

# 城市密集环境下基于不确定度的无人机频谱地图 主动感知方法

张 晗, 韩 宇\*, 孟令欣, 姚孟辰, 陈晗卉

(哈尔滨工程大学信息与通信工程学院, 黑龙江 哈尔滨 150001)

**摘 要:** 在城市密集环境中, 高层建筑遮挡、多系统并发辐射及多源干扰叠加使电磁环境呈现显著时变性和空间异质性, 传统离线构建并低频更新的频谱地图难以及时反映真实信道状态, 难以满足动态频谱管理、无线资源调度及低空通信保障对精度与时效性的要求, 尤其是在复杂任务场景下对在线感知与快速决策能力提出了更高要求。针对上述问题, 本文提出一种基于不确定度驱动的无人机频谱地图主动感知方法。首先, 将无人机在线频谱地图更新与路径规划联合建模为马尔可夫决策过程, 在状态中显式引入地图不确定度统计量、无人机位姿、飞行时间、目标距离和任务进度等信息, 以增强策略对环境变化和任务阶段的综合表征能力。其次, 利用预训练 U-Net 实现 SINR 频谱地图在线重构, 并输出像素级不确定度估计, 以不确定度下降表征信息增益, 引导无人机优先感知高价值未知区域。进一步, 设计基于特征线性调制的 FiLM-D3QN 决策网络, 利用不确定度与任务进程条件动态调制中间特征和价值估计, 实现信息获取与航迹效率之间的自适应权衡。仿真结果表明, 在多种干扰城市场景下, 所提方法在到达率保持 90% 以上的同时, 可有效降低地图重构误差, RMSE 较经典 IPP 方法下降约 7.1%, 较原始 D3QN 下降约 9.8%, 并具有更短的平均回合长度, 验证了该方法在复杂城市电磁环境中开展在线频谱地图更新与主动感知任务的有效性。

**关键词:** 城市电磁环境; 无人机; 频谱地图; 主动感知; 不确定性估计; 深度强化学习

**基金项目:** 国家自然科学基金(No.62201172); 中央高校基本科研业务费专项资金资助项目(No.3072025YY0801)

**中图分类号:** TN929.5

**文献标识码:** A

**文章编号:** 0372-2112(XXXX)XX-0001-12

**电子学报 URL:** <http://www.ejournal.org.cn>

**DOI:** 10.12263/DZXB.20260052

## Uncertainty-Driven UAV Active Sensing for Radio Map Construction in Dense Urban Environments

ZHANG Han, HAN Yu\*, MENG Lingxin, YAO Mengchen, CHEN Hanhui

(College of Information and Communication Engineering, Harbin Engineering University, Harbin, Heilongjiang 150001, China)

**Abstract:** In dense urban environments, blockage by high-rise buildings, concurrent radiation from multiple wireless systems, and the superposition of multi-source interference make the electromagnetic environment highly time-varying and spatially heterogeneous. Traditional radio maps, which are constructed offline and updated at a low frequency, cannot reflect the real channel state in time. As a result, they are unable to meet the requirements of dynamic spectrum management, wireless resource scheduling, and low-altitude communication support in terms of both accuracy and timeliness. This challenge is more pronounced in complex mission scenarios, where online sensing and rapid decision-making are especially important. To address this issue, this paper proposes an uncertainty-driven active sensing method for UAV-based spectrum mapping. First, the online radio map updating and path planning processes are jointly modeled as a Markov decision process. The state explicitly includes statistical measures of map uncertainty, UAV position, flight time, target distance, and mission progress, so as to improve the policy's ability to represent environmental changes and mission stages. Second, a pre-trained U-Net is employed to reconstruct the SINR radio map online and to output pixel-level uncertainty estimates. The reduction in uncertainty is used to characterize information gain, which guides the UAV to sense high-value unknown regions with priority. Furthermore, a FiLM-D3QN decision network based on Feature-wise Linear Modulation is designed. It uses uncertainty and mission-progress conditions to dynamically modulate intermediate features and value estimation, thereby achieving an adaptive balance between information acquisition and trajectory efficiency. Simulation results show that, in urban scenarios with dense interference, the proposed method effectively reduces map reconstruction error while maintaining an arrival rate above 90%. The RMSE is reduced by about 7.1% compared with the classical IPP method and by about 9.8% compared with the original D3QN. The proposed method also achieves a shorter average episode length. These results veri-

fy its effectiveness for online radio map updating and active sensing in complex urban electromagnetic environments.

**Keywords:** urban electromagnetic environment; unmanned aerial vehicle(UAV); radio map; active sensing; uncertainty estimation; deep reinforcement learning

**Foundation Item(s):** National Natural Science Foundation of China (No.62201172), Fundamental Research Funds for the Central Universities (No.3072025YY0801)

## 0 引言

城市密集环境中无线系统密集,遮挡与多径明显,链路质量在秒级快速波动<sup>[1-2]</sup>。传统离线构建或低频更新的频谱地图(Radio Map, RM)易在短时间内与真实环境失配。再考虑地面节点在街谷区域覆盖不足、回传受限, RM难以同时满足通信保障、干扰抑制和动态频谱管理对时效与精度的要求。因此,频谱地图更新需要从离线/低频模式转向在线主动感知,以提升对复杂动态电磁场景的适应能力。

随着“低空经济”产业的快速发展,城市低空飞行正由试验性、间歇性活动迈向常态化运行,无人机(Unmanned Aerial Vehicles, UAVs)在巡检、测绘、应急等领域得到广泛应用<sup>[3-4]</sup>。无人机具备机动观测能力,可在任务执行过程中补充频谱测量,从而支持RM的在线维护<sup>[5-6]</sup>。该类在线感知依托巡检/应急等既定任务飞行,通常受到航时/步长预算约束并具有软终端目标,因此在线主动感知需要在不显著增加任务代价的前提下进行。从具身智能视角看,该过程可表述为“感知-决策-行动”闭环:无人机通过机动飞行改变观测分布,并在软终端目标与预算约束下依据地图状态在线调整航迹,实现对城市电磁环境的按需感知与持续更新<sup>[7]</sup>。

然而,面向城市密集电磁环境的任务执行过程在线主动感知仍面临三方面挑战:(1)非平稳与稀疏观测并存,城市电磁辐射场空间起伏剧烈,而无人机沿轨迹采样稀疏且分布不均,导致未观测区域外推误差与盲区风险突出<sup>[8]</sup>;(2)资源约束显著,地空链路带宽、端侧算力与控制时延限制了在线更新频率与状态规模<sup>[9]</sup>,需要轻量化表征与有效更新机制;(3)终端目标下的阶段性权衡要求在线主动感知需在有限步长/时效预算内尽量到达终点,同时兼顾信息获取与航迹效率,策略需随任务进程自适应调整以避免过度探索或过早收敛。

围绕城市密集电磁环境的主动感知,已有研究从信息路径规划(Informative Path Planning, IPP)与强化学习(Reinforcement Learning, RL)等方向展开探索。IPP类方法通常以信息增益为目标引导采样轨迹,在环境监测中表现突出,但其信息度量往往依赖手工设计的代理模型/协方差更新或简化统计量,难以直接利用深度频谱地图重构模型输出的预测不确定度(如

方差、置信区间)来驱动在线决策;同时在软终端目标与步长预算并存时,多目标耦合会带来求解复杂度与可控性的挑战。深度学习类方法可通过交互学习闭环策略<sup>[10]</sup>并具备融合复杂约束的潜力<sup>[11]</sup>,但在频谱地图驱动的决策任务中,缺少对频谱地图不确定度的显式建模并将其与策略学习深度耦合,如何有效利用地图置信信息并稳定权衡信息获取与航迹效率仍有待进一步研究<sup>[12]</sup>。

为弥补上述不足,本文从“频谱地图不确定度建模—深度强化学习决策”的角度出发,将频谱地图在线更新产生的预测不确定度显式纳入决策闭环:一方面把不确定度统计融入马尔可夫状态与奖励设计,以信息增益驱动探索;另一方面通过条件化的值函数学习机制,使策略能够随任务进程自适应调整探索强度与到达效率,从而实现面向软终端与预算约束的在线主动感知。本文的主要贡献有如下几点。

(1)提出一种不确定度驱动的频谱地图主动感知框架,将频谱地图重构不确定度显式建模并用于决策,区别于仅用预测均值或仅用启发式信息度量的路径规划方法。

(2)提出基于特征线性调制(Feature-wise Linear Modulation, FiLM)的决策网络模型,基于不确定度统计与任务进程量对中层特征与价值估计进行动态调制,使值函数在不同任务阶段自适应调整探索以到达权衡。

(3)基于城市多源干扰场景的实验结果表明,与基线方法相比,在到达率保持同一水平的条件下,所提方法的频谱地图重构误差更低且平均回合长度更短。

## 1 相关工作

### 1.1 基于静态部署的频谱地图构建

频谱地图构建方法通常依赖预先部署的传感节点或基站进行感知,构建方法从传统方法逐渐过渡到深度学习方法。早期研究主要利用克里金插值<sup>[13]</sup>等经典空间统计方法,通过建立变异函数和协方差模型来刻画频谱数据的空间依赖特性。随着深度学习技术的发展和在无线信号领域的广泛应用<sup>[14]</sup>,数据驱动的方法显著提升了对复杂传播环境的适应性<sup>[15]</sup>。文献[16]利用基于U-Net结构的深度卷积神经网络构建频谱地图,实现了利用稀疏测量数据对复杂环境下频谱地图的高精度重构。文献[17]提出一种生成式无线电环境建模方法,能够在有限的条件下实

现高精度的频谱生成。文献[18]提出通过图卷积结构实现了频谱地图多模态特征的深度融合,能够优化样本稀缺与遮挡严重区域的估计偏差;同时,为克服分辨率与参数规模问题,注意力机制被引入频谱地图补充与主动感知中,既兼顾几何变化,又降低推理开销<sup>[19]</sup>。这些研究极大地提高了频谱地图的构建精度。

## 1.2 无人机辅助的动态频谱地图构建

基于静态部署的方法受感知节点部署限制,在城市复杂环境下难以保持精度。因此,结合具备机动性和三维感知能力的无人机辅助的地图构建方法逐渐成为解决上述问题的手段。

无人机辅助的频谱地图构建方法可分为非自适应和自适应两类。非自适应构建方法可以根据是否考虑环境信息进行分类,环境无关的方法按照预定义路径进行采样,环境相关的方法结合环境结构等先验知识,对频谱采样路径进行有针对性的设计与优化<sup>[20]</sup>。文献[21]在采样策略中加入贪心优化机制,通过迭代选择信息量最大的采样位置,提高了频谱的采样效率。然而,上述方法在城市密集环境中往往效率低下,难以实现高效且精确的频谱地图构建。

自适应频谱地图构建方法基于观测数据动态规划路径,克服了非自适应构建方法关键信息获取不足的缺点。文献[22]结合在线贝叶斯估计和深度学习方法,动态更新频谱地图的不确定度指标,并利用动态规划在高不确定度区域实时生成航迹,获得更高质量的频谱地图。文献[23]提出了一种基于无人机的三维频谱地图构建方法,通过对无人机飞行轨迹与采样位置进行合理规划,实现了对有限且不均匀采样数据的高精度重建。文献[24]通过迭代选择采样点来实现3D频谱地图的高效重建,但其采样路径容易受到估计误差的影响。

为进一步提升信息获取效率,研究者开始从主动感知角度引入信息路径规划思想,通过最大化收集到的未知感兴趣区域(Region of Interest, RoI)信息来优化采样路径<sup>[25]</sup>。IPP常以概率建模为基础(如高斯过程回归)刻画空间场的不确定性或估计误差,并据此构造采样价值函数,在预算约束下进行路径优化<sup>[26-27]</sup>。为降低多步非近视规划的计算复杂度,滚动/在线优化成为常用近似策略<sup>[28]</sup>。近年来,一方面通过稀疏高斯过程与连续空间优化提升可扩展性,以支持更大规模或多机器人IPP<sup>[29]</sup>;另一方面,主动学习/贝叶斯实验设计脉络与IPP逐步交叉,出现了将多种采样价值函数融合为可规划地图表征并纳入规划目标的通用框架工作<sup>[30]</sup>。文献[31]使用协方差矩阵进行路径优化,实现了稳定且具有良好的适应性的主动信息采集。

文献[32]将信息路径规划问题建模为RL问题,

是将RL引入IPP的代表工作。文献[1]和文献[33]提出了导航与无线电映射框架,构建了一种多层级频谱地图,包括空间位置和信道特性,利用深度强化学习在飞行过程中一边学习覆盖感知导航策略,一边构建和更新频谱地图,实现了在覆盖空洞环境下显著提升UAV通信与路径规划的效率。

综上,现有研究已在信息路径规划和强化学习等方面对频谱地图主动感知问题进行了探索,并取得了一定的进展。近年来,得益于深度强化学习在复杂状态下的建模优势,这些方法逐渐成为主流方案。然而,现有强化学习方法在频谱地图相关任务中,多以位置或即时观测信息驱动决策,尚缺乏对地图不确定度的显式建模,同时在任务约束下感知与运动策略的协同调节机制仍有待进一步研究。

## 2 系统建模与问题描述

### 2.1 系统模型

如图1所示,本文考虑了一种城市密集电磁环境下的无人机频谱地图主动感知场景。假设一个感兴趣的立方体空域,该空域可以描述为 $[x_{lo}, x_{up}] \times [y_{lo}, y_{up}] \times [z_{lo}, z_{up}]$ ,其中下标lo和up分别表示空域的坐标上下限。区域中存在 $M$ 个基站(Base Station, BS),第 $m$ 个BS的位置用 $\mathbf{g}_m$ 表示。同时,场景中存在 $K$ 个干扰设备,第 $k$ 个干扰设备的位置用 $\mathbf{i}_k$ 表示。这些基站和干扰源共同构建了城市密集电磁环境。

无人机作为该场景中的频谱地图主动感知主体,其运动轨迹贯穿整个感知过程。假设无人机的起始坐标为 $\mathbf{u}_0$ ,结束坐标为 $\mathbf{u}_T$ ,在 $t$ 时刻,无人机的坐标用 $\mathbf{u}(t)$ 表示,其中 $0 \leq t \leq T$ 。为简化实验,假设UAV采用

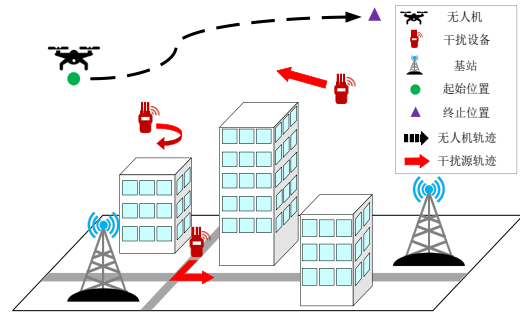


图1 UAV路径规划场景图

Figure 1 Scene diagram for UAV path planning

恒等速度运动。

在 $t$ 时刻,UAV与第 $m$ 个BS的等效信道增益表示为

$$p_m(t) = P_m |h_m(t)|^2 = P_m \beta_m[\mathbf{u}(t)] G_m[\mathbf{u}(t)] \tilde{h}_m(t) \quad (1)$$

其中, $P_m$ 是第 $m$ 个BS的发射功率; $h_m(t)$ 为时刻 $t$ 下第

$m$  个基站到无人机的等效信道功率增益;  $\beta_m(\cdot)$  和  $G_m(\cdot)$  分别代表第  $m$  个 BS 的大尺度信号增益和第  $m$  个 BS 的天线增益, 与无人机和基站的相对位置相关;  $\tilde{h}_m(t)$  表示用于描述小尺度衰落的随机变量。

在本文中, 将 UAV 与相关 BS 之间的信号与干扰加噪声比 (Signal to Interference plus Noise Ratio, SINR) 作为 UAV 辅助构建的频谱地图应用目标, SINR 的计算过程为

$$\text{SINR}(t) = \frac{p_{b(t)}(t)}{\sum_{m \neq b(t)} p_m(t) + N_0(t) + \sum_K I_K(t)} \quad (2)$$

其中,  $N_0(t)$  表示  $t$  时刻环境中的噪声功率,  $b(t)$  为  $t$  时刻与 UAV 关联的服务基站索引,  $I_K(t)$  表示在  $t$  时刻中的  $K$  个干扰设备的功率。后文中的 RM 均指以 SINR 为表征量的频谱地图。

## 2.2 优化问题描述

鉴于利用深度学习构建频谱地图已成为当前的主流范式, 本文设定依托一个通用的频谱服务平台, 该平台使用神经网络来进行 SINR 地图重构, 并且得到网络模型的不确定度, 将城市地图划分成  $N \times N$  的频谱地图  $\mathbf{E}$ 。假设  $\mathbf{S}$  为真实 SINR 地图, 则  $t$  时刻神经网络推理的估计 SINR 图为  $\mathbf{O}_t^{\text{SINR}}$ , 则  $\mathbf{O}_t^{\text{SINR}}$  的计算过程为

$$\mathbf{O}_t^{\text{SINR}} = f_\theta(\mathbf{X}_t) = f_\mu(\mathbf{X}_t) + \varepsilon, \varepsilon \in \mathcal{N}(0, f_\sigma(\mathbf{X}_t)) \quad (3)$$

其中,  $\mathbf{X}_t$  表示  $t$  时刻输入神经网络中的特征图,  $f_\theta(\cdot)$  表示参数为  $\theta$  的神经网络的整体表示,  $f_\mu(\cdot)$  表示神经网络预测的均值,  $f_\sigma(\cdot)$  表示神经网络预测的方差。

在无人机运行的过程中, 需要在通过主动感知降低不确定度, 提高频谱地图构建精度的前提下使用尽量少的无人机总飞行时间, 因此需要设计权重  $\mu$  作为联合优化这两个指标的参数。联合优化函数如下:

$$\begin{aligned} & T + \bar{T}_{\text{out}}(\{\mathbf{u}(t), \mathbf{i}(t)\}) \\ \min_{T, \|\mathbf{S} - \mathbf{O}_t^{\text{SINR}}\|_2^2} & \quad \mathbf{u}(0) = \mathbf{u}_0, \mathbf{u}(T) = \mathbf{u}_F \\ \text{s.t.} & \quad \mathbf{u}(t) \leq [x_{\text{lo}}, x_{\text{up}}] \times [y_{\text{lo}}, y_{\text{up}}] \times [z_{\text{lo}}, z_{\text{up}}] \\ & \quad t \in [0, T] \end{aligned} \quad (4)$$

其中,  $\leq$  代表集合中的每个元素均处于指定的区域范围内。

## 3 基于不确定度的频谱地图主动感知方法

针对城市密集电磁环境中无人机在线主动感知任务存在的观测稀疏非平稳、通信资源受限与任务效能动态权衡这三重挑战, 本文提出了一种不确定度驱动的频谱地图在线主动感知方法, 设计了基于不确定度的主动感知框架, 并改进了决策模型, 从而显著提高原有系统在复杂动态环境下的环境认知精度与任务执行效能。

### 3.1 频谱地图主动感知框架

为实现结合无人机导航的频谱地图主动感知任务, 本文提出以不确定度为核心驱动的频谱地图主动感知框架。该框架通过实时评估环境不确定度, 自主触发模型更新, 并融合多维信息生成决策状态, 以此动态指导无人机的频谱感知与飞行规划。通过在框架中引入不确定度机制, UAV 的规划行为不再只依赖预测结果本身, 而是在每一步都根据不确定度动态调整探索与利用的平衡, 从而实现更稳健的决策。

为了实现基于频谱地图在线更新的无人机主动感知, 本文将无人机的任务建模为一个马尔可夫决策过程 (Markov Decision Process, MDP)。马尔可夫决策过程由四元组  $\langle S, A, P, R \rangle$  组成。其中,  $S$  表示状态空间, 即所有可能观测状态  $s$  的集合。定义  $t$  时刻观测状态  $\mathbf{s}(t) = (\mathbf{O}_t^{\text{map}}, \mathbf{O}_t^{\text{SINR}}, \mathbf{O}_t^{\text{cond}})$ , 其中  $\mathbf{O}_t^{\text{map}}$  表示  $t$  时刻的不确定度地图,  $\mathbf{O}_t^{\text{SINR}}$  则为在线更新的频谱地图, 两者拼接为一个双通道矩阵作为模型输入。 $\mathbf{O}_t^{\text{cond}}$  则是由无人机当前位置坐标  $\mathbf{u}(t)$ 、已观测位置占总位置的比例  $r_{\text{obv}}$ 、剩余时间比例  $r_{\text{time}}$ 、平均不确定度  $\Delta \bar{V}$  以及目标距离  $d$  组成。因此,  $\mathbf{O}_t^{\text{cond}} = [\mathbf{u}(t), r_{\text{obv}}, r_{\text{time}}, \Delta \bar{V}, d] \in \mathbb{R}^6$ 。此外,  $A$  表示无人机的离散飞行动作集合, 本文中 UAV 飞行高度保持不变, 遵循经典四方向移动假设。 $P$  表示状态转移概率函数, 由无人机运动状态和频谱地图更新机制共同决定。 $R$  表示奖励函数, 用于引导 UAV 在探索未知与执行任务之间取得平衡。本文采用有限时域 MDP, 并引入折扣因子  $\gamma \in (0, 1)$  用于累计回报计算 (与式 (10) 一致)。

在 MDP 的状态转移过程中, 环境感知的核心在于构建高置信度的频谱地图  $\mathbf{O}_t^{\text{SINR}}$  及其不确定度评估指标  $\mathbf{O}_t^{\text{map}}$ 。在每一个时间步  $t$ , UAV 移动到新的位置  $\mathbf{u}(t)$  并采集当前位置的真实 SINR 值  $\mathbf{S}(\mathbf{u}(t))$ 。这些实时观测数据被发送到频谱服务平台, 用于触发频谱地图的在线更新。平台则采用预训练的 U-Net, 接受当前的稀疏观测, 输出估计的 SINR 地图和不确定度地图。前者用于评估通信服务质量下游应用, 而后者作为无人机进行主动感知的依据和量化指标, 决定 UAV 接下来应优先探索哪些区域。这种在线更新机制, 使得 MDP 的状态  $\mathbf{s}(t)$  能够动态反映环境认知的最新程度, 并通过不确定度地图对当前模型对各区域的认知程度进行直接量化。在环境动态性较强的条件下, 不确定度地图比单纯的 SINR 预测更能反映模型对局部区域认知的充分程度。

U-Net 离线训练数据由随机抽取的稀疏观测点构成, 输入为观测地图和观测掩码两通道; 网络输出 SINR 预测均值图与对数方差图 (对应式 (3)、式 (8)), 以同时学习预测值与像素级方差。其中, 将对数方差

作为本文的不确定度估计量。

为实现对 UAV 的精准指导航迹规划,需要对状态  $\mathbf{s}(t)$  的不确定度进行解耦。模型的不确定度由认知不确定度  $U_{\text{epi}}$  和偶然不确定度  $U_{\text{alc}}$  组成。认知不确定度是由频谱地图生成模型本身的知识不足导致的,认知不确定度  $U_{\text{epi}}$  的数学表示为

$$U_{\text{epi}} = \mathbb{E}_{p(\theta)} \left[ \mathbb{E}_{p(\mathbf{S}|\mathbf{X},\theta)} [\mathbf{S}|\mathbf{X},\theta]^2 \right] - \mathbb{E}_{p(\theta)} \left[ \mathbb{E}_{p(\mathbf{S}|\mathbf{X},\theta)} [\mathbf{S}|\mathbf{X},\theta] \right]^2 \quad (5)$$

认知不确定度反映了模型对环境的认知局限,是可以通过增加训练数据来减少的。与之相对的,偶然不确定度是由环境固有随机性导致的。偶然不确定度  $U_{\text{alc}}$  的数学表示为

$$U_{\text{alc}} = \mathbb{E}_{p(\theta)} \left[ \text{Var}_{p(\mathbf{S}|\mathbf{X},\theta)} [\mathbf{S}|\mathbf{X},\theta] \right] \quad (6)$$

其中,  $\text{Var}(\cdot)$  表示方差计算公式。偶然不确定度反映环境中的内在噪声,无法通过数据增加的手段消除。总不确定度由认知不确定度与偶然不确定度求和得到:

$$U = U_{\text{epi}} + U_{\text{alc}} \quad (7)$$

两者相加的总不确定度为 UAV 路径规划提供了一个先验信息,用于决定探索优先序。后续本文中用于主动感知路径规划与奖励设计的不确定度均为由式(7)给出的总不确定度。

在本文中,不确定度地图由 U-Net 网络得到。具体来说,  $t$  时刻网络输出中包括对预测 SINR 地图的对数方差  $\log(\sigma^2(t))$ 。不确定度地图  $\mathbf{O}_t^{\text{map}}$  的计算表示为

$$\mathbf{O}_t^{\text{map}} = \exp\left(0.5 \times \log(\sigma^2(t))\right) \quad (8)$$

全局平均不确定度  $U_t$  通过不确定度地图  $\mathbf{O}_t^{\text{map}}$  计算得出:

$$U_t = \frac{\sum_{i=1}^N \sum_{j=1}^N \mathbf{O}_t^{\text{map}}(i,j)}{N \times N} \quad (9)$$

在 MDP 框架下,无人机的最优策略旨在最大化累积奖励。为了实现主动感知与任务导航的双重目标,框架的奖励函数采用双目标优化设计,其主目标是提升环境认知质量,约束目标则是在规定时间步内达到目的地。强化学习的奖励函数如下:

$$R_t = \begin{cases} +R_{\text{goal}}, & \text{if } d_t \leq r_{\text{arrive}} \\ -R_{\text{goal}}, & \text{if } t \geq T \text{ or stop} \\ \lambda \cdot \Delta V_t - 1 - (d_t - d_{t-1}), & \text{otherwise} \end{cases} \quad (10)$$

其中,  $R_{\text{goal}}$  表示到达目标奖励,  $d_t$  表示当前时间步 UAV 到目标点的距离,  $r_{\text{arrive}}$  表示到达半径,  $T$  表示任务的步数限制。当 UAV 实际飞行步数  $t \leq T$ , 且 UAV 与目标位置之间的距离小于或等于到达半径时,认为 UAV 成功抵达目标区域,获得到达目标奖励  $R_{\text{goal}}$ ; 当  $t \geq T$  时,受到惩罚,得到  $-R_{\text{goal}}$ 。在 UAV 飞行过程中,

若出现尝试飞出地图边界的情况,则 UAV 位置  $\mathbf{u}(t)$  不更新,表现为无人机停止,并受到出界惩罚,得到  $-R_{\text{goal}}$ 。  $\lambda$  表示不确定度权重,  $\Delta V_t$  表示当前时间步  $t$  的不确定度改善情况。  $\Delta V_t$  的计算公式为

$$\Delta V_t = \max(U_t - U_{t-1}, 0) \quad (11)$$

其中,  $U_t$  与  $U_{t-1}$  表示当前时间步  $t$  和上一时间步  $t-1$  的全局平均不确定度。当满足  $\Delta V_t$  改善条件时, UAV 获得正向的不确定度奖励,鼓励无人机向高不确定度主动探索。该设计能促使 UAV 主动探索高不确定度区域。步数惩罚为  $-1$ , 目的是鼓励快速完成任务,避免无意义的徘徊。  $\Delta V_t$  将地图不确定度的降低转化为 MDP 的正向奖励,迫使 UAV 主动规划路径去覆盖  $U_{\text{epi}}$  较高的未知区域。同时,位置项奖励确保了任务的最终完成。这种设计保证了 UAV 并非盲目游荡,而是基于对环境不确定度的动态评估,规划出一条兼顾信息采集效率与飞行能耗的最优轨迹。框架整体算法流程如算法 1 所示。

#### 算法 1 基于不确定度的频谱地图主动感知流程

输入: UAV 初始位置  $\mathbf{u}(0)$ ; 目标位置  $\mathbf{u}_{\text{goal}}$ ; 预训练 U-Net 网络  $f_{\theta}(\cdot)$ ; 到达半径  $r_{\text{arrive}}$ ; 单回合步数上限  $T$ 。

输出: UAV 轨迹  $\{\mathbf{u}(t)\}$ ; 最终估计 SINR 地图  $\mathbf{O}^{\text{SINR}}$ 。

初始化: 令  $t=0$ , 若  $\|\mathbf{u}(0) - \mathbf{u}_{\text{goal}}\| \leq r_{\text{arrive}}$ , 则直接终止并输出。

1. FOR  $t=0$  to  $T-1$ , DO。

2. UAV 感知: 基于当前 UAV 位置为  $\mathbf{u}(t)$ , 感知该位置真实 SINR 值  $\mathbf{S}(\mathbf{u}(t))$ , 并将感知结果返回频谱服务平台。

3. 不确定估计: 频谱服务平台基于所有已知观测值, 运行预训练 U-Net 网络  $f_{\theta}(\cdot)$ , 由式(3)更新估计 SINR 地图  $\mathbf{O}_t^{\text{SINR}}$ , 并由式(8)计算不确定度地图  $\mathbf{O}_t^{\text{map}}$ 。

4. 观测状态构造: 将 UAV 当前位置  $\mathbf{u}(t)$ 、观测位置比例  $r_{\text{obv}}$ 、观测时间比例  $r_{\text{time}}$ 、平均不确定度  $\Delta \bar{V}$  和目标距离  $d$  合成构造  $\mathbf{O}_t^{\text{cond}}$ , 并将其与  $\mathbf{O}_t^{\text{SINR}}$  和  $\mathbf{O}_t^{\text{map}}$  共同构造观测状态  $\mathbf{s}(t)$ 。

5. 动作选择: 由式(10)计算奖励函数, 运行强化学习模型, 得到下一步动作和 UAV 下一步位置  $\mathbf{u}(t+1)$ 。

6. 终止判定:

```
IF  $\|\mathbf{u}(t+1) - \mathbf{u}_{\text{goal}}\| \leq r_{\text{arrive}}, R_{t+1} = R_{\text{goal}}, \text{BREAK};$ 
ELSE IF  $t+1 = T$ , 任务超时失败  $R_{t+1} = -R_{\text{goal}}, \text{BREAK};$ 
END IF
```

7. END FOR

8. 输出 UAV 轨迹  $\{\mathbf{u}(t)\}$  和最终估计 SINR 地图  $\mathbf{O}^{\text{SINR}}$ 。

## 3.2 基于 FiLM 条件化的 D3QN 网络

无人机主动感知问题求解的核心目标是寻找针对该任务的最优路径规划策略  $\pi$ , 使无人机在满足可靠通信质量约束的同时, 最小化任务完成时间与频谱地图更新误差。为实现这一目标, 本文采用改进的 D3QN (Dueling Double Deep Q-Network) 作为强化学习的核心求解框架。

D3QN 因将状态价值与动作优势函数解耦并具备较高训练稳定性而被广泛应用于无人机路径规划任务。其通过逼近状态-动作价值函数  $Q(s, a)$  来指导策略更新。与传统 DQN 相比, D3QN 结合了双重 DQN 与对偶网络有效缓解了动作价值估计的正偏差问题。其中, 双重 DQN 在  $Q$  值更新中使用在线网络  $Q_\theta$  选择动作、目标网络  $Q_{\bar{\theta}}$  评估价值, 从而降低动作价值的过估计偏差, 目标值计算公式为

$$y_i = r_i + \gamma Q_{\bar{\theta}}(s_{t+1}, \arg \max_{a'} Q_\theta(s_{t+1}, a')) \quad (12)$$

其中,  $y_i$  为目标值,  $r_i$  表示在状态  $s_t$  执行动作  $a_t$  后获得的即时奖励,  $\gamma$  为折扣因子,  $\theta$  和  $\bar{\theta}$  分别为在线网络与目标网络参数,  $s_{t+1}$  为下一状态,  $Q_\theta(s_{t+1}, a')$  代表在线网络的  $Q$  值计算, 用于选择下一步的最优动作由目标网络计算, 用于评估在线网络选出的动作。而对偶网络将  $Q$  值函数分解为状态价值函数  $V(s)$  与动作优势函数  $A(s, a)$ ,  $V(s)$  表示在给定状态  $s$  下, 遵循当前策略能够获得的整体期望回报,  $A(s, a)$  表示在给定状态  $s$  下, 相比于平均水平, 执行特定动作  $a$  的相对优势, 两者结合使网络在某些状态下无需依赖具体动作即可评估状态优劣:

$$Q(s, a) = V(s) + \left( A(s, a) - \frac{1}{|A|} \sum_{a'} A(s, a') \right) \quad (13)$$

这种结构在状态采样稀疏的情况下提升了泛化能力, 并降低了无关动作对价值估计的影响。然而, 传统 D3QN 仅关注无人机位置、频谱地图等信息, 难以根据任务进程量动态调整无人机主动感知与航迹优化平衡。一方面, 因对已观测规模不敏感而引发重复探测或盲区遗漏; 另一方面, 因缺乏时间约束感知, 难以实现各阶段资源的最优分配, 最终导致路径规划偏离最优与环境认知收敛缓慢。

为应对上述挑战, 本文引入 FiLM 方法, 对 D3QN 网络进行条件化扩展, 模型整体结构如图 2 所示。FiLM-D3QN 将不确定度与任务进程量作为条件输入, 通过一个小型映射网络生成通道级的缩放与偏移参数, 并作用于 D3QN 的卷积特征层。这样可以在不改变主体网络结构的前提下, 动态调整特征表达, 使策略随环境状态与任务阶段自适应平衡探索与利用, 从而提升在非平稳场景下的稳定性。

具体而言, 本文利用无人机横纵坐标、已观测网格数、任务剩余步长、全局平均不确定度以及到终点的距离等多维观测量作为条件输入, 经由可学习的映射函数生成调制参数, 并在网络前向传播过程中施加到各通道特征上。在数学形式上, 首先, FiLM 根据给定的第  $i$  个输入样本的条件信息  $x_i$  生成调制参数:

$$\gamma_i = f(x_i), \beta_i = h(x_i) \quad (14)$$

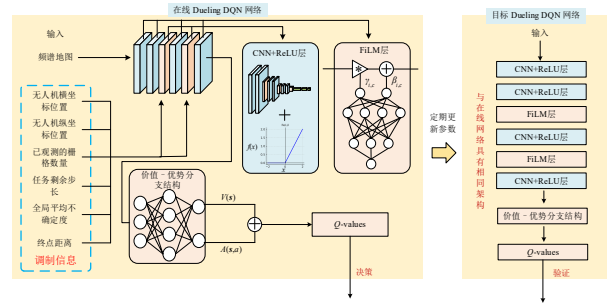


图2 基于FiLM的D3QN模型

Figure 2 D3QN model based on FiLM

其中,  $\gamma_i$  和  $\beta_i$  分别表示缩放系数和平移量输出与特征通道数对应的缩放系数;  $f$  和  $h$  表示可学习的映射函数, 用以将条件输入映射为与特征通道维度一致的参数矩阵。

随后, FiLM 将这些调制参数应用于中间特征  $F_{i,c}$ , 完成逐通道的仿射变换:

$$\text{FiLM}(F_{i,c} | \gamma_{i,c}, \beta_{i,c}) = \gamma_{i,c} \cdot F_{i,c} + \beta_{i,c} \quad (15)$$

其中,  $F_{i,c}$  表示第  $i$  个样本在通道  $c$  的特征表示。通过该操作,  $\gamma_{i,c}$  控制特征强度的放大或抑制, 从而突出与任务相关的模式或削弱无关信息, 而  $\beta_{i,c}$  则调整整体激活水平, 以满足下游任务的需求。

这一机制使得模型具备了细粒度的条件化调节能力。相比于拼接或注意力等传统融合方式, FiLM 在保持较低计算开销的同时, 可以实现高效而灵活的特征控制, 从而显著提升了复杂推理任务中的表现。

## 4 仿真结果与分析

本文参考文献[1], 构建一个  $2 \text{ km} \times 2 \text{ km}$  的密集城区测试场景, 仿真生成分辨率为  $10 \text{ m} \times 10 \text{ m}$  的 SINR 地图。区域内高层建筑密集, 分布 5 个固定位置的宏基站与 3 个随机的干扰源。在该仿真场景中, 城区形态参数遵循国际电信联盟建议: 建筑高度服从截尾瑞利分布, 为避免不现实的超高建筑, 在 90 m 处截断。基站设置固定高度为 25 m, 采用三扇区固定覆盖, 方位角每扇区  $120^\circ$ , 扇区天线配置为  $8 \times 1$  垂直线阵, 机械下倾  $10^\circ$ , 垂直/水平 3 dB 波束宽度均为  $65^\circ$ 。基站的等效各向同性辐射功率设为 0.1 W。干扰源设置固定高度为 2 m, 辐射功率设为 0.01 W, 采用全向天线。基站-UAV 与干扰源-UAV 的大尺度路径损耗采用 3GPP 城市宏场景视距 (Line-of-Sight, LoS) 与非视距 (Non-Line-of-Sight, NLoS) 模型, 小尺度衰落对 LoS 段采用 Rician 分布, 对 NLoS 段采用 Rayleigh 分布。

上述仿真参数旨在对标城市密集电磁环境的关键特征: 通过密集建筑分布与高度随机刻画街谷遮挡与覆盖边界带来的空间异质性; 通过大尺度损耗与小尺度衰落刻画城市环境中的视距概率变化与多径

散射效应;通过随机干扰源叠加模拟同频系统共存或外来干扰导致的多源干扰情形,从而形成具有代表性的复杂 SINR 空间分布与不确定度热点。基于该设置,可检验所提方法在密集城区场景下对频谱地图精度与航迹决策的综合优化能力。

UAV 以 10 m/s 的恒定速度在高度 100 m 巡航,初始位置设置为(500,500),目标点设为(1 500,1 500)。控制器以 1 Hz 的频率执行,单回合决策步数上限为 500(对应真实时间 500 s)。

实验在搭载 Intel Core i9-13980HX、32 GB 内存与 NVIDIA RTX 4080 显卡的工作站上完成,全部算法基于 PyTorch 2.7 实现,优化器采用 Adam,激活函数为 ReLU,采用批归一化。初始学习率设为 0.005。主要仿真参数汇总见表 1。

表 1 仿真参数表

Table 1 Simulation parameter table

参数		数值
环境参数	地图尺寸	2 km×2 km
	地图分辨率	10 m×10 m
	基站辐射功率	0.1 W
	基站天线高度	25 m
	干扰源辐射功率	0.01 W
	干扰源天线高度	2 m
	决策间隔	1 Hz
	到达半径	3 m
预训练参数	到达奖励	1 000
	学习率	0.000 1
	batch size	16
模型参数	训练轮数	50
	最大回合数	2 000
	单回合步数上限	500
	经验回放容量	100 000
	初始探索率	0.5
探索衰减系数	0.99	
多步学习步数	30	

图 3 给出了该仿真场景中一个随机实例下的真实 SINR 空间分布,用于作为建图误差评估基准,无人机测得的 SINR 为在该分布上的稀疏采样点。

受遮挡、覆盖边界与干扰热区影响,该实例中高不确定度区域在空间上的位置、规模与连通性存在明显差异,如图 4 中红框所示。在本文方法中,这类差异通过在线更新的不确定度地图反映到状态与奖励中,从而间接影响策略的路径选择。

后续实验中将 UAV 任务终止后的 SINR 地图均方根误差(Root Mean Square Error, RMSE)作为评估主动感知性能的指标,并同步对比任务到达率和平均回合长度,以评估智能体对信息获取与航迹效率的权衡能

力。为降低偶然波动的影响,实验结果展示为对多次独立运行随机环境实例的均值结果。

首先,将本文提出的不确定度驱动主动感知方法与随机游走、基于不确定度的贪心策略以及经典 IPP 框架进行对比。为保证对比公平,对贪心与 IPP 中的到达软约束权重采用网格扫描,在保持到达率与本文方法处于同一水平的前提下,选择 RMSE 更低且回合长度更短的权重作为最终基线参数。随机游走不引入额外超参数,按相同终止条件直接统计指标。

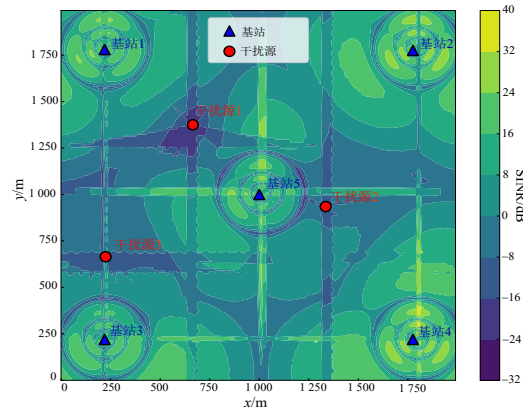


图 3 仿真场景的真实 SINR 分布示例图

Figure 3 Example diagram of the real SINR distribution in a simulation scenario

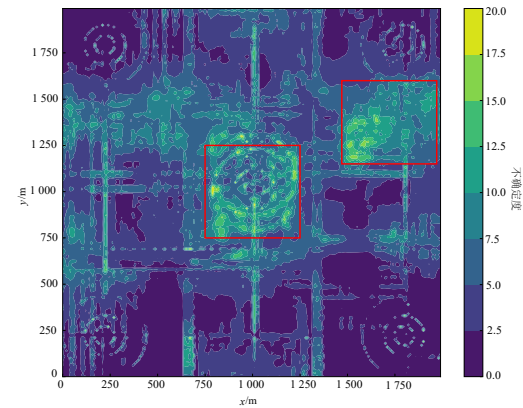


图 4 仿真场景的不确定度分布示例

Figure 4 Example of uncertainty distribution in a simulation scenario

如表 2 所示,本文方法 RMSE 最低,IPP 次之,贪心与随机游走较差。原因在于随机游走缺乏信息引导,贪心策略容易重复采样;IPP 通过滚动规划缓解短视;本文方法则通过学习在到达约束下自适应分配探索强度,因此在相近到达率下获得更低 RMSE。

为分析所提出方法在不同建图权重下的表现,引入式(10)中的权衡参数 $\lambda$ (不确定度权重),在其余设置保持一致的条件下,仅改变奖励项的 $\lambda$ 取值,分析方法的性能。图 5 给出了 $\lambda$ 对主动感知策略综合性能

表2 主动感知策略性能对比表

Table 2 Performance comparison table of active perception strategies

模型	平均RMSE/dB	平均到达率/%	平均回合长度/步
随机游走	4.88	91.0	277.0
贪心算法	4.72	89.4	289.7
经典IPP框架	4.48	90.1	286.0
本文方法	4.16	90.7	269.4

的影响,其中图5(a)为频谱建图RMSE,图5(b)为UAV到达率,图5(c)为平均回合长度。整体来看, $\lambda$ 决定了策略在信息采样任务完成之间的偏好强度,从而形成典型的精度—效率权衡。图5(a)显示,随着 $\lambda$ 从0增大至0.25/0.33/0.5, RMSE曲线整体下移且收敛更快,说明提高不确定度改善奖励能够驱动策略更积极地访问高不确定度区域并获得更具信息增益的观测,从而更有效地提升建图精度;而当 $\lambda$ 进一步增大到1时, RMSE虽然可达到最低,但后期不同曲线间差距趋于缩小,表明精度收益存在边际递减的情况。与此同时,图5(b)和图5(c)揭示了这一精度提升所付出的任务代价:较小 $\lambda$ (如0或0.25)时到达率更快达到并稳定在较高水平且回合长度较短,策略更接近最短路径规划;  $\lambda$ 增大(如0.33或0.5)后到达率出现一定回落且平均回合长度上升,表明策略为降低不确定度而更频繁偏离主通道并增加绕行;当 $\lambda=1$ 时,到达率显

著降低且回合长度明显更长,且伴随更大波动,反映出策略持续被高不确定度区域牵引而牺牲终点推进。

从航迹效率角度看,本文在奖励中引入了步数惩罚以鼓励快速完成任务、抑制无意义徘徊(见式(8)~式(9)的设计说明),因此在保证精度的同时,仍可通过适当降低 $\lambda$ 或提高步数惩罚系数,使策略在同等约束下进一步偏向短航迹解。与此同时,表3也显示出FiLM-D3QN在RMSE下降的同时将平均回合长度由278.3步降至269.4步(约下降3.2%),说明本框架不仅能提升建图精度,还具备进一步优化航迹效率的空间。

图5也反映了不同随机场景的综合表现,不同场景主要通过 $U_i$ 分布差异影响路径:高不确定度区域越分散、越偏离主通道,降低RMSE的代价越高,因而更易带来回合长度上升和到达率下降。

综合来看,在合适的建图权重下( $\lambda=0.5$ ),所提出的不确定驱动主动感知方法能够在显著降低频谱地图误差的同时,保持较高的到达率,且回合长度仅小幅提升,整体收敛平稳。权重过小则难以有效提高主动感知性能,权重过大又会引入明显的路径开销。

下面进一步分析本文提出的FiLM-D3QN模型,表3给出了FiLM-D3QN在不同的超参数scale下与主流强化学习模型的性能对比结果。同时,表3补充了若干消融设置,用于分析FiLM调制方式及条件信息组成对性能的影响。

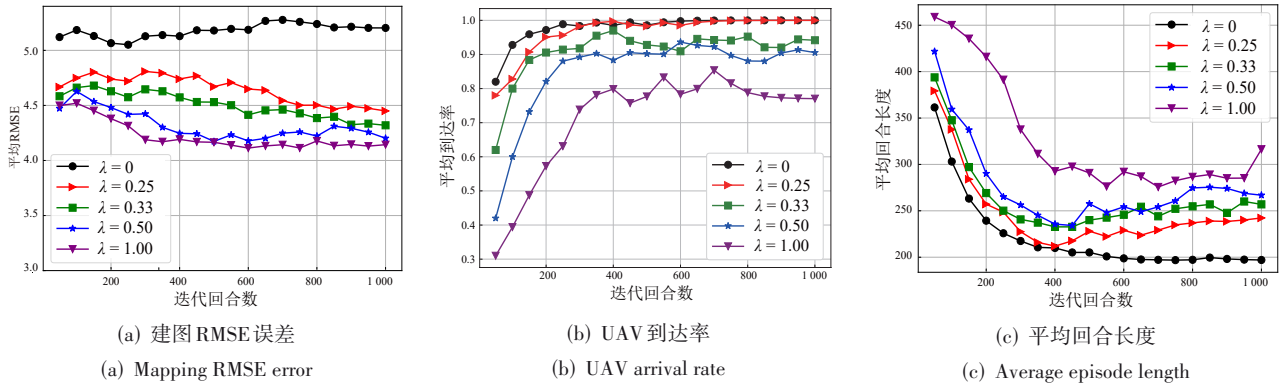


图5 不确定度权重对主动感知策略综合性能的影响

Figure 5 The impact of uncertainty weighting on the comprehensive performance of active perception strategies

从表3结果可见,FiLM-D3QN在合适的调制强度下能够同时提升建图精度与航迹效率。其中,  $scale=0.10$ 时综合效果最佳: RMSE由原始D3QN的4.61降至4.16(下降约9.8%),到达率由90.3%小幅提升至90.7%,平均回合长度由278.3步降至269.4步(下降约3.2%)。这表明FiLM条件化不仅增强了策略对不确定度分布与任务进程的响应能力,还减少了不必要的绕行与重复采样,使得精度提升并未以更长航迹为代价。

进一步的敏感性对比表明,  $scale$ 取值过小或过大

都会削弱综合性能:当 $scale=0.05$ 时调制偏弱,FiLM对中间特征的影响不足,条件信息难以有效改变价值估计和动作偏好,因此RMSE与回合长度改善有限;当 $scale$ 增至0.15时调制偏强。

条件信息对特征的缩放与偏移更剧烈,虽然能继续驱动对高不确定度区域的采样,从而保持较低RMSE,但也更容易引起到达率下降和回合长度回升。上述结果说明FiLM调制存在合理强度区间,后续实验中本文据此选取 $scale=0.10$ 作为默认设置,以

表3 RL模型性能对比表

Table 3 Performance comparison table of RL models

模型	平均RMSE/ dB	平均 到达率/%	平均回合 长度/步
FiLM-D3QN(scale=0.05)	4.43	90.4	275.8
FiLM-D3QN(scale=0.10)	4.16	90.7	269.4
FiLM-D3QN(scale=0.15)	4.19	87.8	279.1
FiLM-D3QN(仅进程度)	4.49	90.5	271.2
FiLM-D3QN(仅不确定度)	4.31	89.7	274.3
D3QN(拼接特征)	4.48	90.1	276.8
原始D3QN <sup>[2]</sup>	4.61	90.3	278.3
SAC <sup>[34]</sup>	4.45	90.7	273.7
DDPG <sup>[35]</sup>	4.37	88.0	281.3
PPO <sup>[36]</sup>	4.74	81.7	337.2

兼顾精度、效率与训练稳定性。

与此同时,消融结果显示:仅使用进程度或仅使用不确定度均能一定程度改善性能,但二者同时使用时效果更优,表明两类条件信息具有互补性;而在不使用FiLM、仅将特征进行拼接的基线设置下,性能虽优于原始D3QN,但仍弱于FiLM-D3QN(scale=0.10),说明FiLM的特征调制机制相比简单拼接更能有效利用条件信息并减少不必要绕行。综合来看,适度的FiLM调制与合理的条件信息组合能够在信息获取与航迹效率之间形成更均衡的收益,从而在保证建图精度的同时进一步提升航迹效率。

表4给出了FiLM-D3QN与原始D3QN的复杂度对比。可以看出,FiLM-D3QN的参数数量相比原始D3QN仅增加 $27.84 \times 10^3$ (约1.0%)。同时,FiLM-D3QN的平均推理时延较原始D3QN的增加0.07 ms(约7.2%)。总体而言,FiLM引入的时延增长有限,不会显著加重在线推理负担,计算开销能够满足实时决策需求。

表4 模型复杂度对比表

Table 4 Model complexity comparison table

模型	参数量/ $10^3$	平均推理时延/ms
FiLM-D3QN	2 776.52	1.04
原始D3QN <sup>[2]</sup>	2 748.68	0.97

图6展示了不同算法在相同任务条件下的平均回合长度变化趋势。结果表明,FiLM-D3QN收敛后平均回合长度约为269.4步,低于D3QN的278.3步、PPO的337.2步和DDPG的281.3步,同时其收敛阶段波动更小。这说明该方法在相同任务条件下能以更短航迹完成任务且训练更稳定。从趋势上看,FiLM-D3QN在早期阶段的下降幅度较大,而在收敛阶段曲线振荡幅度则较小,体现出该模型在复杂环境中具备更强的收敛鲁棒性与泛化性。综合分析可知,FiLM-D3QN在到达时间指标上表现最优,不仅平均步数最短、收敛

速度最快,而且波动最小,充分证明了其在特征调制与值函数估计结合下的有效性与稳定性。

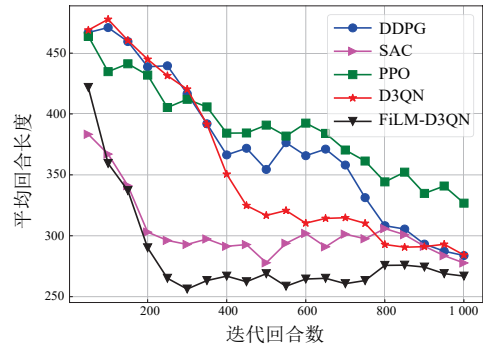


图6 不同算法迭代过程中的平均回合长度图

Figure 6 Graph of average episode length during iterative process for different algorithms

图7进一步展示了不同算法在训练过程中平均RMSE随迭代回合数变化的情况。从结果看,FiLM-D3QN算法的综合性能显著优于DDPG、SAC、PPO及传统D3QN算法。在迭代回合数从200增加至1000的过程中,FiLM-D3QN的RMSE始终保持最低,1000回合时其RMSE降至4.16,相对传统D3QN降低约10%,相对DDPG降低约5%,且收敛速度更快。相比之下,PPO和SAC的RMSE下降斜率较平缓,且局部存在微小波动。这一结果说明特征线性调制能提升特征判别能力,减少特征冗余与噪声干扰。

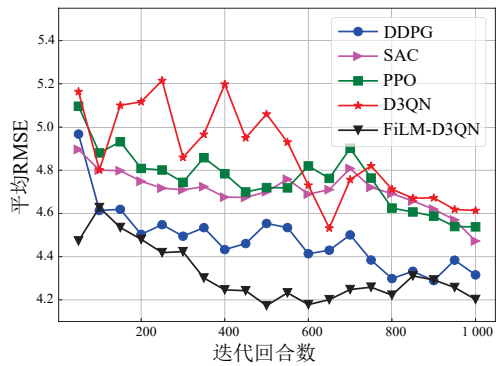


图7 不同算法迭代过程中的平均RMSE图

Figure 7 Graph of average RMSE during iterative process for different algorithms

为直观展示不同方法的效果,图8给出了在图3示例下的真实SINR分布以及不同方法重构结果的对比。可以观察到,本文方法在整体空间结构与局部干扰峰值区域的重构上更接近真实分布。相比之下,原始D3QN与SAC在部分区域存在更明显的局部失真。该结果与RMSE指标一致,进一步说明本文方法能够在有限采样条件下获得更高质量的频谱地图重构。

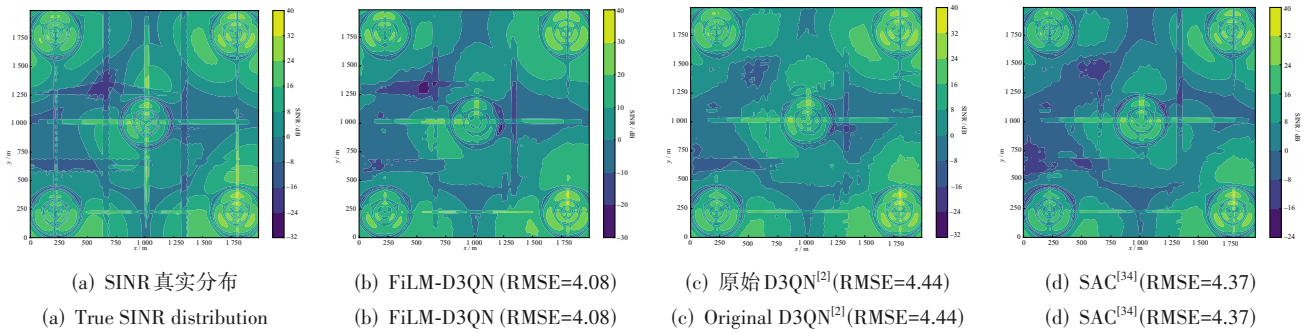


图8 真实SINR地图与不同方法重构结果可视化

Figure 8 Visualization of true SINR map and reconstruction results from different methods

最后,图9给出了本文提出方法在图3示例场景下的UAV主动感知路径的可视化结果。该轨迹可反映策略在该实例中倾向于先覆盖图4中的两个高不确定度区域再推进终点,在没有引入过多的路径探索情况下达到了良好的频谱地图精度,与图5所反映的精度—效率权衡。该示例可定性说明本文方法的闭环机制在单实例中的体现,而跨场景情况下路径形态的差异则由各场景下 $U_i$ 分布不同所诱导,其总体影响已在图5的统计结果中体现。

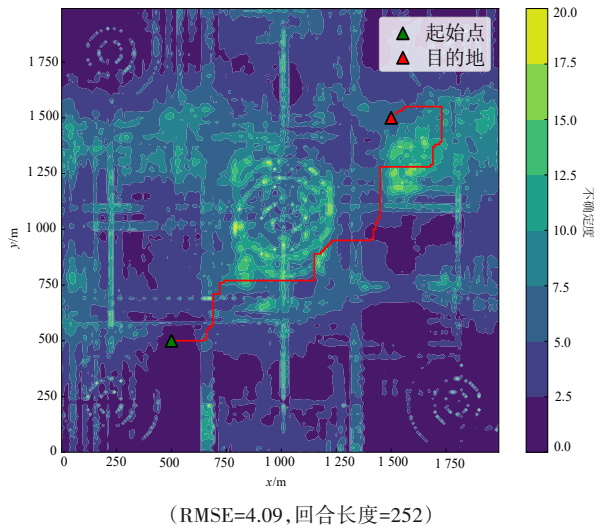


图9 提出方法的主动感知示例轨迹

Figure 9 Example trajectory of active perception using the proposed method (RMSE=4.09, Episode Length=252)

## 5 结论

本文针对城市密集电磁环境下频谱地图易失配、难以在线维护的问题,提出不确定度驱动的无人机主动感知方法,并将不确定度纳入航迹决策。进一步设计FiLM-D3QN条件化模型,使策略能够结合不确定度与任务进程量在建图精度与航迹代价之间自适应权衡。仿真结果表明,在多干扰城区场景中,本文方

法在到达率保持90%以上时RMSE较经典IPP下降约7.1%,且FiLM-D3QN相对原始D3QN的RMSE降低约9.8%。未来,将面向多无人机与实测数据,开展泛化与不确定度校准研究。

## 参考文献

- [1] Ding Guoru, Jiao Yutao, Wang Jinlong, et al. Spectrum inference in cognitive radio networks: Algorithms and applications[J]. IEEE Communications Surveys & Tutorials, 2018, 20(1): 150-182.
- [2] Zeng Yong, Xu Xiaoli, Jin Shi, et al. Simultaneous navigation and radio mapping for cellular-connected UAV with deep reinforcement learning[J]. IEEE Transactions on Wireless Communications, 2021, 20(7): 4205-4220.
- [3] Shi Zhiguo, Chang Xianyu, Yang Chaoqun, et al. An acoustic-based surveillance system for amateur drones detection and localization[J]. IEEE Transactions on Vehicular Technology, 2020, 69(3): 2731-2739.
- [4] 饶凌风, 耿娜, 张勇, 等. 不确定环境下无人机任务分配的种群交互式粒子群算法[J]. 电子学报, 2025, 53(8): 2678-2690.  
Rao Lingfeng, Geng Na, Zhang Yong, et al. Population interactive particle swarm optimization algorithm for UAV task allocation in uncertain environments[J]. Acta Electronica Sinica, 2025, 53(8): 2678-2690. (in Chinese)
- [5] Sun Lu, Wang Jiashuai, Wan Liangtian, et al. Human-UAV interaction assisted heterogeneous UAV swarm scheduling for target searching in communication denial environment[J]. IEEE Transactions on Automation Science and Engineering, 2025, 22: 4457-4472.
- [6] Li Jiapeng, Zhang Qixun, Ma Dingyou, et al. UAV target reconstruction and imaging algorithm design and performance evaluation using CSI of ISAC signal[J]. IEEE Transactions on Wireless Communications, 2025, 24(11): 9422-9434.

- [7] 陈进朝, 王洋, 张营, 等. 基于时空密度聚类的异构无人机集群覆盖路径规划方法[J]. 电子学报, 2025, 53(3): 705-715.  
Chen Jinchao, Wang Yang, Zhang Ying, et al. Coverage path planning for heterogeneous UAVs based on temporal-spatial density clustering[J]. Acta Electronica Sinica, 2025, 53(3): 705-715. (in Chinese)
- [8] 沈锋, 丁国如, 李婕, 等. 电磁频谱多维态势压缩测绘技术研究进展[J]. 通信学报, 2023, 44(11): 25-42.  
Shen Feng, Ding Guoru, Li Jie, et al. Research progress on electromagnetic spectrum multidimensional situation compressed mapping technology[J]. Journal on Communications, 2023, 44(11): 25-42. (in Chinese)
- [9] Lin Yun, Tu Ya, Dou Zheng. An improved neural network pruning technology for automatic modulation classification in edge devices[J]. IEEE Transactions on Vehicular Technology, 2020, 69(5): 5703-5706.
- [10] Kang Jiawen, Zhong Yue, Xu Minrui, et al. Tiny multiagent DRL for twins migration in UAV metaverses: A multileader multifollower Stackelberg game approach[J]. IEEE Internet of Things Journal, 2024, 11(12): 21021-21036.
- [11] Gao Qian, Zhong Ruikang, Hyundong Shin, et al. MARL-based UAV trajectory and beamforming optimization for ISAC system[J]. IEEE Internet of Things Journal, 2024, 11(24): 40492-40505.
- [12] Zhou Yifan, Zhou Fuhui, Zhou Huilin, et al. Robust trajectory and transmit power optimization for secure UAV-enabled cognitive radio networks[J]. IEEE Transactions on Communications, 2020, 68(7): 4022-4034.
- [13] 王红军, 徐浩然, 赵文元. 电磁频谱地图构建中的空间插值方法综述[J]. 电讯技术, 2025, 65(1): 141-151.  
Wang Hongjun, Xu Haoran, Zhao Wenyuan. Survey on spatial interpolation methods for electromagnetic spectrum map construction[J]. Telecommunication Engineering, 2025, 65(1): 141-151. (in Chinese)
- [14] Tu Ya, Lin Yun, Zha Haoran, et al. Large-scale real-world radio signal recognition with deep learning[J]. Chinese Journal of Aeronautics, 2022, 35(9): 35-48.
- [15] Romero D, Kim S J. Radio map estimation: A data-driven approach to spectrum cartography[J]. IEEE Signal Processing Magazine, 2022, 39(6): 53-72.
- [16] Levie R, Yapar Ç, Kutyniok G, et al. RadioUNet: Fast radio map estimation with convolutional neural networks[J]. IEEE Transactions on Wireless Communications, 2021, 20(6): 4001-4015.
- [17] Zhang Songyang, Wijesinghe A, Ding Zhi. RME-GAN: A learning framework for radio map estimation based on conditional generative adversarial network[J]. IEEE Internet of Things Journal, 2023, 10(20): 18016-18027.
- [18] Zhang Han, Han Yu, Meng Lingxin, et al. MFFGCN: Multimodal feature fusion graph convolution network for radio map estimation with uneven spatial sampling[J]. IEEE Transactions on Mobile Computing, 2025, 24(12): 13366-13382.
- [19] 谭海东, 杨晶晶, 黄铭. 基于卷积神经网络和注意力机制的无线电地图构建方法研究[J]. 电波科学学报, 2025, 40(6): 1069-1077.  
Tan Haidong, Yang Jingjing, Huang Ming. A study on the methodology for constructing radio maps utilizing convolutional neural network and attention mechanisms[J]. Chinese Journal of Radio Science, 2025, 40(6): 1069-1077. (in Chinese)
- [20] Paonessa F, Virone G, Bolli P, et al. UAV-based antenna measurements: Scan strategies[C]//Proceedings of 2017 11th European Conference on Antennas and Propagation (EUCAP). Piscataway: IEEE, 2017: 1303-1305.
- [21] Liu Wenjie, Chen Junting. UAV-aided radio map construction exploiting environment semantics[J]. IEEE Transactions on Wireless Communications, 2023, 22(9): 6341-6355.
- [22] Wang Jie, Zhu Qiuming, Lin Zhipeng, et al. Sparse Bayesian learning-based hierarchical construction for 3D radio environment maps incorporating channel shadowing[J]. IEEE Transactions on Wireless Communications, 2024, 23(10): 14560-14574.
- [23] Shrestha R, Romero D, Chepuri S P. Spectrum surveying: Active radio map estimation with autonomous UAVs[J]. IEEE Transactions on Wireless Communications, 2023, 22(1): 627-641.
- [24] Du Xiaofu, Zhu Qiuming, Ding Guoru, et al. UAV-assisted three-dimensional spectrum mapping driven by spectrum data and channel model[J]. Symmetry, 2021, 13(12): 2308.
- [25] Wu Qihui, Shen Feng, Wang Zheng, et al. 3D spectrum mapping based on ROI-driven UAV deployment[J]. IEEE Network, 2020, 34(5): 24-31.
- [26] Chen Yiran, Zhu Qiuming, Wang Jie, et al. UAV-aided efficient informative path planning for autonomous 3D spectrum mapping[J]. IEEE Transactions on Cognitive Communications and Networking, 2026, 12: 1664-1677.
- [27] Dutta S, Wilde N, Smith S L. Informative path planning for active regression with Gaussian processes via sparse

- optimization[J]. IEEE Transactions on Robotics, 2025, 41: 2184-2199.
- [28] Zhu Hai, Chung J J, Lawrance N R J, et al. Online informative path planning for active information gathering of a 3D surface[C]//Proceedings of 2021 IEEE International Conference on Robotics and Automation (ICRA). Piscataway: IEEE, 2021: 1488-1494.
- [29] Jakkala K, Akella S. Multi-robot informative path planning from regression with sparse Gaussian processes[C]//Proceedings of 2024 IEEE International Conference on Robotics and Automation (ICRA). Piscataway: IEEE, 2024: 12382-12388.
- [30] Rückin J, Magistri F, Stachniss C, et al. An informative path planning framework for active learning in UAV-based semantic mapping[J]. IEEE Transactions on Robotics, 2023, 39(6): 4279-4296.
- [31] Schlotfeldt B, Tzoumas V, Pappas G J. Resilient active information acquisition with teams of robots[J]. IEEE Transactions on Robotics, 2021, 38(1): 244-261.
- [32] Wei Yongyong, Zheng Rong. Informative path planning for mobile sensing with reinforcement learning[C]//Proceedings of IEEE INFOCOM 2020-IEEE Conference on Computer Communications. Piscataway: IEEE, 2020: 864-873.
- [33] Zhao Haitao, Hao Qing, Huang Hao, et al. Online trajectory optimization for energy-efficient cellular-connected UAVs with map reconstruction[J]. IEEE Transactions on Vehicular Technology, 2024, 73(3): 3445-3456.
- [34] Deng Tan, Wang Yanping, Li Jin, et al. Entropy normalization sac-based task offloading for UAV-assisted mobile-edge computing[J]. IEEE Internet of Things Journal, 2024, 11(15): 26220-26233.
- [35] Zhang Lijuan, Peng Jiabin, Yi Weiguo, et al. A state-decomposition DDPG algorithm for UAV autonomous navigation in 3-D complex environments[J]. IEEE Internet of Things Journal, 2024, 11(6): 10778-10790.
- [36] Guan Yue, Zou Sai, Peng Haixia, et al. Cooperative UAV trajectory design for disaster area emergency communications: A multiagent PPO method[J]. IEEE Internet of Things Journal, 2024, 11(5): 8848-8859.

#### 作者简介



张 晗 男,1997年8月出生于辽宁省葫芦岛市。现为哈尔滨工程大学信息与通信工程学院博士研究生。主要研究方向为频谱地图生成、无人机端侧应用。  
E-mail: hldzh1997@hrbeu.edu.cn



韩 宇 男,1986年10月出生于黑龙江省哈尔滨市。现为哈尔滨工程大学副教授。主要研究方向为电磁环境数据挖掘与智能模型建模、电磁频谱地图生成等。  
E-mail: yu.han1001@hrbeu.edu.cn