

人形机器人,何时跨越“幼儿期”

随着“具身智能”等关键词首次写入政府工作报告,2025年被业界称为人形机器人量产元年,全行业正迎来快速发展阶段。一方面,它们能跑能跳,可以成为高效劳作的得力干将;另一方面,它们“反应迟钝”“不够智能”,智慧犹如处在“幼儿期”。从实验室走向日常生活的过程中,人形机器人的“成人”之路刚起步。

■ 本报记者 何冬健

8月,全球首个以人形机器人为参赛主体的综合性赛事——“2025世界人形机器人运动会”在国家速滑馆“冰丝带”开幕。形形色色的“钢铁选手”被推到聚光灯中央。这些拥有成人体格的“运动员”,看起来“不太聪明”。它们会毫无征兆地偏航、摔倒、看不懂参赛规则。网友不禁调侃它们“像极了我家学走路的娃”。但对行业专家而言,每一个摔倒的瞬间,都暴露出人形机器人在感知、遥控、零部件等方面的诸多缺陷。美国加州大学伯克利分校讲座教授肯·戈德堡(Ken Goldberg)做过一个有趣的对比:“人类预测两百万公里之外的小行星运动,远比预测机器人抓取物体时的运动要准确得多。”谁承想,比起宇宙尺度的运算,让机器器完成“抓杯子”这样的日常动作,竟成了机器人发展的一道难关。站在风口上的人形机器人,它们当前不过是一群刚学会站立的“幼儿”——四肢发达,头脑简单。它们如何才能跨过“幼儿期”?

成人般的外形 幼儿般的笨拙

宁波市海曙区的浙江人形机器人创新中心,几十台机器人和平均年龄不到35岁的研发团队相互围绕,充满未来感。实验室中,一台尚处于研发阶段的人形机器人正尝试独立行走。它的膝盖微微发颤,关节发出机械嗡嗡声,在科研团队的“使坏”干扰下,每迈出几步都要停顿片刻,笨拙的姿态像极了刚学会走路的孩子。 “幼儿”般的笨拙,在科学层面恰与人类进化历程形成呼应——百万年前,古猿“露西”(记者注:在现今的埃塞俄比亚发现的拥有318万年历史的著名古人类,被称为“人类祖母”)的枕骨大孔位置与现代人一致,股骨向内侧倾斜,足弓清晰可见,这正是人类直立行走的最早实证。仿人形而生的机器人承载着人类的骄傲——我们坚信自身躯体是世间最精密的“机械”,人类所有造物皆以此为蓝图;它们也寄托着人类的想象——倘若借钢铁之躯辅以具身智能,机器人又将如何塑造世界?

后空翻、跳舞、跑步……今天,人形机器人的动作能力令人惊叹。但这些精彩表演背后是高度可控的环境和反复调试的结果。 “我们最大的愿望是让机器人干活。但现在,距离家里、工厂里大规模应用,技术还差得远。”宇树科技创始人王兴兴在接受媒体采访时坦言。这位曾因制造春晚“扭秧歌”机器人刷屏的青年科学企业家,对技术边界保持着清醒。在进行机器人研发时,科学家们当然不会满足于扫地机器人、自动炒菜机、外卖机械狗这样功能相对单一的产品。在他们的设想中,人形机器人的双腿能像人类一样稳健、快速、灵活,使其能够到达任何地点,同时双手也要足够灵巧,能够处理从家庭到工厂的各种任务,具备“上厅堂、下厨房、进工厂”的能力;更重要的是,脑袋要足够聪明,真正向一个能够感知、推理并与物理世界互动的具身智能演进。

“现实情况是,大多数机器人仍然依赖几十年前的工业机器人的技术,这极大限制了其发展。”杭州旷行科技创始人舒江鹏说。舒江鹏团队多年来专注基建领域巡检机器人的研发。他介绍,当前浙江市场真正在各种场景上应用了巡检机器人的,甚至不足1%。 “市场占有率最能说明问题,效率低的工具自然无人问津。”他说,人形机器人在马拉松赛场上的表现很直观,要在无遥控状态下自主跑一分钟不摔跤,目前都比较困难。未来要让机器人真正“干活”,必然不能依赖遥控。

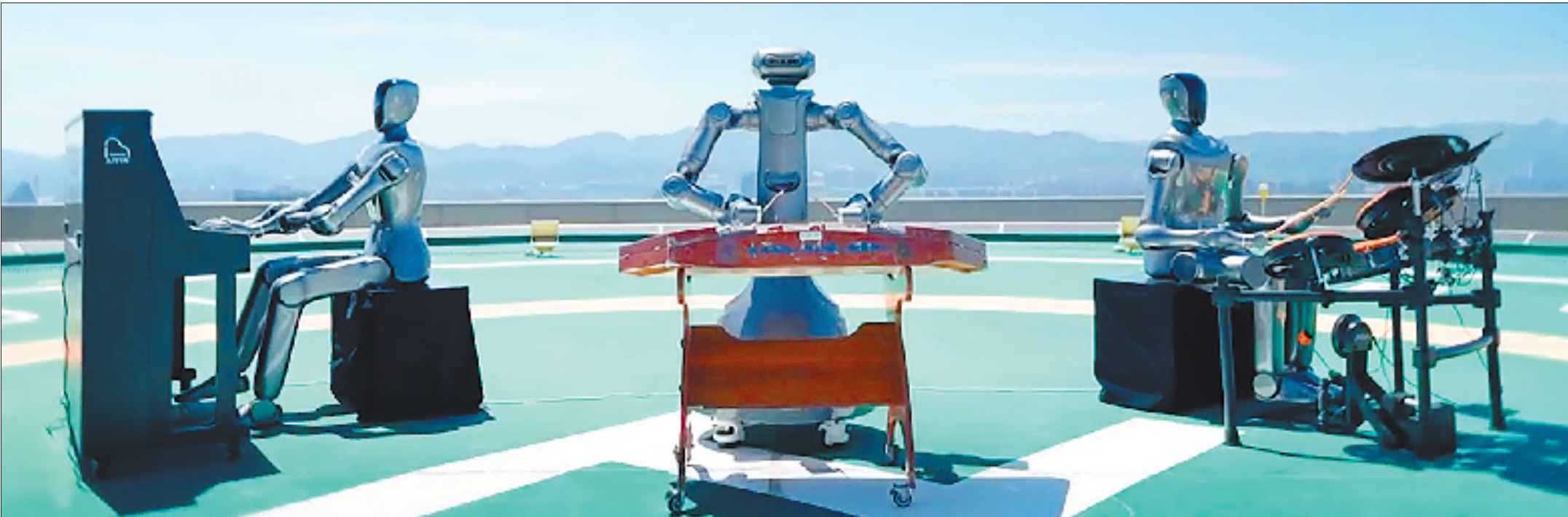
这种笨拙在工业场景中被无限放大。轮式机器人不用考虑平衡,重心低,碰到台阶就着走就行。人形机器人要直立行走,地面稍有不平就容易失稳——摔一次,身上的传感器等精密部件可能就会损坏需要更换,维修成本比请几个工人还高。

但这或许正是技术最迷人的地方。就像幼儿从只能匍匐前行到独立行走需要18个月,从跌跌撞撞到健步如飞需要数年,人形机器人,其发展仍面临诸多亟待解决的困难。

变“聪明” 需要学会用“脑”

如果把机器人的身体比作一个人,那么藏在关节里的运动控制系统就像人类的小脑,负责将指令转化为具体动作。搭载的深度学习算法则是大脑,承担着感知、决策与认知的重任。

然而,这对“大小脑”组合的发育速度,远未达到科学家的期待。舒江鹏表示,如果说“大脑”还停留在“刚学会说话的孩童”阶段,那么“小脑”



“和璇”机器人乐队在演奏。

受访者供图



第四届中国-中东欧国家博览会暨国际消费品博览会上,浙江人形机器人创新中心工作人员在“数智中东欧”展区展示一款零售机器人。

新华社发

成熟度也仅仅是“小学一年级新生水平”。

在杭州海创人形机器人创新中心展厅,由三台人形机器人组成的“和璇”乐队通过机械手和机械臂演奏钢琴、扬琴、架子鼓,流畅弹奏着《青花瓷》《沧海一声笑》《孤勇者》等歌曲。

演奏的“秘诀”,在于灵巧手、机械臂与算法的协同。浙江省机器人产业发展协会秘书长、杭州海创人形机器人创新中心副主任宋伟介绍,程序会先解析曲谱,确定每个音符对应的“手指”与时机,再通过驱动器件将算法决策转化为动作。团队当前研发出来的机器人乐队在该领域已具备国际领先水平。

然而,这种模式似乎只是在执行预设程序,就像一个记忆力惊人却缺乏思考能力的孩子。“真正的挑战在于,当乐谱突然出现变奏,它不能像人类演奏家那样,靠触觉、听觉等多种感知即时调整力度与节奏。”杭州海创人形机器人创新中心总工程师谢安桓说。

问题的核心之一,在于机器人的“大脑”太“年轻”。

与人一样,机器人的运行方式是持续循环的过程——感知、思考,然后运动。感知、运动可以在传感器、关节等硬件上下功夫,比如宋伟团队将“和璇”的机械臂从12千克的重量减轻到6千克,机械臂越轻巧就越容易控制。但教会机器人思考,甚至协同,可比让它更好感知、运动难得多。

宋伟告诉记者,人类的大脑能同时处理视觉、听觉、触觉等多模态信息,还能调用记

忆、经验甚至直觉,可现在的人形机器人“大脑”,遇到需要综合判断的场景就容易卡壳。通俗地讲,以神经网络训练为代表的人工智能大模型让机器人变聪明了,但距离真正融入人类社会还远远不够。

当前,如果要实现人形机器人的即兴演奏,需突破三大核心层面。在“大脑”层面,目前依托成熟的人工智能作曲大模型,已可实现即兴的旋律生成与变化设计;在“小脑”层面,若运动控制算法无法快速精准求解,便会出现“指令落地失效”的断层问题;在硬件性能上,对电机的响应速度、扭矩精度等核心指标提出了更高要求。

“我们的核心目标,是推动人形机器人与人类实现协同创作及演奏,如同人类艺术家间的即兴互动。当人类演奏者完成一段旋律后,机器人能即时捕捉并快速回应,以即兴演奏形成互动。”宋伟说。

舒江鹏的团队核心由土木建筑与人工智能交叉领域人才组成。舒江鹏深有感触:“现在开发机器人的多是计算机、自动化专业人员,若若想让机器人干基建工程领域的活,它得先懂专业知识——这是计算机工程师教不了的。”培养一个工程师通常需要4年本科的专业学习,机器人也需要“行业老师”的教导。

可目前,既懂专业知识又懂人工智能的复合型人才寥寥无几。“让写代码的人教机器人写代码没问题,但教它学修桥梁、梁隧道,跨专业难度很大。”他说。

“摔”得多 就能“长”得快吗

就像孩子学走路,人形机器人今天“摔”得多,“成长”也许就会更快。

今年4月,“天工Ultra”参加人形机器人半程马拉松时,还需依赖领跑员发射无线信号跟随,采用半自主奔跑模式。在4个月后的世界人形机器人运动会上,它已实现全自主奔跑,更在百米“飞人大战”中以21.50秒夺冠。

“今年人形机器人火爆。从长远来看,没准这只是一个小火苗,就像当初互联网的诞生一样。”接受媒体采访时,王兴兴表示。

人形机器人的走红,实则与一个诞生已久的概念紧密相连——具身智能。早在1950年,计算机科学之父艾伦·图灵就在经典论文《计算机与智能》中首次提出“智能体需要与物理世界互动”的设想。

作为具身智能最受关注的关键载体,人形机器人正从实验室快步走向大众视野,成

为科技行业的焦点。那么,机器人进化的终点,真的非“人形”不可吗?

前不久在北京开张的全球首个具身智能机器人4S店里,有超100家机器人产业链上下游企业明确表达了入驻意向。其中的大多数企业专注领域是非人形机器人。

业内普遍认为,相较具身智能的外在形态,机器人的内核更关注。毕竟,无论是四足机器人、轮式机器人,还是人形机器人,归根结底还是要切实解决实际问题、提升生产生活效率。

那么,距离机器人真正“敲门”进入人们的生活还需要多久呢?

在人形机器人的未来发展路径上,行业正上演着多种思路的激烈角力。

比如,有学者坚信“数据万能论”,试图通过给机器人提供足够多的训练场景,让它在虚拟世界里“摔”上亿次,在真实环境中“试”上万回,终有一天,机器人能通过海量数据学会运动控制的精髓。还有的专家更强调“工程智慧”,认为机器人的运动控制本质是对物理规律的数学建模,需要回到基础力学、控制理论的底层逻辑,用工程师的巧思,而非单纯的数据堆砌来突破限制。

不过,从纯技术的角度来看,解决现存的难点问题或许指日可待。

目前,为整合不同形态与平台的机器人数据,业界已推出迄今为止全球最大的开源真实机器人数据集。由于机器人领域的数据集普遍存在规模偏小、多样性不足的问题,如何将人类动作捕捉数据转化为机器人可利用的动作空间,至今仍是该领域的前沿研究课题。

宋伟大胆预测,未来五到十年,全球至少有一家企业可以把握比较通用的机器人大模型开发出来。这个基础大模型就像是一个完整的积木,大语言模型只是其中一块,其他组成部分还包括视觉感知、力觉感知、决策和交互等。

行业的突围正在悄然展开。输电塔的检修现场,杭州旷行科技研发的机器人沿着钢架攀爬检修;警务人员的执法现场,杭州逻腾科技的球形机器人可以完成巡逻、侦测、救援等任务……

“我们更希望让机器人去干人类不想干的活,而不是抢人类的饭碗——矿山井下、电力高塔、隧道深处这些恶劣环境下的工作。”舒江鹏说,施工现场哪怕一道简单工序,都需要机器人“看懂文字、听懂指令、识别三维环境”。他们正在尝试通过大模型训练,让机器人掌握3到5个固定工序,从“会干活”开始,再追求“干好活”。

这些钢铁“幼儿”的跨越,也需要技术的沉淀、场景的打磨,以及对“成长周期”的耐心。当它们真正摆脱“幼儿期”的那天,会带着更聪明的大脑、更灵活的双手,成为我们身边不可或缺的伙伴。

专家观点

人类与机器的关系 将发生颠覆性变革

■ 翟雪松 钱春雨

具身智能(Embodied Intelligence),当今科技界最热门的话题之一。它不仅是技术概念的飞跃,更预示着我们与机器的关系将发生颠覆性变革。

简单来说,具身智能就是赋予AI一个物理“身体”。它不再仅仅是处理数字信息的程序,而是能够像人类一样,通过感知、决策和行动,与真实世界进行深度互动的一种智能实体。

这背后是多项前沿技术的集成:如生物传感(Biosensing),指通过可穿戴设备等载体,实时捕捉和感知人体的生理信号,例如眼动、心率等,从而理解个体的认知状态;空间计算(Spatial Computing),指利用同步定位与建图(SLAM)等技术,让机器能够像人一样感知和理解三维物理空间中的环境、物体和结构;生成式人工智能(GAI),指大模型不再只停留在生成文本或图像,而是能够根据环境变化和任务需求,实时生成多模态的学习资源或行动方案。

具身智能赋予了智能体“身体在场”的能力。它能够将抽象的概念转化为可感知、可操作的学习内容,增强知识与实践的深度连接,并能够根据环境和任务的变化实时调整决策。正如图灵奖得主杨立昆所预言,未来的趋势将是从小语言模型转向由具身智能驱动的“世界模型”,因为后者才能真正理解多变复杂的物理世界。

在所有具身智能的物理载体中,人形机器人无疑最受瞩目。这是因为人形机器人的设计初衷,就是模仿人类的形态,从而更好地融入人类为自己设计的社会环境。无论是操作工具、打开房门,还是在复杂的环境中行走,人形机器人的类人形态使其能够适应现有的基础设施,无需对环境进行大规模改造。

具身智能的飞速发展,正在让人形机器人真正成为能够自主思考、决策和行动的“智能体”。它将大模型的理解能力、常识推理能力与机器人的物理执行能力相结合,使得机器人能够根据语言指令理解任务意图,并在开放、非结构化的环境中完成复杂任务。

具身智能驱动下的人形机器人,正在加速从科幻走向现实,其未来的作用将渗透到社会生产与生活的方方面面。

如,工业与制造业的“新劳力”。过去,工业机器人主要在标准化的流水线上进行重复性工作。而具身智能机器人则能够胜任更多非标准化、柔性化的任务,填补传统自动化无法覆盖的空白领域。这不仅能够提升效率,还能解放人类从事更具创造性的工作。

如,家庭与服务领域的“新伙伴”。当人形机器人的成本降至消费级别价格,并具备更强的认知能力时,它们将有望走进千家万户。它们可以成为我们的家庭管家、朋友和护理员。此外,它们还能在家庭安防、娱乐和教育等多个场景中发挥作用。

教育领域的“新模式”具身智能应用也已初见端倪。利用集成生物传感、空间计算等技术的可穿戴式教育装备,未来的学习将不再受限于教室或书本。例如,学生可以佩戴智能眼镜,在博物馆里与虚拟的恐龙进行实时互动,或是将化学实验中的虚拟分子模型叠加在现实世界中进行操作。这种“身体—环境—认知”的深度耦合,将极大增强学习的沉浸感和效果,真正实现“随时随地、无处不在”的泛在学习。

(作者分别系浙江大学科学与技术教育学位点负责人、浙江大学科学与技术教育硕士生)



人形机器人在重庆举行的2025世界智能产业博览会上与观众互动。

新华社发



视觉中国