面向低空无线网络的具身智能频谱地图构建

Hongtao Liang，研究生会员，*IEEE*，Yiyao Wan，Jiahuan Ji，会员，*IEEE*，
Lingda Meng，Fuhui Zhou，高级会员，*IEEE*，以及 Qihui Wu，会士，*IEEE*

arXiv:2611.32799v1 [eess.SP] 24 Nov 2026

摘要—**利用无人机（UAV）准确构建频谱地图，对于低空无线网络（LAWNs）中的智能资源管理至关重要。尽管已有多种构建方法得到研究，但在极度稀疏采样条件下，这些方法的性能往往会下降，因为纯数据驱动模型可能无法捕捉潜在的传播物理。为应对这些挑战，本文提出了一种具身智能框架，通过协同整合闭环感知—认知—执行（PCE）耦合与具身探索，实现精准且高效的频谱构建。在此框架基础上，设计了一种相应的具身认知与决策网络（ECDNet），该网络包含三个紧密耦合的模块。具体而言，ECDNet利用距离感知卷积、语义桥和自适应特征对齐机制，建立了构建与决策之间的双向信息流。仿真结果表明，在非重复访问螺旋协议下，ECDNet在4.00%的采样率下实现了2.787 dBm的均方根误差（RMSE）。此外，本文开发并进行现场部署了一套实用的基于无人机的频谱构建平台，为原型级证据提供了支持，证明感知、认知和动作模块能够在受尺寸、重量和功耗约束的无人机边缘设备上作为闭环执行。**

索引词—**具身智能、频谱地图构建、无人机、无线电地图估计、主动探索、低空无线网络。**

I. 引言

大规模无线网络

OW-ALTITUDE

（LAWNs）正逐渐成为下一代智能系统的关键基础设施 [1]–[4]。尽管密集部署的无人机能够带来显著的应用效益，但其日益普及也造成了前所未有的频谱管理挑战。为了在此类复杂的大规模无线网络中实现智能资源分配和干扰抑制，利用无人机进行准确的频谱地图构建至关重要，从而提供电磁环境的全面时空频率剖面 [5]。然而，无人机平台受到严格的尺寸、重量和功耗约束，严重限制了其运行续航时间。这些约束使无人机只能在少量空间位置采集测量值，相对于目标地图分辨率，其采样密度通常低于 5%

H. Liang、L. Meng 和 Q. Wu 就职于南京航空航天大学电子与信息工程学院，中国南京 210000。（电子邮箱：ceie.lht@nuaa.edu.cn、042300229@nuaa.edu.cn、wuqihui2014@sina.com）。Y. Wan、F. Zhou 和 J. Ji 就职于南京航空航天大学人工智能学院，中国南京 210000。（电子邮箱：yiyaowan@nuaa.edu.cn、zhoufuhui@ieee.org、jiahuan.ji@nuaa.edu.cn）。

[6]。因此，利用此类稀疏测量值构建高保真频谱地图至关重要。

传统的频谱地图构建方法主要采用纯数据驱动的解耦范式 [7]。该范式将频谱地图构建分为两个相互独立的模块，即根据有限数据推断连续频谱图的构建模块 [8]–[18]，以及选择信息丰富的感知位置的采样模块 [19]–[24]。尽管这种模块化设计简化了构建与采样的独立实现，但当缺乏将不断演化的构建状态与采样策略进行交互的机制时，可能会限制联合优化。具体而言，在没有在线构建到采样反馈的方法中，当前构建状态无法直接指导下一次感知决策 [19]，[20]。此外，在数据极度稀缺的情况下，数据驱动的构建方法可能难以生成符合物理规律的结果，从而产生无法反映底层传播特性的非真实伪影 [10]–[12]。

具身智能（EI）通过建立感知、认知与执行（PCE）之间的闭环耦合，为应对这些挑战提供了一种有前景的框架 [7], [25]。这一机制将被动观测转变为主动探索，使构建过程中的认知洞见能够直接指导采样决策。这使无人机智能体能够根据不断演化的构建状态动态调整其采样决策并收集额外的测量值，以减少构建误差，从而支持更高效地利用有限的采样预算。此外，物理学引导的方法通过架构先验或正则化先验，将无线传播知识融入构建过程，从而能够在稀疏测量条件下促进符合物理规律的频谱图估计 [14], [15]。

受这些优势的启发，将具身智能应用于无人机频谱图构建，是打破解耦范式并缓解数据稀缺问题的一个颇具前景的方向。然而，由于低空无线网络的独特特性，将这一统一方法应用于频谱构建仍面临特定挑战。具体而言，与接收密集像素级反馈的传统视觉型具身智能系统不同，频谱构建依赖于从不连续空间位置采集的稀疏标量接收信号强度指示（RSSI）值，这阻碍了有效的 PCE 耦合。此外，纯数据驱动方法可能难以从极其稀疏的测量值中捕捉

复杂的电磁传播模式。因此，数据稀缺可能增加频谱图估计在物理上不合理以及泛化能力较差的风险 [5]。此外，复杂的 LAWN 环境容易受到高水平噪声和干扰的影响。由此产生的低信噪比（SNR）进一步加剧了构建难度，因此需要一种能够从稀疏且含噪的测量值中恢复可靠信号的稳健方法 [26]。

为应对这些挑战，我们提出了一种用于无人机频谱地图构建的基于EI的框架。该框架通过相应的具身认知与决策网络（ECDNet）实现。本文的主要贡献总结如下。

- 本文提出了一种用于频谱地图构建的基于EI的框架。具体而言，我们建立了协同的闭环PCE耦合作为核心机制，以弥合构建与决策之间的信息鸿沟。此外，我们将物理信息先验作为辅助技术加以整合，以增强模型在稀疏测量下的泛化能力。基于自由能原理（FEP），该框架提供了统一的优化视角，通过特定任务的构建目标函数和决策目标函数，协同组织构建过程与探索策略。

- 我们设计了ECDNet，通过三个紧密耦合的组件有效实现这一统一框架。具体而言，环境感知组件将多模态传感器测量融合为结构化输入。此外，情境认知组件利用距离感知卷积对与传播相关的结构进行编码，以稀疏数据构建频谱图。主动探索组件则利用语义桥提取认知见解，以作出智能采样决策。这些组件构成了闭环PCE循环，使无人机能够在资源受限条件下自主完成频谱地图构建。

- 开发了一种实用的基于无人机的频谱构建平台，用于支持受控仿真和原型级现场演示中的评估。具体而言，仿真结果表明，在非重复访问螺旋协议下，该平台具有更高的样本效率，在稀疏的 4.00% 采样率下达到了 2.787 dBm 的均方根误差。此外，现场测试提供了在受尺寸、重量和功耗（SWaP）约束的无人机边缘设备上实现实时闭环运行的原型级证据。

本文其余部分的组织结构如下。第二节回顾相关工作。第三节介绍基于 EI 的频谱地图构建框架。第四节详细介绍 ECDNet 架构。第五节给出实验结果并介绍现场部署。第六节总结全文。

II. 相关工作

为实现精确的频谱地图构建，研究人员采用各种构建策略开展了广泛研究。具体而言，当前方法可拼接为基于静态观测的方法和基于动态探索的方法 [8]–[18]。

A. 基于静态观测的构建方法

基于静态观测的方法部署传感器网络或预先确定的测量活动，重点研究如何从分布式静态观测中推断连续场的构建算法。这些方法包括经典的地统计学技术 [8], [9],深度学习方法 [10]–[12],以及融合物理或语义信息的框架 [13]–[18]。具体而言，早期的地统计学方法在平稳性假设下，将频谱图估计视为空间回归 [8], [9]。文献 [8]提出了空间-频率插值无线电地图，以联合利用空间相关性和跨频率相关性，并提供有理论依据的不确定性量化。然而，当平稳性假设无法准确近似异质城市传播时，其精度可能会下降。

标准批量克里金的计算成本可能随观测值数量呈立方增长。将其估计结果与采集或规划机制结合后，可以支持主动测量。

为克服这些局限，研究人员广泛探索了数据驱动深度学习方法 [10]–[12]。具体而言， [10]开发了带有跳跃连接的编码器—解码器架构，以实现端到端学习，但其生成的估计结果过度平滑，缺乏细粒度的空间细节。在此架构基础上， [11] 提出了聚合上下文变换生成对抗网络，通过上下文感知注意力提升感知质量。 [12] 中的作者引入了空间变换器，从而实现无网格插值，并增强了对空间变换的稳健性。然而，极度稀疏的测量值可能降低构建精度，并增加生成不符合物理规律的估计结果的风险。

物理信息和语义增强方法应运而生，以解决这一根本局限 [13]–[18]。物理信息神经网络在 [14] 中被提出，将偏微分方程残差作为软约束嵌入损失函数。在此基础上，针对无线场景， [15] 中的作者引入了由路径损耗模型推导出的RSSI偏微分方程。作为对物理先验的补充， [16] 提出了频率—空间推理，通过利用跨频率相关性推断缺失频段；与此同时， [17] 中的作者采用多任务学习，对地图构建和源定位进行联合优化。信道知识地图这一概念在 [18] 中被提出，作为一种统一的环境感知表示框架。尽管方法日益复杂，所有静态方法都存在一个根本局限：构建质量严重依赖于采样位置的空间分布 [9]。当这些方法作为独立的构建模块使用时，它们并未指定后续的测量位置，因此需要配套的主动采集策略。在目标区域内采集密集测量值可能成本高昂，这推动了自适应采样策略的发展 [21]。

B. 基于动态探索的构建方法

基于动态探索的方法将被动测量转变为主动数据采集，无人机不再沿固定路径飞行，而是动态优化采样轨迹，以最大化信息增益。相关研究已从地图辅助规划 [19], [20] 和不确定性驱动的主动学习 [21], [22] 发展到用于联合优化的深度强化学习（DRL） [23], [24], ；近期观点 [7] 则强调紧密的PCE耦合。

基于地图的方法利用预先计算的频谱地图进行频谱感知轨迹设计，以优化通信目标 [19], [20]。具体而言， [19] 制定了三维路径规划，以确保蜂窝连接无人机的连通性。同样， [20] 利用无线电地图设计抗干扰轨迹，在服务质量与弹性之间取得平衡。尽管这些方法展示了频谱信息的价值，但它们依赖静态地图，缺乏在线更新。因此，这些方法表现出单向信息流，代表了一种解耦范式，无法支持主动地图构建。

为解决地图质量与采样策略之间的相互依赖关系，研究人员开发了不确定性驱动的主动学习方法 [21], [22]。文献 [21] 提出了贝叶斯主动学习，在高斯过程模型上设计基于信息增益的采集函数。为将该框架扩展到空中平台， [22]中的作者利用信息论目标制定了无人机频谱测绘方法，对路径与采样进行协同设计。这些方法显著减少了测量值。然而，构建模块与决策模块仍然彼此分离，不确定性作为显式的估计器输出提供，而不是从网络内部表征中提取。

深度强化学习方法应运而生，通过端到端策略优化实现更紧密的集成 [23], [24]。 [23]中的作者研究了同步导航与无线电测绘，智能体在学习运动和测量策略的同时逐步构建地图。为提高样本效率， [24]提出了模型辅助方法，将近似物理模型纳入其中。这些方法与我们的工作最为接近。然而，它们存在耦合浅层的问题：只有低维几何特征被用作智能体状态，交互则通过标量奖励信号进行。与能实现密集像素级反馈的机器人主动SLAM [27] 相比，频谱构建面临独特挑战：稀疏的标量接收信号强度指示测量值造成了严重的信息瓶颈，而电磁传播物理要求采用特定领域的先验知识。近期观点强调，紧密的PCE耦合是具身系统的标志 [25]。针对这些不足，我们提出一种EI框架，通过语义桥，在自由能原理启发的优化视角下，以闭环PCE耦合统筹频谱构建与主动探索。

III. 基于EI的无人机频谱地图构建框架

本节介绍基于具身智能（EI）的无人机频谱构建框架。随后，我们给出数学公式，用于联合PCE优化。

A. 系统模型

如图1所示，考虑大规模无线网络中的无人机频谱地图构建系统。无人机作为具身智能体，在三维监测区域 $\Omega \subset \mathbb{R}^3$ 内运行，以构建能够捕获多个频带上RSSI空间分布的精确频谱图。为此，将连续空间离散化为规则网格结构 $\mathcal{G} = \{(i,\ j,\ k) \mid 0 \le i < N_x,\ 0 \le j < N_y,\ 0 \le k < N_z\}$，分辨率为 $\Delta_x,\ \Delta_y,\ \Delta_z$，其中 $|\mathcal{G}| = N_x N_y N_z$。在本研究中，无人机以固定高度进行巡测，因此评估网格缩减为相应的水平切片，并在所有监测频段上对该切片进行构建。因此，真实频谱地图表示为完整的多频场$\mathbf{C} \in \mathbb{R}^{|\mathcal{G}| \times M}$，其中 M 为监测频段的数量。目标是根据稀疏测量值精确构建完整的稠密频谱图$\mathbf{C}$，通过最小化$\mathbf{C}$ 与$\mathbf{C}$之间在适当距离量度下的构建误差来实现。由于严格的SWaP约束严重限制了飞行续航能力，无人机只能采集极其稀疏的测量值，总采样预算上限为 $N_{budget} \ll |\mathcal{G}|$。

频谱场值由无线传播物理学决定。具体而言，图$\mathbf{C}$的每个条目 $C_{i,f}$ 表示网格单元 $i \in \mathcal{G}$ 和频率 $f \in \mathcal{F} = \{f_1, \ldots, f_M\}$处的接收信号强度指示。该频谱场由一个或多个静态发射机生成，其中位置$\mathbf{p}$和频率 f 处的接收信号强度指示遵循传播模型

$$P_{rx}(\mathbf{p}, f) = P_{tx}(f) - PL(d, f) - X_\sigma(\mathbf{p}) - L_{obs}(\mathbf{p}, f), \quad (1)$$

其中， $P_{tx}(f)$为发射机功率， $PL(d, f)$表示距离发射机 d 处的路径损耗， $X_\sigma(\mathbf{p}) \sim \mathcal{N}(0, \sigma^2)$对阴影效应进行建模，而 $L_{obs}(\mathbf{p}\ f)$则计及障碍物衰减。

B. 框架构建

如图1所示，所提出的框架遵循闭环PCE循环，通过三个紧密耦合的组件实现EI范式。该循环从时间步t 的环境感知模块开始，该模块将多模态传感器测量聚合为稀疏测量数据集 $\mathbf{Y}_t$ 和障碍物地图 $\mathbf{O}$。随后，情境认知模块利用神经网络 $\mathcal{N}_\psi$ 处理这些输入，并嵌入物理知识以构建稠密频谱图ˆ估计 $\mathbf{C}_t$，作为表示估计频谱分布的主要输出。与此同时，该认知模块从编码器中提取能够捕捉空间结构和采样覆盖范围的内部网络表示，生成语义嵌入 $\mathbf{z}_{sem}^t \in \mathbb{R}^{d_{sem}}$，用于编码全局构建状态。随后，这些认知输出被馈送至主动探索模块，该模块根据当前位置、语义嵌入以及从 $\mathbf{C}_t$中提取的局部地图观测值，确定无人机的下一个航路点。执行该导航指令会触发新的

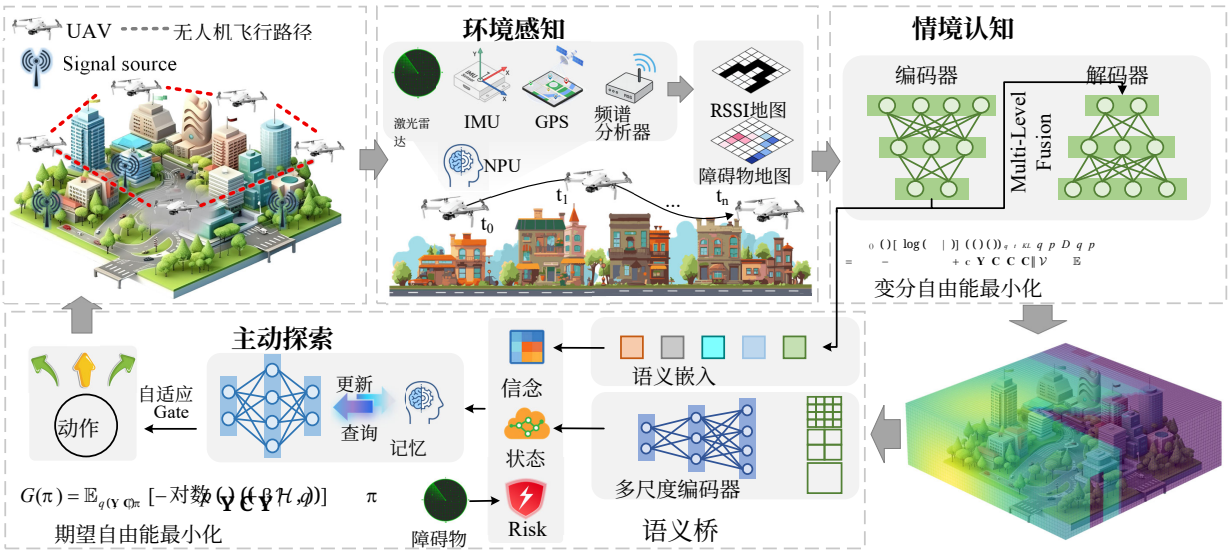

图1。所提出的　　用于频谱图构建的无人机感知、认知与主动探索闭环框架　　n.

感知过程，在新位置收集RSSI测量值 $\mathbf{m}^{t+1}$ ，并将数据集更新为 $\mathbf{Y}^{t+1}$，从而闭合循环。构建与导航之间的这种双向信息流被形式化为

$$\mathbf{Y}_t \xrightarrow{\mathcal{N}_\psi} (\hat{\mathbf{C}}_t, \mathbf{z}_{sem}^t) \xrightarrow{\pi_\theta} a_t \xrightarrow{\mathcal{S}} \mathbf{Y}_{t+1}, \tag{2}$$

其中， $\mathcal{N}_\psi$ 表示同时输出频谱图和语义特征的认知网络， π_θ 表示决策策略，该策略基于全局语义嵌入和局部地图观测值，从离散动作空间中选择导航动作 a_t ； $\mathcal{S}$ 表示感知过程，该过程依据公式（1）中的传播模型，将动作 a_t 映射为新的测量值 $\mathbf{m}^{t+1}$ 。

嵌套的观测历史赋予该闭环信息状态性质。由于 $\mathbf{Y}_{t+1} = \mathbf{Y}_t \cup \{(\mathbf{p}^{t+1}, \mathbf{m}^{t+1})\}$， σ-代数满足 $\sigma(\mathbf{Y}_t) \subseteq \sigma(\mathbf{Y}_{t+1})$，因此，当这些熵有限时，场的条件熵平均而言是非递增的，

$$h(\mathbf{C} \mid \mathbf{Y}_{t+1}, \mathbf{O}) \le h(\mathbf{C} \mid \mathbf{Y}_t, \mathbf{O}), \tag{3}$$

其中，该差值等于每步的信息增益 $I(\mathbf{C}; \mathbf{m}_{t+1} \mid \mathbf{Y}_t, \mathbf{p}^{t+1}, \mathbf{O}) \ge 0$。因此，策略能够控制该增益的幅度，但不能改变其符号。此外，解耦导航状态 $(\mathbf{p}_t, N_t)$ 是增广 PCE 状态 $(\mathbf{p}_t, \theta_t, \mathbf{z}_{sem}^t, \hat{\mathbf{C}}_t^{local}, N_t)$ 的可测函数，因此，在前者上定义的每个策略在后者上仍然是可容许的，并且在增广状态下能够达到的最优价值不低于前者。平台和仿真器提供相应的接收信号强度指示、障碍物几何结构和位置变量。在平台上，这些数据分别来自频谱分析仪、激光雷达和GPS/IMU；仿真器则通过其场景生成器提供这些数据。

1) 环境感知：如图1所示，PCE循环始于环境感知，它融合机载传感器的测量值，为后续认知提供两个关键输出。第一个输出是稀疏RSSI图，以累积测量数据集 $\mathbf{Y}_t$ 表示；第二个输出是障碍物地图 $\mathbf{O}$，用于定义可导航空间和传播屏障。具体而言，激光雷达扫描环境以构建障碍物地图 $\mathbf{O}$ ： $\mathcal{G} \to \{0, 1\}$，其中 $\mathbf{O}(\mathbf{p}) = 1$ 表示位置 $\mathbf{p}$ 处存在障碍物，而 $\mathbf{O}(\mathbf{p}) = 0$ 表示自由空间。该二值地图具有双重作用：既能识别不可通行区域以实现碰撞规避，也能为频谱构建提供结构先验，因为障碍物会造成信号遮蔽和衍射，从而影响RSSI的空间分布 [18]，[28]。此外，包括惯性测量单元和GPS在内的导航传感器可提供精准定位，从而在每个时间步 t 确定无人机的位置 $\mathbf{p}_t$ 和航向 θ_t 。

频谱分析仪是捕获多频RSSI测量值的关键感知组件。在位置 $\mathbf{p}$ $[m_t^{(1)}, \ldots, m_t^{(M)}]^T \in \mathbb{R}^M$ ，分析仪会观测监测频率集 $\mathcal{F}$ 上的RSSI向量 $\mathbf{m}$。观测测量值 $\mathbf{m}_t$ 与真实频谱地图 $\mathbf{C}$ 之间的关系可形式化表示为

$$\mathbf{m}_t = \mathbf{C}_{\mathcal{I}(\mathbf{p}_t)} + \mathbf{n}_t, \tag{4}$$

其中， $\mathcal{I}(\mathbf{p}_t)$ 通过最近邻分配将连续位置 $\mathbf{p}_t$ 映射为离散网格索引 $i \in \mathcal{G}$ ， $\mathbf{C}_i = [C_{i,f_1}, \ldots, C_{i,f_M}]^T$ 表示网格单元 i 处的真实频谱向量，而 $\mathbf{n}_t = [n_t^{(1)}, \ldots, n_t^{(M)}]^T \sim \mathcal{N}(\mathbf{0}, \sigma_n^2 \mathbf{I}_M)$ 表示由硬件缺陷和环境干扰引起的测量噪声向量。观测元组 $(\mathbf{p}_t, \mathbf{m}_t)$ 随时间累积，形成稀疏数据集 $\mathbf{Y}_t = \mathbf{Y}_{t-1} \cup \{(\mathbf{p}_t, \mathbf{m}_t)\}$。由于飞行续航限制严格，整个任务期间的数据集都极其稀疏， $N_t = |\mathbf{Y}_t| \ll |\mathcal{G}|$ 。稀疏RSSI地图仅能为地图构建提供有限的测量值，而障碍物地图则具有双重作用：既能实现无碰撞导航，也能通过明确的障碍物几何结构，为物理信息构建提供结构先验，以支持对衰减和阴影效应的学习表征。

2）情境认知：情境认知模块接收来自感知模块的稀疏测量数据集$\mathbf{Y}_t$和障碍物地图$\mathbf{O}$，旨在根据有限的测量值构建覆盖所有空间网格单元和频带的完整稠密频谱图$\hat{\mathbf{C}} \in \mathbb{R}^{|\mathcal{G}|\times M}$。由于测量值数量 $N_t = |\mathbf{Y}_t|$ 远小于未知场值的数量 $|\mathcal{G}| \times M$，这一构建任务本质上是病态的，因为有无穷多个解都能拟合稀疏数据。

为解决这一病态问题，地图构建被表述为一个贝叶斯推断问题，旨在求得后验分布 $p(\mathbf{C}|\mathbf{Y}_t)$，该分布表示在给定测量值 [9], [21] 的情况下对真实场的信念。然而，由于计算证据$p(\mathbf{Y}_t) = \int p(\mathbf{Y}_t|\mathbf{C})p(\mathbf{C})d\mathbf{C}$ 需要进行高维积分，计算该后验分布在计算上难以实现。为克服这一难题，我们采用自由能原理（FEP），从贝叶斯视角利用近似后验 $q(\mathbf{C})$ [29], [30] 来构建频谱图。变分自由能（VFE）定义为

$$\mathcal{V}(q) = \mathbb{E}_{q(\mathbf{C})}[-\log p(\mathbf{Y}_t|\mathbf{C})] + D_{KL}(q(\mathbf{C})\|p(\mathbf{C})), \quad (5)$$

其中，$\mathbb{E}_{q(\mathbf{C})}$ 表示对近似后验的期望，$p(\mathbf{Y}_t|\mathbf{C})$ 是描述测量值如何通过式（4）由真实场产生的似然函数，而$p(\mathbf{C})$ 是编码关于合理频谱图的领域知识的先验分布。最小化 $\mathcal{V}(q)$ 等价于最小化近似后验与真实后验 $D_{KL}(q(\mathbf{C})\|p(\mathbf{C}|\mathbf{Y}_t))$ [31] 之间的 KL散度。

网络 $\mathcal{N}_\psi$ 生成单个点估计 $\hat{\mathbf{C}}_t = \mathcal{N}_\psi(\mathbf{Y}_t, \mathbf{O})$。数据保真度项和先验项促成了以下实用的重建目标，其形式化表示为

$$\hat{\mathbf{C}}_t^* = \arg\min_{\hat{\mathbf{C}}} \mathcal{L}_{data}(\hat{\mathbf{C}}, \mathbf{Y}_t) + \lambda \mathcal{L}_{phys}(\hat{\mathbf{C}}), \quad (6)$$

其中，数据保真度损失$\mathcal{L}_{data}(\hat{\mathbf{C}}, \mathbf{Y}) = \sum_{\mathbf{p}_i, \mathbf{m}_i \in \mathbf{Y}_t} \| \mathcal{I}(\mathbf{p}_i, \mathbf{m}_i) - \hat{\mathbf{C}}(\mathbf{p}_i, \mathbf{m}_i) \|^2$ 对应于负对数似然项，用于衡量构建的场如何解释观测到的测量值；物理学正则化 $\mathcal{L}_{phys}(\mathbf{C})$则通过惩罚违反物理信息先验的构建结果，近似表示负对数先验。此外，加权参数 λ 用于平衡数据保真度与物理信息先验。

引入物理信息先验项的动机在于，纯粹依赖$\mathcal{L}_{data}$ 的数据驱动方法在极端稀疏性条件下可能表现不佳。当 $N_t \ll |\mathcal{G}|\times M$时，数据保真项提供的约束不足，网络可能产生无穷多个能够拟合稀疏测量值、却违背物理信息先验的解，从而产生突变不连续或错误传播模式等伪影 [32]。为解决这一局限，先验 $p(\mathbf{C})$ 根据公式（1）编码无线传播物理，包括距离相关功率衰减、频率选择性衰落以及障碍物阴影效应。该先验通过编码器架构实现，其中频率相关、距离相关和空间相关操作的有序组合，结合显式障碍物和采样支持输入，使估计结果偏向物理上合理的场。因此，$\mathcal{L}_{phys}$ 并未作为独立的残差项实现，而是由这种架构偏置与训练过程中采用的结构正则化共同提供其替代形式。

该模块的主要输出是构建的频谱图 $\hat{\mathbf{C}}_t \in \mathbb{R}^{|\mathcal{G}|\times M}$，它提供频谱图的完整空间频率表征，覆盖所有网格单元和频带。除这一主要重建输出外，该模块还会同时从内部编码器表示中提取语义嵌入 $\mathbf{z}_{sem}^t \in \mathbb{R}^{d_{sem}}$，其中编码了全局空间模式、测量值覆盖分布以及潜在的信息丰富区域。该语义特征通过全局空间模式和采样覆盖范围总结构建状态，为后续测量选择提供上下文信息。频谱图 $\mathbf{C}_t$ 和语义嵌入 $\mathbf{z}_{sem}^t$ 会一同传递至主动探索模块，以指导智能采样决策，从而闭合 PCE循环。因此，构建状态能够为后续测量选择提供依据，使策略在可用预算内根据不断演化的地图估计和空间覆盖率调整采样决策。

3) 主动探索：主动探索模块利用情境认知模块的认知输出，引导无人机导航以完成频谱地图构建，从而闭合PCE循环。如图1所示，该模块接收语义嵌入$\mathbf{z}_{sem}^t$和构建的频谱图$\hat{\mathbf{C}}_t$，以确定下一步动作，采集信息丰富的测量数据，从而提升构建质量。其核心挑战是在能量约束下，最大化频谱地图构建的信息增益。无人机必须采集能够降低构建不确定性的测量值，但飞行时间受到电池容量的限制。低效的轨迹会在收集到足够测量值之前耗尽有限的能量预算，从而导致较差的构建质量 [3]。因此，策略必须在信息丰富的采样与导航成本之间取得平衡，以便在有限的动作预算内支持频谱地图构建。

为形式化这种平衡，我们采用主动推断中的预期自由能（EFE） [31], [33]。一种受EFE启发、用于策略设计的参考目标函数可写为

$$G(\pi) = \mathbb{E}_{q(\mathbf{Y},\mathbf{C}|\pi)}[-\log p(\mathbf{Y}) + \beta \mathcal{H}(q(\mathbf{C}|\mathbf{Y}, \pi))], \quad (7)$$

其中，$q(\mathbf{Y}, \mathbf{C}|\pi)$ 表示在策略 π 下未来测量值和频谱图的预测分布；$p(\mathbf{Y})$ 对策略诱导的结果编码先验偏好，这些偏好促使实际奖励中产生导航相关反馈；$\mathcal{H}(q(\mathbf{C}|\mathbf{Y}, \pi))$ 量化所构建频谱图的不确定性；β 则平衡这两项。第一项体现与导航任务相关的偏好结果。第二项倾向于选择能够降低频谱图预期不确定性的观测值。该目标函数指导策略设计，使其在构建质量与导航成本之间取得平衡。

然而，直接最小化公式（7）在计算上难以处理，因为这需要对所有可能的未来轨迹和测量值进行积分。因此，将主动探索建模为一个通过深度强化学习（DRL）学习并由单步准则指导的马尔可夫决策过程（MDP），其表达式为

$$G_t(a_t) \approx -\log p(a_t) + \beta\, \mathbb{E}_{q(\mathbf{m}_{t+1}|\mathbf{Y}_t,a_t)}\left[\mathcal{H}\left(q(\mathbf{C} \mid \mathbf{Y}_{t+1}, a_t)\right)\right]. \tag{8}$$

第一项表示动作成本，第二项则对下一次测量后剩余的不确定性取平均。

策略使用一种增强表示，将导航信息与构建上下文相结合。元组$\mathbf{s}_t^a = (\mathbf{p}_t,\ \theta_t,\ \mathbf{z}_{sem}^t,\ \hat{\mathbf{C}}_t^{local},\ N_t)$包含位置$\mathbf{p}_t$、航向$\theta_t$、语义特征$\mathbf{z}_{sem}^t$、局部地图观测$\hat{\mathbf{C}}_t^{local}$以及测量次数 N_t。策略利用这种表示来选择导航动作。

构建误差为式（8）中的信息寻求目标提供了实用的训练信号。局部地图观测值和语义特征为策略提供构建上下文。在第二阶段优化期间，奖励反映构建精度和导航进展的变化。令E_t 表示全地图均方根误差， D_t 表示来自训练环境的导航距离信号。奖励函数可表示为

$$r_t = w_\Delta(E_{t-1} - E_t) - w_E E_t + w_D(D_{t-1} - D_t) + r_{aux}, \tag{9}$$

其中， w_Δ、 w_E和 w_D 分别对误差降低、当前构建误差和导航进展赋予权重。辅助项 r_{aux} 包括每步时间惩罚、碰撞惩罚，以及对成功或未成功终止的反馈。这种奖励结构在地图构建改进与导航进展之间取得平衡，同时纳入时间、碰撞和终止结果的惩罚。

我们注意到，只有在基于仿真的策略训练期间才需要真实地图$\mathbf{C}$来评估公式（9）；而在真实世界部署中，训练好的策略直接根据观测状态$\mathbf{s}_t^a$ 选择动作，无需访问$\mathbf{C}$。

策略 π_θ 通过与环境交互进行训练，以最大化期望累积奖励。训练期间对动作进行采样以实现探索，而在部署期间采用贪心方式选择动作。沿轨迹持续收集测量值，用于频谱地图构建。执行动作 a_t 会触发感知模块更新数据集$\mathbf{Y}^{t+1}$，该数据集随后输入认知模块，以生成更新后的地图$\hat{\mathbf{C}}^{t+1}$ 和特征$\mathbf{z}_{sem}^{t+1}$，从而闭合PCE循环，并在极端稀疏性和能量约束下实现高效构建。

式（8）表达了FEP框架内单步过程中期望信息增益与动作成本之间的平衡。在式（9）中，地图误差降低表示观测价值，而导航、碰撞、时间和终止项则体现了移动与安全偏好。结合式（6）中的认知目标，这些特定于任务的目标函数将数据保真度、先验知识和面向构建的探索联系起来，并纳入FEP框架。VFE和EFE表达式为其设计与解释提供指导，而无需在训练过程中显式评估后验熵。

C. 优化目标

该框架优化处理的是频谱地图构建问题，通过协同优化无人机轨迹$\boldsymbol{\tau} = \{\mathbf{p}_0, \mathbf{p}_1, \ldots, \mathbf{p}_T\}$和频谱地图估计 $\hat{\mathbf{C}}$，在满足稀疏采样约束的同时最小化构建误差。该联合优化形式化为

$$\min_{\boldsymbol{\tau},\ \psi,\ \theta} \quad \mathbb{E}_{\mathbf{C}\sim p_{env}}[\|\mathcal{N}_\psi(\mathbf{Y}(\boldsymbol{\tau}), \mathbf{O}) - \mathbf{C}\|^2] \tag{10a}$$

$$\text{s.t.} \quad \boldsymbol{\tau} = \{\mathbf{p}_t\}_{t=0}^T, \quad , \tag{10b}$$

$$\mathbf{p}_{t+1} = \mathbf{p}_t + \Delta\mathbf{p}(a_t, \theta_t), \quad 0 \le t < T$$

$$a_t = \pi_\theta(\mathbf{s}_t^a), \quad \mathbf{s}_t^a = (\mathbf{p}_t, \theta_t, \mathbf{z}_{sem}^t, \hat{\mathbf{C}}_t^{local}, N_t), \tag{10c}$$

$$N_T \le N_{budget}, \quad \mathbf{p}_t \notin \mathbf{O} \quad \forall t, \tag{10d}$$

其中，主要目标函数公式（10a）最小化了从环境分布 p_{env}中采样的环境实例$\mathbf{C}$ 上的构建误差，测量数据集$\mathbf{Y}(\boldsymbol{\tau}) = \{(\mathbf{p}_t、\ \mathbf{m}_t)\}_{t=0}^T$显式依赖于轨迹$\boldsymbol{\tau}$。该表述捕捉了双向耦合关系：轨迹决定测量位置，进而定义数据集；数据集又决定地图质量和提取的认知特征，并通过增广状态$\mathbf{s}_t^a$指导未来的轨迹决策。

公式（10b）和公式（10c）中的约束形式化描述了轨迹如何通过策略依作用于包含构建网络认知特征的增广状态上所选择的连续动作生成。该策略利用公式（9）中的构建项和导航项最大化期望累积奖励。公式（10d）中的实际约束确保满足采样预算和碰撞避免要求。该表述在任务层面将轨迹规划与地图构建相耦合。网络按照算法1所述依次进行训练。

IV. 我们提出的 ECDNet

本节介绍图2所示 ECDNet 的详细架构设计，该网络通过三个紧密耦合的模块实例化了所提出的框架：环境感知、物理信息驱动的态势认知和主动探索。

A. 环境感知模块

环境感知模块作为 ECDNet 与物理世界之间的接口，将稀疏测量数据集 $\mathbf{Y}_t$ 和障碍物地图 $\mathbf{O}$ 转换为用于神经网络处理的固定大小张量表示。如图2所示，该模块构建了一个多通道输入张量 $\mathbf{X}^{in}$，封装认知网络所需的全部信息。

输入张量 $\mathbf{X}^{in}$ 由四个互补分量构成。稀疏采样地图 $\mathbf{I}^{sparse}$ 携带监测频段在其网格位置处测得的接收信号强度指示，并将未测量的单元填充为零。插值接收信号强度通道 $\mathbf{I}^{interp}$ 为相同频段提供稠密初步场。二值采样掩码 $\mathbf{M}_{sam}$ 区分已观测单元与填零的未观测单元，并记录测量支持。障碍物地图$\mathbf{O}$ 提供几何信息

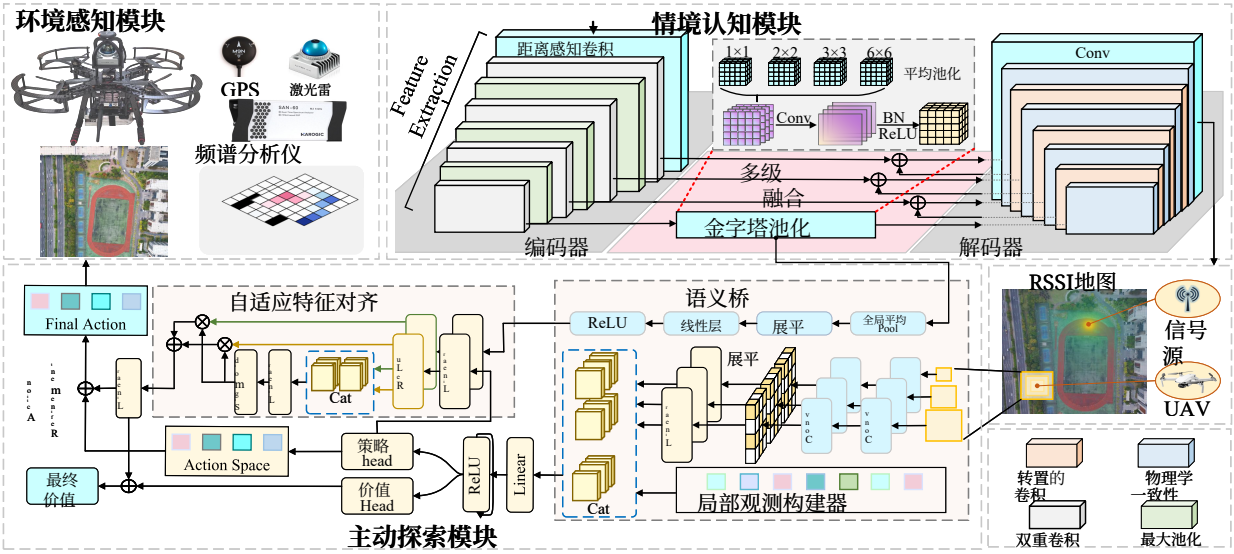

图2. E的架构　　CDNet由环境感知、物理信息驱动的态势认知和主动探索模块组成。

用于传播感知构建和运动可行性。完整输入为

$$\mathbf{X}^{in} = \mathrm{Concat}(\mathbf{I}_{sparse}, \mathbf{I}_{interp}, \mathbf{M}_{sam}, \mathbf{O}), \tag{11}$$

其中，稀疏信道和插值信道提供测量证据和临时空间上下文，采样掩码保留观测来源，障碍物地图则提供明确的环境几何结构。

B. 情境认知模块

情境认知模块实现认知网络 $\mathcal{N}_\psi$，该网络将$\mathbf{X}^{in}$ 转换为稠密频谱图估计值$\mathbf{C}_t$。其物理信息先验分布于整个编码器中。编码器实现了物理信息驱动的变换

$$\Phi_{phys} = \Phi_{spatial} \circ \Phi_{dist} \circ \Phi_{freq}, \tag{12}$$

其中，Φ_{dist} 执行输入条件化空间加权，Φ_{freq}对特定频率响应进行建模，$\Phi_{spatial}$ 聚合空间上下文。如图3所示，距离感知卷积结合输入条件化空间权重和学习得到的通道偏置，以及特定频率子网络和注意力机制。距离感知层不接收显式的发射机距离坐标，也不施加解析形式的路径损耗定律。障碍物几何结构和测量支持通过$\mathbf{O}$ 和$\mathbf{M}_{sam}$显式输入，而障碍物引起的阴影效应则根据由式（1）所启发的受控训练分布进行学习。由此得到的架构形成了偏向物理上合理的频谱图的归纳偏置。

如图2所示，完整的前向传播分为三个阶段。编码器生成特征图和跳跃连接张量。随后，金字塔池化模块从编码器特征中聚合多尺度上下文。最后，解码器将池化上下文与跳跃连接融合，以构建图。该过程形式化表示为

$$\mathbf{Z}_{enc}, \mathbf{Z}_{skip} = \mathcal{E}_\psi(\mathbf{X}^{in}), \tag{13a}$$
$$\mathbf{Z}_{pool} = \mathcal{P}_{pool}(\mathbf{Z}_{enc}), \tag{13b}$$
$$\hat{\mathbf{C}}_t = \mathcal{D}_\psi(\mathbf{Z}_{pool}, \mathbf{Z}_{skip}), \tag{13c}$$

其中，$\mathcal{E}_\psi$、$\mathcal{P}_{pool}$和 $\mathcal{D}_\psi$ 分别表示编码器、金字塔池化模块和解码器组件，$\mathbf{Z}_{enc}$ 表示编码器输出特征，$\mathbf{Z}_{pool}$ 表示池化上下文，$\mathbf{Z}_{skip}$ 表示多级跳跃连接张量。

网络采用受式 (6) 中重建目标启发的复合损失进行训练。该损失由三个互补项组成，用于数据拟合和结构正则化，其形式如下：

$$\mathcal{L}_{cog}(\psi) = \alpha \mathbb{E}_{\mathbf{p}\in\mathcal{G}}[\mathbf{M}_{sam}(\mathbf{p}) \cdot \|\hat{\mathbf{C}}(\mathbf{p}) - \mathbf{C}(\mathbf{p})\|^2] \tag{14a}$$
$$+ \mathbb{E}_{\mathbf{p}\in\mathcal{G}}[\|\hat{\mathbf{C}}(\mathbf{p}) - \mathbf{C}(\mathbf{p})\|^2] \tag{14b}$$
$$+ \lambda_{perc} \sum_l \|\Phi_l^{VGG}(\hat{\mathbf{C}}) - \Phi_l^{VGG}(\mathbf{C})\|_1, \tag{14c}$$

其中，ψ 表示可训练参数，α 是用于强调采样位置的大权重，$\mathbf{M}_{sam}$ 是式 (11) 中的二值采样掩码，λ_{perc} 控制结构正则化强度，Φ_l^{VGG}从预训练 VGG网络的第 l 层提取特征。式 (14a) 中的第一项通过为 $\mathbf{M}_{sam}$所指示的单元赋予权重 α，实现采样位置监督。在离线训练过程中，该项近似实现式 (6) 中的数据保真度作用，同时不强制与含噪声的场测量值完全一致。

第二项 eq. (14b) 在离线训练期间提供有监督的全局重建引导，第三项 eq. (14c) 则充当多尺度结构正则项，用于抑制空间不连贯伪影。$\mathcal{L}_{phys}$ 中的物理先验由网络中基于传播原理的特征处理来表示，而结构正则项则约束重建输出。这些架构和损失

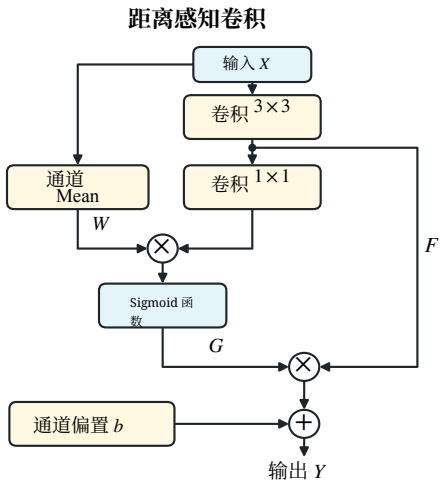

图3. 具有输入条件空间加权和学习得到的通道偏置的距离感知卷积层内部结构。

函数组件提供了式 (6) 中概念先验的替代实现。

除$\mathbf{C}_t$之外，该模块还通过金字塔池化和全局平均池化，从全局瓶颈表征中提取语义嵌入 $\mathbf{z}_{sem}^t = \Psi_{extract}(\mathbf{Z}_{enc})$。由于 $\mathbf{M}_{sam}$ 是显式输入，观测支持和覆盖率被编码在 $\mathbf{Z}_{enc}$ 中，并可传递给策略。局部裁剪图 $\hat{\mathbf{C}}_t^{local}$提供即时的邻域细节，而$\mathbf{z}_{sem}^t$ 则提供构建状态的互补全局摘要，包括覆盖率、大尺度频谱变化和障碍物相关结构。

C. 主动探索模块

主动探索模块实现了策略 π_θ，该策略将增广状态$\mathbf{s}_t^a$映射为用于频谱地图构建的导航动作 a_t 。在具体实现中，π_θ 分为两个阶段。如图2所示，在评估配置中，第一阶段基线 $\pi_{\theta_1}^{(1)}$ 采用前馈策略进行导航。第二阶段精炼器冻结 $\pi_{\theta_1^*}^{(1)}$ 和 $\mathcal{N}_{\psi^*}$，并学习对第一阶段逻辑值和价值估计的残差修正，同时将残差头权重初始化为较小值。

该模块架构包含两个互补的策略阶段，各自具有不同的输入范围。第一阶段基线策略侧重于根据物理观测和局部构建地图观测实现稳定导航。具体而言，它使用局部处理路径 Φ_{local} ，将物理位置$\mathbf{p}_t$、航向 θ_t、从构建的频谱图$\mathbf{C}$中当前位置周围提取的局部地图观测$\hat{\mathbf{C}}_t^{local}$、障碍物地图$\mathbf{O}$以及测量次数 N_t 编码为局部特征向量$\mathbf{z}_{local}$。该局部路径处理构建的RSSI场及其空间一致性模式，以表征局部构建状态。因此，第一阶段的输入为

$$\mathbf{z}_t^{(1)} = \Phi_{local}(\mathbf{p}_t, \theta_t, \hat{\mathbf{C}}_t^{local}, \mathbf{O}, N_t), \tag{15}$$

从而为基线导航策略提供局部观测，而无需直接使用全局语义嵌入。该

算法1 顺序训练流程 1：输入：离线数据集 $\mathcal{D}_{train}$、并行环境$\{\mathrm{Env}_k\}_{k=1}^{K}$ 2：输出：冻结网络 $\mathcal{N}_{\psi^*}$、 $\pi_{\theta_1^*}^{(1)}$ ，精炼策略π_{refine}3：// 认知网络训练 4：对于 $epoch = 1$，...，N_{cog} 执行 5：从 $\mathcal{D}_{train}$ 中采样小批量（Y O C）：使用公式（13）计算 $\hat{\mathbf{C}} \leftarrow \mathcal{N}_\psi(\mathbf{Y}, \mathbf{O})$ 7：使用公式（14）计算 $\psi \leftarrow \psi - \nabla_\psi \mathcal{L}_{cog}(\psi)$ 8：结束循环 9：保存 ψ^* 并冻结 10：// 基线策略训练 11：初始化 $\pi_{\theta_1}^{(1)}$ 和 V_{ϕ_1}12：对于 $iter = 1$，.....，N_{nav} 执行 13：对于每个环境 $k \in \{1, \ldots, K\}$ 执行 14：使用 $\pi_{\theta_1}^{(1)}$ 和 $\mathcal{N}_{\psi^*}$ 15：收集轨迹 $r_t^{(1)}$：根据地图误差改进、距离、安全性和辅助项计算 16：结束循环 17：根据所有轨迹计算优势函数值 $\{A_t\}$ 18：使用 PPO 更新（θ_1，ϕ_1）$\leftarrow$ 19：结束循环 20：保存 θ_1^* 并冻结 21：// 策略网络精炼 22：使用 $\pi_{\theta_1^*}^{(1)}$ 和 $\mathcal{N}_{\psi^*}$ 23：在冻结的 $iter = 1$ ... 上初始化细化器头：对于，，N_{refine} 执行 24：对于每个环境 $k \in \{1, \ldots, K\}$ 执行 25：使用π_{refine}收集轨迹 26：对于轨迹中的每个状态执行 27：从冻结的 $\mathcal{N}_{\psi^*}$ 中提取 $\hat{\mathbf{C}}$、z 和 $\hat{\mathbf{C}}$：s（p、，z、$\hat{\mathbf{C}}$、） 28：$a_t \leftarrow \theta_t(\mathbf{s}_t, \mathbf{z}_{sem}^t, \hat{\mathbf{C}}_t^{local}, N_t)$ 29：结束循环 30：根据公式（9）计算 r_t 31：结束循环 32：根据所有轨迹计算优势函数值 $\{A_t\}$ 33：通过 PPO 更新 θ_{refine} 34：结束循环 35：返回 $\mathcal{N}_{\psi^*}$、 $\pi_{\theta_1^*}^{(1)}$ 、π_{refine}

第二阶段精炼器随后引入语义桥。全局处理通路 Φ_{global} 将认知衍生语义嵌入 $\mathbf{z}_{sem}^t$ 编码为全局特征向量，用于概括与构造相关的空间模式和全局观测上下文。该全局特征通过下文所述的门控残差机制与冻结的第一阶段逻辑值结合，使全局语义嵌入仅在精炼阶段进入策略。

第二个组件是策略网络。第一阶段基线策略 $\pi_{\theta_1}^{(1)}$ 使用前馈层处理 $\mathbf{z}_t^{(1)}$ ，并输出基线动作逻辑值 $\mathbf{l}_t^{(1)}$ 和价值 $v_t^{(1)}$ 。第二阶段精炼器冻结了 $\pi_{\theta_1^*}^{(1)}$ 和 $\mathcal{N}_{\psi^*}$，并通过自适应混合冻结的策略表示与认知衍生表示来细化基线逻辑值：

$$(\mathbf{l}_t^{(1)}, v_t^{(1)}) = \pi_{\theta_1^*}^{(1)}(\mathbf{z}_t^{(1)}), \tag{16a}$$
$$\mathbf{z}_{sem}^t = \Psi_{extract}(\mathbf{Z}_{enc}), \tag{16b}$$
$$\mathbf{r}_t^{stage} = \Phi_{stage}\big(\mathrm{sg}(\mathbf{l}_t^{(1)})\big), \tag{16c}$$
$$\mathbf{r}_t^{enc} = \Phi_{enc}(\mathbf{z}_{sem}^t), \tag{16d}$$
$$\mathbf{g}_t = \sigma\big(\mathrm{MLP}_g([\mathbf{r}_t^{stage}, \mathbf{r}_t^{enc}])\big), \tag{16e}$$
$$\mathbf{u}_t = \mathbf{g}_t \odot \mathbf{r}_t^{stage} + (1 - \mathbf{g}_t) \odot \mathbf{r}_t^{enc}, \tag{16f}$$
$$\mathbf{l}_{final} = \mathrm{sg}(\mathbf{l}_t^{(1)}) + \Delta_\pi(\mathbf{u}_t), \tag{16g}$$
$$v_{final} = \mathrm{sg}(v_t^{(1)}) + V_\Delta(\mathbf{u}_t). \tag{16h}$$

其中，$\mathbf{Z}_{enc}$ 由冻结认知编码器提取，Φ_{stage} 和 Φ_{enc} 分别表示阶段分支和编码器分支，$\mathbf{g}_t$ 是一个由两种表示共同决定的逐元素门控。因此，$\mathbf{u}_t$ 会在生成残差策略校正之前，自适应地混合冻结策略嵌入与认知派生嵌入。停止梯度算子 sg(·)表示精炼过程中不会更新第一阶段策略。与策略逻辑值类似，价值估计以冻结的第一阶段价值头为基础进行残差校正。在所评估的配置中，不使用循环隐藏状态。两个阶段均使用 PPO 训练：第一阶段更新前馈策略和价值头，第二阶段仅更新细化器头。待最大化的 PPO 目标函数可写为

$$\mathcal{L}_{dec}(\theta, \phi) = \mathbb{E}_t[\mathcal{L}_{CLIP}(\theta) - c_1\mathcal{L}_{VF}(\phi) + c_2\mathcal{S}_\pi(\theta)], \tag{17}$$

其中，$\mathcal{L}_{CLIP}$、$\mathcal{L}_{VF}$和 $\mathcal{S}_\pi$ 分别表示裁剪替代目标、价值损失和熵奖励。

在实际应用中，局部地图观测值 $\hat{\mathbf{C}}_t^{local}$ 通过从完整构建地图$\hat{\mathbf{C}}_t$ 中裁剪以当前位置 $\hat{\mathbf{p}}_t$ 为中心的局部区域，并通过多尺度卷积层进行处理而获得。该局部观测值编码了构建的RSSI及其空间模式，为频谱地图构建过程中的动作选择提供学习到的邻域上下文。该模块完成了闭环框架。情境认知模块构建频谱地图并提取语义嵌入。主动探索模块利用这些嵌入以及局部地图观测值，根据当前构建状态和空间覆盖率选择动作。执行动作 a_t 会触发环境感知模块收集测量值$\mathbf{m}^{t+1}$ 并更新数据集，从而重新启动PCE循环；随着通过具身探索收集到更多测量值，构建地图也会随之更新。

D. 训练与部署策略

该框架采用三阶段训练策略，如算法1所总结。首先，通过最小化公式（14），在 $\mathcal{D}_{train}$ 上对认知网络 $\mathcal{N}_\psi$进行预训练，并冻结收敛后的参数 ψ^*。其次，使用PPO训练第一阶段基线策略 $\pi_{\theta_1}^{(1)}$，其基线奖励结合了初步的地图误差改进、距离效率、安全代价和辅助项。所得 θ_1^*被冻结，作为稳定的导航骨干网络。第三，在冻结的认知网络之上训练第二阶段细化器

算法 2 实时部署流程

要求： 网络（$\mathcal{N}_{\psi^*}$，π_{policy}）、障碍物地图$\mathbf{O}$、步长预获取 T_{budget}，配置的最小覆盖率为 ρ_{min}
确保： 构建地图$\mathbf{C}_{final}$，轨迹 $\boldsymbol{\tau}$

在初始位置 $\mathbf{p}_0$ 部署无人机，航向为 θ_0
$\mathbf{Y}_{-1} \leftarrow \emptyset$, $N_{-1} \leftarrow 0$, $t \leftarrow 0$, $\boldsymbol{\tau} \leftarrow (\mathbf{p}_0)$
$\mathcal{G}_{free} \leftarrow \{\mathbf{p} \in \mathcal{G} \mid \mathbf{O}(\mathbf{p}) = 0\}$
当 $t < T_{budget}$ **时执行**
  $\mathbf{m}_t \leftarrow$ 在 $\mathbf{p}_t$ 处测量接收信号强度指示
  $\mathbf{Y}_t \leftarrow \mathbf{Y}_{t-1} \cup \{(\mathbf{p}_t, \mathbf{m}_t)\}$，$N_t \leftarrow N_{t-1} + 1$
  $\hat{\mathbf{C}}_t \leftarrow \mathcal{N}_{\psi^*}(\mathbf{Y}_t, \mathbf{O})$ 使用公式（13）
  $\hat{\mathbf{C}}_{final} \leftarrow \hat{\mathbf{C}}_t$
  $\hat{\mathbf{C}}_t^{local} \leftarrow$来自 $\mathbf{p}_t$ 周围的局部区域 $\hat{\mathbf{C}}_t$
  如果 $\pi_{policy} = \pi_{refine}$ **则**
    $\mathbf{Z}_{enc}, \mathbf{Z}_{skip} \leftarrow$ 公式（13a）的输出
    $\mathbf{z}_{sem}^t \leftarrow \Psi_{extract}(\mathbf{Z}_{enc})$
  else
    $\mathbf{z}_{sem}^t \leftarrow \mathbf{0}_{d_{sem}}$
  结束如果
  $\mathbf{s}_t^a \leftarrow (\mathbf{p}_t, \theta_t, \mathbf{z}_{sem}^t, \hat{\mathbf{C}}_t^{local}, N_t)$
  如果 $\pi_{policy} = \pi_{refine}$ **则**
    $a_t \leftarrow \pi_{refine}(\mathbf{s}_t^a)$
  else
    $a_t \leftarrow \pi_{\theta_1^*}^{(1)}(\mathbf{s}_t^a)$
  结束 if
  使用 a_t将 $(\mathbf{p}_t, \theta_t)$ 更新为 $(\mathbf{p}^{t+1}, \theta_{t+1})$
  将$\mathbf{p}^{t+1}$ 追加到 $\boldsymbol{\tau}$
  如果 $|\{\mathbf{p} \mid (\mathbf{p}, \mathbf{m}) \in \mathbf{Y}_t\}| / |\mathcal{G}_{free}| \geq \rho_{min}$ **则**
    跳出
  结束 if
$t \leftarrow t + 1$
结束 while
返回 $\hat{\mathbf{C}}_{final}$，$\boldsymbol{\tau}$

以及冻结的基线策略，此时仅使用公式（9）中的完整奖励更新细化器头。该课程学习在固定基线参数的同时，学习具有构建感知能力的策略对数几率残差修正。

在算法 2 所述的部署过程中，框架在没有真实值的情况下运行，其中 $\mathcal{G}_{free}$ 表示可通行单元格。在每个时间步，无人机测量频谱 $\mathbf{m}_t$，使用冻结的 $\mathcal{N}_{\psi^*}$ 构建 $\hat{\mathbf{C}}_t$，并根据局部观测和语义特征构成 $\mathbf{s}_t^a$。随后，所选策略计算动作 a_t。轨迹展开在达到步数预算 T_{budget} 或设定的覆盖率阈值 ρ_{min}时终止。覆盖率按不同的已采样空闲单元格计算，重复访问不增加覆盖率。由于允许重复访问，步数预算为一次轨迹展开中收集的不同采样单元格数量提供了上限。

V. 实验结果

本节对所提出的 ECDNet 进行了全面评估。评估涵盖受控仿真，以及在嵌入式无人机平台上的原型级演示。为验证在严格硬件约束下自主构建频谱图的可行性，

表 I 仿真与部署平台规格。

仿真平台		部署平台	
操作系统	Ubuntu 22.04 LTS	硬件	Jetson Orin NX，16 GB
GPU	NVIDIA RTX 4090	GPU 核心	1024 CUDA
CPU	Intel i9-13900K	认知延迟	27.63 毫秒（p95 28.22）
RAM	64 GB DDR5	决策延迟	8.16 毫秒（p95 9.07）
框架	PyTorch 2.0.1	掩码与动作生成	1.54 毫秒（p95 1.70）
Python	3.10.12	模型循环延迟	37.33 毫秒（p95 为 38.48）
CUDA	12.1	计时范围	驻留于GPU的模型循环
批量大小	32	精度与批次	FP32，批次 1
学习率	1e-4 和 3e-4	功耗模式	25 W
优化器	Adam	控制周期	100 毫秒（10 赫兹）
		续航时间	18 分钟



图4. 所开发的无人机具身频谱构建平台的硬件架构，集成了感知与计算模块。

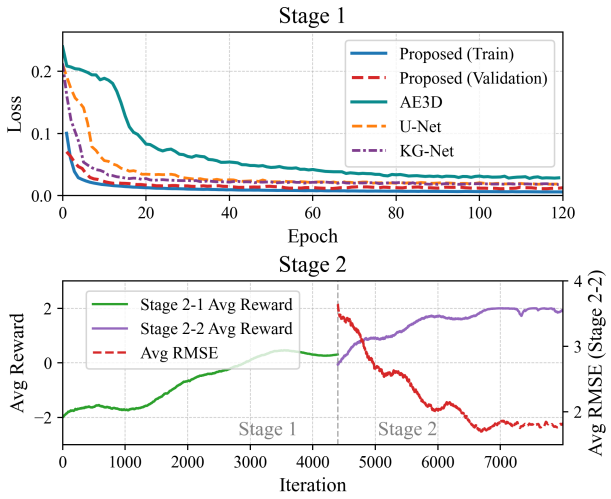

图5. 顺序训练策略下认知网络预训练与策略网络学习的收敛分析。下方图表中的两条奖励曲线分别标记为阶段2-1和阶段2-2，对应阶段1和阶段2策略训练；均方根误差曲线对应精炼策略的阶段2评估。

我们开发了图 4 所示的定制无人机平台。如图 4 所示，该系统在紧凑的机身内集成了 Livox Mid-360 激光雷达和 HAROGICSAN-60 M2 频谱分析仪。Jetson Orin NX 在机载处理数据。然而，在此类设备受尺寸、重量和功耗（SWaP）限制，受硬件限制影响，面临显著的计算挑战。

表 I 报告了详细的系统效率指标。Jetson 的延迟是在 GPU 驻留输入、 FP32 推理以及每个计时边界处执行 CUDA同步的条件下测得的。认知、决策和动作掩码阶段的平均耗时分别为 27.63、8.16 和 1.54 ms，完整模型循环仅需 37.33 ms（p95 为 38.48 ms），不到配置的 100 ms 控制周期的 40%。这些结果表明，在受 SWaP 约束的边缘设备上，ECDNet 的完整认知—决策链能够在单个控制周期内于设备端执行。

除上述硬件演示外，我们还在传播复杂度逐步增加的三种合成城市场景中开展了全面的仿真实验，以系统评估重建性能。真实频谱地图由校准传播模型生成，该模型结合了距离相关路径损耗、频率选择性衰减，以及通过发射机与接收机之间的视线栅格化计算得到的障碍物引起的遮挡损耗。尽管这一简化模型没有显式纳入反射和衍射核，但它能生成非平稳场结构，捕捉障碍物密集环境中的主要阴影衰落模式，并作为评估稀疏构建方法的受控基准。在所执行的配置中，无人机对一个离散为 64×64 网格的固定高度切片进行勘测，覆盖四个监测频段；认知输入是一个十通道张量，由四个稀疏 RSSI 信道、四个插值信道、采样掩码和障碍物地图组成；提取的语义嵌入维度为 1024。为评估性能，我们将 ECDNet 与四种代表性基线方法进行比较：Kriging [8]（经典地统计插值）、AE3D [16]（纯数据驱动自编码器）、U-Net [10]（语义辅助空间架构）和 KG-Net [17]（模型引导学习，通过采用金字塔空间池化的 U-Net 主干网络实现）。AE3D、U-Net 和 KG-Net 与 ECDNet 使用相同的 23,009 张地图划分进行训练，并根据验证性能进行选择；所有重建方法采用相同的输出网格、不重复访问的螺旋测量位置和均方根误差定义。

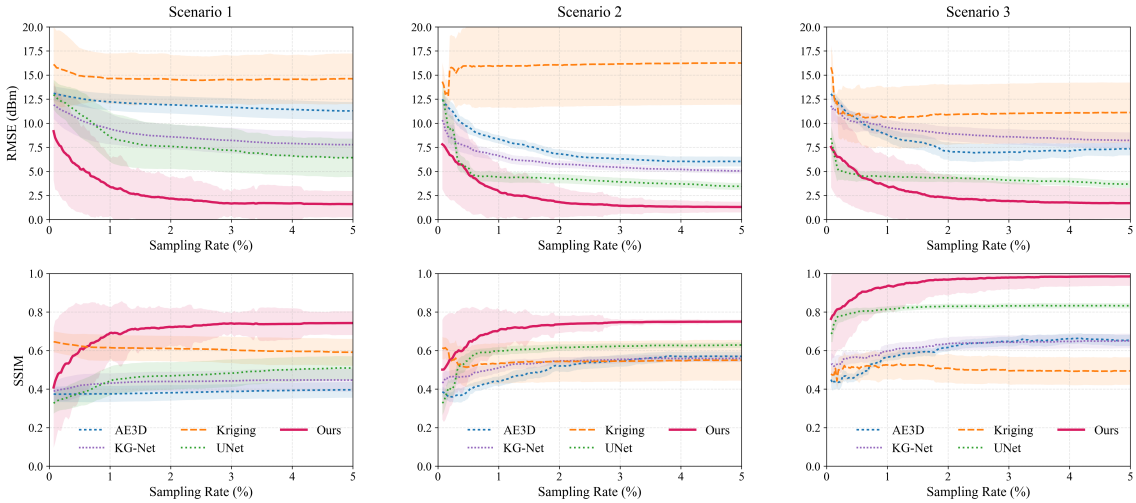

图6. 稀疏采样下 RMSE 和 SSIM 的定量比较。场景1、场景2和场景3分别包含 186、189 和 181 个留出的频谱图实例。在每个采样级别，曲线表示各场景中所有图的均值，阴影区域表示各图之间的 ±1 样本标准差，而非独立训练运行之间的标准差。绘制的横轴为连续采样率扫描，不局限于表II中的六个运行点。

图5展示了训练动态：上方面板显示认知预训练情况，下方面板显示两个策略训练阶段的情况。在认知预训练阶段，纯数据驱动方法通常难以在高维参数空间中找到最优解，并且往往收敛缓慢。相比之下，我们的方法通过引入物理信息变换 Φ_{phys}加快了收敛速度，损失在前 20 个训练轮次内迅速下降。这表明，距离感知卷积提供了归纳偏置，可在优化早期约束解空间。在策略优化中，第二阶段优化器通过 PPO 在 $K = 4$ 个并行环境中更新。这里的 K 表示用于轨迹收集的环境数量，而非训练集大小。策略训练池包含来自六个场景族的 540 个场景实例；第二阶段优化使用无障碍单发射机族（single-free）的 90 个实例，同时冻结第一阶段策略和认知编码器。第二阶段奖励稳步上升，评估均方根误差持续下降，表明构建感知残差精炼有效改善了探索策略。为确保公平比较，所有后续测试时评估均严格在独立的留出实例上进行。

不同采样率下的频谱构建性能如图6和表II所示。由图6可见，当采样率低于2%时，在极端稀疏条件下，频谱构建性能显著下降。在这些情况下，由于缺乏足够的样本，克里金法和AE3D无法捕捉阴影效应的非平稳变化，因此均方根误差较高。相比之下，当采样率超过约0.6%后，ECDNet在三种场景下均取得最低的均方根误差；当采样率达到约1%时，其结构相似性指数超过所有基线方法。此外，随着测量值不断累积，ECDNet的均方根误差持续降低，而基线曲线下降缓慢，或在较高的误差水平趋于平稳，表明在稀疏采样条件下，ECDNet具有更高的样本效率。

表II进一步量化了离散运行点下的构建精度，其中采样率为 $n/4096$网格上的 64×64，而 n则统计不同的采样单元格。如表 II 所示，ECDNet 在所有六个运行点上均取得最低的 RMSE 和最高的 SSIM。具体而言，在 4.00% 的采样率下，ECDNet 的 RMSE 为 2.787 dBm，SSIM 为 0.822；与表现最优的基线 U-Net 相比，RMSE 降低了 47.8%，SSIM 提高了 0.149。此外，ECDNet 仅使用 $n = 41$ 个采样单元（1.00%），便取得了 4.042 dBm 的 RMSE，仍低于 U-Net 使用 $n = 164$ 个采样单元（4.00%）时的 5.335 dBm。这一对运行点对应的采样位置减少了 75%。这些结果表明，ECDNet 在极端稀疏采样下具有更高的样本效率。

图7展示了采样率从0.24%到4.00%时具有代表性的频谱图估计。如图7的顶行所示，包括U-Net和克里金法在内的基线方法，由于稀疏重建的病态性，往往会产生过度平滑的结果或块状伪影。例如，采样率为0.5%时，AE3D模型无法捕捉清晰的信号传播方向，而采样率为4%时，克里金法在采样点周围呈现出典型的平滑伪影。相比之下，如中间行所示，ECDNet即使在较低采样率下也能恢复主要信号源和障碍物诱导的阴影边界，其重建结果与真实值高度一致。不同于螺旋采样基线方法，图7中的ECDNet图沿着叠加显示的学习策略轨迹进行重建，这些轨迹会主动经过阴影区域和信号源附近，因此同时体现了传播感知编码器和重建感知采样，而表II比较了

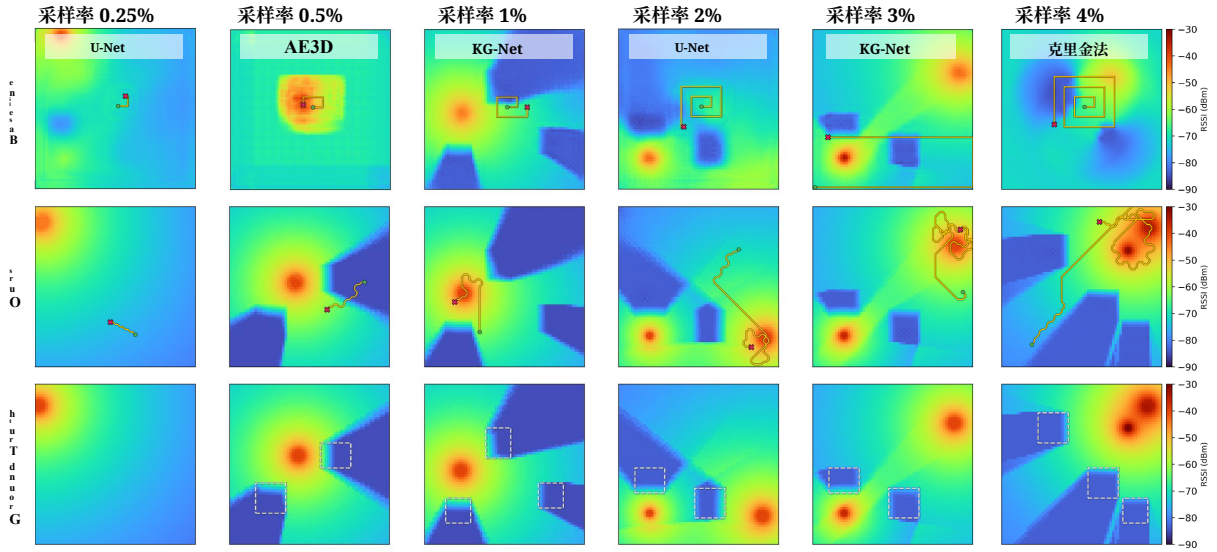

图7。不同稀疏度级别下各方法在频谱地图构建质量和采样轨迹方面的可视化比较。面板标题为标称采样率，其中 0.25% 和 4% 面板分别对应表II中的 0.24% 和 4.00%。基线图使用非重复螺旋采样点，而我们的方法图使用沿叠加的学习策略轨迹分布的采样点。

表 II 非重复螺旋采样模式下的定量性能比较。

方法	0.24%（n=10）		1.00%（n=41）		2.00%（n=82）		3.00%（n=123）		4.00%（n=164）		4.88% (n=200)	
	RMSE	SSIM	RMSE	SSIM	RMSE	SSIM	RMSE	SSIM	RMSE	SSIM	RMSE	SSIM
AE3D [16]	12.058	0.454	11.670	0.454	11.296	0.457	10.896	0.460	10.525	0.475	10.093	0.482
克里金法 [8]	13.096	0.605	12.468	0.585	12.048	0.565	11.827	0.552	12.117	0.529	12.164	0.512
U-Net [10]	11.005	0.483	8.078	0.589	6.335	0.646	5.781	0.665	5.335	0.673	4.924	0.684
KG-Net [17]	11.428	0.508	10.118	0.547	9.114	0.570	8.527	0.582	8.098	0.588	7.769	0.593
Ours	**6.689**	**0.644**	**4.042**	**0.780**	**3.155**	**0.810**	**2.847**	**0.820**	**2.787**	**0.822**	**2.746**	**0.823**

表 III 60 回合验证协议下的等预算主动探索对照方法。所有对照方法均使用 205 步和相同的合法动作掩码。

控制器	评估随机种子	最小均方根误差 (dBm)	最终均方根误差 (dBm)
学习策略	7471	1.622	2.877
随机合法动作	7531	2.392	4.411
均匀合法动作循环	7551	2.539	4.046

在相同采样位置进行的重建结果。

表 III 比较了学习策略与随机和均匀合法动作控制在共同预算为 205 步时的表现。由于允许重复访问，该步数最多覆盖网格的 $205/4096 = 5.005\%$。所有控制均使用相同的基于历史的合法动作掩码：若某个动作会离开网格、进入障碍物，或在近期访问窗口内超过重访阈值，则该动作不合法。最小均方根误差是一个回合内达到的最低地图均方根误差，最终均方根误差则是最后一步的数值。如表 III 所示，学习策略取得了最低的最小均方根误差和最终均方根误差，分别为 1.622 和 2.877 dBm；与随机合法动作控制相比，分别低 32.2% 和 34.8%，与均匀合法动作循环相比，则分别低 36.1% 和 28.9%。这些结果表明，构建感知策略采集了更多在相同动作预算下，信息丰富的测量值多于非学习型探索。

表 IV 报告了选定检查点的家族级性能。尽管第二阶段优化仅使用单自由实例，但该检查点是在由 60 个回合组成的六个家族验证聚合结果上选定的，随后在由 120 个回合组成的独立锁定测试集划分上进行一次评估；后者从未用于模型选择。如表 IV 所示，选定检查点在验证集和锁定测试集划分上的聚合最小均方根误差分别达到 1.492 和 1.850 dBm，并且六个家族中有五个在锁定测试集划分上的均方根误差低于 1.65 dBm。表 III 使用不同的随机种子和掩码设置对同一检查点进行了评估，即 (8，0。50) 而非 (4，0。75)，因此每项比较均在各自的协议组内进行。这些结果表明，学习策略在不同数量的活动

表 IV 选定检查点的家族级性能。认知编码器和第一阶段策略被冻结，因此不同候选检查点之间仅第二阶段精炼头存在差异。验证使用每个家族 10 个回合，基础随机种子为 31415；独立锁定测试集划分使用每个家族 20 个回合，基础随机种子为 271828。聚合值为六个家族的等权平均值。

场景家族	验证最小值	验证最终值	锁定最小值	锁定最终值
双自由　，固定弱	0.838	1.226	0.892	1.113
双障碍布局 A	1.679	1.992	1.543	2.363
双障碍布局 B	0.964	1.227	1.643	2.627
单自由完整网格	2.712	6.590	4.239	8.160
单障碍物，3个矩形	1.620	2.088	1.485	1.912
单障碍物，2个矩形	1.143	1.492	1.300	1.543
聚合	1.492	2.436	1.850	2.953

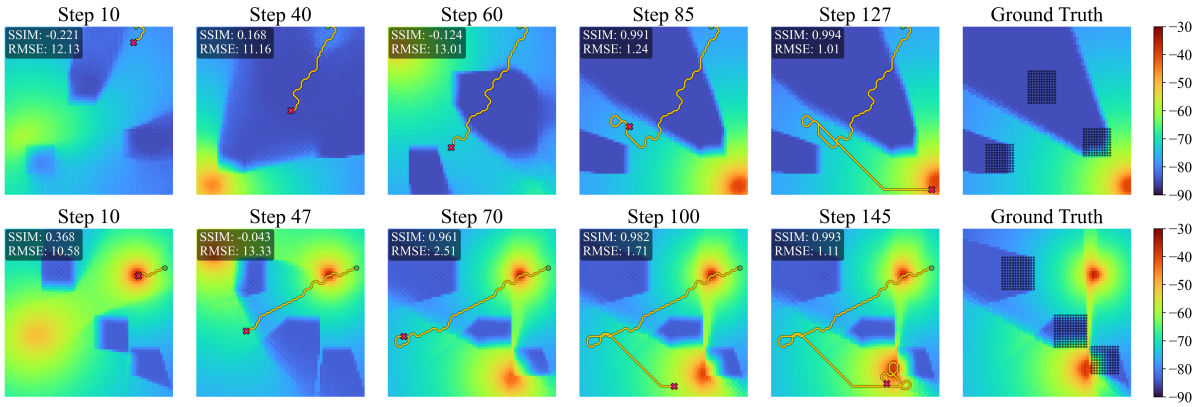

图8. 在线任务中代表性案例的频谱图渐进演化过程及无人机主动探索轨迹。

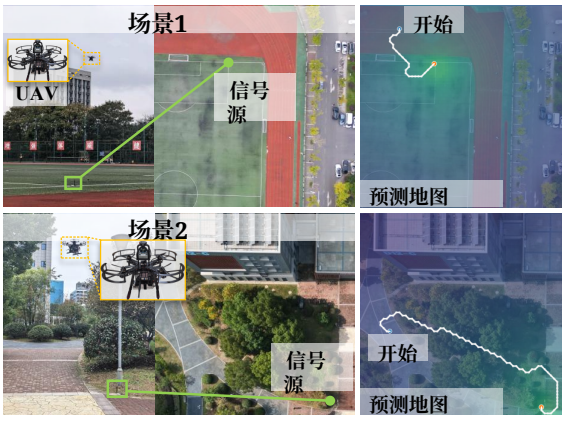

图9. 开放校园环境（场景1）以及有植被和附近障碍物的森林环境（场景2）中的原型级现场演示。

发射机和障碍物布局下仍然有效，而单自由空间族的要求最高，因为完整的自由网格没有提供障碍物边界作为空间锚点。此外，每个家族都在 205 步步数预算耗尽之前达到其最小均方根误差，这表明最佳图是在步数预算内获得的。

为进一步揭示 PCE循环的内部耦合机制，图 8 展示了频谱图的渐进演变，以及相应的无人机轨迹在主动探索任务期间的变化。在图示的早期探索阶段（第 10 步至第 40 步），无人机大范围移动以覆盖测绘区域，而构建地图仍然模糊，误差率较高。随着无人机沿障碍物边界移动，上方回合的均方根误差从第 60 步的 13.01 dBm 急剧下降至第 85 步的 1.24 dBm。在下方回合中，无人机在第 100 步至第 145 步期间于两个信号源之间进行局部移动，最终地图的均方根误差达到 1.11 dBm，结构相似性指数达到 0.993。

最后，图9展示了在真实校园环境和森林环境中的现场部署结果。真实世界部署引入了域偏移这一挑战，未知干扰和硬件噪声可能会降低模型在实际环境中的性能。如图9顶行的场景1所示，无人机在开放的校园环境中运行，运动场上放置了一个视距信号源。预测地图呈现出与信号源已知位置一致的高接收信号强度指示区域。图9底行展示了场景2，该场景的环境更具挑战性，包含茂密植被和附近障碍物。预计这一环境会引入更强的阴影效应和多径效应，从而遮蔽信号。尽管如此，无人机仍成功穿行于障碍物密集的森林环境中，并持续更新频谱图。在两种环境中，从感知到动作选择的完整链路均在 Jetson Orin NX 上运行，并生成 64×64 频谱图输出，证明了原型

级 ECDNet 在受尺寸、重量和功耗约束的无人机平台上进行闭环频谱地图构建的可行性。

VI. 结论

本文提出了一种用于频谱地图构建的具身智能框架，通过闭环PCE耦合与具身探索实现精确的频谱构建。该框架以 ECDNet 的形式实现，通过距离感知卷积、语义桥接提取和认知引导决策模块，有效整合了 PCE 循环。此外，我们还开发了一个实用的基于无人机的频谱构建平台，用于在两种现场环境中进行原型评估。实验结果表明，在非重复访问螺旋协议下，ECDNet 在 4.00% 的采样率下实现了 2.787 dBm 的均方根误差；而在仅 1.00% 的采样率下，其均方根误差低于 U-Net 在 4.00% 采样率下的表现，相应地，在这两个运行点下，采样位置减少了 75%。此外，现场测试为机载闭环运行提供了原型级证据。这些发现支持在评估的六类仿真场景系列和两种现场环境中进行基于具身智能的频谱地图构建，并为低空无线网络中的干扰抑制和频谱感知轨迹规划提供了基础。

参考文献

[1] Y. Bai, H. Zhao, X. Zhang, Z. Chang, R. J¨antti 和 K. Yang，“面向自主多无人机无线网络：基于强化学习方法的综述”，*IEEECommun. Surveys Tuts.*，第25卷，第4期，第3038–3067页，2023年第4季度。[2] J. Fan, P. Ren, J. Chen, J. Qian, J. Wang 和 C. Jiang，“基于扩散的语义通信辅助低空智能物联网服务”，*IEEE Internet Things J.*，第12卷，第10期，第13568–13580页，2025年5月。[3] Y. Zeng, R. Zhang 和 T. J. Lim，“无人机无线通信：机遇与挑战”，*IEEE Commun. Mag.*，第54卷，第5期，第36–42页，2016年5月。[4] K. K. Nguyen, A. Masaracchia, V. Sharma, H. V. Poor 和 T. Q. Duong，“利用深度强化学习进行面向物联网和无线能量传输的 RIS辅助无人机通信”，*IEEE J. Sel. Topics Signal Process.*，第16卷，第5期，第1086–1096页，2022年8月。[5] S. Shrestha, X. Fu 和 M. Hong，“深度频谱制图：利用学习型神经模型补全无线电地图张量”，*IEEE Trans. Signal Process.*，第70卷，第1170–1184页，2022年。[6] S. Timilsina, S. Shrestha 和 X. Fu，“利用张量和深度生成模型进行量化无线电地图估计”，*IEEE Trans. Signal Process.*，第72卷，第173–189页，2024年。[7] H. Liang, Y. Diao, Y. Wu, F. Zhou 和 Q. Wu，“协同赋能：无线通信 meets 具身智能”，*IEEE Commun. Mag.*，提前访问，2026年。[8] K. Sato, K. Suto, K. Inage, K. Adachi 和 T. Fujii，“空间—频率插值无线电地图”，*IEEE Trans. Veh. Technol.*，第70卷，第1期，第714–725页，2021年1月。[9] D. Romero 和 S.-J. Kim，“无线电地图估计：一种频谱制图的数据驱动方法”，*IEEE Signal Process. Mag.*，第39卷，第6期，第53–72页，2022年11月。[10] J. Liu, F. Zhou, X. Liu, R. Ding, L. Yuan 和 Q. Wu，“数据与语义双驱动的6G频谱管理频谱地图构建”，载于*Proc. IEEE Global Commun. Conf. (GLOBECOM)*，2024年12月，第139–144页。[11] Q. Chen, J. Yang, M. Huang 和 Q. Zhou，“ACT-GAN：基于带有ACT模块的生成对抗网络进行无线电地图构建”，*IET Commun.*，第18卷，第19期，第1541–1550页，2024年12月。[12] V. Q. Pham 和 D. Romero，“用于无线电地图估计的空间变换器”，载于*Proc. IEEE Int. Conf. Commun. (ICC)*，2025年6月，第6155–6160页。[13] H. Lu, C. Vattheuer, B. Mirzasoleiman 和 O. Abari，“NeWRF：用于无线辐射场重建和信道预测的深度学习框架”，载于机器学习国际会议（ICML）论文集，2024年7月，第33147–33159页。[14] M. Raissi, P. Perdikaris 和 G. E. Karniadakis，“物理信息神经网络：用于求解涉及非线性偏微分方程的正问题和逆问题的深度学习框架”，《计算物理学杂志》，第378卷，第686–707页，2019年。[15] M. Shahid, K. Das, H. Ushaq, H. Zhang, J. Song, D. Qiao, S. Babu, Y. Guan, Z. Zhu 和 A. Ahmad，“ReVeal：用于高保真无线电环境映射的物理信息神经网络”，载于IEEE动态频谱接入网络国际研讨会（DySPAN）论文集，2025年5月，第1–10页。[16] F. Zhou, C. Wang, G. Wu, Y. Wu, Q. Wu 和 N. Al-Dhahir，“通过智能频率—空间推理实现面向频谱管理的高精度频谱地图构建”，《IEEE通信汇刊》，第71卷，第7期，第3932–3945页，2023年7月。[17] J. Liu, X. Liu, H. Liang, L. Yuan, F. Zhou 和 Q. Wu，“用于同步进行频谱地图构建和信号源定位的数据与语义双驱动智能推断框架”，《IEEE物联网杂志》，第12卷，第12期，第18750–18764页，2025年6月。[18] Y. Zeng 和 X. Xu，“通过信道知识地图迈向环境感知的6G通信”，《IEEE无线通信》，第28卷，第3期，第84–91页，2021年6月。[19] S. Zhang 和 R. Zhang，“基于无线电地图的蜂窝连接无人机三维路径规划”，《IEEE无线通信汇刊》，第20卷，第3期，第1975–1989页，2021年3月。[20] Y. Dong, C. He, Z. Wang 和 L. Zhang，“用于无人机抗干扰通信的无线电地图辅助路径规划”，《IEEE信号处理快报》，第29卷，第607–611页，2022年。[21] K. D. Polyzos, A. Sadeghi, W. Ye, S. Sleder, K. Houssou, J. Calder, Z.-L. Zhang 和 G. B. Giannakis，“面向样本高效5G无线电地图重建的贝叶斯主动学习”，《IEEE无线通信汇刊》，第23卷，第12期，第19382–19396页，2024年12月。[22] R. Shrestha, D. Romero 和 S. P. Chepuri，“频谱测绘：利用自主无人机进行主动无线电地图估计”，《IEEE无线通信汇刊》，第22卷，第1期，第627–641页，2023年1月。[23] Y. Zeng, X. Xu, S. Jin 和 R. Zhang，“利用深度强化学习实现蜂窝连接无人机的同步导航与无线电测绘”，《IEEE无线通信汇刊》，第20卷，第7期，第4205–4220页，2021年7月。[24] O. Esrafilian, H. Bayerlein 和 D. Gesbert，“用于物联网网络中样本高效无人机轨迹设计的模型辅助深度强化学习”，载于IEEE全球通信大会（GLOBECOM）论文集，2021年12月，第1–6页。[25] H. Liu, D. Guo 和 A. Cangelosi，“具身智能：形态学、动作、感知与学习的协同作用”，《ACM计算综述》，第57卷，第7期，文章编号186，2025年。[26] Z. Shao, H. Yang, L. Xiao, W. Su, Y. Chen 和 Z. Xiong，“面向抗干扰无人机辅助移动边缘计算的深度强化学习资源管理”，《IEEE移动计算汇刊》，第23卷，第12期，第13358–13374页，2024年12月。[27] M. F. Ahmed, K. Masood, V. Fr´emont 和 I. Fantoni，“主动SLAM：过去十年的回顾”，《传感器》，第23卷，第19期，文章编号8097，2023年。[28] Z. Lai, C. Zhang, Y. Zhang 和 D. Yuan，“DEMO：通过基于GPU的光线追踪实现动态环境中无线信号传播的实时仿真”，载于ACM SIGCOMM海报与演示论文集，2024年8月，第98–100页。[29] T. Isomura, K. Kotani, Y. Jimbo 和 K. J. Friston，“利用体外神经网络对自由能原理进行实验验证”，《自然·通信》，第14卷，文章编号4547，2023年。[30] K. Friston, L. Da Costa, N. Sajid, C. Heins, K. Ueltzh¨offer, G. A. Pavliotis 和 T. Parr，“更简单但又不过分简单的自由能原理”，《物理学报告》，第1024卷，第1–29页，2023年。[31] B. Millidge, A. Tschantz 和 C. L. Buckley，“期望自由能从何而来？”，《神经计算》，第33卷，第2期，第447–482页，2021年2月。[32] X. Wang, P. Zheng, H. Jia, N. Cheng, R. Sun, C. Zhou 和 X. Shen，“RadioDiff-Flux：通过生成式去噪扩散模型和轨迹中点复用实现高效无线电地图构建”，《IEEE认知通信与网络汇刊》，第12卷，第4882–4895页，2026年。[33] T. Isomura, H. Shimazaki 和 K. J. Friston，“规范神经网络执行主动推断”，《通信生物学》，第5卷，文章编号55，2022年。