

放宽静稳定度飞行器过失速重定向控制

崔颢, 张科, 黄煦晨, 李浩宇, 谭明虎, 王靖宇

(西北工业大学 航天学院, 陕西 西安 710072)

摘要:传统飞行器受离轴机动能力限制,难以有效满足后半球目标跟踪需求。提出利用近端策略优化(PPO)智能控制方法实现放宽静稳定度飞行器在气动控制下的过失速机动重定向。基于Jorgensen工程方法分析了飞行器在超大攻角下的非线性气动特性,揭示了静不稳定度对自翻转能力的影响,并对重定向过程进行了分段划分。针对超大攻角非线性姿态控制问题进行了马尔可夫决策过程建模,结合势能奖励函数设计引导智能体有效学习最优舵偏控制策略。在低马赫数条件下进行了重定向仿真实验,结果表明仅依靠空气舵即可控制飞行器完成快速重定向,并在速度过零段维持飞行器姿态稳定,蒙特卡洛实验进一步验证了该控制方法在随机初始条件下的鲁棒性。

关键词:放宽静稳定度;过失速机动;重定向;速度过零;智能控制;强化学习

中图分类号:V448

文献标志码:A

文章编号:1000-2758(2026)01-0151-09

在复杂空中环境中传统飞行器受离轴机动能力限制,难以实现对后方目标的轨迹跟踪^[1-2]。超大离轴轨迹跟踪要求飞行器通过快速重定向调整姿态和飞行轨迹,实现对后方目标的追踪,从根本上提升动态场景下的追踪性能。发展全向追踪能力已成为高动态飞行器姿态、轨迹控制的核心研究方向^[3]。

目前快速重定向技术研究聚焦于快速转弯飞行器的气动特性分析与姿态翻转控制。飞行器在重定向过程中,尤其是在超大攻角下,其气动特性表现出强非线性和非定常特征^[4-5]。现有研究基于细长体理论与黏性横流理论对飞行器重定向过程中的气动特性进行了建模,揭示了细长气动布局飞行器在转弯全过程中的俯仰力矩系数、轴向力系数、法向力系数等随攻角的变化规律^[6-7]。例如,文献[8-9]利用计算流体力学(CFD)数值模拟技术、动网格技术以及Jorgensen工程方法,对飞行器重定向过程中的气动特性进行了建模和数值计算,为后续姿态控制设计提供了重要理论依据和数据支撑。基于对飞行器

过失速翻转气动特性的深入理解,其他文献进一步提出多种重定向过程中的姿态控制方案。文献[10-11]利用最优控制理论,将快速转弯问题转化为轨迹优化问题,通过设置合理的性能指标函数求解最优控制指令,从而借助空气舵、矢量推力发动机、燃气舵等装置实现了飞行器的快速转弯控制。然而该方法需要精确的初始和终止条件,实际场景中载机可能以任意的角度、速度和角速度投放飞行器,最优控制指令的解算时间过长,实时性较差,难以应对复杂的空中环境。另一方面,文献[12-14]通过在飞行器上加装直接侧向力装置,利用滑模控制方法设计了直接力/气动力复合控制系统,实现了飞行器在铅垂平面内的180°转弯。这种方法能够快速调整飞行器姿态,但难以抑制超大攻角状态下气动参数摄动引发的控制抖振,且直接侧向力装置增加了飞行器的质量和结构复杂性。以上方法均存在执行机构过多、转弯半径过大的问题,会大幅增加快速转弯时的能量消耗,降低后续跟踪阶段的动能储备,削弱过载能力。

文献[6]提出一种放宽静稳定度的飞行器,仅通过纯空气舵实现飞行器低马赫数下的过失速机动翻转。这种转弯方案无需额外的直接侧向力装置,避免系统质量增加和可靠性降低等问题。进一步地,飞行器通过过失速机动直接调整纵轴指向,无转

收稿日期:2025-05-21

基金项目:空基信息感知与融合全国重点实验室和航空科学基金(ASFC-20230001053004, ASFC-20230001053006, ASFC-20240001053003, ASFC-20220001053001, 202468)联合资助

作者简介:崔颢(1977—),博士研究生

通信作者:李浩宇(1990—),副教授 e-mail:haoyuli@nwpu.edu.cn

弯半径,从而极大降低动能损失,具有广阔的应用前景。然而,在低马赫数状态下由于脱体涡的影响,飞行器过失速翻转过程中会经历强非线性的气动区域,并伴随有非正常迟滞效应,故难以进行动力学建模,依赖精确动力学模型的控制方法不适用于该方案。相较于传统控制方法,数据驱动控制在该场景中具有显著优势:不依赖精确的数学模型;能够通过与环境交互自主学习最优控制策略;适应复杂动态环境下的非线性、非正常气动特性。

本文以放宽静稳定度飞行器为对象,结合数据驱动控制中的近端策略优化(PPO)算法,提出一种仅通过空气舵控制实现快速转弯的智能控制方法,克服现有方法在姿态控制精确性和转弯快速性上的缺陷^[15-17]。首先,基于 Jorgensen 工程方法分析不同静稳定度飞行器在过失速机动时的气动特性,对重定向过程进行分段划分。随后,以放宽静稳定度飞行器为对象,对其快速转弯姿态控制问题进行马尔可夫决策过程建模,针对过失速转弯问题特性设计势能奖励函数。最后,训练和部署强化学习智能体,仿真验证该方法在提升转弯快速性、降低飞行器动能损失以及实现末端姿态角稳定控制上的有效性。

1 问题分析

1.1 放宽静稳定度飞行器气动特性

本文使用文献[18]中给出的一种圆柱形细长体鸭翼布局飞行器模型。该模型的直径为40 mm,细长比为18,鸭翼距离机体顶部396 mm,鸭式前翼布局呈“X”型,基于该模型的飞行器重定向建模为

$$\begin{aligned} m \frac{dV}{dt} &= P \cos \alpha - C_a \cos \alpha - C_n \sin \alpha - mg \sin \theta \\ mV \frac{d\theta}{dt} &= P \sin \alpha + C_n \cos \alpha - C_a \sin \alpha - mg \cos \theta \\ J_z \frac{d\omega_z}{dt} &= M_z \\ \frac{dx}{dt} &= V \cos \theta \\ \frac{dy}{dt} &= V \sin \theta \\ \frac{d\vartheta}{dt} &= \omega_z \end{aligned} \quad (1)$$

式中: V 为飞行器速度; C_n, C_a 分别为机体坐标系下飞行器所受法向力和轴向力; J_z 为飞行器对机体坐

标系 z 轴的转动惯量; M_z 为俯仰力矩; $\alpha, \vartheta, \theta$ 分别为攻角、俯仰角和轨迹倾角。利用 Jorgensen 工程方法可计算不同攻角范围($0^\circ \sim 180^\circ$)内飞行器的静态气动数据,具体计算公式参考文献[19]。

1.2 重定向过程

飞行器过失速翻转过程主要经历5个阶段,示意图如图1所示。

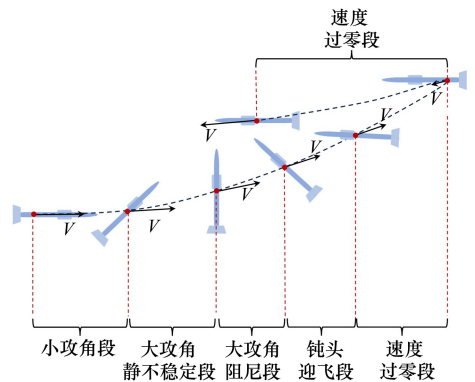


图1 飞行器过失速翻转的5个阶段

1) 小攻角段:俯仰力矩系数的正负由飞行器静不稳定度大小决定。若飞行器静不稳定度小,则无法实现翻转,需要预置舵偏为其提供额外的俯仰力矩^[9]。若静不稳定度较大,飞行器能够自翻转。

2) 大攻角静不稳定段:攻角增大导致气动中心前移,俯仰力矩系数变为正。因此飞行器在大攻角静不稳定阶段翻转速度会进一步加快。

3) 大攻角阻尼段:此时飞行器翻转姿态超过 90° ,空气舵完全失效,机体存在翻转阻尼,如果飞行器之前累积的转动惯性无法克服该阻尼,则无法完成过失速翻转。

4) 钝头迎头段:飞行器成功穿越超大攻角静稳定阶段,攻角达到 140° 以上,此时尾部迎风朝向前方,可等效为钝头“鸭式布局”,飞行器为纵向静不稳定,需施加空气舵偏控制使机体稳定指向期望方向来完成重定向。

5) 速度过零段:飞行器完成过失速重定向后还会经历速度衰减与反向动能重建,该过程被称为速度过零^[20]。由于速度过零段飞行器速度较低,难以利用空气舵施加有效控制,因此需要飞行器在钝头迎头段末端精确控制飞行器状态至指定的攻角和俯仰角速度。通过预留一定的攻角裕度和一个使攻角继续减小的翻转角速度,使飞行器在穿越速度过零段过程中依靠翻转惯性使自身的姿态运动状态继续

朝期望的状态点移动,这样可以解决飞行器在速度过零段由于自身静不稳定特性及空气舵失效导致的攻角发散问题。

2 智能控制模型

2.1 马尔可夫决策过程建模

马尔可夫决策过程(MDP)是描述强化学习的理论模型,本文将飞行器快速控制问题建模为连续空间、连续动作的有限马尔可夫决策过程,记为元组 $(\mathbf{S}, \mathbf{A}, P, r, \gamma)$ 。在快速转弯场景中,MDP智能体为飞行器的姿态控制器,负责学习控制飞行器姿态以及在实际的快速转弯过程中给出姿态控制指令。智能体所处环境为控制器以外的机体。姿态控制器通过选择舵偏大小作为动作来改变机体运动状态,即影响环境,环境状态发生变化后,会产生一个奖励信号,用于评判控制器当下动作的好坏程度,控制器收到奖励信号和环境状态改变的反馈后,会进一步做出下一步动作。智能姿态控制器的目标是在整个过程中最大化奖励信号的累积值,需通过人为设计奖励信号,向智能体传达想要实现的控制目标,即实现快速稳定的机体姿态翻转。

在元组 $(\mathbf{S}, \mathbf{A}, P, r, \gamma)$ 中,取俯仰角、俯仰角速度以及速度3个维度作为状态空间,有 $\mathbf{S} = [\vartheta, \dot{\vartheta}, V]$ 。注意,由于翻转过程中轨迹倾角变化极小,可近似认为攻角等于俯仰角。此外,为避免MDP状态维度过高,微分方程组中除 $\vartheta, \dot{\vartheta}, V$ 之外的其他变量均被视为仿真环境中用于数值计算的微分方程变量,而不是MDP建模中与智能体交互的环境变量。

\mathbf{A} 为动作空间,本文场景动作是一维舵偏角,故有 $\mathbf{A} = [\delta]$ 。环境状态转移函数 P 由强化学习数值仿真环境直接模拟给出,可视为黑盒模型。 $r: \mathbf{S} \rightarrow \mathcal{R}$ 为奖励信号函数; $\gamma \in [0, 1)$ 为奖励信号折扣因子。

与跟踪控制不同,这里智能体的控制目标是飞行器在一段尽可能短的时间段内达到期望的俯仰角和俯仰角速度。该目标作为一种终端约束,无法在控制过程中给予智能体姿态变化相关的奖励信号,引发奖励稀疏问题。本文设计一种基于势能的曲面奖励函数,通过提供密集的奖励信号,引导智能体向控制目标的方向优化动作策略,如图2所示。图2中红色圆点为曲面插值点,由手动调参得到,势能奖励函数曲面则由红色点进行插值得到。该曲面在靠

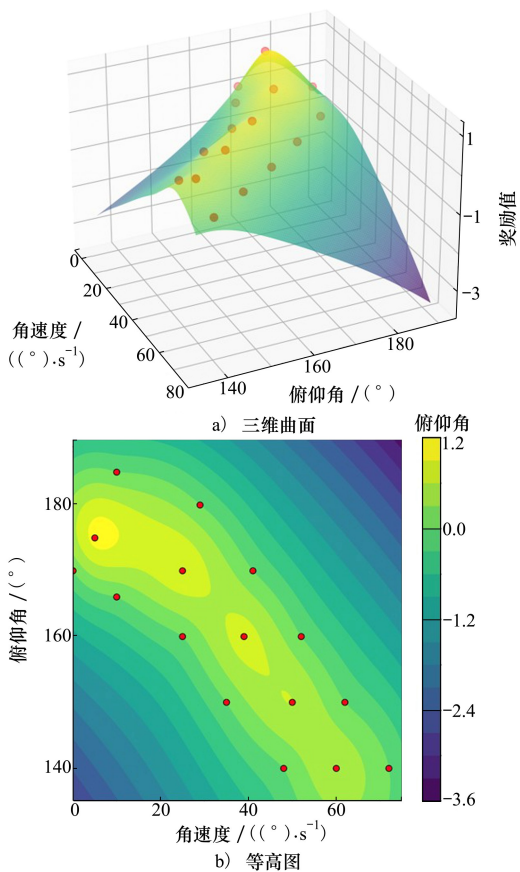


图2 奖励函数曲面

近终端目标状态处奖励势能最高且为正,在无法完成控制反转的状态区域势能为负,在翻转的初始状态处势能则接近0。强化学习控制器在训练过程中会受到势能的引导,向着势能更高的位置移动来积累奖励,从而实现控制目标。为增加控制的快速性,同时避免强化学习训练过程中可能出现的奖励漏洞攻击,本文在势能奖励的基础上设置了时间惩罚项,因此总的奖励函数可以表示为

$$r(t) = r_{\text{potential}}(s_{t+1}) - r_{\text{time}} \quad (2)$$

式中: $r_{\text{potential}}(s): s \rightarrow \mathcal{R}$ 为图2所示势能奖励函数; $r_{\text{time}} \in \mathcal{R}^+$ 为常数项。

2.2 近端策略优化

强化学习中状态价值函数 v_{π} 衡量当前策略 π 下状态 s 的未来预期奖励累积值,定义为

$$v_{\pi}(s) \doteq E_{\pi} \left[\sum_{k=0}^{\infty} \gamma^k R_{t+k+1} \mid S_t = s \right], \forall s \in \mathbf{S} \quad (3)$$

动作价值函数 q_{π} 代表策略 π 在状态 s 时采取动作 a 后所有可能的决策序列的期望奖励累积,定义为

$$q_{\pi}(s, a) \doteq \mathbb{E}_{\pi} \left[\sum_{k=0}^{\infty} \gamma^k R_{t+k+1} \mid S_t = s, A_t = a \right] \quad (4)$$

传统策略梯度算法通过梯度上升来优化策略时,可能会因为步长过大而导致策略性能突然下降,从而影响训练效果。PPO 策略梯度算法基于 Actor-Critic 框架,通过截断来限制策略更新的幅度从而优化了梯度更新过程^[21]。PPO 中存在新旧 2 个策略: $\pi_{\gamma_{\text{old}}}$ 和 π_{λ} 。 $\pi_{\gamma_{\text{old}}}$ 为已有策略, π_{γ} 是待学习策略。PPO 算法利用新旧策略动作概率分布之间的 KL 散度来限制梯度更新的步长,其优化目标为

$$\lambda = \underset{\lambda}{\operatorname{argmax}} \mathbb{E}_{\tau \sim \pi_{\text{old}}} \left[\frac{\pi_{\lambda}(a_t | s_t)}{\pi_{\lambda_{\text{old}}}(a_t | s_t)} A_{\pi_{\text{old}}}(s_t, a_t) \right],$$

s.t. $D_{\text{KL}}(\pi_{\lambda} \parallel \pi_{\lambda_{\text{old}}}), \Delta$ (5)

式中: Δ 为 KL 散度约束; $r_{\lambda} = \frac{\pi_{\lambda}(a_t | s_t)}{\pi_{\lambda_{\text{old}}}(a_t | s_t)}$ 是重要性采样权重; $D_{\text{KL}}(\pi_{\lambda} \parallel \pi_{\lambda_{\text{old}}})$ 为策略间的 KL 散度^[23]; $A_{\pi_{\text{old}}}(s_t, a_t)$ 为策略的优势函数,有

$$A_{\pi}(s, a) = Q_{\pi}(s, a) - V_{\pi}(s) \quad (6)$$

$A_{\pi_{\text{old}}}(s_t, a_t)$ 需要通过神经网络来估计,记为 A_{ϕ} ,可以使用 Critic 网络通过估计 V_{ϕ} 得到 A_{ϕ} ,网络通过时序差分来更新,即

$$V_{\phi}(s_t) = r_t + \gamma V_{\phi}(s_{t+1}) \quad (7)$$

记(5)式为 $L = \mathbb{E}[r_t A_{\phi}]$,PPO 提出裁剪优化目标为

$$L^{\text{clip}} = \mathbb{E}[\min(r_t A_{\phi}, \text{clip}(r_t, 1 - \varepsilon, 1 + \varepsilon) A_{\phi})] \quad (8)$$

(8)式中含有 2 项,一个是原始的优化目标,另一个则是裁剪的优化目标,PPO 从两者中取极小值用于策略网络参数的迭代。

3 智能体训练方法

3.1 智能控制训练框架

智能体在每个时间步长更新前从强化学习环境中获取飞行器当前的运动状态 s_t ,进而通过采样 Actor 策略网络输出的动作概率分布得到 a_t ,即空气舵偏控制指令。 a_t 产生的力矩作用于机体,使其俯仰角速度改变。基于四阶龙格库塔积分器,强化学习环境,利用(1)式对飞行器的运动状态进行一个时间步长的更新,从而得到 s_{t+1} ,同时环境给出状态转移产生的奖励信号 r_t 。智能体在每一步更新中都会产生一组状态转移数据 (s_t, a_t, r_t, s_{t+1}) 。所有数据会被存储在经验回放池 D 中,用于训练更新 Actor

网络和 Critic 网络的参数。快速转弯智能控制器的训练过程总结如下。

算法 1 基于 PPO 的快速转弯控制策略训练

输入: ϑ_{exp} —— 期望俯仰角; ω —— 期望俯仰角速度

输出: λ^* —— 最优动作网络参数

初始化策略网络参数和价值网络参数

for each episode

重置训练环境: $\text{env.reset}()$

初始化经验池 D

对每个时间步长 $t_k (k = 1, 2, \dots)$ 执行

由当前状态 s_t , 策略网络 π_{λ} 采样得到动作 a_t

由 a_t 更新当前状态,得到 s_{t+1}

计算奖励信号 r_t

if 终止条件为真

当前回合结束

else

将 $s_t, a_t, s_{t+1}, r_t, \text{done}$ 存入经验池 D 中

if 经验池存满

for $j = 1, 2, \dots, |D|$

计算 $\pi_{\lambda_{\text{old}}}(s_j, a_j)$ 和期望价值 $V_i(s_i)$

通过最小化价值损失更新价值网络参数:

$$\underset{\phi}{\operatorname{argmin}} \frac{1}{|D|} \sum_{j=1}^{|D|} (V_i s_j - V_{\phi} s_j)^2$$

计算新策略采样概率 $\pi_{\lambda}(s_j, a_j)$, 得到新

旧策略概率比 $r_{\lambda} = \frac{\pi_{\lambda}(a_t | s_t)}{\pi_{\lambda_{\text{old}}}(a_t | s_t)}$

更新策略网络参数:

$$\underset{\lambda}{\operatorname{argmax}} \frac{1}{|D|} \sum_{j=1}^{|D|} \min(r_{\lambda} A_{\phi}, \text{clip}(r_{\lambda}, 1 - \varepsilon,$$

$1 + \varepsilon) A_{\phi})$

清空经验池 D

$\lambda^* \leftarrow \lambda$

3.2 训练设置

Actor 和 Critic 网络均采用 4 层全连接神经网络,每个网络包含 2 个隐藏层,隐藏层激活函数选用 ReLU。Actor 网络利用 Tanh 输出策略分布的均值, Sigmoid 输出策略分布的方差。Critic 网络由 ReLU 输出当前状态的价值。网络结构设置如表 1 所示。

训练参数与初值设置如表 2 所示。

表 1 网络结构设置

神经网络	结构	激活函数
Actor	(2, 128)	ReLU
	(128, 256)	ReLU
	(256, 128)	ReLU
	(128, 1)	Tanh
Critic	(2, 128)	ReLU
	(128, 256)	ReLU
	(256, 256)	ReLU
	(256, 128)	ReLU
	(128, 64)	ReLU
	(64, 1)	Tanh

表 2 训练参数与初值设置

参数	取值
飞行器速度 $V/(m \cdot s^{-1})$	200
飞行器初始角度 $\vartheta, \theta/(^{\circ})$	0, 0
飞行器初始角速度 $\omega/(^{\circ} \cdot s^{-1})$	0
经验池大小	10 000
学习率 $\alpha_{\lambda}, \alpha_{\phi}$	0.000 1, 0.000 1
奖励折扣因子 γ	0.99
折扣因子 γ_{GAE}	0.95
裁剪率 ϵ	0.2
更新时间步长 dt/s	0.01
训练回合数	2 000

4 训练与仿真分析

4.1 无控自翻转仿真结果

利用 1.1 节给出的几组不同静稳定度飞行器的气动数据,针对不同静稳定度以及不同预置舵偏下飞行器的自翻转能力进行仿真分析。

图 3 展示了静稳定、静中立稳定和静不稳定 3 类飞行器在主发动机未点火时,无舵偏控制自翻转过程中轨迹及姿态角的变化情况。图 3 俯仰角与角速度变化曲线中的紫线代表静稳定飞行器俯仰角和角速度随时间的变化情况,可以看出该类飞行器具备静稳定性,姿态平衡被破坏后会恢复到初始平衡状态。由黄色曲线可知静中立稳定飞行器则具备一定的自翻转能力,飞行器翻转时可穿越 1.2 节中提到的小攻角段和大攻角静不稳定段。但由于其静不稳定度较小,飞行器进入超大攻角静稳定段时所累积的转动能量不足以克服该阶段的静稳定性,因此俯仰角速度会降为负值,俯仰角在达到 140° 左右后开始减小,飞行器无法完成自翻转。对于静不稳定飞行器,如图 3 中红色线所示,在穿越小攻角段和大攻角静不稳定段时能够累积足够的转动能量,因此可以穿越超大攻角静稳定段,完成整个自翻转过程。但可以看到其俯仰角在转弯结束时趋向于发散。

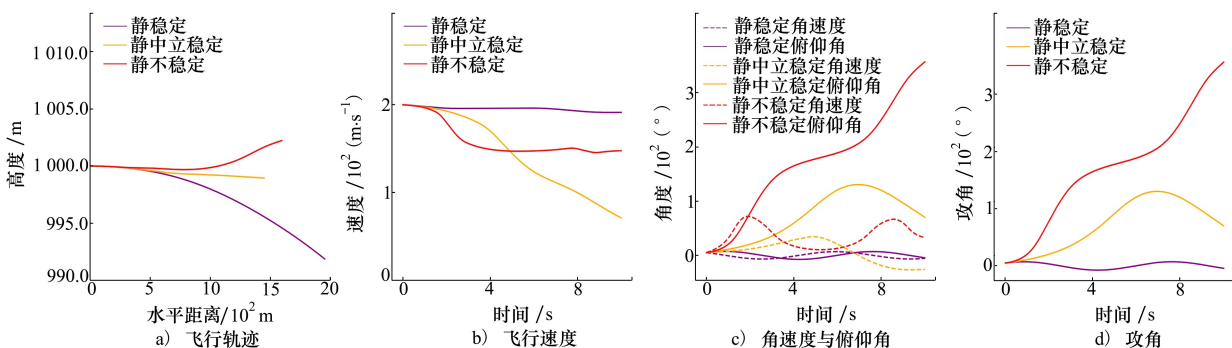


图 3 飞行器自翻转运动状态

预置舵偏可提供额外的俯仰力矩使飞行器抬头,从而加快翻转过程。图 4 显示了不同预置舵偏角对不同静稳定度飞行器翻转特性的影响。由图 4a)可知,静稳定飞行器,但未能穿越小攻角静稳定段,表明即便有往外翻转力矩也无法完成快速转弯。

观察图 4b)可知,静中立稳定飞行器在预置舵偏的影响下,其最大的翻转幅度可达到 160° 。而图 4c)则表明,静不稳定度过大时飞行器在无控制状态下俯仰角会发散。

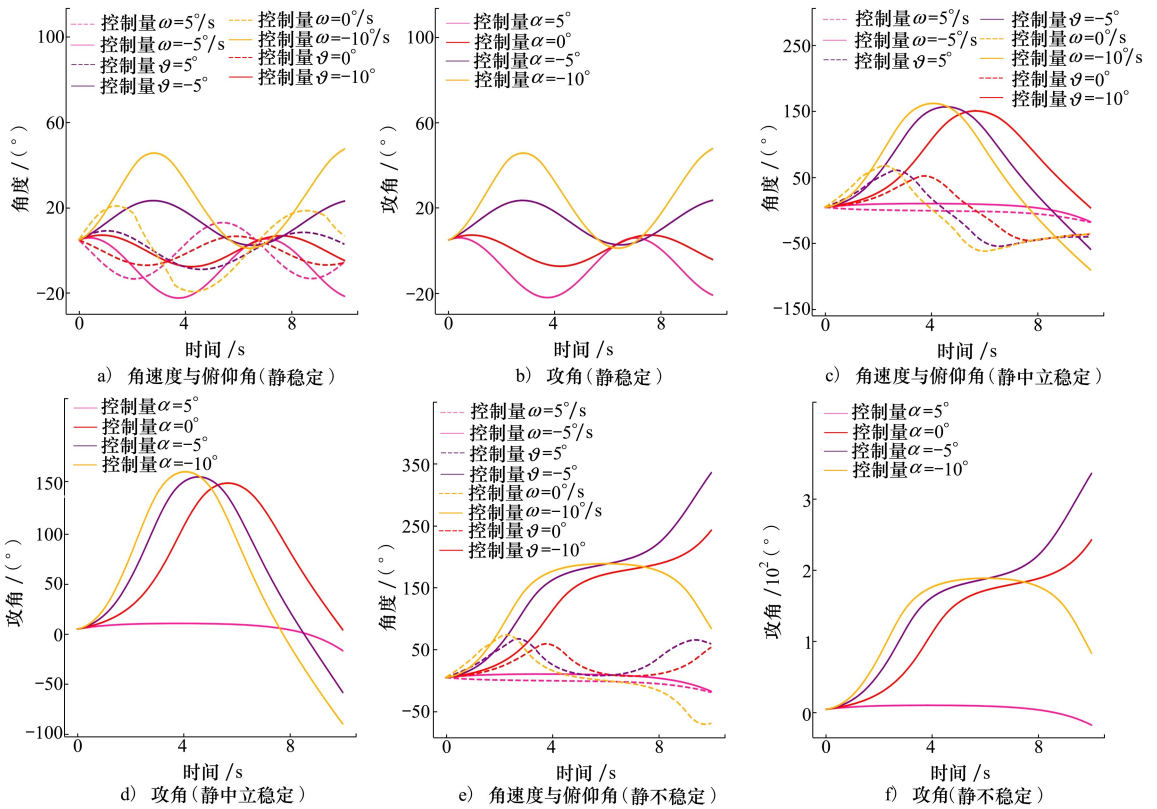


图 4 不同预置舵偏下飞行器翻转特性

4.2 智能体训练

基于 4.1 节分析可知,静不稳定飞行器具备自翻转能力,但翻转末端飞行器姿态不稳定。此外,小攻角阶段空气舵的作用是加快翻转速度,无需对飞行器姿态进行控制。因此,本文以静不稳定飞行器为控制对象,在无控制翻转至攻角 140° 时开始进行智能姿态控制。智能体在训练阶段尝试控制飞行器翻转至 175° 俯仰角,并为速度过零阶段预留 $5^\circ/\text{s}$ 俯仰角速度,该阶段飞行器主发动机仍未点火,仅依靠惯性向前滑行。因此,强化学习训练环境的初始状态为飞行器无控翻转至攻角 140° 时的运动状态,每回合训练终止的条件为俯仰角达到 175° ,训练目标是使得终止时刻的俯仰角速度尽可能接近 $5^\circ/\text{s}$ 。

图 5 记录了训练过程中飞行器在俯仰角和俯仰角速度组成的相平面中状态轨迹的变化情况,由淡蓝色实线表示。翻转结束时的期望终端状态则由红色圆点表示 ($\omega = 5^\circ/\text{s}, \vartheta = 175^\circ$)。智能体每训练 1 000 回合绘制 1 幅由多条轨迹组成的轨迹簇图。随着训练的进行,轨迹簇中轨迹线的末端越来越接近终端期望状态点,说明控制策略在训练的过程中

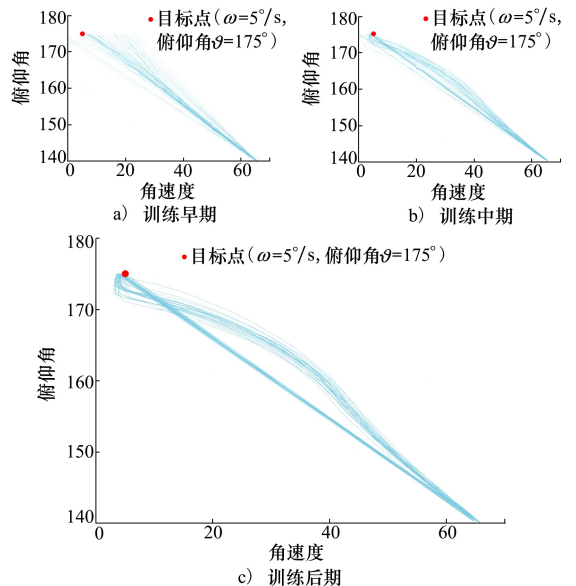


图 5 相平面状态轨迹

逐渐收敛,实现了对飞行器快速转弯末端姿态角与姿态角速度的精确控制。观察图 5c) 中的轨迹线可以看到,智能体在穿越相平面的过程中选择尽量维

持较高的平均俯仰角速度,仅在接近终端点时才开始加大减速力度。因此轨迹线在 $10^\circ/\text{s}$ 至 $50^\circ/\text{s}$ 区间呈现凸起状。而在终端点附近,智能体为了尽量接近终端状态,选择快速减小减速力度,使得俯仰角快速收敛到期望值,故轨迹线末段呈现凹陷状。由此可见,智能体成功地学习到了一种快速转弯策略。

图 6 为奖励学习曲线,深蓝色色线为滑动平均值。其收敛情况与图 5 中轨迹簇的收敛情况一致。

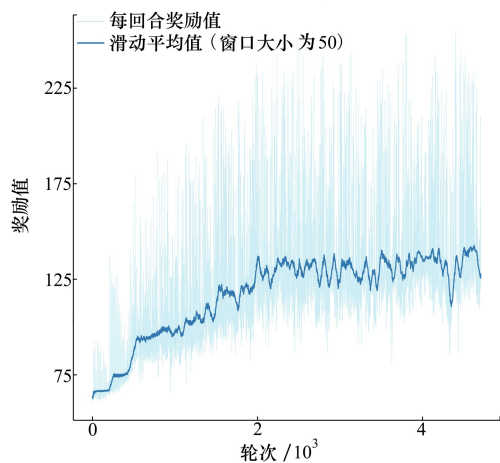


图 6 奖励学习曲线

4.3 快速转弯全过程仿真

转弯过程共分为预置舵偏翻转、无舵效自翻转、智能姿态控制和速度过零这 4 个阶段。为最大程度上节省主发动机推进能量,飞行器在前 3 个阶段依靠惯性向前飞行并完成超机动重定向,仅在进入速度过零段时,主发动机才开始点火,在轨迹倾角发生改变前迅速完成飞行器的反向动能建立。

仿真初始飞行器运动状态变量设置如下:高度 1 000 m,初速度 300 m/s,轨迹水平,预置舵偏角为 -10° 。仿真终止条件是飞行器在反方向完成动能重建,飞行器速度达到 250 m/s。仿真假定飞行器在无舵效自翻转段和速度过零段空气舵完全失效。

图 7 分别显示单次快速转弯仿真中飞行器的轨迹和姿态信息。从图 7a) 可以看到,飞行器在转弯全过程中轨迹保持水平,飞行器穿越速度过零段时以 180° 锐角转弯。飞行器动能重建后攻角收敛至 5° 以内。因而可以认为飞行器在惯性飞行阶段成功实现了超机动重定向,并在反向速度重建过程中维持了姿态的稳定,从而成功完成了快速转弯。以往方法快速转弯半径在 100 m 以上^[14],而本方法通过过失速机动实现姿态掉转,且延迟开启主推力发动

机,这不仅节省了推进能量,而且在后续中制导阶段无需调整轨迹,可迅速以最佳速度与角度指向后半球目标,飞行器后向转弯的快速性得到大幅度增强。

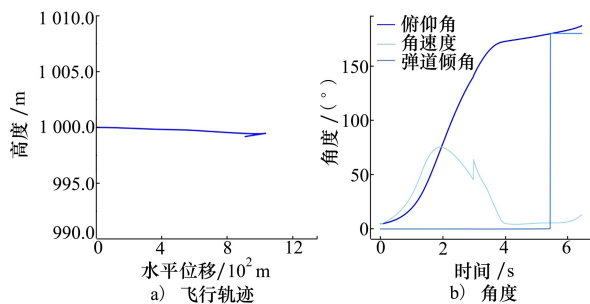


图 7 飞行器仿真轨迹与姿态信息

在随机初始条件下进行了 20 次的快速转弯全过程仿真,仿真初始时刻的俯仰角与俯仰角速度分别取 0° 至 5° 和 $0^\circ/\text{s}$ 至 $5^\circ/\text{s}$ 内的随机值。结果如图 8 所示,可以看到智能控制器在一定的投放范围内具备良好的鲁棒性。

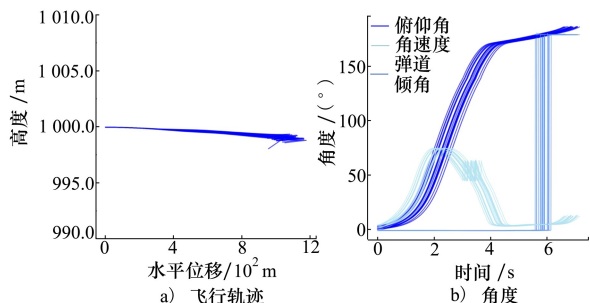


图 8 快速转弯蒙特卡洛仿真结果

5 结 论

本文针对载机后半球目标快速追踪问题提出了一种基于近端策略优化(PPO)的智能控制方法,实现了放宽静稳定度飞行器的过失速机动重定向。研究表明,通过放宽飞行器静稳定度,结合空气舵的纯气动控制,能够有效克服传统方法对额外执行机构的依赖以及复杂气动特性下的建模难题。基于 Jorgensen 工程方法的气动特性分析表明,静不稳定度是飞行器自翻转能力的决定性因素,其重定向过程可划分为小攻角段、大攻角静不稳定段、大攻角阻尼段、钝头迎飞段及速度过零段。通过马尔可夫决策过程建模与势能奖励函数设计,强化学习智能体成功实现了对飞行器末端俯仰角(误差 $<1^\circ$)和角速度(误差 $<0.5^\circ/\text{s}$)的精确控制。该方法在低马赫数下

可驱动飞行器完成 180° 快速翻转, 动能损失降低 40% 以上, 且蒙特卡洛实验验证了其在随机初始条件下的鲁棒性。此外, 无转弯半径的特性可显著提

升飞行器的快速性与能量效率, 为后半球目标追踪技术的发展提供了新思路。

参考文献:

- [1] 樊会涛, 闫俊. 空战体系的演变及发展趋势[J]. 航空学报, 2022, 43(10): 296-305.
Fan Huitao, Yan Jun. Evolution and development trend of air combat system[J]. Acta Aeronautica et Astronautica Sinica, 2022, 43(10): 296-305. (in Chinese)
- [2] Zhang Ke, Zhang Zhenlin, Han Zhiguo. Optimal guidance law design for three-dimensional trajectory tracking of missiles based on fixed-gain disturbance observer[J]. ISA Transactions, 2023, 143: 298-312.
- [3] 谢岚凤, 陈军, 焦璐, 等. 未来空战全域火力场研究[J]. 航空学报, 2024, 45(5): 296-313.
Xie Lanfeng, Chen Jun, Jiao Lu, et al. All-domain fire field in future air combat[J]. Acta Aeronautica et Astronautica Sinica, 2024, 45(5): 296-313. (in Chinese)
- [4] Wu Gang, Zhang Ke. A novel fixed-time convergence three-dimensional guidance law for intercepting highly maneuvering targets[J]. International Journal of Aeronautical and Space Sciences, 2023, 23: 595-608.
- [5] 王方剑, 王宏伟, 李晓辉, 等. 细长旋成体亚声速超大攻角非定常流动特性研究[J]. 力学学报, 2022, 54(2): 379-395.
Wang Fangjian, Wang Hongwei, Li Xiaohui, et al. Subsonic unsteady aerodynamic characteristics on slender revolutionary body at extra-wide angle-of-attack[J]. Chinese Journal of Theoretical and Applied Mechanics, 2022, 54(2): 379-395. (in Chinese)
- [6] 李斌, 郭正玉. 基于放宽静稳定度空空飞行器重定向研究[J]. 航空兵器, 2023, 30(6): 37-43.
Li Bin, Guo Zhengyu. Study on heading reversal of air-to-air missile based on relaxed static stability[J]. Aero Weaponry, 2023, 30(6): 37-43. (in Chinese)
- [7] 关为群, 殷兴良. 飞行器大攻角高机动飞行特性分析与仿真[J]. 现代防御技术, 2006(4): 43-47.
Guan Weiqun, Yin Xingliang. Analysis and simulation of missile's high maneuver flying at high angles of attack[J]. Modern Defence Technology, 2006(4): 43-47. (in Chinese)
- [8] 孔轶男, 伍彬, 汪清, 等. 过失速重新定向机动过程气动特性建模[J]. 空气动力学学报, 2024, 42(5): 72-80.
Kong Yinan, Wu Bin, Wang Qing, et al. Aerodynamic characteristics modeling of post-stall re-orientation maneuver[J]. Acta Aerodynamica Sinica, 2024, 42(5): 72-80. (in Chinese)
- [9] 高昌昊, 宋文萍, 韩少强, 等. 空空飞行器过失速重新定向技术研究[J]. 空天防御, 2022, 5(3): 17-26.
Gao Changhao, Song Wenping, Han Shaoqiang, et al. Research on poststall re-orientation technology of air-to-air missile[J]. Air & Space Defense, 2022, 5(3): 17-26. (in Chinese)
- [10] 彭继平, 霍鑫, 杨宝庆, 等. 不同攻角约束下空空飞行器快速转弯过程轨迹优化[J]. 航空兵器, 2017, 24(1): 39-44.
Peng Jiping, Huo Xin, Yang Baoqing, et al. Trajectory optimization for air-to-air missile's agile turn with different angle of attack constraints[J]. Aero Weaponry, 2017, 24(1): 39-44. (in Chinese)
- [11] 韦亚利, 殷玮, 刘印田, 等. 多约束下的快速飞行器大离轴转弯规律研究[J]. 空天防御, 2020, 3(3): 96-102.
Wei Yali, Yin Wei, Liu Yintian, et al. Research on big off-axis turning law under multiple constraints for agile missile[J]. Air & Space Defense, 2020, 3(3): 96-102. (in Chinese)
- [12] Thukral A, Innocenti M. A sliding mode missile pitch autopilot synthesis for high angle of attack maneuvering[J]. IEEE Trans on Control Systems Technology, 1998, 6(3): 359-371.
- [13] 李政, 于剑桥, 赵新运. 空空飞行器快速转弯固定时间收敛滑模控制[J]. 航空学报, 2023, 44(8): 208-221.
Li Zheng, Yu Jianqiao, Zhao Xinyun. Fixed-time convergent sliding mode control for agile turn of air-to-air missiles[J]. Acta Aeronautica et Astronautica Sinica, 2023, 44(8): 208-221. (in Chinese)
- [14] 赵新运, 于剑桥. 飞行器快速转弯段的新型非奇异终端滑模控制[J]. 宇航学报, 2022, 43(4): 454-464.
Zhao Xinyun, Yu Jianqiao. Novel non-singular terminal sliding mode control for missile's agile turn[J]. Journal of Astronautics, 2022, 43(4): 454-464. (in Chinese)
- [15] 黄旭, 柳嘉润, 贾晨辉, 等. 深度确定性策略梯度算法用于无人飞行器控制[J]. 航空学报, 2021, 42(11): 404-414.
Huang Xu, Liu Jiarun, Jia Chenhui, et al. Deep deterministic policy gradient algorithm for UAV control[J]. Acta Aeronautica et Astronautica Sinica, 2021, 42(11): 404-414. (in Chinese)

- [16] 黄旭, 柳嘉润, 张远, 等. 知识与数据混合驱动的高速飞行控制方法综述[J]. 宇航学报, 2023, 44(8): 1113-1126.
Huang Xu, Liu Jiarun, Zhang Yuan, et al. Review on knowledge-based and data-driven cooperating control methods of high-speed vehicle[J]. Journal of Astronautics, 2023, 44(8): 1113-1126. (in Chinese)
- [17] 裴培, 何绍溟, 王江, 等. 一种深度强化学习制导控制一体化算法[J]. 宇航学报, 2021, 42(10): 1293-1304.
Pei Pei, He Shaoming, Wang Jiang, et al. Integrated guidance and control for missile using deep reinforcement learning[J]. Journal of Astronautics, 2021, 42(10): 1293-1304. (in Chinese)
- [18] Wang Fangjian, Liu Jin, Qin Han, et al. Unsteady aerodynamic characteristics of slender body at extra-wide angle-of-attack range[J]. Aerospace Science and Technology, 2021, 110: 106477.
- [19] Jorgensen L H. Prediction of static aerodynamic characteristics for space-shuttle-like and other bodies at angles of attack from 0° to 180°[R]. NASA TN-D-6996, 1973.
- [20] 张华明, 姜玉宪. 后射空空飞行器稳定性问题研究[J]. 北京航空航天大学学报, 2004, 30(6): 572-576.
Zhang Huaming, Jiang Yuxian. Study on stability of air-to-air missile in firing-back attack[J]. Journal of Beijing University of Aeronautics and Astronautics, 2004, 30(6): 572-576. (in Chinese)
- [21] 先苏杰, 王康, 曾鑫, 等. 有限视场角下基于深度强化学习的落角约束制导律[J]. 兵工学报, 2024, 46(4): 1-14.
Xian Sujie, Wang Kang, Zeng Xin, et al. Guidance law with impact angle constraints based on deep reinforcement learning under limited field of view[J]. Acta Armamentarii, 2024, 46(4): 1-14. (in Chinese)
- [22] Schulman J, Levine S, Moritz P, et al. Trust region policy optimization [C] // International Conference on Machine Learning, 2015.

Stalled redirection control of relaxed static stability aircraft

Cui Hao, Zhang Ke, Huang Xuchen, Li Haoyu, Tan Minghu, Wang Jingyu

(School of Astronautics, Northwestern Polytechnical University, Xi'an 710072, China)

Abstract: Traditional aircraft face limitations in off-boresight maneuverability, making it challenging to effectively meet rear hemisphere target tracking requirements. This paper proposes a proximal policy optimization (PPO)-based intelligent control method to achieve aerodynamic-controlled post-stall maneuver redirection for relaxed static stability aircraft. First, the nonlinear aerodynamic characteristics of the aircraft at extremely high angles of attack are analyzed using the Jorgensen engineering method, revealing the impact of static instability on autorotation capability and segmenting the redirection process. Second, a Markov decision process (MDP) model is established to address nonlinear attitude control under ultra-high angles of attack, incorporating a potential energy-based reward function to guide the intelligent agent in learning optimal rudder deflection strategies. Finally, redirection simulations under low Mach number conditions demonstrate that aerodynamic control surfaces alone can enable rapid aircraft redirection while maintaining attitude stability during velocity zero-crossing phases. Monte Carlo experiments further validate the robustness of this control method under random initial conditions.

Keywords: stalled redirection; relaxed static stability; reinforcement learning; intelligent control

引用格式: 崔颖, 张科, 黄煦晨, 等. 放宽静稳定度飞行器过失速重定向控制[J]. 西北工业大学学报, 2026, 44(1): 151-159.
Cui Hao, Zhang Ke, Huang Xuchen, et al. Stalled redirection control of relaxed static stability aircraft[J]. Journal of Northwestern Polytechnical University, 2026, 44(1): 151-159. (in Chinese)