基于深度强化学习的变形滑翔飞行器智能变形决策方法

王滔鋆 孟凡一 陈刚
（西安交通大学航天航空学院，710049，西安）

摘要：为了充分发挥变形滑翔飞行器（MGV）相较于传统固定构型飞行器在航程、速度及环境适应性方面的显著优势，解决多物理场耦合下的最优变形构型决策难题，本文提出了一种基于深度强化学习的智能变形决策方法。首先，利用深度神经网络构建翼面变形与气动性能的代理模型，实现了从变形动作到气动数据的端到端快速映射；然后，结合飞行器纵向运动模型建立强化学习框架，并通过算法性能对比，选用近端策略优化（PPO）算法构建智能决策模型，实现了气动/变形/弹道交叉耦合下的自主策略学习；最后，在训练范围外的拉偏环境中进行全流程仿真实验验证。结果表明：该方法能够综合气动/变形/弹道耦合影响进行智能变形决策，使MGV终端射程增加了7.304%，并能有效维持较高的升阻比飞行状态；同时，拉偏测试证明了该方法在不确定环境下依然具备良好的可行性与泛化能力。

关键词：变形滑翔飞行器；智能变形决策；深度强化学习
中图分类号：V448 **文献标志码：**AA
文章编号：0253－987X(XXXX)XX－0001－18

Intelligent Deformation Decision-making Method for Morphing Glide Vehicle with Deep Reinforcement Learning

WANG Taojun,MENG Fanyi,CHEN Gang
(School of Aerospace Engineering, Xi'an Jiaotong University, Xi'an, Shaanxi 710049, China)

Abstract: To fully exploit the significant advantages of Morphing Gliding Vehicles (MGVs) over traditional fixed-configuration vehicles in terms of range, speed, and environmental adaptability, and to address the challenge of optimal configuration decision-making under multi-physics coupling, this paper proposes an intelligent morphing decision-making method based on deep reinforcement learning. Firstly, a deep neural network is employed to construct a surrogate model for wing deformation and aerodynamic performance, facilitating a rapid end-to-end mapping from morphing actions to aerodynamic data. Secondly, a decision-making framework is established by combining the longitudinal motion model with reinforcement learning. Through comparative analysis of algorithm performance, the Proximal Policy Optimization (PPO) algorithm is selected to construct the intelligent model, enabling autonomous strategy learning under the cross-coupling of aerodynamics, morphing, and

收稿日期：2025-06-05。

trajectory. Finally, full-process simulation experiments are conducted in deviated environments outside the training envelope. The results indicate that the proposed method achieves intelligent decision-making by comprehensively considering the cross-coupling of aerodynamics, morphing, and trajectory, increasing the terminal range of the MGV by 7.304% , while effectively maintaining a high lift-to-drag ratio. Furthermore, deviation tests confirm the method's feasibility and generalization capability in adapting to uncertain environmental challenges.

Keywords: Morphing Glide Vehicle; Intelligent deformation decision-making; Deep reinforcement learning

随着科学技术的发展和飞行任务需求的日趋复杂化，固定构型飞行器有限的任务执行区间和执行能力，严重制约了飞行器生存能力和实际飞行效能的提升，因此亟需发展先进的气动布局及控制方式。与传统飞行器不同，变形滑翔飞行器(MGV)具有灵活、连续、多自由度及大幅度变形的能力[1-5]，可以根据飞行指令改变自身的气动外形和结构，主动修改自身的气动特性，能够适应更广泛的飞行范围和速域[2-3]。通过控制机翼变形，可以保持变形飞行器时刻处于最大升阻比状态，减少飞行能量损失，在各种环境和条件下实现优越的飞行性能和宽广的飞行包线。

对于变形滑翔飞行器来说，变形不仅改变了气动性能的变化，同时对飞行弹道也产生了较大的影响。由于翼面变形引起气动性能的显著变化，变形滑翔飞行器的运动状态难以用精确的解析动力学模型来描述 [6-7]。近年来有关变形飞行器变形策略方面的研究主要集中于如何通过变形来提高飞行器的气动性能，尚未考虑其翼面变形与弹道性能之间的耦合影响。Chen 等[8]针对变形飞行器设计沿指定飞行轨迹的变形策略，以保持良好的升阻力特性。Yue 等[9-10]采用滑模控制方法设计了跟踪机动参考命令的控制器，实现了无尾伸缩机翼变形飞机的横向稳定控制。Gong 等[11]将变形作为系统状态的函数，研究了一种基于Q-learning 的变后掠翼飞行器高度运动切换控制策略。Bao 等[12]基于自适应块动态曲面反演方法设计了变形滑翔飞行器的制导、控制和变形耦合的控制系统。然而，该控制器的设计依赖于一个精准的被控对象模型，只适用于特定类型的变形滑翔飞行器。

近年来，强化学习(RL)得到了迅速发展[13-15]，其基于动作-评价网络选择动作，并通过动作执行后获得的奖励更新网络参数[16]。目前，深度强化学习(DRL)算法在智能控制和决策的各个领域有着广泛的应用[17-19]，为实现变形飞行器的智能变形决策提供了新的思路与途径。Lampton 等[20-21]基于Q-learning 算法设计了奖励函数，结合升力和阻力系数对变形飞行器的机翼厚度和后掠角度进行优化设计。Tanadel 等[22-23]采用简化的椭球变形飞行器，通过Q-learning 算

法获得适合整个飞行过程的变形策略。然而，Q-learning算法的动作空间是离散的，无法实现变形飞行器的连续变形。Xu[24]和Vinicius[25]采用了深度确定性策略梯度(DDPG)算法来控制机翼的连续自主变形。这种方法允许对称和非对称变形，提高了飞行性能并验证了训练模型的鲁棒性。然而，现有研究主要集中在采用强化学习算法实现变形控制的可行性上，将翼面变形和飞行弹道视为相互独立的过程，忽略了翼面变形和飞行弹道之间的耦合关系。

本文重点研究如何在气动/变形/弹道交叉耦合的情况下实现变形滑翔飞行器的智能变形决策，提出适用于各类变形飞行器智能变形决策的通用方法。以高速变形滑翔飞行器为例，研究了无动力滑翔弹道下的智能变形决策。通过神经网络技术构建变形滑翔飞行器气动代理模型，用以快速获取不同飞行条件下各翼面变形构型的气动数据。在此基础上，提出了一种基于深度强化学习算法的变形决策方法，实现了变形飞行器在无动力滑翔过程中的自主变形，消除了对精确被控对象模型的依赖，充分发挥变形滑翔飞行器的潜在气动性能，提高其滑翔阶段的弹道性能，为变形飞行器智能变形决策提供了新的通用解决方法。

1 变形滑翔飞行器弹道模型

1.1 变形滑翔飞行器外观参数

本文所讨论的研究模型是带变形翼的变形滑翔飞行器。轮廓和变形模式如图1所示。变形滑翔飞行器由三部分组成：机身、两个可变机翼和两个不变尾翼。机翼的变形在机身两侧同步进行变化。变形滑翔飞行器相关参数的定义和取值如表1所示。

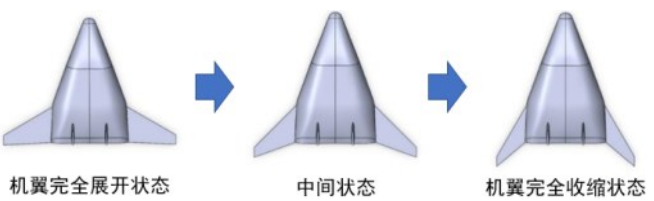


图1 变形滑翔飞行器的外形轮廓和变形模式

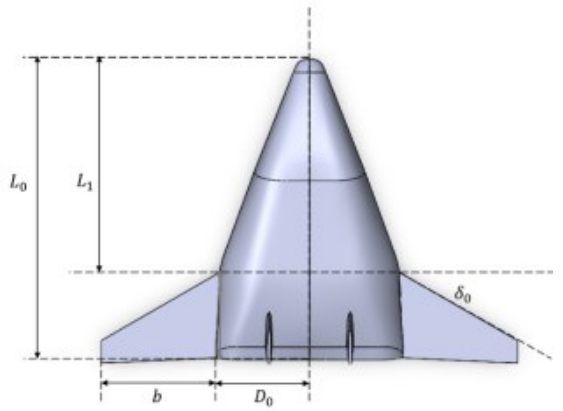
Fig. 1 Outline and morphing modes of MGV

其中，δ_0表示变形滑翔飞行器机翼完全收缩时的最小变形角度；δ_1表示机翼完全展开时的最大变形角度；L_0表示变形滑翔飞行器头部尖端到尾部的纵向参考长度；L_1表示头部尖端到机翼前端的纵向长度；b表示变形飞行器的翼展长度；D_0表示尾部横向参考长度。

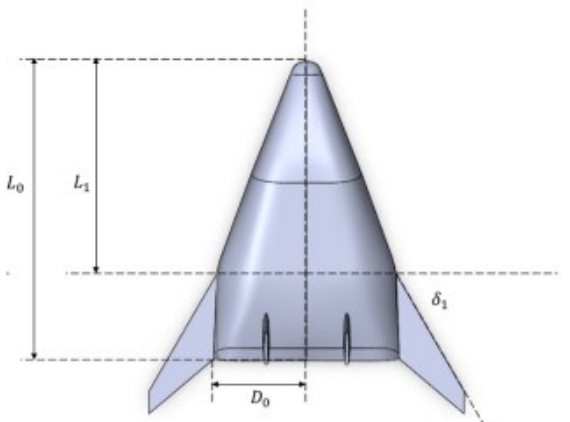
1.2 变形滑翔飞行器运动建模

由于变形滑翔飞行器翼面变形通常为对称变形，只在纵向飞行平面内通过升阻力变化对运动进行控制，并不影响

横侧向平面的运动变化。为简化模型复杂度，本研究选择变形滑翔飞行器的纵向飞行平面开展方法研究。变形滑翔飞行器在无动力滑翔阶段的纵向运动方程为



（a） 机翼完全展开状态下的外观参数



（b） 机翼完全收缩状态下的外观参数

图2 变形滑翔飞行器外形参数

Fig. 2 Shape parameters of MGV

表 1 变形滑翔飞行器相关参数的取值

Table 1 The definitions and values of MGV characteristic parameters

参数	数值
质量/kg	2500
最小变形角度 δ_0/°	30
最大变形角度 δ_1/°	60
纵向参考长度 L_0/m	1.2
头部到机翼的长度 L_1/m	0.8
翼展长度 b/m	0.5
尾部横向参考长度 D_0/m	0.35

$$\begin{cases} \dfrac{dV}{dt} = -\dfrac{D}{m} - g\sin\gamma \\ \dfrac{d\gamma}{dt} = \dfrac{L}{mV} + \dfrac{1}{V}\left(\dfrac{V^2}{r} - g\right)\cos\gamma \\ \dfrac{dr}{dt} = V\sin\gamma \\ \dfrac{dX}{dt} = R_e\dfrac{V}{r}\cos\gamma \\ r = r_0 + H \\ \varepsilon(\alpha,\delta) = 0 \end{cases} \tag{1}$$

式中：V、γ、r、H、X、α、δ分别为速度、弹道倾角、地心距离、飞行高度、航程、攻角、翼面变形量；$\varepsilon(\alpha,\delta)=0$是攻角和机翼变形的控制方程；$g$ 和 R_e 分别表示重力加速度和地球半径；控制向量为变形翼的变形 $U=\boldsymbol{\delta}$，为强化学习中的动作向量；D和L分别代表阻力和升力，分别为

$$\begin{aligned} D &= \frac{C_D\rho V^2 S}{2} \\ L &= \frac{C_L\rho V^2 S}{2} \end{aligned} \tag{2}$$

其中：S是飞机机翼的表面积，ρ是大气密度，C_l是升力系数，C_d为阻力系数。为保证飞行器的安全飞行，需要满足驻点热流密度、动压和过载等过程约束。同时为限制滑翔弹道的跳跃，一般还需要考虑施加拟平衡滑翔约束条件。为此变形滑翔飞行器的过程约束条件为

$$\begin{cases} \left(g - \dfrac{V^2}{r}\right) - L \le 0 \\ q = 0.5\rho^2 < q_{\max} \\ Q = K_n\sqrt{\rho}\,V^{3.15} < Q_{\max} \\ n = \sqrt{L^2 + D^2}\,/mg_0 \le n_{\max} \end{cases} \tag{3}$$

式中：$q_{\max}$和$Q_{\max}$分别表示滑翔过程中允许的最大动压和热流密度；K_n表示相关系数，一般与飞行器头部半径和材料有关；g_0为海平面重力加速度；$n_{\max}$为最大允许过载。

为降低气动加热，保护滑翔飞行器，攻角一般采用分段函数模型。为使攻角控制曲线平滑，采用改进的攻角控制律为：

$$\alpha = \begin{cases} \alpha_{\max}, V_1 < V \\ \alpha_{\mathrm{mid}} + \alpha_{\mathrm{bal}}\sin\left[\dfrac{(V - V_{\mathrm{mid}})\pi}{V_1 - V_2}\right], V_2 \le V \le V_1 \\ \alpha_{\max(K)}, V < V_2 \end{cases} \tag{4}$$

式中，　　为最大攻角，$\alpha_{max(K)}$为最大升阻比对应的攻角。V_1和V_2表示分段点的速度。其中，$\alpha_{\mathrm{mid}} = (\alpha_{\max} + \alpha_{\max(K)})/2$，$\alpha_{\mathrm{bal}} = (\alpha_{\max} + \alpha_{\max(K)})/2 V_{\mathrm{mid}} = (V_1 + V_2)/2$。

1.3 气动代理模型构建

为了快速获得变形滑翔飞行器变形过程中的气动力，本研究针对变形滑翔飞行器所覆盖的空域与速域，构建滑翔阶段的变形滑翔飞行器翼面变形-气动

参数的映射代理模型。通过选取变形过程中几种典型的离散构型，运用CFD数值仿真获得变形过程中的气动数据集。基于神经网络的数据回归技术，结合上述数据集，构建翼面变形-气动参数的映射代理模型。该气动代理模型可用于生成滑翔阶段内最小变形角度δ_0到最大变形角度δ_1之间的气动力系数。

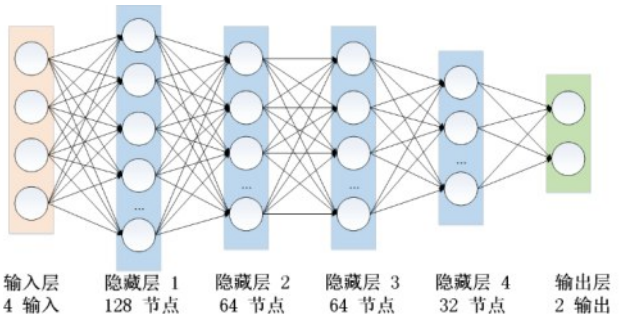


图3 映射模型的网络结构

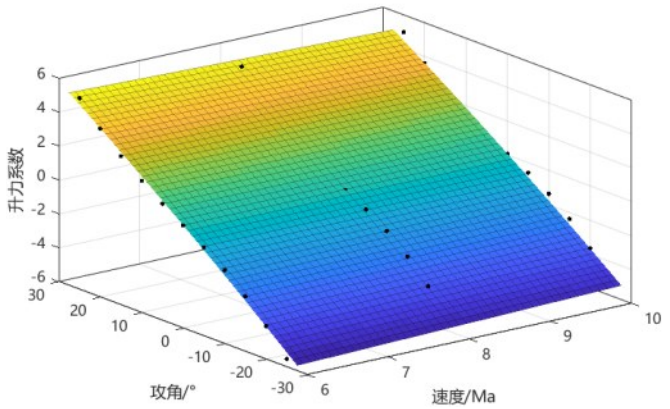
Fig. 3 The structure of the neural network

如图3所示，本研究构建了一个神经网络映射架构，并使用变形滑翔飞行器的CFD气动数据集进行训练。在模型训练前，需要对数据集进行检验，剔除部分有缺陷的气动性能系数异常的数据。

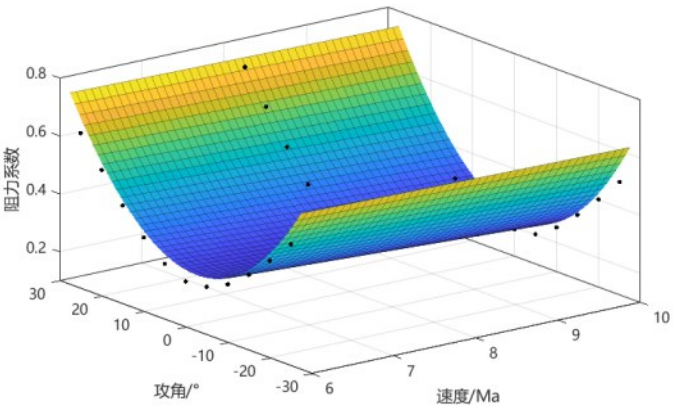
其中气动代理模型的输入量为变形滑翔飞行器的攻角、飞行高度、变形量与马赫数，而输出量为对应飞行状态与变形量下的升阻力系数。样本覆盖的飞行高度范围为20~60km，攻角范围为-25°~24°，变形量范围为δ_0至δ_1，速度范围为0~10马赫。将数据集随机分为3组，分别用于测试、训练和验证。其中80%的数据集被保留用于训练，其余20%被用于测试。随后，从训练集中随机选择20%用于验证。为防止因输入数据的幅度差异导致的收敛问题，需要将数据集归一化处理，公式为

$$x_i^{NL}=\frac{x_i-\min(x)}{\max(x)-\min(x)} \tag{5}$$

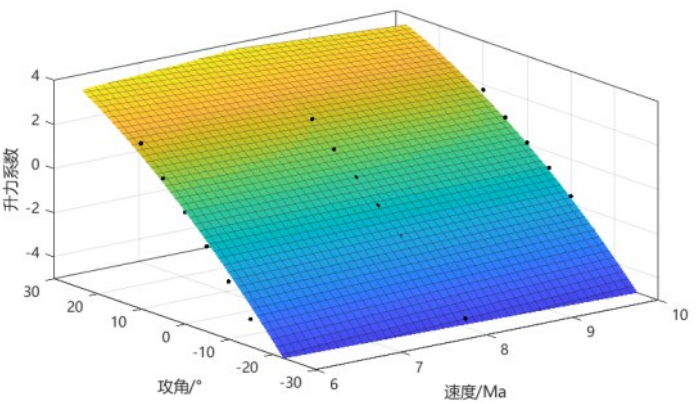
式中：$\max(x)$和$\min(x)$分别表示数据集内的最大值和最小值；x_i^{NL}表示归一化样本。



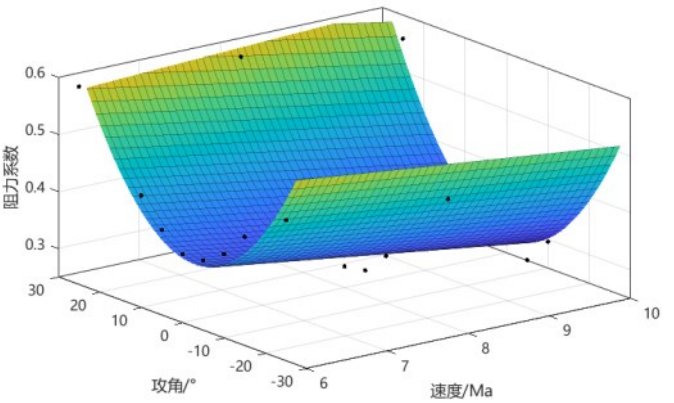
（a） 变形角度为δ_0的升力系数C_{l0}预测结果



（b） 变形角度为δ_0的阻力系数C_{d0}预测结果



（c） 变形角度为δ_1的阻力系数C_{l1}预测结果



（d） 变形角度为δ_1的阻力系数C_{d1}预测结果

图 4 气动性能预测

Fig. 4 Aerodynamic properties prediction

图4中，最小变形角度δ_0处的C_{l0}和C_{d0}的预测值与测试值的R^2分别为0.9904和0.9556，最大变形角度δ_1处的C_{l1}和C_{d1}的预测值与测试值的R^2分别为0.9954和0.9724。上述结果表明，该变形滑翔飞行器翼面变形-气动参数映射代理模型具有良好的精度与可靠性。由于飞行器变形速率相对于飞行速度较慢，流场建立时间远小于轨迹演化时间，暂态气动效应对弹道性能影响微弱，因此本研究采用稳态气动数据进行建模是合理且有效的。

2 变形决策方法框架

2.1 变形决策框架

变形决策系统的强化学习框架如图5所示。

对于变形滑翔飞行器而言，变形决策的目标是让其能够准确选择出最佳的变形量，在整个飞行包络内保持出色的气动力性能的同时，优化飞行弹道。因此，本研究基于气动/变形/弹道的强耦合关系，提出了一种基于深度强化学习技术的变形滑翔飞行器智能变形决策方法。

本方法核心在于构建一个高效闭环的“代理模型-飞行环境-决策智能体”训练与决策框架。首先，为解决强化学习训练中海量气动力数据调用的效率瓶颈，我们开展了基于数据驱动的变形滑翔飞行器气动力代理模型构建。通过CFD数值模拟获取覆盖宽速域、大攻角范围的翼面变形-气动力样本数据集，并对变形参数进行特征提取，进而训练一个深度神经网络模型（DNN）以精确实现飞行状态与变形参数到气动力系数的复杂非线性映射。该模型可替代耗时的数值模拟，提出了一种将变形决策与弹道控制相结合的为智能体与环境的高频交互提供实时、可靠的气动力数据支撑。

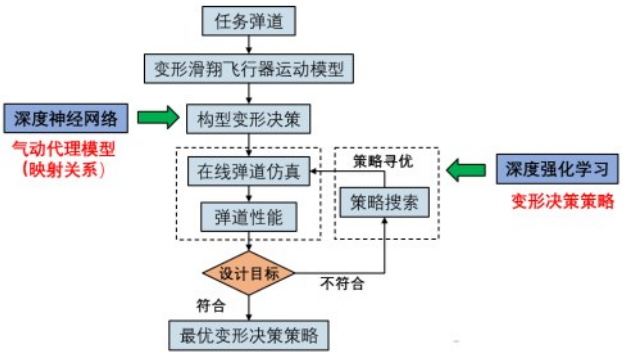


图5　变形决策的强化学习框架

Fig. 5　Reinforcement learning for morphing decision

在此基础上，本文构建了一个面向强化学习的在线弹道模拟训练环境，该环境严格遵循马尔可夫决策过程。我们将飞行器的高度、速度、当前弹道倾角与射程作为关键运动学参数定义为状态空间，并将连续变化的翼面变形量参数定义为动作空间。环境的状态转移函数则由耦合了变形飞行器气动代理模型的飞行器纵向运动学方程组构成，并采用四阶龙格-库塔法进行高精度数值积分，从而模拟智能体执行变形动作后的弹道变化。为了有效引导策略搜索，我们设计了一个多目标的精细化奖励函数，该函数不仅包含指向最大航程的任务目标的终端奖励，更加入了鼓励高升阻比的气动性能正奖励项以及对动压、热流密度和过载等关键过程约束的负惩罚项。

为提升策略的泛化能力，每个训练回合均在预设空域内随机生成初始飞行状态。鉴于翼面变形属于连续动作空间范畴，为了实现平滑的控制输出，本研究选用能够处理连续动作空间的近端策略优化（PPO）算法作为强化学习策略优化器。PPO采用Actor-Critic架构，策略网络（Actor）负责输出变形动作的概率分布，价值网络（Critic）则评估当前状态的优劣。在训练过程中，智能体与环境交互收集大量的弹道轨迹数据存入缓冲池，随后利用泛化优势估计（GAE）计算优势函数，并通过PPO特有的裁剪代理目标函数在多个周期内迭代更新网络参数。这种机制有效平衡了样本利用效率与训练稳定性，防止了破坏性的策略更新。通过在仿真环境中进行充分的“探索-试错”式训练，最终收敛得到能够根据实时飞行状态进行在线生成最优连续变形策略的变形决策网络模型，从而实现变形滑翔飞行器的智能变形决策。

2.2 训练环境设置

考虑到模型训练的收敛性，状态空间 □ 的设计应在能充分描述变形滑翔飞行器飞行状态的同时，限制其维度规模。因此，根据变形飞行器的纵向运动模型，其状态空间$S=[V,\gamma,H,X]^{\mathrm{T}}$，其中

包括飞行速度(V)、当前弹道倾角(γ)、飞行高度(H)和射程(X)。动作空间A的变量为变形翼的翼面变形量(δ)。为了统一维度和提高计算速度，在将状态和动作输入神经网络之前，需要对其进行归一化处理。

对于动作网络，隐藏层采用*Leaky-ReLU*函数来扩展策略搜索的范围。随后，*Tanh*函数被用于计算动作的期望μ，*Softplus*函数被用于计算动作的标准差δ。对于评价网络，隐藏层采用*ReLU*函数，以用于计算评价值。为使得变形决策模型得到充分的学习与收敛，训练最大集数设置为1000，每集的最大步数设置为90。为兼顾模型训练速度、精度与泛化能力，隐藏层节点数应在满足精度同时，尽可能采用紧凑的结构。经过调整，两层隐藏层节点数分别设置为256和64。动作网络和评价网络的学习率分别为0.0001和0.0002。批处理量设置为32，折扣因子为0.99。

2.3 奖赏函数设置

在计算奖励函数时，需要获得完整变形飞行过程下的弹道相关数据。其中，终端高度下的最大结束射程为$X_{\max}$，最小结束射程为$X_{\min}$；最大终端速度为$V_{\max}$，最小终端速度为$V_{\min}$。利用动作网络确定的变形量，积分计算出当前飞行环境下变形后构型的终端速度V与终端射程X。α和β为影响系数。为了充分考虑变形滑翔飞行器的滑翔弹道性能与终端状态下的剩余能量，本文选取α=0.4和β=0.6。奖励函数R设置如下：

$$R=\alpha\frac{V^2-V_{min}^{\ 2}}{V_{max}^{\ 2}-V_{min}^{\ 2}}+\beta e^{\frac{X-X_{min}}{X_{max}-X_{min}}} \tag{6}$$

3 仿真算例

3.1 算法对比

本节使用了两种强化学习算法，即近端策略优化 (PPO)算法和深度确定性策略梯度 (DDPG)算法，并比较了不同算法对变形决策训练的影响。两种算法均采用式（6）的奖励函数与相同的超参数设置进行强化学习训练。通过对比奖励学习曲线和纵向滑翔弹道性能，选择更适合解决该问题的强化学习算法。弹道训练范围如表2所示。

表2 训练环境范围

Table 2 Conditions of the train environment

参数名称	最大值	最小值
初始高度H_0/km	52	48
初始速度V_0/m/s	3000	2500
弹道倾角γ/°	10	-10
终端高度H_f/km	25	25
攻角α/°	-15	15

学习曲线对比如图6所示，可以看出PPO算法具有更好的训练性能。其训练奖励在前250集内迅速增加，在300集后奖赏值保持在85左右。与DDPG算法相比，PPO算法具有收敛更快、更稳定的学习曲线，且最终奖励值更高。对比结果说明PPO算法更适合本研究的训练场景，表现出了较强的学

习效率。

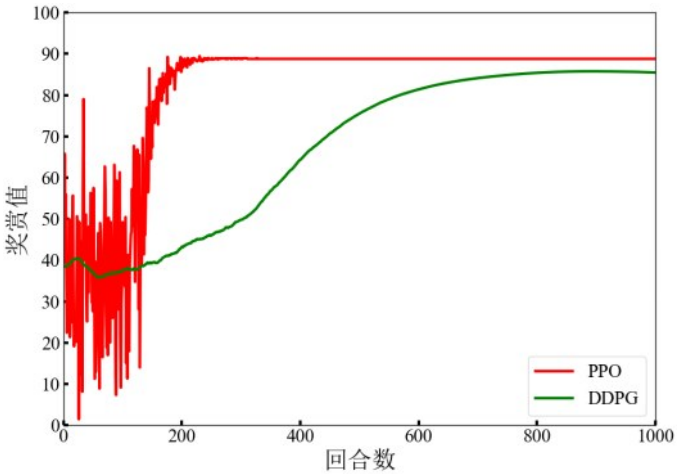


图 6　PPO 和 DDPG 算法的奖励曲线

Fig. 6　Reward curves of PPO and DDPG algorithm

为了验证上述算法选择的结果，通过在相同的测试弹道初始约束条件下(如表 3 所示)，对比 PPO 算法与 DDPG 算法的变形决策弹道，并选择基准弹道作为参照项验证变形决策模型的性能。

表 3　测试弹道初始条件

Table 3　Initial conditions of the test trajectory

参数名称	数值
质量 m/kg	2500
初始速度 V_0/m/s	2863
初始弹道倾角 γ_0 /°	-2
初始高度 H_0/km	50

基于图 7 与表 4 的弹道性能对比结果，采用 PPO 算法的变形滑翔飞行器变形决策模型性能更好，其终端射程离较基准弹道增加了 7.304%，滑翔时间增加了 9.404%。基于滑翔弹道仿真数据可以看出，在滑翔阶段中采用变形决策模型的弹道较基准弹道在终端射程和剩余能量等方面都具有更好的性能，验证了该变形决策方法的可行性。基于强化学习技术构建的智能变形决策模型可以根据升阻比、滑翔段弹道总体性能智能选择翼面变形角度，使变形滑翔飞行器在飞行过程中保持较高的升阻比，维持更高的飞行能量以获得更好的飞行机动能力，优化滑翔弹道以获得更远的航程，实现变形滑翔飞行器的智能变形决策。

3.2 泛化性能测试

在 3.1 节中对变形决策模型进行了测试以证明其可行性，本节将重点研究变形决策模型在训练范围之外的场景下的泛化性。在本节中，通过拉偏试验弹道的初始位置、速度等约束条件，验证

所训练的变形决策模型的泛化性。

场景1：初始速度拉偏的测试案例。其中，case1的初速度拉偏为2000m/s，case2的初速度拉偏为3500m/s。

场景2：初始高度拉偏的的测试案例。其中，case1的初始高度拉偏为55km, case2的初始高度拉偏为45km。

训练范围外的拉偏场景下弹道对比结果如图8、图9所示。在初始条件拉偏情况下，采用变形决策模型的弹道与基准弹道相比在滑翔高度、速度等过程状态以及结束射程、能量等终端状态方面仍具有更好的性能。对比结果表明，基于PPO算法的变形滑翔飞行器变形决策模型对于训练范围外的弹道拉偏情况仍具有一定的泛化性，能够一定程度上实现对未知飞行弹道的智能变形决策。

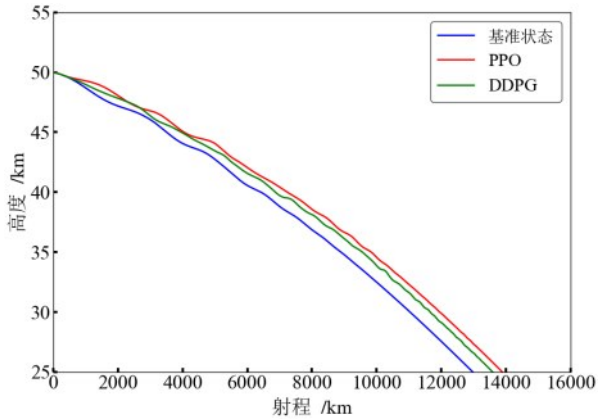
4 结论

本文聚焦于变形滑翔飞行器的智能变形决策难题，提出了一种基于深度强化学习的变形决策方法。研究以无动力变形滑翔飞行器为例，开展了变形决策策略的训练与验证工作。通过对比分析，优选了适配的强化学习算法，并验证了该方法的通用性与可行性。主要研究结论如下：

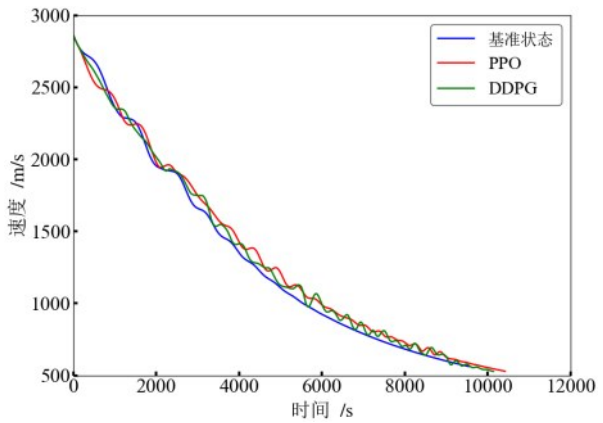
1. 仿真结果表明，本方法能够实现变形滑翔飞行器的智能变形决策，有效维持了高升阻比飞行状态，显著改善了纵向弹道性能；相较于基准状态，飞行器终端射程提升了7.304%，滑翔时间延长了9.404%。

2. 通过对决策模型进行训练域外的参数拉偏测试，证实了本方法在面临模型参数摄动时依然具备良好的泛化性能，验证了其在不确定环境下的鲁棒性与通用性

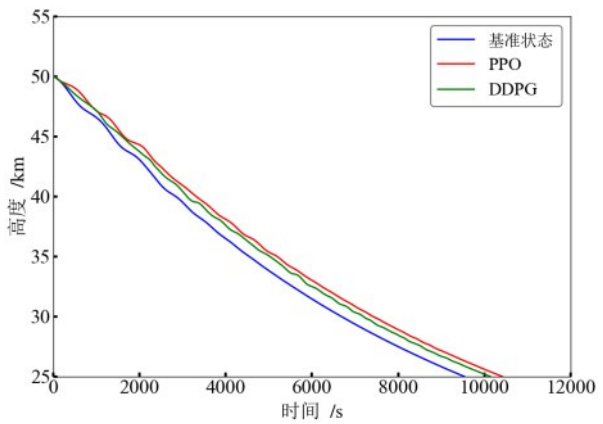
3. 本文提出了一种基于深度强化学习的通用变形决策框架，突破了传统依赖特定对象模型设计固定变形策略的局限，为解决复杂任务场景下各类变形飞行器的智能决策问题提供了新的研究视角与技术途径。



（a） 高度-射程曲线



（b） 速度-时间曲线



（c） 高度-时间曲线

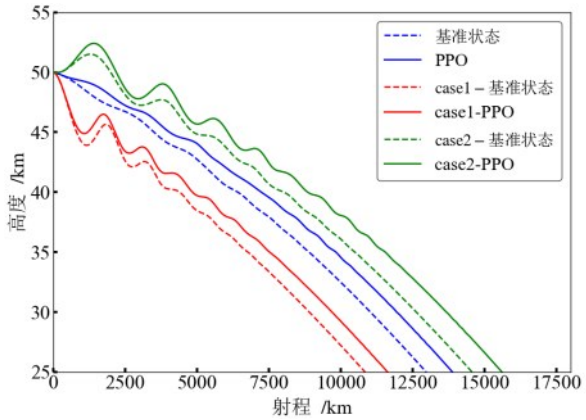
图 7 PPO、DDPG 和基准弹道性能比较

Fig. 7 The trajectory performance comparison of the PPO， DDPG， and basic

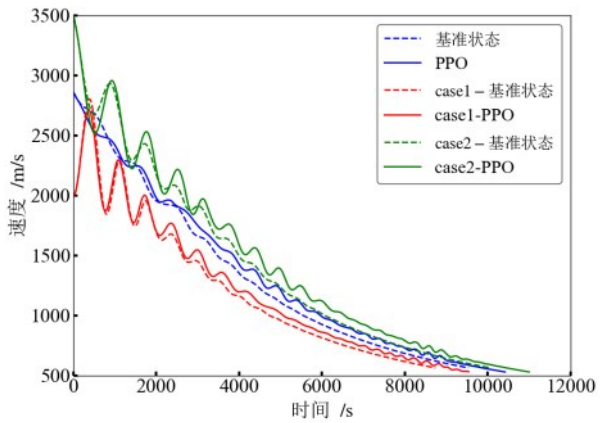
表4 PPO、DDPG和基准弹道的终端性能对比

Table 4 Terminal state comparison of the PPO, DDPG, and basic trajectory

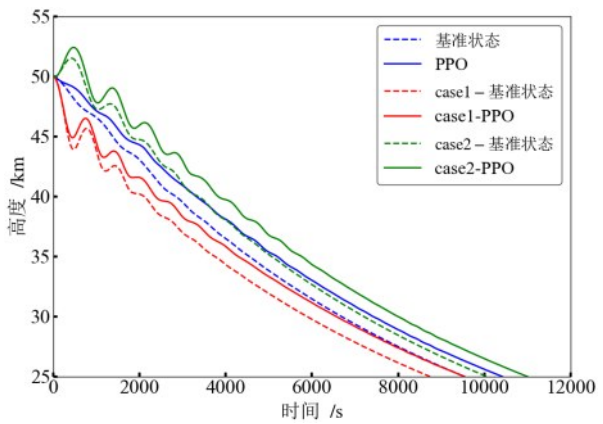
名称	PPO	DDPG	基准弹道
结束射程/km	13923.8	13622.4	12976.1
飞行时间/s	10426.6	10147.6	9530.4



（a） 高度-射程曲线



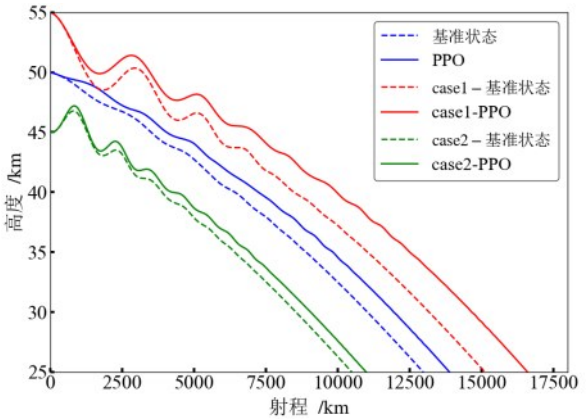
（b） 速度-时间曲线



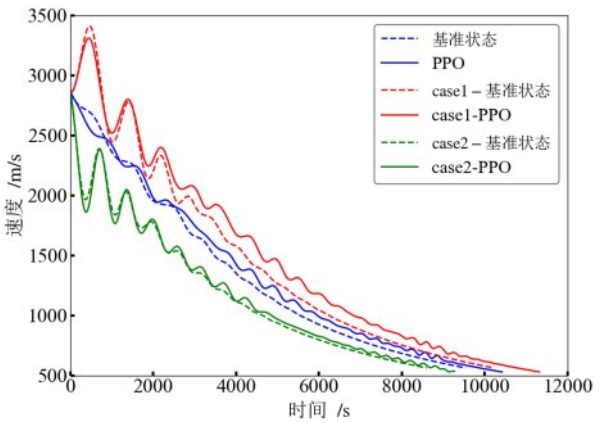
（c） 高度-时间曲线

图 8 场景 1 中初始速度拉偏的弹道对比

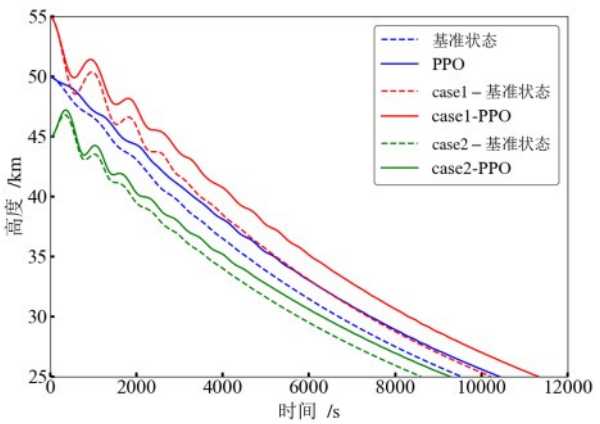
Fig. 8 Comparison results for deflected initial velocities in Scenario 1



（a） 高度-射程曲线



（b） 速度-时间曲线



（c） 高度-时间曲线

图9 场景2中初始高度拉偏的弹道对比

Fig. 9 Comparison results for deflected initial positions in Scenario 2

参考文献

[1] SHI R, SONG J. Dynamics and control for an in-plane morphing wing [J]. Aircraft Engineering and Aerospace Technology, 2013, 85 (1): 24-31.

[2] 王青, 王通, 董朝阳, 等. 变体飞行器链式平滑切换控制[J]. 控制理论与应用, 2015, 32(7): 949-954.
WANG Q, WANG T, DONG C, et al. Chained smooth switching control for morphing aircraft [J]. Control Theory and Applications, 2015, 32(7): 949-954.

[3] 郭雷, 王恩美, 卢昊, 等. 仿生变构型飞行器智能控制技术: 进展与展望[J]. 上海航天(中英文), 2022, 39(4): 25-32.
GUO L, WANG E, LU H, et al. Intelligent Control Technologies for Bio-inspired Morphing Flight Vehicles: Progress and Prospect [J]. Aerospace Shanghai (Chinese & English), 2022, 39(4): 25-32.

[4] 王子健, 张书宇, 侯明哲. 基于在线参数辨识的变体飞行器控制[J]. 兵器装备工程学报, 2022, 43(10): 60-65.
WANG Z, ZHANG S, HOU M. Morphing aircraft control based on online parameter identification [J]. Journal of Ordnance Equipment Engineering, 2022, 43(10): 60-65.

[5] BAO W, LU X, BAI G, et al. Exploration on the frontiers of flexible and deformable cross-domain intelligent flight [J]. Aerospace China, 2021, 22(1): 5-11.

[6] PALACIOS R, MURUA J, COOK R. Structural and aerodynamic models in nonlinear flight dynamics of very flexible aircraft [J]. AIAA Journal, 2010, 48(11): 2648-2659.

[7] 郭东, 徐敏, 陈士橹. 弹性飞行器飞行动力学建模研究[J]. 空气动力学学报, 2013, 31(04): 413-419, 436.
GUO L, XU M, CHEN S. Research on flight dynamic modeling of highly flexible aircrafts [J]. Acta Aerodynamica Sinica, 2013, 31(04): 413-419, 436.

[8] CHEN X, LI C, GONG C, et al. A study of morphing aircraft on morphing rules along trajectory [J]. Chinese Journal of Aeronautics, 2012, 34(7): 232–243.

[9] YUE T, ZHANG X, WANG L, et al. Flight dynamic modeling and control for a telescopic wing morphing aircraft via asymmetric wing morphing [J]. Aerospace Science and Technology, 2017, 70: 328–338.

[10] YUE T, WANG L, AI J. Longitudinal Linear Parameter Varying Modeling and Simulation of Morphing Aircraft [J]. Journal of Aircraft, 2013, 50(6): 1673–1681.

[11] GONG L, WANG Q, HU C, et al. Switching control of morphing aircraft based on Q-learning [J]. Chinese Journal of Aeronautics, 2020, 33(2): 672–687.

[12] BAO C, WANG P, TANG G. Integrated method of guidance, control and morphing for hypersonic morphing vehicle in glide phase [J]. Chinese Journal of Aeronautics, 2021, 34(5): 535–553.

[13] SILVER D, LEVER G, HEESS N, et al. Deterministic Policy Gradient Algorithms [C]. Proceedings of the International Conference on Machine Learning, Beijing, China, 2014.

[14] SCHULMAN J, WOLSKI F, DHARIWAL P, et al. Proximal Policy Optimization Algorithms [A/OL]. arXiv, 2017 [2023-06-30]. http://arxiv. org/abs/1707.06347. DOI: 10.48550/arXiv.1707.06347.

[15] LILLICRAP T, HUNT J, PRITZEL A, et al. Continuous control with deep reinforcement learning [C]. Proceedings of The Twelfth International Conference on Learning Representations, San Juan, Puerto Rico, United States, 2015.

[16] BHATNAGAR S, SUTTON R, GHAVAMZADEH M, LEE M. Natural actor – critic algorithms [J]. Automatica, 2009, 45(11): 2471 – 2482.

[17] CHEN W, GAO C, JING W. Proximal policy optimization guidance algorithm for intercepting near-space maneuvering targets [J]. Aerospace Science and Technology, 2023, 132: 108031.

[18] 闫斌斌, 李勇, 戴沛, 等. 基于增强学习的变体飞行器自适应变体策略与飞行控制方法研究[J]. 西北工业大学学报, 2019, 37(4): 656-663.
YAN B, LI Y, DAI P, et al. Adaptive Wing Morphing Strategy and Flight Control Method of a Morphing Aircraft Based on Reinforcement Learning [J]. Journal of Northwestern Polytechnical University, 2019, 37(4): 656-663.

[19] XU W, LI Y, PEI B, et al. Coordinated intelligent control of the flight control system and shape change of variable sweep morphing aircraft based on dueling-DQN [J]. Aerospace Science and Technology, 2022, 130: 107898.

[20] LAMPTON A, NIKSCH A, VALASEK J. Reinforcement learning of morphing airfoils with aerodynamic and structural effects [J]. Journal of Aerospace Computing, Information, and Communication, 2009, 6(1): 30-50.

[21] LAMPTON A, NIKSCH A, VALASEK J. Reinforcement Learning of a Morphing Airfoil-Policy and Discrete Learning Analysis [C]. AIAA Guidance, Navigation and Control Conference and Exhibit, Honolulu, Hawaii, 2008.

[22] TANDALE M, RONG J, VALASEK J, Preliminary Results of Adaptive Reinforcement Learning Control for Morphing Aircraft [C]. AIAA Guidance, Navigation, and Control Conference and Exhibit, Providence, Rhode Island, 2004.

[23] VALASEK J, TANDALE M, RONG J. A reinforcement learning-adaptive control architecture for morphing [J]. Journal of Aerospace Computing, Information, and Communication, 2005, 2(4): 174-195.

[24] XU D, HUI Z, Liu Y, CHEN G. Morphing control of a new bionic morphing UAV with deep reinforcement learning [J]. Aerospace Science and Technology, 2019, 92: 232-243.

[25] GOECKS V, LEAL P, WHITE T, et al. Control of morphing wing shapes with deep reinforcement learning [C]. Proceedings of AIAA Scitech 2018 forum, Kissimmee, Florida, United States, 2018.