移概率 P_{ij}	1	2	3	4
1	3/10	4/10	2/10	1/10
2	3/10	1/10	4/10	2/10
3	4/11	4/11	1/11	2/11
4	2/7	1/7	4/7	0

由此可得出状态转移矩阵 P_{ij} 如下所示:

$$P_{ij} = \begin{bmatrix} \frac{3}{10} & \frac{2}{5} & \frac{1}{5} & \frac{1}{10} \\ \frac{3}{10} & \frac{1}{10} & \frac{2}{5} & \frac{1}{5} \\ \frac{4}{11} & \frac{4}{11} & \frac{1}{11} & \frac{2}{11} \\ \frac{2}{7} & \frac{1}{7} & \frac{4}{7} & 0 \end{bmatrix}$$

2.2 栅格地图

栅格地图, 即将环境细分成若干个栅格的一种地图构造, 每一个栅格均被赋予一个特定的数值, 此数值用以表示该栅格内是否存在障碍物。作为一种相对简明且稳健的无人机二维环境建模手段, 栅格图法广受青睐。目前, 在机器人路径规划和无人机航迹规划领域, 占据栅格地图已成为备受推崇的环境建模方式之一^[11]。栅格法要求对搜索区域进行合理的栅格划分成多个独立的单元格。每个单元格都代表着无人机的一个搜索步长且满足最小搜索航迹长度大于设定单元格边长条件^[12]。随后为单元格依次对应分配序号, 并借助一系列序号来呈现无人机飞行航迹。最后将序号列对应变换成实际坐标序列, 进而在地图清楚地表现轨迹^[13]。

2.3 马尔科夫移动模型及仿真

假设, 一架无人机在无阻碍区域内以一定的转移概率进行马尔科夫移动。已知转移概率是恒定的, 不随时间而改变, 且整个运动过程中不存在后效性。这一点在相关文献 [14-16] 有详细阐述。同时, 也获知了无人机在区域内的初始投放位置。基于该模型, 成功地模拟无人机的搜索范围, 得到了任意时刻下的位置概率分布^[17-19]。

如果无人机在 t 时刻位于序号 5 即中心的位置, 如图 2 (a) 所示。则在 $t+1$ 时刻有概率移动到与之相邻的 1, 2, 3, 4, 6, 7, 8, 9 的位置 (分别对应着: 西北、正北、正北、正西, 正中、正东、西南 正南和东南共 8 个方向), 又或者可能停在 5 的初始位置, 如图 2 (b) 所示。任意栅格都与一个 3×3 的转移矩阵相对应。这个转移矩阵精确地界定了无人机在接下来的一刻可能移动到的位置及其相应的概率。而无人机的整个运动轨迹, 便是由这些转移概率矩阵的集合所共同构成的。

假设 $t+1$ 时刻位于序号为 1, 2, 3, 4, 5, 6, 7, 8, 9 的概率分别为 $P_{51}, P_{52}, P_{53}, P_{54}, P_{55}, P_{56}, P_{57}, P_{58}$,

P_{59} , 同时满足概率之和始终为 1 这一条件。如图 2 (c) 所示一个转移概率与之对应一个栅格。在本次模拟实验中, 已知向各个方向的转移概率, 如图 2 (d) 所示。 P_{55} 值为 0, 意味着马尔科夫移动目标将不会在原地停留, 必须选择周围 8 个相邻间隔分布方向之一移动。

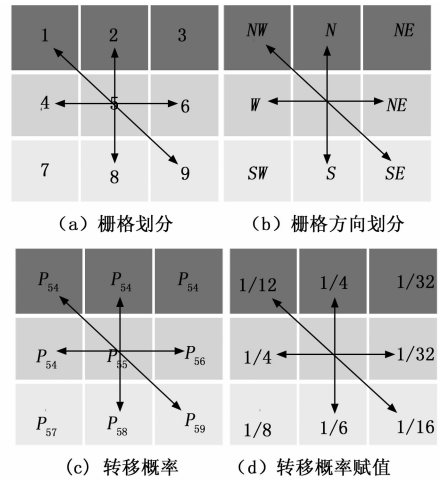


图 2 进行栅格化前后目标的位置

在研究无障碍条件下无人机移动规律的仿真实验中, 采用了 20×20 的栅格来适当地划分无人机运动区域。设定无人机的起始位置为该栅格化环境的中心。同时, 以图 3 (d) 数据作为西北、正北、正北、正西, 正中、正东、西南 正南和东南共 8 个方向的转移概率。本次仿真实验采用了某仿真软件 2022b 版本作为实验平台。通过该平台, 首先给出马尔科夫移动的初始位置所在, 接着进一步模拟目标在经过一步、十步和二十步移动后的位置概率分布。图 4 (a) 展示了初始位置概率分布。而图 3 (b) 则揭示了在一步转移后的位置概率分布。图 3 (c) 描绘了在十步转移后的位置存在概率分布, 图 3 (d) 则呈现了在二十步转移后的位置概率分布。从图中可以明显看出, 随着马尔科夫移动步数的递增, 其在运动区域内的位置概率分布逐渐变得更为分散。这意味着无人机可能出现的地点范围在不断扩大, 极大地提高了搜索工作的复杂性和困难系数。

3 基于“风险井”原理的负奖励机制

3.1 无人机限制约束模型

3.1.1 雷达威胁

目前可用的雷达监测设备具备探测无人机在空中的位置和相应的速度信息的能力, 一旦无人机进入雷达设定的禁飞区域, 有可能激活地面的防御系统, 对违规无人机实施打击。雷达在执行任务时, 会进行 360° 的扫描。设定无人机与雷达的探测装置之间的间距为 d_R , 其最大有效探测距离为 R_{max} 。雷达探测范围如图 4 所示。

3.1.2 防空火力威胁

地面防空火力系统配备了可灵活转动的炮台, 使其能够在一定的射程内对防空目标实施精确的攻击。防空火力的有效威胁范围如图 5 所示。

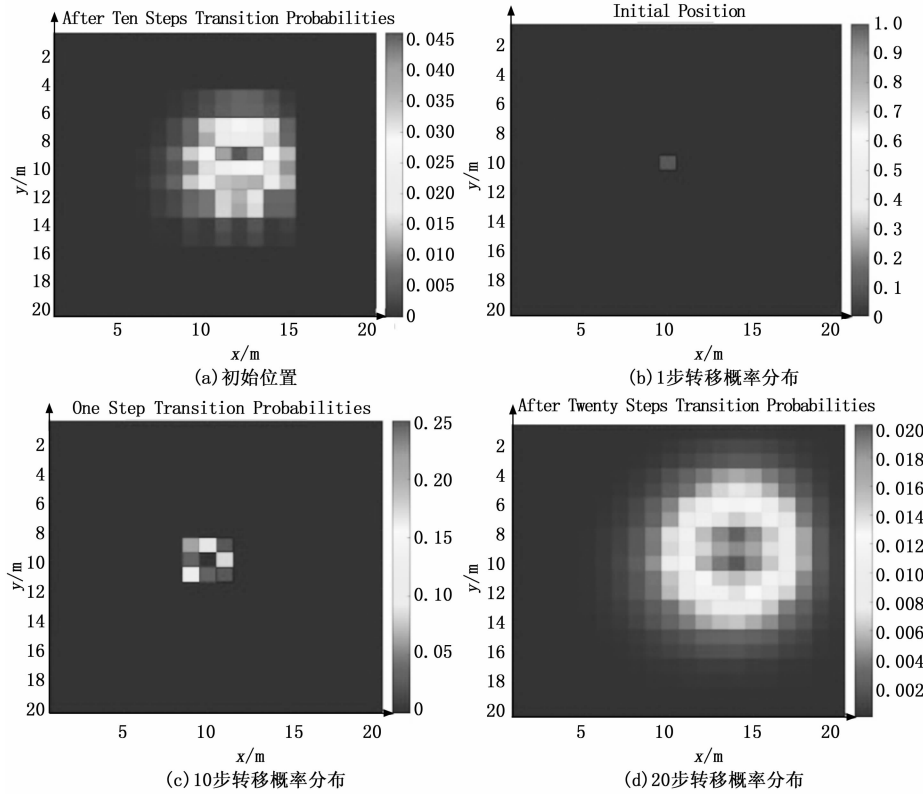


图 3 马尔科夫移动概率分布

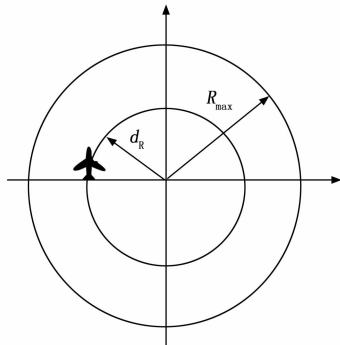


图 4 雷达威胁模型

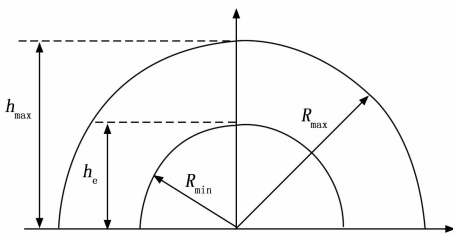


图 5 防空火力威胁模型

图 5 中, h_{\max} 代表的是地面防空火力可达到的最大攻击高度。然而, 受到地面防空火力自身特性的限制, 其实际有效的攻击高度被界定为 h_e 。同时, R_{\max} 和 R_{\min} 分别代表了地面防空火力的最大有效射程和最小有效射程, 从而定义了其有效的射击范围。当无人机在有效射击高度 h 之外、最大射击高度 h_{\max} 之内, 并且在有效的射程 R_{\max} 和 R_{\min} 之间

飞行时, 地面防空火力能够成功地击毁目标。但值得注意的是, 由于地面防空火力的固有特性, 当目标飞行在最大射击高度 h_{\max} 附近或有效射击高度 h 的边缘区域外时, 其摧毁目标的能力会显著降低。

3.1.3 大气威胁

由于空中环境多变且复杂, 雷雨、台风等不稳定天气的频繁出现, 这些天气变化会直接影响空中大气压强以及云层厚度等状况。导致无人机在飞行过程中受到空气或水滴云摩擦而发生颠簸, 在严重情况下可能造成坠机的结果。大气威胁是导致无人机失灵或者坠机结果最常见的原因之一, 直接影响到无人机执行任务的成功率。

3.2 负奖励机制

基于以及 Bertram 和 Wei 等人^[19-20]研究, 尝试在 MDP 奖励函数 R 中引入负奖励对状态值函数的影响, 虽然无法找到精确计算负奖励的方法。但可以用另一种方法将负奖励以某种近似形式进行表示。基于近似

形式在合理且适当的情况下模拟负奖励的作用, 提出了一种基于精确和无记忆的算法, 可以很好地用来计算价值函数。

在状态空间映射到基于 MDP 的底层度量空间中, 例如对于无人机在二维平面环境的航迹规划中, 负奖励的表现与正奖励有很大区别。负奖励主要作用于一个单一的状态。在此状态下, 所设定的负奖励能够以简单绕过它的形式来避免取得所设定的负奖励。这样可以通过类比障碍避撞的直观形式实现负奖励的近似转换。

3.2.1 负奖励近似转换

在无人机进行搜索任务过程中常常会遇到山峰自然环境, 在这里假设山峰的外形类似圆锥形, 如图 6 所示。红色部分表示山峰形状, 蓝色部分表示山底。很显然当无人机在山底附近执行目标搜索任务时, 所受到的外部环境威胁较小; 而处于山峰位置时, 所处环境复杂且恶劣, 不利于搜索任务执行。通过图 6 清楚地认识到山峰对无人机的威胁系数从山顶到山底依次减小。结合上面二维栅格环境中障碍物的模拟表现形式, 尝试将山从俯视角度进行二维投影构想为一个圆形的障碍物, 从圆心 (也就是山顶) 到圆周边 (也就是山底) 沿半径方向对无人机飞行威胁约束越来越小。

结合负奖励自身传播特性, 它们会在价值函数中产生负的“尖峰”。这点性质恰好和自然环境中山对无人机飞行的威胁不谋而合。综合二者的共同之处, 尝试将搜索区域中存在的诸多约束障碍例如雷达威胁、防空火力威胁、自

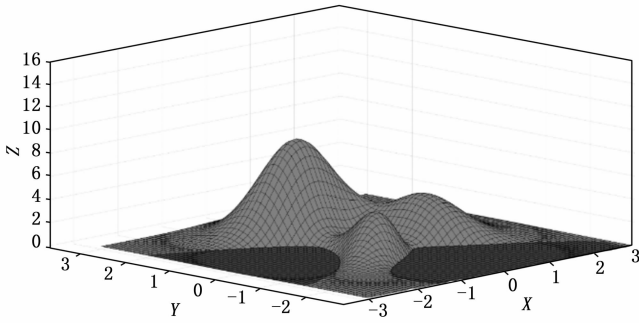
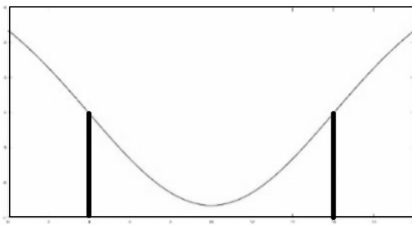


图 6 三维山峰示意图

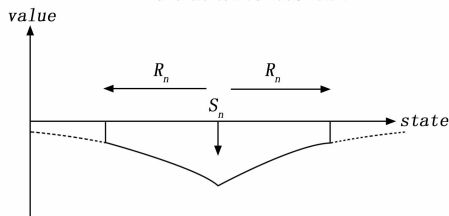
然环境障碍物等进行转换为二维环境中的障碍设置。通过对二维栅格环境中的圆形障碍物设置实现对无人机在环境中的负奖励转换。越靠近障碍物圆心，无人机所受威胁越大；越靠近圆周，无人机所受威胁越小。当无人机已经不在圆形障碍物范围内时，无人机将不会受到障碍物的任何威胁和约束。

3.2.2 风险井原理

基于上面的思路对负奖励机制转换寻求一种方法，尽可能地将飞行区域中涉及的灯塔、自然地形和其他飞行物以及雷达威胁、防空火力威胁等进行建模。很明显，无人机靠近障碍物应该比远离障碍物更危险。基于这种思想，为此在这里将这种情况进行命名，称为“风险井”。图 8 展示了一个“风险井”的存在形式。



(a) 负奖励作用机制效果图



(b) 值函数随状态变化示意图

图 7 “风险井”作用机制效果及值函数变化图

如图 7 (a) 所示，负奖励机制两条竖直黑线所夹区域对值函数发生效用，值函数发生迭代，类似于之前提到的山峰剖面图，随着距离底部波峰的距离增大，负奖励机制对值函数的影响逐渐减小。这也正类似于山峰到山底对无人机的威胁程度。当位于两条黑线向外区域时，负奖励对状态值函数不会发挥作用。图 7 (b) 以更直观的形式说明“风险井”中当位于 r 范围内时，负奖励才会对值函数迭代产生影响。并且随着 r 的不断增大，其影响越来越小；当位

于 r 范围外时，负奖励在值函数迭代中不会产生任何影响。负奖励的期望形式，在 S_i (例如障碍物的定位) 状态下，此刻获得的负奖励最多。当从 S_i 进一步得到负奖励衰减。当位于半径 r 之外，如果不存在影响，将指数衰减截断为 0。这种形式在此称为“风险井”。

3.2.3 引入负奖励机制下的实际奖励

当奖励以正数形式存在，它代表收益；当它以负数形式存在，它代表损失。无人机在此刻状态和动作共同作用下，以即时奖励函数形式作为评判在每个栅格中取得奖励优劣的准则。以定量的方法估算无人机搜索过程中对搜索目标的贡献程度。随后将根据无人机相对目标所处位置的转移概率，尝试创建从初始点开始搜索至找到目标后停止搜索的各状态下奖励函数。假设无人机在无障碍条件下进行马尔科夫移动的奖励函数为 R ，在风险井区域进行搜索的奖励函数为 $R_{riskwell}$ ，则整个目标搜索航迹规划过程中的真正奖励函数 R_{real} ，如式 (2) 所示：

$$R_{real} = R - R_{riskwell} \quad (2)$$

4 仿真实验实验结果与分析

4.1 搜索区域建模

假设在无人机无障碍区域进行马尔科夫移动，在初始状态下无人机在划分为 9 个序号方向的 3×3 栅格区域中的移动是随机且平等概率的。换言之，其在任意栅格处，其靠近相邻的 8 个栅格移动或者停住不动的概率是相同的，都是 $1/9$ 。设定该状态下的奖励函数 R 如式 (3) 所示：

$$R = \begin{cases} p_{target}(i) & \text{未搜索到目标} \\ 100 & \text{搜索到目标} \end{cases}, i \in S \quad (3)$$

尝试在上面设计的栅格环境中将灯塔、自然环境地形和其他飞行物以及雷达威胁、防空火力威胁等威胁在合适的情况下设定为数个半径不一的圆形障碍物。利用不同的半径圆来模拟“风险井”以及实际环境中如灯塔，山峰以及其他飞行物等障碍物对无人机栅格环境下做马尔科夫运动搜索目标产生的影响。关于风险井的负奖励函数设定如式 (4) 所示：

$$R_{riskwell} = \begin{cases} -100 + 10i(2 \leq i \leq 5), & \text{状态 } i \text{ 到风险井圆心距离} \\ 0, & \text{状态 } i \text{ 未处于风险井区域} \end{cases}, i \in S \quad (4)$$

进行仿真前设置相关参数，当无人机在 $5 \text{ km} \times 5 \text{ km}$ 的区域环境里搜索目标，则该区域拆解划分成 50×50 大小相同的单元格，每一格单元格代表一个 100 m 移动步长。给定无人机整个过程中截止步长存在最大值 k ，并对进行赋值 100。无人机于 50×50 的栅格环境里任何地点都能执行马尔科夫移动飞行，同时获悉目标的具体位置以及无人机每一步移动的转移概率。在实验过程，已经部署了一架无人机来对目标进行精确搜索。为了使计算量和时间得到有效控制，首先在搜索过程中，无人机可依据设定的转移概率矩阵，持续预测接下来的航迹分布。无人机一旦移至目标位

置, 搜索便会成功完成并终止。如果在给定最大搜索步长范围里无人机无法到达位置, 视作任务执行失败。给定无人机起点位置和目标位置并搭建好的栅格环境, 如图 8 所示, 右上角灰点为无人机初始位置, 左下角白色点为静态目标位置。

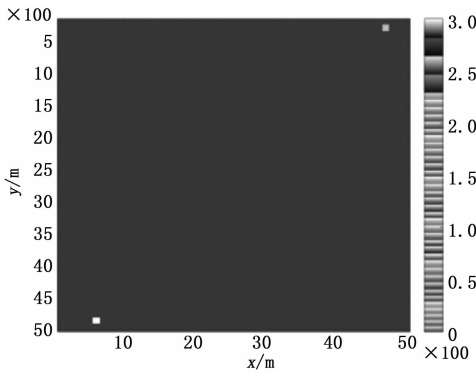


图 8 栅格环境下无人机及目标起始位置设定

4.2 奖励函数设定

在本次仿真实验中, 使用的强化学习算法是 Q-learning 算法, Q-learning 算法是一种基于动作函数 (即 Q 函数) 的强化学习方法, 即通过判断每一步动作的价值奖励来进行下一步的动作。在实际算法中, Q 函数的具体的更新函数如式 (5) 所示:

$$Q(S_t, A_t) = Q(S_t, A_t) + \alpha [R + \gamma * Max Q(S_{t+1}, A_{t+1}) - Q(S_t, A_t)] \quad (5)$$

为了体现无人机在目标搜索任务中对复杂环境中障碍物的模拟, 现对上面小节中已经设定好无人机和目标初始位置的栅格环境随机进行添加 7 个半径不同且位置不重合的“风险井”障碍物的效果图, 如图 9 所示。

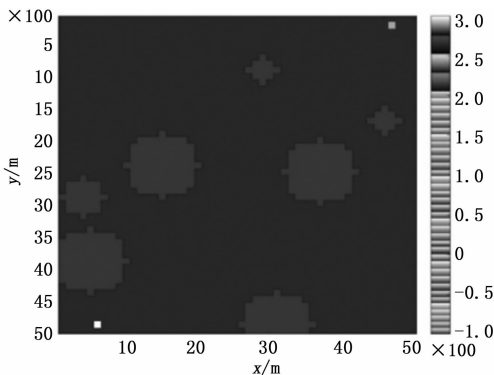
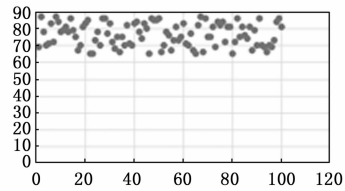


图 9 随机添加风险井效果图

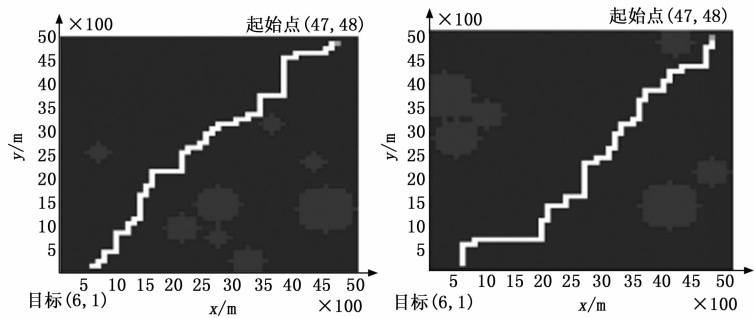
4.3 实验仿真结果分析

本次仿真实验进行循环进行 100 次后记录, 记录每次完成目标搜索任务所需步长。实验结果和航迹规划轨迹如图 10 所示。图 10 (a) 为得到的 100 次重复仿真实验, 成功找到目标的搜索步长数据记录; 图 10 (b) 是随机 2 次实验中完成目标搜索的航迹规划轨迹图。

如图 10 (a) 所示, 统计 100 次重复实验搜索步长散点



(a) 100 次重复实验搜索步长散点图



(b) 2 次完成目标搜索航迹规划轨迹图

图 10 搜索步长散点图和航迹规划轨迹图

图的总步长为 7 554 次, 平均每次搜索步长为 75.54 次, 这样的目标搜索策略是十分快速且高效的。实验证明了设计提出的 MDP 多奖励协同控制迭代算法能够在复杂障碍约束的环境下能够很好地执行目标搜索任务。

5 结束语

本文研究了基于马尔科夫移动分布模型的单无人机执行目标搜索任务的避撞航迹规划策略。基于类比“风险井”原理引入一种新型的负奖励机制, 将障碍约束环境对无人机的作用直观地表现出来, 提出了一种 MDP 多奖励协同控制迭代算法。借助仿真模拟试验, 验证了基于复杂障碍约束环境下, 通过调整负奖励函数的迭代策略算法进行避撞并完成目标搜索任务的可行性。

参考文献:

[1] 蓝星辉. 无人机路径规划算法研究 [D]. 桂林: 桂林电子科技大学, 2023.

[2] 张志荣. 无人机, 智能化装备的“空中多面手” [J]. 工业设计, 2011 (5): 16-21.

[3] SZCZERBA J R, GALKOWSKI P, GLICKTEIN S I, et al. Robust algorithm for real-time route planning [J]. IEEE Trans. Aerospace and Electronic Systems, 2000, 36 (1): 869-878.

[4] 宋 超, 李 波, 马云红, 等. 基于优化 A* 和 MPC 融合算法的三维无人机航迹规划 [J]. 系统工程与电子技术, 2023, 45 (12): 3995-4004.

[5] 陈晓蒙. 基于 DirectX 的无人机航迹规划模拟系统的设计与实现 [D]. 哈尔滨: 哈尔滨工业大学, 2014.

[6] 汪晓群. 基于马尔科夫移动目标的无人机最优搜索策略研究 [D]. 郑州: 河南大学, 2018.

[7] ZHU P X, FANG X. Multi-UAV cooperative task assignment based on half random Q-learning [J]. Symmetry, 2021, 13 (12): 2417-2417.

- [8] 赵 雅. 需求演变情况下应急物资动态分配优化模型与算法设计 [D]. 济南: 济南大学, 2023.
- [9] 马 庆, 丁文将, 邓关军. 农民工行业转移就业研究——基于马尔可夫链模型的分析 [J]. 中国人力科学, 2023 (9): 82-92.
- [10] 彭志行, 马红霞, 赵 杨, 等. 马尔可夫链模型在医学高等教育教学质量评价中的应用研究 [J]. 中国高等医学教育, 2009 (5): 22-24.
- [11] 王楨发. 无人车巡检路径规划研究 [D]. 南京: 南京邮电大学, 2022.
- [12] 谷志珉, 岳伟韬, 苏 婧, 等. 占据栅格地图的最佳栅格大小与地图精度 [J]. 机器人, 2020, 42 (2): 199-206.
- [13] 叶文敏. 有限位置马尔科夫移动目标的最优搜索策略研究 [D]. 武汉: 武汉理工大学, 2013.
- [14] 孙 涛. 无人机路径规划算法优化研究 [D]. 天津: 中国民航大学, 2022.
- [15] 郭 军, 朱 凡, 刘远飞. 基于马尔科夫链预测的多无人机协同搜索控制 [J]. 火箭与制导学报, 2007, 27 (5): 315-318.
- (上接第 291 页)
- [2] 祁铎颖, 贺 萍. 跨模态数据融合综述 [J]. 软件工程, 2022, 25 (10): 1-7.
- [3] 李高勇, 刘 露. 多模态大模型如何助力企业数字化转型 [J]. 清华管理评论, 2024 (z1): 105-109.
- [4] 杜锦丰, 王海荣, 梁 焕, 等. 基于表示学习的跨模态检索方法研究进展 [J]. 广西师范大学学报 (自然科学版), 2022, 40 (3): 1-12.
- [5] 徐文婉, 周小平, 王 佳, 等. 跨模态检索技术研究综述 [J]. 计算机工程与应用, 2022, 58 (23): 12-23.
- [6] CAO M, LI S, LI J, et al. Image-text retrieval: a survey on recent research and development [J]. IJCAI, 2022: 5410-5417.
- [7] WANG S, ZHU L, SHI L, et al. A survey of full-cycle cross-modal retrieval: from a representation learning perspective [J]. Applied Sciences, 2023, 13 (7): 4571.
- [8] ROUSSOPOULOS N, KELLEY S, Vincent F, et al. Nearest neighbor queries [C] //Proceedings of the 1995 ACM SIGMOD International Conference on Management of Data, 1995: 71-79.
- [9] XU H, WANG J, LI Z, et al. Complementary hashing for approximate nearest neighbor search [C] //2011 International Conference on Computer Vision. IEEE, 2011: 1631-1638.
- [10] LUO X, WANG H, WU D, et al. A survey on deep hashing methods [J]. ACM Transactions on Knowledge Discovery from Data, 2023, 17 (1): 1-50.
- [11] CAO W, FENG W, LIN Q, et al. A review of hashing methods for multimodal retrieval [J]. IEEE Access, 2020, 8: 15377-15391.
- [12] CAO Y, LONG M, WANG J, et al. Deep visual-semantic hashing for cross-modal retrieval [C] //Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2016: 1445-1454.
- [13] JIANG Q Y, LI W J, et al. Deep cross-modal hashing [C] //2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017: 3270-3278.
- [14] LI C, DENG C, LI N, et al. Self-supervised adversarial hashing networks for cross-modal retrieval [C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 4242-4251.
- [15] DENG C, CHEN Z, LIU X, et al. Triplet-based deep hashing network for cross-modal retrieval [J]. IEEE Transactions on Image Processing, 2018, 27 (8): 3893-3903.
- [16] LIU X, LI Z, WANG J, et al. Cross-modal zero-shot hashing [C] //2019 IEEE International Conference on Data Mining (ICDM), 2019: 449-458.
- [17] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need [J]. Advances in Neural Information Processing Systems, 2017: 30.
- [18] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770-778.
- [19] HINTON G E, SALAKHUTDINOY R R, et al. Reducing the dimensionality of data with neural networks [J]. Science, 2006, 313 (5786): 504-507.
- [20] KAUFMAN L, ROUSSEEUW P J, et al. Finding groups in data: an introduction to cluster analysis [M]. John Wiley & Sons, 2009.
- [21] MURTAGH F, CONTRERAS P, et al. Algorithms for hierarchical clustering: an overview [J]. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 2012, 2 (1): 86-97.
- [22] LIN T Y, MAIRE M, BELONGIE S, et al. Microsoft coco: Common objects in context [C] //Computer Vision-ECCV 2014: 13th European Conference, Zurich, Switzerland, Proceedings, Part V 13. Springer International Publishing, 2014: 740-755.
- [23] PLUMMER B A, WANG L, CERVANTES C M, et al. Flickr30k entities: Collecting region-to-phrase correspondences for richer image-to-sentence models [C] //Proceedings of the IEEE International Conference on Computer Vision, 2015: 2641-2649.