doi:10.3969/j.issn.1003-2029.2025.06.003

融合梯度自适应 A* 与DA-Q 学习的多无人船亚中尺度涡边缘搜索算法

李雯雯[1,2]，马 毅[1]，张飞飞[1,2]，李忠伟[2]

（1. 自然资源部第一海洋研究所，山东 青岛 266061；2. 中国石油大学（华东）海洋与空间信息学院，山东 青岛 266580）

摘 要：亚中尺度涡在海洋能量传输和生态调控中具有重要作用，其精准实时监测面临挑战。针对传统监测手段时效性不足、空间分辨率有限和观测平台灵活性差等问题，本文基于高机动性的多无人船平台，提出一种融合梯度自适应 A* 与动态奖励驱动的强化学习算法（Dynamic-Award Q-learning，DA-Q 学习）的多无人船亚中尺度涡边缘搜索算法。针对亚中尺度涡边缘形态复杂、环境动态变化快且信息不完整等具象化场景，算法利用梯度自适应 A* 引导多无人船迅速向涡旋区域收敛，并在搜索边缘过程中，构建温度梯度变化与历史路径信息融合的动态奖励机制，驱动 DA-Q 学习算法实现无人船自主适应环境变化并灵活调整航向。基于模拟数据与仿真海表面温度数据，设计了涵盖不同季节、无人船数量和初始配置的多组对比实验，分析了不同环境、不同涡旋类型和涡旋畸变条件下的算法性能和鲁棒性。结果表明：多无人船搜索亚中尺度涡边缘策略在路径长度与搜索效率方面优于单船方案；在部署方式上，多点分布式部署进一步提升了搜索覆盖能力与任务响应效率，优于单点母船释放式部署。本文结果可为多无人船在复杂海洋环境中的亚中尺度涡监测应用提供技术支撑。

关键词：亚中尺度涡；无人船；A* 算法；DA-Q 学习

中图分类号：P714 **文献标识码：**A **文章编号：**1003-2029（2025）06-0020-16

亚中尺度涡是一种重要的海洋动力过程，其空间尺度约为 0.1~10 km，时间尺度约为 0.1~10 d[1]，较中尺度涡小 1 至 2 个数量级[2]。亚中尺度涡通过打破地转平衡激发显著垂向流动，进而影响能量串级和垂向物质输运[3-5]，同时在调节生物地球化学过程及维持浮游生态系统方面发挥关键作用[5-6]。开展亚中尺度涡监测是深化区域海洋动力过程的基础。

近年来，亚中尺度涡成为海洋科学研究的热点，其监测方式主要包括卫星遥感和海基现场。卫星遥感能够大范围同步测量海面高度、海表温度等海洋参数。BECKER F 等[7]提出基于光学遥感影像的亚中尺度涡自动识别与特征提取方法。HOU H 等[8]使用卫星高度计数据提取了南海亚中尺度涡时空分布。然而，光学遥感影像易受云层覆盖、海面耀斑区反射等因素干扰，导致数据缺失或质量下降；卫星高度计数据空间分辨率有限，难以对小尺度、短生命期的亚中尺度涡精准监测[9-12]。海基平台主要有漂流浮标、水下滑翔机和自主式水下航行器等。崔虎山等[13]基于 Argos 漂流浮标数据识别亚中尺度涡，但在监测节点稀疏区或空白区易出现遗漏或误判问题。无人船是一种强机动监测节点，可以根据环境变化灵活调整航行路径，持续获取海洋观测数据。因此，无人船是开展亚中尺度涡边缘搜索的重要监测平台。与单无人船相比，多无人船作业可以

收稿日期：2025-06-11

基金项目：山东省自然科学基金资助项目（ZR2021QD006）

作者简介：李雯雯（2000—），女，硕士研究生，主要从事海洋无人系统智能监测研究。E-mail：li_wenwen11@163.com

通信作者：马毅（1973—），男，博士，研究员，主要从事海洋遥感与应用研究。E-mail：mayimail@fio.org.cn

提升空间覆盖能力与搜索鲁棒性。

多无人船高效精确地搜索亚中尺度涡，关键包括两个步骤：一是路径规划，即如何引导无人船到达亚中尺度涡边缘；二是边缘搜索，即如何引导无人船进行精准的边缘搜索。由于海洋环境复杂多变、目标区域不固定，传统路径规划算法难以适应实际需求。A* 算法是经典的启发式搜索算法，常用于寻找两个节点之间的最短路径[14]。ZHANG H 等[15]通过引入双向搜索策略和路径平滑方法，提升了 A* 算法的路径质量。然而，经典 A* 算法的代价函数不适应无人船搜索亚中尺度涡的场景，易导致搜索效率下降[16]。在边缘搜索阶段，涡旋边缘结构具有较强的不确定性和动态变化特征，传统搜索算法难以保证实时性与适应性。Q 学习是基于价值的强化学习算法[17]。JIANG J 等[18]通过模糊化状态变量结合集成学习策略，克服了传统 Q 学习在自由空间环境中的性能瓶颈。传统 Q 学习在边缘搜索中依赖固定的即时奖励更新 Q 值，难以有效响应环境中关键特征的变化，缺乏对路径方向性的引导[17,19]。这种机制易导致无人船在搜索亚中尺度涡过程中频繁切换航向，难以实现稳定连贯的边缘搜索。

针对上述应用场景和技术问题，本文提出融合梯度自适应 A* 与 DA-Q 学习的多无人船亚中尺度涡边缘搜索算法。首先，采用梯度自适应 A* 算法进行路径规划，实现对亚中尺度涡边缘的快速定位；然后，融合动态奖励驱动的 Q 学习算法，利用无人船实时获取的温度信息动态修正航向，实现高效且精确的亚中尺度涡边缘搜索。本文在模拟实验的基础上，利用仿真多无人船实时监测的高分辨率海温数据，开展亚中尺度涡边缘搜索实验。通过分析提出的算法在不同环境、不同涡旋类型及涡旋畸变条件下的表现，对其性能和鲁棒性进行了验证。研究结果可为多无人船在复杂海洋环境中的亚中尺度涡监测应用提供技术支撑。

1　数　据

1.1　区　域

南海是西北太平洋最大的边缘海，在季风、黑潮入侵等外源强迫，以及地形、地貌等内源机制的共同驱动下，形成了典型的多尺度动力结构，孕育出大量亚中尺度涡[20-22]。因此，本文以南海及其邻近海域（5°N—25°N，105°E—120°E）作为研究区域，开展亚中尺度涡边缘搜索实验。

1.2　数据与处理

目前，缺乏无人船实时监测的海表面温度数据。因此，本文基于遥感监测的海表面温度数据，构建仿真实验环境，模拟无人船在真实海洋环境中实时获取监测数据的过程。在仿真过程中，严格限制无人船获取信息的方式，使其仅能访问当前位置及历史路径上采集的海表温度。该设置模拟了无人船实际作业时获取数据的情境，确保算法在不依赖完整场景信息条件下进行路径规划和边缘搜索。

仿真数据基于高分辨率海表温度组（Group for High Resolution Sea Surface Temperature，GHRSST）框架下的 v4.1 高分辨率海表面温度产品构建(https://cmr.earthdata.nasa.gov/)[23]。首先，将获取的GHRSST海表面温度数据进行格式转换、区域裁剪及开尔文至摄氏度转换；其次，考虑到亚中尺度涡初定位和边缘搜索需要依赖温度梯度数据，采用高斯滤波处理去除数据高频噪声，提升数据的平滑性和边缘连续性；然后，为匹配无人船感知尺度，对数据进行重采样，重采样的空间分辨率为 10 m，将处理后的数据用于仿真多无人船亚中尺度涡边缘搜索实验。

为了提高亚中尺度涡边缘搜索效率，本文在多无人船搜索亚中尺度涡边缘之前，应用遥感数据进行引导，在 GHRSST 遥感数据中使用中心差分法计算水平温度梯度，通过设定阈值，获取亚中尺度涡分布的初步位置，从而确定多无人船亚中尺度涡边缘初始搜索方向。

2　研究方法

2.1　融合梯度自适应 A* 与 DA-Q 学习的多无人船亚中尺度涡边缘搜索算法

本文构建了一种多无人船亚中尺度涡边缘搜索算法，该算法主要包括两个步骤：首先，基于梯度自适应 A* 进行路径规划，通过融合局部温度梯度信息构建动态代价函数，引导无人船从分散初始位

置自主朝亚中尺度涡方向靠近，初步抵达边缘区域；然后，开展 DA-Q 学习的亚中尺度涡边缘搜索，该阶段设计了以梯度突变为驱动的自适应奖励机制，结合实时海表面温度数据动态修正无人船航向，引导无人船完成边缘的闭合搜索。整体算法结构如图 1 所示。

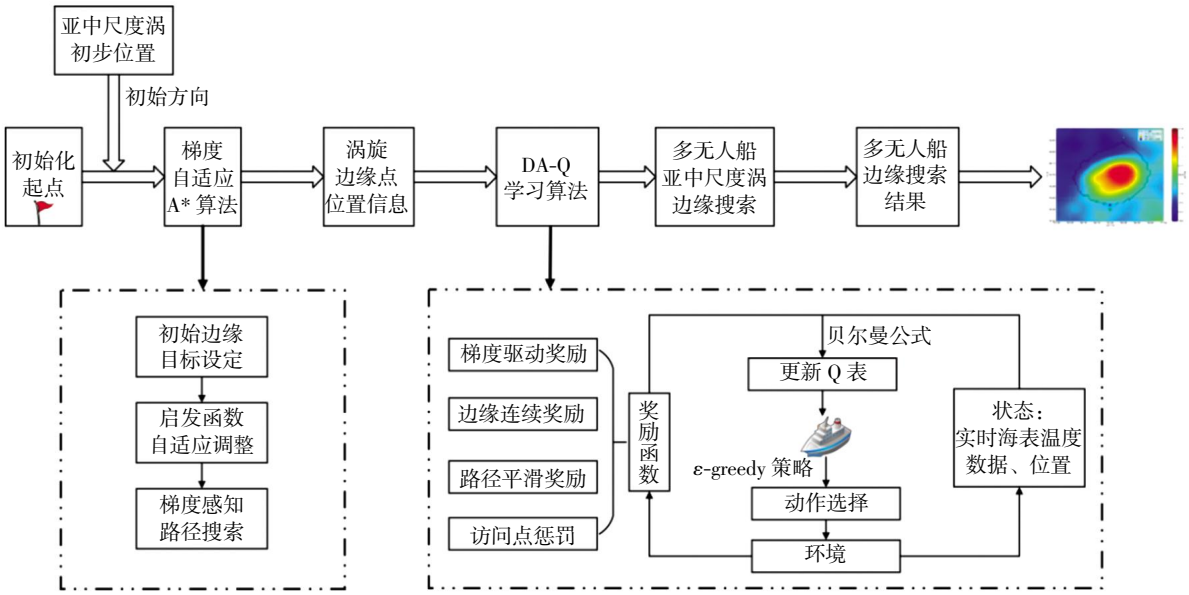


图 1 融合梯度自适应 A* 与 DA-Q 学习的多无人船亚中尺度涡边缘搜索算法结构

2.2 梯度自适应 A* 算法

经典 A* 算法通常以欧几里得距离作为代价函数，计算从起点到目标点的最短路径，但在实际海洋环境中，亚中尺度涡边界往往具有不规则形态，并可能包含多个潜在目标区域，这对路径规划提出了更高要求，须具备更强的环境适应性与目标判别能力。

为了解决上述问题，本文引入基于海表温度梯度方向的动态代价函数，使路径规划过程能够结合局部监测信息自适应调整方向，引导无人船优先朝向温度变化显著区域航行，逐步逼近潜在的亚中尺度涡边缘。具体地，在代价函数评估过程中，增加了基于当前位置温度梯度方向与路径方向一致的引导项，使路径不仅考虑几何最短距离，同时兼顾边缘特征的空间分布。温度梯度向量 $\boldsymbol{t}$ 可以表示如下。

$$\boldsymbol{t} = \left(\frac{\partial T}{\partial x}, \frac{\partial T}{\partial y} \right) \tag{1}$$

式中，T 为无人船当前位置获取的海表面温度数据；x 和 y 分别为无人船当前位置的经度和纬度。

对于当前节点 n 与某个邻节点 n'，其移动方向角 $\boldsymbol{d}$ 计算如下。

$$\boldsymbol{d} = (y_{n'} - y_n, x_{n'} - x_n) \tag{2}$$

式中，x_n 和 y_n 分别为当前位置的经度和纬度；$x_{n'}$ 和 $y_{n'}$ 分别为邻节点的经度和纬度。

由式(1)和式(2)可得当前位置温度梯度方向与邻接节点方向之间的夹角 θ 满足式(3)。

$$\cos\theta = \frac{\boldsymbol{t} \cdot \boldsymbol{d}}{\|\boldsymbol{t}\| \cdot \|\boldsymbol{d}\|} \tag{3}$$

本文提出了基于海表温度梯度方向的动态代价函数。

$$f(n) = \lambda_1 \times g(n) + \lambda_2 \times \cos\theta \tag{4}$$

式中，λ_1、λ_2 为调整权重，$\lambda_1 + \lambda_2 = 1$；$g(n)$ 为起点到当前位置的欧几里得距离。

梯度自适应 A* 算法流程如图 2 所示，展示了无人船从初始位置逐步逼近亚中尺度涡边缘的过程。

2.3 DA-Q 学习算法

传统 Q 学习通常采用离散状态空间与固定奖励机制。在实际的亚中尺度涡边缘搜索任务中，环境状态呈现非线性和动态变化特征，传统 Q 学习在状态建模和奖励设计方面存在泛化能力弱、收敛速度慢等问题。因此，本文基于温度梯度与路径信息融合，提出了动态奖励驱动的 Q 学习算法，通过改进状态表示、方向性动作策略与动态奖励函数，有效

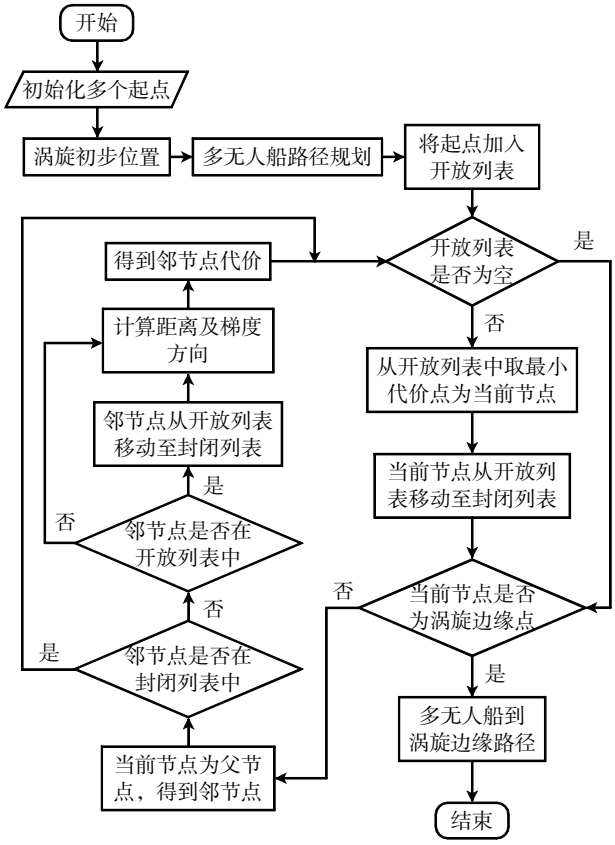


图 2　梯度自适应 A* 算法流程

提升了多无人船在边界搜索任务中的收敛速度与策略鲁棒性。

2.3.1　状态空间

状态空间是所有可能的状态集合。在本文中，状态空间表示在搜索亚中尺度涡边缘的过程中，无人船所处的当前位置，状态 s_t 定义如下。

$$s_t = [(x,y), T(x,y), G_t] \tag{5}$$

式中，(x,y)为无人船所处当前位置的经纬度坐标；$T(x,y)$为当前位置的海表面温度；G_t 为当前位置的温度梯度。

2.3.2　动作空间

动作空间描述了在每个状态下，无人船可以采取的所有可能动作。与传统的 4 个方向相比，本文定义了 8 个方向，显著提升了无人船在复杂环境的探索能力。由于无人船能够沿对角线进行移动，这不仅能缩短路径长度，还减少了由于仅沿 4 个方向航行可能产生的冗余步骤，从而提升算法的效率和边缘搜索的精度。动作空间 a_t 具体定义如图 3 所示。

2.3.3　多模态动态奖励函数

奖励函数定义了每个状态–动作对应的即时奖

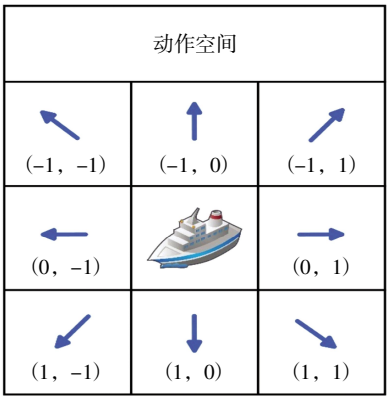


图 3　栅格地图上的动作空间

励，奖励为正时，鼓励智能体采取某个动作，奖励为负时，惩罚智能体采取某个动作。针对传统 Q 学习的奖励函数在复杂海洋环境中的单一距离反馈和缺乏场景化引导的局限，本文提出多模态动态奖励机制，通过融合梯度突变、边缘连续性、访问点惩罚和路径平滑性 4 个关键因子，实现多无人船的精细化引导，提出的奖励函数设计如下。

$$R = R_{\text{grad}} + R_{\text{cont}} + R_{\text{revisit}} + R_{\text{smooth}} \tag{6}$$

式中，R_{grad}、R_{cont}、R_{revisit} 和 R_{smooth} 分别为梯度驱动奖励、边缘连续性奖励、访问点惩罚和路径平滑性奖励函数。

（1）梯度驱动奖励

利用当前监测温度与无人船已访问节点的温度记录，基于局部温度变化趋势构建动态梯度感知，引导无人船进行路径决策，设 $T_t(x,y)$为当前位置的海表面温度；G_t 为当前位置的温度梯度。

$$G_t = \sqrt{\left(\frac{\partial T}{\partial x}\right)^2 + \left(\frac{\partial T}{\partial y}\right)^2} \tag{7}$$

则梯度驱动奖励函数定义如下。

$$R_{\text{grad}} = \begin{cases} +g_1, & G_t \geqslant T_g \\ -g_2, & G_t < T_g \end{cases} \tag{8}$$

式中，T_g 为温度梯度阈值；g_1 和 g_2 分别为奖励和惩罚权重参数。该设计不仅利用了边缘空间特征，还引入时间演化信息，使得奖励机制更具有时序感知能力与边缘适应性。

温度梯度阈值 T_g 的选取对亚中尺度涡边缘的搜索结果有显著影响。由于不同海域及不同时段的亚中尺度涡存在明显差异，固定阈值难以适应所有场景，进而导致亚中尺度涡搜索结果存在差异。因此，本文提出一种基于统计先验与在线反馈的动态

自适应阈值机制：基于南海近十年海表温度数据统计得到的温度梯度分布，并参考相关文献中的典型阈值范围[24-26]，设定初始梯度阈值范围为0.001~0.006；在搜索过程中，算法综合分析无人船路径上的温度梯度变化与局部边缘闭合性，若涡旋边缘闭合率不足，则提升阈值以增强边缘约束。每轮路径评估后，阈值在设定区间内以固定步长（$\Delta = 0.001$）进行微调，直至满足边缘搜索的稳定性判据。该机制能够在不同海域和不同背景梯度环境下实现自适应阈值选择，提升算法的鲁棒性与泛化能力，避免了固定阈值带来的适应性不足问题。

（2）边缘连续性奖励

考虑到亚中尺度涡边缘的空间连续性特征，本文设计了边缘连贯性判别函数，统计当前位置的八邻域中是否至少含有三个相同的邻节点。若满足该条件，则给予奖励。

$$R_{\text{cont}} = \begin{cases} +c_1, & \text{邻节点} \geqslant 3 \\ 0, & \text{邻节点} < 3 \end{cases} \tag{9}$$

式中，c_1 为奖励参数。该设计鼓励无人船沿连续的边缘结构前进，避免跳跃式、断裂式的路径生成。

（3）访问点惩罚

为了提高搜索效率，避免在有限步长内频繁访问同一位置，设定对已有访问点的惩罚与新探索点的正向激励。

$$R_{\text{revisit}} = \begin{cases} +v_1, & s_t \text{不在} visited \text{集合里} \\ -v_2, & s_t \text{在} visited \text{集合里} \end{cases} \tag{10}$$

式中，v_1 和 v_2 分别为奖励和惩罚参数。$visited$ 为已走过的路径集合，该机制能有效降低重复搜索概率，提升路径覆盖率。

（4）路径平滑性奖励

亚中尺度涡边缘通常呈现环状或者流线结构，若移动方向频繁剧烈变化，容易导致路径折返或者偏离边缘。因此，设计了方向一致性奖励变量。定义当前动作方向(x,y)与前一步方向$(x_{\text{prev}},y_{\text{prev}})$的变化程度 Δ_{angle} 如下。

$$\Delta_{\text{angle}} = |\boldsymbol{dx} - \boldsymbol{dx}_{\text{prev}}| + |\boldsymbol{dy} - \boldsymbol{dy}_{\text{prev}}| \tag{11}$$

该度量采用 L1 范数形式。其中，$(\boldsymbol{dx},\boldsymbol{dy})$为当前时间步无人船的动作方向分量；$(\boldsymbol{dx}_{\text{prev}},\boldsymbol{dy}_{\text{prev}})$为前一时刻无人船的动作方向分量。

则路径平滑性奖励函数 R_{smooth} 定义如下。

$$R_{\text{smooth}} = -s_1 \times \Delta_{\text{angle}} \tag{12}$$

式中，s_1 为平滑奖励参数。该奖励鼓励无人船以最小的方向持续移动，有助于生成更合理、光滑的路径。

2.3.4 动态探索因子

传统 Q 学习采用 ε-greedy 策略进行动作选择，ε-greedy 策略中固定的 ε 值会因训练初期探索不足或后期探索过多影响策略收敛。在亚中尺度涡边缘搜索任务中，边缘具有梯度突变与形态复杂等特征。因此，本文设置每五次迭代更新一次 ε 值，每次衰减 10%并设置下限约束。动态 ε 策略有助于无人船在关键区域保持适度探索，提升边缘搜索的完整性与鲁棒性。

2.4 实验结果评价指标

采用路径长度、搜索时间作为评价指标，对不同环境条件下的亚中尺度涡搜索算法的性能进行评估。

（1）路径长度

在实验中，每艘无人船的路径由两段组成：第一段是无人船航行到亚中尺度涡边缘的路径，第二段是边缘搜索的路径。路径总长度 L_{path} 通过式(13)计算。

$$L_{\text{path}} = N_{\text{straight}} \times d + N_{\text{diag}} \times d \times \sqrt{2} \tag{13}$$

式中，N_{straight} 为直线移动次数；N_{diag} 为斜线移动次数；d 为每次直线移动的欧式距离。

（2）搜索时间

无人船的搜索时间是指从起点到完成边缘搜索所消耗的时间。在计算搜索时间时，假设无人船的速度是恒定的，每艘无人船搜索时间的计算公式表示如下。

$$T_{\text{search}} = \frac{L_{\text{path}}}{v} \tag{14}$$

式中，T_{search} 为搜索时间；L_{path} 为路径长度；v 为无人船的平均速度。

对于亚中尺度涡边缘搜索任务完成时间 T_{total}，由多艘无人船中最慢单船完成边缘搜索的时间决定。

$$T_{\text{total}} = \max(T_{\text{search}(i)}),\ i \in \{1,\cdots,N\} \tag{15}$$

式中，N 为无人船数量；$T_{\text{search}(i)}$为第 i 艘无人船搜索时间。

（3）搜索成功率

采用搜索成功率 R_{sucess} 衡量算法在亚中尺度涡

边缘搜索任务中的稳定性和有效性，其定义如下。

$$R_{\text{sucess}} = \frac{N_{\text{sucess}}}{N_{\text{total}}} \times 100\% \tag{16}$$

式中，N_{sucess} 为在给定条件下无人船成功搜索涡旋边缘闭合路径的实验次数；N_{total} 为总实验次数。

3　实验结果与分析

为验证本文算法在不同环境条件下的准确性与有效性，本文设计了两类实验：第一类是基于模拟数据开展的边缘搜索实验，第二类是基于仿真无人船监测数据的南海亚中尺度涡边缘搜索实验。

3.1　模拟数据实验

3.1.1　实验环境与参数设置

采用离散化网格方式构建理想化的亚中尺度涡形状，模拟涡旋形状为圆形、椭圆形和不规则圆形。模拟数据实验参数设置如表 1 所示。

表 1　模拟数据实验参数设置

参数	数值
栅格地图的大小	1 600×1 600
网格的分辨率/m	10×10
无人船数量/艘	2
无人船速度/(m·s^{-1})	10
无人船初始位置	多船从分散位置出发

3.1.2　不同涡旋形态的边缘搜索结果与分析

为了测试不同涡旋形态对搜索效率的影响，本文设置了三种典型的理想涡旋边缘形态：圆形涡旋、椭圆形涡旋及不规则圆形涡旋。通过表 2 的实验结果可以得出，不同的涡旋形态对边缘搜索的效率和准确性影响不同。随着亚中尺度涡边缘形态的复杂性增加，边缘搜索效率逐渐降低。

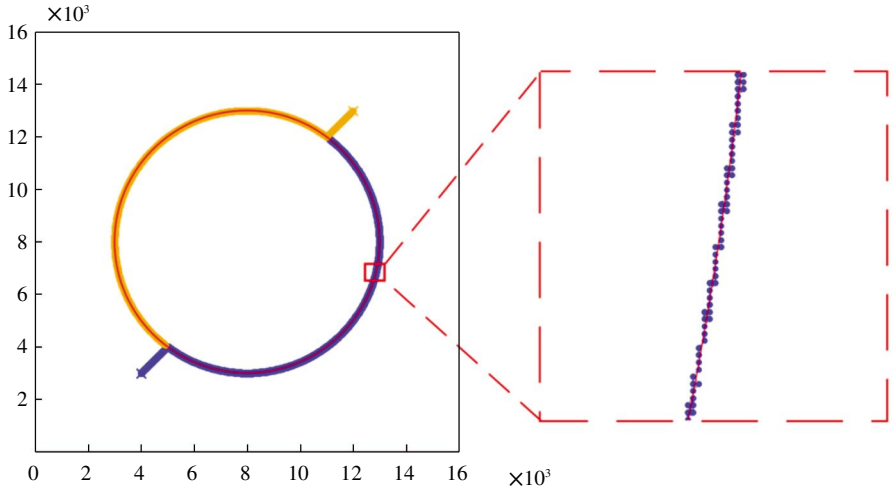
表 2　涡旋参数设置

涡旋形态	圆形	椭圆形	不规则圆形
参数/km	直径：10	长半轴：6 短半轴：4	由长半轴：6，短半轴：4的椭圆引入扰动生成
起点 1 坐标	（400，300）	（100，800）	（100，850）
起点 2 坐标	（1 200，1 300）	（1 500，1 000）	（1 500，1 000）
起点 1 路程/km	70.93	82.23	87.24
起点 2 路程/km	49.79	40.38	41.94
起点 1 搜索时间/h	1.97	2.28	2.42
起点 2 搜索时间/h	1.38	1.12	1.17

如图 4 所示，在圆形涡旋的实验场景中，由于边缘规则且对称，本文算法能够高效地进行边缘搜索，几乎无冗余动作。相较于圆形涡旋，椭圆形涡旋搜索时间稍微延长，但边缘识别的准确性保持较高。对于不规则圆形涡旋，涡旋的边缘不再是规则的几何图形，变得更加复杂，增加了搜索过程的不确定性，本文算法通过自适应调整，能够有效地搜索不规则边缘，表现出较强的鲁棒性。

3.1.3　多无人船初始位置分布对边缘搜索结果的影响

为进一步验证边缘搜索算法的鲁棒性，设定三种典型的无人船初始位置分布方式：多船全部位于涡旋外部、全部位于内部和分散部署在涡旋内外区域。旨在评估算法在多无人船起点空间布局相异条件下的适应能力，验证算法是否具有良好的起点无关性，能够在不依赖预设优选部署条件的前提下，实现边缘的准确搜索与协同到达。



（a）圆形涡旋

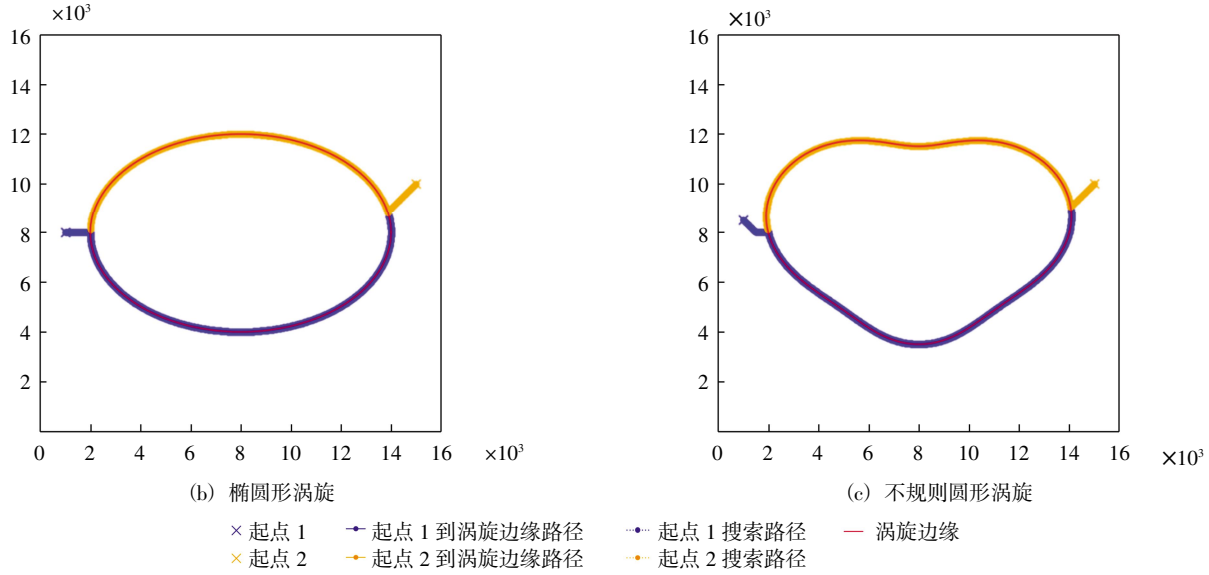


图 4 不同环境形状实验结果

根据表 3 的定量统计结果可知，多无人船的初始位置分布直接影响边缘搜索的效率和准确性，合理地部署可以显著提高搜索性能。

表 3 多无人船实验结果统计

无人船初始位置	涡旋内部	涡旋外部	涡旋内外区域
起点 1 坐标	(500,700)	(200,800)	(200,800)
起点 2 坐标	(1 100,1 000)	(1 400,1 200)	(1 100,1 000)
起点 1 路程/km	82.03	88.10	81.61
起点 2 路程/km	49.29	42.81	36.63
起点 1 搜索时间/h	2.28	2.45	2.27
起点 2 搜索时间/h	1.37	1.19	1.02

如图 5 所示，针对不同初始位置分布的实验结果表明，算法在多种部署场景下均展现出较强的适应性与协同能力。在不同初始部署方式下，无人船可能位于涡旋内、外或随机分布，所提算法通过融

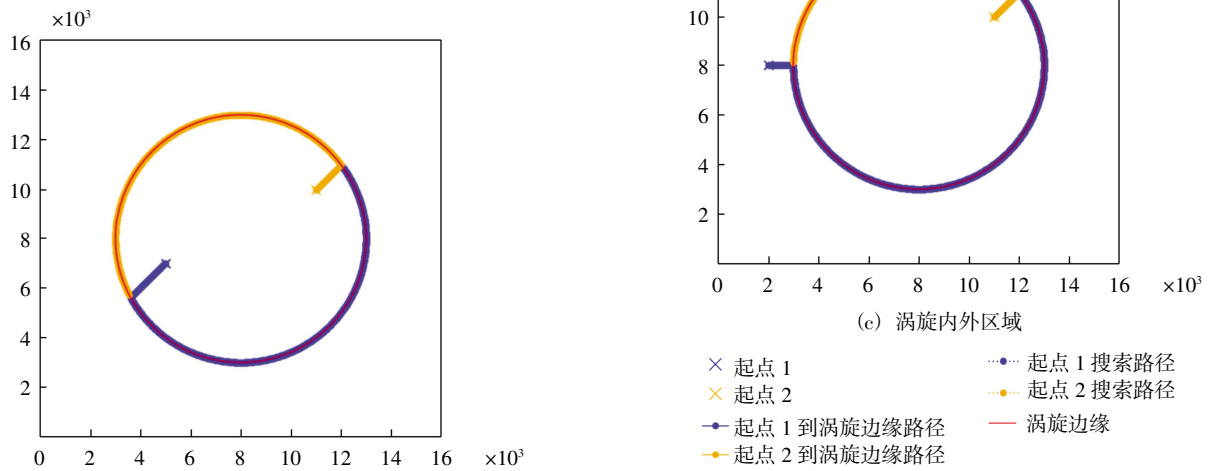


图 5 不同初始位置实验结果

合梯度自适应 A* 的全局引导能力与 DA-Q 学习的局部策略优化，实现了对边缘方向的快速判断、路径重复的有效规避与搜索效率的持续提升，展现出良好的鲁棒性与协同搜索能力。

3.2　仿真多无人船监测数据实验

3.2.1　实验场景与参数设置

基于高分辨率海表温度数据构建的仿真实验场景，开展仿真多无人船边缘搜索实验。实验中，多无人船以历史路径信息与当前位置温度监测数据为基础，采用融合梯度自适应 A* 与 DA-Q 学习算法进行亚中尺度涡边缘搜索。仿真实验的参数设定如表 4 所示。

表 4　仿真实验参数设置

参数	数值
最大迭代轮数/轮	50
每轮迭代最大步数/步	20 000
学习率 α	0.1
折扣因子 γ	0.95
初始探索率 ε	0.3
无人船数量/艘	1、2
无人船速度/(m·s^{-1})	10

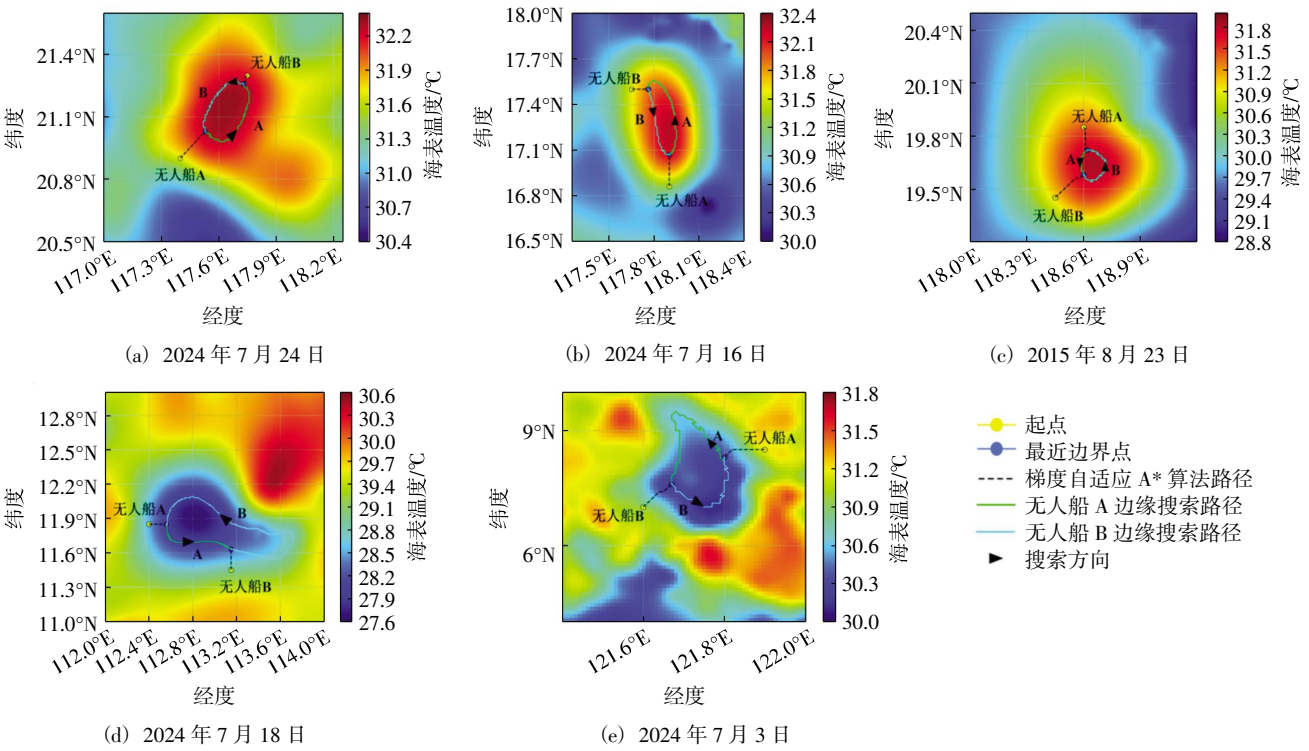
3.2.2　不同季节的亚中尺度涡边缘搜索结果与分析

通过选择 2015—2025 年夏季和冬季的数据，测试多无人船边缘搜索算法在不同季节海洋环境下的适应性和鲁棒性。夏季和冬季亚中尺度涡的搜索结果分别如图 6 和图 7 所示，实验结果如表 5 所示。

通过图 6、图 7 和表 5 可以得出，夏季和冬季的暖涡梯度分布更为集中，搜索路径较短且闭合性更强；而冷涡边界梯度过渡相对平缓，导致搜索路径和时间增加。进一步可以得出，若无人船起始点位于涡旋远侧或梯度弱区，容易导致搜索路径回绕或增加，影响整体效率，因此需要设置合理的起始部署策略，从而有效提升搜索鲁棒性与任务协同效率。

3.2.3　亚中尺度涡边缘搜索结果验证与分析

为评估本文方法与传统水平温度梯度法对边缘搜索的效果，选取多个典型样本进行对比分析，结果如图 8 所示，可以看出，本文方法在所有样本中均实现路径闭合，且不依赖于整场数据，适合实时监测任务；而水平梯度法需全图数据输入。两者适用于不同应用场景，本文方法在自主性与部署灵活性方面具有明显优势。此外，从梯度法实验结果来

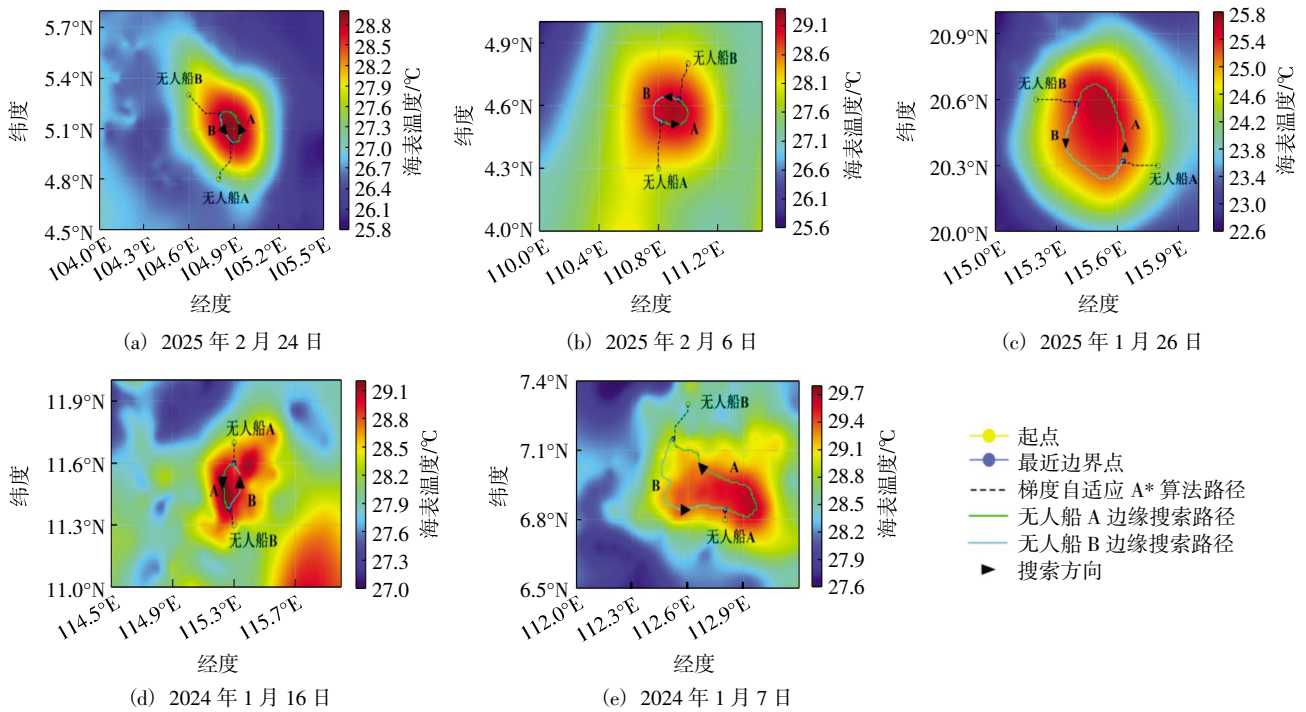


(a) 2024 年 7 月 24 日　(b) 2024 年 7 月 16 日　(c) 2015 年 8 月 23 日

(d) 2024 年 7 月 18 日　(e) 2024 年 7 月 3 日

路线是两艘无人船的搜索航行闭合路径，去除梯度自适应 A* 算法路径后，为搜索得到的亚中尺度涡边界；背景栅格图是海表面温度数据，其中有亚中尺度涡，该图与无人船搜索路径叠加，用于表现无人船搜索的效果。

图 6　夏季亚中尺度涡边缘搜索结果



(a) 2025 年 2 月 24 日　(b) 2025 年 2 月 6 日　(c) 2025 年 1 月 26 日

(d) 2024 年 1 月 16 日　(e) 2024 年 1 月 7 日

图 7　冬季亚中尺度涡边缘搜索结果

表 5　夏季和冬季多无人船实验结果统计

季节	涡旋属性	日期	无人船	路程统计/km	搜索时间/h
夏季	暖涡	2024-07-24	船 A	86.33	2.40
			船 B	50.29	1.41
		2024-07-16	船 A	135.41	3.76
			船 B	75.72	2.10
	冷涡	2015-08-23	船 A	32.91	0.91
			船 B	73.20	2.03
		2024-07-18	船 A	104.32	2.90
			船 B	305.59	8.49
		2024-07-03	船 A	43.26	1.20
			船 B	57.85	1.61
		2025-02-24	船 A	72.74	2.02
			船 B	53.79	1.49
冬季	暖涡	2025-02-06	船 A	64.35	1.79
			船 B	57.5	1.60
		2025-01-26	船 A	97.89	2.72
			船 B	102.91	2.86
		2024-01-16	船 A	43.18	1.20
			船 B	48.35	1.34
		2024-01-07	船 A	112.06	3.11
			船 B	102.34	2.84

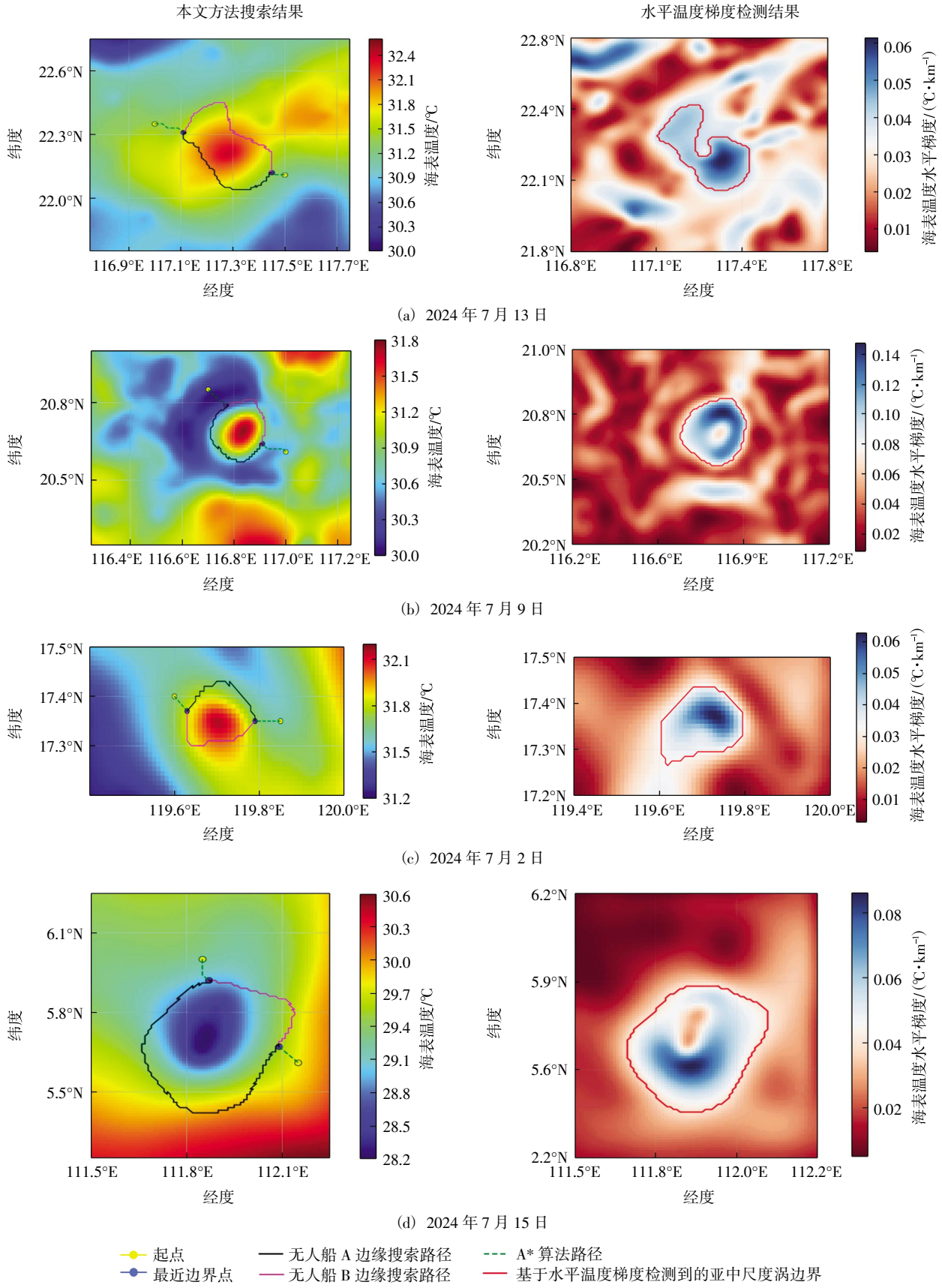
看，本文方法在边缘搜索路径上与梯度变化趋势高度一致，验证了其在边缘搜索中的合理性，为算法的有效性提供了理论依据。

由表 6 可知，本文方法基于无人船当前位置及局部温度梯度进行动态搜索，能够完整地搜索出亚中尺度涡边缘，与基于全图数据的传统水平温度梯度法检测结果一致，说明本文方法能有效、准确地搜索亚中尺度涡。

3.2.4　无人船初始部署对搜索结果的影响

多无人船搜索的效果不仅依赖于所采用的搜索算法，还与初始的无人船部署密切相关。因此，为评估算法在不同部署条件下的适应性和搜索效率，设计了两种典型的多无人船初始部署方式：一是多点分布部署，多艘无人船从不同位置同时出发，从不同方向覆盖搜索区域；二是单点母船释放式部署，一艘母船作为中央控制平台独立搜索涡旋区域，待母船搜索到亚中尺度涡边缘点后，再释放两艘子船，由子船接力执行精细化边缘搜索。如图 9 所示，多点分布部署在路径闭合能力、路径优化方面优于单点部署。

实验结果如表 7 所示，多点分布部署在路径优化与搜索效率方面优于单点母船释放式部署。多点部署具备广域覆盖与并行搜索优势，无人船从多个方向同时逼近涡旋区域，整体搜索路径更短，多船协同效率更高。相比之下，单点部署虽能集中资源由母船主导搜索，但在母船引导与子船释放过程中存在阶段性延迟，且路径高度依赖母船的先验判断，若初始搜索存在偏差，则整体搜索效率显著下降。此外，子船投放后的路径规划仍需重新适应局部梯度环境，降低搜索效率。因此，在多



左侧图为本文方法的搜索结果，其中闭合轨迹为不依赖全局数据搜索获得，表示亚中尺度涡边缘。右侧图为传统水平梯度法，基于整幅图的温度场数据检测的亚中尺度涡边缘结果。

图 8　边缘检测效果对比

表 6　两种算法性能对比

涡旋属性	日期	本文方法涡旋长度/km	梯度检测涡旋长度/km	梯度检测涡旋长度/本文方法
暖涡	2024-07-13	158.8	159.74	100.59%
暖涡	2024-07-09	99.84	91.64	91.77%
暖涡	2024-07-02	70.17	65.52	93.37%
冷涡	2024-07-15	207.28	166.15	80.18%

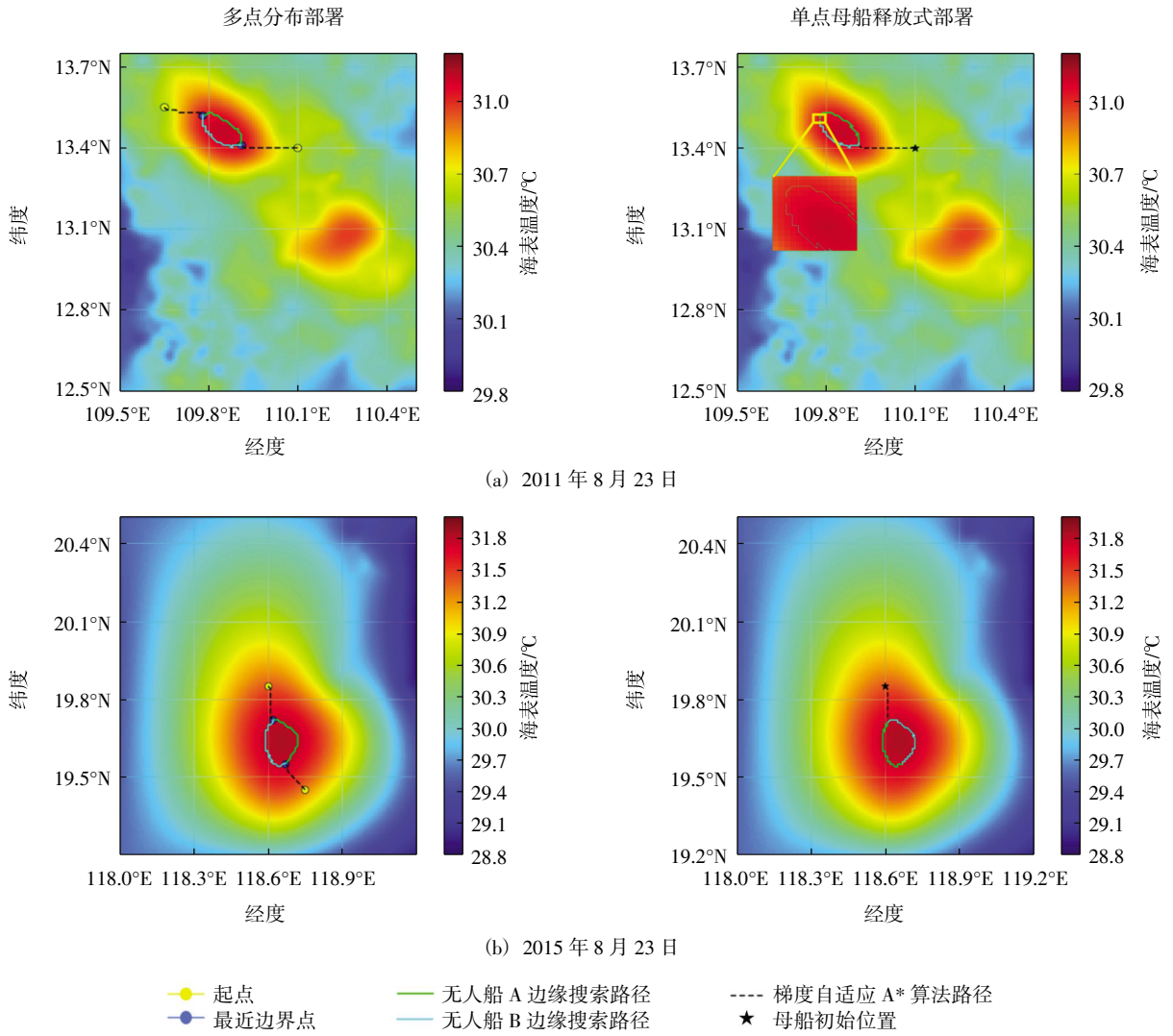


图 9　不同部署方式实验效果

表 7　多无人船实验结果统计

日期	部署方式	搜索完成时间/h
2011-08-23	多点分布式	1.13
	单点母船释放式	1.42
2015-08-23	多点分布式	1.21
	单点母船释放式	1.33

无人船搜索亚中尺度涡任务中，宜采用多点分布部署策略。

3.2.5　不同数量无人船对比

为验证多无人船在亚中尺度涡边缘搜索中的优势，分析了不同数量无人船对相同亚中尺度涡旋边缘搜索效率的影响。实验结果如表 8 和图 10 所示，多无人船配置在整体搜索效率上明显优于单船方案，搜索完成时间约为单船的一半。同时，多船协同能够实现更均匀的边缘覆盖，减少路径重叠与盲区，提高了资源利用率和算法稳定性。

表 8　不同无人船数量实验结果统计

日期	本文方法（多船）搜索时间/h	单无人船搜索时间/h
2024-07-05	1.97	3.38
2024-01-14	0.64	1.09
2022-01-19	1.14	2.19
2011-08-23	1.42	2.05

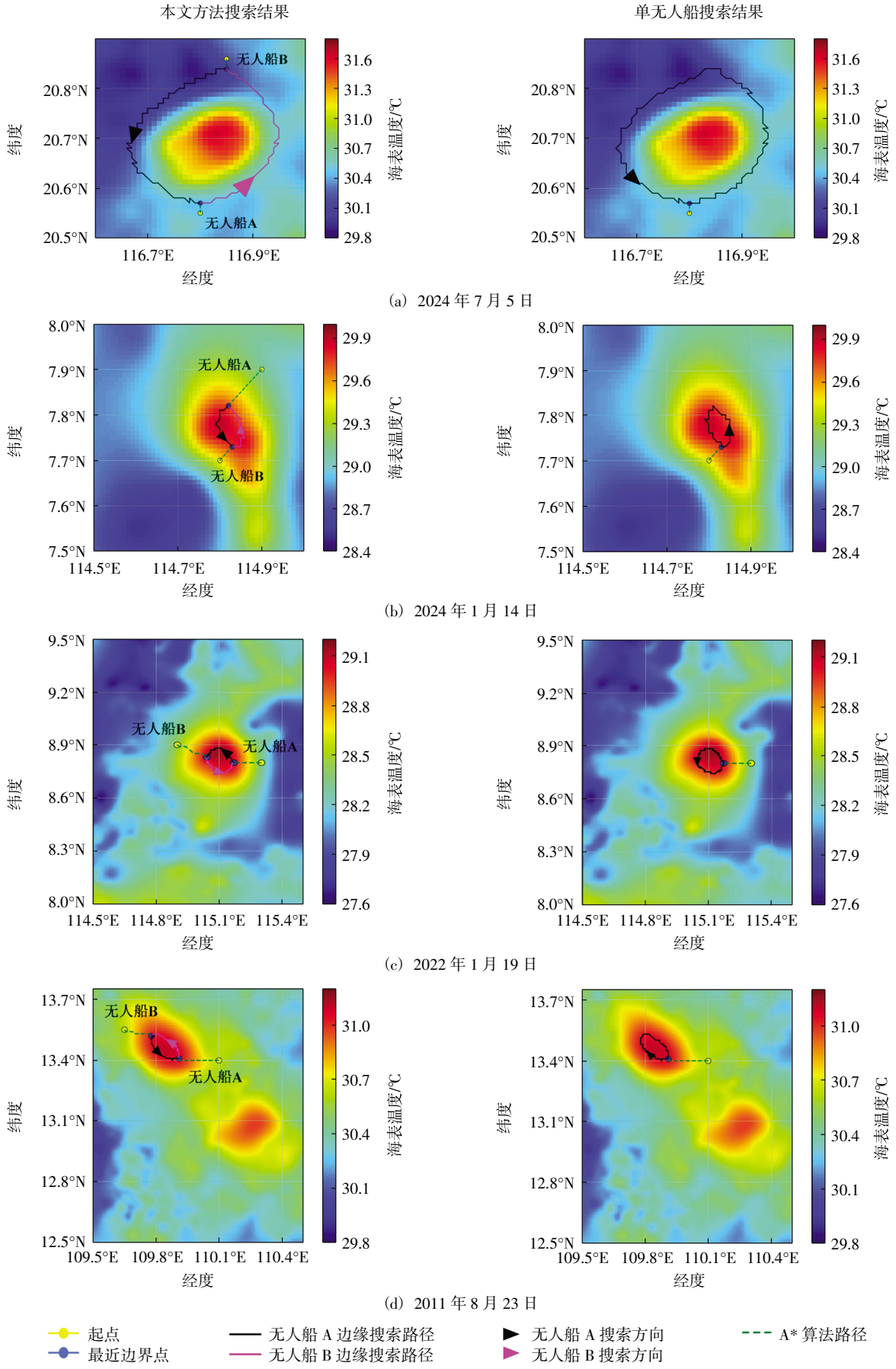


图 10　不同无人船数量实验效果

3.2.6 不同边缘搜索算法对比

为进一步验证融合梯度自适应 A* 与 DA-Q 学习算法在亚中尺度涡边缘搜索任务中的有效性，本文选取典型的深度确定性策略梯度算法（Deep Deterministic Policy Gradient，DDPG）与近端策略优化算法（Proximal Policy Optimization，PPO）作为对比算法。综合评估了不同算法在搜索成功率、搜索完成时间等方面的表现。实验设置每种算法在相同初始条件下分别独立运行 20 次，并以公式(16)定义的搜索成功率作为关键指标，统计闭合路径形成的稳定性。实验结果如图 11 所示，DDPG 算法在部分场景下出现航迹震荡和偏离现象；PPO 通常能够形成闭合，整体收敛较稳，但相较 DA-Q 学习在部分样本中存在较多绕行与路径冗余。

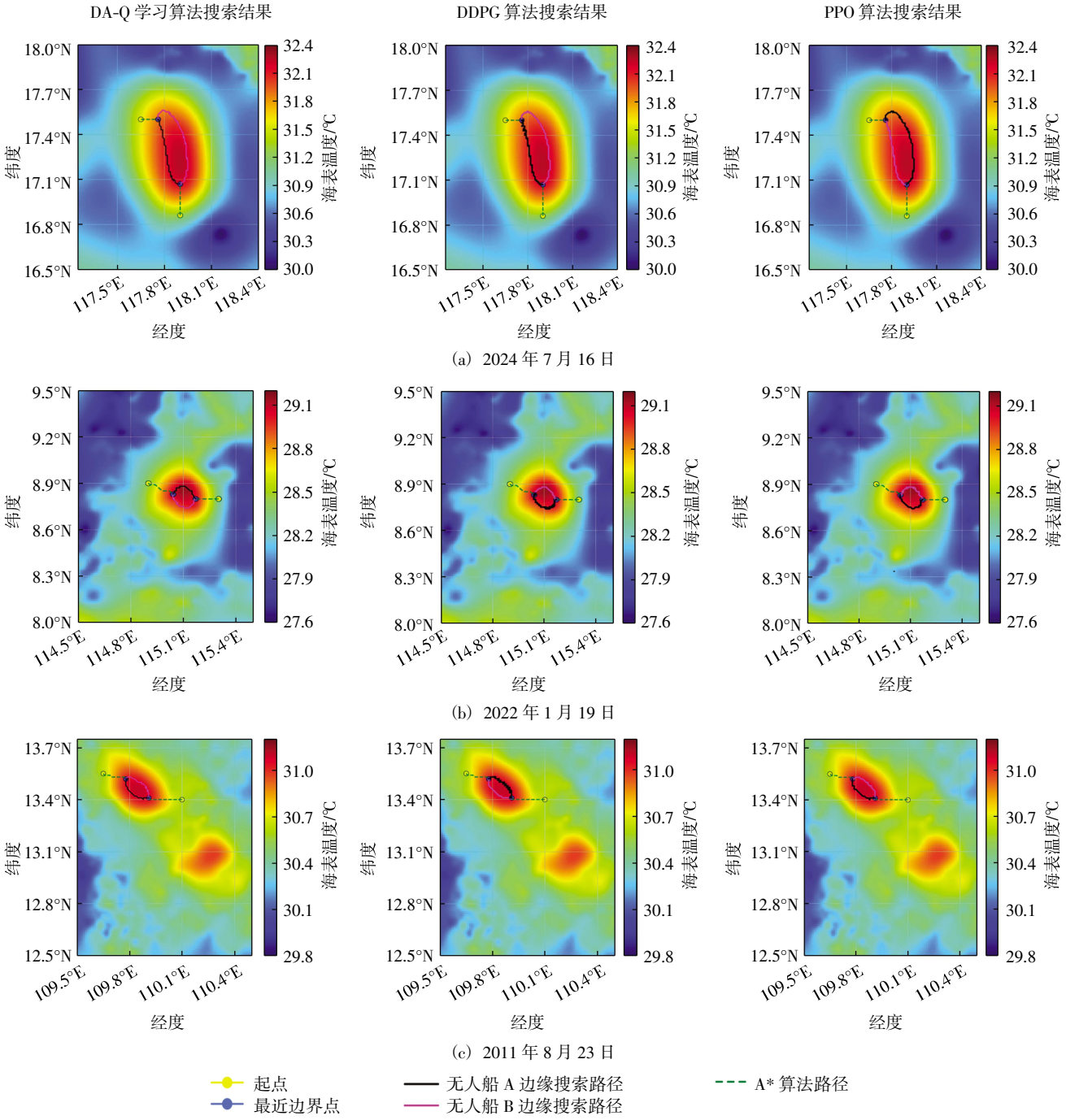


图 11 不同搜索算法实验效果

由表 9 可知，DDPG 与 PPO 虽在部分实验中具有较短的搜索时间，但整体搜索成功率较低，反映出其策略稳定性不足。特别是 DDPG 在个别场景下搜索成功率下降，并伴随搜索时间过长的问题；而

PPO 虽然收敛过程相对平稳，但整体搜索成功率偏低，存在一定的不确定性。而 DA-Q 学习在所有实验配置下搜索成功率较高，同时搜索时间与路径长度处于较优水平。这说明 DA-Q 在缺乏全局信息、需快速响应的动态环境中具有更高的鲁棒性和泛化能力，凸显了其在亚中尺度涡边缘搜索任务中的特异性优势。

表 9　不同边缘搜索算法对比

日期	搜索算法	搜索成功率/%	搜索完成时间/h
2024-07-16	DA-Q 学习算法	100	3.76
	DDPG 算法	65	3.05
	PPO 算法	10	3.22
2022-01-19	DA-Q 学习算法	100	1.14
	DDPG 算法	35	1.38
	PPO 算法	10	1.38
2011-08-23	DA-Q 学习算法	100	1.42
	DDPG 算法	75	3.85
	PPO 算法	70	1.31

4　结　论

本文针对亚中尺度涡边缘监测中存在的实时性和监测精度不足等问题，提出了一种融合梯度自适应 A* 与 DA-Q 学习的多无人船亚中尺度涡边缘搜索算法。首先，利用梯度信息引导多无人船快速逼近涡旋区域，实现初始路径的高效规划；然后，通过引入基于温度梯度变化与历史路径构建的动态奖励机制，驱动 DA-Q 学习算法完成复杂边缘结构下的多无人船自适应搜索。通过模拟实验与仿真实验相结合的方式，全面评估了算法在不同涡旋边缘形态、不同季节、不同涡旋类型及不同部署方式下的综合性能。

研究结果表明，在上述实验条件下，所提算法在搜索路径长度和搜索效率等方面均优于单船搜索策略。此外，与基于全图数据的传统水平温度梯度法检测结果相比，本文算法能有效、准确地搜索亚中尺度涡。综上所述，本文提出的多无人船边缘搜索算法在多种复杂海洋环境下均展现出较强的鲁棒性和实用价值，为构建高效、智能、实时的海洋边缘监测体系提供了有效的技术支持。未来工作将加强现场数据采集的获取，在复杂真实海洋环境中进一步验证算法的稳定性与实用性，从而提升其应用价值。

参考文献：

[1] LÉVY M，FERRARI R，FRANKS P J S，et al. Bringing physics to life at the submesoscale[J]. Geophysical Research Letters，2012，39(14)：14602.

[2] ZHANG Z. Submesoscale dynamic processes in the south China Sea[J]. Ocean-Land-Atmosphere Research，2024，2(2)：103-120.

[3] WANG Q，DONG C，DONG J，et al. Submesoscale processes-induced vertical heat transport modulated by oceanic mesoscale eddies[J]. Deep-sea Research, Part II. Topical Studies in Oceanography，2022，202(8)：105138.

[4] 王芮芮，孙忠斌. 海洋亚中尺度非地转运动的季节变化及对地转能量串级的影响研究[J]. 中国海洋大学学报（自然科学版），2022，52(7)：14-21.

[5] ZHANG Z，ZHANG X，QIU B，et al. Submesoscale currents in the subtropical upper ocean observed by long-term high-resolution mooring arrays[J]. Journal of Physical Oceanography，2021，51(1)：187-206.

[6] LÉVY M，FRANKS P J S，SMITH K S. The role of submesoscale currents in structuring marine ecosystems[J]. Nature Communications，2018，9(1)：4758.

[7] BECKER F，ROMERO S I，PISONI J P. Detection and characterization of submesoscale eddies from optical images: A case study in the Argentine continental shelf[J]. International Journal of Remote Sensing，2023，44(10)：3146-3159.

[8] HOU H，XU T，WANG Y，et al. Spatiotemporal Distribution of submesoscale eddies with updated interferometric imaging radar altimeter data[J]. Journal of Remote Sensing，2025，5：475.

[9] LI R，XIA L，YANG J，et al. Towards submesoscale eddy detection in SDGSAT-1 data through deep learning[J]. International Journal of Digital Earth，2024，17(1)：2369625.

[10] 朱金彤，范陈清，万勇，等. 基于 AVISO 卫星测高数据的海洋中尺度涡识别算法对比分析[J]. 海洋技术学报，2025，44

(2)：1–14.

[11] 陈津乐，张锦水，段雅鸣，等. 中分辨率遥感影像云检测与厚云去除研究综述[J]. 遥感技术与应用，2023，38(1)：143–155.

[12] ARCHER M，WANG J，KLEIN P，et al. Wide-swath satellite altimetry unveils global submesoscale ocean dynamics[J]. Nature，2025，640(8059)：691–696.

[13] 崔虎山，孙佳，郭延良，等. 基于 Argos 漂流浮标资料的南海及邻近海域次中尺度涡识别及特征分析[J]. 海洋科学进展，2024，42(3)：490–500.

[14] HART P E，NILSSON N J，RAPHAEL B. A formal basis for the heuristic determination of minimum cost paths[J]. IEEE transactions on Systems Science and Cybernetics，1968，4(2)：100–107.

[15] ZHANG H，TAO Y，ZHU W. Global path planning of unmanned surface vehicle based on improved A-Star algorithm[J]. Sensors，2023，23(14)：6647.

[16] SANG T，XIAO J，XIONG J，et al. Path planning method of unmanned surface vehicles formation based on improved A* algorithm[J]. Journal of Marine Science and Engineering，2023，11(1)：176.

[17] WATKINS C J C H，DAYAN P. Technical note: Q-learning[J]. Machine Learning，1992，8(3–4)：279–292.

[18] JIANG J，XIN J. Path planning of a mobile robot in a free-space environment using Q-learning[J]. Progress in Artificial Intelligence，2019，8：133–142.

[19] ZHAO M，LU H，YANG S，et al. The experience-memory Q-learning algorithm for robot path planning in unknown environment[J]. IEEE Access，2020，8：47824–47844.

[20] HU J，HO C R，XIE L，et al. Regional oceanography of the South China Sea[M]. Singapore：World Scientific，2020.

[21] ZHENG Q，XIE L，XIONG X，et al. Progress in research of submesoscale processes in the South China Sea[J]. Acta Oceanologica Sinica，2020，39：1–13.

[22] DONG J，ZHONG Y. Submesoscale fronts observed by satellites over the northern South China Sea shelf[J]. Dynamics of Atmospheres and Oceans，2020，91：101161.

[23] VAZQUEZ-CUERVO J，CASTRO S L，STEELE M，et al. Comparison of GHRSST SST analysis in the Arctic Ocean and Alaskan coastal waters using saildrones[J]. Remote Sensing，2022，14(3)：692.

[24] 李延志，孙伟富. 南海海洋温度锋面时空特征遥感分析[J]. 数字海洋与水下攻防，2024，7(1)：37–46.

[25] 许素芹，陈标，陶荣华，等. 中国近海温度锋时空分布特征及验证[J]. 遥测遥控，2015，36(3)：62–69.

[26] 尹汉军，谢波涛，黄必桂，等. 基于南海温盐流数值产品的南海气候态温跃层、温度锋和中尺度涡评估[J]. 海洋预报，2023，40(4)：34–46.

A Submesoscale Eddy Boundary Search Algorithm for USVs Based on Gradient-Adaptive A-Star and Dynamic-Award Q-Learning

LI Wenwen[1,2], MA Yi[1], ZHANG Feifei[1,2], LI Zhongwei[2]

(1. First Institute of Oceanography, Ministry of Natural Resources, Qingdao 266061, China; 2. College of Oceanography and Space Informatics, China University of Petroleum (East China), Qingdao 266580, China)

Abstract: Submesoscale eddies play significant roles in oceanic energy transfer and ecological regulation, yet their accurate and real-time monitoring remains challenging. To address the limitations of conventional monitoring methods-such as insufficient timeliness, limited spatial resolution, and poor flexibility of observation platforms, this paper proposes a multi-Unmanned Surface Vehicle (multi-USV) edge search algorithm for submesoscale eddies based on a highly maneuverable multi-USV platform. The algorithm integrates a gradient-adaptive A-star algorithm and a dynamic reward-driven reinforcement learning method (Dynamic-Award Q-learning, DA-Q learning). Targeting the complex morphology of submesoscale eddy edges, rapidly changing dynamic environments, and incomplete information in realistic scenarios, the algorithm employs gradient-adaptive A-star to guide multiple USVs to quickly converge toward the eddy region. During the edge search process, a dynamic reward mechanism combining temperature gradient variations and historical path information is established to drive the DA-Q learning algorithm, enabling USVs to autonomously adapt to environmental changes and flexibly adjust their courses. Using both simulated data and simulated sea surface temperature data, multiple sets of comparative experiments were designed, covering different seasons, numbers of USVs, and initial configurations. The performance and robustness of the algorithm under various environmental conditions, eddy types, and eddy distortion scenarios were analyzed. Results show that the multi-USV strategy for submesoscale eddy edge search outperforms the single-USV approach in both path length and search efficiency. In terms of deployment strategy, distributed multi-point deployment further enhances search coverage and mission response efficiency compared to single mother-vessel release deployment. The findings of this study can provide technical support for multi-USV monitoring of submesoscale eddies in complex marine environments.

Keywords: submesoscale eddy; Unmanned Surface Vehicle; A-star algorithm; DA-Q learning