http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2024.0569

自适应斥力势场的无人机航迹规划

郭靖[1]，鲜勇[1,*]，任乐亮[1]，李少朋[1,2]

(1. 火箭军工程大学 导弹工程学院，西安 710025；2. 清华大学 自动化学院，北京 100083)

摘 要：针对多障碍复杂场景下传统人工势场(APF)存在的势场函数设计困难、场景适应性差的难题，提出一种基于深度确定性策略梯度(DDPG)算法的三维自适应斥力势场的航迹规划(TAPF)方法。建立基于强化学习的斥力势场学习智能体，实现“端到端”的斥力输出，提高了斥力场的适应性；设计动态三角形描述斥力场范围，有效解决传统圆形斥力场导引的多余路径问题；根据无人机(UAV)航迹规划的特点，设计TAPF方法的状态空间、动作空间及奖励函数。仿真结果表明：所提方法规划出的路径更短，规划效果更优，对于复杂场景适应性具有更好的鲁棒性。

关 键 词：无人机；航迹规划；强化学习；人工势场；神经网络；避障

中图分类号：TB391.9

文献标志码：A **文章编号**：1001-5965(2026)11-3995-14

随着科技的不断进步和无人机(unmanned aerial vehicle, UAV)技术的迅猛发展，无人机在军事侦察[1]、灾害救援[2]、物流配送[3]、农业勘察[4]等多个领域的应用日益广泛，成为现代社会中不可或缺的重要工具。无人机航迹规划技术作为其核心组成部分，对无人机技术的重要性日益显著。该技术不仅关乎任务执行的效率和安全性，还直接影响无人机在复杂环境中的自主飞行能力和适应性。深入研究无人机航迹规划，旨在实现避障同时保证航迹最短，对于推动无人机技术发展，拓展其应用领域，具有重要的理论意义和实践价值。

针对无人机的航迹规划算法，国内外很多研究人员进行了广泛而深入的研究。目前大致可将算法分为两大类：一类是传统算法如Dijkstra算法[5]、A*算法[6]、模拟退火算法[7]、人工势场(artificial potential field, APF)法[8]等；另一类是智能优化算法如遗传算法[9]、蚁群算法[10]、粒子群算法[11]、人工蜂群算法[12]等。而在实际应用过程中，APF法因其实时性高、计算量小、运算速度快等优势，被广泛运用在无人机航迹规划任务中。尽管APF法存在诸如势场函数难以设计、参数凭经验调节等问题，但凭借其实时性的优势得到了许多研究者的关注，众多研究者对其缺陷进行了相应的改进，主要包括以下3方面：①势场函数的改进，文献[13]基于Frenet坐标系统改进斥力场函数，通过引入安全椭圆模型和预测距离概念调整势场影响区域，使车辆能在静态和动态条件下成功避障，文献[14]通过将引力影响因子添加到斥力场的生成中，改进斥力场函数，从而避免引力与斥力合力为零的情况，克服了目标不可达问题；②参数优化与调整，文献[15]利用反向传播(back propagation, BP)神经网络对APF函数的系数进行调整，得到能够自适应环境的系数，再将APF作为神经网络输出端的特征进行提取，提高了算法的适应性，文献[16]通过研究APF引力参数对运动路径的影响，构建了APF因子的模型，利用梯度下降法求解最短路径；③与其他算法结合，文献[17]结合A*算法和APF法，通过调整代价函数和引入偏航引力增益系数，提高了路径规划的安全

收稿日期：2024-07-25；**录用日期**：2024-11-15；**网络出版时间**：2024-12-06 10:59
网络出版地址：link.cnki.net/urlid/11.2625.V.20241205.1253.002
基金项目：国家自然科学基金(62103432)；陕西省高校科协青年人才支持计划(20210108)
*通信作者. E-mail：xy603xy@163.com

引用格式：郭靖，鲜勇，任乐亮，等. 自适应斥力势场的无人机航迹规划[J]. 北京航空航天大学学报，2026，52(11)：3995-4008.
Guo J，Xian Y，Ren L L，et al. Adaptive repulsive potential field for UAV trajectory planning[J]. Journal of Beijing University of Aeronautics and Astronautics，2026，52(11)：3995-4008 (in Chinese).

性和效率，文献[18]通过精选路径中间点，并仅保留APF法中对目标点的引力作用，为蚁群算法提供路径规划启发，还显著降低了初始阶段路径搜索的随机性，从而实现了算法的快速收敛。

随着人工智能技术的持续进步，机器学习算法已经成为无人机航迹规划领域的关键驱动力。这些算法通过从大量数据中学习并优化决策过程，显著提升了无人机在复杂环境中的航迹规划能力。许多学者结合强化学习算法和APF法进行了相关研究，文献[19]提出了一种基于改进近端策略优化(proximal policy optimization, PPO)算法的智能决策算法，通过引入自适应学习率和状态空间压缩技术，提高了算法在复杂环境中的收敛速度、稳定性和适应性。文献[20]提出了一种改进深度确定性策略梯度(deep deterministic policy gradient, DDPG)算法，通过将势场信息引入到奖励函数中，提高了无人机在复杂三维环境中在线轨迹规划的能力。与传统算法相比，强化学习算法在路径规划中提供了显著的优势，包括更强的环境适应性和灵活性，能够自主学习并优化复杂的决策过程，同时在不断变化的环境中实时调整策略以避免障碍和减少风险。不少研究人员利用这些优势对APF法进行改进，通过学习调整增益系数来提高APF法面对复杂环境时的适应能力。文献[21]采用深度学习优化斥力系数以适应动态环境。结合遗传算法和APF法，以找出特定环境下的最佳斥力系数。随后用这些系数训练一个残差神经网络，使其能够预测出变化环境中所需的斥力系数，生成更加高效和适应性强的无人机飞行路径。文献[22]通过双延迟深度确定性策略梯度(twin delayed deep deterministic policy gradient, TD3)算法对类似APF法的扰动流体算法系数进行训练，得到适应环境的系数组合，再结合扰动流体法进行规划，在满足任务需求的前提下提高了算法的环境适应性。

用强化学习算法学习得到适应环境的系数虽然一定程度上能够解决APF法的缺陷，但在面对复杂环境时，自适应系数能够调整的能力与范围有限，且APF法的势场函数大多靠经验进行设计，环境适应性较差。而使用强化学习算法直接学习APF的势场函数相比于学习斥力增益系数，有2点优势：① 摒弃了依赖人工设定和调整势场参数的繁琐过程，转而采用端到端的自适应函数学习方法，该方法使算法能够直接从环境数据中学习并适应最佳的斥力势场，无需外部干预，显著提升了算法的适应性和泛化能力；② 通过利用神经网络强大的非线性拟合能力，不仅学习到了具体的势场函数，还隐式地优化了势场中的参数配置，相比传统的学习参数方法，这种方式虽然可能在训练阶段需要更多的计算资源，但其在处理复杂多变的场景时，展现出了更高的效率和更强的适应性。

针对上述分析，本文提出一种基于深度确定性策略梯度算法的三维自适应斥力势场的航迹规划(three-dimensional self-adaptive repulsive potential field trajectory planning, TAPF)方法，通过建立端到端的强化学习模型直接输出斥力的大小，摆脱了对人工设计的依赖，增强了对环境的适应能力。另外，提出并设计了三角斥力势场，解决了避障结束后持续受到斥力使路径过长的问题。实验结果表明，本文方法能更有效地应对复杂飞行场景，生成更高质量的安全航迹，满足巡航导弹对航迹规划的快速性和准确性要求。

1 无人机航迹规划问题描述

无人机的航迹规划就是在给定无人机起始点、目标点、飞行环境信息(包括地形、障碍物、禁飞区等)及无人机的性能参数的条件下，寻找一条从起始点到目标点的最优或可行飞行路径。

1.1 无人机运动学约束

由于无人机在飞行过程中受自身性能参数和气动特性的限制，需要考虑的性能约束主要包括最大航程、最大转弯角、最短直行距离、最大俯仰角和最小飞行高度等约束条件。

1) 最大航程约束。无人机所携带的能量是有限的，这就直接限制了其续航能力，即无人机能够达到的最大航程$S_{\max}$。因此，在进行航迹规划时，必须考虑到这一关键的航迹距离约束，假定无人机的任务航程为L。假设航迹点共有n个，无人机的任务航程L满足：

$$\begin{cases} L \leqslant S_{\max} \\ \sum_{i=1}^{n} l_i \leqslant L \\ l_i = V\Delta t \end{cases} \tag{1}$$

式中：l_i为单个节点的航程；V为无人机的飞行速度；Δt为决策周期。

2) 最大转弯角约束。假设无人机实际转弯角为α_i，限制条件下的最大转弯角为$\alpha_{\max}$，这种限制条件可以用数学公式来表达，即

$$\begin{cases} \alpha_i = \dfrac{\boldsymbol{A}_i^{\mathrm{T}}\boldsymbol{A}_{i+1}}{\|\boldsymbol{A}_i\|_2\|\boldsymbol{A}_{i+1}\|_2} \\ -\alpha_{\max} \leqslant \alpha_i \leqslant \alpha_{\max} \end{cases} \tag{2}$$

式中：$\boldsymbol{A}_i$为第i段航程在水平地面上的投影向量，

$\boldsymbol{A}_i=(x_i-x_{i-1},y_i-y_{i-1})$; $\alpha_{\max}$为最大转弯角。转弯角α_i约束如图 1 所示。

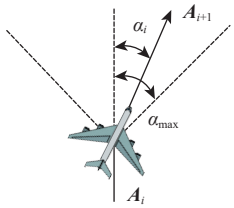


图 1　转弯角约束

Fig. 1　Yaw angle constraint

3) 飞行高度约束。综合考虑安全性和性能因素, 无人机在飞行过程中需严格遵守一定的高度范围。飞行高度约束为

$$h_{\min}<h<h_{\max} \tag{3}$$

式中: h为无人机飞行高度; $h_{\min}$为无人机离地面最低飞行高度; $h_{\max}$为无人机最大可飞高度。

4) 最大俯仰角约束。为确保飞行安全, 无人机在飞行过程中的俯仰角必须控制在一定范围内, 即不得超过其最大允许俯仰角$\gamma_{\max}$。转弯角与俯仰角的关系如图 2 所示, 俯仰角的约束条件为

$$\frac{|z_i-z_{i-1}|}{\sqrt{|(x_i-x_{i-1})^2+(y_i-y_{i-1})^2|}}\leqslant\tan\gamma_{\max}\quad i=2,3,\cdots,n \tag{4}$$

式中: z_i为第i段航程起始高度; $\gamma_{\max}$为最大俯仰角。

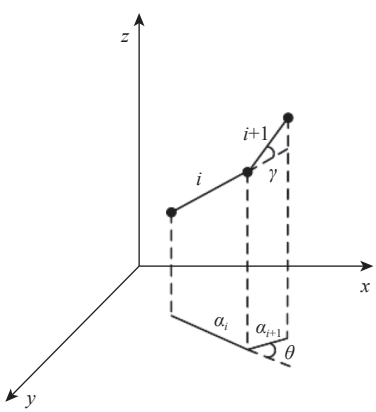


图 2　俯仰角和转弯角关系示意图

Fig. 2　Schematic diagram of relationship between pitch and deflection angles

1.2　无人机飞行环境建模

假设$A(x,y)$为规划空间中的任意一点, x,y分别为横、纵坐标, 则二维规划空间可以表示为

$$\{(x,y)|x_{\min}\leqslant x\leqslant x_{\max},y_{\min}\leqslant y\leqslant y_{\max}\} \tag{5}$$

本文提出的基于 DDPG 算法的 TAPF 方法建立在连续环境空间和动作空间的基础上, 智能体可以位于环境中的任何位置, 在满足动作约束的区间范围可以朝任何方向运动, 采用矢量图方法进行建模更为简便。如图 3 所示, 将平面投影为圆形的防空阵地和预警探测雷达分别以最大打击半径和最大探测半径处理为不可触碰的圆形障碍物; 将不规则障碍物向外进行膨化扩充, 并以其最小外界圆为半径, 同样处理为圆形障碍物, 以保证航迹与威胁源存在一定间隔。

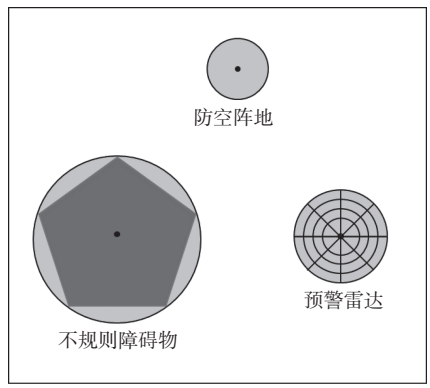


图 3　矢量环境地图

Fig. 3　Vectorized environmental map

2　本文算法原理

2.1　APF 法

APF 法是路径规划任务中常用的一种算法。其核心思想是在空间中引入势场, 即目标点产生引力势场, 障碍物产生斥力势场, 通过 2 个势场力的作用引导无人机避开障碍物, 到达目标点。

2.1.1　引力势场

在 APF 法中, 引力场源只有目标点, 目标点对轨迹上的点产生引力作用从而使轨迹向其汇合。一种常见的引力场计算方式为

$$U_{\text{att}}(\boldsymbol{P})=\frac{1}{2}\xi\|\boldsymbol{P}-\boldsymbol{P}_{\text{g}}\|_2^2 \tag{6}$$

式中: U_{att}为引力场; ξ为引力增益系数; $\boldsymbol{P}$为当前位置; $\boldsymbol{P}_{\text{g}}$为目标点位置。

根据引力势场表达式, 可以求出引力为

$$F_{\text{att}}=-\nabla U_{\text{att}}=\xi\|\boldsymbol{P}-\boldsymbol{P}_{\text{g}}\|_2 \tag{7}$$

当空间中一点坐标$\boldsymbol{P}$与目标点坐标$\boldsymbol{P}_{\text{g}}$的欧氏距离越大时, 引力越大, 引力场模型如图 4 所示。随着引力增益系数的增加, 引力场模型示意图由原来平缓的曲面变为较为陡峭的曲面, 有利于引导无人机更快地到达目标点。

2.1.2　斥力势场

针对空间中存在的每一个障碍物, 引入斥力场来修正原始轨迹, 轨迹与障碍物表面距离越小斥力场越大, 当距离超过一定范围时可以不用考虑斥力的影响, 使轨迹沿直线向目标点移动, 从而可以使

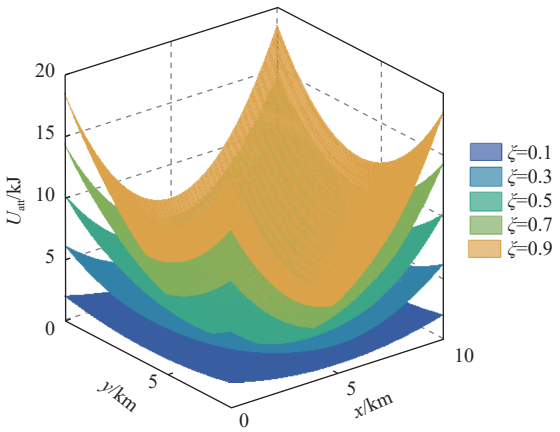


图4 不同引力增益系数下引力场模型示意图

Fig. 4 Schematic diagram of gravitational field model with different gravitational gain coefficients

轨迹长度减小。根据以上描述，一种常见的斥力场计算方式为

$$U_{rep}=\begin{cases}\dfrac{1}{2}\eta\left(\dfrac{1}{\|\boldsymbol{P}-\boldsymbol{P}_{obs}\|_2}-\dfrac{1}{\rho_0}\right)^2 & \|\boldsymbol{P}-\boldsymbol{P}_{obs}\|_2\leqslant\rho_0\\ 0 & \|\boldsymbol{P}-\boldsymbol{P}_{obs}\|_2>\rho_0\end{cases}\quad(8)$$

式中：U_{rep}为斥力场；η为斥力增益系数；$\boldsymbol{P}_{obs}$为障碍物中心坐标；ρ_0为斥力作用范围。

根据斥力势场表达式，可以求出斥力为

$$F_{rep}=-\mathrm{grad}(U_{rep}(\boldsymbol{P}))=\begin{cases}\eta\left(\dfrac{1}{\|\boldsymbol{P}-\boldsymbol{P}_{obs}\|_2}-\dfrac{1}{\rho_0}\right)^2 & \|\boldsymbol{P}-\boldsymbol{P}_{obs}\|_2\leqslant\rho_0\\ 0 & \|\boldsymbol{P}-\boldsymbol{P}_{obs}\|_2>\rho_0\end{cases}\quad(9)$$

斥力场模型如图5所示，随着斥力增益系数的增大，斥力场模型示意图中的斥力场强度显著增强。从较低的增益系数开始，斥力场的峰值逐渐升高，当增益系数达到较高水平时，斥力场的峰值也达到较高水平。

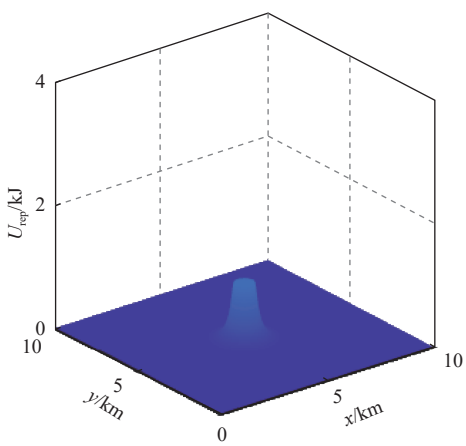
通常情况下，在无人机的路径规划问题中，存在一个目标点和一个或多个障碍物，因此，无人机所受到的合势场为

$$U_{all}=U_{att}(\boldsymbol{P})+\sum U_{rep}(\boldsymbol{P})\quad(10)$$

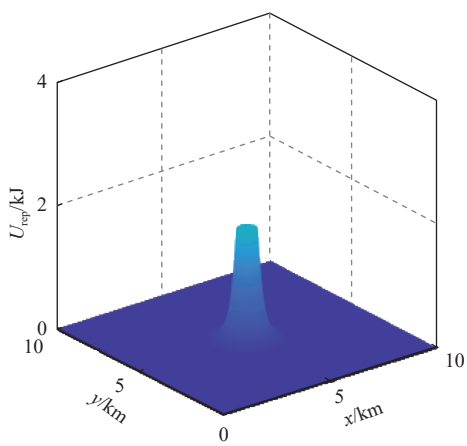
同样合力也为引力与斥力的叠加，无人机受到的合力为

$$F_{all}(\boldsymbol{P})=F_{att}(\boldsymbol{P})+\sum F_{rep}(\boldsymbol{P})\quad(11)$$

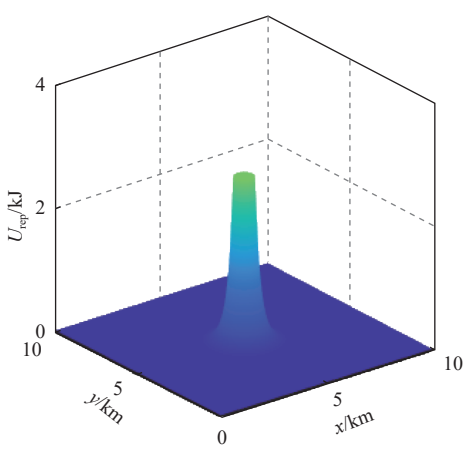
合势场模型如图6所示。可以看到，合势场在



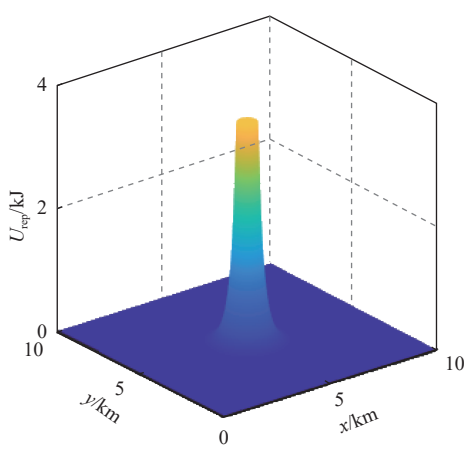
(a) η=0.5, $U_{rep,max}$=0.9 kJ



(b) η=1, $U_{rep,max}$=1.8 kJ



(c) η=1.5, $U_{rep,max}$=2.7 kJ



(d) η=2, $U_{rep,max}$=3.5 kJ

图5 不同斥力增益系数下斥力场模型示意图

Fig. 5 Schematic diagram of repulsive force field model with different repulsive force gain coefficients

障碍物附近隆起, 并同时向目标点位置凹陷, 从而引导轨迹在避障的同时向目标点移动。

2.2　DDPG 算法理论

DDPG 算法是一种确定性策略强化学习算法。相较于随机策略算法学习在某个状态下采取动作的概率分布而言, DDPG 算法会直接给出在该状态下的动作。确定性策略具有收敛速度快、学习效率高等优势, 但另一方面, 确定性策略相比随机策略缺少一些对环境的探索性, 因此, 确定性策略往往需要在采集数据阶段引入噪声来增强算法探索性。本文在测试了各种主流强化学习算法解决本文提出的航迹规划问题后, 综合考虑最终选择 DDPG 算法作为改进算法的底层算法。图 7 给出了本文算法的框图。

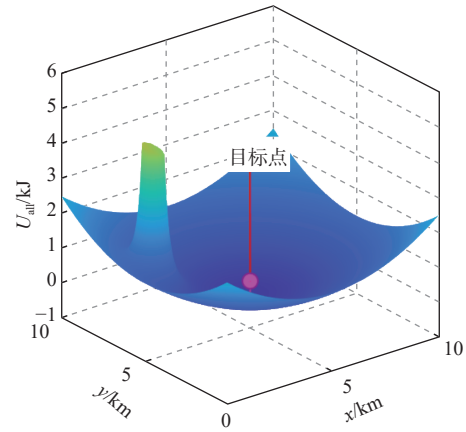


图 6　引力场与斥力场的叠加场模型示意图(单个斥力场)

Fig. 6　Schematic diagram of superimposed field model of gravitational and repulsive fields (single repulsive field)

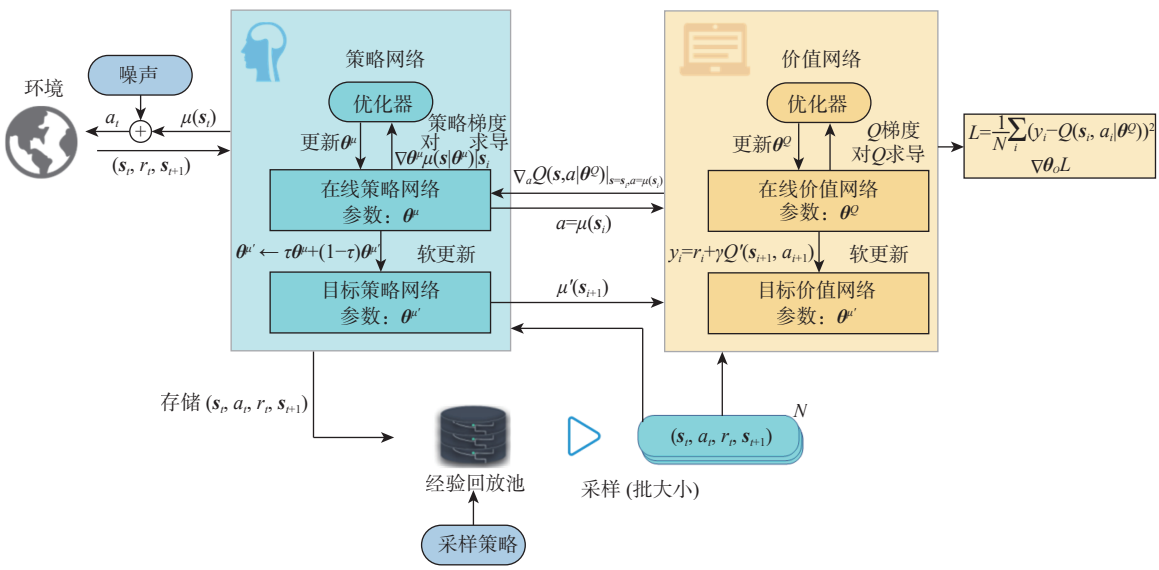


图 7　本文 DDPG 算法框图

Fig. 7　Detailed block diagram of the DDPG algorithm

3　自适应势场航迹规划

本文将障碍物看作强化学习的智能体, 通过在环境中与 UAV 的交互数据学习如何控制其产生的斥力场大小, 并学习障碍物之间如何通过合作产生叠加斥力场使 UAV 的规划航迹达到理想的最佳状态。将深度强化学习算法与 APF 算法结合得到本文提出的三维自适应势场航迹规划方法。值得注意的是, 本文 TAPF 方法由于其自身的斥力场会随着 UAV 的位置不同而不同, 因此, 其局部最小值点位也会随之改变, 且在训练模型时落入局部最小值区域后, 智能体会受到一定的惩罚, 因此, 本文 TAPF 方法有效避免了各类势场算法都存在的共性问题——局部最小值问题。另一方面, 通过学习训练直接输出斥力的大小, 解决了势场函数设计及参数调节困难的问题。

3.1　本文方法主要框架

在本文 TAPF 方法中, APF 为规划层, 负责航路的最终规划, 强化学习模型为学习与决策层, 通过大量智能体与环境交互的数据学习如何控制斥力场的大小。本文 TAPF 方法框架如图 8 所示, 其中, $\boldsymbol{P}$ 表示位置坐标, $\boldsymbol{s}$ 表示相对位置。UAV 在三维空间中某一点通过雷达等手段接收到障碍物的相对位置 $\boldsymbol{s}_{obsi}$ 及目标的相对位置 $\boldsymbol{s}_{goal}$, 将其输入对应障碍物的强化学习策略网络, 可得到对应障碍物在当前环境状态下决策出的斥力 F_{repi}, 重复该步骤可在 t 时刻得到所有障碍物的斥力 F_{rep}, 将其作为强化学习的动作输入规划层, 计算得到 UAV 的下一航路点, 并返回从状态 $\boldsymbol{s}_t$ 到 $\boldsymbol{s}_{t+1}$ 的奖励值 r_t, TAPF 方法将动作、状态、奖励存储到对应障碍物的数据缓存中, 以便模型之后的学习。为简化不必要的问题, 本文采用无限高圆柱体模拟障碍物。

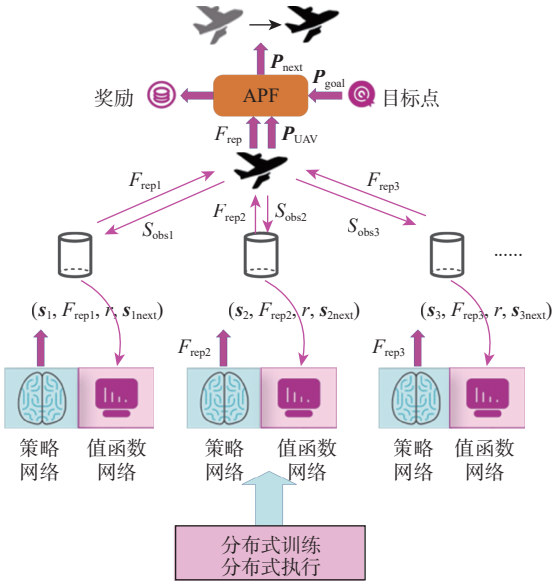


图8 本文TAPF方法实现框架

Fig. 8 The TAPF method implementation framework

3.2 马尔可夫决策过程设计

无人机路径规划中的强化学习要素主要包括状态空间、动作空间和奖励函数。

1）状态空间。作为强化学习模型策略的输入，定义状态$\boldsymbol{s}$为三维空间中障碍物相对UAV的位置及目标点相对于UAV的位置：

$$\boldsymbol{s}=(\Delta x_{\text{obs}},\Delta y_{\text{obs}},\Delta z_{\text{obs}},\Delta x_{\text{goal}},\Delta y_{\text{goal}},\Delta z_{\text{goal}}) \tag{12}$$

式中：$(\Delta x_{\text{obs}},\Delta y_{\text{obs}},\Delta z_{\text{obs}})$为障碍物与UAV的相对位置；$(\Delta x_{\text{goal}},\Delta y_{\text{goal}},\Delta z_{\text{goal}})$为目标点与UAV的相对位置。

2）动作空间。基于无人机的运动学模型和APF法，本文主要学习APF法的斥力势场，因此，输出为斥力，可得

$$a=F_{\text{rep}}(\boldsymbol{P}) \tag{13}$$

式中：$F_{\text{rep}}(\boldsymbol{P})$为当前位置所受斥力的大小。

3）奖励函数。强化学习的奖励机制是影响算法学习效果至关重要的因素，奖励值需体现待解决问题的关键表征。另外，多个奖励值之间需统一数值范围，奖励方差应该尽量小，这可以避免累计奖励值的大幅度振荡，提升学习效率。一条拥有高性能指标的航路往往需满足轨迹平滑、距离较短、不与障碍物有交点等条件，而平滑性问题在基础的APF算法中已经基本保证，且经过大量实验，在航路距离较短的情况下往往能兼顾平滑性的要求，因此，奖励中将不再考虑平滑性约束。综上所述，本文设计了如下奖励机制：

$$R_t(a_t|\boldsymbol{s}_t)=\lambda_1R_1+\lambda_2R_2+\lambda_3R_3 \tag{14}$$

式中：λ_1、λ_2、$\lambda_3>0$为加权因子；R_1为避障奖励；R_2为航迹奖励；R_3为航向奖励。

若下一航路点$\boldsymbol{P}_{t+1}$与障碍物有交点时，对智能体进行惩罚：

$$R_1=-\frac{\|\boldsymbol{P}_{t+1}-\boldsymbol{O}_{\text{obs}}\|_2}{R_{\text{obs}}}-C_0 \quad \|\boldsymbol{P}_{t+1}-\boldsymbol{O}_{\text{obs}}\|_2\leqslant R_{\text{obs}} \tag{15}$$

式中：$\boldsymbol{O}_{\text{obs}}$为与航路有交点的障碍物“中心”，圆柱体障碍物取其水平截面圆的圆心；R_{obs}为障碍物半径，圆柱体取其水平截面圆的半径；C_0为常数惩罚项$(C_0>0)$，有助于智能体快速认识到其规划航路不满足碰撞检测。

若下一航路点$\boldsymbol{P}_{t+1}$与障碍物无交点时，设置的奖励值应该倾向于优化航路长度：

$$R_2=\begin{cases}-\dfrac{\|\boldsymbol{P}_{t+1}-\boldsymbol{P}_{\text{g}}\|_2}{\|\boldsymbol{P}_0-\boldsymbol{P}_{\text{g}}\|_2} & \|\boldsymbol{P}_{t+1}-\boldsymbol{P}_{\text{g}}\|_2>\varepsilon\\ -\dfrac{\|\boldsymbol{P}_{t+1}-\boldsymbol{P}_{\text{g}}\|_2}{\|\boldsymbol{P}_0-\boldsymbol{P}_{\text{g}}\|_2}+C_1 & \|\boldsymbol{P}_{t+1}-\boldsymbol{P}_{\text{g}}\|_2\leqslant\varepsilon\end{cases} \tag{16}$$

式中：$\|\boldsymbol{P}_{t+1}-\boldsymbol{P}_{\text{g}}\|_2$表示下一航路点与目标点的距离；$\|\boldsymbol{P}_0-\boldsymbol{P}_{\text{g}}\|_2$表示从起始点到目标点的距离；$\varepsilon$为小常数，用于判断$\boldsymbol{P}_{t+1}$是否已经抵达目标点附近，若到达，则额外奖励$C_1(C_1>0)$。

为避免无人机在避障时偏转角度过大、飞行距离增加，设计了航向奖励，当无人机偏离目标方向时，给予惩罚。表达式为

$$R_3=-0.1\frac{|\varphi|}{\pi} \tag{17}$$

式中：$|\varphi|$为无人机航向与目标点方向之间的夹角。

3.3 斥力势场改进

在传统的APF中，由斥力势场的表达式可知，无人机与障碍物之间的距离越近，所受到的斥力越大。当无人机进入斥力势场影响范围内时即受到障碍物的斥力，直到无人机飞离斥力影响范围，这会导致无人机在绕开障碍物时产生多余的避障路径。如图9所示，无人机此时已完成避障，但仍处于斥力影响范围内，会受到斥力的作用，产生多余的飞行距离。

本文对斥力势场影响范围进行改进，提出一种三角势场，即当无人机进入障碍物斥力影响范围时，判断无人机飞行方向是否位于当前位置与障碍物2条切线的夹角之间。若飞行方向位于切线夹角间，无人机受到斥力作用，绕开障碍物；若飞行方向在切线夹角之外，无人机尽管处于斥力影响范围内，但位于三角区外，此时不受斥力，仅受目标引力作用。与传统的斥力势场相比，该三角势场使无人机的路径明显缩短。三角势场如图10和图11所示。

当无人机进入障碍物斥力范围内时，判断此时飞行方向与当前位置到障碍物圆心向量的夹角是否小于90°。若夹角大于90°，则不受障碍物斥力，若夹角小于90°，则判定为进入三角区斥力势场。设无人机的方向向量为$\boldsymbol{A}$，当前位置点到障碍物圆

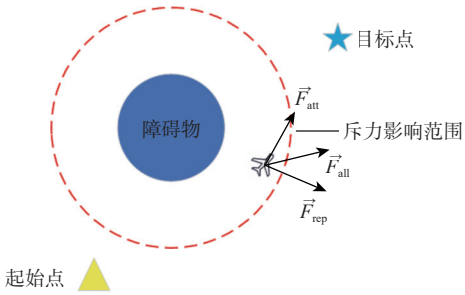


图 9　APF 算法原理图

Fig. 9　Schematic diagram of APF algorithm

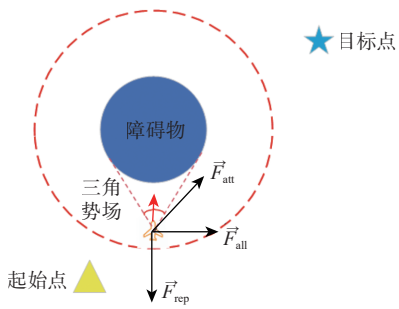


图 10　无人机进入三角势场

Fig. 10　UAV entry triangle potential field

的两条切线向量分别为$\boldsymbol{B}$和$\boldsymbol{C}$, 若航向角在两条切线形成的夹角之间, 则受到三角区势场的斥力, 若不在切线夹角之间, 则无人机不受斥力, 如下:

$$F_{rep}=\begin{cases}\eta\left(\dfrac{1}{\|\boldsymbol{P}-\boldsymbol{P}_{obs}\|_2}-\dfrac{1}{\rho_0}\right)^2 & (\boldsymbol{A}\times\boldsymbol{B})(\boldsymbol{A}\times\boldsymbol{C})>0\\ 0 & (\boldsymbol{A}\times\boldsymbol{B})(\boldsymbol{A}\times\boldsymbol{C})\leqslant 0\end{cases} \tag{18}$$

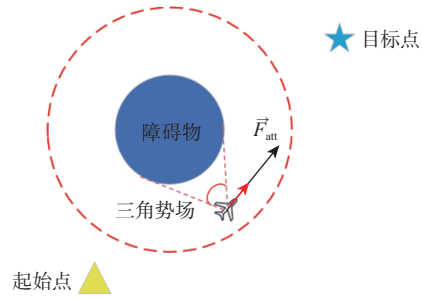


图 11　无人机离开三角势场

Fig. 11　UAV departure triangle potential field

3.4　训练流程设计

在本文 TAPF 方法的训练过程中, 无人机通过在环境中进行探索, 根据环境的反馈不断学习更加正确的策略, 训练流程如图 12 所示。无人机首先将自身位置、障碍物位置、目标点和地形等数据作为状态空间, 通过学习到的策略和噪声生成动作。在执行完动作后, 会对无人机是否到达目标点或发生碰撞进行判定, 若发生碰撞或到达目标点, 则初始化环境后重新开始训练, 若未发生碰撞或未到达目标点, 则通过奖励函数获得奖惩值, 用于更新神经网络, 重复训练过程, 直至网络收敛。

3.5　相关参数设计

在飞行参数方面, 三维航迹规划需要考虑无人机飞行的最小飞行高度h_{min}, 结合实验中地形模型数据与无人机的性能参数, 设定$h_{min}=1$ km, 同时, 无人机最大离地飞行高度的期望值h_{max}=2 km。假设无人机全程以速度大小V做定速飞行, 且每个时间步为定值Δt, 因此, 各航迹段的长度l_i亦为定值,

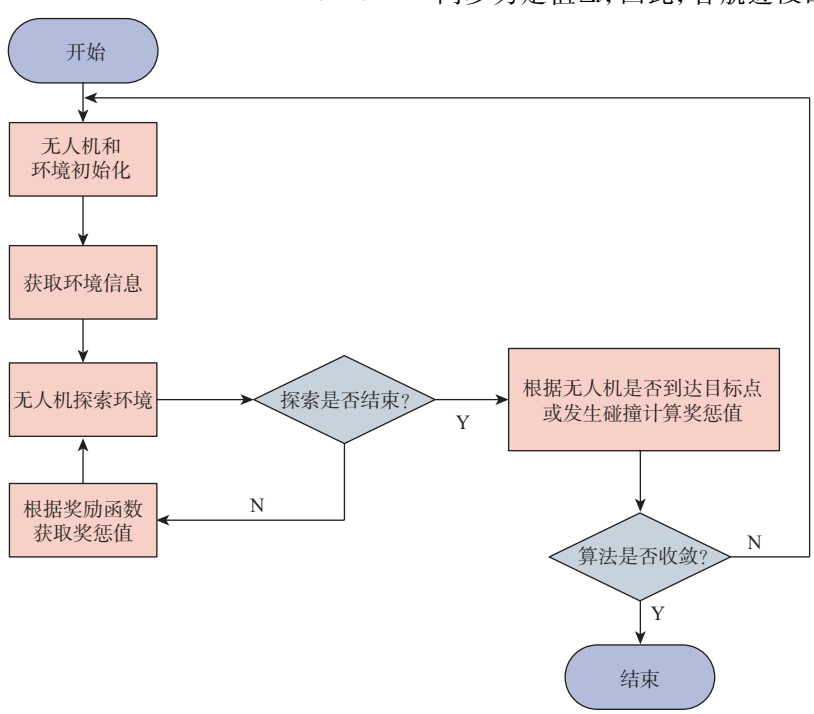


图 12　本文方法流程图

Fig. 12　The proposed method flow chart

即智能体的搜索步长一定，本节实验中设定为500 m。综合考虑航迹的平滑性、机动性及无人机的自身性能特性，设定转弯角α的取值范围为$[-15°,15°]$，俯仰角γ的取值范围为$[-6°,6°]$。

在APF参数设计方面，目标引力势场与障碍物斥力势场同样存在于整个三维规划空间中。设引力势场增益系数$k_{att}=2$，障碍物的斥力势场影响范围$\rho_0=1$ km。

在强化学习算法方面，本文使用基于PyTorch框架搭建的深度强化学习算法进行实验优化，主要参数设定如表1所示。

表1 本文DDPG算法参数

Table 1 Parameters of the DDPG algorithm

学习率	折扣系数	经验池容量	批处理数据	训练总步数	测试回合数
0.000 3	0.99	10^6	256	1 000	1 000

4 仿真实验与分析

4.1 实验设计

在三维在线航迹规划的实验中，设有训练环境和测试环境，并分别使用本文TAPF方法和自适应斥力参数人工势场(adaptive repulsive parameter artificial potential field, ARP-APF)方法[23]进行训练和测试，分析对比实验结果，同时指出算法改进后的优势所在。

4.1.1 训练环境

实验中地形模型构建采用数字高层模型(digital elevation model, DEM)，并采用双线性插值的方法计算非格网点的高程数据。如图13所示，飞行环境的大小为100 km×100 km，其水平维度上相邻网格数据点的间隔距离为20 km，垂直方向上相邻网格点间隔距离为1 km。

如图14所示，本实验以地形曲面为底面、以环境限高为高度随机构建2～5个位置不同的品红色圆柱障碍物模型，其截面圆的半径大小等于各威胁源的半径r_o，本实验中统一设定为1 km。起始点为红色圆，坐标值为(0, 0, 2) km，目标点为红色三角形，坐标值为(100, 100, 1) km。

4.1.2 测试环境

测试环境有3个，测试环境设置与训练实验环境相似。其中，地形环境模型都是以DEM为基础构建的100 km×100 km三维地形模型，但环境中威胁圆柱设置的大小、数量和位置有所不同。

测试环境1为固定半径障碍物的环境，如图15所示，该环节中共设置500次包含6个位置不同、半径相同的固定品红色圆柱，其中一个场景如图15所示，各圆柱中心线的水平二维坐标分别为：(55, 55) km、(63, 72) km、(32, 64) km、(30, 22) km、(45, 39) km、(58, 36) km，半径均为6 km。该环境的目的在于测试算法在路径规划上相较于其他算法具有更优的性能。

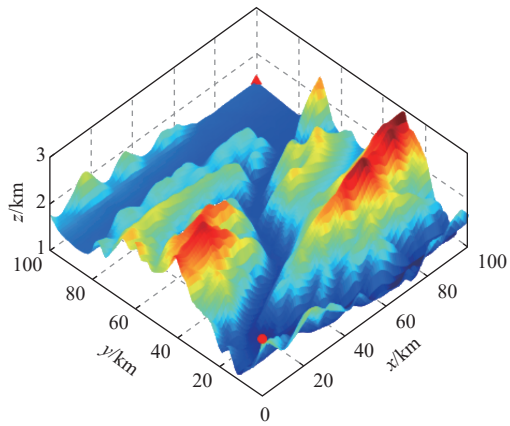


图13 三维地形模型

Fig. 13 Three-dimensional terrain model

测试环境2为与训练环境不同的随机半径障碍物的陌生环境。部分场景如图16所示，该环境中共设置500次包含6个位置不同、半径大小不一的圆柱障碍物。该环境的目的在于测试算法对于存在不同半径障碍物的陌生环境的泛化性和可靠性。

测试环境3为随机位置、固定半径障碍物的陌生环境。该环境中共设置500次包含2～6个位置不同的圆柱障碍物。该环境的目的在于测试三角区势场对于减少APF法多余路径的效果。

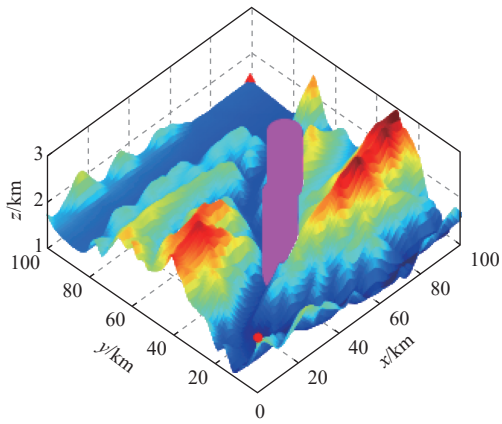
4.2 训练实验

实验中使用ARP-APF方法和本文提出的TAPF方法在随机2～6个障碍物环境中训练1 000个回合，以50回合为单位进行滑动平均处理后，得到奖励值变化曲线，如图17所示。可以看出，本文TAPF方法能够生成曲线的奖励值相对更高，且本文TAPF方法收敛速度比ARP-APF方法更快，接近200回合时就能够收敛。另外，本文TAPF方法的收敛稳定性更好，其收敛后曲线的波动相对原APF算法更小。综上所述，在训练实验中，相对原APF算法而言，本文TAPF方法的收敛稳定性更好，航迹规划能力有所提升。

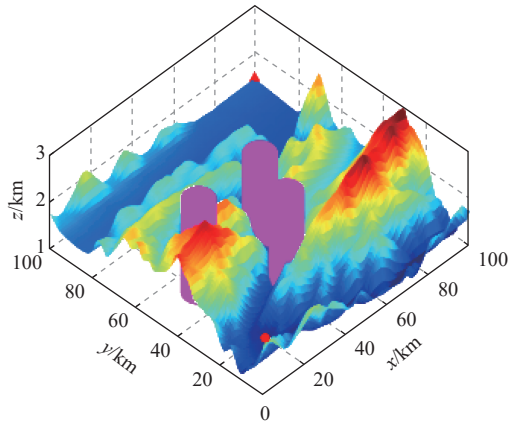
4.3 测试实验

4.3.1 固定半径障碍物环境的测试实验

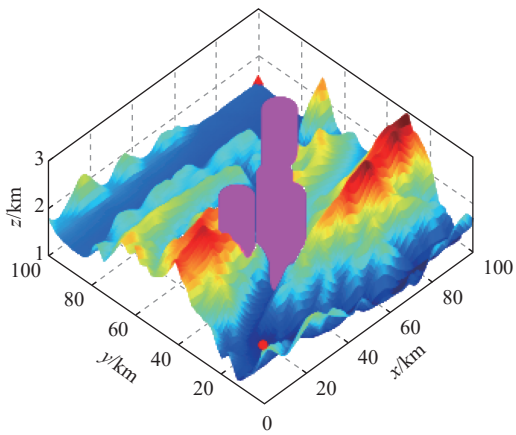
在陌生固定障碍物环境的测试实验中，分别将本文TAPF方法、ARP-APF方法及改进的APF(improve APF, IAPF)法[24]在测试环境1中测试500次，分别以平均成功率M_s、平均奖励M_r、平均距离M_l、平均规划用时M_t、平均水平机动角M_{turn}为指标统计实验结果，同时输出测试的航迹图。其



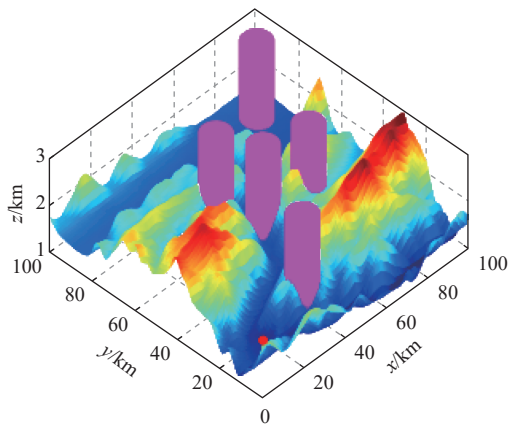
(a) 训练环境三维模型1



(b) 训练环境三维模型2



(c) 训练环境三维模型3



(d) 训练环境三维模型4

图 14　训练环境三维模型

Fig. 14　Three-dimensional model of training environment

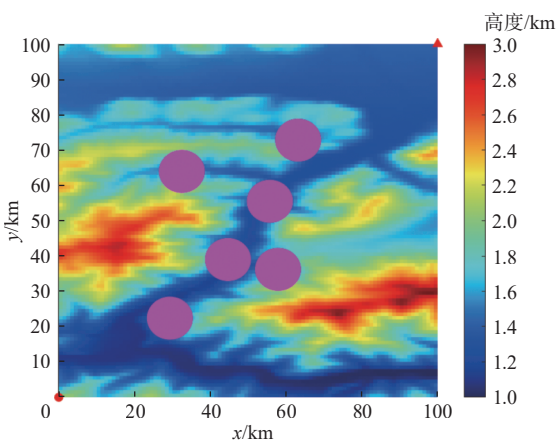


图 15　测试环境 1 模型俯视图

Fig. 15　Top view of test environment 1 model

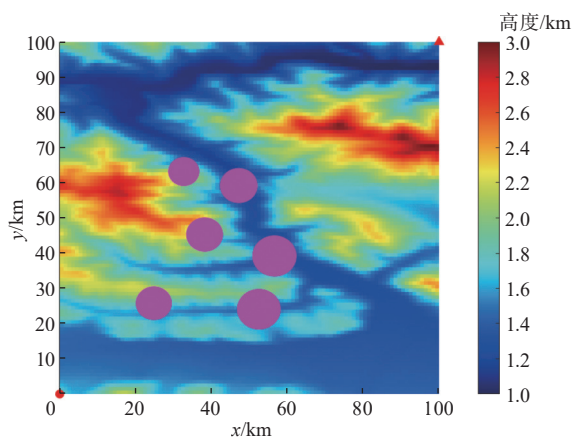


图 16　测试环境 2 模型俯视图

Fig. 16　Top view of test environment 2 model

中, 测试实验中各方法的指标如表 2 所示。

由表 2 可以看出, 在平均转弯角方面, 本文 TAPF 方法生成的航迹机动值明显小于 ARP-APF 方法和传统的 IAPF 法; 同时, 由于机动幅度变小, 生成的航迹更短, ARP-APF 方法平均需要 363.3 km 才能到达目标, 而本文的 TAPF 方法平均只需要 304.5 km 即可到达, 整体降低约 16.2%。此外, 从平均奖励值指标可以看出, 本文 TAPF 方法能够根据飞行环境的实际情况得出更为优秀的动作, 进而航迹的总奖励值也更大。由此, 在测试实验的陌生障碍物环境条件下, 本文 TAPF 方法进行三维在线航迹规划时比 ARP-APF方法和 IAPF 方法具有更优秀的表现。

从图 18 和图 19 的无人机三维航迹中可以看出, 本文 TAPF 方法机动幅度更小, 生成的路径更为平滑, 该结果与表 2 中的数据相符。

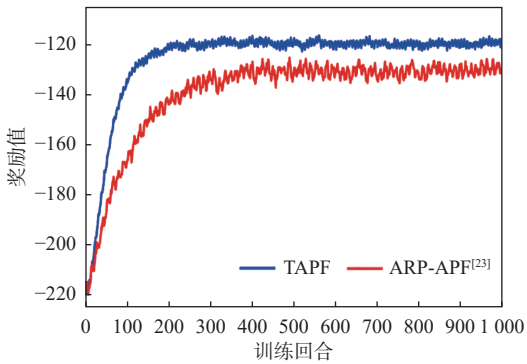


图 17 训练奖励值变化曲线

Fig. 17 Change curves of training reward value

4.3.2 随机半径障碍物环境的测试实验

测试环境 2 中的障碍物半径随机，在该随机半径障碍物的环境中测试 500 次后，统计各实验指标结果，如表 3 所示。可以看出，ARP-APF 方法和 IAPF 方法的成功率有所下降，但本文 TAPF 方法的成功率依然是 100%，具有较好的泛化性。由于障碍物影响，智能体在 2 种方法下都需要进行避障，因此，2 种方法的航迹长度、规划时间和机动值大小都有所增加。但本文 TAPF 方法规划出的航迹无论在航迹长度、奖励值大小还是平滑性方面都更好。

表 2 测试环境 1 中指标对比

Table 2 Comparison of metrics in test environment I

方法	M_s/%	M_r	M_l/km	M_t/s	M_{turn}/(°)
TAPF	100.0	−123.4	304.5	0.70	43.52
ARP-APF[23]	99.4	−137.3	363.3	0.80	59.01
IAPF[24]	97.0		424.5	0.65	75.32

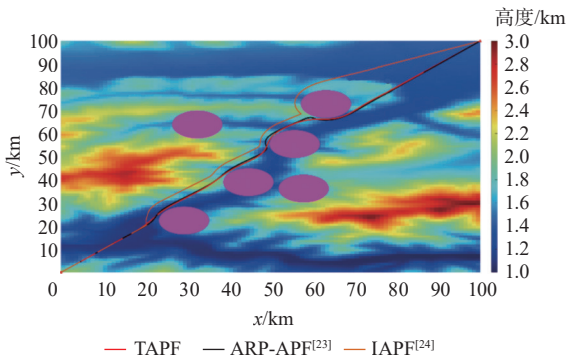


图 18 测试环境 1 中三维航迹俯视图 (测试 1)

Fig. 18 Top view of 3D trajectory in test environment 1

综合上述结果分析，本文 TAPF 方法在陌生环境中的在线航迹规划能力优于 ARP-APF 方法和 IAPF 方法，且能够稳定快速地生成更优的航迹。

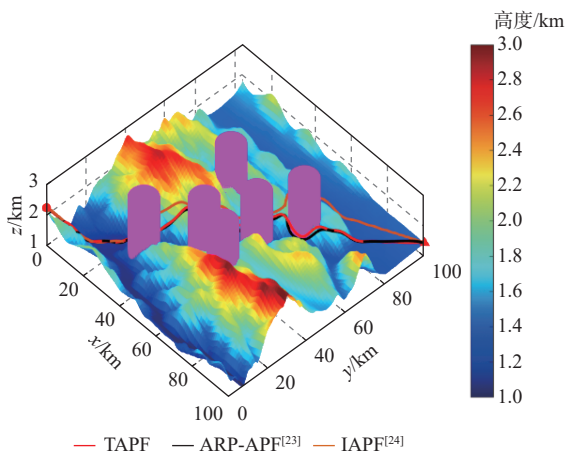


图 19 测试环境 1 中三维航迹模型图 (测试 1)

Fig. 19 Three-dimensional trajectory model map in test environment 1

表 3 测试环境 2 中指标对比

Table 3 Comparison of metrics in test environment 2

算法	M_s/%	M_r	M_l/km	M_t/s	M_{turn}/(°)
TAPF	100.0	423.40	327.2	0.70	21.62
ARP-APF[23]	99.4	337.35	384.3	0.80	37.82
IAPF[24]	95.2		456.4	0.63	68.40

3 种算法的航迹图如图 20 所示。可以明显看出，无论是从竖直维度还是从水平维度来看，TAPF 方法的机动幅度明显更小，路径具有更好的平滑性。

4.3.3 三角区斥力势场的测试实验

为测试三角区势场对于减少 APF 法多余路径的效果，设计了测试实验 3。测试环境 3 中随机生成 3～6 个半径相同的障碍物，分别采用本文 TAPF 方法和 IAPF 方法在同一环境中进行测试实验，统计各实验指标结果，如表 4 所示。可以看出，本文 TAPF 方法通过改进的三角区斥力势场规划出的路径长度较优化的 APF 法更短，且平均水平机动角更小。本文 TAPF 方法规划出的航迹无论在航迹

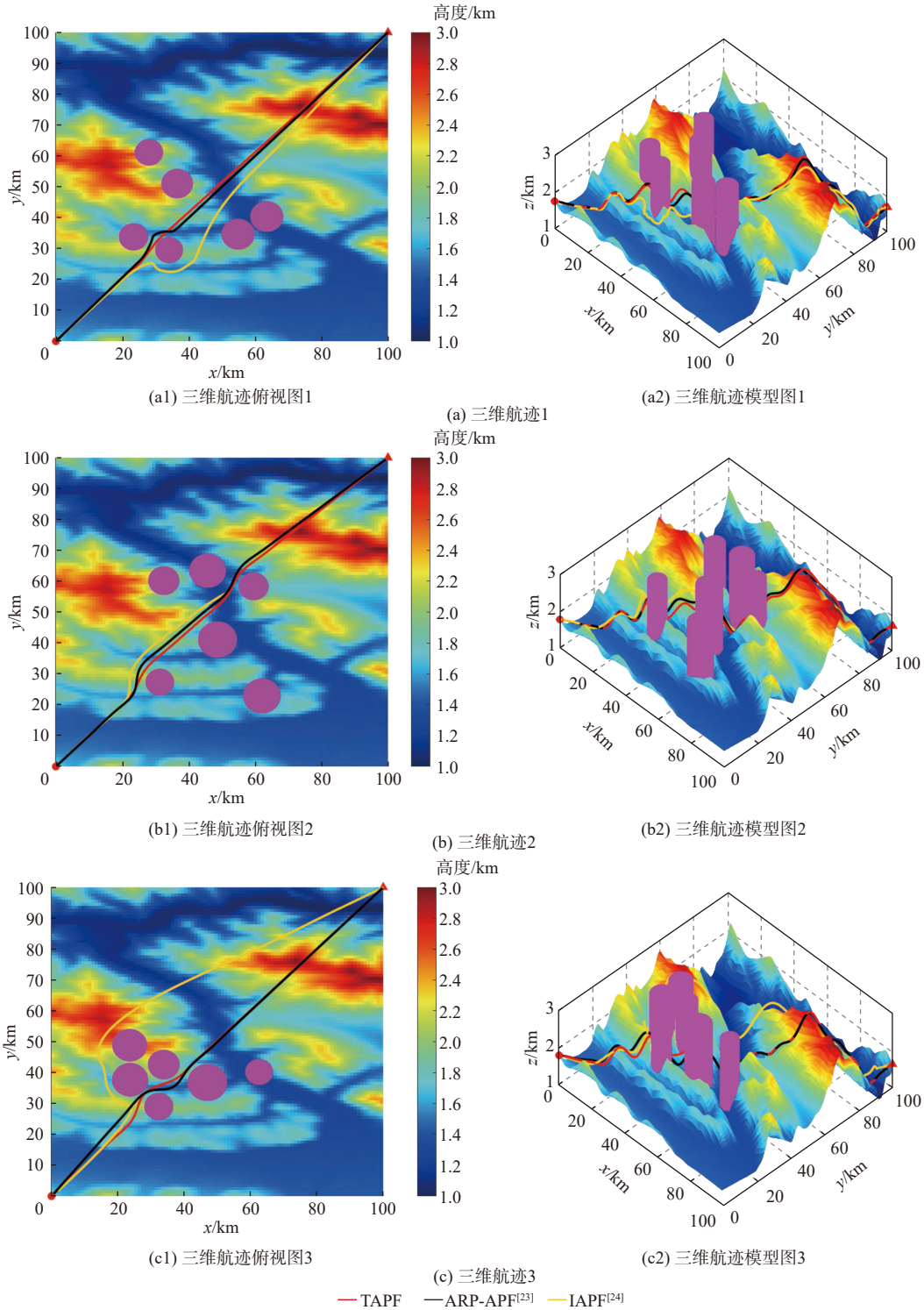


图 20　测试环境 2 中三维航迹图

Fig. 20　Three-dimensional trajectory map in test environment 2

表 4　测试环境 3 中指标对比

Table 4　Comparison of metrics in test environment 3

方法	M_s/%	M_l/km	M_t/s	M_{turn}/(°)
TAPF	100.0	294.5	0.74	31.10
IAPF[24]	97.4	337.1	0.62	67.03

长度还是平滑性方面都更好。

2 种方法对比的俯视图如图 21 所示。可以明显看出, 当进入到障碍物斥力势场的影响范围内时, 传统的斥力场将会一直对无人机产生斥力作用, 即使当无人机已经绕开障碍物, 前进方向已经没有威胁, 仍然会受到斥力影响, 导致产生多余的航迹。而对于改进的三角区势场当无人机绕开障碍物后, 前进方向没有障碍物时, 无人机即使处于斥力影响范围内, 但不受斥力影响, 保持直接飞向目标点的状态。因此, 从水平维度来看, 本文 TAPF

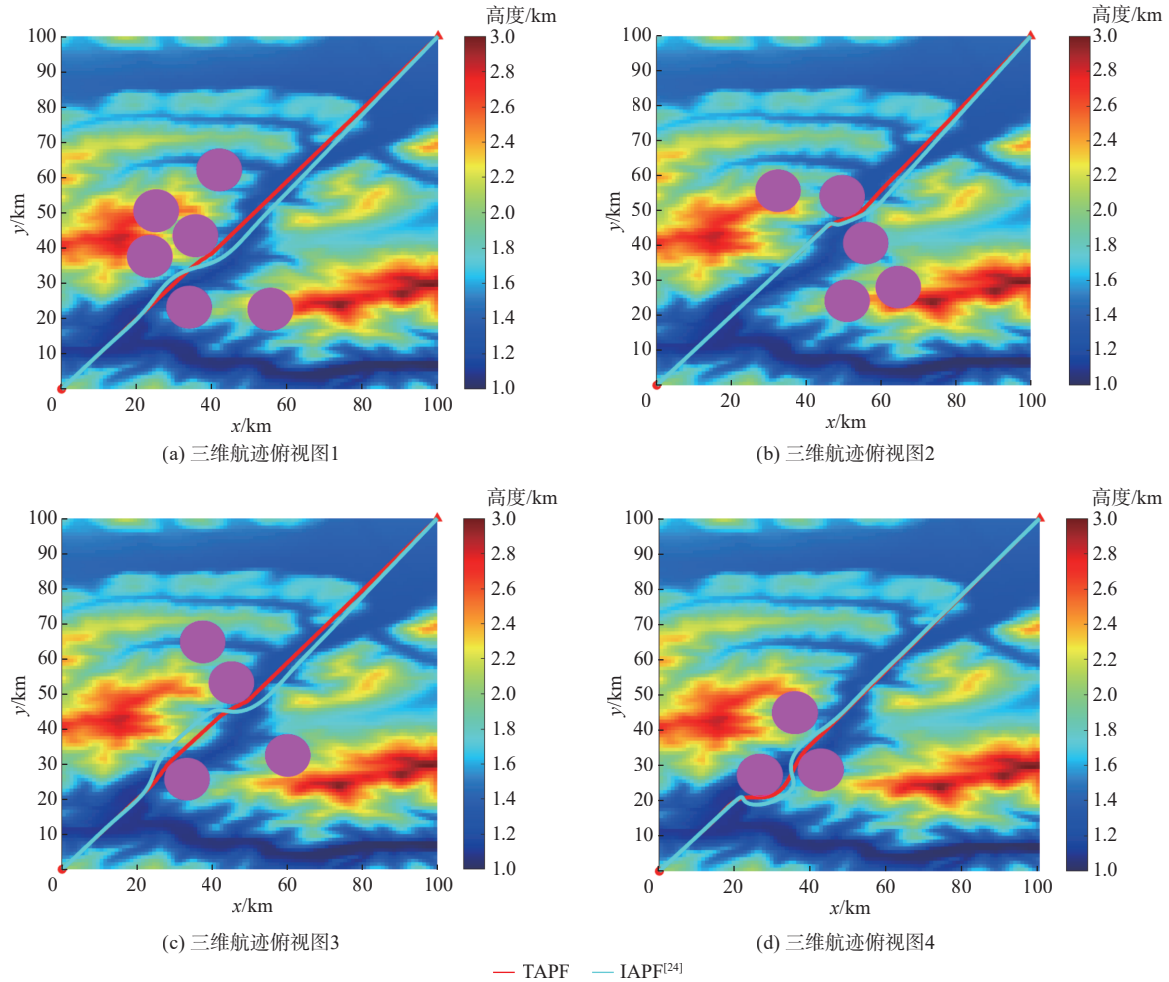


图 21 测试环境 3 中三维航迹俯视图

Fig. 21 Top view of 3D trajectory in test environment 3

方法中采用的三角区斥力势场机动幅度明显更小，路径具有更好的平滑性。

5 结 论

1）本文方法针对 APF 法势场函数设计困难的问题，利用强化学习学习斥力势场，消除了靠经验设计势场函数的不确定性；针对 APF 法适应性差的问题，通过强化学习得到适应性更广的斥力势场，提高了方法的适应性和优化性能；针对当前斥力势场使无人机绕飞过远，提出了三角区斥力势场，解决了无人机在已经绕开障碍物后无效避障增加航迹长度的问题。

2）仿真结果表明，本文方法提高了无人机在复杂环境中规划航迹的能力，在保证飞行成功率的前提下，同时缩短了航迹路程。在忽略自然干扰因素影响条件下，本文方法相较 ARP-APF 方法和 IAPF 法，在无人机航迹规划领域更具可行性。

参考文献（References）

[1] 杨孝鹏, 李伟春, 孙锐. 基于国家安全情报的海岸警卫队态势感知研究[J]. 飞航导弹, 2020(6): 87-91.
Yang X P, Li W C, Sun R. Research on situation awareness of coast guard based on national security intelligence[J]. Aerodynamic Missile Journal, 2020(6): 87-91(in Chinese).

[2] 白文超, 班明飞, 宋梦, 等. 电动汽车-无人机联合救援系统协调调度模型[J]. 上海交通大学学报, 2024, 58(9): 1443-1453.
Bai W C, Ban M F, Song M, et al. Coordinate scheduling model of electric vehicle-unmanned aerial vehicle joint rescue system[J]. Journal of Shanghai Jiao Tong University, 2024, 58(9): 1443-1453(in Chinese).

[3] 王飞, 杨清平. 面向多无人机物流配送的双层任务规划方法[J]. 北京航空航天大学学报, 2026, 52(1): 94-103.
Wang F, Yang Q P. Two-layer task planning method for multi-UAV logistics distribution[J]. Journal of Beijing University of Aeronautics and Astronautics, 2026, 52(1): 94-103(in Chinese).

[4] 武锦龙, 吴虹麒, 李浩, 等. 基于改进 DeepLabV^{3+}的荞麦苗期无人机遥感图像分割识别方法研究[J]. 农业机械学报, 2024, 55(5): 186-195.
Wu J L, Wu H Q, Li H, et al. Segmentation of buckwheat by UAV based on improved lightweight DeepLabV^{3+} at seedling stage[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(5): 186-195(in Chinese).

[5] 王超, 王银花. 一种改进 Dijkstra 算法的 UAV 路径规划[J]. 信息技术与信息化, 2021(10): 217-219.

Wang C, Wang Y H. An improved Dijkstra algorithm for UAV path planning[J]. Information Technology and Informatization, 2021(10): 217-219(in Chinese).

[6] 赵烈海, 李大鹏. 高密度场景下基于改进 A*算法的无人机路径规划[J]. 无线电通信技术, 2024, 50(4): 713-719.
Zhao L H, Li D P. Unmanned aerial vehicle path planning based on improved A* algorithm in high-density scenarios[J]. Radio Communications Technology, 2024, 50(4): 713-719(in Chinese).

[7] 杨玉, 金敏, 鲁华祥. 融合简化稀疏 A*算法与模拟退火算法的无人机航迹规划[J]. 计算机系统应用, 2019, 28(4): 25-31.
Yang Y, Jin M, Lu H X. UAV path planning based on fusion of simplified sparse A* algorithm and simulated annealing algorithm[J]. Computer Systems & Applications, 2019, 28(4): 25-31(in Chinese).

[8] Chen Y B, Luo G C, Mei Y S, et al. UAV path planning using artificial potential field method updated by optimal control theory[J]. International Journal of Systems Science, 2016, 47(6): 1407-1420.

[9] 程泽新, 李东生, 高杨. 一种改进遗传算法的无人机航迹规划[J]. 计算机仿真, 2019, 36(12): 31-35.
Cheng Z X, Li D S, Gao Y. GASA drone path planning to improve mutation strategy[J]. Computer Simulation, 2019, 36(12): 31-35(in Chinese).

[10] Meng X L, Zhu X J, Zhao J. Obstacle avoidance path planning using the elite ant colony algorithm for parameter optimization of unmanned aerial vehicles[J]. Arabian Journal for Science and Engineering, 2023, 48(2): 2261-2275.

[11] Yu Z H, Si Z J, Li X B, et al. A novel hybrid particle swarm optimization algorithm for path planning of UAVs[J]. IEEE Internet of Things Journal, 2022, 9(22): 22547-22558.

[12] 王庆海, 刘广瑞, 郭珂甫. 基于改进人工蜂群算法的无人机航迹规划研究[J]. 机床与液压, 2017, 45(21): 68-72.
Wang Q H, Liu G R, Guo K F. Research on path planning of UAV based on improved ABC algorithm[J]. Machine Tool & Hydraulics, 2017, 45(21): 68-72(in Chinese).

[13] 姬鹏, 郭明皓. 基于 Frenet 坐标下改进人工势场法的无人车局部路径规划[J]. 兵工学报, 2024, 45(7): 2097-2109.
Ji P, Guo M H. Local path planning for unmanned ground vehicles based on improved artificial potential field method in frenet coordinate system[J]. Acta Armamentarii, 2024, 45(7): 2097-2109(in Chinese).

[14] 张铮, 薛波, 柯子鹏, 等. 改进人工势场算法的路径规划[J]. 西安理工大学学报, 2024, 40(1): 27-35.
Zhang Z, Xue B, Ke Z P, et al. Path planning based on improved artificial potential field algorithm[J]. Journal of Xi'an University of Technology, 2024, 40(1): 27-35(in Chinese).

[15] 张菁, 何友, 彭应宁, 等. 基于神经网络和人工势场的协同博弈路径规划[J]. 航空学报, 2019, 40(3): 322493.
Zhang J, He Y, Peng Y N, et al. Neural network and artificial potential field based cooperative and adversarial path planning[J]. Acta Aeronautica et Astronautica Sinica, 2019, 40(3): 322493(in Chinese).

[16] 苗田恬, 方厚招, 安康, 等. 基于人工势场引力因子的机械臂避障规划算法[J]. 计算机工程与设计, 2024, 45(2): 578-586.
Miao T T, Fang H Z, An K, et al. Obstacle avoidance planning algorithm of robot arm based on attraction factor of artificial potential field[J]. Computer Engineering and Design, 2024, 45(2): 578-586(in Chinese).

[17] 刘光才, 马寅松, 齐福强, 等. 基于改进 A*-人工势场法的城市物流无人机路径规划[J]. 飞行力学, 2022, 40(6): 16-23.
Liu G C, Ma Y S, Qi F Q, et al. Flight path planning for urban logistics UAV based on improved A*-APF algorithm[J]. Flight Dynamics, 2022, 40(6): 16-23(in Chinese).

[18] 董炫良, 赵桂清. 人工势场引导蚁群算法的机器人导航路径规划[J]. 机械设计与制造, 2021(6): 169-173.
Dong X L, Zhao G Q. Robot navigation path planning based on ant colony algorithm guided by artificial potential field[J]. Machinery Design & Manufacture, 2021(6): 169-173(in Chinese).

[19] Guan W, Cui Z W, Zhang X K. Intelligent smart marine autonomous surface ship decision system based on improved PPO algorithm[J]. Sensors, 2022, 22(15): 5732.

[20] Zhou Y X, Shu J S, Hao H, et al. UAV 3D online track planning based on improved SAC algorithm[J]. Journal of the Brazilian Society of Mechanical Sciences and Engineering, 2023, 46(1): 12.

[21] 曹馨文, 时宏伟. 自适应斥力系数的无人机路径规划[J]. 计算机系统应用, 2023, 32(5): 36-44.
Cao X W, Shi H W. Path planning of UAV based on adaptive repulsion coefficient[J]. Computer Systems & Applications, 2023, 32(5): 36-44(in Chinese).

[22] 陈康雄, 刘磊. 基于扰动流体与 TD3 的无人机路径规划算法[J]. 电光与控制, 2024, 31(1): 57-62.
Chen K X, Liu L. A UAV path planning method based on interfered fluid and TD3 algorithm[J]. Electronics Optics & Control, 2024, 31(1): 57-62(in Chinese).

[23] Hao G Q, Lv Q, Huang Z, et al. UAV path planning based on improved artificial potential field method[J]. Aerospace, 2023, 10(6): 562.

[24] 王庆禄, 吴冯国, 郑成辰, 等. 基于优化人工势场法的无人机航迹规划[J]. 系统工程与电子技术, 2023, 45(5): 1461-1468.
Wang Q L, Wu F G, Zheng C C, et al. UAV path planning based on optimized artificial potential field method[J]. Systems Engineering and Electronics, 2023, 45(5): 1461-1468(in Chinese).

Adaptive repulsive potential field for UAV trajectory planning

GUO Jing[1], XIAN Yong[1,*], REN Leliang[1], LI Shaopeng[1,2]

(1. School of Missile Engineering, Rocket Force Engineering University, Xi'an 710025, China;
2. School of Automation, Tsinghua University, Beijing 100083, China)

Abstract: Aiming at the difficulty of potential field function design and the poor scene adaptability of the traditional artificial potential field (APF) method in multi-obstacle complex scenes, a three-dimensional self-adaptive repulsive potential field trajectory planning (TAPF) method based on the deep deterministic policy gradient (DDPG) algorithm is proposed. The adaptive repulsive potential field is used for trajectory planning. Firstly, a reinforcement learning-based repulsive potential field learning intelligence is established, which achieves "end-to-end" repulsive force output and improves the adaptability of the repulsive field. Secondly, dynamic triangles are designed to describe the range of the repulsive field, which effectively solves the problem of redundant paths in the traditional circular repulsive field guidance. Lastly, according to the characteristics of UAV trajectory planning, the state space, action space, and action space are designed for the TAPF method, and the proposed algorithm is based on the characteristics of unmanned aerial vehicle (UAV) trajectory planning. The algorithm's state space, action space, and reward function are created. According to simulation results, the proposed method has superior robustness to complicated scene adaptation and plans shorter courses with a better planning effect.

Keywords: unmanned aerial vehicle; trajectory planning; reinforcement learning; artificial potential field; neural network; obstacle avoidance

Received: 2024-07-25; **Accepted**: 2024-11-15; **Published Online**: 2024-12-06 10:59

URL: link.cnki.net/urlid/11.2625.V.20241205.1253.002

Foundation items: National Natural Science Foundation of China (62103432); Shaanxi University Science and Technology Association Young Talent Support Program (20210108)

* **Corresponding author**. E-mail: xy603xy@163.com