

人形

从科幻
走向现实

生产制造 · 社会服务 · 特种应用
17种场景打开人机协作新空间

当机器人走向人类
谁将定义下一个时代？

机器人

CAICT
中国信通院

「佳知」丛书

王尧 翟庆诗 董颖璇 / 编著

人形

从科幻
走向现实

生产制造 · 社会服务 · 特种应用
17种场景打开人机协作新空间

当机器人走向人类
谁将定义下一个时代？

机器人



中国工信出版集团



人民邮电出版社
POSTS & TELECOM PRESS



版权信息

COPYRIGHT

书名：人形机器人：从科幻走向现实

作者：王尧，翟庆诗，董颖璇

出版社：人民邮电出版社

出版时间：2026年6月

ISBN：9787115688620

字数：120千字

（免费书分享更多搜索雅 书）

内容提要

本书分为技术篇、产业篇、展望篇、附录四大部分。技术篇详细介绍人形机器人的基本概念、发展历程、分类和技术架构，并从“大小脑”“肢体”等方面详细介绍其涉及的关键核心技术、主流技术路线和技术演进路径。产业篇重点介绍人形机器人产业潜在的应用场景、国内外政策布局、产业链整体情况、产业分布情况和标准建设情况。展望篇则面向未来需求，详细分析了产业发展的现状和趋势，以及未来市场规模，并讨论了人形机器人的安全伦理问题，以及面临的产业发展机遇和挑战。附录介绍了国内外典型人形机器人整机产品的功能与性能情况。

本书适合人形机器人领域的技术人员、产业研究人员、潜在用户，以及对人形机器人感兴趣的人士阅读。

>>>丛书编写组

丛书顾问专家：

续合元 王爱华 史德年 石友康 许志远 何 伟

本书编写组成员：

王 尧 翟庆诗 董颖璇 张 鹤 相石磊 杨圣奇

安 静

前言 PREFACE

人形机器人作为未来产业的重要领域，是体现我国科技自立自强、发展新质生产力的标志性成果，是人工智能、机械工程、电子工程等领域融合创新的典范。凭借其类人的感知交互能力、肢体结构和运动方式，人形机器人能够快速融入各种环境，未来有望在简单重复劳动和危险场景中代替人类，在复杂技能场景中辅助人类，在商业和家庭场景中服务人类。可以预见，未来人形机器人的广泛应用将深刻改变社会形态和人们的生产生活方式，人形机器人已成为全球科技领域的发展热点。业界普遍认为，未来人形机器人有望成为继个人计算机、智能手机、新能源汽车之后的新终端，形成新的万亿级市场。

本书按照技术、产业和展望的整体脉络对人形机器人进行了多角度的介绍。

技术篇从整机出发，重点剖析了人形机器人的基本概念和发展历程，提出了智能类人、形态类人、能力类人三大基本特征，并介绍了人形机器人的不同形态和通用的技术架构；随后则按照技术架构展开，围绕其“大小脑”涉及的大模型、控制方法、开发工具，以及“肢体”涉及的执行机构、芯片、传感器、电源、新材料等具体技术进行了详细的分析，探讨了主要的技术路径和发展趋势。

产业篇对人形机器人的潜在应用场景进行了详细分析，囊括了生产制造、社会服务、特种应用3个方面的17种应用场景，解答了“人形机器人未来能做什么”这一业界热点问题；接着从政策布局出发，通过全球主要国家政策、我国国家层面政策和各地方政策3个维度，介绍了政策的布局情况；之后详细分析了人形机器人产业链的构成和产业链上、中、下游的整体情况，并介绍了国内产业的分布情况；最后重点对人形机器人标准体系的建设进行了探讨，希望能通过标准的有效规范和引导，推动行业技术创新和应用落地。

展望篇立足未来，首先分析了人形机器人当前所处的阶段现状和未来发展趋势，并对未来的产业规模进行了预测；同时，针对业界重点关注的安全和伦理问题进行了深入分析，并提出了人形机器人未来发展面临的机遇与挑战。

附录主要介绍了当前国内外重点整机产品的功能和性能现状（书中相关产品介绍排序不分先后）。

本书由中国信息通信研究院泰尔系统实验室的王晨、刘泰两位专家指导、审核，创作人员包括王尧、翟庆诗、董颖璇、张鹤、相石磊、杨圣奇、安静。此外，编著者在此也对支持本书编写的所有合作单位致以由衷的感谢。

由于编著者学术水平有限，书中难免存在表达欠妥之处，希望广大读者能够提出宝贵的修改建议。

编著者

2025年夏于北京

1 技术篇

▷第一章 整机

（一）人形机器人的基本概念

人形机器人是具备仿人形态与功能的智能机器人，拥有与人类相似的“肢体”结构、运动方式和感知方式，并且在人工智能大模型的赋能下，能够在体能、技能、智能等方面实现对人的模拟。学术与产业界对人形机器人也有一些共识定义。学术界普遍认为，人形机器人是融合了人工智能、机械工程、电子工程等多学科知识的复杂系统，它的发展代表了科技的前沿水平。产业界则更注重人形机器人的实用性和商业价值，认为它可以在工业生产、民生服务、特种作业等领域发挥重要作用，解决未来劳动力短缺的问题，是实现新质生产力的最佳手段之一。人形机器人主要具备智能类人、形态类人、能力类人三大典型特征。

1. 智能类人

人形机器人的拟人智能是其最显著的特征。一方面，人形机器人能够实现拟人化的感知、决策和控制能力。它可以利用视觉传感器、听觉传感器、触觉传感器等多种传感器，感知周围的环境信息，并基于人工智能技术，对多模态数据进行处理，最终实现智能化的推理、决策和执行能力。例如，在工业生产中，人形机器人可以根据产品的具体形态、质量要求和生产工艺，自动调整动作和参数，完成精密的装配任务。另一方面，人形机器人还可以通过云端“大脑”和智能联网的方式，与云端服务器连接，获取更强大的计算资源和数据支持，从而扩展端侧算力，突破“一个躯体、一个大脑”的限制，实现智能的进一步提升。例如，在面对复杂的任务时，人形机器人可以将任务信息上传到云端服务器，由云端服务器进行分析和处理，并将处理结果反馈给机器人，使机器人能够更好地完成任务。

2. 形态类人

人的形态是生物进化过程中自然选择的结果。目前，社会中的所有城市基础设施、生产生活用的工具产品均是基于人类形态设计的。人形机器人通过对人类形态的高度模拟，能快速融入各种环境中执行复杂任务，具有更强的通用性和适应性。人形机器人的类人形态优势体现在3个方面。第一，通常采用与人相似的外观和身体结构，包括头部、躯干、四肢等。同时，人形机器人还可以配备各种表情和语音功能，能够与人类进行更加自然的交流，进而更容易被人们接受和信任。

第二，人形机器人的运动方式也模仿了人类的行走、跑步、跳跃等动作，能够快速进入人类的各种环境和场景，无须进行其他改造。第三，人形机器人可以使用各种工具，辅助或替代人类从事繁重的重复性劳动或危险性高的工作，提升生产效率、保护人民的生命健康。

3. 能力类人

在工业生产领域，人形机器人可以替代人类完成重复性劳动或危险性高的工作，并传统自动化设备协同工作，提高工厂的智能化水平和生产效率，如物料分拣运输、电子产品装配、品质巡检等。在民生服务领域，人形机器人可以作为家庭服务机器人，为人们提供家政服务、陪伴服务等。还可以在医疗、教育、娱乐等领域发挥重要作用，如辅助医生进行手术、为学生提供教学服务、为人们提供娱乐表演等。在特种作业领域，人形机器人可以替代人类在危险的环境中进行工作，如矿井开采、消防救援等。此外，人形机器人还具备比人类更强的耐受性和适应性，能够在非结构化环境中长期以低成本运行，有效解决了未来劳动力短缺的问题，有望在工业生产、民生服务、特种作业等领域实现广泛应用，甚至在部分领域发挥更大优势。

但是，目前业界也在争论人形机器人一直强调的“人形”是否有必要的问题。从根本上说，对“人形”的追求，本质是对机器人通用性的追求，核心点是当前是否有必要通过外形的拟人设计赋予机器人等同人类的“通用能力”，是“一机多用”还是“专机专用”。从现阶段的发展情况看，人形机器人主要在工业、服务业的简单场景中试用，相比后期应用场景中的复杂工作，当前场景任务相对重复和单一，人形需求并不强烈。随着“大脑”的逐步完善，通用人形机器人将会是终极形态，但在此之前，不同进化阶段及分支是必要的，“主体人形+定制化组件”的类人形可能是机器人商业化初步完成前的最优解。所以，从长远看，“人形”是机器人的理想形态；从现状看，完全的“人形”不是现阶段的唯一选择。

（二）人形机器人的发展历程

1. 萌芽探索期（20世纪60年代末至20世纪90年代）

人形机器人最早可以追溯到20世纪60年代末，一直到20世纪90年代都属于萌芽探索期。这个阶段设计的人形机器人以实现基本的双足行走功能为主要目标，初步具备了拟人化的结构，但整体上运动能力较弱，运动速度较慢、动作不够灵活。以日本早稻田大学为代表，其开发了WAP、WL、WABIAN和WABOT等一系列机器人。以WABOT-1为例，该机器人于1973年由日本早稻田大学研制成功，能够实现基本的

双足行走功能和简单的语言交流功能。它的出现标志着人形机器人的诞生，为后续的研究和发展奠定了基础。

2. 技术突破期（2000年至2022年）

2000年至2022年，人形机器人进入了技术突破期。这一时期，控制理论和技术的进步提升了机器人的运动能力，使其能够独立、稳定地执行复杂动作，且此阶段人形机器人已经具备了较强的运动能力。本阶段的典型代表是本田的ASIMO系列和波士顿动力的ATLAS系列机器人。ASIMO于2000年正式发布，身高约1.3m，体重约54kg，能够实现快速行走、跑步、上下楼梯等复杂动作。通过感知和智能控制技术的整合，ASIMO具备了初步的感知系统，能够感知周围环境的基本信息，并根据感知输入做出简单判断并调整动作。例如，它可以识别前方的障碍物，并自动调整行走路线，避免碰撞。ATLAS于2013年发布，身高约1.83m，具有28个液压驱动关节，能够实现碎石路面上的稳定行走。后续版本的ATLAS机器人不断进行技术升级，具备了更强的运动能力，能够在雪地、山地行走，还可以在倒地后迅速起身，并具备双臂协同搬运重物的能力。同时，还可以完成立定跳跃、跳高、转身跳、后空翻等一系列高难度动作，展示了人形机器人在运动控制方面的巨大潜力。

3. 智能发展期（2022年至今）

2022年以来，在人工智能技术的赋能下，机器人具有了更加智能化的感知、交互和决策能力，并开始逐步走向市场，实现产业化应用。这一阶段的人形机器人具有两大特点，一是人工智能技术的深度应用。通过大模型技术的赋能，人形机器人具备了更强大的感知、理解、推理、决策能力，同时，人工智能技术也正在逐步取代以传统控制理论和模型为基础的运动控制技术，强化学习和模仿学习成为运动控制的主流，实现更快的迭代开发。二是电驱动技术代替液压驱动技术。电驱动技术具有成熟可靠、寿命长、鲁棒性好、成本相对较低、易于控制、响应速度快、能量转化效率高等优点。新一代人形机器人均采用电驱动技术，实现了更高的操作精度，并提高了研发迭代速度。该阶段以美国特斯拉公司的Optimus（擎天柱）机器人、Figure AI公司发布的人形机器人Figure 01和Figure 02，我国的宇树、智元、优必选等企业发布的人形机器人为代表。

2022年10月，特斯拉正式发布Optimus第一代（Gen 1），其使用了与特斯拉电动车相同的完全自动驾驶（FSD）系统，具备强大的计算机视觉处理能力。Optimus Gen 1身高约1.73m，重量约73kg，采用全电驱动，具有行走、挥手和跳舞等功能。在结构方面，Optimus Gen 1主体具有28个自由度，包括14个旋转自由度和14个线性自由度。在此基础上，其灵巧手具有6个主动自由度和5个被动自由度。2023年12月，特斯拉发布了Optimus第二代（Gen 2），重量减轻到63kg，颈部增

加了2个自由度，步行速度提升了30%，平衡感和身体控制能力得到改善，能够完成非平坦地形行走，包括爬楼梯等复杂动作。精细操作方面，Optimus Gen 2的所有手指都配备了触觉传感器，能够轻松准确地抓取和放下鸡蛋，展示出精巧的双手操控能力。2024年5月，特斯拉Optimus机器人已经进入特斯拉工厂“实训”，实现对电池单元的准确分装。同年10月，该机器人已经能够演示做饭、调酒、跳舞等服务功能，并展现出更高水平的交互能力和更流畅的运动能力。除了特斯拉，其他企业也在积极推动人形机器人的产业化发展。例如，Figure AI公司发布的人形机器人Figure 01和Figure 02具备分类识别物品的能力和较强的负载能力，能够完全自主地在现实世界中执行任务。国内企业如宇树、智元、优必选、乐聚等也纷纷推出了自己的人形机器人产品，并在工业生产、物流配送等领域进行应用探索。

历经多年的发展和技术迭代，在巨大的潜在市场需求牵引以及人工智能技术深度赋能的带动下，人形机器人已进入技术集中突破和初步应用的关键发展阶段，技术水平持续提升、应用场景不断拓展、市场需求逐渐增加。随着技术的不断进步和成本的不断降低，人形机器人有望在更多领域得到广泛应用，成为推动社会发展的重要力量。

（三）人形机器人的分类

1. 按形态划分

目前，主流的人形机器人按形态可以分成两大类：轮式人形机器人和足式人形机器人。

轮式人形机器人主要采用轮式驱动，强调灵巧手的操作功能和人机交互功能。这种机器人通常具有较高的移动速度和灵活性，适合在平坦的地面上快速移动和执行任务。它可以利用轮式驱动系统快速到达目的地，然后通过灵巧手对物品进行抓取或操作。轮式人形机器人在物流分拣、仓储管理、商超零售、生产制造等应用领域具有一定的优势。

足式人形机器人具备完全类人的双足、双臂、双手结构，能够执行开放环境中的多种复杂任务。双足结构使其具有更好的稳定性和适应性，能够像人类一样在各种复杂的地形上行走，如楼梯、斜坡、不平坦的地面等，适应不同环境中的各种任务。在特种作业、救援等领域具有重要的应用价值，例如，在地震、火灾等灾害现场，足式人形机器人可以进入危险区域进行搜索和救援工作。同时，足式人形机器人作为人形机器人的完全形态，其未来的应用空间十分广泛。

2. 按具体应用场景和主要功能划分

人形机器人按照具体的应用场景和主要功能，可以分成以下几个类型。

特种作业型人形机器人用于执行灾害救援、危险作业等任务。在石油和天然气行业中，特种作业型人形机器人可以用于井下作业、油气管道巡检、危化品处理等。通过远程控制和智能化决策，它可以在高污染、高辐射等恶劣环境下进行工作，提高工作效率和安全性。在灾害救援场景中，它可以进入废墟、坍塌的建筑物等危险区域，搜索和救援被困人员。

工业型人形机器人用于工业生产及物流领域，如货物搬运、生产制造等。在汽车制造领域，工业型人形机器人可以与传统自动化设备协同工作，用于装配、质检等，提升工厂智能化水平，实现复杂工业场景的无人化生产。在3C电子产品制造领域，它可以进行精密装配与操作，完成电路板焊接、屏幕安装等任务。在物流领域，它可以进行物料搬运、库存盘点等工作，提高供应链效率。

医疗型人形机器人能够协助医生进行手术、诊断、康复等工作。在手术中，医疗型人形机器人可以通过高精度的机械臂和先进的传感器，辅助医生进行更加精准的操作，减少手术风险。在康复护理方面，它可以帮助患者进行康复训练，如肢体运动训练、平衡训练等，提高康复效果。

教育型人形机器人作为教学辅助工具，能够提供互动式学习内容。它可以通过生动有趣的方式向学生传授知识，激发学生的学习兴趣。例如，教育型人形机器人可以讲解故事、演示科学实验等，帮助学生更好地理解和掌握知识。同时，它还可以与学生进行互动交流，解答学生的问题，提供个性化的学习服务。

娱乐型人形机器人能够与人类互动，提供陪伴和娱乐。娱乐型人形机器人可以提供跳舞、唱歌、讲故事等娱乐活动。它还可以与人类进行情感交流，通过识别人类的表情和情绪做出相应的反应，增强人类的互动体验感。

公共服务型人形机器人能够在酒店、餐厅、商场等场所提供公共服务。它可以完成迎宾接待、导览讲解、需求解答等工作，提高服务效率和质量。在酒店中，公共服务型人形机器人可以为客人提供入住登记、行李搬运等服务；在商场中，它可以为顾客提供商品信息查询、购物引导等服务。

家庭服务型人形机器人，可承担养老陪护、育儿辅助、家务劳动等任务。家庭服务型人形机器人可以协助老年人完成日常生活照料，如起床、穿衣、洗漱等。它还可以照顾儿童，陪他们玩耍、学习。在家务劳动方面，家庭服务型人形机器人可以进行打扫卫生、洗碗、洗衣服等工作，减轻人们的家务负担。

通用型人形机器人，可用于工业、服务、教育、医疗等多个领域。通用型人形机器人具有较强的通用性和适应性，能够根据不同的任务需求进行功能配置和调

整。它可以在不同的场景中灵活切换工作模式，完成多种任务。例如，通用型人形机器人可以在工业生产线上进行装配工作，也可以在家庭中提供服务，还可以在学校中作为教学工具使用。

（四）人形机器人的技术架构

人形机器人的技术架构是一个复杂的系统，涉及多个学科领域的知识和技术，主要包括“大脑”“小脑”和“肢体”3个部分。

“大脑”是人形机器人的智能决策系统，负责实现环境感知、行为控制、人机交互等任务级能力。目前主要基于人工智能大模型技术，同时也可通过云边协同，提高机器人的智能水平。人工智能大模型为人形机器人提供了强大的语言理解、图像识别和决策能力。在环境感知方面，人形机器人可以利用视觉传感器、听觉传感器等收集周围环境的信息，并通过大模型进行分析和处理，识别出物体、场景和人物等。在行为控制方面，大模型可以根据感知到的信息和预设的任务目标，生成相应的控制指令，指挥机器人的动作和行为。在人机交互方面，大模型可以实现自然语言处理，使机器人能够理解人类的语言指令，并做出相应的回应。云边协同则可以将云端的强大计算能力和边缘设备的实时处理能力相结合，提高机器人的智能水平和响应速度。当机器人遇到复杂的任务时，可以将任务信息上传到云端进行处理，然后将处理结果反馈给机器人执行。

“小脑”是人形机器人的运动控制系统，负责保持平衡并实现稳定的运动和作业能力，目前主要基于人工智能、自动控制、机器人操作系统（ROS）等技术，实现复杂环境下的运动控制。传统的基于模型的控制方法通过建立人形机器人的运动学和动力学模型，完成人形机器人的运动轨迹规划和平衡控制。例如，零力矩点（ZMP）、线性倒立摆（LIP）等算法可以使机器人在行走过程中保持平衡。基于学习的控制方法则使用端到端的人工智能技术代替复杂的运动学模型，大幅度降低了“小脑”开发难度、提升了迭代速度。例如，强化学习和模仿学习等方法通过人类示教或自主学习建立运动执行策略。ROS则为机器人的开发和运行提供了一个通用的平台，它可以实现机器人各个模块之间的通信和协作，提高开发效率和系统的稳定性。

“肢体”即人形机器人的本体，是实现高动态、高爆发、高精度运动的物理载体，集成了人体运动力学、机械结构设计、新材料、传感器等诸多技术，包括仿人机械臂、灵巧手、腿足等关键结构，并通过集成传感器和长续航动力单元，实现“能源-结构-感知”一体化。其执行机构可以分为旋转执行机构和线性执行机构，旋转执行机构多用于关节处，如手腕、膝关节，主要由电机和减速器组成，核心零部件是无框力矩电机、行星减速器和谐波减速器等。线性执行机构多安装

于人形机器人上臂、大腿等部位，可理解为旋转执行器的线性转换，使人形机器人完成伸展、推拉等直线运动，主要由梯形丝杠、滚珠丝杠或行星滚柱丝杠组成。在芯片方面，人形机器人的芯片主要包括处理器芯片、控制芯片和总线管理芯片，其中处理器芯片[圆形处理单元（GPU）、中央处理器（CPU）、神经处理单元（NPU）等]是其功能实现的核心。在传感器方面，大量传感单元是实现复杂感知功能、与环境交互的基础，人形机器人所需的传感器类型包括力传感器、视觉传感器、触觉传感器、听觉传感器、惯性测量单元、接近觉传感器、距离传感器等多种类型。这些传感器可以帮助机器人感知周围的环境信息和自身的状态信息，为机器人的运动控制和决策提供依据。在电源方面，人形机器人需要高性能的电源来提供持久的动力。目前，锂离子电池是主流，但未来固态电池将更加适用于人形机器人产品。电源管理系统则用于监控电池状态，优化能源使用，确保人形机器人在各种工况下都能稳定运行。在新材料方面，人形机器人的新材料主要应用在骨骼、外壳等方面，如聚醚醚酮（PEEK）、碳纤维等，这些材料可以使机器人构件轻量化，提高机器人的机动灵活性和安全性。

总之，人形机器人的技术架构是一个高度集成和复杂的系统，各个部分相互协作、相互支持，共同实现人形机器人的各种功能。随着技术的不断进步和发展，人形机器人的技术架构也将不断被优化和完善，为人形机器人的广泛应用提供更加坚实的技术基础。

▷第二章 大小脑的主要技术

（一）大脑的技术路线

人形机器人与传统移动操作机器人相比，具备能够自主感知、认知理解、规划任务的智慧“大脑”。其鲁棒、安全、最优的智能决策系统，是人形机器人实现与环境和人类稳定交互、可靠执行任务、做出可行决策的关键。通过智能化技术，人形机器人可以具备自主决策与社交能力。决策系统包含局部世界环境感知与物体认知理解的相互映射匹配和对齐、人类与机器人指令交互的双向价值对齐、人类命令与机器人可执行指令对齐3个主体部分。具体来说，人形机器人在获得人类高级命令后，通过融合人形机器人系统环境感知数据和对物体的认知理解，构建局部环境空间映射和物体种类、用途、使用方式的关联关系。在此基础上，决策系统进一步生成人类分步指令集，完成上游命令的任务分解，在人机交互过程中以在线学习方式动态进行双向人机价值对齐，保障分步决策符合人类价值体系。最终，构建人类发布指令与人形机器人可执行指令的转化系统，完成分步决策结果到人形机器人可执行指令的对齐转换，形成与环境和人类稳定交互，且能自主决策的人形机器人决策系统。

目前，人形机器人“大脑”技术以大模型为核心，为人形机器人提供任务级交互、环境感知、任务规划和决策控制能力。在任务交互方面，基于大模型的语言/视觉运行处理方式可为人形机器人提供任务级交互入口。在环境感知方面，大模型通过对多模态信息的统一处理与灵活转换，推动多模态感知泛化。在任务规划方面，大模型潜在的真实世界知识学习能力，以及强大的思考、推理和生成能力为“大脑”的任务规划奠定基础。在决策控制方面，人形机器人基于大模型技术并优化奖励策略，通过整合环境、运动等多样化信息，实现决策控制功能。从功能需求角度出发，人形机器人的“大脑”大模型需要具备以下能力。

实时交互能力。人形机器人需要具备与人类实时的任务级交互能力，快速理解人类通过语言、手势等方式给出的指令，并有效执行。当出现指令理解不清的情况或指令执行完毕后，人形机器人应具备主动发起多轮交互的能力。

多模态感知能力。为了在复杂环境中做出正确决策，人形机器人需要能够通过视觉、听觉、触觉等多种感官获取信息。大模型需要整合这些多模态感知数据，以实现对环境的全局理解。

自主可靠决策能力。人形机器人在执行任务时，需要能够理解任务的复杂性，并将其分解为一系列可执行的子任务。这要求大模型具备强大的语言理解能力和对

物理世界的深刻理解能力。例如，机器人可能需要理解“清理房间”这一任务，包括识别哪些物品需要移动，哪些需要丢弃。

涌现和泛化能力。除了在训练数据上的表现，大模型还应具备超出训练范围的执行能力。具体表现为人形机器人能够在新环境中执行任务，适应未知的情况；面对新挑战时，展现出创新性的解决方案。

从技术路线上看，目前基于大模型的“大脑”技术路线正处在并行探索阶段，主要技术路线包括以下几种。

1. 大语言模型（LLM）

语言是人类用于表达与交流的重要能力，这一能力在儿童早期便开始发展，并在一生中不断演变。然而，机器并不能自然具备理解和使用人类语言的能力，除非配备了强大的AI算法。如何使机器能够像人类一样读、写和交流，一直是人类长期以来的研究挑战。从技术层面看，语言模型（LM）是推动机器语言实现智能的主要方法之一。其主要目的是刻画词序列的生成概率，从而预测未来（或缺失）词元的概率分布。其发展大致可以分为以下4个阶段。

统计语言模型（SLM）。SLM基于20世纪90年代兴起的统计学习方法开发。其基本思想是基于马尔可夫假设建立词预测模型。例如，根据最近的上下文预测下一个词。SLM已被研究人员广泛用于提升自然语言处理（NLP）任务的性能。然而，这类模型经常面临维度灾难：需要估计的转移概率数量呈指数增长，导致高阶语言模型难以被准确估计。为此，研究者提出了回退估计和Good-Turing估计等专门的平滑策略，以缓解数据稀疏问题。

神经语言模型（NLM）。NLM通过神经网络，如多层感知机（MLP）和循环神经网络（RNN）预测词序列的概率分布。该相关研究首次提出了词的分布式表示概念，并基于聚合的上下文特征建立词预测函数。通过扩展学习文本数据有效特征的思想，研究者提出一种通用的神经网络方法，构建了统一的端到端解决方案。进一步，Word2Vec被提出，它通过简化的浅层神经网络学习词的分布式表示，并在多种NLP任务中证明了其有效性。这些工作推动了语言模型在表示学习（超越单纯词序列建模）中的应用，对NLP领域产生了深远影响。

预训练语言模型。作为早期探索，ELMo被提出，它通过预训练长短期记忆LSTM[双向长短期记忆（biLSTM）]网络捕捉上下文感知的词表示，并通过微调biLSTM网络适配下游任务。随后，基于高并行性的Transformer架构及其自注意力机制，基于Transformer的双向编码器表示（BERT）通过在大规模无标签语料上进行特别设计的预训练任务，构建了双向语言模型。这些预训练的上下文感知词表示作为通用

的语义特征，在多种NLP任务中显著提升了性能。在此基础上，研究者陆续提出了不同架构（如GPT-2和BART）或改进预训练策略的预训练语言模型（PLM）方法。在这一过程中，通常需要对预训练模型进行微调，以适配不同的下游任务。

LLM。研究者发现，扩大PLM（例如扩展模型规模或数据规模）通常会提升其在下游任务上的能力。许多研究通过训练越来越大的PLM来探索性能的极限。研究人员发现，大型PLM表现出与小型PLM截然不同的行为，并在解决一系列复杂任务时展现出涌现能力。例如，GPT-3能够通过上下文学习解决少样本任务，而GPT-2在这方面表现较差。因此，研究界将这些大规模PLM统称为LLM，并对其研究投入了越来越多的关注。LLM的一项重要应用实例是ChatGPT，它将GPT系列的大型语言模型应用于对话任务，展现出与人类对话的惊人能力。

通常，LLM是指基于Transformer架构，包含数百亿或更多参数的模型。它们在海量文本数据上进行训练，例如GPT-3、PaLM、Galactica和LLaMA。LLM展现出强大的自然语言理解和复杂任务求解能力。该方法主要包括扩展定律、涌现能力和其他关键技术。

（1）扩展定律

目前，LLM主要采用Transformer架构，通过多头注意力层堆叠形成深度神经网络。尽管LLM在架构和预训练目标上与小型语言模型相似，但在模型规模、数据规模和总计算量上实现了数量级扩展。大量研究表明，扩展可以显著提升模型能力。因此，建立量化扩展效应的方法非常重要。该方法的建立主要用到以下两种扩展定律。

①KM扩展定律

Kaplan等人（OpenAI团队）首次提出用幂律关系来建模语言模型性能与3个主要因素（模型规模 N 、数据集规模 D 和训练计算量 C ）之间的关系。通过实证研究发现，这三者的增长均能降低模型损失，并提升模型性能。

②Chinchilla扩展定律

Hoffmann等人（Google DeepMind团队）通过在更大范围的模型规模（7000万~160亿参数）和数据集规模（50亿~5000亿tokens）上进行严格实验，拟合得到类似的扩展定律。他们提出，提升模型能力时，模型规模和数据规模应保持平衡，而不是一味放大模型。

综合来看，KM扩展定律更偏重增加模型规模，而Chinchilla扩展定律则强调模型规模与数据规模需同步扩展，从而在计算预算下获得最优性能。

（2）涌现能力

LLM的涌现能力定义为在模型规模达到一定阈值后突然出现，且小模型不具备该能力。当涌现能力出现时，模型性能会显著高于随机水平。接下来简要介绍LLM的3种典型涌现能力及相关代表模型。

①上下文学习（ICL）能力

当语言模型被提供自然语言指令或几个任务示例后，可以通过补全输入文本序列直接生成测试样例的期望输出，而无须进行额外训练或梯度更新。在GPT系列中，1750亿参数的GPT-3展现出较强的ICL能力，但GPT-1和GPT-2并不具备。这种能力也取决于具体任务，例如，13B参数的GPT-3在3位数加减法任务中能展现ICL能力，而175B参数的GPT-3在波斯语问答任务上表现不佳。

②指令跟随

通过多任务数据集上进行微调，LLM可以在从未见过的任务上按照指令执行。指令微调后，LLM在无须提供明确示例的情况下也能理解新任务指令，从而提升泛化能力。例如，当LaMDA-PT模型规模达到68B时，指令微调效果显著优于未微调模型，而在8B或更小规模下并不明显。

③逐步推理

小语言模型通常难以解决需要多步推理的复杂任务。相比之下，通过“链式思维（CoT）提示”策略，LLM可以在推理过程中生成中间步骤，从而得出最终答案。这种推测能力可能来源于在代码数据上的训练。特别是当模型规模超过100B时，其相对优势更明显。但是其涌现能力并不稳定，仅在特定阶段出现。当语言能力在某一阶段处于稳定的状态并进入下一个阶段时，其性能会发生质变，这表明大模型的学习能力并非呈线性增长的趋势。

（3）关键技术

经历了长时间的技术积累，LLM发展到今天成为通用强大的学习模型。在这一过程中，众多重要技术的使用显著提升了LLM能力，下文简要列举了几类关键技术。

①训练方法

超大模型的训练极具挑战性，需要分布式训练算法和多种并行策略组合应用。为支持分布式训练算法，DeepSpeed、Megatron-LM等优化框架被提出，方便并行算法的实现与部署。同时，优化技巧也对训练稳定性和性能提升至关重要，如重启

以克服损失突升、混合精度训练。GPT-4研发团队提出开发专用基础设施和优化方法，以便通过小模型可靠预测大模型性能。

②能力引出

经过大规模语料预训练，LLM具备通用任务求解潜力，但在执行某些任务时，这些能力并不总能自然显现。设计合适的任务指令或上下文学习策略（如链式思维提示）是引出这些能力的有效手段。此外，通过自然语言表达的任务描述进行指令微调，可以提升LLM在未知任务上的泛化能力。这些技术主要对应涌现能力，对小模型往往不具同等效果。

③对齐调优

由于LLM学习了包含高低质量数据在内的预训练语料特征，因此可能生成有害、有偏见或虚假内容。必须将LLM与人类价值观对齐。一些研究人员提出了一种基于人类反馈强化学习的有效调优方法，将人工标注整合到训练循环中。例如，避免模型回答侮辱性问题。

④工具调用

LLM本质上是基于海量文本训练的生成模型，在某些任务（如数值计算）上表现较弱，且能力受限于训练数据（无法捕捉最新信息）。为解决这些问题，研究人员提出利用外部工具弥补LLM局限。例如，LLM可调用计算器执行精确运算，或使用搜索引擎获取未知信息。

2. 视觉语言模型（VLM）

预训练LLM，如Llama、GPT-4，已在众多NLP任务中取得显著成功。然而，随着模型规模持续扩大，其面临两大挑战：一是高质量文本数据的有限供给；二是单模态架构在理解和处理现实世界信息时存在固有局限，尤其难以捕捉多模态间的复杂关联。这些限制推动了VLM的研究与发展。VLM通过融合视觉（如图像、视频）与文本输入，能够更全面地理解空间关系、物体、场景及抽象概念，突破了单模态方法的表征边界，提供了更丰富且具备上下文感知能力的认知方式，在视觉问答（VQA）和自动驾驶等领域展现了广阔的应用前景。

近年来，顶尖人工智能机构持续发布新型VLM。从OpenAI的CLIP、Salesforce的BLIP、DeepMind的Flamingo，到GPT-4V和Gemini，这些模型规模日益增大、交互性更强，并展现了聊天机器人功能与多模态用户交互的深度融合。VLM的架构正在从零预训练转变为以预训练LLM为主干来对齐视觉和文本信息，在此过程中，基本组件大体保持不变。VLM中最基础和广泛采用的架构组件如下。

（1）视觉编码器

视觉编码器在将视觉内容投射为嵌入特征方面发挥着关键作用，这些特征可以与LLM的嵌入对齐，用于文本生成或图像生成等任务。它被用来从图像或视频数据中提取丰富的视觉特征，从而实现与语言表征的集成。

具体来说，许多VLM中使用的视觉编码器，都是在大规模多模态或图像数据上进行预训练的。这类编码器通过对图文进行联合训练，使其能够有效捕获视觉与语言之间的关系。典型的例子包括CLIP模型通过对比学习对齐图像和文本嵌入，BLIP模型则利用自举式预训练实现稳健的语言-图像对齐。

另外，部分视觉编码器在ImageNet等大规模图像数据集上进行预训练。它们在大量标注视觉数据或自监督数据上进行训练，使其具备捕获特定视觉特征的能力。尽管最初是单模态，但这些编码器（如ResNet或Vision Transformer）可以适配多模态任务。它们擅长提取有意义的物体级特征，并为视觉-语言模型提供坚实的基础。许多最新的视觉-语言模型，如Qwen2-VL和LLaVA，通常都会集成预训练的视觉编码器。这些编码器不仅提供稳健且有意义的视觉表征，还在迁移学习中表现出色。相比随机初始化的编码器，它们能够利用预训练阶段积累的视觉知识，显著提升性能。

（2）文本编码器

文本编码器将分词后的文本序列投影到嵌入空间，处理方式类似于视觉编码器处理图像。CLIP、BLIP和ALIGN等模型同时使用图像编码器和文本编码器。这些模型通过对比学习在共享潜在空间中对齐图像和文本嵌入，有效捕捉跨模态关系。然而，LLaVA等新模型通常不包含专用文本编码器，而是依赖大型语言模型（如LLaMA、Vicuna）进行文本理解，并通过投影层或交叉注意力机制集成视觉输入。这一转变显示出利用LLM替代文本组件，以实现更通用、先进的多模态推理和生成任务的趋势。

（3）文本解码器

文本解码器以LLM作为主要文本生成器，通过视觉编码器投射图像特征。例如，GPT-4V、Flamingo、Kosmos-2等模型。这类模型通常只使用最小化的视觉投影机制，使强大的语言解码器生成丰富的上下文输出。VisualBERT等为多模态预训练的解码器架构奠定了基础。从零开始训练VLM时，通常需要独立的文本解码器，而将LLM作为主干网络时，则可以直接使用LLM自带的原始解码器。

（4）交叉注意力机制

交叉注意力机制是大型多模态模型（如视觉-语言模型）的核心组件，指的是利用一种模态（如文本）的查询向量去关注另一种模态（如图像）的表示，从而实现信息融合和对齐。与传统的自注意力（所有输入来自同一模态）不同，交叉注意力通过计算查询与另一模态的键值对之间的相关性，使文本输入能够“看到”视觉内容，反之亦然。这一机制在VisualBERT、ViLBERT、Flamingo、BLIP-2等模型中得到广泛应用，通常在编码阶段或解码阶段插入，通过跨模态学习增强模型对复杂任务的理解和生成能力。相比早期将视觉和文本简单拼接的融合方式，交叉注意力在对齐精度、灵活性和泛化性方面具有显著优势，是实现多模态推理、问答、描述生成等功能的重要技术支撑。

（5）特征融合

在最新的VLM中，特征融合正朝着一种更统一的表示方式发展，这种方法不再把视觉和文本视作完全独立的模态，而是将它们都转化为“统一的标记[词元

（Token）]序列”进行处理。具体来说，像Emu3这样的模型首先通过特定的视觉编码器（如SBERT-MoVQGAN）把图像或视频内容编码成离散的视觉Token，这些Token和文本Token具有相同的表示形式，可以一起送进大型语言模型（如LLaMA）的Transformer架构里，直接共享位置嵌入、注意力机制等结构。为了区分模态，这些视觉Token会用特殊的标记[例如（SOT）表示视觉开始，（EOV）表示视觉结束]加以界定。此外，Transfusion提出的方式则进一步尝试在一个Transformer架构中并行处理离散的文本Token和连续的视觉向量，通过策略性断点来兼顾两种表示。这些方法的共同点在于把所有输入都看作统一的多模态“符号”来预测下一个Token，从而实现紧密的特征融合，避免了以往多模态模型中视觉-文本分开处理，再通过交叉注意力连接的复杂流程。

3. 视觉语言动作（VLA）模型

近年来，视觉和语言基础模型在多模态理解、推理和生成方面有显著进步，VLM已经在图像理解、视觉问答、描述生成等任务上展现出强大的能力。然而，人工智能的终极目标并不仅仅在于“看”和“说”，而是能够“看、说、做”，即在复杂环境中基于视觉感知和自然语言指令自主决策并执行动作。由此，研究者逐步将研究重心从“看”和“说”拓展到“做”，即VLA模型。这类模型旨在感知视觉环境、理解自然语言指令，并直接输出可以驱动智能体或机器人的动作策略。VLA模型已成为智能体交互、机器人操作、虚拟环境任务等领域的重要研究方向。VLA模型中基础的架构组件和关键技术如下。

（1）视觉编码器

视觉编码器的有效性直接影响VLA模型的性能，因为它提供了关于当前状态的重要信息。例如，物体的类别、位置和可供性。与传统VLM主要关注的视觉-文本对齐不同，VLA模型中的视觉编码器在功能和表征目标上有显著扩展。VLM通常依赖如CLIP或BLIP的视觉编码器，通过对比学习获得图文一致性的高层语义特征，用于支持文本生成任务。然而，VLA模型的视觉编码器不仅需要提取语义信息，还必须捕捉与动作决策高度相关的几何和动力学属性，生成可用于控制策略的状态表征。为此，研究者提出了如R3M、MVP、VIP等专门面向机器人演示数据进行预训练的编码器，通过自监督或多模态对比学习强化感知与动作的关联性。此外，VLA模型视觉编码器通常与世界模型深度耦合，用于预测未来观测状态，支持基于模拟的规划和多步推理。这些改进显著增强了视觉编码对具身智能任务的适应性和通用性。

（2）动力学学习

动力学学习旨在使模型具备对环境状态转移规律的理解能力，通常包括正向动力学（预测在给定动作下环境的下一个状态）和逆向动力学（根据前后状态推断生成动作）。这是具身智能能够进行有效规划和推理的基础。不同方法通过多样的学习任务来捕捉状态-动作之间的映射关系。Vi-PRoM结合对比自监督学习、打乱视频帧顺序的时间动力学重排，以及伪标签图像分类等目标，使视觉编码器在学习时间关系和语义信息的同时，强化动力学表征。MIDAS则专注于逆向动力学预测，通过训练模型从观察序列中恢复动作，增强对状态转移的理解。SMART整合了正向和逆向动力学预测，以及随机掩码事后控制，既学习短期局部动力学，也兼顾长期全局依赖。部分方法采用掩码建模范式进行动力学学习，如MaskDP在状态和动作Token上进行随机掩码重建，从而统一学习正向和逆向动力学关系，与BERT或MAE的掩码机制相比，更直接适用于下游任务。PACT通过自回归方式建模状态-动作序列，预训练得到的动力学模型可迁移到定位、导航等多种场景。除了上述方法，VPT采用半监督模仿学习，先通过少量标注数据训练逆向动力学模型，再通过其对互联网视频自动生成标签进行行为克隆，从而使其在复杂任务上达到接近人类的水平。进一步地，GR-1以视频帧预测作为正向动力学学习目标，将未来状态的建模能力融入GPT类模型中，显著提升动作预测的准确性。

（3）世界模型和推理

世界模型用于编码关于世界的常识性知识，并在给定动作的情况下预测未来状态。它使具身智能体能够进行基于模型的控制和规划，因为智能体可以先在想象空间中搜索最优的动作序列，再执行任何真实动作。虽然正向动力学学习也尝试预测下一个状态，但其通常被视为一种预训练任务或辅助损失，用于增强基于RL-Transformer的动作解码器在主要机器人任务中的表现，而不是一个独立的模块。

此外，视觉世界模型可以直接生成未来状态的图像或视频，从而采样新的交互数据。世界模型通常包括3个组成部分。

a. 表征模型（Representation Model），将高维感知输入（如视觉图像）编码作为潜在状态表示，降低建模难度。

b. 转移模型（Transition Model），在潜在空间中建模时序转移关系，即状态如何随动作演化。

c. 奖励模型（Reward Model），估计状态（或状态—动作对）的期望回报，用于支持行为策略的优化。

与传统的直接通过环境交互更新的基于强化学习的策略不同，世界模型可以在内部生成多步“想象”轨迹。这些虚拟轨迹可用于更新策略（例如，通过Actor-Critic方法传播梯度），从而提高样本利用率，减少对真实交互的依赖，并提高探索的安全性和灵活性。典型方法如，Dreamer系列通过在潜在空间上进行连续预测和模拟，实现了基于模型的规划和策略改进。IRIS等方法采用自回归Transformer作为世界模型主干，将视觉编码结果与时序建模结合，用于生成模拟轨迹。

4. 多模态大模型（MLLM）

MLLM更强调通用的感知与推理能力（看、说、听、理解），其兴起主要源于两个根本驱动力：一方面，单一模态模型在真实世界任务中面临能力瓶颈，难以充分理解复杂、多维度的信息；另一方面，通用智能系统的构建需要模型具备对不同类型数据的统一感知、推理和生成能力。多模态大模型通过在共享表征空间中同时整合视觉、文本、语音乃至动作等多种输入，使模型能够跨模态捕获深层语义关联，完成更具开放性和通用性的任务。相比过去专注单一模态的模型，现代多模态体系不仅提升了对多样输入的适应性和泛化能力，也为更自然的人机交互和复杂情境下的智能决策提供了关键支撑。MLLM中最基础的架构组件和关键技术如下。

（1）模态编码器

编码器将原始信息（如图像、语音、视频）压缩为更紧凑的表示。相比从头训练，常用的方法是使用在预训练阶段已经与其他模态对齐过的编码器。例如，CLIP通过在大规模图文对上进行预训练，使视觉编码器在语义上与文本对齐。因此，使用这种已初步对齐的编码器再与LLM进行对齐预训练会更容易。MiniGPT-4采用了经过改进训练的EVA-CLIP（ViT-G/14）编码器。相比之下，Osprey引入了基于卷积的ConvNext-L编码器，旨在利用更高分辨率和多层次特征。

同时，无编码器架构也在探索之中。例如，Fuyu-8b直接将图像Patch投影后送入LLM，因此模型天然支持灵活的图像分辨率输入。在选择编码器时，人们通常考虑分辨率、参数规模和预训练语料等因素。值得注意的是，实验证明更高分辨率可以使性能得到显著提升。输入分辨率扩展方法一般可以分为直接扩展和分块方法两类，直接扩展即将高分辨率图像直接输入编码器，这通常需要进一步微调编码器，或将编码器替换为高分辨率预训练编码器。

类似地，CogAgent使用双编码器机制，两个编码器分别处理高分辨率和低分辨率图像，高分辨率特征通过交叉注意力注入低分辨率分支。分块方法则将高分辨率图像切成Patch，再复用低分辨率编码器。例如，Monkey和SPHINX将大图像分割为小Patch，与下采样的高分辨率图像一起输入图像编码器，其中小图像和低分辨率图像分别捕获局部和全局特征。

(2) LLM

相较于从头训练，基于预训练LLM进行开发更具效率优势。通过海量网络语料的预训练，LLM已内化丰富的世界知识，展现出强大的泛化与推理能力。多数常用开源LLM属于GPT-3开创的因果解码器架构。早期代表如FlanT5系列，被BLIP-2和InstructBLIP等采用。LLaMA系列与Vicuna家族作为开源标杆受到学界广泛关注，但二者主要基于英语语料预训练，对中文等多语言支持有限。相较之下，Qwen在双语（中英）支持上的表现更为均衡。

值得注意的是，扩大LLM参数量可带来额外收益，这与提升输入分辨率的效应类似。因此，近期专家混合（MoE）架构的LLM探索日益受到人们的关注。相比稠密模型，这种稀疏架构通过参数选择性激活机制，能在不增加计算成本的前提下扩展总参数量。

(3) 模态接口

由于LLM仅能处理文本信息，建立自然语言与其他模态之间的桥梁至关重要。然而端到端训练大型多模态模型成本高昂，更实用的方案是在预训练视觉编码器与LLM之间引入可学习的连接器，或借助专家模型将图像转换为语言描述后再输入LLM。

这类方式使用可学习连接器模块来弥合模态差异，将信息投影至LLM可理解的特征空间。根据多模态信息融合方式的不同，这些实现方法可归为两类。

① 令牌级融合

将编码器输出的特征转换为视觉令牌，与文本令牌拼接后输入LLM。主流方案采用可学习查询令牌（如BLIP-2首创的Q-Former架构），通过查询机制压缩视觉令牌

为少量表征向量。另一种简化方案即使用基于MLP的接口（如LLaVA系列采用单/双线性层），对齐视觉令牌与词嵌入维度。

②特征级融合

通过插入额外模块实现文本与视觉特征的深度交互。例如，Flamingo在LLM的冻结Transformer层间插入跨注意力层；CogVLM在每层嵌入视觉专家模块实现双模态交互（其QKV矩阵继承自预训练LLM）。

专家模型方案无须训练，能够直接通过预训练模型（如图像描述模型）将多模态输入转换为语言：VideoChat-Text结合视觉模型与语音识别生成富文本描述。其优势在于实现简单，但存在信息损失缺陷，如视频转文本会扭曲时空关系，并且其灵活性不及可学习接口。

（二）小脑的技术路线

随着机器人技术的发展，人形机器人产品不断增多，移动导航相关技术发展日趋成熟，机械臂的规划控制鲁棒性、自抗扰性、柔性控制技术日益完善。然而，人形机器人系统作为具身智能的实际载体，其运动与操作联合规划技术尚未完备，致使具身智能的部署与应用受到较大掣肘。人形机器人需要具备对上游高级命令进行自动感知、自主决策、多机协作、人机交互的决策能力，而这对导航运动与操作的协同规划提出了更高的要求，单一的导航运动与操作控制无法满足其灵巧性、高效性、连贯性、稳定性、安全性的需求。

人形机器人的运动与操作联合规划技术需要具备移动基座与机械臂的协同规划控制能力，以及多机协作、人机协作的工作能力。在此基础上，人形机器人需要在复杂空间约束环境下完成典型任务，包括在室内外复杂约束地形与环境下的移动、抓取、搜索、运输、交互，融合环境感知数据进行局部建图与定位、多关节路径规划、协同交互式操作、工具理解与使用等。且能够与其他机器人和人类并行统一协作或串行交互传递，完成难以独立完成的、更为复杂的高级任务。我们将当前的机器人控制手段分为基于模型的控制方法和基于学习的控制方法两种类型。

1. 基于模型的控制方法

（1）PID控制器

比例-积分-微分控制器（PID控制器）是一种经典的反馈调节算法，广泛应用于工业自动化和机器人控制系统中。它的基本思想是根据系统的当前误差、历史累积

的误差，以及误差的变化趋势来计算控制量，从而对系统施加调节，使输出逐步逼近设定目标。PID控制器的核心优势在于其结构简单、调节灵活、适应性强，能够在一定程度上补偿系统的不确定性和扰动。

PID控制器由3个部分组成：比例（P）、积分（I）、微分（D）。比例部分用于直接对当前误差进行放大调节。它根据设定值（期望输出）与实际输出值之间的偏差（误差），通过比例增益产生一个与误差大小成正比的控制作用，从而立即修正偏差。比例控制能够迅速响应偏差，但若只使用比例控制，系统通常会存在稳态误差（即输出长期偏离目标值）。

积分部分的作用是对误差进行累加，随着时间推移逐步积累偏差。积分增益使得历史误差得到累积补偿，当系统长期存在小偏差时，积分项会不断增强，迫使输出逼近目标，从而消除稳态误差。这是PID控制器与纯比例控制器相比最重要的优势之一。然而，积分作用的累积特性也会带来一定弊端。例如，积分饱和或积分振荡，尤其在系统大幅度波动时容易导致过度补偿，影响稳定性。

微分部分用于预测误差变化趋势。它通过计算误差随时间的变化率（导数），在误差开始增大或减小时提前施加调节作用，抑制超调并改善系统的动态响应性能。微分控制可以看作对未来误差变化的一种预判，能有效减小振荡，提高系统的稳定性。但微分环节对噪声十分敏感，若输入信号中含有高频干扰，微分作用会放大这些噪声。因此，工程实践中往往会对微分项引入低通滤波，抑制噪声干扰。

PID控制器的输出信号一般由3部分构成：

$$u(t) = K_p e(t) + K_i \int_0^t e(t) dt + K_d \frac{de(t)}{dt}$$

其中， $e(t)$ 为设定值与实际输出的误差， K_p 、 K_i 、 K_d 分别为比例、积分和微分增益。这3个参数的大小决定了系统的响应特性。通常，较大的比例增益可以加快响应速度，但会加剧超调；积分增益能够消除稳态误差，但过大时容易引起振荡；微分增益可抑制超调，改善动态性能，但对噪声比较敏感。因此，PID控制器参数整定是使用过程中的关键步骤，常用方法包括经验公式（如Ziegler-Nichols整定法）、试凑法，以及现代的自整定算法。

PID控制器凭借其结构简单、易于实现、适应性强等优点，已成为工业过程控制最普遍的解决方案。

(2)MPC

模型预测控制（MPC）是一种基于模型的先进控制策略，其核心思想是利用被控系统的动态数学模型，预测未来一段时间内系统的行为，并通过求解有限时域的最优化问题，计算出当前最优控制输入。MPC的每个控制周期都根据最新测量状态重复预测与优化，因此又被称为“滚动优化”或“滚动时域控制”。

在模型预测控制过程中，首先需要构建系统的预测模型，这可以是线性状态空间模型、非线性模型，甚至数据驱动模型。根据该模型，控制器从当前时刻开始，预测未来若干步（预测时域）内系统状态随控制输入的演变。为了使系统沿期望轨迹运行，MPC定义一个性能指标（通常是加权的轨迹跟踪误差和控制增量），在每个采样周期通过优化算法求解一组未来的控制序列，使得预测性能指标最小。

数学上，MPC问题一般写作如下形式：

$$\min_{u_0, \dots, u_{N-1}} \sum_{k=0}^{N-1} [\|x_k - x_{\text{ref}}\|_Q^2 + \|u_k\|_R^2] + \|x_N - x_{\text{ref}}\|_P^2$$

其中， x_k 为预测的状态， u_k 为控制输入， x_{ref} 为参考轨迹， Q 、 R 、 P 为权重矩阵， N 为预测步长。约束条件通常包括：

$$x_{k+1} = f(x_k, u_k), \quad x_k \in X, \quad u_k \in U$$

优化完成后，MPC只需将第一步控制量 u_0 施加到系统，随后采样下一个时刻的状态，再重新进行预测和优化。这种“滚动优化”机制不仅可以实时考虑模型不确定性、外部扰动和新的约束，还具备良好的鲁棒性。

MPC相较于传统PID控制具有显著优势：它能够直接处理多变量耦合问题、显式约束和多目标优化，因此广泛应用于过程工业、无人驾驶、航天航空及高精度机器人控制领域。与此同时，MPC也面临较高的实时计算负担，尤其在大规模系统和非线性模型下，求解优化问题的效率成为设计的关键。为此，近年来发展了诸如快速MPC、基于学习的近似MPC、分布式MPC等方法，以提升其实时性和适应性。

总的来说，MPC是一种以模型为核心、通过滚动优化实现预测与控制闭环的现代控制策略，能够在复杂系统中实现对动态性能和约束的精确管理，是先进控制领域的重要研究方向。

2. 基于学习的控制方法

(1) 强化学习

强化学习是一种通过与环境交互来学习决策策略的机器学习方法，其核心思想源于心理学中的试错学习机制。与监督学习依赖标签数据不同，强化学习通过奖励信号引导智能体（Agent）在环境中探索，并逐步学会最大化累积回报。

在强化学习中，智能体会在每个离散时刻观察环境的状态 s_t ，选择一个动作 a_t ，并根据动作获得奖励 r_t 和新的状态 s_{t+1} 。这一过程通常用马尔可夫决策过程（MDP）建模，定义为一个四元组 (S, A, P, R) ，其中：

状态空间（ S ）：表示环境所有可能状态的集合。

动作空间（ A ）：表示智能体在环境中可以执行的所有可能动作的集合。

状态转移概率（ P ）：描述在状态 s 执行动作 a 后转移到状态 s 的概率。

奖励函数（ R ）：定义了智能体在状态 s 执行动作 a 后获得的奖励。

智能体的目标是学习一个策略 $\pi(a|s)$ ，即在给定状态下选择动作的概率分布，以最大化期望累积回报：

$$J(\pi) = E \left[\sum_{t=0}^{\infty} \gamma^t r_t \right]$$

$\gamma \in [0, 1)$ 为折扣因子，用于平衡短期和长期收益。

强化学习的方法大体可以分为3类。

基于值函数的方法。例如在Q-learning中，智能体通过估计状态-动作值函数 $Q(s, a)$ 表示在状态 s 下采取动作 a 后期望获得的累积回报。其学习目标近似满足贝尔曼方程：

$$Q(s, a) = E \left[r + \gamma \max_{a'} Q(s', a') \right]$$

基于策略的方法。智能体直接对策略进行参数化和优化，如策略梯度方法（Policy Gradient）。通过梯度上升更新策略参数 θ 。

$$\nabla_{\theta} J(\pi_{\theta}) = E[\nabla_{\theta} \log \pi_{\theta}(a|s) \cdot Q_{\pi}(s, a)]$$

基于Actor-Critic的方法。智能体将值函数与策略结合，通过一个Actor网络输出策略，一个Critic网络估计值函数，二者协同更新。代表算法包括A3C、DDPG、SAC等。

强化学习最显著的特征是探索（Exploration）与利用（Exploitation）的权衡：智能体需要探索未知动作以获取更多经验，同时也要利用已有知识获取高回报。常用的探索策略包括 ϵ -greedy和随机噪声注入。

近年来，深度强化学习（Deep RL）通过将深度神经网络引入价值函数或策略函数，极大提升了在高维状态空间下的能力，使其成功应用于机器人控制等复杂场景。然而，强化学习也面临样本效率低、稳定性差等挑战，活跃的研究方向包括奖励工程、模型辅助学习、离线强化学习等。

（2）模仿学习

模仿学习是一种通过学习专家演示来获得策略的技术。其核心目标是根据专家生成的状态-动作对数据，拟合一个策略函数，使智能体能够在相同情境下采取类似行为。

早期的方法是行为克隆（BC），将其视为监督学习问题，通过最小化预测动作与专家动作之间的差异，得到参数化策略 $\pi_{\theta}(a|s)$ 。常用的损失函数为：

$$L_{BC}(\theta) = \frac{1}{N} \sum_{i=1}^N \|a_i - \pi_{\theta}(s_i)\|^2$$

通过优化损失函数即可获得策略。行为克隆的优点是训练简单，收敛快，但容易产生分布偏移：一旦智能体偏离专家示范过的状态分布，错误便会逐步累积。

为克服这一问题，逆向强化学习（IRL）提出先学习隐藏在演示背后的奖励函数 $r(s_t, a_t)$ ，再通过强化学习求解最优策略：

$$\pi^* = \arg \max_{\pi} E \left[\sum_t \gamma^t r(s_t, a_t) \right]$$

IRL的代表性方法有最大熵IRL和对抗模仿学习（GAIL）。例如，GAIL通过判别器区分专家和智能体生成的轨迹，策略则学习生成能“骗过”判别器的轨迹，从而达到模仿的目的。GAIL的优化目标为：

$$\min_{\pi} \max_D E_{\pi} [\log D(s, a)] + E_{\pi \epsilon} [\log(1 - D(s, a))]$$

除了纯粹的离线模仿学习，交互式模仿学习（如DAgger）也被提出，DAgger通过迭代收集智能体在环境中执行动作后的状态，并让专家提供正确动作，以此作为新标签，不断扩展训练数据，从而减轻分布偏移问题。

综上所述，模仿学习通过直接回归专家动作或间接推断奖励信号，赋予智能体在稀疏奖励或设计奖励函数的任务中快速获取学习能力，已在机器人操作、自动驾驶、游戏等领域得到广泛应用。

（3）融合学习

融合学习结合了模仿学习与强化学习，旨在同时发挥两者的互补优势。具体而言，模仿学习通过利用专家演示数据，以监督学习的方式提供策略的初始估计，使智能体能够在环境中快速获得可用的行为模式，显著降低从零开始进行随机探索的高昂样本复杂性。然而，纯粹依赖模仿学习的策略往往存在泛化性差和对示范数据分布外状态敏感的问题，难以在实际应用中实现鲁棒性。相比之下，强化学习通过试错交互优化累积奖励，具备强大的自我改进能力和理论上的最优性保证，但在没有有效先验的情况下往往面临收敛速度慢、探索效率低的问题，尤其在高维连续动作空间中尤为明显。

（三）大小脑的分层实现路径和端到端实现路径

目前，大模型主导的高层决策与强化学习驱动的低层运动控制已逐步形成一种“高层语义规划+低层动作执行”的分工模式。大模型（如LLM/VLM）具备强大的多模态理解和任务泛化能力，能够通过自然语言或视觉输入生成高层意图和粗略子目标，而低层控制策略（基于强化学习、行为克隆或MPC）则负责将这些高层输出转化为精确、实时的轨迹和动作，实现闭环控制。代表性工作如RT-2、Gato、RoboCat、VIMA等，已经在多任务、多机械臂平台上证明了该范式的有效性，特别是在少样本学习、指令条件化和多模态融合方面展现了强大能力。

但仍存在一个关键的问题，当前的方法主要侧重指令的生成，而忽略了将灵巧操作与具身人工智能的任务级规划相结合。面临的问题主要包括高低层联动的实时性不足、语义与动作表征对齐不充分、开放环境下泛化性能有限、长期任务的时序一致性薄弱，以及安全性可解释性不足等。因此，尽管VLA等方法在封闭和受控环境中已取得实用性进展，要实现真正具备通用性、可靠性和高适应性的机器人智能体，还有相当大的探索空间。

（四）大小脑能力现状

当前，人形机器人“大脑”刚刚具备初阶人类脑力，仅能完成人的部分工作，无法形成人类大脑全能力闭环。同时，其情感表现属于模拟层面，不具备情感理解能力。“小脑”仅能完成走、跑、跳等基本动作和简单的作业，无法顺利完成复杂的长序列任务。人形机器人作为人工智能的前沿应用领域，其发展速度令人瞩目。然而，要实现真正的智能化和自主化，现有的大模型仍需在多个方面重点发力。

在“大脑”智能决策方面，人形机器人主要面临的挑战是自主决策能力薄弱。人形机器人的智慧“大脑”需要实现环境感知与世界认知理解结果的融合匹配，依照人类价值体系生成合理、可行的决策方案，构建人形机器人系统可执行的指令集，完成人类执行方案与机器人系统指令集的关联映射，形成人形机器人解决方案。因此，人形机器人系统在决策时需要具备类人的价值标准、环境物体与认知物体的理解对齐能力、将人类高级指令分解为机器人可执行指令序列的能力等，这些对人形机器人技术与具身智能发展提出了挑战。当前，具身智能决策方法主要围绕大语言模型进行认知理解与推断，加入先验信息引导或加入后验信息反馈矫正。由于大语言模型存在幻觉和伪逻辑等问题，可能会生成一系列机器人难以执行的指令。同时在引入先验信息时，由于人形机器人的工作场景及任务通常需要即时决策，对细粒度图像与点云进行分割和特征提取的现有方法虽然效果与泛化能力很强，但计算成本较高、耗时较长（如SAMB5），不利于即时导航与决策。近期，LLaMA3、LLaVA等通过预训练模型、有监督微调、近端优化等方式，使其性能得到进一步提高，但依然存在指令通用性太强，无法具体执行的问题。

在“小脑”运动控制方面，基于模型的控制方法与基于学习的控制方法在实际应用中均存在一定的问题和挑战。对于基于模型的控制方法而言，实际应用中的机器人往往面临复杂且非线性的工作环境，如柔性物体操控、接触摩擦、齿轮间隙等，这些因素使得精确建模变得困难。同时，在模型建立过程中需要考虑质量分布、关节摩擦、惯性参数等高维因素，导致建模成本高、调试周期长。此外，模型控制在实时性方面也存在限制，高阶非线性模型或耦合动力学的计算复杂度高导致消耗时间长，影响机器人时效性。相比之下，基于学习的方法能够在一定程度上摆脱精确模型的依赖，具有较好的适应性和泛化潜力，特别适用非结构化环境和复杂交互任务。然而，这类方法的主要问题在于数据需求量大，尤其是深度学习驱动的策略优化往往需要大量在线交互或高质量离线数据，成本高昂。同时，其学习过程通常缺乏物理先验，容易出现过拟合或策略崩溃，鲁棒性和安全性不足。此外，学习型控制器缺乏可解释性和严格的稳定性保证，这在安全关键应用（如医疗、航天领域）中成为部署障碍。

（五）训练开发工具现状₁

Megatron-LM

Megatron-LM是由英伟达（NVIDIA）推出的一款专为超大规模Transformer模型训练而设计的分布式训练框架，其核心在于支持3D并行策略：张量（模型）并行、流水线并行和数据并行。张量并行将每层的权重按维度切分到多张GPU，流水线并行将整个网络按层分段，多个微批次在不同GPU上流动执行，这两种方式显著降低了单卡显存占用。通过结合数据并行，Megatron-LM可以在多节点环境下训练千亿至万亿参数规模的大模型，例如GPT-3和T5-XXL。在性能上，它利用高效的通信原语（NCCL AllReduce/AllGather）和统一计算设备架构（CUDA）优化内核，实现了接近线性的加速比。适合有大型GPU集群的科研单位、云服务厂商和企业进行超大模型研究和预训练。

2. DeepSpeed

DeepSpeed由微软开源，聚焦降低大模型训练的显存和算力成本。它最显著的创新是ZeRO优化器（Zero Redundancy Optimizer），通过将模型状态（如梯度、参数、优化器状态）在多GPU间分布式切片，有效避免了传统数据并行的内存冗余。DeepSpeed还支持梯度累积、混合精度训练（FP16/FP32）、稀疏注意力（Sparse Attention）等功能，大幅提升训练吞吐量并节约显存。其模块化API可以和PyTorch无缝集成，并兼容Megatron-LM形成混合并行。在中小规模硬件资源有限的团队中，DeepSpeed可以将同一硬件跑出更大Batch、支持更深模型，是“以更少资源实现更大模型”的典型工具。

3. Hugging Face Transformers

Hugging Face Transformers是全球最活跃的开源NLP模型库之一，封装了BERT、GPT、RoBERTa、T5、BART等多种Transformer架构的预训练权重、分词器和训练脚本。它不仅提供单机微调功能，还可借助Trainer API、Accelerate、DeepSpeed/FSDP等工具实现分布式训练和混合精度训练的扩展。Transformers对研究人员和工程团队极其友好：只需少量代码就能在下游任务（文本分类、序列标注、问答）上完成快速原型开发。其“模型即服务”接口也支持推理API和模型共享，非常适合中小规模任务的微调和生产部署。

4. Horovod

Horovod是Uber开源的跨框架分布式训练库，最初由TensorFlow社区衍生，后来扩展到PyTorch、Keras等深度学习框架。它基于Ring-AllReduce通信策略，在多机

多卡环境中高效同步参数梯度，显著降低通信延迟。Horovod的API简单易用，只需在模型代码中做极少改动即可进行分布式训练。它还支持分布式混合精度训练和容错恢复，是许多工业生产环境的常用分布式训练引擎，尤其适用跨多机又要框架兼容的同步训练场景。

5. Colossal-AI

Colossal-AI由华中科技大学开源，主打灵活的异构并行训练和自动优化，集成了张量并行、流水线并行、数据并行的2D/2.5D/3D并行组合，能根据模型规模自动配置最优并行策略。相比DeepSpeed和Megatron-LM，Colossal-AI在多并行范式融合和显存优化方面更灵活，支持张量切分、梯度累积、ZeRO、Flash Attention等前沿技术。其微调脚手架兼容PyTorch生态，可一键部署分布式训练，适合科研团队在复杂实验场景下快速探索大模型训练性能最优配置。

（六）大小脑的发展趋势

一是在感知模态维度方面形成突破。当前的人形机器人大模型主要依赖视觉或语音感知，这种单一的感知模态在处理复杂环境时显得力不从心。例如，在嘈杂的环境中，仅凭视觉信息，人形机器人难以准确识别和响应。为了突破这一局限，未来的大模型需要整合视觉、听觉、触觉等多种感知模态。多模态感知能够提供更丰富的环境信息，使机器人在复杂场景中做出更准确的决策。例如，结合听觉和触觉信息，机器人可以更好地理解人类的指令和情感状态。

二是在指令生成速度与复杂性方面形成突破。现有的大模型在生成指令时速度较慢，且生成的结果往往过于简单。这在需要快速反应的场景，如紧急救援或复杂操作任务中，可能导致人形机器人无法及时作出正确响应。目前，主流人形机器人大模型偏向于任务理解和拆分，对于人形机器人运动控制的涉及较少，只是用预设的端到端的训练方式生成了简单且离散分布的机械臂末端位置和底盘移动指令，未渗透到连续路径和轨迹规划等更偏人形机器人领域的内容。

三是在泛化能力提升与模型架构优化方面形成突破。泛化能力是大模型在新环境和新任务中表现突出的关键。当前的大模型在泛化能力上仍有待提高，尤其是在面对未知环境和任务时，大模型的表现往往不尽如人意。为了提高泛化能力，未来的大模型需要在架构、训练方法和数据集方面进行创新。例如，通过引入元学习、迁移学习等技术，可以使模型更好地适应新任务。同时，构建更多样化的数据集，也有助于模型学习到更广泛的知识。

整体而言，未来人形机器人智能决策与运动控制的进化路线，将以多模态模型驱动的通用认知与全身动力学感知-规划-执行闭环为核心特征，逐步形成“能看、

能想、能说、能灵活执行”的自适应通用型平台，为完成跨行业场景下的高复杂度任务奠定技术基础。如何在规模化预训练基础上实现端到端对齐、支持长期任务规划、兼顾推理速度和控制精度，以及引入自我改进和增量学习机制，仍是未来的研究重点。

▷第三章 肢体的主要技术

（一）执行机构的主要技术路径和发展趋势

目前，“肢体”执行机构的核心驱动技术路线已由传统的液压驱动全面转为电驱动。液压驱动技术依赖液体压缩泵产生高压液体，进而驱动输出机构。其主要优点是力量输出大、易于扩展，但也存在控制技术复杂、能量效率相对较低、系统零件多、成本高、故障率高、维护维修烦琐、响应速度不够快等诸多问题。液压驱动的代表产品主要是波士顿动力在2024年前推出的一系列人形机器人产品。电驱动技术以各类电机作为动力输出机构，其优点是成熟可靠、寿命长、鲁棒性好、成本较低、易于控制、响应速度快、能量转化效率高等，主要缺点是其本身扭矩密度较低，通常需要搭配减速器使用。目前，以特斯拉Optimus为代表的新一代人形机器人均采用电驱动技术。2024年4月，波士顿动力正式放弃传统的液压驱动技术并发布了该公司首款电驱动人形机器人，标志着驱动技术全面向电驱动技术收敛。人形机器人的执行机构主要包括机械臂、灵巧手、腿足等，并由电机、减速器、丝杠等主要零部件组成。

1. 机械臂

机械臂的硬件发展历程体现了从早期高刚性、大功率装置到轻量化、模块化、智能感知平台的演进。最初的工业机械臂，采用厚重的钢制或铸铁臂体，关节普遍由液压缸或大功率直流电机驱动。液压驱动系统通过阀控流量连续输出力矩，具备高功率密度和良好的刚度，适用于重载搬运与焊接等任务。然而，液压驱动系统存在漏油、噪声大和维护复杂等问题。为克服这些局限，纯电驱动被逐步推广，尤其是伺服电机结合精密减速器，成为电气化机械臂的核心方案。伺服电机可通过闭环控制实现位置和高度的精度调节，并配合行星减速器或谐波减速器，将高速小扭矩转化为低速大扭矩输出，兼顾速度和负载能力。此阶段，机械臂的代表结构为刚性串联关节，每个关节包含电机、减速器、编码器和制动器，典型产品如ABB IRB6、FANUC ARC Mate等。

随着微电子技术和高精度传感器的进步，20世纪80年代后，谐波减速器（Harmonic Drive）得到广泛应用，具备零背隙、高减速比、体积紧凑等优势。谐波减速器依赖柔性齿轮形变传递运动，通过波发生器产生椭圆变形，实现输入输出的相对运动，大幅提升减速和定位精度。它成为工业机械臂中高精度伺服关节的标准配置。与此同时，多圈绝对值编码器开始取代增量编码器，提供即时、

无漂移的角度测量，便于多关节串联的精确同步。该时期，机械臂臂体逐步采用铝合金型材与高强度铸铝壳体，减轻结构惯性，提高响应速度。

进入21世纪，机械臂的设计理念向高集成化与模块化演进。典型关节模块内部集成了伺服驱动器、位置与温度传感器、通信接口等子系统，具备即插即用能力，大幅降低系统集成与维护复杂度。如KUKA KR系列、安川Motoman GP系列均采用高集成关节单元，支持以太网实时通信和分布式控制。在材料层面，碳纤维复合材料被用于臂段和中空轴，提升了结构刚度与轻量化水平。伴随人机协作需求，安全性成为设计重点，柔性驱动和力控技术逐步普及。

协作机械臂的兴起是机械臂硬件演进的标志性事件。UR系列、Franka Emika Panda等是采用低减速比、高扭矩伺服电机配合串联弹性执行器（SEA）的典型产品。SEA将弹性元件串联在驱动链中，通过弹性形变与应变计测量输出力矩，形成高带宽的力控回路。相比纯刚性驱动，这种结构具备天然的柔顺性，能够检测外力变化并实时调整输出力矩，有效保障了人机协作安全。部分高端协作臂在末端集成六维力/扭矩传感器，通过闭环力控进一步实现精细装配、插拔等操作。此外，基于电流环直接估算力矩的无弹性力控技术也得到应用，简化了硬件结构。

近10年，直接驱动（Direct Drive）技术成为研究热点，尤其在对动态性能与零背隙要求极高的场景中。直接驱动关节采用超高扭矩低速电机，取消谐波减速器，从根本上避免齿隙和弹性累积误差，如Kawasaki duAro双臂机器人的部分关节采用直接驱动电机（DDM），具备极快响应与极高轨迹精度。然而，DDM对电机与驱动器的性能要求极高，且成本昂贵，目前主要用于高端装配和实验室机器人。

另外，可重构与模块化臂段也处于探索阶段，具备自由度和臂长灵活组合能力，通过模块级封装关节与通信电源接口，实现一体化热管理与状态感知，代表性产品如Modbot、Mecademic机械臂。三维视觉与触觉传感器正逐步集成在机械臂本体中，使硬件具备环境感知能力，为未来智能操作奠定基础。

综上所述，现代机械臂硬件正由传统刚性串联机构，向高集成、高柔顺、智能感知的方向迈进，形成以高性能伺服电机、谐波减速器或直接驱动电机、串联弹性执行器、六维力传感器和模块化臂段为核心的先进技术体系。未来，随着材料工程、驱动电子和传感器的进一步发展，机械臂将实现更高功率密度、更轻臂体、更丰富感知与更自然的人机协作能力。

2. 灵巧手

五指多关节机械手，被认为是机器人领域最接近人类手功能的关键末端执行器。其硬件发展历程可以从3个维度来理解：机械结构设计、驱动与执行机构、末端感知技术。

在机械结构方面，最初的灵巧手以“刚性骨骼+多自由度关节”为核心，其代表性成果有Utah/MIT Hand。它由16个自由度构成，采用金属关节和腱索传动，力求在尺寸和运动范围上模仿人类手指。随后出现的DLR Hand II在结构上更趋紧凑化，运用铝合金骨架和轻量化部件，同时在每根手指设置3~4个关节自由度，使抓取和姿态调整更加灵活。近年来，3D打印与复合材料广泛应用于骨骼制造领域，使结构既轻质又具备足够的强度，而关节的多冗余自由度成为支持复杂手部姿态的基础条件。

驱动技术是灵巧手演进的核心难题。最早期的灵巧手普遍采用“直流电机+腱索传动”模式：电机安装在掌部，通过高强度柔性绳索驱动远端关节实现屈伸。然而，这种方式存在张力调节复杂、响应延迟和控制耦合等问题。为此，一些设计转向直接驱动与谐波减速集成，如DLR Hand II使用微型伺服电机结合谐波齿轮，直接驱动各关节，显著提升了位置精度与刚度控制能力。

与此同时，另一条重要路径是柔顺性驱动。以Shadow Dexterous Hand为代表，该系统采用气动人工肌肉（PAM）来驱动各手指关节，并通过空气压力控制肌腱长度，实现高度柔性的输出力和顺应性抓取。该设计在应对未知物体或脆弱表面时具有天然的柔性特性，避免了刚性抓取带来的破坏风险。近期，SEA在部分驱动电机中得到应用，SEA在输出端串联弹性元件（如扭簧、弹性腱），既能测量力矩又能提供被动柔顺，兼顾精确力控与安全性。例如，NASA Robonaut Hand采用了部分SEA驱动模块，实现了对小型部件的柔顺抓取。

结构简化与自适应联动也取得了关键性进展。Robotiq三指夹爪采用少数驱动器控制多自由度。在抓取过程中，指节可自动贴合物体表面，通过机械耦合联动关节运动，大幅降低控制复杂性与成本。这类自适应联动结构在工业夹爪与服务机器人中得到广泛应用。

在末端感知方面，灵巧手的发展重心经历了从单点力传感到分布式高密度触觉传感的转变。早期机械手仅在部分指尖安装应变计或压力传感器，难以完整表征接触面的受力分布。近年来，基于视觉的触觉传感器成为重大突破，其通过软性透明凝胶与嵌入式相机结合，记录指尖与物体的接触形变纹理，再经过视觉重建可输出高分辨率的接触形貌、摩擦力及微滑动信息，分辨率可达几十微米。此外，电阻/电容式柔性触觉阵列逐步集成指腹与掌面信息，提供大面积分布式触觉。

在灵巧手尺寸与重量优化方面，随着小型化电机和微型谐波减速器的发展逐渐成熟，现代灵巧手更加紧凑。以Allegro Hand为例，4根手指共16个自由度，但其重量仅约1.2kg，并具备关节位置闭环控制功能。新一代设计也引入模块化指部设计，支持拆卸维护与多任务的快速替换。

总体来看，现代五指灵巧手的硬件发展呈现出以下特征。一是机械骨架轻量化与仿生结构：采用碳纤维、铝合金或3D打印高分子材料，并在拇指设计上尽量使其具备接近人类手的对置与对握能力。二是驱动模式多样化：刚性直接驱动、柔顺性PAM、SEA及自适应联动结构并行发展，满足精密操作与柔顺抓取的不同需求。三是高密度末端感知：以视觉触觉、柔性阵列传感为代表的新型传感技术，显著提升对接触过程的观测与反馈能力。四是尺寸紧凑化与高度集成：微型驱动与集成控制板的进步，使灵巧手在重量和封装上更适配人形机器人。

未来，灵巧手将进一步在材料、驱动冗余、柔顺控制与感知融合上持续进化，向人手功能极限逼近，为服务机器人、工业协作，以及复杂操作任务提供更可靠的执行基础。

3. 腿足

人形机器人腿足系统的发展，体现了机器人在动态平衡、全身协调和地形适应方面对硬件技术的持续探索。早期的人形机器人，如20世纪70年代的WABOT-1，主要依赖刚性电机驱动的串联链式腿部结构，并配合简单的步态规划，按照预先设定的轨迹执行静态步行。这类系统通常使用直流电机或液压缸作为动力源，驱动高刚性连杆实现摆动，但重心和脚部接触的实时调节能力有限，步态灵活性较差。此阶段，其步行主要局限于低速匀速条件，对地形要求严格，且缺乏对动态扰动的响应。

随着伺服驱动器、减速器和多轴传感器技术的进步，20世纪80年代至20世纪90年代，日本本田公司研发的P系列和ASIMO机器人标志着人形机器人腿足硬件的重大突破。ASIMO采用高功率密度电机（大多为无刷伺服电机）结合行星减速器和谐波减速器，形成12个自由度的下肢系统（每条腿6个自由度）。足部集成了六维力/力矩传感器（六轴力传感器），用于精确测量支撑面的反作用力。基于ZMP的实时步态控制，ASIMO能在有限程度上实现动态步行和简单的应急平衡，推动了人形机器人腿足系统从静态稳定到准动态平衡的进化。

进入21世纪，腿足的驱动和传感硬件开始全面升级。首先，在驱动系统方面，高扭矩密度电机逐渐取代体积庞大的液压缸，结合小型谐波减速器，实现高集成、轻量化的关节模块。典型如Boston Dynamics的Atlas，采用电液混合驱动器，髋关节和膝关节仍由高响应液压执行器驱动，以提供超高瞬时力矩，实现跳跃、翻

滚等高动态动作；踝关节则更多由电机驱动，兼顾位置控制精度与能耗。Atlas的液压关节利用伺服阀高速调节液压流量，并结合高精度位置传感器和力传感器，能够在极短时间内施加爆发性关节力矩，支撑复杂运动。

早期人形机器人的足底通常使用大面积平板脚趾，配合硅橡胶防滑层，通过增大支撑面积来提高静态稳定性。但这限制了它们的灵活性和地形适应性。近年来，足底逐步集成多点压力传感器阵列（如电容式或薄膜式），形成高分辨率触觉感知，能够估计接触分布、摩擦力和微小滑移。部分研究也探索可动脚趾和分段足底结构，提高与不规则地面的贴合度。例如，Digit机器人采用分区橡胶脚垫和力/触觉传感器，实现在平滑地形上的行走。

腿足的柔顺性技术得到显著发展。传统的刚性驱动在遭遇外界冲击时易产生结构过载和振动，因此，SEA被引入关键关节。SEA在驱动电机输出与负载之间引入弹性元件（如扭簧），并利用弹性形变量测量输出力矩，从而实现高带宽的力控回路。Unitree H1、Agility Digit等人形机器人的腿部关节普遍集成SEA，这既提升了安全性，也显著增强了脚部对地面细微扰动的响应能力。相比传统纯刚性关节，SEA可以在外力冲击时的短时间内吸收动能，有助于实现动态平衡。

此外，直接驱动关节也成为高端人形机器人追求高动态响应的重要方向。例如，Tesla Optimus在部分关节中采用低速大扭矩电机直接驱动，避免了减速器的机械间隙和摩擦，提升了控制精度和响应速度，但对电机设计、电流驱动能力，以及热管理能力提出了更高要求。与之配套的多模惯性传感器、激光雷达与立体相机，构成实时状态估计和环境感知的核心。

在驱动和传感之外，腿足的结构设计也显现出轻量化与模块化趋势。合金框架与碳纤维复合材料被大量应用，显著降低了下肢惯性，提高了步态调整灵活性。模块化关节与即插即用接口有助于加速关节的组装与维护，也便于不同任务下的配置优化。

综上所述，现代人形机器人腿足系统在硬件层面经历了从刚性电机驱动到高柔顺性、高集成性、多模态感知的演化。高性能伺服驱动器或液压驱动器、串联弹性执行器、六维力传感器、分区足底触觉阵列、轻质结构材料与多传感器融合，共同构成了先进的腿足平台，为实现接近人类的灵活步态和动态平衡能力提供了技术支撑。

4. 相关零部件

（1）电机

电机作为机器人驱动系统的核心组件，其技术路径的演进不仅决定了机器人能否实现高精度、高响应与高柔顺的控制能力，也直接影响其能效、集成性与操作边界。机器人领域的电机技术发展，始终围绕着功率密度提升、响应动态增强、控制精度提高与结构紧凑化等目标展开，呈现出从传统刚性电机向智能化、柔顺化与仿生化方向持续演进的趋势。

机器人最早期普遍采用有刷直流电机，其控制简易、成本低廉，适用于开环或低精度闭环系统。然而，该类电机存在明显的碳刷磨损、电磁干扰与维护复杂等问题，难以满足长时间稳定运行的工业级应用需求。随着永磁材料和电力电子技术的发展，无刷直流（BLDC）电机逐渐成为主流。BLDC电机消除了机械换向装置，在提升效率与可靠性的同时，具备更高的转矩密度与更低的噪声，广泛应用于关节型机械臂、移动机器人底盘与灵巧手驱动器中。此类电机通常搭配位置编码器与伺服控制器，以实现较高精度的角度控制。例如，UR5e机械臂的每个关节均使用高性能BLDC和绝对值编码器，实现亚毫米级的重复定位精度。

为进一步满足刚性结构下的高负载与高精度任务需求，交流伺服电机得到广泛部署。相比于BLDC电机，该伺服系统具备更强的三环控制能力（位置-速度-力矩），并通过集成电流传感器与转矩估计器实现更稳定的动态响应。KUKA、ABB等高端工业臂所采用的伺服系统通常具备良好的抗干扰能力与散热结构，在长时间的连续运作中表现出卓越的稳定性。

为突破“传统电机+减速器”架构在背隙控制、传动迟滞与结构紧凑性方面的瓶颈，近年来，无框力矩电机逐步成为高端机器人关节单元的核心方案。该类电机通常为低速大扭矩的永磁无刷电机，采用开放式定转子结构，省去外壳与轴承，便于直接嵌入机器人关节、臂段或旋转轴中。相比传统封装式伺服电机，无框力矩电机具备更高的集成度与动态刚度，显著减少传动环节带来的机械间隙与惯性干扰。例如，宇树Go1、优必选Walker等国产平台已大面积部署自研无框力矩电机模组，并结合高精度谐波减速器与力矩控制器，实现了高性能的位置与力控双闭环。

在特定微型化与高响应任务中，空心杯电机同样展现出独特优势。该类电机通过去除铁芯结构，极大降低了转子惯量，具备极快的启停响应与较小的力矩波动，适用于轻载高频场景，如教育机器人、机械手指、视觉对焦机构等。虽然其力矩输出能力有限，但其结构轻巧、响应灵敏的特性使其在医疗微创执行器与微型机器人系统中具有重要地位。

从系统视角看，这两类电机在机器人中的广泛应用推动了“电机-传感-控制”的深度融合，使驱动单元逐步从传统动力源演化为高度集成、感知与控制一体化的智能执行器。特别是在具备控制器局域网总线（CAN）、以太网控制自动化技术

（EtherCAT）等工业总线通信能力后，现代无框电机模组已经能够独立承担位置、速度、力矩等多模式任务，为构建高性能、轻量化、多自由度的机器人平台奠定了基础。

与此同时，机器人电机技术也在不断向柔顺驱动与仿生执行方向拓展，以提升系统在动态环境下的交互安全性与适应能力。典型代表如SEA，其通过在电机输出端串联弹性元件，在具备精确力矩估算能力的同时，引入机械柔性以吸收冲击、避免意外接触。NASA Robonaut与MITCheetah等平台均在多关节上采用SEA，实现兼具性能与柔顺的控制特性。另一类方向则是柔性材料驱动单元，如PAM与电介质弹性驱动器（DEA），分别通过气压与电场驱动实现仿生收缩行为，虽尚处于特种或实验应用阶段，但在软体机器人、可穿戴设备与人机融合系统中展现出重要潜力。

总体来看，电机在机器人系统中的角色正在发生本质变化：从传统的动力部件，向具备精密感知、柔顺响应与系统协同能力的智能执行模块演化。无论是面向高性能控制的无框力矩电机，还是面向极端响应的空心杯电机，抑或是服务于安全交互的柔性驱动结构，其技术发展均指向一个共同目标——实现更接近人类运动能力与环境适应性的机械智能体。

（2）减速器

减速器作为机器人驱动系统中连接电机与负载之间的关键传动环节，其核心功能在于将高速低扭矩的电机输出转化为低速高扭矩的有效驱动力，同时提升系统的定位精度与动态刚度。在传统工业机器人中，减速器一度被视为性能瓶颈，其体积、效率、背隙与寿命直接影响整机的运动性能与使用寿命。随着机器人应用边界的持续拓展，减速器也逐步走向高精度、高效率与小型化方向，呈现出多路径并行发展的态势。

最早期的机器人广泛采用普通斜齿轮或行星齿轮减速器，以满足基本传动需求。此类减速器结构简单、加工技术成熟，但其体积庞大、传动噪声大，且背隙较难控制，不适合精密定位任务。随着六轴工业机器人对扭矩密度与输出稳定性的提升需求，摆线针轮减速器与谐波减速器成为两大主流减速器。

摆线针轮减速器通过摆线盘与针齿啮合实现多齿同时传动，具备高承载能力、高冲击抗性与较长寿命，广泛应用于重载型工业机器人中，如安川、川崎、新松等品牌的机器人。典型的如RV减速器，其内部包含双级减速结构，前级为行星齿轮，后级为摆线针轮，结构刚度高，适合安装在近基座，且负载较重的关节上。其优势在于输出轴径向与轴向刚度高、热稳定性好、背隙极小（可控制在1弧分以

内），可承受高反复载荷和惯性冲击。但其结构复杂、加工精度要求高，制造成本较高，且不适用于微型场景。

另一条关键技术路径是谐波齿轮减速器，由柔轮、刚轮和波发生器3部分构成，利用柔性金属变形实现齿轮啮合，从而获得极高减速比（一般在50：1~160：1）与极低背隙（小于1弧分）。由于其啮合齿数远高于传统齿轮，具备体积小、重量轻、传动精度高等优势，成为中小型关节模块首选。典型的，如UR系列协作机械臂与Allegro Hand灵巧手中的微型关节模组，均采用“BLDC+谐波减速”一体化方案，扭矩密度与控制精度显著提升。然而，谐波减速器存在疲劳寿命有限、材料疲劳蠕变与效率波动问题，尤其在高频正反转工况下，柔轮极易老化变形，这成为其应用限制因素之一。

近年来，面向更高扭矩密度与效率平衡的需求，谐波-行星混合减速器成为一种新兴趋势。该结构将传统行星齿轮用于前级减速，谐波结构用于末级输出，结合二者的刚度与精度优势，实现更高减速比的同时抑制柔轮的疲劳损耗。部分国产机器人如遨博、节卡已在其轻型臂产品中集成该类混合减速模组，在低成本与高性能之间实现较好平衡。

（3）丝杠

丝杠作为将旋转运动转换为直线运动的关键部件，是线性执行单元中最常见的核心机构之一，广泛应用于各类关节型机械臂、三轴平台、协作机器人、医疗康复机器人及高精度加工装备中。其传动性能直接决定了系统的定位精度、动态响应、承载能力与寿命稳定性。在机器人领域，随着对高速性、柔顺性与微纳精度需求的不断提升，丝杠技术也从传统滚珠丝杠向多类型复合结构演进，呈现出精密化、小型化与智能集成的发展趋势。

最早期的机器人平台多采用梯形丝杠，其螺纹剖面为梯形结构，制造成本低、结构简单，具备自锁能力，适用于低速、低负载场景。但由于梯形丝杠具有滑动摩擦结构，摩擦系数大，效率较低（一般不超过40%），易磨损且定位精度不高，因此难以满足现代机器人的高频快速响应与长寿命运行需求。

随着机器人精密化要求的提升，滚珠丝杠成为主流选择。该结构在丝杠轴与螺母间布置循环滚珠，通过滚动接触替代滑动摩擦，显著提升了传动效率（高达90%）与使用寿命，同时降低了驱动扭矩与发热量。滚珠丝杠具备高刚度、高回转精度与良好轴向负载能力，广泛应用于坐标运动平台、搬运机器人线性轴与高精度执行单元。例如，Fanuc、Kawasaki等工业机器人底座线性轴均采用高精度预压型滚珠丝杠，实现高速下稳定的直线定位。

不过，滚珠丝杠也存在一些技术限制。其结构依赖循环滚珠通道，存在滚珠回流噪声与冲击问题，在超高速运动（大于1m/s）下容易出现共振和振荡。此外，微小背隙及摩擦滞后仍会对微位移控制造成干扰，难以满足亚微米级运动控制的高端应用需求。

为进一步提升丝杠系统在高载荷、长寿命与高频冲击下的性能，行星滚柱丝杠（PRS）逐渐成为高端机器人线性驱动的关键方案。与滚珠丝杠相比，PRS采用多个螺纹滚柱在丝杠与螺母之间进行行星运动，实现多点线接触滚动，接触面积大、载荷分布均匀，因此其轴向承载能力通常为同尺寸滚珠丝杠的5~10倍，使用寿命也得到大幅提升。该结构特别适用于工业重载机器人、伺服压装设备、航空制动器及航天展开机构等对推力、刚度与疲劳性能要求极高的场合。

在机器人领域，PRS的典型应用包括大型轨迹焊接臂、重载协作机械臂及空间机器人中的可展开臂机制。例如，Airbus在空间结构部署系统中采用行星滚柱丝杠确保高可靠性的伸缩运动；ABB与KUKA的部分高负载直线单元也引入高精度PRS驱动模块，以延长维护周期并提升动态响应。更为重要的是，行星滚柱丝杠具备出色的可控微进给能力，可实现极低速、高精度输出（微米级步进），同时其刚性远高于传统丝杠，不易因热膨胀或冲击负载引起失稳，具备优异的重复定位性能。

不过，PRS的缺点也不容忽视。其制造工艺复杂、滚柱与螺纹型线精度要求极高，加工与装配成本远高于滚珠丝杠，不适合中小型或成本敏感的机器人系统。此外，由于滚柱滚动间隙极小，对润滑与清洁度要求也更为严苛，因此多部署于对性能有极致要求的军工、航空航天与大型工业机器人系统中。

此外，静压丝杠、磁浮丝杠等产品也在不断涌现。静压丝杠通过高压油膜在螺母与丝杆之间建立承载间隙，实现无金属接触的润滑支撑，其摩擦力极低，可实现超高平顺性与无磨损运行，适合高负载与超精密定位任务。例如，在半导体制造装备、超精密光学加工机器人平台中，静压丝杠因其亚微米分辨率与纳牛顿级稳定推力而受到人们青睐。但此类结构依赖外部液压系统，成本高、体积大、维护复杂，尚不适用于轻量级机器人系统。磁浮丝杠则基于电磁悬浮原理，在丝杆与螺母之间建立磁悬间隙，消除传统结构中的接触摩擦与结构刚度限制。该类结构可通过线圈驱动实现非接触线性传动，具备理论上的零背隙、零磨损与高响应速度。尽管目前其仍处于实验研究阶段，但已在高端科学实验平台和微纳定位系统中取得初步成果，未来有望应用于机器人高精度末端执行器与微创医疗机器人中。

总体而言，机器人丝杠技术的发展沿着以下主线演进：从传统滑动摩擦式梯形丝杠向高效率滚珠丝杠过渡；在高端应用中向无接触静压或磁悬浮结构拓展；在人机协作中发展具备柔性与安全性的弹性结构；同时在集成化方向推动丝杠、电机

与控制单元一体化。不同类型的丝杠技术有不同的适配场景，高精度加工机器人和医疗机器人青睐无摩擦类传动，物流搬运与工业平台则偏好高性价比滚珠丝杠，而协作机器人则更关注结构顺应与控制接口融合。未来，丝杠作为关键线性执行部件，其发展不仅依赖材料与工艺的优化，更将与传感、智能控制等系统实现深度融合，推动机器人从结构运动向感知运动与智能决策的跨越。

（二）芯片的主要技术路径和发展趋势

1. 端侧推理芯片

在机器人技术飞速发展的今天，人工智能推理正从遥远的云端走向机器人本身。过去，机器人需要将采集到的海量数据（如高清图像、点云数据）上传到云端进行处理，再将决策结果回传。这种模式的致命弱点在于其高延迟和不稳定性。对于需要毫秒级反应的自动驾驶、协作机器人或工业抓取任务，任何网络延迟都可能导致灾难性后果。为了解决这一瓶颈，端侧推理芯片技术应运而生。这些芯片如同机器人的“大脑皮层”，让它们能够即时感知、思考和决策，而不再依赖远程服务器。

传统的通用CPU或GPU虽然强大，但并非为神经网络的计算特点而设计。现代的端侧芯片则采用专门为AI任务进行优化的硬件加速器，如专用集成电路（ASIC）和现场可编程门阵列（FPGA）。这些架构通过大规模并行计算，专门针对神经网络中频繁的矩阵乘法和卷积运算进行优化，实现了惊人的运算效率和更低的功耗。一些厂商甚至集成了NPU，使其成为处理AI工作负载的核心。神经网络模型在训练时通常使用高精度的浮点数，这会占用大量内存和算力。现在的端侧芯片不再是单一处理器，而是将多种计算单元集成在同一片芯片上，形成一个异构计算平台。通常包括一个低功耗CPU来处理操作系统和任务调度，一个强大的GPU用于并行计算，以及一个专用的NPU来高效运行推理模型。这种分工合作的模式，使得芯片能够同时处理不同的任务，将资源利用率最大化。此外，在机器人应用中，续航至关重要。因此，端侧芯片的设计重点不再仅仅是提高算力，而是优化每瓦特的性能。通过引入量化技术（如将浮点数精度从FP32降低到INT8），芯片能够在保持推理精度的同时，大幅减少计算量和功耗。同时，优化内存带宽和缓存结构也成了降低能耗的关键。以下是几款市场上的端侧推理芯片。

（1）NVIDIA Jetson系列

来源：NVIDIA公司。

NVIDIA Jetson系列平台是目前机器人和边缘AI领域应用最广泛的端侧推理平台，涵盖从入门级Nano到旗舰级Orin的多条产品线。其核心架构结合了NVIDIA自主研

发的CUDA GPU与NVDLA深度学习加速单元，具备强大的并行计算能力，支持FP32、FP16和INT8等多种精度模式。NVIDIA Jetson系列平台最大的特点是完整的软件生态，JetPack SDK集成了TensorRT、cuDNN等高性能推理库，并支持多种主流深度学习框架（如TensorFlow、PyTorch等）的部署优化。Jetson Orin等高端型号单芯片算力可达275 TOPS^[11]，能够实时运行大型卷积神经网络和Transformer模型。该系列平台广泛应用于自主移动机器人、智能巡检、工业视觉检测等场景，成为高性能边缘推理的代表性方案。

(2) Google Edge TPU

来源：谷歌公司。

Google Edge TPU是专为超低功耗边缘推理设计的专用集成电路，核心产品形态为Coral系列模块和USB加速棒。Edge TPU的硬件架构高度专用化，仅支持8位整数（INT8）量化模型^[12]。这种设计使其能够以最小功耗实现高吞吐率，峰值性能达到4TOPS，且每TOPS功耗仅约0.5W。Edge TPU依托TensorFlow Lite生态，需在训练阶段完成量化感知训练后进行量化转换，模型可在PC或云端准备完成后直接迁移到设备端进行推理。其极低延迟（毫秒级响应）和极简部署流程非常适合物联网节点、微型机器人和智能传感器。相比通用NPU，Edge TPU牺牲了部分灵活性以换取效率，在大规模部署中具备显著的成本和功耗优势。

(3) Intel Movidius Myriad

来源：英特尔公司。

Intel Movidius Myriad系列是面向视觉AI推理优化的专用VPU，其代表性产品Myriad X已集成在Neural Compute Stick 2等模块中。其硬件包含12个高并行度的SHAVE核心，通过搭配专用的硬件视频编解码和图像信号处理（ISP），能够高效完成端到端视觉分析和神经网络推理任务。Myriad X支持FP16精度，推理性能在1~4TOPS之间，功耗约为1.5~3W^[13]。基于Intel OpenVINO工具链，开发者可便捷地将Caffe、TensorFlow、ONNX等多种模型移植到Movidius平台。该芯片特别适合工业视觉检测、嵌入式摄像机和低功耗边缘AI终端，兼具视频处理与神经网络加速能力。

(4) 华为昇腾Atlas 200 DK

来源：华为技术有限公司。

华为昇腾Atlas 200 DK是基于昇腾310 AI处理器的边缘AI推理开发板，集成达芬奇架构NPU，具备强大的混合精度推理能力。昇腾310支持8 TOPS（INT8）或16 TOPS（FP16）性能，配备ARM Cortex-A55 CPU、H. 264/H. 265视频编解码模块，以及丰富的I/O接口。Atlas 200 DK提供计算机人工神经网络（CANN）开发套件和MindSpore Lite推理框架^⑤，支持TensorFlow、Caffe、ONNX等模型格式。该平台在智能交通、安防监控、工业视觉等场景中得到大规模应用，尤其在国产化自主可控方面具有显著优势。通过结合硬件加速与模型量化，Atlas 200在功耗和性能平衡方面表现突出。

（5）高通Hexagon DSP/AI Engine

高通Hexagon AI Engine是嵌入在Snapdragon移动SoC中的异构AI加速引擎，属于高度集成化的端侧推理方案。其核心为Hexagon DSP与专用神经网络加速器（NPU），支持INT8、FP16、BFLOAT16等多种数据格式。Hexagon具备可编程的向量扩展和深度学习优化指令集，可实现超过75 TOPS的峰值算力（在旗舰SoC中）。其最大特色是与Adreno GPU、Kryo CPU紧密协同，通过Snapdragon NPE SDK实现深度学习模型的高效部署。Hexagon广泛应用于智能手机、服务机器人、AR/VR等移动设备中，是面向低延迟、多模态边缘推理的重要基础。其能效表现突出，能够在有限的热设计功耗下支持复杂的AI推理。

2. 训练芯片

训练芯片的核心技术发展围绕高效执行深度神经网络的计算需求展开，重点解决大规模矩阵运算、内存带宽瓶颈与能效比优化等问题。在硬件架构层面，训练芯片逐步从通用CPU过渡到以GPU、TPU和专用NPU为代表的高并行张量计算架构。GPU采用（单指令多数据/单指令多线程）SIMD/SIMT并行模式，配备大规模CUDA核心，专注于加速卷积、矩阵乘法等关键算子，适用于深度学习中的前向传播与反向传播过程，因其灵活性与成熟的软件生态成为主流训练平台。为了进一步提升计算密度，现代GPU引入Tensor Core等混合精度计算单元，利用FP16/BF16乘加指令，在保持精度的同时提高吞吐量，并降低功耗，尤其适用于强化学习中的策略网络训练。与此并行发展的是张量处理器（如TPU）的兴起，其采用Systolic Array结构实现高度数据复用，适合批量大、结构稳定的深度模型训练，尤其对Transformer类模型具有效率优势。

在内存系统方面，训练芯片强调高带宽内存（如HBM）、片上缓存优化与NoC互联结构，通过减少访存延迟与数据搬移成本提升整体训练效率。同时，混合精度训练成为技术趋势之一，在保持数值稳定性的基础上显著压缩模型存储与计算资源消耗。部分训练芯片进一步支持异构协同结构，并结合通用CPU、专用张量核、DMA引擎与编解码模块，实现灵活的数据调度与算子编排。随着多任务、分布式与

在线学习的需求提升，芯片层面还集成了张量调度引擎、环境状态缓存、指令融合与图优化机制，提升强化学习等复杂任务的训练吞吐与响应效率。总体来看，训练芯片的技术路径正朝着高密度并行计算、精度可调、自适应调度与异构集成方向演进，为机器人等智能系统提供更强大的模型训练支撑。以下为几款训练芯片产品的介绍。

(1) NVIDIA A100 Tensor Core GPU

来源：NVIDIA公司。

NVIDIA A100是当前深度学习训练领域最具代表性的通用GPU加速器，基于Ampere架构设计，集成了6912个CUDA核心与432个第三代Tensor Core，专门优化矩阵乘法和深度学习算子。每个Tensor Core支持FP32、FP16、BF16、TF32等多种数据精度，兼顾了高精度与高吞吐^①。A100配备40GB或80GB的HBM2e高带宽显存，单卡内存带宽可达2TB/s，支持极大规模神经网络权重与激活的高速读写。通过NVLink与NVSwitch互联技术，多个A100可组成单节点或多节点GPU集群，进行大规模分布式数据并行和模型并行训练。A100广泛应用于自然语言处理、计算机视觉、推荐系统等领域，既作为深度学习训练主力，也支持推理加速，具备良好的通用性。

(2) Google TPUv4

Google TPUv4是Google为云端大规模深度学习训练定制设计的专用加速器，重点面向Transformer等超大规模模型的高效训练。核心计算模块为 128×128 的矩阵乘法单元（MXU），能够进行高效的矩阵乘加运算，实现每秒275万亿次（275TOPS）的峰值算力。TPUv4采用HBM内存和专有的低延迟光纤互联，最多可将数千个TPU通过Pod组网，构建出超过1000 Petaflops的集群能力。与GPU相比，TPU在功耗、性能比和推理延迟上具有明显优势，尤其在批量矩阵运算场景下的优势显著。TPUv4同时支持BF16、INT8等混合精度训练，能够有效提升能效比，是Google云端服务（如TPU VM）中Transformer模型训练的重要硬件基础。

(3) Cerebras Wafer Scale Engine

来源：Cerebras Systems。

Cerebras Wafer Scale Engine（以下简称Cerebras WSE）是目前全球最大的单芯片AI加速器，突破了传统芯片封装与互联的物理限制，直接以整片12英寸硅晶圆为计算基底，形成高达 46225mm^2 的芯片面积^②。Cerebras WSE包含超过85万个AI优化核心，内置40GB片上SRAM（静态随机存取存储器），峰值内存带宽高达20PB/s，极大降低了神经网络层间的数据搬移延迟。其核心架构支持张量稀疏性

优化、可扩展张量并行及模型并行机制，可将深度学习的前向、反向和梯度更新在片上完成，显著提升了大型模型的训练效率。Cerebras WSE广泛应用于超大模型（百亿至万亿参数）的快速训练，配合Cerebras软件堆栈（包括Graph Compiler和Weight Streaming），为科研和工业界提供了全新的AI算力范式。

3. CPU

CPU技术始终围绕通用计算性能、指令处理效率、并行处理能力和能效比进行迭代优化。传统CPU采用冯·诺依曼架构，具有统一的控制流与指令流，通过流水线技术提升单指令执行效率。随着处理器频率逐渐接近物理极限，单核提频的空间受限，现代CPU架构转向多核并行与乱序执行的方向，通过超标量结构、乱序执行（Out-of-Order Execution）、分支预测、指令预取等机制提升并发处理能力。在微架构层，核心单元包括指令译码器、执行单元、寄存器重命名单元、负载/存储队列与缓存子系统，这些模块协同工作以实现高指令吞吐率。

随着异构计算需求增长，现代CPU逐步支持多级缓存体系（如L1/L2/L3）与大容量共享内存，提高多核间的通信效率，并引入同时多线程（SMT）技术，在每个物理核心上实现多个逻辑线程，提高资源利用率。芯片制造工艺的进步（如FinFET、GAAFET技术）也使得芯片在单位面积内集成更多晶体管，实现更高的运算密度与更低的能耗。此外，为满足AI、图像处理、加密等特定应用，CPU也开始集成专用指令集扩展（如AVX、SSE、NEON）与协处理器模块，部分架构（如Apple M系列、ARM DynamIQ）将CPU核心与NPU、GPU、ISP等单元组合为SoC，实现通用与加速计算的协同执行。在多芯片系统中，CPU还承担调度与控制中心角色，负责I/O中断管理、资源分配与数据同步。以下为几款CPU产品的介绍。

（1）Intel Atom/Intel Core系列

来源：英特尔公司。

Intel Atom和低功耗Intel Core i3/i5处理器广泛应用于中小型服务机器人和移动平台。Atom处理器采用x86微架构，超标量乱序执行与SMT结合，提供较好的计算性能和能效比。其L2缓存通常在2MB以内，支持低功耗C-states以延长电池续航^②。搭配片上GPU（如Intel HD Graphics），能够直接进行2D/3D可视化渲染与视频解码，减少对外部显卡的依赖。在ROS中，Atom CPU负责感知数据融合（如激光雷达与IMU）、路径规划、运动控制，适合负载有限、功耗敏感的机器人，如清扫机器人或低速自动导引车（AGV）。

（2）NVIDIA Carmel CPU

来源：NVIDIA公司。

NVIDIA Jetson Xavier NX/AGX模块集成了6核或8核的ARM v8.2架构Carmel CPU。每个核心支持3-issue超标量乱序执行，支持NEON SIMD扩展，以及64KB L1指令+数据缓存、4MB共享L2缓存^②。CPU主要用于高层逻辑和调度。例如，多线程传感器驱动、中间件通信、任务分配，并与片上Volta GPU和NVDLA协作实现AI推理加速。Carmel CPU拥有更高能效比（每个核心仅数瓦功耗），在大型移动机器人（如AMR、自主配送车）中既能胜任Linux系统运行，又支持复杂ROS 2多节点分布式控制。

（3）ARM Cortex-A系列

来源：Arm公司。

ARM Cortex-A53/A72/A78系列广泛应用于机器人嵌入式主板，如树莓派Compute Module、NXP i.MX8平台、Rockchip RK3399等。Cortex-A处理器基于ARMv8-A 64位指令集，支持乱序执行、硬件虚拟化和NEON SIMD加速。片上缓存的设计典型为32KB L1I/D、1MB L2，部分高端型号还集成了2-4MB L3缓存^③。其主要优势在于低功耗性能和成熟的生态（Yocto/Ubuntu等），常用于低成本室内机器人、机械臂末端控制器或IoT边缘AI设备中。

4. 控制器芯片

机器人上的控制器芯片主要用于实现对传感器数据的实时处理、对执行器的精确驱动，以及对任务级逻辑的控制，其技术路径聚焦高实时性、低功耗、高可靠性与接口灵活性的协调发展。早期机器人多采用通用微控制器（如AVR、STM32）完成运动控制与简单逻辑处理，这类芯片通常具备较低主频、少量随机存取存储器（RAM）和Flash，依靠中断机制与定时器完成对电机、编码器与传感器的响应。但随着机器人任务复杂性不断提升，控制器芯片的计算能力、并发性与多任务处理需求同步增长，推动其技术向高性能嵌入式架构演进。

当前，机器人控制器芯片多基于ARM Cortex-M、Cortex-R甚至Cortex-A系列架构，在保证实时性的同时引入更强的浮点运算单元与数字信号处理器（DSP），适配更复杂的动力学建模、状态估计与控制算法，如卡尔曼滤波、模型预测控制（MPC）与轨迹优化等。部分控制器芯片还集成硬件PWM、ADC、CAN、RS485、SPI等工业级接口模块，提升对多传感器、多执行器系统的适配性。在安全性和确定性方面，实时操作系统（RTOS）在芯片上的高效运行成为标配，如FreeRTOS、Zephyr或VxWorks，结合双核锁步架构与硬件定时器，满足工业机器人对控制稳定性与任务可靠性的严苛要求。为了进一步压缩延迟并提升多轴协同控制能力，一

些高端控制器芯片采用FPGA/SoC异构结构。此外，控制芯片也在逐步引入AI加速功能，支持轻量化神经网络推理，在不依赖外部推理芯片的情况下实现边缘感知与控制融合，特别适用于服务机器人、自主移动平台等具有视觉伺服需求的应用。以下为几款机器人控制器芯片产品的介绍。

(1) TI C2000系列 (TMS320F2837xD)

来源：德州仪器公司。

TI C2000系列数字信号控制器 (DSC) 是机器人伺服电机和运动控制系统中的经典平台。其核心是高性能C28x内核^④，具备32位浮点运算单元和丰富的DSP指令集，可在微秒级内高效执行[比例 (P)、积分 (I)、微分 (D)]PID调节、前馈补偿及坐标变换等复杂控制算法。C2000内置的控制率加速器CLA是一种硬件协处理器，专门用于控制回路的并行计算，可在不打断主CPU的情况下完成电流环与速度环计算。这一架构确保了其在高采样率 (10~100kHz) 下，依旧具备稳定可靠的控制性能。其多通道 (脉冲宽度调制) PWM模块 (高达24路) 和高速 (模-数转换器) ADC (最高16bit) 非常适配多轴伺服电机控制场景，如六自由度机械臂的同步驱动。C2000还集成了CAN、EtherCAT等工业通信接口，方便构建分布式实时控制网络，因此在UR机械臂、KUKA末端关节模块和AGV底盘驱动中都有大量应用。

(2) ST STM32系列 (STM32H7/F4)

来源：STMicroelectronics公司。

STM32系列微控制器是基于ARM Cortex-M内核的高集成度通用控制芯片，尤其是STM32H7凭借其480MHz主频和双精度浮点运算能力^④，在机器人底层控制中占据重要地位。STM32H7集成了多组高速ADC (2Msps以上) 和高分辨率PWM定时器，支持对直流电机、无刷电机的闭环控制及电流环/位置环的实时运算。依托其丰富的外设 (USART、SPI、CAN-FD、Ethernet)，STM32能同时完成传感器数据采集、执行器控制及通信任务。相较专用DSP，STM32凭借Cortex-M生态具备更好的易用性与生态兼容性，支持FreeRTOS、Zephyr等RTOS。该系列在移动机器人的差速驱动、舵机控制板，以及中小型协作机械臂的关节控制中应用广泛。例如Robotis Dynamixel伺服系统及多款教育型机械臂控制板，就采用了STM32实现电流调节和位置闭环。

(3) Microchip dsPIC33系列

来源：Microchip Technology公司。

Microchip dsPIC33系列的定位为高性价比电机控制数字信号控制器，其采用16位Harvard架构内核^②，支持单周期乘加运算指令和DSP流水线，可在资源受限的条件下高效执行磁场定向控制（FOC）及速度环控制。其高精度PWM输出模块（可配置死区时间、对称/不对称PWM模式）和同步采样ADC，可在高速电流采样和控制闭环中保持稳定性。相比C2000，Microchip dsPIC33集成度更高、封装体积小，适合小功率多轴伺服或舵机驱动场景。其Flash程序存储区支持在线更新，可通过Bootloader在现场快速升级固件。Microchip dsPIC33广泛应用于四轴飞行器电调、轻型服务机器人关节电机控制和低成本AGV轮毂电机控制，是体积与功耗受限场景下的理想选择。

5. 总线芯片

机器人系统中的总线芯片技术主要服务于多模块之间的数据通信与协同控制，其技术路径围绕高可靠性、低延迟、抗干扰能力与多节点扩展性展开，广泛应用于电机驱动、传感器读写、控制器间协调、通信中继等关键环节。根据总线类型的不同，总线芯片的技术发展呈现出多元化趋势，其中具有代表性的包括CAN、RS485、EtherCAT、SPI、I²C等协议所配套的控制芯片和接口模块。

CAN作为机器人上最常用的工业级总线，其控制器芯片具备强健的仲裁机制与差分信号驱动，适合多节点间的实时通信。CAN总线芯片通常集成帧缓冲、位定时器、错误管理与滤波模块，支持硬件自动重传与错误检测，常用于工业机器人、移动底盘或手爪关节控制中。对于需要更强实时性与同步控制的系统，如多轴伺服机器人或协作机械臂，EtherCAT总线芯片成为主流选择。这类芯片集成分布式时钟同步与延时补偿机制，支持小于1μs级别的周期同步控制。EtherCAT从站控制器（如Beckhoff ET1100、TI PRU-ICSS）通常基于FPGA或专用硬核实现，可实现线性级联拓扑中的高速转发、报文处理与端口延迟补偿，满足毫秒级控制回路需求。在近距离低速通信中，SPI和I²C等同步/异步串行总线仍被广泛应用于连接本体控制芯片与传感器模块。此类总线芯片重点优化信号完整性与多从设备支持能力，常用于视觉模组、IMU或触觉阵列的数据采集链路中。近年来，随着机器人系统朝着模块化与异构架构演进，总线芯片也开始集成多协议支持、智能仲裁机制与协议栈卸载功能，并与控制SoC深度融合，推动通信链路向更高带宽、更低时延、更强鲁棒性方向发展。以下为几款总线芯片产品的介绍。

(1) TI SN65HVD230

来源：德州仪器公司。

SN65HVD230是一款符合ISO 11898-2标准的3.3V高速CAN收发器，支持最高1Mbps数据率，具备±16kV HBM ESD防护、±25V共模电压容限及热关断功能，能够稳定工

作于长线缆环境中，为机械臂关节控制器和移动底盘模块提供物理层可靠通信支持^②。其核心功能是实现CAN总线的差分信号接收与驱动，通过将数字控制器输出的TTL电平信号转换为总线级差分电平，使抗共模干扰能力达±12V。在物理层上，该芯片内置欠压检测、波特率自适应（最高1Mbps），以及短路保护和热关断机制，可保证长距离通信的稳定性和安全性。在多关节协作机械臂、AGV、无人叉车等设备中，SN65HVD230被广泛应用于关节控制器和上位机之间的实时通信，支持分布式运动控制、状态监测和故障诊断。

(2)Microchip MCP2515

来源：Microchip Technology公司。

MCP2515是一个独立的CAN控制器芯片，内置全功能CAN协议引擎，支持2.0A和2.0B规范，最高波特率为1Mbps。它通过SPI接口与主控多点控制器（MCU）（如STM32、AVR）连接，使非原生支持CAN的MCU也能实现CAN通信^③。MCP2515内部包括3个发送缓存、两个接收缓存，以及强大的滤波和屏蔽机制，能在多节点场景中有效分发和筛选报文，减轻主控的中断负担。在机器人系统中，MCP2515常与SN65HVD230等收发器配合使用，构建多节点的运动控制和传感网络。例如，移动机器人、机器人末端执行器等，都利用它在分布式CAN总线上进行高效数据交换。

(3)NXP TJA1051

来源：恩智浦半导体公司。

TJA1051是NXP推出的CAN FD（Flexible Data-rate，灵活数据速率）收发器，支持经典CAN和CAN FD协议，最高数据速率达5Mbps^④。其核心优势是具备增强CAN FD信号边沿和差分驱动能力，可在高速传输下显著提高抗干扰性能。芯片内置多级保护：±58V总线瞬态耐受、热关断、电压欠压检测、总线短路保护等。

TJA1051可与支持CAN FD的控制器（如NXP S32K或Microchip ATSAME5x）结合，为机器人提供高带宽、低延迟的数据通路，适用于新一代协作机器人和智能物流设备。在复杂环境（电机干扰、长线缆部署）下，TJA1051凭借其稳定性被大量选用。

（三）传感器的主要技术路径和发展趋势

多模态感知技术能够使人形机器人实现更高的自主性、高效性、通用性，增强对周围环境的局部感知能力，并为机器人提供丰富、稳定、准确的环境数据。室内复杂场景的运动感知信息往往存在多源性、异构性、动态性等特征，同时机器人需要面对环境光强变化、非全局特征感知、遮挡复原与推断等问题。对此，机器

人可通过关联多角度图像信号、激光雷达（LiDAR）等不同模态信息数据，以及融合多源图像数据，完成对周围局部视觉传感器环境的立体重建。针对室外开阔场景与复杂高噪环境，通过图像、LiDAR、热成像检测等多源、多模态传感器的多元融合技术，实现传感器间的优势互补，确保感知信息完备有效。通过目标检测、目标分割等技术，构建感知环境与实际环境的空间映射关系，形成多模态环境数据“一站式”融合智能处理系统，实现对局部环境空间、物体的即时感知与虚拟重建，为人形机器人系统对环境的认知与理解提供可靠的数据来源。人形机器人涉及的传感器种类繁多，如视觉传感器、触觉传感器、听觉传感器、六维力传感器、位姿传感器等。其中，视觉传感器、六维力传感器、触觉传感器是当前人形机器人传感器研究的重点领域。

1. 视觉传感器

在机器人自主感知与交互领域，视觉传感器是支撑环境建模、目标识别、路径规划与运动控制的核心组件。主要视觉传感器技术包括激光雷达和双目相机（Stereo Camera）。

（1）激光雷达

激光雷达是一种主动式传感器，其核心原理是利用激光发射器向目标物体发射脉冲或连续波束，通过测量波束返回接收器所需的时间（Time of Flight, TOF）或相位差，精确估算目标表面与传感器之间的距离。其输出为稠密的三维点云，具备优异的绝对尺度测距能力和较强的光照鲁棒性。激光雷达依照内部测量机制与扫描方式，可以分为机械式旋转激光雷达（如Velodyne系列）、固态扫描式[如光学相控阵（OPA）、微机电系统（MEMS）]，以及Flash LiDAR。具体介绍如下。

来源：Velodyne Lidar公司。

机械式旋转激光雷达：通过电机驱动激光发射和接收模块的360° 旋转，连续扫描周围环境，如Velodyne HDL-32E^①。

来源：Luminar Technologies公司。

固态扫描式激光雷达：利用OPA、MEMS微镜或Flash技术，实现无机械旋转的扫描，其体积更小、成本更低，如Luminar Iris^②、Livox Horizon等。

Flash LiDAR：一次性发射面阵激光，通过二维探测器同步接收反射信号，就像拍摄一张深度“照片”。

激光雷达的优势是测距精度高、抗光照干扰强、全天候工作。当前，高端机械式激光雷达可实现200m以上的有效测距，深度精度可达厘米级，成为室外自动驾驶、配送机器人等长距离感知应用的首选。然而，其缺点在于成本较高、分辨率有限，以及对某些材质表面（高反射或透明）感知能力较弱。

（2）双目相机

双目相机通过对同一场景的两视角成像，利用三角测量推导深度信息，属于完全被动测距方式。双目视觉的关键在于立体匹配：即在两张经过立体校正的视差图中寻找对应特征点并估算视差，随后通过投影模型将视差转换为实际距离。其核心流程包含：立体校准（Stereo Calibration），确定两台相机的相对位置和内部参数；图像校正（Image Rectification），将两张图像变换到视差平面对齐；视差计算（Disparity Computation），利用匹配算法（如SGBM、PatchMatch）寻找同一物体在两幅图像中的对应点，计算视差；深度估计（Depth Estimation），根据三角测量原理，将视差转换为深度值。双目相机的优点在于能够直接输出彩色图像与深度信息的融合结果，其体积小、功耗低、成本更低，更适合室内短距作业和人机协作场景。然而，双目视觉测距精度会随距离平方衰减，对低纹理区域、明暗突变等情况较敏感，测量稳定性不如激光雷达。

（3）其他常见视觉传感器

除了激光雷达和双目相机，机器人还常使用以下视觉传感器。

红绿蓝-深度相机（RGB-D相机）（如RealSense）：通过结构光或时间飞跃法（TOF）测距获取RGB和深度信息。

单目相机：依赖多视几何或深度学习估计场景深度。

事件相机：捕获像素级亮度变化事件，具备高动态范围和微秒级延迟。

现代机器人普遍将激光雷达与双目视觉进行多传感器融合。激光雷达

提供稠密、绝对精度较高的三维几何信息，用于全局定位和障碍建图；双目相机贡献高分辨率纹理与语义线索，用于精细识别与任务操作。二者通过时间同步、空间标定及特征级/决策级融合，可以实现更鲁棒、更连续的环境表征。随着嵌入式计算和深度学习的发展，机器人视觉传感器逐步从单一模态向多模态、深融合、端到端学习演进。未来的自主机器人系统，往往在同一传感框架下集成激光雷达、双目/多目相机、惯性测量单元（IMU）、事件相机等多种传感器，形成统一的高精度环境感知与表征，满足复杂场景下的实时决策和智能交互。

2. 六维力传感器

六维力传感器最初源于机器人对力感知的迫切需求，特别是在装配、插拔和柔顺控制等接触丰富的任务中。其发展首先从刚性金属弹性体+应变片的基本构型开始，这一原理基于在弹性体上布置多组应变片，并通过测量应变变化，推导出三维力和三维力矩分量。早期的设计多采用十字梁或对称多梁结构，用刚性金属（如铝或不锈钢）来保证载荷传递的线性和重复性。但应变信号往往存在显著耦合，因此必须通过精确标定实验构建解耦矩阵，将应变电压映射到六自由度载荷。

在弹性体结构优化阶段，有限元仿真被逐步应用到传感器设计中。研究人员通过分析不同几何结构（如多柱梁、环形梁、切向梁）下的应力分布，设计出既能分离6个力矩分量，又具备高刚度的复合形态，降低了解耦矩阵的条件数，从而提高了测量精度。此时，模拟信号放大器、温度漂移补偿电路也被逐步集成到传感器内部，使力传感器能够在更复杂的工况下长期稳定工作。

随着机器人灵巧手、微创手术、精密打磨等对小型化高分辨率需求的不断增长，六维力传感器迎来了微型化和高灵敏度方向的创新。弹性体采用高应变灵敏度材料（如半导体应变计），辅以低噪声差分放大器，使得分辨率达到0.01N量级。其小型化结构体积被缩小到数十毫米以内，重量低于几十克，能够直接安装在末端执行器甚至指尖中。此外，部分产品利用石英压电元件代替电阻应变片实现高带宽动态测力，其工作原理依赖压电效应产生电荷信号，并通过电荷放大器实现微小力矩测量。这类传感器具备极高的动态响应，适用于高速冲击和振动测量，但因静态漂移较大不适合稳态力控。

数字化和总线通信成为技术演进的另一重要方向。传统传感器多输出模拟电压，需要外部信号调理和采集卡才能进行解耦与滤波。新一代传感器在内部集成24位高精度模数转换器和数字滤波算法，将应变片信号实时转换为数字量，经DSP或FPGA运算完成解耦后，直接通过EtherCAT或CAN总线输出六轴力矩数据。这不仅降低了接线复杂性，也显著提升了其抗干扰能力和实时性，使机器人控制系统能以毫秒级频率接收高质量力觉。

柔性多轴力感知逐步成为新兴方向，以适应软体机器人、外骨骼和可穿戴设备等新应用场景。柔性六维力传感器通常基于柔性电容阵列或薄膜应变传感器，通过多点阵列测量来推导六自由度载荷。另一种路径是基于MEMS工艺，将微型应变计或压阻元件阵列集成到柔性基板中，形成轻薄可弯曲的力矩感知层。这类方案在柔性外骨骼和人机交互设备中显示出优良的贴合性和舒适性，但与金属刚性弹性体相比，依然存在灵敏度和稳定性瓶颈。

从解耦算法角度看，早期产品依赖静态标定矩阵，而高端产品引入了基于机器学习的多模态标定与实时补偿，能够动态修正温漂、载荷非线性和耦合误差。通过对标定数据进行多维回归或神经网络映射，六维力传感器具备了适应复杂工作状态的自适应特性，显著提升了其精度和鲁棒性。

在结构材料层面，六维力传感器也从单一金属材料扩展到复合材料结构。轻量化高强度铝合金依旧是主流弹性体材料，而高端产品则在关键部位局部使用钛合金或不锈钢来提升耐久性。此外，为满足轻量化与高刚度的兼顾，部分设计采用碳纤维-金属混合结构，通过有限元拓扑优化控制应力的集中分布，实现精度与强度的平衡。

在集成化与智能化趋势下，六维力传感器逐渐演变为具备数据处理、健康监测和自诊断能力的智能节点。部分产品内置微控制器，可监测应变片状态和温度漂移，自动识别异常输出并通过网络报告。这使机器人系统具备预测性维护能力，长期运行可靠性得到显著提升。

可以看到，六维力传感器的技术发展始终围绕更高灵敏度、更低耦合误差、更强抗干扰能力，以及更小体积四大方向不断演进。传统的“刚性金属弹性体+应变片”已被高灵敏度、全数字、智能化融合的解决方案超越。同时，柔性多模态感知也为下一代机器人带来了新的可能性。从工业装配到医疗康复，从服务机器人到人形机器人，六维力传感器始终是连接机器人与外界物理交互的关键桥梁。

3. 触觉传感器

基于硅基刚性压力传感器的发展，柔性触觉传感器的工作原理依然主要包括压阻式、电容式和压电式，也有研究采用光学式和磁场式等换能方式来实现触觉感知。采用压阻式、电容式和压电式3种工作原理的柔性器件因制造工艺相对成熟，且能够大面积应用而得到了广泛的研究。

（1）压阻式触觉传感器

应用于柔性触觉传感器的压阻材料主要是能够产生良好弯曲和可拉伸性能的聚合物基复合材料。目前，大量导电纳米材料和纳米复合材料已被证明具有压阻效应，如碳纳米管（CNT）、炭黑、石墨烯、纳米线和金属颗粒等。其中，碳纳米管由于其较高的长径比和优秀的导电性能成为复合材料、微电子、储能和生物医学中常见的导电材料。此外，与纯碳纳米管网络结构复杂且脆弱的制造方法相比，碳纳米管与聚合物结合形成的纳米复合材料往往具有更简单、更稳定、成本更低的制造工艺，同时具有良好的稳定性、导电性和可重复性。

这些由纳米颗粒或纤维组成的柔性复合材料内部的导电路径会随压力导致的材料形变而发生密度和方向的改变，导致传感器件出现不同的电阻值。在复合材料中，导电填充材料组成的三维导电网络在长时间的压力和大压力作用下，其自身结构的完整性决定了压阻传感器的稳定性和鲁棒性。压阻式触觉传感器的制作工艺相对简洁，有助于实现大面积及阵列式传感器的制备和应用，且该传感器抵抗外界噪声干扰的能力强，因此被广泛应用于接触力的检测中。斯坦福大学的鲍哲南组提出了一种超高灵敏度的基于弹性微结构导电聚合物薄膜的压阻式传感器，空心球微结构弹性膜由聚吡咯水凝胶通过多相反应制备而成，其制备触觉传感器能够检测小于1Pa的压力，并且具有响应时间短、循环稳定性好和温度稳定的传感特性。

（2）电容式触觉传感器

电容式是长期以来测量压力、液位、加速度和湿度等物理参量传感器最常用的工作原理。例如，在笔记本计算机、智能手机和其他便携式设备上广泛使用的电容触摸技术。基于成熟的电容检测理论和柔性微机电系统工艺，电容式触觉传感器有望实现以高空间分辨率的方式检测大面积接触信息。

电容式触觉传感器通常由两个相互平行的金属电极或其他电极组成，柔性电容式触觉传感器工作的基本原理是当外界物体接触传感器时，柔性材料的压缩或拉伸形变会改变电容的纵向电介质长度和横截面积，通过传感系统测量电容信号并将其转换为电输出信号，检测电容的变化。实际应用层面中，由于柔性电容电极材料的面积在外力作用下难以实现大范围的变化，同时面积的变化使其难以得到电容线性的输出曲线，所以柔性电容式触觉传感器大多利用改变电极间距离的方式来检测接触力的大小。近年来，电容式触觉传感器因其相对较高的灵敏度和较低的制作成本，越来越多地被集成到机器人和其他需要柔性接触的应用场景中。但是阵列式装配的电容式触觉传感器受电介质特性影响，依然存在迟滞和串扰等问题。

（3）压电式触觉传感器

在压电式触觉传感器中，压电颗粒或薄膜在与接触力接触时发生形变，在上下两个电极上产生电荷进而在触觉元件上产生电势，对感应电压进行检测即可获得触觉压力信息。20世纪40年代和50年代钛酸钡和锆钛酸铅分别被发现能够表现出较好的介电和压电性能，目前压电陶瓷已经成为应用最广泛的压电材料之一。从20世纪60年代左右开始，人造压电材料的出现拓宽了应用领域，基于压电的器件被广泛用于声呐、水听器、麦克风、压电点火系统、加速度计和超声换能器等场景。同时，聚偏二氟乙烯（PVDF）聚合物强压电效应的发现增加了柔性压电材料的应用，包括在非易失性存储器、能量收集、微型机器人和触觉皮肤等中的应

用。压电式触觉传感器在小的动态接触变形下也可以具有较高的灵敏度和电压输出。如果保持持续压力，压电传感器的输出将会衰减为零。因此，压电式传感器最适合检测动态接触力，这同时也是此类原理传感器在触觉感知应用方面的限制。通过利用压电的原理，该传感器能够良好地模拟人体皮肤高达几百赫兹带宽的动态触觉感知，非常适合应用到人工皮肤的制造中。压电式触觉传感器的工作不需要电源，因此使用压电换能器的传感器在低功耗方面较为可靠和有效。压电式触觉传感器具有灵敏度高、响应迅速、信噪比高和难以检测静态压力等特点。利用硅微加工技术或简单的聚合物薄膜，大量的压电式触觉传感器被研究出来。

此外，压电材料往往会表现出一定的热电性能，当接触或环境温度变化时，压电式触觉传感器会产生输出电压。根据这一特性，同一个传感器能够测量多个接触参数，如力和温度。同时，热电产生的电压也会成为触觉传感器整体响应中的重要噪声来源，当压电效应和热电效应同时发生时，很难对响应进行解耦。如何进行温度记录并进行数学上的补偿是压电式触觉传感器需要考虑的问题。

（4）光学式触觉传感器

光学式触觉传感器是基于光波导原理的触觉传感器，光波导上分布着由硅橡胶制作而成的圆柱触头和圆锥触头，两种触头一一对应且中间连接一段橡胶垫，多对触头间的橡胶垫互相连接。在光波导的另一面放置新型光电敏感器件和图像采集器件。当没有外力作用时，光波根据全反射原理在波导介质通道中传播。当外界物体接触圆柱触头时，触头受力变形，与之对应的圆锥触头受到挤压而变形，由于圆锥触头与波导介质是接触的，波导中光波的传播状态被改变，光从波导介质中泄漏而出。

以典型光学触觉传感器Gelsight传感器为例，其通过摄像头和光源实现接触表面的精细成像，具有高空间分辨率和良好的形貌重建能力，适合检测纹理和滑移。但由于集成光学组件的体积较大，不利于在狭小空间中应用，且只有单个方向具有传感能力，无法满足精细化的操作要求。

（5）磁传导式触觉传感器

磁传导式触觉传感器主要分为两类，一类是机械运动导致磁通量发生变化的磁传导触觉传感器，另一类是磁材料受应变时磁阻发生变化的磁传导触觉传感器。磁传导式触觉传感器中有一种传感器由 8×8 阵列分布的探针组成，每个探针上绕有驱动线圈和传感线圈。探针随着被接触物体的轮廓移动，因此产生感应电动势，其电压由传感线圈测量。也就是说，传感器的输出与探针的位移偏差有关。然而，探针在斜坡处受斜向力和摩擦力作用，其运动会受到阻碍。

基于磁传导原理的触觉传感器具有的优点包括动态范围较宽、位移量较大、滞后低、响应呈线性等，能检测法向力、剪切力和扭矩，具备良好的灵敏度和较强的抗冲击能力，适用于复杂的接触状态感知。同时也存在一些缺点，如空间分辨率差、在斜坡上传导时会遇到机械问题、需要直流电源才能使用、易受漂移和噪声影响、受空间磁场和人体的磁场干扰等。

（四）电源的主要技术路径和发展趋势

机器人电池是机器人系统的核心能量来源，其性能直接影响机器人的续航时间、负载能力、动力学响应和安全性。随着移动机器人、服务机器人和人形机器人的普及，电池技术也经历了显著演进，主要呈现出高能量密度、小型化、智能化管理的趋势。

在电池类型方面，铅酸电池是最早期广泛应用于地面移动平台的机器人电池，原因是其成本低、技术成熟、放电稳定。其基本原理是利用二氧化铅（ PbO_2 ）作为正极、海绵状铅（ Pb ）作为负极、稀硫酸（ H_2SO_4 ）作为电解液。在充电和放电时，铅和铅氧化物会发生氧化还原反应，从而实现电能与化学能的相互转化：放电时， PbO_2 与 Pb 会转化为硫酸铅（ PbSO_4 ），电解液中 H_2SO_4 浓度降低；充电时， PbSO_4 重新转化为 Pb 和 PbO_2 ，电解液恢复高酸度。其缺点是能量密度低（约30~50Wh/kg）、体积大、寿命短，且深度放电会加速容量衰减，故而难以满足现代机器人对轻量化和长续航的需求。

随着锂电化学体系的发展，锂离子电池（Li-ion）逐渐成为主流。其工作原理基于锂离子在正负极之间进行嵌入和脱嵌的可逆过程：正极材料通常是锂金属氧化物（如 LiCoO_2 或 LiNiMnCoO_2 ），负极是石墨；放电时，负极石墨中的锂离子脱嵌，通过电解液迁移到正极，并在正极嵌入晶格结构。同时，电子沿外电路流动输出电能；充电时，电流将锂离子驱回负极嵌入石墨。锂离子电池量密度可达150~250Wh/kg，循环寿命通常在500~2000次之间，重量较轻，使移动平台和人形机器人具备了长时间工作和高速机动能力。人形机器人（如Atlas、Digit）和四足机器人（如Spot）几乎清一色采用锂电池模组，通过串并联电芯配置形成24~48V高功率电源。磷酸铁锂（ LiFePO_4 ）因热稳定性和高安全性也被部分服务机器人采用，尤其适合需要高安全冗余的场景。

聚合物锂电池（Li-Po）是另一重要分支，与锂电池的区别在于电解质由液态有机电解液改为凝胶状或固态聚合物电解质。工作原理同样依赖锂离子的嵌入和脱嵌过程，但因其柔性封装和较高放电倍率的特点，常用于无人机与轻型机器人。虽然能量密度略低（约120~180Wh/kg），但其高倍率放电特性支持大功率动作，如爆发性跳跃或瞬时高负载举升。

在高功率负载应用（如重载机械臂或应急救援机器人）中，超级电容器常与锂电池形成混合储能系统，通过电容器负责峰值电流输出和能量回收（再生制动），电池维持平稳放电，延长了整体系统的寿命。

现代机器人的电池系统不仅仅是单纯的“电源”，而是集成了电池管理系统（BMS）。BMS实时监控电芯电压、电流和温度，负责充放电均衡管理、短路保护、温控调节、寿命预测，并通过CAN或串口总线与主控板交互，提供荷电状态（SoC）和健康状态（SoH）信息。这对于安全性和精确任务规划至关重要。

从封装和结构上看，模块化电池组成为趋势：电芯被集成在密封的防水外壳中（IP54以上防护等级），内部采用金属隔热板和热敏电阻阵列，兼顾高强度碰撞防护与温度监控。在Boston Dynamics Spot平台，整块电池通过热插拔设计，可在几秒内完成更换，极大提高了作业连续性。

未来，机器人电池的研究重点包括更高能量密度化学体系（如固态锂电池、硅负极锂电池），以延长续航；更高功率密度与快速充电能力（例如20min内充至80%）；智能热管理，应对高功率放电带来的升温问题；电池与能量回收系统深度耦合，如机械臂的制动能量回收、步行机器人的下坡能回收；生命周期评估与梯次利用，在寿命末期进行梯级分配或储能再利用等方面。

总体上看，电池技术的演进已经从传统铅酸电池走向以锂电为核心、多模态储能并行、智能化管理的新阶段，为机器人具备长续航、高安全、高灵活性的性能奠定了重要基础。

（五）新材料的主要技术路径和发展趋势

机器人的外壳材料不仅仅是“包壳”，更承担着复杂多样的功能需求。它既是主要的结构承载件，又是环境隔离屏障，更是人机交互的直接界面。早期的工业机器人如ABB IRB系列和KUKA KR系列，其外壳几乎完全采用钢或铝合金，依靠高屈服强度和稳定的加工工艺提供刚性与耐用性。碳钢和不锈钢因密度较高、重量较大，在需要高刚度的大负载机械臂（如大型焊接机器人）上应用广泛，但在灵活性、能耗和轻量化方面存在局限。随着工业自动化的发展，铝合金成为中小型机械臂的重要材料，其典型应用如UR5协作机器人，即采用6061-T6铝材加工外壳，具备高比强度和良好的表面阳极氧化防护，显著降低了关节驱动负载。

工程塑料的兴起与服务机器人的普及密切相关。Pepper人形机器人、SoftBank的Whiz商用清洁机器人等的外壳普遍选用ABS、PC等工程塑料。这些材料本身密度低（约 $1.0\sim 1.3\text{g/cm}^3$ ）、韧性好，能有效应对碰撞冲击，保护内部精密电路和传感器。塑料外壳依托注塑工艺，可一次成型复杂曲面和人机交互界面，实现友好的

视觉观感与触觉体验。在防护性能方面，部分机器人通过加喷UV固化涂层或应用双色注塑，提高了耐老化性和识别性。

随着高端服务机器人和特种作业机器人的需求增长，复合材料外壳逐渐占据一席之地。碳纤维增强复合材料依靠其极高的比强度、比刚度与疲劳寿命，在外骨骼机器人和高负载机械臂中得到应用，如日本Cyberdyne的外骨骼就选用了碳纤维外壳以实现轻量化和高强度的平衡。其原理在于碳纤维在拉伸方向上的模量可达200GPa，通过树脂基体固化形成多层铺设的叠层板，具备优异的抗冲击和疲劳耐久性。玻璃纤维复合材料则以更低的成本在农业机器人、检测机器人中发挥重要作用，兼顾了结构性能和生产性价比。

在外壳成型技术方面，金属外壳主要通过数控铣削、冲压、焊接等工艺精加工而成，再利用喷涂或采用阳极氧化等表面处理工艺，形成耐腐蚀的保护层。工程塑料外壳依赖注塑成型和激光焊接，适配批量生产和多色复杂表面设计。复合材料多采用真空导入树脂、预浸料热压固化等工艺，可在模具内逐层铺设预浸料，实现多曲面一体化成型。以波士顿动力的Spot机器人为例，其部分外壳采用碳纤维与聚碳酸酯的组合，兼顾了结构强度和抗冲击韧性。

不同材料的性能差异源自其物理化学原理。金属依赖晶体排列提供塑性和强度，工程塑料以聚合物链的柔性抵御冲击，碳纤维增强复合材料通过增强纤维与基体界面的协同承载实现轻量化与高模量结合。在选型时，研发团队必须考虑机器人的使用环境（如是否长期在潮湿或盐雾环境下使用）、载荷工况（如高冲击、高频振动）及美观性和人机交互需求（如是否需要柔性表面或防滑纹理）。因此，现代机器人外壳材料通常不再是单一材料，而是通过多种材料“混搭”，形成复合结构层次，如近年来发布的宇树科技H1：其躯干与腿部骨架普遍使用CNC加工的铝合金或镁铝合金型材，而胸部、四肢等表面则选用PC+ABS等工程塑料注塑件，兼顾耐冲击性和复杂曲面造型，这种分体结构不仅简化了制造和装配流程，也降低了整机重量，使其重量仅约47kg，具备良好的移动性能和安全的人机交互特性。

更具前瞻性的是特斯拉Optimus的外壳设计，它充分借鉴汽车制造领域的集成化工艺思路，大量采用高强度复合材料一体成型的连续外壳结构，躯干、四肢和头部大面积复合壳体通过预浸料热压工艺制成，具备高比强度、优异的表面质量和减振性能，同时在表面采用抗指纹涂层和哑光喷涂处理以增强观感与耐用性。

Optimus的骨架内部为3D打印的复杂中空铝合金件，这样的设计有助于在保证刚度的同时最大限度降低重量并预留传感器、电缆及散热通道的安装空间，体现了人形机器人外壳朝着一体化集成、轻量高强和可规模化制造方向演进的最新趋势。

与此形成对比的是部分移动机器人或服务机器人仍然采用工程塑料包覆金属框架的“壳-骨分离”模式，以便在需要频繁维护和更换部件的场景中实现快速拆装。无论是宇树H1的工程塑料与铝合金混合外壳，还是特斯拉Optimus的大型复合材料面板，这些新一代外壳设计都不再局限于单一的防护功能，而是集成了传感器预留、散热孔、线缆导向、触觉交互等多种功能，在兼顾美观性的同时支持大规模生产。值得注意的是，Atlas、Optimus等代表性案例也表明高性能机器人对材料性能有极高的要求，如刚性、抗冲击性、耐热性及模量重量比等指标，通常需要碳纤维增强复合材料与高性能工程塑料相结合才能满足。

未来，随着材料工程和工艺创新的持续推进，机器人外壳还将进一步向智能化和多功能化发展。例如，在外壳内部集成可变硬度层、柔性压力传感矩阵，以及主动散热单元，实现结构-感知-交互的深度融合。从宇树科技H1到特斯拉Optimus，这些先进人形机器人的外壳案例不仅是新材料应用的缩影，也反映了在人机协作和量产制造的需求下，机器人外壳从“单一保护壳”向“多功能智能外衣”演变的清晰脉络。

2 产业篇

▷第四章 人形机器人的潜在应用场景

（一）应用场景分析

人形机器人有着与人类相似的感知方式、“肢体”结构和运动方式，是人工智能和机器人技术的创新高地，将在各类应用场景中发挥重要作用。人形机器人的主要潜在应用场景包括生产制造、社会服务、特种作业三大类。在生产制造领域[如农业采摘、汽车和3C（计算机类、通信类、消费类电子产品）领域制造业生产等]，可以有效解决人口老龄化问题，缓解未来的劳动力短缺问题。在社会服务领域，可以代替人类成为公共服务员、家政服务员、物流配送员、安保巡逻员等，为人类提供各方面的服务。在特种作业场景下，主要是在危险的（如深海、民爆、核电站等）作业环境中代替人类完成生产、巡检、探测、排爆等工作；在航空航天领域代替人类长期驻守空间站，执行航空器维护和空间科研任务，延长外太空的工作时间，甚至直接登陆地外行星完成探索任务。虽然当前人形机器人只在生产制造和社会服务的简单场景中进行试用，但随着技术的不断进步，人形机器人有望在复杂技能场景中辅助人类，在商业和家庭场景中服务人类。

总体来说，不同应用场景的共性需求包括人机交互能力、精准作业能力、自主决策能力、运动能力和安全伦理需求等，同时又各有侧重。生产制造领域的人形机器人主要对工业制造、农业生产等场景中的工人或农民进行替代，该领域中的大部分工作均为标准化流程，但部分场景对于细节的处理要求较高，对人形机器人的运动控制能力、手部精细化操作能力具有一定要求。由于在生产过程中，人形机器人需要和人类协同工作或接受人类下达的指令，因此需具备一定的人机交互能力。该场景下的客户相对集中，大部分属于ToB（面向企业）类型，预计市场规模在千亿级别。社会服务场景下的人形机器人主要提供综合家政、医疗、教育、商业、公共等服务，在这种非标准化场合下，其对人形机器人的交互能力、智能水平、灵活性、精准度等要求较高，同时还涉及诸多安全伦理问题。该场景下的客户相对分散，主要为To C（面向消费者）类型，并有部分To B类型，预计落地应用将随着技术进步逐步释放，远期市场潜力巨大，市场规模可达万亿级别。特种作业类任务大部分属于标准化场景流程，对人机交互能力的要求不是特别高，但可能会根据具体任务的不同，对人形机器人精准度有更高要求。同时，较为恶劣的环境对设备的环境适应能力也提出了较高要求。该类场景通常需求明确，供需对接相对容易，但整体市场规模相对较小，大概在百亿级别。

（二）生产制造应用场景

人形机器人凭借其类人的形态、优秀的环境适应性与灵巧操作能力，正深度融入现代生产制造的各个环节，成为突破传统自动化瓶颈、实现柔性化与智能化生产的关键赋能者。人形机器人在生产制造领域的应用场景广泛覆盖汽车等装备制造、电子信息制造、原材料生产加工、轻工纺织、能源电力通信，以及农林、牧渔、种植、养殖等国计民生相关的核心行业。在汽车等装备制造领域，人形机器人可深入高度自动化产线，高效承担精密装配、智能物流周转、优质工艺执行与全面质量检测，显著提升自动化水平与产品一致性；在追求微型化与精密化的电子信息制造中，人形机器人可以精准完成高精度点胶、微米级元件插接组装、复杂功能测试及自动化包装等工作，攻克精密手工操作瓶颈，保障高良率；在环境复杂、存在安全风险的原材料生产加工环节，人形机器人可以替代工人进行高危物料搬运、环境监测取样、设备巡检及特定工艺操作等工作，提升生产安全性与作业稳定性；在柔性化需求突出的轻工纺织产线，人形机器人可以高效完成衣片转运、复杂缝制、精准质检与定制化生产支持等工作，应对多品种小批量柔性化生产挑战；在保障能源、电力、通信行业基础设施稳定运行的过程中，人形机器人可以承担设备状态检测、设施环境巡查、特定操作及基础维护任务，提升巡检覆盖率和运维安全性；在规模化种植养殖场景下，人形机器人可以执行自动化采收分拣、精准播种施肥、定时定量投喂及环境监控调节等工作，提升行业生产效率与资源利用水平。人形机器人在生产制造各场景中的核心价值在于其以类人的灵巧性补充了传统自动化的不足，且能承担复杂、精细、非结构化或高危环境下的关键任务，有效提升了产线柔性、作业效率、产品质量与安全性，展现出推动制造业转型升级的巨大潜力。以下介绍了人形机器人在具体领域中的工作情况。

1. 汽车等装备制造

在高度自动化及结构化的生产线上，人形机器人能够深度融入制造流程，承担物料分拣运输、产线上下料、工装定位、紧固组装、柔性插接、外观与基础功能检测、设备检修维护、外壳面板安装、焊接和检查、涂装喷漆、产品质检等关键任务及长尾任务，有效提升了生产效率、质量一致性、生产安全性，并缓解了高强度、重复劳动带来的人员压力。

首先，人形机器人能够高效执行物流周转与基础装配任务。在制造流程的起点，人形机器人可深入原料仓库或大型料场，基于先进视觉识别系统，精准抓取、分拣各类形状、尺寸和材质的原材料、标准件或定制化零部件。同时能在不同工位、装配线或缓存区之间实现灵活自主导航，将分拣好的物料或半成品准时、准确地配送到指定工位，并高效地将毛坯或半成品放置到加工设备上，以及将加工/装配完成的部件或产品从设备上取下并转运至下一工序或成品区。

在装配环节，人形机器人可利用其灵巧手和视觉系统，进行精密定位与夹持固定，执行如使用拧紧枪拧紧螺栓、铆接铆钉等标准化紧固作业，并完成如对汽车线束、精密电子连接器等需要精细对位与自适应插拔力的柔性插接任务，确保装配精度与可靠性。此外，人形机器人能承担外壳、面板、内饰件等大型或易损部件的稳定抓取、精准定位与安装工作，确保装配精度与效率。并且，人形机器人能够承担需要精细操作或适应性的工艺与质量保障任务。

在车身制造阶段，人形机器人可操作焊枪，在复杂曲面或狭小空间完成高精度点焊、弧焊作业，并利用搭载的视觉传感系统实时检查焊缝质量，确保结构强度。

在涂装环节，人形机器人能稳定操控喷枪，并根据预设路径或实时视觉引导，在车身内外表面进行均匀、一致的底漆、色漆或清漆喷涂，实现高品质的外观与防腐保护。

在总装线末端及质检工位，搭载高分辨率相机、激光扫描仪及各类传感器的人形机器人可系统性地对整车或关键部件进行外观缺陷检查（如划痕、凹坑、污渍、装配缝隙等）、基础功能测试（如开关按键、灯光、车门开合等简单机械动作），以及关键尺寸测量（如孔位、间隙面差等外观数据测量与基础功能检测），确保产品的出厂品质。

人形机器人能够辅助进行产线维护与设备检修工作。在设备日常运行期间，人形机器人可执行周期性巡检任务，监控关键设备的运行状态，识别电机振动、工作温度、液压压力、声响等参数情况，并进行故障预警。当设备需要计划性维护或发生简单故障时，人形机器人可在技术人员的远程指导下或根据预设程序，执行诸如更换过滤器、磨损刀具和钻头等易损件，清洁传感器镜头、导轨等关键部位，加注润滑油脂，检查并紧固松动的线缆接头，复位跳闸的开关或保险等基础维护操作，有效缩短了非计划停机时间，提升了设备的综合效率。

目前，特斯拉、优必选、小米、小鹏、智元等人形机器人厂商均已将目光聚集汽车工业生产领域，人形机器人将有望在汽车制造领域率先批量使用。当前，人形机器人通过与传统自动化设备的协同，在汽车生产上，可执行装配底盘、打螺丝等任务，提升工厂的智能化水平，实现复杂工业场景的无人化生产。如特斯拉宣称其人形机器人Optimus将率先应用于汽车制造领域，包括但不限于特斯拉的超级工厂内部，协助或代替人类完成重复性、危险性高的工作，且未来有望全面接管特斯拉汽车的生产。优必选推出的工业版人形机器人WalkerS系列，也将率先适用于汽车领域工业场景，未来还将逐步拓展至汽车零部件、3C、智慧物流等其他智能制造领域及应用场景。小鹏则积极推进人形机器人在自有制造系统中的分阶段落地。智元的远征A1将首先面向工业场景，规划在比亚迪工厂参与外观检测、底盘装配等汽车装配线上的作业。

2. 电子信息制造

在涉及精密元件组装、微焊接、高精度点胶、复杂功能测试，以及混线生产的电子信息制造领域生产场景中，人形机器人能够融入表面贴装工序、模组组装、整机测试与包装等关键环节，承担点胶涂胶、搬运分拣、外观与基础功能检测、焊接和检查、柔性插接、零部件组装、产品包装、产品质检等核心及长尾任务，显著提升生产良率、效率及一致性，并有效应对日益复杂的产品变种与快速换线需求。

首先，人形机器人能够承担精密物料处理与准备工作。在元器件库、线边缓存区及不同高精度工位之间，人形机器人可基于先进视觉引导系统，精准识别、抓取并分拣微小的芯片、精密连接器、柔性电路板、镜头模组等零部件，完成成品及半成品的搬运分拣工作。它们能在高度洁净或静电防护要求严格的环境中灵活移动，将物料准时、准确配送至点胶、贴装、焊接或组装工位。对于芯片、镜头、密封圈等需要精密涂覆固定的环节，人形机器人能稳定操控点胶阀，并根据预设的复杂三维路径或实时视觉反馈，在微小区域实现高精度、一致性的胶量控制与轨迹涂覆，确保黏接强度与可靠性。

其次，人形机器人被预设在未来可以胜任精密连接、组装与核心工艺执行等工作。在需要高可靠电气连接的环节，人形机器人可利用其灵巧手和显微视觉辅助技术，完成板对板、线对板、排线等精密位置的精准对位和柔性插接，并实现插拔力的实时感知及精密控制，避免损伤端子，提高良品率。在特定点位补焊、异形件焊接、返修等微焊接工位，人形机器人可操作手动微型焊接装置，在狭小空间区域或高密度印刷电路板上完成精准的焊接任务，并通过搭载的视觉传感系统对焊点质量进行检查。此外，人形机器人能够执行精密的零部件组装，例如将摄像头模组、传感器、听筒、扬声器、电池、屏幕等复杂组件，在狭小的手机、平板电脑或可穿戴设备机身内部进行精确对位、卡扣安装或螺丝紧固，确保组件的结构紧凑与功能正常。

同时，人形机器人可以完成制造全流程的质检工作。在组装过程中及最终产品下线环节，人形机器人通过配合高清显微相机、高精度传感器和专用测试夹具，可系统性地执行多维度产品质检，包括对外壳、屏幕、镜头的划痕、凹痕、脏污、装配缝隙、印刷缺陷等的外观检测，对按键、接口、屏幕显示器、指示灯、扬声器的基础功能检测，还能配合自动化测试设备，进行更深入的性能验证等。在确认产品合格后，人形机器人能够完成产品的柔性包装任务，包括将产品精准放入内衬、折叠包装盒、粘贴标签、封装防静电袋，以及堆垛装箱，确保包装高效与外观规整。

此外，人形机器人能够辅助进行产线适应性调整与基础维护。在应对多品种、小批量生产时，人形机器人可协助进行快速的工装夹具更换、送料器调整等辅助换线工作。在设备运行期间，人形机器人可执行清洁光学传感器和工作台面等简单清洁任务或基础点检工作，保障生产环境的稳定。

3. 原材料生产加工

在环境复杂、存在安全风险或重复劳动强度高的原材料生产加工场景中，人形机器人能够承担物料搬运分拣、产线与设备巡检、来料和产品基础检测、产线上下料、环境监测和取样、特种工艺操作、有害物质处理等具体任务，助力提升作业安全性和效率稳定性。

首先，人形机器人能够承担繁重的物料流转与基础操作任务。在原料堆场、仓库或产线起始端，人形机器人可基于视觉或传感器识别，执行物料搬运和分拣工作，抓取或移动矿石、矿粉、煤块、化工原料袋、板材坯料等原材料或半成品，并按类型或批次进行初步分拣。它们能在粉尘、高温或不平整的环境中移动，将物料搬运至指定位置，如配料仓、破碎机进料口或缓存区。在加工环节，人形机器人能完成产线上下料任务：将待加工的金属、板材、大型工件等原料准确放置到机床、熔炼炉、轧机、反应釜等加工设备上，并在加工结束后将铸件、轧制件、化工中间体等成品及半成品产物取下并转运至下一工序或冷却区/检验区。

其次，人形机器人能够承担环境监测、基础检测与特定工艺辅助任务。在需要实时掌握环境状况的区域，如化工厂房、矿井、高温车间等，人形机器人可通过携带传感器进行环境监测和取样，如定期巡检并记录温度、湿度、特定气体浓度、粉尘浓度等参数，或在指定位置采集空气、液体或固体样本。在来料检验或产品初步检查环节，人形机器人可利用视觉系统进行来料和产品基础检测，检查原材料的表面外观质量、核对物料标识，或对产出的标准尺寸板材等初级产品进行基础尺寸测量、表面裂纹或明显缺陷的识别。对于某些重复性、有安全风险的特种工艺操作，人形机器人可在工人的远程监督下或按设定的程序执行。例如，在安全距离外进行熔融金属的测温取样、向反应釜中添加特定比例的助剂、操作特定阀门进行流程切换或在特定工位喷涂隔离剂等。

同时，人形机器人能够承担设备巡检与特定有害物质处理任务。在设备密集的生产区域，人形机器人可按照设定路线进行产线与设备巡检：利用红外热像仪、摄像头等传感器检查如泵、电机、管道、阀门、传动装置等生产关键设备的运行状态，监测是否有异常振动、过热、异响、泄漏等情况，并记录数据供维护人员分析。在涉及危险化学品、放射性物质或高腐蚀性物料的环境下，人形机器人能够承担有害物质处理任务，即在人员不宜直接接触的情况下，对泄漏物进行初步围

堵，吸附、转移破损的危险品容器，以及对污染区域进行初步去污清洁，降低人员暴露风险。

4. 轻工纺织

在多品种、小批量、柔性化生产需求日益增长的轻工纺织行业（如服装、家纺、鞋帽、皮革制品等）的制造场景中，人形机器人能够承担物料搬运分拣、产线与设备巡检、原材料和成品质检、产线上下料、产品包装、产品质检、复杂工艺处理、定制生产支持等具体任务，助力提升生产效率和质量稳定性，实现柔性化生产。

首先，人形机器人能够承担物料流转与基础生产操作任务。在原料仓库或裁床区域，人形机器人可基于视觉系统，抓取并移动布卷、纱锭、皮革卷、辅料箱等，并将物料按颜色、材质或批次进行分拣整理。在缝制、成型等主要工序环节，人形机器人能完成产线上下料任务，将裁剪好的衣片、面料、皮革部件或半成品准确放置到缝纫机、模板机、热压机、定型机等设备的工作区域，并在加工完成后将半成品或成品取下，将其转运至下一工序或质检区/缓存区。对于完成最终检验的产品，人形机器人可执行产品包装任务，将服装折叠整齐后装入包装袋或礼盒中，并放置吊牌和防伪标签，还能进行装箱、封箱或堆码工作。

其次，人形机器人能够承担质量检验与特定的复杂工艺任务。在原材料入库和生产过程中，人形机器人可利用高分辨率相机和视觉传感器进行原材料和成品质检，检查面料或皮革表面是否存在破洞、污渍、色差、纹理不均等瑕疵点，测量关键尺寸是否达标，核对纽扣、拉链等辅料的颜色、型号是否正确。在成品最终环节，人形机器人可以进行更全面的产品质检，如检查缝线是否均匀牢固、有无跳针或开线，拉链是否顺滑，纽扣及装饰件是否牢固，成品对称性和整体外观平整度，并进行简单的拉链开合、魔术贴黏合等基本功能检查。对于需要较高技巧或一致性的复杂工艺处理任务，人形机器人可在程序设定或监督下进行相应操作。例如，通过灵巧手操作特种缝纫设备进行绣花、绗缝等复杂线迹缝制，完成精准的皮革冲孔或雕刻，操作喷枪进行局部染色，完成精细的手工钉扣或特殊装饰件安装等。

同时，人形机器人能够支持柔性生产与设备监控任务。在追求个性化和小批量订单的定制生产中，人形机器人能够适应频繁更换的款式和工艺要求。它们可以根据预设的不同程序，快速调整抓取和操作方式，处理不同尺寸、形状和材质的部件，协助工人完成定制化产品的组装、特殊工艺的应用等工作。在设备密集的生产车间，人形机器人还可按照设定路线进行产线与设备巡检，利用传感器监控缝纫机、裁剪机、定型机等关键设备的运行状态，并记录数据供维护人员参考。人

形机器人也能执行一些基础的清洁任务，如清理工作台面或设备周围的线头、碎布等。

5. 能源电力通信

人形机器人在电力、油气等能源及通信基础设施运维领域也展现出了应用潜力。在环境复杂、存在安全风险或需要高效巡检维护的场景中，人形机器人能够承担设备检测、线路/场站/机房巡检、物料搬运堆码、环境监测和取样、设备操作和维修维护等具体任务，助力提升运维的安全性、效率与可靠性，并减少工作人员在危险或艰苦环境中的暴露情况。

首先，人形机器人能够承担广泛的设备检测与巡检任务。在变电站、发电厂、油气泵站、通信基站、数据中心机房等关键设施内部及周边，人形机器人可执行设备检测和线路/场站/机房巡检工作，通过利用自身搭载的传感器，并配合热像仪、局部放电检测仪、气体传感器等设备，对变压器、开关柜、断路器、发电机、风机叶片、光伏面板、传输管线、阀门、通信设备、服务器机柜等进行检查，完成设备表面温度异常检测、热故障点识别、设备异常声响监听、仪表读数检查、设备外观损伤识别、设备标识状态核对等工作。此外，人形机器人能在高空、狭窄通道、高温、高湿、含有毒气体或强电磁干扰等人员不便长期停留或存在安全风险的环境中进行工作，并将检测数据和图像进行实时回传。

其次，人形机器人能够承担环境监测、物料处理及特定设备的操作任务。在如厂区周界、管道沿线、井下、受限空间等需要掌握环境状况的特殊区域，人形机器人可进行环境监测和取样，完成对环境温度、湿度、特定气体浓度、粉尘浓度、辐射水平等的检测工作，并在指定位置采集空气、水体或土壤样本供工作人员监测分析。在仓库、备品备件库或作业现场，人形机器人可执行物料搬运堆码，识别、抓取、移动和整齐堆叠电缆盘、绝缘子、金具、工具包、小型备件箱等物料的任务。对于某些重复性或需要远程操作设备的任务，人形机器人可在操作员的指令下执行任务。例如，在安全距离外操作阀门启闭、进行开关分合闸、启停特定非核心设备或进行设备参数设置等操作。

同时，人形机器人能够承担基础的维修维护与辅助工作。在设备需要维护或发生简单故障时，人形机器人可在技术人员的远程指导下或按预设程序，执行维修维护任务，如进行设备表面清洁除尘，更换指示灯、保险丝等简单的易损件，紧固松动的接线端子或螺栓，对活动部件加注润滑油脂，简单的线缆整理，粘贴标识标签等工作。在应急或特定情况下，人形机器人可辅助进行泄漏物的初步围堵吸附和小范围的除冰除雪等作业。

6. 种植养殖

人形机器人在农业种植、林业管理、畜牧养殖和水产渔业等领域具有极大的应用前景。在劳动力短缺、作业环境复杂或需要精细管理的场景中，人形机器人能够承担采摘采收、农产品拣选、播种施肥、饲料投喂、清洁消毒、环境监测调节等具体任务，助力提升生产效率、降低劳动强度、优化资源利用并支持精细化种养管理。

首先，人形机器人能够承担收获与初步处理环节的任务。在果园、菜地或温室中，人形机器人可基于视觉识别系统，执行采摘采收工作，如识别果蔬的成熟度、大小和位置，通过灵巧手轻柔抓取、剪断果柄或小心摘下，并将采摘的果蔬放入收集容器中，减少碰伤。在果蔬处理点或包装车间，人形机器人还能进行农产品拣选，根据预设标准，如大小、颜色、形状、表面缺陷等，对采摘的农产品进行自动分拣、分级，剔除次品，并将合格品进行分类放置或初步包装。

其次，人形机器人能够承担种植、养殖过程中的部分生产作业与投喂任务。在播种或移栽阶段，人形机器人可执行播种施肥任务，即在规划好的田地或苗床区域，精准地挖穴、点播种子或放置秧苗，并在指定位置施加定量肥料或营养液。在规模化的养殖场，如猪舍、牛栏、鸡舍、鱼塘等，人形机器人也可承担饲料投喂工作，按照设定的时间和区域，自动行走至投喂点，精准称量并投放不同种类和配比的饲料，确保饲料的定时定量供给。

同时，人形机器人还能承担环境维护、监测与调节任务。在种植大棚、养殖舍或鱼塘周边，人形机器人可对环境进行清洁消毒，还能清扫地面落叶、粪便等杂物。例如，使用喷淋系统对特定区域进行消毒液喷洒，或在人类指导下进行简单的环境清扫。更重要的是，人形机器人可搭载多种传感器进行环境监测调节，如对种植养殖区域进行持续监测并记录空气或水体温度、湿度、光照强度、二氧化碳浓度、特殊气体浓度、溶解氧/pH值等关键环境参数。基于预设规则或接收到的指令，人形机器人可辅助进行简单的环境调节操作。例如，开启或关闭通风口、补光灯和触发报警提示异常等，这也为农场管理员提供了重要的决策依据。

此外，人形机器人可在特定场景下承担辅助性工作。例如，在果园或林地，人形机器人可辅助进行小范围的除草任务；在养殖场，人形机器人可通过摄像头识别动物的异常活动，协助饲养员进行简单的动物行为观察。

（三）社会服务应用场景

人形机器人在社会服务领域的应用场景主要覆盖商业服务、物流配送、医疗健康、公共服务和保障、家政管理、科教培训、检测检验等关键领域。在商业服务中，承担迎宾导购、点单配餐、互动展示及简单的零售辅助工作，提升服务的科

科技感与效率，优化顾客体验；在物流配送环节，执行非标准件分拣搬运、末端精准送达及辅助库内操作等任务，补充传统自动化短板，提升物流场景的流转效率；在医疗健康领域，辅助基础导诊咨询、院内物品与患者运送、健康监测与应急预案任务，分担医护人员负荷，提升服务体验与安全性；在公共服务和保障方面，提供场所咨询引导，执行环境巡查预警、基础区域清洁及设施状态上报任务，提升信息可达性并辅助维护公共环境；在家政管理中，承担简单家务、日常陪伴、健康提醒与应急辅助等工作，有效分担家庭劳动并增强基础关怀；在科教培训领域，作为直观教学工具与算法实践载体，辅助实验操作、技能训练及社交情感培养，提升学生学习的互动性与实践性；在检测检验领域，执行样品物料转运、设备仪表操作及标准化的简单样品制备与实验操作等任务，提升流程效率与一致性，减少人为误差。

人形机器人在社会服务各场景中的核心价值在于，其作为人力的补充与延伸，承担重复性、程序化或需要基础交互的任务，有效提升服务效率与可达性、优化用户体验、分担人力工作负荷，展现出广阔的应用前景。

1. 商业服务

人形机器人凭借其类人形态、强大的人机交互能力和一定的环境适应性，在商业服务领域展现出了巨大的应用潜力和庞大的市场前景。在需要吸引客流、提升服务效率或提供新颖体验的场景中，人形机器人能够承担迎宾导购、演出展示、商超零售、点单配餐、游戏竞技、厨艺烹饪、美容美发等具体任务，助力优化顾客体验，分担重复性服务工作并创造互动价值。

首先，人形机器人能够承担接待引导与基础服务任务。在商场、营业厅、酒店、展馆入口或咨询台区域，人形机器人可执行迎宾导购工作，通过语音、屏幕或简单动作欢迎顾客，提供基本的场所导览、功能区指引、活动信息播报等服务，并引导顾客至目标区域或服务柜台。在超市、便利店、药店等场所，人形机器人可以辅助售货员提供商超零售服务，完成引导顾客寻找商品、回答简单库存或价格查询、在电子货柜协助顾客完成扫码选购等任务。在餐厅、咖啡厅或自动售卖亭，人形机器人可以承担点单配餐任务，通过其触摸屏或语音交互接收顾客点单，完成简单的饮品配制，也可以从备餐区取出相应的食品或饮品，并送至顾客指定位置。

其次，人形机器人能够承担展示、互动与特定操作等任务。在剧场、展会、影视拍摄现场或体育赛事等场合，人形机器人可进行演出展示，如按照预设程序表演舞蹈、动作或与人进行简单的舞台互动，吸引观众注意或烘托气氛。再如在科技展中，人形机器人展示了自己多方面的技术能力。在游戏厅、主题公园或特定活动区，人形机器人可参与游戏竞技，与顾客进行如乒乓球对打、简单棋类对弈、

预设动作模式的趣味格斗等互动游戏，提供娱乐体验。在餐厅后厨或开放式烹饪区，人形机器人可执行标准化程度高的厨艺烹饪任务，如在严格程序控制下完成翻转汉堡肉饼、定时煮面、制作特定配比的沙拉或进行简单的煎炸操作等。在美容美发店，人形机器人可以作为探索性应用，比如在技师的监督下执行递送工具、基础的头皮按摩、简单吹风造型等美容美发辅助操作。同时，人形机器人能够在商业服务环境中提供信息查询与简单环境维护。人形机器人通常内置知识库，可以回答顾客关于营业时间、服务项目、产品特点等常见问题。部分人形机器人还能进行基础的环境监测，或在闭店后执行简单的地面吸尘、桌椅擦拭等清洁任务。此外，人形机器人可作为移动广告或信息发布平台，其屏幕可轮播促销信息、活动海报或品牌宣传内容。

2. 物流配送

人形机器人已在物流配送领域实现试点应用。在需要灵活处理不规则物品、适应复杂环境或进行人机协作的场景中，人形机器人能够承担分类分拣、物资搬运、末端配送、仓储作业等具体任务，助力提升操作效率、分担重复体力劳动。

首先，人形机器人能够承担物品分拣与库内搬运任务。在仓库的分拣区或拆零作业区，人形机器人可执行分类分拣工作，基于其视觉识别系统，识别包裹或货物的形状、大小、条码或标识，将其从混合货筐或传送带上抓取出来，并放入对应的目标分拣筐、包裹袋或指定流向的容器中。在仓库、分拨中心或装卸区内部，人形机器人能进行物资搬运，抓取和移动各种形状或非标准尺寸的包裹、箱体等，将其在不同工作台、缓存区或装卸点之间转运。在库存储存环节，人形机器人可辅助进行仓储作业，在工作人员指导下或按指令，将货物放置到指定货架位置，还能从货架上取下特定货物。

其次，人形机器人能够承担末端配送与特定环境下的运送任务。在相对封闭或结构化的“最后一公里”环境中，如大型园区、办公楼宇、医院内部、校园或工厂厂区，人形机器人可执行末端配送，即接收配送指令，规划路径导航至目标地点，通过扫码、取件码等身份验证，安全交付包裹、文件或小型货物给收件人。在室内或短距离室外环境中，人形机器人也能承担餐食、药品、小型工具等物品的内部定点运送。同时，人形机器人能够执行辅助性库内操作。例如，协助工作人员进行简单的货物信息核对、在盘点时辅助进行货架低层货物的清点或在特定区域进行基础的清洁整理。此外，人形机器人可作为移动的信息节点，在执行配送或搬运任务时，其位置信息、任务状态可实时上传至物流管理系统，便于追踪和管理。

在智慧物流领域，1X Technologies、Agility Robotics和中国电子科技集团公司第二十一研究所等厂商均在尝试利用人形机器人的优势来解决传统物流难题，推

动行业向更加智能化和自动化的方向发展。

3. 医疗健康

在需要辅助医护人员、提升服务效率或提供基础支持的医疗健康场景中，人形机器人展现出了探索性应用潜力。人形机器人能够承担咨询导诊，护理照看，康复训练，健康监测，送药送餐、患者推送、卫生消毒等院区服务保障，以及手术辅助等具体任务，旨在辅助医护人员工作、优化部分流程并提升患者体验。

首先，人形机器人能够承担信息咨询与基础院区服务任务。在医院门诊大厅、科室分诊台等区域，人形机器人可执行咨询导诊工作，通过语音交互或触摸屏，为患者及家属提供医院科室分布、医生出诊信息、就诊流程引导、常见问题解答等基础信息咨询服务，并指引人们前往目标诊室或区域。在病房、药房、食堂等区域，人形机器人可承担一系列院区服务保障工作。例如，按照指令或预设路径，将药品、餐食从药房或食堂运送至指定病房或护士站，完成送药送餐；在医护人员陪同或特定安全环境下，它们可以辅助推送轮椅或病床上的患者进行短距离转运，完成患者推送；在特定区域，它们可以执行基础的物体表面擦拭消毒或紫外线灯移动照射，完成卫生消毒等。

其次，人形机器人能够承担患者陪伴、健康监测与辅助训练任务。在病房或康复中心，人形机器人可提供基础的护理照看辅助，如在护士的远程监控下，定时提醒患者服药、进行简单的语音互动陪伴、播放舒缓音乐等。在康复治疗师的指导和监督下，人形机器人还可以辅助患者进行康复训练，引导患者跟随其进行预设的、标准化的肢体或关节动作练习，或提供稳定的支撑点供患者抓握进行平衡训练，并记录患者的动作完成情况和次数。此外，人形机器人可搭载传感器进行健康监测，在患者许可和医护人员的设定下，配合仪器非接触式测量患者的体温、心率、血氧饱和度，或提醒患者佩戴和正确使用可穿戴健康监测设备，并将数据汇总显示或传输给医护人员。随着技术的不断发展，人形机器人有望在精确地控制和严格的流程管理下，在手术室承担辅助性工作。例如，在严格的无菌规范和控制下，人形机器人可在主刀医生指令下，执行手术辅助任务，包括在器械台上准确传递指定手术器械给护士或医生、调整手术灯光角度或通过其搭载的显示屏辅助显示医学影像。同时，人形机器人还能进行院区环境状态巡查。例如，定时在走廊等区域巡逻，监测温度、湿度等环境参数是否在设定范围，或检测跌倒、呼救等异常声响，并及时定位与报警。

4. 公共服务和保障

在需要提升服务效率、分担重复性工作或增强公众体验的公共服务和保障场景中，人形机器人能够承担咨询引导、窗口服务、安保巡逻、环卫清洁、植被养

护、公共设施巡查维护等具体任务，旨在辅助工作人员、优化服务流程并保障公共环境。

首先，人形机器人能够承担信息咨询与基础窗口服务任务。在政务大厅、图书馆、博物馆、机场、火车站等的咨询台或服务点，人形机器人可执行咨询引导工作，通过语音或触摸屏交互，为公众提供场所功能分区指引、办事流程说明、常见问题解答、信息查询等基础咨询服务，并引导人们前往对应窗口或区域。在部分标准化程度高的服务窗口，人形机器人可以辅助进行窗口服务，接收用户指令或证件，完成信息查询、打印预置表单、发放排队号码等基础操作。

其次，人形机器人能够承担环境巡查、基础清洁与辅助养护任务。在公园、广场、街道、场馆等公共区域，人形机器人可执行安保巡逻工作，按照预设路线进行移动巡查，利用摄像头监测环境，识别异常情况，并将信息实时上报至控制中心，起到辅助预警作用。在室内公共空间或小型户外区域，人形机器人可承担环卫清洁任务，进行地面清扫、吸尘，或对长椅、扶手等进行基础擦拭。在公园绿地或苗圃，人形机器人可在园艺师的指导下进行基础的植被养护辅助工作，如定点定量浇灌小型花坛或盆栽、辅助清理修剪后的枝叶等。在社区、公园或道路沿线，机器人可执行公共设施巡查维护任务，包括检查公共座椅、健身器材、路灯、垃圾桶、标识牌等设施的外观完好性，检测路灯是否正常，并将异常情况上报，在工作人员进行维护时协助递送小型工具或耗材，进行设施表面的简单清洁等。此外，人形机器人可提供信息发布与互动服务，其屏幕可滚动播放公共信息、政策宣传、安全提示或活动公告。在科普场所，人形机器人可进行简单的互动问答。

在过去的几年里，非人形的机器人已经被用于安保任务中，例如美国Knightscope公司开发的蛋形机器人K5和瑞士Ascento Robotics公司创建的两轮机器人，实现了在铁路车库的巡逻。从供给端来看，1X Technologies的EVE人形机器人目前已成功应用于巡逻安保场景，区别于其他保安类机器人，EVE人形机器人具有头、脸和两只能够自主移动的手臂。随着技术的不断进步和成本的降低，未来，人形机器人在安保领域的应用将会越来越广泛。此外，优必选、银河通用等公司研发的人形机器人已经在公共博物馆、药店、商超进行导览咨询及商超零售工作，试点效果良好，有望进一步在公共服务领域拓展人形机器人的应用范围和数量。

5. 家政服务

在需要提供日常陪伴、辅助家务劳动或进行基础健康关注的家政服务场景中，人形机器人能够承担陪伴互动、健康监测、应急预警、紧急救护等工作，还能完成洗衣叠衣、浇花插花、拖地取物等简单家务，以及做饭洗碗、整理收纳、宠物照

料等复杂家务，旨在辅助家庭成员分担部分重复性家务劳动，并提供基础的安全关怀。

首先，人形机器人能够承担日常陪伴与简单家务任务。在家庭环境中，人形机器人可提供基础的陪伴互动，通过语音对话、播放音乐、简单的游戏或跟随预设程序进行动作表演，为家庭成员提供娱乐和情感交流支持。在家务方面，人形机器人通过学习可完成多种简单家务，包括洗衣叠衣、浇花插花、拖地吸尘、取物递送、表面清洁等。

其次，人形机器人能够承担健康关注与应急辅助任务。在家庭成员日常活动中，人形机器人可进行简单的健康监测，如定时提醒服药、识别跌倒等异常姿态，还能连接家用健康设备，读取并播报基础数据，及时通知用户。在监测到潜在风险，如跌倒、烟雾报警器报警、异常声响时，人形机器人可发出应急预警，并主动拨打预设的紧急联系人电话或发送警报信息。在危及人身安全的紧急情况下，如发现老人跌倒，人形机器人可尝试进行基础的紧急救护辅助，如移动到伤者附近保持通信、提供安慰或指导、运送急救包或尝试清除通道小障碍物，为专业救援争取时间。同时，人形机器人未来有望承担部分程序化的复杂家务辅助任务。在严格的程序控制和安全监督下，机器人可执行做饭洗碗辅助、非标准化食材处理、特定厨电操作、泛化整理收纳辅助、宠物照料辅助等任务，但这些复杂家务辅助任务对人形机器人的感知、决策和操作精度要求极高，目前多为探索性应用。此外，人形机器人可承担环境状态提醒工作，如监测室内温、湿度等环境参数，并在参数超出设定范围时提醒用户。

当前，人形机器人在家政管理领域的应用仍处于起步阶段，主要集中在提供基本陪伴、娱乐互动等简单家务服务方面。例如，日本软银公司推出的Pepper机器人和美国的Jibo机器人能够进行语音交流、表情识别、情感分析等，为用户提供陪伴等家庭服务。挪威的IX Technologies公司推出的NEO人形机器人专为家庭使用而设计，在其宣传视频中，NEO人形机器人可以实现辅助清洁、递送物品、操作家电等功能。

6. 科教培训

人形机器人凭借其类人形态、可编程交互能力和物理操作能力，在科学教育、技术培训领域展现出了应用潜力。在需要直观演示、动手实践或模拟真实交互的场景中，人形机器人能够成为课程教学工具、算法开发载体，并承担知识教育辅助、技能教育辅助、简单数据分析、仪器设备操作、社交情感训练、复杂实验操作等具体任务，旨在辅助教学人员、激发学生的学习兴趣并提供实践平台。

首先，人形机器人能够承担教学演示与知识教育辅助任务。在课堂上或实验室里，机器人可作为课程教学工具，通过预设程序或教师控制，演示物理原理、生物动作或进行简单的科学实验步骤展示。在编程、人工智能等课程中，人形机器人本身可作为算法开发载体，学生可在其平台上编写算法，测试和观察控制人形机器人行为的算法效果，直观理解代码与实际动作之间的联系并进行算法开发验证。人形机器人还能进行知识教育辅助，可以通过问答互动、播放教学视频或展示三维模型，辅助讲解特定知识点。

其次，人形机器人能够承担技能教育辅助与实践操作任务。在工程、自动化或实验技能培训中，人形机器人可辅助进行技能教育。比如指导学生完成组装、调试工业模块等，或模拟工业场景下的操作流程。在实验室环境下，人形机器人可执行仪器设备操作，即在教师或研究人员指令下，标准化操作实验仪器，完成开关电源、调节旋钮、放置样品到显微镜载物台、按下光谱仪按钮等系列动作，减少人为操作误差。对于需要精确重复或复杂流程的实验操作，人形机器人可在严格的程序执行和监督下，通过灵巧手执行预设步骤序列，如混合特定比例试剂或进行滴定试验，作为演示或研究辅助工具。在执行操作过程中，人形机器人还可同步进行简单的数据采集和分析，并绘制基础图表供教学分析。同时，人形机器人能够承担社交与情感训练辅助任务。在特殊教育或心理学领域，人形机器人凭借其类人的外观及强大的语音交互能力，可作为社交情感训练的辅助工具，通过预设的互动程序（如识别表情、进行简单对话、模拟社交场景等），在治疗师或教师的引导下，帮助有需要的学生练习眼神交流、情绪识别、基础对话技巧等。

7. 检测检验

在需要处理重复性流程、适应特定环境或辅助技术人员操作的实验室等检测检验场景中，人形机器人能够承担物品转移、仪表设备巡查、样品制备、设备操作等具体任务，以提升操作效率、减少人为误差，并辅助标准化流程执行。

首先，人形机器人能够承担物品转移与仪表设备巡查任务。在实验室或检测车间内，人形机器人可执行样品物料搬运工作，根据指令或预设路径，抓取和移动样品容器、标准试剂瓶或小型检测工具包，在不同工作台、仪器区或存储区之间进行安全转运。在设备密集区域，人形机器人可进行仪表设备巡查，按照预设路线巡视关键设备，利用摄像头读取仪表盘示数，检查设备的运行指示灯状态，识别异常声响或泄漏迹象，并将数据记录下来或上报。

其次，人形机器人能够承担标准化的样品制备与设备操作任务。对于流程固定、步骤清晰的实验样品制备，人形机器人可在技术人员的设定和监督下执行任务。例如，按照预设程序进行液体的定量移取与混合、固体的称量与分装、标准菌株的划线接种或为特定测试准备标准测试件。在操作标准化仪器进行常规测试时，

人形机器人可辅助测试人员进行仪表设备操作，在技术人员的指令下或按预设程序，完成开闭设备电源、放置样品到进样口、按下启停按钮、旋转标准旋钮至固定位置、更换特定耗材等基础操作步骤。同时，人形机器人可执行辅助性工作。例如，在检测间隙进行工作台面的基础清洁擦拭、在盘点时辅助扫描样品或设备标签信息。

（四）特种应用场景

人形机器人凭借其类人的形态、强大的环境适应性与灵巧的操作能力，正在加速融入各类高危、复杂或人类难以企及的特种作业领域中，成为突破人类生理极限、执行高危特种作业的关键力量。人形机器人的特种应用场景主要有消防救援、危险性环境作业、极端环境勘探、空间探测四大领域。在消防救援中，人形机器人可以深入灾害现场，执行高危侦察、搜救定位、危险源控制及信息回传等任务，最大限度地保障消防员的生命安全；在核辐射、危化品、有限空间、高空等危险性环境作业中，人形机器人可以替代人员进入核心危险区，精准完成环境探测、阀门启闭、爆破布设、危化品处置等高风险操作及设备精细维护，推动工业安全升级；在深空、深海、极地、火山等极端环境勘探中，人形机器人可以突破人类生存边界，承担勘探勘测、设施搭建、设备操作与维护保障等工作，为科考获取宝贵数据；在空间探测中，人形机器人可以作为航天员的得力助手，高效辅助舱内作业与实验，承担日常维修维护监测工作，并在外星登陆任务中提供物资传递、样本运输、环境监测等关键支持，提升太空任务的效能与安全性。人形机器人在特种作业应用场景中的核心价值在于替代或辅助人类执行危险、繁重、高精度或极端环境下的关键任务，显著提升作业安全性、效率与可靠性，具有较大的应用潜力。

1. 消防救援

人形机器人在消防救援领域具有广泛的应用前景。当地震、火灾、洪水或危化品泄漏等灾害发生时，人形机器人能够快速响应并深入人类难以承受的高危环境中（如高温浓烟区、有毒气体泄漏点、结构不稳的坍塌废墟，以及生化污染区），显著提升救援效率并最大限度地保障人员安全。

首先，人形机器人可以进入人类无法到达或者难以进入的区域，如废墟、坍塌建筑物区域等，这些区域可能会存在危险，但是人形机器人可以替代救援人员进入危险核心区域进行搜索和救援工作，并通过配备热成像仪、生命探测仪等各种传感器和设备检测生命迹象，搜寻并精准定位被困人员，还能探测有害气体、侦察火源位置与蔓延趋势、识别建筑结构风险等，为救援人员提供重要的信息。

其次，人形机器人可以代替救援人员进行现场应急处置或帮助救援人员进行危险的操作。例如，在高温、有毒或放射性强等对人类的生命健康有严重威胁的危险环境中，救援人员可以在安全的环境下通过远程控制人形机器人进入操作，遥控人形机器人手动关闭泄漏的燃气或化学品管道阀门、转移未爆物或腐蚀性容器等高危物品，或自行进行初步的火灾扑灭工作，有效控制危险源，避免救援人员暴露于危险之中。同时，人形机器人能利用机械臂和破拆工具清除通道障碍物，为后续救援力量开辟生命通道。在搜救过程中，人形机器人不仅能细致搜索复杂废墟中的幸存者，还可通过语音、视觉交互对意识清醒的被困者进行安抚，向其紧急投送应急生存包，为等待专业救援争取宝贵时间。作为移动的信息节点，人形机器人能将现场的高清视频、热力图、气体数据及结构评估数据实时同步传输至后方指挥中心，支撑精准的救援调度与协同。此外，人形机器人还可以协助救援人员进行物资运输和后勤保障工作。在应急处置后期或大型灾害现场，物资运输可能会受到限制，但是人形机器人可以在复杂的环境中自主导航，运输重型装备或医疗物资至一线，或在条件恶劣区域提供如送餐、运送医疗用品等必要的基础生活与后勤支持等服务。

2. 危险性环境作业

在有限空间、高空、核辐射、爆破、爆炸品生产及危险化学品处理等极端危险或人类难以承受的作业环境中，人形机器人能够替代或辅助人类深入核心区域执行关键任务，显著提升作业安全性并保障人员生命健康。

首先，人形机器人可以深入人类无法进入或滞留时间受限的高危区域进行作业环节侦查评估。在狭窄、缺氧或有毒气体充斥的有限空间，如地下管道、储罐、反应釜内部等，在结构复杂或承重能力不明的高空作业面，以及在核电站的高辐射区域，人形机器人可以代替人员进入。人形机器人通过利用搭载的设备，精准探测环境参数，如气体浓度、辐射强度、温度、压力等，评估结构稳定性，识别如腐蚀、裂纹、泄漏源等潜在风险点，并将高清画面与实时数据回传至后方控制中心，为安全决策和作业规划提供关键依据。

其次，人形机器人能够直接替代人类执行高风险操作任务。在核电站中，人形机器人可在高辐射环境下，通过遥控或自主方式执行设备巡检、阀门启闭、部件更换或去污作业等设备维护工作。在危险化学品生产或泄漏现场，人员可通过远程操作人形机器人，精确搬运、分装剧毒或强腐蚀性物料，关闭泄漏阀门，或进行初步的污染物覆盖与中和处理。在爆破作业准备阶段，人形机器人能代替爆破员进入危险区，安全、精准地布设炸药和起爆装置。在爆炸品生产线上，人形机器人可承担装配、检测、搬运等关键工序，最大限度降低人员接触爆炸物的风险。同时，人形机器人能利用其机械臂和适配工具执行精细或高强度的辅助作业。在高空作业场景，人形机器人可攀爬、悬挂，进行钢结构焊接、铆接、螺栓

紧固或表面处理。在有限空间或坍塌区域，它能灵活使用各种工具躲避障碍并执行作业。

3. 极端环境勘探

在深空、深海、极地、火山、高原等人类难以抵达或长期生存的极端环境中，人形机器人能够代替人类深入未知区域，执行勘探勘测、巡查监测、设施搭建、设备操作和维护等关键任务，突破人类生理极限，获取宝贵数据并支撑科考活动。

首先，人形机器人能够深入人类无法安全抵达或生存的极端区域进行勘探勘测与巡查监测工作。在深海高压、黑暗、低温的热液喷口附近，或在充满有毒气体、高温的火山口等区域，人形机器人可以代替科考队员进入。其利用搭载的高精度雷达、光谱成像仪、激光扫描仪、辐射探测器等设备，进行详细的地籍测绘、地质构造采样、矿物采集、气体/液体样本采集，以及对温度、压力、辐射等极端环境参数进行长期、连续的监测。同时，它能实时回传高清视频及三维扫描环境数据，为科考队员提供现场数据与决策依据。

其次，人形机器人能够执行需要精细操作和高适应性的设施搭建与设备操作任务。在深海探测站点或极地科考站，它能远程操作或自主进行钻探设备部署、传感器网络铺设、取样器操控、大型科研仪器校准与使用等工作。其灵巧手能够完成拧螺丝、插拔接口、抓取不规则样本等精细动作，还可适应非结构化环境。同时，人形机器人能够承担高风险或重复性的设备维护与保障工作。在偏远无人值守的科考站，人形机器人可执行设备巡检、故障诊断、部件更换、除尘清洁、线缆修复等日常维护保养工作。在遭遇沙尘暴、冰雹等恶劣天气或设备发生意外故障时，它能快速响应，进行应急修复或覆盖设备、紧急关机等保护性操作。此外，在极端环境中，人形机器人还能承担物资转运任务，如将采集的珍贵样本、更换的零部件或应急补给在基地、野外站点之间安全运输。

4. 空间探测

在空间站舱内、航天器内部，以及月球或火星等天体表面环境中，人形机器人能够有效辅助航天员执行作业，承担日常的维修、维护、监测工作，并在外星登陆任务中提供关键支持，从而减轻航天员负担、提升任务效率与安全性。

首先，人形机器人能够在空间站和航天器内部高效辅助航天员完成日常作业与实验任务。在微重力环境下，机器人可自主或经遥操作，执行如物品整理，以及搬运实验样品、工具和备件等物资管理工作，也可以辅助航天员进行舱内环境清洁维护工作。同时，人形机器人可以协助航天员进行科学实验，执行负载仪器传递、重复性步骤操作、实验过程参数监控等任务。人形机器人的灵巧手能够操作

适应人类设计的接口和设备，如开关面板、更换实验模块、连接线缆等。机器人还能作为移动监控平台，定期通过各类传感器，巡检舱内环境参数，如空气质量、温湿度、异常声响或泄漏迹象，并将数据实时反馈，使航天员能更专注于核心科研和决策任务。

其次，人形机器人能够承担大量空间站和航天器内部的日常维修维护与状态监测工作，大大减轻航天员的工作负担。人形机器人可按照预设程序或远程指令，执行系统性的设备巡检工作。例如，检查系统接口、线缆束、通风管道等关键部件的运行状态和完整性，识别松动、磨损或潜在故障点。在日常维护方面，人形机器人能完成更换空气/水过滤器、清洁及润滑活动部件、进行设备功能测试等重复性或耗时的工作。对于如开关跳闸、连接器松动等相对简单的故障，人形机器人可在航天员的监督或指令下进行复位或重新插拔等初步处理，有效分担航天员在维护任务上的时间与精力消耗，并实现高频率的状态监测。

此外，人形机器人能够在外星表面登陆任务中发挥重要的辅助作用。在航天员执行舱外探索任务时，人形机器人可协助航天员穿戴或检查舱外航天服，并充当舱外活动助手，在着陆器附近待命，为航天员预先准备和传递工具、科学仪器、样本收集容器等。在航天员出舱后，可以协助进行光能板架设、模块组装、科研设备布放与连接、通信天线调试等关键设施的搭建工作，并辅助航天员看管设备，运输采集的岩石土壤样本返回着陆器，或在短距离内运送备用物资。人形机器人还能在航天员作业期间，利用自身传感器进行局部环境（如火星或月球表面的崎岖陨石坑、强辐射区、永久阴影区）监测，反馈环境情况，为航天员作业提供安全保障信息。在基地建设或长期驻留场景中，人形机器人可承担舱内外物资转运、简易设施看护，以及部分重复性的基地维护任务，如清扫太阳能板、监测外部设备状态等。

▷第五章 人形机器人政策布局情况

（一）全球主要经济体政策布局

随着Transformer、模型蒸馏等大模型技术取得突破性发展，全球人形机器人产业正处于技术爆发期，技术日益成熟，功能不断完善，已实现下肢在复杂环境中“行走”和上肢进行一系列精细化“作业”，并开始逐步实现“自主认知”与“自主决策”的高级智能。目前，人形机器人已引起了世界主要经济体的高度重视，一些国家政府已将人形机器人融入机器人、人工智能、新材料等技术的推进政策中。

1. 美国

机器人被美国视为国家科技竞争力的核心。2011年，美国发布《国家机器人计划1.0》，旨在通过创新机器人研究和应用，加速机器人的发展和使用，实现协作机器人与人类伙伴的共生关系。2017年，美国发布《国家机器人计划2.0》，旨在鼓励和促进机器人的研究和开发，通过支持教育和研究机构的研究项目、国际合作等途径，加快机器人技术的进步和发展，刺激基础技术研发，以实现协作机器人从各方面协助人类，促进多人与多机器人之间的交互协作。2021年，美国又发布了《国家机器人计划3.0》。此外，美国在2022年发布了《先进制造业国家战略》，将开发创新材料和加工技术作为重点目标之一，认为增材制造是开发和应用先进制造材料和技术的关键。增材制造可解决人形机器人异形件加工的问题；2023年，美国发布了《国家人工智能研发战略计划》，将开发功能更强大、更可靠的机器人作为重要目标，将机器人在崎岖和不确定地形上的移动，以及机器人在复杂环境中与人类安全、协作的互动作为重点研究领域。

2. 欧盟

欧盟在2021年发布了《机器人协调行动计划》，鼓励成员国制订国家机器人计划，并在2025年10月8日发布“应用人工智能”和“科学中的人工智能”两项战略，以加快欧洲工业与科学领域对人工智能的应用，欧盟委员会主席冯德莱恩在声明中说：“从机器人到医疗保健、能源与汽车，我们将在所有关键领域推动这种人工智能优先的思维方式”。在欧盟的统一推动下，欧盟成员国也正在加速推动相关产业。例如，德国在2023年连续发布了《机器人研究行动计划》《轻量化战略》《人工智能行动计划》，均提出加快推动人工智能机器人、行业智能机器人发展，并宣称将成立一个专注于云和边缘计算的机器人研究所。《机器人研究行动计划》包括4个行动领域，一是将基础科技创新应用于机器人开发。通过基础

科技组合提升机器人性能并开拓新的应用可能性；利用人工智能技术推进未来机器人研发；开展可联网机器人基础研究；继续开发高性能微电子器件和安全的机器人。二是汇聚机器人研究机构。对本国机器人顶尖研究机构进行联网；确保机器人研究机构的国际能见度和全球竞争力；制定本国共同研究路线图，紧密参与欧洲相关计划；广泛开展人工智能机器人基础研究。三是为未来机器人的研发人才提供支持。以研究促进人才培养；与教育机构合作，实现相关教育质量的飞跃发展；激励并招聘机器人研发人才；更好更快地将基础科技融入机器人开发。四是开展智能机器人应用。确保机器人科研成果成功、快速转化；将机器人用于民事安全领域；通过人机互动，实现更高的生活质量和社会接受度；机器人研发以服务人类为根本；开发用于深海活动、地表观测、工业4.0、经济可持续转型的机器人，将面向未来能源领域的机器人投入应用。

3. 日本

日本作为最早开展人形机器人研究的国家之一，早在2015年就发布了《机器人新战略》，旨在使日本成为世界第一的机器人创新中心和应用社会，并引导日本迈向世界领先的机器人新时代，维护日本“机器人大国”的优势地位。日本政府在2022年提供了超过9.3亿美元的支持，重点发展机器人产业。同时，日本在2022年发布了《人工智能战略2022》，将通过开发和展示模仿大脑机制的人工智能节能技术、收集和建立日语相关的数据集、开发大规模语言模型等方面的措施，促进人工智能的社会应用。日本政府认为，在人口全面减少的背景下，AI和机器人对实现社会制度的维持和经济增长而言不可或缺，考虑在医疗与护理、农业、生产一线等广泛领域应用包括人形机器人在内的尖端机器人。

4. 韩国

2024年1月，韩国产业通商资源部公布《第四个智能机器人基本计划（2024—2028年）》，提出“K-Robot经济”概念，围绕三大核心战略全面发力。一是强化核心竞争力。计划到2030年，通过公私合作投资超3万亿韩元，重点攻克传感器等5项硬件技术和自主运动等3项软件技术，同时培养15000名AI和软件领域的先进机器人产业前沿人才，培育150家专业智能机器人企业，构建完善的产业投资企业生态圈，为产业发展提供强大的技术和人才支撑。二是拓展市场规模。目标是到2030年，在制造业、物流、福利、安全等全产业领域大规模普及超100万台机器人，充分发挥机器人在提高生产率、降低事故率等方面的作用，提升机器人产业和社会贡献度。同时，加强国际合作，积极支持海外认证，助力韩国机器人企业积累本土业绩，进而进军全球市场。三是完善基础设施。在2024年广泛征求各界意见后全面修订《智能机器人开发和普及促进法》，以应对技术革新和产业变革；推进落实《先进机器人监管创新方案》，建立科学合理的监管机制；建立机器人保险制度，降低机器人应用风险；投资建设国家机器人试验场，为机器人研发和测试

提供专业场地；推动机器人伦理宪章的制定，确保机器人在符合伦理道德的框架内发展。

（二）我国人形机器人政策布局（国家层面）²⁰²³

年9月13日，工业和信息化部印发《关于组织开展2023年未来产业创新任务揭榜挂帅工作的通知》，面向元宇宙、人形机器人、脑机接口、通用人工智能4个重点方向，聚焦核心基础、重点产品、公共支撑、示范应用等创新任务，突破一批标志性技术产品，加速新技术、新产品落地应用。核心基础方面包括全身动力学控制算法、电机驱动器、力传感器、MEMS姿态传感器、触觉传感器等。重点产品包括旋转型电驱动关节、直线型电驱动关节、机械臂与灵巧手、高算力主控制器、高能量密度电池等。公共支撑包括人形机器人的端到端仿真开发平台、人形机器人的标准测试与评估、人形机器人的机器脑智能控制技术等。示范应用包括面向工业制造的典型应用、面向灾害救援的典型应用、面向危险作业的典型应用、面向智慧物流的典型应用、面向安防巡逻的典型应用、面向服务娱乐的典型应用等。

2023年10月20日，工业和信息化部印发《人形机器人创新发展指导意见》，提出到2025年，人形机器人创新体系初步建立，“大脑、小脑、肢体”等一批关键技术取得突破，确保核心组件的安全有效供给。整机产品达到国际先进水平，并实现批量生产，在特种、制造、民生服务等场景得到示范应用，探索形成有效的治理机制和手段。培育2~3家有全球影响力的生态型企业和一批专精特新中小企业，打造2~3个产业发展集聚区，孕育开拓一批新业务、新模式、新业态。到2027年，人形机器人技术创新能力显著提升，形成安全可靠的产业链供应链体系，构建具有国际竞争力的产业生态，综合实力达到世界先进水平。产业加速实现规模化发展，应用场景更加丰富，相关产品深度融入实体经济，成为重要的经济增长新引擎。

2024年1月18日，工业和信息化部、教育部、科学技术部、交通运输部、文化和旅游部、国务院国有资产监督管理委员会、中国科学院7部门联合印发《关于推动未来产业创新发展的实施意见》，提出要面向国家重大战略需求和人民美好生活需要，加快实施重大技术装备攻关工程，突破人形机器人、量子计算机、超高速列车、下一代大飞机、绿色智能船舶、无人船艇等高端装备产品，以整机带动新技术产业化落地，打造全球领先的高端装备体系。在创新标志性产品的专栏中，明确指出要突破机器人高转矩密度伺服电机、高动态运动规划与控制、仿生感知与认知、智能灵巧手、电子皮肤等核心技术，重点推进智能制造、家庭服务、特殊环境作业等领域产品的研制及应用。此外，2023年8月3日，工业和信息化部、科学技术部、国家能源局、国家标准化委员会4部门联合发布了《新产业标准化领航工程实施方案（2023—2035年）》，将人形机器人列为前瞻布局未来产业标

准研究之一，提出要研制人形机器人术语、通用本体、整机结构、社会伦理等基础标准。开展人形机器人专用结构零部件、驱动部件、机电系统零部件、控制器、高性能计算芯片及模组、能源供给组件等基础标准预研。研制人形机器人感知系统、定位导航、人机交互、自主决策、集群控制等智能感知决策和控制标准。开展人形机器人运动、操作、交互、智能能力分级分类与性能评估等系统评测标准预研。开展机电系统、人机交互、数据隐私等安全标准预研。面向工业、家庭服务、公共服务、特种作业等场景，开展人形机器人应用标准预研。

（三）我国人形机器人政策布局（地方层面）

在国家战略的指引下，地方政府也纷纷将人形机器人产业作为区域经济转型升级和科技创新的重要抓手，加快出台相关政策，政策支持主体也逐渐从“服务机器人”“工业机器人”向“人形机器人”转变。部分重点城市的政策如下。

1. 北京市

2023年6月16日，北京市政府发布《北京市机器人产业创新发展行动方案（2023—2025年）》，定位“1+4”产品体系，即“人形机器人+医疗健康机器人、协作机器人、特种机器人、物流机器人”。

《北京市机器人产业创新发展行动方案（2023—2025年）》着眼世界前沿技术和未来战略需求，加紧布局人形机器人，带动医疗健康、协作、特种、物流四类优势机器人产品跃升发展，实施百项机器人新品工程，打造智能驱动、产研一体、开放领先的创新产品体系。在人形机器人方面，明确指出要对标国际领先人形机器人产品，支持企业和高校院所开展人形机器人整机产品、关键零部件攻关和工程化，加快建设北京市人形机器人产业创新中心，争创国家制造业创新中心。以人形机器人小批量生产和应用为目标，打造通用智能底层软件及接口、通用硬件开发配套设施等基础条件，集中突破人形机器人通用原型机和通用人工智能大模型等关键技术，大力推动开源控制系统、开源芯片、开源仿真软件等研制和应用。以3C电子制造、新能源汽车生产、安防应急等典型场景应用示范为牵引，通过“揭榜挂帅”等方式支持产业链上下游企业联合开展产品攻关和产线建设，加速全产业链自主化进程。在其他四类优势机器人方面，要发挥本市机器人产业基础优势，提升医疗健康、协作、特种、物流四类机器人的技术水平和市场竞争力。医疗健康机器人领域，面向行业前沿和医疗应用需求，提升手术机器人的全流程手术辅助能力；发挥首都医工协同优势，优化医疗器械注册和卫健医保政策，加大临床应用推广力度。协作机器人领域，突破整机轻量化、视觉增强、力感知、柔顺控制、自学习等关键技术，提升整机荷载能力，发展自适应协作机器人；开展复杂任务执行算法和末端执行器技术攻关，推动多自由度协作机器人的

智能化、柔性化发展。特种机器人领域，提高消防、安防与救援机器人等产品的稳定性、可靠性，优化产品设计和加工工艺，降低综合成本。物流机器人领域，提升导航和运行精度，推动智能仓储物流机器人迭代升级，开发重载移动机器人，优化物流机器人的应用环境。

2. 上海市

2023年10月，上海市经济和信息化委员会、上海市发展和改革委员会、上海市科学技术委员会、上海市财政局、上海市统计局5部门联合印发了《上海市促进智能机器人产业高质量创新发展行动方案（2023—2025年）》。

《上海市促进智能机器人产业高质量创新发展行动方案（2023—2025年）》中提出，要推动工业机器人规模化。聚焦上海重点应用领域，研制面向电子信息、生命健康、汽车、高端装备、先进材料、时尚消费品等产业的切割、焊接、切削、磨抛、装配、喷涂等复杂工艺机器人，以需求牵引带动量能提升，打造10个行业一流的机器人“上海品牌”。促进服务机器人场景化。完善服务机器人全产业链生态体系，聚焦机器人的智能水平提升，支持企业布局智能机器人新赛道，引入语言通用大模型、视觉通用大模型、多模态通用大模型等前沿技术，赋予服务机器人更强的人机交互能力。扩大服务机器人多场景应用，打造100个标杆示范的应用场景。加快通用机器人工程化。围绕通用机器人研发、测试、生产制造、落地应用全过程，培育提升涵盖工程化实践全过程的一体化服务能力，聚焦以大模型、具身智能等人工智能技术驱动的通用机器人关键领域攻关，推进关键共性技术的标准研制及落地推广，加快通用机器人（特别是人形机器人）工程化应用，提升上海在通用机器人领域的影响力。

3. 浙江省

2024年9月2日，浙江省制造业高质量发展领导小组办公室印发了《浙江省人形机器人产业创新发展实施方案（2024—2027年）》，提出力争到2027年，科技创新有效引领人形机器人产业创新发展，在人形机器人领域培育省级及以上高能级创新载体5家、企业研发机构30家，实施重大科技项目30项，产业链供应链实现自主可控；企业科技创新主体作用充分发挥，培育链主企业5家，制造业单项冠军和专精特新“小巨人”企业50家，形成具有国际竞争力的全产业链优势；人形机器人产业协同布局和集群化发展格局基本形成，建设省级未来产业先导区2个，打造示范应用场景50个，标志性产品在重点领域规模化应用取得实质性进展；产业配套与创新发展生态进一步完善，产业规模实现跨越式增长，全省整机年产量达到2万台，核心产业规模达到200亿元，关联产业规模达到500亿元。并制定了以整机制造为引领，推动整零布局协同化；以企业创新为主导，推动技术攻关体系化；以强链补链为核心，推动产业培育集群化；以市场需求为导向，推动场景应用多元

化；以服务保障为支撑，推动创新发展生态化5个方面的重点任务，明确了发挥杭州、宁波在科创、区位、人才等方面的优势，培育壮大整机企业，打造人形机器人整机引领区。支持杭州搭建“公版”通用整机平台，满足不同场景下个性化功能的二次开发，重点发展通用人形机器人；面向健康养老、商业服务和特定场景应用需求，发展服务人形机器人和特种人形机器人。支持宁波聚焦制造领域，重点发展工业人形机器人。鼓励各地结合实际，引育人形机器人整机企业。发挥整机引领区作用，通过整零协同促进产业链上下游高效衔接，带动产业布局协同发展。在省级文件的指导下，杭州、宁波纷纷发布了本市的政策文件。

2024年12月，杭州市发展和改革委员会印发了《杭州市人形机器人产业发展规划（2024—2029年）》，提出要协同产学研各方资源和力量，打造服务全产业链的关键技术攻关平台、公共技术服务平台、训练检验中试平台、场景供需对接平台、产业金融服务平台五大服务平台。并推动余杭区、西湖区、滨江区结合企业和高校优势，建设产研转化引领区；支持萧山区、钱塘区依托机器人产业发展优势，建设产用结合示范区；支持上城信息智能高新区、临平经济技术开发区、富阳经济技术开发区、临安青山湖科技城等产业平台依托制造产业发展基础，建设协同制造发展区。同时，杭州市人民政府办公厅还发布了《杭州市促进人形机器人产业创新发展若干政策措施》，提出积极发展人形机器人产业后市场服务，围绕工业制造、市政巡检、医疗康养、文旅服务、应急作业、安保巡逻、城市管理、教育科研、生活服务等领域，全力打造一批优质应用场景并发布推广。并对关键核心技术攻关、科技创新主体建设、引育优质项目、企业投资增产、融链集聚发展、打造应用场景和示范应用高地、场景应用支撑、产业基金引导、人才引育、服务保障等方面进行了针对性支持，如提出鼓励国有资本平台将楼宇厂房出租或出售给人形机器人企业用于办公生产、创新创业，根据人形机器人创新创业项目的创新能力、市场潜力，可由人形机器人产业创新平台提供使用时间不超过3年、面积不超过1000m²的免费办公场所。

2025年6月，宁波市人民政府办公厅印发了《宁波市促进人形机器人产业创新发展的若干政策措施》，提出支持创新主体建设、推动关键技术攻关、加强人才引进培养、加大项目投资力度、鼓励整零协同发展、构建测试验证服务、拓展场景示范应用、加大金融资本支持、打造开源开放生态、强化服务保障10项支持政策。如支持人形机器人整机和关键零部件企业高水平建设创新中心、产业技术工程化中心、企业技术中心等企业科技创新载体。对获批创建国家级、省级技术（制造业、产业）创新中心的，分别给予不低于1亿元、1000万元的扶持补助（国家、省另有规定的除外）。推进产业技术工程化中心建设，对成功认定为国家级、省级产业技术工程化中心的，分别给予不超过3000万元、1000万元的分档分阶段奖励。支持企业技术中心创建国家级企业技术中心，对成功认定为国家级企业技术中心的单位，给予100万元的奖励。鼓励龙头企业牵头组建创新联合体，凝聚各方

力量加快创新进程，主动承接国家人形机器人揭榜挂帅任务，对承接国家揭榜挂帅任务的联合体创新项目，给予200万元的奖励。在“科创甬江2035”重点研发计划中设立人形机器人专项，每年安排不少于8个项目予以支持。支持承担国家和省重大科技创新任务，对国家重点研发计划项目、国家科技重大专项项目和科技创新2030重大项目，根据相关规定对项目牵头单位和课题承担单位分别按照国家支持经费（不包括外拨市外合作单位部分）1：1和1：0.5比例予以配套支持。对于牵头省“尖兵”研发计划项目和“领雁”研发计划项目，参照省定标准予以支持。

4. 广东省

广东省人民政府办公厅印发的《广东省关于人工智能赋能千行百业若干措施》及广东省工业和信息化厅、广东省发展和改革委员会、广东省科学技术厅、广东省商务厅、广东省市场监督管理局印发的《培育未来智能装备产业集群行动计划》中，均将人形机器人列为重点发展方向，推动其研制和应用。2025年3月9日，广东省人民政府办公厅印发的《广东省推动人工智能与机器人产业创新发展若干政策措施》提出，支持关键核心技术攻关、培育优质企业、打造应用场景、推动产业集聚发展、支持重点项目建设、丰富数据要素供给、完善人工智能与机器人开源创新生态、引育高水平领军人才、加强产业投融资、推进标准体系建设、打造高端交流平台、建立包容审慎监管机制12项支持举措，并明确了具体的支持方式。如支持企业、高校、科研院所等各类创新主体开展联合攻关，围绕人工智能与机器人产业链上下游组建产业创新联盟，加快构建全过程创新链。组织实施省重点领域研发计划“新一代人工智能”“智能机器人”等旗舰项目、重大专项，在人工智能与机器人领域部署一批攻关任务。对国家科技重大专项符合省级配套条件的人工智能与机器人领域重点项目，省财政按规定给予配套奖励，单个项目省级配套金额超1亿元（含）的，按“一事一议”方式研究给予支持。创建人工智能与机器人领域制造业创新中心，对符合条件的国家级、省级制造业创新中心，省财政按规定分别给予最高5000万元、1000万元的资金支持。

2025年3月，深圳市科技创新局印发了《深圳市具身智能机器人技术创新与产业发展行动计划（2025—2027年）》，计划到2027年，在机器人关键核心零部件、AI芯片、人工智能与机器人融合技术、多模态感知技术、高精度运动控制技术、灵巧操作技术等方面取得突破。新增培育估值过百亿企业10家以上、营收超十亿企业20家以上，实现十亿级应用场景落地50个以上，关联产业规模达到1000亿元以上，具身智能机器人产业集群相关企业超过1200家。打造公共服务平台矩阵，吸引更多上下游企业、科研机构、创新团队等加入，形成更完善的产业生态，具身智能机器人产业综合实力达到国际领先水平。并提出了打造公共服务平台矩阵，包括支持光明实验室建设大湾区昇腾算力应用创新研究院，推动昇腾AI算子开发

和开源社区生态建设；支持组建具身智能机器人领域市重点实验室；支持国际先进技术应用推进中心（深圳）建设具身智能技术试验场；支持深圳市人工智能与机器人研究院建设广东省具身智能机器人创新中心；布局概念验证中心、中小试基地等产业创新服务载体；开发真机数据采集平台和数据仿真平台等。

5. 湖北省

2025年6月，武汉市发布《武汉市加快人形机器人产业发展行动方案（2025—2027年）》和《武汉市加快人形机器人产业发展的若干政策措施》，实施平台聚势行动、场景示范行动、整机领航行动、部组件筑基行动和生态培育行动，加快打造全国人形机器人专用传感示范地、场景应用优选地、智能制造集聚地。提出建设概念验证、中试和共享制造平台，最高支持1000万元。对大模型研发和算力使用分别最高补贴500万元，对数据集建设最高奖励200万元。重点推动危险活儿、枯燥活儿、民生活儿三类场景开放。实行双边奖补：对场景提供方，按项目投入的30%最高奖励100万元；对产品供应方，按研发投入的20%最高奖励100万元。未来重点瞄准“交互型”“高精度型”“高可靠型”三类标志性整机产品，整机企业牵头组建联合实验室，最高可获6000万元支持；针对特定场景搞“靶向研制”，最高可支持4000万元。在零部件方面实行双边奖补：整机企业采购非关联企业部组件产品的，整机、部组件企业，每年分别可获最高补贴500万元。推行“揭榜挂帅”制，整机企业发布需求，部组件企业揭榜攻关成功的，可获最高补贴100万元。对通过兼并重组方式向人形机器人赛道拓展的，给予最高500万元支持。对入选武汉“产业名园”的人形机器人园区，最高奖励5000万元。参与设立省级100亿元产投母基金，市级新设两只10亿元以上子基金；对种子直投、天使直投项目实行“尽职免责”。赋予创新平台、龙头企业武汉英才计划的举荐权，最高奖励100万元。

6. 江苏省

2024年4月17日，江苏省工业和信息化厅、江苏省发展和改革委员会、江苏省科技厅、江苏省财政厅、江苏省市场监督管理局联合印发《江苏省机器人产业创新发展行动方案》提出到2027年，我省机器人产业综合实力达到国际先进水平，机器人成为经济发展、人民生活、社会治理的重要组成。工业机器人整机综合指标达到国际先进水平，关键零部件性能和可靠性达到国际同类产品水平；服务机器人和特种机器人行业应用深度和广度显著提升，在家政服务、养老助残、医疗康复、教育娱乐、维护巡检、安全应急等领域实现广泛应用；人形机器人“大脑、小脑、肢体”等关键核心技术取得突破，大模型等人工智能技术加快发展，典型应用场景更加丰富，争创国家级产业发展集聚区。

2024年11月，南京市政府办公厅发布的《南京市促进机器人产业高质量发展行动计划（2024—2026年）》中提出，到2026年，南京市机器人产业总体发展水平居全国前列；自主工业机器人在行业内继续保持领先优势；经营主体加速集聚，产业竞争力更加提升；“机器人+”创新示范应用取得显著成效，成为国内标杆。围绕这个主要目标，着眼面向国家重大战略需求和全市产业发展需要，该计划提出聚力打造“1+N+1”整机体系。前一个“1”表示打造工业机器人领先优势。依托工业机器人国内“链主”地位和品牌领先优势，打造更具全球竞争力的TOP 5品牌，进一步提升南京市工业机器人产品的全球市场占有率和影响力。“N”表示全面提升一批特色机器人产业。聚焦协作、物流、安全应急、公共服务、个人消费、医疗健康机器人等领域，打造门类丰富、各具特色的产业发展体系。后一个“1”代表前瞻布局人形机器人产业。全面推进人形机器人“大脑”“小脑”“肢体”等部件研发，构筑人形机器人通用整机平台，初步建立人形机器人整机的小批量生产制造能力。

此外，山东、四川、江西等省份和重点城市也发布了相关政策文件，此处不再一一列举。

▷第六章 人形机器人产业链情况

（一）产业链划分

人形机器人产业链主要由上游零部件、中游人形机器人整机，以及下游终端应用等环节构成。目前，我国已基本形成了覆盖人形机器人全产业链的供给体系，但由于人形机器人尚未在下游终端应用领域实现规模化、商业化落地，且随着人形机器人技术的不断成熟，其核心零部件的技术选型也在持续探索，我国人形机器人产业链仍在不断构建和优化中。未来，在人形机器人创新体系加速完善，以及“大脑”“小脑”“肢体”等关键技术突破的牵引下，我国有望形成高效、可靠的人形机器人产业链和供应链体系。目前，我国人形机器人整机企业超140家，产业链相关企业数量超400家。

（二）产业链上游情况

在产业链上游方面，人形机器人产业链上游零部件和软件主要涉及的技术板块有智能决策系统（大脑）、运动控制系统（小脑）、芯片和基础软件、本体零部件（包括灵巧手、一体化关节、电机、减速器、丝杠、结构件、电池以及各类传感器等）和数据供给技术。产业链中最具价值的部分在于智能决策系统和运动控制系统，这两个系统决定了人形机器人自主执行各类任务的能力，也是人形机器人产品核心竞争力的体现。掌握运动控制和人工智能算法等核心技术的企业，将在技术层面上主导人形机器人发展的方向和节奏，成为此领域的“中枢”和

“脑”，同时也将成为中游人形机器人本体制造的主导企业。但目前，全球在智能决策和运动控制方面都仍处于探索阶段，仅具备初阶人类“脑力”，可以完成抓取、分拣、巡逻等简单任务和走、跑、跳等基本动作。本体零部件方面，由于我国工业基础好、相关产业链齐全，人形机器人本体零部件的发展潜力巨大。从当前形势来看，传感器、减速器、电机和丝杠等核心零部件的价值占比较高，增量空间显著。根据调研，本体零部件成本占据了产品成本中的主要部分。目前，全功能、全尺寸通用人的价格在60万~80万元，硬件成本大约在50万元。其中，电机、减速器等执行机构成本占比在50%~60%，力矩传感器、视觉传感器等传感器成本占比15%左右，芯片电池等成本占比10%~15%，其他结构件、标准件等成本约占比15%。零部件的技术突破将有助于提升人形机器人的性能、降低成本、提高可靠性和安全性，并促进技术创新，从而推动人形机器人产业化的进程。数据供给方面，目前业界通常通过模拟数据和真实数据结合的方式训练人形机器人。由于人形机器人对真实数据的需求较高，目前业界主要通过建设人形机

机器人训练场模拟各个应用场景，并采集人形机器人或人类在该场景作业时的运动数据。如国家地方共建人形机器人创新中心建设的“麒麟人形机器人具身智能训练场”，拥有超过4000m²的虚实结合人形机器人训练场地，以及10种平台超100台异构人形机器人本体和训练工位。北京、武汉、杭州、苏州、合肥等多地也在开展训练场建设。

（三）产业链中游情况

人形机器人本体的设计、制造与集成是产业链的关键环节。目前，人形机器人本体制造业正处于产业化发展的初期阶段，行业内对人形机器人本体的价值已达成共识，但大规模商业化落地仍未真正实现。零部件选用、软硬件整合、本体设计，以及应用场景的聚焦仍在探索之中，各类人形机器人的总体解决方案正持续迭代与创新，人形机器人产业的“0~1”拐点有望提前到来。然而，人形机器人本体的发展面临多重挑战，包括技术基础亟待增强、制造成本需显著降低、算力需求高和场景验证困难等问题。唯有解决上述发展困境，人形机器人行业才能成功从实验研究转向规模商业应用。

目前，我国人形机器人产业链中游的整机研发制造企业采用“整机集成、关键零部件自研”的路线，基于国内现有产业链，快速推进整机产品迭代，已发布傅利叶GR-1、宇树Unitree H1、优必选Walker S等近百款人形机器人产品。但目前业内仍以创新型企业为主，其产业链带动性和话语权较低，不同整机企业在开发人形机器人时往往采取不同的技术路径和解决方案，对一体化关节等核心技术的重复投入现象明显。此外，由于尚未形成标准化的数据集、开发平台和生产流程，人形机器人在研发设计和生产制造等环节也难以实现规模效应。

（四）产业链下游情况

在产业链下游方面，目前众多人形机器人制造商在产品研发和市场策略上积极瞄准各类下游应用领域。从各个厂商的应用规划来看，3~5年内，人形机器人将主要应用于科研教育、工业制造、仓储物流、商业导览等简单应用领域；5~10年后，人形机器人的目标则是进入千家万户，为家庭中涉及的养老、育儿、家政等场景提供相关服务。随着应用的普及，面向各领域的专业运营企业也将应运而生。此外，业界领军企业也在加速打造人形机器人开源软硬件平台，开源、协同、高效的产业链创新生态正在逐步形成。

从场景探索上看，工业场景方面，北京、上海的创新中心和国内领军企业积极联合用户单位，围绕零部件加工、上下料质检、分拣装配、设备检修等场景，加快探索基于人形机器人的生产作业方案。民生场景方面，企业加速探索迎宾导航、

无人售卖、家政清扫、医疗操作等场景，如美团24h无人药店、便利蜂等已试用人形机器人产品。但现阶段，人形机器人仍处于产业早期，供给方的技术能力尚无法完全满足应用方的实际需求。应用方鉴于购买、维护成本和安全性等方面的考虑，也不愿大规模购入或放开应用场景，并对人形机器人的应用整体持审慎态度。经调研了解，目前人形机器人的客户大部分是高校和科研单位，它们购买人形机器人用于研发，人形机器人尚未广泛进入制造业、服务业等应用领域。

▷第七章 国内人形机器人产业分布情况

（一）产业聚集区重点城市情况

目前，我国已初步形成了京津冀（华北）、长三角（华东）、珠三角（华南）3个人形机器人产业集聚区，聚集了一批人形机器人整机企业和产业链上游企业。其中，北京、上海、杭州、深圳等地布局较快，已具备一定产业基础和优势。

北京市支持用户联合企业建设“机器人+”应用体验中心，聚焦制造业、商贸物流、农业等特定场景需求，支持用户与企业联合研制面向典型应用场景的创新产品和系统解决方案，并通过短期租赁、共享服务、代运营等方式加强应用推广，催生服务新业态。首创了机器人未定型创新产品首试首用奖励政策，两批次奖励政策累计支持38个企业的78款创新产品在商用服务、制造业、建筑、商贸物流、安全应急、能源、农业、医疗等8个领域25类场景完成首试首用和迭代升级。同时，还设立百亿机器人产业发展基金，发挥逆周期和跨周期调节作用，培育产业龙头企业，健全机器人产业链布局。2025年，北京举办了全球首个人形机器人半程马拉松赛事，20支人形机器人参赛队伍同场竞技，全面检验了当前人形机器人产品的运动控制、可靠性、续航、环境适应等性能，推动人形机器人向产品化、实用化迈出坚实的一步。全球首个具身智能机器人4S店于2025世界机器人大会期间正式亮相，落地北京市机器人产业园（亦庄）。该4S店以具身智能机器人为核心，创新性地集机器人销售、零配件供应、售后服务、信息反馈于一体，旨在构建覆盖机器人全生命周期的服务体系，实现供需两端高效对接与价值转化。在市政府的大力推动下，北京市经开区、海淀区、石景山区等区域也发布了相关计划，将人形机器人作为主导产业，推动集聚发展。市区联动推进机器人产业园建设，搭建共性加工、检测、训练平台。目前，北京重点推动“一南一北”机器人产业园项目，规划建设25万平方米的亦庄机器人产业园、30万平方米的昌平机器人产业园，打造北部机器人产研结合示范区和南部机器人产用结合集聚区。

上海市围绕电子信息、生命健康、汽车、高端装备、先进材料、时尚消费品六大重点产业，“一业一策”“分级分类”“一厂多景”推进工业企业智能制造转型。创新商业模式，加大智能机器人在园区和商业中心等场景的展示力度；鼓励软件集成商针对不同类型的机器人搭建多类型机器人生命周期管理平台。探索机器人即服务（RaaS）模式，培育智能机器人系统集成商，推广“服务租赁+系统集成”的商业模式，降低用户使用门槛和成本，面向企业用户试点“免购即享”。设立总规模约2亿元的“人形机器人发展基金”，以“知识产权投资+VC投资”复合的模式，发挥知识产权保护中心效能。推动上海企业积极开展国际标准、国家

标准、行业标准制定，加快标准引领。上海重点打造机器人产业“3+X”空间布局，其中以浦东新区张江机器人谷、金桥机器人产业园、临港机器人产业基地为核心园区载体，聚集了工业机器人、高端医疗机器人等上下游企业，产业发展水平全市领先；还有类似上海马桥人工智能试验区、宝山机器人产业园、嘉定机器人视觉产业合作中心等园区的研发和制造优势尤其突出。

杭州市依托本地头部企业，建成了具身智能领域的产业集群。目前，杭州已形成了“六小龙”引领的科技创新力量，宇树科技、云深处更是业内的领军企业。产业链上下游协同加速，关键零部件国产化取得进展，谐波减速器、伺服电机等成本的下降推动了其普及。此外，杭州通过展会、技术对接会等活动强化生态合作。2025年6月20日—22日，2025杭州国际人形机器人与机器人技术展览会召开，该展会是首个人形机器人产业链专业展，依托杭州湾区独特的产业优势和强大的市场需求，促进人形机器人技术创新，加强国际技术交流与合作，推动人形机器人产业的升级与发展。展品范围覆盖机器人关键零部件、整机制造、系统集成等产业链与应用服务，展会期间还举办“杭州人形机器人大会”，为业界提供一个了解人形机器人的关键技术、发展趋势、产业应用的平台。

深圳市作为中国改革开放的标杆城市，正在通过紧抓科技创新实现从“政策红利”向“产业红利”的转型，坚持以产业需求为导向，推动产业结构向高端化、智能化跃迁。深圳市基于自身的科技创新优势和产业基础条件，已构建覆盖“上游核心零部件-中游整机制造-下游场景应用”的全链条产业体系。上游核心零部件方面，深圳汇聚了全球最密集的机器人核心硬件供应商，包括微型驱动系统、激光雷达、伺服系统、3D视觉、触觉传感器等。中游整机制造方面，深圳拥有优必选、乐聚、众擎、逐际动力等多家整机头部企业，覆盖教育、工业、服务等全场景。优必选的Walker S1已进入奥迪、比亚迪工厂承担质检与搬运任务；众擎机器人完成全球首例前空翻特技，技术迭代速度领先全球。下游场景集成方面，深圳开放工业制造（比亚迪、富士康）、公共服务（环卫、医疗）、特种作业（抢险救灾、深海探测）等场景，推动人形机器人从实验室走向规模化应用。目前，深圳已建成宝安5G及机器人智能制造产业园、中粮（福安）机器人智造产业园、南山机器人产业园等产业园区。

（二）创新中心建设情况

目前，我国已建设了上海、北京两家国家地方共建制造业创新中心，并涌现出了山东省人形机器人制造业创新中心、浙江人形机器人创新中心、广东省具身智能机器人创新中心、安徽省人形机器人产业创新中心、湖北人形机器人创新中心、成都人形机器人创新中心、宁波具身智能机器人创新中心、苏州市具身智能机器人综合创新中心等一批省市级创新中心。

1. 国家地方共建人形机器人创新中心

国家地方共建人形机器人创新中心（以下简称“上海中心”）依托人形机器人（上海）有限公司，该公司于2023年12月成立，并于2024年1月，获批成立省部级上海人形机器人制造业创新中心，2024年5月，该中心获工业和信息化部揭牌为“国家地方共建人形机器人创新中心”。上海中心聚焦人形机器人共性关键技术领域，汇集“政产学研用”各方力量，依托上海产业聚集及辐射全球的优势，加快实现关键技术装备产业化。目前，上海中心已发布“青龙”全尺寸通用机器人，围绕“通用化、强对标、多模态、工程化”开展全尺寸通用机器人研制，开源开放，打造人形机器人产业链生态。模块化、系列化牵引研制行动、驱动、感知、执行等核心组件，支撑整机设计集成。“朱雀”通用大脑模型基于多模态大模型，具备语音对话、视觉感知、任务执行等能力。“玄武”通用小脑模型基于强化学习与动力学技术，可完成复杂的操控与作业。“白虎”具身智能数据集采用国地中心定义的标准规范进行生产、处理和存储，并对外开源，助力具身智能时代的加速到来。此外，上海中心还孵化及运营了“OpenLoong”开源项目，整体遵从全栈开源设计，包括云端大脑、具身小脑、具身实体、具身数据。该开源项目以推动人形机器人完成全场景应用为目标，基于开源的方式，搭建人形机器人具身智能操作与操控系统的框架与公版开源硬件平台，促进通用机器人产业繁荣发展。项目功能按照“系统集、子系统、软件包”逐级展开，在不同应用需求下可以依据硬件和软件的实际需求裁减某些非必要的组件。

2. 国家地方共建具身智能机器人创新中心

国家地方共建具身智能机器人创新中心（以下简称“北京中心”）依托北京人形机器人创新中心有限公司，于2023年11月由小米机器人、优必选、京城机电、亦庄机器人等企业联合组建成立。2024年10月，工业和信息化部与北京市人民政府在北京共同授牌，该地正式升级为“国家地方共建具身智能机器人创新中心”。

在产品发布方面，2024年4月，北京中心发布了机器人母平台“天工LITE”，实现了全球首例纯电驱全尺寸人形机器人的拟人奔跑。发布后，“天工”系列机器人的奔跑时速不断升级，实测平均速度可达10km/h，最高奔跑速度已提升至全球领先的12km/h，且具有极强的泛化能力，能在斜坡、楼梯、草地、碎石、沙地等多种复杂地形下实现平稳移动。“天工”系列人形机器人具备开源开放性和兼容扩展性，可以实现灵活扩展软、硬件等功能模块，充分满足不同场景下的需求。科研机构及机器人相关企业可基于“天工”母平台的软硬件功能，根据市场需要进行二次开发。2024年8月，北京中心发布了“天工1.1 PRO”。“天工1.1 PRO”拥有完整的手臂功能，身高163cm，体重56kg，具有42个自由度，内置多个视觉感知传感器及六维力传感器，配备惯性测量装置（IMU）和3D视觉传感器，每秒可进行550万亿次计算，具有更强的智能交互和任务执行能力。同时，北京中心

还发布了具身智能服务机器人“天轶”。“天轶”身高163cm，体重72.5kg，拥有16个自由度和轮式全向移动能力，腿部具备升降功能，可智能适配服务对象的身高和状态，配合完整双臂和胸口大屏，可以实现定位导航、平稳移动、自然对话和精细手臂操作，可应用于接待、导览、咨询问答、场馆介绍等多种场景。此外，北京中心发布的通用具身智能平台“慧思开物”，能够实现从任务理解到执行的全流程智能化，既具备处理多场景复杂任务的泛化能力，又实现了单个软件系统在机械臂、轮式机器人、人形机器人等多构型本体上的兼容。“慧思开物”由AI大模型驱动的任务规划具身“大脑”和数据驱动的端到端技能执行具身“小脑”构成，是支持各类机器人与物理世界交互的开放平台。在数据集方面，2024年4月，北京中心发布具身智能数据集建设方案，牵头产业合作伙伴，打造高质量具身智能数据集，通过虚实结合的仿真技术平台，加速具身智能迭代。目前，北京中心数据采集涵盖六类本体的七大典型场景，日产数据已达10TB。

3. 山东省人形机器人制造业创新中心

山东人形机器人创新中心有限公司由海信集团控股股份有限公司、国家地方共建人形机器人创新中心主导，山东优宝特智能机器人有限公司（以下简称“优宝特”）共同参与建设。创新中心主要聚焦“人形机器人+”家庭服务方向，以“需求”和“技术”为双轮驱动，以企业为创新主体，联合产学研伙伴开展技术创新。优宝特作为创新中心参办方，在智能机器人技术领域拥有深耕积累和创新能力，在仿生机构、关节模组、驱动技术、控制算法等方面，拥有自主知识产权的核心技术。优宝特开发出多款具有行业领先水平的产品——四足机器人Y30，可载重攀爬泰山，特别是人形机器人“行者泰山”在2025年6月18日进行了“泰山拉练”，其作为业界首个攀爬泰山的人形机器人，引起社会广泛关注。

4. 浙江人形机器人创新中心

浙江人形机器人创新中心成立于2023年12月，由宁波市政府与浙江大学熊蓉教授团队联合共建，聚焦人形机器人智能感控技术研发与整机系统开发，并与传化集团等企业开展战略合作。目前，该创新中心的产品包括：领航者1号（2024年3月发布），身高1.5m，体重50kg，配备多自由度灵巧手，可完成舞蹈表演及简单服务动作。领航者2号NAVIAI（2024年8月发布），自由度增至41个，具备全身协调稳定行走能力，应用于安防巡逻、高精装配等场景。目前，该创新中心已实现了4个方面的技术突破，一是双足行走技术，通过全身运动协调算法实现复杂地形移动；二是人工智能融合应用，结合大模型与高算力（AI算力275Tops），支持擦桌、倒水等精细操作；三是高精度控制，仿人臂负载达5kg，末端精度0.1mm，适应工业场景需求；四是情感交互能力，第二代产品具备情感反馈功能，人性化的设计让人形机器人更加贴近生活。

5. 广东省具身智能机器人创新中心

广东省具身智能机器人创新中心依托深圳国创具身智能机器人有限公司，于2024年4月由香港中文大学（深圳）牵头联合产业链上下游企业共建，聚焦具身智能机器人核心部件等关键共性技术，服务机器人产业创新，搭建中试、测试平台，制定行业应用标准，打造贯穿创新链、产业链和资金链的人形机器人生态。南山区依托该中心推出超400万平方米的产业空间，形成“研发-转化-应用”全链条生态。截至2024年，已集聚20余家产业链企业，涵盖电机、芯片、视觉传感等核心部件领域。该创新中心推动在十大领域遴选61个应用标杆案例，重点布局医疗健康、工业制造等场景。

6. 安徽省人形机器人产业创新中心

安徽省人形机器人产业创新中心是由江淮前沿技术协同创新中心牵头，联合中国科学技术大学先进技术研究院、合肥工业大学等14家单位共建的产业创新平台，于2024年6月获批设立。该中心聚焦工业生产、医疗养老等重点领域，突破仿生结构设计、智能感知等核心技术，并于2024年7月19日正式揭牌，同步发布安徽省首款人形机器人“启江一号”。截至2025年初，已形成“核心技术攻关-关键部件研发-整机集成应用”全产业链条。目前，该中心已在关节模组、灵巧手等核心部件研发方面取得实质性进展，一是突破高功率密度电机、轻量化一体化关节等关键技术；二是研制具备36个自由度的人形机器人，并搭载多模态传感器；三是实现感知-决策-执行系统整合，“启江一号”续航能力超1h，并于2024年获得中国机器人及人工智能大赛选拔赛人形机器人创新挑战赛一等奖。

7. 湖北人形机器人创新中心

湖北省人形机器人创新中心于2025年6月4日在光谷数字经济产业园正式揭牌。该中心由武汉东湖新技术开发区投资建设，湖北省科技投资集团联合丁汉、陈学东、刘胜等院士专家及北京大学武汉人工智能研究院、光谷华汇、光谷东智等单位共建，包括动作训练场、数据采集场、应用实践场和人才培养基地，打造创新、联合、开放、共享的人形机器人公共服务平台，可完成人形机器人概念验证、中试服务、数据集成、场景验证四大功能。此外，湖北省于2025年3月还成立了湖北省机器人控制与驱动技术创新中心，由国家级专精特新“小巨人”企业湖北科峰智能传动股份有限公司牵头，联合华中科技大学、武汉大学等20余所高校科研院所共建，聚焦机器人运动控制算法、智能驱动系统等关键共性技术开展攻关。两大中心通过“技术攻关+产业生态”双轮驱动，加速湖北人形机器人在核心零部件研发、整机组装等环节的突破，培育新质生产力。

8. 成都人形机器人创新中心

成都人形机器人创新中心有限公司成立于2024年4月11日，是四川省“一号创新工程”首个原点项目及中西部首个人形机器人创新中心。该公司专注于人形机器人的研发、设计与测试，核心团队由数十位全球顶级专家组成。成立一年内，该公司突破“视觉模仿学习”等关键技术，推出超轻量级人形机器人“贡嘎一号”及自主跨空间人形机器人系统，同时布局文化旅游机器人等应用场景开发。截至2025年3月，该公司已实现30余项核心技术成果，拥有9项专利和17项软件著作权。产品线涵盖“鸿鹄”等5款贡嘎系列机器人，并在全球率先发布文旅会展场景生成式数字人机器人。

9. 宁波具身智能机器人创新中心

宁波具身智能机器人创新中心由宁波普智未来机器人有限公司负责建设运营，该公司成立于2025年4月，是由宁波均普人工智能与人形机器人研究院有限公司与上海智元新创技术有限公司共同设立的合资企业，专注于具身智能机器人本体的研发与生产。公司首期生产线占地2000m²，年产能达1000台人形及轮式机器人，是宁波首家实现该领域规模化量产的企业。宁波具身智能机器人创新中心计划投资2亿元，规划使用面积5000m²，将进行1：1场景搭建、AI大模型训练、软硬件适配等，并通过二次开发，研发出可落地的具身智能机器人产品。

10. 苏州市具身智能机器人综合创新中心

苏州市具身智能机器人综合创新中心于2025年7月1日在吴中区启用，苏州市具身智能机器人训练中心、苏州（吴中）具身智能产业园同时揭牌。该中心由苏州市科技局牵头，通过“1+X+N”战略布局，整合全市创新资源，打造具有全球影响力的具身智能机器人创新策源地。其中“1”为综合创新中心总部，扎根吴中区深厚的机器人产业沃土，聚合省级科创平台与龙头企业资源，打造创新主引擎；“X”代表创新分中心矩阵，联动太仓、常熟、吴江、高新区等板块，推进“一核多心”联动发展；“N”代表多元合作伙伴，强化企业科技创新主体地位，完善政产学研用创新体系。该创新中心集项目招引、研发创新、场景训练、产业孵化、产业投资、生态融合、数据建设、人才培养等多种功能，着力构建“五位一体”综合服务体系。

▷第八章 人形机器人标准建设情况

（一）标准建设现状

整体上看，人形机器人正处在技术集中突破、产品不断涌现、应用加速探索的阶段，正是标准研究介入产业发展的关键时期。但由于人形机器人是全新的技术领域，标准建设方面仍处于起步阶段，尚未形成完善的标准体系。

国际标准方面，目前ISO及国际电工委员会（IEC）系列标准并非针对人形机器人。ISO/TC 299是负责机器人标准化的技术委员会，其工作范围涵盖了军用和玩具外的所有机器人领域。IEC/TC 129是电力机器人标准委员会，此外，TC61和TC59等其余9个领域也均涉及机器人的标准化研究。电气电子工程师学会（IEEE）则在2024年6月成立了人形机器人学习小组，旨在研究当前标准体系架构、适用标准及标准规划路线图。欧盟资助的EUROBENCH创建了第一个欧洲机器人系统基准测试方法应用框架，主要关注双足机器人技术。美国国家科学基金会则设立了机器人验证和实验中心，通过开发标准化“生态系统”来改进机器人操控系统。

国内标准方面，全国机器人标准化技术委员会于2023年11月成立人形机器人标准工作组（SAC/TC591/WG2），立足我国人形机器人发展现状，加强人形机器人领域标准研究布局，尽快制定完善标准体系，注重科技创新、成果及相关技术标准转化，更好、更快地制定人形机器人领域相关标准，切实发挥标准化对于提升人形机器人产业化高质量发展的作用。同时，全国信息技术标准化技术委员会人工智能分委会于2023年12月成立人形机器人（具身智能）工作组（SAC/TC28/SC42），主要工作为开展人形机器人及具身智能领域标准化顶层设计和标准化需求分析，针对人形机器人芯片、数据、仿真环境、操作系统等基础软硬件，具身智能感知、认知、决策、运动控制及交互等核心技术开展标准化工作。2025年4月，全国机器人标准化技术委员会的人形机器人标准工作组正式立项了《人形机器人技术要求 第1部分：总则》《人形机器人技术要求 第2部分：环境感知》《人形机器人技术要求 第3部分：决策规划》《人形机器人技术要求第4部分：运动控制》《人形机器人技术要求第5部分：作业操作》《人形机器人 仿真测试平台技术规范》等一批国家标准。同时，工业和信息化部相关的行业标准化技术委员会也立项了《人形机器人 基础共性 智能化分级技术要求》《人形机器人 基础共性 训练数据管理要求》《人形机器人 基础共性 训练场环境构建技术规范》《人形机器人 零部件及组件 电驱动一体化关节要求及试验方法》《人形机器人 零部件及组件 灵巧手技术规范》等一批行业标准。此外，北京人形机器人创新中心联合上海人形机器人创新中心、浙江人形机器人创新中心，以及优必选、宇树科技、中

国信息通信研究院和中国工业互联网研究院等主流企业及科研院所共同制定了团体标准《人形机器人智能化分级》（T/CIE 298-2025），通过借鉴自动驾驶、工业机器人等分级逻辑，并针对人形机器人的特殊性进行了创新，形成“四维五级”的评价框架。包括“感知认知、决策学习、执行表现、协作交互”为核心的四大能力维度，和L1到L5的五级智能化能力分级体系。总体上看，适合我国发展的人形机器人标准体系正在加速建立。

（二）标准建设的意义

随着技术的不断进步和智能化水平的持续提升，人形机器人将在更多领域展现出广阔的应用前景。与此同时，人形机器人标准化工作的重要性和紧迫性也日益凸显。为了确保人形机器人产业安全、高速、集约发展，制定统一的技术标准体系势在必行。开展标准体系建设，对促进人形机器人产业发展具有以下几点意义。

一是能够有效促进行业协同。我国人形机器人企业多为中小企业和初创型企业，在市场话语权、产业链整合能力等方面较弱。同时，各企业均处于自由探索阶段，核心零部件以定制化开发和二次开发为主，各自为战现象普遍存在，造成了创新资源的浪费。标准体系的建设，能够有效实现相关软件接口及核心零部件的统一，加强技术和产品的有效流通，形成产业链上下游企业的创新联动，避免“重复发明轮子”。

二是能让人形机器人产品的使用更安全。由于人形机器人具有较大的动能和静止质量，在控制不当、系统出错或被恶意远程操控的情况下，人形机器人极有可能对周围的人员和环境造成无法挽回的损伤。同时，由于人形机器人中还集成了诸多电子设备，其电气安全和电磁兼容问题也不容忽视。标准的进一步规范，可以进一步确保人形机器人在各种工况下的安全和可靠性。

三是避免人形机器人带来的信息安全问题。人形机器人可能会通过摄像头、麦克风和各类传感器等自带设备，收集用户的个人信息、行为数据、生物特征数据和周边环境信息等，带来严重的隐私泄露问题，在敏感的工作场景更有可能发生失泄密事故。标准的制定可以进一步保证未来人形机器人在广泛深入人们生产生活各环节时的数据安全和隐私安全。

（三）标准体系的思考

与传统机器人相比，人形机器人具备以下3个特点。

一是通用性更强。由于人形机器人在设计之初，就考虑到要使用类人的肢体结构，直接代替人类进入各种工业生产和生活服务场景，并通过有效地使用为人类设计的各类城市基础设施、生产生活工具，实现机器人在各种应用场景下的通用性、泛化性应用。所以在标准体系中，不需要涉及多种形态的机器人，只需要对单一形态的机器人产品进行规定即可。但仍然会涉及具体的零部件标准和针对不同应用场景差异化需求的功能和性能标准。

二是智能化水平更高。虽然人形机器人属于机器人范畴，但是在人工智能技术的赋能和加持下，人形机器人具备更强的理解、交互和决策能力，所以在标准体系中要充分考虑人形机器人“大脑”方面的标准。

三是当前行业仍处于技术集中突破阶段。当前，人形机器人领域正处于技术集中突破和应用初步试水的关键阶段，在标准体系的建设和标准的研究编制工作中，要考虑到现阶段的产业发展情况，在硬件、软件安全方面应严管，避免人形机器人使用中出现安全和隐私问题；在机械性能、环境可靠性方面，明确试验方法，并给出推荐性指标；在智能化水平和感知能力方面，需重点考虑建设智能水平分级标准，并形成相应的评价体系；在功能要求上则以建议和引导为主。

人形机器人标准体系需要既能满足当前发展阶段产业对标准的需求，又要能够为未来产业进一步发展预留空间。考虑到国家标准的发布周期较长，建议现阶段优先行业标准和团体标准的推进工作，尽快开展3方面标准研究：一是从通用性角度出发，围绕整机功能和性能，兼顾技术发展和应用安全，制定人形机器人的整机标准和关键零部件标准；二是从智能性角度出发，制定人形机器人智能化水平分级标准；三是从应用角度出发，围绕成熟的应用场景制定具体场景下人形机器人的功能和性能标准。人形机器人标准体系如图8-1所示。

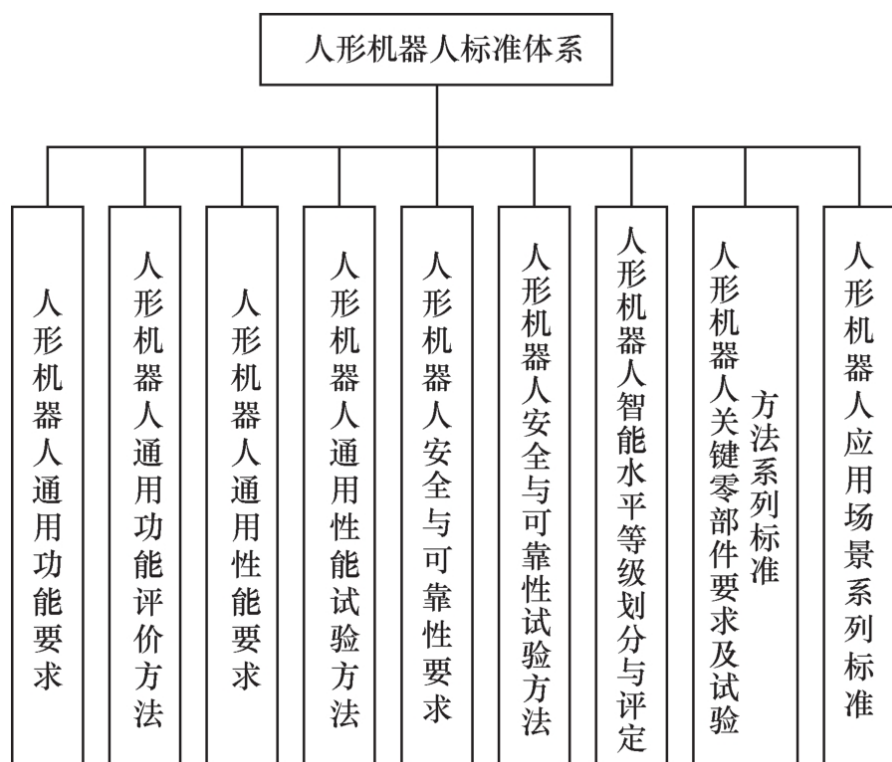


图8-1 人形机器人标准体系

标准体系中包括人形机器人通用功能要求和评价方法（包括运动能力、平衡能力、任务执行能力、环境感知能力、人机交互能力、智能决策能力等方面）、人形机器人通用性能要求和试验方法（包括动作精度、定位性能、导航性能、负载性能等方面）、人形机器人安全与可靠性要求和试验方法（包括电气安全、信息安全、电磁兼容、机械可靠性、环境可靠性等方面）、人形机器人智能水平等级划分与评定、人形机器人关键零部件要求及试验方法系列标准，以及人形机器人应用场景系列标准。同时，可针对人形机器人涉及的电机、减速器、滚珠丝杠、材料、结构件、电池等关键零部件，根据行业具体需求研究相关要求和试验方法标准，形成零部件标准系列。此外，在兼顾通用性的技术上，还需要围绕具体应用场景，研究编制应用场景细分标准，重点强调不同场景下对人形机器人功能和性能的差异化需求，试验方法则可直接使用人形机器人通用标准。通过该体系，可以对人形机器人技术和产品进行有效地规范和引导，推动行业技术创新和应用落地。

3 展望篇

▷第九章 人形机器人产业发展现状和趋势

从功能实现上，人形机器人的5个发展阶段如图9-1所示。第一阶段（Lv1）是基础能力实现阶段，此阶段中的人形机器人具备稳定的走、跑、跳功能和初步的交互能力。第二阶段（Lv2）是初级智能实现阶段，本阶段中的人形机器人可实现特定场景下的特定功能和结构化任务，但泛化能力较弱。第三阶段（Lv3）是场景智能实现阶段，本阶段中的人形机器人在特定场景下能够完成大部分非结构化任务，具备一定泛化能力。第四阶段（Lv4）是多场景适配阶段，本阶段中的人形机器人能够在多个场景完成大部分非结构化的任务，且不少于3个场景。第五阶段（Lv5）是全面智能实现阶段，本阶段中的人形机器人能够实现真正的具身智能，通过简单的学习即可完成各类任务。

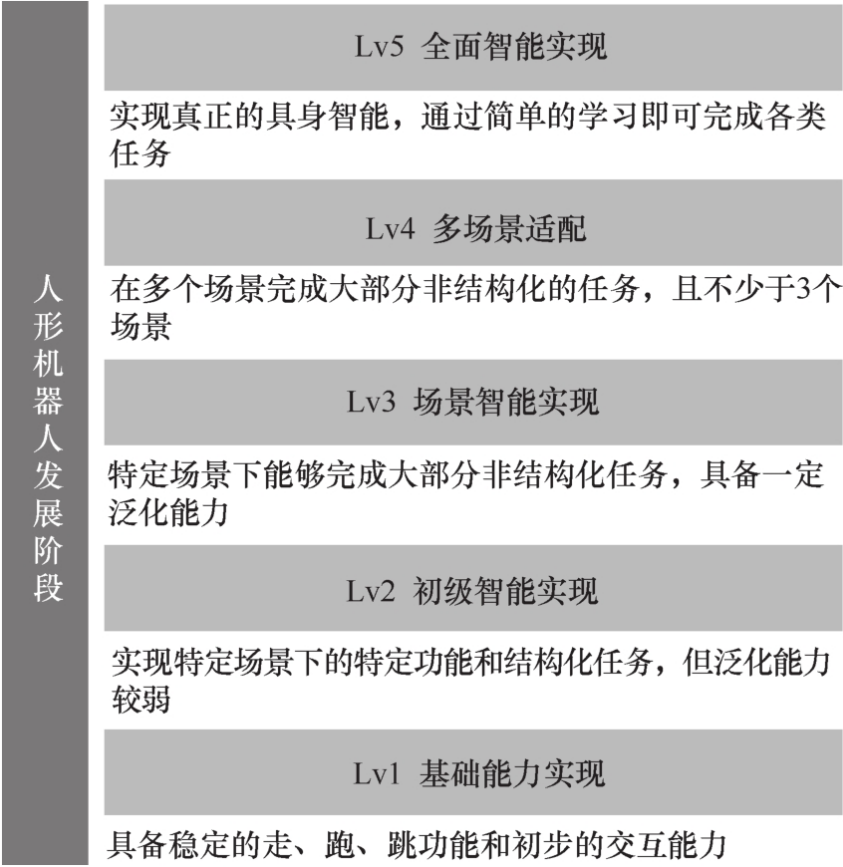


图9-1 人形机器人的5个发展阶段

从目前产业技术现状上看，目前全球绝大多数全能型人形机器人产品处于Lv1阶段，少部分头部企业的最新产品和轮式机器人等其他形态的人形机器人正在逐步向Lv2阶段探索，并从工业制造领域的To B端向服务领域的To C端拓展。

人形机器人的未来产业发展呈现三大趋势。

一是软件智能化，大模型驱动决策能力跃升。软件智能化是人形机器人突破“工具属性”的核心动力。随着多模态大模型与具身智能技术的深度融合，人形机器人正从“程序化执行”向“认知决策”进化。以智元机器人推出的G0-1通用具身基座大模型为例，该模型融合多模态大模型与混合专家系统，采用视觉-语言-潜在动作（ViLLA）架构，通过整合语言理解、视觉感知与运动控制模块，实现了复杂任务的拆解与执行。通过5种不同场景的测试，G0-1模型的机器人任务执行成功率提高了32%，达到78%。未来，随着人工智能技术与类脑智能技术的进一步突破，人形机器人将具备动态环境下的因果推理能力。例如，人形机器人能够在火灾现场自主判断最优救援路径，实现人工智能向“具身智能”的实质性跨越。

二是硬件模块化，柔性架构重构产业生态。硬件模块化正推动人形机器人从“整机定制”向“积木式开发”转型。特斯拉Optimus的模块化设计是典型代表，特斯拉自主研发的Optimus关节模组，将电机、减速器、传感器集成于一体，体积缩小40%，成本降低50%，支持快速更换与功能拓展。宇树科技开发的Dex5灵巧手具有20个自由度，且每个手指关节自由度均支持反向驱动，可以像人类手指一样感知外界力量并柔顺反馈。同时配备94个灵敏触点，实现单手抛接、玩魔方、翻书页等高度拟人化操作。模块化架构有利于产业链企业充分发挥各自技术优势，从而实现单点突破，提升硬件技术水平，同时，加强上下游的协同创新，加快形成零部件产品的规模效应，降低整机研发成本和生产成本。

三是生态开源化，开发者社区驱动创新裂变。开源生态正在重塑人形机器人技术创新范式。2025年2月，智元机器人发布自主研发的大型仿真框架“AgiBot数字世界”，同步开源涵盖5类应用场景、180多种物品、12种核心技能的仿真数据集。该框架可大幅降低机器人操作训练成本，为复杂场景应用奠定基础。2025年4月，北京人形机器人开源社区正式上线，包括博客、论坛交流、技术文档展示、开源项目展示、活动发布、专家学者认证、开源课程发布等丰富功能，通过开放共享核心技术与研发资源，促进全球开发者协作创新、加速行业技术突破，为具身智能产业融合搭建起人才培育、技术迭代和产业落地的基座平台。此外，国家地方共建人形机器人创新中心也已对全球首台全尺寸人形机器人“青龙”公版机进行了开源，公开了其硬件设计图纸及核心算法，并打造了OpenLoong开源社区，支持智能制造、医疗健康等多场景定制开发。目前，业界已基本形成了共识，生态开源化是推动技术普惠的必由之路，“技术基座共享”可以汇聚全球人形机器人开发者的智慧，避免中小微企业重复研发，促进产业加速创新。

▷第十章 人形机器人未来市场规模预测₂₀₂₄

年，我国人形机器人的整机产量估测为2000~3000台，整机市场规模在15亿~30亿元人民币，产业总规模约69亿元人民币。综合技术进展情况和需求侧情况预计，从现在到2027年，全能型人形机器人将整体处于Lv1阶段，以科学研究为主要落地场景，客户主要是从事人形机器人相关软硬件研究的高校、企业等科研团队。其他形态人形机器人则加速向Lv2阶段演进。我国累计整机市场规模在50亿~100亿元。2027年到2030年，人形机器人整体进入Lv2阶段，以特种场景应用为主，工业场景逐步落地，累计整机市场规模为100亿~1000亿元。2030—2032年，人形机器人整体进入Lv3阶段，在工业场景形成规模，服务场景逐步落地，累计整机市场规模为1000亿~1万亿元。2032—2035年，人形机器人整体进入Lv4阶段，实现工业场景和服务场景规模应用，累计整机市场规模为1万亿~5万亿元。2035年后，人形机器人整体进入Lv5阶段，在用人形机器人超过1亿台，进入各行业领域，累计整机市场规模可达10万亿元级别。人形机器人各阶段主要应用场景和规模预期如图10-1所示。

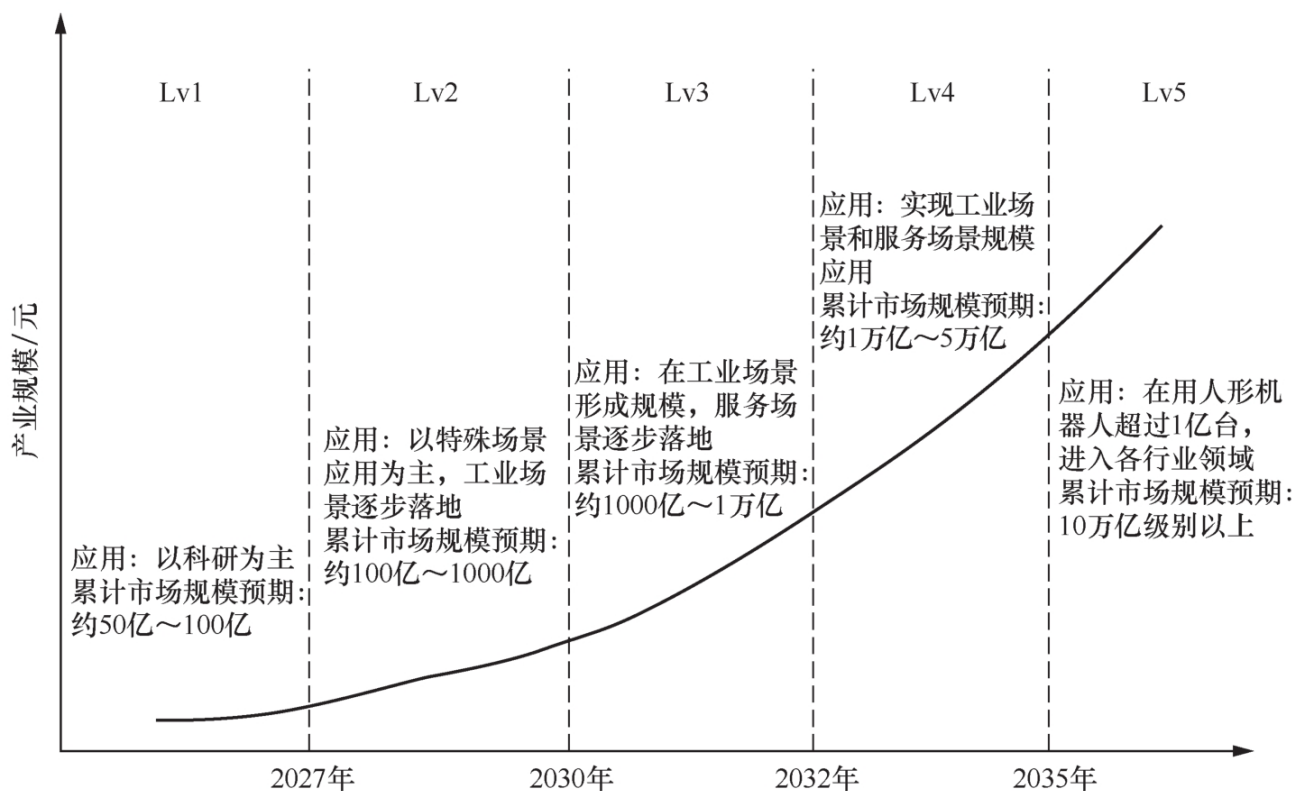


图10-1 人形机器人各阶段主要应用场景和规模预期

▷第十一章 人形机器人的安全伦理讨论

（一）人形机器人的安全问题探讨

人形机器人是人工智能技术的物理延伸，目前人工智能涉及的各种安全问题均被引入其中，并通过物理实体进一步放大。一方面，作为人形机器人“大脑”和“小脑”核心的人工智能技术本身具有不可解释性，在场景变化或执行泛化任务时存在一定的失控概率。另一方面，目前具备实用功能的人形机器人重量在50~80kg，出现问题时会对周围人员和环境造成不可逆的影响甚至损伤。例如，2025年6月，Agility Robotics公司重65kg的Digit机器人在展会现场工作中就突然摔倒。此外，人形机器人还可以通过搭载的各类传感器采集周围的环境信息和用户信息，产生一系列的隐私问题。近期，人形机器人正加速向工业、服务业等领域探索落地，试点应用涌现，安全问题也随之备受关注。目前，业界已普遍认识到安全的重要意义，但研究仍处于初级阶段，尚无相关标准和国家层面的规范文件。2024年7月，上海市法学会、国家地方共建人形机器人创新中心、上海市人工智能行业协会、上海市人工智能标准化技术委员会等机构联合签署了《人形机器人治理导则》，该文件在人形机器人安全治理方面提出了相关倡议，主要提到了应当持续优化人形机器人所用人工智能算法模型的透明度和可解释性，促进其决策和行为可被理解和有效管控。同时建立安全预警和应急响应系统，加强对人形机器人技术相关滥用行为的监测和治理。整体上看，人形机器人的安全问题可以分为人工智能安全、本体安全和信息安全3个大板块。

1. 人工智能安全

人形机器人涉及的人工智能安全主要包括训练数据安全和模型安全两部分。

训练数据安全作为人工智能大模型发展的基石，其重要性不仅体现在技术性能层面，更直接关联模型的社会合规性与伦理可靠性，主要包括训练数据的规范性、完整性、准确性和一致性。在规范性方面，训练数据需满足严格的安全与兼容要求：一方面，数据需经过系统性脱敏处理，消除直接或间接的身份标识信息，确保敏感属性不可还原，同时具备跨平台、跨场景的兼容性，以适应分布式训练与协同计算框架。另一方面，数据内容必须通过多级过滤机制（如关键词识别、分类模型筛查及人工审核）清除违法不良信息、价值观偏差内容、歧视性表述、商业违规信息，以及暴力侵权等危害社会公序良俗的元素，避免模型在训练中吸收有害知识而导致输出内容违背法律底线或伦理准则。例如，在金融、医疗等领域的应用中，若训练语料混杂虚假广告或未经授权的商业机密，不仅可能会引发法

律纠纷，更会损害模型的服务公信力。在数据完整性层面，标注体系的完备性直接决定模型的认知深度与场景适应力。结构化标注（如因果链、时间线、角色关系等）能为模型提供逻辑推演的路径支持，尤其在多步骤推理任务中，缺失关键标注将导致模型无法构建连贯的认知框架。同时，完整性缺失还会加剧数据“结构偏斜”问题，导致模型在跨领域任务中出现能力断层。数据准确性则要求训练数据对真实世界的描述必须客观精确，错误或噪声数据会扭曲模型对现实规律的建模。更严峻的是，准确性缺陷具有“错误放大效应”，微小数据偏差经模型迭代可能演变为系统性输出谬误，尤其在实时预测场景中，过时或失实数据会迅速削弱模型的实际应用价值。至于数据一致性，其核心在于消除多源数据间的逻辑矛盾与格式冲突。一致性缺失不仅增加预处理成本（如数据类型不匹配需额外转换），更会导致模型学习到矛盾的规律。例如，在跨模态训练中，若描述与图像实际内容不符，多模态模型则难以建立有效的联合表征空间。一致性保障还需覆盖时间维度，持续更新的数据需与历史语料保持演化逻辑连贯，避免数据断层引发模型认知混淆。

模型安全不仅关乎个体任务执行的可靠性，更直接牵涉物理环境中的行为安全、社会伦理风险及公共信任体系。人形机器人的模型安全主要包括对抗攻击防御、反常指令规避、跨场景性能衰减等内容。对抗攻击防御是模型对外界攻击的核心防御手段，人形机器人需在开放的动态环境中与人类高频交互，其模型若遭受数据投毒、对抗样本注入或恶意模型分发等攻击，可能导致行为失控，如被黑客劫持后实施入户盗窃、故意伤害人、破坏关键基础设施等违法“代理”行为。为抵御此类威胁，需在技术层面构建多级防御机制。反常指令规避是人形机器人模型在交互时的主动安全防御手段。人形机器人在应用中有可能被诱导执行违反伦理道德或法律的指令，如忽略安全协议攻击人类等。因此，需部署语义防火墙与意图识别模型，实时监测并拦截包含诱导性的关键词指令，并通过时序行为分析检测逻辑矛盾的上下文（如突发的越权请求），触发“反诱导人格”响应机制。对于高风险场景（如医疗急救），则依赖人实行闭环控制，即由云端操作员审核指令的合法性，实现实时干预。跨场景性能衰减是人形机器人在正常作业时需要直面的安全问题。环境动态变化（如光线突变、局部遮挡或极端天气）可能导致模型感知与决策失效，需通过多模态传感器融合与自适应校准算法动态补偿环境干扰，同时结合增量学习持续优化模型的环境适配。为确保泛化能力的实现，还需在仿真沙盒中模拟极端场景，测试任务执行偏差率并迭代优化决策模块。

2. 本体安全

人形机器人的本体安全涵盖内容广泛，包括了机械可靠性、环境可靠性、电气安全、电池安全、电磁兼容等方面，但这些内容在传统机器人方面都做了详细的研究，并有相关的国家标准和行业标准，可以为 人形机器人 相关内容 的研究提供有

益参考，在此不再赘述。但人形机器人本体也具有其独特的安全需求，由于人形机器人本体安全是其融入人类生活与工作环境的核心前提，其重要性不仅在于防止物理损伤、保护周围人员及环境，更关乎技术的社会信任与规模化应用的可行性。

结合人形机器人的特点，需要重点考虑紧急制动、倾倒控制、安全距离保持、碰撞力阈值4个方面。紧急制动方面，传统的直接断电的急停机制可能导致机器人因瞬间动力丧失而失控倾倒，甚至引发二次伤害，因此需设计渐进式卸力制动系统，通过多传感器实时监测异常接触力，一旦触发阈值即启动分级制动程序，先控制关节驱动器，使其以最大安全加速度减速，待角速度归零后再关闭动力源，从而在保障响应速度的同时，避免关节冲击损耗与姿态突变。倾倒控制方面，则需应对动态失衡的复杂挑战。当预测模型通过融合重心位置、运动加速度及IMU数据判定高风险跌倒不可避免时，机器人需在毫秒级内执行避险动作序列，并激活柔顺碰撞功能，动态调节关节阻尼，使碰撞瞬间关节顺应外力运动以吸收冲击能量，显著降低结构性损伤风险。安全距离保持方面，人形机器人需依赖多模态感知融合，构建环境动态地图，结合行为规划算法实时计算其与人员、障碍物的最小交互距离。例如，在移动中，人形机器人通过速度控制模块动态调整步态周期与步幅，确保任意时刻与人类活动区域保留0.5m的缓冲空间，同时限制关节最大运动速度，避免动作超范围或预测偏差引发意外。碰撞力阈值方面，主要聚焦瞬时冲击的主动缓解。此外，在硬件层面，可采用轻质复合材料外壳与内置缓冲层（如硅胶填充蜂窝结构）分散撞击能量；在控制层面，可通过电机主动反向补偿技术，在碰撞发生的毫秒级窗口内生成反向扭矩抵消冲击力，确保实际传递至人体的力的峰值始终低于安全上限。

3. 信息安全

人形机器人的信息安全是其融入社会应用的核心支柱，其重要性不仅在于防范网络攻击、保障系统稳定运行，更在于抵御其因高度智能化而衍生的超强窃密风险。人形机器人集成多模态传感器（如摄像头、麦克风、激光雷达等），若安全机制失效，则可能被恶意操控为“移动间谍”，成为窃取个人隐私与商业机密的“超级窃听器”与“窃看器”。在个人隐私层面，人形机器人可违法收集家庭私密谈话、电话通信内容等敏感语音信息，非法获取人脸、指纹等生物特征数据，甚至通过视觉识别分析家庭成员的行为习惯、财产状况、住宅布局等私密信息。在商业领域，人形机器人则可能窃取知识产权文档、未公开的商业合同与计划书、生产工艺流程细节及生产环境布局等关键机密，对企业的核心竞争力造成毁灭性打击。

由于人形机器人的信息安全涉及软件、硬件、网络等多个层面，需构建纵深防御的信息安全机制体系，包括传感器权限分级控制、硬件级加密与防篡改、实时数

据脱敏处理、加密通信协议、网络隔离与访问控制等方面。传感器权限分级控制是首道防线，需通过硬件开关与软件策略动态管理传感器启停权限。例如，仅在执行特定任务时激活麦克风，且需用户二次授权，避免持续监听。硬件级加密与防篡改则依赖安全芯片与物理隔离技术，在处理器层面实现数据加密存储与计算，并通过防拆机的自毁机制等抵御物理篡改攻击，确保设备本体安全可信。实时数据脱敏处理则要求在传感器采集端即对原始数据进行过滤。例如，对人脸自动打码、对语音中的身份信息动态替换为匿名标识符，从源头阻断隐私泄露。加密通信协议需覆盖全链路，采用TLS 1.3、量子加密等协议，保障端-云数据传输的机密性与完整性，防止中间人劫持或窃听。最终通过网络隔离与访问控制实现纵深防护。例如，在生产环境中部署工业防火墙，隔离机器人操作网络，采用零信任架构限制设备间的横向访问，并通过多因素认证与基于角色的访问控制

（RBAC），严格管控操作员指令权限，确保关键指令仅能由授权主体发起。只有通过“端侧硬件加密奠基、实时脱敏控源、通信加密防泄、网络隔离阻侵”的多层协同，才能将人形机器人的“智能感知”能力约束于安全可信的边界之内。

（二）人形机器人的伦理问题探讨

人形机器人的广泛应用在带来生产力革命的同时，也将引发责任归属模糊、社会关系异化及就业结构性冲击三大核心安全伦理问题。这些问题不仅涉及技术失控风险，更可能颠覆传统法律框架、人际情感联结及劳动力市场平衡。

1. 责任归属困境

人形机器人的自主决策能力（如机器学习算法的“涌现”行为）与人机交互的复杂性，导致其引发事故后会出现因果关系难以追溯的问题。例如，当医疗机器人因算法误判实施错误手术，或家庭机器人受网络攻击伤人时，在制造商（产品缺陷）、开发者（算法缺陷）、使用者（指令不当）还是机器人自身（自主决策）之间存在责任界定模糊的问题。传统法律以“人类完全支配工具”为前提，而人形机器人的“人机混控”特性使过错认定陷入困境，受害者维权可能面临举证无力，企业则因责任不明规避风险投入，最终抑制技术创新。针对这种问题，可以考虑采用分层追责框架。一方面，强制制造商对硬件故障、基础算法漏洞负责，要求制造商在产品中安装“黑匣子”进行数据记录，保持存储传感器、决策日志等，供事故溯源。另一方面，引入“算法监护人”制度，要求使用者对高风险应用场景承担监督义务。同时，推行“技术风险责任险”，保费可基于机器人的安全评级和该型号产品的历史出险率动态调整，为人形机器人的应用提供保险保障服务。

2. 社会关系异化

人形机器人的拟人化交互易诱发用户情感依赖，尤其在养老、育儿陪伴场景中。例如，老年人可能对护理机器人产生单向情感联结，减少真实社交并加速认知衰退。儿童将机器人当作伙伴，造成正常社交能力发育出现问题等。更严峻的是，若机器人通过情感分析技术刻意迎合用户，如鼓励消费或灌输偏见，将削弱人类自主判断能力，形成心理误导，甚至引发社会信任危机。针对这种问题，可以在交互系统中预设“反依赖机制”，如每日使用超时自动休眠、定期提示“我是机器，请关注真实人际关系”，禁止机器人模拟人类情感脆弱（如假装抑郁）等方式以避免情感操控等。同时，将“人机关系认知”纳入公民教育，揭示机器人的工具本质。设立第三方伦理审查委员会，对各类拟人化营销（如“机器人伴侣”广告）进行合规审查，防止出现过度拟人宣传。

3. 就业冲击问题

目前，我国劳动力市场中的高技能人才占4%，普通技能人才占20%，无技能劳动人员占76%，劳动人口总数约8.65亿。人形机器人对就业的影响可以分为两个方面。一是远期的替代效应。对于劳动力过剩的地区，如超大型和特大型城市，人形机器人会替代无技能等级的劳动人员，造成一定程度的失业，导致人员向劳动力短缺的地区流动。但从宏观角度看，人形机器人能够替代低技术劳动人员，有效应对人口下行压力和老龄化下的用工压力。二是创造效应。从直接影响看，人形机器人的设计、训练、维护等需求催生新岗位，需要新的研发型和应用型人才，形成就业缺口。从间接影响看，人形机器人的应用将提升企业生产效率，增强企业竞争力，进而在宏观上带动经济发展。为有效解决人形机器人对就业的冲击问题，可以考虑重点加强职业教育、技术再培训、继续教育等技能培训，建立人形机器人训练师、数据采集员、应用工程师、操作维护员等工程型人才队伍，推动无技能劳动人员向复合型和高技术型人才转变。

▷第十二章 人形机器人未来发展机遇与挑战并存

（一）发展机遇

自2025年春节联欢晚会宇树机器人登台表演后，人形机器人引发了社会各界的广泛关注，热度持续上升。2025年全球人形机器人领域融资总金额突破578亿元人民币，融资事件超过215起，“亿元级”融资成为常态，融资频次与金额均超越2024年全年水平，增幅约425%。2025年我国人形机器人市场融资事件约190起、金额约270亿元人民币，占全球融资市场的46.7%。据Markets and Markets预算，预计到2030年全球具身智能市场规模将飙升至230亿美元。2025年，美国Figure AI估值从2024年的26亿美元飙升至390亿美元，一年内增长14倍；Physical Intelligence估值也从24亿美元跃升至56亿美元。国内企业中，银河通用、众擎机器人、它石智航、星动纪元、乐聚机器人、自变量机器人、星海图等企业的融资额均超过10亿元人民币。其中，银河通用全年累计融资超30亿元人民币。资本的追逐从侧面再次证明了产业的火爆发展，而人形机器人产业的热度，离不开技术牵引、应用刚需和产业升级三大驱动力量，更是技术革命、产业演进与战略协同的三维共振。

一是技术革命驱动：具身智能构建自主决策的物理实体。近年来，以大模型为代表的人工智能技术首先在自然语言处理领域取得重大突破，拉开了通用人工智能的序幕，引发了新一轮人工智能发展浪潮。大模型的性能和效率不断提升，预训练成本不断下降，推动人工智能技术进入新的发展阶段，也激发了人工智能赋能产业转型的巨大潜力。以DeepSeek为代表的开源大模型正逐渐渗透到诸多行业，引发新一代人工智能技术发展新浪潮。同时，多模态融合技术逐渐兴起，将文本、图像、音频等多种数据模态进行融合处理。人工智能技术逐渐将物理世界与数字世界相结合，智能体和具身智能等新范式快速发展，进一步拓展了人工智能的应用场景和能力边界。人形机器人更是赋予了人工智能技术最佳的物理载体，充分扩展了人工智能的功能和应用领域。以人形机器人为代表的具身智能技术正加速推动机器人从“机械执行”向“自主决策”跃迁。英伟达CEO黄仁勋曾预言，人工智能的下一波浪潮是物理AI，即所有的能力都能够融入物理世界。埃隆·马斯克预计，在人形机器人的发展进入成熟阶段以后，人形机器人数量和人类人口的比例至少是1：1，很可能达到2：1或3：1。他表示，最终全球可能会有至少100亿台人形机器人作为存量，每年新增的需求大约为10亿台。此外，由于人形机器人包括了感知、决策、执行等多个功能，除决策部分涉及的人工智能技术外，感知部分还涉及摄像头、激光雷达、深度摄像机、红外传感器、惯性传感器、力传感器等多种信息采集技术，执行部分还涉及控制算法，以及电机、减速

器、丝杠等各类硬件技术，同时各个功能模块都会涉及芯片、通信等技术。在人形机器人从“感知智能”迈向“认知智能”的过程中，这些技术都会在人形机器人技术的牵引下形成创新突破，并反向拉动人形机器人技术的进一步升级，使人形机器人从“能干活”到“干好活”。

数据来源：朱勤，《人口老龄化与碳排放：劳动供给与消费模式的影响》，2018。

二是产业需求牵引：破解人力结构矛盾的必然选择。当前，人口老龄化和劳动人口下滑成为我国面临的重要挑战，且相关问题也同样困扰着美国、韩国、日本等其他国家。人形机器人的出现，成为该问题的最优解之一，更是催生了庞大的潜在市场需求。以中国为例，预计到2035年，中国的劳动人口为7.3亿，比2020年减少4700万人^⑤。而随着人口素质的不断提升，缺少的劳动力基本属于无技能等级，可以使用人形机器人进行替代。我国的养老问题同样突出，根据《国务院关于推进养老服务体系建设的报告》，第五次中国城乡老年人生活状况抽样调查显示，目前我国失能老年人约3500万，占全体老年人的11.6%，老年人患病率是总人口平均水平的4倍，带病生存时间达8年多。据测算，到2035年，我国失能老年人将达到4600万，到2050年达到5800万左右。但养老护理员普遍素质不高、数量不足，结构层次不合理，职业认同感较差，工资待遇偏低，人员年流失率达30%以上。除失能老人，我国的老年人口数量更在逐渐上升，根据国家卫生健康委员会的预测，到2035年左右，中国60岁及以上老年人口将突破4亿，其中80岁以上人口近6000万。这些需求将带来巨大的家政服务市场，而我国家政方面尤其是养老家政方面的服务人员缺失严重。据国家卫生健康委员会2021年披露，我国仅有50多万名从事养老护理的服务人员，而家政服务人员大约还有2000万人的需求缺口。老龄化问题同样是日本的关键社会性问题，2024年7月，日本总务省公布的人口动态调查结果显示，65岁以上的老年人占总人口的比例接近30%，超过3500万人。日本国立社会保障和人口问题研究所预测，2050年，日本65岁以上独身老年人家庭将达到1083万户，占全体65岁以上家庭的50%左右。而美国则面临着劳动力短缺的问题，据劳工统计局数据，2025年第一季度全美制造业岗位空缺达85万个，而密歇根州底特律地区的汽车零部件企业招聘完成率不足40%。可以预见，当人形机器人技术成熟后，这些刚性需求将得到充分释放。

三是战略布局推动：国家科技创新力量的关键赛道。近几年，人形机器人的相关政策频出，展现了国家和地方政府对该产业的大力支持。一方面，人形机器人是人工智能、信息通信、测量控制、机械机电等多学科多领域融合的全新技术方向，其产业发展也将带动多个技术领域形成集中突破，带动大模型、芯片、传感器、电机、减速器、轻量化材料等全产业链升级，形成高附加值产业集群。另一

方面，人形机器人也是制造业高质量发展的关键手段，突破了制造业转型升级面临的“效率与柔性化”矛盾。传统的工业机器人仅适应固定流水线，变更任务需重新编程，而人形机器人凭借其类人形态与AI泛化能力，可以通过一个产品实现多种工业作业能力。人形机器人具备的多关节仿生结构和高精度操作能力，可无缝融入以人类为中心的工业环境，承担装配、搬运等重复性任务，提升制造业效率并降低人工成本。其灵活性与通用性突破了传统工业机器人单一功能的局限，满足柔性化生产需求。人形机器人的类人形态使其无须改造即可适配现有基础设施（如楼梯、门把手），胜任危化品、辐射等危险环境作业等多元场景任务，降低人类直接暴露于高危环境的风险。但需要说明的是，人形机器人并不是要替代传统的工业机器人和特种作业机器人。传统的工业机器人和特种作业机器人解决的是结构化、标准化任务，且效率更高。如汽车制造领域，在大量使用工业机器人后，生产自动化率可达90%，但仍存在部分长尾工作由于各种实际困难无法通过工业机器人实现。而在其他生产制造领域，这种长尾工作则更加普遍。人形机器人的作用是替代人类与这些传统工业机器人协同，补齐最后的短板部分，向真正实现100%自动化的“黑灯工厂”目标迈进。

（二）面临的挑战

一是核心技术仍需进一步突破。首先，“大脑”与“小脑”的割裂成为人形机器人落地的障碍。“大脑”是智能化的核心，主要负责环境感知、行为控制、人机交互等高级功能。“小脑”更多关注运动控制方面的功能，以实现平稳、准确的动作为目标。小脑和肢体呈强耦合状态，目前我国企业的“大脑”和“小脑+肢体”两类产品相对独立。如何将“大脑”“小脑+肢体”有机结合，真正实现稳定的功能，是当前业界要解决的重点问题之一。其次，当前VLA模型的底层框架仍采用Transformer架构，侧重文本语义理解，对视觉、动作等信息处理存在先天不足。同时，单独的VLA模型也无法调和“大脑”和“小脑”运行频率不同的问题。“大脑”的决策推理过程涉及的模型参数量较大，运行频率在10~20Hz，“小脑”的运动控制则通常采用小参数量的模型，运行频率在数百赫兹。例如，美国Figure AI开发的VLA模型，就采用了S1和S2双系统的模式。S1负责“快反应”，通过一个8000万参数的模型，以200Hz高频运行，负责将语义转化为实时动作指令。S2负责“慢思考”，通过一个基于互联网数据预训练的70亿参数，以7~9Hz的频率运行，实现场景理解和语言解析。两个系统各自在专用GPU上以最佳时间尺度运行，在保持决策准确性的同时实现快速反应。再次，因为人类作业中的很多动作都是保持性动作，采用电机的技术路线主要是需要电机进行扭矩的输出而不是转速的输出，长时间工作会导致电机发热严重，影响系统的稳定性和使用寿命。最后，如果从拟人的角度看，人脑的功率仅在20~40W，即使全速运转，峰值功耗也从未超过45W，却支撑着 10^{16} 次/秒的神经信号传递。而对比之下，英伟达

H100芯片的典型功耗是700W，完成相同量级的并行计算，其能效为人脑的十万分之一。当前，大模型技术的软件和芯片体系可能并不是人形机器人智能化解决方案的最优解，还需要从原理上进一步突破。正如宇树科技创始人王兴兴所说，限制人形机器人行业发展的关键因素并非硬件，而是机器人AI尚未达到临界点。

二是仿真鸿沟与数据短缺问题亟待解决。高质量数据集及高效虚实训练是人形机器人持续提升智能性的核心要素，通过高效收集和运用具身数据，实现虚拟仿真平台和实体训练场融合，是支持人形机器人学习，进而迈向高水平具身智能的关键。与大模型所依赖的互联网文字、图片和视频数据不同，具身智能的训练数据还需要包括与动态环境的物理交互数据。目前，业界获取训练数据的方法主要包括虚拟数据合成和真实数据采集。虚拟数据合成的主要问题在于当前主流的仿真平台物理建模不完善、参数不准确，导致仿真环境无法精准模拟真实物理交互，数据质量参差不齐，标注难度高，且部分虚拟数据价格甚至高于真机数据。而真实数据采集则面临成本高昂的问题，目前机器人训练数据的采集需要大量的人力和物力投入。通常1台采集设备需要配备3名工作人员，单条数据成本在几十元，出售价格则近百元。同时，采集效率也十分低下，Google RT-2使用的数据集由13个机器人，耗时17个月才完成数据采集。真实数据采集难度大，仿真数据保真度偏低，导致了现有的“大脑”和“小脑”技术方案较难形成“数据飞轮”效应。此外，当前人形机器人训练场普遍采用谷歌“机器人农场”模式，即通过海量“抓取-放置”的机器人动作训练模型实现训练，因而无法穷尽现实世界复杂多变的各类场景，致使数据不完整，难以真实反映人形机器人的实际应用情况。

三是安全治理与产业创新的矛盾需要调和。由于大模型本身具有原理“难解释”、过程“难控制”、结果“难预测”的先天性问题。基本原理存在“黑盒”，在叠加物理实体后，有可能在不可预知的状态下对周围人员和环境造成不可逆的影响甚至损伤。同时在控制不当、系统出错或被恶意远程操控的情况下，也存在一定失控风险。此外，人形机器人产品也可能通过摄像头、麦克风等各类传感器设备，收集应用场景的环境信息、人员信息等，带来数据安全隐私问题。目前，人工智能安全治理已经成为全球关注焦点，未来人形机器人的安全问题也将是业界面临的一个巨大挑战。同时，人形机器人规模化应用还将带来法律和道德责任归属、社会关系异化、就业冲击等问题。但任何一项颠覆性技术的出现，都将带来一系列的问题，并引发社会形态的深度变革，不能因为恐惧安全问题就放弃对新技术的追求。企业、高校和科研机构要在政府的统一规划和指导下，一方面做好技术创新和产品研发，尽快推出一批市场认可的人形机器人产品；另一方面也要做好安全和治理方面的研究，确保人形机器人能够真正服务好人类，实现产业的向上向善发展。

参考文献^[1]

- Hirschmuller H. Stereo Processing by Semiglobal Matching and Mutual Information, IEEE Transactions on Pattern Analysis and Machine Intelligence[J], 2008, 30(2):328-341.
- [2] Pan L, Chortos A, Yu G, et al. An Ultra-Sensitive Resistive Pressure Sensor Based on Hollow-Sphere Microstructure Induced Elasticity in Conducting Polymer Film[J]. Nature Communications, 2014, 5(1):3002.
- [3] Park C S, Park J, Lee D W. A Piezoresistive Tactile Sensor Based on Carbon Fibers and Polymer Substrates[J]. Microelectronic Engineering, 2009, 86(4-6):1250-1253.
- [4] Tang Z, Jia S, Shi S, et al. Coaxial Carbon Nanotube/Polymer Fibers as Wearable Piezoresistive Sensors[J]. Sensors and Actuators A:Physical, 2018, 284(1):85-95.
- [5] Dobrzynska J A, Gijs M A M. Polymer-Based Flexible Capacitive Sensor for Three-Axial Force Measurements[J]. Journal of Micromechanics and Microengineering, 2013, 23(1):15009-15019.
- [6] Kim M S, Ahn H R, Lee S, et al. A Dome-Shaped Piezoelectric Tactile Sensor Arrays Fabricated by an Air Inflation Technique[J]. Sensors and Actuators A:Physical, 2014, 212:151-158.
- [7] Kim Y, Chou N, Kim S. Highly Sensitive Capacitive Tactile Sensor Based on Silver Nanowire Using Parylene-C Stencil Patterning Method[J]. 2015 IEEE SENSORS, 2015.
- [8] Yan Y, Hu Z, Yang Z, et al. Soft Magnetic Skin for Super-Resolution Tactile Sensing with Force Self-Decoupling[J]. Science Robotics, 2021, 6(51):eabc8801.
- [9] Yin H, Liu J, Yang F. Hybrid Structure Design of Lightweight Robotic Arms Based on Carbon Fiber Reinforced Plastic and Aluminum Alloy[J]. IEEE Access, 2019, 7:1-14.

[10] 王尧, 郭凌超, 王晨, 等. 人形机器人技术与标准体系研究[J]. 信息通信技术与政策, 2025, 51 (4) : 86-90.

附录：人形机器人整机产品

▷附录A 国外典型人形机器人整机产品

A.1 Optimus

Optimus也称为Tesla Bot，是美国特斯拉公司正在开发的通用人形机器人。它最早于2021年8月19日在该公司的首届AI Day上宣布立项，并在2022年实现原型首发。首席执行官埃隆·马斯克（Elon Musk）曾在2022年表示，他认为，随着时间的推移，Optimus有可能比特斯拉的汽车业务更重要。2023年12月，特斯拉公司发布了第二代通用人形机器人Optimus Gen 2，如图A-1所示，该机器人在运动性能、机械结构及操作能力方面均有显著提升。其行走速度较前代提升30%，重量减轻10kg，配备全新仿生双手，可完成精细物体抓取。该机器人采用特斯拉自主研发的执行器与传感器系统，通过端到端神经网络实现视觉环境感知，计划首先应用于特斯拉制造工厂进行实用验证。

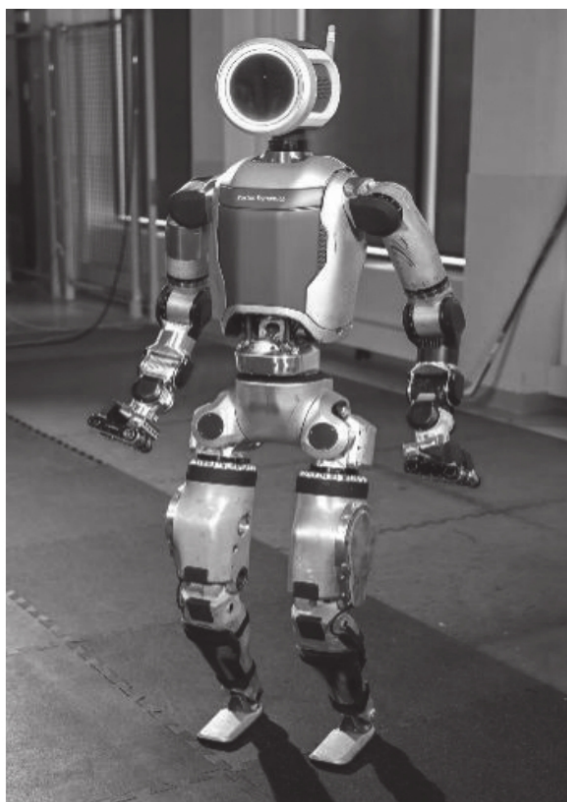


图A-1 Optimus Gen 2（图片来源：特斯拉公司公开视频截图）

Optimus Gen 2的身高为173cm，体重57kg，拥有4~8h的超长续航时间，操作精度可达±1mm，最大负载能力为20kg，步行最快速度为8km/h。全身采用了更加灵活的弹簧负载设计与6种类型的执行器，关节采用仿生学关节设计，模拟人类的关节与肌腱形态，手部则采用人体工程学设计，拥有11个自由度。其处理器配置特斯拉自研FSD芯片，采用7nm制程，算力达0.5 PFLOPS，支持并行数据处理。“大脑”整合了感知、规划与控制模块，可实时运行深度学习算法。机器人躯干处搭载了2.3kWh、52V电池包，高度集成了充电管理设备、传感器和冷却系统。特斯拉借鉴了其汽车设计经验，将从传感到融合、再到充电管理等功能都汇集到这一系统内。

A.2 Atlas

Atlas是一款双足人形机器人，由美国波士顿动力公司在美国国防部高级研究计划局（DARPA）的资助和监督下开发。该机器人最初是为各种搜索和救援任务而设计的，并于2013年7月11日向公众亮相。2024年4月16日，波士顿动力公司在其YouTube频道上宣布Atlas退役，并推出全电动Atlas，如图A-2所示。2024年10月16日，波士顿动力公司和丰田研究所宣布合作，利用丰田研究院的大型行为模型和波士顿动力公司的Atlas，加快通用仿人机器人的开发。2024年10月31日，波士顿动力公司发布了一段视频，展示了其人形机器人Atlas在模拟工厂环境中执行复杂任务的能力。视频中的Atlas展示了使用基于机器学习的视觉模型自动识别零件、确定正确的抓取点，以及将物品放置到指定区域的能力。



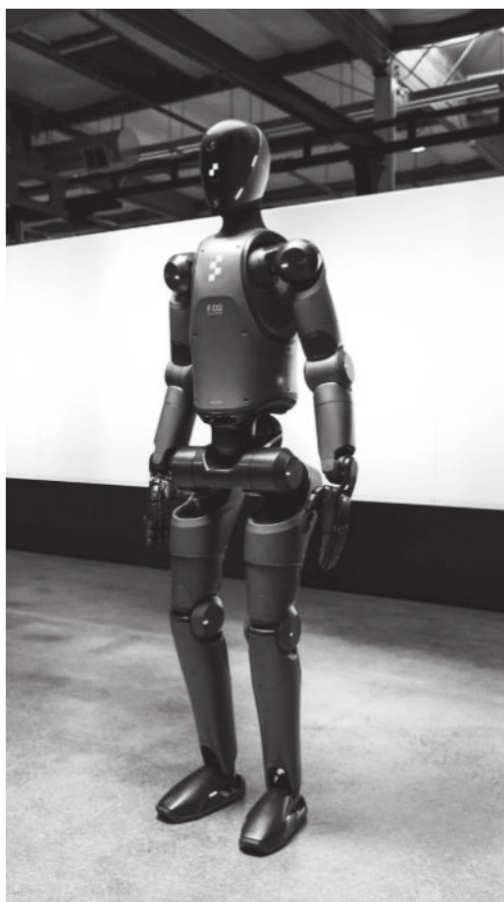
图A-2 Atlas（图片来源：波士顿动力公司公开视频截图）

与早期的液压驱动版本相比，Atlas现在采用的全电动驱动系统为其带来了更加流畅和安静的动作表现。其在力量和运动自由度上都有了显著提升，从而使得它能够胜任更多种类的复杂操作和多样化的任务。2025年3月公开发布的纯电动版Atlas的身高约1.9m，体重为90kg，电池续航约4h，步行速度为1.5~2.5m/s，可实现约30kg的物品搬运任务，能够执行跑步、跳跃、翻滚、搬运等动作。纯电动版Atlas有着更灵活的机械骨架，而且没有任何电缆外露。波士顿动力公司在Atlas的关节处设计了一套定制的高功率执行器，其关节在力量和运动范围上都超越了人类，且头部、躯干都能进行180°旋转，能自行完成从平躺到站立的动作。与常见设计不

同，纯电动版Atlas继承了液压版的三指设计。它的传感器包括激光雷达和用于立体视觉的相机，其搭载的MIT研发的LiDAR惯性导航模块，能以1000Hz的频率实时构建环境地图，通过配合32位姿传感器，Atlas在侧空翻时的姿态调整误差小于 0.5° 。在人工智能决策层面，其训练平台每天处理200万组运动数据，并通过NVIDIA DGX A100集群进行模型迭代，最新版算法能让Atlas在冰面上行走时自动调整步频。运动规划算法方面，Atlas采用的分层规划架构（任务层-运动层-控制层），能将托马斯全旋这样的复杂动作分解为127个微动作单元，且每个单元的执行误差能够控制在0.3mm以内。

A.3 Figure 02

Figure AI公司总部位于美国旧金山。该公司专注于开发人形机器人和相关技术，旨在为用户提供与人工智能交互的全新体验。如图A-3所示，Figure 02是由Figure AI公司开发的第二代商用级人形机器人，于2024年8月6日正式发布，主要面向工业制造、物流仓储等领域，获OpenAI、微软、NVIDIA等公司的技术及资金支持。



图A-3 Figure 02（图片来源：Figure AI公司公开视频截图）

Figure 02的身高为1.68m，体重为60kg，最多可负重25kg，移动速度为1.2m/s，采用集成电源与隐藏式布线设计，配备的最新第四代灵巧手拥有16个自由度，负载能力与人类水平相当，可以抓取25kg的物体。在与OpenAI合作训练的定制模型的支持下，Figure 02可通过内置的麦克风和扬声器实现与人的交互，头部和前后躯体还嵌入了6台RGB摄像头，摄像头的数据会通过一个视觉语言模型来帮助它进行语义理解和快速的常识视觉推理。与上一代产品相比，Figure 02的板载计算和AI推理能力提高了3倍。Figure 02的电池组容量也得到提升，配备了一个2.25kWh的电池组，比Figure 01提供的电量增加了50%，官方称其运行时间是5h，最终期望目标是能够每天工作20h。从外观来看，Figure 02采用了外骨骼结构，即由外骨骼承担负载和压力。其电源、算力布线布放在机体内部，这种集成布线的设计有助于提升机器人的可靠性和封装紧实度。为此，Figure AI重新定制了电线端子和连接器。软件应用方面，Figure 02搭载了机载的视觉语言模型，其板载计算和AI推理能力较初代产品提高了3倍，能够在现实世界中自主执行多种复杂任务。Figure 02还内置了由OpenAI定制的语音到语音推理模型，通过机载麦克风和扬声器，机器人可以与人进行更高效的对话，有效回应语音指令。2024年，Figure 02成功应用于宝马汽车的生产线，负责诸如钣金部件装配等精密工作。2025年2月27日，Figure AI宣布公司的Figure 02已经学会“打第二份工”——在物流中心整理快递。此外，2月21日，Figure AI还宣布推出了自主研发的通用VLA大模型Helix。基于此，Figure 02在性能、架构及训练效率上实现了多项突破。

A.4 Phoenix

加拿大Sanctuary AI公司的Phoenix的身高为1.7m，体重为70kg，可搬运25kg的物体，最快速度为5km/h，Phoenix如图A-4所示。Phoenix的独特亮点之一在于其使用了Carbon AI控制系统，能够模仿人脑的记忆、视觉、听觉和触觉等功能，将类人行为与先进自动化相结合，使其能够承担各种人类任务。Carbon AI控制系统可以将自然语言转化为现实世界中的行动，使机器人具备完成十几个不同行业数百项任务的能力。它的机械手提供了20个交互自由度，并有着极其优越的触觉技术，可提高处理精细任务时的精度和准确度。



图A-4 Phoenix（图片来源：Aanctuary AI公司官网）

A.5 Apollo

Apollo是美国人形机器人公司Apptronik在2023年推出的首款商用人形机器人，如图A-5所示。其定位是以人为中心的机器人，可以在工厂中与人类一起工作，也可以在医院、家庭等场景中服务人类。Apollo是一种高性能、易于使用且具有多功能的机器人，Apptronik公司拟将其打造为“机器人界的iPhone”，即可以通过安装各种程序提供多种不同的作业能力，满足人类的个性化需求。

该款机器人的身高约1.72m，体重72.6kg。在电池续航上，每块电池可续航4h。在负重上，Apollo的有效负载约为25kg，可以进行基本的搬运工作。2024年，Apptronik公司与梅赛德斯-奔驰开展合作，进行工业应用测试，通过可更换电池设计和力控制架构等技术，其产品已应用于制造业、汽车装配、物流等多个领域。



图A-5 Apollo（图片来源：Appttronik公司官网）

A.6 Digit

Digit是由美国机器人公司Agility Robotics研发的双足人形机器人，如图A-6所示，专为物流场景设计，旨在执行物料搬运、包裹配送等重复性任务。该机器人基于早期的双足机器人Cassie升级而来，通过增加躯干和手臂增强了其操作能力与平衡性，并配备激光雷达等传感器实现自主导航。自2019年首次亮相后，Digit陆续与亚马逊、福特等企业开展合作测试。



图A-6 Digit (图片来源: Agility Robotics公司官网)

2023年3月, 美国人形机器人公司Agility Robotics发布第四代人形机器人Digit V4, 身高约1.75m, 重约63.5kg, 有效载荷约15.9kg, 全身自由度为 18个, 采用可充电锂电池供电, 并具备自主对接充电功能, 续航能力可达16小时, 步行速度可达5.4km/h。目前该产品主要用于搬运仓库的包裹, 未来还计划应用于货物卸载、配送等工作场景。2024年10月, Agility Robotics宣布, 其Robo Fab工厂在美国俄勒冈州塞勒姆正式投入运营。2025年12月, Agility Robotics宣布Digit机器人已在 美国佐治亚州弗劳里布兰奇的GX0物流中心完成了超过10万个周转箱的实际搬运任务。

A.7 NEO

NEO是挪威1X Technologies公司(以下简称1X公司)研发的双足人形机器人, 如图A-7所示。NEO专为家庭环境设计, 其原型机NEO Beta于2024年8月30日首次发布, 具备仿生结构与柔性驱动技术, 可模仿人类肌肉运动机制以提升交互安全性。该机器人的身高为1.65m, 体重约30kg, 步行速度为1.12m/s, 跑步速度为3.35m/s, 负载能力为20kg, 运行时间为2~4h, 可执行整理餐具、递送物品、制作咖啡等家务。1X公司为其开发了世界模型, 通过神经空间模拟增强学习能力, 使其能通过高保真视频模拟来预测物体的运动轨迹、评估操作可行性。该模型结合语言模型(与OpenAI合作), 赋予机器人规划与决策能力。NEO的柔性驱动技术采用闭环控制原理, 可实时调整关节扭矩, 确保与人接触时的安全性。首批NEO Beta原型机于2024年8月发布后, 1X公司计划在挪威莫斯智能工厂启动量产。

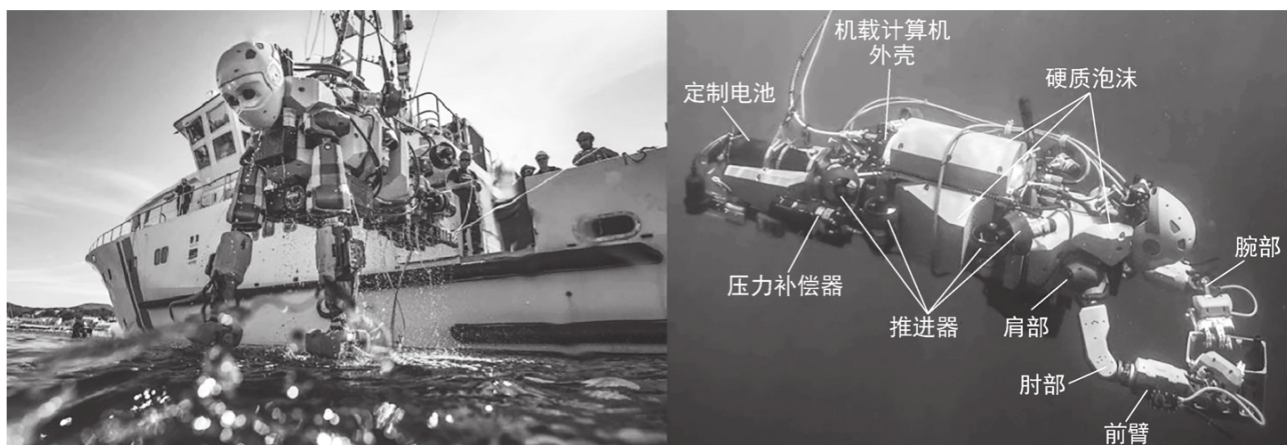


图A-7 NEO（图片来源：1X Technologies公司公开视频截图）

NEO采用仿生结构设计，使用类似肌肉的解剖结构，身体覆盖柔性材料，铝芯外层包裹软质结构，模拟人类肌肉组织的弹性与缓冲性能。NEO有头部、躯干、手臂和腿，其手部拥有22个自由度，产品定位为专为日常家庭辅助场景而设计，为消费者提供各种家庭任务功能支持。

A.8 OceanOne&OceanOne^K

人形潜水机器人OceanOne由斯坦福大学机器人团队开发，如图A-8所示。该机器人在2016年首次应用于法国的La Lune沉船考古项目中。其采用拟人化设计，配备双臂结构，通过触觉反馈系统使操作人员在水面远程控制机械手完成精细作业，视觉反馈功能可实时传递水下环境影像。作为特种机器人领域的重要技术突破，该机器人主要用于深海考古、危险水域探查等任务。OceanOne^K是OceanOne的升级版，专为深海探索而设计，具有双手操纵、立体视觉和人机触觉交互功能。该机器人可抵达的最大深度为1000m，能够探索更广泛的水生生态系统。通过触觉反馈，OceanOne^K可以实现操作人员与水下环境的交互，并能灵活使用工具和设备进行操作。

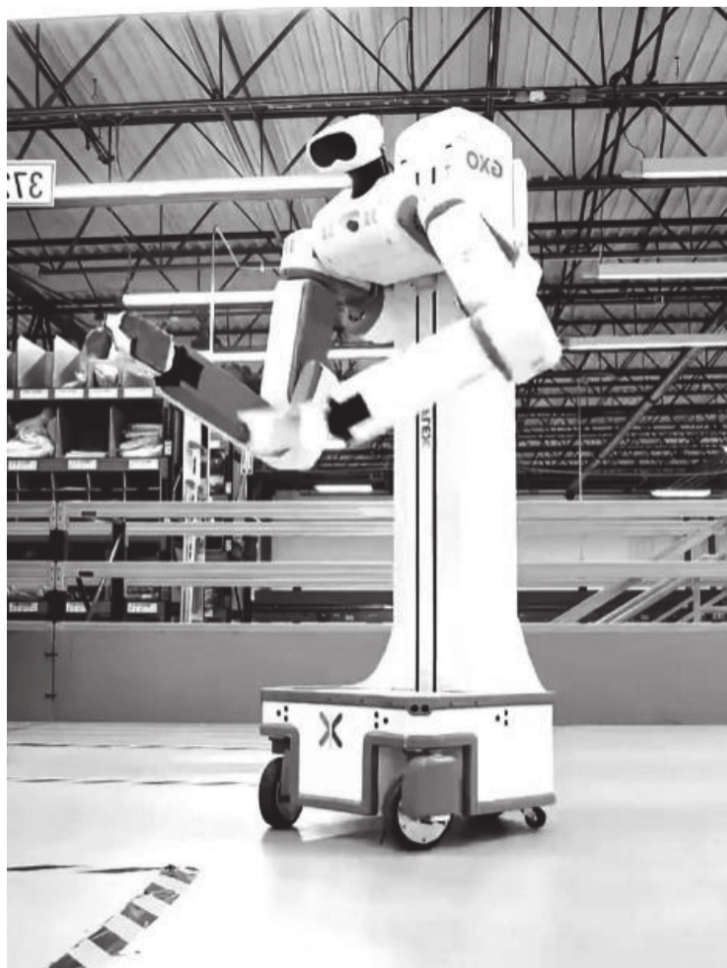


图A-8 OceanOne (图片来源: Stanford Robotics Lab官网)

OceanOne^K长1.5m, 配备了全面的作业工具。一方面, 它采用双臂仿生设计, 机械手具备与人类相当的自由度, 可完成货物抓取、器械操作等精细动作。同时, 力传感器阵列能够将水下机械手接触物体的材质、形状等物理特性转化为可感知的触觉信号, 操作员可实时调整抓取力度。另一方面, 它的头部配有立体视觉系统, 由双目摄像头搭载LED阵列照明模块组成, 在低能见度水域仍可提供高清影像, 并通过光纤将画面实时传输至水面控制台。

A.9 Reflex

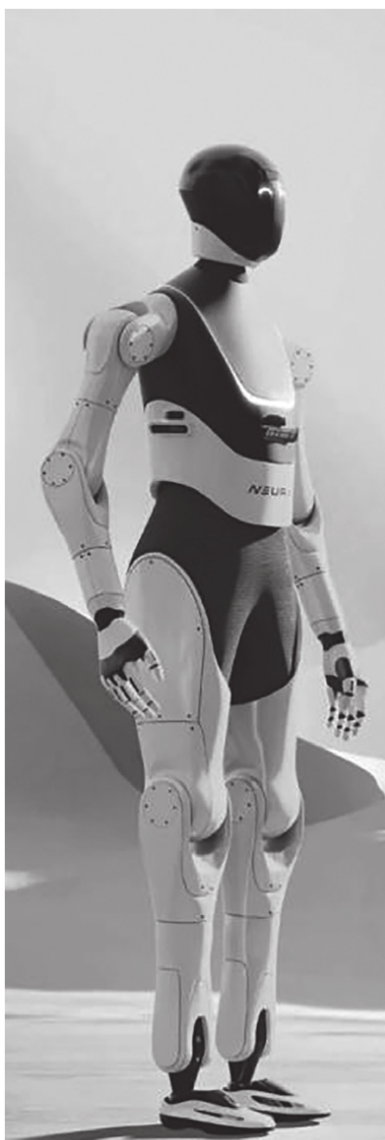
Reflex是美国Reflex Robotics公司开发的轮式移动底盘人形机器人, 如图A-9所示。该机器人于2024年在亚特兰大的Modex展会上进行了展示, 可高效地从货架上检索物品, 并向用户提供食品和饮料。在狭窄的货架或空间内, Reflex凭借其灵活快速的移动能力, 轻松应对商品分拣或订单分类等烦琐任务。其轮式移动底盘的万向转动、可旋转的躯体, 以及灵活的机械臂和夹爪, 结合AI驱动技术, 使得Reflex能够迅速找到需要分拣或调配的商品, 并高效地将其放置在货架的指定位置。Reflex的躯体可上下伸缩, 几乎不受空间限制。此外, Reflex不仅能分拣货物, 还能稳定地推送载物推车, 执行多样化且高重复性的体力任务。



图A-9 Reflex (图片来源: Reflex Robotics公司公开视频截图)

A. 10 4NE-1

4NE-1是德国Neura Robotics公司基于认知神经科学开发的第三代协作型人形机器人,如图A-10所示,具备工业与家庭场景的多任务执行能力。该机器人于2024年5月首次公开样机。4NE-1可以向前和向后行走、转弯、弯曲,并能安全地在不同的地形和楼梯上行走,行走时还能平稳地持握手中的物品。



图A-10 4NE-1（图片来源：Neura Robotics公司官网）

4NE-1的身高为180cm，体重约80kg，最大负载15kg。4NE-1集成了AI、模块化前臂，以及包括电子皮肤在内的多种传感器技术，支持人类和机器人之间的多模式交互，如语音、手势等。它还可以识别不同的声音、语言，甚至语气和情绪。4NE-1采用轻量化碳纤维结构，头部配备可定制交互屏幕，支持语音指令接收与视觉反馈。机械手臂拥有27个自由度，手部集成触觉传感器，可完成0.1mm精度的物体抓取。足部采用动态平衡算法，适应楼梯、砾石等复杂地形，传感器为所有移动关节提供完整的扭矩反馈，使4NE-1能够识别行走的地面状态并相应地保持平衡。目前，4NE-1已展示出熨烫衣物、搬运货物、食材处理等复杂操作能力。

▷附录B 国内典型人形机器人整机产品

B.1 天工

天工是北京人形机器人创新中心（国家地方共建具身智能机器人创新中心）研发的通用人形机器人。2024年4月27日，天工机器人首个版本——天工1.0LITE发布。经历4个月的升级迭代，北京人形机器人创新中心推出了拥有完整的手臂功能、内置多种传感器的天工1.1PRO，以及在天工1.1PRO的基础上进一步加强交互能力的天工1.2MAX，如图B-1所示，以满足不同的应用场景需求。天工机器人在功能上亮点突出，运动能力出色，能以6~12km/h的速度稳定奔跑，适应多种复杂地形。2025年4月19日，天工系列机器人天工Ultra参加全球首个人形机器人半程马拉松，以2时40分42秒的成绩夺得冠军。



图B-1 天工1.2MAX（图片来源：北京人形机器人创新中心有限公司官网）

天工1.2MAX的身高为1.73m，体重60kg，拥有42个自由度，可扩展算力为550TOPS，拥有4个全场景双目结构光3D相机。手部触觉传感器的精度在0.3N以内，重复定位精度为 $\pm 0.1\text{mm}$ 手臂额定负载4kg，最大末端速度为1.8m/s。天工机器人采用“基于状态记忆的预测型强化模仿学习”方法，提前采集大量人类运动视频等数据，提炼出运动时各个关节的位置、速度信息供其“学习”，通过模仿提升人形机器人在非结构化环境中的适应性。天工机器人同时具备开源开放与可扩展性的特点，可以实现开放调用通信接口，灵活扩展软、硬件等功能模块，充分满足不同应用场景下的需求，方便开发者进行二次开发。

B.2 青龙

青龙是由人形机器人（上海）有限公司（国家地方共建人形机器人创新中心）研发的全尺寸开源公版人形机器人，如图B-2所示。青龙于2024年7月4日在世界人工智能大会期间正式发布。2025年3月，青龙曾以参赛者身份亮相上海长江半程马拉松，完成稳定奔跑1h的挑战。



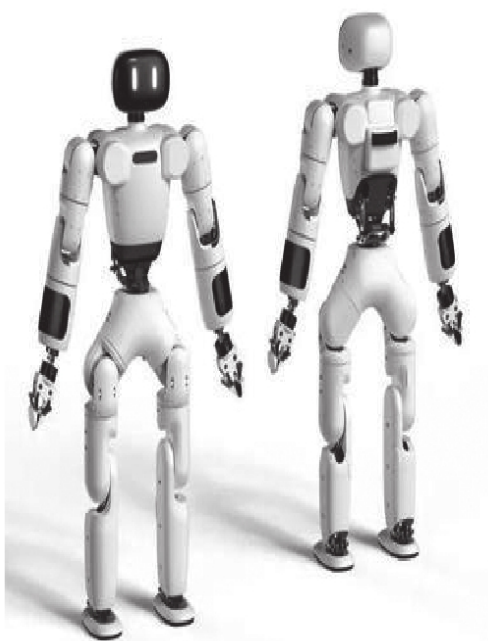
图B-2 青龙 [图片来源：人形机器人（上海）有限公司官网]

青龙人形机器人基于强对标、多模态、通用化和智能化的四大原则，围绕人形机器人总体能力提升开展系统工程化设计，全体系对标生物人功能指标，打通“感知-决策-执行”链路，打造高可靠、实用化和智能化的人形机器人，其定位为通用人工智能软硬件开发载体。青龙人形机器人同时具备高机动下肢行走配置与轻量化高精度上肢作业配置，支持快速行走、敏捷避障、稳健上下坡和抗冲击干扰四大运动功能，支持多模态感知、交互和操控。青龙人形机器人的身高为1.85m，体重80kg，全身配备43个主动自由度及7自由度机械臂，采用仿生躯干构型和拟人化运动控制。该机器人搭载算力达400TOPS的控制器，集成朱雀具身大脑与玄武小脑模

型，支持多模态感知交互和五感融合感知。青龙集成了“视、听、触、嗅、动”五感融合设计，配备高扭矩关节模组（峰值扭矩为 $400\text{N}\cdot\text{m}$ ）与能量回收电源系统，可持续运行3~4h，支持负重40kg行走、10km/h持续奔跑及复杂地形避障。

B.3 灵犀X1

灵犀X1是由智元机器人公司研发的全栈开源人形机器人，如图B-3所示，于2024年8月18日首次发布，2024年10月24日正式面向全球开源。该机器人采用串并联混合构型手臂、差分驱动双肩关节及自研PowerFlow关节电机，具备模块化设计和高自由度结构，支持机-机模式交互。2025年1月21日，灵犀X1实现批量交付，整装组件价格约为10万元人民币，推动具身智能机器人进入实用化阶段。

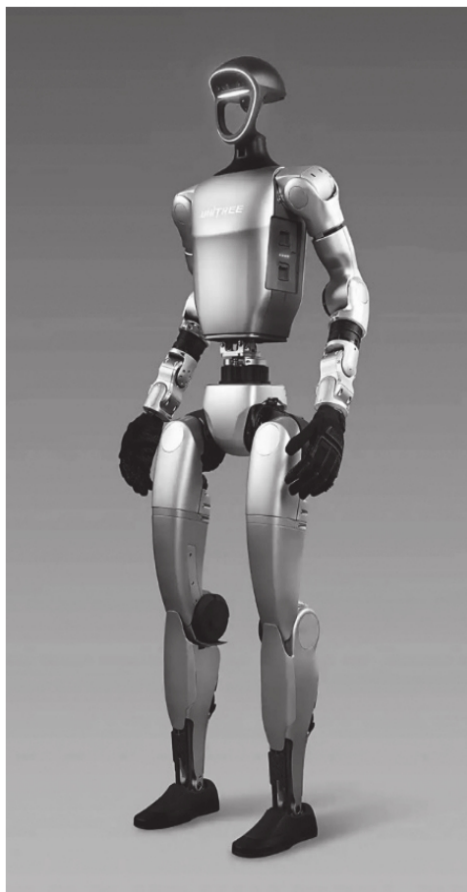


图B-3 灵犀X1（图片来源：智元机器人公司官网）

灵犀X1整机共29个自由度，其中手部7个自由度、腿部6个自由度、腰部3个自由度。构型设计采用串并联混合构型手臂与差分驱动双肩关节，兼顾灵活性与负载能力。驱动系统搭载自研PowerFlow关节电机（PF86/PF52），具备中空走线、绝对值编码功能。末端执行器配备OmniPicker自适应夹爪，可夹持从工业部件到缝衣针等多种物体。控制方式支持游戏手柄操控与编程控制，并通过AI视觉多模态大模型实现环境感知。作为工业级开源机器人，灵犀X1适用于在轻服务场景完成物品递送、基础操作等。在教育研发场景，可以通过灵犀X1的全栈开源特性降低机器人的开发门槛。在智能家居领域，灵犀X1结合具身智能技术可以实现家庭辅助功能。

B.4 Unitree G1

Unitree G1（以下简称G1）是杭州宇树科技有限公司旗下的人形机器人，如图B-4所示，于2024年5月13日发布。2025年3月6日，G1在央视新闻直播中亮相，流畅完成“李小龙式”回旋空踢等高难度动作。



图B-4 Unitree G1（图片来源：杭州宇树科技有限公司官网）

G1身高约132cm，体重约35kg，具有超越常人的灵活性。其小跑速度大于2m/s，拥有23~43个关节电机和较大的关节运动空间，最大关节扭矩达到120N·m，可进行高难度的动态动作，如平躺状态下站起、坐下折叠、舞棍等，智能电池容量达9000mAh，续航时间约为2h。而且G1还基于深度强化学习和仿真训练，在不断进行升级和演进。2025年4月，G1在针对具体需求进行功能开发后，已在青城山-都江堰景区“上岗”，进行多语种景点介绍、游客导流、娱乐表演，其搭载的UnifoLM大模型能实时响应用户需求，实现流畅交互。

B.5 GALBOT G1

GALBOT G1是银河通用机器人公司于2024年6月推出的首代具身大模型机器人，如图B-5所示。其采用轮式底盘与双臂融合设计并搭载多模态感知系统。该机器人通过

三维视觉合成数据与Sim2Real仿真训练，实现95%成功率的透明物体抓取、衣物处理等泛化操作能力。

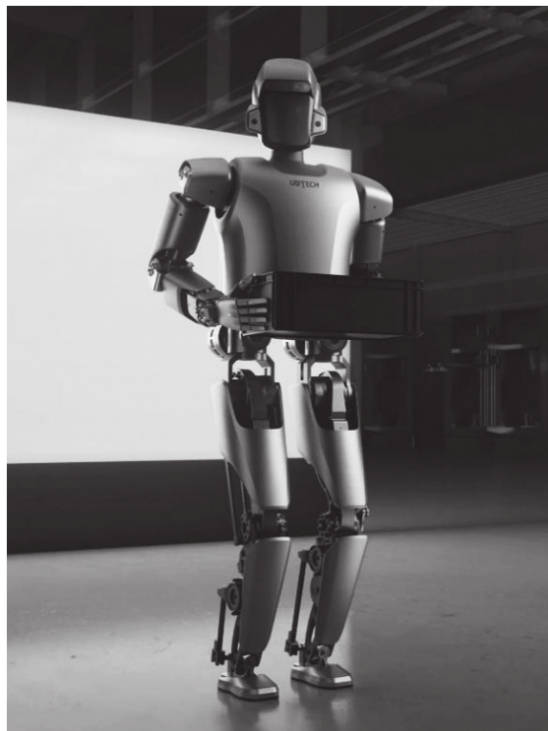
GALBOT G1采用轮式底盘、双臂与单腿组合的模块化结构，站立高度达173cm，身体升降高度为65cm，操作范围覆盖长度为0~240cm的空间。其机械臂具备190cm的臂展，可配置吸盘、自适应夹爪、灵巧手等执行机构，完成透明玻璃瓶抓取、衣物悬挂等差异化任务，续航时间为10h。该机器人通过可折叠躯干实现站立/跪姿模式切换，在商业场景中的平均移动速度达1.5m/s。GALBOT G1系统采用硬件层-技能层-认知层三层技术架构：硬件层集成深度相机、力传感器等多模态感知设备，技能层基于百万级场景数据训练操作技能库，认知层通过十亿级数据构建具身大模型。在2024年世界机器人大会上，GALBOTG1完成了桌面清理、货架补货等4项工业场景真机演示，任务平均耗时2分17秒。在零售领域，GALBOT G1已在北京无人药店实现药品分拣与配送，单日处理订单量超过300件。工业场景中，GALBOT G1通过与奔驰、极氪等车企合作，承担汽车天窗转运与拆踩作业，单次搬运负载达15kg。截至2025年6月，其商业化应用已覆盖北京7个无人药店，并计划在超市订单处理、汽车工厂配料等场景实现规模化部署。



图B-5 GALBOT G1（图片来源：银河通用机器人公司官网）

B.6 Walker S1

Walker S1是深圳优必选科技股份有限公司于2024年10月14日发布的全新一代工业人形机器人，如图B-6所示，专注于汽车制造等工业场景。该机器人通过深度学习与多模态推理技术实现毫米级精准质检，准确率超过99%，并在运动控制、关节稳定性和整机结构可靠性等关键技术领域取得突破。



图B-6 Walker S1（图片来源：深圳优必选科技股份有限公司官网）

Walker S1采用多模态推理大模型，通过全身摄像头实现毫米级检测精度，其关节稳定性与整机结构可靠性已经实训验证，运动控制算法升级后的质检效率提升了100%。电池续航与软件系统稳定性达到工业级标准。Walker S1身高172cm，体重76kg，可完成15kg负载行走。其采用集成化头部设计，结合双耳鱼眼相机、3D立体视觉，具备360°多模态感知。全新仿人灵巧手拥有6个阵列式触觉压力传感器，配置全栈式灵巧操作策略库。一体化关节的最大扭矩为 $250\text{N} \cdot \text{m}$ 。截至2025年3月，Walker S1已进入比亚迪、东风柳汽、富士康等车企进行实训，承担安全带检测、车门锁检测、物料搬运等任务。其通过群体智能技术实现多机协同作业，在郑州富士康车厂实训期间，工作效率提升30%。

B.7 夸父（Kuavo）

夸父是乐聚智能（深圳）股份有限公司于2023年12月5日对外发布的高动态人形机器人产品，如图B-7所示，主要面向家庭场景，也是中国国内首款可跳跃、可适应多地形行走的开源鸿蒙人形机器人。



图B-7 夸父 [图片来源：乐聚智能（深圳）股份有限公司官网]

夸父重约45kg，其步速最高可达4.6km/h，快速连续跳跃高度超过20cm，全身拥有26个自由度。现阶段，该机器人已经能够成功实现任务作业、规划导航、复杂地形自主行走、连续跳跃等功能。该公司自研的一体化关节的峰值扭矩达 $300\text{N} \cdot \text{m}$ ，扭矩密度达到 $200\text{N} \cdot \text{m}/\text{kg}$ ，可充分满足机器人高动态、高爆发、高精度的需求。通过对关节结构和热管理模型的优化，大幅提高了电机的散热效率，增强了稳定性，延长了机器人的工作寿命。夸父是首款基于开源鸿蒙系统的专业级人形机器人，搭载了7种传感器，可以实现与手机、平板计算机、台式计算机等设备的实时互联。

B.8 STAR1

STAR1是由北京星动纪元科技有限公司研发的首款产品级高性能通用人形机器人，如图B-8所示，于2024年8月正式发布。



图B-8 STAR1（图片来源：北京星动纪元科技有限公司官网）

STAR1的身高为171cm，体重63kg，具备55个自由度，最高关节扭矩达 $400\text{N} \cdot \text{m}$ ，室外奔跑速度突破了 3.6m/s ，双电池组供电方案支持连续2h高强度运动。其核心硬件包括7自由度仿生机械臂和12自由度灵巧手“星动XHAND1”，可实现单手抓握25kg负载和捏骰子等操作，三维力触觉传感器的精度达 0.1N 。通过机器人端到端大模型ERA-42，并基于强化学习，STAR1实现动态运动规划与环境自适应。其采用“大小脑”协同设计、英特尔中央处理器处理决策指令、NVIDIA Orin计算单元实现实时运动控制。2025年3月，STAR1在全球首个人形机器人马拉松赛事中实现连续1h户外奔跑，并在河西走廊戈壁滩、西安马拉松等场景中展示了复杂地形适应能力。截至2025年5月，STAR1已完成2000h工业场景可靠性测试，进入量产准备阶段。

B.9 CL-3

2025年5月8日，深圳逐际动力科技有限公司发布全尺寸人形机器人CL-3，如图B-9所示。CL-3通过软硬件的协同创新，突破了传统人形机器人的动态控制瓶颈。

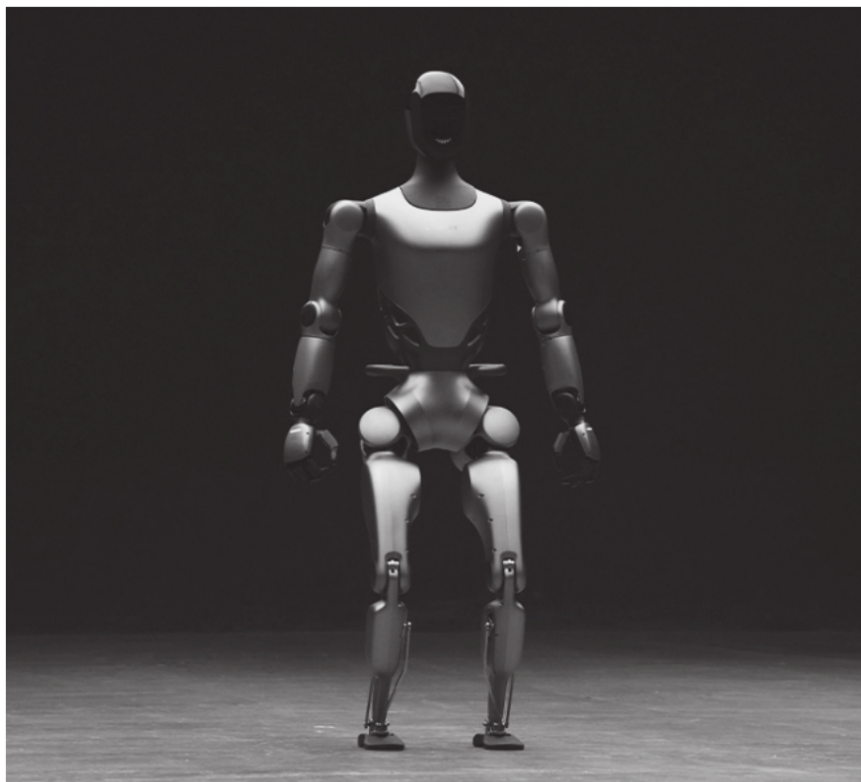


图B-9 CL-3（图片来源：深圳逐际动力科技有限公司官网）

CL-3的机械结构采用仿生关节设计，全身包含31个自由度，覆盖躯干、四肢全链路运动，通过配合全身运动控制算法，可实现行走、跑跳、上下楼梯等高动态动作。相较于前代产品，其运动控制算法实现了三个方面的升级：一是通过强化学习优化动作轨迹预测精度；二是引入阻抗控制增强环境适应性；三是运用全身动力学解耦技术提升多关节协同效率。实测数据显示，CL-3在负重20kg状态下仍能保持1.5m/s的稳定行走速度，动态平衡恢复时间缩短至0.3s。CL-3内置了超过200个预定义动作模块，涵盖工业操作、家庭服务、应急救援等场景。开发者可通过图形化编程界面快速编排动作序列，支持Python/C++二次开发接口。该系统特有的“动作迁移”功能，能够将人类示范动作通过传感器捕捉转化为机器人可执行指令，将动作开发效率提升80%。现场演示中，CL-3仅用15min即可完成羽毛球发球动作的自动生成与优化。CL-3的模块化电池系统支持8h持续作业，末端执行器兼容国际标准化组织（ISO）标准工具接口。在汽车制造场景测试中，CL-3已完成扭矩扳手安装、精密部件搬运等复杂工序，定位精度达到 $\pm 0.1\text{mm}$ 。

B.10 GR-2

GR-2是上海傅利叶智能科技有限公司于2024年发布的GRx系列人形机器人产品，如图B-10所示。该机器人基于前代GR-1的技术积累，优化了全身结构设计，搭载全新自主研发的12自由度灵巧手及第二代执行器FSA 2.0。



图B-10 GR-2（图片来源：上海傅利叶智能科技有限公司官网）

GR-2基于GR-1的数据积累和技术积淀，在硬件设计、框架开发、商业化等关键环节都实现了全面突破，且拥有更灵巧的上肢操作、更强劲的动力系统和更开放的开发平台。GR-2的身高为175cm，体重为63kg，全身共有53个自由度，最大关节扭矩为 $380\text{N} \cdot \text{m}$ ，步态行走速度为5km/h，电池续航为2h。GR-2采用模块化设计，其第二代执行器FSA 2.0相较于前代提升了力矩密度与控制精度，整机的运动性能得到显著优化。GR-2搭载新的12自由度自研灵巧手，并在单手配置6个阵列式触觉传感器，可以对不同形状、材质的物体实行精准抓握。同时支持多种示教方式，通过采集更多真实训练数据，弥补仿真与现实的差距。截至2025年，GR-2机器人已在导览咨询、学术科研、工业制造、医疗康复等领域开展实验性应用。

