

文章编号:1671-1637(2026)03-0215-13

基于值迭代的无人机动态避撞优化方法

魏志强*, 安心

(中国民航大学 空中交通管理学院, 天津 300300)

摘要:针对无人机飞行冲突自主解脱需要,提出了一种基于值迭代方法的马尔可夫决策过程优化模型。首先构建了值迭代动态避撞模型,实现无人机的实时安全避撞;然后针对空域的复杂性和不确定性问题,构建了涵盖两机相对高度、本机与入侵机的垂直速度、历史动作及时间等参数的精细化状态空间集;之后通过构建多因素动态成本函数,综合考虑冲突风险、最接近时间等因素进行动作判断,减少了无人机避撞时的不必要机动操作;最后提出通过引入自适应双层概率融合机制,解决传统确定性决策在复杂动态环境中的脆弱性问题,提高决策的鲁棒性。仿真试验结果表明:提出的动态避撞方法在 3 个仅考虑动态入侵机的冲突场景中可以实现无人机的安全避撞,两机最终相对高度分别为 152.5、188.0、143.7 m;在同时考虑静态障碍物和动态入侵机的混合冲突场景中,本机与静态障碍物的最小垂直相对高度为 174.7 m,两机的垂直相对高度为 230.7 m,可以保证无人机安全飞行;与动态窗口法方法相比,4 个场景下本机执行基于值迭代的避撞策略后,平均过度位置调整高度减少了 62.4%,平均不必要的动作切换次数减少了 88%。提出的基于值迭代的动态规划方法解决无人机避撞场景下的马尔可夫决策过程问题是可行的,无人机可以实现安全避撞。

关键词:航空安全;动态避撞;值迭代;动态规划;无人机;低空交通;马尔可夫决策过程

中图分类号:U8 **文献标志码:**A **DOI:**10.19818/j.cnki.1671-1637.2026.153

A dynamic collision avoidance method for UAVs using value iteration

WEI Zhi-qiang*, AN Xin

(College of Air Traffic Management, Civil Aviation University of China, Tianjin 300300, China)

Abstract: A Markov decision process (MDP) optimization model based on value iteration was proposed for the needs of autonomous conflict resolution of unmanned aerial vehicles. A value iteration-based dynamic collision avoidance model was first constructed to achieve real-time safe collision avoidance. To address the complexity and uncertainty of airspace, a refined state space was formulated, incorporating parameters such as relative altitude between two aircraft, vertical speeds of ownship and intruder, historical actions, and time. A multi-factor dynamic cost function was designed to integrate conflict risk and time to closest approach for action judgement, thereby reducing unnecessary maneuvers during collision avoidance. An adaptive two-layer probabilistic fusion mechanism was introduced to address the vulnerability of traditional deterministic decision-making in complex dynamic environments and improve decision robustness. The results indicate that the proposed dynamic collision avoidance method can achieve safe

出版历程:2025-07-29 收稿,2025-12-11 修回,2026-01-23 录用

基金项目:天津市科技计划项目(23JCZDJC00580)

* 作者简介:魏志强(1979-),男,河南浉池人,教授,E-mail:weizhiqia@sina.com。

引用格式:魏志强,安心. 基于值迭代的无人机动态避撞优化方法[J]. 交通运输工程学报,2026,26(3):215-227.

Citation: WEI Zhi-qiang, AN Xin. A dynamic collision avoidance method for UAVs using value iteration[J]. Journal of Traffic and Transportation Engineering, 2026, 26(3): 215-227.

collision avoidance in three conflict scenarios considering only dynamic intruders, and the final vertical relative heights between two aircraft are 152.5, 188.0, and 143.7 m, respectively. In the mixed conflict scenario considering both static obstacles and dynamic intruders, the minimum vertical relative height between the ownship and static obstacles is 174.7 m, and the vertical relative height between two aircraft is 230.7 m, which ensures the safe flight of unmanned aerial vehicle. Compared with the dynamic window approach (DWA) method, after the ownship executes the collision avoidance strategy based on value iteration in four scenarios, the average excessive altitude adjustment is reduced by 62.4%, and the average number of unnecessary action switches is reduced by 88%. It is indicated that the proposed dynamic programming method based on value iteration is feasible to solve the Markov decision process problem in collision avoidance scenarios, and the unmanned aerial vehicle can achieve safe collision avoidance.

Keywords: aviation safety; dynamic collision avoidance; value iteration; dynamic programming; unmanned aerial vehicle; low-altitude traffic; Markov decision process

Publication history: Received 2025-07-29; Received in revised form 2025-12-11; Accepted 2026-01-23

Funding: Tianjin Science and Technology Program (23JCZDJC00580)

* **Corresponding author:** WEI Zhi-qiang, professor, E-mail: weizhiqia@sina.com.

0 引 言

随着无人机(Unmanned Aerial Vehicle, UAV)应用空间的扩大和需求的增加,与传统有人机在同一空域内的融合运行成为研究热点问题^[1-2]。融合运行机型多样,飞机性能差异大,空域飞行环境复杂,易受物障、天气等因素的影响^[3]。UAV与有人驾驶飞机的融合运行既需要和飞行员能够看到并避开其他飞机(Detect and Avoid, DAA)的功能,也需要像现有空中交通告警和防撞系统(Traffic Alert and Collision Avoidance System, TCAS)和机载防撞系统(Airborne Collision Avoidance System, ACAS X)一样能够提供避碰(Collision Avoidance, CA)的能力。对于无法看到其他飞机的UAV, DAA系统使用传感器来探测其他飞机,并为UAV飞行员提供态势感知^[4]。

在复杂空域下(如城市低空、多机协同场景),如何进一步提升避撞决策的鲁棒性与适应性仍面临挑战,动态避撞能力已成为无人机系统的关键技术瓶颈^[5]。张云燕等^[6]基于深度确定性策略梯度(Deep Deterministic Policy Gradient, DDPG)算法建立了端到端决策控制模型,试验结果表明该模型在无人机避障轨迹优化方面优于三维向量场直方图避障算法。刘钊瑄等^[7]在传统基于固定距离和时间的安全间隔标准设定原理基础上,提出无人机动态安全间隔模型与风险评估方法,在保证运行安全的同时实现空域资源的灵活使用。常绪成等^[8]针对传统无人

机避障算法在未知复杂环境中效能低的问题,提出了一种改进动态窗口(Dynamic Window Approach, DWA)融合算法并进行仿真验证,结果显示无人机避撞的整体性能得到了显著提升,对多无人机自主避障技术的工程应用具有重要意义。祁云等^[9]提出了一种结合智能冠豪猪算法(Crested Porcupine Optimizer, CPO)与DWA算法的路径规划方法,方法在复杂三维障碍环境中表现出高效的避障能力。尽管这些方法在确定性场景中表现良好,但难以适应空域的复杂性和不确定性。Li等^[10]提出一种多尺度图强化学习方法(Multi-scale Graph Reinforcement Learning, MS-GRL),在密集环境中学习无人机避撞策略,利用局部冲突随时间演变的强度来评估冲突中无人机的全局注意力权重,同时通过图嵌入和必要的细化特征来聚合无人机观测数据。

在避撞系统的设计与优化中,若系统配备了高精度且近乎完美的传感器,避撞问题可被看作是一个马尔可夫决策过程(Markov Decision Process, MDP),在MDP框架下,系统的状态演化被假定遵循固定的动态模型,即当前状态仅依赖于前一状态及所采取的动作,而与更早的历史状态无关^[11]。Kochenderfer等^[12]设计了一种基于部分控制的MDP的无人机避撞方法,通过将问题分解为可控和不可控的2个子问题,降低了计算复杂度,该方法在模拟环境中显著优于现有的TCAS系统,冲突率降低,同时减少了不必要的警报和路径偏离。Mueller等^[13]提出了一种基于部分可观测马尔可夫决策过

程(Partially Observable Markov Decision Process, POMDP)的多旋翼无人机动撞算法扩展,通过动态规划优化水平加速度避撞策略,实现了在数千次碰撞模拟中同时满足最小间隔标准和任务轨迹跟踪的目标。Sunberg等^[14]提出了一种基于相遇状态动态调整参数的动态规划方法,试验表明,这种优化方法能够在不增加风险比的情况下减少偏离标称路径的次数。Jiang等^[15]提出了一种基于POMDP和改进灰狼优化器(Grey Wolf Optimizer with Leadership Competition, GWOLC)的三维无人机路径规划与避障算法,通过信息粒子滤波树解决连续域信念更新问题,仿真验证了算法在约束处理方面的优越性。Al-Husseini等^[16]将初期野火攻击监控与抑制建模为多代理部分可观测马尔可夫决策过程,通过分层框架结合无人机监控和有人直升机抑制,提升野火控制效率。由上述文献可知,马尔可夫过程在避撞领域有广泛应用,能够有效建模复杂环境中的不确定性,并通过动态规划或强化学习等方法生成最优避撞策略,提升运行安全性和效率。MDP通过显式建模环境的不确定性和多步决策的长期影响,为动态避撞提供了理论支撑,但是避撞决策的鲁棒性不足,无人机采取的不必要机动动作较多。

美国联邦航空管理局(FAA)开发了ACAS X系列防撞系统,ACAS Xu是为无人机设计的,提供CA和DAA能力,允许多源监视信息输入^[17]。CA决策逻辑基于CA问题的马尔可夫决策过程模型,生成水平和垂直机动以避免碰撞。DAA逻辑使用展开方法来预测较长时间线的碰撞风险,以支持飞行员与空中交通管制员的交互和通信。Owen等^[18]提出了一种ACAS Xu的逻辑选择功能,可以在水平和垂直逻辑之间进行有效选择,不需要同时使用水平和垂直机动。蒙特卡洛模拟表明,该功能可以为各种监视来源提供性能优势。ACAS Xu的监视和跟踪模块集成了多个监视输入,包括自动依赖监视广播(Automatic Dependent Surveillance-broadcast, ADS-B)、主动模式S/C审讯、机载空对空雷达。目前,已有相关组织进行了ACAS Xu的飞行验证。随着无人机在民用空域的日益普及,ACAS Xu将成为确保无人机安全运行的关键技术之一^[19]。Rorie等^[20]的研究表明,飞行员可以有效地利用ACAS Xu管理DAA功能。Stroeve等^[21]提出了一种用于评估无人机在响应ACAS Xu探测与避免系统生成的清晰制导及冲突解脱建议时所执行机动的远程飞行员(Remote Pilot, RP)模型,构建了包含

航空器及其传感系统的仿真环境,结果表明远程飞行员行为特性会影响ACAS Xu建议机动策略的实施效果并增加系统复杂性。当前研究虽验证了MDP在ACAS Xu冲突解脱逻辑中的有效性,但在根据无人机实际位置进行动态调整防止过度机动,以及成本函数的设置等方面仍存在不足。本文针对上述问题有如下创新:

(1)针对复杂空域的不确定性问题,提出重新划分状态空间,采用六维状态空间表示,包括相对高度、本机垂直速度、入侵机垂直速度、前一个动作、当前动作持续时间以及时间。这种多维状态能够更全面地捕捉飞行环境的动态特性,有效防止了频繁的动作切换,提高了模型稳定性和实际可操作性,同时不会导致状态空间太大难以计算。

(2)针对减少无人机避撞时不必要的机动问题,设计了涵盖距离冲突时间、高度范围、动作切换等因素的成本函数。可以实时动态调整不同动作的成本,使无人机做出更优的避撞机动。

(3)针对空域条件动态变化场景下无人机动撞决策鲁棒性不足等问题,提出自适应的马尔可夫决策框架。通过双层概率融合机制来选择最优避撞动作,结合确定性决策和随机性选择,在高概率情况下直接选择最优动作,而在不确定情况下保留一定随机性,提高了系统的鲁棒性。

本文首先描述无人机动撞问题;然后构建无人机动撞场景下的值迭代模型,构建多维状态空间,设计动态成本函数,引入自适应双层概率融合机制;之后进行4个不同场景下的仿真试验与分析,同时与DWA方法进行对比,验证该方法的可行性与优越性。

1 无人机动撞问题描述

无人机动撞场景通常指,2架或多架无人机在运行过程中的特定时间,与其他无人机的水平和垂直方向相对距离小于安全阈值(图1),应及时进行避撞决策机动,防止距离持续缩小至发生碰撞。在城市空中交通场景中,四旋翼无人机凭借其具备垂直起降、悬停效率高、对起飞环境要求低等优点,得到了广泛应用^[22]。由于城市建筑物影响、水平空间限制以及可能的多机协调场景等,无人机动撞决策需要在保证飞行安全的前提下,最小化调整范围和尺度^[23]。

无人机的垂直机动主要有:无冲突(Clear of Conflict, COC)、 $8\text{ m} \cdot \text{s}^{-1}$ 爬升率(Rate of Climb, CL)和 $8\text{ m} \cdot \text{s}^{-1}$ 下降率(Rate of Descend, DES)。

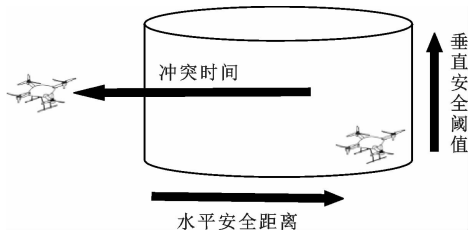


图 1 无人机碰撞示意

Fig. 1 Schematic of UAV collision

这些动作代表了在当前版本的 TCAS 中常用的标准垂直避撞机动,如图 2 所示。

动作的选择直接影响本机的垂直加速度,进而改变本机的垂直速度和高度。COC动作表示飞机

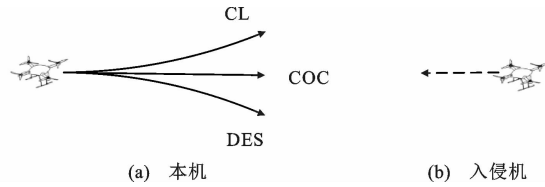


图 2 垂直机动示意

Fig. 2 Schematic of vertical maneuver

无需采取避让措施,垂直速度目标值为 $8 \text{ m} \cdot \text{s}^{-1}$; DES 代表飞机以 $0.25g$ 的垂直加速度下降直至达到 $8 \text{ m} \cdot \text{s}^{-1}$ 的稳定下降率;CL 代表飞机以 $0.25g$ 的垂直加速度爬升直至达到 $8 \text{ m} \cdot \text{s}^{-1}$ 的稳定爬升率,执行选择的动作后效果见图 3。

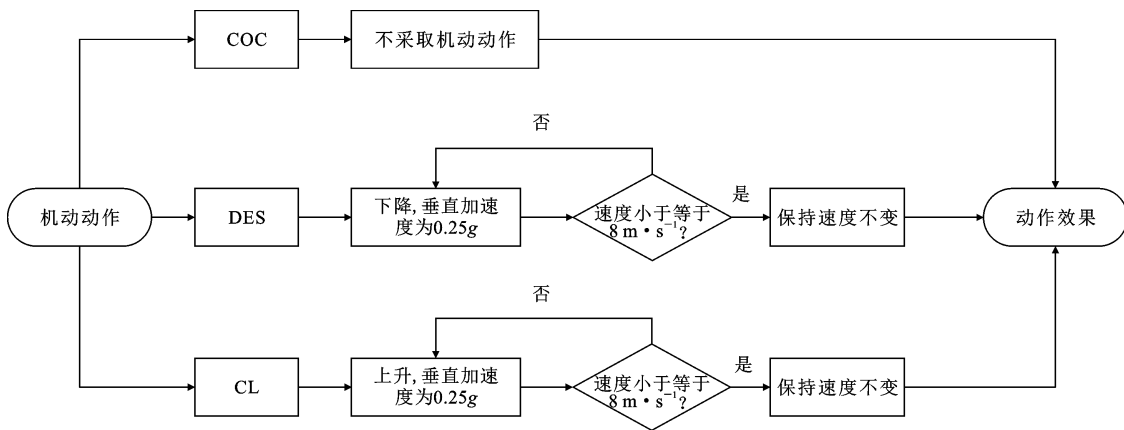


图 3 执行动作效果

Fig. 3 Execute action effect

为了实现上述场景,本文在城市低空空域场景中,通过评估本机当前状态与入侵机的相对位置关系,结合动态规划的值函数和马尔可夫决策过程,使无人机避撞时在每个时间步选择最优动作,以实现安全分离并最小化不必要机动的目标。

为了进一步阐释无人机动态碰撞场景,相关假设如下:

(1)在城市交通场景下进行无人机避撞研究。

(2)每秒钟做一次决策。

(3)无水平机动。入侵机以恒定速度迎头接近本机,双方均未执行任何水平航向改变,因此两机之间的水平接近率在整个过程中保持恒定。

(4)入侵机未配备防撞系统。

2 值迭代动态避撞模型构建

值迭代方法作为动态规划的核心算法,基于贝尔曼最优性原理,能够保证找到马尔可夫决策过程的全局最优策略,从理论上确保了决策的最优性。使用值迭代方法构建动态避撞模型,首先,确定值迭

代函数,计算最小代价值,对应动作即期望得到的最优机动方式;之后,构建无人机状态空间,全面考虑飞行环境的动态特性;其次,构建涵盖距离冲突时间、高度范围、动作切换等因素的成本函数;最后,提出状态转移模型,基于值迭代的动态避撞模型构建流程见图 4。

2.1 值迭代函数

首先定义一个最优期望代价函数作为目标函数,公式如下

$$V(s,t) = \min_{a \in A} \left\{ \left[\sum_{s'} P(s' | s,a) \right] [C(s,a,s',t) + \gamma V(s',t+1)] \right\} \quad (1)$$

$$A = \{COC, DES, CL\}$$

式中: A 为动作集合; s' 为下一状态; a 为当前准备选择的动作; γ 为折扣因子; $V(s,t)$ 为状态 s 在时间步 t 的值函数; $P(s' | s,a)$ 为状态转移概率; $C(s,a,s',t)$ 为即时成本函数。

值迭代方法采用反向递推方式,从终端时刻 t

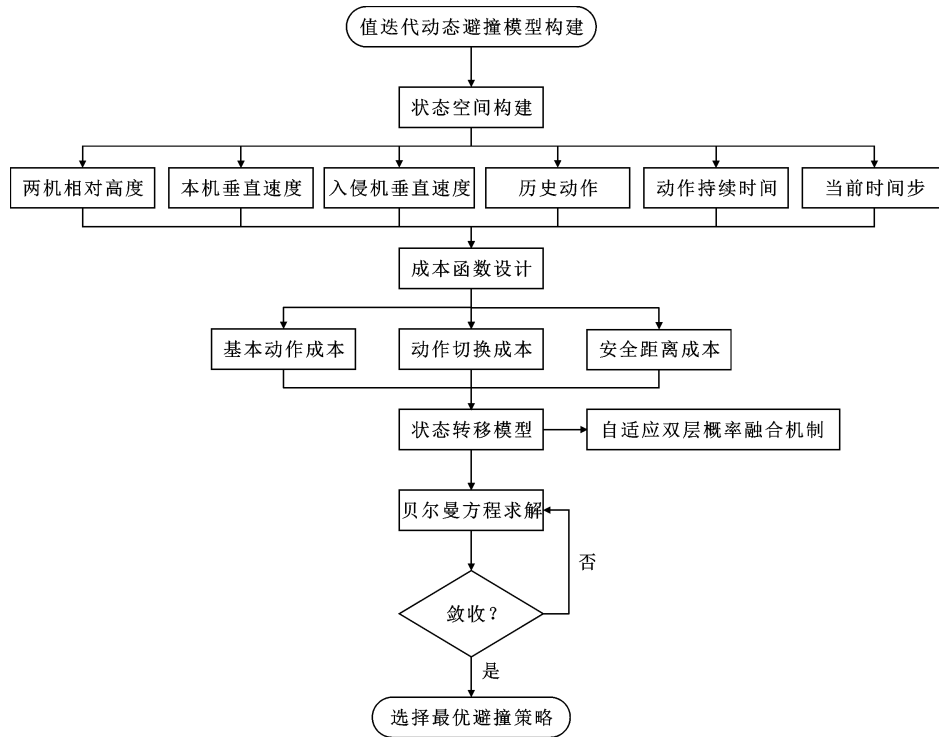


图 4 基于值迭代的动态避撞模型构建流程

Fig. 4 Construction flow of value-iteration-based dynamic collision-avoidance model

为 T 开始,依次计算每个时间步 t 的所有状态值函数。对于每个状态 s ,遍历所有可能动作 a ,通过状态转移函数获得后继状态分布,结合即时成本和折扣未来值函数计算期望代价,选择最小期望代价对应的动作更新值函数,直至满足收敛条件,公式如下

$$\pi^*(s,t) = \arg \min_{a \in A} \{ P(s'|s,a) [C(s,a,s',t) + \gamma V(s',t+1)] \} \quad (2)$$

2.2 状态空间

完整的六维状态空间 s 如下

$$s = (h_r, v_1, v_2, a_p, d, t) \quad (3)$$

式中: h_r 为入侵机相对于本机的高度; v_1 为本机的垂直速度; v_2 为入侵机的垂直速度; a_p 为前一个历史动作; d 为当前动作的持续时间; t 为当前时间步。

在状态空间中,相对高度直接反应两机垂直分离状态,垂直速度表现了避撞动作的趋势,前一动作和动作持续时间可以提供动作连续性约束,防止频繁切换。

2.3 成本函数

本研究在无人机避撞场景下,以最小化成本值的期望作为目标函数。

总成本函数可表示为多个组成部分的和,本文除了考虑基本的动作成本之外,还将其他因素纳入其他安全或操作考虑因素,如下式

$$C(s,a,s',t_c) = C_1(a) + C_2(a,a_p) +$$

$$C_3(h_r) + C_4(a,t_c,h_r) \quad (4)$$

式中: t_c 为距离冲突点的时间; $C_1(a)$ 为基本成本值; $C_2(a,a_p)$ 为动作切换成本; $C_3(h_r)$ 为基于高度范围的成本,包括对不同高度空间的奖励或惩罚; $C_4(a,t_c,h_r)$ 为基于与冲突点的时间和距离的成本。

2.4 状态转移模型

2.4.1 位置更新方程

根据当前状态和选定的动作,更新本机和入侵机的水平位置和高度。式(5)、(6)分别为本机的水平位置和高度更新,式(7)、(8)分别为入侵机的水平位置和高度更新,具体位置更新公式如下

$$x_e(t+1) = x_e(t) + v_h \Delta t \quad (5)$$

$$h_e(t+1) = h_e(t) + v_1(t) \Delta t \quad (6)$$

$$x_i(t+1) = x_i(t) + v_h \Delta t \quad (7)$$

$$h_i(t+1) = h_i(t) + v_2(t) \Delta t \quad (8)$$

式中: $x_e(t)$ 、 $x_i(t)$ 分别为本机和入侵机的水平位置; $h_e(t)$ 、 $h_i(t)$ 分别为本机和入侵机的垂直高度; v_h 为两机的水平相对速度。

2.4.2 自适应双层概率融合机制

设计双层概率融合机制保证决策质量,可以解决传统确定性决策在复杂动态环境中的脆弱性问题。机制根据环境危险程度自适应调整动作选择的概率分布,进而影响动作切换。

(1)在当前状态采取选择的动作后,下一时刻状

态不一定位于离散的状态网格点上,所以需要使用插值的方法对状态进行拟合。三维线性插值提供了一种在离散网格点之间估计任意连续点值的方法,可以保证无人机对实时状态的有效响应。三维线性插值是在三维空间中的线性插值推广,用于在三维网格的单元中估计任意点的值。在本研究中,3个维度分别对应入侵机相对于本机的高度、本机垂直速度和入侵机垂直速度。

假设在三维网格中有以下8个顶点的函数值: $V_{000}=V(x_0, y_0, z_0)$ 、 $V_{001}=V(x_0, y_0, z_1)$ 、 $V_{010}=V(x_0, y_1, z_0)$ 、 $V_{011}=V(x_0, y_1, z_1)$ 、 $V_{100}=V(x_1, y_0, z_0)$ 、 $V_{101}=V(x_1, y_0, z_1)$ 、 $V_{110}=V(x_1, y_1, z_0)$ 、 $V_{111}=V(x_1, y_1, z_1)$,对于网格内任意点 (x, y, z) ,其中 $x_0 \leq x \leq x_1, y_0 \leq y \leq y_1, z_0 \leq z \leq z_1$,首先计算插值权重

$$\begin{cases} x_d = \frac{x-x_0}{x_1-x_0} \\ y_d = \frac{y-y_0}{y_1-y_0} \\ z_d = \frac{z-z_0}{z_1-z_0} \end{cases} \quad (9)$$

三维线性插值的完整公式如下

$$\begin{aligned} V_{xyz} = & (1-x_d)(1-y_d)(1-z_d)V_{000} + \\ & (1-x_d)(1-y_d)z_dV_{001} + (1-x_d) \cdot \\ & y_d(1-z_d)V_{010} + (1-x_d)y_d \cdot \\ & z_dV_{011} + x_d(1-y_d)(1-z_d) \cdot \\ & V_{100} + x_d(1-y_d)z_dV_{101} + x_d \cdot \\ & y_d(1-z_d)V_{110} + x_dy_dz_dV_{111} \end{aligned} \quad (10)$$

(2) 马尔可夫转移概率

马尔可夫转移矩阵 $\mathbf{P}(a' | a_c)$ 表示从当前动作转移至其他动作的概率,依次是 COC、DES、CL,公式如下

$$\mathbf{P}(a' | a_c) = \begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{bmatrix} \quad (11)$$

式中: a_c 为当前已经选择的动作; a' 为所有可能的动作。

(3) 双层概率集成方法

$$\mathbf{P}(a | s, t) = \frac{\mathbf{W}(a | s, t)\mathbf{P}(a | a_c)}{\sum_{a'} [\mathbf{W}(a' | s, t)\mathbf{P}(a' | a_c)]} \quad (12)$$

$$\mathbf{W}(a | s, t) = \sum_{i=1}^8 w_i \cdot I[\pi^*(s_i, t) = a] \quad (13)$$

式中: $\mathbf{W}(a | s, t)$ 为通过状态插值从最优策略中获取的动作权重; s_i 为三维线性插值8个顶点的状态; w_i 为对应的插值权重; $I[\cdot]$ 为指示函数。

$$\mathbf{P}(a | a_c, s) = f[\mathbf{P}(a | a_c), h_r, t_c] \quad (14)$$

式中:函数 $f(\cdot)$ 为根据当前相对高度以及距离冲突点的时间动态调整转移概率,在接近冲突时增强避让的概率,在安全分离充分时鼓励返回正常飞行状态。

(4) 动作选择

选择概率最大的动作为最优动作,如下式

$$a^* = \arg \max_{a \in A} (\mathbf{P}(a | s, t)) \quad (15)$$

(5) 选择机制

在基于概率的动作选择中,引入一种确定性选择机制,以平衡随机性和稳定性

$$a_t = \begin{cases} \arg \max_{a \in A} (\mathbf{P}(a | s, t)) & \max_{a \in A} (\mathbf{P}(a | s, t)) > \theta \\ \text{RW}[\mathbf{P}(a | s, t)] & \text{其他} \end{cases} \quad (16)$$

式中: a_t 为在时间步 t 选择的动作; θ 为确定性阈值 0.6;RW(\cdot) 为基于给定概率分布的轮盘赌选择函数。

当某个动作的选择概率显著高于其他动作(超过阈值 0.6)时,模型会选择该动作。这种确定性选择在置信度高情况下避免了不必要的随机波动;当所有动作的选择概率较为接近(最高概率不超过阈值)时,模型会采用轮盘赌方法进行随机选择。

状态转移模型的流程图见图 5。

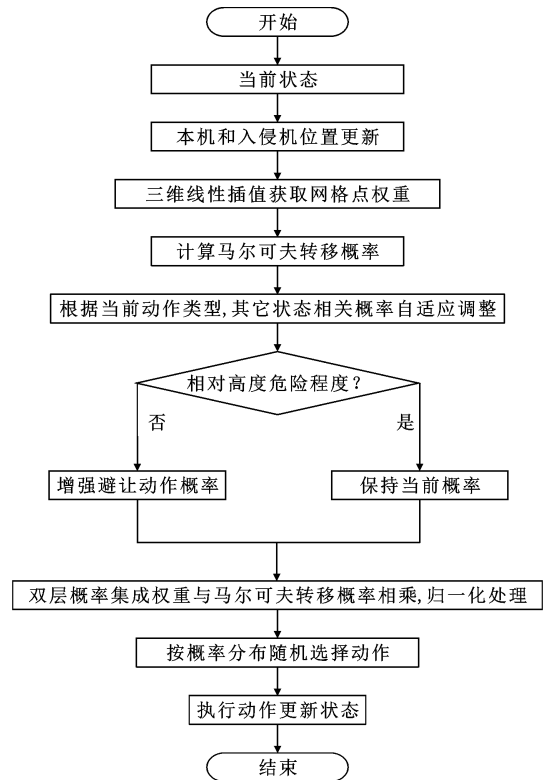


图 5 状态转移模型计算流程

Fig. 5 Process of state transition model calculation

3 验证分析

3.1 试验场景与参数设计

根据 FAA 白皮书中对现有净空定义中的阈值的修改,当 2 架无人机在一定的水平范围内时,相对垂直距离小于 137.0 m 时,则触发警报,应该采取机动^[24]。仿真试验在 MATLAB 2023B 环境下进行,设置无人机的水平速度为 $20 \text{ m} \cdot \text{s}^{-1}$,试验时间 45 s,步长 1 s。

为了验证在本机和入侵机的初始状态以及运动趋势不确定的条件下,策略仍具有有效性和鲁棒性,试验共设置 4 个冲突场景,分别代表了 4 种在低空常见的冲突情况。场景 1,代表两机水平相向冲突的场景,覆盖了一般可能的水平冲突情况;场景 2,本机和入侵机分别具有垂直方向速度,相向冲突,覆盖了可能的不同高度不同速度的无人机冲突常见情况;场景 3,两机航迹无交点,但是两机的垂直距离小于安全阈值,覆盖了可能存在的多种不安全场景;场景 4,在考虑两机不能发生冲突的同时,也考虑了静态障碍物,模拟了现实静态、动态障碍物混合存在的情况。

场景 1 中,两架无人机的初始高度均为 915 m,两机相距 1 800 m,均无初始垂直速率。若两机不采取动作避撞,将在(900,915) m 处发生碰撞;场景 2 中,本机的初始高度为 1 600 m,垂直速度为 $-8 \text{ m} \cdot \text{s}^{-1}$,入侵机的初始高度为 915 m,垂直速度为 $8 \text{ m} \cdot \text{s}^{-1}$,若两机不采取动作避撞,将在(900,1 240) m 处发生碰撞;场景 3 中,本机的初始高度为 1 260 m,无垂直速度,入侵机的初始高度为 915 m,垂直速率为 $8 \text{ m} \cdot \text{s}^{-1}$,两机不会发生碰撞;场景 4 中,本机的初始高度为 1 260 m,无垂直速度,入侵机的初始高度为 900 m,垂直速度为 $8 \text{ m} \cdot \text{s}^{-1}$,若两机不采取动作避撞,将在(900,1 260) m 处发生碰撞,静态障碍物的水平位置为 610 m,高度为 1 220 m,在避撞过程中位置不发生变化。每个场景中均为入侵机保持运动状态不变,本机机动避撞。

表 1 为无人机避撞所采取的动作策略,表 2 为使用的多维离散状态空间,状态网络规模为 2 457 945 个状态,贝尔曼方程中的折扣因子 γ 取值 0.99,可以向后覆盖所有的仿真时间步,注重无人机飞行长期安全。

表 3~6 分别为考虑了基本动作成本、动作切换成本、基于高度范围的成本、基于与冲突点的时间和距离的成本。基本动作成本值的设定基于安全优

表 1 动作空间

Table 1 Action space

动作	垂直速度/($\text{m} \cdot \text{s}^{-1}$)		垂直加速度	前一状态
	最大	最小		
COC	∞	$-\infty$	0	ALL
DES	∞	-8	0.25g	COC
CL	8	$-\infty$	0.25g	COC

表 2 离散状态变量

Table 2 Discrete state variable

状态变量	离散范围	离散尺度	离散值数量
h_r	-300,-270,...,300 m	30 m	21
v_1	-8,-7,...,8 $\text{m} \cdot \text{s}^{-1}$	1 $\text{m} \cdot \text{s}^{-1}$	17
v_2	-8,-7,...,8 $\text{m} \cdot \text{s}^{-1}$	1 $\text{m} \cdot \text{s}^{-1}$	17
a_p	COC, DES, CL		3
d	1,2,3 s	1 s	3
t	1,2,...,45 s	1 s	45

表 3 基本动作成本

Table 3 Basic action cost

动作	$C_1(a)$ 成本/ 10^{-2}
COC	0.1
DES	5.6
CL	5.6

表 4 动作切换成本

Table 4 Action switching cost

切换动作情况	$C_2(a, a_p)$ 成本
$a = a_p$	0.00
$a \neq a_p$	0.01
$a \neq a_p$ 且 a 不等于 COC 且 a_p 不等于 COC	0.06

表 5 基于高度范围的成本

Table 5 Altitude-based cost

相对高度范围/m	$C_3(h_r)$ 成本
$137 \leq h_r$	-0.30
$61 \leq h_r < 137$	-0.01
$30 \leq h_r < 61$	0.00
$15 \leq h_r < 30$	0.05
$h_r < 15$	0.15

表 6 基于与冲突点的时间和距离的成本

Table 6 Time-to-collision based cost

决策时机	$C_4(h_r, a, t_c)$ 成本
a 不等于 COC, $t_c > 30$	0.25
a 不等于 COC, $20 < t_c \leq 30$	0.10
a 等于 COC, $h_r < 61, t_c < 10$	0.15

先,同时减少不必要的动作切换的原则。将 COC 成本设定为极低的 10^{-3} , 目的是奖励模型在无冲突风险时优先保持平飞这一最经济、最舒适的飞行状态。将 DES 与 CL 的成本均设定为 5.6×10^{-2} , 为了抑制模型进行不必要的操作, 确保其仅在冲突风险远超此机动成本时才生成避撞指令。所有成本项指标量纲均为 1, 这是 MDP 中最优策略对成本函数线性缩放不变性的标准做法, 只需通过调节各项成本的相对比例即可平衡安全性与运行效率, 无需引入具有主观性的物理量纲转换^[25]。

表 7 为状态转移模型计算中相对高度的危险程度分级。

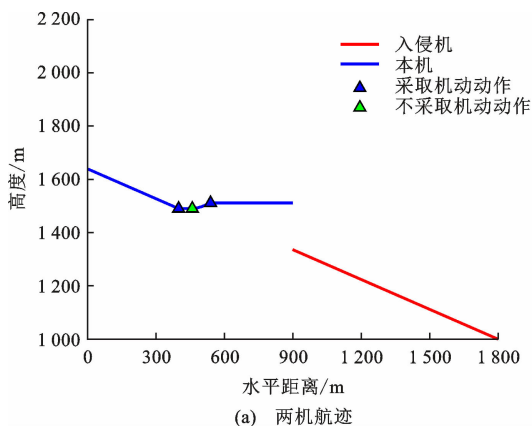
表 7 危险程度分级
Table 7 Risk classification

危险级别	相对高度范围/m
安全	$137 \leq h_r$
较安全	$61 \leq h_r < 137$
警戒	$30 \leq h_r < 61$
危险	$15 \leq h_r < 30$
极危险	$h_r < 15$

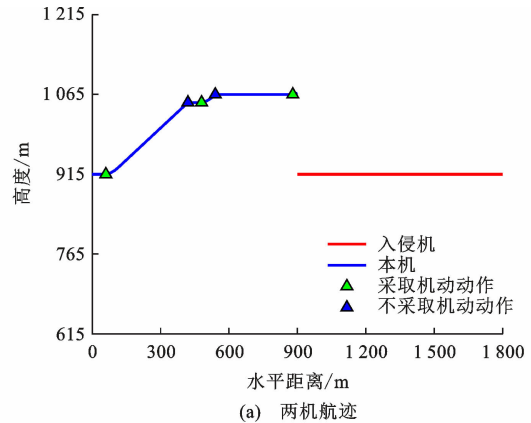
3.2 同高度相向航线避撞

场景 1 代表两机水平相向冲突的场景, 覆盖了一般可能的水平冲突情况, 同高度相向航线的避撞情况见图 6。本机在第 4 s 时开始机动, 选择 CL, 第 22 s 时动作转为 COC, 第 25 s 时动作改为 CL, 第 28 s 时保持 COC, 最终在 45 s 改为 CL, 最后两机的相对高度为 152.5 m。

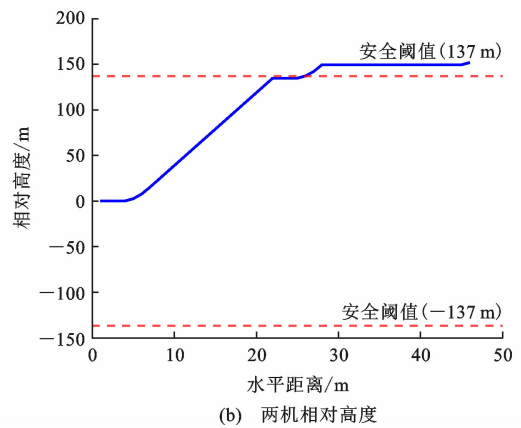
通过试验结果可以看出, 最终两机之间的垂直相对高度超出安全阈值 15.5 m, 两机无碰撞风险。在整个避撞过程中本机仅执行了 5 次动作切换, 减少了不必要的动作切换次数, 体现了方法策略的鲁



(a) 两机航迹



(a) 两机航迹



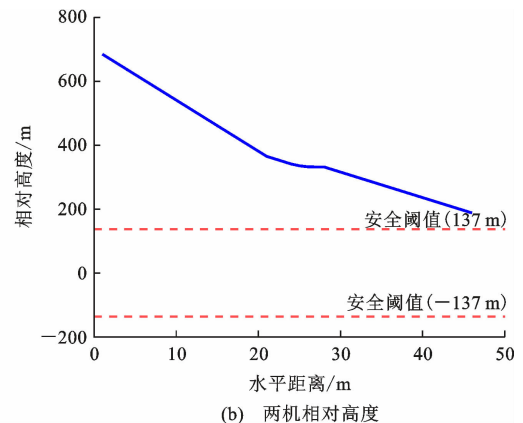
(b) 两机相对高度

图 6 同高度相向航线避撞

Fig. 6 Collision avoidance for head-on aircraft at the same altitude

3.3 不同高度相向航线避撞

场景 2 中本机和入侵机分别具有垂直方向速度, 存在相向冲突, 两机在不同高度的相向航线避撞情况见图 7。本机在前 20 s 综合冲突风险、最接近时间等因素进行动作判断, 保持 DES 动作, 在 21 s 时选择动作 COC 并保持 3 s, 在 24 s 时判断保持 COC 可能有碰撞风险, 切换至 CL 动作, 最后在 28 s 时判断没有冲突风险后, 保持 COC 不变, 最后两机



(b) 两机相对高度

图 7 两机相向航线避撞

Fig. 7 Collision avoidance for aircraft on head-on trajectories

的相对高度为 188.0 m。

通过试验结果可以看出,最终两机的垂直相对高度大于安全阈值 137.0 m,且避免了不必要的动作切换,两机无碰撞风险。

3.4 无交点航线避撞

场景 3 中两机航迹无交点,但是两机垂直距离小于安全阈值,无交点航线避撞情况见图 8。本机在 1~18 s 保持 COC,19 s 时动作选择 CL,持续至 33 s。

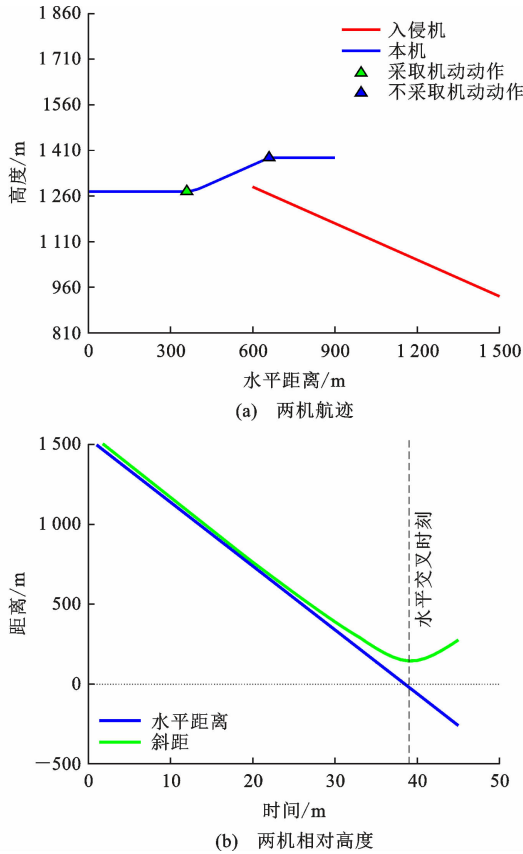


图 8 无交点航线避撞

Fig. 8 Collision avoidance for aircraft on non-intersecting trajectories

通过试验结果可以看出,两机水平交叉发生在约 39 s 处,两机相对高度为 143.7 m,高于 137.0 m 的安全阈值,两机无碰撞风险。同时,没有过多的动作变切换,避撞策略的鲁棒性较好。

3.5 考虑静态障碍物的入侵机爬升航线避撞

场景 4 模拟了现实静态、动态障碍物混合存在的情况,考虑静态障碍物的入侵机爬升航线避撞见图 9。考虑到静态障碍物位置固定的特殊性,采用混合成本函数方法进行处理。将总成本函数设计为入侵机避撞成本与障碍物避撞成本的加权组合,权重分别设为 0.6 和 0.4。障碍物避撞成本基于本机

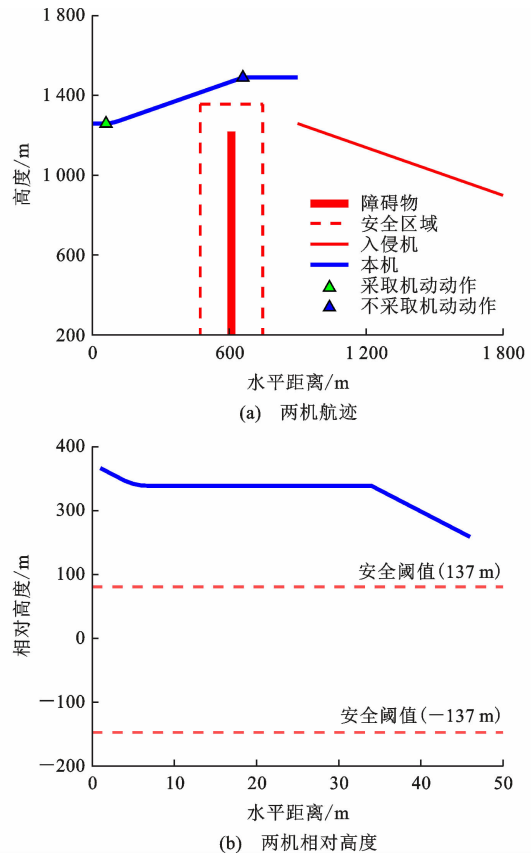


图 9 考虑静态障碍物的入侵机爬升航线避撞

Fig. 9 Collision avoidance for intruder aircraft's climb trajectory in the presence of static obstacles

与障碍物的相对高度计算,当相对高度小于安全距离 137 m 时进行惩罚,惩罚强度随接近程度递增。在值迭代过程中,通过当前时间步直接计算本机的水平位置,从而确定与固定障碍物的空间关系并计算相应成本。当本机同时面临入侵机和障碍物双重威胁时,比较两者的威胁程度,优先避让距离更近的目标。

如图 9 所示,该方法通过综合分析静态障碍物和动态入侵机的水平距离、垂直分离和时间因素,准确识别冲突威胁并确定最优采取机动时机,在前 3 s 保持 COC,之后选择 CL 持续上升 30 s,建立充足的安全分离距离,在 34 s 时选择 COC 动作至遭遇结束。

通过试验结果可以看出,最终两机的垂直相对高度为 230.7 m,与静态障碍物的最小垂直相对高度为 174.7 m,本机可以根据完整的冲突风险作出决策,仅执行 2 次动作切换,体现了方法对不同时刻危险程度的精确判断,同时,说明提出的方法可以在复杂空域中实现无人机安全避撞。

3.6 消融试验

为了验证模型中每个模块的实际贡献,分别针

对动态成本函数和双层概率融合机制进行消融试验。

3.6.1 动态成本函数消融试验

场景 1 中,在不考虑动态成本函数,仅保留概率融合机制时,独立验证动态成本函数模块的具体贡献,结果见图 10。

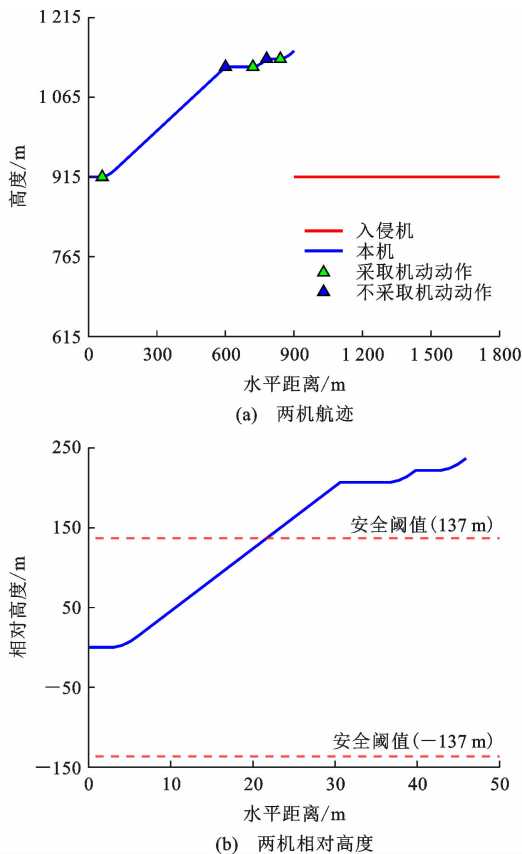


图 10 同高度相向航线避让

Fig. 10 Collision avoidance for head-on aircraft at the same altitude

在不考虑动态成本函数的消融试验中,最终两机之间的垂直相对高度超出安全阈值 100 m,本机执行了 5 次动作切换。从图 6(a)、10(a)的对比可以看出,虽然场景 1 中 2 种方法的机动切换次数均为 5 次,但提出的改进模型避让机动提前,有更充足的安全余量,消融试验在接近冲突点的区间(水平距离 600~900 m)存在更多的局部调整和航迹波动。提出的改进模型与仅考虑双层概率融合机制模型相比,过度位置调整减少了 84.5%,说明了动态成本函数的有效性。

3.6.2 双层概率融合机制消融试验

场景 2 中,在不考虑双层概率融合机制,仅保留动态成本函数的时候,独立验证其具体贡献,如图 11 所示。

在不考虑双层概率融合机制的消融试验中,最

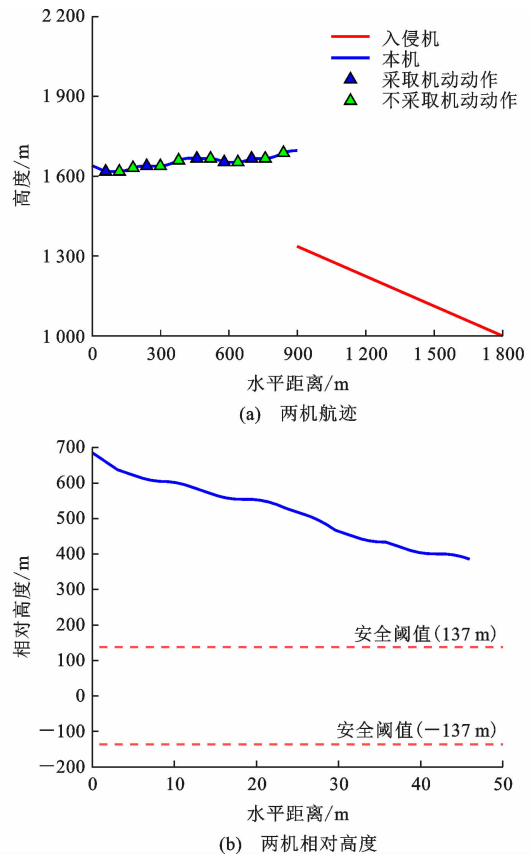


图 11 两机相向航线避让

Fig. 11 Collision avoidance for aircraft on head-on trajectories

终两机的垂直相对高度 385 m,切换动作 13 次。提出的改进模型与仅考虑动态成本函数模型相比,过度位置调整减少了 79%,动作切换次数减少 77%。说明了双层概率融合机制的有效性。

3.7 提出模型与经典值迭代方法优劣对比

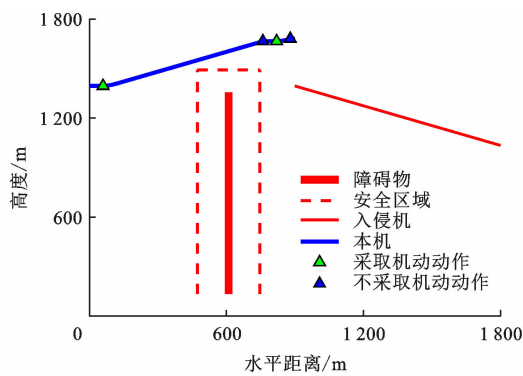
为了验证提出模型在避让效果、决策鲁棒性等方面的有效性,在场景 4 的静态动态障碍物混合场景中,与经典值迭代方法进行对比,具体试验参数设置见表 8,试验结果见图 12。

表 8 经典值迭代方法试验参数设置

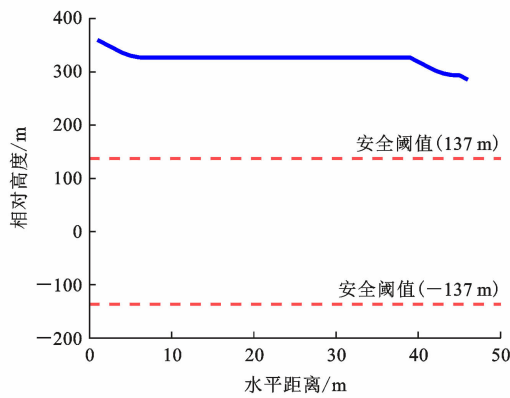
Table 8 Experimental parameter settings for the classical value iterative method

试验设置	状态空间规模/个	垂直速度/ ($m \cdot s^{-1}$)		垂直加速度		成本函数/ 10^{-2}	
		小	大	COC	0.00	COC	0.1
参数	2 457 945	-8	8	DES	0.25g	DES	5.6
				CL	0.25g	CL	5.6

由试验结果看出,经典值迭代方法在整个避让过程中进行了 4 次动作切换。通过试验结果可以看出,最终两机的垂直相对高度为 285.4 m,避让机动完成后,本机与静态障碍物的垂直相对高度为 174.7 m。



(a) 两机航迹



(b) 两机相对高度

图 12 使用经典值迭代方法考虑静态障碍物的入侵机爬升航线避撞
Fig. 12 Collision avoidance for intruder aircraft's climb trajectory in the presence of static obstacles using the classical value iterative method

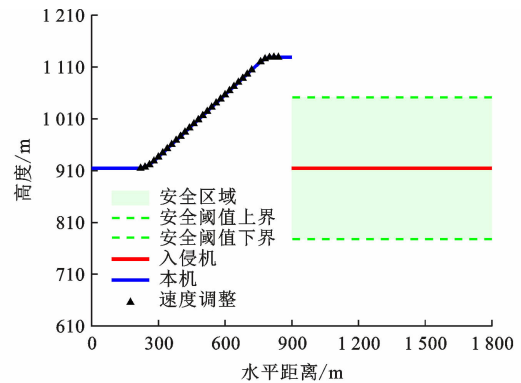
文中改进后的方法较经典方法,过度位置调整高度减少了 19%,不必要的动作切换次数减少了 50%。

3.8 提出模型与 DWA 的优劣对比

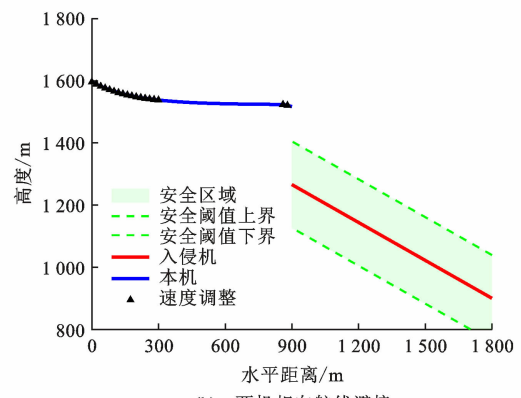
为了验证方法相对于其他避撞算法的优越性,在相同场景下(决策时间步长、最大调整加速度、两机初始状态等),分别与现有动态避撞方法 DWA 方法对比,结果如图 13 所示^[26]。同时从最终两机相对高度、动作切换次数等方面,将文中模型结果与 DWA 结果进行具体对比如表 9、10 和图 14 所示。

在 4 个场景中,2 种方法均可以使无人机安全避撞。本机使用值迭代方法执行避撞策略后与入侵机的相对高度平均超出安全阈值 42.0 m,使用 DWA 方法策略超出安全阈值 111.7 m,过度位置调整高度减少了 62.4%;值迭代方法的平均动作切换次数在 3 次左右,DWA 方法需要执行动作切换 25 次左右,不必要的动作切换次数减少了 88%。

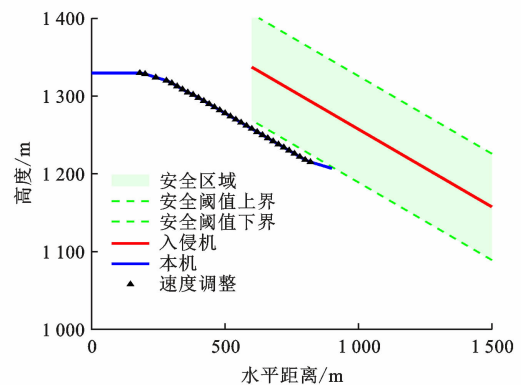
由以上结果可知,基于值迭代的无人机动态避撞方法在避撞效果、决策鲁棒性,以及动作切换次数方面均具有一定的优越性。在动作切换方面,DWA 方法在状态评价函数梯度变化敏感区域容易产生高



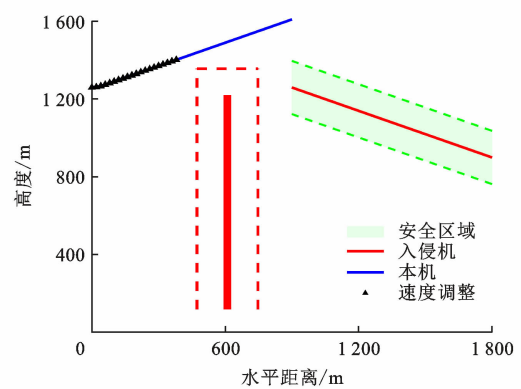
(a) 同高度相向航线避撞



(b) 两机相向航线避撞



(c) 无交点航线避撞



(d) 考虑静态障碍物的入侵机爬升航线避撞

图 13 4 个场景下 DWA 避撞方法

Fig. 13 DWA collision avoidance algorithm in four scenarios
频指令微调现象;而值迭代方法的自适应特点可以减少不必要的动作切换。在最终相对高度方面,

表9 值迭代方法4个场景相关结果

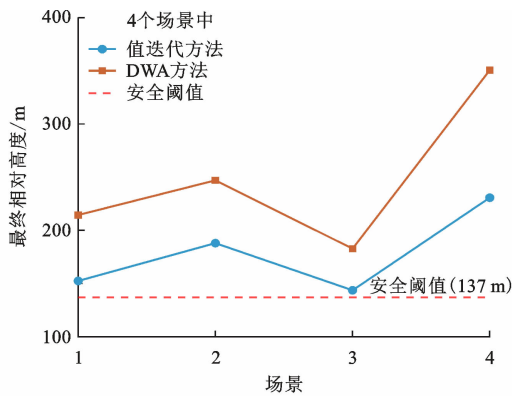
Table 9 Results of value iteration algorithm in four scenarios

场景	场景1	场景2	场景3	场景4
最终相对高度/m	152.5	188.0	143.7	230.7
动作切换次数	5	3	2	2

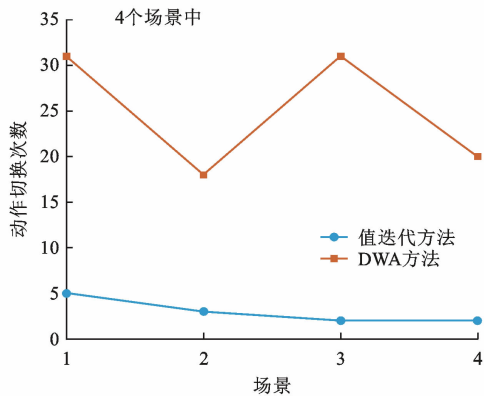
表10 DWA方法4个场景相关结果

Table 10 Results of the DWA method in four scenarios

场景	场景1	场景2	场景3	场景4
最终相对高度/m	214.4	247.0	182.8	350.5
动作切换次数	31	18	31	20



(a) 两种方法最终相对高度



(b) 两种方法动作切换次数

图14 两种方法避撞性能对比

Fig. 14 Performance comparison of the two collision-avoidance methods

DWA方法局部贪婪优化特性在接近安全边界时可能引发过度调整行为;值迭代方法获得的全局最优策略,可以有效避免不必要的大幅度位置调整。

4 结语

(1)提出的方法在3个仅考虑动态入侵机的冲突场景中可以实现无人机安全避撞,最终相对高度分别达152.5、188.0、143.7 m;在同时考虑静态障碍物和动态入侵机的冲突场景中,无人机仅在必要时刻(第4~33 s)执行CL机动,与静态障碍物的最小垂直相对高度为174.7 m,两机的垂直相对高度

为230.7 m,可以保证无人机安全飞行。

(2)在相同避撞场景下,提出的值迭代避撞方法较DWA避撞方法具有优势。在4个场景中,本机使用值迭代方法执行避撞策略后与入侵机的相对高度平均超出安全阈值42.0 m,使用DWA方法策略超出安全阈值111.7 m,过度位置调整高度减少了62.4%,说明提出的方法有利于无人机在复杂空域中进行避撞,不会过度进行位置调整;值迭代方法的动作切换次数在3次左右,DWA方法需要执行动作切换25次左右,不必要的动作切换次数减少了88%,说明提出的方法具有策略鲁棒性,可以减少不必要动作切换。

(3)本研究目前主要考虑2架无人机和障碍物的避撞,下一步研究应综合考虑多架无人机;当前的民航飞机使用的是TCAS II,其核心逻辑是基于垂直机动,只指令爬升或下降,不指令左右转弯,考虑到之后无人机与民航飞机的融合运行,本文也同样只考虑了本机在垂直方向机动。考虑水平机动可能会使避撞效果更好,但复杂度大大增加,后续研究应考虑加入水平机动。

参考文献:

References:

- [1] MANFREDI G, JESTIN Y. An introduction to ACAS Xu and the challenges ahead[C]//IEEE. 2016 IEEE/AIAA 35th Digital Avionics Systems Conference (DASC). New York: IEEE, 2016: 1-9.
- [2] 周志崇,赵顾颢,吴亚荣,等.局部空域高密度无人机冲突解脱算法研究[J/OL].北京航空航天大学学报,2025,https://doi.org/10.13700/j. bh. 1001-5965. 2025. 0157.
ZHOU Zhi-chong, ZHAO Gu-hao, WU Ya-rong, et al. Research on conflict resolution algorithms for high-density UAVs in local airspace[J/OL]. Journal of Beijing University of Aeronautics and Astronautics, 2025, https://doi.org/10.13700/j. bh. 1001-5965. 2025. 0157.
- [3] 张启钱,王中叶,张洪海,等.基于SMILO-VTAC模型的复杂低空多机冲突解脱方法[J].交通运输工程学报,2019,19(6): 125-136.
ZHANG Qi-qian, WANG Zhong-ye, ZHANG Hong-hai, et al. Multi-aircraft conflict resolution method for complex low-altitude airspace based on the SMILO-VTAC model[J]. Journal of Traffic and Transportation Engineering, 2019, 19(6): 125-136.
- [4] 陈思名. ACAS X系统监视跟踪与冲突解脱模块的研究与实现[D].成都:电子科技大学,2022.
CHEN Si-ming. Research and implementation of ACAS X system surveillance and tracking module and threat resolution module[D]. Chengdu: University of Electronic Science and Technology of China, 2022.
- [5] 陈丹,汤程,谢宇,等.面向城市低空物流配送的无人机实时航迹双层规划[J].航空学报,2025,46(16):229-247.
CHEN Dan, TANG Cheng, XIE Yu, et al. Real-time dual-

- layer trajectory planning for UAVs in urban low-altitude logistics delivery [J]. *Acta Aeronautica et Astronautica Sinica*, 2025, 46(16): 229-247.
- [6] 张云燕,魏 瑶,刘 昊,等.基于深度强化学习的端到端无人机避障决策[J].西北工业大学学报,2022,40(5):1055-1064. ZHANG Yun-yan, WEI Yao, LIU Hao, et al. End-to-end UAV obstacle avoidance decision based on deep reinforcement learning [J]. *Journal of Northwestern Polytechnical University*, 2022, 40(5): 1055-1064.
- [7] 刘钊瑄,师 可,许逸凡.融合空域无人机动态安全间隔模型与风险评估[J/OL].西安电子科技大学学报,2025,https://doi.org/10.19665/j.issn1001-2400.20241108. LIU Zhao-xuan, SHI Ke, XU Yi-fan. UAS dynamic separation design and collision risk analysis in integrated airspace[J/OL]. *Journal of Xidian University*, 2025, https://doi.org/10.19665/j.issn1001-2400.20241108.
- [8] 常绪成,王敬宇,李 康,等.基于改进DWA融合算法的多无人机避障方法[J/OL].弹箭与制导学报,2025,http://kns.cnki.net/kcms/detail/61.1234.tj.20250227.1118.002.html. CHANG Xu-cheng, WANG Jing-yu, LI Kang, et al. Multi-UAV obstacle avoidance method based on improved DWA fusion algorithm[J/OL]. *Journal of Projectiles, Rockets, Missiles and Guidance*, 2025, http://kns.cnki.net/kcms/detail/61.1234.tj.20250227.1118.002.html.
- [9] 祁 云,于开旺,李绪萍,等.融合CPO-DWA的矿井应急无人机路径规划[J/OL].安全与环境学报,2025,https://doi.org/10.13637/j.issn.1009-6094.2024.2198. QI Yun, YU Kai-wang, LI Xu-ping, et al. Mine emergency UAV path planning integrating CPO-DWA[J/OL]. *Journal of Safety and Environment*, 2025, https://doi.org/10.13637/j.issn.1009-6094.2024.2198.
- [10] LI Y, LI J, WANG J, et al. Multi-scale graph enhanced reinforcement learning for conflict resolution in dense UAV networks[J]. *IEEE Internet of Things Journal*, 2025: 1-14.
- [11] TEMIZER S, KOCHENDERFER M, KAELBLING L, et al. Collision avoidance for unmanned aircraft using Markov decision processes[C]//AIAA. AIAA Guidance, Navigation, and Control Conference. Reston: AIAA, 2010: 8040.
- [12] KOCHENDERFER M J, CHRYSSANTHACOPOULOS J P. Partially-controlled Markov decision processes for collision avoidance systems[C]//FILIFE J, FRED A. Proceedings of the 3rd International Conference on Agents and Artificial Intelligence (ICAART 2011). Berlin: Springer, 2011: 61-70.
- [13] MUELLER E R, KOCHENDERFER M. Multi-rotor aircraft collision avoidance using partially observable Markov decision processes[C]//AIAA. AIAA Modeling and Simulation Technologies Conference. Reston: AIAA, 2016: 3673.
- [14] SUNBERG Z N, KOCHENDERFER M J, PAVONE M. Optimized and trusted collision avoidance for unmanned aerial vehicles using approximate dynamic programming[C]//IEEE. 2016 IEEE International Conference on Robotics and Automation (ICRA). New York: IEEE, 2016: 1455-1461.
- [15] JIANG W, LYU Y, LI Y, et al. UAV path planning and collision avoidance in 3D environments based on POMPD and improved grey wolf optimizer [J]. *Aerospace Science and Technology*, 2022(121): 1-11.
- [16] AL-HUSSEINI M, WRAY K H, KOCHENDERFER M J. Hierarchical framework for optimizing wildfire surveillance and suppression using human-autonomous teaming[J]. *Journal of Aerospace Information Systems*, 2024, 21(10): 22.
- [17] OWEN M P, PANKEN A, MOSS R, et al. ACAS Xu: Integrated collision avoidance and detect and avoid capability for UAS[C]//IEEE. 2019 IEEE/AIAA 38th Digital Avionics Systems Conference (DASC). New York: IEEE, 2019: 1-10.
- [18] OWEN M P, KOCHENDERFER M J. Dynamic logic selection for unmanned aircraft separation[C]//IEEE. 2016 IEEE/AIAA 35th Digital Avionics Systems Conference (DASC). New York: IEEE, 2016: 1-8.
- [19] 汤新民,顾俊伟,刘 冰,等.低空监视技术及其发展趋势综述[J].南京航空航天大学学报,2024,56(6):973-993. TANG Xin-min, GU Jun-wei, LIU Bing, et al. Review on low-altitude surveillance technology and its development trend [J]. *Journal of Nanjing University of Aeronautics and Astronautics*, 2024, 56(6): 973-993.
- [20] RORIE R C, SMITH C, SADLER G, et al. A human-in-the-loop evaluation of ACAS Xu[C]//IEEE. 2020 AIAA/IEEE 39th Digital Avionics Systems Conference (DASC). New York: IEEE, 2020: 1-10.
- [21] STROEVE S, VILLANUEVA-CANIZARES C J, DEAN G. Remote pilot modelling for evaluation of ACAS Xu [J]. *Proceedings of the SESAR Innovation Days*, 2023, 5: 1-8.
- [22] 张国林,曾喆昭,唐钰淇.四旋翼无人机轨迹跟踪的ACPD控制方法[J].兵器装备工程学报,2025,46(5):185-191,271. ZHANG Guo-lin, ZENG Zhe-zhao, TANG Yu-qi. ACPD control method for trajectory tracking of quadrotor UAVs [J]. *Journal of Ordnance Equipment Engineering*, 2025, 46(5): 185-191, 271.
- [23] 申 炎,张学军,张维东.基于MAD3QN的多无人机协同避撞方法[J/OL].计算机工程与应用,2025,http://kns.cnki.net/kcms/detail/11.2127.tp.20250704.1214.004.html. SHEN Yan, ZHANG Xue-jun, ZHANG Wei-dong. Multi-UAV cooperative collision avoidance method based on MAD3QN[J/OL]. *Computer Engineering and Applications*, 2025, http://kns.cnki.net/kcms/detail/11.2127.tp.20250704.1214.004.html.
- [24] 高雅琪.无人机系统中DAA模块的研究和设计实现[D].成都:电子科技大学,2022. GAO Ya-qi. Research design and implementation on DAA module of UAV System [D]. Chengdu: University of Electronic Science and Technology of China, 2022.
- [25] NEU G, JONSSON A, GÓMEZ V. A unified view of entropy-regularized Markov decision processes [EB/OL] (2017-05-22). https://doi.org/10.48550/arXiv.1705.07798.
- [26] 刘连玉,琨在武,张 雪,等.应急情景下融合改进D* Lite算法和DWA算法的无人驾驶汽车路径规划[J/OL].控制与决策,2025,https://doi.org/10.13195/j.kzyjc.2025.0009. LIU Lian-yu, GONG Zai-wu, ZHANG Xue, et al. Emergency-scenario path planning for autonomous vehicles integrating improved D* Lite and DWA algorithms[J/OL]. *Control and Decision*, 2025, https://doi.org/10.13195/j.kzyjc.2025.0009.