



具身智能的崛起与定义

具身，指的是具有支持感觉和运动的物理身体。具身智能，是有身体并支持物理交互的智能体，如家用服务机器人、无人车等。具身智能机器人，则是满足具身智能的能力的机器人，可以像人类一样通过观察、移动、说话和与世界互动从而完成一系列具身任务。

机器人，是人工智能的最终解决方案。什么意思呢？

具体点说，像人一样能与环境交互感知，自主规划、决策、行动、执行能力的机器人/仿真（指虚拟环境中）是 AI 的终极形态，我们暂且称之为“具身智能机器人”。它的实现包含了人工智能领域内诸多的技术，例如计算机视觉、自然语言处理、机器人学等。

具身机器人 (Embodied Robot) 并非简单的机械装置，而是将人工智能嵌入物理实体，使其具备“身体”与“大脑”协同进化的能力。其核心在于情境感知（如视觉、触觉、环境建模）和自主决策（如运动控制、任务规划），最终通过物理动作与环境互动。

这一概念可追溯至 20 世纪 50 年代图灵对“具身性”的哲学思考，但直到近年，随着深度学习、强化学习与大模型的突破，具身智能才真正从理论走向实践。

2023 年波士顿动力公司发布的 Atlas 人形机器人，以其流畅的后空翻动作震惊世界。这个身高 1.5 米、体重 82 公斤的“钢铁舞者”，通过 3D 视觉和力反馈系统实现了前所未有的动态平衡能力。然而，当工程师试图让它在阿富汗战场废墟中执行搜救任务时，却遭遇了意想不到的困境——复杂地形导致关节磨损加剧，能源续航仅能维持 45 分钟，而士兵们更倾向于使用无人机完成同类任务。

这种“实验室完美，现实中脆弱”的矛盾，暴露出

人形机器人的先天局限。斯坦福大学机器人研究中心主任 Oussama Khatib 指出：“人类形态赋予机器人与人类环境的交互优势，但也限制了其在非结构化场景中的适应性。”数据显示，人形机器人在平坦路面的移动效率仅为轮式机器人的 63%，而在崎岖地形中的故障率则是后者的 4.2 倍。

2024 年达闼科技发布的 Cloud Ginger 智能服务机器人，标志着机器人演进的重要转折点。这个身高 1.4 米的人形机器人不仅具备自然语言交互能力，更通过云端大脑实现了跨场景的任务处理。但真正引发行业震动的是其“具身智能”架构——通过 18 个自由度的关节控制、3D 环境建模和实时决策系统，它能在商场、医院等复杂环境中自主规划路径，规避动态障碍物，并根据用户需求调整行为模式。

具身智能的核心在于“身体与环境的持续互动”。麻省理工学院媒体实验室的研究表明，这种交互机制能使机器人的学习效率提升 300% 以上。与传统机器人依赖预设程序不同，具身机器人通过传感器实时感知环境变化，结合机器学习算法动态调整策略，从而实现“感知—决策—执行”的闭环优化。

具身机器人的核心特征还包括形态适应性，具身机器人不再局限于人类形态，四足、多足、轮式等形态根据应用场景动态调整。例如，美国 Ghost Robotics 的 Vision 60 四足机器人，通过仿生关节设计可在 -40℃ 至 60℃ 的极端环境中运行，负载能力达 32 公斤。

具身机器人更重要的特征是智能决策力，具身机器人可以基于深度学习模型实现自主决策。比如波士顿动力的 Spot 已能通过强化学习自主规划巡逻路线，规避潜在危险。

军工领域的钢铁战士

人形机器人曾是具身智能的典型代表，其设计初衷