

群智网络：从物联网到智联

孙卓¹, 於志文^{1,2}, 郭斌¹, 李玉¹, 杨博¹

(1. 西北工业大学计算机学院, 陕西 西安 710072; 2. 哈尔滨工程大学计算机科学与技术学院, 黑龙江 哈尔滨 150001)

摘要: 随着具身智能从单体自主向群体协作演进, 群具身智能体系统通过信息交互和协作执行, 在复杂开放环境中展现出超越个体能力之和的群体智能。然而, 现有移动通信网络和物联网 (Internet of things, IoT) 分别以提升总速率和扩大连接设备数量为核心目标, 难以支撑基于“感知-交互-决策-行动”闭环的群具身智能体系统在开放物理世界中协作执行任务。因此, 研究以具身智能体为节点、由物理世界任务驱动的群智网络, 是实现群具身智能体系统的核心方向。首先, 阐述了群具身智能体系统的主要特点与关键技术, 并提炼其对通信网络的核心需求, 包括物理世界任务驱动、高价值语义交互、拓扑敏捷重构以及感通算控协同等。其次, 梳理了移动通信网络和物联网的发展概况, 分析了其支撑群具身智能体系统的局限性。在此基础上, 提出了面向群具身智能体系统的群智网络, 明确了语义驱动的轻量化通信、任务为中心的网络调度、自组织的动态拓扑管理以及资源自适应的感通算控协同增强等主要特性。对于群智网络, 从基础效能和适应性能的角度, 构建了由协作任务完成率、闭环响应时长、群体能量效率、协作重构时效与性能保持率和开放场景泛化性组成的评价指标体系, 并阐明了在实际网络设计优化时指标间的折中关系。从安全可信、异构智能体互操作、网络规模扩展、实时控制稳定性以及标准化 5 个方面, 提出主要挑战。最后, 以空地具身智能体协作追踪与群具身智能体协作导航为典型应用场景, 结合生物启发的空地低时延通信和基于信息瓶颈的具身智能体语义通信等代表性研究进展, 初步验证了群智网络中关键技术的可行性。该研究旨在为群智网络的理论研究与产业实践提供系统性的参考框架。

关键词: 群具身智能体系统; 群智网络; 关键特性; 评价指标

中图分类号: TP393

文献标志码: A

doi: 10.11959/j.issn.2096-3750.WLW26082

Embodied swarm-agent network: from things connection to embodied agents collaboration

Sun Zhuo¹, Yu Zhiwen^{1,2}, Guo Bin¹, Li Yu¹, Yang Bo¹

1. School of Computer Science, Northwestern Polytechnical University, Xi'an 710072, China

2. College of Computer Science and Technology, Harbin Engineering University, Harbin 150001, China

Abstract: As embodied intelligence evolves from individual autonomy to collective collaboration, the embodied swarm-agent system demonstrates collective intelligence that surpasses the sum of individual capabilities through multi-agent information interaction and cooperation in complex open environments. However, existing mobile communication networks and Internet of things (IoT), which focus on enhancing the sum rate and the number of connected IoT devices, respectively, struggle to support the loop “perception-interaction-decision-action” based embodied swarm-agent system in collaboratively executing tasks within open physical environments. Therefore, investigating the task-driven embodied swarm-agent network, where embodied agents serve as nodes, represents a core direction for realizing embodied swarm-agent

收稿日期: 2026-06-22; 修回日期: 2026-08-14

通信作者: 於志文, zhiwenyu@nwpu.edu.cn

基金项目: 国家重点研发计划项目 (No. 2024YFB4505502); 国家自然科学基金资助项目 (No. 62532009); 陕西省重点研发计划项目 (No. 2025GH-YBXM-049)

Foundation Items: The National Key Research and Development Program of China (No. 2024YFB4505502), The National Natural Science Foundation of China (No. 62532009), The Key Research and Development Program of Shaanxi (No. 2025GH-YBXM-049)

systems. Firstly, the core characteristics and key technologies of embodied swarm-agent systems were presented, and their key requirements for communication networks were extracted, including physical-world task orientation, high-value semantic interaction, agile topology reconfiguration, and sensing-communication-computation-control coordination. Secondly, the development of mobile communication networks and IoT was reviewed, and their limitations in supporting embodied swarm-agent systems were analyzed. On this basis, the embodied swarm-agent network was proposed to support embodied swarm-agent systems. Its core features were defined, including semantic-driven lightweight communication, task-centric network scheduling, self-organized dynamic topology management, and resource-adaptive coordination enhancement of sensing, communication, computation, and control. For the embodied swarm-agent network, from the perspectives of basic efficiency and adaptability, a performance evaluation system was systematically constructed, comprising the collaborative task completion rate, closed-loop response latency, swarm energy efficiency, collaborative reconfiguration timeliness, and performance retention rate, as well as open-scenario generalization capability. Furthermore, the trade-offs among these metrics were clarified for practical network design and optimization. Moreover, the key challenges were identified in five aspects: security and trustworthiness, heterogeneous agent interoperability, network scalability, real-time control stability, and standardization. Finally, using air-ground embodied agent collaborative tracking and embodied swarm-agent collaborative navigation as typical application scenarios, preliminary verification of the feasibility of two key technologies was provided, such as bio-inspired air-ground low-latency communication and information bottleneck-based embodied agent semantic communication, in the embodied swarm-agent network. This paper aims to provide a systematic reference framework for both theoretical research and industrial practice of embodied swarm-agent networks.

Key words: embodied swarm-agent system, embodied swarm-agent network, key characteristics, performance evaluation metrics

0 引言

依托物理形态在真实环境中持续感知、交互与学习,与物理世界深度融合的具身智能体时代已经到来。利用多模态大模型作为认知决策单元、灵巧机械臂与高动态腿足作为执行载体,具身智能体逐渐从实验室走向真实物理世界^[1-3]。

在现实场景中,很多任务具有广区域、长时程和高动态等复杂特点,例如,广域空间的搜索救援和动态环境中多机对抗博弈等。受限于感知视野、计算资源和物理操作能力,单具身智能体难以高效独立地完成这些复杂任务。因此,类似于人群协作,具身智能体需要从孤立的个体走向高度协作的群体,共同完成任务。通过信息交互达成共识、规划决策实现协作执行以及知识共享完成协同演进,形成群具身智能体系统^[4-6]并涌现出超越个体能力之和的群体智能。

对于群具身智能体系统,高可靠低时延的关键信息交互是实现具身智能体之间共识达成和协同决策的重要基础。不同于智能手机和传感设备,群具身智能体的信息交互以协作完成任务为目标,可与环境感知、行为决策和执行构成闭环。现有的移动通信网络和物联网等无线网络主要作为被动且通用

的数据搬运管道,仅关注总速率、吞吐量等性能指标。这导致现有无线网络存在任务需求感知能力弱、传输信息价值密度低、感知计算控制分离、拓扑自组织重构能力弱等突出问题,难以支撑群具身智能体系统协同任务的高效执行。为此,面向具身智能体的通信网络技术被提出^[7-9]并被广泛关注。例如,文献[7]设计了面向多具身智能体协同感知的务实通信算法,通过优化通信对象选择、信息选择和表示,在保证协同感知性能的同时最小化通信成本;文献[8]提出了融合实时环境状态追踪的语义通信方法,减小通信时延对多具身智能体协同的影响。尽管相关研究已取得了较好的成果,但大多数研究集中在语义通信算法,没有系统地分析面向群具身智能体的通信网络特性,且缺乏相关性能指标的详细讨论。

本文旨在通过分析群具身智能体系统特点和现有无线网络的局限性,提出面向群具身智能体的通信网络,即群智网络,并探讨其特性和性能指标,为未来构建支撑群具身智能体协同感知、决策和执行的通信网络系统提供理论基础。本文的主要贡献包括以下3方面。

(1) 从群具身智能体系统特点和现有无线通信网络的发展概况两个维度,分析现有无线通信网络

在群具身智能体场景中的主要挑战，以论证群智网络设计的必要性。

(2) 提出了群智网络，从通信和网络两方面分析了主要特性，并设计了关键性能指标，包括基础效能和适应性能，探讨了两个指标在实际系统中的折中关系。

(3) 以空地具身智能体协作追踪与群具身智能体协作导航为典型应用场景，介绍了生物启发的空地低时延通信和基于信息瓶颈的具身智能体语义通信等代表性研究进展，验证了群智网络的有效性，为学术界和工业界推动群智网络的发展提供了理论参考与实践指引。

1 群具身智能体系统

本节将针对群具身智能体系统，介绍主要特性和关键技术，并阐述其对通信网络性能的核心需求。

随着具身智能体和多智能体协同技术的发展，具身智能体的研究逐渐由单智能体的主动感知、自主决策与行动，扩展到多个具身智能体之间的交互和协同，以完成现实世界开放环境中的复杂任务。2025年，中国人工智能学会发布《中国人工智能学会系列白皮书——具身智能》^[10]，并将“群体具身智能”列为具身智能十五大重点方向之一。

群具身智能体系统是由多个具备物理实体形态、主动环境感知与自主行动能力的具身智能体组成的分布式协作系统，群具身智能体系统示意如图1所示。在该系统中，各具身智能体通过自适应的关键信息交互，形成对环境态势与任务目标的共同认知。在此基础上，各具身智能体利用分布式规划决策，动态分配并协同执行任务。此外，群具身智能体能够共享经验和迁移知识，实现能力的协同演进，提升系统的自适应性及任务适应性。这一过

程构成了“感知-交互-决策-执行-演进”的完整闭环，使群具身智能体系统能够涌现出超越个体能力之和的群体智能，在复杂开放环境中高效、强鲁棒性地完成长时程协同任务。

1.1 群具身智能体系统特点

传统多智能体系统^[11-12]通常以多智能体强化学习为核心框架，由多个智能体在共享或部分共享的环境中，通过与环境交互及相互间策略博弈，学习如何最大化累积奖励。相比于传统多智能体系统，群体具身智能体系统具有更强的物理性、主动性、异质性和通用性，具体表现为如下4个特点。

1.1.1 异构物理具身与动态协作拓扑

在群具身智能体系统中，各智能体可以是无人机、无人车、机械臂、四足机器人等多样化的载体，它们在物理具身形态、感知模态、算力水平和机动性能等物理属性方面存在显著的差异^[13-14]。具身智能体的形态异构性是实现其群体协同的重要基础。面向复杂开放场景，通常需要依靠异构具身智能体的能力优势互补，协同完成任务。

在任务执行过程中，由于具身智能体在物理空间中的移动性，群具身智能体的相对位置持续变化。同时，具身智能体的感知、计算和运动能力也会随环境实时变化。这些空间和能力等物理约束都将导致具身智能体间的协作拓扑动态变化。例如，在搜索救援场景中，当具身智能体观测到关键目标或发生故障时，群具身智能体的任务分工和协作拓扑会发生变化^[15-16]。

1.1.2 任务式耦合交互

具身智能体间的交互包括以任务为中心的信息交互和动力学驱动的形态耦合。以任务为中心的信息交互是指具身智能体间交互的信息内容、时机与方式均根据协作任务的目标与需求主动生成，以最大化协作任务性能^[17]。具体来说，具身智能体不再盲目地交换原始感知信息，而是基于实时任务状态，主动提取并传输与任务需求最相关的感知特征、意图及其置信度（如“前方障碍物类别及危险程度”“我即将执行的动作序列”），从而在有限带宽下实现高价值信息的定向交互。

动力学驱动的形态耦合是指群具身智能体在物理行为协作过程中，通过实时感知与响应彼此的相对位姿和力学状态，主动调整自身动作与运动模

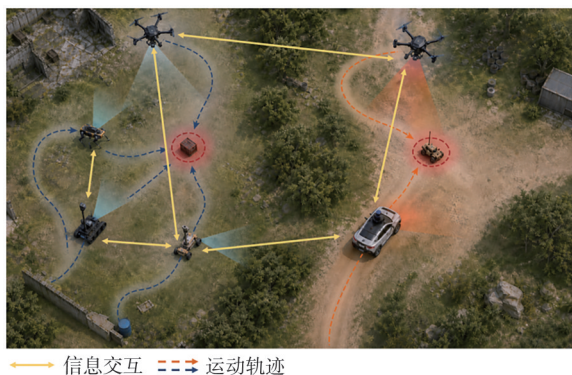


图1 群具身智能体系统示意

式,从而实现物理形态上的动态适配与协同。这种形态耦合表现为具身智能体之间在空间与时间维度的动作协同,是一种自组织、高鲁棒性的物理协同机制,使群具身智能体系统具备在复杂开放环境中执行精密操作与机动变形的能力。

1.1.3 群智感知-决策-行动闭环

在任务完成的过程中,群具身智能体系统持续执行由协同感知、分布决策与群体行动构成的动态闭环。具体来说,在协同感知阶段^[18-20],多个具身智能体通过空间分布与异构感知能力的互补,实时融合各智能体的局部感知信息,对复杂环境进行多视角、多尺度的联合观测,从而形成对复杂开放环境的增强感知和统一认知。根据环境认知和任务目标,群具身智能体通过分布式决策算法进行任务分配与行为规划^[21-22]。在群体执行阶段,各具身智能体根据决策结果,执行相应的物理动作。环境因群体行为而发生变化,这种变化又被群具身智能体协同感知,触发新一轮环境认知的更新,从而形成“感知-决策-行动”的紧耦合循环。在此过程中,群具身智能体与环境的持续交互不仅使其能够自适应地应对动态障碍、未知地形等突发变化,更通过长期反复的闭环迭代,逐渐内化环境规律并优化协同策略,最终实现超越个体能力之和的鲁棒、高效群体智能。

1.1.4 主动协同自成长

主动协同自成长是指群具身智能体系统能够自发地根据任务需求与环境变化,动态调整任务分工和跨具身地迁移能力,实现从“被动响应”到“主动进化”的转变。主动协同自成长包括主动自组织的能力协同和跨具身的协同演进。在主动自组织的能力协同^[23]方面,群具身智能体实时评估当前环境中各智能体的感知、计算和运动等状态,分布式协商出最优的分工组合,实现灵活的能力互补。

在跨具身的协同演进方面,群具身智能体能够在多轮交互与持续任务执行过程中逐步积累经验和演进,通过经验共享与技能迁移实现能力的跨具身演进^[24]。具身智能体能够将其学习到的新技能编码为形态无关的语义抽象知识,并在其他智能体请求时无损传递,使其在不重复试错的前提下快速掌握该技能。通过主动自适应地拓展能力边界,群具身智能体能够相互帮助,成为持续自成长的智能集群。

1.2 群具身智能体系统关键技术

为了应对高动态、不确定的开放物理环境,群具身智能体系统需要突破以下4项关键技术。

1.2.1 精准聚焦的协同感知与环境理解

作为群具身智能体系统实现高效协作的感知基础,精准聚焦的协同感知与场景理解技术旨在以协作任务为牵引,动态识别并交互与任务目标高度相关的语义信息,获得开放环境的增强感知和共识理解。

在群具身智能体系统中,各具身智能体携带异构传感器,形成环境的多视角多模态覆盖。设计按需通信机制,实时评估每个具身智能体感知数据对当前任务及邻居智能体决策的信息贡献度^[7,25],选择并传输能够填补邻居智能体感知盲区且对联合决策最具价值的局部观测^[26-27]。

对于来自不同具身智能体的异构感知数据,需在时空尺度、坐标系和模态语义上进行对齐,从而实现多视角多模态感知数据的融合,生成更完整精确的场景模型^[28-29],包括时间同步校准^[30]、空间坐标变换^[31]和模态语义对齐^[32]。

不同于以识别环境中所有物体语义信息为目标的传统场景理解,任务驱动的场景理解主动压缩信息空间^[33],仅聚焦于与当前协作任务直接相关的语义信息。通过将全局场景理解转化为局部、动态、任务相关的语义解析^[34-35],感知系统能够以最小的计算与通信开销,为群体具身智能体的分布式决策提供精准、实时、低冗余的数据基础。

1.2.2 主动自适应的任务解析与分配

针对现实世界中长时序、多步骤的复杂任务,系统需要对任务指令进行解析,理解任务目标、执行条件和约束关系,将整体任务分解为若干个可执行的子任务序列,并根据具身智能体的位置分布和异构能力进行合理分配。面向群具身智能体系统,主动自适应的任务解析与分配的内涵在于让群体从被动接收指令转变为主动理解意图且动态适配任务,包括基于大语言模型的任务指令解析、子任务序列的自适应生成,以及基于在线评估的协商重调度技术。

对于模糊、非结构化的任务指令,可利用大语言模型作为指令解析器^[36-37]。利用大语言模型强大的语义理解与常识推理能力,将自然语言指令转化为机器可执行的逻辑表示。首先,解析器识别指令中的核心动作、目标对象、约束条件与优先级信

息。之后，结合群具身智能体的感知精度、计算能力、运动速度等能力列表，将模糊指令拆解为具有时序依赖关系的子任务序列，并利用有向无环图等显示建模子任务之间的依赖关系。

基于子任务序列的时序依赖关系、空间约束和具身智能体的能力约束，确定初始任务分配方案，以确保在满足任务执行时空逻辑的情况下，每个子任务都被分配给最适合的具身智能体组合。在实际执行过程中，初始任务分配会因场景变化而无法继续执行，如具身智能体突发故障或任务目标因环境变化而调整等。为此，设计在线评估与协商重调度机制^[24, 38]，使群具身智能体具备持续自适应的能力。在任务执行过程中，各具身智能体持续上报自身的任务进度、状态参数和环境感知变化。通过轻量级的在线评估模块，实时计算当前分配方案的适配度。一旦检测到适配度下降超过阈值或收到智能体主动发起的重协商请求，群具身智能体将启动分布式协商过程，提出重新分配候选方案，并通过共识机制达成一致，实现任务分工的动态调整。整个重调整过程完全由群具身智能体自主完成，实现了任务在动态环境中的自适应拆解、在线评估与主动重分配闭环。

1.2.3 前瞻稳健的分布式决策

前瞻稳健的分布式决策是群具身智能体在开放环境中安全高效执行任务的关键。包括前瞻性决策推演技术、高鲁棒性的分布式决策技术及其在线适应机制。针对物理世界中试错成本高的问题，构建轻量级的具身世界模型^[39]，即一个能够在具身智能体本地运行、高保真模拟物理环境动态及智能体交互后果的神经预测器。在执行动作之前，各具身智能体会将其决策的动作输入该具身世界模型，推演该动作在未来时间窗口内对环境状态、其他智能体状态和任务进度的影响。通过影响评估，具身智能体能够预判式地选择并执行在预期收益大且风险低的决策序列。

传感器噪声或位置遮挡导致的感知误差以及无线通信的时延和丢包，是分布式决策面临的两大现实挑战。为此，设计高鲁棒性的分布式决策方法，使群具身智能体在实际的感知和通信条件下依然能够维持协作的有效性。考虑通信时延导致的信息过时和决策时延问题，构建时延感知的马尔可夫决策模型，并设计时延鲁棒性的分布式决策方法^[40-42]。

此外，面向未知的环境和具身智能体，需要引入在线适应机制^[43]，使分布式决策方法能够边执行边学习。每个具身智能体执行决策后，能够实时收集环境反馈与任务结果，并将其作为训练样本在线更新其决策策略与世界模型参数。对于未知的具身智能体，观测其行为模式并推断意图与能力，并据此动态调整自身的协作策略，实现自组织的合作团队。

1.2.4 跨具身可迁移的知识共享

为提高群具身智能体在开放环境中的泛化能力，跨具身可迁移的知识共享通过具身形态解耦的任务语义表征、技能知识库共享和任务驱动语义通信技术，使具身智能体能够获得其他队友的已有知识，实现跨具身智能体的知识共享、技能复用和能力迁移。

作为知识共享的基础，具身形态解耦的任务语义表征方法将知识抽象为与具体执行器无关的、任务目标驱动的通用语义原语。异构形态的具身智能体通过解析该任务语义表征，理解任务需求并适配自身执行器能力。

存储大量经过验证的任务语义表征，以构建模块化技能知识库^[44]。当具身智能体遇到未处理过的新任务时，它会查询请求技能知识库。通过查询-响应机制，基于语义相似度匹配出最相关的技能模块，并将该技能的语义表征发送给请求的具身智能体。具身智能体结合自身的形态参数，将接收的通用语义表示实例化为其可执行的动作序列与感知策略。其中，通过传输高度压缩的语义信息包，低时延高可靠的任务驱动语义通信是实现群具身智能体实时查询知识和共享技能的关键。

由上述系统特点和关键技术可以看出，群具身智能体间的信息交互以协作完成任务为目标，与其协同感知、分布决策和知识共享等过程紧密关联。群具身智能体系统对通信网络的需求不仅是传统的总速率和吞吐量，更是以协作任务为牵引且支撑集群感知与决策闭环的语义理解能力、信息价值密度和确定性服务保障。

2 移动通信网与物联网发展概述

本节围绕移动通信网和物联网两类通信网络，分别介绍发展历程、功能定位和主要技术，并分析这两类通信网络的性能特点，以及群具身智能体系统对通信网络需求的差异。

2.1 移动通信网

移动通信网络是以人为中心、为移动终端提供无线接入、连续覆盖和稳定数据传输服务的重要通信基础设施^[45-46]。随着无线通信技术的发展,移动通信网络经历了从语音通信到移动数据通信,再到高速宽带接入和智能化网络服务的持续演变。

第一代移动通信(the first generation mobile communication, 1G)和第二代移动通信(the second generation mobile communication, 2G)网络主要面向语音通话和短信等基础通信业务,采用频分多址和时分多址技术,实现了移动通信的初步普及。第三代移动通信(the third generation mobile communication, 3G)网络引入码分多址技术,开始支持移动数据传输。第四代移动通信(the fourth generation mobile communication, 4G)以正交频分复用和多输入多输出为核心技术,显著地提升了宽带接入能力。进入第五代移动通信(the fifth generation mobile communication, 5G)后,除了传统的传输速率,移动通信网络开始关注时延和连接密度等多维度性能,形成了增强型移动宽带、超可靠低时延通信和大规模机器类通信3大应用场景。目前,移动通信网络的研究重点已转向第六代移动通信(the sixth generation mobile communication, 6G)技术。在追求更高传输速率、更低时延和更多连接数的基础上,6G将深度融合空天地一体化通信网络、通感一体化、内生智能网络等前沿技术,实现全球覆盖、多功能的智能移动通信网络。

尽管移动通信网络在传输速率和覆盖连续性等方面取得了显著提升,但其内涵仍然是以连接人为目标、被动通用的数据搬运管道。网络架构以基站为中心的静态或低动态结构为主,且资源调度多依赖基站集中式管理^[47]。因此,面对协作任务驱动、高动态的群具身智能体系统时,移动通信网络存在任务需求感知能力弱、时延高等问题。在上述定性描述基础上,本文考虑多无人机编队协同搜救场景,定量说明基站集中调度难以满足多无人机协同物理闭环的时延要求。具体来说,多架无人机同时产生上行数据并触发基站集中调度的总时延主要由调度请求排队时长、基站调度时长、上行数据传输时长、基站数据处理和路由决策时长、下行数据传输时长组成。根据3GPP标准^[48],计算可得当基站集中调度5架无人机时,端到端总时延在4.5 ms到

17.0 ms,并且端到端时延的抖动范围达12.5 ms。该时延难以满足多无人机协同作业响应时间小于10.0 ms的要求。

2.2 物联网

随着终端数量的迅速增长和感知需求的不断扩展,物联网出现并发展为未来互联网的重要组成部分^[49]。不同于移动通信网,物联网更关注传感设备的低成本接入、低功耗运行和感知数据的可靠收集,完成感知数据采集、传输和汇聚^[50-51]。以通信技术为驱动,物联网经历了从简单连接到万物智联的演进过程。随着2G/3G网络的发展,无线传感器网络和云计算技术推动了物联网环境感知与数据采集能力的提升。随着智能采集终端的出现,物联网开始具备移动计算与设备协同能力。在5G时代,大规模设备通信技术使物联网能够支持每平方公里百万级设备的接入。之后,物联网将由单一地面部署扩展至空、天、地跨域覆盖,深度融合通信、感知、计算与智能4个功能^[52-54]。

然而,群具身智能体系统需要执行的是面向物理世界的协作任务,包括协作搬运和协同搜索等。这使多个具身智能体在开放物理环境中执行感知、交互、决策与行动的紧耦合闭环,并且任务完成质量取决于行动成功率、任务执行时间和群体协调度。现有物联网技术缺乏将感知推理结果转化为群体物理行动并实时接收环境反馈的闭环能力。并且,物联网的资源调度以计算效率而非任务驱动为优先级,难以支撑群具身智能体在动态物理环境中的自组织协同与任务驱动的信息价值动态评估。

值得指出的是,随着通信技术的持续演进,现有移动通信网络和物联网也在不断发展。例如,5G网络已通过设备到设备(device-to-device, D2D)通信^[55]等技术在一定程度上突破了传统基站中心化的局限;物联网通过边缘智能^[56]和任务导向通信^[57],具备了一定的协同感知、协同学习和推理能力。然而,这些进展的核心设计目标仍以保障通信连通性、提升学习和推理精度为导向,尚未将物理动力学规律和时空约束作为协议栈的内生变量,难以具备支撑“感知-交互-决策-行动-演进”完整闭环的能力。因此,本文提出了群智网络,通过将物理动力学规律和时空约束融入协议栈中每层协议的设计,实现任务驱动下网络行为与物理演化的深度耦合与协同演进,支撑具身智能体在物理空间协作完成任务。

3 群智网络

本节重点介绍群智网络的主要特性，分析其关键性能指标。然后，从网络功能、结构和传输内容等方面，将群智网络与移动通信网和物联网进行比较。

3.1 群智网络定义

群智网络是一种以具身智能体为节点、协作任务驱动的自组织网络架构，群智网络愿景如图2所示。面向开放环境的群具身智能体系统，群智网络构建了一个以协作任务为中心的新型智能网络体系：各具身智能体节点实时感知任务需求，动态评估自身信息对邻居节点协同决策的价值，并据此决定传输信息的时机、对象和语义内容，实现智能体间高价值密度的语义信息交互；通过感知并预测网络状态，群具身智能体主动分布式地控制传输，将网络调度与协作任务深度耦合。



图2 群智网络愿景

不同于以总速率或吞吐量为目标的现有网络，群智网络的目标是通过信息交互，使群具身智能体在不确定环境中稳定高效地完成协作任务。为此，群智网络需要根据协同任务需求，将物理动力学规律和时空约束融入协议栈中每层协议的设计，使其参与物理环境变化的感知和预测，主动预判式地调整网络参数，实现任务驱动下网络行为与物理演化的深度耦合与协同演进。群智网络架构如图3所示，具体来说，群智网络由应用层、协同层和网络执行层构成。其中，应用层负责具身智能体的发现和连接，从而获得联网可用的具身智能体集合；在协同层，根据任务需求和具身智能体能力，给智能体分配角色，据此，建立具身智能体拓扑图；网络执行层根据任务需求和具身智能体拓扑图，将物理

动力学规律和时空约束落实到各层协议中。在物理层，利用信道状态信息的时变特性，感知收发智能体之间的相对距离变化率和多径环境中的遮挡物运动，预测未来信道状态，为上层拓扑预切换提供信息；在数据链路层，根据任务需求和物理动力学方程，确定信息有效时长，并据此设计调度策略，保证时延的确定性上界，最大化有效信息数量；在网络层，根据任务需求、节点位置和能力向量，确定目的地址和路由决策，并利用节点的速度、角速度等动力学参数，预测未来网络拓扑，实现路径的提前切换。三者共同支撑以任务为中心的网络调度。在语义层，发送智能体根据具身世界模型等物理环境模拟器，预测并传输任务关键的未来环境状态变化轨迹，接收智能体根据自身运动状态按需解码语义信息，从而实现语义驱动的轻量化通信。随着物理环境的变化，网络执行层状态反馈至协同层；协同层根据网络状态和任务执行阶段调整拓扑和联合资源调度策略，从而进一步影响网络执行层。

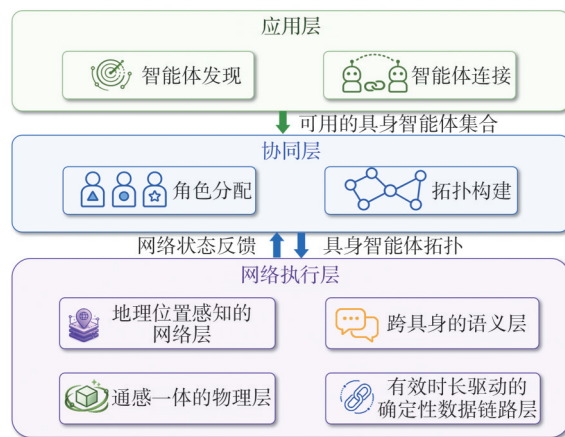


图3 群智网络架构

通过上述机制，群智网络不再是数据的搬运管道，而是支撑群具身智能体系统协同决策与行动的核心——面向异构智能体频繁移动、任务目标动态变化和高度不确定的开放场景，群智网络能够自适应地调整其拓扑结构与通信策略，与群具身智能体的感知、决策和行动形成闭环，保障实时高效地完成物理世界的协作任务。值得注意的是，群智网络以支撑具身智能体节点高效协作完成任务为目标，且节点具有异构的物理具身特性。这使相比于现有的目标导向语义通信技术，群智网络的语义层传输面临跨具身的语义适配、时延感知的语义预测和在线语义价值评估等独特挑战。

3.2 主要特性

为实现从被动承载到主动支撑的转变,群智网络从信息交互、网络调度、拓扑组织与资源协同4个维度重新定义了网络能力,主要特性可概括为4个方面,群智网络的主要特性如图4所示。

3.2.1 语义驱动的轻量化通信

语义驱动的轻量化通信特性的内涵在于群具身智能体间高价值密度信息的精准交互,使群智网络能够实现群具身智能体间极简语义的实时交互,支撑在复杂开放环境中达成共识和协作执行任务。它包括高价值密度的语义通信^[58]和跨具身的生成式语义通信^[5]。高价值密度的语义通信是指提取、编码和传输与协作任务相关且对邻居具身智能体节点决策有用的、高价值且小数据量的语义信息^[59],例如,目标位置变化、环境风险和行动意图等。在此基础上,由于感知模态、运动能力和任务分工上的显著差异,各具身智能体需要将接收的通用语义信息,结合自身形态特征与能力需求进行适配性生成与表达^[60],从而准确理解信息并充分用于行为决策,实现跨具身智能体的信息交互与协同行动。

3.2.2 任务为中心的网络调度

以任务为中心的网络调度,是指以协作任务完成性能为优化目标,群具身智能体能够分布式地动态调度其传输信息的时机、接收具身智能体集合以及优先级,围绕任务性能调度网络中的信息流动,从而使每一次传输调度都能够促进群具身智能体协作任务的推进和完成^[61-62]。

在群智网络中,根据任务当前阶段、具身智能体任务分工和环境状态,各具身智能体节点能够实时评估自身信息的任务价值,即对邻居节点协同决策的影响程度和对任务推进的贡献度,并将其转化为传输意图。高任务价值的信息会触发更强的传输意图^[42],

获得更高的调度优先级;低任务价值的信息则被降级处理或延迟发送,以避免占用有限的通信资源。

3.2.3 自组织的动态拓扑管理

自组织的动态拓扑管理是群智网络在开放环境中保持群具身智能体协作韧性的关键。它的内涵在于网络拓扑能够随着任务进展、具身智能体能力状态变化及其加入退出而自适应变化,包括任务/能力自适应的拓扑敏捷重构和智能体加入/退出的弹性拓扑管理。

根据协作任务的不同阶段以及具身智能体感知范围、计算能力和行动能力的变化,群具身智能体能够敏捷重构其网络拓扑结构,自组织地建立或重组通信关系。例如,当任务阶段需要信息汇聚时,群具身智能体能够形成有利于数据融合的汇聚拓扑;而当任务阶段需要群具身智能体紧密协同行动时,网络快速重构为支持低时延局部交互的连接拓扑。此外,在开放环境中具身智能体因任务需要、故障中断或位置移动而加入或退出系统,群智网络能够及时地感知具身智能体集合的变化,并自适应地调整拓扑连接关系。这种自组织的动态拓扑管理能力使群智网络在具身智能体节点频繁变动和任务动态演进的环境中能够始终维持通信的连续性,为群具身智能体的持续协作提供可靠支撑。

3.2.4 资源自适应的感通算控协同增强

在6G网络中通感一体和通算融合技术的基础上,群智网络能够根据任务状态、具身智能体的行动需求和可用的多维资源,分布式地联合优化感知、通信、计算和控制过程^[63-64],实现感通算控的协同增强,提高群具身智能体的任务协作性能^[65]。

在完成的任务过程中,群具身智能体需要通过协同感知以获得更全面的环境感知信息,协同计算以深度处理感知数据并获得环境的认知理解,以及

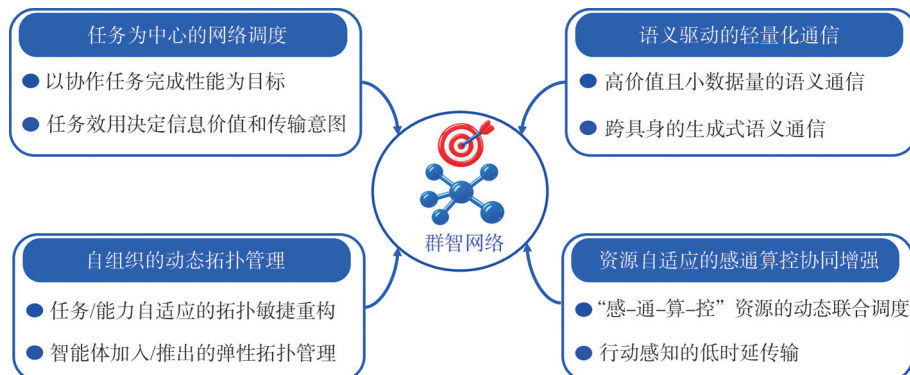


图4 群智网络的主要特性

协同决策以确定动作。协同感知、协同计算和协同决策过程构成不断迭代的闭环。同时，具身智能体间的通信是支撑协同感知、协同计算和协同决策实现的基础。因此，在群智网络中，群具身智能体分布式地联合优化感知、通信、计算和控制过程及其资源调度。当任务阶段或外部环境发生变化时，系统需要同步调整感知重点、通信优先级、计算资源分配以及决策方式，使信息获取、处理与行动反馈能够紧密衔接。资源自适应的感通算控协同增强关注的不是单一过程性能的提升，而是围绕任务需求，将感知什么、如何传输、如何协同计算和怎样快速反馈到行动统一组织起来，从而增强群具身智能体协作的实时性、连续性与执行稳定性。

3.3 与移动通信网和物联网的比较

移动通信网络以高速率为优化目标，物联网以连接设备数量为牵引，而群智网络则以物理世界的协作任务为中心。3类网络在网络功能、结构、管理方式和信息传输等维度上存在差异，与移动通信网和物联网的比较见表1。

从网络功能来看，移动通信网络主要扮演数据管道的角色，服务人与人、人与服务之间的连接。物联网进一步拓展至传感设备互联，支撑感知数据的采集与传输。而群智网络则将连接对象升级为具身智能体，其核心功能是支撑群具身智能体系统的“感知-决策-行动”闭环，使网络直接服务于物理世界的协作任务。在网络结构和资源调度方面，移动通信网络和物联网均采用以基站或接入点为中心的静态或低动态结构，由控制中心集中调度资源^[47]。而群智网络则采用任务驱动的自组织网络结构，具备高动态性，以适应具身智能体位置变化和任务演进。同时，群智网络采用主动分布式资源调度，具身智能体根据任务需求自主协商并动态调整通信关系。在传输内容上，移动通信网络承载任务无关的原始数据，物联网传输设备感知数据，而群智网络

传输任务相关的高价值语义信息。最后，在传输目标上，移动通信网络追求大速率^[45-46]，物联网追求低功耗^[50-51]，群智网络则以信息价值的高密度为核心目标，追求以最小通信开销传递最具价值的语义内容。

4 群智网络的关键性能指标

传统通信网络主要关注速率、时延、吞吐量等性能指标，以量化其数据传输能力。不同于传统通信网络，群智网络的目标已从被动的数据高效传输管道，转变为主动高效地支撑群具身智能体协作完成任务。因此，需要建立面向群智网络的性能评估体系，具体包括5个性能指标。

4.1 协作任务完成率

协作任务完成率是群智网络最根本的性能衡量指标。它定义为在给定测试任务集中，群具身智能体系统成功完成的任务数量与总任务数量的比值，即 $R_{\text{task}} = N_{\text{succ}}/N$ ，其中， N_{succ} 表示成功完成的任务数， N 是任务总数。测试条件为：给定具身智能体数量与类型、环境初始状态、通信带宽等资源预算，多次独立实验后，计算任务完成率。它直接反映了群智网络支撑下的群具身智能体系统能否在给定环境与资源约束下，成功达成预定的物理协作目标。例如，协同搬运的多机器人是否将物体安全送达目的地，无人机集群是否完成指定区域的联合搜索等。该指标能够从任务层面，量化群智网络对群具身智能体协作性能的提升程度。

4.2 群体“感知-决策-行动”闭环响应时长

“感知-决策-行动”闭环响应时长是衡量群智网络对高动态场景中实时协同任务支撑能力的关键性能指标。它刻画了群具身智能体协同感知、高价值语义信息交互、分布式共识决策生成以及物理动作执行完成的总时间，直接影响其在高动态物理环境中的生存能力。设第 k 次闭环过程中，物理环境发生变化（如目标出现）的时刻为 t_k^{sen} ，群具身智能

表1 与移动通信网和物联网的比较

对比项	移动通信网络	物联网	群智网络
网络功能	连接人 数据管道	连接设备 支撑感知数据采集与传输	连接具身智能体 支撑群具身智能体系统的“感知-决策-行动”闭环
网络结构	以基站为中心，支持终端直连 静态/低动态	“端-边-云”融合，支持自组织网络 静态/低动态	任务驱动的自组织网络 高动态
网络管理	集中式为主	集中式为主、辅以多设备协同	主动分布式
传输内容	任务无关数据	感知数据	任务相关语义
传输目标	大速率	低功耗	信息价值的高密度

体首次执行有效动作的时刻为 t_k^{act} ，则群体闭环响应时长定义为 $T_{\text{loop}} = \sum_{k=1}^K (t_k^{\text{act}} - t_k^{\text{sen}})/K$ ，其中， K 为测试过程中触发闭环事件的总数。测试时，采用包含多类物理环境变化（例如，目标出现和障碍物进入）的任务，独立运行多次实验并计算平均闭环响应时长。

4.3 群体能量效率

群体能量效率是指群具身智能体单位焦耳能量下，协作完成的物理任务数量。表达式为 $\eta_E = N_{\text{succ}}/E_{\text{total}}$ ，其中， N_{succ} 为成功完成的任务数， E_{total} 为系统总能耗。它量化了群体的能量利用效率，也反映了群智网络对感通算控资源协同调度的优化程度。同时，高价值密度语义信息的高效传输与利用，能够在保证协作任务性能的情况下，显著减小通信能耗。因此，该指标也间接体现了群智网络中高价值密度语义信息的传输与利用效率。测试时，在相同的任务场景和时间窗口内，对比不同通信协议或调度策略下的能量效率差异。

4.4 协作重构时长与协作性能保持率

协作重构时长是指从检测到具身智能体突然失效、通信链路遮挡或任务重新分配等情况的时刻 t_{det} ，到拓扑和物理协作能力恢复时刻 t_{rec} 的总时间，即 $T_{\text{rec}} = t_{\text{rec}} - t_{\text{det}}$ 。其中，协作能力恢复的阈值应根据具体测试任务而设定。

协作性能保持率是指相同时间内网络重构前后成功任务数量的比值，即 $\rho_{\text{keep}} = \bar{N}_{\text{succ}}/N_{\text{succ}}$ ，其中， N_{succ} 和 $\bar{N}_{\text{succ}}$ 分别是重构前后成功完成的任务数量。测试时，重构前状态应在系统稳定阶段获得，而重构后状态应在系统恢复稳定阶段获得。协作重构时长和协作性能保持率分别从恢复速度和性能两个角度，量化群智网络自组织重构并维持群具身智能体任务协作功能的鲁棒性。

4.5 协作任务泛化成功率

协作任务泛化成功率是指相同时间内陌生场景和熟悉场景中成功完成的任务数量的比值，即 $R_{\text{gen}} = \tilde{N}_{\text{succ}}/N_{\text{succ}}$ ，其中， $\tilde{N}_{\text{succ}}$ 和 N_{succ} 分别是陌生场景和熟悉场景中成功完成的任务数。它衡量群智网络在未知任务、陌生智能体组合或不确定环境等开放场景中，维持群具身智能体高效协作的能力。具体地，它由群智网络对新任务信息需求变化的实时感知和适应能力、跨具身语义理解与共享能力、对在线分布式学习和知识复用的支撑能力等组成。协作

任务泛化成功率越高，说明群智网络从被动数据连接向主动智能体支撑的转变程度越高。值得注意的是，可采用多类陌生场景进行测试，以获得协作任务泛化成功率。

4.6 关键性能指标折中关系

在实际的网络设计与优化中，上述性能指标间存在复杂的折中关系，难以同时使所有指标达到最优，群智网络的关键性能指标如图5所示。面向多样的应用场景需求，理解并利用这些折中关系是设计群智网络的关键，主要包括以下3个方面。

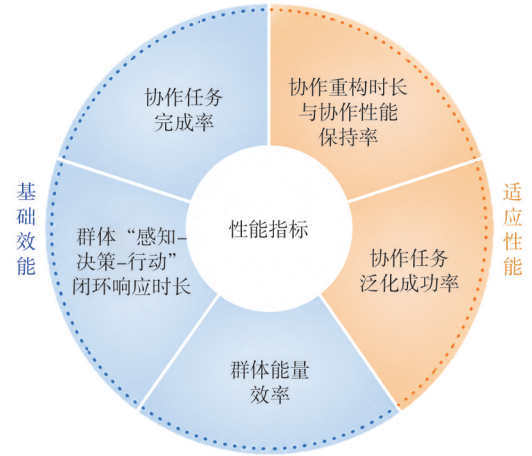


图5 群智网络的关键性能指标

4.6.1 协作任务完成率与协作重构时长的折中

为了在具身智能体突然失效、通信链路遮挡时保持高任务完成率，网络需要实时检测变化并完成重构，这要求持续发送拓扑感知心跳消息并预存冗余路由，从而会增加正常任务执行时的通信开销和决策复杂度，影响任务完成率。同时，若过度追求重构速度，可能因频繁的拓扑调整而干扰正常任务执行，反而降低任务完成率。

4.6.2 闭环响应时长与群体能量效率的折中

为实现极短的任务闭环响应时长，群具身智能体系统通常需要更高的传感器采样频率、更密集的通信交互以及更快速的决策计算，这会导致群体总能耗显著上升。相反，为提升能量效率，降低通信频率或采用低功耗传输策略，会导致响应速度下降。这一矛盾在群体避障、实时对抗等时间敏感场景中尤为突出。

4.6.3 协作任务完成率与泛化成功率的折中

高泛化性要求群智网络具备更强的在线学习和自适应能力，即群智网络策略需要保持一定的灵活

性和探索性，以适应开放场景。但过度探索会降低在已知场景中的确定性执行效率，从而降低特定任务的完成率。反之，专门针对某一类任务优化的网络策略，在该任务上完成率极高，但面对新任务时性能下降明显。因此，根据应用场景的具体需求，例如，工业场景的高完成率需求和灾害救援的强泛化性需求，通过动态权重调整和多目标优化算法，在协作任务完成率与泛化成功率之间实现权衡。

5 关键挑战及未来研究方向

本节聚焦安全可信、异构智能体互操作、网络规模扩展、实时控制稳定性以及标准化5个方面，探讨群智网络从概念走向部署过程中亟待解决的关键科学问题与技术挑战，从而为群智网络的体系化设计与后续深入研究提供方向性参考。

5.1 安全可信问题

群智网络中大量具身智能体需要通过开放的无线信道进行高频次的信息交互、任务协商与决策共享。网络的安全性和可信性不仅会造成信息层面的错误，更会直接引发群具身智能体的灾难性物理执行错误。然而，群智网络处于去中心化且高动态的物理闭环场景中，同时面临恶意节点注入、强电磁干扰与拜占庭故障等多重威胁。因此，群智网络需要在网络层建立多层次的可信共识与容错机制。具体来说，在物理层，利用信道状态信息的时变特性构建物理指纹，实现通信行为的实时异常检测，在恶意信息进入高层协议栈之前进行截断。在网络层，引入轻量级分布式共识协议，借鉴授权拜占庭容错^[6]和并行拜占庭容错^[7]的思想，通过动态选举验证节点子集、分片共识和并行化处理，降低共识的通信复杂度，并确保在部分节点被恶意控制的情况下仍能对关键决策达成一致。此外，建立信誉驱动的动态信任管理机制，每个具身智能体基于历史行为维护其动态信誉评分，并将信誉低于阈值的节点隔离出决策圈层。

5.2 异构具身智能体互操作

异构具身智能体的互操作性是群智网络面临的主要挑战之一。具体而言，互操作性问题体现在两个层面：（1）语法层面的互操作——不同厂商、不同类型的异构具身智能体采用不同的通信协议、数据格式和接口规范，需要在网络层实现协议的动态适配与转换；（2）语义层面的互操作——异构智能体对同一物理世界状态的感知能力、描述方式和

行动能力不同，需要在语义层建立统一的具身形态无关的语义表征框架。在群智网络架构下，需要设计具身智能体能力感知的协议适配机制，使网络能够在运行时自动识别新加入智能体的能力特征，并动态地调整通信协议参数（如编码方式、重传策略、数据速率）以匹配其能力约束等。

5.3 网络规模扩展

在群具身智能体系统的实际应用场景中，参与协作的智能体数量通常需要从数十扩展至数百甚至更大规模，网络的组织方式需要具备与规模同步演进的能力。分布式自组织方式虽然天然具备良好的扩展性，能够避免中心节点因规模增长而产生的通信汇聚瓶颈，但节点间信息同步的滞后也会随规模扩大而加剧，容易导致群体决策出现分歧甚至协作失调。因此，如何设计兼具可扩展性与决策一致性的分层组网架构，是群智网络支撑大规模群体应用的重要前提。

5.4 实时控制稳定性

群智网络支撑的是物理世界中的“感知-决策-行动”闭环任务，网络性能的波动会直接传导至智能体的物理控制回路，因此通信时延的确定性对任务执行的稳定性至关重要。然而，群智网络针对群具身智能体系统的高动态自适应网络架构本身就会带来时延抖动和协作空窗期，如何在保持任务驱动的动态拓扑管理的同时减小其对物理控制稳定性的扰动，是群智网络能否真正支撑物理协作任务的关键。

5.5 标准化问题

群智网络作为新兴网络架构，其性能评估指标和体系架构目前仍处于探索阶段，尚未形成统一的规范，不同研究之间也缺乏统一的评估基准和技术接口。因此，在未来的研究中，需要进一步探索群智网络的架构详细设计、关键使能技术以及性能评估指标体系，并逐步推动相关标准化工作，为不同研究成果的横向对比、跨平台异构智能体的互联互通以及群智网络的规模化应用奠定基础。

6 典型应用

本节聚焦空地具身智能体协作追踪和群具身智能体协作导航两个典型应用场景，初步验证群智网络中“意图驱动的低时延通信”和“任务导向的语义压缩”两个关键技术的可行性。需要指出的是，当前这两项工作仅聚焦于部分具体技术，未对群智网络的完整特性及其性能指标进行全面评估。本文

主要提出群智网络的定义、特性与性能指标。群智网络的系统性验证将是下一步研究的重点方向。

6.1 空地具身智能体协作追踪

群具身智能体协作追踪是指由多具身智能体构成的集群，在开放动态环境中协同完成对移动目标的持续定位和追踪任务。该技术被广泛应用于城市安防和灾害救援等重要领域。然而，无人机集群受限于续航时间和载荷能力，难以实现长时间持续追踪。地面机器人集群因视野受限和地形阻碍，易丢失目标。因此，空地具身智能体协作能够充分发挥无人机集群广域俯瞰与地面机器人集群近距离精确作业的互补优势，显著提升了目标的追踪性能。空地具身智能体的高效协作依赖于实时的关键信息共享。然而，空地具身智能体间通信时延会导致共享信息过时和协同决策时延的问题。针对这个问题，本文进行了面向空地具身智能体协作追踪的低时延通信研究^[42, 68]，提出了一种受生物启发的空地低时延通信方法，面向空地智能体协作追踪的低时延通信如图6所示。该方法将独立强化学习与通信模块相融合，通过模拟自然界中乌鸦与狼的协作捕猎机制，使无人机集群能够利用卷积神经网络-长短期记忆（convolutional neural network-long short-term memory, CNN-LSTM）网络，从长时序观测中推断出目标的未来运动意图，并自主决策通信时机。地面机器人集群根据自身环境状态，仅在必要时向无人机集群主动请求语义信息共享。实验在平面场、方形场、障碍场、动态场和屏障场等复杂度递增的5种环境中进行测试，并与5种多智能体强化学习基线方法进行对比。实验结果表明，所提方法在各类环

境中均优于所有基线方法。在障碍场环境中，所提方法累计奖励比最优基线方法高出约42%。在通信效率方面，该方法将通信频率降低了50%~80%，通信带宽减少了90%以上。在通信时延鲁棒性测试中，当时延达到4~5 s时，该方法仍能维持较高的追踪性能，验证了意图信息传输对时延影响的缓解作用。

6.2 群具身智能体协作导航

具身智能体导航是指具身智能体在未知、动态或部分已知的物理环境中，通过本体传感器实时感知环境信息，结合自身运动学与动力学约束，自主规划并执行从起始位置到目标位置的运动路径的过程。导航是具身智能体实现自主移动与任务执行的基础能力。然而，受限于感知视野、地形通过能力以及负载续航等物理约束，单具身智能体在复杂环境中难以独立完成广域、长时程的导航任务。因此，群具身智能体通过协同感知与决策，共同完成从起点到目标点的路径规划与自主移动，即群具身智能体协作导航，该研究方向具有重要意义。

针对群智能体协作图像导航任务中，智能体之间传输完整的高维感知数据会导致巨大的通信开销，本文进行了面向群智能体协作图像导航的语义通信研究。本文提出了一种基于信息瓶颈的目标驱动语义通信方法，在保证协作导航性能的情况下最小化通信成本，面向智能体协作图像导航的语义通信如图7所示。该方法将率失真理论引入多智能体协作图像导航任务，以刻画通信效率与导航性能之间的折中关系。设计了基于差分信息瓶颈的可变长度语义特征编码方案，能够根据动态信道条件自适应地调整传输特征长度。导航策略网络采用基于多智能体近端策略优化

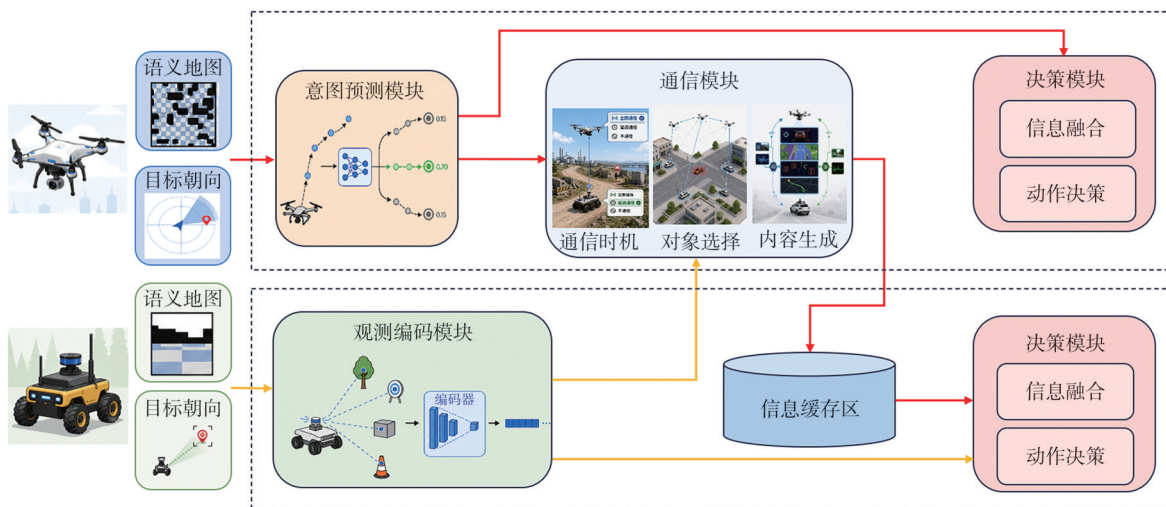


图6 面向空地智能体协作追踪的低时延通信

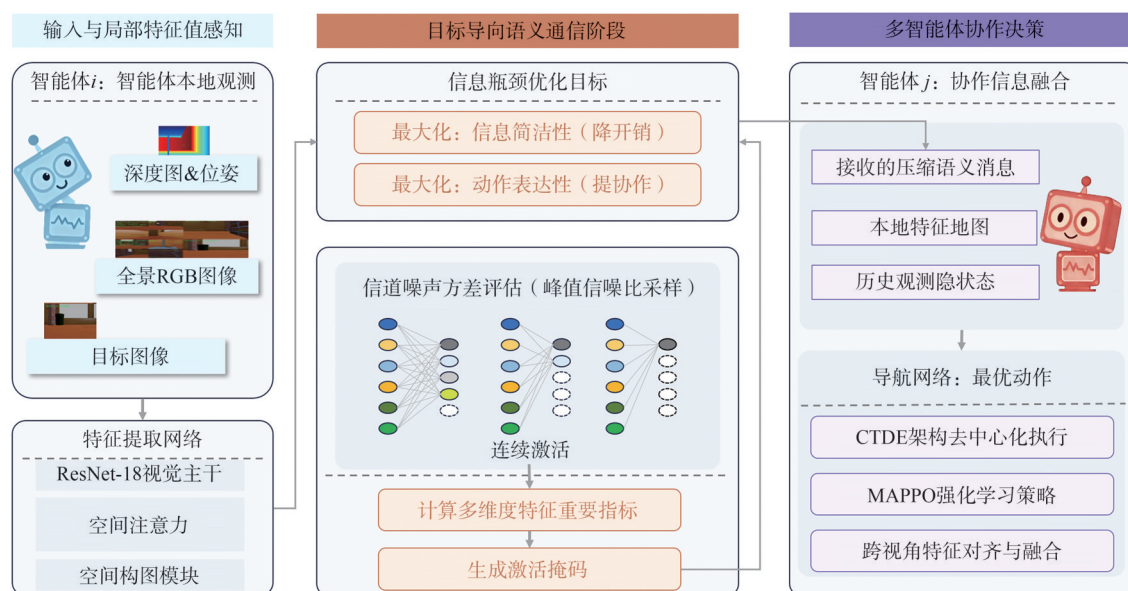


图7 面向智能体协作图像导航的语义通信

(multi-agent proximal policy optimization, MAPPO)的多智能体强化学习方法实现,训练过程遵循集中训练与分散执行(centralized training with decentralized execution, CTDE)框架。在基于Gazebo的仿真平台上构建实验环境,在包含多个房间的12 m×16 m室内环境中执行协作导航任务。实验结果表明,与单智能体基线以及多类多智能体协作方法相比,所提方法在共享目标与特定目标两类任务中均取得了最高的导航成功率,在困难场景中导航成功率提升了68.4%,并使通信开销降低了35.2%。

7 结束语

首先,本文阐述了群具身智能体系统的核心特点与关键技术,提炼出其对通信网络的物理世界的主要需求。其次,通过梳理移动通信网络和物联网的发展概况,介绍了其支撑群具身智能体系统的局限性。在此基础上,本文提出了支撑群具身智能体系统的群智网络,并给出其核心特性。面向群智网络,构建了由协作任务完成率、闭环响应时长、群体能量效率、协作重构时效与性能保持率和开放场景泛化性组成的评价指标体系,并阐明了在实际网络设计优化时指标间的折中关系。最后,以空地具身智能体协作追踪与群具身智能体协作导航为典型应用场景,介绍了生物启发的空地低时延通信和基于信息瓶颈的具身智能体语义通信等代表性研究进展,并以此验证了群智网络的有效性,为后续理论研究提供了系统性参考。在未来的研究中,将进一

步设计群智网络架构并分析其关键使能技术,为群具身智能体系统提供核心支撑。

参考文献:

- [1] 刘云浩. 具身智能: 人工智能的下一个浪潮[M]. 北京: 中信出版社, 2025.
Liu Y H, Embodied intelligence: the next wave of artificial intelligence[M]. Beijing: CITIC Press, 2025.
- [2] Gupta A, Savarese S, Ganguli S, et al. Embodied intelligence via learning and evolution[J]. Nature Communications, 2021, 12: 5721.
- [3] Liu H P, Guo D, Cangelosi A. Embodied intelligence: a synergy of morphology, action, perception and learning[J]. ACM Computing Surveys, 2025, 57(7): 1-36.
- [4] Li Z, Wu W R, Guo Y L, et al. Embodied multi-agent systems: a review[J]. IEEE/CAA Journal of Automatica Sinica, 2025, 12(6): 1095-1116.
- [5] Wu D, Wei X, Chen G, et al. Generative multi-agent collaboration in embodied AI: a systematic review[C]//Proceedings of the Thirty-Fourth International Joint Conference on Artificial Intelligence. International Joint Conferences on Artificial Intelligence Organization, 2025: 10723-10732.
- [6] 郑南宁, 杨勐, 姜维周, 等. 具身智能发展趋势与展望[J]. 中国工程科学, 2026, 28(2): 1-13.
Zheng N N, Yang M, Jiang W Z, et al. Embodied intelligence: development trends and prospects[J]. Strategic Study of CAE, 2026, 28(2): 1-13.
- [7] Hu Y, Pang X H, Qin X Q, et al. Pragmatic communication in multi-agent collaborative perception[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2026, 48(8): 9279-9296.
- [8] Li Y L, Zhang R C, Shao J M, et al. Integration of semantic

- communication and state tracking for intelligent agentic collaboration[J]. IEEE Communications Magazine, 2026. DOI: 10.1109/MCOM.001.2500793.
- [9] Liu Y H, Wang X, Liu Y H, et al. EIoT: embodied intelligence of things[J]. Tsinghua Science and Technology, 2026, 31(1): 1-15.
- [10] 中国人工智能学会. 中国人工智能学会系列白皮书-具身智能[M]. 北京: 中国人工智能学会, 2026.
Chinese Association for Artificial Intelligence. Embodied intelligence: CAAI series white paper[M]. Beijing: Chinese Association for Artificial Intelligence, 2026.
- [11] Littman M L. Markov games as a framework for multi-agent reinforcement learning[M]. Machine Learning Proceedings 1994. Amsterdam: Elsevier, 1994: 157-163.
- [12] Hernandez-Leal P, Kartal B, Taylor M E. A survey and critique of multiagent deep reinforcement learning[J]. Autonomous Agents and Multi-Agent Systems, 2019, 33(6): 750-797.
- [13] Liu X Z, Guo D, Zhang X Y, et al. Heterogeneous embodied multi-agent collaboration[J]. IEEE Robotics and Automation Letters, 2024, 9(6): 5377-5384.
- [14] Verma A, Gautam A, Dutta A, et al. CF-HMRTA: coalition formation for heterogeneous multi-robot task allocation[J]. Journal of Intelligent & Robotic Systems, 2025, 111(3): 81.
- [15] Fazal N, Khan M T, Anwar S, et al. Task allocation in multi-robot system using resource sharing with dynamic threshold approach[J]. PLoS One, 2022, 17(5): e0267982.
- [16] Ramachandran R K, Pierpaoli P, Egerstedt M, et al. Resilient monitoring in heterogeneous multi-robot systems through network reconfiguration[J]. IEEE Transactions on Robotics, 2022, 38(1): 126-138.
- [17] Zhou H, Deng Y S, Liu X N, et al. Goal-oriented semantic communications for 6G networks[J]. IEEE Internet of Things Magazine, 2024, 7(5): 104-110.
- [18] Zhou Y, Xiao J H, Zhou Y, et al. Multi-robot collaborative perception with graph neural networks[J]. IEEE Robotics and Automation Letters, 2022, 7(2): 2289-2296.
- [19] Nair B R. Collaborative perception in multi-robot systems: case studies in household cleaning and warehouse operations[C]//Proceedings of the 2024 6th International Conference on Robotics and Computer Vision (ICRCV). Piscataway: IEEE Press, 2024: 195-200.
- [20] Ma D L, Sun Z, Wei Z Q, et al. AdaptTrack: a robust tracking system for complex environments based on WiFi device selection strategy[J]. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 2025, 9(3): 1-29.
- [21] Liu X Z, Li X H, Guo D, et al. Embodied multi-agent task planning from ambiguous instruction[C]//Proceedings of the Robotics: Science and Systems XVIII. Robotics: Science and Systems Foundation, 2022. DOI:10.15607/rss.2022.xviii.032.
- [22] 於志文, 郭斌, 王亮. 群智感知计算[M]. 北京: 清华大学出版社, 2021.
Yu Z W, Guo B, Wang L. Crowd sensing computing[M]. Beijing: Tsinghua University Press, 2021.
- [23] Zhu W X, Oğuz S, Heinrich M K, et al. Self-organizing nervous systems for robot swarms[J]. Science Robotics, 2024, 9(96): ead15161.
- [24] Emam Y, Mayya S, Notomista G, et al. Adaptive task allocation for heterogeneous multi-robot teams with evolving and unknown robot capabilities[C]//Proceedings of the 2020 IEEE International Conference on Robotics and Automation (ICRA). Piscataway: IEEE Press, 2020: 7719-7725.
- [25] Liu Y C, Tian J J, Glaser N, et al. When2com: multi-agent perception via communication graph grouping[C]//Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE Press, 2020: 4105-4114.
- [26] Hu Y, Fang S H, Lei Z X, et al. Where2comm: communication-efficient collaborative perception via spatial confidence maps[C]//Proceedings of the Advances in Neural Information Processing Systems 35. Neural Information Processing Systems Foundation, Inc. (NeurIPS), 2022: 4874-4886.
- [27] Yang D K, Yang K, Wang Y Z, et al. How2comm: communication-efficient and collaboration-pragmatic multi-agent perception[C]//Proceedings of the Advances in Neural Information Processing Systems 36. Neural Information Processing Systems Foundation, Inc. (NeurIPS), 2023: 25151-25164.
- [28] Liu X Z, Guo D, Liu H P, et al. Multi-agent embodied visual semantic navigation with scene prior knowledge[J]. IEEE Robotics and Automation Letters, 2022, 7(2): 3154-3161.
- [29] Song R, Liang C W, Cao H, et al. Collaborative semantic occupancy prediction with hybrid feature fusion in connected automated vehicles[C]//Proceedings of the 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE Press, 2024: 17996-18006.
- [30] Wei S Z, Wei Y X, Hu Y, et al. Asynchrony-robust collaborative perception via bird's eye view flow[C]//Proceedings of the Advances in Neural Information Processing Systems 36. Neural Information Processing Systems Foundation, Inc. (NeurIPS), 2023: 28462-28477.
- [31] Song Z Y, Xie T H, Zhang H L, et al. A spatial calibration method for robust cooperative perception[J]. IEEE Robotics and Automation Letters, 2024, 9(5): 4011-4018.
- [32] Radford A, Kim J W, Hallacy C, et al. Learning transferable visual models from natural language supervision[C]//Proceedings of the International Conference on Machine Learning, 2021: 8748-8763.
- [33] Maggio D, Chang Y, Hughes N, et al. Clio: real-time task-driven open-set 3D scene graphs[J]. IEEE Robotics and Automation Letters, 2024, 9(10): 8921-8928.
- [34] Ze Y, Yan G, Wu Y H, et al. Gnfactor: multi-task real robot learning with generalizable neural feature fields[C]//Proceedings of the Conference on Robot Learning, 2023: 284-301.
- [35] Huang W, Wang C, Zhang R, et al. VoxPoser: composable 3D value maps for robotic manipulation with language models[C]//Proceedings of the Conference on Robot Learning, 2023: 540-562.
- [36] Kannan S S, Venkatesh V L N, Min B C. SMART-LLM: smart multi-agent robot task planning using large language models[C]//

- Proceedings of the 2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Piscataway: IEEE Press, 2024: 12140-12147.
- [37] Obata K, Aoki T, Horii T, et al. LiP-LLM: integrating linear programming and dependency graph with large language models for multi-robot task planning[J]. IEEE Robotics and Automation Letters, 2025, 10(2): 1122-1129.
- [38] Choi H L, Brunet L, How J P. Consensus-based decentralized auctions for robust task allocation[J]. IEEE Transactions on Robotics, 2009, 25(4): 912-926.
- [39] Wu P, Escontrela A, Hafner D, et al. Daydreamer: world models for physical robot learning[C]//Proceedings of the Conference on Robot Learning, 2023: 2226-2240.
- [40] Song S C, Lin Y F, Han S, et al. CoDe: communication delay-tolerant multi-agent collaboration via dual alignment of intent and timeliness[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2025, 39(22): 23304-23312.
- [41] Yuan T T, Chung H M, Yuan J, et al. DACOM: learning delay-aware communication for multi-agent reinforcement learning[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2023, 37(10): 11763-11771.
- [42] Zhang Q, Sun Z, Zhang Y, et al. VIL2C: value-of-information aware low-latency communication for multi-agent reinforcement learning[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2026, 40(35): 29874-29882.
- [43] Wang C, Rahman A, Durugkar I, et al. N-agent ad hoc teamwork[C]//Proceedings of the Advances in Neural Information Processing Systems 37. Neural Information Processing Systems Foundation, Inc. (NeurIPS), 2024: 111832-111862.
- [44] Oneill A, Rehman A, Maddukuri A, et al. Open x-embodiment: robotic learning datasets and RT-x models: open x-embodiment collaboration[C]//Proceedings of the IEEE International Conference on Robotics and Automation, 2024: 6892-6903.
- [45] Andrews J G, Buzzi S, Choi W, et al. What will 5G be?[J]. IEEE Journal on Selected Areas in Communications, 2014, 32(6): 1065-1082.
- [46] 王胡成, 徐晖, 程志密, 等. 5G 网络技术研究现状和发展趋势[J]. 电信科学, 2015, 31(9): 149-155.
- Wang H C, Xu H, Cheng Z M, et al. Current research and development trend of 5G network technologies[J]. Telecommunications Science, 2015, 31(9): 149-155.
- [47] Qian M L, Wang Y Y, Zhou Y Q, et al. A super base station based centralized network architecture for 5G mobile communication systems[J]. Digital Communications and Networks, 2015, 1(2): 152-159.
- [48] 3GPP. NR: medium access control (MAC) protocol specification[S]. 3GPP TS 38.321, V18.10.0, 2026.
- [49] Chen X M, Ng D W K, Yu W, et al. Massive access for 5G and beyond[J]. IEEE Journal on Selected Areas in Communications, 2021, 39(3): 615-637.
- [50] Ikpehai A, Adebisi B, Rabie K M, et al. Low-power wide area network technologies for Internet-of-things: a comparative review[J]. IEEE Internet of Things Journal, 2019, 6(2): 2225-2240.
- [51] Gu F, Niu J W, Jiang L D, et al. Survey of the low power wide area network technologies[J]. Journal of Network and Computer Applications, 2020, 149: 102459.
- [52] 郭斌, 刘思聪, 刘琰, 等. 智能物联网: 概念、体系架构与关键技术[J]. 计算机学报, 2023, 46(11): 2259-2278.
- Guo B, Liu S C, Liu Y, et al. AIoT: the concept, architecture and key techniques[J]. Chinese Journal of Computers, 2023, 46(11): 2259-2278.
- [53] Ren W Q, Qu Y B, Dong C, et al. A survey on collaborative DNN inference for edge intelligence[J]. Machine Intelligence Research, 2023, 20(3): 370-395.
- [54] McMahan B, Moore E, Ramage D, et al. Communication-efficient learning of deep networks from decentralized data[C]//Proceedings of the Artificial Intelligence and Statistics, 2017: 1273-1282.
- [55] Jameel F, Hamid Z, Jabeen F, et al. A survey of device-to-device communications: research issues and challenges[J]. IEEE Communications Surveys & Tutorials, 2018, 20(3): 2133-2168.
- [56] Zhou Z, Chen X, Li E, et al. Edge intelligence: paving the last mile of artificial intelligence with edge computing[J]. Proceedings of the IEEE, 2019, 107(8): 1738-1762.
- [57] Guo J, Chen H, Song B, et al. Distributed task-oriented communication networks with multimodal semantic relay and edge intelligence[J]. IEEE Communications Magazine, 2024, 62(6): 82-89.
- [58] Xie H Q, Qin Z J, Tao X M, et al. Task-oriented multi-user semantic communications[J]. IEEE Journal on Selected Areas in Communications, 2022, 40(9): 2584-2597.
- [59] Zhang C, Zou H, Lasaulce S, et al. Goal-oriented communications for the IoT and application to data compression[J]. IEEE Internet of Things Magazine, 2022, 5(4): 58-63.
- [60] Brohan A, Chebotar Y, Finn C, et al. Do as I can, not as I say: grounding language in robotic affordances[C]//Proceedings of the Conference on Robot Learning, 2023: 287-318.
- [61] Feng C, Zheng K D, Wang Y, et al. Goal-oriented wireless communication resource allocation for cyber-physical systems[J]. IEEE Transactions on Wireless Communications, 2024, 23(11): 15768-15783.
- [62] Shao J W, Mao Y Y, Zhang J. Learning task-oriented communication for edge inference: an information bottleneck approach[J]. IEEE Journal on Selected Areas in Communications, 2022, 40(1): 197-211.
- [63] Luo J H, Fei Z S, Wang X Y, et al. Toward low-altitude embodied intelligence: a sensing-communication-computation-control closed-loop perspective[J]. IEEE Communications Magazine, 2026: 1-8.
- [64] Jin H J, Wu J, Yuan W J, et al. Co-design of sensing, communications, and control for low-altitude wireless networks[J]. IEEE Transactions on Mobile Computing, 2025, 24(11): 12035-12048.
- [65] Sun Z, Yu Z W, Guo B, et al. Integrated sensing and communication for effective multi-agent cooperation systems[J]. IEEE Communications Magazine, 2024, 62(9): 68-73.
- [66] Zhan Y, Wang B C, Lu R X, et al. DRBFT: delegated randomization Byzantine fault tolerance consensus protocol for blockchains[J].

Information Sciences, 2021, 559: 8-21.

- [67] Chen X, Ma T J, Er-Rahmadi B, et al. Parallel Byzantine consensus based on hierarchical architecture and trusted hardware[J]. IEEE Transactions on Dependable and Secure Computing, 2024, 21(6): 5493-5508.
- [68] Hao Z T, Guo B, Zhao K X, et al. AGCoTrack: a communication-efficient independent reinforcement learning method for aerial-ground collaborative tracking[J]. IEEE Internet of Things Journal, 2024, 11(23): 37758-37769.

[作者简介]



孙卓(1990-), 女, 博士, 西北工业大学计算机学院教授、博士生导师, 国家级青年人才。2019年于澳大利亚新南威尔士大学获得博士学位, 之后在澳大利亚国立大学担任博士后研究员。主要研究方向为多智能体通信网络、任务驱动语义通信等。曾获2024年《航空学报》最具传播力文章奖。担任IEEE Transactions on Mobile Computing编委。



於志文(1977-), 男, 博士, 哈尔滨工程大学副校长, 西北工业大学计算机学院教授, 教育部“长江学者”特聘教授, 国家杰出青年科学基金获得者。主要研究方向为普适计算、人机系统、社会感知等。曾获CCF优秀博士学位论文奖、CCF青年科学家奖、国家级教学成果二等奖、教育部自然科学一等奖和陕西省科学技术一等奖。



郭斌(1980-), 男, 博士, 西北工业大学计算机学院教授, 国家杰出青年科学基金获得者。主要研究方向为人机物融合智能计算、泛在感知计算与群智计算。曾获教育部自然科学一等奖、陕西省自然科学一等奖、军事科技进步一等奖、CCF自然科学二等奖、IEEE UIC17、BIBM20等国际会议“最佳论文奖”8次。担任CCF理事、普适计算专委会秘书长。



李玉(1999-), 男, 西北工业大学计算机学院博士生, 主要研究方向为演化博弈论、多智能体合作演化。



杨博(1986-), 男, 博士, 西北工业大学计算机学院教授、博士生导师, 2017年3月在西北工业大学获得博士学位, 2017年10月至2022年2月, 分别在美国和新加坡开展博士后研究。主要研究方向为群智网络。曾获得2023年CCF物联网大会(CWSN)优秀论文奖和IEEE VTC2025-Spring最佳学生论文奖。担任IEEE VTC-2018 Fall、EAI IoTaas 2020、IEEE PIMRC 2026分会主席, SN Computer Science副编辑。