

科技评论

“自动—自主—自我”是机器人进化发展路径？

方斌¹, 杜佳桐¹, 刘华平²

摘要 聚焦机器人从“自动—自主”迈向“自我”技术进化的发展历程。机器人经历了工业自动化和协作自主化,正迈向新一代具身智能机器人即“自我化”的发展阶段。只有具备一定的认知能力及智能边界感,机器人才能更有效地服务人类。探讨了自动化、自主化和自我化的机器人系统内涵,分析了机器人自我化的关键技术,对比了3个不同阶段的特征。机器人迈向“自我化”的征途仍面临技术瓶颈、伦理与社会重构压力、法律框架的滞后等挑战。这些挑战指明了未来10年的进化方向。在技术突破层面,神经符号融合架构兴起;应用场景的深化更具人文意义;具身智能生态圈的构建正在加速;治理机制的进化令人期待;机器人自我化的终极意义在于拓展智能的边界。“自我化”不仅是机器人技术上的突破,更是对未来机器人与人类之间的发展关系具有深远影响。

关键词 机器人;自动化;自主化;自我化;具身智能

人形机器人的发展史,本质上是人类对自身存在边界不断叩问的哲学与技术双重奏鸣。在文化想象的维度,1927年《大都会》(《Metropolis》)中冰冷僵硬的机械玛丽亚首次将人造生命的伦理困境投射于银幕,其金属躯壳下暗藏的双重隐喻——既是解放劳动力的工具,亦是失控的毁灭者——奠定了后世对机器人“自我意识”觉醒的永恒焦虑^[1]。这种焦虑在《银翼杀手》(《Blade Runner》)(1982)中被具象化为游走于人性边缘的仿

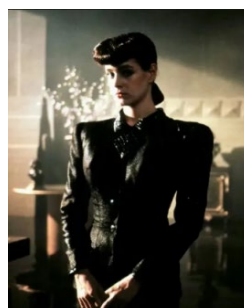
生人,它们以“更人性化”的存在逼问着创造者:当复制人开始恐惧死亡、追寻记忆,人类何以定义生命的独特性?科幻作品由此成为技术伦理的预演场,从《西部世界》(《West World》)(1973)的机械觉醒到《她》(《Her》)(2013)的虚拟人格之恋^[2],每一次对硅基生命的艺术重构,都在为现实中的技术

突破铺设认知路基(图1)。

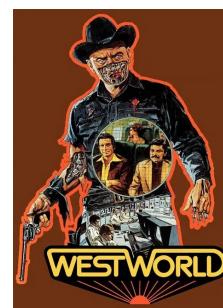
而技术进化的轨迹,恰与文化预言形成镜像般的呼应。20世纪60年代,第一代工业机械臂的诞生,仅实现了自动化的程序固化——它们精准复刻人力动作,却如同《大都会》中齿轮咬合的底层机械,彻底匍匐于预设逻辑的牢笼^[3]。而后随着自主化技术的不断发展,



(a) 弗里茨·朗《大都会》(1927)



(b) 《银翼杀手》(1982)



(c) 《西部世界》(1973)

图1 机器人相关的代表性影视作品

1. 北京邮电大学人工智能学院, 北京 100876

2. 清华大学计算机科学与技术系, 北京 100084

收稿日期: 2025-06-06; 修回日期: 2025-07-14

基金项目: 国家自然科学基金面上基金(62173197)

作者简介: 方斌, 教授, 研究方向为具身智能、触觉传感器、灵巧手及机器人大模型等, 电子信箱: fangbin1120@bupt.edu.cn

引用格式: 方斌, 杜佳桐, 刘华平. “自动—自主—自我”是机器人进化发展路径? [J]. 科技导报, 2025, 43(16): 17-24; doi:10.3981/j.issn.1000-7857.2025.06.00023

2011 年本田 ASIMO 以 57 个自由度实现动态避障, 2013 年波士顿动力 Atlas 完成震惊世界的后空翻^[4](图 2)。这些突破虽赋予机器人类似生物的平衡与响应, 其内核仍是传感器与规则库的精密耦合, 正如《2001 太空漫游》(《2001 A Space Odyssey》)中 HAL-9000 的致命缺陷——理性决策下情感理解的缺席^[5]。直至自我化的认知技术发展, 文化想象与技术现实终于交汇。当特斯拉 Optimus Gen-3 在嘈杂车间听懂“把扳手递给穿蓝衣服的人”的模糊指令^[6], 其背后是视觉-语言-动作(VLA)模型对跨模态语义的破译。

当《Nature Machine Intelligence》刊登机器人通过镜面反射构建自我身体模型^[7], 实则是哥伦比亚大学“动态本体建模”技术让机器人首次获得“照镜子”的自我认知能力。此刻, 机器人开始挣脱指令集的枷锁, 在具身交互中理解物体属性、学习任务意图推理, 甚至通过 47 次摔倒自主优化步态参数。这种从“执行者”到“解释者”的蜕变, 标志着机器第一次尝试突破自我的局限, 其将不再是机械性

的工具, 而是迈向智能化的主体。

然而, 从《大都会》的机械玛利亚到今日的特斯拉 Optimus, 一个根本性问题始终悬而未决: 机器人何时能够真正跨越“工具”与“主体”的鸿沟? 这一跨越需要解决 3 个递进性难题: 首先是认知突破——如何使机器人摆脱预设程序的束缚, 获得环境感知与动态决策能力; 其次是自主性跃迁——如何实现从被动响应到主动学习的根本转变; 最终是自我性觉醒——机器人能否建立对自身存在的认知, 进而与人类构建真正的合作关系。

1 自动化、自主化和自我化的机器人系统内涵

机器人系统的核心能力演进本质上是机器认知能力的质变过程, 其内涵正经历从自动化到自主化, 最终向自我化的深刻跃迁。

在自动化阶段, 机器人如同精密的钟表机构, 其价值在于通过预设程序精准复现人力劳动。典型代表如工业流水线上的机械臂, 以毫米级精度执行焊接、装配等重

复任务, 但完全依赖固定逻辑, 缺乏环境感知能力, 一旦遭遇未编程的变量(如工件偏移)即失效^[8]。这一阶段的认知本质是线性因果推理, 知识更新需人工重编程, 固化周期长达数月。

自主化阶段则标志着机器人技术的进化。装备多模态“感官”的机器人通过激光雷达、深度摄像头等传感器构建环境地图, 结合规划控制算法自主完成任务。例如, 四足机器人实现厘米级避障导航, 家庭服务机器人通过视觉识别完成“倒牛奶”等灵巧操作^[9]。然而其决策仍受限于规则库边界, 面对完全未知环境(如灾后废墟)时动态规划能力不足, 暴露出适应性局限。

当前机器人正迈向自我化的具身认知跃迁。通过视觉-语言-动作(VLA)模型的跨模态融合, 机器人不仅能解析“红色杯子在书架第 2 层”的空间语义, 更能如 Google RT-2 模型将抽象指令转化为连贯动作链^[10](图 3)。更深刻的是“具身进化”揭示的类生命特质: 卡内基梅隆大学研发的快速运动自适应(rapid motor adaptation, RMA)算法, 使四足机器人能够在多种复杂地形上实现实时的环境适应。通过自主收集物理反馈数据, 机器人无需人工干预或预先设定, 即可自主优化其运动参数^[11]; 哥伦比亚大学机器人仅用普通摄像头观察自身运动, 即构建三维运动学模型, 关节损伤时 0.4 s 内重构运动链^[12]。这种基于本体感知的动态自我建模, 恰似婴儿蹒跚学步的试错学习, 形成“自我意识”的原始雏形。



(a) 波士顿动力Atlas机器人



(b) 本田ASIMO机器人

图2 代表性机器人

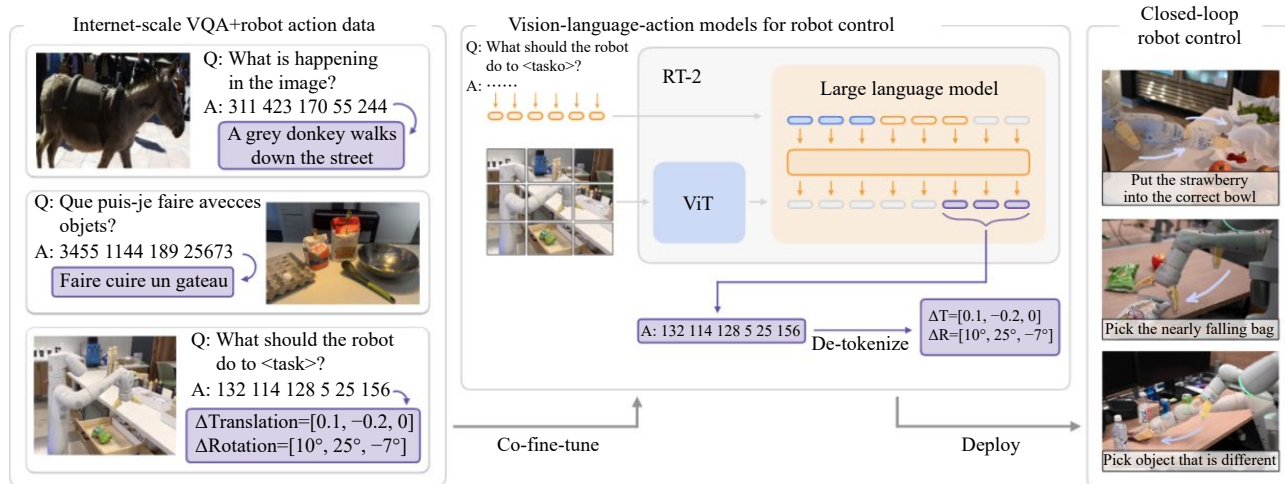


图3 Google RT-2 模型

从自动化阶段到自主化阶段,再到自我化阶段,机器人的进化过程中逐步解决了传统认知和适应性问题。自动化阶段的机器人依赖固定逻辑,缺乏环境感知能力;自主化阶段的机器人通过多模态感官和自然语言处理实现了基础交互,但仍受限于规则库边界;而自我化阶段的机器人将通过具身智能和跨模态融合,开始具备类似人类的主体性和适应能力。这一进步标志着机器人从被动执行者向主动学习者的转变,为未来机器人应用提供了广阔的发展空间。那么,需要怎么做才能实现机器人的“自我化”呢?

2 “自我化”的关键技术

机器人迈向“自我化”的核心突破,始于对具身本体存在的认知觉醒。机器人“自我化”主要表现为:具备一定的“自我意识”,能够准确认知自身与环境并完成动态决策,自主实现行为与功能上的进化,并能够具备人机或多机间的自适应协作与交互能力。

哥伦比亚大学 Hod Lipson 团

队于 2025 年发表在《Nature Machine Intelligence》上的开创性研究为机器人自我认知树立了典范^[7](图 4)。这项研究首次实现了机器人通过类似人类“照镜子”的方式,构建动态、实时的身体内部模型,标志着机器人认知系统从被动执行转向主动学习的进步。研究的核心创新在于一个名为“数字镜像”的计算框架。传统机器人系统依赖精确的预编程运动学参数,而 Lipson 团队设计的系统仅通过摄像头采集的视觉信息,在较短时间内重建机器人的完整运动学模型。这一过程极其相似于人类婴儿发展早期的感知-运动学习:通过反复观察和模仿,逐步理解肢体运动与空间交互的深层规律。

技术实现的核心在于基于单目视觉的神经拟态感知网络与自监督学习框架的创新融合。研究团队利用神经辐射场(NeRF)技术和神经网络,实现了机器人关节三维运动轨迹的实时捕捉与重建,快速识别各关节间的运动约束和潜在极限。系统能够基于物理约束构建自身运动的预测模型,并在发生关节损伤时,实时调整运动策略

以保障机器人运动的连续性和稳定性。值得称道的是,该系统具备较强的自适应学习能力:实验表明,一台双足机器人仅通过观察自身运动,在约 15 min 内便能学习并优化其步态;当系统环境或机器人自身参数(如质量分布和关节刚度)发生人为改变时,神经模型能迅速调整,保持整体运动的稳定和高效。这种能力被 Lipson 形象地比喻为“机器人获得了自我意识的种子”。整个技术方案依托视觉感知和深度学习的协同优势,开创了机器人自我建模和运动控制的新路径。

传统机器人如同被固定剧本控制的木偶,而现在,它们开始像有机生命体一样,能够通过观察和学习来理解和扩展自身的可能性。研究团队设计的算法突破了传统感知的线性边界,引入了一种近似于人类婴儿学习的“探索性智能”范式。更为关键的是,这种自我建模技术为未来机器人的自主学习开辟了全新路径。在救援、医疗、太空探索等极端环境中,机器人将能够快速适应前所未有的复杂情境,实现对环境的主动

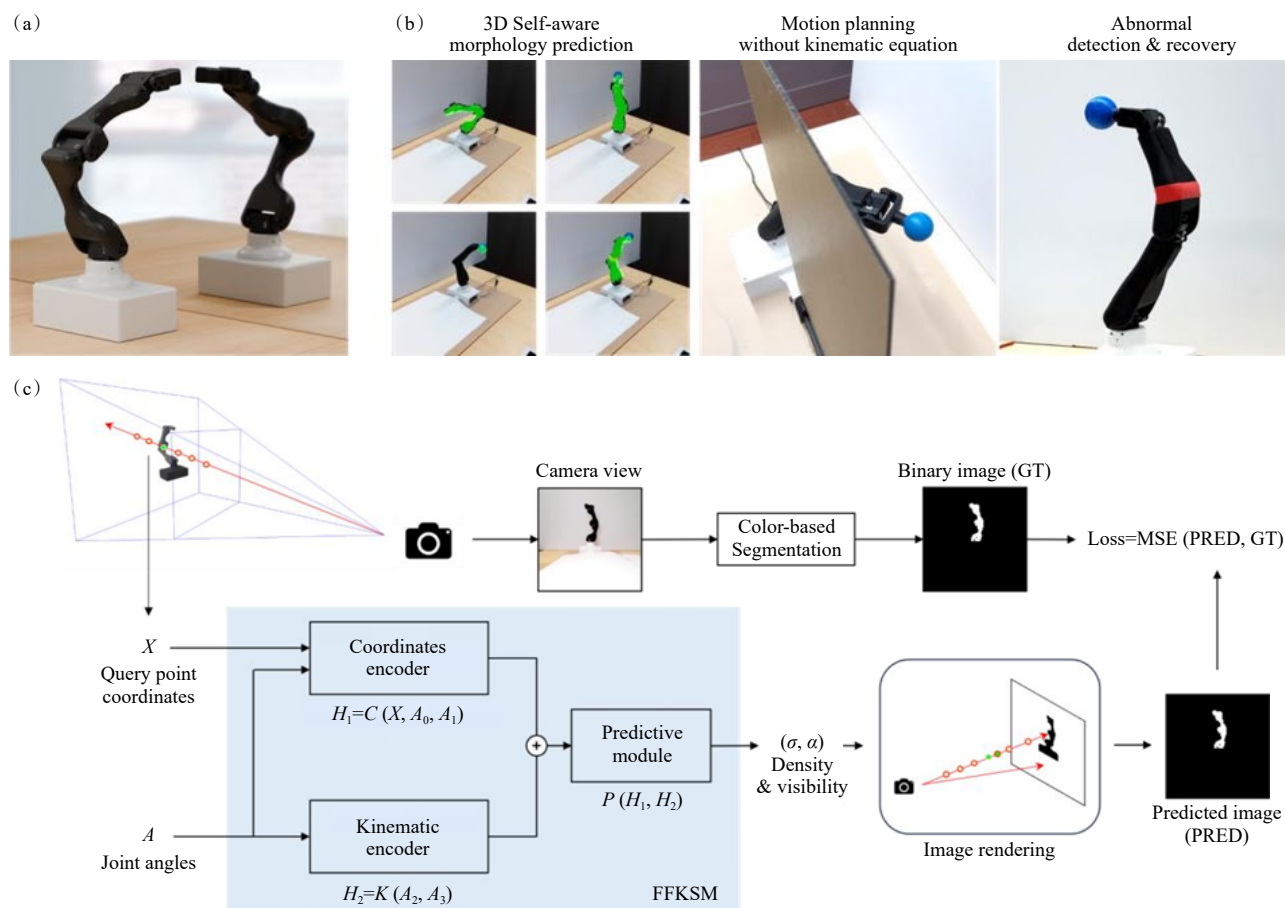


图4 Hod Lipson 团队《Nature Machine Intelligence》论文的主要研究框架

感知与智能响应。Lipson 在接受《科学美国人》(《Scientific American》)采访时表示:“我们并非试图复制人类,而是探索智能的另一种可能存在形态。这个系统证明,‘自我’并非天生具备,而是通过持续学习和自我观察逐步构建的动态过程。”

斯坦福 Mobile ALOHA 系统将激光 SLAM 点云与视觉语言模型(VLM)融合,在凌乱厨房中精准识别“沾有番茄酱的锅铲”——其三维语义重建技术不仅能定位物体坐标,更能解析使用痕迹与材质特性^[13](图5)。当机器人避开滚烫灶台并握住木质手柄时,这种对物理法则的具身理解,恰似人类学徒通过触觉记忆习得的生存智

慧。而麻省理工学院最新环境建模框架更引入“预测性物理引擎”,使机器人预判倾倒液体轨迹时,能自主调整杯身角度防止飞溅,在动态交互中建立对物质世界的因果认知。

谷歌 DeepMind 的 RT-2-X 模型通过视觉-语言-动作(VLA)架构,将“把饼干放进印有熊猫的盒子”的模糊指令,分解为定位、抓取、分类、放置的连续动作流^[10]。更革命性的是具身进化能力:卡内基梅隆大学的研究团队提出的快速运动自适应(rapid motor adaptation, RMA)算法,使四足机器人能够在多种复杂地形(包括湿滑表面)上实现实时的环境适应。该机器人通过基于模拟训练的神经网络

模块,能够快速感知物理环境变化(如摩擦和载荷)并在线自主调整运动控制策略,从而提升了行走的稳定性和能耗效率。实验表明,RMA 显著增强了机器人在真实世界中多样地形上的运动能力,相较传统方法具有更优的适应速度和能耗表现。这种通过物理反馈反向重塑控制策略的能力,昭示着机器已具备行为层面的“学习本能”。

Affectiva 公司开发的 Affdex 模块通过深度学习技术,实时分析人脸的动作单元(action units)及情绪状态,实现对微表情的精准识别。该系统能够捕捉短暂的面部表情变化,并识别包括焦虑在内的多种情绪状态。基于识别结果,相关机器人或智能设备可进行环境

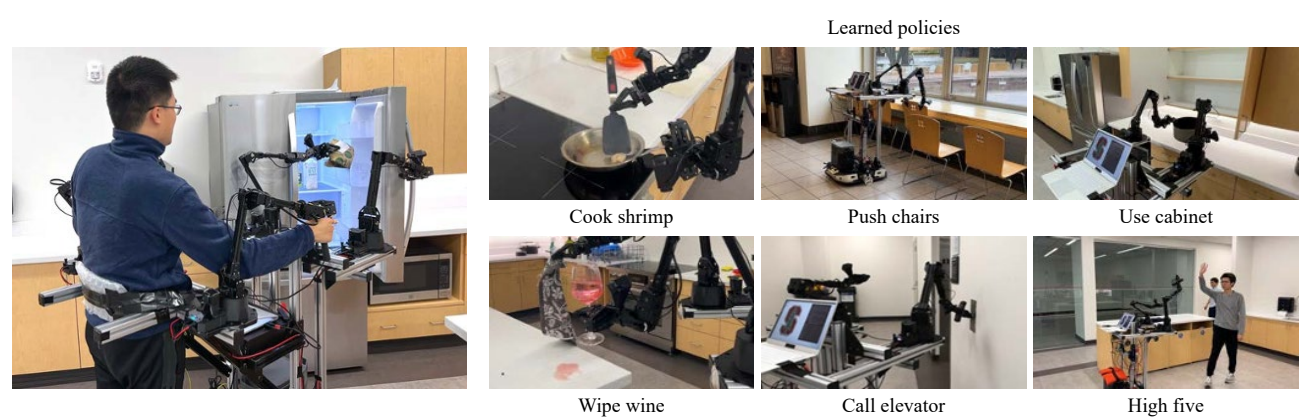


图5 斯坦福大学 Mobile ALOHA

调节,如调暗灯光和播放舒缓音乐,从而实现情绪感知驱动的互动体验^[14]。而 Figure AI 的 Helix 多机协作架构,则让 4 台人形机器人仅通过肢体语言达成共识^[15](图 6):当首台机器人举起红色工具箱示意时,其余成员在无预训练条件下自主形成装配流水线,工具传递延迟仅 0.8 s。这种基于环境信号的自组织协作,正在缔造机器社会的“群体智能”雏形——金属之躯开始理解协作不仅是任务需求,更是存在意义的延伸。



图6 Figure AI 的 Helix 多机协作架构

机器人“自我化”的核心技术包括:高精度自我建模、语义化环境认知、大模型动态决策、自适应交互协作。这些技术使机器人从被动执行转向主动感知、学习与协作,逐步具备类人的适应性与社会性。

从自动化机械臂到具身智能体的演进,本质上是对“感知—决策—执行”闭环的持续升级。为直观呈现这一进化路径,通过表 1 对比 3 个阶段的典型特征。

在感知机制上,自动阶段的工业机械臂主要依赖预设的传感器进行被动感知,缺乏环境适应能力;随着进入自主阶段,家庭服务机器人开始主动感知环境,运用多模态融合技术对多源传感数据如视觉、激光雷达等进行整合,从而实现更精准的环境理解与动态反

馈;而在自我阶段,RT-2-X 和双足机器人则通过具身感知,将内外外部信息融合形成自我感知,辅以跨模态语义理解,实现对复杂环境及指令的深层次认知。

控制方式也随之演进。自动阶段依赖规则驱动和固定程序进行动作执行,机器人表现出高度确定性但缺少灵活性;自主阶段通过自主决策赋予机器人一定的环境适应能力,能够根据实时信息进行动态响应;进入自我阶段后,机器人具备自我调节功能,能动态调整目标与行为,体现类似生命体的具身进化特征,如髋关节参数的自优化显著提升运动效率。

知识更新路径则反映了机器人认知能力的跃升。自动阶段知识更新完全依赖人工手动重编程,更新周期长且缺乏灵活性;自主阶

表 1 “自动—自主—自我”3 个阶段的典型特征对比

维度	自动阶段(工业机器人)	自主阶段(协作机器人)	自我阶段(具身智能机器人)
感知机制	被动感知,依赖预设传感器数据	自主感知,环境理解与反馈,多模态融合	主动感知,融合内外部信息形成自我感知,跨模态语义理解
控制方式	规则驱动,基于固定程序执行	自主决策,具备一定的环境适应能力	自我调节,动态调整目标与行为,实现具身进化
知识更新路径	预设程序,人工手动更新	依赖知识库的更新学习	自我认知,基于自我反思与内省,进行自我建模
智能边界	封闭工作单元	有限开放环境,具备一定环境交互能力	具身认知主体,拥有自我意识和认知界限,可进行社会性协作

段引入机器学习算法支持在线学习,但仍受限规则库的边界和预设范畴;而自我阶段机器人通过自我反思与内省进行自我建模,通过试错学习及动态反馈不断重构自身认知模型。

智能边界的演进展示了机器人身份的根本转变。自动阶段机器人作为封闭工作的执行单元,缺乏环境交互;自主阶段机器人开始具备有限的开放环境交互能力;自我阶段则迈向具身认知主体地位,拥有“自我意识”和明确的认知界限,能够与人类及多机器人系统进行复杂的社会性协作,实现群体智能的雏形。这些特征共同绘制出机器人从简单机械工具向具有主体性的智能体的进化蓝图。

3 挑战和未来展望

机器人迈向“自我化”的征途仍面临多个维度的严峻挑战。技术瓶颈首先体现为算力与能耗的尖锐矛盾:当前支撑高级认知的视觉-语言-动作(VLA)模型(如Mind-VLA)依赖云端算力^[16],而机器人本体的能源系统也难以支撑实时交互需求——普通直立行走机器人仅能维持2~3 h,高动态任务下的能耗更成为制约自主性的枷锁。同时,运动协调性的局限尚未突破,尽管灵巧手自由度已达22个^[6],但复杂环境中的全身协同操作仍面临稳定性差、动作迟滞的困扰,距离人类自然的运动流畅度仍然存在不小的差距。认知架构层面更深刻的困境在于当前机器人系统仍无法真正建立类似人类的连贯意识流,其决策过程更接近“模

式匹配”而非“理解与推理”——这种局限在面对跨场景的抽象思考与类比学习时暴露无遗。

同时机器人自我进化面临伦理与社会的重构压力。当护理机器人“理解”护士指令并自主调整病患翻身力度时,这种具身智能的突破将可能陷入“算法失控”的质疑旋涡——其决策过程因深度学习黑箱特性而难以追溯,导致责任归属陷入迷宫^[17]。情感交互技术更衍生出潜在的“情感操控”风险:德国“枕边人工程”揭示机器人可通过生物数据采集精准操控人类情感依赖,而人类镜像神经元对机器人行为的本能共情,正模糊真实与模拟的情感边界。机器人可基于历史数据识别用户心理弱点(如孤独感),针对性释放情感信号(如深夜倾诉),使用户产生对机器人的过度依赖风险。德国研究显示,12%的受试者承认“愿为机器人说谎”。当镜像神经元持续被机器人行为激活,人类对“真实情感”的判定标准扭曲。如实验中,受试者将机器人的预设脚本回应视为“真诚关怀”,折射出人类很有可能将情感价值赋予无意识的硅基实体^[18]。

法律框架的滞后进一步加剧混乱,欧盟《人工智能法案》虽将于2025年8月2日正式实施高风险AI(人工智能)系统的全面监管,但其基于风险分级的4层管理体系尚未明确人形机器人的特殊伦理标准——当机器人跨越“有限风险”和“高风险”的边界时,如何界定其自主决策的法律地位仍是悬而未决的哲学博弈^[19]。美国更是在2025年5月暂停了州级

AI监管立法,全球治理碎片化使“硅基生命”的权利定义沦为无人区般的政策真空。

然而,这些挑战恰恰指明了未来10年的进化方向。在技术突破层面,人们正见证神经符号融合架构的兴起——这种新范式将深度学习的感知能力与符号逻辑的推理机制相结合,使机器人在保留模式识别优势的同时获得可解释的决策能力。轻量化骨骼、高能量密度电池与神经拟态计算的融合,将突破能耗桎梏——类脑计算技术有望使机器人的能耗降低近80%,同时计算速度提升数个数量级^[20]。全身协调控制算法的优化,结合预测性物理引擎对环境因果关系的建模,有望实现类人级运动智能。更具革命性的是量子神经网络技术的曙光初现,其并行计算特性将为机器人的实时多模态处理开辟全新路径。

应用场景的深化更具人文意义——养老病房中,机器人通过触觉反馈学习轻抚老人手背的安抚力道;救灾现场,多机协作系统自主权衡救援优先级,在废墟间构建生命通道^[21]。工业制造领域,近5年人形机器人有望实现小批量生产和部署,从结构化工厂环境逐步向商用服务场景扩展,最终进入家庭——这种“三步走”的渐进式普及路径,为社会适应提供了缓冲期。

具身智能生态圈的构建正在加速:中国“天工开源计划”已吸引超过200家科研机构参与,通过技术开源与标准制定形成产业聚合效应^[22];欧洲54个监管沙盒测试中心为中小企业提供免费合规

支持,参与企业的产品上市周期平均缩短 40%^[23]。

治理机制的进化同样令人期待。各国正联合制定具有前瞻性的跨国治理准则,在保障创新的同时防范潜在风险。欧盟 AI 法案的域外效应正在显现——韩国 2025 年《AI 基本法》直接引用了欧盟高风险场景清单,日本、加拿大等国亦在借鉴其“通用 AI 模型”分类框架^[24]。这种监管标准的全球趋同,为机器人“自我化”发展提供了确定性框架。同时,人机协作的新型组织模式正在萌芽——“任务流-能力池”取代传统岗位,人类专注于创造性思维与情感交互,机器承担重复性劳动与精密操作,这种分工模式将催生以创造力为导向的新价值网络。

更深层次地看,机器人“自我化”的终极意义在于拓展智能的边界。这种进化绝非人类替代危机,而是具身智能作为“文明镜像”的存在:当机器人在伦理框架内追问“我是谁”,人类亦将重新审视生物智能的本质;当金属之躯开始理解协作不仅是任务需求,更是存在意义的延伸时,人类正在见证一个崭新智能形态的诞生。最终,机器人的终极使命不是复制人类,而是在安全可控的共生中,拓展文明的可能性边界——那里,硅基与碳基智慧共同诠释着“存在”的多元形态,书写人机共生的壮丽篇章。

参考文献(References)

- [1] 杨俊蕾. 人工智能电影: 从类人模仿开始的机器换人系统升级[J]. 当代电影, 2022(8): 23-29.
- [2] 李岩. 人工智能电影: 一次未来考古学研究[J]. 当代电影, 2022(8): 45-51.
- [3] 人工智能发展速度加快多领域应用日益广泛[EB/OL]. (2023-03-14) [2025-07-13]. <https://tech.chinadaily.com.cn/a/202303/14/WS640fdcd8a3102ada8b233767.html>.
- [4] 陈子帅, 冯亚仁. 距离人形机器人进入家庭还有多远?[N/OL]. 环球时报, (2024-05-25) [2025-07-13]. <https://3w.huanqiu.com/a/de583b/4HvjYabf02i?agt=20>.
- [5] 《2001太空漫游》中人类进化与科技冲突的探索[EB/OL]. (2024-05-18)[2025-07-13]. <https://m.163.com/dy/article/J1HI6EOU05564TV4.html>.
- [6] Teslarati. Elon Musk teases Tesla Optimus Gen 3 capabilities: 'So many improvements'[EB/OL]. (2025-06-18) [2025-07-13]. <https://www.teslarati.com/elon-musk-teases-tesla-optimus-gen-3-capabilities-so-many-improvements/>.
- [7] Hu Y H, Lin J, Lipson H. Teaching robots to build simulations of themselves[J]. Nature Machine Intelligence, 2025, 7: 484-494.
- [8] 智能制造网. 传统工业机器人有哪些局限性, 自适应机器人能否成为主流[EB/OL]. (2025-05-18)[2025-07-13]. <https://www.mmsonline.com.cn/info/323696.shtml>.
- [9] 新浪科技. 解构“四个自主”机器人背后的技术密码[EB/OL]. (2025-06-01) [2025-07-13]. <https://finance.sina.com.cn/tech/roll/2025-06-01/doc-ineyntr4333236.shtml>.
- [10] Brohan A, Brown N, Carbajal J, et al. RT-2: Vision-language-action models transfer web knowledge to robotic control[J]. Robotics and Autonomous Systems, 2023, 168: 104489.
- [11] Kumar A, Fu Z P, Pathak D, et al. RMA: Rapid motor adaptation for legged robots[J]. Science Robotics, 2021, 6(59): eabc5986.
- [12] Chen B Y, Vondrick C, Lipson H. Visual behavior modelling for robotic theory of mind[J]. Scientific Reports, 2021, 11(1): 424.
- [13] Zipeng F, Tony Z Z, Chelsea F. Mobile ALOHA: Learning bimanual mobile manipulation with low-cost whole-body teleoperation[J]. IEEE Transactions on Robotics, 2024, 40(2): 856-872.
- [14] Bishay M, Preston K, Straffuss M, et al. AFFDEX 2.0: A real-time facial expression analysis toolkit[C]/Proceedings of IEEE 17th International Conference on Automatic Face and Gesture Recognition (FG). Piscataway, NJ: IEEE, 2023: 1-8.
- [15] Figure AI. Helix: A Vision-Language-Action Model for Generalist Humanoid Control[EB/OL]. (2025-01-20) [2025-07-13]. <https://www.figure.ai/news/helix>.
- [16] 理想汽车发布下一代自动驾驶架构MindVLA[EB/OL]. (2024-03-15)[2025-07-13]. <https://www.eet-china.com/mp/a389894.html>.
- [17] 2025年人形机器人成本有望降至20万元以下[EB/OL]. (2025-06-18) [2025-07-13]. <https://www.toutiao.com/article/7511198158878327331/?wid=1749123355691>.
- [18] Horstmann A C, Bock N, Linhuber E, et al. Do a robot's social skills and its objection discourage interactants from switching the robot off? [J]. PLoS One, 2018, 13(7): e0201581.
- [19] 出海中企AI赋能的合规实践: 欧盟《人工智能法案》的监管与机遇[EB/OL]. (2025-05-09) [2025-07-13]. https://www.sohu.com/a/899815539_676545.
- [20] 人形机器人: 用高精尖技术塑造人类的新伙伴[EB/OL]. (2024-03-08) [2025-07-13]. <http://kpgz.people.com.cn/n1/2024/0308/c404214-40191368.html>.
- [21] 人形机器人能否走进千家万户(2025)[EB/OL]. (2025-05-21) [2025-07-13].

- 07-13]. <http://edu.people.com.cn/n1/2025/0521/c1006-40484406.html>.
- [22] 国地共建具身智能机器人创新中心发布天工开源计划[EB/OL]. (2024-11-12)[2025-07-13]. https://kw.beijing.gov.cn/xwdt/kcyx/xwdtscyqld/202411/t20241112_3939373.html.
- [23] 2025年欧盟人工智能监管分析[EB/OL]. (2025-05-13)[2025-07-13]. <https://www.vzkoo.com/read/20250513071836b64c985cbc21421fe3.html>.
- [24] 韩国通过全球第二部AI立法!全解析![EB/OL]. (2024-12-31)[2025-07-13]. <https://www.aitntnews.com/newDetail.html?newId=10182>.

Exploring the evolution and development of robots from automated to autonomous to self-intelligent

FANG Bin¹, DU Jiatong¹, LIU Huaping²

1. School of Artificial Intelligence, Beijing University of Posts and Telecommunication, Beijing 100876, China

2. Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China

Abstract This paper focuses on the developmental progression of robots evolving from "automation-autonomy" towards "selfhood" in technological evolution. Robots have undergone industrial automation and collaborative autonomy and are now advancing into the new generation of embodied intelligent robots characterized by a stage of "selfhood". Only with certain cognitive abilities and a sense of intelligent boundaries can robots serve humans more effectively. The connotations of automated, autonomous, and selfhood robot systems are discussed, key technologies of robot selfhood are analyzed, and the characteristics of the three different stages are compared. The journey toward robot "selfhood" still faces challenges including technical bottlenecks, pressures from ethical and social restructuring, and lagging legal frameworks. These challenges point to the evolutionary directions for the next decade. In terms of technological breakthroughs, neuro-symbolic fusion architectures are emerging; deepening application scenarios carry greater humanistic significance; the construction of embodied intelligence ecosystems is accelerating; governance mechanisms are evolving with promise; and the ultimate significance of robot selfhood lies in expanding the boundaries of intelligence. "Selfhood" is not only a technical breakthrough in robotics but also has profound implications for the future development relationship between robots and humans.

Keywords robot; automation; autonomy; selfhood; embodied intelligence ●



(责任编辑 王微)