

doi: 10.7690/bgzdh.2022.12.013

无人系统自主性技术研究现状与发展趋势

刘佩林¹, 陈祥¹, 牛小明^{2,3}

(1. 上海交通大学电子信息与电气工程学院, 上海 200240;

2. 中国兵器装备集团自动化研究所有限公司武器装备信息与控制技术创新中心, 四川 绵阳 621000;

3. 重庆大学数学与统计学院, 重庆 400044)

摘要: 鉴于无人系统自主性技术影响未来战争的发展方向, 对其研究现状及发展趋势进行分析。阐述无人系统自主性的内涵, 并以面向不同单一导航任务为例介绍无人系统自主性技术的框架, 重点揭示内驱力和外驱力在环境模型构建和行为决策中的作用; 对无人系统自主性技术研究现状进行分析, 包括态势感知、自动规划、策略学习和环境模型学习 4 方面, 据此判定无人系统自主性技术的研究趋势; 讨论无人系统自主性技术在实际应用中存在的问题与面临的挑战。结果表明, 该分析可为我国无人系统的发展提供参考。

关键词: 无人系统; 自主性技术; 态势感知; 自动规划; 策略学习; 环境模型学习

中图分类号: TJ99 **文献标志码:** A

Research Status and Development Trend of Autonomous Technology for Unmanned System

Liu Peilin¹, Chen Xiang¹, Niu Xiaoming^{2,3}

(1. School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China;

2. Weapon Equipment Information and Control Technology Innovation Center, Automation Research Institute Co., Ltd. Of China South Industries Group Corporation, Mianyang 621000, China;

3. College of Mathematics and Statistics, Chongqing University, Chongqing 400044, China)

Abstract: In view of the influence of autonomous technology of unmanned system on the development direction of future warfare, the research status and development trend of autonomous technology of unmanned system are analyzed. The autonomy of unmanned system is elaborated, and the framework of autonomous technology of unmanned system is introduced by taking different single navigation tasks as an example, focusing on the role of internal and external drives in environment model construction and behavior decision-making; The research status of autonomous technology of unmanned system is analyzed, including situation awareness, automatic planning, policy learning and environment model learning. Based on this, the research trend of autonomous technology of unmanned system is determined. The problems and challenges in the practical application of the technology are discussed. The results show that the analysis can provide reference for the development of unmanned system in China.

Keywords: unmanned system; autonomous technology; situation awareness; automated planning; policy learning; environment model learning

0 引言

现阶段, 无人系统呈现出快速发展的态势, 在军事领域尤其复杂战场环境中的应用价值日益凸显, 不仅能够执行情报侦察、战场监视、目标指示等任务, 而且还在电子干扰、防空压制、空中对抗、对地攻击等方面大显身手。无人系统获得世界各军事强国的广泛关注, 自 2007 年以来美国先后发布了 5 版“无人系统综合路线图”, 最新版涵盖了在 2017—2042 年间将无人系统整合到作战系统中的整体规划和关键技术^[1]; “欧盟永久结构性合作”中也开展了多项无人系统相关的科研项

目^[2]。无人系统在现代作战和国家安全领域中承担愈加重要的角色, 并最终有望代替有人系统, 成为未来作战和攻防的主力。基于此, 对于无人系统中的核心支撑技术—自主性技术, 各军事强国均加大投资, 积极开展自主性关键技术研究, 试图攻克技术难点, 以期在未来战场新作战模式和武器装备的探索中占领先机。笔者对无人系统自主性技术的内涵和外延进行阐述, 分析无人系统自主性技术的研究现状, 并对其研究趋势以及面临的问题与挑战进行探讨, 以期为无人系统的相关研究提供一定参考。

收稿日期: 2022-08-15; 修回日期: 2022-09-14

作者简介: 刘佩林(1965—), 女, 天津人, 博士, 教授, 从事基于信号环境特性的智能信号处理、深度视觉、多媒体处理、类脑导航和低功耗 SoC 设计研究。E-mail: liupeilin@sjtu.edu.cn。

1 无人系统自主性技术

1.1 无人系统自主性概念

无人系统是指不需要人工干预的系统，由无人平台及若干辅助部分组成，具有自主完成预定任务的能力。无人系统包括无人机系统(unmanned aerial systems, UAS)、无人地面车辆(unmanned ground vehicles, UGV)和无人海上系统(unmanned marine systems, UMS)3 部分。无人系统自主性是指无人系统实体对象根据对外部环境的感知以及对环境与自身的理解，在外部任务驱动下，实现独立制定并选择不同行动方案的能力。

1.2 无人系统自主性技术框架

以实现不同单一导航任务为例，介绍如图 1 所示的无人系统自主性技术框架。首先，利用感知模块采集环境与机器人的状态信息。然后，构建环境

模型，包括对感知信息进行编码，识别出环境和机器人的状态；基于图模型构建知识存储，为状态识别提供先验信息，并与其共同预测不同行为的状态/收益。最后，对任务进行决策与执行，同时感知模块获取变化后的环境信息，并循环上述流程。相比常规导航框架，自主系统框架具有更强的推理能力，表现在行为决策中具有 2 个驱动力：1) 外驱动力，包括导航中使用的训练数据、训练算法与收益函数等，系统在外驱动力的作用下进行环境模型和执行策略的学习更新，“发掘”执行任务的思路以完成目标；2) 内驱动力，其有别于外驱动力，在更好地完成任务的同时可构建出认知模型，实现环境模型中以最小代价存储知识。此外，内驱动力能够对感知模块进行补充，实现未知环境的预测，即构建出“梦境”，有利于替代真实环境与无人系统中智能体进行交互，提高策略的学习效率。

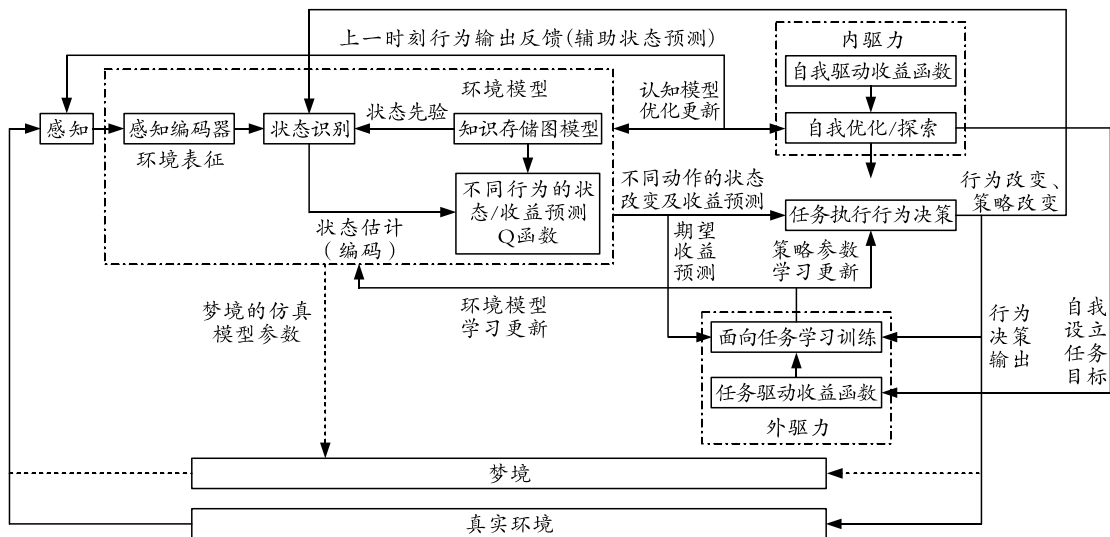


图 1 无人系统自主性技术框架

内驱动力使无人系统更加智能，特别是针对感知信息不完备、存在干扰等场景。无人系统可利用内驱动力构建环境模型如图 2 所示。为实现内驱动力的“触类旁通”，在环境模型中将环境与知识解耦合，而外驱动力则为两者提供了记忆关联绑定(例如将地图上各场景节点绑定到图结构，由推理引擎搜索物体)。为使环境模型实现最小代价遍历，基于图结构构建知识体系。利用图推理实现预测前推，通过内驱学习实现因果推理，获得反事实前推，从而构建出环境模型。

1.3 无人系统自主性特征

无人系统自主性主要具有 2 种特征：1) 自学习

特征。根据外部任务环境的性能反馈可修正并归纳自身行为，实现独立分析、解决问题，从而完成目标。2) 自演化特征。不仅以外部任务为驱动对环境模型进行更新，而且能够根据自身情况进行主动学习与提升，获取外部环境知识并提高任务完成效率。

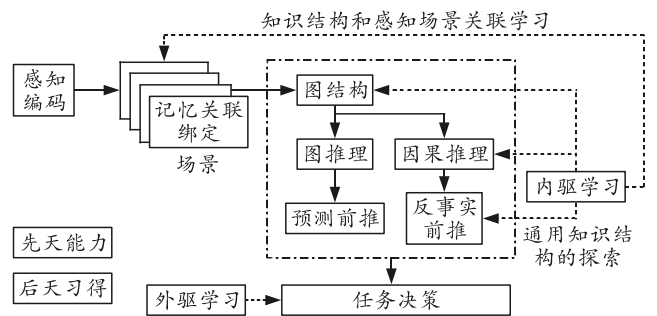


图 2 环境模型构建框架

1.4 无人系统自主性层级

根据无人系统对环境模型的理解以及执行任务的能力,可将无人系统自主性层级由低到高分为 5 个级别。笔者提出的无人系统自主性技术框架可归类到第 5 个级别。

- 1) 按照人工规划的任务执行步骤完成任务;
- 2) 根据给定规则(基于“动作-状态变化”生成的规则)自动规划任务执行步骤,完成任务;
- 3) 探索环境以构建“动作-状态变化”集合,自主规划完成任务;
- 4) 主动“发掘”环境模型,构建“状态-对策”集合,自主规划完成任务;
- 5) 以 4)为基础提取因果机制,获取“反事实”的“动作-状态变化”规则,自主规划完成任务。

2 无人系统自主性技术研究现状

根据无人系统自主性技术框架可归纳出其主要包括态势感知、自动规划、策略学习和环境模型学习 4 项技术,下面具体介绍各项技术的研究现状。

2.1 态势感知

态势感知的主要作用是针对无人系统所感知的多维信息进行处理,理解与预测目标的未来态势^[3]。目前,态势感知方法主要分为:

1) 基于专家系统的态势感知方法:专家系统主要通过对战场地事件进行分析研判,进而预测未来战场态势。研究人员已基于该方法利用卫星获取的信息实现了战场态势评估,提升了无人系统在复杂任务环境中面对异常情况的快速处理能力^[4]。也有学者利用该方法进行无人潜航器研究,评估了任务执行过程威胁性等级,自主生成了合理的行为策略^[5]。

2) 基于模板匹配的态势感知方法:模板匹配算法根据目标的特征信息,从知识库模板中选择合适的模板进行相似度匹配,从而实现目标特征的定性或定量分析。目前,已有研究人员提出了基于改进模板匹配的态势推理算法,将态势模板表示成图,通过最大化公共子图算法解决模板和观测数据之间的匹配问题,实现更加准确的态势模式评估^[6]。

3) 基于贝叶斯推断的态势感知方法:该方法主要包括 4 步:①对照战场中的事件构造变量节点;②构造由因到果的有向边,建立不同节点之间的依赖关系;③构造条件概率表以表示某事件发生的条件下采取某项行动的概率值;④利用贝叶斯推断原

理对战场地事件进行态势评估。当前,已有研究者利用该方法实现了弹道导弹态势的评估,为弹道导弹拦截任务提供了快速智能化的态势感知支持^[7]。

4) 基于深度学习的态势感知方法:该方法利用深度神经网络进行训练,获得由目标原始信息到目标属性的复杂映射关系。当前,已有研究者提出了一种基于 LSTM 的空中目标作战意图识别方法,在作战意图识别问题中加入时序特征,提升了识别准确率^[8]。Gerald 等利用卷积神经网络进行多人检测与动作识别,提升了空间态势感知范围^[9]。

2.2 自动规划

无人系统中的自动规划从全局角度为智能体提供了全局目标和路线,目前的研究聚焦于如何通过地图规划合理路线为智能体提供子目标。这方面比较成熟的研究方案主要分为 2 类。

1) 基于启发式搜索的方法。

在小规模地图上能够在较短的时间为智能体规划出较优的路径,如 RRT*、A*等算法已成功广泛应用于单智能体及多智能体的路径规划和导航。

2) 基于深度强化学习的方法。

更侧重于复杂动态场景中的路径规划,有效解决启发式搜索算法在动态场景中重规划速度慢、发生碰撞等问题,产生了诸如 ActiveNeuralSLAM、SemExp 等优秀的算法^[10-11]。以深度强化学习为代表的自动规划方案更符合具有自主性的无人系统,能够为无人系统中的智能体提供类似指挥官的作用,以类似课程学习的方式引导智能体在较短的时间内学到更好的局部策略,形成“整体-局部”的多层次鲁棒策略方案,增强无人系统的灵活性和可靠性。

2.3 策略学习

策略学习是实现无人系统自主性的关键技术之一。目前基于深度强化学习的策略因其与生物奖励驱动的决策具有高度相似性,使其在无人系统多智能体编队、复杂场景军事打击等领域极具应用价值。

根据无人系统所需智能体的数量可将决策划分为单智能体和多智能体决策。其中基于深度强化学习的单智能体算法发展迅速,产生了能够应对 UGV 自主规划、避障和导航等任务的算法,如 DDPG、A3C、PPO 和 SAC 等^[12]。这些算法通过与场景进行交互构建出环境状态到行为决策的映射,具有较强的学习能力。近年来,在单智能体强化学习算法的

基础上,产生了基于值分解的 VDN、QMIX 以及基于策略梯度的 MADDPG、COMA、VMIX 等多智能体决策算法。其中 MADDPG、COMA 等通过集中式训练、分布式执行的方式使无人系统能够应对多智能体协作和博弈对抗等任务^[13]。

无人系统对环境的认知存在外驱力和内驱力 2 种决策学习方式。外驱力以完成特定任务给予智能体奖励的方式,促使智能体针对具体问题学习完成任务的行为决策;内驱力不依赖场景给予的特定奖励,而是通过衡量智能体对场景的不确定性来增强对环境的探索能力。该方面研究包括基于最大熵原理的 SQL、SAC 等算法以及基于好奇心机制的 ICM 等算法^[14]。内驱力使无人系统进行自主学习,对环境具有较强的认知能力。在无人系统中,有效的感知信息决定决策的准确性,将基于知识图谱的认知地图作为状态信息,采用基于深度强化学习的算法进行导航,有利于提升策略的可靠性和扩展性。

2.4 环境模型学习

目前基于深度强化学习的策略学习方式,虽然可满足复杂场景的决策,但其以“试错”方式进行学习,该过程学习效率较低。通过构建环境模型以替代真实环境进行交互,从而实现策略学习,是提高学习效率的关键技术之一。

环境模型试图利用深度神经网络模拟环境的状态表示和状态间的概率转移,以取代环境进行策略学习。无人系统中智能体与环境大量交互会对智能体硬件带来负担,同时消耗大量能源。监督学习方式以少量交互样本学习到丰富的环境信息,可替代真实的交互样本,提升智能体的策略学习效率,降低硬件依赖和能源消耗。此外,学习到的环境模型在相似场景下具有较强的泛化能力,通过微调方式便能模拟新场景并进行策略学习^[15]。

当前针对环境模型学习的研究产生了大量研究成果,如 MBPO、MuZero 等,实现了低样本开销下的有效策略学习^[16-17]。其中 MuZero 算法在未知规则的情况下,从零开始学习环境模型和策略,并在围棋这一场景中战胜了 AlphaGo、AlphaZero 等系列算法,成为目前最强的围棋 AI。这类算法在交互代价较高的真实场景中极具应用前景。

3 无人系统自主性技术研究趋势

当前,各军事强国针对无人系统自主性技术开展了深入研究,推动了无人系统在军事领域的应用。

作为无人系统的重要组成部分,态势感知系统正朝着自主化和协同化方向不断发展,提升了全域响应和多维感知能力,对于实现跨域协同侦察、空地一体打击和多目标作战等复杂任务具有重要意义。

针对具体技术而言,以深度强化学习为基础的多智能体协作、博弈将成为推动无人系统发展的核心决策方案,而态势感知所采集的信息作为产生决策的依据,其相关的信息表征学习技术也会成为未来无人系统自主性技术发展的重点。

4 应用中存在的问题与挑战

实际战场中由于电磁环境十分恶劣,常规的通信方式难以正常工作,无人系统各节点之间的通信将受到严重影响,如何解决复杂电磁环境下的无人系统通信问题是态势感知技术在实际战场应用中需要解决的重要问题。此外,由于无人系统体积和功耗的限制,如何分配各模块的计算资源也是要解决的难题。

在实际战场环境中,无人系统中智能体的交互学习面临严重的采样效率问题,在样本量不足的情况下以深度强化学习为基础的算法无法学到有效策略,将对无人系统自主决策产生较大影响。如何解决在复杂场景下通过有限样本学习有效策略是无人系统自主性技术面临的重要挑战。

5 结束语

无人系统初步形成了以态势感知、自动规划、策略学习与环境模型学习为支撑的自主系统,表现出一定的自学习和自演化特征,进一步向更高层级的自主系统发展。目前,无人系统自主性技术形成了以内驱力和外驱力为核心的技术框架,在民用和军用等领域都表现出巨大的应用前景。同时,无人系统自主性技术也面临一系列技术瓶颈以及应用中存在的安全问题、伦理问题等。这既是机遇又是挑战,应集中科研资源,增加无人系统自主性技术储备,加强顶层设计与制定专项规划,关注各军事强国的无人系统自主性技术进展,确保我国在国防领域高端技术的先进性。

参考文献:

- [1] Unmanned Systems Integrated Roadmap 2017-2042[EB/OL]. [2018-8-28]. https://www.defensedaily.com/wp-content/uploads/post_attachment/206477.pdf.
- [2] MOGHERINIF. PESCO Member States Driven[EB/OL].

- [2019-11-12]. <https://www.pesco.europa.eu>.
- [3] 符杰凯. 基于无人系统集群的多维数据态势感知技术研究[D]. 成都: 电子科技大学, 2021.
- [4] 李婷. 基于专家决策系统的空间战场态势感知研究[D]. 沈阳: 沈阳建筑大学, 2017.
- [5] 么洪飞. 不确定条件下 UUV 态势感知、威胁评估与自主决策方法研究[D]. 哈尔滨: 哈尔滨工程大学, 2020.
- [6] 余源丰. 群目标态势推理关键技术研究[D]. 西安: 西安电子科技大学, 2020.
- [7] 宁宇, 张峰, 张增辉, 等. 弹道导弹多维态势快速构建与智能处理[J]. 信息系统工程, 2021(4): 121-124.
- [8] XUE J, ZHU J, XIAO J, et al. Panoramic convolutional long short-term memory networks for combat intension recognition of aerial targets[J]. IEEE Access, 2020, 8: 183312-183323.
- [9] GERALDES R, GONCALVES A, LAI T, et al. UAV-based situational awareness system using deep learning[J]. IEEE Access, 2019, 7: 122583-122594.
- [10] CHAPLOT D S, GANDHI D, GUPTA S, et al. Learning to explore using active neural slam[J]. arXiv, 2020(4): 2004.
- [11] CHAPLOT D S, GANDHI D P, GUPTA A, et al. Object goal navigation using goal-oriented semantic exploration[J]. Advances in Neural Information Processing Systems, 2020, 33: 4247-4258.
- [12] HAARNOJA T, ZHOU A, ABBEEL P, et al. Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor[C]//International conference on machine learning. New York: ACM, 2018: 1861-1870.
- [13] GRONAUER S, DIEPOLD K. Multi-agent deep reinforcement learning: a survey[J]. Artificial Intelligence Review, 2022, 55: 895-943.
- [14] PATHAK D, AGRAWAL P, EFROS A A, et al. Curiosity-driven exploration by self-supervised prediction[C]//International conference on machine learning. New York: ACM, 2017: 2778-2787.
- [15] MOERLAND T M, BROEKENS J, JONKER C M. Model-based reinforcement learning: A survey[J]. arXiv, 2020(6): 2006. 16712.
- [16] JANNER M, FU J, ZHANG M, et al. When to trust your model: Model-based policy optimization[J]. Advances in Neural Information Processing Systems, 2019, 32: 1-47.
- [17] SCHRITTWIESER J, ANTONOGLOU I, HUBERT T, et al. Mastering atari, go, chess and shogi by planning with a learned model[J]. Nature, 2020, 588(7839): 604-609.
- *****
- [5] 黄双得, 周仿荣, 普凯, 等. 电缆隧道架空式巡检机器人行走机构研究与应用[J]. 电气应用, 2020, 39(4): 71-76.
- [6] 虞鸿江. 电缆隧道智能巡检机器人设计探析[J]. 通信电源技术, 2020, 37(2): 142-144.
- [7] 吴伟. 高压电缆隧道智能巡检机器人控制系统研发[D]. 天津: 天津职业技术师范大学, 2016.
- [8] 黄荣辉, 向真, 姜勇, 等. 一种高压室巡检机器人移动机构设计[J]. 现代机械, 2018(6): 9-15.
- [9] 曹涛, 王明瑞, 孙大庆, 等. 变电站巡检机器人云台设计及动力学仿真[J]. 制造业自动化, 2012, 34(9): 37-40.
- [10] 董龙飞, 陶卫军. 变电站巡检机器人结构设计与分析[J]. 兵工自动化, 2019, 38(9): 78-81, 96.
- [11] 宋庆军, 胡程量, 姜海燕, 等. 巡检机器人软土地面行进动力学分析与仿真[J]. 煤矿机械, 2021, 42(11): 88-90.
- [12] 汪鹏. 变电站巡检机器人系统结构设计与控制研究[D]. 合肥: 安徽大学, 2020.
- [13] 陶陶. 吊轨式巡检机器人动态特性研究[D]. 徐州: 中国矿业大学, 2019.
- [14] 谢志坤, 韩军, 冯虎田. 轨道摄影机器人瞬时转弯半径分析[J]. 机械设计与制造工程, 2021, 50(10): 19-23.
- [15] 王川伟, 马宏伟, 马琨, 等. 带式输送机巡检机器人行驶力学及其仿真研究[J]. 煤炭技术, 2018, 37(10): 258-261.

(上接第 55 页)

5 结束语

笔者设计一种适用于电缆隧道的吊轨式巡检机器人, 结合隧道内实际环境, 确定了机器人的工作原理, 提出了整体的结构设计与控制系统方案。该机器人能够对隧道内的 3 层电缆进行红外测温, 对环境中的危险气体进行检测, 同时能实现自主充电功能。此外, 针对机器人建立了相应的力学模型, 并分析了机器人在水平直线轨道、爬坡阶段以及转弯轨道上的运动情况。利用 Matlab 软件对机器人运动过程中输出力矩的变化情况进行了仿真与分析, 根据机器人性能指标对驱动电机进行了合适的选型, 可为巡检机器人的设计与应用提供一定参考。

参考文献:

- [1] 钩达, 胡新建. 论地下电缆发展的必然趋势[J]. 城市照明, 2012, 16(2): 14-16.
- [2] 赵晖, 钱瑞明. 新型电缆隧道巡检机器人机构设计与轨道优化仿真[J]. 机械设计与制造工程, 2018, 47(3): 61-65.
- [3] 罗军东. 试论武汉城市电缆隧道建设[J]. 供电企业管理, 2014(4): 4-6.
- [4] 吴锴, 左兆陆, 窦少校. 我国轨道式巡检机器人研究及发展现状[J]. 软件, 2018, 39(11): 88-91.

