

# 动态电磁环境中具身频谱感知技术研究

查浩然<sup>1</sup>, 孙露<sup>1</sup>, 尚佳颖<sup>1</sup>, 王美玉<sup>2</sup>, 林云<sup>1</sup>

(1. 哈尔滨工程大学信息与通信工程学院, 黑龙江 哈尔滨 150001; 2. 杭州电子科技大学通信工程学院, 浙江 杭州 310018)

**摘要:** 针对动态电磁环境下传统频谱感知依赖固定采集与离线训练、在分布偏移时易出现识别性能退化与决策僵化的问题, 提出一种具身频谱感知建模与策略求解方法。构建融合环境状态、观测、知识表示与具身行动的统一系统模型, 将频谱感知抽象为基于知识状态的序贯决策过程; 以识别正确性、时延以及资源开销为核心设计即时回报, 定义值函数与状态行动值函数, 推导满足 Bellman 结构的最优性方程, 给出最优策略的存在形式与结构特征, 并提出基于值函数更新的策略求解方法。结合调制识别任务构建具身决策框架, 通过多模型选择实现对不同环境的自适应匹配。仿真结果表明, 相比静态单模型方案, 所提方案在整体识别准确率、环境适应性与错误恢复能力等方面均显著提升。

**关键词:** 具身智能; 频谱感知; 动态环境; 闭环决策; Bellman 结构

中图分类号: TP393.0

文献标志码: A

DOI: 10.11959/j.issn.1000-436x.2026022

## Embodied intelligent spectrum sensing techniques in dynamic electromagnetic environments

Zha Haoran<sup>1</sup>, Sun Lu<sup>1</sup>, Shang Jiaying<sup>1</sup>, Wang Meiyu<sup>2</sup>, Lin Yun<sup>1</sup>

1. College of Information and Communication Engineering, Harbin Engineering University, Harbin 150001, China

2. College of Communication Engineering, Hangzhou Dianzi University, Hangzhou 310018, China

**Abstract:** The performance degradation and decision rigidity of conventional spectrum sensing in dynamic electromagnetic environments, arising from reliance on fixed acquisition and offline training under distribution shifts, were addressed. An embodied spectrum sensing modeling and policy computation approach was proposed. A unified system model integrating environment state, observation, knowledge representation, and embodied actions was constructed, and spectrum sensing was formulated as a sequential decision process over knowledge states. An immediate reward was designed with recognition of correctness, latency, and resource cost as the core factors. The value function and the state action value function were defined, the Bellman optimality equations were derived, the existence form and structural characteristics of the optimal policy were characterized, and a policy computation method based on value function updates was developed. An embodied decision framework for modulation recognition was established, and adaptive matching to diverse environments was enabled through multi model selection. Simulation results show that, compared with a static single model baseline, the proposed scheme significantly improves overall recognition accuracy, environmental adaptability, and error recovery capability. The proposed scheme provides an effective approach for robust spectrum sensing in dynamic electromagnetic environments.

**Keywords:** embodied intelligence, spectrum sensing, dynamic environment, closed loop decision, Bellman structure

收稿日期: 2025-11-29; 修回日期: 2026-01-21

通信作者: 林云, linyun@hrbeu.edu.cn

基金项目: 国家自然科学基金资助项目 (No.62501184, No.U23A20271); 中央高校基本科研业务费专项资金资助项目 (No.3072025YY0801)

**Foundation Items:** The National Natural Science Foundation of China (No.62501184, No.U23A20271), The Fundamental Research Funds for the Central Universities (No.3072025YY0801)

## 0 引言

电磁频谱感知作为新一代信息基础设施中连通物理空间与信息空间的关键技术,通过对环境中各类发射源、业务信号和干扰行为进行监测与识别,构建了一个覆盖“发现-感知-评估”的综合体系<sup>[1]</sup>。该体系为低空通信、物联网、车联网等应用提供了必要的电磁环境认知能力,并为频谱精细化管理和业务安全可靠运行提供了重要支撑,能够提升频谱利用效率并降低电磁冲突风险<sup>[2-3]</sup>。随着低空经济、泛在物联和智能感知网络的快速发展,频谱占用主体数量和类型不断增加,业务体制和干扰方式愈发多样,电磁环境在时间和空间上呈强非平稳特性<sup>[4-5]</sup>。在此背景下,传统的基于固定采集策略和离线训练模型的频谱感知方法难以适应环境的快速演化,不能及时反映新体制信号和异常行为,频谱感知的实时性与可靠性面临新的挑战。

具身智能作为数据驱动决策一体化的重要发展方向,通过在同一载体内融合观测、处理和执行能力,使智能体能够主动调节自身观测方式并在闭环中持续更新内部知识,为电磁频谱感知提供了一种新的实现范式<sup>[6]</sup>。在具身频谱感知场景中,感知主体不再被动接收外界信号,而是根据当前对环境的认识选择观测频段、采样时长和处理路径,并结合观测结果不断修正决策策略,从而在动态环境中形成自适应的感知决策闭环<sup>[7]</sup>。将具身智能思想引入频谱感知,有望在有限观测与计算资源下提升未知信号发现能力和异常识别鲁棒性,缓解训练环境与运行环境不一致带来的性能退化问题。

目前,已有研究从多个方面对电磁频谱智能感知展开探索。一方面,围绕特定辐射源识别、调制识别和协议识别等任务,提出了迁移学习、增量学习和开集识别等方法,以缓解设备漂移、目标集合扩展和噪声增强带来的性能下降<sup>[8-11]</sup>。另一方面,在动态频谱接入和主动感知方面,已有工作尝试利用序贯决策方法调整观测频段和访问策略,以提升频谱利用率和业务服务质量<sup>[12-13]</sup>。此外,部分研究结合具身智能与低空电磁应用,探讨了在低空场景中通过具身智能体开展频谱监测与管控的可行性,初步体现了感知-决策-执行一体化的潜力<sup>[14]</sup>。这些工作为提升电磁频谱感知的智能化水平奠定了基础。

然而,现有研究大多面向某一具体任务或局部

模块开展优化,缺乏从整体视角出发对具身频谱感知进行统一建模与分析的系统性工作。一方面,观测、知识和行动三者之间的内在关系尚未在统一数学框架下清晰刻画,不利于从理论层面解释具身频谱感知的运行机理和性能边界。另一方面,将具身频谱感知视为一个随时间演化的决策过程进行分析仍然较少,缺乏对不同决策策略在长期运行中的优劣进行统一评价的工具,这使策略设计更多地依赖经验调参,理论依据不足。特别是在复杂电磁环境中,频谱感知主体需要在有限观测机会下持续做出影响后续观测条件和识别效果的决策,如何在兼顾识别性能与时延开销的前提下构建有明确最优性条件的具身频谱感知策略,仍是一个有待深入研究的课题。

针对上述问题,本文结合动态电磁环境特点与具身智能体运行机制,提出一种面向电磁频谱感知的具身感知决策架构,以实现环境观测、内部知识更新与决策感知的有机协同。基于该架构,本文从知识状态出发对具身频谱感知过程进行建模,将频谱感知视为一个在知识空间上演化的决策过程,建立观测、知识更新与行动之间的闭环关系,给出满足贝尔曼(Bellman)结构的最优性条件,从统一视角刻画具身频谱感知策略的结构特征。为弥补现有研究在具身频谱感知整体建模与策略最优性分析方面的不足,本文进一步设计具身决策方法并开展仿真验证。仿真结果表明,所提方案在多种典型动态频谱环境中能够显著提升识别准确率与错误恢复能力,同时保持较好的时延表现,验证了具身频谱感知闭环建模与决策方法的有效性。

本文的主要工作如下。

1) 提出一种面向电磁频谱感知的具身感知决策架构,从环境状态、观测、知识表示和行动4个方面出发,建立感知主体与动态电磁环境的闭环交互关系,为具身频谱感知的整体分析提供统一建模基础。

2) 在具身感知决策架构基础上,将频谱感知过程抽象为以知识状态为核心的序贯决策问题,构建即时回报函数、值函数和状态行动值函数,推导满足Bellman结构的最优性方程,给出具身频谱感知策略的状态依赖形式与基本性质。

3) 结合调制识别任务构建具身频谱感知方法框架,通过多模型选择与观测配置决策实现具体具

身感知流程。仿真结果表明,在不同动态环境和信噪条件下,所提方案相较静态单模型方案在识别准确率、环境适应性和错误恢复能力等方面具有明显优势,展现出较好的鲁棒性与实际应用潜力。

## 1 具身频谱感知的理论建模

### 1.1 基本定义

在动态电磁环境中,频谱感知不再仅仅是对外界信号的被动采集,而是涵盖观测、内部认知和外部决策的完整过程。具身频谱感知强调感知主体通过自身行动主动调节观测方式,并在持续交互中更新对环境的认知状态,使感知行为与环境演化形成闭环结构<sup>[15-16]</sup>。为刻画这一过程,本文建立具身频谱感知理论模型,将电磁环境视为随时间演化的状态域,将感知主体的内部状态划分为观测、知识与行动3类基本变量。

设离散时间步  $t = 0, 1, 2, \dots$ , 在时间  $t$ , 电磁环境处于某一真实状态  $X_t \in \mathcal{X}$ , 其中  $\mathcal{X}$  为环境状态集合, 可包含频谱占用情况、信号体制分布以及干扰活动等信息。为完整刻画动态电磁环境中恶意干扰行为对频谱感知的影响, 本文将干扰源活动显式纳入环境状态建模。设目标信号相关的环境状态为  $S_t \in \mathcal{S}$ , 干扰源状态为  $J_t \in \mathcal{J}$ , 则环境状态可表示为复合形式  $X_t = (S_t, J_t)$ , 其中,  $\mathcal{S}$  为目标信号状态空间, 涵盖调制类型、信噪比等级、信道条件等属性;  $\mathcal{J}$  为干扰状态空间, 包含干扰功率  $P_J$ 、干扰带宽  $B_J$ 、干扰中心频率  $f_J$  以及干扰样式  $\tau_J$  等参数。

考虑到实际电磁对抗场景中恶意干扰往往呈动态变化特性, 干扰参数的时域演化可建模为

$$J_{t+1} = \Gamma(J_t, X_t^{(S)}, A_t, \zeta_t) \quad (1)$$

其中,  $\Gamma: \mathcal{J} \times \mathcal{S} \times \mathcal{A} \times \mathcal{E} \rightarrow \mathcal{J}$  为干扰状态转移映射,  $\zeta_t \in \mathcal{E}$  为随机扰动项, 用于刻画干扰行为的随机性与不确定性。上述建模具有良好的通用性, 能够涵盖多种典型恶意干扰模式: 当  $\Gamma$  退化为恒等映射时, 对应固定参数干扰; 当  $\Gamma$  仅依赖  $\zeta_t$  时, 对应随机跳变干扰或扫频干扰; 当  $\Gamma$  显式依赖感知主体的行动  $A_t$  时, 则可刻画具有对抗性的自适应智能干扰, 即干扰方根据感知方的策略调整自身干扰参数。这一建模框架为分析具身频谱感知策略在对抗环境下的性能提供了理论基础。

考虑到实际电磁环境的复杂性, 环境状态  $X_t$  还应涵盖多种典型影响因素, 包括但不限于: 信噪

比的时变特性、多径传播引起的频率选择性衰落、载波频率偏移、相位噪声以及接收机硬件非理想特性(如 IQ 不平衡)等。这些因素在时间和空间上呈非平稳特性, 共同构成了动态复杂电磁环境的核心要素。频谱感知主体在时间  $t$  的外部观测记为  $O_t \in \mathcal{O}$ , 内部知识记为  $K_t \in \mathcal{K}$ , 行动决策记为  $A_t \in \mathcal{A}$ , 分别属于观测空间  $\mathcal{O}$ 、知识空间  $\mathcal{K}$  与行动空间  $\mathcal{A}$ 。观测  $O_t$  反映在当前行动配置下可获得的频谱信息, 知识  $K_t$  用于凝练历史观测与决策结果, 行动  $A_t$  决定下一步的观测配置与处理策略。在存在恶意干扰的场景中, 知识状态  $K_t$  还需整合对干扰行为的认知信息, 包括干扰存在性的检测置信度、干扰参数的在线估计和干扰模式的历史特征等, 从而为感知主体的抗干扰决策提供依据。具身频谱感知的核心在于观测、知识与行动之间的相互作用。首先, 观测由环境状态和当前行动共同决定, 可表示为

$$O_t = \Phi(X_t, A_t) \quad (2)$$

其中,  $\Phi$  为从环境状态与行动到观测空间的映射, 体现了行动对可观测性的影响。在该映射中, 干扰状态  $J_t$  通过叠加于目标信号或改变信道特性等方式影响观测质量, 感知主体可通过调整行动  $A_t$  (如改变观测频段、增加采样时长或启用抗干扰预处理) 来缓解干扰影响。这种观测、干扰与行动之间的耦合关系, 使具身频谱感知框架天然具备通过主动决策实现抗干扰的能力。在后续的 Bellman 最优性结构中, 策略优化过程会自动驱动感知主体学习能够有效应对干扰的决策行为, 干扰对识别性能的负面影响通过即时回报函数中的识别正确性指标得以体现, 不需要额外引入专门的抗干扰回报项。其次, 知识随观测不断演化, 可表示为

$$K_{t+1} = \Psi(K_t, O_{t+1}) \quad (3)$$

其中,  $\Psi$  为知识更新映射, 用于将新观测与既有知识进行融合, 使知识在时间上形成递推结构。在频谱感知场景中, 知识状态  $K_t$  通常包含3类信息: 一是环境特征估计, 如当前信噪比水平、信道衰落程度、干扰强度等统计量; 二是历史决策效果, 如各候选模型在近期的识别准确率、错误类型分布等; 三是任务相关的累积统计, 如连续正确或错误的次数、置信度变化趋势等。相应地, 知识更新映射  $\Psi$  可采用多种实现形式: 对于环境特征估计, 可通过指数滑动平均或贝叶斯后验更新进行递推; 对于历

史决策效果,可采用滑动窗口统计追踪各模型近期表现;对于累积统计量,可通过计数器递增或重置机制进行更新。上述方式的共同特点是仅依赖当前知识与新观测即可完成计算,不需要存储完整历史序列。最后,行动由当前知识驱动,可表示为

$$A_t = \pi(K_t) \quad (4)$$

其中,  $\pi$  为感知主体的决策映射,用以在知识空间上选择合适的观测与处理方式。在上述定义下,具身频谱感知可以抽象为一个以知识状态为中心的闭环系统。环境状态通过观测映射投影到观测空间,观测经由知识更新形成内部认知,再由决策映射生成行动,行动反过来影响未来可获得的观测,从而构成  $O_t \rightarrow K_t \rightarrow A_t \rightarrow O_{t+1}$  的循环结构。

## 1.2 闭环感知结构

结合式(2)~式(4),在给定策略  $\pi$  时,系统的运行轨迹可以表示为  $\{X_0, O_0, K_0, A_0, X_1, O_1, K_1, A_1, \dots, X_t, O_t, K_t, A_t\}$ , 其联合分布可分解为

$$P(X_0, O_0, K_0, A_0, X_1, O_1, K_1, A_1, \dots) = P(X_0) \prod_{t=0}^{\infty} P(A_t | K_t) P(O_t | X_t, A_t) \cdot P(K_{t+1} | K_t, O_{t+1}) P(X_{t+1} | X_t, A_t) \quad (5)$$

其中,  $P(A_t | K_t)$  由策略  $\pi$  决定,  $P(O_t | X_t, A_t)$  由观测映射  $\Phi$  对应的观测通道决定,  $P(K_{t+1} | K_t, O_{t+1})$  对应知识更新映射  $\Psi$ ,  $P(X_{t+1} | X_t, A_t)$  则反映环境自身的演化规律及其对行动响应。在复杂电磁环境中,观测条件分布  $P(O_t | X_t, A_t)$  受到多种因素的联合影响:信噪比变化直接决定观测信号的可用信息量;多径衰落引起信号幅度和相位的随机起伏;频率偏移和相位噪声导致接收信号的频谱扩展与畸变。上述因素通过观测映射  $\Phi$  综合作用于观测结果,使感知主体在不同环境条件下获得的有效信息存在显著差异。具身感知闭环结构的优势在于,感知主体能够通过行动  $A_t$  主动调整观测策略,并依据知识状态  $K_t$  对上述环境因素进行间接估计与自适应响应。

在具身频谱感知模型中,知识状态  $K_t$  承担了两个关键角色。第一,从信息凝练角度看,  $K_t$  可被视为观测序列  $\{O_0, \dots, O_t\}$  的压缩表示。若存在某种知识表示方式,使得对于任意  $t$  和未来决策序列  $\{A_t, A_{t+1}, \dots, A_{\infty}\}$  均有

$$P(X_{t+1:\infty} | O_{0:t}, A_{0:t}) = P(X_{t+1:\infty} | K_t, A_{0:t}) \quad (6)$$

则认为  $K_t$  在预测未来环境演化方面是一个充分统计量,即在给定  $K_t$  的条件下,观测历史中的冗余信息已经被压缩。第二,从决策依赖角度看,若对任意策略  $\pi$  与性能评价函数  $f(\cdot)$  有

$$\begin{aligned} \mathbb{E}[f(X_{t:\infty}, O_{t:\infty}, A_{t:\infty}) | O_{0:t}, A_{0:t}] = \\ \mathbb{E}[f(X_{t:\infty}, O_{t:\infty}, A_{t:\infty}) | K_t] \end{aligned} \quad (7)$$

则意味着在优化长期性能时仅依赖  $K_t$  即可,不必显式保留完整历史观测序列,这为将具身频谱感知过程抽象为以知识状态为核心的决策过程提供了理论依据。

行动选择不仅影响当前识别结果,还决定未来可获得的观测类型与质量,因此具身频谱感知不再是单次判决问题,而是一个通过行动设计观测序列的动态过程,如图1所示。上述闭环结构为后续从序贯决策视角分析具身频谱感知提供了明确的理论支撑。

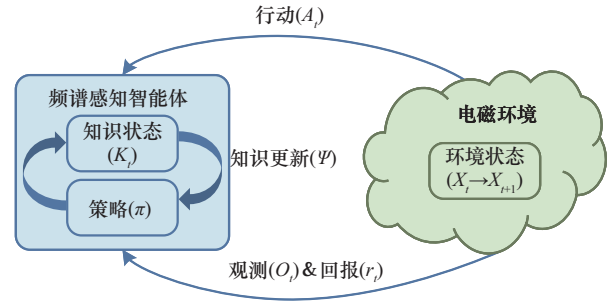


图1 具身感知闭环交互模型

## 1.3 性能指标与问题建模

为了在统一理论框架下评价具身频谱感知策略的优劣,需要构建与任务目标相匹配的性能指标体系。对于频谱感知任务而言,典型关注点包括识别正确性、误警与漏警情况、感知时延以及观测成本等方面。

设在时间  $t$  的频谱感知任务输出为  $\hat{y}_t$ , 与其对应的真实标签为  $y_t$ 。对于占用检测或体制识别任务,使用指示函数  $\mathbb{I}(\hat{y}_t = y_t)$  刻画识别正确性。为进一步细化识别性能的评价,在频谱感知任务中引入误警率与漏警率两项指标。设在时间  $t$  的真实频谱状态为  $H_t \in \{H_0, H_1\}$ , 其中,  $H_0$  表示频谱空闲,  $H_1$  表示频谱占用,对应的判决输出为  $\hat{H}_t$ 。误警率定义为在频谱实际空闲条件下误判为占用的概率,即

$$P_{\text{FA}} = P(\hat{H}_t = H_1 | H_t = H_0) \quad (8)$$

漏警率定义为在频谱实际占用条件下误判为空闲的概率, 即

$$P_{\text{MD}} = P(\hat{H}_t = H_0 | H_t = H_1) \quad (9)$$

在具身频谱感知的序贯决策框架下, 误警率与漏警率可进一步表示为知识状态与策略的函数。设在知识状态  $K_t$  下采取行动  $A_t$  后的判决输出为  $\hat{H}_t(K_t, A_t)$ , 则策略  $\pi$  的平均误警率与漏警率分别为

$$\bar{P}_{\text{MD}}^\pi = \mathbb{E}^\pi [P(\hat{H}_t = H_0 | H_t = H_1, K_t)] \quad (10)$$

$$\bar{P}_{\text{FA}}^\pi = \mathbb{E}^\pi [P(\hat{H}_t = H_1 | H_t = H_0, K_t)] \quad (11)$$

上述定义将误警率与漏警率纳入以知识状态为核心的决策评价体系, 为后续构建综合性能指标提供了基础。在具身频谱感知场景中, 观测与决策是序贯进行的, 过多的重复观测和调用复杂模型会增加整体感知时延。为反映这一开销, 记与行动  $A_t$  相关的时间成本为  $d_t(A_t)$ , 与观测相关的资源成本为  $c_t(A_t)$ 。在上述基础上, 定义即时回报函数为

$$r_t = R(K_t, A_t, O_{t+1}) \quad (12)$$

典型形式为

$$r_t = \alpha \cdot \mathbb{I}(\hat{y}_t = y_t) - \beta \cdot d_t(A_t) - \eta \cdot c_t(A_t) - \lambda_{\text{FA}} \cdot \mathbb{I}_{\text{FA},t} - \lambda_{\text{MD}} \cdot \mathbb{I}_{\text{MD},t} \quad (13)$$

其中,  $\alpha$ 、 $\beta$ 与 $\eta$ 为权重系数, 用于平衡识别正确性、时延开销与观测成本;  $\mathbb{I}_{\text{FA},t}$ 与 $\mathbb{I}_{\text{MD},t}$ 分别为时间  $t$  发生误警与漏警的指示函数, 用于平衡识别正确性、时延开销、观测成本以及误警与漏警带来的代价。式(13)表明, 若当前决策既能做出正确判决又不会引入过大的时延与观测开销, 即时回报较高; 反之, 则在回报中体现惩罚。值得注意的是, 在复杂电磁环境下, 不同环境因素对即时回报的影响存在差异: 低信噪比和强干扰环境通常导致识别正确性下降, 从而降低即时回报中的正向分量; 频率选择性衰落和相位噪声等因素可能需要感知主体采用更复杂的处理模型或延长观测时间以维持识别性能, 进而增加时延开销  $d_t(A_t)$ ; 当环境条件剧烈变化时, 感知主体可能需要更频繁地切换观测配置, 导致观测成本  $c_t(A_t)$  上升。因此, 本文定义的即时回报函数能够综合反映复杂电磁环境因素对具身频谱感知性能的多维度影响。

为了评价具身频谱感知策略在长期运行中的表现, 引入折扣因子  $\gamma \in (0, 1)$  定义策略  $\pi$  在初始知识状态  $K_0$  下的长期性能指标为

$$J^\pi(K_0) = \mathbb{E}^\pi \left[ \sum_{t=0}^{\infty} \gamma^t r_t | K_0 \right] \quad (14)$$

其中, 期望运算对环境演化、观测随机性以及策略引起的随机性同时取值。式(14)反映了在给定策略下, 具身频谱感知在整个运行周期内综合考虑识别正确性与开销的整体表现。

在上述性能指标定义下, 可以将具身频谱感知策略设计问题刻画为

$$\max_{\pi} J^\pi(K_0) \quad (15)$$

即在所有可行策略中寻找长期折扣回报最大的策略。优化目标不仅涉及单一时刻的识别性能, 还综合考虑了未来观测条件与决策空间的变化, 因此具身频谱感知本质上是一个面向长期效果的序贯优化问题。

## 2 具身频谱感知决策建模

本节从序贯决策角度引入值函数与状态行动值函数, 推导具身频谱感知策略所满足的 Bellman 最优性结构。

### 2.1 知识状态与序贯决策

由式(10)和式(11)可知, 在合适的知识表示下,  $K_t$  对未来环境演化与性能评价具有充分性。直观而言, 知识状态同时汇集了历史观测信息与既有决策结果, 在给定  $K_t$  的条件下, 可以不再显式关注完整的观测与行动历史。这一性质使具身频谱感知过程可以抽象为以  $K_t$  为状态的序贯决策过程。

在策略  $\pi$  下, 具身频谱感知过程可视作知识状态空间  $\mathcal{K}$  上的受控随机过程。每一时间步的演化遵循

$$\begin{aligned} A_t &= \pi(K_t) \\ O_{t+1} &\sim P(O_{t+1} | X_{t+1}, A_t) \\ K_{t+1} &= \Psi(K_t, O_{t+1}) \end{aligned} \quad (16)$$

其中, 环境状态  $\{X_t\}$  的演化由  $P(X_{t+1} | X_t, A_t)$  决定, 知识状态  $\{K_t\}$  的演化由知识更新映射  $\Psi$  与观测过程共同决定。由于  $K_t$  已经对历史信息进行了压缩与整合, 在给定当前知识状态与行动的情况下, 下一知识状态的分布可以写为

$$P(K_{t+1}|K_t, A_t) = \iint P(K_{t+1}|K_t, O_{t+1})P(O_{t+1}|X_{t+1}, A_t) \cdot P(X_{t+1}|X_t, A_t) dO_{t+1} dX_{t+1} \quad (17)$$

其中, 对  $X_{t+1}$  与  $O_{t+1}$  的积分可以理解为对环境与观测随机性的综合。若将环境状态边缘化, 则对决策而言, 可以将  $\{K_t\}$  看作满足

$$P(K_{t+1}|K_{0:t}, A_{0:t}) = P(K_{t+1}|K_t, A_t) \quad (18)$$

由此,  $K_t$  在给定行动下构成马尔可夫决策过程, 具身频谱感知可以建模为在知识状态空间上的马尔可夫决策过程, 式(14)给出的长期性能指标可以重写为

$$J^\pi(K_0) = \mathbb{E} \left[ \sum_{t=0}^{\infty} \gamma^t R(K_t, A_t, O_{t+1}) \mid K_0, A_t = \pi(K_t) \right] \quad (19)$$

其中, 期望仅依赖于初始知识状态  $K_0$  与策略  $\pi$ 。如图 2 所示, 具身频谱感知策略优化问题式(19)等价于在知识状态空间上求解一个标准的序贯决策优化问题。

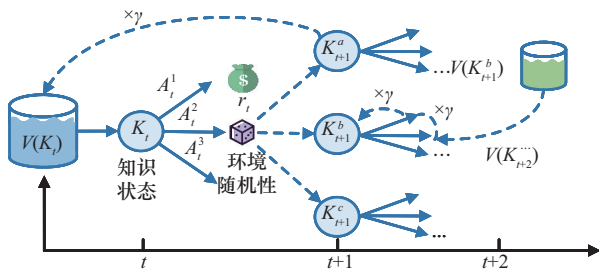


图 2 基于知识状态的具身序贯决策演化过程

### 2.2 值函数定义

为了刻画不同策略在不同知识状态下的表现, 引入具身频谱感知的值函数。对于给定策略  $\pi$ , 定义值函数为

$$V^\pi(K) = \mathbb{E}^\pi \left[ \sum_{t=0}^{\infty} \gamma^t r_t \mid K_0 = K \right] \quad (20)$$

其中,  $V^\pi(K)$  表示以知识状态  $K$  作为起点, 在整个后续运行过程中遵循策略  $\pi$  所能获得的期望折扣回报。

在值函数基础上, 进一步定义状态行动值函数为

$$Q^\pi(K, A) = \mathbb{E} \left[ r_0 + \gamma V^\pi(K_1) \mid K_0 = K, A_0 = A \right] \quad (21)$$

其中,  $K_1$  的分布由  $P(K_1|K_0, A_0)$  决定, 如式(17)所示;  $Q^\pi(K, A)$  表示在当前知识状态为  $K$  且采取行动

$A$  时, 当前步即时回报与后续在策略  $\pi$  下累积回报的期望和。

由值函数定义可得, 值函数与状态行动值函数之间满足

$$V^\pi(K) = \mathbb{E}_{A \sim \pi(\cdot|K)} [Q^\pi(K, A)] \quad (22)$$

同时, 利用式(12)与式(17), 可将式(22)展开为

$$Q^\pi(K, A) = \mathbb{E} \left[ R(K, A, O_1) + \gamma \mathbb{E}^\pi \left[ \sum_{t=1}^{\infty} \gamma^{t-1} r_t \mid K_1 \right] \right] = \mathbb{E} \left[ R(K, A, O_1) + \gamma V^\pi(K_1) \right] \quad (23)$$

其中, 外层期望对  $O_1$  与  $K_1$  的随机性取值。式(20)~式(23)给出了具身频谱感知在策略  $\pi$  下的值函数与状态行动值函数的基本关系。

### 2.3 Bellman 最优性方程

在值函数定义的基础上, 首先明确最优值函数与最优状态行动值函数的核心内涵。最优值函数为

$$V^*(K) = \sup_{\pi} V^\pi(K) \quad (24)$$

其中,  $V^*(K)$  代表在知识状态  $K$  下, 通过选择最优策略能获得的最大长期折扣回报, 刻画了该状态的性能上限。最优状态行动值函数为

$$Q^*(K, A) = \sup_{\pi} Q^\pi(K, A) \quad (25)$$

其中,  $Q^*(K, A)$  聚焦“状态行动”对, 描述了在知识状态  $K$  下采取行动  $A$  后, 后续遵循最优策略的最大期望回报, 将状态最优性与具体行动选择直接关联。

基于动态规划的递推逻辑, 可推导具身频谱感知的 Bellman 最优性方程, 具体可通过式(26)和式(27)共同体现。

$$V^*(K) = \sup_{A \in \mathcal{A}} Q^*(K, A) \quad (26)$$

$$Q^*(K, A) = \mathbb{E} \left[ R(K, A, O_1) + \gamma V^*(K_1) \right] \quad (27)$$

式(26)和式(27)的核心思想是“当前决策的最优性依赖未来决策的最优性”。只要即时回报有界且折扣因子  $\gamma$  严格小于 1, 可以证明上述方程存在唯一解, 并且对应某一个最优策略集合。

由式(26)可直接构造一类最优策略, 即对任意知识状态  $K$ , 若选取策略

$$\pi^*(K) \in \arg \max_{A \in \mathcal{A}} Q^*(K, A) \quad (28)$$

则有  $V^{\pi^*}(K) = V^*(K)$ , 即  $\pi^*$  在所有知识状态上均

达到最优值函数。换言之，具身频谱感知的最优策略可以表示为对最优状态行动值函数在行动维度上的逐点极大选择。

从算法实现角度看，若能构造一系列近似值函数  $\{V_n(K)\}$  或状态行动值函数  $\{Q_n(K,A)\}$ ，并依据式(29)和式(30)进行迭代

$$V_{n+1}(K) = \sup_{A \in \mathcal{A}} \mathbb{E} [R(K,A,O_1) + \gamma V_n(K_1)] \quad (29)$$

$$Q_{n+1}(K,A) = \mathbb{E} \left[ R(K,A,O_1) + \gamma \sup_{A' \in \mathcal{A}} Q_n(K_1,A') \right] \quad (30)$$

则在适当条件下， $\{V_n\}$  与  $\{Q_n\}$  将分别收敛于  $V^*$  与  $Q^*$ 。由此可以得到近似最优策略

$$\pi_n(K) \in \arg \max_A Q_n(K,A) \quad (31)$$

完成从 Bellman 最优性结构到具体决策策略的连接。

### 3 具身感知策略设计与实现

在前两节中，具身频谱感知已被建模为知识状态空间上的序贯决策问题，并给出了满足 Bellman 结构的最优性方程。本节在此基础上，从方法与实现角度构建具身频谱感知策略设计框架，给出感知主体的闭环决策流程，如图 3 所示，结合多模型调制识别任务设计具体行动空间，并采用基于深度 Q 网络的值函数近似方法，将理论模型落地为可实现的决策方案。

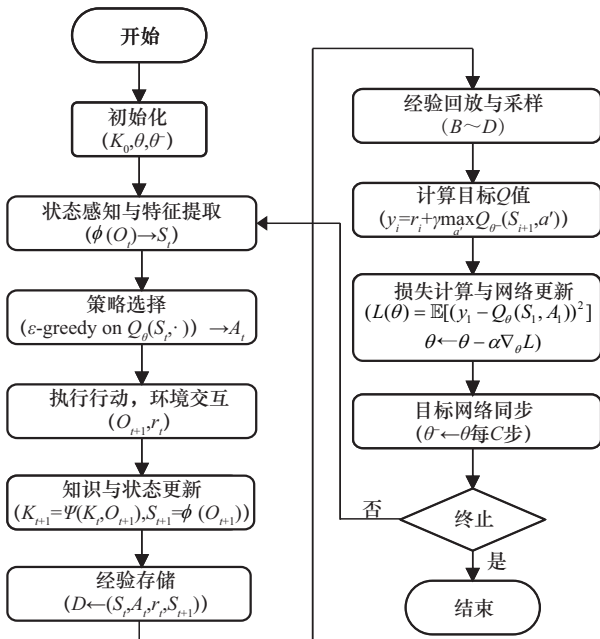


图3 具身频谱感知决策与知识更新流程

### 3.1 闭环决策流程

具身频谱感知主体在动态电磁环境中运行时，按照固定的感知决策闭环不断重复以下步骤。

1) 初始知识设定。系统启动时给定初始知识状态  $K_0$ ，该状态可以来自先验统计信息，也可以通过短时间预扫描得到。初始设定主要影响早期行为，在长期运行中将被持续更新的知识状态逐步修正。

2) 观测阶段。在时间步  $t$ ，主体根据上一时刻的知识状态生成行动  $A_t = \pi(K_t)$ ，确定当前的观测配置，如观测频段、采样时长以及是否启用复杂识别模型等。行动执行后，从电磁环境中获得观测  $O_t$ ，其统计特性由环境状态与当前行动共同决定。

3) 知识更新阶段。主体利用知识更新映射，将新观测与已有知识进行融合，得到更新后的知识状态  $K_{t+1}$ 。这一过程可以体现为模型参数的递推更新、统计量的滑动估计或隐状态的递归推断，目标是用紧凑的知识表示尽可能地保留对未来决策有用的信息。

4) 决策与行动生成阶段。在获得当前知识状态后，主体依据策略  $\pi$  生成下一步行动。在学习阶段中，策略会随着值函数估计不断调整；当运行收敛后，策略逐渐接近根据状态行动值函数近似实现的贪心结构，使行动选择与 Bellman 最优性条件保持一致。

5) 即时回报评估与样本记录。在完成一次观测、识别与决策后，根据识别结果与真实标签以及时延和资源开销，计算即时回报  $r_t$ ，并将  $(K_t, A_t, r_t, K_{t+1})$  记录为一条经验样本，用于后续值函数近似的更新。

通过在时间轴上反复执行上述步骤，具身频谱感知主体在动态环境中持续积累经验、修正知识状态，并在 Bellman 结构的指导下逐步改进策略，从而在长期运行中获得更优的频谱感知性能。

### 3.2 多模型调制识别策略

为具体展示具身频谱感知策略设计过程，考虑动态环境下的调制识别任务。感知主体需要对接收信号进行调制类型判决，并在不同信噪比、衰落程度和干扰条件下兼顾识别精度与决策时延。

设主体预先训练了若干结构与复杂度不同的识别模型，构成候选集合  $\mathcal{M} = \{M_1, M_2, \dots, M_K\}$ ，其

中既包含适用于高信噪比场景的轻量模型,也包含针对低信噪比和强衰落场景的复杂深度模型,各模型在计算开销与识别性能之间有不同折中。

在具身频谱感知框架下,行动不仅需要确定观测配置,还需要从候选模型集合中选择合适的识别路径。可以将行动表示为  $A_t = (m_t, c_t)$ , 其中,  $m_t \in \mathcal{M}$  表示当前采用的识别模型编号,  $c_t$  表示观测相关配置,如观测频段索引、采样时长以及是否进行重复检测等。

知识状态  $K_t$  综合刻画了与决策相关的历史信息,可以包括不同环境下各模型的历史识别表现、近期错误类型及其频度、当前信噪条件的统计估计以及观测开销等。策略在知识空间上选择  $A_t$ , 本质上是在不同识别路径和观测配置之间进行联合权衡,使长期折扣回报最大化。

在这一建模下,具身频谱感知任务被自然地嵌入多模型调制识别场景中,感知主体能够依据知识状态自适应选择识别模型与观测配置,实现对多种环境条件的动态匹配。

### 3.3 基于值函数近似的策略学习

在实际应用中,很难显式获得环境状态转移分布和观测分布的解析形式,值函数与状态行动值函数也难以通过解析方式直接求解。因此,本文采用基于经验样本的值函数近似方法,对 Bellman 方程进行数值实现,并据此构造近似最优策略。

具体地,在知识状态空间上引入参数化函数族  $\{Q_\theta(K,A)\}$ , 该函数族近似最优状态行动值函数  $Q^*(K,A)$ , 其中  $\theta$  表示待学习的参数向量。如图4所示,本文采用深度Q网络作为逼近器,将知识状态与行动映射到相应的值估计中。具身频谱感知主体在运行过程中不断生成经验样本  $(K_t, A_t, r_t, K_{t+1})$ , 通过最小化时序差分误差来更新网络参数,使  $Q_\theta(K,A)$  在样本分布上逐步逼近  $Q^*(K,A)$ 。

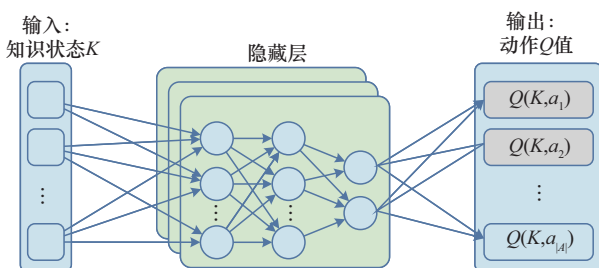


图4 DQN策略求解网络结构

策略选择方面,采用基于值函数的贪心结构,并在学习阶段加入适度探索,使主体在知识空间上既能利用当前值函数估计,又能持续发现更优的决策区域。随着经验样本数量的增加和值函数近似的收敛,策略将从以探索为主逐渐过渡为以利用为主。

为提升学习过程的稳定性与效率,实现中结合了经验回放和目标网络等常用机制。经验回放通过在缓冲区中存储历史样本并随机采样更新,降低相邻样本的时间相关性,改善梯度估计稳定性。目标网络通过缓慢更新用于构造时序差分目标的网络参数,减缓目标值随在线网络变化产生的振荡,使值函数近似在复杂动态电磁环境下更易收敛。

需要指出的是,DQN在动态环境中易受分布漂移影响而出现训练振荡,探索策略与超参数的合理设计对缓解该问题具有重要作用。探索策略方面,通常采用  $\epsilon$ -贪心机制并设置衰减策略,其设计需权衡两方面因素:初期探索率应足够高以覆盖多样化的环境状态与行动组合,后期则应逐步降低以利用已学得的价值函数估计,衰减速度需根据环境切换频率和状态空间规模进行调整。超参数设置方面,折扣因子  $\gamma$  的选取需平衡长期回报关注度与值函数估计方差,动态环境下通常不宜过高;学习率应配合衰减机制使用,避免环境切换时参数剧烈波动;目标网络更新频率需在估计稳定性与环境跟踪能力之间折中;经验回放缓冲区容量则应足以涵盖多种环境条件下的样本,以保证采样的代表性。上述设计原则可为具身频谱感知场景下的DQN训练提供参考。

## 4 实验与结果分析

### 4.1 实验设置

仿真实验基于公开数据集RML2016.10a。该数据集包含11种典型调制类型,分别是八相位移键控(8PSK)、双边带调幅(AM-DSB)、单边带调幅(AM-SSB)、二进制相移键控(BPSK)、连续相位频移键控(CPSK)、高斯频移键控(GFSK)、四电平脉冲幅度调制(4PAM)、16阶正交幅度调制(16QAM)、64阶正交幅度调制(64QAM)、四相位移键控(QPSK)和宽带调频(WBFM)等,覆盖从-20 dB至18 dB的宽信噪比区间,能够较好地反映实际电磁环境中多体制信号

并存的情况。

为体现电磁环境的动态性与复杂性，仿真中构造了 10 种不同复杂程度的测试场景，从近似理想信道到极端恶劣环境逐级增强环境挑战程度。记环境集合为

$$\mathcal{P} = \{P_1, P_2, \dots, P_{10}\} \quad (32)$$

其中，第  $i$  个环境  $P_i$  由四元组  $(\Delta\text{SNR}_i, n_i, j_i, f_i)$  表征，分别表示在该环境下相对基准的信噪比衰减量、噪声因子、干扰强度和衰落因子，具体配置如表 1 所示，每一类环境对应不同的主要测试目标。动态环境切换场景示意如图 5 所示，系统在运行过程中可以在这些环境之间按照预设规则切换，用以模拟实际电磁环境中条件随时间变化的特性。

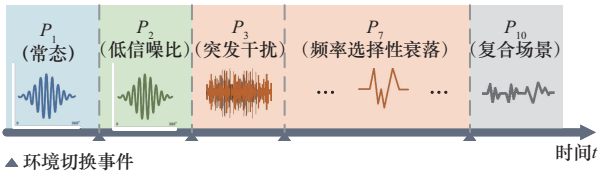


图 5 动态环境切换场景示意

#### 4.2 评估指标

为了全面刻画具身频谱感知策略的性能，本节在整体准确率、环境适应性、模型选择智能度等方面构建了多维度评估指标体系。

##### 1) 整体准确率

整体准确率定义为在所有环境与样本上的平均正确判决比例。设在环境  $P_i$  下有  $N_i$  个测试样本，第  $j$  个样本的真实标签和预测标签分别为  $y_{ij}$  和  $\hat{y}_{ij}$ ，则整体准确率定义为

$$\text{Acc} = \frac{\sum_{i=1}^M \sum_{j=1}^{N_i} \mathbb{I}(\hat{y}_{ij} = y_{ij})}{\sum_{i=1}^M N_i} \quad (33)$$

##### 2) 环境适应性

环境适应性定义为各环境准确率的标准差，即

$$\sigma_{\text{env}} = \sqrt{\frac{1}{M} \sum_{i=1}^M (\text{Acc}_i - \overline{\text{Acc}})^2} \quad (34)$$

其中， $\text{Acc}_i$  为每个环境的准确率， $\overline{\text{Acc}}$  为整体平均准确率， $\sigma_{\text{env}}$  越小，表示策略在不同环境下性能越均衡，适应性越强。

##### 3) 模型选择智能度

为刻画具身策略对候选模型的利用情况，统计在长期运行过程中各模型被选中的频率  $p_k$ ，并基于信息熵构造模型选择智能度

$$H_{\text{scl}} = - \sum_{k=1}^K p_k \text{lb} p_k \quad (35)$$

##### 4) 鲁棒性指标

为分析策略对环境复杂度变化的敏感性，引入环境挑战等级序列  $\{d_i\}_{i=1}^M$ ，并对  $(d_i, \text{Acc}_i)$  进行线性回归，得到回归斜率  $k_{\text{rob}}$ 。鲁棒性指标定义为

$$\kappa_{\text{rob}} = -k_{\text{rob}} \quad (36)$$

当环境复杂度增加时，准确率下降较缓，即斜率绝对值较小，则  $\kappa_{\text{rob}}$  较大，表示策略在复杂环境下具有更好的鲁棒性。

##### 5) 错误恢复能力

为评价具身策略在出现连续错误后恢复正确判决的速度，记录测试过程中所有连续错误序列的长

表 1 10 种测试环境参数

环境编号	环境名称	SNR 衰减/dB	噪声因子	干扰强度	衰落因子	主要测试目标
$P_1$	理想信道条件	0	1.0	0	1.0	基准性能、上限评估
$P_2$	低信噪比环境	-8	2.5	0.1	1.0	抗噪声性能
$P_3$	频率选择性衰落	-3	1.5	0.15	1.8	多径衰落适应性
$P_4$	频率偏移环境	-2	1.3	0.05	1.0	载频偏移容忍度
$P_5$	时序同步误差	-1	1.2	0.08	1.0	定时同步鲁棒性
$P_6$	相位噪声环境	-2	1.4	0.1	1.0	相位稳定性要求
$P_7$	IQ 不平衡环境	-1	1.3	0.12	1.0	硬件非理想适应
$P_8$	强干扰环境	-4	2.0	0.5	1.2	抗干扰能力
$P_9$	多径深度衰落环境	-6	1.8	0.3	2.2	复合衰落处理
$P_{10}$	极低信噪比环境	-10	3.0	0.4	2.5	极限工作能力

度。设一共有  $M$  段连续错误序列，第  $m$  段序列的长度为  $T_m$ ，则平均错误恢复时间为

$$\bar{T} = \frac{1}{M} \sum_{m=1}^M T_m \quad (37)$$

错误恢复能力定义为平均错误恢复时间的倒数。该指标越大，表示平均错误恢复时间越短，具身策略的自修正能力越强。

### 6) 置信度稳定性

为衡量识别结果置信度的时间稳定性，定义置信度稳定性为

$$S_{\text{conf}} = \frac{\mu_c}{\sigma_c + \varepsilon} \quad (38)$$

其中， $\mu_c$  和  $\sigma_c$  为测试过程中预测置信度的均值和标准差， $\varepsilon$  为防止分母为0的微小常数。该指标越大，表示策略在时间上表现更稳定。

上述指标从不同侧面反映具身频谱感知策略在动态环境下的综合表现，有助于从 Bellman 结构定义的最优性角度量化验证策略对“识别正确性-时延开销-环境适应性”多目标的平衡能力。

## 4.3 结果与分析

### 4.3.1 整体识别性能提升

图6展示了具身智能模型与静态模型的整体平均准确率对比。具身智能模型实现了49.27%的平均准确率，相比静态模型的44.22%提升了5.05%，相对提升幅度达11.4%。这一整体性能提升在调制识别领域具有重要的实际意义。

整体性能的显著提升主要源于具身智能模型的自适应能力。通过实时感知环境状态变化并动态调整模型选择策略，具身智能模型能够在各种复杂条

件下保持相对稳定的识别性能，避免单一静态模型在特定环境下性能急剧下降的问题。该结果表明，具身决策框架成功地将多个异构模型的互补优势整合为统一的智能决策系统。

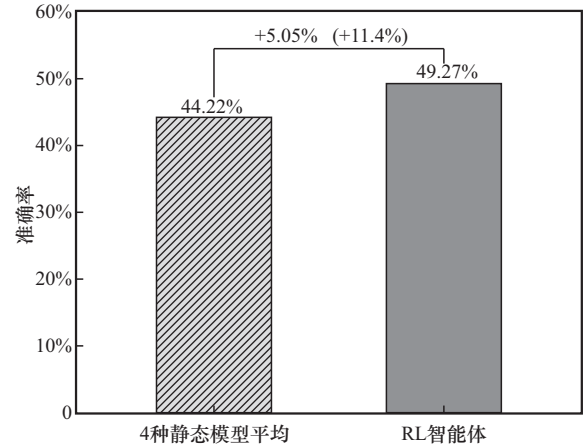


图6 具身智能模型与静态模型的整体平均准确率对比

### 4.3.2 环境适应性分析

环境特异性的准确率对比如图7所示。从图7可以看出，具身智能模型在大部分测试环境中都实现了显著的性能提升。在理想环境  $P_1$  中，具身智能模型达到了84.0%的识别准确率，相比静态基准的73.7%提升了10.3%，展现了在最优条件下的性能上限。

在中等挑战环境  $P_4 \sim P_7$  中，具身智能模型保持了相对稳定的性能优势，提升幅度为4.0%~6.2%。特别是在频率选择性衰落环境  $P_7$  中，具身智能模型准确率达到71.8%，相比基准提升了5.8%，表明系统能够针对多径衰落特性动态调整识别策略。

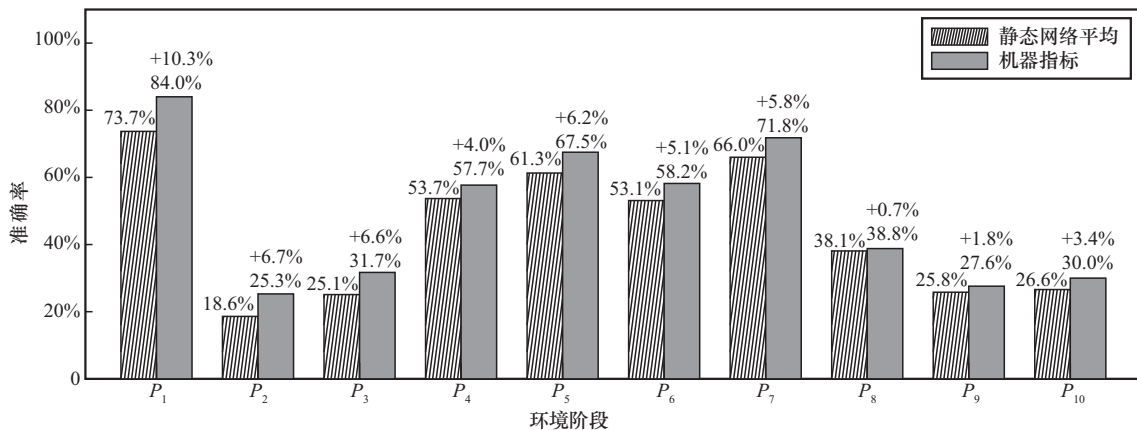


图7 环境特异性的准确率对比

在高挑战环境  $P_8 \sim P_{10}$  中，虽然绝对性能水平受到严重信道损伤的限制，但具身智能模型仍能实现 0.7%~3.4% 的性能改善。值得注意的是，具身智能模型在低信噪比环境  $P_2$  中表现出了 6.7% 的显著提升，这主要得益于其能够动态识别并调用具有较强抗噪能力的模型。

### 4.3.3 多维性能对比

图 8 的综合性能雷达图从 6 个维度全面刻画了具身智能模型相对于静态基准的性能优势。结果显示，具身智能模型在所有维度上都实现了不同程度的性能提升，呈现出较均衡的能力分布特征。

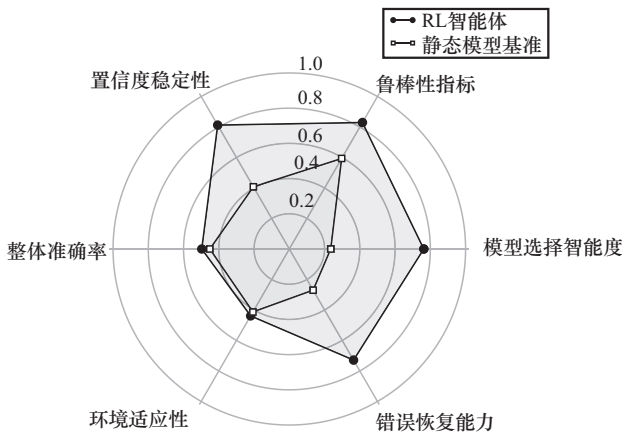


图 8 综合性能雷达图

在模型选择智能度维度上，具身智能模型表现最突出，显著超越静态基准，验证了具身智能模型在学习复杂决策策略方面的有效性。在鲁棒性指标和环境适应性维度上，具身智能模型也展现出明显优势，表明通过与环境的持续交互学习，具身智能模型获得了更强的环境适应能力。

置信度稳定性的提升反映了具身智能模型决策的一致性和可靠性。通过多目标奖励机制的引导，具身智能模型不仅追求识别准确性，还注重决策的稳定性，避免了过度频繁的模型切换。错误恢复能力的显著改善体现了具身智能系统的自我修正特性，具身智能模型能够从决策失误中快速学习并调整策略。

### 4.3.4 错误恢复能力分析

图 9 展示了具身智能模型错误恢复能力的详细分析结果。统计数据表明，具身智能模型的平均错误恢复时间为 2.4 步，这一指标在实时应用场景中具有重要意义。

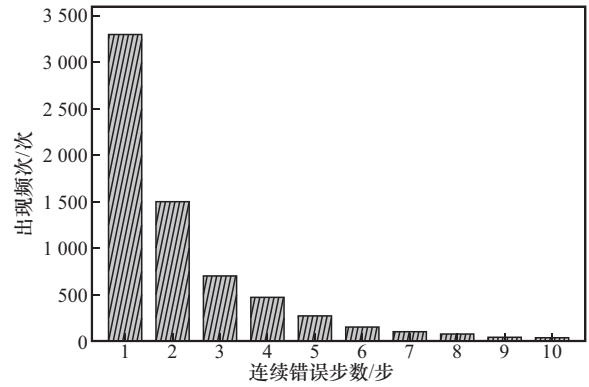


图 9 具身智能模型错误恢复能力的详细分析结果

从连续错误步数的分布特征可以看出，单步错误占绝对主导地位，出现频次超过 3 300 次，且连续错误的发生概率随着步数增加呈指数衰减趋势，连续 2 步错误约为 1 500 次，连续 3 步错误约为 700 次，连续 4 步及以上错误则相对罕见。这一分布模式表明具身智能模型具备良好的错误抑制和快速恢复能力。

短平均恢复时间的实现主要得益于多目标奖励机制中的即时反馈、注意力机制增强的状态表示以及经验回放机制的策略优化。连续错误的快速衰减特性表明具身智能模型具备有效的错误传播抑制能力，这种特性对维持系统的稳定运行具有重要价值。

## 5 结束语

本文围绕动态电磁环境下的具身频谱感知问题，从理论建模、策略求解和实验验证 3 个层面开展了系统研究。首先，构建了包含环境状态、观测、知识表示与具身行动的统一感知决策架构，将频谱感知过程抽象为以知识状态为核心的序贯决策问题，明确了观测、知识更新与行动之间的闭环关系，为具身频谱感知的整体分析提供了统一理论基础。其次，在折扣回报评价准则下引入值函数与状态行动值函数，推导了满足 Bellman 结构的最优性方程，给出了具身频谱感知策略的状态依赖形式与基本性质，并提出了基于值函数更新的近似最优策略构造方法。再次，结合多模型调制识别任务设计具身感知决策流程，将模型选择与观测配置联合纳入行动空间，并基于深度 Q 网络实现状态行动值函数近似与策略学习。仿真结果表明，所提方案在典型动态电磁环境下能够有效整合多种识别模型的互

补优势,在整体识别准确率、环境适应性、模型选择智能度、错误恢复能力和置信度稳定性等多个维度均优于静态单模型基准,同时在时延与观测开销方面保持较为可接受的水平,验证了具身感知决策框架以及基于 Bellman 结构的策略求解方法的有效性与实用价值。

未来工作中,一方面,可进一步结合更丰富的知识表示形式和不完全观测条件,探索更贴近实际系统约束的具身频谱感知建模方法,完善对感知开销、安全约束等因素的刻画;另一方面,本文重点关注动态环境下的识别性能与适应能力验证,对计算开销、决策时延及能耗等效率指标的系统评估尚不充分,后续研究将针对资源受限的实际部署场景,探索轻量化网络结构、知识蒸馏以及增量式策略更新等方法,在保持环境适应性优势的同时降低计算复杂度;此外,有望将本文方法推广到多智能体协同感知场景与更复杂的电磁任务体系,引入迁移学习、元学习等机制提升跨场景泛化能力,并在软硬件一体化平台上开展在线闭环实验验证,推动具身频谱感知技术向工程应用落地。

## 参考文献:

- [1] 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 T C, Wu Q H, Zhao S J, et al. Cognitive learning: a new paradigm for machine learning in electromagnetic spectrum environment[J]. Acta Electronica Sinica, 2023, 51(6): 1430-1442.
- [2] 李泓余,韩路,李婕,等. 电磁空间态势研究现状综述[J]. 太赫兹科学与电子信息学报, 2021, 19(4): 549-555, 595.  
Li H Y, Han L, Li J, et al. A summary of the present situation of electromagnetic space situation research[J]. Journal of Terahertz Science and Electronic Information Technology, 2021, 19(4): 549-555, 595.
- [3] 杨学军,苏金树. 关于我国网络电磁空间安全战略的思考[J]. 国防科技, 2010, 31(4): 1-3.  
Yang X J, Su J S. Security strategy of network and electromagnetism in China: perspectives and suggestions[J]. National Defense Science & Technology, 2010, 31(4): 1-3.
- [4] 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 F, Ding G R, Li J, et al. Research progress on electromagnetic spectrum multidimensional situation compressed mapping technology[J]. Journal on Communications, 2023, 44(11): 25-42.
- [5] Su Z Y, Teh K C, Xie Y H, et al. Signal enhancement aided end-to-end deep learning approach for joint denoising and spectrum sensing[J]. IEEE Transactions on Vehicular Technology, 2024, 73(3): 4424-4428.
- [6] 沈甜雨,陶子锐,王亚东,等. 具身智能研究的关键问题:自主感知、行动与进化[J]. 自动化学报, 2025, 51(1): 43-71.  
Shen T Y, Tao Z R, Wang Y D, et al. Key problems of embodied intelligence research: autonomous perception, action, and evolution[J]. Acta Automatica Sinica, 2025, 51(1): 43-71.
- [7] Ren L, Dong J B, Liu S, et al. Embodied intelligence toward future smart manufacturing in the era of AI foundation model[J]. IEEE/ASME Transactions on Mechatronics, 2025, 30(4): 2632-2642.
- [8] Huan X T, Wu C F, Lei Y L, et al. LLRF: towards long-term LoRa radio frequency fingerprint identification based on transfer learning[J]. IEEE Transactions on Vehicular Technology, 2025, 74(12): 19921-19926.
- [9] Sun L, Xue R, Zha H R, et al. AFD-IL: a long-term incremental learning approach with adaptive feature distillation for specific emitter identification[J]. IEEE Transactions on Cognitive Communications and Networking, 2026, 12: 2769-2781.
- [10] Huang S N, Guo L T, Fu X, et al. Open-set specific emitter identification leveraging enhanced metric denoising autoencoders[J]. IEEE Internet of Things Journal, 2025, 12(4): 3453-3462.
- [11] 查浩然,刘畅,王巨震,等. 面向无人机辐射源个体识别的域适应模型设计[J]. 信号处理, 2024, 40(4): 650-660.  
Zha H R, Liu C, Wang J Z, et al. Design of domain-adaptation model for specific emitter identification of UAV signal[J]. Journal of Signal Processing, 2024, 40(4): 650-660.
- [12] Bhattarai S, Park J J, Gao B, et al. An overview of dynamic spectrum sharing: ongoing initiatives, challenges, and a roadmap for future research[J]. IEEE Transactions on Cognitive Communications and Networking, 2016, 2(2): 110-128.
- [13] 杨玉好,袁国伦. 主动感知下的集群车辆数据交换网络通信协议栈优化算法[J]. 计算技术与自动化, 2025, 44(2): 15-19.  
Yang Y H, Yuan G L. Optimization algorithm for communication protocol stack of cluster vehicle data exchange network under active perception[J]. Computing Technology and Automation, 2025, 44(2): 15-19.
- [14] 金立民,王海超,谷江春,等. 低空具身智能频谱管控技术研究[J]. 数据采集与处理, 2025, 40(1): 45-55.  
Jin L M, Wang H C, Gu J C, et al. Research on low altitude embodied artificial intelligence enabled spectrum management and control technology[J]. Journal of Data Acquisition and Processing, 2025, 40(1): 45-55.
- [15] 李颂元,朱祥维,李玺. 基座模型技术背景下的具身智能体综述[J]. 浙江大学学报(工学版), 2025, 59(2): 213-226, 268.  
Li S Y, Zhu X W, Li X. Survey of embodied agent in context of foundation model[J]. Journal of Zhejiang University (Engineering Science),

2025, 59(2): 213-226, 268.

[16] 王文晟, 谭宁, 黄凯, 等. 基于大模型的具身智能系统综述[J]. 自动化学报, 2025, 51(1): 1-19.

Wang W S, Tan N, Huang K, et al. Embodied intelligence systems based on large models: a survey[J]. Acta Automatica Sinica, 2025, 51(1): 1-19.

[作者简介]



查浩然 (1996-), 男, 安徽枞阳人, 哈尔滨工程大学博士生, 主要研究方向为射频指纹、信号处理。



孙露 (2000-), 女, 吉林长春人, 哈尔滨工程大学博士生, 主要研究方向为信号处理、特定辐射源识别、增量学习。



尚佳颖 (2002-), 女, 河南平顶山人, 哈尔滨工程大学硕士生, 主要研究方向为调制识别、强化学习、电磁具身感知等。



王美玉 (1992-), 女, 黑龙江鹤岗人, 博士, 杭州电子科技大学特聘副研究员, 主要研究方向为数据受限下特定辐射源识别、电磁大数据挖掘、射频域迁移学习等。



林云 (1980-), 男, 黑龙江哈尔滨人, 博士, 哈尔滨工程大学教授、博士生导师, 主要研究方向为智能无线电技术、人工智能和机器学习、大数据分析挖掘、软件和认知无线电、信息安全与对抗、智能信息处理。