

# 从“思维智能”到“行动智能”

## ——具身智能的未来发展趋势与挑战

周玄

北京国谦投资咨询有限公司 北京

**【摘要】**具身智能是人工智能向物理世界延伸的重要发展方向,使 AI 具备自主感知、决策与执行能力,实现从数字智能向“知行合一”的演进。近年来,具身智能在工业制造、医疗健康、智能交通、家用服务等领域展现出广阔的应用前景,推动智能系统从单一任务执行向多任务协同和环境自适应能力提升的方向发展。然而,具身智能仍面临技术体系完善、数据获取与泛化、软硬件协同优化以及产业化落地等多重挑战。本文系统梳理了具身智能的核心技术架构,分析其关键应用场景,探讨当前发展瓶颈,并展望未来发展趋势及其对社会的深远影响,以期为相关研究和产业发展提供参考。

**【关键词】**具身智能; 人工智能; 多模态感知; 自主决策; 人机协作; 产业应用

**【收稿日期】**2025 年 2 月 15 日

**【出刊日期】**2025 年 3 月 28 日

**【DOI】**10.12208/j.sdr.20250004

### From "Cognitive intelligence" to "Embodied intelligence" future development trends and challenges of embodied intelligence

*Xuan Zhou*

*Stable Fund, Beijing*

**【Abstract】** As a crucial advancement in artificial intelligence (AI), embodied intelligence extends AI capabilities into the physical world, enabling autonomous perception, decision-making, and execution. This evolution marks a transition from purely digital intelligence to an integrated "perception and action" paradigm. In recent years, embodied intelligence has demonstrated significant potential across various domains, including industrial manufacturing, healthcare, intelligent transportation, and home services. However, it still faces multiple challenges in technology, data acquisition, engineering, security, and industrialization. This paper systematically reviews the core technical framework of embodied intelligence, analyzes its key application scenarios, and explores the existing development bottlenecks. Finally, it presents an outlook on future trends, including the integration of cognitive and action intelligence, the construction of human-machine collaboration ecosystems, and the broader societal impact of embodied AI. This study aims to provide insights for future research and industry development in the field of embodied intelligence.

**【Keywords】** Embodied intelligence; Artificial intelligence; Autonomous learning; Multimodal perception; Industrial automation; Human-machine interaction

#### 1 引言

具身智能 (Embodied Intelligence) 是人工智能 (AI) 向物理世界拓展的重要方向,旨在赋予 AI 实体形态,使其具备自主感知、决策和执行能力,实现从“思维智能”到“行动智能”的转变。随着深度学习、大模型、计算机视觉等技术的快速发展, AI 在

数据处理和智能推理方面已表现出超越人类的潜力,但其在真实物理环境中的感知与交互能力仍存在诸多局限。具身智能的提出,正是为了突破这一瓶颈,使 AI 能够从单纯依赖数据训练,进一步发展出在现实世界中感知环境、执行任务,并通过持续交互实现自我进化的能力<sup>[1]</sup>。

目前，具身智能已在多个行业展现出广泛的应用前景，包括工业制造、医疗健康、智能交通、家庭服务和灾难救援等。例如，智能协作机器人能够与人类协同作业，提升生产精度和效率；智能护理机器人可以辅助病患恢复，提高医疗服务质量。然而，其落地仍然面临一系列挑战。从技术角度来看，多模态感知、环境理解、自主学习和运动控制等核心能力仍需进一步突破；从工程实现角度而言，硬件适配性、能耗管理、实时交互等问题限制了其大规模应用。此外，数据安全、隐私保护、伦理规范和法律监管等问题，也成为影响具身智能推广的重要因素。本文将从具身智能的核心技术体系入手，梳理其发展路径，探讨其在工业、医疗、交通、服务等领域的应用潜力，分析其当前面临的技术和产业化挑战，并展望其未来的发展趋势。

## 2 具身智能的理论基础与学科融合

具身智能的兴起得益于多个学科的交叉融合，涵盖人工智能、机器人技术、认知科学、机械工程和神经科学等领域。其核心目标是赋予人工智能“身体”，使其不仅具备思维和推理能力，还能感知、适应并与物理世界互动，实现从“思维智能”向“行动智能”的跃迁。尽管具身智能已成为人工智能发展的重要方向，不同领域的研究者对其定义和应用仍持有不同理解，导致该概念在学术界尚未形成完全统一的共识。

### 2.1 具身智能的概念与核心观点

目前，关于具身智能的研究主要存在两种学术观点。第一类观点强调智能并非仅依赖计算能力，而是在与环境的持续交互中不断发展和进化。例如，清华大学刘华平教授等人在《基于形态的具身智能研究：历史回顾与前沿进展》中指出，具身智能通过身体与环境的相互作用，在信息感知和物理操作过程中不断形成和优化智能<sup>[2]</sup>。同样，上海交通大学卢策吾教授认为，智能体的适应性来源于其与环境的交互过程，这种交互不仅提升了行为智能，还增强了其生存和发展能力<sup>[3]</sup>。

第二类观点则更关注具身智能的实际应用价值，强调其在解决现实问题中的重要性。斯坦福大学李飞飞教授提出，具身智能的关键在于 AI 如何在物理环境中执行任务，与环境的交互需求直接决定了其智能水平<sup>[4]</sup>。此外，中国科学院院士姚期智认为，通

用人工智能（AGI）的未来发展离不开具身智能，只有能够与现实世界交互的 AI，才能真正实现通用智能并完成复杂任务<sup>[5]</sup>。

尽管研究者的关注点有所不同，但他们普遍认同：智能不仅是对数据的处理和推理能力，还包括对环境的感知、理解和自主操作能力。具身智能的提出，正是为了突破传统人工智能在物理交互方面的局限，使 AI 具备更强的环境适应能力和自主行动能力。

### 2.2 具身智能的学科交叉融合

具身智能的发展不仅依赖于人工智能技术的进步，还涉及多个学科深度融合。其核心技术体系主要来源于以下几个关键领域：

**人工智能与机器学习：**AI 为具身智能提供了决策推理、环境感知和自主学习能力。深度学习、强化学习及大模型等技术，使智能体能够在复杂环境中不断优化自身的行为模式，并逐步提升适应能力。**机器人技术与自动化：**具身智能的“身体”依赖于机器人技术的发展。先进的机械结构、运动控制系统和传感技术，使 AI 能够精准地执行物理操作，实现自主导航、操控和人机交互。**认知科学与神经科学：**具身智能不仅是技术的集成，更涉及智能的本质。认知科学研究表明，智能的形成依赖于感知、学习和行动的闭环，而具身智能正是这一理论的具体实践。神经科学的研究也为具身智能提供了关于生物智能如何运作的理论基础，有助于 AI 系统模拟生物大脑的学习与决策机制。**嵌入式系统与边缘计算：**为了提升实时响应能力，具身智能设备通常需要在本地进行数据处理和决策。嵌入式系统与边缘计算技术的结合，可以降低延迟，提高能效，使具身智能体能够更高效地执行任务。

正是这些学科的交叉融合，使具身智能成为 AI 迈向现实世界的重要桥梁。随着各领域研究的不断深入，具身智能的应用边界正在持续拓展，其智能水平也在稳步提升，为未来人工智能的发展提供了新的可能性。

## 3 具身智能的应用场景

具身智能正逐步成为人工智能发展的重要方向，使 AI 不仅具备推理和学习能力，还能在真实物理环境中自主感知、决策和执行任务。随着人形机器人、多模态感知和自主决策技术的发展，具身智能在多

个行业的应用日趋成熟。2024年，OpenAI与Figure公司联合推出Figure 01机器人，该机器人能够听、说、与人类进行交互，并执行多种任务。同年，升级版Figure 02借助GPT-4o大模型进一步优化了感知、运动控制和操作能力，使其在人机交互和任务执行方面更加接近人类<sup>[6]</sup>。与此同时，无人驾驶技术的商业化进程也在加快，百度的萝卜快跑无人驾驶出租车已在多个城市投入运营，其单车日订单量已接近传统出租车水平。具身智能正在从实验室研究走向实际应用，并逐步融入各行各业，为人工智能拓展更广阔的发展空间。

### 3.1 工业制造：提升自动化与人机协作能力

在工业制造领域，具身智能机器人正在逐步取代传统的固定式机器人，实现更加灵活和智能的生产模式。与依赖预设程序执行任务的传统机器人不同，具身智能机器人能够自主感知环境并适应不同的任务需求。例如，协作机器人可以在装配、物料搬运等环节辅助工人工作，不仅提高了生产效率，还降低了工人的劳动强度，并优化了生产流程的安全性和灵活性。

### 3.2 自动驾驶：增强环境适应性与决策能力

具身智能技术的突破为自动驾驶领域带来了更高的安全性和环境适应能力。相比传统自动驾驶系统，具身智能利用多模态传感器融合技术，能够更精准地分析周围环境，并根据实时变化调整驾驶策略。例如，百度的萝卜快跑无人出租车具备高度自主的驾驶能力，能够应对复杂的城市交通环境，提高用户体验和整体运营效率。

### 3.3 家庭服务：智能化生活助手

在家庭场景中，具身智能机器人正逐步实现多功能化，不仅能够执行家务，还可提供情感陪护，提升生活便利性。例如，现代家务机器人已能够自主学习用户习惯，完成扫地、洗碗等日常任务。同时，智能陪护机器人正逐步应用于老年人和特殊人群的护理，具备语音交互、健康监测等功能，以满足用户个性化需求。

## 4 具身智能技术面临的挑战

尽管具身智能被视为迈向通用人工智能（AGI）的重要路径，但在技术实现、数据采集、软件生态和硬件适配等方面仍面临诸多瓶颈。目前的具身智能系统虽然能够执行特定任务，但在自主学习、泛化

能力和环境适应性上仍存在明显不足。此外，高昂的数据成本、软硬件融合的复杂性以及能效优化的限制，也对具身智能的进一步发展构成障碍。

### 4.1 技术挑战：自主智能能力有限

当前，具身智能主要依赖大模型驱动，虽然能够执行较为复杂的任务，但仍未达到真正的自主智能水平。目前的系统仍属于“辅助智能”（Assisted Intelligence），而非完全独立的“自主智能”（Autonomous Intelligence）<sup>[7]</sup>。现阶段，具身智能系统仍需要大量人工干预，例如数据采集、算法优化和模型训练等，无法像人类一样自主理解任务、设定目标并通过交互不断学习新技能。

此外，具身智能尚未完全实现“感知-决策-行动”的闭环智能。人类思维可分为“快思考”（快速、直觉式决策）和“慢思考”（深度推理与分析）<sup>[7]</sup>。然而，当前的具身智能系统主要依赖大模型的推理能力，即“慢思考”，在应对突发情况或实时调整行为时仍然较为滞后。例如，在复杂的交通环境中，自动驾驶汽车在遇到突发状况时仍可能需要人工接管，而不像人类驾驶员那样能够迅速做出本能反应。这种限制使得具身智能在高动态环境中的应用仍然受限。

### 4.2 数据挑战：高质量数据获取困难

训练具身智能需要海量高质量的数据，但获取这些数据的成本极高。例如，采集一小时的自动驾驶数据成本高达180美元，是模拟数据成本的100倍<sup>[9]</sup>。为了降低数据成本，行业内开始探索合成数据（如仿真环境生成的数据）来补充真实数据。然而，合成数据面临“现实差距”（Reality Gap）问题，即模拟数据与真实世界的数据仍存在较大差异，这可能导致模型在实验环境中表现良好，但在真实应用场景中失效。例如，在实验室训练的机器人可能能够精准执行任务，但在实际应用中，由于光照、噪音、材质等因素的不同，可能导致机器人识别错误或行为不稳定。

### 4.3 软件挑战：缺乏统一标准，开发门槛高

具身智能的软件生态尚处于发展初期，面临缺乏统一操作系统和开发标准的问题。目前，不同厂商的具身智能设备往往采用各自独立的软件架构，使得开发者需要针对不同平台进行适配，增加了技术实现的难度，并制约了整个行业的协同发展。

此外，软件如何更好的服务硬件仍是一个难题。具身智能系统需要处理大量数据，并做出实时决策，但现有的软件架构难以满足这一需求。许多强化学习算法在仿真环境中表现优秀，但在真实环境中的适应性较差，导致机器人难以灵活应对现实场景。当前的软件开发仍然偏向于理论研究，而真正适用于商业化应用的软件体系尚未成熟。

#### 4.4 硬件挑战：耐用性与能效问题

具身智能的硬件发展仍面临诸多挑战，主要涉及耐用性、能源效率和软硬件协同等方面。

耐用性问题仍是具身智能在长期应用中的一大瓶颈。机器人往往需要在复杂环境中长时间运行，但现有硬件设备的稳定性和耐久性仍待提升。例如，工业机器人在高强度生产环境下持续运作时，机械臂、传感器等关键部件的寿命可能低于预期，长期使用后可能出现精度下降或磨损，影响整体系统的可靠性和效率。

能源效率问题限制了具身智能系统的长时间运行能力。由于具身智能需要大量计算资源进行实时感知、决策和执行，目前的电池技术和计算架构尚无法有效支撑高功耗设备的长时间自主运作。例如，人形机器人在执行任务时往往需要频繁充电，这在实际应用中造成了较大的使用限制。低功耗 AI 芯片、能源管理系统，新型电池技术将成为未来重要的技术发力点。

此外，许多具身智能系统采用的是“先开发硬件，再适配软件”的模式，导致软硬件之间的协同性较低，影响整体系统的响应速度和性能。例如，传感器采集的环境数据与执行机构的控制信号可能存在延迟，导致机器人在复杂环境中难以及时作出精准决策。未来应聚焦于加强软件与硬件的深度融合优化数据传输和处理机制，提高系统的实时性和稳定性，从而增强具身智能在高动态环境下的适应能力。

### 5 具身智能的发展趋势

随着算法优化、数据获取方式改进、标准化软件生态建设、低功耗硬件研发等方面的持续突破，具身智能的技术瓶颈有望逐步得到解决，加速其向通用人工智能（AGI）的演进，并在日常生活中发挥更广泛的作用。具身智能的发展将主要围绕以下几个方向展开：

增强自主学习能力。具身智能将更加依赖大模

型的能力，减少对人工设定规则的依赖，实现更高水平的自主学习。例如，英伟达推出的 VIMA 系统已经能够通过视觉和文本提示自主完成复杂任务，这种技术的突破将极大提升具身智能的泛化能力，使其在不同环境下具备更强的适应性。

提升多模态感知与交互能力。未来的具身智能系统将融合视觉、听觉、触觉等多种感知方式，使其能够更加精准地理解环境并作出相应决策。例如，谷歌 DeepMind 推出的 RT-2 模型结合了视觉-语言-动作（VLA）技术，使机器人能够更自然地进行人机交互，并根据具体语境调整行为模式。这种多模态感知能力的提升，将进一步缩小机器人与人类在认知和行动上的差距。

加速商业化落地。随着技术的成熟，具身智能的商业化应用进程正在不断推进。例如，百度与优必选合作，将百度大模型接入 Walker S 人形机器人，实现任务调度和智能交互。未来，具身智能将在制造、医疗、交通、家庭等多个领域实现规模化落地，为社会带来更高效、安全和便捷的智能化服务，进一步加快智能社会的建设步伐。

具身智能正推动人工智能迈向更高水平，使 AI 从数字空间深入融合至现实世界，实现从认知智能到行动智能的转变。随着技术创新和行业生态的不断完善，具身智能有望成为智能社会的核心驱动力，促进人机协作、智能化生产和智慧生活的发展。未来，随着自主学习、多模态感知、商业化落地的持续推进，具身智能将进一步释放其潜能，为各行业带来深远影响，并推动人类迈向更加智能化的未来。

### 6 结论

近年来，随着大模型、多模态感知、自主学习等技术的发展，具身智能在工业制造、自动驾驶、家庭服务等领域的应用不断拓展，展现出广阔的商业化前景。然而，当前的具身智能仍然面临一系列挑战，包括自主智能能力不足、数据获取成本高、软件生态缺乏标准化、硬件耐用性和能效问题等，限制了其在现实世界的广泛应用。

未来，具身智能的发展将围绕几个核心方向展开：自主学习能力的提升，减少对预设规则的依赖，使智能体具备更强的环境适应性；多模态感知与决策能力的增强，推动机器人整合视觉、听觉、触觉等多种信息，提高任务执行的精准度；商业化落地的

加速，随着技术突破和成本优化，具身智能将在制造、医疗、交通等行业逐步实现规模化应用。此外，行业需要建立更完善的软硬件协同体系，推动标准化软件开发环境的建设，提高系统的兼容性和可扩展性。

总体来看，具身智能正处于从实验探索向产业落地的关键阶段，尽管仍存在技术和工程上的挑战，但其在智能社会中的应用潜力不可忽视。随着技术的不断进步和生态的完善，具身智能有望突破现有瓶颈，实现更高效、更灵活、更智能的应用，为未来社会的智能化变革提供重要支撑。

### 参考文献

- [1] 中国信息通信研究院. 具身智能发展报告(2024年)[R]. 北京: 中国信息通信研究院, 2024.
- [2] 刘华平, 等. 基于形态的具身智能研究: 历史回顾与前沿进展[J]. 计算机科学, 2023, 50(6): 23-30.
- [3] 卢策吾. 具身智能的环境交互与自适应性[J]. 机器人学报, 2023, 39(7): 56-63.
- [4] 李飞飞. AI 与具身智能的未来发展[J]. 人工智能研究, 2023, 11(3): 27-34.
- [5] 姚期智. 通用人工智能的发展趋势与挑战[J]. 计算机工程与应用, 2023, 59(10): 39-45.
- [6] OpenAI, Figure Robotics. Figure 01 机器人发布[R]. 2024.
- [7] Yann Lecun. Autonomous Intelligence vs. Assisted Intelligence[R]. 2023.
- [8] Daniel Kahneman. Thinking, Fast and Slow[M]. 2011.
- [9] 北京具身智能机器人创新中心. 具身智能技术发展与挑战研究报告[R]. 2024.

**版权声明:** ©2025 作者与开放获取期刊研究中心(OAJRC)所有。本文章按照知识共享署名许可条款发表。

<https://creativecommons.org/licenses/by/4.0/>



**OPEN ACCESS**