

基于智能体交互的无线电地图构建与 非合作辐射源定位

周 陬^{1,2}, 邓添荣¹, 仇 琛^{3*}, 仇洪冰^{1,2}

(1. 桂林电子科技大学认知无线电与信息处理省部共建教育部重点实验室, 广西桂林 541004;
2. 广西人工智能学院, 广西南宁 530201; 3. 鹏城实验室宽带通信部, 广东深圳 518055)

摘要: 无线电地图构建与非合作辐射源定位是电磁环境精确定位与动态管控的核心任务,二者高度耦合、相互制约。本文借鉴认知学习与具身智能理论,提出基于智能体协同交互、智能内生的无线电地图构建与辐射源定位双任务统一架构,实现感知-认知-行为的反馈闭环。首先利用众包概率共识思想,结合高斯回归过程与二维位置编码,将多智能体的离散稀疏采样数据转换为包含连续信号场的预测分布、几何拓扑及辐射源目标空间位置存在概率估计的空间注意力图,建立多智能体关于辐射源目标的共识感知先验;在此基础上,构建多智能体统一语义空间中的辐射源目标一致性联合表征,利用多头自注意力机制设计任务驱动和目标驱动的智能体交互策略,以同时完成无线电地图构建和非合作辐射源定位双任务;基于近端策略优化强化学习方法,设计双任务信息交互反馈策略,实现自组织与自学习,以形成感知、认知与交互决策的具身行为闭环。仿真结果验证了所提方案的可行性:在5%稀疏采样条件下,本系统对无线电地图构建与辐射源定位双任务进行协同优化,相较于独立执行单任务,无线电地图构建与辐射源定位的精度分别提升18.2%和43.5%。其中,无线电地图构建精度(RMSE)相比克里金算法和全卷积深度补全自动编码器分别提升了60.95%和32.55%;辐射源定位精度相较于传统的RSS定位基准算法,定位误差降低了65.7%,扩展到10 km × 10 km的典型定位场景时,辐射源定位任务仍能稳定维持在1.2%以内的相对精度。

关键词: 智能体交互;具身智能;共识感知;无线电地图;非合作辐射源定位

基金项目: 国家自然科学基金(No.62101293);广西自然科学基金(No.2023GXNSFAA026058);认知无线电与信息处理省部共建教育部重点实验室主任基金(No.CRKL220110)

中图分类号: TN92;TP18

文献标识码: A

文章编号: 0372-2112(XXXX)XX-0001-14

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

DOI: 10.12263/DZXB.20251269

Radio Map Construction and Non-cooperative Radiation Source Localization Based on Agents Interaction

ZHOU Zou^{1,2}, DENG Tianrong¹, QIU Chen^{3*}, QIU Hongbing^{1,2}

(1. Provincial Ministry of Education Key Laboratory of Cognitive Radio and Signal Processing, Guilin University of Electronic Technology, Guilin, Guangxi 541004, China;

2. Guangxi Academy of Artificial Intelligence, Nanning, Guangxi 530201, China;

3. Department of broadband communication, Peng Cheng Laboratory, Shenzhen, Guangdong 518055, China)

Abstract: Radio map construction and non-cooperative emitter localization are pivotal tasks for accurate perception and dynamic control of the electromagnetic environment, characterized by high coupling and mutual constraints. Drawing on cognitive learning and embodied intelligence theories, this paper proposes a unified dual-task framework for radio map construction and emitter localization featuring multi-agent collaborative interaction and endogenous intelligence, thereby establishing a "perception-cognition-action" feedback loop. First, leveraging the concept of crowdsourced probabilistic consensus, we integrate Gaussian Process Regression (GPR) with 2D-positional encoding to transform discrete sparse sampling data from multiple agent into spatial attention maps. These maps encompass continuous signal field predictive distributions, geometric topologies, and probability estimates of emitter spatial presence, establishing a consensus perceptual prior regarding the target emitters. Building upon this, a consistent joint representation for emitters is constructed within a unified multi-agent semantic space. A multi-head self-attention mechanism is then utilized to design task-driven and target-driven agent interaction strategies, enabling the simultaneous execution of both tasks. Furthermore, a dual-task information interaction feedback strategy based on Proximal Policy Optimization (PPO) is designed to achieve self-organization and self-learning,

forming an embodied behavioral loop of perception, cognition, and interactive decision-making. Simulation results validate the feasibility and superiority of the proposed scheme. Under a 5% sparse sampling condition, the synergistic optimization of both tasks achieves accuracy improvements of 18.2% and 43.5% for radio map construction and emitter localization, respectively, compared to independent task execution. Specifically, the RMSE of radio map construction improves by 60.95% and 32.55% over Kriging and Full Convolutional Auto-Encoder, respectively. Meanwhile, the localization error is reduced by 65.7% compared to the RSS-based baseline, maintaining a stable relative accuracy within 1.2% even when extended to a 10 km×10 km large-scale scenario.

Keywords: agent interaction; embodied intelligence; consensus perception; radio map; non-cooperative radiation source localization

Foundation Item(s): National Natural Science Foundation of China (No.62101293); Guangxi Natural Science Foundation of China (No.2023GXNSFAA026058); The Director Fund project of the Key Laboratory of Cognitive Radio and Information Processing of Ministry of Education (No.CRKL220110)

0 引言

无线电地图是对某一感兴趣区域内电磁环境的可视化表示^[1],而辐射源定位旨在确定空间中多辐射源的精确位置分布^[2]。无线电地图构建与非合作辐射源定位是实现电磁环境精确认知与动态管控的关键。例如,2023年某城市高峰时段因多源干扰导致5G掉线率高达18%,经济损失超亿元^[3]。若能构建区域内无线电地图并定位干扰源的具体位置,即可实现干扰源的精确定位与频谱资源的高效利用。这凸显了两个任务在网络规划^[4]、智能交通^[5]及智慧城市^[6]等领域的重要价值。

辐射源定位是一种通过测量辐射信号特征参数来确定辐射源空间位置的技术,包括基于角度测量(AOA/DOA)的定位方法^[7]、基于时间测量(TOA/TDOA)的定位方法^[8]、基于频率测量(FDOA)的定位方法^[9]和基于信号强度RSSI的定位方法^[10],其中RSS定位方法凭借低成本和低硬件复杂度的优势,在实际应用中备受关注^[2]。常见的位置解算方法有模型驱动类方法(如贝叶斯推断法^[11])和数据驱动类方法(如基于机器学习的端到端模型^[12])。在多源信号叠加的复杂环境下,理论传播模型与实际信道的匹配度下降,导致模型驱动类方法精度受限;新兴的数据驱动类方法通过数据学习信号与位置之间的复杂映射关系,但不同源在特征空间中极易混淆^[13],通常需要引入额外的信号分选或目标辨识机制。近年来,随着技术发展,利用分布式传感器节点协作,通过多节点间关于辐射源目标的信息交互与协同推理来实现辐射源定位,最为典型的是基于无线电地图的定位方法^[2,14],通过将定位转化为信号空间分布图中的搜索匹配过程,增强了对多源混叠与干扰的鲁棒性,逐渐成为该领域的研究热点。然而,其定位精度依赖于无线电地图的构建精度^[15]。

无线电地图构建是获取电磁空间态势、实现频谱

资源高效利用的关键技术,包括假设传播特性已知^[15]的模型驱动方法和不依赖于信号传播模型^[16]的数据驱动类方法(克里金法^[17]、张量补全^[18]及贝叶斯函数^[19]等插值方法和深度学习)。在实际应用中,模型驱动类方法易出现实际信号观测结果与模型计算结果不匹配的情况,导致模型失配^[20]、边缘模糊^[21]等问题;数据驱动类方法中的插值类方法受多源混叠与多径干扰影响,构建的无线电地图往往精度受限^[15];而深度学习类方法虽具备强大的特征提取能力,如文献^[22]通过构建深度补全自动编码器,学习信号反射、绕射与阴影的基本规律,完成端到端构图,但其构图的分辨率与离散采样节点的数量成正比^[15],稀疏采样时易出现边缘伪影现象。为进一步提升低采样率下的构图精度,文献^[23-24]利用物理先验指导神经网络优化,改善了稀疏采样下的特征缺失问题。可见,无线电地图构建正朝着模型驱动与数据驱动融合的方向演进,但解决稀疏采样且采样位置分布不均导致的“看不全”和“看不准”问题,仍是提升构图精度的瓶颈。

随着物联网的普及,众包协同感知成为解决上述瓶颈的有效途径^[25]。对此,近年来涌现出多种基于智能体选择性交互的优化方案。在源定位方面,本团队分别从误差均衡性^[26]、目标适应性^[27]出发,提出了传感器选择策略;相关学者通过引入加权评估^[28]、联合误差递归优化^[29]等方法,有效提升了定位感知的精度。在构图方面,利用众包感知数据进行空间区域统一表征逐渐成为主流^[30-31]。例如,针对动态频谱环境的传感器布局优化策略^[32]通过主动优选与环境特征匹配的观测位置,有效降低了稀疏采样下的构图误差。因此,探索基于任务驱动和目标驱动的选择性信息交互机制,已成为优化无线电地图构建和辐射源定位精度的核心。在多辐射源混叠的场景中,构图与源定位任务之间存在高度的内在关联与相互依赖关系:高精度构图需要利用源位置分离混叠信号^[15],而精

确的源定位又需要借助无线电地图的数据分布信息先验,为求解源的空间位置缩小搜索范围、提供初始解方向^[2]。两个任务本质上高度耦合,难以割裂。在两个任务协同优化方面,两个任务均依赖于对信号传播物理模型的深刻理解,离不开对特定区域电磁环境中源目标空间相关特征的认知与推理。例如,文献[33]在统一语义空间中实现对电磁频谱与空间特征的提取,分别完成无线电地图构建和辐射源定位任务;文献[15]从有限的采样数据和多维数据特征出发实现数据和语义双驱动的全频段频谱地图与信号源位置的联合推理,进一步验证了利用智能化的信息交互组织,可以实现双任务统一认知的可行性。然而,传统的开环协同感知模式往往难以应对未知电磁环境的多源干扰问题。以某地区爆发大规模GNSS干扰事件^[34]和某机场的干扰事件^[35]为例,尽管部署了大量监测节点,但受限于协同感知信息汇聚的单向特性,无法根据未知环境自适应调节信息交互策略,难以实现对实际复杂电磁环境的精确认知与动态管控。

从认知物理学^[36]的视角看,两个任务均可基于传感器集群对区域电磁环境的众包离散采样、独立稀疏感知、多智能体联合表征,从空间场连续恢复与信号源位置估计这两个不同任务目标出发,利用不同的策略组织智能体间的信息交互而实现。其过程恰好构成具身智能体“感知-认知-行为”的反馈闭环^[37]。在该闭环中,智能体通过持续的信息交互与动作反馈,利用负反馈寻求熵减,从而形成会学习、自纠错、自组织、自成长的迭代交互智能^[36]。由此可见,在具身智能的理论框架下,面向辐射源定位和无线电地图构建任务,设计可解释的智能体分布式概率共识感知模型,构建任务目标导向的语义信息交互统一学习架构,并引入多智能体强化学习的自适应迭代更新能力,将有可能构建感知-认知-行为一体化的具身智能交互系统,为同时完成高精度的无线电地图构建和辐射源定位任务提供可行方案。

值得注意的是,具身智能体反馈闭环机制的研究已在自动驾驶领域得到了有效验证。文献[38]指出,智能体通过与环境及其他智能体的持续交互,在闭环中不断迭代反馈,这是解决未知环境下多任务智能体协同决策与自适应执行的关键。同时,电磁频谱认知学习理论指出,现有深度学习方法本质上仍属于感知智能,难以很好地适用于电磁频谱空间,而涵盖训练、执行、反馈与自学习全过程的闭环架构为实现从感知智能向认知智能的跨越提供了重要的理论框架^[39]。与之相比,国外前沿工作^[40]通过探索具备自我意识的具身智能范式,验证了多智能体协同完成任

务中利用交互闭环消除认知偏差的必要性;而文献[41]则进一步证实,在多智能体协同场景下,建立一体化的具身闭环系统对提升系统自适应性能的必要性。可见,全过程闭环反馈系统为构建基于智能交互的电磁频谱具身智能系统提供了可行的途径。

鉴于此,本文提出一种基于智能体交互的电磁频谱具身智能系统,通过构建“感知-认知-行为”的反馈闭环,将辐射源定位和无线电地图构建的任务系统转化为智能体自主交互的群体具身系统。其具身闭环特性具体体现在感知、认知和基于行为决策的反馈三个层面。

(1)具身感知层。利用高斯过程回归将离散稀疏采样数据建模为连续信号场的预测分布,并将其与融合了全局几何拓扑及辐射源目标空间位置存在概率估计的二维位置编码相结合,生成具备物理可解释性的空间注意力权重图。

(2)具身认知交互层。以共识感知先验为引导,在多智能体初始表征的基础上,通过智能体间的信息交互,在统一语义空间中建立多智能体联合表征,设计任务驱动和目标驱动的选择性语义交互策略,以同时实现无线电地图构建与非合作辐射源定位任务。

(3)基于行为决策的反馈层。设计基于近端策略优化(PPO)强化学习的双任务信息交互反馈优化策略。该策略能够实时根据双任务误差,动态调整感知模型、认知交互网络和关于双任务行为决策的核心超参数,形成端到端的具身智能闭环反馈。

1 系统模型

1.1 场景描述与问题建模

考虑采用多旋翼无人机(UAV)作为空中移动智能体,对电磁空间信号进行动态离散化采样^[33]。为便于对采样到的原始信号数据进行预处理,将感兴趣区域 $\Gamma \subset \mathbb{R}^{L \times W}$ 沿着 x 和 y 轴离散划分为空间集合为 $P^{L \times W}$ 的二维网格。其中,任意一个具体网格可记为 p 和 p' ,满足 $p, p' \in P$ 。基于此空间划分,可将采样数据以多维张量的形式作为系统的输入。区域内有 N 个位置集为 $\{\chi_n\}_{n=1}^N$ 的辐射源,沿无人机轨迹的采样点位置集合表示为 $\{\gamma_u\}_{u=1}^U$ 。其中, U 为采样点的个数。真实无线电地图可表示为 $\mathbf{M}_\Gamma \subset \mathbb{R}^{L \times W \times 1}$ 。采样点处的无人机采样数据为

$$m(\gamma, f) = \sum_{n=1}^N \alpha_n(f) |\eta_n(\gamma, f)|^2 + v(\gamma, f) \quad (1)$$

其中, $\alpha_n(f)$ 表示第 n 个辐射源的发送数据, $\eta_n(\gamma, f)$ 表示第 n 个辐射源与无人机之间的信道的频率响应,

$v(\gamma, f)$ 表示模型的噪声和干扰。采样信号的映射表示为 $M = M_\Gamma \cdot M_{\text{mask}}$ 。其中, M_{mask} 是一个掩模矩阵, M_{mask} 中的每个元素都可以设置为 1 和 0, 表示网格点是否被采样。采样数据可以定义为 $M = \{M^S, M^T\}$ 。其中, $M^S = \{m^s(\gamma_u, f)\}_{u=1}^U$ 表示采样点上的数据集, M^T 表示未采样点的数据集。

1.2 系统框架

基于智能体交互的电磁频谱具身智能系统通过构建“感知-认知-行为”的反馈闭环,使无人机具备会学习、自纠错、自组织、自成长的一体化具身智能特性。如图 1 所示,系统由共识感知模型、认知交互网络和任务导向的智能体强化学习反馈组成。

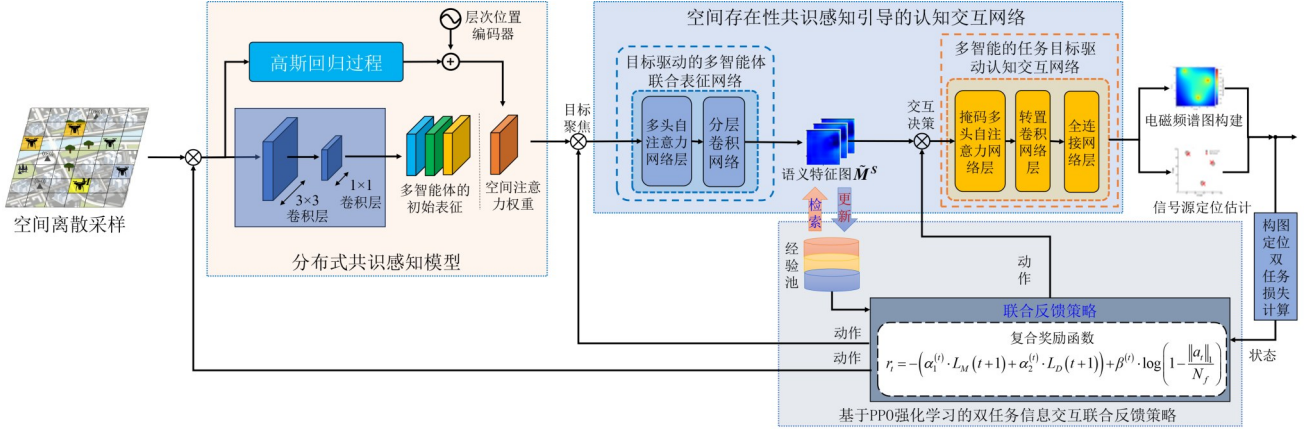


图 1 基于智能体交互的电磁频谱具身智能系统框架图

Figure 1 Framework of electromagnetic spectrum embodied intelligence system based on agent interaction

共识感知模型旨在建立具备物理可解释性的目标辨识先验。该模型利用高斯过程回归将多智能体的稀疏采样数据 M^S 建模为均值为 $\mu(p)$ 、方差为 $k(p, p')$ 的连续信号场的预测分布场;在此基础上,通过与位置编码相融合生成空间权重注意力图 $A(p) \in (0, 1)^{L_h \times J_h}$, 用于刻画辐射源目标在预测分布上的时空存在概率估计, 以实现采样信号、特征信息与辐射源目标的简单数据关联, 引导系统后续对辐射源目标属性、状态、动态事件的深度认知与推理, 为辐射源目标精确定位和区域无线电地图整体构建, 建立相对准确的目标先验。

认知交互网络旨在构建多智能体关于任务目标的一致性联合表征。具体而言, 该网络在空间权重注意力图 $A(p)$ 的约束和引导下, 对多智能体的独立感知结果 E_0 进行联合表征, 得到语义特征图 \tilde{M}^S :

$$\tilde{M}^S = F(A(p) \oplus E_0) \quad (2)$$

其中, $F(\cdot)$ 表示多智能体联合表征操作, \oplus 为先验嵌入操作。

多智能体需要交互并共享信息以协同完成双任务。然而, 在通信信道进行交互传输的过程中, 不可避免地引入语义特征恢复过程中产生的语义噪声以及通信信道固有的信道噪声。为了抑制这些噪声, 我们引入掩码多头自注意力机制。该机制通过对特征间的长程依赖关系进行建模, 并对低置信度特征进行

动态掩码, 抑制了非关键特征与传输噪声。处理后的特征信息经转置卷积, 上采样映射至真实无线电地图相匹配的空间维度, 最终输入到并行全连接网络 n_o 。协同输出无线电地图 \hat{M}_Γ 和辐射源定位 \hat{D}_Γ 的估计如下:

$$(\hat{M}_\Gamma, \hat{D}_\Gamma) = n_o(\hat{M}) \quad (3)$$

其中, n_o 表示参数集为 σ 的全连接网络。

真实的电磁环境具有未知、动态变化等不确定因素, 因此, 系统必须具备自学习能力与自适应调整的能力, 关键挑战在于如何根据动态变化的环境, 自适应地调整系统中多个模块的关键超参数。为此, 在感知-认知-行为的基础上, 设计了任务目标导向的智能体强化学习反馈策略。该策略的核心在于有效地根据实际环境, 对感知模型、认知交互网络和关于双任务的行为决策中的关键超参数进行动态更新。

2 基于智能体认知交互的电磁频谱具身智能系统

本文提出的基于智能体认知交互的电磁频谱具身智能系统由感知模型引导的认知交互网络和基于 PPO 强化学习的双任务信息交互联合反馈策略组成, 构建了一个“感知-认知-行为”的反馈闭环, 实现无线电地图构建和非合作辐射源定位双任务的同时完成和协同优化。

2.1 基于多智能目标联合表征的任务驱动认知交互

2.1.1 辐射源目标空间存在共识感知模型

受随机噪声及多径干扰等影响,在缺乏信源和信道先验信息时难以有效分离多源信号。而无线电地图构建与辐射源定位任务是典型的高度非凸非线性

参数估计问题,难以直接求解。因此,通过目标辨识降低信号混叠引起的参数估计偏差,去除其导致的采样野值和定位鬼点,建立观测数据与辐射源目标之间的内在关联是无线电地图构建与辐射源定位任务的首要任务。为此,本文提出了一个共识感知模型。如图2所示,具体包括连续信号场的预测分布图表示和二维位置编码模块。

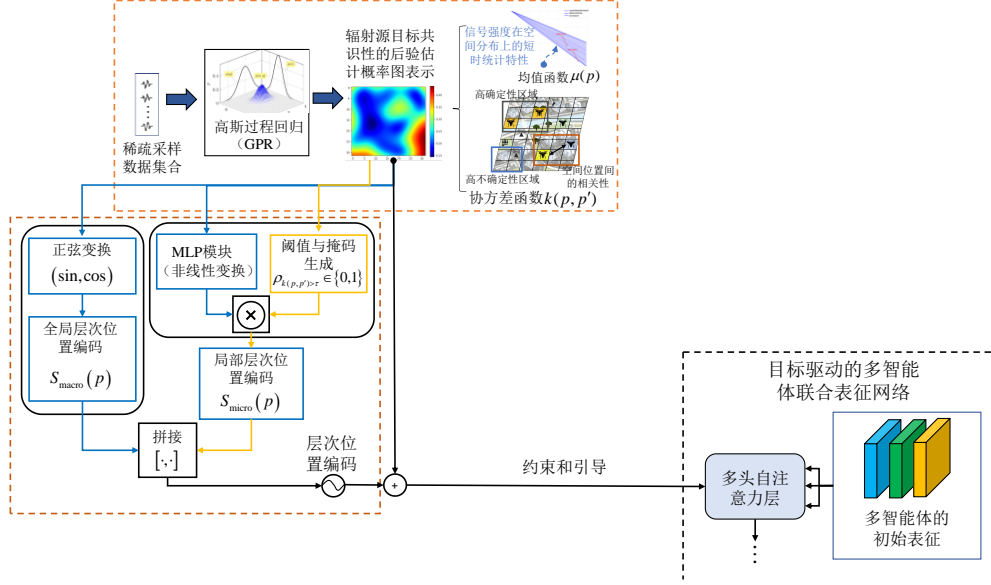


图2 辐射源目标空间存在性共识感知模型图

Figure 2 Consensus sensing model for the spatial presence of emitter targets

(1)连续信号场的预测分布图表示。根据区域内网格的空间集合 P ,将多智能体的稀疏采样数据 M^S 建模为一个高斯回归过程。该过程本质上是利用信号传播衰落特性作为先验,对稀疏观测数据进行回归分析,推断在区域内各个网格的接收信号强度(RSS)的后验分布,该预测分布图 $F(p)$ 定义为

$$F(p) \sim \text{GPR}(\mu(p), k(p, p')) \quad (4)$$

其中,均值函数 $\mu(p)$ 旨在刻画信号强度在空间分布上的短时统计特性。通过引入对数距离路径损耗模型 $\|\nabla E_{\text{FSL}}\|$ 的物理先验,可反映信号强度随距离衰减的确定性规律,具体计算方法如下:

$$\mu(p) = \beta_p \|\nabla E_{\text{FSL}}\| + (1 - \beta_p) M^S \quad (5)$$

其中, β_p 为平衡权重,用于调节模型先验与数据驱动估计之间的约束和引导比例。

为了捕捉信号传播过程因噪声干扰所带来的微观随机扰动与不确定性,选用对应于一阶可微的随机过程Matern3/2核函数^[42]来计算信号协方差-空间采样位置函数 $k(p, p')$,以量化空间中任意两点间的信号相关性:

$$k(p, p') = \sigma_c^2 \left[\left(1 + \frac{\sqrt{3} \|p - p'\|}{L} \right) \times \exp \left(-\frac{\sqrt{3} \|p - p'\|}{L} \right) \right] \quad (6)$$

其中, σ_c^2 为信号方差, L 为距离尺度,决定了任意两点间信号相关性的数值。

(2)二维位置编码。为了在连续信号场的预测分布的基础上构建多智能体关于辐射源目标的辨识先验,本文通过改进传统位置编码方法,构建全局几何拓扑空间信息与多智能体共时感知辐射源目标空间存在概率估计的复合表示 $\mathbf{S}(p)$:

$$\mathbf{S}(p) = \text{Concat}[\mathbf{S}_{\text{macro}}(p), \mathbf{S}_{\text{micro}}(p)] \quad (7)$$

其中,全局几何拓扑编码器 $\mathbf{S}_{\text{macro}}(p)$ 采用正弦-余弦函数形式,生成平滑且连续的空间坐标参考,以建立网格间的几何拓扑结构与距离依赖:

$$\mathbf{S}_{\text{macro}}(p) = \left[\sin \left(\frac{p}{Z^{2/d}} \right), \cos \left(\frac{p}{Z^{2/d}} \right) \right]_{l=1}^{d/4} \quad (8)$$

其中, l 表示维度索引, d 表示位置编码的维度, Z 则是控制特征间空间相关性显示度的超参数,可通过缩放调整特征间的维度来实现。

共识感知编码器 $\mathbf{S}_{\text{micro}}(p)$ 旨在构建辐射源目标空间位置存在概率估计。针对高斯回归输出的均值为单通道标量的特征,该模块利用多层感知机(MLP)对其进行高维投影与非线性映射,并引入方差激活掩码 $\rho_{k(p,p') > \tau}$ 进行动态门控:

$$\mathbf{S}_{\text{micro}}(p) = \text{MLP}\left(\mu(p) \cdot \rho_{k(p,p') > \tau}\right) \quad (9)$$

其中,掩码函数 $\rho_{k(p,p') > \tau}$ 定义为

$$\rho_{k(p,p') > \tau} = \begin{cases} 1, & \text{if } k(p,p') > \tau \\ 0, & \text{otherwise} \end{cases} \quad (10)$$

其中, τ 是位置 p 对应网格辐射源目标空间位置存在概率的置信度阈值, $k(p,p') > \tau$ 表示位置 p 对应网格采样数据之间的空间相关性满足置信度阈值要求。

在此基础上,通过卷积操作对二维位置编码 $\mathbf{S}(p)$ 与连续信号场的预测分布图 $\mathbf{F}(p)$ 在通道维度进行初步的拼接融合:

$$\mathbf{F}_{\text{fuse}} = \text{Conv}_{1 \times 1}\left([\mathbf{F}(p), \mathbf{S}(p)]\right) \quad (11)$$

其中, $[\cdot, \cdot]$ 表示特征拼接操作。随后,经非线性激活与第二层卷积映射,并通过 Sigmoid 函数归一化,生成空间注意力权重图 $\mathbf{A}(p)$:

$$\mathbf{A}(p) = \text{Sigmoid}\left(\text{Conv}_{1 \times 1}\left(\text{ReLU}\left(\mathbf{F}_{\text{fuse}}\right)\right)\right) \quad (12)$$

综上,空间注意力权重图融合了信号场的预测分布、几何拓扑及辐射源目标空间位置存在概率估计,构建了具备物理可解释性的复合表征先验,为后续的多智能体一致性联合表征提供了具备物理可解释性的目标辨识引导。

2.1.2 空间存在性共识感知引导的认知交互网络

认知交互网络旨在利用多智能体间的信息交互,在统一的语义空间中构建任务目标的多智能体联合表征。然而,由于受限于非均匀分布多智能体的稀疏观测和智能体间的认知偏差,智能体局部观测难以在统一特征空间中实现对任务目标特征的一致性表示。为此,本文设计了目标驱动的多智能体联合表征网络,以改善多智能体间联合表征的一致性;并且构建多智能体认知交互网络,依据任务的本质性差异,执行智能体间的选择性语义交互,以补全局部观测信息,并抑制智能体间认知偏差,从而实现无线电地图构建和非合作辐射源定位的协同估计。

(1) 目标驱动的多智能体联合表征网络。对多智能体稀疏观测数据 \mathbf{M}^s 进行多层卷积的特征投影与非线性映射操作,形成多智能体关于辐射源目标的初始表征 \mathbf{E}_0 :

$$\mathbf{E}_0 = \text{Swish}\left(\text{Conv}_{1 \times 1}\left(\text{BN}\left(\text{Conv}_{3 \times 3}\left(\mathbf{M}^s\right)\right); \mathbf{W}_0, \mathbf{b}_0\right)\right) \quad (13)$$

其中,投影矩阵 $\mathbf{W}_0 \in \mathbb{R}^{3 \times d}$, d 为嵌入维度,激活函数 $\text{Swish}(\cdot) = (\cdot) \times \text{Sigmoid}(\cdot)$,在表征初期有助于促进多智能体间信息有效融合。

随后,将第 3.1.1 节生成 $\mathbf{A}(P)$ 转换为注意力计算的偏置项,引导多智能体间聚焦于联合表征辐射源目标的关键特征。相比于全连接或卷积机制,该方法使得各智能体能够相互关注其他智能体关于目标辐射源的语义信息,从而形成多智能体间关于辐射源目标特征的一致性表征:

$$\text{Attention}\left(\mathbf{Q}_h, \mathbf{K}_h, \mathbf{V}_h; \mathbf{A}\right) = \text{Softmax}\left(\frac{\mathbf{Q}_h \mathbf{K}_h^T}{\sqrt{d_k}} \cdot G\left(\mathbf{A}(P)\right)\right) \mathbf{V}_h \quad (14)$$

其中, $\mathbf{Q}_h, \mathbf{K}_h, \mathbf{V}_h$ 分别表示第 h 层自注意力机制中,初始表征 \mathbf{E}_0 经层归一化与线性变换后的查询(Query)、键(Key)和值(Value)矩阵。该机制将智能体的局部观测特征映射为查询向量,将特征间的关联索引与属性标识映射为键向量,将提取的智能体间任务目标相关语义特征集合映射为值向量。 \cdot 表示元素相乘, $G(\cdot)$ 是一个将空间注意力权重图 $\mathbf{A}(P)$ 适配到注意力矩阵维度的映射函数。

按式(15)计算的辐射源目标一致性表征 \mathbf{E}_H 经线性投影后,与初始表征 \mathbf{E}_0 进行残差连接并送入前馈网络。其中,残差连接旨在保留智能体的局部观测信息,而前馈网络则通过非线性映射对聚合后的特征进行语义重构。该过程实现了智能体间关于同一辐射源目标中不同特征间的长程依赖关系的捕捉,挖掘了多智能体间关于同一辐射源目标的特征关联。本质上,这种形成一致性联合表征的信息交互过程,是空间一致性表征学习^[43]的实现路径。其核心数学目标是确保多智能体在认知交互后,满足如下空间一致性收敛约束:

$$\lim_{t \rightarrow \infty} \left\| \mathbf{h}_i^{(t)} - \mathbf{h}_j^{(t)} \right\|_2 = 0, \quad \forall i, j \in U_s \quad (15)$$

其中, t 表示交互迭代轮次, $\mathbf{h}_i^{(t)}$ 与 $\mathbf{h}_j^{(t)}$ 分别表示智能体 i 与 j 在 t 次交互后对同一目标的联合表征向量, U_s 为智能体集合。该式表明,随着交互过程的深入(即 $t \rightarrow \infty$),系统最终收敛至全局一致性的联合表征。

(2) 任务驱动的多智能体认知交互网络。为了实现双任务的协同估计,需要根据无线电地图构建与辐射源定位任务需求,设计多智能体间的语义交互机制,用于实现语义特征的动态筛选,从而生成包含双任务关键信息的语义特征图 $\tilde{\mathbf{M}}^s$ 。

具体而言,初始化模块利用自适应卷积核配合 He-normal 权重初始化及批归一化,建立多智能体联合表征的特征映射基底,并引入 LeakyReLU 激活函数

($\alpha=0.3$)增强非线性表达。第 s 层卷积运算定义为

$$X_s = \delta(W_s * X_{s-1} + b_s) \quad (16)$$

其中, $*$ 表示卷积操作, δ 为LeakyReLU激活函数, W_s 与 b_s 分别为第 s 层的权重矩阵与偏置项。随后,级联的中间模块通过堆叠恒定滤波器进一步深化多智能体间关于双任务的语义交互深度,并利用后置平均池化(步长为2)实现关于任务语义特征的选择与降维:

$$Y_s = \frac{1}{p_r p_j} \sum_{r,j \in p} Y_{s-1}^{r,j} \quad (17)$$

最后,终端模块依据特征图尺度 $r \times j$ 动态调整滤波器数量 $f_{\text{final}} = \lfloor c_{\text{len}}/r \cdot j \rfloor$,以完成多智能体间对深度语义特征的选择性交互,最终输出包含无线电地图构建与辐射源定位任务语义信息的语义特征图 \tilde{M}^s 。并且针对多智能体交互过程中精细空间拓扑信息丢失的问题,引入类似式(8)的几何位置编码 $E_{(i,j,2k)}$ 。其中, (i,j) 为二维坐标, k 为维度索引。将 $E_{(i,j,2k)}$ 与语义特征图 \tilde{M}^s 进行拼接操作,能够有效补充智能体间精确的相对空间拓扑信息,为协同交互选择策略提供参考。

在系统执行任务驱动的选择性语义交互得到语义特征图 \tilde{M}^s 的过程中,智能体间的语义噪声可能会导致信息融合后出现任务精度的损失;同时语义特征仍需要编码成比特流并通过信道传输,可能会受到信道噪声的影响。因此,在接收端有效抑制上述两类噪声非常重要^[31]。为此,利用掩码多头自注意力网络层对低置信度特征进行动态过滤以抑制噪声,由转置卷积网络完成任务相关的语义特征恢复,利用并行全连接层输出无线电地图构建和非合作辐射源定位任务相关的语义特征交互结果,设计损失函数如下:

$$L_{\text{RES}} = \omega_{\hat{M}} L_{\text{MSE}}(\hat{M}_{\Gamma}, M_{\Gamma}) + \omega_{\hat{D}} L_{\text{MAE}}(\hat{D}_{\Gamma}, D_{\Gamma}) \quad (18)$$

其中, M_{Γ}, D_{Γ} 分别为真实无线电地图和真实辐射源位置坐标, $\omega_{\hat{M}}, \omega_{\hat{D}}$ 分别为无线电地图构建和非合作辐射源定位任务的计算权重。具体任务相关的损失函数分别为

$$L_{\text{MSE}} = \frac{1}{U} \sum_{u=1}^U (m_u - \hat{m}_u)^2 \quad (19)$$

$$L_{\text{MAE}} = \frac{1}{N} \sum_{n=1}^N (|l_n - \hat{l}_n| + |w_n - \hat{w}_n|) \quad (20)$$

其中, m_u 和 \hat{m}_u 表示第 u 个采样点的真实值和多智能体认知交互网络的估计结果。 l_n, w_n, \hat{l}_n 和 \hat{w}_n 分别为第 n 个辐射源的真实横纵坐标和多智能体认知交互网络估计的横纵坐标。

2.2 基于PPO强化学习的双任务信息交互反馈策略

在感知模型与认知交互网络基础上,本文引入

PPO强化学习设计了双任务信息交互反馈策略。该策略通过引入经验记忆的反馈更新机制,构建“感知-认知-行为”的闭环智能体交互系统。这使得多智能体能够通过持续的交互迭代,形成具备记忆演化与经验累积能力的自学习、自纠错、自组织的具身智能闭环,如图3所示。

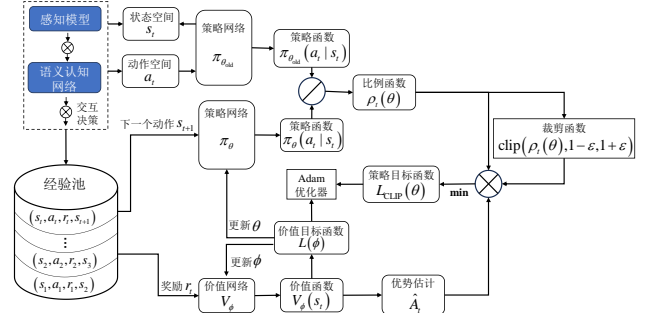


图3 基于PPO强化学习的双任务信息交互反馈策略图

Figure 3 Dual-task information interaction feedback strategy based on PPO reinforcement learning

2.2.1 基于PPO强化学习的信息交互反馈策略

本文设计了基于PPO强化学习的信息交互反馈机制,将多智能体交互过程建模为当前交互策略、平均交互策略和输出最优交互策略的协同演化机制,以实现“感知-认知-行为”的具身演化。在每一个训练的时间步 t 中,智能体根据当前状态 s_t 执行动作 a_t (即调整特征筛选权重),并获得反馈的奖励 r_t 。这些由 (s_t, a_t, r_t) 构成的状态-动作-奖励被实时存储于经验回放缓冲区,形成系统的短时记忆。随后,系统利用广义优势估计(GAE)计算优势函数 \hat{A}_t ,量化当前动作相较于平均交互策略的优劣,通过价值评估识别出哪些能够有效降低双任务损失、控制系统熵增。这一迭代过程通过遍历策略网络参数,旨在最小化无线电地图构建与辐射源定位的综合损失。为了确保策略更新的稳定性与探索性,采用策略熵约束的裁剪目标函数:

$$L_{\text{CLIP}}(\theta) = \mathbf{E}_t \left[\min(\rho_t(\theta) \hat{A}_t, \text{clip}(\rho_t(\theta), 1 - \epsilon, 1 + \epsilon) \hat{A}_t) + C \cdot H(\pi_\theta(\cdot | s_t)) \right] \quad (21)$$

其中, $\rho_t(\theta) = \pi_\theta(a_t | s_t) / \pi_{\theta_{\text{old}}}(a_t | s_t)$ 为比例函数, ϵ 为裁剪超参数, H 为策略熵, C 为熵系数。最终,通过时序差分误差最小化价值函数 $V_\phi(s_t)$,系统将迭代产生的最优交互策略转化为长时记忆,实现复杂电磁环境下双任务精度的持续优化。

2.2.2 状态、动作与奖励函数设计

为了构建“感知-认知-行为”的具身闭环,以多智

智能体根据双任务估计精度相关的状态空间为基础,设计感知、联合表征与交互决策的关键超参数动态调整交互策略,并基于奖励函数量化系统熵减收益。本文设计了关于无线电地图构建和辐射源定位任务估计精度的状态空间、具备精细调节能力的动作空间以及兼顾双任务精度与交互效率的多目标奖励函数。智能体在 t 时刻的状态 \mathbf{s}_t 需反映当前训练性能与优化趋势,定义为多维状态向量:

$$\mathbf{s}_t = \left[L_M(t), L_D(t), \bar{L}_{\text{total}}(t-1:t), \Delta L(t), \omega_{\tilde{M}}, \omega_{\tilde{D}}, \eta_t \right]^T \quad (22)$$

其中, L_M 与 L_D 分别为当前批次的双任务损失, $\bar{L}_{\text{total}}(t-1:t)$ 为 t 个批次的任务估计精度对比于上个时刻任务估计精度的滑动平均值, $\Delta L(t)$ 为损失的一阶差分。为实现对超参数的精细调控,定位动作空间为 $\mathbf{a}_t \in \mathbb{R}^{N_f}$, 其中 N_f 为语义特征图 $\tilde{\mathbf{M}}^S$ 的通道数。动作向量的每个分量 $a_t^{(i)} \in [0, 1]$ 作为软注意力权重,作用于对应通道的语义特征 $f_t^{(i)}$, 使传输前的语义特征被调制为 $\tilde{f}_t^{(i)} = a_t^{(i)} \cdot f_t^{(i)}$ 。当权重更接近 1 时,表示该通道语义特征对当前任务更关键,需充分保留以支持后续语义恢复;当权重较小时,表示该通道信息在当前阶段更可能冗余或与任务相关性较低,可在传输前抑制。相较于二值化(0/1)的硬选择,该连续软筛选机制赋予了系统精细化的平滑调制能力。它允许智能体在特征保留与抑制之间进行非线性的自适应权衡,从而更灵活地设计交互策略,适应多任务估计的电磁环境场景。为协同优化双任务精度并提升交互效率,本文构建了复合奖励函数:

$$r_t = -\left(\alpha_1^{(t)} \cdot L_M(t+1) + \alpha_2^{(t)} \cdot L_D(t+1) \right) + \beta^{(t)} \times \log \left(1 - \frac{\|\mathbf{a}_t\|}{N_f} \right) \quad (23)$$

该奖励在鼓励智能体有效降低下一时刻双任务精度损失的同时,激励智能体压缩传输数据量,其中 $\alpha_1^{(t)}$ 与 $\alpha_2^{(t)}$ 为平衡无线电地图构建与辐射源定位重要性的系数; $\|\mathbf{a}_t\|$ 表示动作向量分量之和,其与语义特征通道数 N_f 的比值可近似表征语义特征传输的激活率。对压缩率采用对数变换 $\log(1 - (\cdot))$, 可在高效压缩区间提供更明显的正向激励,并避免在压缩率接近上限时奖励过快增长,从而减少因过度追求效率而显著牺牲精度的风险。超参数 $\beta^{(t)}$ 控制效率奖励的尺度,用于调节精度与效率的全局权衡。

3 仿真结果及分析

为验证系统框架的有效性,本文数据采用文献[22]发布的开源数据集。该数据集基于 Gudmundson 相关模型^[44]通过 Wireless InSite 软件生成。在仿真实验

中,尺寸为 $L \times W = 100 \text{ m} \times 100 \text{ m}$ 的目标区域 Γ 划分为 32×32 的网格,网格间距为 3.125 m,中心频率设置为 1.4 GHz。为增强结果的可信度,具有 1.5 m 高度的辐射源的位置、数量被均匀地随机放置,并且在每个信道中以在 18 dBm~23 dBm 之间随机的功率进行发射。部署 6 架无人机组成的传感器集群,沿优化后的随机轨迹执行稀疏采样。初始采样率设置为 5%。训练集包含了 40 000 个样本,测试集和训练集按 9:1 比例划分。训练超参数设置为 100 个训练周期,批量大小为 27,初始学习率为 0.000 4,详细仿真参数设置如表 1 所示。

表 1 仿真参数设置

Table 1 Simulation parameters

参数	取值	描述/依据
带宽	5 MHz	中低频段设置,支持光线追踪算法模拟多径效应和频率响应
中心频率	1.4 GHz	模型的路径损耗和阴影衰落建模
发射功率	[18,23] dBm	辐射源功率设置
路径损耗指数	3.0	城市环境典型值
阴影衰落方差	8 dB	相关阴影模型参数
阴影相关系数	0.95	相关阴影模型参数
噪声标准差	1.0	实际感知物理噪声干扰

图 4 展示了系统在无线电地图构建与非合作辐射源定位双任务上随训练轮次变化的训练损失曲线与验证损失曲线。其中,无线电地图构建任务采用式(19)所示的均方误差(MSE)量化构图偏差;辐射源定位采用式(20)所示的平均绝对误差(MAE)衡量定位精度。

如图 4 所示,两项任务的训练与验证损失曲线均快速下降并最终趋于稳定,表明系统具有良好的收敛性。图 4(a)为构图任务训练阶段与验证阶段的损失曲线对比,两者均在约 40 个训练周期后稳定收敛,曲线接近且未出现明显发散,说明模型在该任务上泛化性能良好。图 4(b)为定位任务训练阶段与验证阶段的损失曲线对比,两者均在约 35 个训练周期后收敛,验证损失略高于训练损失,但差距不大,且未观察到明显的过拟合现象。

图 5 展示了系统在无线电地图构建与辐射源定位任务上的可视化结果。其中,图 5(a)为真实无线电地图和辐射源真实位置(以白色叉标注),图 5(b)为稀疏、非均匀采样的采样图 \mathbf{M} (采样点数为 51,对应采样率为 5%),图 5(c)为系统的无线电地图和辐射源定位估计结果(以黑色叉标记)。

可视化结果表明,在 5% 的稀疏采样条件下,系统仍能够较好地恢复无线电地图的整体结构与局部细节;辐射源信号分布边界较清晰,未出现明显伪影;

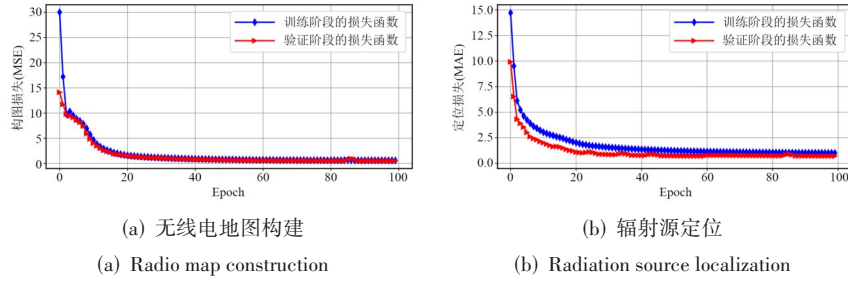


图4 系统所得双任务的训练和验证损失图

Figure 4 Training and validation loss curves of the dual-tasks obtained by the system

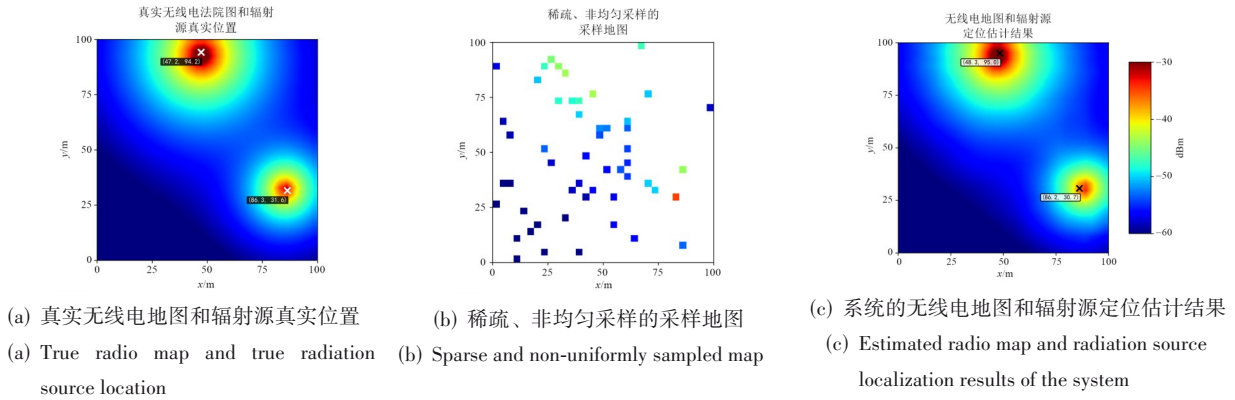


图5 无线电地图构建和辐射源定位可视化图

Figure 5 Visualization of radio map construction and radiation source localization

辐射源定位估计结果与真实位置高度吻合,这表明该系统能够有效同步执行无线电地图构建与辐射源定位任务。

图6展示了在不同信噪比条件下,系统同时优化构图和定位双任务,以及分别独立执行构图、定位任务时的误差变化趋势对比图。其中,图6(a)为构图任务的均方根误差(RMSE)^[22]对比,图6(b)为定位任务的平均绝对误差(MAE)对比。实验基于 $100\text{ m} \times 100\text{ m}$ 区域的1000次随机生成场景,信噪比由10 dB逐渐增

至30 dB,每个测试场景独立重复测试100次后对结果取平均值,以降低随机性误差的影响。实验结果表明,随着信噪比从10 dB增加至30 dB,通过建立双任务协同优化机制,无线电地图构建与辐射源定位的精度均显著优于系统独立执行单任务的方案。这一结果充分证明了在5%稀疏采样条件下,联合优化机制能有效提升系统性能。具体而言,该系统通过协同优化,将无线电地图构建和辐射源定位的精度分别平均提升了18.2%和43.5%。



(a) 不同任务模式下无线电地图RMSE随SNR变化曲线

(a) RMSE of radio map versus SNR under different task modes

(b) 不同任务模式下辐射源定位误差随SNR变化曲线

(b) Localization error of radiation source versus SNR under different task modes

图6 比较系统同时完成双任务和分别完成双任务的误差随信噪比由10 dB增至30 dB时的变化趋势图

Figure 6 Error comparison between simultaneous and separate execution of dual-tasks with SNR increasing from 10 dB to 30 dB

图7展示了在观测噪声和视角差异的条件下,不同数量的智能体对目标特征空间的一致性估计结果。实验模拟4~6个智能体对感兴趣区域电磁环境进行观测,通过信息交互实现2个辐射源目标特征空间的一致性联合表征。我们采用平均特征分歧(MFD)^[43]作为量化指标,计算智能体在特征空间中语义表征向量的欧氏距离,以衡量智能体间的认知差异。实验结果表明,随着交互迭代次数的增加,多智能体的平均特征分歧(MFD)均呈现指数级下降趋势,并最终在8轮次内收敛至0.001以下,验证了系统中多智能体达成一致性联合表征的能力。

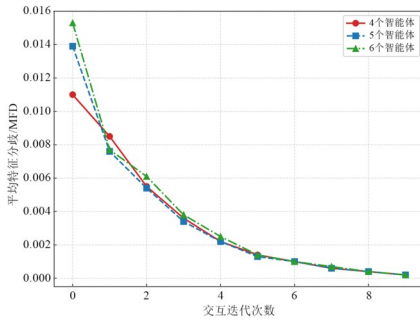


图7 随着交互迭代过程,智能体间特征分歧的演化趋势图

Figure 7 Evolution of feature divergence among agents over interaction iterations

为进一步评估所提系统在无线电地图构建任务上的性能优势,本文将所提方法与多种基线算法进行对比。

(1) 克里金插值算法,其正则化参数为 1×10^{-5} , 高斯径向基函数的参数为 $\sigma_k = 5\sqrt{\Delta_y N_y \Delta_x N_x}$, 其大约是采样观测值中两个点之间平均距离的5倍。

(2) 具有正则化参数 1×10^{-4} 的拉普拉斯多核算法,其参数均匀分布在 $[0.1\sigma_k, \sigma_k]$ 之间。

(3) 具有正则化参数 1×10^{-4} 的高斯多核算法,包含20个高斯核,其参数分布在0.005~100 m之间。

(4) 通过正则化参数 1×10^{-5} 的核范数最小化来完成矩阵求解,正则化参数为 3×10^{-1} 和高斯径向基函数作为基准。

(5) $K=5$ 的K近邻(KNN)算法。

(6) 全卷积深度补全自动编码器算法。文献[22]中通过一种全卷积深度补全自动编码器的编码器-解码器。

(7) 语义通信。文献[33]中通过量化语义提取策略提取语义特征的补全算法。图8展示了不同采样率(不同采样率对应的采样点的数量)下无线电地图构建的RMSE,以及在5%采样率(采样点数量为51)条件下,各算法在信噪比变化时的RMSE。

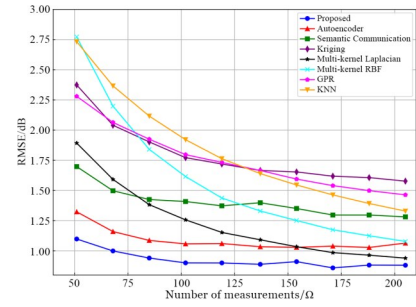


图8 采样率对不同算法在无线电地图构建任务下的精度对比

Figure 8 Accuracy comparison of different algorithms for radio map construction under varying sampling rates

图8表明,在不同采样率(对应采样点数量)下,所提方法的无线电地图构建的RMSE均低于各基线方法。以5%采样率(采样点数量为51)为例,相比克里金插值、拉普拉斯核与RBF核两种多核方法、高斯过程回归、KNN、全卷积深度补全自动编码器以及语义通信方法,本文所提方法的RMSE分别降低60.95%、50.81%、65.69%、57.27%、63.93%、32.55%、43.31%。

图9表明,在不同信噪比条件下,本文所提方法的无线电地图构建的RMSE仍低于各基线方法。以10 dB为例,相比克里金插值、拉普拉斯核与RBF核两种多核方法、高斯过程回归、KNN、全卷积深度补全自动编码器以及语义通信方法,所提方法的RMSE分别降低44.5%、33.5%、49.4%、32.3%、40.5%、10.6%和15.7%。

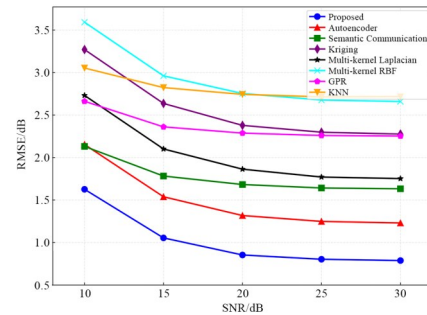


图9 信噪比变化对不同无线电地图构建任务算法的精度对比

Figure 9 Accuracy comparison of different algorithms for radio map construction under varying SNRs

为评估所提基于PPO强化学习的双任务信息交互联合反馈策略在无线电地图构建与辐射源定位双任务上的性能优势,本文开展了不同交互反馈策略的对比实验。以固定策略为基准,与线性衰减与启发式策略进行对比,实验结果如图10所示,可见线性衰减与启发式策略的奖励性能提升有限,分别为28.7%和36.0%。相比之下,本文方法凭借对双任务协同反馈增益的动态优化,获得了52.2%的性能提升。

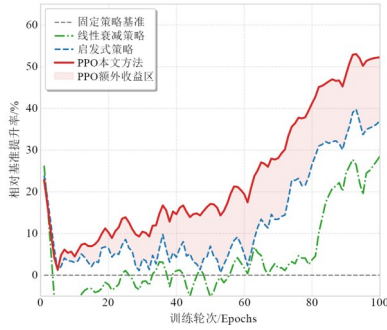


图10 本文方法与三种基准调度策略的相对性能提升率对比图

Figure 10 Relative performance improvement of the proposed method over three baseline scheduling strategies

为进一步验证系统在非合作辐射源定位任务上的可行性,本文将数据集扩展至 $10 \text{ km} \times 10 \text{ km}$ 的经典辐射源定位场景进行单独的定位验证。实验基于该测试场景在不同采样率条件下独立重复测试 500 次,得到辐射源位置估计在 X - Y 平面上的分布,以进一步验证该系统在辐射源定位任务中的性能优劣,结果如图 11 所示(红色圆点为真实位置,蓝色圆点为估计位置)。

由图 11 可知,在四种采样率条件下,估计位置总体集中在真实辐射源位置附近;随着采样率提高,估

计分布进一步收敛,并且定位的相对精度(定位误差以感知范围的百分比表示)^[45]稳定维持在 1.2% 以内。这说明该系统在大范围场景下仍能获得较稳定的定位结果,从而验证了其在非合作辐射源定位任务上的可行性。统计结果表明,相较于传统的 RSS 定位基准算法,本系统的定位误差降低了 65.7%。

为了验证所提具身智能交互系统在无人机智能体嵌入式设备上的实际部署能力,本文结合理论复杂度量与边缘算力约束进行了深入分析。在模型空间复杂度方面,系统参数的统计范围严格覆盖了从感知前端到决策后端的完整映射结构。在具身感知层并非简单的特征提取,而是包含了用于构建共识感知辐射源目标空间位置存在概率估计的 MLP 映射与空间注意力融合卷积层,配合独立感知特征提取模块,该部分权重参数量约为 0.59 万个;具身认知交互层依赖于多头自注意力机制实现多智能体的联合表征,其参数统计为 1.03 万个;具身反馈层采用 Actor-Critic 双决策网络,参数量合计 0.92 万个。将上述环节精确累加,系统整体的模型参数总量约为 2.54 万个。若以 Float64 高精度格式存储,该系统仅需占用约 203 KB 的内存空间。考虑到当前主流嵌入式 MCU 芯片(如 STM32H7 系列)通常具备 1 MB 以上的片上静态随机存取存储器(SRAM),该模型完全满足嵌入

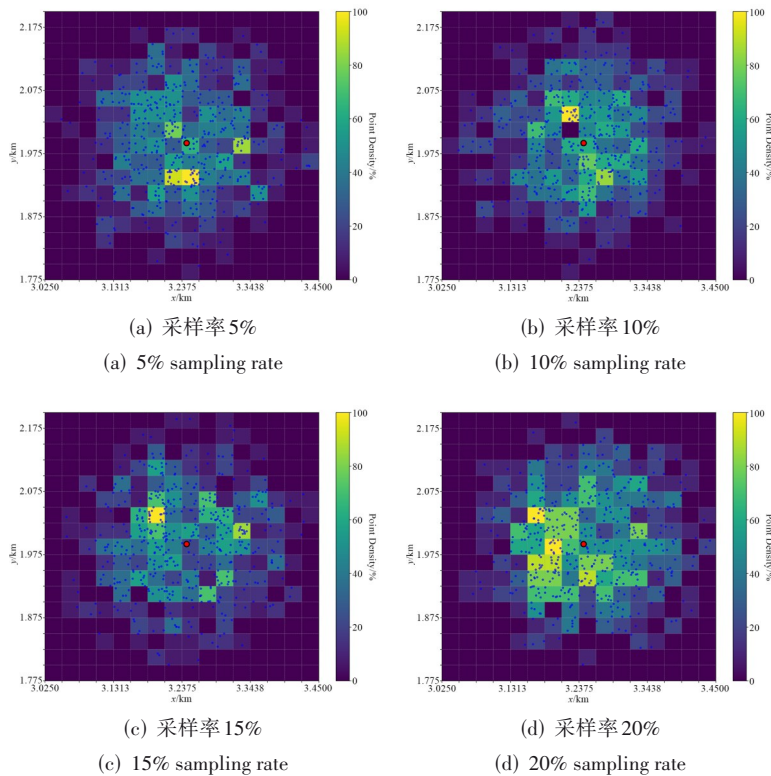


图11 系统在定位任务中四个采样率下的结果分布图

Figure 11 Result distribution of the system in the localization task under four sampling rates

式芯片严苛的片上存储限制,无需依赖外部存储器即可高效运行。

在时间复杂度与实时性方面,计算开销的核算严格遵循数据在具身闭环中的流动路径。其涵盖了从稀疏采样点的物理先验推断(GPR)、二维位置编码生成到最终决策输出的全过程。得益于特征通道被严格限制在12以内,系统处理单帧数据的全链路总浮点运算数被精确控制在 0.055×10^9 左右。对比经典计算平台 NVIDIA Jetson Nano 的 GFLOPs 为 472 算力峰值,本系统的计算负载仅占其理论上限的 0.012%;即使在低功耗的 Raspberry Pi 4B 平台上,其算力占用率也仅维持在 0.41%。实验实测表明,该模型在通用 CPU 上的单帧端到端推断耗时稳定在 20 ms 以内,具备充足的实时性余量,表明本系统计算成本可控,能支撑采样率从 20% 降低至 5% 的非对称优化策略,具有显著的工程实用价值。

4 结论

本文面向无线电地图构建与非合作辐射源定位的协同优化,提出一种基于智能体交互的电磁频谱多任务具身智能系统框架。通过层次化改进的位置编码器,并结合高斯回归过程,设计了辐射源目标空间存在性概率感知模型,充分利用多智能体的共识先验为智能体信息交互提供具有物理可解释性的约束和引导;该框架采用多智能体统一语义空间中的辐射源目标一致性联合表征网络,同时构建任务和目标驱动的智能体协同交互策略,将任务、目标与信息解耦,并结合 PPO 强化学习设计双任务的反馈策略,重构感知-认知-交互决策闭环,将原本相互制约的无线电地图和辐射源定位双任务转化为相互促进的自组织、自学习、自成长的迭代交互具身智能系统。基于 Gudmundson 开源数据集的试验结果验证了该框架在双任务同时优化方面的可行性和提升无线电地图构建和辐射源定位精度方面的有效性。在无线电地图构建和非合作辐射源定位精度上,该框架均优于现有基线方法,尤其是在存在多个信号源且低采样密度的场景中表现突出。

致谢 本文研究工作受国家自然科学基金项目(No.62101293)、广西自然科学基金面上项目(No.2023GXNS-FAA026058)、认知无线电与信息处理省部共建教育部重点实验室主任基金项目(No.CRKL220110)资助,受广西人工智能学院提供平台支持。

参考文献

- [1] Zhang Shoubin, Li Zhimeng, Li Haojin, et al. Novel radio environment map construction scheme for 3-D and full band for modern internet of things applications[J]. IEEE Internet of Things Journal, 2025, 12(9): 12419-12432.
- [2] Zhang Shoubin, Wang Hongjun, Shen Zhexian, et al. Multiple radiation source localization in IoT: A radio-map-assisted scheme[J]. IEEE Internet of Things Journal, 2025, 12(20): 42484-42500.
- [3] Tsoulos G, Athanasiadou G, Zarbouti D, et al. Empirical analysis of 5G deployments: A comparative assessment of network performance with 4G[J]. AEU-International Journal of Electronics and Communications, 2024, 186: 155479.
- [4] Jaiswal R K, Elnourani M, Deshmukh S, et al. Leveraging transfer learning for radio map estimation via mixture of experts[J]. IEEE Transactions on Cognitive Communications and Networking, 2026, 12: 846-863.
- [5] Yu Weijia, Du Jianhe, Chen Yuanzhi, et al. Vehicle localization based on Bayesian tensor decomposition in intelligent transportation systems[J]. IEEE Transactions on Intelligent Transportation Systems, 2025, 26(10): 16204-16221.
- [6] Kiouvrekis Y, Panagiotakopoulos T, Nousi E, et al. Explainable machine learning for radio environment mapping: An intelligent system for electric field strength monitoring[J]. IEEE Access, 2025, 13: 75104-75122.
- [7] Xhafa A, Fabra F, Egea-Roca D, et al. Experimental investigation of AoA estimation and antenna calibration with 5G NR signals using USRP devices[J]. IEEE Transactions on Instrumentation and Measurement, 2025, 74: 5504515.
- [8] Zhang Chengjie, Han Xinyang, Kong He, et al. Optimal sensor placement for full-set TDOA localization accounting for sensor location errors[J]. IEEE Transactions on Aerospace and Electronic Systems, 2025, 61(4): 10944-10950.
- [9] Kim J. Source localization based on FDOA and TDOA with bounded sensor location errors and changing carrier frequency[J]. IEEE Internet of Things Journal, 2025, 12(12): 22478-22488.
- [10] Fabris A, Rayel O K, Rebelatto J L, et al. AoA and RSSI-based BLE indoor positioning system with Kalman filter and data fusion[J]. IEEE Internet of Things Journal, 2025, 12(11): 15348-15359.
- [11] Ko C H, Wu S H. A framework for proactive indoor positioning in densely deployed WiFi networks[J]. IEEE Transactions on Mobile Computing, 2020, 21(1): 1-15.
- [12] Yoo J. Multiple fingerprinting localization by an artificial neural network[J]. Sensors, 2022, 22(19): 7505.
- [13] Alrabaee S, Debbabi M, Wang Lingyu. A survey of bina-

- ry code fingerprinting approaches: Taxonomy, methodologies, and features[J]. *ACM Computing Surveys (CSUR)*, 2023, 55(1): 19.1-19.41.
- [14] Chen Qidong, Liu Rui, Yan Qiuzhen, et al. Localization of multiple GNSS interference sources based on target detection in C/N_0 distribution maps[J]. *Remote Sensing*, 2025, 17(15): 2627.
- [15] Liu Jiayu, Liu Xiaodong, Liang Hongtao, et al. A data-and-semantic dual-driven intelligent inference framework for simultaneously spectrum map construction and signal source localization[J]. *IEEE Internet of Things Journal*, 2025, 12(12): 18750-18764.
- [16] Zhang Songyang, Yu Tianhang, Choi B, et al. Radiomap inpainting for restricted areas based on propagation priority and depth map[J]. *IEEE Transactions on Wireless Communications*, 2024, 23(8): 9330-9344.
- [17] Feng Bin, Zheng Meng, Liang Wei, et al. A recent survey on radio map estimation methods for wireless networks[J]. *Electronics*, 2025, 14(8): 1564.
- [18] Zhao Jianli, Feng Xingzhao, Fang Sheng, et al. Tensor completion via nonlocal tensor wheel decomposition for hyperspectral image recovery[J]. *IEEE Geoscience and Remote Sensing Letters*, 2025, 22: 5504905.
- [19] Chen Guokai, Liu Yongxiang, Zhang Jianzhao, et al. GPRT: A Gaussian process regression-based radio map construction method for rugged terrain[J]. *IEEE Internet of Things Journal*, 2025, 12(13): 23905-23920.
- [20] Ma Yilin, Zhang Chiya, He Chunlong, et al. Radio map estimation using a CycleGAN-based learning framework for 6G wireless communication[J]. *Digital Communications and Networks*, 2025, 11(6): 1822-1830.
- [21] Maiti P, Mitra D. Ordinary kriging interpolation for indoor 3D REM[J]. *Journal of Ambient Intelligence and Humanized Computing*, 2023, 14(10): 13285-13299.
- [22] Teganya Y, Romero D. Deep completion autoencoders for radio map estimation[J]. *IEEE Transactions on Wireless Communications*, 2022, 21(3): 1710-1724.
- [23] Wang Jie, Zhu Qiuming, Lin Zhipeng, et al. Sparse Bayesian learning-based 3-D radio environment map construction: Sampling optimization, scenario-dependent dictionary construction, and sparse recovery[J]. *IEEE Transactions on Cognitive Communications and Networking*, 2024, 10(1): 80-93.
- [24] Chen Qi, Huang Ming, Yang Jingjing. REM-Net+: Quantified 3D radio environment map construction guided by radio propagation model[J]. *IEEE Transactions on Vehicular Technology*, 2026, 75(4): 6351-6362.
- [25] Čolaković A. IoT systems modeling and performance evaluation[J]. *Computer Science Review*, 2023, 50: 100598.
- [26] 陈兆, 王玫, 周瞰, 等. 基于多参数融合的超宽带基站布设优化方法[J]. *电子学报*, 2023, 51(10): 2855-2865.
- Chen Zhao, Wang Mei, Zhou Zou, et al. Optimization method for ultra-wideband base station configuration based on multi-parameter fusion[J]. *Acta Electronica Sinica*, 2023, 51(10): 2855-2865. (in Chinese)
- [27] 周瞰, 王若楠, 纪冰辉, 等. 一种基于感知阵型的无人机蜂群动态侦察任务调度系统及方法: CN202110700086.9[P]. 2024-01-09.
- Zhou Zou, Wang Ruonan, Ji Binghui, et al. Unmanned aerial vehicle swarm dynamic reconnaissance task scheduling system and method based on perceptual formation: CN202110700086.9[P]. 2024-01-09. (in Chinese)
- [28] Zhao Yue, Cheng Nan, Li Zan, et al. An efficient sensor selection algorithm for TDOA localization with estimated source position[C]//*ICC 2022-IEEE International Conference on Communications*. Piscataway: IEEE, 2022: 871-876.
- [29] Li Xianliang, Xu Sheng, Ho K C. Time-of-arrival simultaneous sensor and target localization with dynamic optimal sensor placements[J]. *IEEE Transactions on Aerospace and Electronic Systems*, 2025, 61(5): 12663-12677.
- [30] Liu Wenjie, Chen Junting. UAV-aided radio map construction exploiting environment semantics[J]. *IEEE Transactions on Wireless Communications*, 2023, 22(9): 6341-6355.
- [31] Wang Xuegang, Wang Fanggang, He Boxiang. Radio environment map reconstruction via tensor completion: Bayesian and semantic approaches[J]. *IEEE Transactions on Vehicular Technology*, 2025, 74(5): 7897-7913.
- [32] Gao Qianhao, Zhu Qiuming, Lin Zhipeng, et al. Time-variant radio map reconstruction with optimized distributed sensors in dynamic spectrum environments[J]. *IEEE Internet of Things Journal*, 2025, 12(12): 20927-20941.
- [33] Wu Wei, Chen Yu, Zhou Fuhui, et al. Semantic communications for efficient spectrum maps construction and signal source location[J]. *IEEE Wireless Communications Letters*, 2024, 13(12): 3266-3270.
- [34] Juang J C, Tsai Y F. Analysis of GNSS interference events based on TRITON GNSS-R measurements[J]. *NAVIGATION: Journal of the Institute of Navigation*, 2025, 72(4): navi.720.
- [35] Ramchandra A R, Skurdal A, Ranganathan P, et al. De-

- tecting GPS interference using automatic dependent surveillance-broadcast data[J]. Electronics, 2024, 13(16): 3145.
- [36] Li Deyi. Cognitive physics: The enlightenment by Schrödinger, Turing, and Wiener and beyond[J]. Intelligent Computing, 2023, 2: 0009.
- [37] An Zhulin, Yu Xinqing, Wang Chu, et al. Embodied intelligence: Recent advances and future perspectives[J]. The Innovation Informatics, 2025, 1(1): 100008.
- [38] Ma Nan, Pan Jia, Liu Yongjin, et al. Embodied interactive intelligence towards autonomous driving[J]. Engineering, 2026, 59: 337-351.
- [39] 阮天宸, 吴启晖, 赵世瑾, 等. 认知学习: 电磁频谱空间机器学习新范式[J]. 电子学报, 2023, 51(6): 1430-1442. Ruan Tianchen, Wu Qihui, Zhao Shijin, et al. Cognitive learning: A new paradigm for machine learning in electromagnetic spectrum environment[J]. Acta Electronica Sinica, 2023, 51(6): 1430-1442. (in Chinese)
- [40] Obayashi N, Abdulali A, Iida F, et al. Embodied intelligence paradigm for human-robot communication[J]. Science Robotics, 2025, 10(105): eads8528.
- [41] Wan Zishen, Qian Jiayi, Du Yuhang, et al. Generative AI in embodied systems: System-level analysis of performance, efficiency and scalability[C]//2025 IEEE International Symposium on Performance Analysis of Systems and Software (ISPASS). Piscataway: IEEE, 2025: 26-37.
- [42] Palar P S, Shimoyama K. Efficient global optimization with ensemble and selection of kernel functions for engineering design[J]. Structural and Multidisciplinary Optimization, 2019, 59(1): 93-116.
- [43] Roh B, Shin W, Kim I, et al. Spatially consistent representation learning[C]//Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2021: 1144-1153.
- [44] Gudmundson M. Correlation model for shadow fading in mobile radio systems[J]. Electronics Letters, 1991, 27(23): 2145-2146.
- [45] Mao Guoqiang, Fidan B, Anderson B D O. Wireless sensor network localization techniques[J]. Computer Networks, 2007, 51(10): 2529-2553.

作者简介



周 颀 男, 1983年3月出生于湖北省松滋市。现为桂林电子科技大学教授、博士生导师。主要研究方向为多模态感知与语义认知通信、电磁频谱具身智能、智能体互联网、无人集群/低轨卫星通感一体化、物联网与大数据技术应用等。中国电子学会会员编号为E19006051M。
E-mail: zhouzou@guet.edu.cn



邓添荣 男, 2001年2月出生于广东省东莞市。现为桂林电子科技大学硕士研究生。主要研究方向为电磁频谱空间认知学习的无线电地图构建。
E-mail: bitzhzzz@163.com



仇 琛 男, 1990年11月出生于江苏省如皋市。现任鹏城实验室助理研究员。主要研究方向为无人机辅助无线网络、通感一体化、语义通信技术。
E-mail: qiuch@pcl.a.cn



仇洪冰 男, 1963年2月出生于江苏省如皋市。现为桂林电子科技大学信息与通信学院教授、博士生导师。主要研究方向为宽带无线通信、通信信号处理、辐射源定位等。中国电子学会会员编号为E190013474M。
E-mail: qiuhb@guet.edu.cn