文章编号:1002-0640(2020)09-0116-06

军事智能博弈对抗系统设计框架研究

李 航1,刘代金2,刘 禹2*

(1.海军研究院建设发展研究所,北京 102300;2.中国科学院自动化研究所,北京 100190)

摘 要:在军事领域,人工智能提供了一种新路径来建立和维护军事优势。随着深度学习、强化学习算法的巨大突破,如何在复杂多变的战场环境下制定高效的军事策略并进行及时调整成为新的研究方向,人工智能技术在博弈对抗领域的应用价值和前景也越来越高。将人工智能技术集成到计算机兵棋推演和作战模拟仿真系统中形成军事智能博弈对抗系统,提出一种适应于军事博弈对抗系统智能应用研究的通用设计框架,适用于 AI 智能体实现跨系统、跨平台的泛化应用,为指挥训练提供技术支撑。

关键词:人工智能,军事博弈,作战决策,推演引擎

中图分类号:E9;TJ0

文献标识码: A

DOI: 10.3969/j.issn.1002-0640.2020.09.021

引用格式:李航,刘代金,刘禹.军事智能博弈对抗系统设计框架研究[J].火力与指挥控制,2020,45(9):116-121.

Architecture Design Research of Military Intelligent Wargame System

LI Hang¹, LIU Dai-jin², LIU Yu^{2*}

(1.Institute of Construction and Development, Naval Research Academy, Beijing 102300, China)

2.Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China)

Abstract: Artificial Intelligence technology provides a new route to build advantages in the military field. With the great breakthrough of deep learning and reinforcement learning algorithm, how to develop efficient military strategies in the complex and changeable battlefield environment has become an extremely valuable challenge. Integrating AI technology into computer wargame and simulation system to form Military Intelligent Wargame System (MIWS), this paper presents a general architecture adapted to the military intelligence application, which is suitable for AI agent to realize cross—system and cross—platform generalization application, and will provide benefits for decision—making and wargaming training.

Key words: artificial intelligence, military wargame, decision-making, wargame engine

Citation format: LI H, LIU D J, LIU Y. Architecture design research of military intelligent wargame system [J]. Fire Control & Command Control, 2020, 45(9): 116–121.

0 引言

人工智能是研究开发用于模拟、延伸和扩展人的智能的理论、方法、技术及应用系统的一门技术 科学,其本质是依托计算机运用数学算法模仿人类 智力,让机器达到甚至超越人类的分析、推理和思维能力。人工智能技术从诞生以来,已经发展了半个多世纪的时间,在各个领域都取得了重要的成果。在军事领域,人工智能提供了一种新路径来保持己方或打破对手的军事优势,同时还能起到削减

收稿日期:2019-05-24 修回日期:2019-08-10

作者简介:李 航(1969-),男,北京人,硕士,教授。研究方向:军事博弈对抗战略规划。

*通信作者:刘 禹(1979-),男,北京人,博士,教授。研究方向:人工智能博弈对抗。

军费开支、降低战场伤亡的作用[1-2]。20世纪60年 代,美军率先将人工智能技术应用于制导武器的智 能化控制,实现对目标的自动识别、锁定和打击; 1997年,IBM 公司的"深蓝"击败国际象棋大师卡斯 帕罗夫,这是基于知识规则引擎和强大计算机硬件 的人工智能系统的胜利;2008年,美国国防高级研 究计划局(Defense Advanced Research Projects Agency, DARPA) 启动了"深绿(Deep Green)"计划[3],提 出在指挥控制系统中嵌入智能化模块,以提高指挥 员临机决策的速度和质量;2009-2014年,DARPA 又先后启动了大量基础技术研究项目,研究探索从 文本、图像、声音、视频、传感器等多源多模态数据 中自主获取、智能处理、特征抽取、关联挖掘的相关 技术;2011年,IBM 公司的问答机器人"沃森"在美 国智力问答竞赛节目中大胜人类冠军,这是基于自 然语言理解和知识图谱的人工智能系统的胜利;紧 紧依赖常规的知识推理和启发式搜索[7]策略,会有 极高的计算复杂度。2016年的 AlphaGo^[8]与李世石 的围棋大战, AlphaGo 最终以 4:1 的战绩战胜李 世石,这是基于蒙特卡洛树搜索和强化深度学习的 人工智能系统的胜利,为解决智能决策问题提供了 思路[9]。2017年中科院自动化所研发的"先知 1.0" 在兵棋推演挑战赛中以7:1的战绩战胜了经验丰 富的人类高手,这也是我国科研团队首次尝试将人 工智能技术应用于复杂场景非完全信息的博弈对 抗领域。

近年来,随着深度学习算法的巨大突破,如何在复杂多变的战场环境下制定高效的军事策略并进行及时调整成为新的研究方向[4-5]。以美国为代表的世界军事强国已经预见到人工智能技术在军事领域的广阔应用前景,认为未来的军备竞赛是智能化的竞赛,发布"第三次抵消战略"[6],力求在智能化上与潜在对手拉开代差。本文提出军事智能博弈对抗系统的通用设计框架,同时以计算机兵棋推演系统和作战模拟仿真系统的智能化需求为例,实现军事作战决策的泛化应用。

1 军事博弈对抗系统的智能化需求

军事博弈对抗是通过模拟战场环境进行战争 学习、战争推演和战争预判的重要手段,目前计算 机兵棋推演系统和作战模拟仿真系统是军事博弈 对抗的两种典型应用[10-12]。军事博弈对抗系统在一 定程度解决了军事训练过程中固定化模式、流程化 训练、教条式指令的问题。通过模拟真实的战场环 境、多元化的战场对抗要素提供了实战化、强对抗 的人人对抗军事训练方法和平台。

随着战争信息化程度的不断提高、战争维度的不断增加、战场对抗态势的不断复杂、战争样式的不断多样化,如何在日益复杂和信息爆炸的战场环境下快速实现信息收集、信息处理、态势判断及作战决策制定将成为影响战争走向和胜负的关键[13]。伴随人工智能技术的不断发展,特别是以深度学习为代表的智能技术在语音、图像、博弈对抗等方面取得重大突破。人工智能在以科学运算、逻辑处理为核心的计算智能领域,以图像理解、语音识别和语言翻译为代表的感知智能领域已经达到甚至超过人类水平。随着"沃森"战胜人类、谷歌无人驾驶车获得驾驶许可证,标志着以分析、推理和决策为核心的认知智能也取得长足进步。

如何将人工智能技术应用于军事训练和实际 作战当中,将直接影响未来战争的形态。面对现代 信息化联合作战条件下的爆炸式战场信息、瞬息万 变的战场态势情况,如果无法实现智能化信息处理 与态势判断,完全依靠指挥员的人力处理和经验判 断,势必会延长决策周期,从而处处受制于敌。

军事博弈对抗系统模拟整个战场对抗环境、战斗过程、预测战斗的发展和可能结果,通过指战员对战场环境的分析和预测作出战斗决策,从而达到军事作战训练和军事作战方案验证的效果。军事博弈对抗系统对抗要素全面,信息量较大,对抗态势同样瞬息万变,如果不能实现智能感知、快速决策,则指挥员在对抗过程中无法形成有效的 OODA 环,其训练过程无法贴近于真实战斗情况;目前的对抗博弈系统训练过程缺乏变化,主要表现为战术较为固定、对手水平参差不齐、对抗流程固化等,无法展现强对抗、动态变化的军事战场特点。军事博弈对抗系统如果无法实现智能化信息处理、智能态势感知、智能战场认知,则无法实现贴近真实战场的军事训练和作战方案验证。

虽然目前的人工智能技术无法实现全局的智能认知和指挥决策,许多关键技术仍然需要理论突破,但在战场信息处理、智能态势判断等方面人工智能技术已经实现具体应用[14]。将人工智能技术与军事博弈对抗系统结合,发挥人工智能的智能计算优势实现对抗训练的辅助决策,同时发挥人工智能的局部决策能力,形成动态战场对抗过程,提高对抗训练效果。同时可通过人机对抗过程,为人工智能的自主学习提供训练环境和训练数据,在军事智能博弈对抗框架下实现人机共同进步与提高。

为了设计军事智能博弈对抗系统的通用框架,

首先以计算机兵棋推演系统和作战模拟仿真系统为例,对博弈对抗系统的共性和差异进行分析。计算机兵棋推演系统的优势在于指挥、决策的训练与评估,而作战模拟仿真系统的优势在于数学/物理模型对装备级物理效应的展现[1]。不同博弈对抗系统既有共性也有一定差异。

1.1 系统架构

美军近来提出"重振兵棋"的口号,但其行动方案中均采用 HLA(高级体系结构,High Level Architecture)[15]未来建模/仿真的共同技术框架;台湾地区的"汉光兵推"采用了美军的联合战区级模拟系统 JTLS(Joint Theater Level Simulation),其本质也是离散时间作战仿真系统,因此,二者之间在底层架构的输入方面可以统一。另一方面,兵棋推演系统更侧重于人与人之间对抗的思维博弈,对作战决心、指挥过程具有参与和互动,强调"人在回路"的作用;而作战模拟仿真系统则以物理和数学建模为基础,注重模拟结果的精确性与恒定性,成果产出多集中在作战分析、战法研究、型号论证等方面,二者在底层架构的输出方面略有差异。

1.2 基础规则

计算机兵棋推演系统和作战模拟仿真系统的基础规则并不矛盾,二者均采用蒙特卡洛法建立模型计算规则[16],其中计算机兵棋推演系统是将随机数与条件进行组合,形成具有较高抽象和概括能力的"兵棋推演裁决表",以此提供给使用者进行量化分析和决策,其输入参数相对简单,但动作状态空间可以满足作战筹划需要;而作战模拟仿真系统则是将伪随机数与数学模型进行结合实现对作战过程及效果的概率模拟,其输入参数更为复杂,但最终模拟结果也更为精确。

1.3 推进机制

根据系统设计的目标,计算机兵棋推演系统和 作战模拟仿真系统的推进机制可以按照以下 5 种 机制实现:

- 1) 同步决策异步行动回合制:即红蓝双方同时下达命令,再依次执行相关命令的回合推进机制;
- 2) 异步决策异步行动回合制:即红蓝双方交替 下达命令、执行相关命令的回合推进机制;
- 3) 同步决策反应回合制:即红蓝双方同时下达 命令及约束条件、再依次根据命令执行反制命令及 行动的推进机制;
- 4) 异步决策反应回合制:即一方下达命令,另一方根据其命令执行反制命令及行动,并交替执行的推进机制;

5)时间推进即时制:从计算机实现的本质上看,时间推进即时制也是回合制的一种,所有的裁决过程也需要在队列中按顺序执行,只是时间刻度细化到秒级以下,因此,在感官上给人以一种"即时"的感觉。

将人工智能技术集成到计算机兵棋推演和作战模拟仿真系统中与人类进行对抗,并且通过不断训练和自我进化形成"类人"甚至"超人"的智能化军事决策,正成为验证人工智能技术军事作战应用潜力的重要突破方向。

2 军事智能博弈对抗系统的设计框架

基于以上分析,设计军事智能博弈对抗系统的通 用框架如下页图 1 所示。自下而上分为 3 个层次。

2.1 底层基础模型

底层基础模型是系统推演功能实现的基础,主要提供数据服务、模型基础服务以及系统外部接口服务等。数据是模型运行的基础,规则则是系统推演的核心,因此,规则知识库中系统推演规则和交互裁决规则,以及模型库中算子行为关联表就成为影响系统可信性的关键因素。为实现与外部 AI 智能体进行数据交互,还需开放底层元动作 AI 接口。

2.2 核心推演引擎

核心推演引擎起到承上启下的作用,一方面通过自动裁决引擎与底层基础模型进行数据联通,另一方面通过 Web Service 与前端应用系统进行数据交换。其中:1) 任务管理子模块实时接收前端人类或 AI 智能体下达的指令数据,解析命令并加入到消息队列中依次执行;2) 内存数据库子模块是为实现快速裁决、快速响应而基于内存运行的数据存储模块,完成对作战计划和指挥作业等活动事物的实时处理;3) 模型运行管理子模块提取任务算子实体模型、环境模型和行为规则模型,根据消息队列中的指令次序和推演时间推进博弈对抗进程,此子模块中还集成代表高级战略 / 战术的宏动作 AI 接口,为建立层次型任务规划 AI 智能体奠定基础。

2.3 前端应用系统

前端应用系统分为推演准备、推演实施、智能 快速推演、推演分析 4 个子模块。其推演准备、推演 实施和推演分析子模块是完成常规推演任务的主 体,包括想定管理、推演配置、指挥作业、导调控制、 态势显示,战报分析、复盘评估、综合研讨,而智能 快速推演子模块则提供了人机博弈、机机博弈等基 础环境,可以在作战筹划阶段提供多分枝态势分

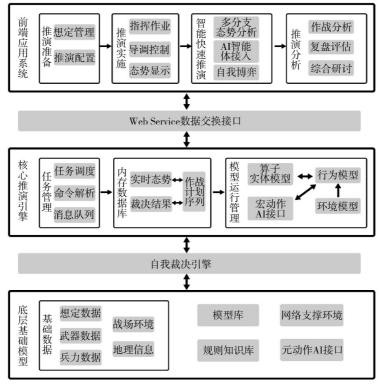


图 1 智能博弈对抗系统通用设计架构

析,强化学习过程中的自我博弈,以及 AI 智能体接入功能。

3 智能博弈对抗系统的 AI 智能体设计

从图 1 可以看出,与传统计算机兵棋推演系统 和作战模拟仿真系统相比,能够与 AI 智能体实现 无缝对接是军事智能博弈对抗系统最大的特点。这 就需要设计能够具有针对不同战场环境、具有泛化 能力的AI智能体。这种泛化能力可以从狭义和广 义两方面来理解。狭义上的泛化能力是指代码层面 具有通用性,其前提条件是平台具有统一的数据接 口标准,包括行动控制指令接口、态势数据接口、裁 决查询接口等。在此前提下,通过一个设计良好的 AI 框架和算法模型可以实现在不同想定条件下的 AI 算法迁移,提高博弈对抗系统的智能性。而从广 义上, AI 的泛化能力是指 AI 的设计技术路线及算 法基础的通用性,例如采用策略网络和价值网络结 合设计的围棋 AI 阿尔法狗, 迁移到国际象棋上训 练出来的模型同样可以战胜顶级人类选手,虽然其 基础规则代码要做大量适应性调整,但是算法实现 流程都是一致的。

AI 智能体的泛化应用能力与其算法实现过程 及模型训练方法有直接联系。当前应用的 AI 智能 体主要包含以下 3 个类型。

3.1 知识驱动型 AI

知识驱动型 AI 智能体以运筹模型计算、经验

规则推理甚至产生式规则表作为驱动行动决策。设计良好的知识驱动 AI 模型,通过对同一类决策问题(宏动作)的思维抽象建模,可具备解决同一类问题的泛化能力,例如:六角格离散化棋盘上,以通视观察、机动力、射程判断是否进行坦克行进间射击。而不经抽象,直接基于元动作构建的规则系统(如"02 驱逐舰编队于 14:00 前到达 3 号海域执行反潜任务")不具备泛化能力。这种类型的 AI 智能体是人类经验的总结,其水平的上限取决于知识本体所具有的能力,或是在有限时间内能够完成的推理和搜索深度。其适用于动作空间相对有限、领域知识比较完备的军事博弈对抗中。

3.2 深度学习驱动型 AI

深度学习驱动型 AI 智能体以深度神经网络学习海量有标注的对抗数据,从而训练出可能超越人类现有认知水平的决策网络模型。这种类型的 AI 在结构模型和算法基础方面具有较高的泛化性,只要有充足和良好的训练数据集,就可在不同场景下训练出与给定数据逼近的模型。但其有两个不足,一是数据输入稍有变化,就需要重新进行训练,例如调整某一项武器参数指标,就需要对整个模型进行重新训练;二是模型的可解释性不强,网络参数的微小改变会对结果产生巨大影响,但无法解释物理原因。因此,这类模型适用于在开放环境的某些固定场景下,通过集中获取大量对抗数据,从而完成最优策略求解,而场景一旦改变模型就将

失效。

3.3 强化学习自适应型 AI

强化学习自适应型 AI 智能体通过构建行为动 作的环境回报函数,基于多分枝态势分析不断随机 "试错"以训练出适应环境的决策网络模型,这是目 前最受关注的 AI 模型。强化学习 AI 的回报函数设 计直接决定了算法的收敛性,它不需要标注好的数 据集, 而是根据基础规则利用自博弈产生数据,再 通过海量"试错"的方法优化产生行为的模型。类似 深度学习,强化学习的框架和算法也是具有较强泛 化能力的,只要设定网络的输入、输出和回报函数, 理论上即可训练出适应于该环境的决策模型。例如 系统中增加了一种新式武器,采用强化学习自我博 弈的 AI 智能体,将会更快检验出该武器的作战效 能。相对于前两种 AI,强化学习自适应 AI 具有更强 的迁移能力,但也需要进行重新训练。并且两种想 定条件下的情报获取及算子操作动作相似度越高, 则强化学习迁移模型收敛的可能性就越大。

本文采用强化学习自适应型 AI,将深度神经网络和强化学习博弈优化决策理论和作战指挥行为知识相结合,结合数据驱动和知识驱动的学习,进行对抗空间特征提取、态势认知和策略优化,并利用生成对抗网络自动生成大量高质量对抗数据,通过强化学习自我博弈提高对弈水平。其实现结构如图 2 所示。

4 军事智能博弈对抗系统的发展思路

针对军事智能博弈对抗系统的发展,提出几点思路:

1)加强相关军事活动的数据采集。数据是支撑军事智能博弈对抗系统的基础,加快训练数据采集系统研发,在各种武器平台和训练设施设备中嵌入数据采集模块,是智能博弈对抗系统的推广基础。为了获得比较全面的数据,数据采集通常要贯穿对

抗的整个过程。

- 2) 研发适合我军使用的智能博弈对抗模型。着 眼未来战场和我军实际,制定和完善相应规则,融 人云计算、人工智能、虚拟现实等先进技术理论,走 出具有我军特色的智能博弈对抗发展之路。
- 3) 重视 AI 智能体培养和相关军事数据积累。 开展 AI 智能体研究,通过持续的对抗推演对人工 智能系统和各类无人化作战平台进行培养训练, 是提升战斗力的重要方式。长期实测数据的积累 同样不可或缺,为训练和培养智能体提供真实的 数据源。
- 4) 建立符合我军实际的博弈对抗想定设计、算子数学模型和量化赋值体系,注重在兵器性能、作战要素等诸多方面进行长期的数据量化和信息积累,实现各类军事数据采集、存储和应用。
- 5)运用推演结果验证计划方案,助力完善战略战术。将人工智能技术运用到博弈对抗,结合深度强化学习、大数据等技术,培养"智能蓝军",通过大量的人机对抗、机机对抗,产生创新的战法训法,提高指挥员作战决策水平。
- 6) 拓展博弈对抗在新型装备研发领域应用范围。结合使命任务设定不同条件进行推演,以获取装备发展指标验证,运用验证结论修正装备研发方向,辅助装备体系设计。

5 结论

智能化是军事博弈对抗系统发展的必然趋势,随着深度学习、强化学习技术的不断进步,人工智能技术在博弈对抗领域的应用价值也越来越大。一直以来,由于不同军事博弈对抗系统在推进机制、数据结构、系统架构等方面具有明显的差异,因此,AI 智能体很难直接实现跨系统、跨平台的泛化应用。基于 AI 智能体的自主学习能力和算法模型训练方法,具有在不同战场环境下生成高水平作战决

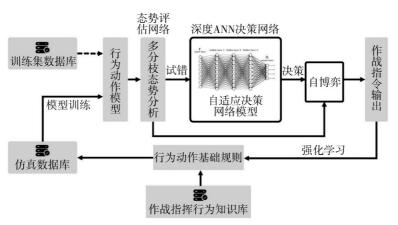


图 2 基于自我博弈的深度强化学习方法

策的迁移能力,因此,设计一套集成 AI 智能体的军事智能博弈对抗系统具有重要的意义和价值。本文提出一种将人工智能技术集成到计算机兵棋推演和作战模拟仿真系统为代表的博弈对抗系统中与人类进行对抗,并能通过不断地训练和自我进化,形成智能化军事决策的军事智能博弈对抗系统通用设计框架,并分析了该框架下 AI 智能体的设计思路。该框架在算法学习和模型训练过程中体现出较强的泛化应用能力,可以支撑各种类型的 AI 智能体无缝接入,最终建立一个军事智能博弈对抗生态链,为指挥训练提供技术支撑,促进智能博弈对抗技术向军事应用快速转化。

参考文献:

- [1] 胡晓峰,杨镜宇,司光亚.战争复杂系统仿真分析与实验 [M].北京:国防大学出版社,2008.
- [2] 杨镜宇,司光亚,胡晓峰.信息化战争体系对抗探索性仿 真分析方法研究 [J]. 系统仿真学报,2005,17(6),1469-1472.
- [3] 蔡亚梅.人工智能在军事领域中的应用及其发展[J].智能物联技术,2018,1(3):41-48.
- [4] 周海瑞,李皓昱,介冲. 美军联合作战指挥体制及其指挥控制系统[J]. 指挥信息系统与技术,2016,7(5):10-18.
- [5] 王振宇, 马亚平, 李柯. 基于作战模拟的联合作战效能评估研究[J]. 军事运筹与系统工程, 2005, 19(4):62-66.

- [6] 李頔飞.美"第三次抵消战略"对我国科技创新的启示[J]. 数码设计,2018,17(4):114.
- [7] ALLIS L V.Searching for solutions in games and artificial intelligence [D]. Maastricht Netherlands; University Limburg, 1994.
- [8] TAO J Y, WU L, HU X F.Principle analysis on A1phaGo and perspective in military application of artificial intelligence [J]. Journal of Command and Control, 2016, 2(2):114–120.
- [9] WANG F Y, ZHANG J J, ZHANG X H. Where does AlphaGo go: from church-turing thesis to alphaGo thesis and beyond [J].IEEE/CAA Journal of Automatica Sinica, 2016, 3 (2): 113-120.
- [10] 韩志军,柳少军,唐宇波.计算机兵棋推演系统研究[J]. 计真机仿真,2011,28(4):10-13.
- [11] 彭春光, 鞠儒生, 杨建池, 等. 现代兵棋推演技术分析 [J]. 系统仿真学报,2009,21(S2):97-100.
- [12] 胡晓峰, 范嘉宾. 兵棋演习对抗概论[M]. 北京: 国防大学出版社, 2012.
- [13] 郭圣明, 贺筱媛, 胡晓峰, 等. 军用信息系统智能化的挑战 与趋势 [J]. 控制 理论 与应用, 2016, 33 (12): 1562-1571.
- [14] 张晓海,操新文.基于深度学习的军事智能决策支持系统[J]. 指挥控制与仿真,2018,40(2):1-7.
- [15] 钟蔚,龚建兴,郝建国. HLA Evolved 规范研究分析[J]. 系统仿真学报,2011,23(4):691-696.
- [16] 吴永波,何晓晔,谭东风.军事建模仿真中概念模型定义 比较[J].火力与指挥控制,2007,32(11):49-53.

(上接第 115 页)

直升机在遇到突发威胁时能快速地为运输直升机 重新规划一条新航线,在武装直升机将威胁消除后 能快速引导其与运输直升机再次会合问题,提出一 种改进的人工势场法,该方法通过在传统的人工势 场方法基础上,引入 Mamdani 型模糊控制算法解决 传统人工势场法会出现局部极小值的问题,能够适 应多障碍复杂的环境,且可以安全避障。以改进的 人工势场方法为基础,采用时间协同引导控制策 略,将武装直升机与运输直升机会和协同引导问题 简化为求解会合时间和会合点问题。通过仿真试验 表明,该方法算法简单、有效,具有很好的工程性, 能很好地解决两栖作战复杂环境下的武装、运输直 升机协同引导问题。

参考文献:

- [1] 佚名.武装直升机的作战[J].兵器知识,2011,33(11):25-30.
- [2] SALICHS M A, MORENO L. Navigation of mobile robot: Open questions[J].Robotica, 2000, 18(3):227-234.
- [3] ARIA T, PAGELLO E, PARKER L E. Advances in mul-

- ti-robot systems [J].IEEE Transact ions on Robotics and Automation, 2002, 18(5);655-661.
- [4] KOHTARO S, MASAKI F, JENSSTEFFEN G, et al. Obstacle avoidance and path planning for humanoid robots using stereo vision [C]// Proceedings of the 2004 IEEE International Conference on Robots & Automation, 2004: 592-597.
- [5] 祁晓明,魏瑞轩,沈东,等.基于运动目标预测的多无人机分布式协同搜索 [J]. 系统工程与电子技术,2014,36 (12):2417-2425.
- [6] 许源. 结合粒子群算法和改进人工势场法的移动机器人混合路径规划 [D]. 杭州:浙江大学工业控制研究所,2013:12-16.
- [7] KHATIB O. Real-time obstacle avoidance for manipulators and mobile robots [J]. The International Journal of Robotics Research, 1986, 5(1):90-98.
- [8] 杨丽春. 基于改进人工势场法的无人机在线航路规划算法 [J].电子技术应用,2018,44(4):5-9.
- [9] 刘琨.基于改进人工势场法的无人船路径规划算法[J].海南大学学报(自然科学版),2016,34(2):99-104.
- [10] 韩卫华.Mamdani 型常用模糊控制算法间的常用算子关系[J]. 四川师范大学学报(自然科学版),2008,31(2): 150-154.